

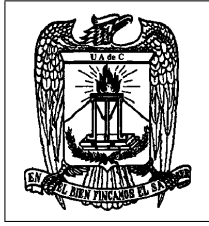
UNIVERSIDAD AUTÓNOMA DE COAHUILA  
CENTRO DE INVESTIGACIONES SOCIOECONÓMICAS

TESIS

TESIS  
2012

HADA MELISSA SÁENZ VELA

“Análisis Exploratorio de Características  
Socioeconómicas de México mediante la  
Aplicación de Métodos de Agregación Espacial, 2010”



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

TESIS

**“Análisis Exploratorio de Características Socioeconómicas de México  
mediante la Aplicación de Métodos de Agregación Espacial, 2010”**

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

**HADA MELISSA SÁENZ VELA**

**Comité Evaluador:**

**Director:** Dra. Myrna Leticia Sastré Gutiérrez

**Lector:** Dr. Daniel Arribas-Bel

**Lector:** Dr. Alejandro Dávila Flores

Saltillo, Coahuila

Septiembre, 2012

## AGRADECIMIENTOS

---

En la elaboración de esta tesis participaron muchas personas, directa e indirectamente, así que lo correcto es dar gracias a todos ellos.

En primera instancia agradezco al Consejo Nacional de Ciencia y Tecnología por el apoyo económico a través del cual me fue posible continuar con mis estudios.

Al Centro de Investigaciones Socioeconómicas de la Universidad Autónoma de Coahuila, por haberme aceptado en la Maestría en Economía Regional, y porque sus investigadores y personal administrativo nos apoyó a lo largo de estos dos últimos años.

En particular, quiero agradecer a la Dra. Myrna Sastré, por su confianza y su apoyo, por su paciencia y por todo el tiempo que dedicó a mi trabajo.

A mis lectores, el Dr. Arribas-Bel y el Dr. Dávila, por el tiempo dedicado a la revisión y corrección de esta tesis.

Quiero agradecer a mis amigos y amigas, que son muchos para ser nombrados aquí, pues siempre me dieron ánimo y confiaron en que podría terminar exitosamente la maestría.

A los compañeros de generación, pues de todos aprendí algo y fue grato convivir con ustedes este tiempo.

Pero sobre todo quiero agradecer a mi mamá, que estuvo conmigo en mis buenos y, sobre todo, en mis malos momentos, por ser la fuerza que me impulsa cada día a seguir adelante y no dejarme nunca vencer.

A todos, muchas gracias!!

# ÍNDICE GENERAL

---

<b>Introducción</b>	<b>1</b>
<b>1. Marco Contextual</b>	<b>4</b>
1.1. La perspectiva de disparidades regionales y el uso de agregaciones geográficas . . . . .	4
1.2. La perspectiva del Ordenamiento Territorial y la importancia de identificar características regionales . . . . .	7
<b>2. Las clasificaciones de entornos regionales desde la perspectiva institucional</b>	<b>10</b>
2.1. Definición de áreas rurales . . . . .	10
2.1.1. Estudios realizados para áreas rurales . . . . .	14
2.2. Definición de zonas metropolitanas . . . . .	18
2.2.1. Estudios realizados para zonas metropolitanas . . . . .	22
<b>3. Marco Teórico y Conceptual</b>	<b>26</b>
3.1. Problema de Unidad de Área Modificable y la Falacia Ecológica . . . . .	26
<b>4. Marco Metodológico</b>	<b>29</b>
4.1. Métodos de Agregación . . . . .	29
4.1.1. Métodos sin restricción espacial explícita . . . . .	30

---

4.1.2. Métodos con restricción espacial explícita . . . . .	32
4.2. Aplicaciones de los Métodos de Agregación . . . . .	34
4.3. Algoritmo de Agregación <i>MaxP</i> . . . . .	38
4.4. Datos y fuentes de información . . . . .	39
4.4.1. Diseño de la Investigación . . . . .	39
4.4.2. Datos . . . . .	40
<b>5. Resultados</b>	<b>44</b>
5.1. Visualizando soluciones: Casos particulares de las ZM de Tijuana, Mon- terrey y Mexicali . . . . .	47
<b>6. Conclusiones</b>	<b>55</b>
<b>A. Anexos Metodológicos</b>	<b>58</b>
A.1. Algoritmo <i>MaxP</i> . . . . .	58
A.2. Variables utilizadas en estudios previos . . . . .	59
<b>Bibliografía</b>	<b>61</b>

## ÍNDICE DE TABLAS

---

2.1. Evolución del número de Zonas Metropolitanas en México. . . . .	19
4.1. Variables y dimensiones a utilizar en la aplicación. . . . .	43
5.1. Resultados usando 3 variables de distintas dimensiones. . . . .	44
5.2. Resultados usando 3 variables de la misma dimensión. . . . .	45
5.3. Resultados usando 1 variable. . . . .	45
A.1. Variables analizadas en estudios sobre áreas rurales en México. . . . .	60
A.2. Variables analizadas en estudios sobre zonas metropolitanas en México. . . . .	60

## ÍNDICE DE FIGURAS

---

2.1. Municipios de México bajo la clasificación de la OCDE, 2005. . . . .	13
2.2. Zonas Metropolitanas de México, 2005. . . . .	21
4.1. Clasificación de Métodos de Agregación Espacial. . . . .	30
4.2. Estados y municipios de México, 2010. . . . .	41
5.1. Análisis del caso particular de la ZM de Tijuana. . . . .	48
5.2. Análisis del caso particular de la ZM de Monterrey. . . . .	50
5.3. Análisis del caso particular de la ZM de Mexicali. . . . .	52

# INTRODUCCIÓN

---

Para México la literatura señala la existencia de evidentes disparidades regionales. Por ejemplo, Rey y Sastré-Gutiérrez (2010) encuentran regiones formalmente definidas al utilizar la variable Producto Interno Bruto (PIB) per cápita por estado utilizando el algoritmo *MaxP*. Sin embargo, la mayoría de las agrupaciones propuestas en otras investigaciones no están basadas en los resultados de la aplicación de algún método de agregación, sino que han seguido criterios de los autores, por lo general no explicados (Esquivel (1999), Bouillon *et al.* (2003), Chiquiar (2005), entre otros).

El ordenamiento territorial ha sido una herramienta utilizada con el fin de reducir los desequilibrios regionales. Para ello, requiere que las agregaciones sobre las cuales actuará estén bien definidas, ya sea porque serán beneficiarias de algún programa o porque son vistas como potenciales generadoras de crecimiento regional. En el caso de planes de desarrollo y política pública, el mayor conocimiento de las distintas áreas y zonas definidas en el territorio permitiría una mejor implementación de los mismos. Sin embargo, estas zonas se forman a partir de delimitaciones administrativas que pueden no coincidir con los procesos a estudiar.<sup>1</sup> En México se han definido diferentes tipos de áreas, por ejemplo aquellas con altos niveles de integración y aquellas que se encuentran rezagadas, pero gran parte de estas definiciones no cuentan con un sustento técnico sólido que las respalde.

Lo explicado anteriormente muestra la necesidad de complementar la literatura actual del país respecto a la temática de las agregaciones de unidades geográficas y la inclusión de la perspectiva espacial. La consideración de diferentes efectos espaciales, tales como dependencia, heterogeneidad, escala, etc., en estudios sobre agrupación de áreas es un aspecto importante a considerar (Duque *et al.* (2007), Sastré-Gutiérrez y

---

<sup>1</sup>De ahora en adelante se entenderá por delimitaciones administrativas a las delimitaciones estatales, municipales, de localidades, etc.; mientras que delimitaciones institucionales serán aquellas que se refieran a zonas metropolitanas, áreas rurales, entre otras.

Rey (2008) y Rey *et al.* (2011)). La omisión de estos efectos en las técnicas de agregación de áreas aumenta la probabilidad de caer en errores al realizar interpretaciones de los resultados. Esto último es conocido en la literatura como falacia ecológica, y se debe al Problema de Unidad de Área Modificable.

El Problema de Unidad de Área Modificable (PUAM) puede deberse al problema de escala, es decir, variaciones al considerar diferentes escalas; o bien, al problema de agregación, que es la variación en resultados al considerar distintas agrupaciones de las unidades espaciales. El segundo problema es el que ha llevado a un amplio desarrollo de metodologías para la agregación espacial. El proceso de agregación de áreas es importante pues se encuentran un gran número de aplicaciones en la vida real, como diseño de áreas de mercado (Ríos-Mercado y Fernández, 2009), la conformación de distritos políticos (Hojati, 1996), incluso para el análisis de problemáticas particulares (Rey *et al.*, 2011).

La presente investigación tiene como fin brindar un mayor conocimiento de algunas definiciones institucionales, por ejemplo, zonas metropolitanas y áreas rurales. El estudio es de carácter exploratorio y emplea una herramienta explícitamente espacial como lo es el algoritmo *MaxP* sobre un conjunto de variables. Las definiciones de áreas rurales y zonas metropolitanas se utilizan como base para determinar aquellas características socioeconómicas de la población que han sido de interés en trabajos previos. Se eligió el nivel municipal ya que, por lo general, es la escala utilizada tanto institucionalmente como por investigadores.

## PREGUNTAS DE INVESTIGACIÓN

¿Existe sensibilidad en los resultados obtenidos por el método de agregación espacial a partir de la consideración de diferentes factores composicionales de la población mexicana?

¿Existe sensibilidad en los resultados del algoritmo al considerar diferentes números de variables?

¿Existe sensibilidad en los resultados del algoritmo al considerar diferentes mínimos de municipios para formar las agrupaciones?

#### OBJETIVO GENERAL Y ESPECÍFICOS

El objetivo general de esta investigación es brindar mayor conocimiento sobre las diferentes áreas en el territorio nacional utilizando técnicas recientes de análisis espacial. Esta información pretende complementar estudios individuales realizados hasta ahora para México. Se plantea determinar los niveles de homogeneidad de las agregaciones formadas por el *MaxP* a partir de las diferentes combinaciones de variables socioeconómicas.

##### Objetivos Específicos:

1. Con base en investigaciones previas elegir variables a utilizar que reflejen características socioeconómicas de la población mexicana con información del Censo de Población y Vivienda 2010.
2. Explorar diferentes combinaciones de variables para delinear zonas que consideren la composición socioeconómica y el espacio a través del uso del algoritmo *MaxP* de agregación espacial.

# CAPÍTULO 1

## MARCO CONTEXTUAL

---

Las disparidades regionales en México son un problema persistente como lo indica la literatura. Una de las herramientas empleadas para inducir la reducción de estos desequilibrios ha sido el ordenamiento territorial, por lo que se ha convertido en un elemento frecuente en los planes de desarrollo de algunos países. Dentro de las disparidades territoriales se encuentra la del tipo “rural-urbano”, que se ha ampliado aún más en los últimos años debido al creciente papel de los centros urbanos.

### 1.1 LA PERSPECTIVA DE DISPARIDADES REGIONALES Y EL USO DE AGREGACIONES GEOGRÁFICAS

Uno de los principales enfoques para analizar el desarrollo regional ha sido desde la perspectiva de la convergencia. Uno de los trabajos más reconocidos es el de Barro y Sala-i Martin (1992), sin embargo Quah (1993) advierte que este tipo de análisis no toma en cuenta los movimientos del ingreso entre los estados (países) y no proporciona información sobre la dimensión espacial.

Quah (1996), analizando regiones europeas, sugiere que las unidades espaciales (sean estados, países, regiones) no deben ser consideradas como entidades aisladas. El autor utiliza Cadenas de Markov y examina la distribución del ingreso per cápita de las regiones, encontrando que el factor localización es más relevante que los factores nacionales en la explicación de las disparidades regionales. Concluye que el desempeño de una región se explica mejor en base al comportamiento del ingreso de las áreas vecinas, que respecto a otras regiones del mismo territorio.

En general, los efectos espaciales no son abordados en la mayoría de los estudios. La literatura sobre convergencia sigue presentando un rezago en cuanto a la aplicación de técnicas explícitamente espaciales (Rey y Janikas, 2005). Otra forma abordar el desarrollo de las regiones es con la utilización de medidas de desigualdad, sin embargo éstas pueden ser afectadas por el PUAM. El cambio de escala espacial puede modificar el número de unidades de observación; por otro lado, las agregaciones consideradas impactan sobre la medida global a través de sus componentes intra e interregional (Rey, 2004).

Para México también se han realizado estudios sobre disparidades regionales. Por ejemplo, Esquivel (1999) utiliza un esquema de siete agrupaciones de los estados de México con información de 1940 a 1995. El autor encuentra que los estados pertenecientes a las regiones Norte, Pacífico, Golfo y Capital tendieron a crecer más rápido que los estados de las zonas Sur, Centro y Centro-Norte. Así, sugiere que estados con un nivel similar de ingreso inicial presentarán diferentes tasas de convergencia, ésto debido a la magnitud de las disparidades regionales existentes.

Hanson (1998) a partir de un esquema de cinco agregaciones concluye que la liberación del comercio tuvo considerables efectos sobre las decisiones de localización de las industrias en el país. Sus resultados para el período 1980 a 1993 muestran que el crecimiento del empleo fue mayor en las regiones cercanas a la frontera con Estados Unidos, además de una considerable disparidad en los niveles de ingreso de los trabajadores calificados que se encuentran en la frontera respecto de los otros. Posteriormente Hanson (2003) propone un diseño de agregación de seis regiones, en el que encuentra nuevamente fuertes disparidades regionales. Los estados cercanos a la frontera tuvieron incrementos en los salarios, lo que llevó a una mayor desigualdad salarial entre los trabajadores más y menos capacitados.

El estudio de Bouillon *et al.* (2003) considera ocho regiones para analizar información del año 1984 a 1994. Analizan los efectos regionales de las desigualdades en los

ingresos, encontrando que son las dinámicas del sur las que mayor participación tienen sobre la disparidades. También señalan que las áreas rurales, sobre todo de las zonas centro y sur, fueron las más deterioradas durante el período de estudio. Con datos de 1970 a 2001, Chiquiar (2005) forma cinco regiones con los estados del país (aunque elimina dos estados debido a los excedentes petroleros) y refrenda la importancia de considerar regiones más que países como unidades de estudio. El autor observa fuertes disparidades regionales en la convergencia de los ingresos entre norte y sur; además de las diferencias en los niveles de capital humano e industrial y las mejores infraestructuras de comunicación y transporte que se encuentran en la frontera.

Por su parte Aroca *et al.* (2005), aplicando técnicas espaciales, buscan la existencia de patrones geográficos en el procesos de convergencia en México: hay evidencia de un mayor desempeño de los estados fronterizos; además, mientras que el sur si puede ser identificado, no ocurre lo mismo para el centro. Lo anterior contradice la suposición usual de un gradiente que irradia desde el centro, o la que supone un gradiente que surge del norte hacia el sur, ya que el “norte” parece limitarse a los estados de la frontera con EU.

Asuad *et al.* (2007) se proponen investigar las disparidades del ingreso agrícola con un enfoque de convergencia espacial, realizan una agregación que se adecua más a las características necesarias para su estudio. Asuad *et al.* (2007) observan divergencia entre las 4 regiones definidas, pero una convergencia del ingreso entre las regiones más pobres y entre las más ricas, aunque empíricamente no encuentran evidencia de dependencia espacial, es decir, existe un patrón aleatorio y no una aglomeración considerando el ingreso agrícola.

Un trabajo pionero en la aplicación de métodos de agregación espacial para México, es el de Rey y Sastré-Gutiérrez (2010). En este trabajo se brinda importancia al papel del esquema territorial bajo el cual se realiza la medición de las disparidades. Su principal resultado, tras la aplicación de métodos de agregación espacial, es que

la desigualdad no ha presentado reducciones considerables desde 1960. Asimismo, la aportación a la desigualdad global atribuible al componente interregional es sensible al esquema de regionalización que se seleccione, ya que puede haber diferencias estadísticamente significativas en las magnitudes de la medida de desigualdad.

Las agregaciones de unidades geográficas han sido utilizadas para efectuar análisis de disparidades regionales. Por ello, que los esquemas de agregación y los criterios o métodos detrás de ellos sean relevantes. En este sentido, la presente investigación puede proporcionar información útil sobre las diferencias regionales al aplicar métodos de agregación espacial, y al ofrecer agregaciones bajo conceptos particulares de áreas rezagadas y otras con mayor desarrollo. El ordenamiento territorial tiene entre sus objetivos la reducción (y eventual eliminación) de estos desequilibrios. En la siguiente sección se presenta una introducción al tema.

## 1.2 LA PERSPECTIVA DEL ORDENAMIENTO TERRITORIAL Y LA IMPORTANCIA DE IDENTIFICAR CARACTERÍSTICAS REGIONALES

El ordenamiento territorial (OT) tiene distintas acepciones, algunas de las cuales son abordadas por Massiris (2002). Del conjunto de definiciones, el autor resalta la comprensión del ordenamiento como disciplina, técnica y política que requiere de la interdisciplinariedad, ya que eso permite encontrar criterios para enfrentar los retos que plantea la organización óptima del espacio.

El OT es un concepto que se expresa a través de la planificación, ya que ésta es la definición de objetivos y su aplicación a través de planes y estrategias (Bustos,

1998). Para Delgadillo y Torres (2006), existe además el campo de la política regional, el cual busca corregir los desequilibrios interterritoriales atendiendo tanto a regiones atrasadas como a aquellas con mayor nivel de desarrollo. Así, el ordenamiento pudiera ser entendido como la planificación física de las regiones.

Montes (2001) hace una recomendación para los países latinoamericanos, y es que la política de ordenamiento debe conjugar dos enfoques. El primero es el de integrar las políticas sectoriales a los diferentes niveles regionales considerando sus interacciones; el segundo aspecto debe ser la atención a las distintas formas de organización espacial y política de la región. Por ejemplo, en México las políticas han tenido un enfoque sectorial más que regional (Delgadillo y Torres, 2006).

A partir del análisis de experiencias internacionales de éxito, Ferreira (2005) muestra que las políticas regionales en México debieran adecuarse a los nuevos contextos nacionales e internacionales. Este autor, al igual que Wong-González (2010) destacan que en el país no se cuenta con una visión articulada del desarrollo regional, lo que ha provocado desfases y pérdida de oportunidades para ciertos territorios. Resultando claro que el OT ha quedado rezagado como elemento de política regional, ya sea porque ésta no se han adecuado a las necesidades y problemas actuales, o porque ha sido implementada pero carece de continuidad en los planes de desarrollo.

Pujadas y Font (1998) remarcan que las disparidades interterritoriales tienen causas desencadenantes distintas y que por ello, requieren de instrumentos de actuación diferentes. Dentro de estas disparidades señala las del tipo “campo-ciudad”, en las que no se tiene como objetivo la convergencia sino el correcto funcionamiento de sus interrelaciones. Por lo que la planificación tiene un papel significativo, pues debe buscar la mejora de accesibilidad, de calidad de vida, mayores oportunidades económicas. Incluso en el análisis del crecimiento urbano y la formación de áreas metropolitanas se necesita de la planificación territorial, pues se debe recurrir a instrumentos que escapan a la planificación urbana convencional.

---

Un dilema usual es cómo definir las áreas rurales y las áreas urbanas. Romero y Vásquez (2005) mencionan que en ocasiones se considera al territorio como el suelo ocupado por las ciudades, y las áreas rurales son espacios residuales y un soporte para la expansión urbana, sin reconocer que el medio físico influye en las aptitudes territoriales. Uno de los problemas del espacio rural es su considerable heterogeneidad (Ashley y Maxwell, 2001), mientras que el urbano enfrenta una creciente complejidad ante la gran variedad de usos de suelo (Pujadas y Font, 1998).

En el capítulo siguiente, a partir de las clasificaciones usuales de áreas rurales y zonas metropolitanas se analizan las variables que mayor interés han generado en estudios previos. Se describen también algunas problemáticas identificadas en estudios previos que se relacionan con la elección de las definiciones a utilizar en una investigación, lo cual tiene que ver con los temas que se abordarán en el Marco Teórico y Conceptual.

## CAPÍTULO 2

# LAS CLASIFICACIONES DE ENTORNOS REGIONALES DESDE LA PERSPECTIVA INSTITUCIONAL

---

En las siguientes secciones se presentan las definiciones más utilizadas para caracterizar una área como rural o metropolitana. Cada una de estas zonas ha sido objeto de diferentes líneas de investigación, por lo que se presenta una breve reseña de estudios que tuvieron como fin el analizar con más detalle determinadas características y su comportamiento en estos entornos. El objetivo es presentar una descripción de las variables que han sido objeto de estudio en análisis previos, para posteriormente definir aquellas a utilizar con el algoritmo *MaxP* en esta investigación.

### 2.1 DEFINICIÓN DE ÁREAS RURALES

La Organización para la Cooperación y el Desarrollo Económicos (OCDE) sugiere que el término rural no tiene una definición precisa, pues la frontera entre lo rural y lo urbano no es fácil de identificar en la mayoría de los países (OCDE, 2007). Los criterios para describir áreas rurales son usualmente relacionados con un bajo nivel de población en una localidad; aunque en cada país pueden tenerse diferentes definiciones institucionales dependiendo del objetivo para el cual se requiera considerar estas áreas. Usualmente lo rural se relaciona con lugares donde la mayoría de la población se dedica a la agricultura; altos costos de transacción asociados a las grandes distancias y pobre infraestructura; entre otras características (Ashley y Maxwell, 2001).

Coburn *et al.* (2007) describen algunas de las definiciones usuales de localidades rurales en EU. Los autores recalcan la importancia de la elección de los conceptos que

mejor se adecuen a los objetivos de la investigación; pues incrementos en la población o cambios en los límites geográficos pueden conducir a una aplicación errónea de programas públicos. Hart *et al.* (2005) también sugieren que antes de iniciar una investigación, se debe familiarizar con diversas definiciones de “rural” con el fin de elegir la que mejor se ajuste a las necesidades del estudio, y así evitar problemas de escala.

Por su parte, du Plessis *et al.* (2002) empleando seis clasificaciones institucionales con datos de Canadá exhiben problemáticas surgidas a partir de la definición seleccionada. En primer lugar describen que el tamaño de la población rural depende de la definición elegida, efecto que se reproduce incluso dentro de la misma provincia/territorio. Por otro lado, diferentes definiciones clasifican diferentes grupos de población como rural; en ocasiones una misma población puede pertenecer a dos categorías distintas (es decir, población definida como rural por el censo puede vivir en áreas predominantemente urbanas según la de OCDE). El efecto anterior afecta los resultados al considerar características socioeconómicas de la población (siguiendo el ejemplo, se tendría que la población tiene en promedio un ingreso mayor comparado con otros grupos “rurales”).

En OCDE (2006) se describen los siguientes criterios para determinar si una región se debe considerar rural, urbana o intermedia:

- Densidad poblacional. Una comunidad se define como rural si la densidad poblacional es menor a 150 habitantes por  $km^2$  (500 habitantes para el caso de Japón).
- Regiones por porcentaje de población en comunidades rurales. Una comunidad se cataloga como “predominantemente rural” si más del 50% de su población vive en comunidades rurales; “predominantemente urbana” si menos del 15% de la población vive en comunidades rurales; e intermedias si parte de la población que vive en comunidades rurales está entre 15% y 50%.

- Una región que se podría clasificar como rural con base en la regla general, se clasificará como “intermedia” si cuenta con un centro urbano con más de 200,000 habitantes (500,000 para Japón), lo que representa no menos del 25 % de la población regional. Una región que se pudiera clasificar como intermedia, se clasificará como predominantemente urbana si cuenta con un centro urbano de más de 500,000 habitantes (1'000,000 para Japón), representando no menos del 25 % de la población regional.

De acuerdo a la definición de la OCDE, más del 75 % del territorio total de los países miembros se define como predominantemente rural, y alrededor del 25 % de la población total vive en dichas áreas rurales (OCDE, 2006). Se utilizan criterios relacionados con la baja densidad poblacional, y lo remoto de algunas comunidades para delimitar áreas rurales, bajo el supuesto de que ello puede generar un conjunto de problemas que impactan sobre el bienestar y un incremento en los niveles de disparidad entre regiones. Por ejemplo, en México las disparidades en el PIB, población y territorio evidencian la existencia de obstáculos para la integración de estas áreas en la economía nacional (OCDE, 2007).

Para el caso de México existen distintas definiciones de rural dependiendo de la instancia que busca delimitar este tipo de áreas (OCDE, 2007):

- El Instituto Nacional de Estadística y Geografía (INEGI) distingue entre áreas urbanas y rurales. Define en primera instancia a las urbanas y el resto son consideradas como rurales. Todas las localidades de más de 2,500 habitantes y las “cabeceras municipales”, independientemente del número de habitantes, se consideran urbanas.
- La Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación (SAGARPA) define como rurales a aquellas localidades con menos de 2,500 habitantes o donde la agricultura represente más del 50 % de la producción local.

- La Secretaría de Desarrollo Social (SEDESOL) considera una localidad rural cualquier asentamiento humano, con por lo menos tres casas y un máximo de 2,500 habitantes.

Los umbrales de población para diferenciar entre rural y urbano varían incluso para una misma institución. Por ejemplo, en SEDESOL (2010) se utiliza la definición que brinda el Consejo Nacional de Evaluación de la Política de Desarrollo Social (CONEVAL), donde una localidad se especifica como rural si en ella habitan 15,000 habitantes o menos. Mientras que en SEDESOL (2007) se recurre a la definición de INEGI.

Siguiendo a OCDE (2007) de los 2,454 municipios registrados en 2005, 991 se considerarían como rurales siguiendo la definición de localidades con menos de 2,500 habitantes (y 1,978 si se consideran localidades de 15,000 habitantes o menos), 354 como urbanos intermedios, y 122 como urbanos en un sentido más estricto. En la Figura 2.1 se presenta el mapa con dicha información.

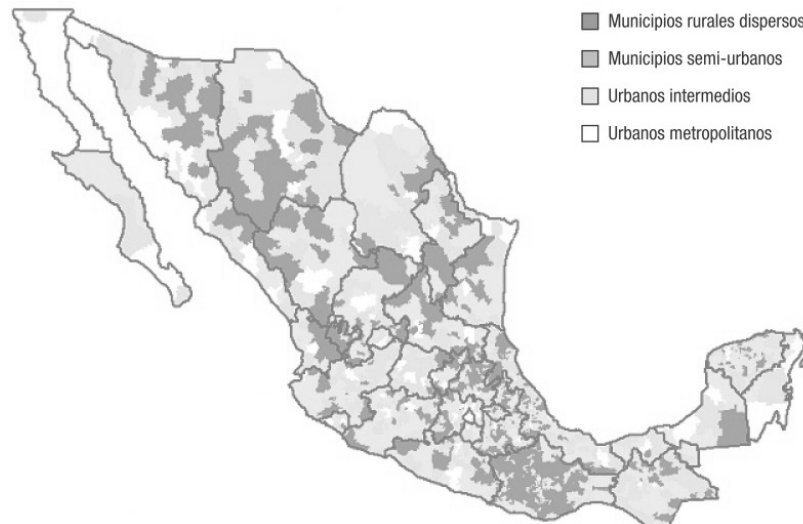


Figura 2.1: Municipios de México bajo la clasificación de la OCDE, 2005.

Fuente: OCDE (2007).

Puede distinguirse de lo anterior que en México no se cuenta con una variedad

de definiciones, ya que las existentes están en función de la brindada por INEGI. Si bien ello permite una mayor comparabilidad entre resultados de investigaciones, a la vez reduce la posibilidad de utilizar otras clasificaciones que se adapten mejor a ciertas características de interés de las áreas rurales.

### 2.1.1 ESTUDIOS REALIZADOS PARA ÁREAS RURALES

Dentro de la literatura que se desarrolla en torno a las áreas rurales se pueden distinguir distintos enfoques de investigación. Una de las perspectivas ha sido considerar las diferencias intrerregionales. Aunque de Poel *et al.* (2009) analiza seis países africanos, concluye que son las diferencias en características de las viviendas, más que las de comunidad, las que tienen un mayor efecto sobre la mortalidad infantil. Por lo que recomiendan que uno de los objetivos de los programas públicos sea el de la mejora de los materiales de las viviendas, sobre todo en las comunidades rurales.

Por su parte, Spellerberg *et al.* (2007) encuentran una fuerte divergencia entre las regiones de Alemania, sobre todo en lo que se refiere a los distritos rurales. Observan que las áreas rurales del Este tienen un mayor número de dificultades, incluso considerando otros países de la Unión Europea. Fallah y Partridge (2007) examinan la relación entre la desigualdad y el crecimiento económico, sugiriendo que los resultados contradictorios en investigaciones empíricas se deben a las disparidades regionales. Los autores encuentran que la desigualdad en los ingresos tiene un impacto negativo significativamente mayor en el crecimiento de los condados rurales. Por ello invitan a considerar la importancia de la escala a seleccionar en los estudios, pues si se presenta variación a nivel rural-urbano con mayor razón existen diferencias en niveles territoriales más agregados.

China ha estudiado ampliamente las brechas rural-urbano. Un ejemplo es el trabajo de Zimmer *et al.* (2007) acerca de la mortalidad en adultos mayores en áreas

rurales, la cual es 30 % mayor que en áreas urbanas. Esto se explica parcialmente por la escasa inversión en infraestructura en zonas en desventaja, lo que conlleva a una reducción en los servicios de salud. Entre las características que explican las diferencias en este índice se encuentran las referentes a la población, como educación o niveles de ingreso, y las amenidades de las comunidades urbanas. Y es que como explican Lin y Chen (2011), las políticas enfocadas en el desarrollo industrial son las que tienen mayor efecto sobre este tipo de brecha. Aunque solo exponen de manera teórica algunas proposiciones, logran obtener interesantes conclusiones acerca de la necesidad de mejorar las políticas impuestas para así reducir las brechas.

Las disparidades en los ingresos han sido analizadas en diversos países. Pradhan *et al.* (2000) realiza una revisión más exhaustiva para el caso de India. Los autores indica que la proporción de los ingresos en zonas rurales se redujo considerablemente en el tiempo, lo cual condujo a la ampliación de las disparidades económicas y sociales. La brecha en China también es significativa como revela Sicular *et al.* (2007), subrayando que el lugar de residencia sigue siendo un factor explicativo, existiendo grandes diferencias regionales en esta brecha, sobre todo al Oeste.

Bergolo y Carbajal (2010) explican que para Uruguay las disparidades regionales son el segundo factor explicativo, tras la educación, de las diferencias en los niveles de ingresos. Destacan que existe una mayor dispersión de ingresos entre los trabajadores urbanos; mientras que las diferencias en los rendimientos de trabajadores con atributos similares fueron mayores en el entorno rural. Con el fin de reducir las brechas en los ingresos a nivel país, aconsejan el diseño de políticas diferenciadas para los entornos rural y urbano.

Como se explicó anteriormente, es usual relacionar lo rural con la agricultura, sin embargo una línea de investigación ha sido el incremento en los ingresos rurales que no provienen de actividades agrícolas. Reardon *et al.* (2001) explora datos para América Latina, encontrando que el empleo no agrícola es clave para las viviendas rurales aun-

que la agricultura siga siendo la principal fuente. Destacan que los ingresos salariales no agrícolas exceden a los agrícolas, y además los ingresos en el sector servicios son mayores a los manufactureros. Finalmente hacen una serie de recomendaciones de política pública, planteando que es necesario eliminar el sesgo hacia el sector agrícola de las políticas rurales; asimismo la promoción de empleos no agrícolas, pues ello beneficiaría de manera indirecta la agricultura.

Para México, de Janvry y Sadoulet (2001) elaboraron un estudio similar. De éste se concluyó que las actividades no relacionadas a la agricultura juegan un papel importante en los ingresos de los ejidatarios, ya que más de la mitad del ingreso promedio proviene del empleo por cuenta propia, salarios no agrícolas y de las remesas. Los autores sugieren que impulsar la educación es crucial, pues al comparar grupos étnicos surgen diferencias aún más significativas en los niveles de ingreso. Esta tendencia persiste en el tiempo y a nivel mundial. Davis *et al.* (2010) analizan una muestra de 15 países, resultando que son las comunidades rurales más pobres las que tienen mayor dependencia de la agricultura, mientras que aquellas que cuentan con mejores condiciones se caracterizan por la diversificación en empleos no agrícolas y el empleo por cuenta propia. Concluyen que son estos recursos no agrícolas los que se asocian a la creciente desigualdad en los ingresos.

Otra línea de investigación usual ha sido lo referente a las definiciones de lo que es “rural”. En ese sentido se han desarrollado algunos índices, como el propuesto para México por Zamudio *et al.* (2008). En el que se intenta examinar la relación entre el grado de ruralidad con el índice de marginación, el índice de desarrollo humano y el concepto de ruralidad de INEGI. Entre sus resultados destacan que lo rural no sólo se relaciona con las actividades agrícolas, pues en algunos casos su cercanía a centros urbanos les ha permitido entrar en actividades secundarias y terciarias; así como tampoco es conveniente asociarlo solamente a su nivel poblacional.

Algunos estudios buscan mostrar las problemáticas que surgen de considerar dis-

tintas definiciones de rural. Por ejemplo en Cossman *et al.* (2008) se sugiere que las delimitaciones usuales de comunidades rurales no son adecuadas para medir efectos sobre la salud. Señalan que dichas desagregaciones alteran las direcciones de los efectos, complicando la interpretación de los resultados. Al realizar pruebas empíricas con diversas clasificaciones (esquemas de 2, 5 y 10 categorías) para analizar índices de mortalidad, demuestran que los resultados varían según las delimitaciones utilizadas, lo que provoca errores de interpretación (sobre todo al aumentar el número de clases). Resaltan también que el sesgo en los resultados puede llevar a que los programas políticos no tengan el efecto deseado, dado el error resultante de utilizar agregaciones no adecuadas de las áreas objetivo.

En una investigación similar, West *et al.* (2010) plantean que las implicaciones de políticas sobre el servicio que presta la *Veterans Health Administration* (VHA) varían con la categorización geográfica utilizada. Lo anterior porque la clasificación VHA enrola a más individuos como rurales que las otras dos definiciones que analizan, lo que tiene efectos directos sobre las políticas que requieren ser aplicadas. Por ejemplo la elección de localización de clínicas, de manera que se reciba a la mayor cantidad posible de derechohabientes sin que tengan que viajar grandes distancias.

Un tema recurrente en la literatura revisada es la importancia de las características con que se busca definir a una comunidad como rural. Como lo muestran los dos últimos trabajos comentados, la selección de una definición tiene fuertes implicaciones de política pública. La aplicación de técnicas espaciales puede brindar nueva información sobre estas áreas en desventaja, de manera que pueda darse una mayor comprensión de sus procesos internos y así lograr un mayor aprovechamiento de los recursos. Resulta interesante que los efectos espaciales parecen tener un fuerte impacto sobre los resultados en las investigaciones del entorno rural, por lo que la presente investigación puede arrojar luz al respecto.

## 2.2 DEFINICIÓN DE ZONAS METROPOLITANAS

Los centros urbanos han recibido mayor atención por parte de los investigadores, y no solamente desde el punto de vista institucional. Patacchini y Zenou (2009) describen que en Estados Unidos actualmente se asocia a la ciudad con las zonas metropolitanas, mientras que en Europa aún se utilizan los conceptos administrativos. Sin embargo, al igual que las áreas rurales no existe una definición única para identificar las zonas urbanas, variando en cada país.

El concepto de Zona Metropolitana (ZM) se acuñó en EU partir de la década de 1920. Es una forma de definir a la ciudad cuando ésta rebasa los límites de su unidad político-administrativa original. Negrete y Salazar (1986) describen que en una primera etapa de la formación de una ciudad, la población y la actividad económica tienden a concentrarse en el centro de la ciudad; en la segunda fase se da la expansión física de la ciudad y su radio de influencia, generando la suburbanización. Si dicho proceso asocia una o más unidades administrativas, aparece una ZM. En Canadá un conjunto de condados que forman un centro que se conecta a otras áreas urbanas se califica como un área metropolitana (Puderer, 2008), mientras que en Europa la definición encaja más en el concepto de región urbana funcional, esto se debe en parte a las diferencias en los sistemas de nomenclaturas estadísticas (Patacchini y Zenou, 2009).

Luo *et al.* (2008) hacen énfasis en que a pesar de no existir consenso sobre la definición de áreas metropolitanas, éstas no deben ser delimitadas de manera vaga. Dado que la estructura de estas áreas y la importancia relativa de sus componentes deben ser la base para la planeación de sistemas urbanos. Las zonas metropolitanas generalmente son vistas como potenciales fuentes de desarrollo, y como motores de crecimiento hacia sus áreas vecinas, por lo que en México también se han propuesto definiciones que buscan agrupar aquellos municipios con mayor integración económica y en términos de empleo.

Año	Número de ZM
1940	5
1950	7
1960	12
1970	27
1980	26
1990	37
2000	55
2005	56

Tabla 2.1: Evolución del número de Zonas Metropolitanas en México.

Fuente: Elaboración propia con base en Zentella (2007).

En la Tabla 2.1 se muestra la evolución del número de Zonas Metropolitanas en México a partir de 1940, evidenciando el fuerte crecimiento de ciudades. Según la comisión interinstitucional SEDESOL-INEGI-CONAPO éstas zonas representan una oportunidad de propagar el desarrollo social y económico debido a que se desempeñan como centros de actividad económica a nivel regional.

#### DEFINICIÓN Y CRITERIOS

En 2003, SEDESOL-INEGI-CONAPO buscaron formular una definición de las zonas metropolitanas del país. Realizaron en 2004 una primera delimitación, misma que apoyándose en el II Censo de Población y Vivienda 2005 fue actualizada en el año 2007. A continuación se presentan los criterios utilizados en la última delimitación ofrecida por este grupo en el documento del año 2007.

Para delimitar zonas metropolitanas se definieron tres grupos de municipios metropolitanos, cada grupo con sus respectivos criterios.

1. **MUNICIPIOS CENTRALES.** Municipios donde se encuentra la ciudad principal, los cuales se identificaron a partir de las siguientes características:

Unión física entre dos o más localidades censales de diferentes municipios y cuya

población en conjunto asciende a 50 mil o más habitantes; municipios con localidades de 50 mil o más habitantes que muestren un alto grado de integración física y funcional con municipios vecinos predominantemente urbanos; municipios con ciudades de un millón o más habitantes; o municipios con ciudades que forman parte de una zona metropolitana transfronteriza, con 250 mil o más habitantes.

2. MUNICIPIOS EXTERIORES DEFINIDOS CON BASE EN CRITERIOS ESTADÍSTICOS Y GEOGRÁFICOS. Municipios contiguos a los anteriores, cuyas localidades no están conurbadas a la ciudad principal, de carácter predominantemente urbano, y que mantienen un alto grado de integración funcional con los municipios centrales, lo que se determina con el cumplimiento de cada una de las siguientes condiciones: La localidad principal está ubicada a no más de 10 km por carretera pavimentada y de doble carril, de la localidad o conurbación que dio origen a la zona metropolitana; al menos 15 % de su población ocupada residente trabaja en los municipios centrales de la ZM, o bien, 10 % o más de la población que trabaja en el municipio reside en los municipios centrales de esta última; tienen un porcentaje de población económicamente activa ocupada en actividades industriales, comerciales y de servicios mayor o igual a 75 %; y tienen una densidad media urbana de por lo menos 20 habitantes por hectárea.

3. MUNICIPIOS EXTERIORES DEFINIDOS CON BASE EN CRITERIOS DE PLANEACIÓN Y POLÍTICA URBANA. Municipios que se encuentran reconocidos por los gobiernos federal y locales como parte de una ZM, a través de una serie de instrumentos que regulan su desarrollo urbano y la ordenación de su territorio, independientemente del cumplimiento de los criterios antes señalados. Estos municipios cumplen al menos una de las siguientes condiciones:

Estar incluidos en la declaratoria de zona conurbada o zona metropolitana correspondiente; estar considerados en el programa de ordenación de zona conurbada o zona metropolitana respectivo; o estar reconocidos en el Plan Nacional de Desa-

rollo Urbano y Ordenación del Territorio vigente.



Figura 2.2: Zonas Metropolitanas de México, 2005.

Fuente: SEDESOL, INEGI y CONAPO (2007).

## UNIVERSO DE ZONAS METROPOLITANAS

Se definieron un total de 251 unidades como municipios centrales dentro de las 56 zonas metropolitanas. En base a los criterios de distancia, integración funcional y carácter urbano se incorporaron 25 municipios; mientras que 69 se agregaron bajo los criterios de planeación y política urbana.

El documento aclara que algunos municipios fueron catalogados dentro de las zonas metropolitanas siguiendo otros criterios: algunos municipios fueron integrados a una ZM solo por su cercanía, aunque su mayor localidad en realidad estaba más integrada a una área conurbada distinta; otro caso es el de un municipio definido como exterior a pesar de no cumplir con el porcentaje mínimo de población ocupada. Además, se presentaron cambios: tres zonas cambiaron de definición, por ejemplo una de ellas

pasó de Conurbación intermunicipal a definida por Distancia. Hubo dos casos de fusión de zonas: la ZM de Texmelucan se anexó a la de Puebla-Tlaxcala y se fusionaron la de Apizaco y Tlaxcala en una sola. Finalmente, se crearon tres zonas, una de ellas únicamente por criterio de Política Urbana.

De estas últimas observaciones resulta evidente que las zonas metropolitanas del país atravesaron por considerables cambios en un corto período de tiempo. Además es interesante que municipios centrales que ya estaban definidos dentro de una ZM se agregaran a otra, como en el caso de las zonas fusionadas. De manera similar debe analizarse el caso de los municipios agregados mediante criterios distintos a los originalmente establecidos.

Dichas agregaciones se realizaron considerando que existe una importante integración física y funcional entre sus municipios, sin embargo, no se realizaron otras pruebas para corroborar tales afirmaciones. La aplicación de métodos espaciales puede arrojar luz sobre las debilidades antes mencionadas, y brindar información útil sobre el comportamiento de las agrupaciones sujetas a las variables en las que ellos basan sus definiciones.

### 2.2.1 ESTUDIOS REALIZADOS PARA ZONAS METROPOLITANAS

El análisis de patrones espaciales va tomando importancia en los estudios regional-urbanos y más específicamente de las Zonas Metropolitanas (ZM). Sobre todo porque estas áreas tienden a ser consideradas centros de desarrollo, por lo que sus dinámicas internas sean de interés. Ejemplo de ello son los estudios sobre dispersión del empleo, o sobre el comportamiento de algunas variables de interés en estas zonas. Debido a su papel en el crecimiento, las pautas que se den al interior de las áreas se vuelven determinantes en el desarrollo de las regiones (Glaeser, 2011).

Para las ZM se han realizado estudios sobre concentración y dispersión del em-

pleo. Lee (2007) encontró evidencia de una continua descentralización de los centros metropolitanos a los suburbios y la dispersión del empleo fue más común que la formación de subcentros en seis ZM de Estados Unidos. Los resultados muestran que, la evolución metropolitana de pasar de una forma monocéntrica a una policéntrica, y entonces a una estructura dispersa no es un proceso lineal. Por su parte, Garcia-López y Muñiz (2010) también estudiaron los patrones del empleo para la Región Metropolitana de Barcelona, sugiriendo que el empleo se ha descentralizado y desconcentrado como lo indica el Índice de Gini, a la vez que el número de subcentros se ha incrementado y son más intensos que en el pasado.

Para el caso de México, Suárez y Delgado (2009) intentaron responder si la Zona Metropolitana de la Ciudad de México (ZMCM) es una metrópolis policéntrica. Los resultados mostraron una gran aglomeración central que se alarga de norte a sur y que se convierte en tres formas de corredor. Considerando el empleo formal se encontraron ocho subcentros, cuatro de ellos muy cercanos a la aglomeración central y cuatro a grandes distancias de ésta. Los autores plantean que este caso es una forma urbana híbrida pues el centro se encuentra disperso y existen subcentros locales, sin embargo, aún se puede definir como monocéntrica.

También se ha analizado el comportamiento de una sola variable en determinadas zonas. Por ejemplo, Montes y Ortega (2003) elaboraron un análisis del índice de marginación para la Zona Metropolitana de Monterrey. Los autores determinan que en la década de los noventa se incrementó el número de personas que viven en condiciones de alta marginación; lo que se debe en parte a la disminución de población de la clase media. Macías *et al.* (2009) elaboraron un estudio similar para la ZM de Guadalajara; comparando los años 2000 y 2005 los resultados exhibieron un descenso general del nivel de vida. Los resultados muestran que el índice pasó de ser mayormente bajo a medio; es decir, la clase media aumentó y la media alta se redujo.

Algunos otros estudios consideran un conjunto de variables para una Zona Me-

tropolitana. Pasternak y Machado (2008) se propusieron examinar si existe un patrón de segmentación social en el espacio intrametropolitano de São Paulo, encontrando que a pesar del patrón periférico tradicional, de centro rico y periferia pobre, se dieron cambios al interior del centro. Rubalcava y Schteingart (1985) estudiaron la ZMCM, concluyendo que los principales factores que explican la diferenciación intraurbana se mantienen en fuerza y sentido a través del tiempo. Además se verificó que algunos de los factores explicativos seguían un esquema concéntrico, mientras que otros seguían el esquema de sectores; por lo que describen que se dio un proceso de consolidación diferenciado: más heterogéneo en zonas centrales y más homogéneo en zonas de la periferia.

Más recientemente, Pérez y Santos (2011) examinaron la existencia de segregación socioespacial en la ZMCM. Recurrieron al uso del Índice de Morán y del Indicador Local de Asociación Espacial para conocer si existe relación espacial entre los factores explicativos. Concluyen que existe un efecto debido a la posición que ocupa cada AGEB en la ZM y se puede afirmar que existe un patrón de segregación espacial claro. Dicho patrón es la división de la zona en tres grandes sectores: el primero va del sur al poniente y una parte del norponiente con los mayores niveles educativos y de consolidación de la vivienda; el segundo es un anillo intermedio con una mezcla de condiciones sociales; y la periferia formada por los municipios limítrofes y conurbados al D.F. con los mayores rezagos sociales.

Otra vertiente es la de analizar un conjunto de variables para varias zonas metropolitanas. Rey *et al.* (2011) propusieron una nueva aproximación para medir el cambio en las Zonas Metropolitanas de Estados Unidos. Examinaron a los *neighborhoods* como unidades con una composición socioeconómica y una impresión geográfica que se va modificando con el tiempo, para ello seleccionaron 22 variables censales y se apoyaron en métodos de agregación espacial. Lo más interesante fue que los resultados mostraron

que el cambio espacial fue lo más común, pues dos tercios de los *census tracts* experimentaron alguna reconfiguración en lo que constituía sus *neighborhoods* entre 1990 y 2000; además aquellos *neighborhoods* que en 1990 tenían una alta densidad y estaban localizados más centralmente en las ZM sufrieron grandes cambios espaciales para el 2000, a diferencia de aquellos de menor densidad y más periféricos.

Ariza y Solís (2009) examinan las Zonas Metropolitanas de Monterrey, Guadalajara y de la Ciudad de México, con el fin de constatar si ha habido cambios en la segregación espacial.<sup>1</sup> Los autores concluyeron que el ingreso individual y familiar mostró un incremento considerable de la segregación; mientras que la ocupación y la escolaridad evidenciaron que los individuos que carecen de formación escolar superior se encuentran más aislados espacialmente de los que sí la poseen. Del estudio resulta evidente que los cambios sufridos por las tres ZM no han reducido las distancias sociales, llevando incluso a una creciente polarización en cuanto a ingresos.

En este capítulo no se desarrolla alguna discusión sobre la forma de delimitar las zonas metropolitanas o áreas rurales. La mayoría de las investigaciones consideran las definiciones institucionales y a partir de ellas realizan los análisis. Se debe destacar el creciente interés por determinar la existencia de patrones espaciales en estas zonas. Sin embargo poco se ha abordado con herramientas y técnicas espaciales para el caso de México, por tanto la actual investigación busca brindar mayor información en este sentido, y adherirse a un reciente cuerpo de literatura sobre el tema (Rey *et al.* (2011), Rey y Sastré-Gutiérrez (2010)).

---

<sup>1</sup>Se entiende por segregación espacial el fenómeno de la ocupación del espacio urbano por distintos grupos sociales, que tienden a agruparse de acuerdo a características comunes, por ejemplo los *ghettos* (Machado, 2001).

## CAPÍTULO 3

### MARCO TEÓRICO Y CONCEPTUAL

---

En este capítulo se aborda el marco teórico que da sustento a la necesidad de utilizar técnicas explícitamente espaciales. El problema de agregación conduce al Problema de Unidad de Área Modificable que se refiere a la sensibilidad de los resultados ante la consideración de distintos niveles de agregación (Openshaw, 1984). Cuanto más homogénea es la agrupación se reduce el riesgo de incurrir en la falacia ecológica, entendiendo ésta como la inferencia errónea sobre el comportamiento de las variables a nivel individual a partir de los resultados de las agregaciones (Freedman, 2001).

### 3.1 PROBLEMA DE UNIDAD DE ÁREA MODIFICABLE Y LA FALACIA ECOLÓGICA

Siguiendo a Openshaw (1984), a nivel individual no existe autocorrelación espacial; así, el grado de ésta depende de las unidades de área consideradas. Además, sugiere que el nivel espacial a utilizar depende en parte de la investigación; por ello que no existan reglas generales para realizar el agregado. En este sentido las unidades son arbitrarias y modificables. Por lo anterior, las unidades territoriales pueden ser agregadas jerárquicamente en diferentes conjuntos de unidades espaciales más grandes (Gauvin *et al.*, 2007); por ejemplo, un conjunto de municipios en un estado, o estos municipios dentro de unidades más grandes que contengan un nivel similar en términos de empleo.

El Problema de Unidad de Área Modificable (PUAM) se relaciona con el hecho de que medidas para datos transversales son sensibles a los niveles de agregación y a

las combinaciones de unidades contiguas que se realizan (Anselin, 1988). Ejemplo de lo anterior es el caso del estudio de Snow (Brody *et al.*, 2000) ; si Snow en lugar de utilizar puntos hubiera usado áreas de densidad, quizá no hubiese sido capaz de identificar los clusters de cólera alrededor de la bomba de agua contaminada (Gauvin *et al.*, 2007). Rey y Folch (2011) utilizando diversos índices de segregación espacial y una serie de simulaciones encuentran, entre otros resultados, que dichos índices se afectan ante la consideración de diversas escalas de agregación.

El PUAM tiene dos componentes distintos pero relacionados. De acuerdo con Openshaw (1984), el primero de ellos es el problema de escala; éste se refiere a las variaciones producidas en los resultados cuando se consideran agregaciones de unidades espaciales dentro de otras más grandes. Mientras que el segundo problema es el de agregación, éste se da cuando al utilizar diversas formaciones de unidades de tamaño similar se generan diferentes resultados; es decir, para un número dado de zonas se utilizan diversas alternativas de agrupación de las mismas.

El PUAM es un problema habitual cuando se consideran datos espacialmente agregados, y resulta claro que los resultados de los estudios dependerán de las áreas que están siendo analizadas (Openshaw, 1984). Este problema es conocido en econometría como el problema de micro-macro agregación, y en sociología como el problema de la falacia ecológica (Anselin, 1988). Si a partir de la información disponible a nivel agregado se infiere sobre conductas individuales, se dice que es una inferencia ecológica; por tanto, la falacia ecológica consiste en pensar que las relaciones observadas para grupos se mantienen para los individuos (Freedman, 2001). Cuando una agrupación de unidades es homogénea en su interior, no existirá este problema; sin embargo eso no ocurre en la mayoría de los casos (Openshaw, 1984).

Robinson (1950) analiza la diferencia entre las correlaciones ecológicas y las individuales. El autor utilizando información de 48 estados de Estados Unidos para 1930, encuentra que la autocorrelación entre haber nacido en la misma ciudad y ser analfabe-

ta es de -0.526, mientras que la correlación a nivel individual es de 0.118. Este caso es una muestra de cómo puede afectar la escala sobre la inferencia, al reducirse en número de observaciones se incrementó el valor absoluto de la correlación. El autor también demuestra matemáticamente la correspondencia entre los coeficientes de correlación de nivel individual y agregado.

Morgenstern (1982) desglosa la falacia ecológica en dos componentes: sesgo de agregación y sesgo de especificación. El primero, sugiere, se debe a la agrupación de individuos, mientras que el segundo aparece por la distribución diferenciada de ciertas variables de confusión creada por la agrupación. La suma de estos componentes la llama sesgo *cross-level*, el cual puede provocar que la asociación ecológica sea más fuerte o débil que a nivel individual. Los componentes antes explicados son análogos a los efectos de escala y agrupación en el PUAM (Gotway y Young, 2002).

Estos problemas conllevan a la necesidad de ser considerados en cualquier estudio, dado que la escala puede tener efectos considerables sobre los resultados. Y además han conducido a una línea de investigación sobre la mejora de métodos de agregación espacial, dado que las regiones cada vez adquieren un mayor papel en los estudios socioeconómicos, como se describirá en el capítulo siguiente.

## CAPÍTULO 4

# MARCO METODOLÓGICO

---

En este capítulo se explican algunas de las aplicaciones que han tenido los métodos de agregación, además de una descripción de la tipología de los mismos. Además se hace una descripción del algoritmo *MaxP* y de los datos y fuentes de información a utilizar en la presente investigación.

### 4.1 MÉTODOS DE AGREGACIÓN

Por lo explicado respecto al Problema de Unidad de Área Modificable es que los métodos de agregación han sido ampliamente investigados. Estos métodos en la literatura han recibido diferentes nombres, dependiendo de los objetivos y métodos aplicados; entre ellos *regionalization*, *zone design*, *zoning*, *clustering*, *districting and redistricting*, *territory design*, *territory alignment*, *spatial allocation*, *partitioning* y *spatial tessellations* (Moreno y García, 2011).

Duque *et al.* (2007) plantean que la importancia de los diferentes métodos de agregación, y su continua mejora, radica en que la inferencia estadística basada en las regiones o áreas definidas administrativamente tienden a estar afectadas por el Problema de Unidad de Área Modificable o el de falacia ecológica. Los autores también describen una serie de características que comparten los métodos de agregación:

- Los métodos generalmente agrupan áreas en un número predeterminado de regiones, al tiempo que optimizan un criterio particular de agregación.
- Las áreas dentro de una región están geográficamente conectadas (restricción de contigüidad espacial).

- El número de regiones debe ser menor o igual al número de áreas.
- Cada área se asigna a una y sólo una región.
- Cada región debe contener al menos una área.

Dado el alcance que tiene la estructura espacial subyacente a los datos en los análisis, resulta relevante la consideración de las restricciones de contigüidad espacial en los métodos de agregación. Así, estos métodos se pueden clasificar en dos grandes grupos: métodos sin una restricción espacial explícita, y aquellos que sí la incluyen explícitamente.

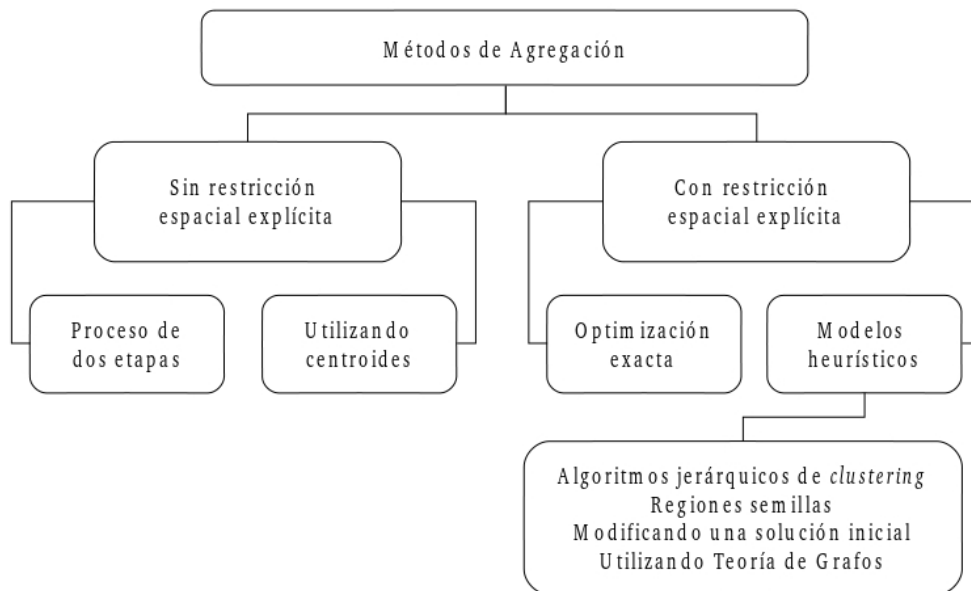


Figura 4.1: Clasificación de Métodos de Agregación Espacial.

Fuente: Elaboración propia con base en Duque *et al.* (2007)

#### 4.1.1 MÉTODOS SIN RESTRICCIÓN ESPACIAL EXPLÍCITA

Estos métodos satisfacen indirectamente la restricción de contigüidad, y ésta, por lo general, se aplica posteriormente. Estos métodos a su vez se dividen en dos

subgrupos, a través de un proceso de dos etapas y utilizando centroides.

#### 4.1.1.1. REGIONALIZACIÓN A TRAVÉS DE UN PROCESO DE DOS ETAPAS

Openshaw (1977) describe un algoritmo desarrollado previamente por él en 1973 que consta de dos etapas; en la primer etapa cada área es asignada a una única zona, de la cual es el único miembro; entonces cada par de zonas cuya agregación produzca la mayor mejora en el valor de la función objetivo son unidas. El proceso de agregación de optimización local continúa hasta que sólo hay una zona, o las discontinuidades en la matriz ya no permiten más fusiones; en cada etapa, la nueva solución se deriva de manera directa de la solución anterior. Este grupo de métodos no impone la restricción de compacidad regional.

#### 4.1.1.2. REGIONALIZACIÓN UTILIZANDO CENTROIDES

En algunas ocasiones se requiere que las regiones formadas cumplan el criterio de compacidad, el cual implica que una zona sea lo más “circular” posible, evitando el fenómeno conocido como *gerrymandering*.<sup>1</sup> Esta problemática ha sido estudiada mayormente para la formación de distritos electorales, los cuales enfrentan dos restricciones básicas extras: igual población para cada zona y contigüidad. Murray y Shyy (2000) intentan construir regiones homogéneas. Con las distancias métricas y las medidas de similitud (e.g. características económicas de la población) entre áreas conforman un criterio que busca minimizar la diferencia total intragrupos. Lo más importante en este criterio es la decisión del número de agrupaciones a formar y el peso asignado a los componentes, de manera tal que la elección de la importancia de la proximidad espacial lleva a formar soluciones contiguas y más compactas. Bação *et al.* (2005) utilizando algoritmo genéticos propone una metodología para hacer frente a las restric-

---

<sup>1</sup>El término *gerrymandering* se refiere a la manipulación de los distritos electorales con el fin de afectar los resultados electorales, provocando que algunos estén sobre o subrepresentados (Bozkaya *et al.*, 2003).

ciones; encuentran que se obtienen mejores resultados cuando no se fuerzan las áreas a los centroides.

#### 4.1.2 MÉTODOS CON RESTRICCIÓN ESPACIAL EXPLÍCITA

Estos métodos incorporan mecanismos que aseguran la contigüidad espacial en cada región. Es decir, se requiere información sobre las relaciones espaciales entre las zonas vecinas. A continuación se muestran los grandes conjuntos en que se pueden agrupar estos métodos.

##### 4.1.2.1. MODELOS DE OPTIMIZACIÓN EXACTA

Se ha tratado de encontrar el número de soluciones factibles de un problema de regionalización de manera exacta, tarea que se dificulta al incluir una restricción de contigüidad espacial. Por ejemplo, Zolterns y Sinha (1983) incorporan la contigüidad a través de las redes de carreteras. Cova y Church (2000) indican que para incluir la contigüidad en un modelo se pueden seguir dos estrategias: la primera es restringir un perímetro predefinido, de manera que un número pequeño resultará en un solución contigua; la segunda opción es establecer un área de influencia alrededor de un sitio central, si se busca minimizar costos entonces la solución serán sitios contiguos. En la mayoría de estas técnicas exactas a pesar de lograr una propuesta teóricamente adecuada, en la práctica no es eficiente pues la restricciones de contigüidad se incrementa exponencialmente con el número de áreas y regiones a considerar.

##### 4.1.2.2. MODELOS HEURÍSTICOS PARA EL PROBLEMA DE REGIONALIZACIÓN

Dada la dificultad de resolver esta problemática de manera exacta se ha incrementado el campo de desarrollo de heurísticas.<sup>2</sup> Una característica importante de estas

---

<sup>2</sup>Una heurística es una técnica que busca soluciones de calidad a través de un conjunto de reglas de costo computacional razonable, ofreciendo una buena solución que se aproxima a la optimalidad

aproximaciones es la habilidad de salir de un óptimo local.

#### ALGORITMOS JERÁRQUICOS DE CLUSTERING ADAPTADOS

Los algoritmos jerárquicos son modificados de manera que sólo se permiten regiones conectadas espacialmente, y al inicio del algoritmo cada área es una región por sí misma (Duque *et al.*, 2007). Entre los primeros trabajos se encuentra el de Lankford (1969), el cual propone un método que considera de forma explícita la contigüidad, y realiza algunas comparaciones contra otros dos diferentes métodos de *clustering*. En su algoritmo se enfrenta al problema técnico de definir a los vecinos; sin embargo este método durante el proceso únicamente considera a las unidades contiguas, mientras que los otros analizan tanto a las continuas como a las que no lo son.

#### REGIONES SEMILLAS

Lo que caracteriza estas heurísticas es que cada región resulta de seleccionar un área semilla sobre la cual se van asignando áreas vecinas. Openshaw (1977) propone una metodología que va mejorando una partición inicial dada. El autor toma aleatoriamente una zona, realiza movimientos con las áreas contiguas a ella y verifica que haya mejoras en la función objetivo. Este procedimiento se realiza hasta que la mejora sea menor a una tolerancia especificada previamente.

#### MODIFICACIÓN DE UNA SOLUCIÓN FACTIBLE INICIAL

Estos métodos parten de una solución factible que de forma iterativa busca mejoras en el criterio de agregación, lo importante es que las iteraciones deben ser factibles en términos de la restricción de contigüidad espacial (Duque *et al.*, 2007). Openshaw y Rao (1995) y Macmillan (2001) utilizan heurísticas de búsqueda de mejoras a una solución porque los procedimientos podrían estancarse en una solución subóptima. La diferencia radica en que el algoritmo de Openshaw y Rao (1995) toma una región y verifica qué áreas vecinas pueden ser anexadas a una región siempre que mejore el criterio de evaluación y no se afecte la estructura de contigüidad; mientras que el de (Voß, 2001).

Macmillan (2001) analiza cuáles movimientos permiten mejorar los criterios en la región donante como la de recepción, sin alterar las estructuras de contigüidad de ambas.

#### MODELOS BASADOS EN TEORÍA DE GRAFOS

Hansen *et al.* (2003) sugiere que entidades muy diferentes en los extremos de una cadena consecutiva de entidades similares suelen ser asignadas a la misma categoría. Así, maximizando los *splits* (los cuales se define como la menor disimilitud entre entidades de diferentes *clusters*) de un grafo que representa las unidades y su contigüidad logra determinar una partición óptima. Assunção *et al.* (2006) proponen un método que se apoya en grafos, sin embargo es costoso computacionalmente y desarrollan entonces una heurística, la cual se repite mientras que el número de sub-árboles sea menor al número de *clusters* deseados.

## 4.2 APLICACIONES DE LOS MÉTODOS DE AGREGACIÓN

Los métodos de agregación buscan formar agrupamientos de unidades espaciales a partir de un cierto criterio. Dichos criterios pueden ser de corte económico o geográfico, por mencionar algunos; sin embargo, el considerar las restricciones espaciales es cada vez más utilizado (Kalcsics *et al.*, 2005). Siguiendo a Moreno y García (2011) y a Kalcsics *et al.* (2005) a continuación se presentan algunas de las aplicaciones que han tenido diferentes métodos de agregación.

#### DELIMITACIÓN DE DISTRITOS POLÍTICOS

Para Hojati (1996) determinar territorios políticos se refiere a la división de una determinada área en secciones censales, por ejemplo, de manera que cada distrito tenga una población de votantes semejante. El autor sugiere, además, que es deseable que

los distritos sean contiguos y conexos; las áreas compactas son más convenientes para el electorado y los representantes elegidos en términos de distancias. El algoritmo de Bozkaya *et al.* (2003) asigna pesos a los criterios a considerar para formar las agrupaciones; entre los criterios usuales en el diseño de distritos políticos se encuentran los siguientes: Los distritos posean un número similar de votantes con el fin de seguir el principio “un hombre-un voto”; los distritos deben ser contiguos; deben respetarse límites naturales, tales como lagos; respetar la existencia de subdivisiones administrativas, como los *census tracts*; homogeneidad socioeconómica, lo cual asegura una mejor representación de residentes que comparten intereses; debe respetarse la integridad de las comunidades; evitar dividir comunidades entre diferentes distritos; e igual probabilidad de representación, ello para lograr que grupos minoritarios tengan sus representantes.

#### DISEÑO DE ÁREAS DE MERCADO

Ríos-Mercado y Fernández (2009) establecen que este problema se puede ver como la agrupación de unidades geográficas para conformar zonas de venta, de manera que éstas sean óptimas de acuerdo con criterios de planificación pertinentes como: contigüidad geográfica, para que las unidades de venta puedan cubrir el territorio; compacidad, para que los clientes estén cercanos; y un número fijo de territorios. El hecho de que el número de regiones está dado llevo a los autores a plantear el problema como la localización de  $p$  centros (uno para cada territorio), considerando la dispersión basada en una medida de distancia, y que las zonas deben estar balanceadas y cumplir un valor umbral en los tres criterios antes mencionados.

#### ZONAS PARA INSTALACIONES PROVEEDORAS DE SERVICIOS EN UNA LOCALIZACIÓN FIJA

En ocasiones, se requiere acudir a un establecimiento para la prestación de un servicio, como colegios u hospitales, y se desea obtener regiones conexas de manera que los habitantes en ellas reciban el servicio de forma equitativa. Un estudio dentro de este campo es el de Armstrong *et al.* (1993), proponen una metodología que permitió re-

portar información sobre diferentes agregaciones para conjuntos de demanda dispersos. Su propuesta permitía que también los administrativos escolares determinaran áreas que probablemente presentaran crecimiento posteriormente, y con ello planear posibles expansiones de los establecimientos ya existentes, o incluso la construcción de nuevos.

#### ZONAS DE SERVICIO DE EMERGENCIAS

Ha sido de interés establecer regiones equilibradas respecto a la asignación de servicios. Un trabajo en este ámbito es el de D'Amico *et al.* (2002), el cual reporta un caso de estudio para el diseño de distritos policiales; los autores señalan que el problema de estos diseños recae en la consideración de las cargas de trabajo de los oficiales y los tiempos de respuesta. Usualmente, para asignar patrullas se dividían los automóviles en proporción directa al total de llamadas para el servicio; mientras que la propuesta de los autores es que tras fijar los distritos se asigne un número óptimo de patrullas que debieran estar en servicio en cada distrito. Sugieren que una mejora al algoritmo sería iniciar con una partición que minimice la frontera promedio de cada distrito geográfico, lo que llevaría a una forma más compacta y en la que se esperaría un menor tiempo de respuesta.

#### ZONAS RECEPTORAS DE RECURSOS ENERGÉTICOS

Siguiendo a Bergey *et al.* (2003), se explica que el problema consiste en lograr una partición física de la red, la cual debe generar zonas de distribución económicamente viables para las compañías eléctricas. Además, refieren que el interés es que las agrupaciones tengan aproximadamente el mismo potencial de ganancias; mientras que la compacidad geográfica permite un mejor manejo y un mantenimiento más económico. Tiede y Strobl (2006) tienen otro objetivo, buscan que las regiones formadas estén balanceadas en cuanto a fuentes y consumo de energía. Así, proponen construir regiones de polígonos con altos potenciales de energía unidos a polígonos vecinos con altos niveles de consumo. Para formar las regiones, los autores parten de cinco áreas densamente pobladas para iniciar el algoritmo.

#### ZONAS PARA LA ASIGNACIÓN DE USOS DEL SUELO

Entre las aplicaciones que sugieren Cova y Church (2000) se encuentran la localización de reservas biológicas, rellenos sanitarios, subdivisión residencial y sitios para residuos peligrosos. Explican que su propuesta se enfrenta a la necesidad de que el área del sitio se reduzca y se desee mayor compacidad; en caso contrario, se requiere de la mayor experimentación con algoritmos heurísticos. Aerts y Heuvelink (2002) analizan el caso de un zona minera de lignito; la cual se forma a su vez de dos áreas de distinto uso: una para la extracción y una para el vertedero de desechos. Se enfrentaban a la necesidad de proponer un plan en el que los usos de tierra perdidos por la minería en ese lugar deberían ser restaurados a la brevedad. Así, las restricciones principales de las asignaciones de uso de tierra restauradas fueron que deberían ser similares en tipo y tamaño a las existentes antes de iniciar la actividad minera.

#### ANÁLISIS DE PROBLEMÁTICAS PARTICULARES

Rey y Sastré-Gutiérrez (2010) analizan las disparidades interregionales sobre un período de sesenta años para los estados de México. Encuentran que las disparidades en el ingreso tuvieron reducciones significativas de 1940 a 1960, sin embargo desde entonces las disminuciones en las disparidades no han sido significativamente altas. Además, muestran que la utilización de diferentes regionalizaciones afecta los resultados de la medición de las disparidades interregionales. Rey *et al.* (2011) utilizando el algoritmo de agregación *MaxP* elaboran indicadores micro y macro de cambio espacial. Dicho cambio se da en el aspecto composicional que se conforma por el conjunto de variables utilizadas en el análisis para los años 1990 y 2000; así como en la escala espacial al reconfigurar las fronteras de los *census tracts*. Este estudio señala la importancia de analizar simultáneamente el cambio en la composición de las unidades y el cambio de sus fronteras, con el fin de mejorar la efectividad de las políticas formuladas.

### 4.3 ALGORITMO DE AGREGACIÓN *MaxP*

De la descripción hecha previamente, puede notarse que gran parte del interés por encontrar agrupaciones de áreas es para que dichas agregaciones sean adecuadas para un determinado análisis. Por ejemplo Rey y Sastré-Gutiérrez (2010) contrastan diversas regionalizaciones propuestas para México, tanto por instituciones como por investigadores; con la ofrecida por el algoritmo *MaxP*. Los autores encuentran que la agregación obtenida por este algoritmo es la única en la que la desigualdad interregional es estadísticamente significativa para todos los años bajo estudio.

El algoritmo *MaxP* por su parte puede endogeneizar el número de regiones a construir. A diferencia de la mayoría de los métodos, en este método no se requiere establecer previamente un número de regiones en las que se deben agrupar las áreas, lo cual es difícil ya que por lo general el investigador desconoce esa condición (Rey y Sastré-Gutiérrez, 2010 : 290). El método básicamente agrega una determinada cantidad de áreas en un número máximo desconocido de regiones, las cuales satisfacen un valor predefinido de un atributo espacial (e.g. población por región, ingreso por región, etc.). El procedimiento no requiere imponer la compacidad porque la restricción de la contigüidad espacial es explícita y además maximiza la homogeneidad intrarregional (Duque *et al.*, 2011).

Dada la problemática a estudiar en la presente investigación, el uso del algoritmo resulta adecuado por diversas cuestiones. En primera instancia, de manera previa se desconoce el número de particiones en que debe fraccionar el territorio. Por otro lado, el hecho de que minimice la heterogeneidad interna de las agregaciones permite aportar conocimiento sobre agrupaciones que pudieran ser definidas como áreas rurales o urbanas en México, dado que como se indicó en el Capítulo 2 las definiciones actuales no cuentan con soportes técnicos que permitan analizarlas desde esta perspectiva.

El problema de las regiones *MaxP* busca la partición óptima  $P_p^*$ . Siendo  $\Pi$  el conjunto de las posibles particiones  $P_p$  de un territorio y  $H$  una función que mide la heterogeneidad total de la partición, el problema se formula así:

$$\text{Determinar } P_p^* \text{ tal que } |P_p^*| = \max(|P_p| : P_p \in \Pi), \text{ y}$$

$$\nexists P_p \in \Pi : |P_p| = |P_p^*| \wedge H(P_p) < H(P_p^*)$$

La resolución exacta del problema *MaxP* es computacionalmente costosa, por lo tanto se recurre una metaheurística para aproximar las soluciones. El algoritmo computacional se divide en dos fases básicas, la de construcción y la de búsqueda local: la primera genera un conjunto de soluciones factibles, mientras que la segunda intenta mejorar el criterio de evaluación.<sup>3</sup>

## 4.4 DATOS Y FUENTES DE INFORMACIÓN

### 4.4.1 DISEÑO DE LA INVESTIGACIÓN

Para objeto de la presente investigación, en la utilización del algoritmo *MaxP* se utilizan los criterios que a continuación se explican.

Se utiliza una matriz de pesos espaciales tipo *queen* que representa la estructura espacial de las unidades definiendo como contiguas a aquellas que comparten una frontera mayor a cero o un vértice.<sup>4</sup> Siguiendo a Anselin (1988) esta matriz surge de la noción de contigüidad binaria entre unidades espaciales; si comparten una frontera común de longitud mayor a cero se dice que son contiguas y se asigna el valor de uno. La matriz de contigüidad espacial es entonces un arreglo bidimensional, donde los elementos no nulos de las columnas representan a las unidades espaciales contiguas al

<sup>3</sup>Para una descripción más detallada revisar el Apéndice A.1.

<sup>4</sup>Vértices son los puntos nodales que definen las esquinas límite de un polígono. Consultado en <https://geodacenter.asu.edu/node/390#vertex>.

elemento fila.

El arreglo de observaciones de los  $m$  atributos a considerar de las  $n$  áreas se formará con diferentes conjuntos de características socioeconómicas de la población de México que se detallan en el apartado 4.4.2. El umbral puede establecerse sobre el mínimo de unidades que deba contener una agregación, o bien, sobre una determinada variable, mismas que se justifican en la siguiente sección y pretenden ser exploratorias.

Por último, la decisión sobre el número de soluciones iniciales a generar, es decir, el número de veces que se iterará el algoritmo a fin de encontrar una mejor solución, se decidió establecer en 99 soluciones iniciales como un primer acercamiento exploratorio.

Por otro lado, la función permite sugerir elementos denominados “semillas”. Estos elementos son aquellas áreas con las que se desearía iniciar agregaciones, esto es, que cada una de estas semillas genere una agrupación distinta. En caso de no sugerirse elementos iniciales, el algoritmo selecciona áreas de manera aleatoria para iniciar las agregaciones. De nuevo, para la realización de este estudio no se utilizan semillas debido a su carácter exploratorio.

Para la aplicación de todos los ejercicios se recurrió al lenguaje *Python v.2.7.3* y a la librería *PySAL v.1.3.0*.

#### 4.4.2 DATOS

En México, el territorio se divide en 31 estados y el Distrito Federal (DF). De ellos, cada estado se divide en municipios y el DF en delegaciones, tal como lo ilustra la Figura 4.2. La escala a utilizar será el municipio (incluyendo delegaciones suman 2456 unidades espaciales) debido a la posibilidad de utilizar información disponible en el Censo de Población y Vivienda 2010 de México. Con la ayuda de PySAL se construyó una matriz *queen* de contigüidad espacial de los municipios mexicanos.

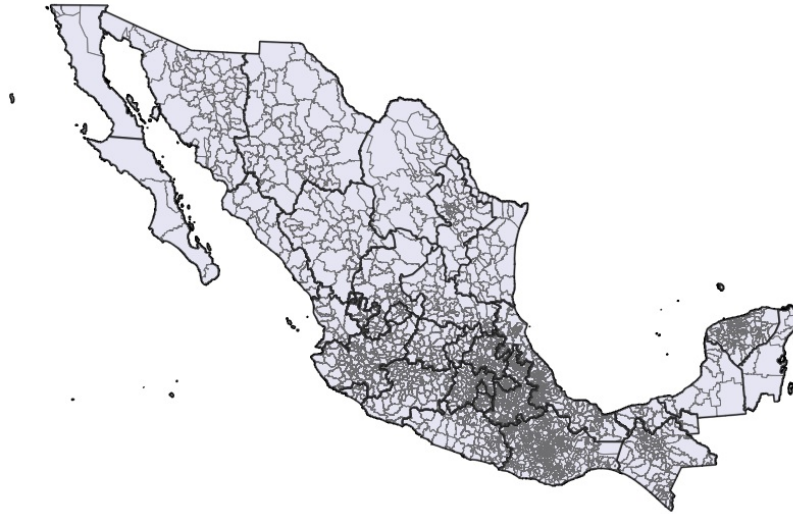


Figura 4.2: Estados y municipios de México, 2010.

Elaboración propia.

En la aplicación se incluyeron distintos atributos de la población. La inclusión de un conjunto de atributos es posible debido a que el algoritmo obtiene la varianza del arreglo de variables a utilizar; así, la función busca el esquema con mayor número de agregaciones y que presente la menor varianza total (suma de las varianzas de cada agrupación). Las combinaciones de las variables se eligieron apoyándose en la revisión de estudios previos en los que son consideradas como representativas de las características socioeconómicas de la población (para una revisión más detallada consultar el Apéndice A.2).

Las variables elegidas para la investigación fueron agrupadas en dimensiones, partiendo inicialmente de los criterios de INEGI, y en seguida se agruparon para el ejercicio de acuerdo a características similares.<sup>5</sup> Así, las variables a utilizar de cada dimensión se presentan en la Tabla 4.1.

---

<sup>5</sup>Se siguió el formato de Temáticas Censales de INEGI (INEGI, 2011), aunque no todas fueron utilizadas en este ejercicio.

Dimensión	Código	Variable
Población	D1_POB_TOT	-Población total por municipio
	D1_MEN_2500_PRC	-Porcentaje de población residente en localidades con menos de 2500 habitantes
	D1_MEN_15000_PRC	-Porcentaje de población residente en localidades con menos de 15000 habitantes
Fecundidad	D2_PROM_HIJ_VIV	-Promedio de hijos nacidos vivos por municipio
Mortalidad	D3_HIJ_FALL_PRC	-Porcentaje de hijos fallecidos por municipio
Lengua Indígena	D5 LENG_IND	-Población de 3 años y más que habla alguna lengua indígena <sup>6</sup>
Educación	D7_ANALFAB	-Población de 15 años y más por municipio analfabeta
	D7_PRIM_TERM	-Población de 15 años y más por municipio con grado de escolaridad primaria
	D7_ED_MED_SUP	-Población de 15 años y más por municipio con educación media superior <sup>7</sup>
Características Económicas	D8_POB_OC	-Población Ocupada por municipio
	D8_AGROP_PRC	-Porcentaje de población ocupada como trabajadores agropecuarios
	D8_SEC_TERC_PRC	-Porcentaje de población ocupada en actividades secundarias y terciarias, como profesionistas, técnicos y administrativos <sup>8</sup> , como trabajadores en la industria <sup>9</sup> y como comerciantes y trabajadores en servicios diversos. <sup>10</sup>
	D8_1SM_PRC	-Porcentaje de población ocupada que recibe hasta 1 salario mínimo mensual
	D8_MASDE1SM_PRC	-Porcentaje que recibe más de 1 salario mínimo mensual <sup>11</sup>

En base a los criterios de la aplicación antes descritos y las variables seleccionadas, se realizaron un conjunto de pruebas. Los resultados obtenidos y su análisis se encuentran descritos en el capítulo siguiente.

<sup>6</sup>Población que habla alguna lengua indígena ya sea que hable o no español, además de aquella que habla alguna lengua indígena y no especificó si habla español.

<sup>7</sup>Incluye a la población que tiene al menos un grado aprobado en estudios técnicos o comerciales con secundaria terminada, preparatoria o bachillerato y normal básica.

<sup>8</sup>Funcionarios, directores y jefes, profesionistas y técnicos, trabajadores auxiliares en administrativas.

<sup>9</sup>Mecánicos y trabajadores industriales y artesanales; operadores de maquinaria industrial, ensambladores, choferes y conductores de transporte.

<sup>10</sup>Comerciantes, empleados y agentes de ventas; trabajadores en servicios personales, vigilancia y fuerzas armadas; y trabajadores en actividades elementales y de apoyo.

<sup>11</sup>Porcentaje de población que recibe de 1 a 2 salarios mínimos mensuales y porcentaje de población que recibe más de 2 salarios mínimos mensuales.

<sup>12</sup>Incluye las viviendas particulares habitadas con agua entubada dentro de la vivienda o en el terreno, la que se acarrea de otra vivienda o de un hidrante o llave pública.

Dimensión	Código	Variable
Servicios de Salud	D9_SERV_SAL	-Población derechohabiente por municipio
Características de la Vivienda	D14_MAS25OC_PRC	-Porcentaje de viviendas por municipio con más de 2.5 ocupantes por cuarto
	D14_PTIERRA_PRC	-Porcentaje de viviendas por municipio con piso de tierra
	D14_AGUA_PRC	-Porcentaje de ocupantes en viviendas por municipio que disponen de agua entubada <sup>12</sup>

Tabla 4.1: Variables y dimensiones a utilizar en la aplicación.

## CAPÍTULO 5

### RESULTADOS

---

A continuación se presentan los resultados obtenidos de la aplicación a una selección de variables. Como se explicó en el capítulo previo se utiliza una matriz *queen* de contigüidad, se estableció en 99 el número de soluciones iniciales y no se sugirieron unidades geográficas como semillas.

Los resultados obtenidos se resumen en las siguientes tablas. En la Tabla 5.1 se consideraron 3 variables de diferentes dimensiones, en la Tabla 5.2 se eligieron 3 variables que pertenecieran a una misma dimensión, y en la Tabla 5.3 se consideró una sola variable. Se aplicaron diferentes umbrales, en las tablas se especifican aquellos que fueron utilizados en cada caso.

Variables/Umbral	50'000 sobre D1_POB_TOT	5 municipios	3 municipios
D3_HIJ_FALL_PRC D7_ED_MED_SUP D8_MASDE1SM_PRC	685	395	688
D1_MEN_15000_PRC D8_1SM_PRC D14_MAS25OC_PRC	685	395	690
D1_MEN_2500_PRC D5_LENG_IND D14_MAS25OC_PRC	695	398	702

Tabla 5.1: Resultados usando 3 variables de distintas dimensiones.

Se eligieron diferentes umbrales, algunos sobre población y otros sobre número de municipios. El umbral sobre la variable de población se eligió siguiendo el criterio para definir a un municipio como central dentro de una zona metropolitana (SEDESOL, INEGI y CONAPO, 2007). El criterio de 75 % de la población ocupada en actividades secundarias y terciarias es utilizado para definir municipios exteriores en una ZM (SEDESOL, INEGI y CONAPO, 2007). El umbral del 50 % de población ocupada en

Variables/Umbral	50'000 sobre D1_POB_TOT	5 municipios	3 municipios
D7_ANALFAB D7_PRIM_TERM D7_ED_MED_SUP	683	395	696
D8_POB_OC D8_SEC_TERC_PRC D8_MASDE1SM_PRC	676	397	699
D14_MAS25OC_PRC D14_PTIERRA_PRC D14_AGUA_PRC	688	395	695

Tabla 5.2: Resultados usando 3 variables de la misma dimensión.

Variables/Umbral	Umbral	5 municipios	3 municipios
D1_MEN_15000_PRC	50'000 sobre D1_POB_TOT 680	394	700
D8_MASDE1SM_PRC	50'000 sobre D1_POB_TOT 688	390	696
D8_AGROP_PRC	50 sobre D8_AGROP_PRC 1081	395	698
D8_SEC_TERC_PRC	75 sobre D8_SEC_TERC_PRC 1427	400	693

Tabla 5.3: Resultados usando 1 variable.

actividades agropecuarias se tomó siguiendo el criterio de la SAGARPA, que define como rurales a aquellas localidades donde la agricultura representa más del 50% de la producción local (OCDE, 2007). Por otra parte, los umbrales de 3 y 5 municipios fueron elegidos para conocer las diferencias en los resultados al definir distintos niveles de agregación.

La variable Población Total en las Tablas 5.1 y 5.2 únicamente es la característica sobre la cual se desea se cumpla un mínimo, y no forma parte de los atributos a homogeneizar. En la Tabla 5.3 las dos primeras combinaciones tienen este mismo umbral, no así la tercera y cuarta; pues en esos casos el atributo a homogeneizar es el mismo que funciona como umbral (Porcentaje de Población Ocupada en actividades agropecuarias y Población Ocupada en actividades secundarias y terciarias, respectivamente).

De la Tabla 5.1, la segunda y tercera combinación son intentos por definir características de la población rural, por lo que el umbral sobre población puede no ser representativo. Sin embargo, se nota que el número de agregaciones a pesar de ello no resultan ser considerablemente distintos de la primera combinación.

La Tabla 5.2 presenta tres combinaciones de variables que pertenecen a la misma dimensión. Lo cual se realizó con el fin de analizar si los resultados variaban respecto al ejercicio anterior, ya que la consideración de variables que pudieran representar un mismo aspecto de la población pudieran lograr un número menor de agregaciones. Sin embargo, resalta el hecho de que el número de agregaciones es similar.

La Tabla 5.3 considera una sola variable y diferentes umbrales. En el caso de la primer variable, el umbral sobre población una vez más puede no ser representativo debido a que se están considerando municipios donde el nivel de población es bajo. Pero deben destacarse las agregaciones de la tercer y cuarta variable con sus respectivos umbrales, pues en esos casos se dio un incremento significativo en el número de agregaciones, mientras que el número se mantiene similar a los resultados previos al considerar mínimos de 3 y 5 unidades.

De las tres tablas resulta relevante el hecho de que a pesar de las variaciones en las combinaciones seleccionadas se mantienen ciertos niveles en el número de agregaciones, salvo en los dos últimos casos de la Tabla 5.3. La variación en el total de agregaciones se explica en parte por la lógica del algoritmo, recordando que el objetivo es encontrar el mayor número de agregaciones con la mínima heterogeneidad posible.

Debe añadirse que en las pruebas en las que no se solicita un mínimo de unidades en la agregación, el algoritmo arroja resultados en los que se presentan numerosas agrupaciones formadas solamente por uno o dos municipios. Esto último resalta que, si únicamente se parte de las características de la población entonces resulta una cifra alta de enclaves, y es el umbral sobre el número de municipios a agregar el que de

manera artificial logra una mayor asociación de unidades.

## 5.1 VISUALIZANDO SOLUCIONES: CASOS PARTICULARES DE LAS ZM DE TIJUANA, MONTERREY Y MEXICALI

A continuación se realiza un ejercicio de visualización como parte de las técnicas exploratorias, en donde se analizan las zonas metropolitanas a fin de sustentar los hallazgos antes descritos. Se tomaron tres diferentes ZM, para analizar qué había ocurrido en esos casos particulares al considerar las diferentes composiciones y umbrales, a continuación se presenta un resumen de los resultados.

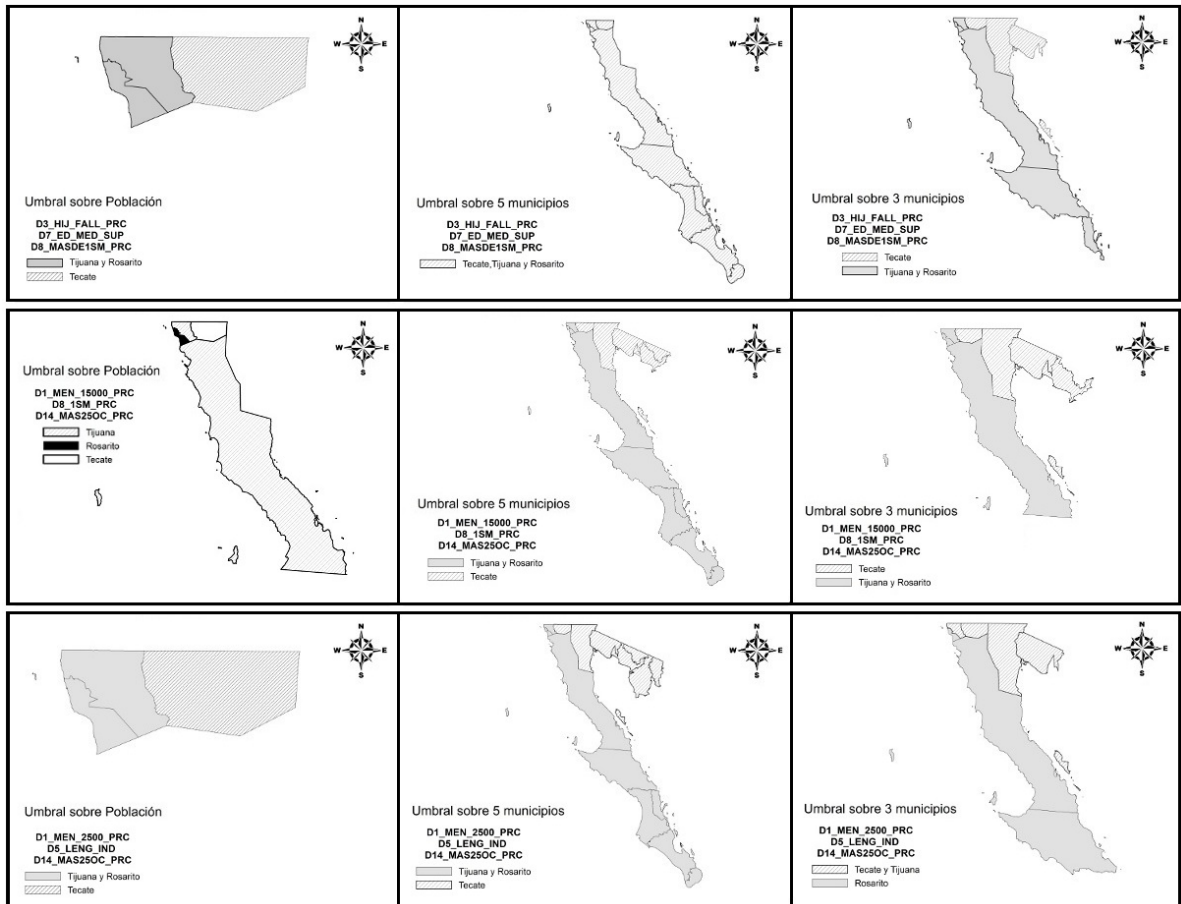


Figura 5.1: Análisis del caso particular de la ZM de Tijuana.

Elaboración propia en *ArcGis 10.0*.

Laboratorio de Análisis Socioeconómico Espacial, Spatial-SEALab.

Se tomó la ZM de Tijuana que está formada por los municipios de Tecate, Tijuana y Playas de Rosarito, donde el primero fue agregado tras la delimitación de 2007 bajo el criterio de planeación y política urbana. Bajo el umbral sobre población y la primera combinación, de las 685 agregaciones totales, los municipios de Tijuana (02004) y Rosarito (02005) se mantienen unidos en la agregación 311, mientras que Tecate (02003) forma la agrupación 508 por sí solo. En el caso del mismo umbral y tercera combinación, los municipios se configuran de la misma manera en las agrupaciones 207 y 211 respectivamente, de las 695 resultantes. Es en el segundo esquema de variables

cada municipio pertenece a diferentes agregaciones, Tecate a la 587, Tijuana a la 456 (junto a Ensenada) y Rosarito a la 285. Siendo este último caso el único en que los tres municipios se encuentran en distintas agrupaciones cada uno.

Cuando se considera el umbral de 5 unidades, de las 395 agregaciones totales es la primera combinación de variables la que logra unir a los tres municipios en el mismo grupo 327(02001, 02003, 02004, 02005, 03002, 03003, 03008 y 03009). En las otras dos combinaciones, una vez más los municipios Tijuana y Rosarito se mantienen juntos (agrupación 298 para la segunda combinación: 02004, 02005, 03001, 03002, 03003, 03008, 03009; y 231 para la tercera: 02004, 02005, 03001, 03002, 03003, 03008 y 03009) mientras que Tecate forma parte de una agregación distinta (agregación 68 para el segundo esquema: 02002, 02003, 26048, 26055, 26070 y 51 para el tercero: 02002, 02003, 26004, 26017, 26048, 26055, 26070).

Ocurre algo similar para la primera y tercera combinación de variables tomando el umbral de 3 municipios. De la primera combinación se forma el grupo 272 que contiene a las unidades 02001, 02004, 02005, 03002 y 03009 y Tecate dentro del grupo 488 (02002, 02003 y 26055). En la segunda combinación es al grupo 66 al que pertenecen Tijuana y Rosarito (02001, 02004, 02005) y Tecate al 163 (02002, 02003, 26048 y 26055). Sólo en la tercera combinación de este umbral se reúnen Tecate y Tijuana en el agrupamiento 30 (02002, 02003, 02004, 26055) y Rosarito queda separado en el 337 (02001, 02005 y 03002).

Es interesante que en seis de los nueve esquemas se mantienen juntos Tijuana y Rosarito y se excluye a Tecate (311, 272, 298, 66, 207 y 231). Lo anterior lleva a pensar que la unión de esa unidad de manera oficial debiera ser analizada con más detalle.

Para el caso de la ZM de Monterrey no se delinearon completas todas las diferentes agregaciones en cada una de las nueve combinaciones de variables de una misma

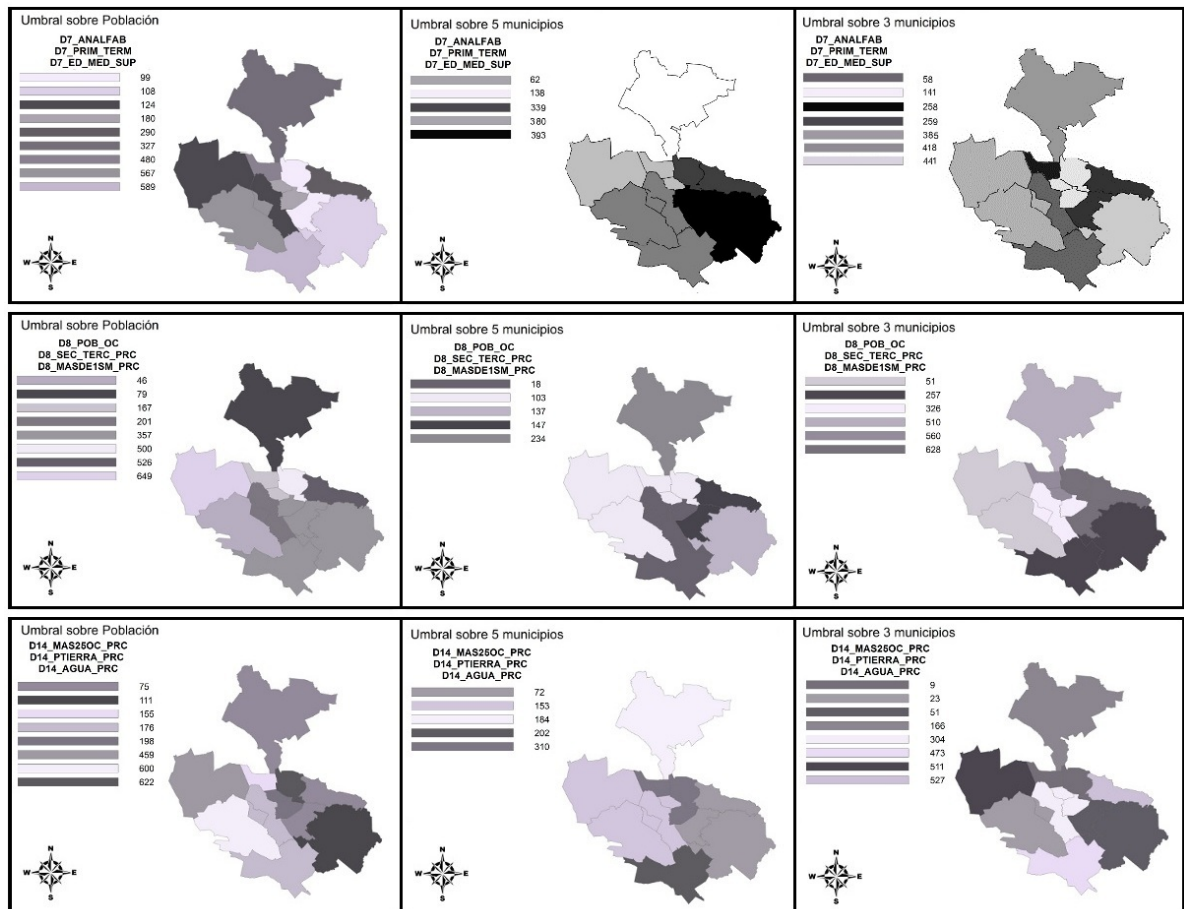


Figura 5.2: Análisis del caso particular de la ZM de Monterrey.

Elaboración propia en *ArcGis 10.0*.

Laboratorio de Análisis Socioeconómico Espacial, Spatial-SEALab.

dimensión. Únicamente se delimitan los trece municipios que la conforman y se realiza con la escala de color el número de agregación a la que pertenecen, aunque en algunas ocasiones éstas se encuentran formadas solamente por el municipio en cuestión. De esta manera se ilustran las unidades de la ZM que pertenecen a una misma agrupación, para facilitar la visualización.

La primera combinación con umbral sobre población divide a los municipios en nueve diferentes agregaciones, los municipios 19009, 19021, 19041, 19045 y 19049 per-

tenecen a las agrupaciones 108, 480, 290, 337 y 589, respectivamente; mientras que las unidades 19018 y 19039 forman parte de la 124, 19026 y 19046 de la 180, 19006 y 19031 de la 99, y 19019 y 19048 de la 567. La segunda y tercera combinaciones con este umbral dividen a los municipios en ocho agregaciones.

Bajo el umbral de 5 unidades, la primera combinación logra que doce de los trece municipios se distribuyan en agrupaciones internas: 19019, 19039, 19048 y 19049 están en el grupo 380; mientras que el 19019, 19026 y 19031 forman el grupo 62, el 19018, 19021 y 19046 forman el 393; las unidades 19006 y 19041 se encuentran en el número 339 y el 19045 está dentro de la 138. La segunda y tercera combinaciones con este mismo umbral consiguen que al menos cinco municipios formen parte de una misma agregación. Por ejemplo, con las variables D8 los municipios 19006, 19018, 19021, 19046 y 19049 pertenecen a la agrupación 103, mientras que las variables D14 agrupan en la 153 a los elementos 19018, 19019, 19039, 19046 y 19048.

Considerando el umbral de 3 municipios y las tres combinaciones de variables, la ZM se dividió en 7, 6 y 8 diferentes agregaciones. En estos tres últimos casos, se logra agrupar máximo a tres municipios. En los nueve ejercicios es recurrente que los municipios formen parte de una agregación donde no se incluyen más elementos de la zona, o bien, constituyen una agregación por sí mismos. Los resultados muestran que a pesar de la conurbación municipal en la ZM de Monterrey y del cumplimiento de criterios de integración funcional, ésta presenta diferencias al considerar diferentes esquemas de características socioeconómicas.

La ZM de Mexicali esta conformada únicamente por ese municipio en base al criterio de Política Urbana, con ella se exploró una variable y diferentes umbrales. Se eligió esta zona a fin de estudiar si tiene algún nivel de similitud con otras unidades contiguas.

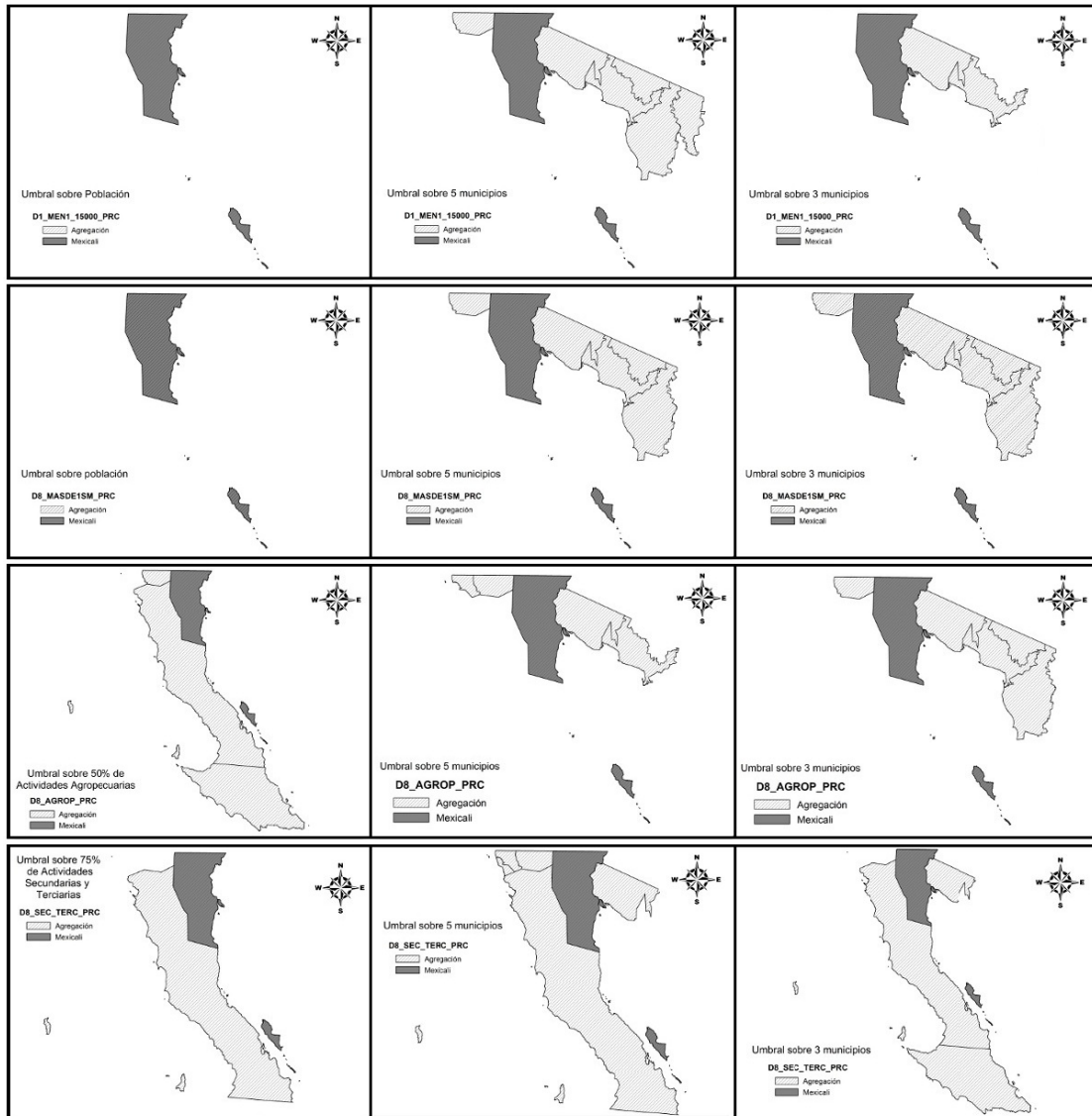


Figura 5.3: Análisis del caso particular de la ZM de Mexicali.

Elaboración propia en *ArcGis 10.0*.

Laboratorio de Análisis Socioeconómico Espacial, Spatial-SEALab.

Quando se utilizaron las variables *D1\_MEN\_15000\_PRC* y *D8\_MASDE1SM\_PRC* bajo el umbral población, el municipio (02002) conforma una agregación por sí mismo. En el esquema donde el umbral se definió sobre porcentaje de actividades agropecuarias se unió a otros tres municipios en el grupo número 226 (02001, 02002, 02003 y

03002); mientras que en el esquema de actividades secundarias y terciarias se une a otro municipio para formar la agrupación 1067 (02001 y 02002).

Al considerar el umbral de 5 unidades y las variables D1\_MEN\_15000\_PRC y D8\_MASDE1SM\_PRC, Mexicali junto a otros seis municipios formaron las agrupaciones 248 (02002, 02003, 26004, 26017, 26048, 26055 y 26070) y 50 (02002, 02003, 26017, 26048, 26055 y 26070), respectivamente. En el caso de la variable D8\_SEC\_TERC\_PRC la agregación 343 se formó con Mexicali y otros cinco municipios (02001, 02002, 02003, 02004, 02005 y 26055). Mientras que la bajo la variable de actividades agropecuarias la agrupación 323 se conformó de Mexicali y otras cuatro unidades (02002, 02003, 02004, 26048 y 26055).

El número de elementos que conforma las agrupaciones cuando el umbral es de 3 municipios varía en cada caso. Tomando la primera variable, la agregación 460 se forma por las tres unidades 02002, 26048 y 26055. Al considerar la segunda variable se tiene que la agregación 214 reúne seis municipios: 02002, 02003, 26017, 26048, 26055 y 26070. El grupo 218 de la tercera variable se conforma también de seis elementos: 02002, 02003, 26017, 26048, 26055 y 26070. Para la variable de actividades secundarias y terciarias los municipios 02001, 02002, 03002 y 26055 conforman la agregación número 640.

Resulta interesante que en siete de las doce combinaciones variable-umbrales (226, 248, 50, 343, 323, 214 y 218) Mexicali (02002) perteneció a la misma agregación que el municipio de Tecate (02003) de la ZM de Tijuana. Y es con los municipios de Sonora con los que tendió a agruparse en mayor número de ocasiones, por lo que esta zona también invita a un análisis más detallado.

El ejercicio de las zonas metropolitanas muestra que a pesar de considerar la misma composición socioeconómica el umbral elegido juega un papel importante. Se muestra una notable sensibilidad al nivel deseado de agregación, pues se presentan

---

diversos resultados al tomar en cuenta un umbral sobre alguna variable y un umbral sobre un mínimo de municipios. El umbral sobre el número de municipios a agregar debe ser un aspecto a cuidar, pues en función de ese número mínimo las agregaciones pueden tener variaciones considerables, como el caso de la ZM de Mexicali.

Estos resultados también exponen la sensibilidad de las agregaciones ante la selección de las variables composicionales. Es importante la selección de variables de la población que serán estudiadas, ya que esto en combinación con los umbrales resulta en diferentes formas de realizar las agregaciones.

El esquema de variables así como el umbral son entonces elementos que requieren de un análisis detallado. Lo anterior permitirá un mejor estudio de las agregaciones y logrará brindar mayor información sobre el comportamiento de los aspectos composicionales entre las unidades geográficas.

## CAPÍTULO 6

# CONCLUSIONES

---

En este capítulo se desarrollan las conclusiones de la revisión de literatura así como de los resultados numéricos obtenidos de la aplicación del algoritmo *MaxP*. Se responde a las preguntas de investigación planteadas a través del seguimiento de los objetivos específicos propuestos al inicio del estudio. Finalmente se exponen las posibles líneas de investigación derivadas del análisis realizado.

La literatura ha mostrado un creciente interés por comprender los fenómenos espaciales que llevan a la persistencia de diferencias entre regiones. Fenómenos como la heterogeneidad, escala y dependencia espaciales tienen un impacto significativo sobre los resultados de los estudios, lo que ha sido una motivación para promover la aplicación de técnicas y métodos adecuados con el fin de mejorar y complementar la visión actual de algunas problemáticas.

En México la cuestión de las disparidades regionales ha sido un área frecuente de estudio. La comprensión de las causas subyacentes o los efectos de reformas políticas ha sido la manera usual de abordar este tema. La literatura para el país sugiere que las diferencias regionales (en ingreso, salud, educación, entre otras) han sido persistentes e incluso se han agudizado a partir de reformas como la apertura comercial. Sin embargo, poco ha sido tratado con un enfoque espacial y ha dominado un enfoque basado solamente en criterios institucionales y de los investigadores.

El presente estudio se ha propuesto adherirse a un naciente cuerpo de literatura para el país que recurre al uso de técnicas explícitamente espaciales. Se intenta aportar conocimiento y nueva información sobre las disparidades regionales a partir del empleo del método de agregación *MaxP*. Estos métodos son una herramienta ampliamente utilizada que, con distintos fines, ha permitido un mayor entendimiento de la composición

interna de las unidades geográficas.

La revisión de estudios previos proporcionó información sobre características de la población usualmente analizadas. Con los datos recabados fue posible proponer una serie de aspectos composicionales que han sido objeto de estudio, particularmente para el caso de México, y que han sido fundamento para definir a su vez áreas tales como las rurales o metropolitanas. Elegidas aquellas características a investigar, se realizaron un conjunto de pruebas exploratorias a partir de posibles combinaciones de las variables y se contrastaron los resultados.

Se partió de diferentes combinaciones y de diferentes umbrales para el algoritmo, lo que ha permitido contestar las preguntas de investigación. Los ejercicios exploratorios permitieron observar que los resultados sí son sensibles al considerar diferentes características socioeconómicas. A pesar de que el número de agregaciones es similar, en los casos estudiados se nota que la forma en que se agregan puede ser distinta.

Finalmente los umbrales sobre mínimo de municipios a agregar sí provocó sensibilidad en los resultados. El número de agregaciones ante un umbral de 5 unidades rondaba las 400, mientras que el umbral de 3 generaba cerca de 700. Y en este sentido, unidades que en el primer caso pertenecían a la misma agregación se separaron en el segundo. Por lo anterior, se puede concluir de los resultados de la variación a escala municipal de las características socioeconómicas de la población es significativa. Sin embargo, esto es explicado en parte porque el algoritmo ofrece el máximo número de agregaciones, reafirmando la importancia de la elección del umbral antes de ser aplicado el método.

La información obtenida de este tipo de ejercicios exploratorios permite arrojar una primera luz sobre un determinado fenómeno, ya que proporciona una visión más amplia de las variables que puedan estar influyendo en el proceso de estudio. La presente

investigación seleccionó un conjunto de variables, las cuales fueron objeto de un análisis exploratorio, con el fin de brindar información relevante sobre su comportamiento y que pudiera además servir a investigaciones futuras sobre disparidades regionales en México.

Este trabajo tiene aún varias extensiones que desarrollar. Entre las cuales se encuentra el realizar una selección más precisa de las variables que resultan relevantes en la explicación de las diferencias en las características de la población. Además, podría mostrarse un análisis a mayor detalle de las combinaciones que pueden resultar de cada composición de variables y la sensibilidad ante estos cambios.

Así, algunas líneas de investigación futura pueden desarrollarse alrededor del análisis particular de ciertas variables, como ingreso, educación, mortalidad, entre otras. Una segunda línea puede ser la extensión del trabajo actual a partir de la incorporación de diferentes métodos de agregación espacial para permitir el contraste de resultados. Y una tercera línea es el tratamiento del factor escala en la aplicación de métodos de agregación, ya que la consideración de diferentes niveles espaciales conduce al refinamiento de los análisis.

## APÉNDICE A

### ANEXOS METODOLÓGICOS

---

#### A.1 ALGORITMO *MaxP*

Siguiendo a Duque *et al.* (2011) se describe a continuación el planteamiento del problema tras el algoritmo *MaxP*. Sea  $A = \{A_1, A_2, \dots, A_n\}$ , que denota el conjunto de  $n = |A|$  áreas. Además sea  $A_{iy}$  que denota el atributo  $y$  del área  $A_i$ , donde  $y \in Y = \{1, 2, \dots, m\}$  con  $m \geq 1$ ; y  $l_i$  denota el atributo espacialmente extensivo del área  $A_i$ .

Sea  $d : A \times A \rightarrow \mathbb{R}^+ \cup \{0\}$  la disimilitud entre áreas basada en el conjunto de atributos  $Y$ , tal que  $d_{ij} \equiv d(A_i, A_j)$  donde  $d_{ij} \geq 0$ ,  $d_{ij} = d_{ji}$  y  $d_{ij} = 0$  si  $i = j = 1, 2, \dots, n$ .<sup>1</sup>

Sea  $W = (V, E)$  que denota el grafo de contigüidad asociado a  $A$ , tal que los vértices  $v_i \in V$  corresponden a las áreas  $A_i \in A$ , y las aristas  $\{v_i, v_j\} \in E$  si y sólo si las áreas  $A_i$  y  $A_j$  comparten una arista. Debe obtenerse un grafo conectado para el modelo  $W$  cuando se utiliza el algoritmo *MaxP*.

Sea  $P_p = \{R_1, R_2, \dots, R_p\}$  que denote una partición de áreas de  $A$  en  $p$  regiones, con  $1 \leq p \leq n$ , tal que:

$$\begin{aligned} &|R_k| > 0 \text{ para } k = 1, 2, \dots, p; \\ &R_k \cap R_{k'} = \emptyset \text{ para } k, k' = 1, 2, \dots, p \wedge k \neq k'; \\ &\cup_{k=1}^p R_k = A; \\ &\sum_{A_i \in R_k} l_i \geq \text{umbral} \begin{cases} \text{para } k = 1, 2, \dots, p, \text{ y} \\ \text{umbral} \in \mathbb{R}^+ \cup \{0\} | 0 \leq \text{umbral} \leq \sum_{A_i \in A} l_i; \end{cases} \\ &W(R_k) \text{ está conectado para } k = 1, 2, \dots, p. \end{aligned}$$

Sea  $\Pi$  el conjunto que denote todas las particiones posibles de  $A$ . Se tienen los siguientes criterios de evaluación para la posible partición  $P_p \in \Pi$ :

$$\begin{aligned} h(R_k) &= \sum_{\substack{ij:A_i, A_j \in R_k, i \leq j \\ p}} d_{ij} \text{ Heterogeneidad de la región } k \text{ con } R_k \in P_p; \\ H(P_p) &= \sum_{k=1}^p h(R_k) \text{ Heterogeneidad total de la partición } P_p \in \Pi. \end{aligned}$$

---

<sup>1</sup>También puede utilizarse la función distancia, que satisface la subaditividad o desigualdad del triángulo, i.e.,  $d_{ij} \leq d_{ik} + d_{kj}$ , con  $i, j, k = 1, \dots, n$ .

Entonces el problema de las regiones *MaxP* se formula así:

$$\begin{aligned} \text{Determine } P_p^* \text{ tal que } |P_p^*| &= \max(|P_p| : P_p \in \Pi), \text{ y} \\ \nexists P_p \in \Pi : |P_p| &= |P_p^*| \wedge H(P_p) < H(P_p^*) \end{aligned}$$

Para la descripción más detallada del problema de optimización, así como de la metaheurística propuesta para su resolución, consulte el documento de Duque *et al.* (2011).

## A.2 VARIABLES UTILIZADAS EN ESTUDIOS PREVIOS

A continuación se describen algunas de las variables utilizadas en estudios previos tanto de áreas rurales como de zonas metropolitanas. Se han agrupado en dimensiones para permitir una mejor comprensión de las características que han sido objeto de análisis.

### VARIABLES PARA ÁREAS RURALES

Dada la dificultad de definir a las áreas rurales, éstas usualmente quedan en función de la población que vive en ellas. Para el caso de México se han establecido umbrales de 2,500 y 15,000 habitantes, según las definiciones de INEGI y CONEVAL, respectivamente. A continuación se presentan algunas de las variables utilizadas a partir de información censal en estudios sobre áreas rurales mexicanas:

### VARIABLES PARA ZONAS METROPOLITANAS

Estudios previos sobre Zonas Metropolitanas realizados para el caso de México utilizan variables censales, entre las más empleadas se encuentran las siguientes:

---

<sup>2</sup>Población Profesionalista - Técnico / Trabajador en la industria / Comerciante - Trabajador en servicios diversos / Actividades Agropecuarias

<sup>3</sup>Estos datos pertenecen a una encuesta realizada por la Secretaría de la Reforma Agraria de México y el Banco Mundial (Economic Adjustment and Institutional Reforms: Mexico's Ejido Sector Responds, 1998).

<sup>4</sup>Elaborado por CONAPO a partir de información censal.

<sup>5</sup>Elaborado por el Programa de las Naciones Unidas para el Desarrollo (PNUD).

Dimensión	Variable
Población	-Número de habitantes
Características del Hogar	-Género y edad del jefe de hogar
Migración	-Migración reciente dentro del país y hacia EU
Población indígena	-Población indígena
Educación	-Años de Educación
Características Económicas	-PEA Desocupada y Ocupada en los diferentes sectores de actividad <sup>2</sup> -Ingresos agrícolas, no agrícolas, por cuenta propia y por remesas <sup>3</sup>
Características de la Vivienda	-Disponibilidad de energía eléctrica, drenaje, agua potable y materiales de la vivienda
Indicadores Socioeconómicos	-Índice de marginación <sup>4</sup> -Índice de Desarrollo Humano <sup>5</sup>

Tabla A.1: Variables analizadas en estudios sobre áreas rurales en México.

Fuente: Elaboración propia.

Dimensión	Variable
Población	-Población por municipio
Fecundidad	-Promedio de hijos nacidos vivos
Migración	-Migración reciente dentro del país
Educación	-Población con grado escolar completo en Primaria, Secundaria, Media Superior, Superior -Sin instrucción -Grado Escolar Promedio -Población Alfabeta y Analfabeta
Características Económicas	-Población económicamente activa y no económicamente activa -PEA Desocupada y Ocupada en los diferentes sectores de actividad -Ingreso hasta 1 salario mínimo mensual (sm) / de 1 a 2 sm / más de 2 sm
Acceso a Servicios de Salud	-Población con acceso a servicios de salud
Características de la Vivienda	-Disponibilidad de energía eléctrica, drenaje, agua potable y materiales de la vivienda -Disponibilidad de bienes y tecnologías de la información (televisor, radio, lavadora, teléfono, computadora, etc.)
Indicadores Socioeconómicos	Índice de marginación

Tabla A.2: Variables analizadas en estudios sobre zonas metropolitanas en México.

Fuente: Elaboración propia.

# BIBLIOGRAFÍA

---

- AERTS, J. y HEUVELINK, G. (2002). Using simulated annealing for resource allocation. *International Journal of Geographical Information Science*, 16, pp. 571–587.
- ANSELIN, L. (1988). *Spatial econometrics: methods and models*. Springer.
- ARIZA, M. y SOLÍS, P. (2009). Dinámica socioeconómica y segregación espacial en tres áreas metropolitanas de México, 1990 y 2000. *Estudios Sociológicos*, 27, pp. 171–209.
- ARMSTRONG, M.; LOLONIS, P. y HONEY, R. (1993). A spatial decision support system for school redistricting. *URISA Journal*, pp. 40–52.
- AROCA, P.; BOSCH, M. y MALONEY, W. (2005). Spatial dimensions of trade liberalization and economic convergence: Mexico 1985-2002. *The World Bank Economic Review*, 19, pp. 345–378.
- ASHLEY, C. y MAXWELL, S. (2001). Rethinking rural development. *Development Policy Review*, 19, pp. 395–425.
- ASSUNÇÃO, R.; CORRÊA, M.; CÂMARA, G. y FREITAS, C. (2006). Efficient regionalization techniques for socio-economic geographical units using minimum spanning trees. *International Journal of Geographical Information Science*, 20, pp. 797–811.
- ASUAD, N.; QUINTANA, L. y RAMÍREZ, R. (2007). Convergencia espacial y concentración regional agrícola en México, 1970-2003. *Problemas del Desarrollo*, 38, pp. 80–111.
- BAÇÃO, F.; LOBO, V. y PAINHO, M. (2005). Applying genetic algorithms to zone design. *Soft Computing*, 9, pp. 341–348.
- BARRO, R. y SALA-I MARTIN, X. (1992). Convergence. *Journal of Political Economy*, 100, pp. 223–251.
- BERGEY, P.; RAGSDALE, C. y HOSKOTE, M. (2003). A simulated annealing genetic algorithm for the electrical power districting problem. *Annals of Operations Research*, 121, pp. 33–55.
- BERGOLO, M. y CARBAJAL, F. (2010). Exploring the urban-rural labor income gap in Uruguay: A quantile regression decomposition. *Revista de Análisis Económico*, 25, pp. 133–168.
- BOUILLON, C.; LEGOVINI, A. y LUSTIG, N. (2003). Rising inequality in Mexico: household characteristics and regional effects. *Journal of Development Studies*, 39, pp. 112–133.

- BOZKAYA, B.; ERKUT, E. y LAPORTE, G. (2003). A tabu search heuristic and adaptive memory procedure for political districting. *European Journal of Operational Research*, 144, pp. 12–26.
- BRODY, H.; RIP, P.; VINTEN-JOHANSEN, P.; PANETH, N. y RACHMAN, S. (2000). Map-making and myth-making in Broad Street: The London cholera epidemic, 1854. *The Lancet*, 356, pp. 64–68.
- BUSTOS, N. (1998). El ordenamiento y planificación territorial en Chile: elementos para su discusión. *Revista de Geografía Norte Grande*, 25, pp. 49–53.
- CHIQUIAR, D. (2005). Why Mexico's regional income convergence broke down? *Journal of Development Economics*, 77, pp. 257–275.
- COBURN, A.; MACKINNEY, A.; MCBRIDE, T.; MUELLER, K.; SLIFKIN, R. y WAKEFIELD, M. (2007). Choosing rural definitions: Implications for health policy. *Rural Policy Research Institute Health Panel*.
- COSSMAN, R.; COSSMAN, J.; COSBY, A. y REAVIS, R. (2008). Reconsidering the rural-urban continuum in Rural Health Research: A test of stable relationships using mortality as a health measure. *Population Research and Policy Review*, 27, pp. 459–476.
- COVA, T. y CHURCH, R. (2000). Contiguity constraints for single-region site search problems. *Geographical Analysis*, 32, pp. 306–329.
- D'AMICO, S.; WANG, S.; BATTÀ, R. y RUMP, C. (2002). A simulated annealing approach to police districting design. *Computer and Operations Research*, 29, pp. 667–684.
- DAVIS, B.; WINTERS, P.; CARLETTO, G.; COVARRUBIAS, K.; QUIÑONES, E.; ZEZZA, A.; STAMOULIS, K.; AZZARI, C. y DIGIUSEPPE, S. (2010). A cross-country comparison of rural income generating activities. *World Development*, 38, pp. 48–63.
- DE JANVRY, A. y SADOULET, E. (2001). Income strategies among rural households in Mexico: The role of off-farm activities. *World Development*, 29, pp. 467–480.
- DE POEL, E.; O'DONNELL, O. y VAN DOORSLAER, E. (2009). What explains the rural-urban gap in infant mortality: Household or community characteristics? *Demography*, 46, pp. 827–850.
- DELGADILLO, J. y TORRES, F. (2006). Dimensiones multicausales del Ordenamiento Territorial. Enfoques y Aplicaciones. En: *Política Territorial en México. Hacia un modelo de desarrollo basado en el territorio*, SEDESOL-UNAM.

- DU PLESSIS, V.; BESHIRI, R. y BOLLMAN, R. (2002). Definitions of “rural”. *Statistics Canada Agriculture and Rural, Working Paper*.
- DUQUE, J.; ANSELIN, L. y REY, S. (2011). The max-p regions problem. *Journal of Regional Science*, 20, pp. 1–23.
- DUQUE, J.; RAMOS, R. y SURINACH, J. (2007). Supervised regionalization methods: A survey. *International Regional Science Review*, 30, pp. 195–220.
- ESQUIVEL, G. (1999). Convergencia regional en México, 1940-1995. *El Trimestre Económico*, LXVI, pp. 725–761.
- FALLAH, B. y PARTRIDGE, M. (2007). The elusive inequality-economic growth relationship: are there differences between cities and the countryside? *Annals of Regional Science*, 41, pp. 375–400.
- FERREIRA, H. (2005). *Construir las regiones. Por una aproximación regional a la formulación y gestión de políticas públicas en México*. Coedición Centro Lindavista, Integración Editorial.
- FREEDMAN, D. (2001). Ecological inference and the ecological fallacy. En: *International Encyclopedia for the Social and Behavioral Sciences*, Elsevier Amsterdam.
- GARCIA-LÓPEZ, M. y MUÑIZ, I. (2010). Employment decentralisation: Polycentricity or scatteration? The case of Barcelona. *Urban Studies*, 47, pp. 3035–3056.
- GAUVIN, L.; ROBITAILLE, E.; RIVA, M.; MCLAREN, L.; DASSA, C. y POTVIN, L. (2007). Conceptualizing and operationalizing neighbourhoods. The conundrum of identifying territorial units. *Canadian Journal Of Public Health Revue Canadienne De Sante Publique*, 98, pp. S18–S26.
- GLAESER, E. (2011). *Triumph of the city: how our greatest invention makes us richer, smarter, greener, healthier, and happier*. Penguin Press.
- GOTWAY, C. y YOUNG, L. (2002). Combining Incompatible Spatial Data. *Journal of the American Statistical Association*., 97, pp. 632–648.
- HANSEN, P.; JAUMARD, B.; MEYER, C.; SIMEONE, B. y DORING, V. (2003). Maximum split clustering under connectivity constraints. *Journal of Classification*, 20, pp. 143–180.
- HANSON, G. (1998). Regional adjustment to trade liberalization. *Regional Science and Urban Economics*, 23, pp. 419–444.
- HANSON, G. (2003). What happened to wages since NAFTA? Implications for hemispheric free trade. *NBER Working Paper*, 9563.

- HART, L.; LARSON, E. y LISHNER, D. (2005). Rural definitions for health policy and research. *American Journal of Public Health*, 95, pp. 1149–1155.
- HOJATI, M. (1996). Optimal Political Districting. *Computers and Operations Research*, 23, pp. 1147–1161.
- INEGI (2011). *Censo de Población y Vivienda 2010. Marco Conceptual*. INEGI.
- JELINSKI, D. y WU, J. (1996). The modifiable areal unit problem and implications for landscape ecology. *Landscape Ecology*, 11, pp. 129–140.
- KALCSICS, J.; NICKEL, S. y SCHRÖDER, M. (2005). Towards a unified territorial design approach: Applications, algorithms and GIS integration. *TOP: An Official Journal of the Spanish Society of Statistics and Operations Research*, 13, pp. 1–56.
- LANKFORD, P. (1969). Regionalization: Theory and alternative algorithms. *Geographical Analysis*, 1, pp. 196–212.
- LEE, B. (2007). “Edge” or “edgeless” cities? Urban spatial structure in U.S. Metropolitan Areas, 1980 to 2000. *Journal of Regional Science*, 74, pp. 479–515.
- LIN, J. y CHEN, B. (2011). Urbanization and urban-rural inequality in China: A new perspective from the government’s development strategy . *Frontiers of Economics in China*, 6, pp. 1–21.
- LUO, S.; JOHNSTON, A. y CHEN, D. (2008). Approach to delimiting metropolitan regions’ boundary and grading urban hierarchy within a metropolitan region - A case study of Shanghai metropolitan region. *Chinese Geographical Science*, 18, pp. 197–205.
- MACHADO, E. (2001). Urban spatial segregation and social differentiation: foundation for a typological analysis. *Lincoln Institute of Land Policy. Conference Paper*.
- MACÍAS, M.; GUZMÁN, G. y ANDRADE, M. (2009). Distribución territorial de los índices de marginación en la zona metropolitana de Guadalajara (en línea). <http://sincronia.cucsh.udg.mx/maciashuertafall09.htm> (Consultado el 24 de noviembre de 2011).
- MACMILLAN, W. (2001). Redistricting in a GIS environment: An optimisation algorithm using switching-points. *Journal of Geographical Systems*, 3, pp. 167–180.
- MASSIRIS, A. (2002). Ordenación del territorio en América Latina. *Scripta Nova, Revista Electrónica de Geografía y Sociales, Universidad de Barcelona*, 6, pp. 741–798.

- MONTES, P. (2001). *El ordenamiento territorial como opción de políticas urbanas y regionales en América Latina y el Caribe*. Series CEPAL.
- MONTES, V. y ORTEGA, E. (2003). Ubicación de la Marginación en el Área Metropolitana de Monterrey. En: *Memorias Congreso de Investigación y Extensión XXXIII*, México: Instituto Tecnológico y de Estudios Superiores de Monterrey.
- MORENO, P. y GARCÍA, J. (2011). Estado del arte en procesos de zonificación. *Geo-Focus*, 11, pp. 151–181.
- MORGENSTERN, H. (1982). Uses of ecologic analysis in epidemiologic research. *American Journal of Public Health*, 72, pp. 1336–1344.
- MURRAY, A. y SHYY, T. (2000). Integrating attribute and space characteristics in choropleth display and spatial data mining. *International Journal of Geographical Information Science*, 14, pp. 649–667.
- NEGRETE, M. y SALAZAR, H. (1986). Zonas Metropolitanas en México, 1980. *Estudios Demográficos y Urbanos*, 1, pp. 97–124.
- OCDE (2006). The State of Rural Regions. En: *The New Rural Paradigm. Policies and Governance*, OCDE.
- OCDE (2007). Perfil del México Rural. En: *Estudios de Política Rural*. México, OCDE.
- OPENSHAW, S. (1977). A geographical solution to scale and aggregation problems in region-building, partitioning and spatial modeling. *Transactions of the Institute of British Geographers*, 2, pp. 459–472.
- OPENSHAW, S. (1984). The Modifiable Areal Unit Problem. Concepts and Techniques. En: *Modern Geography 38*, Norwich Geo Books.
- OPENSHAW, S. y RAO, L. (1995). Algorithms for reengineering 1991 census geography. *Environment and Planning*, 27, pp. 425–446.
- PASTERNAK, S. y MACHADO, L. (2008). Dinámica espacial de la desigualdad. *Ciudades*, 20, pp. 11–20.
- PATACCHINI, E. y ZENOU, Y. (2009). Urban Sprawl in Europe. *Brookings-Wharton Papers on Urban Affairs*, pp. 125–149.
- PÉREZ, E. y SANTOS, C. (2011). Diferenciación socioespacial en la Zona Metropolitana de la Ciudad de México. *Investigaciones Geográficas, Boletín del Instituto de Geografía, UNAM*, (74), pp. 92–106.

- PRADHAN, B.; ROY, P.; SALUJA, M. y VENKATRAM, S. (2000). Rural-urban disparities: Income distribution, expenditure pattern and social sector. *Economic and Political Weekly*, 35, pp. 2527–2539.
- PUDERER, H. (2008). Defining and measuring Metropolitan Areas: A comparison between Canada and the United States. *Statistics Canada Catalogue. Geography Working Paper Series*.
- PUJADAS, R. y FONT, J. (1998). *Ordenación y planificación territorial*. Madrid, Síntesis.
- QUAH, D. (1993). Empirical cross-section dynamics in economic growth. *European Economic Review*, 37, pp. 427–443.
- QUAH, D. (1996). Regional convergence clusters across Europe. *European Economic Review*, 40, pp. 951–958.
- REARDON, T.; BERDEGUÉ, J. y ESCOBAR, G. (2001). Rural nonfarm employment and incomes in Latin America: Overview and Policy Implications. *World Development*, 29, pp. 395–409.
- REY, S. (2004). Spatial analysis of regional income inequality. En: M. Goodchild y D. Janelle (Eds.), *Spatially integrated social science*, Oxford: Oxford University Press.
- REY, S.; ANSELIN, L.; FOLCH, D.; ARRIBAS-BEL, D.; SASTRÉ-GUTIÉRREZ, M. y INTERLANTE, L. (2011). Measuring spatial dynamics in metropolitan areas. *Economic Development Quarterly*, 25, pp. 54–64.
- REY, S. y FOLCH, D. (2011). Impact of spatial effects on income segregation indices. *Computers, Environment and Urban Systems*, 35, pp. 431–441.
- REY, S. y JANIKAS, M. (2005). Regional convergence, inequality, and space. *Journal of Economic Geography*, 5, pp. 155–176.
- REY, S. y SASTRÉ-GUTIÉRREZ, M. (2010). Interregional inequality dynamics in Mexico. *Spatial Economic Analysis*, 5, pp. 277–298.
- RÍOS-MERCADO, R. y FERNÁNDEZ, E. (2009). A reactive GRASP for a commercial territory design problem with multiple balancing requirements. *Computers and Operations Research*, 36, pp. 755–776.
- ROBINSON, W. (1950). Ecological correlations and the behavior of individuals. *American Sociological Review*, 15, pp. 351–357.

- ROMERO, H. y VÁSQUEZ, A. (2005). Pertinencia y significado del ordenamiento territorial en Chile. *Urbano*, 8, pp. 91–99.
- RUBALCAVA, R. y SCHTEINGART, M. (1985). Diferenciación socio-espacial intraurbana en el área metropolitana de la ciudad de México. *Estudios Sociológicos*, 3, pp. 481–514.
- SASTRÉ-GUTIÉRREZ, M. y REY, S. (2008). Polarización espacial y dinámicas de la desigualdad interregional en México. *Problemas del Desarrollo*, 39, pp. 181–204.
- SEDESOL (2007). Programa sectorial de Desarrollo Social 2007-2012.
- SEDESOL (2010). Diagnóstico: Alternativas de la población rural en pobreza para generar ingresos sostenibles.
- SICULAR, T.; XIMING, Y.; GUSTAFSSON, B. y SHI, L. (2007). The urban-rural income gap and inequality in China. *Review of Income and Wealth*, 53, pp. 93–126.
- SPELLERBERG, A.; HUSCHKA, D. y HABICH, R. (2007). Quality of life in rural areas: Processes of Divergence and Convergence. *Social Indicators Research*, 83, pp. 283–307.
- SUÁREZ, M. y DELGADO, J. (2009). Is Mexico City polycentric? A trip attraction capacity approach. *Urban Studies*, 46, pp. 2187–2211.
- TIEDE, D. y STROBL, J. (2006). Polygon-based regionalisation in a GIS environment. En: E. Buhmann; S. Ervin; I. Jorgensen y J. Strobl (Eds.), *Trends in Knowledge-Based Landscape Modeling*, Heidelberg Wichmann-Verlag.
- VOß, S. (2001). Metaheuristics: The state of the art. *Lecture Notes in Computer Science*, 2148, pp. 1–43.
- WEST, A.; LEE, R.; SHAMBAUGH-MILLER, M.; BAIR, B.; MUELLER, K.; LILLY, R.; KABOLI, P. y HAWTHORNE, K. (2010). Defining “rural” for veterans’ health care planning. *The Journal of Rural Health*, 26, pp. 301–309.
- WONG-GONZÁLEZ, P. (2010). El nuevo “problema regional”. *Ciudades*, 21, pp. 9–12.
- ZAMUDIO, F.; CORONA, A. y LÓPEZ, I. (2008). Un índice de ruralidad para México. *Espiral, Estudios sobre Estado y Sociedad*, 14, pp. 179–214.
- ZENTELLA, J. (2007). ¿Cómo gobernar las zonas metropolitanas en México? *Este País*, (194), pp. 30–37.
- ZIMMER, Z.; KANEDA, T. y SPESS, L. (2007). An examination of urban versus rural mortality in China using community and individual data. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 62, pp. S349–S357.

- 
- ZOLTERNIS, A. y SINHA, P. (1983). Sales territory alignment: A review and model. *Management Science*, 29, pp. 1237–1256.