

EL ANÁLISIS EXPLORATORIO DE DATOS ESPACIALES COMO HERRAMIENTA DE ESTUDIOS GEOGRÁFICOS

Meza Julio Cesar

julio_4152@hotmail.com

Arias Federico Carlos

fedearias.-@hotmail.com

Facultad de Humanidades, Universidad Nacional del Nordeste

Eje Temático: N° 3: Investigar y enseñar en Geografía: innovaciones metodológicas y experiencias pedagógicas.

Resumen:

El presente trabajo es de carácter exploratorio y partió en el marco del curso de actualización y perfeccionamiento denominado: análisis exploratorio de datos espaciales y estadística espacial dictado por la Dra. Liliana Ramírez, Prof. Sivia Ferreyra, Prof. Norma Monzón, Prof. Romina Claret.

Los objetivos propuestos son los siguientes: identificar las potencialidades que tienen los análisis exploratorios de datos espaciales y advertir la utilidad de diferentes técnicas estadísticas para la obtención de resultados confiables en la investigación geográfica.

En cuanto a la metodología empleada se basó en: a)- Selección de variables socio-demográficas de la República Argentina y la Provincia del Chaco en los años 2001-2010 b)- Aplicación de técnicas estadísticas descriptivas (medidas de tendencia central, medidas de posición no centrales y las medidas de dispersión), medidas de distribución espacial (centro medio, distancia estándar y elipse de desviación estándar), asociación entre variables y autocorrelación espacial, a partir de un software SIG.

En este sentido, como resultado parcial a fin de contribuir en la investigación geográfica, ya sea para planificación, gestión u ordenamiento territorial, se puede destacar que los análisis exploratorios de datos espaciales nos permiten encontrar patrones de comportamientos o anomalías en la estructura de los datos, tener una visión general de la localización de las variables y comprobar los supuestos necesarios para la aplicación de test estadísticos. Pero desde el punto de vista geográfico este análisis no es suficiente, porque se lo debe contextualizar con información del espacio en estudio para así lograr una mejor interpretación de la realidad y hacer más factible la toma de decisiones y/o intervenciones. En cuanto a las variables utilizadas se puede destacar que existe (según las técnicas estadísticas aplicadas) un cambio en las condiciones socio-demográficas de la República Argentina y la Provincia del Chaco entre ambos censos.

Palabras claves: Estadística - Datos Espaciales – SIG - Argentina/Chaco.

INTRODUCCIÓN

*“El análisis exploratorio de datos (EDA, Exploratory Data Analysis) se considera un instrumento indispensable al momento de realizar las primeras aproximaciones al estudio de la estructura de la información socio-espacial en una determinada área de estudio”.*¹

“A partir del trabajo de Tukey (1977) se han comenzado a afianzar y difundir las técnicas del análisis exploratorio de datos como herramientas que permiten realizar un estudio inicial de los mismos en aproximaciones gráficas previas a la realización de procedimientos estadísticos de mayor complejidad”. Según Fotheringham et al. (2000), “*constituye una tendencia importante del desarrollo metodológico cuantitativo actual, potenciada por los nuevos desarrollos digitales en materia de representación”*.²

Las aplicaciones incluidas en el Análisis Exploratorio de Datos Espaciales tienen por objetivo brindar información estructural del comportamiento de una o más variables. También *permite descubrir errores en la codificación de los datos, determinar los casos anómalos (outliers) y la posibilidad de comprobar supuestos necesarios para la aplicación de la mayoría de los test estadísticos.*³ Asimismo actúa como intermediario entre las bases de datos alfanuméricas y la cartografía digital, ampliando sus posibilidades de manera notable” (Anselin, 1998.)⁴

En función de lo antes expuesto el presente trabajo posee carácter exploratorio el cual partió en el marco del curso de actualización y perfeccionamiento denominado: análisis exploratorio de datos espaciales y estadística espacial dictado por la Dra. Liliana Ramírez, Prof. Sivia Ferreyra, Prof. Norma Monzón, Prof. Romina Claret. Es por ello que se busca un acercamiento al conocimiento de las potencialidades que posee la metodología cuantitativa dentro de la investigación geográfica influenciada por el desarrollo de la tecnología.

Para ello se presentan algunos ejemplos de aplicación de Análisis Exploratorio de Datos Espaciales a partir de la utilización del software GeoDa y ArcGis.

OBJETIVOS

- Identificar las potencialidades que tienen los Análisis Exploratorios de Datos Espaciales.
- Advertir la utilidad de diferentes técnicas estadísticas para la obtención de resultados confiables en la investigación geográfica.

¹ BUZAI, Gustavo D. y BAXENDALE Claudia A. (2012). *ANALISIS SOCIOESPACIAL CON SISTEMAS DE INFORMACIÓN GEOGRAFICA*. Ordenamiento Territorial. Temáticas de base Vectorial. Buenos Aires. Lugar Editorial. Pág. 315.

²³⁴ BUZAI, Gustavo D. y BAXENDALE Claudia A. (2009). *ANÁLISIS EXPLORATORIO DE DATOS ESPACIALES*. Revista digital del Grupo de Estudios sobre Geografía y Análisis Espacial con Sistemas de Información Geográfica (GESIG). Universidad Nacional de Lujan.11. Disponible en Internet: <http://www.gesig-proeg.com.ar/documentos/revista-geosig/BUZAI-BAXENDALE2009.pdf>

METODOLOGÍA

a)- Selección de variables socio-demográficas de la República Argentina y la Provincia del Chaco en los años 2001-2010:

- Necesidades Básicas Insatisfechas (NBI) año 2001-2010. República Argentina.
- Superficie Cosechada de soja y maíz campañas 1999/2000- 2008/2009. Provincia del Chaco.
- Tipo de vivienda: Porcentaje de Ranchos y Porcentajes de Casas años 2001-2010. Provincia del Chaco.

b)- Aplicación de técnicas estadísticas descriptivas (medidas de tendencia central, medidas de posición no centrales y las medidas de dispersión), medidas de distribución espacial (centro medio, distancia estándar y elipse de desviación estándar), asociación entre variables y autocorrelación espacial, a partir de un software SIG.

- 1) Empleo de medidas de tendencia central (media, mediana y moda), las medidas de posición no centrales y las medidas de dispersión.
 - Histograma de frecuencias.
 - Índices de apuntalamiento y el sesgo (Kurtosis y Skewness).
 - El diagrama/mapa de caja.
- 2) Empleo de Estadísticas Espaciales Básicas de Localización Puntual.
 - Centro medio.
 - Distancia Estándar.
 - Elipse de Desviación.
- 3) Aplicación del Coeficiente de Correlación de Pearson y Diagrama de Dispersión
- 4) Aplicación del Índice Global de Moran y el Índice Local de Moran.

RESULTADOS OBTENIDOS

1) A partir del módulo de Geoestadística de ArcGis se elaboró el histograma de frecuencias de la variable Porcentaje de viviendas con NBI 2001 y 2010 a nivel departamental en la República Argentina.

Teniendo en cuenta los estadísticos descriptivos considerados para el AEDE los histogramas nos muestran una agrupación de datos en torno a valores bajos, el cuadro estadístico que acompaña los histogramas (ver figura N°1 y 2) nos brinda información estadística de la centralidad, la dispersión y forma de los datos.

Las **medidas de centralidad** nos dan la idea de donde está el centro y otras partes de importancia dentro de la distribución, ellas son: la media, es decir, el promedio de todos los valores,

siendo en el año 2001 de 16,1% y de 12,3 % en 2010; la mediana, que representa un 14,6% en 2001 y 11% en 2010, es el valor tal que ordenados los datos de menor a mayor, deja por debajo y por arriba el 50 % de éstos.

Las **medidas de posición no centrales**, entendidas como aquellas que nos dividen el conjunto ordenado en partes de igual cantidad de datos, tenemos en el año 2001: el 1º cuartil 8,0% y 3º cuartil 21,9 % y 2010: 1º cuartil 5,6%, 3º cuartil 16,8%, los primero dejan el 25% de los datos por debajo de éstos y los segundos el 75% de los mismos por debajo de aquellos. Entre ambos se encuentra el 50% de los valores.

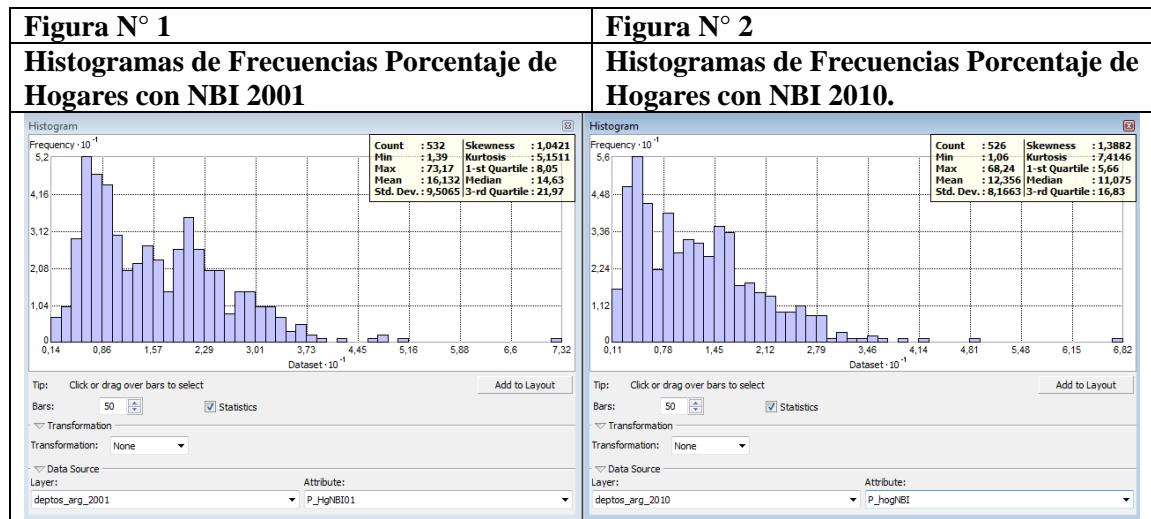
En cuanto a las **medidas de dispersión** proporcionan una información parcial de la distribución de la variable, indican la mayor o menor agrupación de los datos alrededor del valor medio, la desviación estándar es de 9,5% para el primer año y 8,1 % para el segundo, es decir, una desviación de los datos respectos a las medias. Finalmente para medir la asimetría utilizamos el parámetro denominado sesgo, cuyo valor nos da idea de la concentración de datos a un lado y al otro de la media, y para medir el grado de apuntamiento de la curva, si es muy aguda hacia arriba o muy aplanada utilizamos el valor de curtosis, en este caso tenemos un sesgo de 1,04 % en 2001 y 1,38 % en 2010, ambos casos con un agrupamiento asimétrico de datos hacia la derecha. La distribución concentra valores a la izquierda de la distribución y se denomina distribución asimétrica hacia la derecha. Cuando la distribución es asimétrica se está ante la misma situación que cuando hablamos de valores atípicos de la variable. De acuerdo a la dirección de la asimetría existen valores atípicos muy pequeños o muy grandes, estos valores distorsionan el valor que toma la media aritmética, la que pierde representatividad.

Curtosis para el año 2001 es de 5, 1 % y 2010 es de 7, 4 %, teniendo en cuenta el valor de este parámetro las curvas reciben diferentes denominaciones: si el valor es mayo a 3: leptocúrtica, igual a 3: mesocúrtica y menor a 3 platicúrtica. Dicho esto, podemos afirmar para ambos años que la distribución es de tipo leptocúrtica, esto indica una gran concentración de los datos en torno sus valores medios.

Otros datos de interés que nos brinda el cuadro estadístico es el valor mínimo y máximo de la serie de datos que nos permite inferior si hubo un aumento o disminución de valores, en 2001 y 2010 los valores mínimos son: 1,39 % y 1,06% y los máximos son: 73,17% y 68,24% respectivamente, lo que nos muestra que las condiciones de los hogares a mejorado entre ambos censos.

En síntesis la utilización de histogramas presenta la ventaja de mostrar una distribución de datos ordenados y agrupados de manera rápida, debido al impacto visual que provoca. Las columnas representan la frecuencia del número de elementos por estar estos agrupados en intervalos. Sin embargo, no se pueden apreciar valores exactos y tampoco se consideran las ubicaciones de las

entidades con las que se asocian los valores, ni ninguna relación espacial entre las entidades pero a la hora de examinar la distribución de datos referentes a una variable son sumamente útiles.



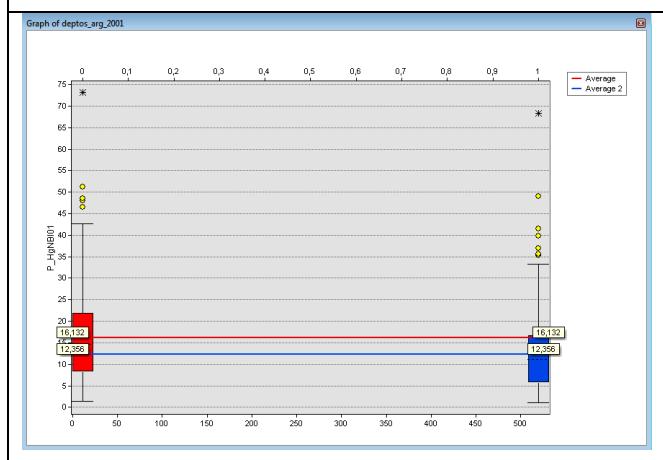
En cuanto al diagrama de caja es *un método de representación univariante utilizado por la perspectiva reticular o “lattice” que considera que los datos espaciales son realizaciones individuales de un proceso estocástico (como en el análisis de series temporales), haciendo posible la identificación de puntos atípicos o “outliers” espaciales. La construcción del diagrama de caja parte del cálculo de los cuartiles y la media de una variable, así como de la obtención de las llamadas cotas o valores adyacentes superior e inferior.*⁵ En la Figura N°3 se advierte en la caja izquierda el porcentaje de hogares con NBI para el año 2001 (caja roja) y en la caja de la derecha la misma variable pero para el año 2010 (caja Azul). Comparativamente podemos destacar que para el año 2010 existe un descenso del valor promedio (12,3%) en relación al año 2001 (16,1%) lo que nos da un indicio que en el país al menos en la calidad de la vivienda ha mejorado, aunque es bien sabido que el NBI presenta sus limitaciones a la hora de conocer las características de éstas.

Otro aspecto a observar es el alto de la caja (que muestra la dispersión del 50% de los datos), en el caso del 2010 los valores estaban más concentrados alrededor del 7 y 16% (aprox.) mientras que en 2001 los valores estaban más dispersos lo que nos da un indicio de que se podían encontrar situaciones dispares, además ese 50% de los datos se concentraban alrededor del 8 y 22% (aprox.).

En cuanto a los extremos encontramos para ambos años el mismo mínimo, es decir, un 1% de hogares con NBI pero la diferencia notable está en el máximo, ya que para el año 2010 fue de alrededor el 34% pero para el año 2001 era del 43%, es decir la situación mejoró considerablemente ya que los departamentos con condiciones más críticas han mejorado en la calidad de la vivienda.

⁵ CHASCO YRIGOYEN, Coro. (Año. S/D) *MÉTODOS GRÁFICOS DEL ANÁLISIS EXPLORATORIO DE DATOS ESPACIALES*. Instituto L.R. Klein-Dpto. de Economía Aplicada. Universidad Autónoma de Madrid. Pág. 25. Disponible en Internet: https://www.uam.es/personal_pdi/economicas/coro/investigacion/aede03.pdf.

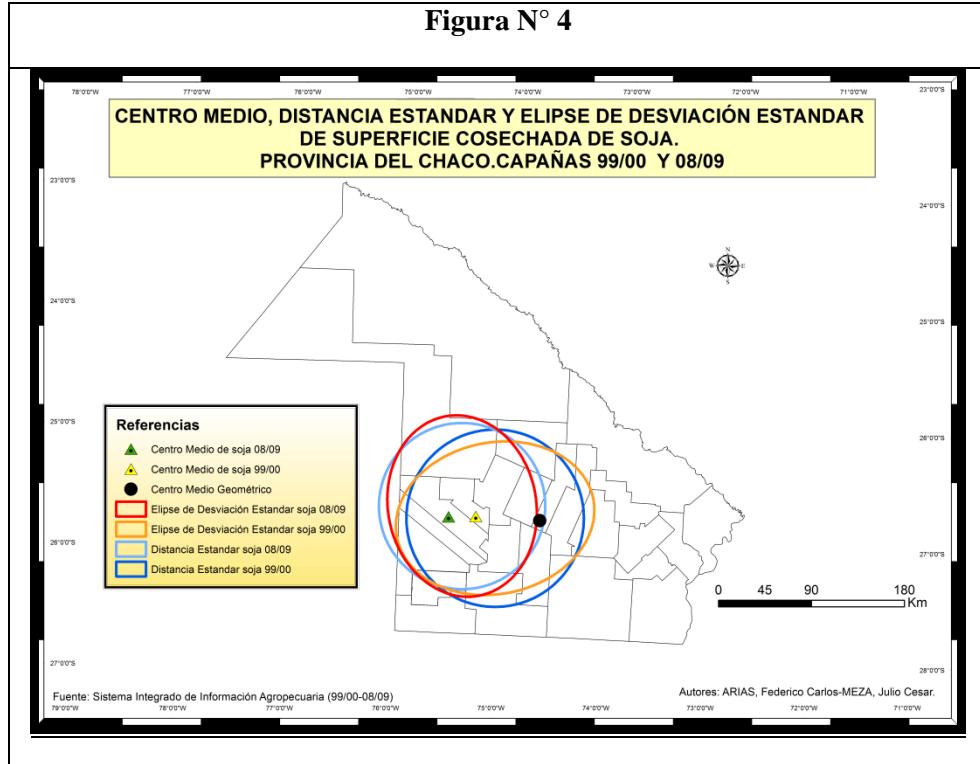
Figura N° 3 Diagrama de caja: NBI años 2001 y 2010



Como se puede advertir el diagrama tiene una utilidad muy significativa, ya que agrupa una gran cantidad de información en un solo gráfico, además permite la comparación entre diferentes años de la misma variable tratando de identificar simetrías o asimetrías en la distribución del fenómeno que se estudie.

2) Distribución espacial de localizaciones puntuales: Cálculo del Centro Medio, Distancia Estándar y Elipse de Desviación Estándar en las siguientes variables: Superficie Cosechada de Soja campañas 99/00 y 08/09- Superficie Cosechada de Maíz campañas 99/00 y 08/09. Provincia del Chaco. (Mediante el módulo de herramientas de estadísticas espaciales de ArcGis 10.1: Spatial Statistics Tools).

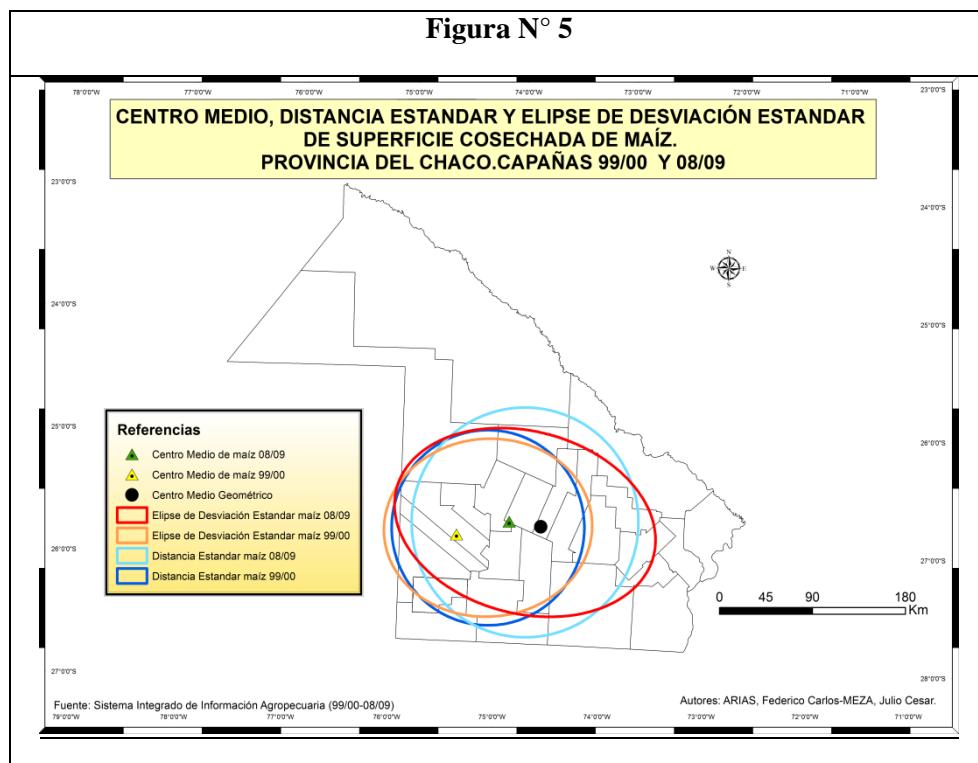
Figura N° 4



En la Figura N°4 se encuentra representada la superficie cosechada de soja para las campañas 99/00 y 08/09 en la Provincia del Chaco. Si se observa la ubicación del centro medio Geométrico (Media Aritmética de la localización de los distintos puntos) y el centro medio de la Sup. Cosechada de Soja para ambas campañas existe un desajuste en sentido este-oeste. Asimismo entre ambas campañas tambien hay un desajuste en el mismo sentido, producto que la actividad en el último tiempo esta experimentando una reubicacion sobre el espacio.

Por otro lado en cuanto a la distancia estandar (distancia media que existen entre la desviación típica de todos los puntos con respecto al centro medio) si observamos el comportamiento de ambas campañas se advierte un alto grado de concentracion de las entidades con respecto a su centro medio hacia el suroeste de la Provincia del Chaco, aunque en la segunda campaña en estudio es mas marcada la concentracion (en el mismo sentido) ya que el tamaño de la circunferencia es menor.

Por ultimo tenemos la elipse de desviacion estandar la cual mide la variabilidad espacial de un mapa de puntos a partir de sus dispersiones en dos direcciones en el plano, una horizontal y otra vertical, buscando las direcciones en las que más se dispersan los puntos. Comparando ambas campañas se advierte una variabilidad espacial notoria de la superficie cosechada, ya que para la primera (99/00) tenia una orientación en sentido NE-SO, mientras que para la segunda (08/09) es de SSE-NNO, es decir, hay un redireccionamiento apparente hacia la zona del impenetrable chaqueño.



En la Figura N°5 se encuentra representada la superficie cosechada de maíz para las campañas 99/00 y 08/09 en la Provincia del Chaco. Si se observa la ubicación del centro medio Geométrico y el centro medio de la Sup. Cosechada de Maíz para la campaña 99/00 el desajuste se da en sentido ENE-

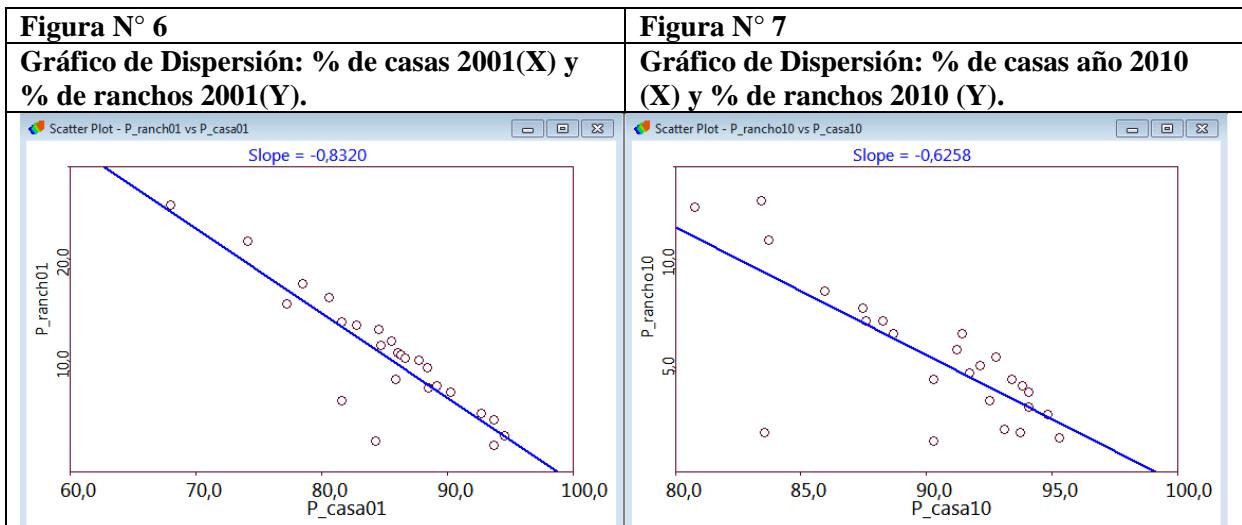
OSO, mientras que en la campaña 08/09 el desajuste se da en sentido E-O. comprando los centros medios de las dos campañas se destaca la reubicación del cultivo dejando de ocupar la zona SO para dirigirse hacia el centro provincial.

Por otro lado la distancia estandar advierte una diferencia de tamaño en las circunferencias entre ambas campañas, lo que denota que para la primera existia una mayor concentracion de la actividad (circulferencia menor) en la zona del SO, pero para la segunda al aumentar el tamaño nos da indicios de una mayor dispersion de la actividad.

Por ultimo tenemos la elipse de desviacion estandar que comparando ambas campañas se advierte una variabilidad espacial notoria de la superficie cosechada de maíz, ya que el mismo se dirige en la actualidad hacia el centro de la provincia (dirección E) dejando de ocupar la zona del SO .

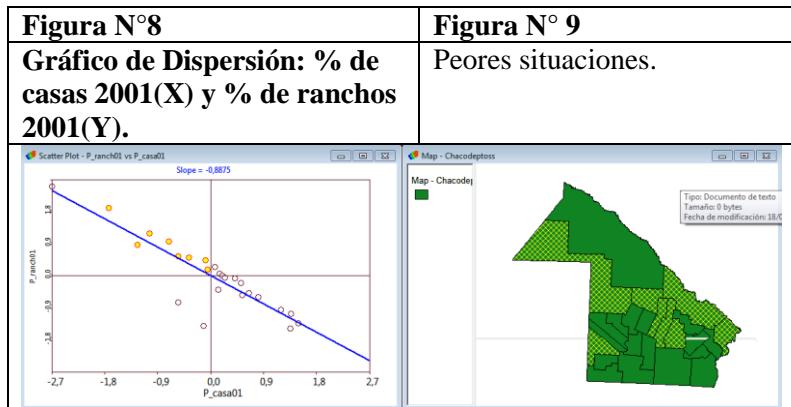
Como se puede advertir estas herramientas: Centro Medio, Distancia Estándar y Elipse de Desviación Estándar si bien en este caso se lo aplicó para estas variables, lo significante es denotar la importancia metodológica que tiene para la geografía, ya que estas técnicas estadísticas nos dan una visión de la localización del fenómeno, el grado de concentración que posee, es decir, si es homogéneo o heterogéneo su distribución y ha presentado variaciones espaciales.

3) Coeficiente de Correlación Lineal de Pearson (mide *el grado de relación lineal entre variables*)⁶ y Diagrama de Dispersion (representación en un sistema de coordenadas puntos correspondientes a pares de datos observados.)⁶ a partir del Sofware Geodata con las variables porcentaje de casas año 2001-2010 y porcentaje de ranchos para los mismos años.

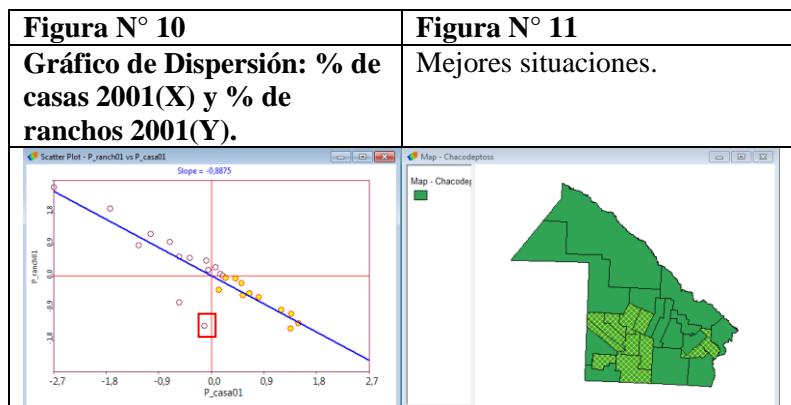


⁶ CHICA DE GALASSI, Nélida et al. (2011). *NOCIONES DE ESTADÍSTICA Y PROBABILIDAD*. Resistencia. Ediciones DE LA PAZ. Pág. 167.

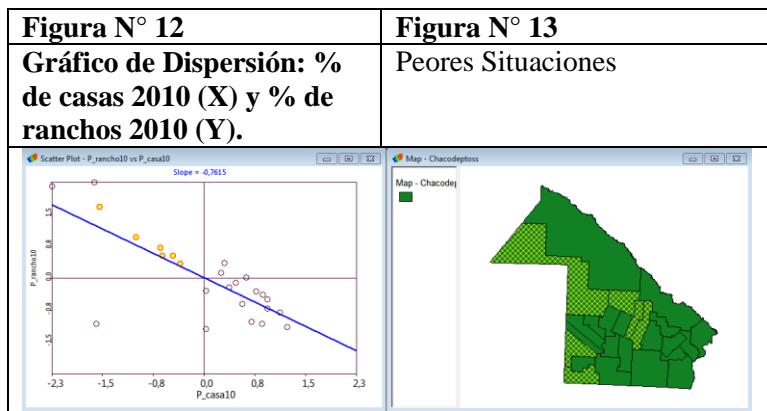
Si comparamos la correlación que existe entre las dos variables en estudio para los distintos años, es inversa ya que los valores son negativos: -0,83 en la figura N° 6 y -0,62 en la figura N°7. En cuanto a la intensidad de la relación, existe una variación entre ambos años ya que en el 2001 es fuerte (-0,8 y -1) pero para el 2010 es moderada.



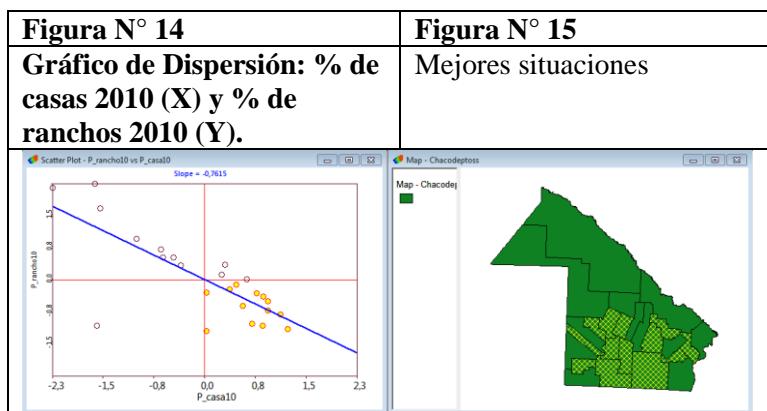
Ahora bien si observamos las figuras N°8 y N°9 se advierte cuáles son los departamentos para el año 2001 dentro de la Provincia del Chaco que tienen las peores condiciones (en líneas generales al NO y NE), es decir, que tiene una predominancia de tipos de viviendas ranchos por sobre casa.



En cuanto a los departamentos que presentan las mejores situaciones se los puede apreciar en la figura N°11 ubicadas al SO provincial, es decir donde tenemos la predominancia de casas por sobre los ranchos. Una cuestión llamativa es que no aparece representado en este caso el departamento San Fernando, lo cual es producto de que es un Outlier (valor atípico) identificado en la figura N°10.



En el caso del año 2010 las peores situaciones se mantienen hacia el NO y O pero los departamentos del NE mejoran sus situaciones.



En cuanto a las mejores situaciones (figura N°15) se dan hacia el sur de la provincia, donde vuelve a no estar representado porque es un outlier (valor atípico) identificado en la figura N°14.

*“El desciframiento y la interpretación de todas las relaciones que contribuyen a la constitución de un complejo o de un equilibrio entre la dinámica y la inercia... de las colectividades humanas que residen en ese medio o que en él aplican sus iniciativas. Como plantea la expresión el geógrafo debe a través de las interacciones entre elementos explicar las relaciones espaciales”.*⁷ Es por ello que en función de este marco el Coeficiente de Correlación de Pearson y el Diagrama de Dispersion son muy importante para identificar si existe algún tipo de relación entre ciertos tipos de variables la intensidad de la misma y donde se localizan las mejores o peores situaciones de la esa relación en el caso de que sea significativa.

- 4) Como señala Buzai y Baxendale (2012) “El estudio de la diferenciación de espacios sobre la superficie terrestre es una aproximación tradicional de los estudios geográficos, por tal motivo, los

⁷ MARTIN HERNANDO, Miguel A. (1991) *EL ANÁLISIS DE CORRELACIÓN EN EL ESTUDIO GEOGRÁFICO DEL COMERCIO. PROYECCIÓN EN EL SURESTE DE LA REGIÓN DE MURCIA*. *Papeles de Geografía N.*º 17, 1991, Pág. 187-200. Disponible en Internet: <http://revistas.um.es/geografia/article/view/43791/41891>.

mapas temáticos de distribuciones espaciales de variables y luego el cálculo de índices de concentración han ocupado un lugar de importancia”.⁸

La Autocorrelación Espacial “*es la concentración o dispersión de los valores de una variable en un mapa. Refleja el grado en que objetos o actividades en una unidad geográfica son similares a otros objetos o actividades en unidades geográficas próximas*” (Goodchild, 1987).⁹ Chasco Irigoyen la define como “*el efecto que consiste en la existencia de una relación funcional entre lo que ocurre en un punto determinado del espacio y en lugares cercanos o vecinos. Es decir, una variable se encontrará espacialmente autocorrelacionada cuando los valores observados en un punto o región dependan de los valores observados en regiones vecinas, de forma que se produzca una cierta continuidad geográfica en la distribución de esta variable*”.¹⁰

Como ejemplo de aplicación se utilizó la variable Porcentajes de hogares con NBI 2001 y 2010. Las medidas de autocorrelación espacial se pueden diferenciar por el alcance o la escala de análisis. Tradicionalmente, se separan en categorías: Índice Global de Moran y el Índice Local de Moran.

Utilizando las Herramientas de estadísticas espaciales de ArcGis (Spatial Statistics Tools) se calcularon ambos parámetros.

El índice Global de Moran ofrece una medida resumen de la intensidad de la autocorrelación de los territorios considerados. Varía entre los valores -1 y +1; cuanto más cercano a 1 sea el indicador, mayor será el nivel de autocorrelación espacial, en este caso para el año 2001 fue de 0,63 y de 0,47 en 2010 (Observar figuras N°16-17) esto nos da un indicio que existía una mayor autocorrelación entre las unidades espaciales para el año 2001 por encontrarse más cercano al valor 1. Sin embargo, carece del detalle de las correlaciones entre las unidades geográficas constituyentes del territorio (vecinas o no).

Índice de moran local: permite identificar la localización de los conglomerados espaciales, cuyas presencias fueron definidas por el índice de moran global, para poder construir los mapas de conglomerados calientes (hot spots).

Al igual que el indicador global, el I. de moran local varía entre -1 y +1, representando el grado de correlación del indicador de una unidad territorial con los indicadores de sus vecinas. Como resultado, el índice identifica unidades territoriales donde valores de análisis altos o bajos se agrupan

⁸ BUZAI, Gustavo D. y BAXENDALE Claudia A. (2012). *ANALISIS SOCIOESPACIAL CON SISTEMAS DE INFORMACIÓN GEOGRAFICA*. Ordenamiento Territorial. Temáticas de base Vectorial. Buenos Aires. Lugar Editorial. Pág. 315.

⁹ CEPAL – NACIONES UNIDAS. *GUIA PARA ESTIMAR LA POBREZA INFANTIL*. Disponible en Internet: <HTTP://DDS.CEPAL.ORG/INFANCIA/GUIA-PARA-ESTIMAR-LA-POBREZA-INFANTIL/GUIA> CONTENIDO-442.PHP

¹⁰ CHASCO YRIGOYEN, Coro. (Año. S/D) *MÉTODOS GRÁFICOS DEL ANÁLISIS EXPLORATORIO DE DATOS ESPACIALES*. Instituto L.R. Klein-Dpto. de Economía Aplicada. Universidad Autónoma de Madrid. Pág. 25. Disponible en Internet: https://www.uam.es/personal_pdi/economicas/coro/investigacion/aede03.pdf.

espacialmente, así como también unidades territoriales con valores muy distintos a los de las áreas circundantes.

Por tanto, este índice puede reconocer cinco tipos de conglomerados espaciales:

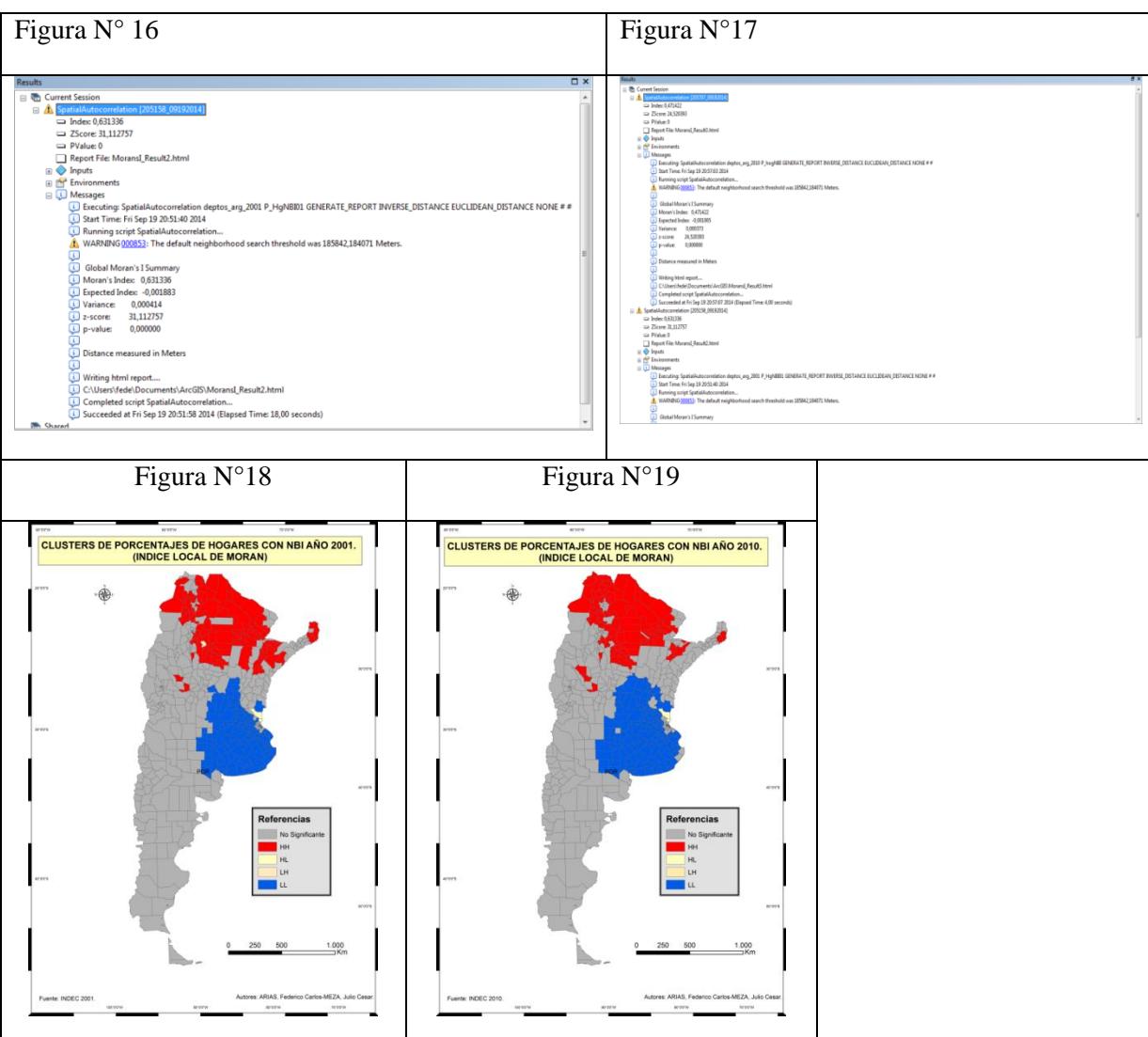
Alto-alto: una unidad territorial con un valor de análisis por encima del promedio, rodeada significativamente por áreas vecinas que también se encuentran por sobre la media con respecto a la variable de interés. hot spots.

Bajo-bajo: una unidad territorial con un valor de análisis inferior al promedio, rodeada por áreas vecinas que también se encuentran bajo la media en relación con la variable de interés. cold spots.

Bajo-alto: presencia de una unidad territorial con un valor de análisis bajo, rodeada significativamente por áreas vecinas con valores que se encuentran por sobre la media de la variable de interés.

Alto-bajo: presencia de una unidad territorial con un valor de análisis alto, rodeada significativamente por áreas vecinas con valores que se encuentran bajo la media de la variable de interés.

No Significante: No existe Autocorrelación.



Teniendo en cuenta la figura N°18 se aprecia que las áreas de **alto-alto**, es decir, unidades territoriales con valores de análisis por encima del promedio, se localizan en las provincias del norte del país, y los valores **bajo-bajo**, valores inferiores al promedio, se encuentran en el centro-este. En el caso de la figura N°19 se mantiene en líneas generales el mismo comportamiento solo modificando sus condiciones algunos departamentos pero que no alcanzan a ser muy representativos espacialmente (en función de la escala de estudio)

(Buzai y Baxendale, 2012) señala: “todas las posibilidades indudablemente amplian las visiones espaciales que se pueden tener de la diferenciación de espacios dentro de un área de estudio, logran una aproximación más detallada al análisis de las relaciones de concentración espacial y generan nuevas posibilidades para el análisis de difusión por contagio espacial.”¹¹

Como se percibe la autocorrelación espacial demuestra utilidad para ver como influye el espacio en el comportamiento correspondiente a la distribución de una variable, en este ejemplo el porcentaje de hogares con NBI. Este procedimiento tiene como principal objetivo comprobar si las unidades espaciales próximas tienen similitud en cuanto a su comportamiento respecto a las más alejadas y es allí donde recae la importancia de este procedimiento metodológico.

CONSIDERACIONES FINALES:

El análisis exploratorio de datos resulta para la ciencia geográfica una herramienta trascendental para acercarse a la interpretación de los hechos o fenómenos que se dan en el espacio geográfico. El mismo nos permite encontrar patrones de comportamientos o anomalías en la estructura de los datos, tener una visión general de la localización de las variables y comprobar los supuestos necesarios para la aplicación de test estadísticos, se pueden obtener resultados significativos que luego habría que ponerlos en contexto para un mejor análisis. Asimismo toma trascendencia relacionar gráficos con cartografía para la identificación de distribuciones espaciales.

En este sentido las variables que se tomaron tanto en la Provincia del Chaco como en la República Argentina, han presentado modificaciones a considerar, aunque es cierto que el dinamismo se encuentra fundado en el tipo de variable, es decir, no presentan los mismos cambios el NBI entre dos censos que el tipo de vivienda o un tipo de cultivo, lo cual es vital tener en cuenta para estudios de esta índole, y así poder inferir posteriormente en las estrategias que se consideren necesarias en pos de aprovechamientos u ordenamientos territoriales.

¹¹ BUZAI, Gustavo D. y BAXENDALE Claudia A. (2012). *ANALISIS SOCIOESPACIAL CON SISTEMAS DE INFORMACIÓN GEOGRAFICA*. Ordenamiento Territorial. Temáticas de base Vectorial. Buenos Aires. Lugar Editorial. Pág. 315.

BIBLIOGRAFÍA

- BUZAI, Gustavo D. y BAXENDALE Claudia A. (2012). *ANALISIS SOCIOESPACIAL CON SISTEMAS DE INFORMACIÓN GEOGRAFICA*. Ordenamiento Territorial. Temáticas de base Vectorial. Buenos Aires. Lugar Editorial. Pág. 315
- BUZAI, Gustavo D. (2008). *SISTEMAS DE INFORMACIÓN GEOGRÁFICA Y CARTOGRAFÍA TEMÁTICA*. Métodos y Técnicas para el trabajo en el aula. Buenos Aires. Lugar Editorial. Pág. 131.
- BUZAI, Gustavo D. y BAXENDALE Claudia A. (2009). *ANÁLISIS EXPLORATORIO DE DATOS ESPACIALES*. Revista digital del Grupo de Estudios sobre Geografía y Análisis Espacial con Sistemas de Información Geográfica (GESIG). Universidad Nacional de Lujan.11. Disponible en Internet: <http://www.gesig-proeg.com.ar/documentos/revista-geosig/BUZAI-BAXENDALE2009.pdf>
- CEPAL – NACIONES UNIDAS. *GUIA PARA ESTIMAR LA POBREZA INFANTIL*. Disponible en Internet: <HTTP://DDS.CEPAL.ORG/INFANCIA/GUIA-PARA-ESTIMAR-LA-POBREZAINFANTIL/GUIA CONTENIDO-442.PHP>
- CHASCO YRIGOYEN, Coro. (Año. S/D) *MÉTODOS GRÁFICOS DEL ANÁLISIS EXPLORATORIO DE DATOS ESPACIALES*. Instituto L.R. Klein-Dpto. de Economía Aplicada. Universidad Autónoma de Madrid. Pág. 25. Disponible en Internet: https://www.uam.es/personal_pdi/economicas/coro/investigacion/aede03.pdf.
- CHICA DE GALASSI, Nélida et al. (2011). *NOCIONES DE ESTADÍSTICA Y PROBABILIDAD*. Resistencia. Ediciones DE LA PAZ. Pág. 167.
- MARTIN HERNANDO, Miguel A. (1991) *EL ANÁLISIS DE CORRELACIÓN EN EL ESTUDIO GEOGRÁFICO DEL COMERCIO. PROYECCIÓN EN EL SURESTE DE LA REGIÓN DE MURCIA*. Papeles de Geografía N." 17, 1991, Pág. 187-200. Disponible en Internet: <http://revistas.um.es/geografia/article/view/43791/41891>.
- ESRI Help. <http://help.arcgis.com/es/arcgisdesktop/10.0/> Redland, USA.
- GOODCHILD Michael F. y HAINING Robert P. (Año: S/D) *SIG Y ANÁLISIS ESPACIAL DE DATOS: PERSPECTIVAS CONVERGENTES*. *Investigaciones Regionales. 6 – Sección PANORAMA Y DEBATES* Pág. 175 a 201. Disponible en Internet: <http://www.aecr.org/images/ImatgesArticles/2007/09%20Goodchild%20y%20Haining.pdf>
- SÁNCHEZ OGALLAR, Antonio. (1999) .*CONOCIMIENTO GEOGRÁFICO. Procedimientos y Técnicas para el aula en Secundaria*. Narcea. S.A Ediciones NE. Pág. S/D. Disponible en Internet:
http://books.google.com.ar/books?id=Hq0P62_vxVoC&pg=PA132&lpg=PA132&dq=coeficiente+de+correlacione+en+la+geografia&source=bl&ots=FV9Uyrzuie&sig=zMw-OQRRX3iLYCkG1Sr0IXEg3Q4&hl=es-419&sa=X&ei=UykeVLHfHq3bsATSw4IY&ved=0CDAQ6AEwAg#v=onepage&q=coefficient&e%20de%20correlacione%20en%20la%20geografia&f=false