

# **ANÁLISIS ESPACIAL DE LA DESERCIÓN ESTUDIANTIL EN COLEGIOS PÚBLICOS Y PRIVADOS DE COSTA RICA EN EL AÑO 2016**

Moisés Coto Hernández  
Melissa Valverde Hernández  
Universidad de Costa Rica

## **RESUMEN**

En el presente estudio se realiza un análisis estadístico espacial de la deserción escolar en colegios públicos y privado. El objetivo del estudio consiste en observar si la deserción escolar en colegios públicos y privados existe una asociación espacial por Dirección Regional y Cantón. Se utilizan los datos del Ministerio de Educación Pública para la deserción en colegios en el año 2016. El procedimiento estadístico utilizado consistió en realizar un análisis espacial mediante modelos de estadísticas de áreas, inicialmente se muestra que la deserción presenta una asociación espacial con los cantones y dirección regional, sin embargo, en los modelos únicamente a nivel cantonal la distribución de los residuos presentan asociación espacial. Se estiman los modelos espaciales CAR y SAR, los coeficientes relacionados a las variables al tipo de zona de centro educativo y modalidad horaria nocturna son significativos al 10%. Finalmente, el modelo de mejor ajuste es el lineal multivariante por mínimos cuadrados ordinarios y los cantones con menos deserción se encuentran en la zona costera y fronteriza del país.

## **INTRODUCCIÓN**

La deserción estudiantil plantea al sistema educativo la necesidad de buscar alternativas que permitan disminuirla y dar respuesta a las necesidades económicas, sociales, afectivas y psicológicas de los estudiantes. Para el logro de estos objetivos es indispensable que el esfuerzo institucional organizado de manera sistemática, se una a la puesta en práctica de acciones educativas e innovadoras dirigidas a satisfacer las necesidades básicas que favorezcan la construcción de aprendizaje significativos que amplíen las oportunidades para que los estudiantes se desenvuelvan en la sociedad.

El abandono estudiantil, es un problema que enfrentan muchos países y teniendo en consideración que la educación es un agente de movilidad social es importante comprender las razones que inciden en los porcentajes de deserción y de esa forma tratar de implementar medidas que no solo prevengan el abandono, si no más bien que motive el desarrollo integral de los estudiantes.

Este fenómeno implica que una porción importante de la población que se ha valorado como un factor de desarrollo humano fundamental, para forjarse una calidad de vida óptima por un

lado y favorecer el desarrollo nacional interrumpe sus estudios o se retira permanentemente de ellos. En nuestro país los colegios Académico Nocturnos presentan la mayor problemática, con índices de deserción de hasta un 27% alcanzado en el 2007, Jiménez et al, (2010) y lo más preocupante es que alrededor de dos terceras partes de los jóvenes en edad escolar no logra terminar la secundaria, como lo reconoce el Plan Nacional de Desarrollo 2006-2010 (PND-MIDEPLAN, 2006):

“Costa Rica mantiene a dos terceras partes de sus jóvenes por debajo de ese piso: no logran terminar la secundaria. De esta forma, la incapacidad del sistema educativo provoca tanto un impacto negativo sobre el potencial de crecimiento del país, al no generar el tipo de recurso humano capaz de acceder a los empleos más productivos que requiere el país; pero provoca un impacto aún más peligroso en los procesos de integración social, ya que solamente abre oportunidades de un trabajo y una vida mejor para una tercera parte de la población, excluyendo de esas oportunidades –de esos derechos al resto. En tales condiciones, la educación amenaza convertirse en un sinsentido: frena el crecimiento y promueve la desigualdad”

Los avances tecnológicos han provocado cambiar la forma de medición en algunos ámbitos. El análisis de datos espaciales se ha vuelto muy común en diversas áreas de las ciencias, el realizar pronósticos climatológicos, comportamientos de especies o animales y consumo o rendimiento de la agricultura hasta en aspectos sociales, son algunos ejemplos en los que se considera la estadística espacial. El objetivo del presente informe consiste en conocer si la deserción estudiantil en colegios públicos y privados tienen una distribución espacial por Dirección Regional y Cantón.

En el tema escolar, Cervera L et al (2008), presentan una investigación donde el objetivo principal “es la descripción de la base de datos obtenida de la aplicación de ENLACE en el Municipio de Juárez y la relación del nivel de aprovechamiento de los alumnos examinados con las variables socioeconómicas que componen el entorno de los centros escolares. El propósito es analizar si esta evaluación es estadísticamente robusta para representar los problemas de educación en este municipio y de proveer elementos analíticos que sirvan para mejorar sus condiciones de educación. Lo anterior se llevó a cabo mediante la georreferenciación de los datos de ENLACE por centro escolar: primaria y secundaria. Esto incluye la producción de cartografía temática de las variables principales.”

Sobre este tema, Steinberg C (2010), en Buenos Aires, Argentina, desarrolló un trabajo que presenta los resultados de un ejercicio realizado en el marco del Programa de Estudios sobre la heterogeneidad de contextos educativos, el ejercicio estuvo orientado a construir una tipología de escenarios educativos en función del nivel de abandono que se registra entre las escuelas secundarias y en nivel de dispersión/concentración que se observa del fenómeno estudiado entre la oferta educativa en un territorio dado, utilizando estadística espacial.

También en Argentina, pero en Córdoba, Herrero V et al (2007), desarrolla un análisis espacial sobre la deserción estudiantil. “La información censal permite encontrar patrones de vulnerabilidad identificados como de vivienda precaria, carencia de capital humano, hogares con baja integración social y hogares compuestos o con jefas adolescentes. Los aglomerados obtenidos a partir de estos factores permiten observar una distribución espacial en forma de

anillos irregulares, donde la vulnerabilidad social crece desde el área central hacia la periferia.”

## MATERIALES Y MÉTODOS

El objetivo de esta investigación consiste en conocer si la deserción estudiantil en colegios públicos y privados de Costa Rica en 2016, existe una asociación espacial por Dirección Regional y/o Cantones. Para el desarrollo de la investigación se obtienen los datos del Ministerio de Educación Pública, correspondiente a la matrícula inicial y deserción estudiantil, en el año 2016, los cuales los centros educativos están georeferenciados en longitud y latitud.

Las variables por utilizar en la investigación son las siguientes:

Variable	Descripción
Región	Se refiere al nombre de la Dirección Regional a la que pertenece el colegio
Cantón	Se refiere al nombre del cantón al que pertenece el colegio
Sector	Tiene que ver con que si el colegio es público o privado
Zona	En donde se ubica el colegio, en zona rural o urbana
Rama	Tiene que ver con que si el colegio es académico o técnico
Horario	En este caso si el colegio tiene horario diurno o nocturno
Total Matrícula	Total de la matrícula para el 2016
Deserción	Deserción para el año 2016

Adicional a estas variables, se crearon dos más, la primera que tiene que ver con la diferencia entre la matrícula final e inicial y posteriormente se crea la variable tasa de deserción la cuál es un porcentaje, la diferencia entre matricula inicial por cien.

Posteriormente se realizó un análisis espacial de los datos, utilizando áreas donde se creó un archivo de polígonos espaciales que reúne las representaciones espaciales de los polígonos con datos. Las etiquetas de identificación de los polígonos en la ranura se combinan con los nombres de las filas del marco de datos para asegurarse de que las filas de datos correctas estén asociadas con el objeto espacial correcto. Como en este análisis espacial se realiza por áreas, se calcula las variables por el porcentaje de centros que son del sector público, de rama

académica, de horario nocturno y de zona urbana, y con la deserción se calcula el promedio por cada área.

Se realizó un análisis de dependencia espacial mediante la autocorrelación la cual mide el grado en que un fenómeno de interés se correlaciona consigo mismo en el espacio cuando los valores de los vecinos cercanos son similares o se observa un agrupamiento en el espacio, si la autocorrelación espacial es positiva, y en caso contrario cuando los valores vecinos son diferentes la autocorrelación espacial es negativa, y si la autocorrelación espacial es nula indica que el patrón espacial es aleatorio. Además, se crearon diferentes matrices de pesos, utilizando vecinos más cercanos utilizando criterios como la reina, el cual indica que dos unidades están cerca si comparten un lado o un borde, el criterio de la torre (dos unidades están cerca una de la otra si comparten un lado), además se utilizó la matriz de pesos basada en la distancia euclidiana tomando como referencia los vecinos que están dentro de un radio de 25 kilómetros, y finalmente se utilizó la distancia inversa.

Con relación a la correlación espacial, calculó la I de Moran, donde los valores oscilan entre -1 (indicando dispersión perfecta) a 1 (correlación perfecta). Un valor de cero indica un patrón espacial aleatorio y se realizó un contraste de las siguientes hipótesis:

$H_0$ : Los datos son espacialmente independientes, los datos observados se asignan al azar entre localidades.

$H_1$ : Los datos no son espacialmente independientes.

Esta medida de autocorrelación espacial se calcula para estudiar fenómenos estocásticos que se distribuyen en el espacio en dos o más dimensiones. La medida I de Moran viene dada por la siguiente expresión:

$$I = \frac{\sum_{i=1}^n \sum_{j=1}^n W_{ij} (z_i - \bar{z})(z_j - \bar{z})}{\sum_{i=1}^n \sum_{j=1}^n W_{ij} \sum_{i=1}^n (z_i - \bar{z})^2}$$

Donde  $n$  es el número de áreas y  $W_{ij}$  una matriz de pesos. En este caso se utilizaron las áreas a nivel cantonal y por Dirección Regional Escolar, se seleccionan estas dos áreas pues la primera es utilizada mayoritariamente para decisión política e indicadores de interés socioeconómico en el país y la segunda la establece el Ministerio de Educación Pública como parte de la toma de decisión administrativa en el sistema educativo nacional.

Posteriormente se estimaron los modelos espaciales autorregresivos (SAR) y el modelo autoregresivos condicionales (CAR), y se utilizó el índice de ajuste del AIC y BIC para determinar el modelo que mejor se ajusta a los datos.

Con respecto al modelo SAR, los términos de error se modelan con otra regresión y este modelo, la variable dependiente se distribuye como una normal multivariante, los errores y el modelo se denotan:

$$\epsilon_i = \sum_{j=1}^m \beta_{ij} \epsilon_j + \gamma_i. \quad Y = X^T \beta + B(Y - X^T \beta) + \gamma$$

$$(I - B)(Y - X^T \beta) = \gamma$$

Asimismo, el modelo CAR se basa en un modelo condicional por lo que es su estimación se centra en la mayoría en el enfoque bayesiano, la distribución condicional de los errores se denota como:

$$\epsilon_i | \epsilon_{i-1} \sim N \left( \frac{\sum_{j \sim i} c_{ij} \epsilon_j}{\sum_{j \sim i} c_{ij}}, \frac{\sigma_{\epsilon_i}^2}{\sum_{j \sim i} c_{ij}} \right)$$

Para este análisis estadístico se utilizó el lenguaje estadístico R versión 3.6.1, algunas de las librerías utilizadas fueron: sp, Rgdal, raster, spatstat, dplyr y maptools.

## RESULTADOS

### Análisis descriptivo estadístico y espacial

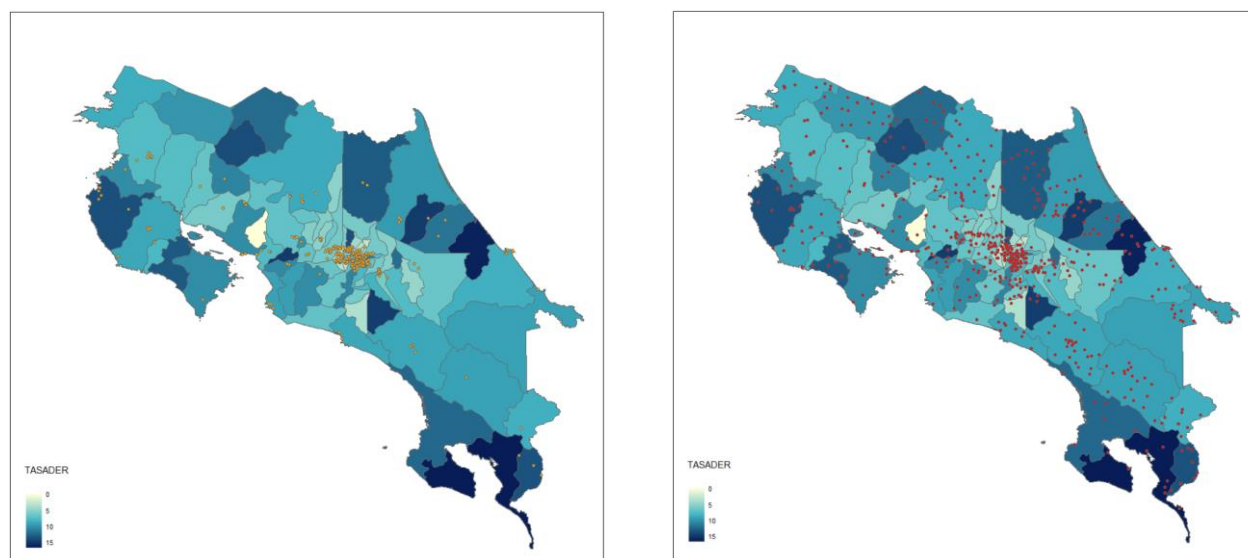
A partir de los datos de deserción del Ministerio de Educación Pública, la Tabla 1, presenta los principales descriptivos de las variables a utilizar, la mayoría de los colegios se encuentran en la provincia de San José, son del sector público, de la zona urbana, de la rama académica y modalidad de horario diurna.

**Tabla 1**  
**Costa Rica 2016: Frecuencia y porcentaje de centros educativos de secundaria según características del centro.**

<b>Provincia</b>	<b>n</b>	<b>%</b>
Alajuela	187	19,5
Cartago	82	8,6
Guanacaste	100	10,4
Heredia	85	8,9
Limón	97	10,1
Puntarenas	133	13,9
San José	275	28,7
<b>Sector</b>	<b>n</b>	<b>%</b>
Privado	208	21,7
Publico	751	78,3
<b>Zona</b>	<b>n</b>	<b>%</b>
Urbana	615	64,1
Rural	344	35,9
<b>Rama</b>	<b>n</b>	<b>%</b>
Académica	738	77,0
Técnica	221	23,0
<b>Horario</b>	<b>n</b>	<b>%</b>
Diurno	817	85,2
Nocturno	142	14,8
<b>Total</b>	<b>959</b>	<b>100,0</b>

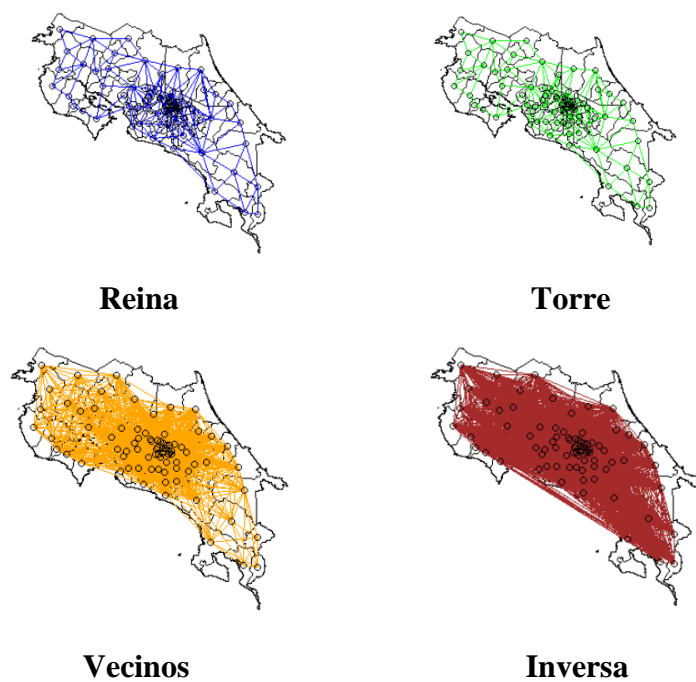
Fuente: Departamento de análisis estadístico del Ministerio de Educación Pública de Costa Rica.

**Figura 1**  
**Costa Rica: Mapas nivel de tasa de deserción estudiantil por Dirección Regional**  
**Escolar según ubicación de colegios privados y públicos**



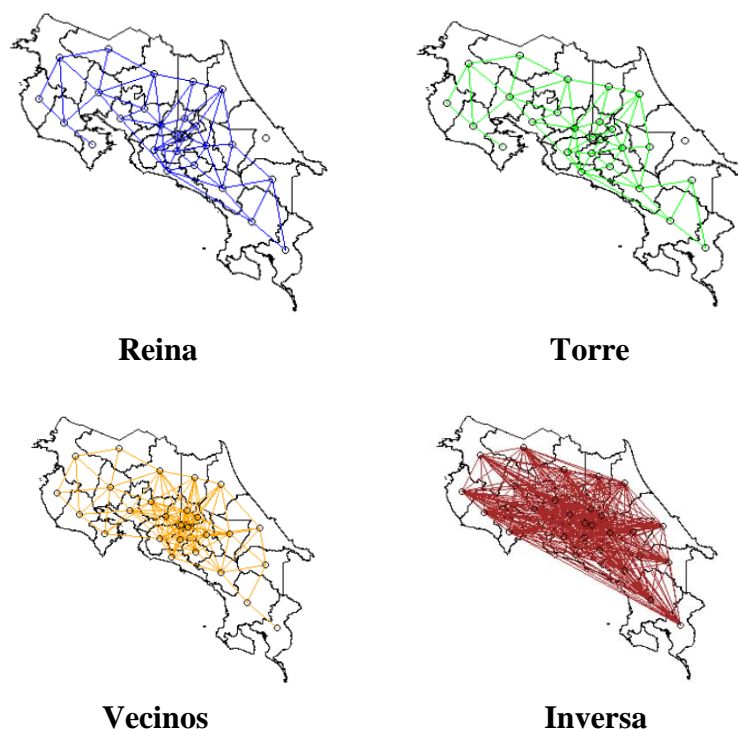
En la Figura 1, se muestra la distribución geoespacial de los centro públicos y privados por dirección regional escolar, se muestra que las tasas de deserción más altas se encuentran en las direcciones de provincias costeras y fronterizas, en particular la zona sur, caribe central y zona norte, presentan tasas entre 10% y 15%.

**Figura 2**  
**Costa Rica: Mapas matriz de ponderación de vecinos a nivel cantonal**



Para estimar la matriz de ponderaciones con respecto a las distancias, se realizaron varios criterios, se realizó a nivel cantonal y dirección regional, en ambas áreas por la información que muestra la matriz de distancia inversa y su significancia en la con la asociación espacial, se elige esta para estimación de la I de Moran y los modelos espaciales (Ver Figuras 2 y 3).

**Figura 3**  
**Costa Rica: Mapas matriz de ponderación de vecinos a nivel de Dirección Regional Escolar**

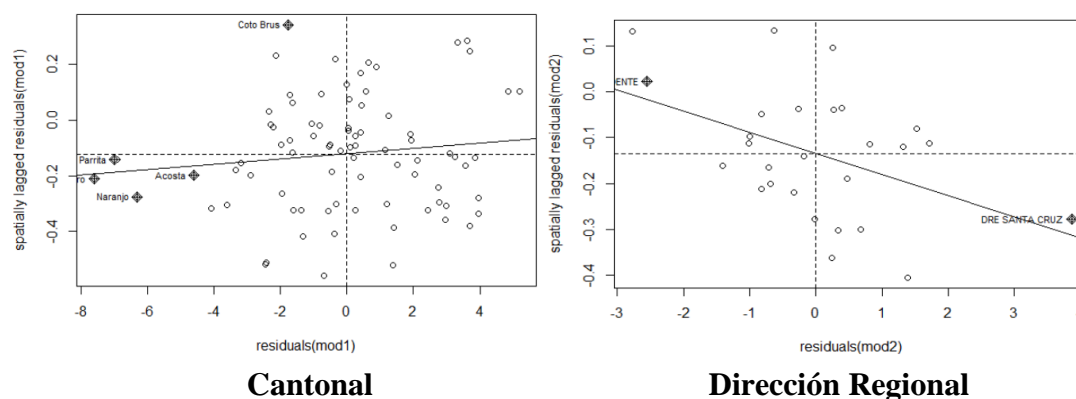


### **Análisis modelos**

Inicialmente se estima I de Moran a la variable deserción para contrastar la independencia espacial, a nivel cantonal se obtiene un estadístico de I Moran de 0.08 con  $p < 0.05$  y por dirección regional I Moran de 0.11 con  $p < 0.05$ . Estos resultados dan hincapié a estimar modelos lineales en donde se contemple el componente espacial, pues se rechaza la hipótesis que la tasa deserción esta distribuida aleatoriamente a nivel cantonal y dirección regional.

Por lo que posteriormente, se estima dos modelos lineales multivariante por mínimos cuadrados ordinarios y luego con la matriz de ponderación de distancia inversa se estima la I de Moran a los residuos de los modelos, uno donde se contemplan las áreas cantonales y el otro las direcciones regionales. Los resultados indican que a un nivel de significancia al 5% los residuos de la variable deserción presenta una asociación espacial a nivel cantonal, no así para las direcciones regionales escolares, los gráficos de la correlación se muestran en la Figura 4. A nivel cantonal se obtiene una correlación positiva de 0.09 con  $p < 0.05$ .

**Figura 4**  
**Correlación espacial I Moran de los residuos de los modelos para estimar la deserción**  
**estudiantil a nivel cantonal y Dirección Regional Escolar**



Por lo que a nivel cantonal se estiman los modelos espaciales CAR y SAR, en la Tabla 2 se muestran los coeficientes estimados, en los tres modelos el porcentaje de centros de zona urbana en el cantón el coeficiente es significativo al 10%, presenta un coeficiente negativo por lo que disminuye la tasa de deserción entre más centros urbanos se encuentren en el cantón. La variable relacionada al porcentaje de centros en modalidad nocturna el coeficiente es significativo al 5%, presenta un coeficiente con magnitud positiva por lo que cada aumento de un punto porcentual la deserción aumenta 0.19 puntos porcentuales. Asimismo, los tres modelos presentan magnitudes en los coeficientes muy similares (Ver Tabla 2). Además, con respecto a los criterios de ajuste AIC y BIC en los tres modelos son similares, en particular el modelo lineal multivariante presenta la menor magnitud entre los criterios, por lo que es el modelo que presenta mejor ajuste con AIC de 395,78 y BIC de 410,15.

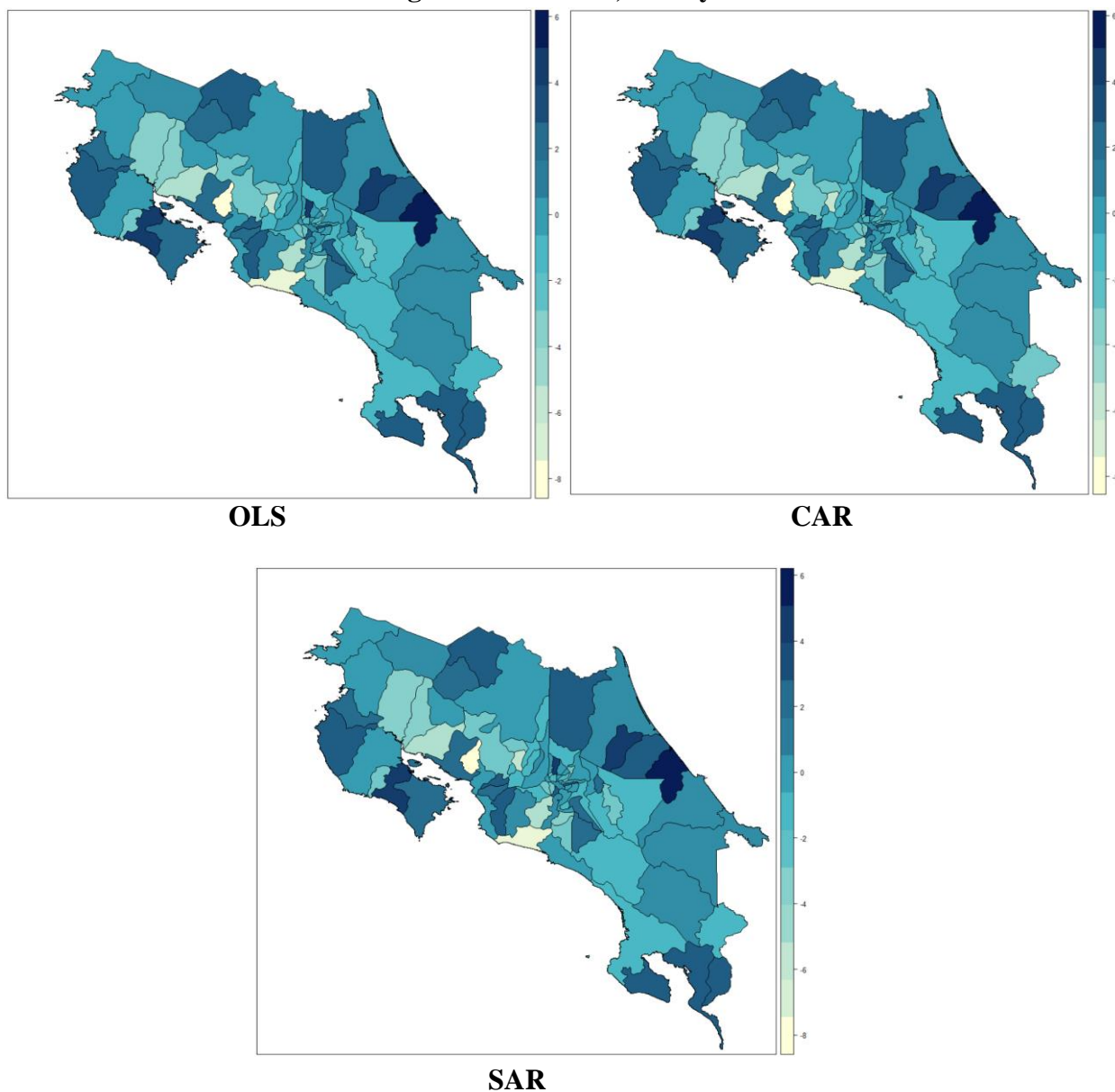
**Tabla 2**  
**Resultado de los modelos OLS, CAR y SAR, coeficientes estimados y criterios de**  
**ajuste**

Variable	Modelo		
	OLS	CAR	SAR
Intercepto	2,777	2,831	3,058
Zona Urbana (PZONAU)	-0,021*	-0,021*	-0,021*
Rama Académica (PRAMAAC)	0,028	0,027	0,025
Horario Nocturno (PHORANOC)	0,196**	0,195**	0,193**
Dependencia Pública (PPUB)	0,020	0,019	0,020
<b>Criterio</b>			
AIC	395,78	397,78	397,51
BIC	410,15	414,54	414,28

Nota: significancia al \* 0,1 , \*\* 0,05.

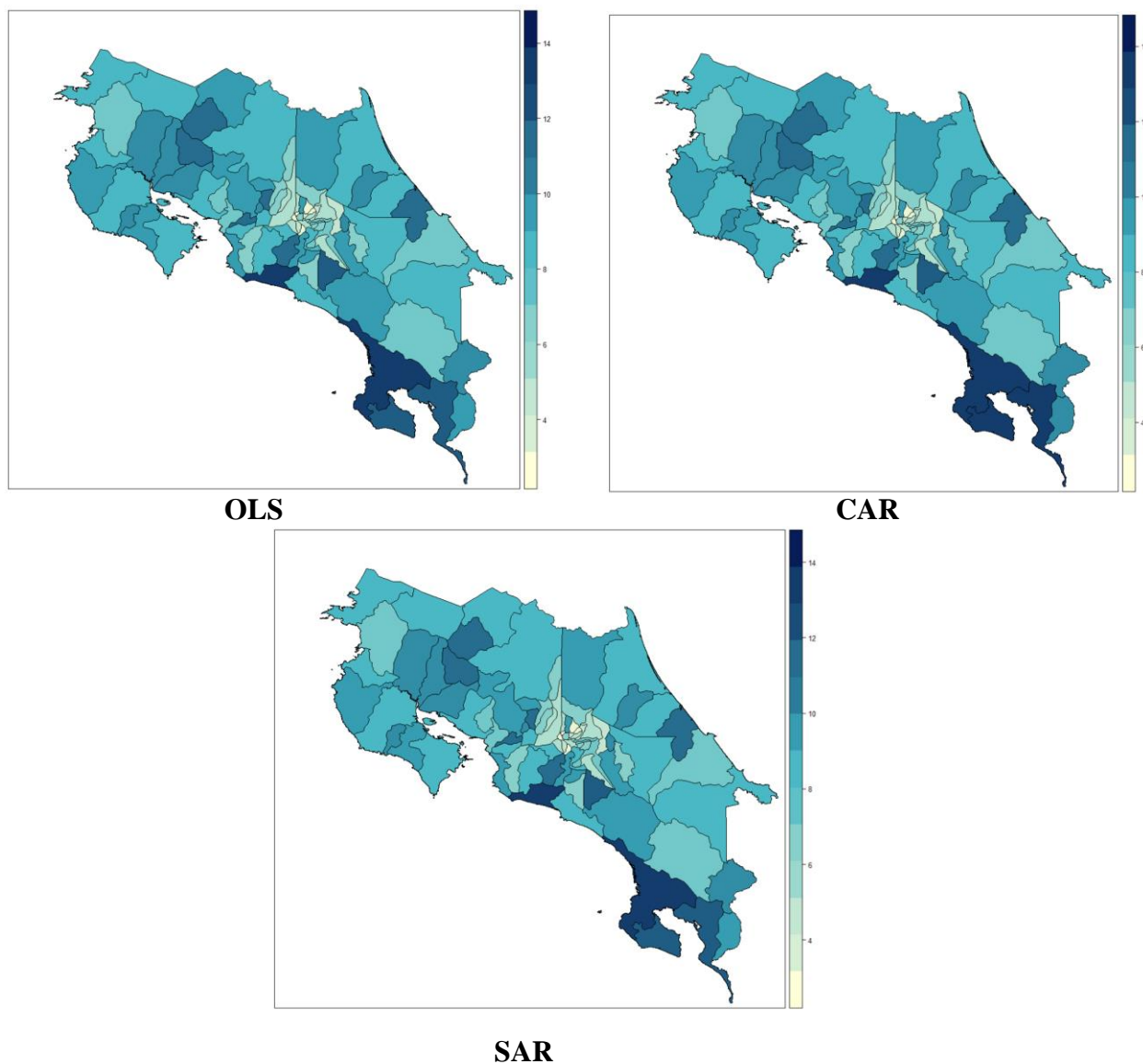


**Figura 5**  
**Costa Rica 2016: Mapas residuales de la tasa de deserción estudiantil estimada por**  
**cantón según modelo OLS, CAR y SAR**



Con respecto al análisis de residuos de los modelos estimados a nivel cantonal se muestran que los residuos con mayor diferencia se muestran en los cantones de la periferia del país, provincias costeras y fronterizas, en particular son las zonas donde se presenta la mayor tasa de deserción estudiantil. En los tres modelos presentan magnitudes de los residuos similares (Ver Figura 5).

**Figura 6**  
**Costa Rica 2016: Mapas tasa de deserción estudiantil estimada por cantón según**  
**modelo OLS, CAR y SAR**



En la Figura 6 muestran las tasas de deserción estimadas por cada uno de los modelos a nivel cantonal. Los tres modelos conservan la tasa de deserción observada a nivel cantonal, de igual forma se muestran tasas de deserción altas en cantones pertenecientes a las provincias costeras y fronterizas, en particular las menores tasas se muestran en los cantones centrales del país.

## CONCLUSIONES

Entre los principales hallazgos del análisis espacial se muestra que existe una asociación espacial leve de la deserción estudiantil de secundaria con los cantones. Además, se muestra que la deserción no presenta asociación espacial con las direcciones regionales. Por lo que si se tomaran acciones para reducir la deserción sería realizarlo en zonas focalizadas en áreas cantonales.

Además, se obtiene que la matriz de pesos con mejores resultados es la de distancia inversa. Con respecto a los coeficientes de los modelos, las variables relacionadas a la zona urbana y en la modalidad de horario nocturno presentan coeficientes significativos.

En análisis de los residuos no se presentan diferencias entre los modelos. Asimismo, con las magnitudes de los valores ajustados de los modelos y conserva el comportamiento de las tasas de deserción observadas en magnitud y en la distribución a nivel de áreas. De acuerdo con los criterios AIC y BIC los modelos presentan un ajuste similar entre sí, sin embargo, el modelo lineal multivariante de mínimos cuadrados ordinarios presenta el mejor ajuste, aunque por cuestiones de toma de decisión es importante considerar el componente espacial.

Para estudios futuros sería conveniente incorporar otras variables asociadas al contexto cantonal y a la dirección regional. Además, sería interesante incorporar modelos espacio temporales pues el Ministerio de Educación en algunos años ha realizado unas acciones para la permanencia de la población estudiantil, otro aspecto a considerar es contemplar estructuras jerárquicas a nivel de los datos dado que dependiendo de algunas divisiones administrativas en los centros educativos se puede llegar a niveles inferiores como los circuitos escolares o el mismo centro educativo como unidad georreferenciada.

## BIBLIOGRAFÍA

Cervera Gómez, L. E., Lizárraga Bustamante, G. M., & Sánchez Guillén, C. P. (2008). Estudio georreferencial de la Evaluación Nacional de Logro Académico en Centros Escolares (ENLACE) en el Municipio de Juárez, Chihuahua: análisis espacial. *Revista electrónica de investigación educativa*, 10(1), 1-23.

Herrero, V., Peláez, E., & González, L. (2007). Vulnerabilidad social, rezago y deserción escolar en el gran Córdoba (Argentina) 2001. In *IX Jornadas Argentinas de Estudios de Población*. Asociación de Estudios de Población de la Argentina.

Mideplan (2006). Plan Nacional de Desarrollo 2006-2010. Recuperado de: <https://www.mideplan.go.cr/plan-nacional-desarrollo-2006-2010>

Steinberg, C. (2010). Acceso y permanencia en una educación de calidad. La dimensión espacial del abandono escolar en las escuelas secundarias urbanas en Argentina: insumos para el planeamiento de las políticas de inclusión educativa. In *Trabajo presentado en el Congreso Iberoamericano de Educación. Metas* (Vol. 2021).