

*APLICACIÓN DE SISTEMAS DE INFORMACIÓN GEOGRÁFICA
Y TÉCNICAS DE AUTO-CORRELACIÓN ESPACIAL PARA ANALIZAR
LA DEMANDA POR EMPLEO INDUSTRIAL EN LOS 81 CANTONES
DE COSTA RICA EN 2011*

Rafael Arias Ramírez¹
Leonardo Sánchez Hernández²

Recibido: 17/06/2014

Aprobado: 13/10/2014

RESUMEN

Este artículo trata de demostrar que existe una elevada concentración del empleo industrial en algunas áreas geográficas de Costa Rica. Para ello se dispone de una base de datos novedosa y completa derivada del Censo de Población y Vivienda del año 2011, el cual permite crear una matriz origen-destino de empleo y con ello tener por primera vez para el país información sobre la demanda por empleo industrial que genera cada cantón. El uso de índices de dependencia espacial, de Sistemas de Información Geográfica y coeficientes de localización industrial permiten demostrar la existencia de conglomerados de empleo industrial en algunos cantones de Costa Rica.

PALABRAS CLAVES: SISTEMAS DE INFORMACIÓN GEOGRÁFICA, EMPLEO INDUSTRIAL, CORRELACIÓN ESPACIAL, INDUSTRIAL CONGLOMERADOS, COEFICIENTES DE LOCALIZACIÓN INDUSTRIAL.

ABSTRACT

This article is aimed to show high concentration of industrial employment in some geographic areas of Costa Rica. In order to do that we use a very new and complete data set from the Census of Population and Housing of the year 2011, which allows us to create an origin-destination matrix of employment and to have, for the first time in the country, information on industrial employment generated by every county. The use of the spatial dependency index technique, the geographic information system and the coefficient of industrial localization allowed us to demonstrate the existence of strong correlation between industrial employment and the concentration of industrial activities at the county level.

KEY WORDS: GEOGRAPHIC INFORMATION SYSTEMS, INDUSTRIAL EMPLOYMENT, SPATIAL CORRELATION, INDUSTRIAL CONGLOMERATES, COEFFICIENTS OF INDUSTRIAL LOCALIZATION.

1 Universidad de Costa Rica, Código Postal 11501-2060, Costa Rica; rarias@una.ac.cr

2 Universidad de Costa Rica, Código Postal 11501-2060, Costa Rica; joseleonardo.sanchez@ucr.ac.cr

I. INTRODUCCIÓN

La demanda por empleo industrial, su localización y concentración a lo largo de la geografía cantonal costarricense ha sido muy poco explorada, estudios de Arias y Sánchez (2011, 2011 y 2012) así como de Trejos (2002) y Sánchez, Murillo y Arguedas (2008) han tratado de aproximar el tema a niveles de desagregaciones mayores o analizando regiones específicas. El instrumento idóneo para analizar la mayor o menor uniformidad en la distribución de las actividades en el espacio son los distintos índices de concentración geográfica disponibles en la literatura relacionada con la economía urbana y regional. Este artículo trata de demostrar la hipótesis de que la demanda por empleo industrial muestra niveles de concentración muy dispares a lo largo de la geografía costarricense. Asimismo, se demostrará que el tipo de empleo manufacturero, presenta desigualdades en su distribución en el territorio.

La literatura económica señala diversos factores explicativos de la concentración geográfica en la demanda por empleo industrial. Entre ellos está el costo de los factores productivos, así como el tamaño de mercado o la dotación de infraestructura, recursos humanos, regulaciones territoriales, entre otros; pero también se señala las economías de aglomeración. De acuerdo con la literatura, este último factor actúa en ámbitos geográficos locales, motivo por el cual parece razonable elaborar un análisis detallado de los patrones de localización y concentración a este nivel de desagregación. De lo contrario, los resultados obtenidos no pueden considerarse del todo satisfactorios.

Las investigaciones aplicadas a escala internacional han incorporado las unidades geográficas de menor tamaño en los análisis empíricos. Sin embargo, para el caso de Costa Rica hasta el momento únicamente se han llevado a cabo aplicaciones empíricas que utilizaban la provincia, las regiones, subregiones y cantones como unidades de medida pero sin contemplar un análisis comprehensivo de todos los cantones del país.

Para este estudio, tal como se explica en la sección sobre la metodología aplicada,

se dispone de una base de datos cantonal que permite el desarrollo de un análisis mucho más preciso de la distribución geográfica de la demanda por empleo industrial. Para ello, se aplican un conjunto de indicadores de concentración e índices de auto-correlación espacial contemplados en la técnica de la econometría espacial. Estos índices incorporan las áreas vecinas al cantón en el cálculo del nivel de concentración geográfica de las actividades generadoras de empleo industrial y, por tanto, permiten contrastar si al analizar la distribución espacial de la actividad, así como sus determinantes, debe considerarse el entorno geográfico inmediato al cantón.

El artículo se encuentra estructurado de la siguiente manera. En la primera sección se esbozan algunos elementos teóricos sobre la localización y aglomeración del empleo industrial así como instrumentos metodológicos aplicados para analizarlos. La segunda sección muestra la metodología a utilizar, los datos y la zona de estudio. Los resultados y la discusión de los mismos se desarrollan en la tercera sección. Por último, la cuarta sección presenta una serie de consideraciones finales.

II. MARCO TEÓRICO

Algunos elementos teóricos sobre la concentración espacial de actividades económicas

Krugman, P (1999) y Glaeser, E (1992) plantean la importancia de las interacciones espaciales en los procesos económicos y como estos elementos han contribuido a trazar las propuestas teóricas de la nueva economía geográfica. De acuerdo con Anselin, L (1999) en las teorías de economía espacial, la atención se centra en modelos que capturan la interacción directa entre los agentes, plasmada en normas sociales, efectos por vecindad como la imitación, spillovers, externalidades y contagio, las cuales arrojan, al mismo tiempo, luz acerca de cómo las interacciones individuales pueden dar lugar de forma agregada a comportamientos colectivos donde son reconocibles patrones comunes. Por su

parte, Gutiérrez P, y J.-Gould, M (1994) explican que para lograr todos los objetivos trazados en esta nueva área de la economía se han tenido que desarrollar nuevas metodologías de análisis, lo que ha llevado a cambios en los modelos econométricos convencionales. Para estos autores, este hecho, hace que surja la nueva Econometría Espacial, que fija al menos cuatro ejes de interés: *a) la evaluación del efecto de la auto-correlación espacial en los modelos econométricos, b) la estimación de modelos que incorporen su efecto, c) la creación de test y herramientas de diagnóstico para reconocer su presencia, y c) la predicción espacial.*

En Costa Rica la concentración de la demanda de empleo industrial a lo largo de la geografía costarricense ha sido poco estudiada y los pocos estudios que existen parten de análisis de unidades territoriales muy agregadas, perdiéndose la perspectiva del comportamiento de las variables a niveles más desagregados territorialmente. De igual forma, los análisis realizados en zonas específicas describen la situación de un área geográfica, sea el distrito, el cantón, la provincia o la región, sin tener en cuenta la localización espacial; es decir, tratando las unidades territoriales como unidades aisladas sin ningún tipo de conexión con las áreas vecinas. En este sentido el Instituto de Investigaciones en Ciencias Económicas de la Universidad de Costa Rica (IICE-UCR) ha venido desarrollando y aplicando una serie de metodologías e instrumentos para cubrir este vacío, entre ellos están los trabajos de Arias y Sánchez (2011, 2012 y 2013) así como los de Trejos (2002) y otros estudios de investigadores de la Universidad de Costa Rica (Pujol, Pérez y Sánchez, 2012 y 2013).

Como se mencionó anteriormente, dentro de los avances más importantes en los últimos años en la ciencia económica se encuentra la reincorporación explícita del efecto del espacio geográfico en el análisis de los problemas económicos. En este sentido, los trabajos de Krugman (1991a y 1991b) sobre lo que se ha llamado la “nueva geografía económica”, resaltan el papel de las externalidades espaciales en los modelos que estudian la influencia del espacio sobre la localización de empresas, el desarrollo

de complejos industriales, los clusters y la difusión del conocimiento y la tecnología.

En este mismo sentido, otro grupo de economistas entre quienes destacan Anselin (1988,1992), Florax, (1995) y Rey (1997, 1999), han desarrollado un conjunto de técnicas para trabajar con datos geo-referenciados y estimar modelos que incorporan explícitamente la dimensión espacial. Este conjunto de técnicas que se utilizaban principalmente en economía regional y urbana, está abarcando cada vez más espacios y es fácil encontrar aplicaciones en las principales revistas científicas de economía general (ver por ejemplo Case, 1991 y Pinsky and Slade, 1998; entre otros).

Los desarrollos metodológicos mencionados anteriormente, utilizan los principales modelos de economía regional y urbana y los combinan con la estadística espacial, lo cual permite tener un análisis descriptivo e inferencial de datos geográficos. En este sentido, el trabajo de Ripley (1981), Cressie (1991), Fotheringham et al (2000) y el más reciente de Haining (2003), introduce y generaliza para diversas disciplinas el término Estadística Espacial. Estos avances metodológicos no solo son importantes para aplicar técnicas estadísticas a datos geográficos, sino que introducen el espacio como elemento fundamental del análisis económico.

Un concepto básico que se desprende del desarrollo de estas nuevas metodologías, que integran el análisis espacial con el análisis económico, es el de dependencia o autocorrelación espacial, el cual analiza la falta de independencia que se produce entre las observaciones de una variable para sus diferentes localizaciones. Es un punto donde la estadística espacial se conecta con la geografía en la línea de los trabajos de Tobler (1979) y su “primera ley de la geografía” en la que se afirma que en el análisis geográfico todo está relacionado con todo, pero las cosas cercanas están más relacionadas entre sí que las cosas lejanas.

Los primeros índices formales para detectar la presencia de autocorrelación espacial se deben a Moran (1948) y Geary (1954). La aplicación de este concepto en economía regional junto a nuevos desarrollos matemáticos se

formalizaron inicialmente con los trabajos de los geógrafos Cliff y Ord (1972, 1973, 1981) y posteriormente con los trabajos y metodologías desarrolladas por Anselin (1988), Anselin y Florax (1995), Anselin et al (2004), Getis et al (2004) y Lesage et al (2004).

Actualmente, una de las pruebas más ampliamente utilizadas para la detección de correlación espacial es aquella basada en la prueba estadística I de Moran. En esencia, esta prueba está formulada como una forma cuadrática de las variables a las que se les está midiendo la correlación espacial. Su definición original estandariza las variables restándoles la media muestral y reduciéndolas a través de la aplicación de un factor adecuado. Los trabajos de Cliff y Ord (1972, 1973, 1981), generalizaron el estadístico de prueba I de Moran con el objetivo de derivar una prueba para correlación espacial que funcionara en modelos de regresión lineal. La generalización que desarrollaron está formulada en términos de una forma cuadrática de los residuos estimados e incluye la fórmula original del índice como un caso especial correspondiente a un error donde el único regresor se relaciona con el intercepto. De esta manera, suponiendo que las innovaciones son independientes e idénticamente distribuidas siguiendo una normal, Cliff y Ord derivaron la distribución del índice de prueba I de Moran para muestras grandes al igual que los momentos para muestras pequeñas.

La prueba mencionada anteriormente permite indagar sobre las asociaciones existentes en una zona a nivel global. Sin embargo, en muchas ocasiones la asociación significativa puede no darse en toda el área que se está analizando, sino sólo en determinadas zonas. Para ello Anselin (1995) propuso una metodología bajo la cual se estiman los denominados indicadores locales de asociación espacial o indicadores LISA (Local Indicator of Spatial Association, por sus siglas en inglés). El indicador LISA es una transformación de la I de Moran, que mide para un conjunto de datos el nivel de influencia entre datos cercanos. La diferencia es que la I de Moran tradicional (global) evalúa el nivel de dependencia en toda la base de datos en tanto que el indicador LISA evalúa localmente:

para cada punto con valor conocido, se evalúa el grado de concordancia o discordancia con los valores más cercanos. Decimos que la I de Moran asume homogeneidad en los datos (que la distribución en el espacio de la demanda por empleo industrial, es similar para toda el área) en tanto que el indicador LISA permite identificar zonas heterogéneas.

III. METODOLOGÍA

La metodología utilizada en este trabajo para analizar la concentración espacial de la demanda por empleo industrial en los 81 cantones de Costa Rica en 2011, se realizará a través de la utilización de sistemas de información geográfica (SIG), los cuales consisten en una integración de datos estadísticos geográficamente referenciados. En este estudio, estos sistemas de información serán procesados mediante técnicas estadísticas espaciales. La metodología utilizada en el análisis espacial propuesto en este artículo consiste en la evaluación de la aglomeración de las unidades geográficas de observación (cantones) en torno a los valores exhibidos por un atributo de interés (empleo industrial), entendido éste como una variable que describe características específicas de estas unidades.

Para tal efecto se analizarán los datos de empleo industrial obtenidos del Censo de Población y vivienda del año 2011 y algunos construidos para efectos del trabajo. De acuerdo con Haining (2001), el análisis espacial consta de tres elementos principales: a) el modelo cartográfico, donde cada base de datos está representada como un mapa; b) los modelos matemáticos, donde los resultados dependen de la forma de interacción espacial entre los objetos en el modelo, o de relaciones espaciales, o bien de la posición geográfica de los objetos dentro del modelo; c) el desarrollo y la aplicación de técnicas estadísticas para el adecuado análisis de los datos espaciales y el cual, como consecuencia, hace uso de la referencia espacial de los datos.

Normalmente los métodos estadísticos tradicionales no involucran coordenadas geográficas ni distancias físicas y asumen que no hay interacción entre las unidades de

observación, es decir, que los valores de los atributos son independientes entre sí. Cuando se realiza inferencia estadística se asume que hay aleatoriedad en el comportamiento de los valores de los atributos, comúnmente dicha aleatoriedad se asocia a una distribución estadística que da origen a los datos (una distribución normal, una *poisson*,).

Sin embargo, es importante cuestionarse ¿Qué supone el análisis estadístico tradicional para estudiar unidades geográficas? El análisis estadístico tradicional utilizado para analizar variables asociadas a un espacio geográfico, se caracteriza por construirse sobre supuestos en los que no hay auto-correlación entre las unidades de observación. La dependencia espacial significa que los valores de una misma variable, que son medidos en localidades que son cercanas entre sí, tienden a ser similares; es decir, a mayor cercanía geográfica corresponde una mayor similitud en los valores. Esto significa que la dependencia espacial se produce cuando el valor de la variable dependiente en una unidad espacial es parcialmente función del valor de la misma variable en unidades vecinas. La dependencia espacial hace diferente a la estadística espacial de la tradicional debido a que considera la distancia y la contigüidad existente entre unidades de observación.

Dado lo anterior, la metodología propuesta, considera la dependencia espacial entre las unidades de observación. Con ello, las características geográficas asociadas a la demanda por empleo industrial y su distribución cantonal pueden ser analizadas a través de técnicas que consideren la dependencia espacial y la auto-correlación espacial.

1. Método empleado para medir la auto-correlación espacial en los cantones de Costa Rica

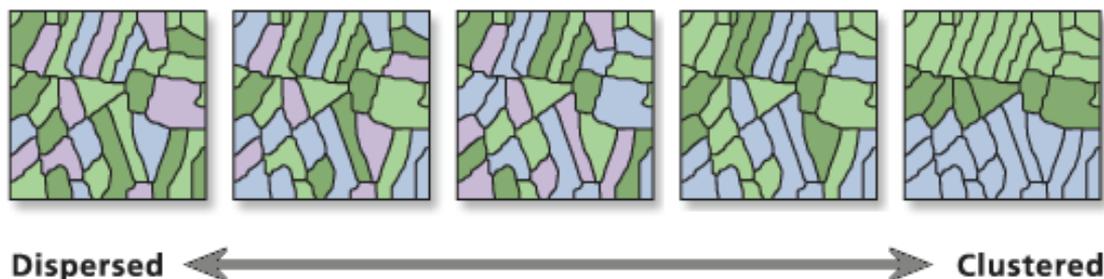
Siguiendo a Anselin (2001) en este artículo se definirá la auto-correlación espacial

como la “coincidencia entre valores similares (empleo industrial) y localizaciones cercanas” (cantones). Aunque la dependencia espacial puede parecer similar a la dependencia en las series temporales, este parecido sólo es real, en parte, debido a la naturaleza multidireccional de la dependencia en el espacio frente a la clara situación unidireccional del tiempo.

La idea es analizar si la generación de empleo industrial en los cantones de Costa Rica se distribuye de forma sistemática en el espacio geográfico, en cuyo caso se probaría la hipótesis de existencia de auto-correlación espacial de la variable. El efecto que podemos encontrar al evaluar la auto-correlación espacial puede ser de signo positivo o negativo, así como nulo.

- a) Sera *positivo* cuando se presenta asociaciones de valores similares entre empleo industrial y localizaciones cercanas (cantones); es decir, cuando, en el espacio geográfico, los valores altos de una variable están rodeados por valores altos y viceversa. Éste sería el caso, del llamado efecto contagio o desbordamiento (“*spillover*”) que se produce en muchos fenómenos socioeconómicos, en general, en los que su presencia en una región es causa de su extensión a regiones vecinas, favoreciendo la concentración del fenómeno en la zona (ver recuadro de la derecha en la figura 1).
- b) Se presenta auto-correlación espacial negativa en un espacio cuando los valores altos de una variable (empleo industrial) se encuentran rodeados por valores bajos de la misma, y viceversa. Esta configuración, en la que se produce una mayor disimilitud entre unidades geográficas cercanas que entre las lejanas, por ejemplo, la que se produciría en fenómenos de jerarquías espaciales del tipo centro-periferia (ver recuadro de la izquierda en la figura 1).

FIGURA 1
POSIBLES EFECTOS DE AUTO-CORRELACIÓN ESPACIAL
(IZQUIERDA EFECTO NEGATIVO Y DERECHA EFECTO POSITIVO)



Fuente: ArcGIS Resources.

c) Por último, se produce ausencia de auto-correlación espacial en una variable geográfica cuando ésta se distribuye de manera aleatoria sobre el espacio.

• *Formulación matemática para la detección de auto-correlación espacial*

Las interacciones espaciales que suelen producirse en muchos fenómenos humanos son la causa verdadera del efecto de dependencia o autocorrelación espacial y podrían expresarse matemáticamente como una relación funcional entre los valores que adopta una variable Y en la localización i (y_i) y los valores de dicha variable en un conjunto de localizaciones situadas en el territorio en estudio, del modo siguiente:

$$y_i = f(y_1, y_2, \dots, y_N)$$

El problema es que esta expresión no resulta operativa en la práctica porque da lugar a un sistema no identificable, con muchos más parámetros a estimar, hasta un total de $(N^2 - N)$, de observaciones disponibles (N). Por este motivo, como plantea Anselin (2001) deben establecerse estructuras comunes, subyacentes al fenómeno de interacción que se estudia, de forma que sólo se estime un pequeño número de características propias de la dependencia espacial existente. En los fenómenos humanos, la influencia que ejercen ciertas unidades geográficas sobre una dada,

suele expresarse matemáticamente a través de conceptos topológicos de vecindad, en los que juega un papel fundamental la distancia métrica (euclídea, bloque de Manhattan o general de Minskovski).

Para el análisis propuesto es importante que las observaciones se encuentran organizadas en unidades espaciales discretas, tanto puntos situados en una red regular (o irregular) como cantones de un mapa. El conjunto de unidades vecinas a una dada suele representarse gráficamente como una estructura gráfica reticular y, de forma matemática, como una matriz de interacciones. A continuación se muestra la metodología empleada para la estimación de matrices de interacción y pesos espaciales.

a) *Estimación de la matriz de interacciones espaciales en los cantones de Costa Rica*

Los contrastes de dependencia o auto-correlación espacial pueden basarse en una noción de contigüidad binaria entre las unidades espaciales, en nuestro caso cantones. De acuerdo con este concepto, una situación de vecindad entre dos unidades espaciales se podría expresar mediante valores de tipo 0-1. Es decir, si dos unidades espaciales tienen una frontera común de longitud no nula, se considera que son contiguas y se les asigna el valor de 1.

Para la definición de contigüidad anterior se generó un mapa de polígonos de los 81 cantones de Costa Rica, a partir del cual, una vez transformados los polígonos a líneas se pudo obtener las fronteras entre las 81 unidades espaciales utilizando el software ArcGIS 10.2 y programando el proceso para que genere los resultados de manera automática. Al ser los 81 cantones unidades geográficas que se posicionan de forma irregular, resulta más factible conocer las fronteras entre las distintos cantones.

De acuerdo con Vayá y Suriñach, (1996) la estructura espacial de estos resultados suele expresarse formalmente a través de una matriz de interacciones espaciales, también llamada “matriz de pesos, ponderaciones, distancias o contactos espaciales”. En esta matriz, cada unidad espacial (cantón) se representa a la vez mediante una fila y una columna. En cada fila, los elementos no nulos de las columnas se corresponden con las unidades espaciales contiguas. En el caso de los cantones de Costa Rica se formó una matriz de (81 x 81) cuyas celdas están numeradas de izquierda a derecha y de arriba abajo, utilizando el criterio de contigüidad de la torre.

Esta matriz estaría relacionando la vinculación entre cantones si estos comparten una frontera en común asignándole un “1” y “0” en caso contrario. Sin embargo, al analizar la realidad espacial de los cantones de Costa Rica es posible que no sea conveniente el plantear una matriz de interacciones espaciales binaria ya que es simétrica y, por tanto, incapaz de incorporar influencias no recíprocas (no siempre la influencia que j recibe de i es la misma que la que i recibe de j) y su diagonal principal está constituida por ceros. Además, considera como único determinante de las interdependencias cantonales la adyacencia física, descuidando con ello posibles influencias mutuas entre cantones que, aun estando alejados, mantienen, por ejemplo, estrechas relaciones comerciales. Pese a todo, esta matriz de interacciones espaciales es muy utilizada habitualmente por su simplicidad, sin embargo con el afán de mejorar los resultados y corregir la simplificación de la matriz de interacciones binaria, se estimaran

matrices más complejas siguiendo el criterio de proximidad física que diversos autores han propuesto en otros estudios, a continuación se detallan.

b) Estimación de la matriz de pesos espaciales en los cantones de Costa Rica

De acuerdo con Anselin (1988) la matriz de pesos espaciales es una generalización de la matriz de interacciones o contigüidades, por lo que suele ser designada también con estos nombres, aunque se la conoce habitualmente como matriz de pesos, ponderaciones, retardos o contactos espaciales. Otros autores también la denominan “matriz de ponderaciones Cliff-Ord”, en honor las personas que la formularon inicialmente.

El uso de las matrices de pesos espaciales a diferencia de la matriz binaria de interacciones, permite elegir del conjunto de ponderaciones la más apropiada para cada fenómeno, lo que supone una mayor flexibilidad en la definición de la estructura de interdependencias de un sistema en este caso cantonal y permite considerar cuestiones como las barreras naturales o el tamaño de los cantones. Anselin (1988) plantea que cuando sea necesaria la consideración de hipótesis acerca del grado de vinculación existente entre áreas vecinas, deben utilizarse distintos conjuntos de ponderaciones que permitan contrastar dichas hipótesis.

En este artículo se utilizará una matriz de pesos generalizada (W), en lugar de utilizar las ponderaciones binarias δ_{ij} como cuantificación del concepto de vínculo, ya que recoge el efecto de cantón i sobre el cantón j a través de un peso o ponderación W_{ij} , de forma que $W = (W_{ij})$. La interpretación de estos elementos es la siguiente:

- $W_{ij} = 0$ sería indicativo de ausencia de auto-correlación espacial entre las observaciones i, j (por convenio, los elementos de la diagonal principal de la matriz de pesos serán, como en la matriz de interacciones, igual a cero).
- $W_{ij} \neq 0$ sería indicativo de existencia de una interacción espacial entre las observaciones

i, j, que podría ser expresada como simple contigüidad binaria (teniendo una frontera común), como contigüidad de distancias, a través de centroides, dentro de una banda de distancia mínima (considerándose contiguos dos puntos cuando estén situados a una distancia inferior a la prefijada como mínima), o como función inversa de la distancia simple o cuadrática.

La especificación adecuada de los elementos de esta matriz, w_{ij} , es uno de los puntos metodológicos más difíciles y controvertidos en la econometría espacial. Para efectos de este artículo y como lo plantean Stetzer (1982); Anselin y Rey (1991) y Florax y Rey, (1995) se seguirán los siguientes criterios: 1) el grado de vinculación existente entre dos cantones que dependerá de la distancia entre sus centros geográficos o demográficos (en este caso centro urbano) (estimando la distancia más cercana por carretera en km^2), y 2) la relación entre dos cantones midiendo la longitud de la frontera común de dichos cantones (dicho procedimiento se realizó con sistemas de información geográfica convirtiendo los polígonos de los cantones a líneas y luego programando en ArcGIS 10.2 para que estime el valor en km^2 de la frontera que comparten los cantones.

c) Estimación de auto-correlación espacial global y local en los 81 cantones de Costa Rica

Anselin (1998) y Wise (1999) plantean la existencia de una doble perspectiva en el análisis del fenómeno de asociación o dependencia espacial, lo que han denominado “perspectiva global” y “perspectiva local”. En el primer caso (perspectiva global) tiene por objeto el contraste de la presencia de tendencias o estructuras espaciales generales en la distribución de una variable sobre un espacio geográfico completo, mientras que el segundo caso (dependencia local) se definen concentraciones, en un lugar del espacio global analizado, de valores especialmente altos o bajos de una variable en comparación con el valor medio de la misma.

En el caso de la dependencia global, Moreno y Vayá (2000) exponen que los estadís-

ticos de auto-correlación espacial global son las primeras formulaciones propuestas en la literatura como medida estadística del efecto de auto-correlación espacial. De acuerdo con estos autores dichos contrastes tienen la capacidad de resumir el esquema general de dependencia presente en una variable espacial en un único indicador. La literatura muestra que existen varios indicadores para medir este fenómeno, entre ellos destacan los tests I de Moran (1948) y c Geary (1954) que son probablemente los contrastes de autocorrelación espacial más conocidos, aunque también puede añadirse el estadístico G(d), que fue propuesto por Getis y Ord (1992).

Para el caso de la dependencia local, Anselin (1995) plantea que los tests de auto-correlación permiten contrastar la presencia de subzonas de dependencia espacial dentro de un espacio general dado. El autor denomina a estos indicadores “Indicadores LISA” (Indicadores Locales de Asociación Espacial), para los que establece una serie de buenas propiedades estadísticas. A continuación se detalla la forma en que se estimarán los contrastes de auto-correlación (Global y Local) para los cantones de Costa Rica en el caso de la demanda por empleo industrial).

• *Contrastes de auto-correlación espacial global*

La perspectiva global del fenómeno de auto-correlación espacial tiene por objeto el contraste de la presencia de tendencias o estructuras espaciales generales en la distribución de una variable (empleo industrial) sobre un espacio geográfico completo (Costa Rica). En otras palabras, se trata de contrastar la hipótesis de que el empleo industrial se encuentre distribuido de forma totalmente aleatoria en los 81 cantones de Costa Rica o, si por el contrario, existe algún tipo de asociación significativa de valores similares de empleo o distintos entre cantones vecinos. Para efectos del artículo, se estimará el tests I de Moran (1948), del cual se pasa a detallar su formulación matemática:

Seguendo a Cliff y Ord (1973 y 1981) y Anselin (1995), el test I de Moran fue inicialmente formulado como función de una variable (Y), considerada en los puntos del espacio (i,j), en desviaciones a la media, y los elementos de la matriz binaria de interacciones espaciales [δ_{ij}]. Esta expresión inicial de Moran podría ser generalizada, sustituyendo la matriz de interacciones por la más general matriz de pesos espaciales, W_{ij} de la siguiente forma:

$$I = \frac{N \sum_{(2)} W_{ij} (Y_i - \bar{Y})(Y_j - \bar{Y})}{S_0 \sum_{i=1}^N (Y_i - \bar{Y})^2}$$

Donde

W_{ij} : son los elementos de la matriz de pesos espaciales correspondientes al par (i, j)

$$S_0 = \sum_i \sum_j w_{ij} = \sum_{(2)} W_{ij}$$

es la suma de los pesos espaciales.

\bar{Y} : es el valor medio o esperado de la variable y.

N: número de observaciones o tamaño muestral

Al utilizar una matriz de interacciones espaciales estandarizada por filas, que es la situación óptima de aplicación de este test, el término $S_0 = N$, dado que la suma de los valores de cada fila es igual a la unidad. De este modo, el estadístico I queda reducido al cociente del producto espacial cruzado de los valores de la variable partido por la varianza:

$$I = \frac{N \sum_{(2)} W_{ij} (Y_i - \bar{Y})(Y_j - \bar{Y})}{S_0 \sum_{i=1}^N (Y_i - \bar{Y})^2}$$

En este caso I está basado en los productos cruzados de las desviaciones de Y_i respecto de \bar{y} . También es evidente que el test I de Moran es similar al coeficiente de auto-correlación temporal: el término del numerador es una medida de la covarianza entre valores de Y en dos localizaciones distintas (i,j) y el denominador expresa la varianza de Y en el punto i. Sin embargo, aunque parecido, el test I no es

equivalente al clásico coeficiente de correlación, fundamentalmente porque no se encuentra centrado en el valor cero. De hecho, la media teórica de la I de Moran es el cociente $\frac{-1}{N-1}$.

Dado lo anterior, el valor esperado de I es negativo y en función únicamente del tamaño de la muestra (N), aunque esta media tiende a cero a medida que el tamaño de la muestra aumenta. En cuanto a la varianza teórica del coeficiente I, se verá más adelante que depende de determinados supuestos estocásticos. Un coeficiente I de Moran mayor que su valor esperado sería indicativo de auto-correlación espacial positiva, mientras que un valor de I inferior a la media pondría de manifiesto la existencia de auto-correlación espacial negativa. Este estadístico estará muy afectado por aquellos puntos vecinos sensiblemente distintos de la media de la variable en estudio.

Según Cliff y Ord (1981) respecto a la distribución del contraste I, cuando el tamaño muestral es suficientemente amplio, la expresión estandarizada del test I se distribuye como una normal tipificada, $N(0,1)$, como el test de recuento de vínculos de Moran. Por eso, también suele considerarse el estadístico inicial I, el proceso inferencial suele utilizar los valores estandarizados (z) de cada uno de ellos, obtenidos a través del cociente entre la diferencia del valor inicial y la media teórica, y la desviación típica teórica, de la manera siguiente:

$$Z = \frac{I - E[I]}{SD[I]}$$

Siendo $E[I]$: La media teórica del estadístico I y $SD [I]$ la desviación típica del estadístico I. La interpretación de los valores estadísticamente significativos de la variable tipificada Z_1 sería la siguiente:

- I. Valores no significativos del test I estandarizado, Z_1 , correspondiente a una variable Y, conducirían a aceptar la hipótesis nula de no auto-correlación espacial o inexistencia de patrones de comportamiento de dicha variable sobre el espacio.
- II. Valores significativos de $Z_1 > 0$ serían indicativos de auto-correlación espacial positiva, es decir, que es posible encontrar valores

parecidos (altos o bajos) de la variable Y, espacialmente agrupados, en mayor medida de como estarían por casualidad.

III. Valores significativos de $Z_I < 0$ serían indicadores de auto-correlación espacial negativa,

es decir, que se produce una no-agrupación de valores similares (altos o bajos) de la variable Y superior a lo normal en un patrón espacial aleatorio. Se trata de un concepto algo más difícil de captar.

CUADRO 2
INTERPRETACIÓN DE LOS VALORES ESTANDARIZADOS DE LOS ESTADÍSTICOS DE
AUTO-CORRELACIÓN ESPACIAL GLOBAL

| Test | Hipótesis nula (z no significativo) | Hipótesis alternativa (z significativo) $Z > 0$ | Hipótesis alternativa (z significativo) $Z < 0$ |
|---------|-------------------------------------|---|---|
| I Moran | No auto-correlación espacial | Auto-correlación espacial POSITIVA | Auto-correlación espacial NEGATIVA |

Fuente: Elaboración propia.

• *Contrastes de auto-correlación espacial Local*

De acuerdo con varios autores (Getis y Ord (1992); Openshaw (1993); Anselin (1993) y (1995); Ord y Getis (1995) y (2001); Vayá y Suriñach (1996); Tiefelsdorf y Boots (1997) y Sokal (1998)) los estadísticos de auto-correlación global, centrados en el análisis de dependencia general propia de todas las unidades de un espacio geográfico, no son capaces de detectar la inestabilidad o deriva espacial de ciertas estructuras locales de asociación o inestabilidades locales que pueden estar, a su vez, presentes o no en una estructura global de dependencia.

Para Vayá y Suriñach (1996) el problema de la dependencia espacial local puede plantearse desde dos puntos de vista:

- Existe la posibilidad de que, en un espacio dado, no se detecte la presencia de auto-correlación espacial global en la distribución de una variable aunque, de hecho, existan pequeños "clusters" espaciales en los que dicha variable experimenta una concentración (o escasez) importante.
- Existe también la posibilidad de que, habiéndose detectado dependencia a nivel global en una variable, no todas las regiones (cantones en este caso) del espacio considerado

contribuyan con igual peso en el indicador global, es decir, que coexistan unas zonas en las que la variable se distribuya de forma aleatoria junto a otras con una importante contribución a la dependencia existente.

Para analizar lo anterior en los cantones de Costa Rica, se define un contraste de asociación local que indica hasta qué punto un cantón se encuentra rodeado por otros con valores altos o bajos de empleo industrial; para ello se utilizará uno de los indicadores locales de asociación espacial (LISA) (Estadístico Local de Moran), propuestos por Anselin (1995).

El estadístico Local de Moran se puede definir como:

$$I = Z_i \sum_{j=1}^{J_i} W_{ij} Z_j$$

donde Z_i, Z_j , son variables y_i estandarizadas

\sum_j es la sumatoria que unánimemente incluye los valores vecinos a $i: j \in J_i$

Para lograr una mejor interpretación de este estadístico, la matriz de pesos W_{ij} debería estar estandarizada por filas siendo, por convenio, cada elemento $W_{ij}=0$. También, en este caso, es fácil comprobar que la suma de estadísticos locales $I_i=I_1$ es el test I de Moran:

$$\sum_i I_i = \sum_i Z_i \sum_j w_{ij} Z_j = \sum_i \sum_j w_{ij} Z_i Z_j$$

La ecuación anterior se puede expresar como:

$$I = \frac{N \sum_{(2)} W_{ij} Z_i Z_j}{S_0 \sum_{i=1}^N Z_i^2} \Rightarrow I = \frac{\sum_i I_i}{S_0 \sum_i \frac{Z_i^2}{N}}$$

Donde $\sum_{(2)} = \sum_i \sum_j$

$$\sum_{(2)} = \sum_i \sum_j$$

De este modo, se cumple que el sumatorio de valores del estadístico local I_i es equivalente al test global I , siendo el factor de proporcionalidad $\gamma = S_0 m_2$:

$$\sum_i I_i = \frac{I}{S_0 m_2} I = \gamma \cdot I$$

donde: $m_2 = \sum_i \frac{Z_i^2}{N}$, momento de 2º orden de la variable z_i .

En el caso de que la matriz W esté estandarizada por filas, $S_0 = N$, de manera tal que el factor $\gamma = \sum_i z_i^2$ y, en el caso de trabajar con variables estandarizadas, $m_2=1$, lo que daría que dicho factor $\gamma = S_0$. De esta manera el test I local de Moran podría también expresarse del modo siguiente:

$$I_i = \frac{Z_i}{m_2} \sum_j W_{ij} Z_j$$

Anselin (1995), plantea que es posible calcular los momentos de I_i , bajo la hipótesis nula de ausencia de asociación espacial, para el supuesto inferencial de aleatoriedad o muestreo aleatorio. Cualquier contraste de significación de asociación espacial local puede basarse en estos momentos, aunque la distribución exacta de un estadístico de este tipo aún se desconoce. Para el autor, una forma de facilitar su interpretación es utilizando una distribución normal.

De acuerdo con Anselin (1995) el método anterior puede ofrecer, mediante su representación cartográfica y el scatterplot de Moran, información sobre clusters y outliers de unidades con presencia alta de miembros de un grupo. Utilizando el mapa de la significación de los indicadores locales asociado al scatterplot podemos identificar zonas con presencia alta de miembros de un grupo rodeadas de zonas con presencia también alta (situación High-High en el scatterplot de Moran), o bien zonas con presencia alta rodeadas de unidades con presencia baja (situación High-Low en el scatterplot de Moran).

De igual forma permite la detección de zonas con presencia baja rodeadas de unidades también con presencia baja (situación Low-Low), o bien zonas de presencia baja rodeadas de unidades con presencia alta de la variable analizada (situación Low-High). Finalmente, se pueden detectar zonas sin asociación espacial significativa.

El análisis de este artículo se centra en aquellos cantones donde hay presencia significativa de demanda por empleo industrial, que se encuentren rodeados de cantones con igual situación o bien, sea un cantón rodeado de situaciones diferentes; en ambos casos nos referiremos a ellas como zonas cluster. Según Anselin (2003), para llevar a cabo los contrastes de autocorrelación global y local es necesaria la utilización de la matriz de contacto binaria que permite expresar la vecindad entre dos zonas mediante valores 0-1. Si dos zonas tienen una frontera común se las considera contiguas y se les asigna el valor uno.

2. Datos

La fuente de información para estimar la demanda por empleo industrial en los 81 cantones del país se obtuvo de los datos generados por el último Censo de Población y Vivienda del 2011, elaborado por el Instituto de Estadística y Censos (INEC). En dicho censo, se pregunta por primera vez en la historia del país el lugar de residencia y el lugar de trabajo de los empleados, con ello se puede generar una matriz de origen destino

de empleo por rama de actividad económica, que para efectos de este artículo será el empleo en industria. Con dicha matriz, es posible obtener a nivel cantonal el empleo real que genera cada cantón y de donde provienen los trabajadores así como sus características y perfil socioeconómico.

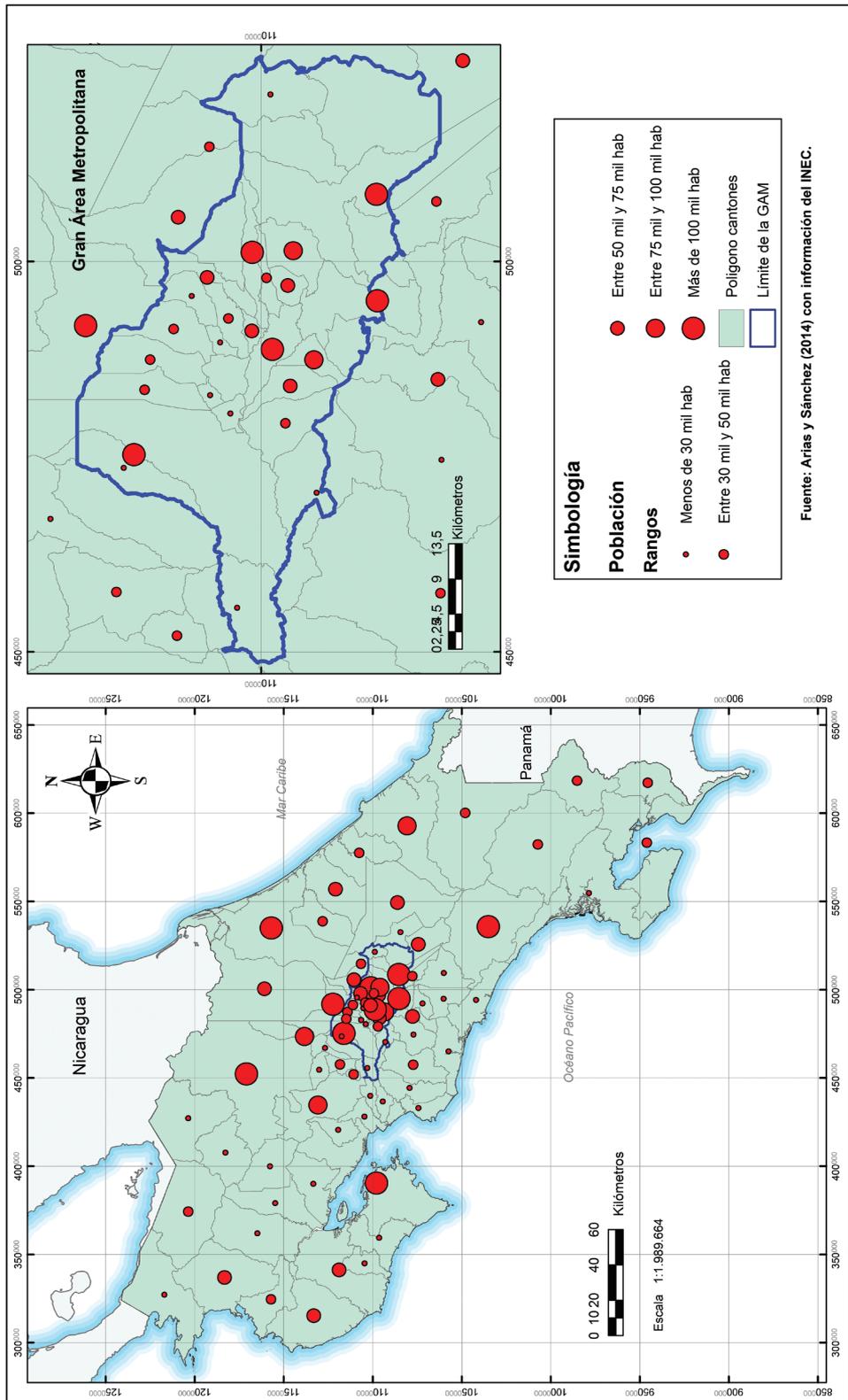
Los datos geo-referenciados de carreteras, distancias y tiempos de viajes así como la información cartográfica de Costa Rica, en especial los polígonos de cantones, se obtuvieron de las bases de datos generadas en el Instituto de Investigaciones en Ciencias Económicas de la Universidad de Costa Rica de proyectos

elaborados en la rama de la economía urbana y regional, así como también del Instituto Geográfico Nacional (IGN) y del Programa de Investigación en Desarrollo Urbano Sostenible (ProDUS-UCR).

3. Zona de Estudio

Como se mencionó anteriormente, la zona de estudio comprende los 81 cantones que conforman el territorio de Costa Rica (51.100km²). La figura 2 muestra el mapa de cantones de Costa Rica con su respectiva población al año 2011.

FIGURA 2
ZONA DE ESTUDIO (CANTONES DE COSTA RICA)



Fuente: Arias y Sánchez (2014) con información del INEC.

Fuente: Elaboración propia con datos del INEC (2012).

IV. RESULTADOS

Esta sección se estructura de la siguiente manera, primero se realiza un análisis general del empleo generado por cantón, posteriormente se analiza la demanda por trabajadores en la industria manufacturera a nivel de cantón, discutiendo como se distribuyen en el espacio y el perfil de dichos empleados. Luego se aplican los indicadores de auto-correlación espacial global y local y por último se generan análisis de mapas *cluster* para identificar algunos conglomerados.

1. Distribución del empleo en los cantones de Costa Rica

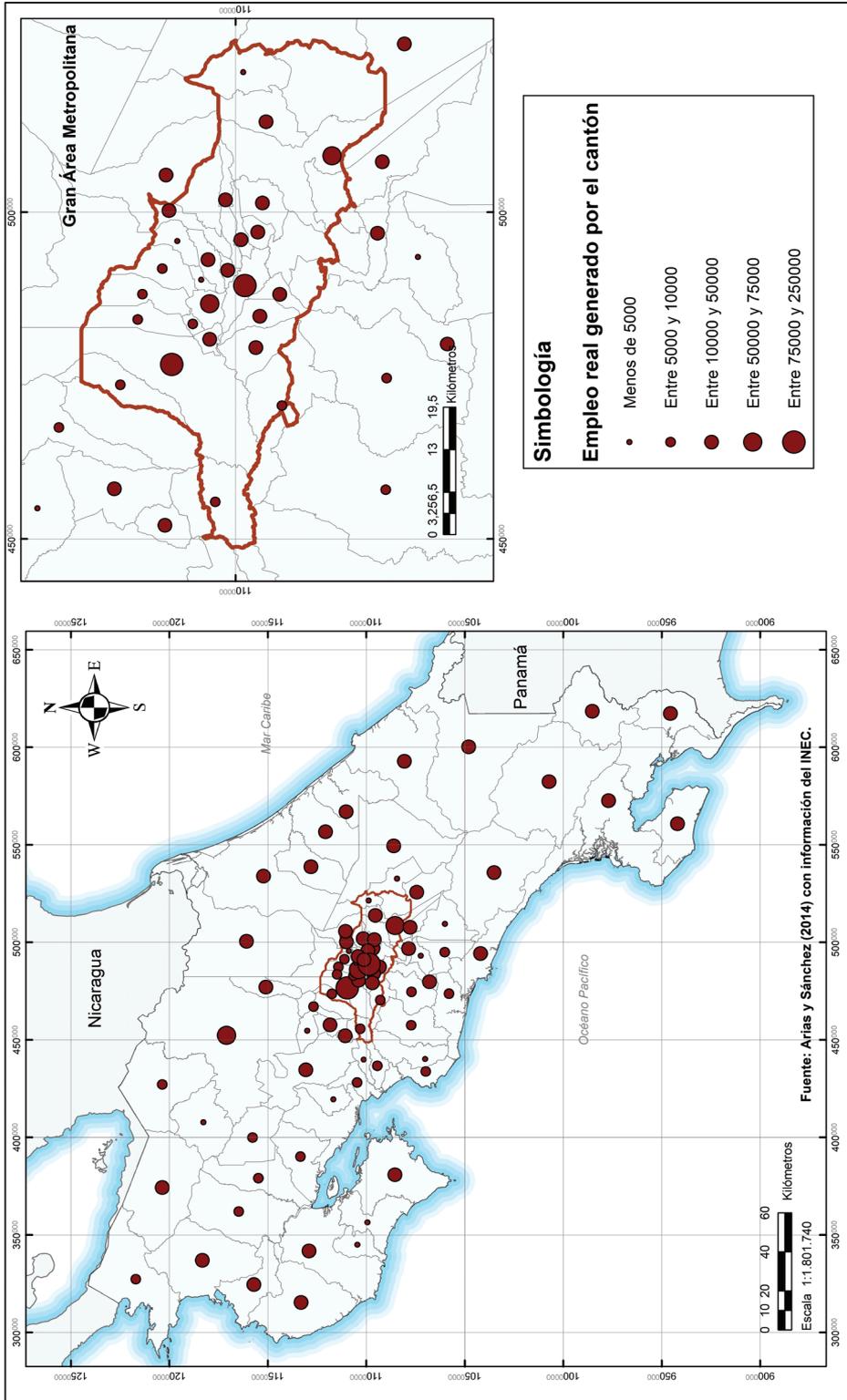
De la figura 3 es posible establecer que el país muestra una estructura muy heterogénea en lo que respecta a la generación de empleo entre los 81 cantones. Algunos aspectos relevantes son:

- Costa Rica en el año 2011 tenía cerca de 1,67 millones de personas ocupadas de las cuales 13,9% lo hacían en el sector primario, 17,9% en el secundario y el sector terciario agrupaba el 68,2%.
- De los casi 1.67 millones de personas que trabajaban cerca el 51% del empleo se genera en 12 cantones.
- Costa Rica cuenta con una Gran Área Metropolitana (GAM) (recuadro figura 2 borde

rojo), que presenta una extensión territorial equivalente al 3,83% del área nacional (31 cantones), y que se convierte en un megamercado de empleo al generar alrededor del 60% de los empleos del país.

- Al interior de la GAM el 50% del empleo que se genera proviene de los 4 principales cantones metropolitanos como los son San José (26,9%), Alajuela (10,3%), Heredia (7,6%) y Cartago (5,7%).
- Fuera de la GAM, los cantones que más generan empleo son los que cuentan con ciudades intermedias, es decir aquellos municipios cabeceras regionales, conformado por San Carlos (3,7%), Pococí (2,7%), Pérez Zeledón (2,7%), Puntarenas (2,5%), Limón (1,9%), Grecia (1,7%), San Ramón (1,6%), Liberia (1,5%) y Turrialba (1,3%); agrupando el 18% del total de empleo generado.
- Más de la mitad de los cantones del país crean una cantidad de empleo muy pequeña, los datos muestran que el 56% de estos municipios solo concentran el 20% del empleo generado en el país.
- Como se muestra en el cuadro 3, de los tres sectores de la economía, es en el sector servicios donde se presenta las mayores diferencias cantonales en la distribución del empleo con una desviación estándar de 27.193 empleos con respecto al promedio, en segundo lugar está el secundario con (4.288) y en tercer lugar el primario con (2.438).

FIGURA 3
EMPLEO TOTAL GENERADO POR CADA CANTÓN DE COSTA RICA



Fuente: Elaboración propia con datos del INEC (2012).

CUADRO 1
ESTADÍSTICAS DESCRIPTIVAS PARA EL EMPLEO POR SECTOR
EN LOS 81 CANTONES DE LA GAM

| | Primario | Secundario | Terciario |
|-------------------------|----------|------------|-----------|
| Media | 2438 | 14715 | 2855 |
| Error típico | 476 | 3021 | 363 |
| Mediana | 1076 | 7130 | 1843 |
| Desviación estándar | 4288 | 27193 | 3265 |
| Curtosis | 17 | 48 | 6 |
| Coficiente de asimetría | 4 | 6 | 2 |
| Rango | 25015 | 227233 | 16153 |
| Mínimo | 86 | 864 | 75 |
| Máximo | 25101 | 228097 | 16228 |

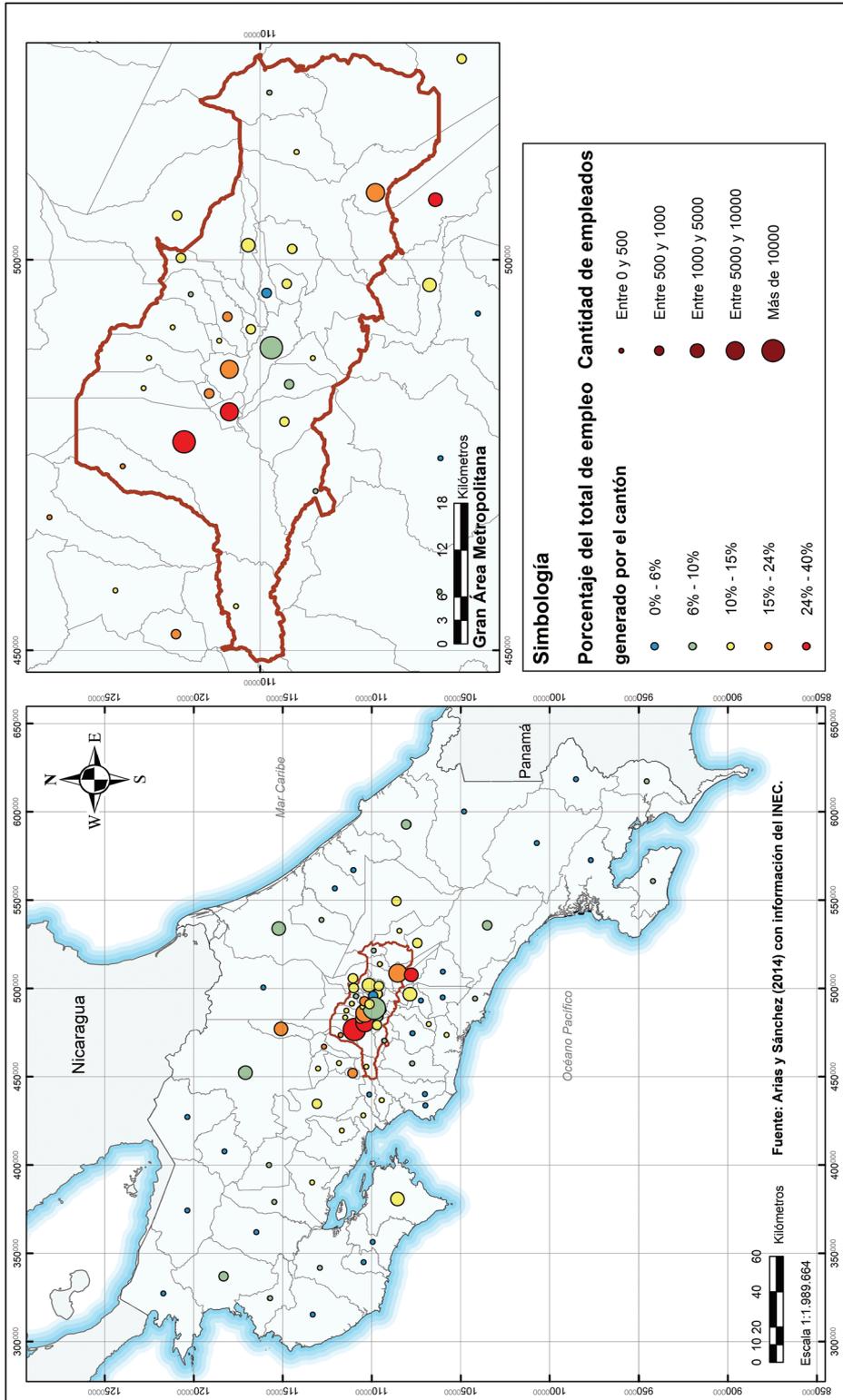
Fuente: Elaboración propia con datos del INEC (2012).

2. Demanda por empleo industrial en los cantones de Costa Rica

Este análisis solo contempla el empleo manufacturero definido en el Censo de Población y vivienda elaborado por el INEC en el año 2011. El análisis no incluye el empleo en cons-

trucción por tanto no es un estudio del empleo en el sector secundario. La figura 4 muestra cómo se distribuye espacialmente el empleo en el sector industrial en los 81 cantones del país. El tamaño del círculo muestra la cantidad absoluta de empleo y el color el aporte del sector al total de empleo generado en el cantón.

FIGURA 4
DEMANDA POR EMPLEO INDUSTRIAL EN LOS CANTONES DE COSTA RICA

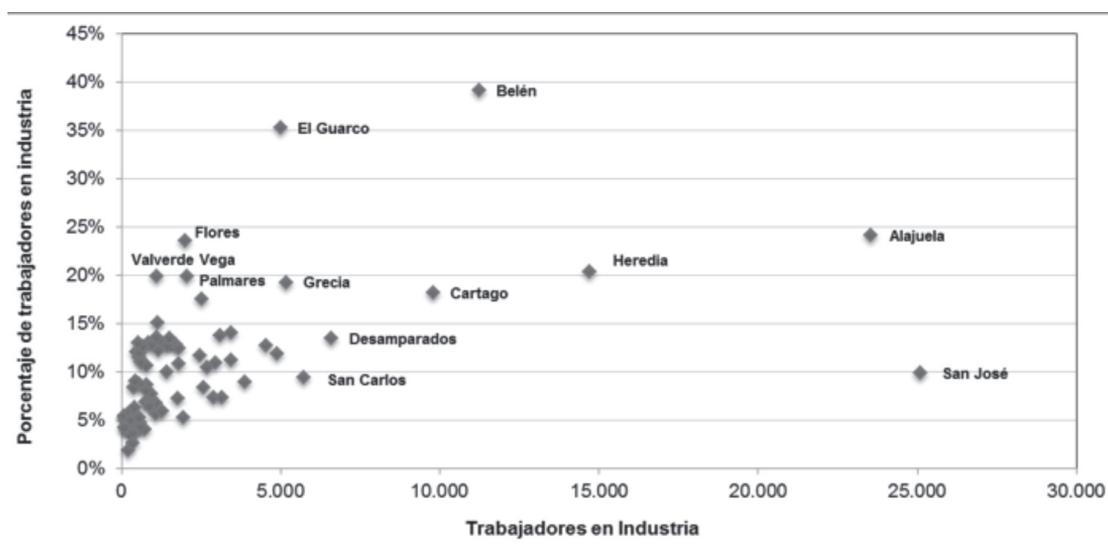


Fuente: Elaboración propia con datos del INEC (2012).

De la lectura del mapa en la figura 4 se deducen los siguientes aspectos:

- Una cuarta parte del empleo industrial generado en Costa Rica proviene de los cantones de San José (12,7%) y Alajuela (11,9%). Si le sumamos el empleo creado en Heredia (7,4%), Belén (5,7%), Cartago (5%), Desamparados (3,3%) y San Carlos (2,9%) se agrupa cerca del 50% de todos los empleos industriales del país. Lo anterior demuestra la gran concentración de esta actividad principalmente en la Gran Área Metropolitana (GAM) donde existen zonas industriales bien definidas.
- Un segundo grupo de cantones que generan en términos absolutos cada uno entre 3.000 y 5.000 empleos (segundo tamaño más grande de los círculos en el mapa) son Grecia (5.169), El Guarco (4.995), Puntarenas (4.890), Goicoechea (4.529), Pococí (3.883), Tibás (3.441), Santa Ana (3.425), Pérez Zeledón (3.146) y la Unión (3.100). Estos cantones se caracterizan por ser parte de corredores industriales dentro de la GAM
- En la parte inferior, una cuarta parte de los cantones del país generan en términos absolutos una cantidad muy pequeña del empleo industrial que se genera en el país (2,7%). En estos cantones, en su mayoría rurales, algunos con población indígena y otros costeros, la cantidad de empleo que se demanda en industria es menor a 500 empleos por municipio.
- En términos de importancia relativa para el cantón, el sector industrial toma protagonismo en Belén (40%) y el Guarco (35%), donde poco más de uno de cada tres empleos que se generan, son dentro del sector manufactura (círculos de color rojo en el mapa). También es importante en los cantones de Alajuela (24,1%), Flores (23,6%), Heredia (20,4%), Palmares (20%), Valverde Vega (20%) y Grecia (19,3%) donde la industria manufacturera aporta uno de cada 5 empleos que genera el cantón.

GRÁFICO 1
DEMANDA POR EMPLEO INDUSTRIAL EN LOS CANTONES
DE COSTA RICA Y PORCENTAJE QUE REPRESENTA DENTRO DEL TOTAL DE EMPLEO DEL CANTÓN, 2011



Fuente: Elaboración propia con datos del INEC (2012).

De los 82 cantones del país, en 41 de ellos menos del 10% del empleo que se genera localmente es creado por la industria manufacturera, es decir la participación relativa de este sector es muy reducida. Esto por el protagonismo que tienen el sector servicios en zonas urbanas o el sector primario en zonas más rurales.

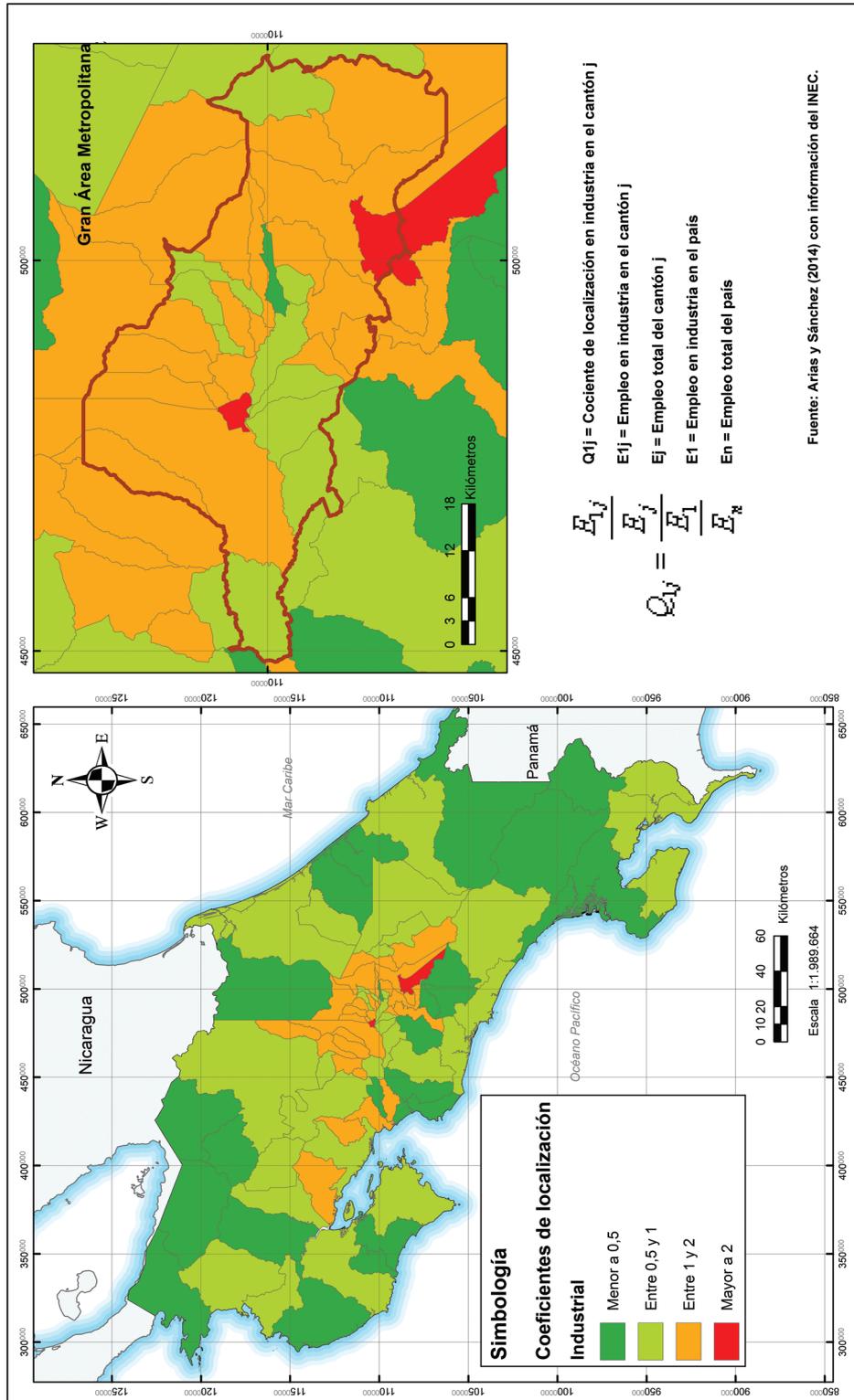
Este bajo protagonismo queda más claro al analizar los coeficientes de localización industrial a nivel cantonal para la industria. Estos coeficientes se obtienen comparando el peso relativo que tiene una actividad en un cantón (medido en términos de empleo industrial) con el peso relativo que tiene esa actividad en el país en su conjunto (Arias y Sánchez, 2012). Si el coeficiente supera la unidad existirá una especialización relativa en ese cantón y en esa medida cierta ventaja competitiva. La figura 5 muestra los coeficientes localización industrial obtenidos para el año 2011. Solamente 28 cantones de los 81 muestra coeficientes mayores a la unidad, es decir la importancia de la industria en estos municipios se ubica por encima de la media nacional.

Los cantones en colores verde claro y oscuro muestran una baja participación de la

industria y coeficientes menores a la unidad; es decir no muestran presencia de especialización. Como se nota en el mapa de la figura 5, estos cantones están, en su mayoría, fuera de la GAM, estando los valores del coeficiente más bajos asociados a cantones con fuerte actividad agrícola (color verde oscuro).

Por el contrario, los cantones en colores naranja y rojo muestran coeficientes mayores a la unidad y por tanto cierto grado de concentración o especialización. Estos cantones se ubican principalmente en la GAM, lo cual no es extraño ya que en esta zona es donde se ubican las principales zonas industriales del país. Los mayores valores se concentran en los cantones de Belén y el Guarco (color rojo oscuro), así como Alajuela (1,98), Flores(1,93) y Heredia (1,67). También destaca la zona de occidente de la región Central, especialmente los cantones de Palmares, Valverde Vega, Grecia, y Naranjo, asociado a la agroindustria, industria textil y a la de madera y muebles. Fuera de la región Central se ubican los cantones de Abangares, Esparza, Orotina y Montes de Oro asociados a la industria alimentaria.

FIGURA 5
COEFICIENTES DE LOCALIZACIÓN INDUSTRIAL EN LOS 81 CANTONES
DE COSTA RICA, 2011



Fuente: Elaboración propia con datos del INEC (2012).

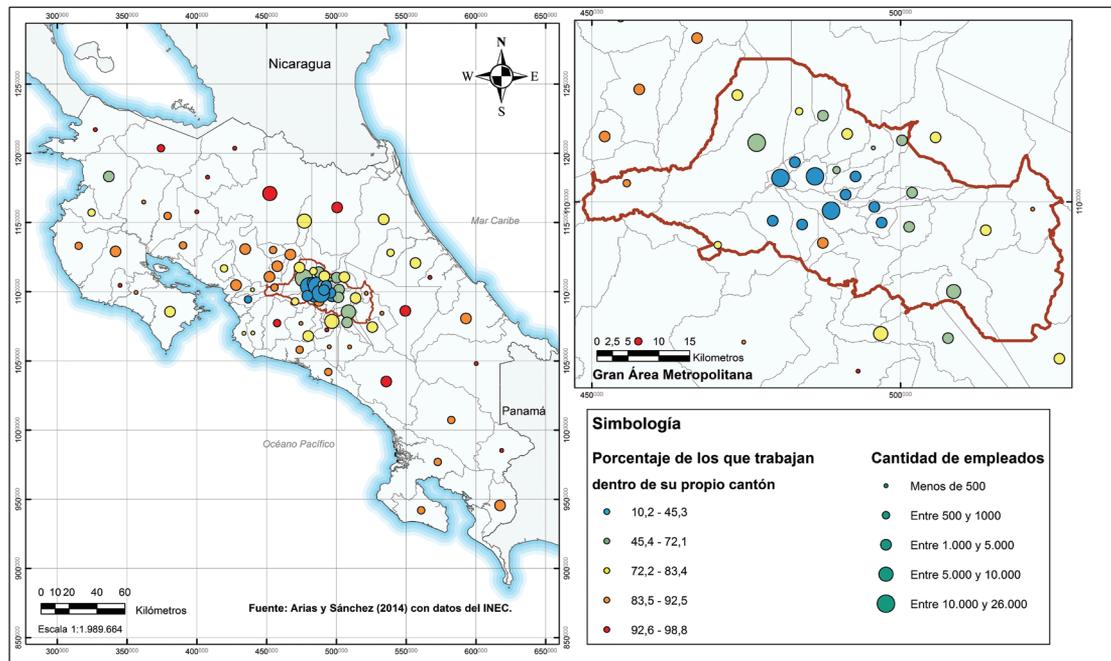
En economía regional es importante identificar si el empleo que genera el cantón es realizado por las personas que viven en el municipio o si por el contrario gran parte es desempeñado por personas que provienen de otros lugares fuera del cantón. La matriz origen destino de empleo desarrollada permite evaluar esta situación, la cual se trató de mapear en la figura 6. Algunos aspectos relevantes son:

En 28 de los 81 cantones del país más del 90% de los trabajos que se demandan para industria son desarrollados por personas que residen en el mismo municipio. Los valores porcentuales más altos se presentan en Hojan-cha (98,8%), Coto Brus (97,6%), Pérez Zeledón

(97,4%), Los Chiles (97,3%), Tilarán (97,2%), Guatuso (97,1%), San Carlos (96,7%), La Cruz (96%), Puriscal (96%), Upala (95,8%) y Sarapiquí (95,7%), (puntos color rojo en el mapa de la figura 6).

Por el contrario, solo en 11 cantones menos del 50% de la demanda de empleo industrial es cubierta por personas que residen en el mismo cantón. Con excepción de Orotina (43,5%), los restantes 10 municipios pertenecen a la GAM. Destacan el cantón de Belén, donde solo el 10,2% del empleo demandado en industria es cubierto por residentes del mismo cantón y en menor medida, el cantón de Flores (29,2%) y Heredia (35,7%).

FIGURA 6
DEMANDA DE EMPLEO INDUSTRIAL PARA PERSONAS DEL MISMO CANTÓN, 2011



Fuente: Elaboración propia con datos del INEC 2012.

La escolaridad promedio de las personas que trabajan en industria es de 8,9 años. Sin embargo, este valor es muy heterogéneo entre los 81 cantones ya que 35 de ellos se ubican por encima del promedio nacional y la mayoría están localizados dentro del Gran Área Metropolitana (GAM). Algunas excepciones a lo anterior son Hojancha (14,1), Talamanca (11,3), San Mateo (11,1) y Nandayure (10,5). Por el contrario, los valores más bajos se ubican en zonas rurales fronterizas como Los Chiles (6,4), Upala (6,5) y Guatuso (7,2) o en cantones con ciudades intermedias y de gran población como San Carlos (7,4), Pérez Zeledón (7,8) y Turrialba (7,9).

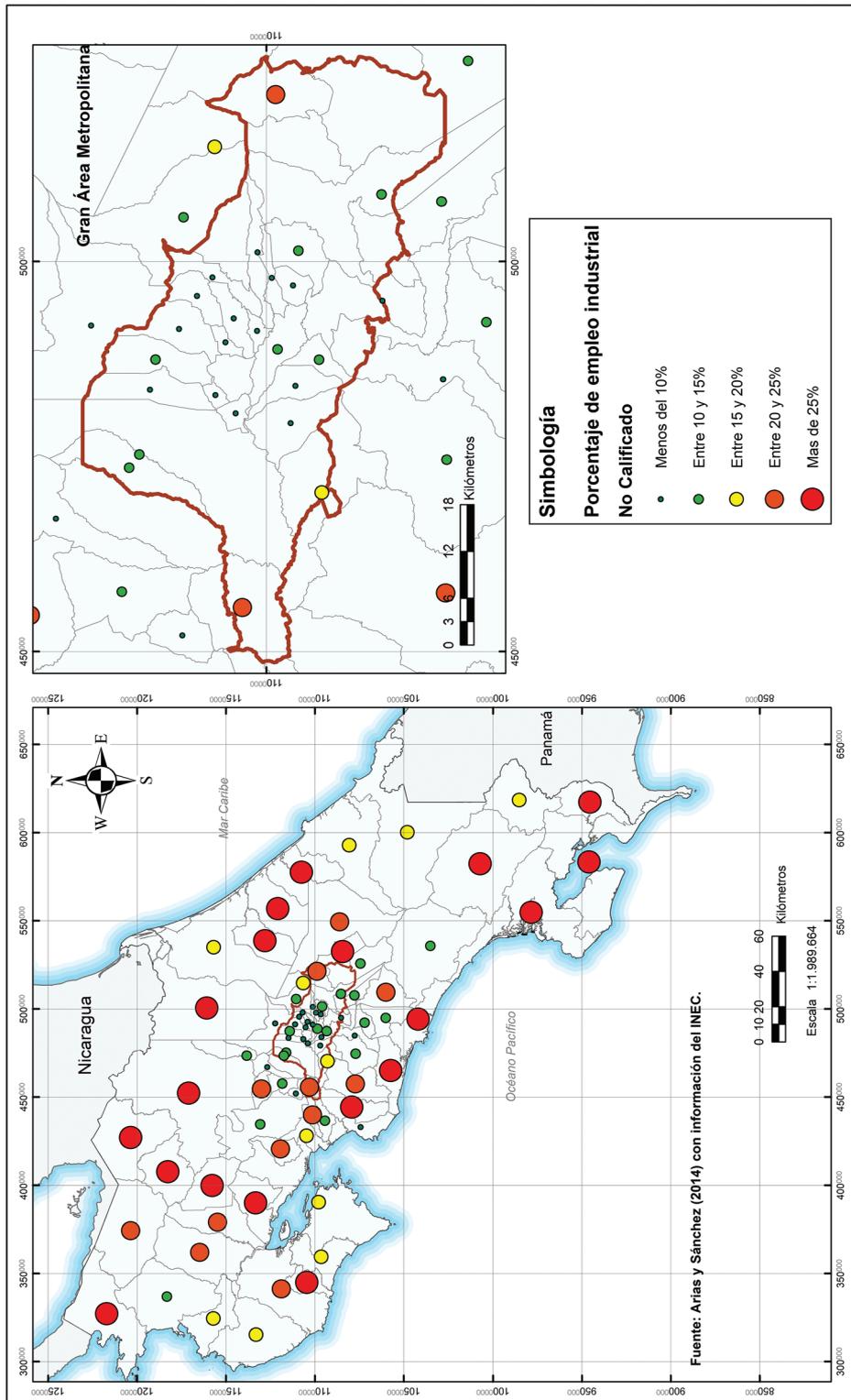
La figura 7 muestra la distribución espacial del empleo industrial no calificado por cantón. A nivel nacional los datos censales muestran que alrededor del 13% del empleo industrial era no calificado en el año 2011. Sin embargo, a nivel espacial varía mucho entre un cantón y otro. Algunos aspectos importantes que resaltan del mapa en la figura 7 son:

- Solo en 20 cantones de los 81 el porcentaje de empleo industrial no calificado es menor al 10% (bolas de menor tamaño color

verde en el mapa). Con excepción de Garabito y Palmares, el resto de estos cantones se localizan en la GAM, es decir, la GAM no solo acumula la mayor cantidad de empleos industriales sino que también los de mejor calificación.

- En 19 cantones (bolas de mayor tamaño color rojo en el mapa), al menos uno de cada 4 empleos es no calificado. La situación más crítica se presenta en los cantones de Turrubares donde alcanza el 41%, Los Chiles (37,6%), Parrita (37,4%), Guatuso (37,4%), Jiménez (36,4%), Sarapiquí (34,8%), Aguirre (34,4%), La Cruz (33%) y Corredores (33%). En estos municipios uno de cada tres empleos generados en industria es no calificado.
- Existe un grupo de 8 cantones que no solo generan un cantidad absoluta de empleo industrial muy baja (menor a 500 trabajos) sino que ese empleo que generan es de muy baja calificación, ya que uno de cada tres empleos es no calificado. Entre ellos están Turrubares, Hojancha, Guatuso, La Cruz, Los Chiles, Matina, Tillarón y Jiménez.

FIGURA 7
 PORCENTAJE DE EMPLEO INDUSTRIAL NO CALIFICADO POR CANTÓN,
 AÑO 2011



Fuente: Elaboración propia con datos del INEC (2012).

3. Resultados de auto-correlación espacial global

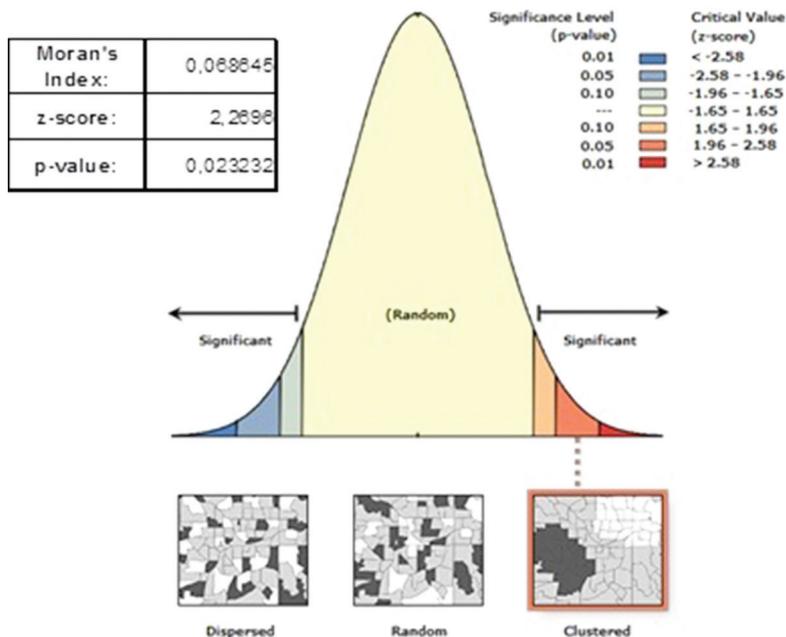
Como se mencionó anteriormente, una matriz de contigüidad se puede construir utilizando diferentes criterios. Para el cálculo de la I de Moran se utiliza una matriz booleana W basada en criterios de adyacencia. Tal que el valor de w_{ij} es 1 cuando dos cantones comparten una frontera común y cero en caso contrario. Los elementos de la diagonal principal son nulos.

Siguiendo a Mayor y López (2005) para determinar la significancia estadística de la I de Moran se calcula un estadístico $z(I)$ bajo el supuesto de aleatoriedad. La normalidad de este estadístico “depende del número de vínculos considerados y de cómo están conectados, es decir, de la estructura de la matriz de pesos espaciales, de forma que con 20 localizaciones puede asumirse normalidad.”

La matriz de pesos binaria presenta algunas limitaciones. Entre ellas la no inclusión de relaciones asimétricas, que es un requisito incluido en los cinco principios establecidos por Paelink y Klaasen, que se completan con los siguientes: interdependencia, asimetría, alotopía, no linealidad e inclusión de variables topológicas. Para corregir este problema se introduce una matriz de pesos espaciales.

De esta manera los resultados de la I de Moran para el caso de Costa Rica vienen dados por la figura 9. Los resultados de la I de Moran a nivel cantonal sugieren la existencia de auto-correlación espacial positiva. El estadístico $z(I)$ es significativo a un nivel de significancia del 3%. Esto sugiere que los cantones con un elevado (bajo) empleo industrial se encuentran cerca de otros cantones con un empleo relativamente alto (bajo).

FIGURA 8
RESULTADOS TEST DE AUTO-CORRELACIÓN ESPACIAL PARA COSTA RICA (EMPLEO INDUSTRIAL)



Fuente: Elaboración propia con datos del INEC (2012).

4. Resultados de auto-correlación espacial local

A continuación se muestran los resultados asociados a los Mapas de Cluster según la metodología de Anselin (1995). Como se mencionó anteriormente, este método sirve para poder identificar, cuantitativa y gráficamente, clusters de empleo industrial a nivel local, que son difíciles de apreciar a nivel global. El atributo considerado en este caso fue la cantidad de trabajadores en el sector industrial por cantón.

Es decir, se analiza la correlación geográfica entre el valor de este atributo en cierta zona versus el valor promedio de la vecindad, ponderado a través de la matriz *W*. Como se va a analizar el comportamiento propio de cada rubro, la matriz *W* que se usa debe contener valores nulos en la diagonal, para así no duplicar valores.

Para construir estos Mapas de Cluster, fue necesario obtener antes la *I* de Moran Local, que fue analizada en la sección metodológica. Para este análisis fue considerada la variable *Z* (incluida en la metodología) asociada al número de empleos de cada cantón, la cual fue normalizada dividiendo cada valor por la desviación estándar de la serie.

Luego, para cada atributo (número de empleo industrial) se estimó la variable *Z* en cada cantón versus los valores promedio de esta variable en la vecindad de cada cantón. Con esta información es posible saber lo que está ocurriendo en cada cantón según el valor *Z* obtenido al estimar la *I* de Moran Local. De esta forma, cantones con altos valores del atributo en estudio y que a su vez estén rodeados por cantones que también tienen altos valores, se les ha asignado color negro en el Mapa de Cluster.

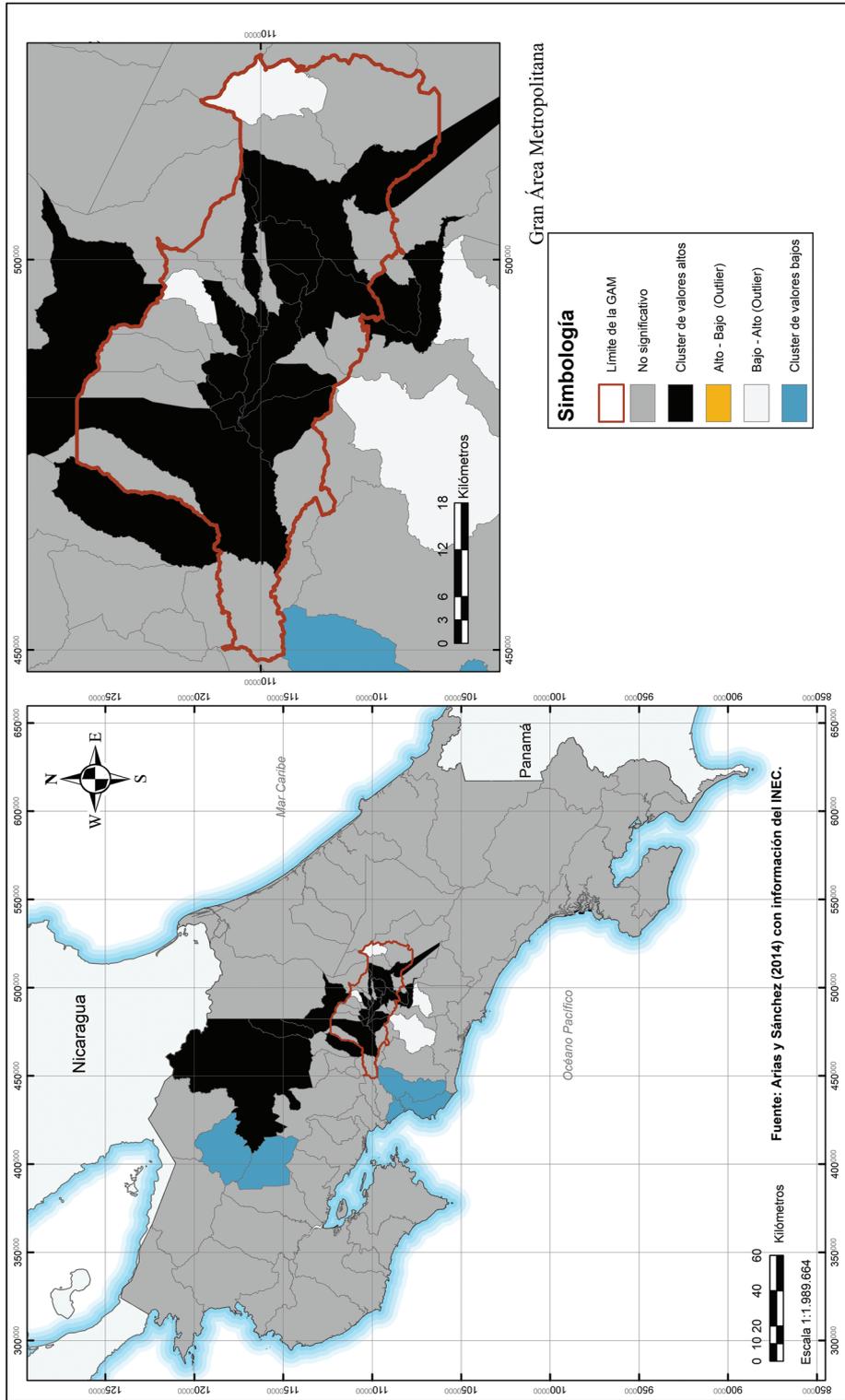
Por su parte, el color azul se asigna a cantones de bajos valores rodeados por cantones de bajos valores también. El resto de los cantones que muestran casos aislados que no siguen un patrón definido, por ejemplo cantones de

bajos valores rodeadas por altos valores (color blanco), o zonas de altos valores rodeadas por bajos valores son coloreados en amarillo.

Los resultados se muestran en el mapa de la figura 9 y en el cuadro 2. Algunos aspectos a resaltar son los siguientes:

- a) La forma en cómo se distribuye la demanda por empleo industrial en los cantones de Costa Rica no es aleatoria, sino que sigue un patrón espacial definido. Existen claros conglomerados de cantones con *alta* generación de empleo industrial en Costa Rica. Los cantones en color negro en el mapa de la figura 10 muestran esta situación.
- b) Estos conglomerados de valores altos están conformados por 16 cantones los cuales en su mayoría (14 cantones) están localizados dentro del Gran Área Metropolitana.
- c) El cluster de valores altos de generación de empleo industrial lo conforman los cantones de Alajuela, Heredia, San José, Belén, Santa Ana, Escazú, Flores, San Isidro, San Pablo, Goicohechea, La Unión, Curridabat, Desamparados y Cartago dentro de la GAM así como Grecia y San Carlos fuera de esta.
- d) Existen dos conglomerados de valores bajos (valor celeste en el mapa), es decir asociaciones estadísticamente significativas de cantones con bajos valores en generación de empleo industrial. Estas zonas están compuestas por los cantones de Guatuso y Tilarrán en la zona norte del país y por los cantones de Garabito, Turrrubares y Esparza en el Pacífico Central.
- e) Se localizan 4 zonas de valores bajos rodeados de valores altos (color blanco en el mapa). Estos cantones son Acosta y León Cortes, así como Alvarado en Cartago y San Isidro en Heredia.
- f) Por último, se puede identificar que cerca de dos terceras partes de los cantones no muestran asociación espacial significativa dada la variabilidad de valores que se presentan.

FIGURA 9
CONGLOMERADO DE EMPLEO INDUSTRIAL EN LOS 81 CANTONES DE COSTA RICA



Fuente: Elaboración propia con datos del INEC (2012).

CUADRO 2
VALORES DEL TEST DE AUTO-CORRELACIÓN LOCAL

| Z-Score (local) | P-Value (local) | Tipo de cluster |
|-----------------|-----------------|--------------------------------------|
| 0.4470 | 0,6549 | No significativo |
| -2.1998 | 0,0278 | Valores bajos cerca de valores altos |
| 0.8645 | 0,3873 | No significativo |
| 3.7247 | 0,0002 | Valores altos |
| 0.8830 | 0,3772 | No significativo |
| -1.1288 | 0,2590 | No significativo |
| -3.1342 | 0,0017 | Valores bajos cerca de valores altos |
| 0.3229 | 0,7468 | No significativo |
| -0.7425 | 0,4578 | No significativo |
| 1.1527 | 0,2491 | No significativo |
| 0.6345 | 0,5258 | No significativo |
| 7.5453 | 0,0000 | Valores altos |
| 0.2706 | 0,7867 | No significativo |
| 0.9153 | 0,3600 | No significativo |
| 0.4160 | 0,6774 | No significativo |
| 3.1683 | 0,0015 | Valores altos |
| -0.2207 | 0,8253 | No significativo |
| -0.2326 | 0,8161 | No significativo |
| 2.5245 | 0,0116 | Valores altos |
| 4.5003 | 0,0000 | Valores altos |
| -0.2181 | 0,8273 | No significativo |
| 1.7916 | 0,0732 | No significativo |
| 3.3677 | 0,0008 | Valores altos |
| -0.2020 | 0,8399 | No significativo |
| 3.6637 | 0,0002 | Valores altos |
| 2.2171 | 0,0266 | Valores bajos |
| 2.4279 | 0,0152 | Valores altos |
| -0.0897 | 0,9285 | No significativo |
| 3.9969 | 0,0001 | Valores altos |
| -0.0861 | 0,9314 | No significativo |
| 2.2008 | 0,0278 | Valores bajos |
| 6.5631 | 0,0000 | Valores altos |
| 0.8217 | 0,4112 | No significativo |
| -1.7365 | 0,0825 | No significativo |
| -0.2326 | 0,8161 | No significativo |
| 2.6382 | 0,0083 | Valores altos |
| -1.9993 | 0,0456 | Valores bajos cerca de valores altos |
| -0.3159 | 0,7521 | No significativo |
| -1.5586 | 0,1191 | No significativo |
| 1.4946 | 0,1350 | No significativo |
| -1.2980 | 0,1943 | No significativo |
| 0.6798 | 0,4966 | No significativo |
| -0.0279 | 0,9777 | No significativo |
| -1.5214 | 0,1282 | No significativo |
| 0.4684 | 0,6395 | No significativo |
| 0.6134 | 0,5396 | No significativo |
| 0.5528 | 0,5804 | No significativo |
| -0.3258 | 0,7446 | No significativo |
| 0.5524 | 0,5807 | No significativo |
| 0.8014 | 0,4229 | No significativo |
| 0.2706 | 0,7867 | No significativo |
| 0.7657 | 0,4438 | No significativo |
| 0.0828 | 0,9340 | No significativo |
| 1.3266 | 0,1846 | No significativo |
| -1.0962 | 0,2730 | No significativo |
| 0.7590 | 0,4478 | No significativo |
| -0.2204 | 0,8255 | No significativo |
| -0.9842 | 0,3250 | No significativo |
| 0.3339 | 0,7384 | No significativo |
| 2.6377 | 0,0083 | Valores altos |
| -3.5300 | 0,0004 | Valores bajos cerca de valores altos |
| 6.8687 | 0,0000 | Valores altos |
| 0.3842 | 0,7008 | No significativo |
| -1.7864 | 0,0740 | No significativo |
| 0.6055 | 0,5448 | No significativo |
| -0.1720 | 0,8634 | No significativo |
| 3.3358 | 0,0009 | Valores altos |
| -1.8254 | 0,0679 | No significativo |
| 0.1465 | 0,8835 | No significativo |
| 2.3206 | 0,0203 | Valores altos |
| 0.5830 | 0,5599 | No significativo |
| -0.1217 | 0,9031 | No significativo |
| -1.0962 | 0,2730 | No significativo |
| -0.3330 | 0,7392 | No significativo |
| 3.0957 | 0,0020 | Valores altos |
| 1.9965 | 0,0459 | Valores bajos |
| -1.3930 | 0,1636 | No significativo |
| 2.1183 | 0,0342 | Valores bajos |
| 0.6310 | 0,5281 | No significativo |
| 0.6892 | 0,4907 | No significativo |
| 0.6704 | 0,5026 | No significativo |

Fuente: Elaboración propia con datos del INEC (2012).

V. CONCLUSIONES

El estudio demuestra que la dimensión espacial es de gran relevancia para entender la generación de empleo industrial en los cantones de Costa Rica. Se concluye que el territorio definido como el Gran Área Metropolitana (GAM) juega un papel preponderante en la localización y concentración del empleo industrial. Por otra parte, el gran tamaño territorial y la escasa población de las regiones periféricas, limita las posibilidades de obtener economías de aglomeración. Esto es claro al constatar que en estas regiones, un porcentaje mayoritario reside en zonas rurales y sobre todo en zonas rurales dispersas.

En concordancia con la literatura económica, el estudio muestra que en algunas ocasiones parece razonable realizar el análisis a escala local y, por tanto, el cantón puede ser la unidad más adecuada. Sin embargo, por otro lado, es evidente que los límites administrativos municipales pueden ser, en algunos casos, demasiado restrictivos y no recoger el área local económicamente representativa. En ese sentido se podría indagar con unidades geográficas intermedias que pueda superar estas limitaciones. Para solucionar en parte el problema se han presentado los índices de dependencia espacial, enmarcados en las técnicas de Econometría Espacial, que permiten contrastar la existencia de agrupaciones de cantones en las que se detecta una elevada concentración del empleo industrial. En los casos en los que estos índices resultan significativos se puede interpretar que el municipio no es la unidad de análisis adecuada ya que la concentración de la actividad va más allá de los límites administrativos municipales y, por tanto, debe considerarse un área de tamaño superior. Sin embargo, para aquellas actividades en las que los resultados indican una ausencia de dependencia o auto-correlación espacial, puede afirmarse que no existen agrupaciones cantonales especializadas y que, por tanto, en estos casos el área municipal podría ser la más idónea.

Un resultado que tiene conformidad con un planteamiento teórico está relacionado con la detección de fuerzas centrífugas propuestas

por Fujita y Thisse (1996). En este sentido, no sorprendieron algunas correlaciones encontradas entre cantones de alta generación de empleo industrial, lo cual puede sugerir el fácil acceso a mano de obra calificada o acceso a proveedores e infraestructura. Otra fuerza centrífuga presente es la cercanía a las residencias, ya que al comparar localización del empleo industrial con el mapa de población se nota una fuerte relación.

Los análisis efectuados por medio de los Índices de Moran, y los Mapas de Cluster, muestran que varios cantones experimentan una fuerte aglomeración de empleo industrial. Sin embargo, hay que realizar nuevas investigaciones en este sentido para analizar la dinámica entre estos cantones.

VI. REFERENCIAS

- Anselin, L., Florax, R.J.G.M., Rey, S. (eds). (2004). *Advances in Spatial Econometrics, Methodology, Tools and Applications*. Berlin: Springer.
- Anselin, L., Florax, R.J.G.M. (eds). (1995). *New Directions in Spatial Econometrics*. Berlin: Springer.
- Anselin, L. (1988). *L. Spatial Econometrics: Methods and Models*. Boston, MA: Kluwer Academic.
- Anselin, L. (1992). "Space Stat tutorial. A workbook for using SpaceStat in the analysis of spatial data". Technical Report S-92-1, National Center for Geographic Information and Analysis, University of California. Santa Barbara, CA.
- Anselin, L. (1995). Local indicators of spatial Association-LISA. *Geographical Analysis*, vol. 27, n° 2, p. 93-115.
- Anselin, L. (2002) "Under the hood. Issues in the specification and interpretation of spatial regression models". *Agricultural Economics* 27; pp. 247-267.
- Anselin, L. (2003). *GeoDa 0.9 User's Guide*. Spatial Analysis Laboratory, University of Illinois, Urbana-Champaign, IL.
- Anselin, L. y R. Florax (1995). *New Directions in Spatial Econometrics*. Springer- Dordrecht: Reidel. Verlag, Berlin, Alemania.

- Arias R y Sánchez L. (2012). *Dinámica espacial del empleo industrial en la Gran Área Metropolitana de Costa Rica*. Documento publicado en la Revista de Ciencias Económicas de la Universidad de Costa Rica. 2012. Volumen 30, número 2, Julio-Diciembre 2012.
- Arias R. y Sánchez L. (2011). *Análisis de la dinámica regional del empleo utilizando el modelo shift share espacialmente modificado: el caso de la Región Chorotega, 1990-2009*. Documento Publicado en la Revista de Ciencias Económicas de la Universidad de Costa Rica. 2011. Volumen 29, número 2, Julio-Diciembre 2011.
- Arias R. y Sánchez L. (2011). *Concentración territorial de las empresas en la región chorotega: un análisis mediante técnicas de economía espacial*. Documento Publicado en la Revista de Ciencias Económicas de la Universidad de Costa Rica. 2011. Volumen 29, número 2, Julio-Diciembre 2011.
- Arias R; Sánchez L y Oviedo L. (2012). *Encuesta sobre producción, perfil del recurso humano y necesidades de capacitación en la región Chorotega*. Documento publicado en la Revista de Ciencias Económicas de la Universidad de Costa Rica. 2012. Volumen 30, número 1, Enero-Junio 2012.
- Cliff, A., Ord, J. (1972). Testing for spatial autocorrelation among regression residuals. *Geographical Analysis*, vol. 4, p. 267-284.
- Cliff, A., Ord, J. (1973). *Spatial autocorrelation*. London: Pion.
- Cliff, A., Ord, J. (1981). *Spatial processes, models and applications*. London: Pion.
- Cressie, N.A.C. (1991). *Statistics for Spatial Data*. John Wiley and Sons, New York.
- Fotheringham, S., Brundson, C. y Charlton, M., 2000, *Quantitative Geography: Perspectives on spatial data analysis*, Sage, Gran Bretaña
- Fujita, M y Thisse, J.F. (1996). Economics of Agglomeration. *Journal of the Japanese and International Economies* 10, 339 – 378.
- Geary, R. (1954). The contiguity ratio and statistical mapping, *The Incorporated Statistician*, vol. 5, p 115-145.
- Getis, A., Mur, J., Zoller, H.G. (eds). (2004). *Spatial econometrics and spatial statistics*. New York: Palgrave Macmillan.
- Goodchild, M. (1987). A spatial analytical perspective on geographical information systems. *International Journal of Geographical Information Systems*, 1, 327-334.
- Growth and Change*, nº 18, spring, p. 69-78.
- Haining, R. *Spatial Data Analysis. Theory and Practice*. Cambridge: Cambridge University Press, 2003.
- Instituto Nacional de Estadística y Censo. (2012). X Censo Nacional de Población y VI de Vivienda Resultados Generales.
- Krugman, P. (1991a). Increasing Return and Economic Geography. *Journal of Political Economy*.
- Krugman, P. (1991b). *Geography and Trade*. MIT Press, Cambridge, MA, USA.
- Krugman, P. (1998). Space: the Final Frontier. *Journal of Economic Perspectives* 12, pp.161-174
- Moran, P. (1948). The interpretation of statistical maps. *Journal of the Royal Statistical Society B*, vol. 10, p. 243-251.
- Moreno, R y Vayá. E. (2000): *Técnicas Econométricas para el tratamiento de datos espaciales: La econometría espacial*. UB 44 manuals, Edicions Univeritat de Barcelona.
- Pinkse J. y M.E. Slade (1998). Contracting in space: An application of spatial statistics to discrete-choice models. *Journal of Econometrics*, Vol. 85, 1, pp. 125-154.
- Pujol R, Sánchez L y Pérez E. (2012). *Growth patterns and concentration of urban activities in the Greater Metropolitan Area of Costa Rica, 1993-2010*. Documento publicado en la Revista Reflexiones Facultad de Ciencias Sociales de la Universidad de Costa Rica. 2012. Volumen 91, número 1. 2012.

- Rey, S. J. (1999), "Spatial Empirics for Economic Growth and Convergence", Mimeo, UCSD, USA.
- Rey, S. J. and Montouri B. D. (1999), "US Regional Income Convergence: A Spatial Econometric Perspective". *Regional Studies*, 33(2), 143-156.
- Ripley, B.D. (1981). *Spatial Statistics*. New York: John Wiley & Sons.
- Sánchez, L. Murillo, D y Arguedas, J (2008). *Análisis de Localización, Competitividad y Especialización de las Actividades Económicas a nivel subregional*. Trabajo Final de Graduación. San José, Escuela de Economía, Universidad de Costa Rica.
- Tobler, W. (1979). *Cellular Geography. Philosophy in Geography* (ed. Gale, S., Olsson, G).
- Trejos, J (2002). *Mercado de Trabajo y estructura productiva regional: una descripción a partir de los censos de población*.

