

Determinantes de la producción agrícola en Colombia: una aproximación aplicando
modelos de datos panel espaciales.

Jeymy Carolina Vargas Suárez
Administradora Pública

Diego Felipe López Ospina
Economista

Fundación Universitaria los Libertadores
Facultad de Ingeniería y Ciencias Básicas
Especialización en Estadística Aplicada
Bogotá D.C.
2017

Determinantes de la producción agrícola en Colombia: una aproximación aplicando
modelos de datos panel espaciales.

Jeymy Carolina Vargas Suárez
Administradora Pública

Diego Felipe López Ospina
Economista

Asesor
MSc. Juan Camilo Santana Contreras

Fundación Universitaria los Libertadores
Facultad de Ingeniería y Ciencias Básicas
Especialización en Estadística Aplicada
Bogotá D.C.
2017

Nota de Aceptación

Firma del presidente del jurado

Firma del Jurado

Firma del Jurado

Bogotá, D.C., 2 de diciembre de 2017

Resumen

En Colombia, la producción agrícola es un elemento importante en la dinámica económica del país. Si bien su participación como porcentaje del producto interno bruto cayó entre los años 2011 y 2013, las toneladas producidas no tienen fuertes variaciones entre los años 2007 y 2013. Abordando la producción agrícola agregada a nivel departamental, se observa que presenta algunos patrones de concentración en zonas como la pacífico y centro. Por esta razón se hace importante abordar el espacio dentro de las determinantes de la producción agrícola en el país.

El uso de datos panel como técnica de medición, constituye una buena herramienta para estudiar dinámicas de ajuste, así como para identificar y medir efectos que no son detectables en ejercicios puros transversales o en series de tiempo puros. No obstante, los modelos tradicionales datos panel no incorporan el espacio en la estimación de las variables dependientes sobre la independiente. Por ello se aplican modelos de datos panel espaciales tipos SAR, SEM y SARAR para determinar la incidencia del espacio sobre la producción agrícola. Se obtuvo como mejor modelo el panel espacial tipo SEM donde el parámetro de autocorrelación espacial fue significativo al 5% y 3 de las 4 variables regresoras también pasan las pruebas individuales. Se concluye que la producción agrícola en el período de estudio se ve afectada positivamente por el área sembrada, el rendimiento medio de los cultivos y la cantidad de población rural de la unidad geográfica. Finalmente, la producción agrícola departamental, en presencia de un choque positivo una determinada ubicación, no solo crecerá más rápido la producción del departamento que recibe este impulso, sino también los otros cercanos a él.

Palabras clave: modelos con datos panel espaciales, producción agrícola, factores de producción, departamentos.

Índice general

1.	Introducción	8
2.	Objetivos	11
2.1.	General.....	11
2.2.	Específicos	11
3.	Justificación.	11
4.	Marco de referencia.	13
4.1.	Producción agrícola.	13
5.	Marco teórico.....	17
5.1.	Modelo de datos panel tradicionales.....	18
5.1.1.	Modelo de efectos fijos	20
5.1.2.	Modelo de efectos aleatorios.....	22
5.2.	Modelos de datos panel espaciales.....	23
6.	Metodología.	28
6.1.	Diseño de investigación.	28
6.2.	Alcance y tipo de investigación.	28
6.3.	Fuentes, herramientas de recolección y procesamiento de datos	29
7.	Análisis y resultados.	30
7.1.	Análisis espacial de los factores.....	34
7.2.	Resultados modelos panel espaciales.....	36
8.	Conclusiones.	42
9.	Referencias bibliográficas.....	43

Índice de gráficos.

Gráfico 1. Producción agrícola nacional entre 2007 y 2015 (en miles).....	8
Gráfico 2. Participación del sector agropecuario sobre del PIB entre 2007 y 2015	9
Gráfico 3. Producción Departamental promedio años 2007 a 2013	30
Gráfico 4. Comportamiento promedio departamental en variables independientes	31

Índice de ilustraciones.

Ilustración 1. Concentración promedio de la producción agrícola en Colombia	10
Ilustración 2. Tipos de contigüidad queen y rook.	24
Ilustración 3. Departamentos y magnitud promedio de producción agrícola	32
Ilustración 4. Producción promedio departamental por aglomeración.	33
Ilustración 5. Conectividad departamental en unidades seleccionadas.....	33
Ilustración 6. Concentración espacial promedio de área cultivada y asignación de crédito.	34
Ilustración 7. Concentración espacial de la población rural y rendimiento agrícola promedio ...	35
Ilustración 8. Residuales años 2007 y 2013 del panel SEM	40
Ilustración 9. Efectos fijos por departamentos panel SEM	41

Índice de tablas.

Tabla 1. Identificación y descripción de variables.....	17
Tabla 2. Especificación y resultado del test de hausman espacial	38
Tabla 3. Resultados paneles espaciales de efectos fijos.....	38

Índice de ecuaciones

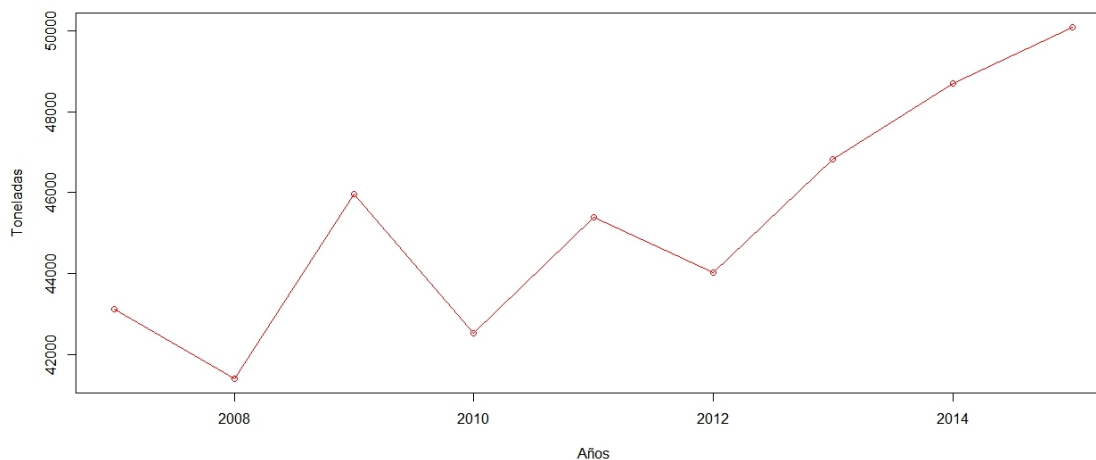
Ecuación 1. Función estándar de producción.....	15
Ecuación 2. Función de producción agrícola.....	16
Ecuación 3. Modelo propuesto producción agrícola.....	17
Ecuación 4. Modelo base datos panel	19
Ecuación 5. Estimador datos panel efectos fijos.....	21
Ecuación 6. Modelo datos panel efectos fijos.....	21
Ecuación 7. Modelo panel de efectos no observados.....	22
Ecuación 8. Modelo con término de error compuesto	23
Ecuación 9. Modelo datos panel efectos aleatorios.	23
Ecuación 10. Modelo estándar para panel espacial.....	24
Ecuación 11. Modelo de panel espacial con rezago y error espacial	25
Ecuación 12. Modelo panel espacial SAR.....	26
Ecuación 13. Modelo panel espacial SEM.....	26
Ecuación 14. Modelo panel espacial estimado SAR ($\rho = 0$).....	37
Ecuación 15. Modelo panel espacial estimado SEM ($\lambda = 0$)	37
Ecuación 16. Modelo panel espacial estimado SARAR (ρ y $\lambda \neq 0$)	37

1. Introducción

La producción agrícola en Colombia desde el año 2012 ha tenido un crecimiento sostenido hasta el año 2015. De acuerdo a cifras consolidadas del Ministerio de Agricultura y Desarrollo Rural a partir de las Evaluaciones Agropecuarias Municipales (EVA), la producción total del país pasó de 4.4 millones de toneladas en 2012 a poco más de 5 millones para el año 2015. Esta tendencia de crecimiento difiere de la producción entre 2007 y 2012. El año 2007 inició con una producción de 4.3 millones de toneladas, bajando para el año 2008 a 4.1 millones de toneladas; en 2009 subió a 4.6 millones para caer a 4.2 en 2010; ya entre 2011 y 2012 la oscilación de la producción fue más pequeña, pues de 4.5 millones de partida, bajó a 4.4 millones (MinAgricultura, 2016).

El comportamiento de la producción en los años de referencia se puede apreciar en el gráfico 1.

Gráfico 1. Producción agrícola nacional entre 2007 y 2015 (en miles)

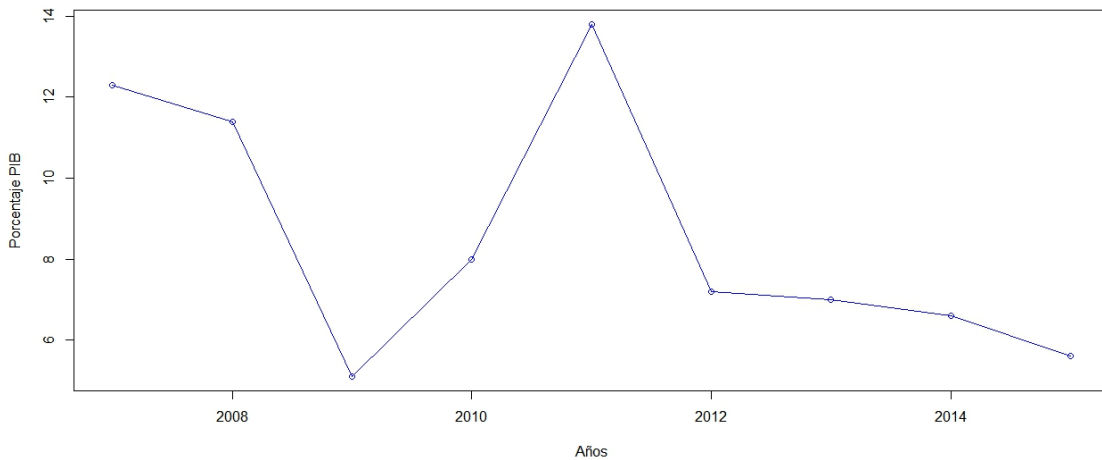


Fuente: elaboración propia.

De acuerdo al Departamento Nacional de Estadística (2017), el aporte de la producción agrícola, junto a la pecuaria, osciló entre los años 2007 y 2015 sobre el 5.1% y el 13.8%

del PIB, siendo el año 2009 el de menor aporte y el 2011 el de mayor aporte. Sin embargo, si bien la producción agrícola ha tenido un crecimiento sostenido entre 2012 y 2015, para el mismo periodo de tiempo, la participación agropecuaria sobre el PIB del país ha sido cada año menor. Desde el valor de 13.8% de aporte al PIB para el año 2011, ha bajado su importancia relativa en la economía nacional con valores que llegan a ser de 5.6% para el año 2015. El gráfico 2 muestra el comportamiento del sector agropecuario como parte del PIB nacional en el periodo de 2007 a 2015.

Gráfico 2. Participación del sector agropecuario sobre del PIB entre 2007 y 2015

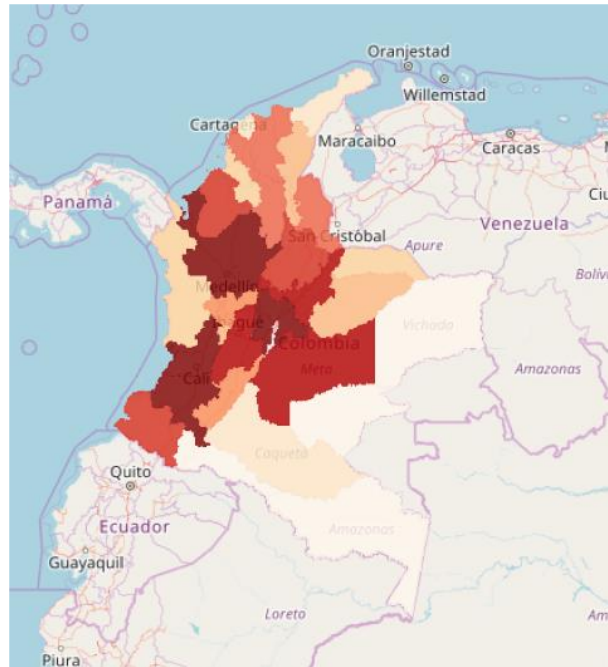


Fuente: elaboración propia.

Históricamente, la producción agrícola y el uso del suelo en actividades pecuarias han estado concentradas en ciertos departamentos del país. La condición geográfica, biofísica y el retorno de estos tipos de producción agrícolas hacen que existan unidades geográficas con mejores aptitudes, generando concentración espacial en ello.

Utilizando los paquetes ggmap y rgdal de R, se trabajó con un shapefile que en su tabla de atributos cuenta con la producción promedio por departamento en Colombia para el período 2007 a 2013. El mapa 1 muestra la concentración espacial de la producción departamental dividida en 5 grupos.

Ilustración 1. Concentración promedio de la producción agrícola en Colombia



Fuente: elaboración propia

Se observa que las mayores producciones promedio (colores más oscuros del mapa) están en la zona pacífica y centro de la geografía del país.

De esta forma, como pregunta de investigación, se tiene la siguiente: ¿de qué manera la producción agrícola de los departamentos país se ve afectada por su ubicación geográfica y dinámicas de producción de sus vecinos contiguos?

2. Objetivos

2.1. General.

Determinar la incidencia de la localización y contigüidad espacial en la producción agrícola de los departamentos en Colombia a través de modelos de datos panel espaciales.

2.2. Específicos

Aplicar un modelo de datos panel espacial sobre la producción agrícola departamental respecto al rendimiento, área sembrada, crédito y población rural como variables independientes.

Estimar, de acuerdo al tipo de modelo panel espacial, el efecto del rezago espacial, así como su autocorrelación espacial.

Analizar el comportamiento de las variables independientes sobre la producción agrícola y los efectos por departamento, de acuerdo al modelo panel espacial elegido.

3. Justificación.

Los resultados obtenidos en el presente trabajo resultan relevantes a nivel estadístico y espacial para explicar el comportamiento de las variables observadas sobre la producción en los departamentos del país. Para el caso particular, a partir de los paneles espaciales se logra una captura importante sobre los efectos en el comportamiento de las variables de un departamento a otro, así como los niveles de producción nos muestran el efecto de contagio o difusión frente a los departamentos contiguos, lo cual nos lleva a resaltar que este tipo de análisis proponen el espacio como elemento determinante sobre los factores de producción.

Así mismo, este trabajo hace aportes significativos para los análisis sobre las estrategias implementadas en el sector agrícola como la colocación de créditos o las acciones de política pública, las cuales al ser analizadas desde esta perspectiva pueden enriquecer los análisis académicos y de política pública.

4. Marco de referencia.

La diversidad de los tipos de suelo, así como la morfología, clima y características propias de la biodiversidad del territorio colombiano son determinantes fundamentales para la vocación y el uso del suelo en relación a los factores productivos, que a su vez son heterogéneos para cada entidad territorial. No obstante, de los principales problemas en la agricultura, es la utilización de tierras en cultivos para los cuales no se tiene vocación o utilización del suelo por encima de las capacidades de producción (Fajardo, 2002).

Como se evidencia en un informe suministrado por la OCDE, el sector agrícola ha enfrentado las consecuencias derivadas de una mala implementación de las políticas públicas, las cuales están reflejadas por la baja productividad, que termina afectando la competitividad del sector (OCDE, 2015). Conocer el comportamiento de la producción en Colombia es relevante para la consolidación de políticas públicas asociadas a la seguridad alimentaria y la lucha contra la pobreza (FAO, 2004).

4.1. Producción agrícola.

Jaramillo (1998) explica que, para el periodo comprendido entre los años 1990 y 1997, se registró un pobre desempeño en el sector, siendo este resultado de tendencias macroeconómicas negativas que afectaron los precios principalmente. Confluyeron en su momento la reforma constitucional de 1991 así como la apertura económica del país, condicionando de esta manera el crecimiento y el fortalecimiento del sector.

De otra parte, Perfeti et al (2003) sostienen que las políticas del sector agropecuario impulsadas por los gobiernos a lo largo de la historia colombiana, han estado ligados profundamente a los vaivenes políticos del país y la región y con ello la transformación territorial con sus dinámicas sociales, políticas y económicas que se desarrollaron como consecuencia de la ocupación de los territorios.

Para el inicio de los años dos mil y con un comportamiento similar ajustado a las políticas macroeconómicas por parte de los gobiernos, el sector agropecuario entre el año 1999 y 2000 evidenció cumplimiento sobresaliente en relación a las tasas de crecimiento con respecto al PIB, en adelante se comportó similar al primer periodo de los años noventa, donde curiosamente se dio también un incremento del 1,5% del sector para el periodo 2002-2006, es decir, desmejoró con relación al incremento en el primer periodo del milenio.

Según Beltrán y Piñeros (2013) las hectáreas cosechadas disminuyeron en un -3,48% y el comportamiento de factores productivos como el bovino, no fueron los determinantes en dicho incremento, por el contrario y coincidentalmente para los inicios de los años noventa, fueron los precios quienes apalancaron el crecimiento poco representativo del sector, es decir, desde el primer periodo de Uribe, el esquema programático del plan de desarrollo para el sector no fue el determinante para dar cuenta de este periodo.

La agricultura juega un papel fundamental en el desarrollo económico de un país. Produce la oferta de alimentos y materias primas, por lo cual cambios en la magnitud de producción puede llevar a condiciones adversas para la demanda de alimentos del país e incluso a generar efectos en la inflación. De otra parte, es un sector que absorbe una importante cantidad de mano de obra, de acuerdo con el tipo de cultivo y su producción dinamiza proceso de transformación industrial, con los efectos que en materia de ingreso y empleo genera (Álvarez, Mamani y Bonilla, 2004).

La magnitud y combinación de elementos para producción en el sector agrícola responde a una función de producción. Por una función de producción se entiende a la relación técnica que transforma insumos o factores en productos. A nivel general, pueden encontrarse diferentes tipos de funciones de producción, como por ejemplo cuadrática, cúbica, raíz cuadrada, Cobb-Douglas, Leontief, CES, transcendental y translogarítmica, entre otras (Álvarez, Mamani y Bonilla, 2004); no obstante, encontrar la función que mejor represente la producción de un producto, se requiere elementos teóricos y

conocimiento empírico que determinen la forma óptima de relación entre factores e insumos.

La producción agrícola departamental recoge una alta cantidad de distintos cultivos que se pueden dar a lo largo del país, de acuerdo con las condiciones biofísicas, de altitud, temperatura, entre otros. Durante los años 2007 y 2015 se reportaron al Ministerio más de 200 cultivos agrícolas distintos (MinAgricultura, 2017). Esta diversidad de tipos de cultivos implica diferentes combinaciones de factores de producción para su logro. Por ejemplo, luego de esfuerzos institucionales, el gobierno nacional ha zonificado algunos departamentos y desagregado la producción de al menos 2 cultivos, especificando las necesidades para la producción, siendo diferentes en todos los casos, por las necesidades de los factores y otros por la escala de los mismos (UPRA, 2013).

Por lo anterior, determinar una función de producción estándar es generalizar a 222 cultivos en una combinación de factores que cambian entre uno y otro, por presencia o magnitud. Es así que, para efectos de este trabajo, se harán los análisis a partir de algunos factores que sobre la producción departamental.

Una estructura de producción estándar puede tener la siguiente forma:

Ecuación 1. Función estándar de producción

$$y = f (X_1, X_2, \dots, X_n)$$

En donde y es el producto y X_k es el factor o insumo, para $k = 1, \dots, n$.

Para el caso de estudio, se contará, de acuerdo con la información disponible, con una producción agrícola que estará basado en la combinación de 3 factores, que, para el caso, constituye la forma básica de producción, tierra (L), trabajo (W) y capital (K).

Ecuación 2. Función de producción agrícola

$$y = f(L, W, K)$$

Para desarrollar estos factores frente a la producción agrícola (y), la cual se relaciona con la producción agrícola total por departamento en toneladas, se toman como variables proxy las siguientes: el área sembrada por departamento (L), la proporción de población rural sobre el total del departamento (W) y el monto de créditos agropecuarios por departamento (K).

5. Marco teórico.

El modelo teórico sobre el cual se trabaja supone que en cada momento (t), para cada departamento, la producción agrícola de los departamentos (prod) depende de 4 variables: área agrícola sembrada (area), población rural (pobr), colocación de crédito agropecuario (cre) y el rendimiento promedio de los cultivos (rend).

Ecuación 3. Modelo propuesto producción agrícola

$$prod_{it} = f (area_{it}, pobr_{it}, cred_{it}, rend_{it})$$

Las condiciones y descripción de las variables a utilizar se encuentran en la tabla 1.

Tabla 1. Identificación y descripción de variables.

Variable	Tipo	Descripción	Fuente
Producción agrícola (prod)	Dependiente	Producción agrícola total por departamento medido en toneladas	MinAgricultura (2016)
Área agrícola sembrada (area)	Independiente	Área agrícola total sembrada por departamento medida en hectáreas	MinAgricultura (2016)
Población rural (pobr)	Independiente	Total de la población rural por departamento (resto DANE)	DANE (2017)
Crédito Agropecuario (cre)	Independiente	Monto total de créditos agropecuarios por departamento	FINAGRO (2016)
Rendimiento (rend)	Independiente	Relación entre el área cosechada y la producción por departamento	MinAgricultura (2016)

Fuente: elaboración propia

La información se encuentra disponible para cada unidad geográfica (departamentos) año a año entre 2007 y 2013. Por lo anterior, se considera el uso de datos panel espaciales apropiado para encontrar las relaciones entre la producción agrícola departamental, sus vecinos y las variables dependientes.

5.1. Modelo de datos panel tradicionales.

El uso de los datos panel, como herramienta, combina elementos de medición transversal y de series de tiempo; de esta forma, los datos panel utilizan información de un conjunto de individuos medidos repetidamente a través del tiempo. De otra parte, se asume que los individuos son mutuamente independientes entre sí, pero las observaciones de un individuo son mutuamente dependientes (Hansen, 2017).

Como es usual en cada herramienta para estimación estadística, el uso de los datos panel trae consigo algunas limitaciones relacionadas con la dificultad de obtener información de un mismo individuo a través del tiempo, distorsiones en la estimación de los errores, problemas de selectividad y la dependencia a elementos de corte transversal (Baltagi, 2005).

No obstante lo anterior, los datos panel tienen muchos beneficios, como, por ejemplo, se controla la heterogeneidad individual, su aplicación brinda más información, más variabilidad, menor colinealidad entre las variables, más grados de libertad y mayor eficiencia. También son mejores para estudiar dinámicas de ajuste, así como para identificar y medir efectos que no son detectables en ejercicios puros transversales o en series de tiempo puros. Finalmente, los sesgos resultantes de agregación de individuos o firmas pueden ser reducidas o eliminadas (Baltagi, 2005) (Hsiao, 2003).

Baltagi (2005) remarca la importancia que tiene la venta del uso de los modelos de datos panel respecto a la capacidad de capturar la heterogeneidad no observable, pues explica que, ya sea entre agentes económicos o de estudio, así como también en el tiempo, la heterogeneidad no se puede detectar ni con estudios de series temporales ni tampoco con los de corte transversal.

La información organizada en los datos panel puede ser balanceada o no balanceada. Los datos panel son balanceados cuando se cuenta con la totalidad de los datos respecto a las

variables de estudio para cada individuo. De forma diferenciada, el dato panel no balanceado se da cuando se presenta carencia de datos, es decir no se encuentran completas toda la serie de datos para un individuo, o faltan individuos para ciertos años de una misma variable.

De acuerdo a Hansen (2017), si una observación es el par $\{y_{it}, x_{it}\}$, donde i denota al individuo y t el tiempo un panel balanceado sería :

$$\{y_{it}, x_{it}\} : t = 1, \dots, T; \quad i = 1, \dots, n$$

De otra parte, un panel no balanceado sería:

$$\{y_{it}, x_{it}\} : \text{Para } i = 1, \dots, n, \quad t = \underline{t}_i, \dots, \bar{t}_i$$

Ya definida la importancia del uso de los datos panel y los tipos de panel que pueden existir, Montero (2011) explica que el modelo básico para al que responde el dato panel es el siguiente:

Ecuación 4. Modelo base datos panel

$$y_{it} = X_{it}\beta + \varepsilon_{it}$$

En donde β es el vector de parámetros, X_{it} es un vector de k variables explicativas, i es la cantidad de individuos, t responde a la cantidad de periodos, ε_{it} son los errores aleatorios.

Agrega Montero (2011) que “si no se disponen de todas las variables de influencia entonces $cov(x_{it}, \varepsilon_{it}) \neq 0$, es decir, los residuos no son independientes de las

observaciones por lo que MCO¹ estará sesgado” (p. 1). Para su solución, se pueden aplicar modelos de efectos fijos y de efectos aleatorios.

Según Lasso (2015) para el caso de los datos panel el supuesto fundamental es que el término de error del modelo básico ε_{it} se puede descomponer como la suma de $\varepsilon_{it} = \alpha_i + e_{it}$, donde α_i es la heterogeneidad inobservada permanente en el tiempo y e_{it} se le denomina término de error idiosincrático que se asume independiente e idénticamente distribuido entre individuos. Añade Lasso (2015) que:

Los modelos estáticos de Datos de Panel asumen dos tipos de supuestos sobre el comportamiento del error: un caso en que la heterogeneidad no observada es independiente del error idiosincrático y otro caso en el cual ambos errores se correlacionan. El método de estimación más eficiente para el primero de los casos, es el de los Efectos Aleatorios (RE –*Random Effect*- por sus siglas en inglés) y para el segundo el método de los efectos fijos (FE –*Fixed Effect*- por sus siglas en inglés). (p.27)

5.1.1. Modelo de efectos fijos

El modelo de efectos fijos se emplea cuando el comportamiento del error tiene tanto a α_i como e_{it} correlacionados y por lo tanto no será posible la estimación consistente de una matriz de varianzas y covarianzas.

Para su manejo, Lasso (2015) indica que de acuerdo a (Baltagi, 2005) y (Wooldridge, 2010) es necesario aplicar una transformación a los datos, eliminando de esta manera la heterogeneidad no observada, permitiendo que el estimador de efectos fijos tome la forma de un estimador de MCO de la siguiente forma:

¹ Acrónimo de Mínimos Cuadrados Ordinarios; se incluye porque la estimación a través de MCO a la información del dato panel viola uno de los supuestos del Modelo Clásico de Regresión Lineal al estar sesgado (Gujarati y Porter, 2010).

Ecuación 5. Estimador datos panel efectos fijos

$$\hat{\beta}_{FE} = (\tilde{X}'\tilde{X})^{-1} \tilde{X}'\tilde{Y}$$

Siendo \tilde{X} e \tilde{Y} matrices con las transformaciones realizadas.

Para aplicar la transformación hay 3 formas de llevarse a cabo, a saber: *between*, *within* y *difference*.

La transformación *within* transforma las variables tomando las diferencias respecto de las medias de cada uno de los individuos. La transformación *between* lo hace por el promedio, mientras que la Transformación de Diferencias, como indica su nombre diferencia las variables con el objetivo de eliminar la heteroscedasticidad. (Lasso, 2015, p. 28)

Para Labra y Torrecillas (2014) el estimador *within* es el indicado para datos panel de efectos fijos, puesto que “asume que el efecto individual está correlacionado con las variables explicativas. Este supuesto relaja la condición impuesta por el estimador de efectos aleatorios, tratando el efecto individual separadamente del término de error”. (p. 16)

De esta manera la correlación entre α_i y X es diferente de cero. El modelo queda entonces de la siguiente manera:

Ecuación 6. Modelo datos panel efectos fijos

$$y_{it} = \alpha_i + \beta X_{it} + \mu_{it}$$

Explica Labra y Torrecillas (2014) que este estimador “tiene la ventaja de que permite conocer los α_i separadamente, lo que contribuye a entender de mejor forma el modelo.

Además, evita una sobrestimación del parámetro β , lo que ocurre cuando se aplica el estimador de efectos aleatorios”. (p. 16)

Finalmente, otra manera de realizar el análisis es a través del uso de variables dicótomas para cada efecto individual que se quiera conocer y controlar. De esta manera “el estimador de MCO entregará separadamente los β de estos efectos”. (Labra y Torrecillas, 2014, p. 17)

5.1.2. Modelo de efectos aleatorios

Wooldridge (2010) para explicar el modelo de datos panel de efectos aleatorios, parte del modelo de efectos no observados. Tal modelo tiene la siguiente configuración:

Ecuación 7. Modelo panel de efectos no observados

$$y_{it} = \beta_0 + \beta_1 x_{it1} + \dots + \beta_k x_{itk} + \alpha_i + \mu_{it}$$

Al usar efectos fijos, el objetivo es eliminar α_i porque se cree que está correlacionado con uno o más de los x_{itj} . Pero supongamos que pensamos que α_i no está correlacionado con cada variable explicativa en todos los periodos de tiempo. Por lo anterior, usar una transformación para eliminar α_i resulta en estimadores ineficientes.

Cuando se asume que el efecto no observado α_i no está correlacionado con la variable explicativa, es decir, que $Cov(x_{itj}, \alpha_i) = 0$, la ecuación del modelo de efectos no observados se convierte a una de efectos aleatorios; “De hecho, los supuestos ideales del efecto aleatorio incluye todos los supuestos del efecto fijo más el requerimiento adicional que α_i es independiente de todas las variables explicativas para todos los periodos de tiempo”. (Wooldridge, 2010, p. 492)

Si se define la composición del término de error como $v_{it} = \alpha_i + \mu_{it}$, entonces se puede escribir la ecuación como:

Ecuación 8. Modelo con término de error compuesto

$$y_{it} = \beta_0 + \beta_1 x_{it1} + \dots + \beta_k x_{itk} + v_{it}$$

Conociendo la composición del término de error, siguiendo a Labra y Torrecillas (2014) la ecuación para efectos aleatorios se puede escribir como:

Ecuación 9. Modelo datos panel efectos aleatorios.

$$y_{it} = \beta x_{it} + (\alpha_i + \mu_{it})$$

Para su tratamiento, se emplea el método generalizado de momentos.

5.2. Modelos de datos panel espaciales

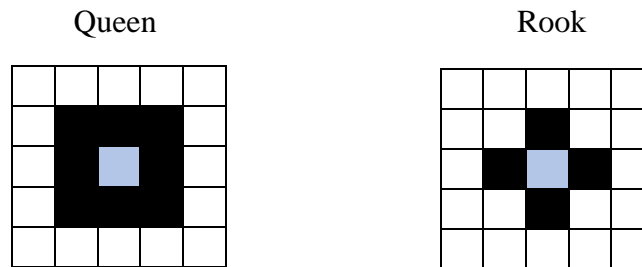
El uso de los modelos espaciales parte de la necesidad de incorporar el espacio como uno de los determinantes de las muchas interacciones que suceden a nivel territorial. En este sentido, el uso de matrices de contigüidad entre unidades geográficas objeto de análisis, permite el relacionamiento del espacio en el fenómeno que se desea explicar.

5.2.1. Matriz de contigüidad.

Las matrices de contigüidad (W) sirven como elemento de captura del espacio y los efectos entre las unidades de análisis geográficas. Estas matrices son binarias, compuestas de 1 y 0 que marcan la contigüidad entre unidades geográficas por contar límites en común. El elemento w_{ij} de una matriz de contigüidad será 1 si se cumple la condición que la unidad geográfica i es vecina de la unidad j, de no serlo, será 0; Lo anterior identifica una simetría en la matriz W. Dado que ninguna unidad geográfica puede ser vecina de si

mismo, 0 será el valor de la diagonal de la matriz W (Kuscevic y Del Río, 2013) y (Getis y Aldstadt, 2010). Esta forma de relación de contigüidad entre las unidades geográficas corresponde al método *queen*².

Ilustración 2. Tipos de contigüidad queen y rook.



Fuente: elaboración propia

La ilustración 2 muestra la contigüidad *queen* frente a otro método para la construcción de la matriz W, *rook*.

5.2.2. Modelos SAR, SEM y SARAR en panel espaciales

Los modelos panel espaciales, a diferencia de los tradicionales, incorporan el espacio en su análisis. Elhorst (2010) explica que el modelo estándar para panel espacial es el siguiente:

Ecuación 10. Modelo estándar para panel espacial

$$y_{it} = X_{it}\beta + \mu_i + \varepsilon_{it}$$

Donde i representa las unidades espaciales, t la dimensión del tiempo, y la variable dependiente, X es un vector de observaciones de variables independientes, β corresponde a los parámetros de las variables independientes, ε_{it} es el término de error distribuido independiente e idénticamente distribuido para i y t con media cero y varianza σ^2 y μ_i

² Se pueden encontrar otras formas de relación como *Rook*, *Inverse Distance* así como *Inverse Distance Squared* (Getis y Aldstadt, 2010).

denota el efecto espacial específico. Afirma Elhorst (2010) que “el razonamiento estandar detrás de los efectos espaciales específicos es que ellos controlan todas las variables espacio-tiempo específicos y tiempo-invariante que se pudieron omitir.” (p. 378)

Cuando se especifica la interacción entre unidades espaciales, el modelo puede contener un variables rezagada de dependencia espacial o un proceso espacial autorregresivo en el término de error, conocidos como el rezago espacial y el modelo de error espacial, respectivamente. (Elhorst, 2010, p. 378)

De esta forma, se cuenta con formas generales de incluir el espacio, a través de la matriz de contiguidad W en los datos panel tradicionales.

Millo y Piras (2012) explican que los modelos de datos panel espaciales capturan las interacciones espaciales entre unidades espaciales y el tiempo, que pueden tener aproximaciones dinámicas y estáticas, de acuerdo al tipo de panel a aplicar. Basandose en elementos estáticos, en general un panel espacial que incorpore tanto el rezago espacial como el error espacial tiene la siguiente especificación:

Ecuación 11. Modelo de panel espacial con rezago y error espacial

$$y = \lambda(I_T \otimes W_N)y + X\beta + u$$

$$u = (\iota_T \otimes I_N)\mu + \varepsilon$$

$$\varepsilon = \rho(I_T \otimes W_N)\varepsilon + v$$

Donde y es la variable dependiente, X un vector de variables independientes, I_T es una matriz identidad de dimensión T , W_N es la matriz de pesos espaciales de tamaño $N \times N$, λ es el parámetro espacial autorregresivo (rezago) y ρ es el parámetro de autocorrelación espacial.

Sobre la ecuación base, se puede imponer algunas restricciones sobre los parámetros ρ y λ . De esta forma, se obtienen variaciones de la ecuación base que dan a lugar a modelos SAR, SEM y SARAR.

Los modelos SAR o *Spatial Autoregressive Models* son aquellos en donde el parámetro ρ es igual a cero. De esta forma la ecuación queda de la siguiente forma:

Ecuación 12. Modelo panel espacial SAR

$$y = \lambda(I_T \otimes W_N)y + X\beta + \varepsilon$$

Gómez et al (2014) define que la característica del SAR es que “la relación espacial se da directamente por la presencia de la variable endógena rezagada, mientras que no se considera la relación directa entre las características propias de las vecinas con la variable endógena.” (p. 8)

En los modelos SEM o *Spatial Error Models*, λ es igual a cero. Por lo tanto, solo queda incorporado el parámetro ρ asociado al error. La ecuación queda así:

Ecuación 13. Modelo panel espacial SEM

$$\begin{aligned} y &= X\beta + u \\ u &= (I_T \otimes I_N)\mu + \varepsilon \\ \varepsilon &= \rho(I_T \otimes W_N)\varepsilon + v \end{aligned}$$

“En este modelo la relación espacial se da por medio del denominado contagio, donde los choques exógenos se propagan entre las unidades vecinas.” (Gomez et al, 2014, p. 8)

Finalmente, los modelos SARAR o *Spatial AutoRegressive with additional AutoRegressive error structure* asumen que tanto λ como ρ existen. De esta manera, la ecuación de un SARAR corresponde a la misma ecuación base, incorporando tanto el

parámetro de rezago espacial como el de autocorrelación espacial. En este caso, de acuerdo con Gomez et al (2014)

La relación espacial se presenta por dos mecanismos, el primero se da por medio de la relación directa entre la variable endógena y su rezago espacial. El segundo por choques aleatorios en las unidades vecinas, los cuales tienen un efecto contagio sobre la unidad i . (p. 8)

6. Metodología.

El sustento metodológico del trabajo está enfocado en la selección del diseño de investigación, alcance, tipo y fuentes de datos a utilizar de acuerdo a los objetivos propuestos.

6.1. Diseño de investigación.

Siguiendo el esquema de investigación propuesto por Baptista, Fernández-Collado y Hernández (2006) corresponde a un trabajo de enfoque cuantitativo, el cual mide el fenómeno a través de la recolección de datos para probar una hipótesis; a partir de la medición numérica y el análisis estadístico, se busca establecer patrones de comportamiento que puedan comprobar el sustento teórico.

6.2. Alcance y tipo de investigación.

Se trata de una investigación de tipo explicativa, la cual se enfoca en explicar por qué ocurre un fenómeno y en qué condiciones se manifiesta, o por qué se relacionan dos o más variables. Para el caso, se mide el grado de asociación entre la producción agrícola departamental, la dependencia espacial de los departamentos vecinos, junto con la disponibilidad de mano de obra, así como la tierra destinada a la producción, el rendimiento de los cultivos y el valor de los créditos agropecuarios girados por cada unidad geográfica. De esta forma se busca mostrar el grado de incidencia que estas variables tienen sobre la producción agrícola.

El tipo de investigación es no experimental, es decir, en donde no se manipulan las variables de manera intencional; se toman los datos de cada variable sin variarlas para estimar su efecto en otras. De acuerdo con la información existente y con el objetivo del trabajo de investigación, se usará la categoría de investigación longitudinal panel, pues se estima la incidencia de las variables independientes (área agrícola sembrada, población

rural, monto de créditos y rendimiento) en la variable dependiente (producción en toneladas) de los departamentos del país entre los años de 2007 y 2013, desde un enfoque espacial.

6.3. Fuentes, herramientas de recolección y procesamiento de datos

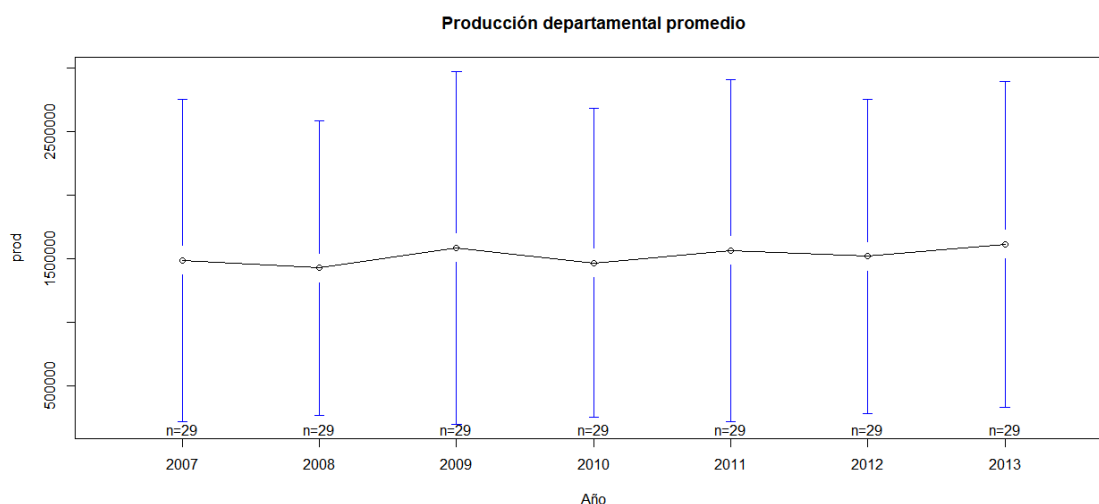
Como fuentes de información primaria se tuvo al Ministerio de Agricultura y Desarrollo Rural, en lo relativo a producción en toneladas, área agrícola sembrada y rendimiento de los cultivos; información relacionada con créditos, proviene del Fondo para el financiamiento del sector agropecuario (FINAGRO). Finalmente, la información relacionada con la población rural está disponible por el Departamento Nacional de Estadística (DANE).

Esta información se analizó y procesó utilizando para ello programas como ArcGIS 10.4.1, GeoDa 1.10.0.8 y R *Foundation for Statistical Computing*. La información secundaria se construyó a partir de material bibliográfico. En ello se revisaron revistas, diarios, artículos y documentos disponibles en bases de datos como SCIELO, SIARE, JSTOR y EBSCO.

7. Análisis y resultados.

Para el período entre los años 2007 y 2013, el comportamiento de la producción con respecto a la media es similar para todos los años, sin embargo, hay algunas variaciones sobre el total de la producción para departamentos como el Vichada y el Amazonas con muy baja producción y caso contrario en el departamento del Valle con una producción bastante más elevada para todos los años con respecto a los demás departamentos del país. El gráfico 3 muestra el comportamiento promedio.

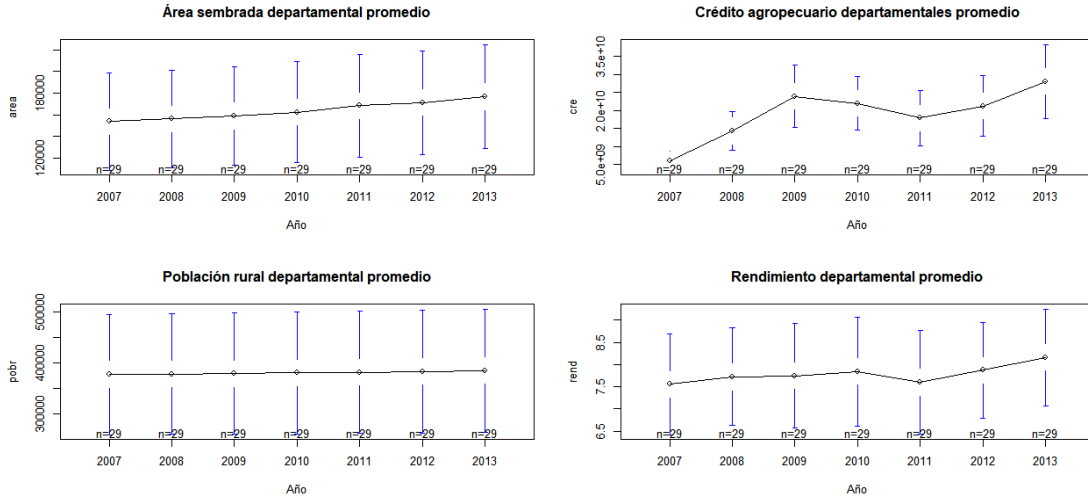
Gráfico 3. Producción Departamental promedio años 2007 a 2013



Fuente: elaboración Propia

Respecto a las variables independientes, se puede observar (Gráfico 4) que el promedio de área sembrada para los años (2007-2013) se comporta de manera similar hasta el año 2010 y a partir de este año empieza a incrementar de manera constante hasta el periodo 2013. No obstante, como se observa en la relación del crédito para el total de los departamentos en cada uno de los años analizados, se evidencia un incremento importante para los años 2007 a 2009 siendo este último su pico y una caída importante hasta el 2011, donde posteriormente se recupera y nuevamente e incrementa hacia el año 2013

Gráfico 4. Comportamiento promedio departamental en variables independientes

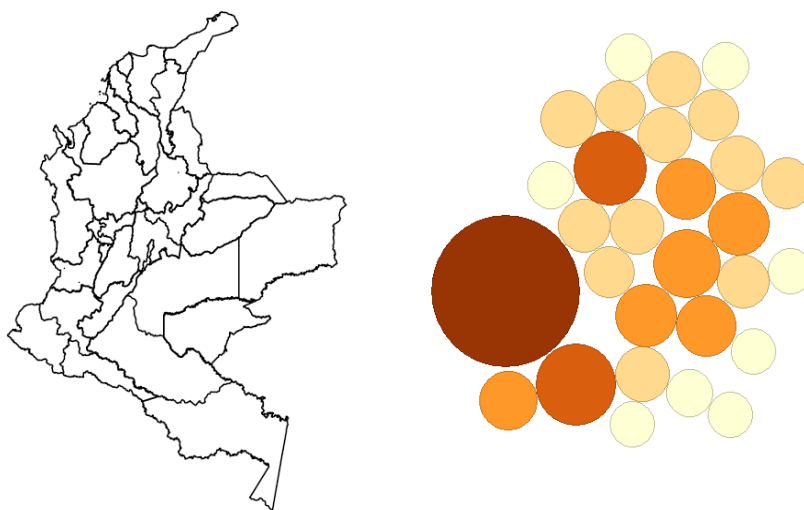


Fuente: elaboración Propia

La población por su parte se comporta similar de manera similar, pues sus cambios son leves; entre tanto, el rendimiento departamental promedio en los departamentos presenta una distribución constante en el promedio y disminuye de manera similar al crédito en el periodo de 2009 a 2011, posiblemente como consecuencia del uso de recursos de crédito con rendimiento tardío como lo afirma el Banco de la República.

La producción agrícola en Colombia cuenta con una concentración espacial en la zona pacífica y centro del país. La ilustración 3 contiene 2 mapas que cuentan con salidas de GeoDa y R, muestra la distribución de los departamentos objeto del análisis y la cartografía en círculos por magnitud de la producción promedio entre los años 2007 y 2013.

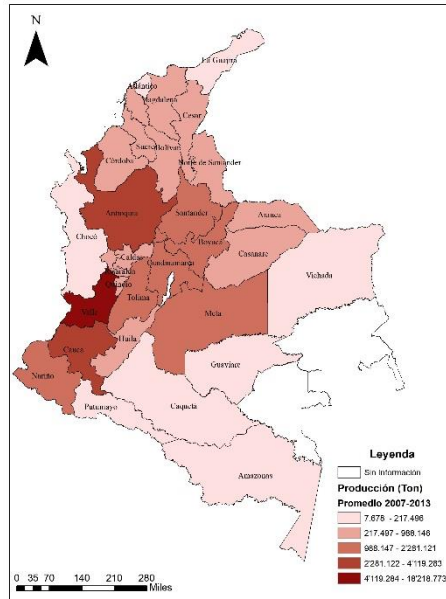
Ilustración 3. Departamentos y magnitud promedio de producción agrícola



Fuente: elaboración propia

Al realizar aglomeraciones de la producción por cortes naturales en GeoDa y ArcGIS, se observa que el Departamento del Valle del Cauca tiene una producción atípica para el país. Es el único departamento que cuenta con una escala de 18 millones de toneladas producidas. En un segundo grupo de producción se cuenta con los departamentos de Antioquia y Cauca, cuya producción promedio oscila entre las 2.2 y 4.1 millones de toneladas en el periodo de estudio. La región central con los departamentos de Tolima, Cundinamarca, Santander, Boyacá y Meta produjeron en promedio entre 965.000 y 2.2 millones de toneladas. Los demás departamentos se encuentran con producciones menores a 965.000 toneladas. La ilustración 4 contiene un mapa realizado en ArcGIS y muestra la división de las agrupaciones para la producción de manera espacial.

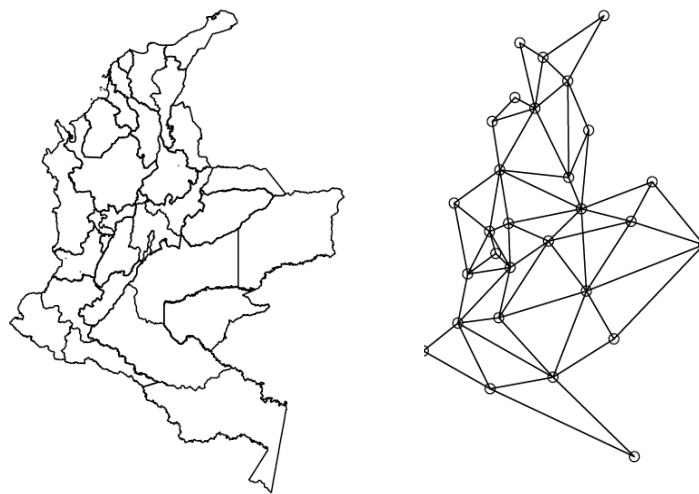
Ilustración 4. Producción promedio departamental por aglomeración.



Fuente: elaboración propia

Para generar los análisis espaciales, se requiere contar con una matriz de contigüidad. De acuerdo con las clases de matriz W analizadas, se optó por generar una de tipo *queen*.

Ilustración 5. Conectividad departamental en unidades seleccionadas



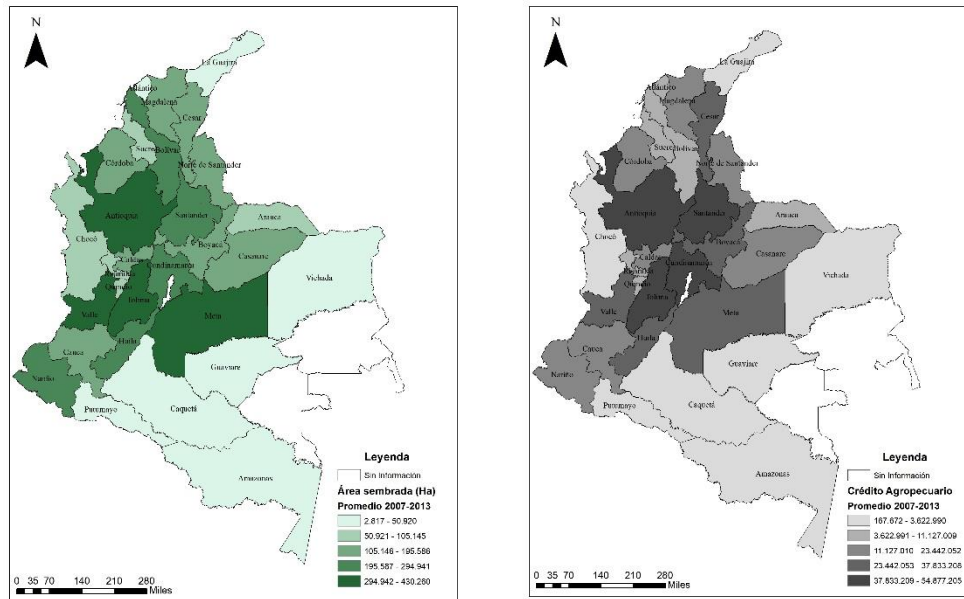
Fuente: elaboración propia.

La ilustración 5 muestra dos mapas; el primer (izquierda) contiene la distribución de los departamentos objeto de análisis. el segundo (derecha) grafica los centroides de los polígonos, para el caso los departamentos, y traza una línea entre aquellos departamentos que cuentan con conexión o contigüidad de tipo *queen*. Ambas figuras fueron realizadas en R con el paquete *rgdal*.

7.1. Análisis espacial de los factores.

Los análisis de las variables independientes, para el caso factores relacionados con la producción agrícola, se realiza utilizando la información geográfica y los promedios del periodo de estudio.

Ilustración 6. Concentración espacial promedio de área cultivada y asignación de crédito.



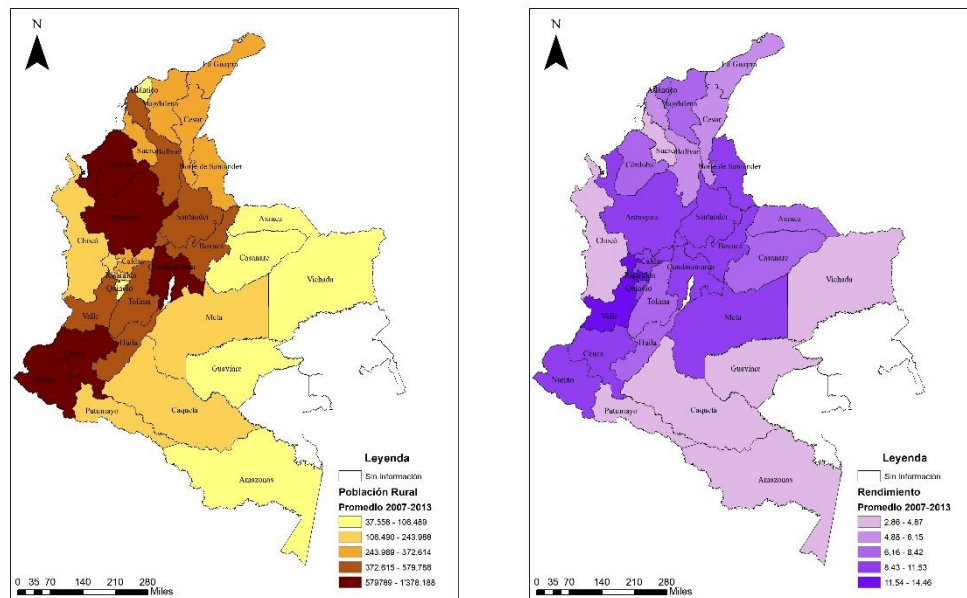
Fuente: Elaboración propia

La ilustración 6 muestra una comparación entre la asignación del crédito agropecuario para cada departamento con respecto al área sembrada. En el primer mapa (izquierda), se

observa que los departamentos del Valle, Tolima, Meta y Antioquia son los que presentan una mayor proporción de área sembrada, en promedio 294.000 y 420.000 Ha; seguido de otro grupo compuesto por los departamentos de Nariño, Huila, Cundinamarca, Santander y Bolívar con un promedio de 195.000 a 294.000 Has. de área sembrada, aquellos que representan una menor cantidad de área sembrada son los departamentos del sur del País, entre los cuales se encuentran, Amazonas, Caquetá, Guaviare, Putumayo y Vichada con áreas que oscilan entre las 2.800 y 50.900 Has.

Lo anterior con relación al mapa de crédito Agropecuario, (parte derecha) nos permite evidenciar una estrecha relación entre la colocación del crédito para aquellos departamentos que presentan una mayor proporción de área sembrada, es decir, Tolima, Antioquia, Santander, Cundinamarca, son los que presentan mayor cantidad de créditos asignados y que a su vez se presentan mayor proporción de área sembrada.

Ilustración 7. Concentración espacial de la población rural y rendimiento agrícola promedio



Fuente: elaboración propia en ArcGIS.

Con respecto a la población rural, observamos que en la ilustración 7, la concentración de la población rural se encuentra en los departamentos de Antioquia, Córdoba, Cauca, Nariño y Cundinamarca (entre 570.000 y 1'360.000 habitantes). Seguido de otros departamentos de la zona centro oriental del país como Valle, Huila, Tolima, Boyacá y Córdoba con población entre los 370.000 y 565.000 habitantes. Por otra parte, los departamentos con menor población rural entre otros, son Amazonas, Guaviare, Vichada, Arauca y muy cerca a esta cifra; Caldas, Chocó, Norte de Santander, Cesar, con poblaciones que oscilan entre los 150.000 y 350.000 habitantes de población rural.

Al contrastar lo anterior con el rendimiento promedio de la producción para los departamentos, se puede observar un comportamiento relativamente proporcional de relación entre la cantidad de población rural con respecto a rendimiento. En el mapa derecho se observa una especie de cordón en los que se encuentran los departamentos de la zona centro y hacia el oriente del territorio colombiano, es decir, departamentos como Nariño, Cauca, Valle, Cundinamarca, Antioquia, Santander, muestran un rendimiento elevado y proporcionalmente una población rural alta con respecto a los demás territorios. En este caso, además el Valle es el único departamento cuyo rendimiento está por encima de los demás.

7.2. Resultados modelos panel espaciales

Dado que el trabajo plantea utilizar información espacial a través de la matriz de contigüidad W y estimar su efecto, se realiza la estimación de los paneles espaciales SAR, SEM, SARAR y elegir entre ellos el más adecuado. A continuación se especifican las formas de relación de las variables dependiente e independientes frente a los tipos de panel espaciales a estimar y los parámetros espaciales que surgen de cada modelo.

Ecuación 14. Modelo panel espacial estimado SAR ($\rho = 0$)

$$\ln prod_{it} = \lambda W_i \ln prod_{it} + \beta_1 \ln area_{it} + \beta_2 \ln pobr_{it} + \beta_3 \ln cred_{it} + \beta_4 rend_{it} + \varepsilon_{it}$$

Ecuación 15. Modelo panel espacial estimado SEM ($\lambda = 0$)

$$\ln prod_{it} = \beta_1 \ln area_{it} + \beta_2 \ln pobr_{it} + \beta_3 \ln cred_{it} + \beta_4 rend_{it} + u_{it}$$
$$u_{it} = \rho W_i u_{it} + \varepsilon_{it}$$

Ecuación 16. Modelo panel espacial estimado SARAR (ρ y $\lambda \neq 0$)

$$\ln prod_{it} = \lambda W_i \ln prod_{it} + \beta_1 \ln area_{it} + \beta_2 \ln pobr_{it} + \beta_3 \ln cred_{it} + \beta_4 rend_{it} + u_{it}$$
$$u_{it} = \rho W_i u_{it} + \varepsilon_{it}$$

Para determinar la existencia o no de autocorrelación espacial, recurrimos al test de Baltagi et al. (2003) usando el multiplicador de Lagrange. La hipótesis alternativa del test es la existencia de autocorrelación espacial. Se obtuvo un LM2³ de 4.2969 y un valor p de 1.732e-05; de esta manera se rechaza la hipótesis nula y se acepta la alternativa, determinando la existencia del parámetro de autocorrelación espacial.

Se aplicó también una prueba para determinar la existencia o no de dependencia transversal en el modelo de panel espacial. El resultado de la prueba CD Pesaran sobre dependencia transversal local en datos panel fue de $z = -1.1597$ con un valor p de 0.2462, aceptando la hipótesis nula de ausencia de dependencia transversal.

Para identificar para cada tipo de panel qué tipo de modelo de agrupamiento aplicar, se estima el test de hausman espacial. Los resultados están en la tabla 2.

³ La notación del LM está relacionada con los diferentes test que permite aplicar el comando `bsktest` del paquete `splm` en R

Tabla 2. Especificación y resultado del test de hausman espacial

Modelo Espacial	Método	χ^2	Df	p-valor
ERROR	Máxima Verosimilitud	356.65	4	<2.2e-16
SARAR	Máxima Verosimilitud	561.46	4	<2.2e-16

Fuente: elaboración propia

El resultado del test de hausman espacial rechaza la hipótesis nula, por lo cual sugiere la aplicación de un modelo SARAR de efectos fijos, el cual, como se comentó anteriormente, incorpora la estimación de los parámetros λ y ρ . La misma estimación por efectos fijos plantea el test al plantear un modelo tipo SEM.

En la tabla 3 se presentan los resultados de la estimación de los paneles espaciales por efectos fijos. El parámetro autorregresivo espacial, tanto en la estimación SAR como SARAR, no es significativo al 5%. Por el contrario, el parámetro error espacial si es significativo en los resultados del SEM y SARAR. Dado que el SARAR tiene su fundamento en la existencia de ambos parámetros espaciales, se puede concluir que el modelo SEM es el que recoge de mejor manera el comportamiento espacial de los datos.

Tabla 3. Resultados paneles espaciales de efectos fijos

	SAR	SEM	SARAR
Ln (pobr)	0.247931 (-0.142189)	0.350765*** (0.053922)	0.351652*** (0.053947)
Rend	0.135673*** (0.035113)	0.149284*** (0.012978)	0.149548*** (0.013000)
Ln (area)	0.911211*** (0.150019)	0.876581*** (0.052442)	0.879193*** (0.053414)
Ln (cre)	-0.011664 (0.077139)	-0.032142 (0.025685)	-0.032689 (0.02574)
Lamda (λ)	0.022156 (0.066979)		-0.0099285 (0.0362627)
Rho (ρ)		0.438762*** (0.076712)	0.442649*** (0.081006)

* p<0.05, ** p<0.01, *** p<0.001

Fuente: elaboración propia

3 de los 4 coeficientes de las variables regresoras en el panel SEM son significativas al 5%. La variable crédito no fue significativa, lo cual puede presentarse porque las inversiones del capital colocado en los departamentos no inciden de manera contemporánea en la producción.

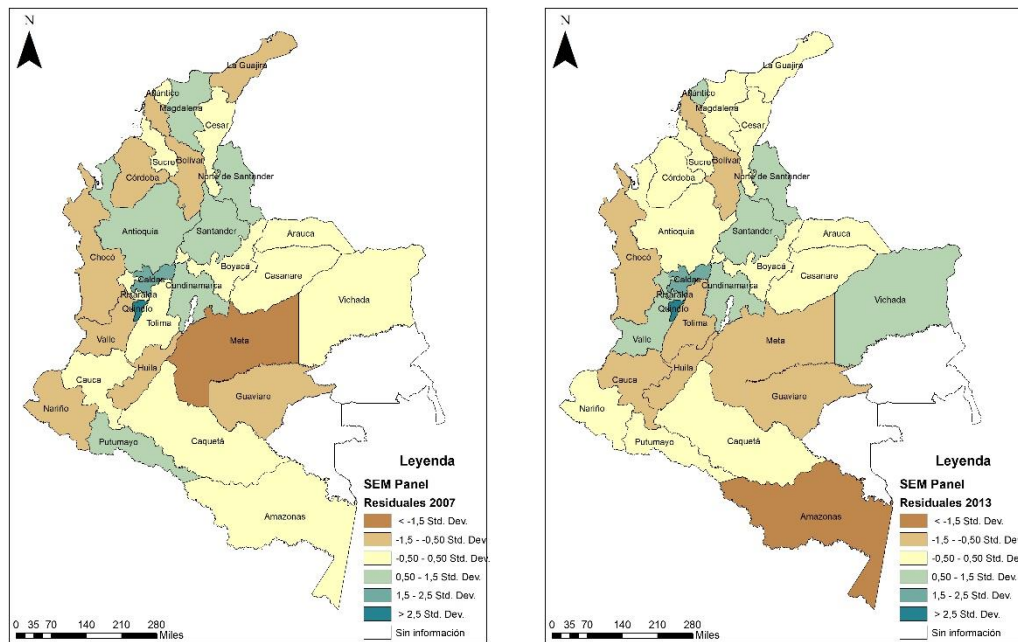
La magnitud de la población rural en los departamentos tuvo un coeficiente de 0.247931, siendo significativa al 5%. Denota la importancia de la población rural sobre el proceso de producción, y cómo podría verse afectada su continua la migración rural – urbano en el país.

El área sembrada y el rendimiento tuvieron coeficientes de 0.135673 y 0.911211 respectivamente. Ambos coeficientes son significativos al 5%. Si bien usualmente se busca un rendimiento cada vez más alto, tiene una mayor importancia el factor tierra en el proceso productivo. El R2 estimado para el modelo SEM es de 0.93.

Para determinar la presencia de heteroscedasticidad en el modelo SEM se aplicó una prueba Breush – Pagan, teniendo como resultado un p-valor de 1, no rechazando la hipótesis nula, implicando homecedasticidad.

Para el análisis de los residuales del modelo panel SEM, se tomaron en cuenta los resultados de los valores para los años 2007 y 2013 para su comparación. La ilustración 8 contiene los mapas para cada año.

Ilustración 8. Residuales años 2007 y 2013 del panel SEM



Fuente: elaboración propia

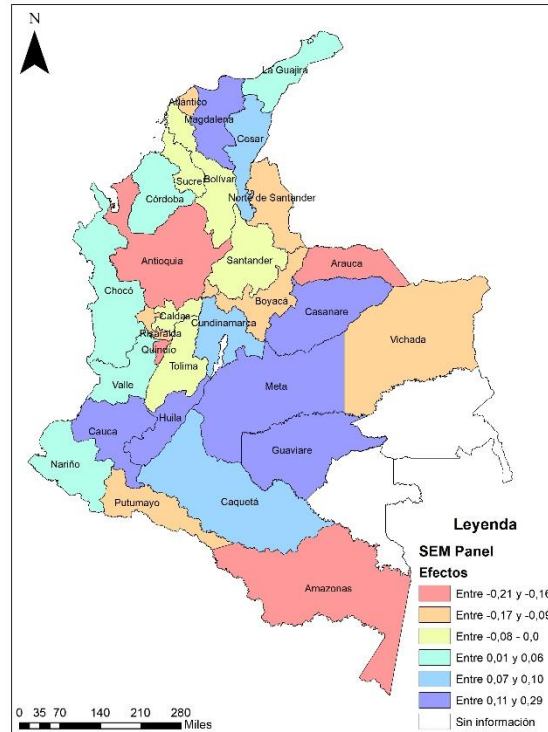
Los colores de los mapas reflejan en qué ubicación se encuentran respecto a la desviación típica. De esta manera, colores como café oscuro y claro se ubican entre -1.5 y -0.5 desviaciones. Los colores azules se ubican, de acuerdo a la intensidad del color, entre 0.50 y más de 2.5 desviaciones.

De acuerdo con los resultados, en el año 2007 (mapa izquierdo) las subestimaciones del modelo se dan en departamentos como Meta, Guaviare, Tolima y Cauca. Las sobreestimaciones se dieron en los departamentos de Quindío, Valle, Vichada y Santander, entre otros.

Para el año 2013 (mapa derecho), las subestimaciones se dan en Meta, Valle, Huila, Nariño, entre otros. Las sobreestimaciones tuvieron lugar en los departamentos de Quindío, Cundinamarca, Antioquia, entre otros.

De otra parte, al extraer los efectos individuales del panel SEM, se obtienen los efectos departamentales sobre la producción. Con base a los resultados se espacializan los efectos estimados para cada una de las unidades geográficas, obteniendo así la ilustración 9.

Ilustración 9. Efectos fijos por departamentos panel SEM



Fuente: elaboración propia

Los departamentos con colores rosado, naranja suave y amarillo son aquellos cuyo efecto es negativo, siendo el rosado el más fuerte y el amarillo el de menor. En este rango de efecto negativo se ubican departamentos como Tolima, Santander y Antioquia. En azul claro al oscuro hay 3 niveles de departamentos cuyo efecto es positivo. Meta, Casanare, Huila, Cauca y Guaviare está en el nivel de mayor efecto; otros departamentos en el rango positivo de efecto son Nariño y Valle del Cauca.

8. Conclusiones y recomendaciones.

La aplicación de modelos de datos panel espacial buscan estimar, además de la incidencia de las variables independientes sobre la dependiente, el efecto de la localización y contigüidad entre las unidades espaciales.

Aplicados 3 tipos de datos panel, a nivel espacial solo el parámetro de autocorrelación espacial fue significativo, determinando que el mejor modelo que se ajusta a los datos es un tipo SEM. Este parámetro (ρ) es positivo, por lo cual se puede afirmar que existen externalidades positivas asociadas a la cercanía con otros departamentos. En presencia de un choque positivo una determinada ubicación, no solo crecerá más rápido la producción del departamento que recibe este impulso, sino también los otros cercanos a él. Hay un efecto contagio.

Los coeficientes de las variables regresoras en el modelo SEM fueron significativas, a excepción de la variable crédito. Es muy relevante el resultado relacionado con la población rural frente a la producción, pues marca una necesidad de intervención a nivel de política pública sobre la intensidad del flujo de población de lo rural hacia lo urbano.

Los efectos individuales de los departamentos muestran que hay una contribución positiva en departamentos como Valle del Cauca, Meta, Nariño y Caquetá.

Se recomienda para futuros ejercicios, diversificar por municipio y por tipo de producto. Es muy posible que el parámetro espacial autorregresivo si sea significativo al especificar aún más la unidad geográfica, además de capturar mejor las relaciones frente a una cadena productiva diferenciada y no agregada.

9. Referencias bibliográficas.

- Álvarez, R., Mamani, E., & Bonilla, J. (2004). *Economía de la producción de bienes agrícolas: teorías y aplicaciones*. Bogotá: CEDE Uniandes.
- Baltagi, B. H. (2005). *Econometric Analysis of Panel Data*. New Jersey: Wiley.
- Baptista, P., Fernández-Collado, C., & Hernández, R. (2006). *Metodología de la Investigación*. Ciudad de México: McGraw-Hill.
- Beltrán, J., & Piñeros, A. (2013). Sector agropecuario colombiano: realidad y perspectiva. *Trabajo de grado*, EAN.
- Departamento Nacional de Estadística. (18 de Octubre de 2017). PIB 2016 - Participación por rama de actividad. [Conjunto de datos]. Bogotá: DANE.
- Elhorst, J. (2010). Spatial Panel Data Models. En F. Manfred, & A. Getis, *Handbook of Applied Spatial Analysis* (págs. 377 - 407). Berlin: Springer.
- Elhorst, J. (2014). *Spatial Econometrics From Cross-Sectional Data to Spatial Panels*. New York: Springer.
- Fajardo, D. (2002). *Tierra, poder político y reformas agraria y rural*. Bogotá: ILSA.
- Getis, A., & Aldstadt, J. (2010). Constructing the Spatial Weights Matrix Using a Local Statistic. En L. Anselin, & S. Rey, *Perspectives on Spatial Data Analysis* (págs. 147-163). Berlin: Springer.
- Gómez, C., Velásquez, H., Urrego, J., & Valderrama, J. (2014). *Efecto de los ingresos permanentes sobre el delito: un enfoque espacial y un caso de aplicación*. Medellín: Documentos de Trabajo: Universidad EAFIT.
- Gujarati, D., & Porter, D. (2010). *Econometría*. Ciudad de México: McGraw-Hill.
- Hansen, B. (2017). *Econometrics*. Wisconsin: University of Wisconsin.
- Hsiao, C. (2003). *Analisis of Panel Data*. Cambridge: University of Cambridge.
- Jaramillo, C. (1998). La agricultura colombiana en la década de los noventa. *Revista de economía de la Universidad del Rosario*, 9 - 39.
- Kuscevic, C., & Del Río, M. (2013). Convergencia en Bolivia: un enfoque espacial con datos panel dinámicos. *Revista de economía del Rosario*, 233 - 256.
- Labra, R., & Torrecillas, C. (2014). *Guía cero para datos panel. Un enfoque práctico*. Working Paper: Univeridad Autónoma de Madrid.

- Lasso, F. (2015). *¿Cuáles son las variables que determinan la estructura de endeudamiento de las empresas latinoamericanas, y específicamente las colombianas?* Tesis de Maestría: Universidad de Buenos Aires.
- Millo, G., & Piras, G. (2012). splm: Spatial Panel Data Models in R. *Journal of Statistical Software*, 1 - 47.
- Ministerio de Agricultura y Desarrollo Rural. (25 de Agosto de 2017). Evaluaciones Agropecuarias Municipales 2007 - 2015 . [Conjunto de Datos]. Bogotá: MinAgricultura.
- Montero, R. (2011). *Efectos fijos o aleatorios: test de especificación*. Granada: Documento de Trabajo - Universidad de Granada.
- Organización de las Naciones Unidas para la Alimentación y la Agricultura. (2004). *Política de desarrollo agrícola: conceptos y principios*. Roma: FAO.
- Organización para la Cooperación y el Desarrollo Económico. (2015). *Revisión de la OCDE de las políticas agrícolas: Colombia 2015*. Bogotá: OCDE.
- Perfeti, J., Hernández, A., Balcázar, A., & Leibovich, J. (2013). *Políticas para el desarrollo de la agricultura en Colombia*. Bogotá: Fedesarrollo.
- Rosales, R., Perdomo, J., Morales, C., & Urrego, J. (2013). *Fundamentos de econometría intermedia*. Bogotá: Uniandes.
- Unidad de Planificación Rural Agropecuaria. (2013). *Evaluación de tierras para la zonificación con interés agropecuario. Metodología a Escala General (1:100.000)*. Bogotá: UPRA.
- Wooldridge, J. (2010). *Introductory Econometrics: A Modern Approach*. Mason: CENGAGE Learning.