

CARACTERIZACIÓN DE LOS PROCESOS ESPACIALES Y TEMPORALES Y SUS INTERRELACIONES EN ESTACIONES DE PRECIPITACIÓN MENSUAL EN LA ZONA CENTRO OCCIDENTAL DE LA REPÚBLICA BOLIVARIANA DE VENEZUELA

JESÚS ENRIQUE ANDRADES GRASSI

Universidad de Los Andes, Mérida Venezuela. República Bolivariana de Venezuela
koflasjesus@gmail.com

RESUMEN

Venezuela dispone de una red de estaciones de precipitación mensual producto de geocodificaciones (escala y densidad espacial desconocidas) con datos discontinuos en el espacio y en el tiempo (datos faltantes y con distinta fecha de instalación y desinstalación). En este trabajo se caracterizaron las propiedades de primer orden (escala y densidad espacial) estimando el análisis del vecino más cercano y la función K de Ripley a través de su parámetro $L(d)$, así como las propiedades de segundo orden (Autocorrelación Espacio Temporal) utilizando el I de Moran y su variante de Clúster LISA, ambas en su versión espacio temporal, en 961 estaciones con climas diversos en 11 estados de la zona centro occidental de la República Bolivariana de Venezuela. Se determinó que las estaciones poseen al menos tres escalas, un proceso puntual de tipo agregado y un mecanismo estocástico denominado Autocorrelación Espacio Temporal de tipo positiva que condiciona cualquier modelamiento de los datos. El Análisis Clúster permitió categorizar las estaciones utilizando el análisis de las K-Medias. Sin embargo, en este estudio, se incorpora un mecanismo de orden espacio temporal que deberá ser tomado en cuenta en estudios posteriores.

Palabras clave: Datos Espacio Temporales, Autocorrelación, Densidad espacial, Escala, Precipitación.

CHARACTERIZATION OF THE SPATIO TEMPORAL PROCESSES AND THEIR INTERRELATIONS IN MONTHLY RAINFALL DATA OF THE WEST CENTRAL REGION OF REPUBLICA BOLIVARIANA DE VENEZUELA

ABSTRACT

Venezuela has a network of monthly precipitation stations obtained from geocoding (unknown scale and spatial density) with discontinuous data in space time (missing data, and different date of installation and uninstallation). First order properties were characterized (scale and spatial density) estimating the nearest neighbour analysis, and Ripley's K function through the $L(d)$ parameter. The second-order properties (Spatio Temporal Autocorrelation) were characterized

Andrades, J. E. (2015): "Caracterización de los procesos espaciales y temporales y sus interrelaciones en estaciones de precipitación mensual en la zona centro occidental de la República Bolivariana de Venezuela", *GeoFocus (Artículos)*, n° 16, p. 153-180. ISSN: 1578-5157

using Moran's I and its variant LISA Cluster, both in their spatio-temporal versions. The data was obtained from 961 stations with different climates in 11 states of the West Central Region of República Bolivariana de Venezuela. Results indicate that the stations have at least three scales, a Cluster point process and a stochastic mechanism called positive Space Temporal Autocorrelation which conditions any data modelling. The Cluster Analysis using K-Means allowed the organization of the stations in five categories. However, the present study incorporates a mechanism of spatio temporal order that should be taken into account in future studies.

Keywords: Spatio Temporal Data, Autocorrelation, Spatial Density, Scale, Precipitation.

1. Introducción.

Manejar datos espacio temporales, es un tema común en las variables climáticas y en los procesos de recolección de datos. Una gran cantidad de observaciones se pueden obtener a partir de numerosas ubicaciones espaciales a través del tiempo. Según Gujarati (2004) este tipo de datos se denomina como datos agrupados (series de tiempo con datos transversales en el espacio). Con los avances tecnológicos, se han generado nuevos retos computacionales para manejar este tipo de datos altamente volumétricos (Androva y Boland, 2011). Normalmente para su modelamiento se evita obviar alguno de los dos componentes, sea espacial o temporal o incluso ambos (en este último caso se supone la independencia y la aleatoriedad de los datos). Sin embargo, ello implica la posibilidad de estar obviando información relevante del fenómeno que se desea caracterizar (Cressie y Winkle, 2011). Cocchi y Bruno (2010) mencionan que la estadística moderna está introduciendo progresivamente el concepto de que los datos no provienen de un proceso estocástico aleatorio e independiente. En los datos espaciales es común que se manifieste la Primera Ley Geográfica de Tobler (1970), que establece que *todo está relacionado con todo, pero las cosas próximas entre sí, están más relacionadas que las distantes*. Desde una perspectiva univariante, esto significa que una variable (z) puede estar correlacionada consigo misma, pero en otra posición en el espacio (x, y), lo que se denomina *Autocorrelación Espacial* (AE). Si se considera un fenómeno espacio temporal como la precipitación u otra variable climática, definirla por la AE puede ser limitado. Es posible que el fenómeno que va a ser caracterizado posea un arreglo más allá de lo espacial. En otras palabras que el grado de asociación unívoca de la variable (es decir consigo misma) sea tridimensional. Es posible que ocurra un conjunto de datos para una variable (z) con localizaciones espaciales (x, y) y temporales (t) que en su conjunto (x, y, t), posea un orden definido (Stojanova, 2012), lo que se denomina *Autocorrelación Espacio Temporal* (AET). La AE y la AET, llamadas por los autores como Propiedades de Segundo Orden, pueden clasificarse en tres tipos (véase figura 1):

Positiva: Los valores altos suelen tener en su entorno cercanos valores altos, los valores bajos suelen tener en su entorno valores bajos. **Negativa:** los valores altos se rodean de valores bajos y viceversa y **Nula:** los valores de los datos recogidos en una serie de puntos son independientes entre sí y no se ordenan como las dos correlaciones anteriores. En este último caso, se espera que la distancia entre los valores no conduzca a una percepción de correlación obvia, lo que significa que los valores poseen independencia y aleatoriedad espacial y/o temporal.

Estos conceptos implican automáticamente la definición de una vecindad de los distintos elementos geográficos, ya que se tiene que los valores de una variable registrados en aquellos

Andrades, J. E. (2015): “Caracterización de los procesos espaciales y temporales y sus interrelaciones en estaciones de precipitación mensual en la zona centro occidental de la República Bolivariana de Venezuela”, *GeoFocus (Artículos)*, n° 16, p. 153-180. ISSN: 1578-5157

elementos vecinos, ejercen una influencia sobre los valores de dicha variable en un punto dado. Por ello es importante definir cuándo dos elementos son vecinos o no (Olaya, 2011), a esto se le llama matriz de ponderación espacial.

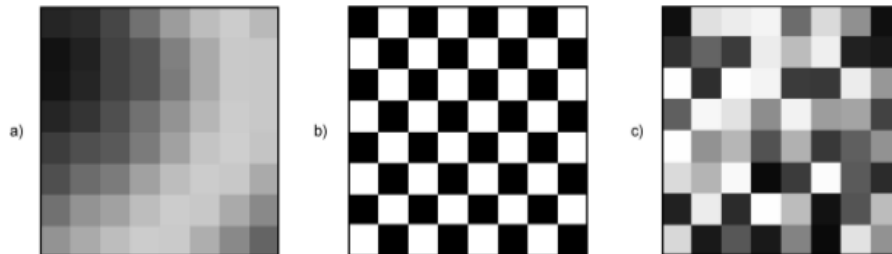


Figura 1: a) Autocorrelación espacial positiva. b) Autocorrelación espacial negativa. c) Ausencia de autocorrelación espacial (independencia) (Olaya, 2011).

Otro elemento importante a considerar es el tipo de vecindad (espacial y espacio temporal) que tienen los valores de la variable. Celemín (2009), Lloyd (2010), Toral (2001), Reynolds y Madden (1988) y Herrera *et al.* (2012) definen múltiples tipos de contigüidad espacio temporal que pueden ser utilizadas.

Para caracterizar si una variable Z posee un proceso AE o AET (propiedades de segundo orden) se pueden enfocar dos alternativas. La primera es modelar la AE o AET; la segunda es medir su intensidad a través del uso de índices (Olaya, 2011). Al estudiar las Propiedades de Segundo Orden, se debe entender que éstas pueden estar influenciadas por las Propiedades de Primer Orden, estas últimas se definen a continuación:

- 1) La escala: un cambio en el distanciamiento de las muestras puede influenciar el signo de AE y AET, sobre todo si la data se creó a partir de coordenadas, es decir, la digitalización directa de valores y coordenadas, sin dispositivos especializados o elementos gráficos. Para este tipo de datos no existe un mapa o documento cartográfico, sino una serie de datos espaciales expresados de forma alfanumérica que son susceptibles de convertirse en una Capa y emplearse dentro de un SIG (ejemplo: muestras en campo) de los cuales se desconoce su escala, este es el caso de la Geocodificación (Olaya, 2011).
- 2) La distribución espacial de las muestras de la variable: a esto lo llamaremos proceso puntual (Proceso Estocástico con Probabilidades Definidas). Formalmente, las muestras del proceso puntual se denominan “Eventos del Proceso” y un patrón puntual Espacial Completamente Aleatorio (CSR, por sus siglas en inglés) se define formalmente si en su función de intensidad (que en lo sucesivo denominaremos “densidad”) tal que, los eventos se distribuyen al azar y de forma independiente sobre el área de estudio. Esto implica que no hay regiones donde los eventos son más o menos probables, y que la presencia de un evento dado no modifica la probabilidad de que aparezcan otros eventos (Bivand *et al.*, 2008). Formalmente, esto constituye la base de la hipótesis nula de los procesos puntuales. Existen derivaciones de un CSR que pueden ser agregados (agrupados) o regulares (véase figura 2), la razón de esta diferenciación es el comportamiento de la función de densidad, ya que la distribución agrupada genera múltiples escalas de trabajo y por tanto resulta de interés determinar el tipo de distribución espacial de los eventos para determinar la escala (Olaya, 2011).

Andrades, J. E. (2015): “Caracterización de los procesos espaciales y temporales y sus interrelaciones en estaciones de precipitación mensual en la zona centro occidental de la República Bolivariana de Venezuela”, *GeoFocus (Artículos)*, n° 16, p. 153-180. ISSN: 1578-5157



Figura 2. De izquierda a derecha, patrones de puntos agregado, aleatorio (CSR) y regular: (Olaya, 2011).

Diversos autores (Lloyd, 2010; Olaya, 2011) mencionan que las propiedades de primer orden del proceso puntual influyen sobre la AE y AET (propiedades de segundo orden) ya que si el patrón de los puntos es agregado, se puede estar recogiendo información redundante (sobremuestreo) en la zonas donde se ubican los conglomerados e información faltante (submuestreo) donde no se recolectó la información, en comparación con la distribución CSR o regular. El efecto de la distribución espacial de los datos, puede ocasionar un fenómeno de *Heterogeneidad Espacial o Espacio Temporal* que se refiere a la variación de las relaciones entre las variables en el espacio tiempo (varianzas no constantes en el espacio tiempo) (LeSage, 1999).

En este trabajo se exploran las propiedades de primer orden para poder establecer la existencia e intensidad del proceso de segundo orden en datos espacio temporales de precipitación mensual en el Centro Occidente de la República Bolivariana de Venezuela.

2. Antecedentes

Venezuela es un país tropical con extensión marítima y continental que posee una gran riqueza hídrica con climas muy diversos y comportamientos variados, influidos por la Zona de Confluencia Intertropical (ZCIT). Un ejemplo se encuentra en Silva (2010) quien explica que *el centro poblado más septentrional es Pueblo Nuevo, que está situado en la península de Paraguaná, Estado Falcón, á 11° 57' N y 69° 55' O; mientras que el más meridional es San Carlos de Río Negro, que está ubicado en el Estado Amazonas á 1° 55' N y 67° 03' O. Ambas son capitales municipales pequeñas y calurosas pero llueve diez veces más en la localidad amazónica que en la falconiana.* Pese a su condición tropical el clima de Venezuela no es uniforme, debido a la Zona de Confluencia Intertropical (ZCIT) y al control climático que ejerce el relieve de los sistemas montañosos andino, costero, coriano y guayanés, generando rangos amplios de temperatura y precipitación media anual que son indicativos de variados tipos climáticos venezolanos. La precipitación se caracteriza por un régimen unimodal principalmente en el centro y este del territorio y bimodal (con picos en Mayo - Junio y en Septiembre - Noviembre, con un mínimo en Julio - Agosto) al noroeste del territorio (Martelo, 2003; Silva, 2010 y Pérez y Soto, 2004).

En este punto surge una serie de preguntas en torno a las propiedades de primer orden (Escala y Distribución Espacial de los Eventos) y de segundo orden (AET) y su influencia sobre la precipitación mensual que de acuerdo con Silva (2010), no tienen respuestas de consenso o definitivas, tales como: *¿Qué es un mes seco y qué es un mes lluvioso? ¿Cómo definir estaciones secas y lluviosas y cuántas puede haber en una localidad?* Una de las primeras consideraciones que nosotros debemos tomar en cuenta para responder estas preguntas es que el clima está sujeto a fenómenos de transición, es decir, que entre una zona de alta concentración

Andrades, J. E. (2015): "Caracterización de los procesos espaciales y temporales y sus interrelaciones en estaciones de precipitación mensual en la zona centro occidental de la República Bolivariana de Venezuela", *GeoFocus (Artículos)*, n° 16, p. 153-180. ISSN: 1578-5157

de pluviosidad (A) y una zona de baja concentración de pluviosidad (B) ocurre un fenómeno de disminución paulatina de A en dirección a B y viceversa. Para definir cuáles estaciones de precipitación son de zona de alta concentración pluviométrica, baja concentración pluviométrica y cuáles son sometidas a un fenómeno de transición, se debe categorizar los datos de precipitación mensual a través de una comparación entre las muestras con respecto a sus vecinas en una localización espacio temporal definida como (x, y, t) y establecer cómo estas se relacionan entre sí.

En Venezuela se han hecho esfuerzos para caracterizar espacio temporalmente la precipitación. Autores como Olivares *et al.* (2012) utilizaron técnicas de Análisis de Componentes Principales (ACP) para grupos de estaciones con precipitaciones homogéneas en el Estado Anzoátegui, a partir de la precipitación mensual de las 40 estaciones climáticas ubicadas en el estado, para el período 1970-2000. El ACP generó cinco grupos coherentes de precipitación constituidos por estaciones meteorológicas con un patrón de precipitación específico. Cortez *et al.* (2011) caracterizaron espacio-temporalmente la precipitación a partir de datos promedio mensuales y anuales del período 1971-2000 de las estaciones meteorológicas del Estado Guárico y sus alrededores. A las series de registro se les realizó un Análisis Exploratorio de Datos (AED) y se efectuaron análisis de correlación de la precipitación con longitud, latitud y altura. Para el análisis espacial se utilizó la interpolación con el método de Kriging Ordinario. Toledo y Hernández-Szczurek (2005) estudiaron la variabilidad temporal y espacial de las lluvias anuales de la Depresión de Carora, Estado Lara, determinando que las precipitaciones son erráticas y la evaporación alta, lo que causa un desbalance hídrico marcado durante todo el año, por ello requirieron el análisis de la distribución de la precipitación y verificación de la periodicidad o ciclos estadísticamente significativos con registros cronológicos de precipitaciones anuales y mensuales, en el lapso 1965-1996 a través de series temporales. Hernández (2008) estudió la vulnerabilidad a la sequía meteorológica de las zonas secas de los Llanos Centro Orientales para la contribución al diseño de estrategias sobre el manejo integral del recurso hídrico, utilizando el Índice de Precipitación Estandarizada (SPI), que usa como única variable la lluvia, y que permitió identificar y describir convenientemente el fenómeno, precisando su inicio, intensidad de la lluvia en términos de mm/tiempo, duración, magnitud y cobertura espacial; se determinó además el impacto que varios eventos de sequía produjeron sobre los recursos hídricos existentes en la zona, a través de los niveles de los embalses y aguas subterráneas. Otro aporte obtenido fue la estimación de los umbrales críticos mensuales de precipitación que indican sequía, los cuales serán base y plataforma para el establecimiento futuro de Sistemas de Alerta Temprana, útiles para detectar inicio de sequías y para el manejo de los recursos hídricos. Todos estos valiosos estudios no llegan a incluir la AET ni a determinar la escala de los datos utilizados.

Los indicadores de AE y AET (I de Moran y su variante LISA (Local Indicators of Spatial Association)) son técnicas que han sido ampliamente usadas en los datos espaciales, de hecho diversos autores tales como Alcaldía Mayor de Bogotá D.C. (2012); Fu *et al.* (2014); Othman *et al.* (2011); Toral (2001) y Huo *et al.* (2012), las han utilizado en trabajos de análisis de precios de la vivienda, análisis de suelos, materia orgánica, caracterización del crecimiento económico y análisis de lluvia, respectivamente. Las propiedades de primer orden han sido ampliamente utilizadas en ecología del paisaje por diversos autores (Khanhasani *et al.*, 2013 y Fisher *et al.*, 2007) de hecho Dixon (2012) la describe como un análisis que puede ser utilizado para resumir un patrón de puntos, probar hipótesis sobre el patrón y estimar parámetros y modelos de ajuste.

Andrades, J. E. (2015): “Caracterización de los procesos espaciales y temporales y sus interrelaciones en estaciones de precipitación mensual en la zona centro occidental de la República Bolivariana de Venezuela”, *GeoFocus (Artículos)*, n° 16, p. 153-180. ISSN: 1578-5157

Para poder llegar eventualmente a la caracterización de un modelo que incluya la AET, se debe realizar el estudio exploratorio previo que es el objetivo del presente trabajo. Para ello, a partir de los datos de precipitación mensual se obtuvieron: los estadísticos descriptivos básicos, las propiedades de primer orden (distribución y escala), las propiedades de segundo orden (AET) y la interrelación entre las propiedades de primer y segundo orden. Todo esto en la zona Centro Occidental de la República Bolivariana de Venezuela.

3. Materiales, datos y métodos

3.1 Delimitación espacial

Este trabajo se realizó en la zona Centro Occidental de la República Bolivariana de Venezuela, en 11 estados (Zulia, Mérida, Táchira, Barinas, Portuguesa, Lara, Falcón, Cojedes, Yaracuy, Trujillo y Carabobo) para un área de estudio de un poco más de 200.000 Km², (véase figura 3), que se encuentran ubicados la zona centro occidental, andina y de los llanos occidentales de Venezuela en los que se manifiestan diversos tipos climáticos característicos y diferenciados debido al control que ejerce el relieve, y que concentra gran cantidad de la población y actividades productivas de Venezuela.

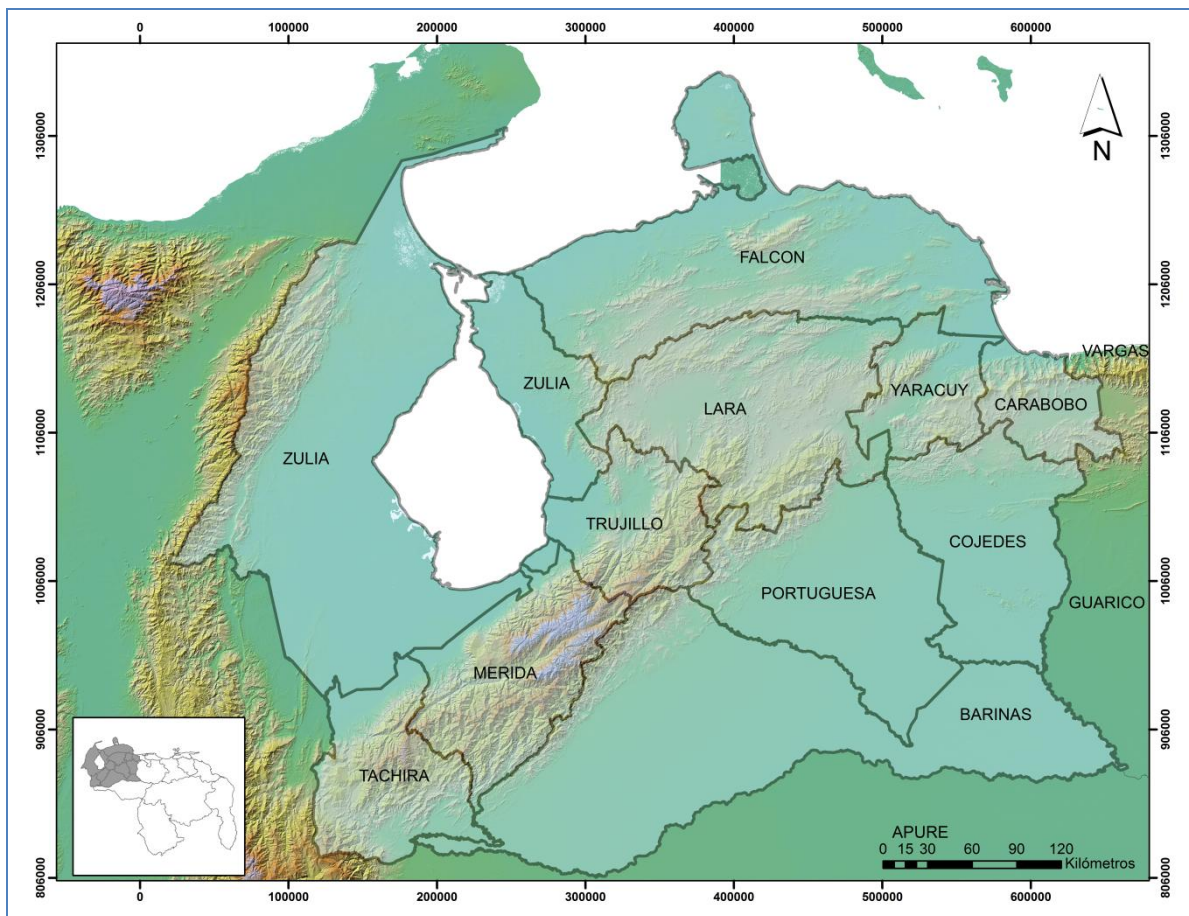


Figura 3. Estados de trabajo seleccionados.

Andrades, J. E. (2015): “Caracterización de los procesos espaciales y temporales y sus interrelaciones en estaciones de precipitación mensual en la zona centro occidental de la República Bolivariana de Venezuela”, *GeoFocus (Artículos)*, n° 16, p. 153-180. ISSN: 1578-5157

3.2 Fuentes

En Venezuela el Instituto Nacional de Meteorología e Hidrología (INAMEH) dispone a nivel nacional, de 2471 estaciones y el portal oficial del Instituto Nacional de Investigaciones Agrícolas (INIA) tiene 20 estaciones adicionales, lo que da un total de 2491 estaciones con data oficial de tipo Geocodificación. Utilizando los datos oficiales de precipitación mensual disponibles en estos portales oficiales se trabajó con un total de 961 estaciones de precipitación mensual (véase figura 4). La data según Gujarati (2004) corresponde estadísticamente con una estructura de datos agrupados (series de tiempo con datos transversales en el espacio). Se excluyeron en el análisis los datos faltantes y englobados con lo que se dispuso de un total 290105 datos correspondientes a los 11 estados de Venezuela.

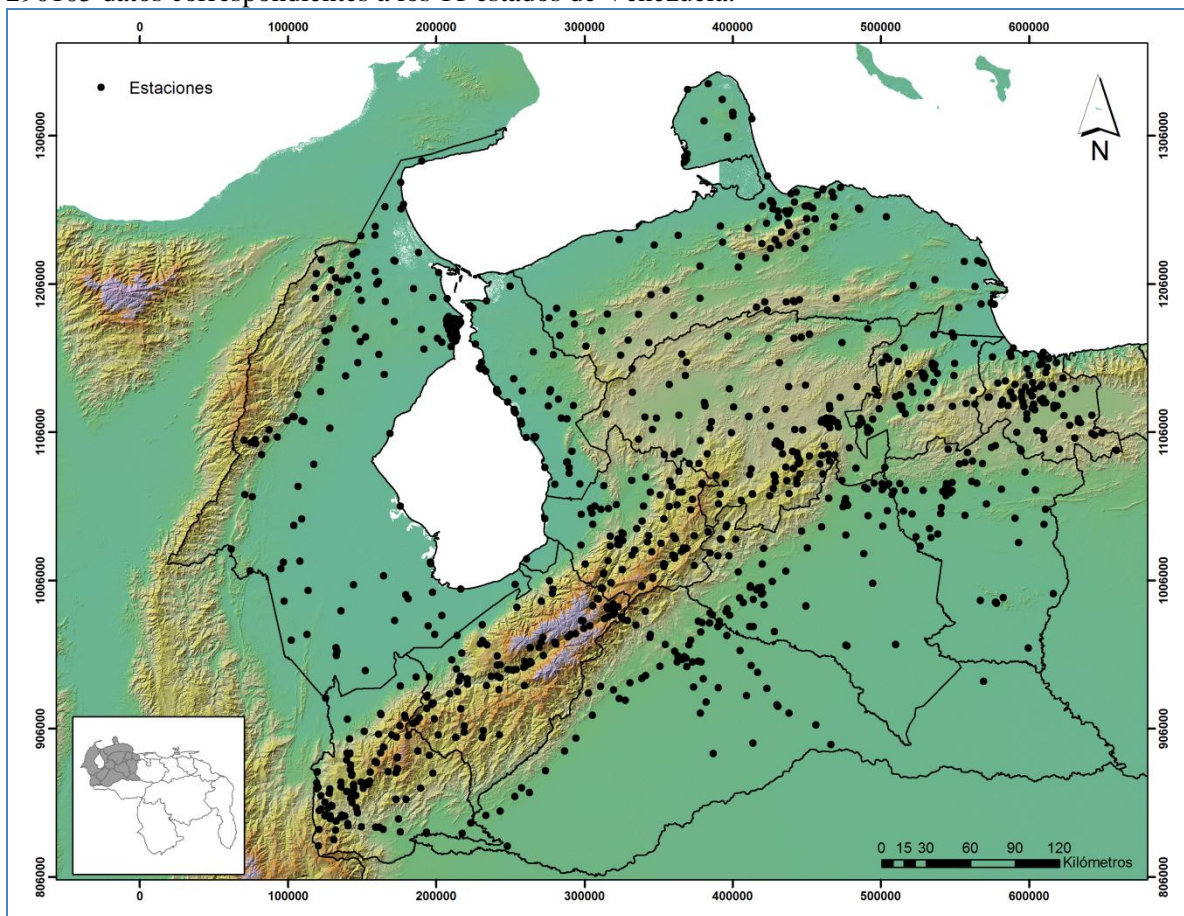


Figura 4. Estaciones de precipitación mensual seleccionadas.

3.3 Metodología

Se estimaron los estadísticos descriptivos básicos a la totalidad de los datos, sin separación espacio temporal, (Mínimo, 1^{er} Cuartil, Mediana, Media, Máxima, Desviación Estándar, Curtosis y Coeficiente de Asimetría) y posteriormente, puesto que los índices de AE y AET más frecuentemente usados asumen el supuesto de normalidad de los datos (es óptimo

Andrades, J. E. (2015): "Caracterización de los procesos espaciales y temporales y sus interrelaciones en estaciones de precipitación mensual en la zona centro occidental de la República Bolivariana de Venezuela", *GeoFocus (Artículos)*, n° 16, p. 153-180. ISSN: 1578-5157

cuando se cumple esta condición), se aplicó el test de normalidad de Shapiro-Wilk el cual es sugerido por Mohd y Bee (2011) y por Gomorov y Gienko (2011) ya que lo definen como el más poderoso de los diversos tests de normalidad. Se trabajó con un nivel de significancia de 5%. Ya que existía la limitante de que el número de datos superaba la máxima capacidad del test, se realizó un muestreo aleatorio de 5000 datos a los que se les ejecutó el mismo, estos se complementaron con técnicas gráficas (aplicadas a la totalidad de los datos) con el fin de evaluar la distribución normal (histograma de los datos, el diagrama de cajas y el diagrama QQ de distribución normal, sugeridos por Figueras y Gallardo, (2003); Chasco (2003); Chasco (2009) y Rocco (2012)).

Para caracterizar las propiedades de primer orden se definieron las siguientes hipótesis:

- a) H_0 : El proceso puntual espacial subyacente en las estaciones de precipitación mensual es de tipo CSR.
- b) H_1 : El proceso puntual espacial subyacente en las estaciones de precipitación mensual es distinto del tipo CSR.

De rechazar la hipótesis nula se debe evaluar qué tipo de derivación del CSR se está manifestando (agregados o regulares), bajo el supuesto de que las estaciones no poseen datos discontinuos en el espacio tiempo.

Para comprobar estas hipótesis se tomó la recomendación de Lloyd (2010) que propone el uso de técnicas como el análisis del vecino más cercano (se usó un nivel de significancia del 5%) y la función K de Ripley. El primero es una prueba de hipótesis que caracteriza el tipo de proceso puntual en base a la hipótesis nula de CSR espacial. Fue desarrollado por Clark y Evans (1954). Olaya (2011) y Lloyd (2010) describen este análisis como el índice de vecino más cercano, que se calcula como:

$$I_{mc} = \frac{\overline{d_{mc}}}{E(\overline{d_{mc}})}$$

donde $\overline{d_{mc}}$ la media de las distancias $\overline{d_{mc}} = \frac{\sum_{i=1}^n d_{mc}}{N}$ y $E(\overline{d_{mc}})$ es la media esperada en

una distribución CRS $\hat{\mu} = \frac{1}{2\sqrt{\lambda}}$, siendo λ la densidad de puntos N por unidad de área A , es decir, $\lambda = \frac{N}{A}$.

El segundo trata de incorporar la escala como un factor más de análisis, siendo su estimador:

$$\hat{K}(d) = \frac{|A|}{n^2} \sum_{i=1}^n \#(C(x_i, d))$$

donde, $\#(C(x_i, d))$ que corresponde con el número de eventos ($\#$), en un círculo $C(x_i, d)$, cuyo centro es x_i con radio d , y $|A|$ corresponde con el área de la región de estudio (Dixon, 2012; Lloyd, 2010). Esta función fue desarrollada por Ripley (1977). Han sugerido diversas variaciones de la Función K de Ripley entre la que se destaca el estimador

Andrades, J. E. (2015): “Caracterización de los procesos espaciales y temporales y sus interrelaciones en estaciones de precipitación mensual en la zona centro occidental de la República Bolivariana de Venezuela”, *GeoFocus (Artículos)*, n° 16, p. 153-180. ISSN: 1578-5157

$$L(d) = \sqrt{\frac{K(d)}{\pi}} - d$$

donde valores por encima de 0 indican un alto grado de agrupación (Clustering), valores menores que 0 indican dispersión y los valores cercanos a 0 indican aleatoriedad espacial (CSR) (Lloyd, 2010). Adicionalmente, en este test se pueden simular intervalos de confianza (véase figura 5).

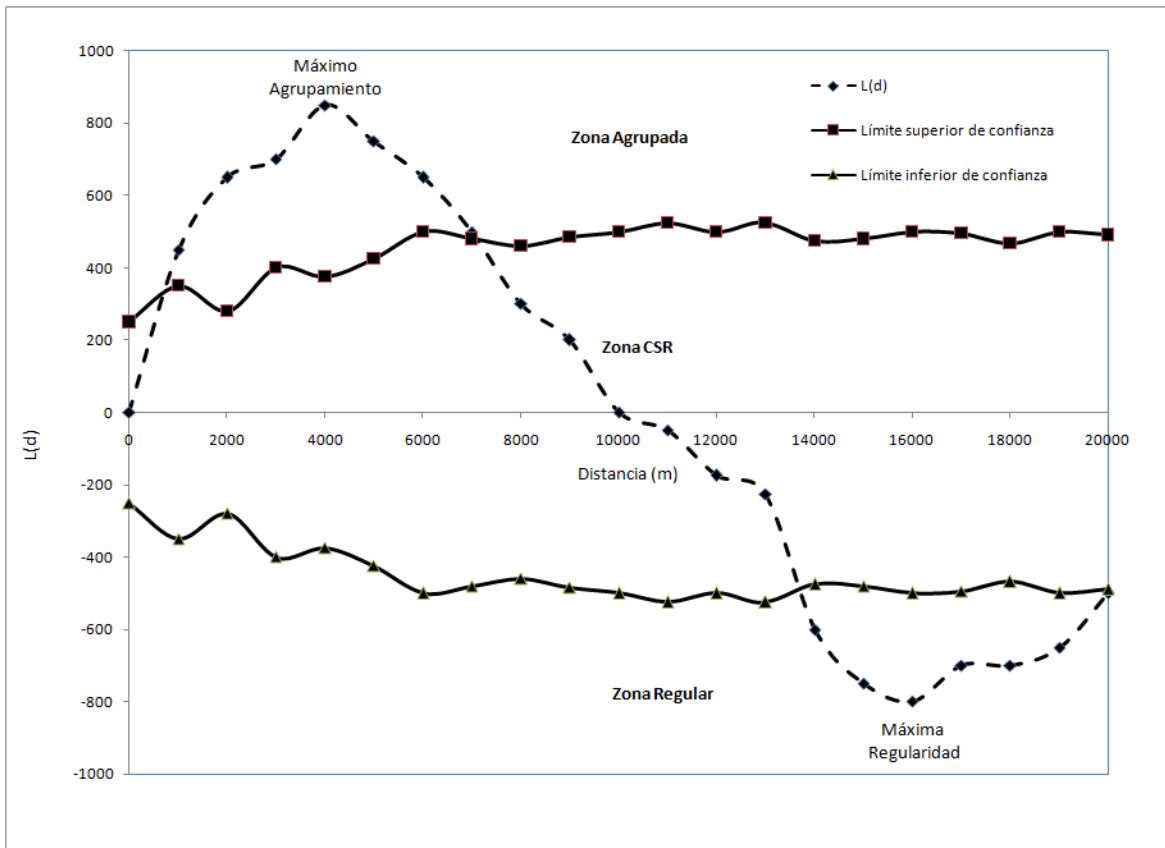


Figura 5. Interpretación del comportamiento del parámetro $L(d)$.

Sin embargo, para evitar la distorsión de la función de densidad causada por el efecto de borde, se utilizó la función K de Ripley a través de su parámetro $L(d)$, en la versión que corrige los efectos de borde:

$$L(d) = \sqrt{\frac{A \sum_{i=1}^N \sum_{j=1, j \neq i}^N k_{ij}}{\pi n(n-1)}}$$

donde d es la distancia, n es el número de muestras, A representa el área total de las muestras y se incluye el valor k_{ij} que pondera los distintos puntos en función de su distancia al

Andrades, J. E. (2015): "Caracterización de los procesos espaciales y temporales y sus interrelaciones en estaciones de precipitación mensual en la zona centro occidental de la República Bolivariana de Venezuela", *GeoFocus (Artículos)*, n° 16, p. 153-180. ISSN: 1578-5157

borde de la zona de estudio (11 estados). Con el fin de representar cartográficamente la densidad, se procedió a calcular una capa de densidad espacial de las estaciones utilizando el núcleo gaussiano, y cuya expresión es:

$$k(h) = \left(1 - \frac{h^2}{r^2}\right)^2 \quad \forall h \leq r$$

donde h es la distancia al punto y r el radio máximo de influencia (Olaya, 2011).

Para caracterizar las propiedades de segundo orden (AET) se definieron las siguientes hipótesis:

- a) H_0 : La precipitación mensual poseen AET lineal nula.
- b) H_1 : La precipitación mensual poseen AET distinta a lineal nula.

De rechazar la hipótesis nula cabe evaluar si la AET es de tipo positiva o negativa. Para ello se procedió a medir su intensidad utilizando técnicas Lattice, (análisis espacial de datos discretos), éstos se separan en dos grupos: medidas globales y medidas locales. Se utilizó la Medida Global desarrollada por Moran (1950), que ha sido ampliamente descrita por diversos autores (Bosque, 1992; Cressie, 1993; Lloyd, 2010; Olaya, 2011; Toral, 2001) y que introduce una variante del coeficiente de correlación lineal de Pearson, denominada I de Moran y está definido como:

$$I = \frac{N}{S_0} \frac{\sum_{i=1}^N \sum_{j=1}^N w_{ij} (x_i - \mu)(x_j - \mu)}{\sum_{i=1}^N (x_i - \mu)^2}$$

donde μ es la media de la variable x , $w_{ij} = d_{ij}^{-k}$ es la matriz de ponderación espacial, d_{ij} corresponde a la distancia entre las muestras (Euclidiana o Manhattan) elevado inversamente proporcional a la potencia k (existen otros métodos para estimar la matriz de ponderación espacial, los cuales están descritos por Getis y Aldstadt (2004)), S_0 es un factor de normalización igual a la suma de todos los elementos de la matriz de ponderación espacial

$S_0 = \sum_{i=1}^N \sum_{j=1}^N w_{ij}$ y N es el número de muestras tomadas, este indicador es funcional tanto para la AE como la AET. El valor esperado de I en el supuesto de AE y AET nula es:

$$E(I) = \frac{-1}{N-1}$$

Valores por debajo de este valor esperado indican autocorrelación negativa, mientras que los situados por encima reflejan autocorrelación positiva. Este indicador es útil ya que da una caracterización más global que otros índices que son sensibles a las variaciones locales a distancia reducida (Olaya, 2011). El I de Moran supone que la correlación es lineal en el espacio o espacio tiempo.

Las medidas locales, corresponden con variantes de las medidas globales y tienen como objeto caracterizar la AE y AET para cada uno de los datos, en ellos se definen "Clúster" o

Andrades, J. E. (2015): “Caracterización de los procesos espaciales y temporales y sus interrelaciones en estaciones de precipitación mensual en la zona centro occidental de la República Bolivariana de Venezuela”, *GeoFocus (Artículos)*, n° 16, p. 153-180. ISSN: 1578-5157

conglomerados de datos en los que los valores están por encima o por debajo de los valores que tienen una distribución con AE y AET nula; entre las medidas locales se usó el indicador local de asociación espacial (LISA por sus siglas en inglés), desarrollado por Anselin (1995) y descrito en Anselin (2005), que estima los valores de la siguiente manera:

$$I_i = \frac{x_i - \mu}{S_i^2} \sum_{j=1, j \neq i}^n [w_{i,j}(x_j - \mu)]; S_i^2 = \frac{\sum_{j=1, j \neq i}^n w_{i,j} [(x_j - \mu)^2]}{n-1} - \mu$$

Las Medidas Globales suelen ser un buen criterio para probar la existencia de AE o AET. Sin embargo, proporcionan poca información acerca de la ubicación de los Clúster. Por lo tanto, el uso de las Medidas Globales debe preceder al uso de las Medidas Locales (Lu, 2000). Se utilizó la versión espacio temporal del I de Moran y la variante LISA con un nivel de significancia del 5%, se generó una matriz de ponderación utilizando los siguientes parámetros: Space-Time Window estandarizada, relación del inverso de la distancia, sin distancia umbral y una ventana de 24 meses, ya que los modelos estocásticos paramétricos de series de tiempo ARMA y ARIMA por sus siglas en inglés, recomiendan para datos estacionales mensuales diferenciaciones de hasta 24 meses (Cryer y Chan, 2008; Gilgen, 2006 y Pruscha, 2013).

La prueba LISA puede arrojar 5 tipos de Clúster de acuerdo a su tipo de AET: 1) Valores altos rodeados de valores altos (HH), equivale a AET positiva de precipitación mensual alta; 2) Valores bajos rodeados de valores bajos (LL), equivale a AET positiva de precipitación mensual baja; 3) Valores altos rodeados de valores bajos (HL), equivale a AET negativa que se describe como zonas de transición espacio-temporal, de bolsones de precipitación mensual alta rodeados de precipitación mensual baja; 4) Valores bajos rodeados de valores altos (LH), equivale a AET negativa con zonas de transición espacio-temporal de bolsones de precipitación mensual baja rodeados de precipitación mensual alta y 5) Valores con AET nula (NS).

Puesto que la precipitación mensual posee una fuerte estacionalidad la cual fue descrita previamente, todas las estaciones cuentan con combinación de los cinco Clúster, es decir, que en una zona de alta precipitación espacio temporalmente hablando, no llueva en una época determinada pero van a ser más dominantes los períodos de lluvias HH que los de LL y de igual manera sucede para las zonas secas pero en este caso dominarán los LL y no los HH y así sucesivamente. Es por ello que para organizar las estaciones en grupos homogéneos en función de “alguna similitud” entre ellos, se realizó un análisis de Agrupación de las K Medias, la cual es una técnica multivariante que permite categorizar un grupo de variables en función del parecido o similitud existente entre ellos, asumiendo la distancia euclídea en el espacio, de un atributo definido (Giménez, 2010). Previamente se estimó el número óptimo de categorías utilizando el método “Elbow” sugerido por Matthew (2011) y desarrollado por Ketchen y Shook (1996). Así se generaron 5 grupos de estaciones de precipitación Alta, Media-Alta, Media, Media-Baja y Baja.

Para cuantificar el grado de afectación que ejercen las propiedades de primer orden sobre las propiedades de segundo orden, se plantearon las siguientes hipótesis:

- a) H_0 : La densidad de las estaciones de precipitación mensual poseen AE lineal nula.
- b) H_1 : La densidad de las estaciones de precipitación mensual poseen AE distinta a la lineal nula.

De rechazar la hipótesis nula cabrá evaluar el tipo de AE, positiva o negativa. Para ello se procedió a medir la intensidad sobre la densidad espacial de la AE utilizando el I de Moran (se

Andrades, J. E. (2015): “Caracterización de los procesos espaciales y temporales y sus interrelaciones en estaciones de precipitación mensual en la zona centro occidental de la República Bolivariana de Venezuela”, *GeoFocus (Artículos)*, n° 16, p. 153-180. ISSN: 1578-5157

utilizó un nivel de significancia 5%), su variante LISA de categorización de los Clúster en su versión espacial y la medición de la autocorrelación espacial autoincremental que permite identificar parcialmente la escala de los datos a partir de la distribución espacial de la intensidad.

Esto se complementó con un ANOVA donde se compararon las medias de las densidades de los grupos de estaciones de precipitación (Alta, Media-Alta, Media, Media-Baja y Baja) y posteriormente se realizó la prueba Post Hoc de Tukey (Garza *et. al.*, 2013), para identificar cuáles de las medias de las densidades de los Grupos, son iguales o diferentes.

4. Resultados

Los resultados de los estadísticos descriptivos (véase tabla 1) y los gráficos exploratorios de la distribución normal (véase figura 6) muestran una clara distribución no normal, esto se corrobora con el test de normalidad de Shapiro-Wilk que dio un resultado de $W = 0,8475$, $p\text{-value} < 2,2e-16$, por lo cual se rechaza la hipótesis nula de que los datos de precipitación mensual poseen normalidad, según esta figura éstos tienen la apariencia de una distribución asimétrica. Esta es muy común en los datos climáticos y es producto de que existe un límite físico sobre los datos restringidos a ser no-negativos (Bidegain y Díaz, 2011). Esto es refrendado por Moran (1969) quien menciona que en muchos experimentos de lluvia la distribución normal no se aplica, y la experiencia frecuentemente muestra que las distribuciones están lejos de ser normal, y por lo general sus densidades de probabilidad son monótonamente decrecientes.

Tabla 1. Estadísticos descriptivos de la precipitación media mensual

Parámetro	Valor
Mínimo (mm)	0,00
1 ^{er} Quartil (mm)	20,70
Mediana (mm)	70,20
Media (mm)	98,34
3 ^{er} Quartil (mm)	146,00
Máxima (mm)	1108,00
Desviación Estándar (mm)	99,66
Curtósis	3,86
Coefficiente de Asimetría	1,63

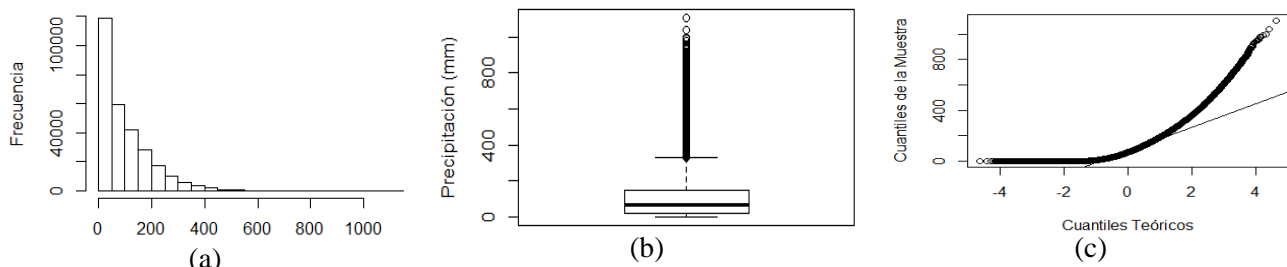


Figura 6. (a) Histograma; (b) Diagrama de Caja y (c) Diagrama QQ de Distribución Normal de precipitación mensual.

Andrades, J. E. (2015): “Caracterización de los procesos espaciales y temporales y sus interrelaciones en estaciones de precipitación mensual en la zona centro occidental de la República Bolivariana de Venezuela”, *GeoFocus (Artículos)*, n° 16, p. 153-180. ISSN: 1578-5157

Al realizar el análisis del vecino más cercano se rechaza la hipótesis nula de CSR espacial, caracterizándose el proceso puntual como agrupado, ya que el valor de la puntuación de z es menor que la media en casi 20 unidades de la desviación estándar (véase tabla 2), esto es corroborado por los resultados obtenidos por el parámetro $L(d)$ (véase figura 7). La curva de la figura es monótona y creciente a medida que aumenta la distancia, no posee ningún punto asintótico (de poseer una asíntota estaría definiendo la escala de los datos, de no poseerla implicaría que los datos pueden poseer múltiples escalas), y posee dos picos suaves, uno cercano a los 40000 m y otro cercano a los 100000 m lo que corrobora las múltiples escalas.

Tabla 2. Resumen del análisis del vecino más cercano de las estaciones de precipitación mensual.

Parámetro	Valor
Distancia media observada (m)	5711,71
Distancia media esperada (m)	8413,25
Relación de vecino más cercano	0,67
Puntuación z	19,40
p-value	0,00

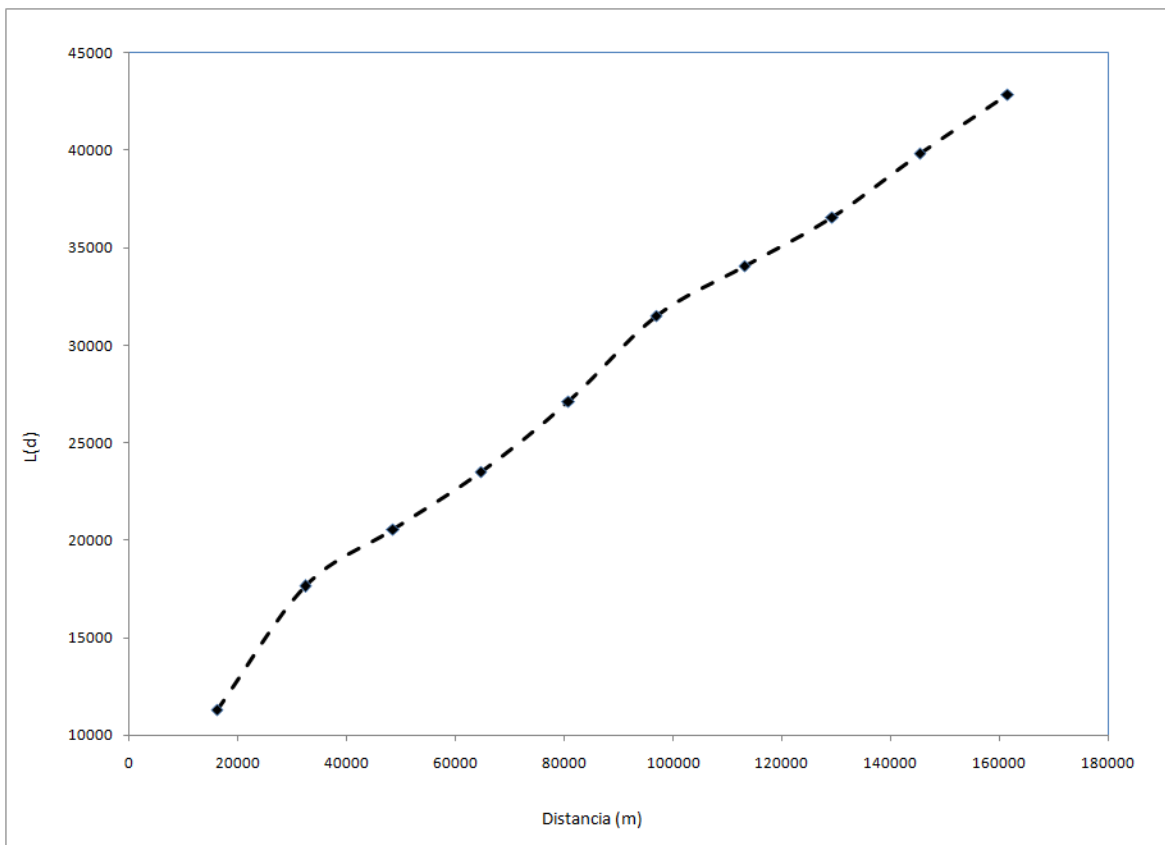


Figura 7. Parámetro $L(d)$ de las estaciones de precipitación mensual.

Al representar cartográficamente la densidad (véase figura 8) se observa que aproximadamente la mitad del área de estudio posee una densidad menor a 0,0024 Estaciones/km², los resultados están categorizados en 5 clases según los intervalos naturales de

Andrades, J. E. (2015): “Caracterización de los procesos espaciales y temporales y sus interrelaciones en estaciones de precipitación mensual en la zona centro occidental de la República Bolivariana de Venezuela”, *GeoFocus (Artículos)*, n° 16, p. 153-180. ISSN: 1578-5157

Jenks (1967), nótese que las zonas con mayor densidad se encuentran en dirección SO-NE y en la zona aledaña a la ciudad de Maracaibo, en la Costa Oriental del Lago de Maracaibo y en la Zona Norte del Estado Falcón.

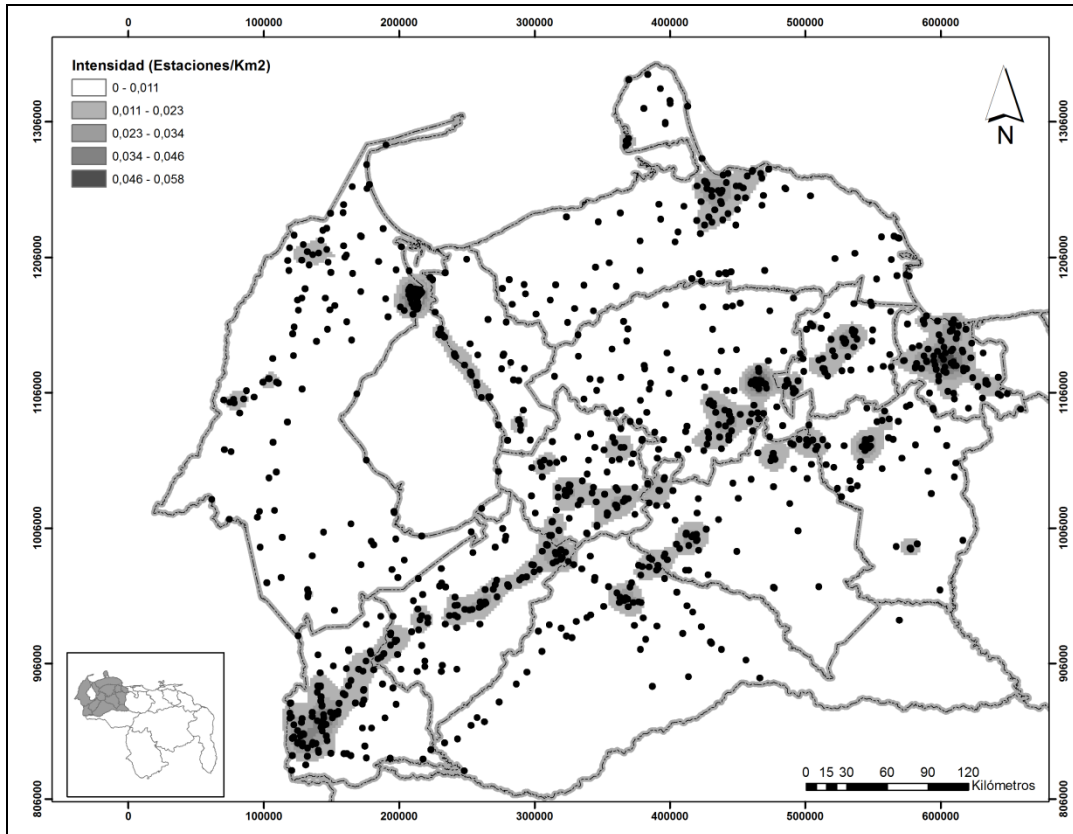


Figura 8. Densidad de las estaciones de precipitación mensual.

Los resultados rechazan la hipótesis nula de aleatoriedad espacio temporal y caracterizan el proceso de segundo orden como de AET positiva y lineal (véase tabla 3). El I de Moran es una variante del coeficiente de correlación lineal de Pearson, sin embargo, el comportamiento no normal de los datos tiene influencia sobre los resultados obtenidos por el I de Moran, ya que éstos no son óptimos, lo que implica que la AET puede ser diferente de lo que se ha estimado. Hay que mencionar que Griffith (2011) argumenta que en presencia de autocorrelación espacial positiva, la interpretación del histograma puede ser distorsionada tanto para distribuciones simétricas como la normal ó asimétricas como la Poisson o semejante. Según los criterios planteados por Li *et al.* (2007) la interpretación del I de Moran debe tomarse con cautela, pues el valor verdadero de la dependencia espacial puede estar subestimando cuando el proceso estocástico subyacente es un modelo SAR (Secuencial Autocorrelation Regression) con parámetros alejados de 0.

Tabla 3. Resumen del I de Moran espacio temporal de la precipitación mensual

Parámetro	Valor
Í de Moran Estimado	0,229100
Índice esperado	-0,000003
Varianza	0

Andrades, J. E. (2015): “Caracterización de los procesos espaciales y temporales y sus interrelaciones en estaciones de precipitación mensual en la zona centro occidental de la República Bolivariana de Venezuela”, *GeoFocus (Artículos)*, n° 16, p. 153-180. ISSN: 1578-5157

Puntuación z	556,53
p-value	0
Umbral de búsqueda (m)	1220,61

Al caracterizar los tipos de Clúster que se generaron se nota que la gran mayoría de los clúster tipificados son NS seguidos por los LL y HH característicos de la AET positiva y en último lugar, la AET negativa (HL y LH). La dominancia de los NS, HH y LL es de esperar dado que la AET es de tipo positiva (véase tabla 4).

Tabla 4. Frecuencia de Clúster espacio temporal

Clúster	Frecuencia (%)	Frecuencia (%) Excluyendo NS
HH	11	36
LL	14	45
HL	4	13
LH	2	6
NS	68	--

Un método para la elección del número apropiado de Clústers es comparar la suma de los cuadrados del error (SSE) para una serie de soluciones de Clúster (SSE se define como la suma de la distancia al cuadrado entre cada miembro de un Clúster y su centroide Clúster, SSE puede ser visto como una medida global de error, a medida que el número de grupos aumenta, la SSE debe disminuir porque las agrupaciones son, por definición, más pequeñas). El número apropiado de Clústers podría definirse como el punto de inflexión donde la pendiente se estabiliza, esto produce un apariencia de “codo” cuando se grafica (Matthew, 2011). Para este caso (véase figura 9) el número apropiado de Clústers es de 5 ya que en este número se alcanza un punto de inflexión definido.

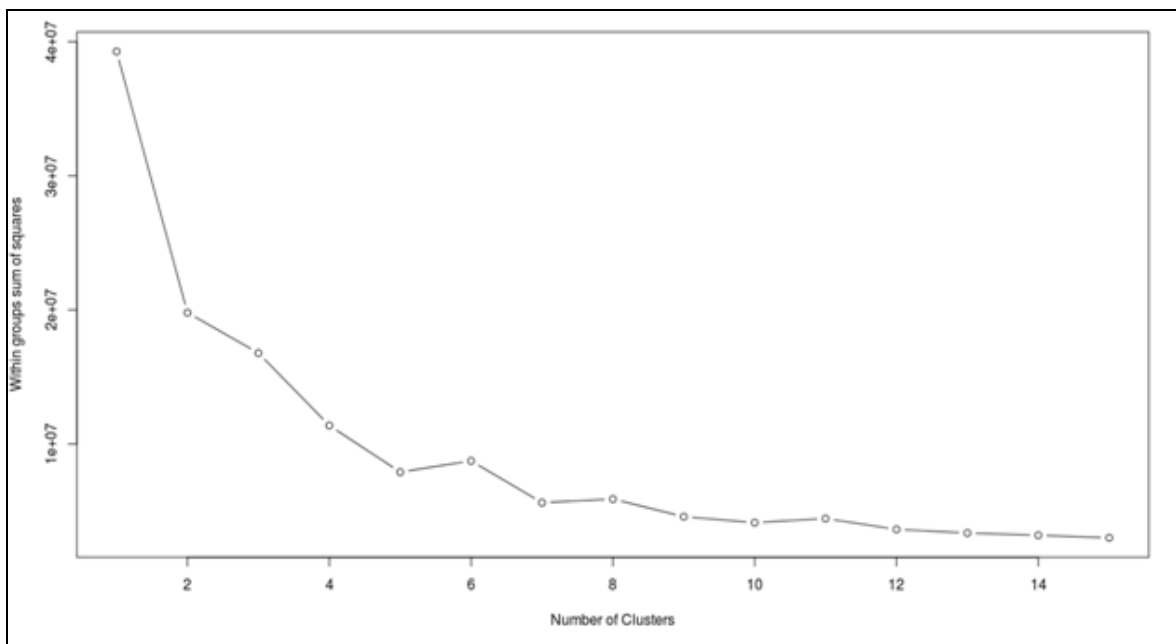


Figura 9. Variación del SSE con número de Clúster.

Andrades, J. E. (2015): “Caracterización de los procesos espaciales y temporales y sus interrelaciones en estaciones de precipitación mensual en la zona centro occidental de la República Bolivariana de Venezuela”, *GeoFocus (Artículos)*, n° 16, p. 153-180. ISSN: 1578-5157

El análisis de agrupación de las K Medias (véase figura 10) revela que los grupos se diferencian por región, y que tienen coincidencia con los mapas de precipitación generados por el Servicio de Meteorología de la Fuerza Aérea Venezolana (1955); Fundación Empresas Polar (2004) y Goldbrunner (1984). De esta manera se obtienen 82 estaciones categorizadas como de Baja precipitación (9%), 279 estaciones como de Media Baja precipitación (29%), 126 como de precipitación Media (13%), 377 estaciones como de precipitación Media Alta (39%) y 96 estaciones como de precipitación Alta (10%).

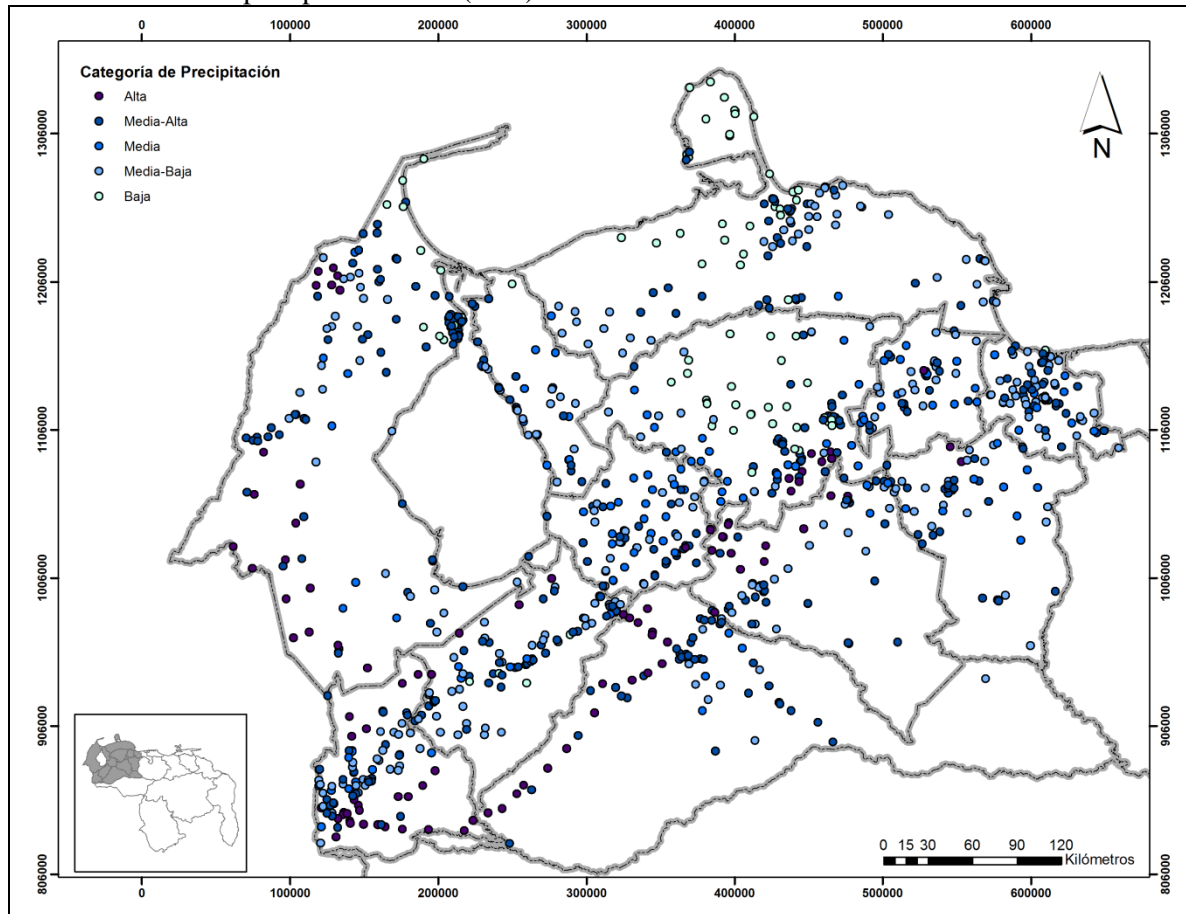


Figura 10. Categorías asignadas por el análisis de Agrupación de las K Medias a las estaciones de precipitación.

Al evaluar el comportamiento de la AE se rechaza la hipótesis nula de aleatoriedad espacial de la densidad de las estaciones de precipitación mensual (véase tabla 5). El indicador LISA manifiesta los 5 grupos definidos con datos de intensidad espacial. Los más dominantes son los valores NS (66%) que se ubican a lo largo de toda el área de estudio, los HH (19%) ubicados en la zona de la ciudad de Maracaibo, el norte del Estado Falcón, el Estado Carabobo, el suroeste del Estado Lara, el norte del Estado Cojedes, el noroeste del Estado Mérida y el sureste del Estado Táchira y los valores LL (12%) se ubican en la zona árida de los Estados Lara y Falcón, el oeste y sur del Estado Zulia y al oeste del estado Barinas, el restante 3% corresponde a valores LH ubicados en los Estados Lara, Norte de Zulia y Carabobo (véase figura 11).

Andrades, J. E. (2015): “Caracterización de los procesos espaciales y temporales y sus interrelaciones en estaciones de precipitación mensual en la zona centro occidental de la República Bolivariana de Venezuela”, *GeoFocus (Artículos)*, n° 16, p. 153-180. ISSN: 1578-5157

Tabla 5. Resumen del I de Moran espacial de la intensidad de distribución de las estaciones de precipitación mensual

Parámetro	Valor
Í de Moran Estimado	1,720340
Índice esperado	-0,001042
Varianza	0,084097
Puntuación z	5,93
p-value	0
Umbral de búsqueda (m)	43143,68

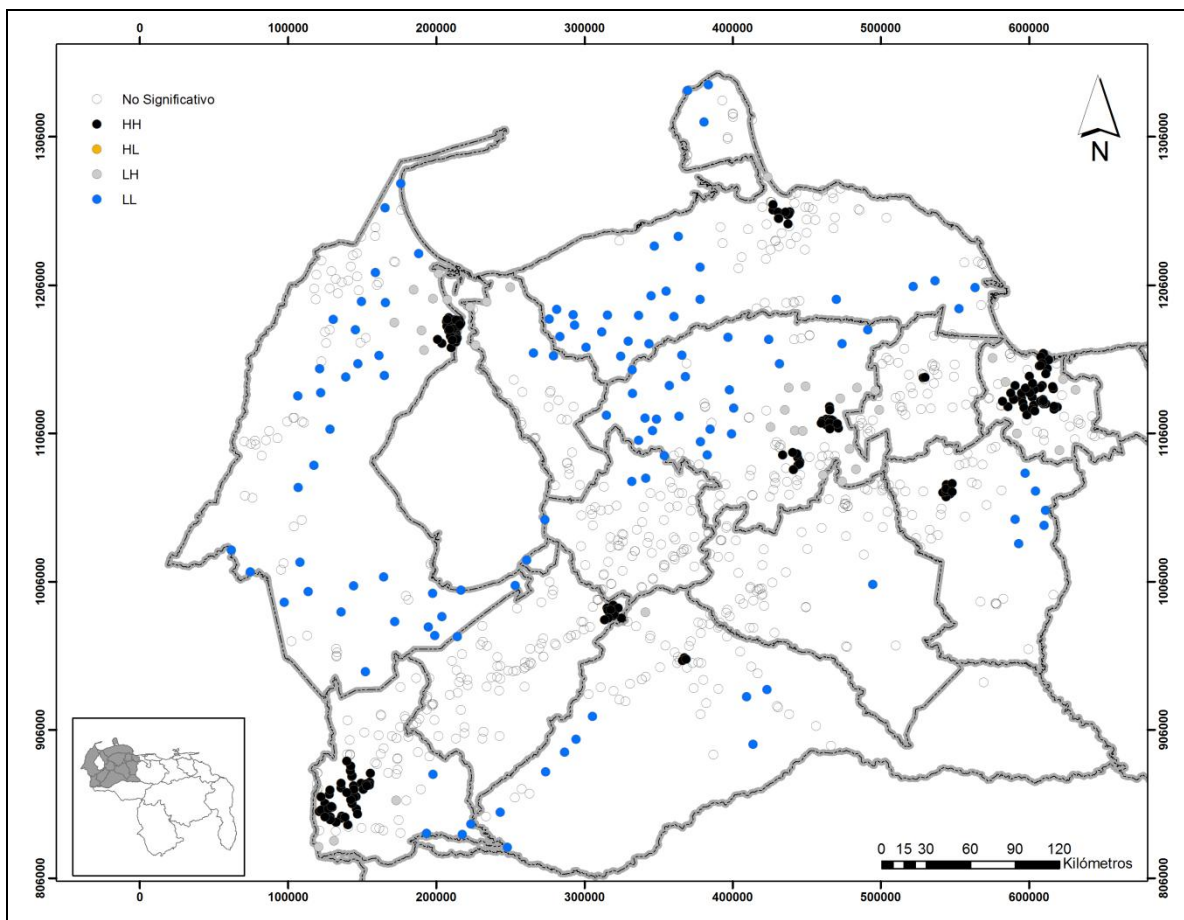


Figura 11. Clúster asignados por el indicador LISA espacial a la densidad de distribución de las estaciones de precipitación mensual.

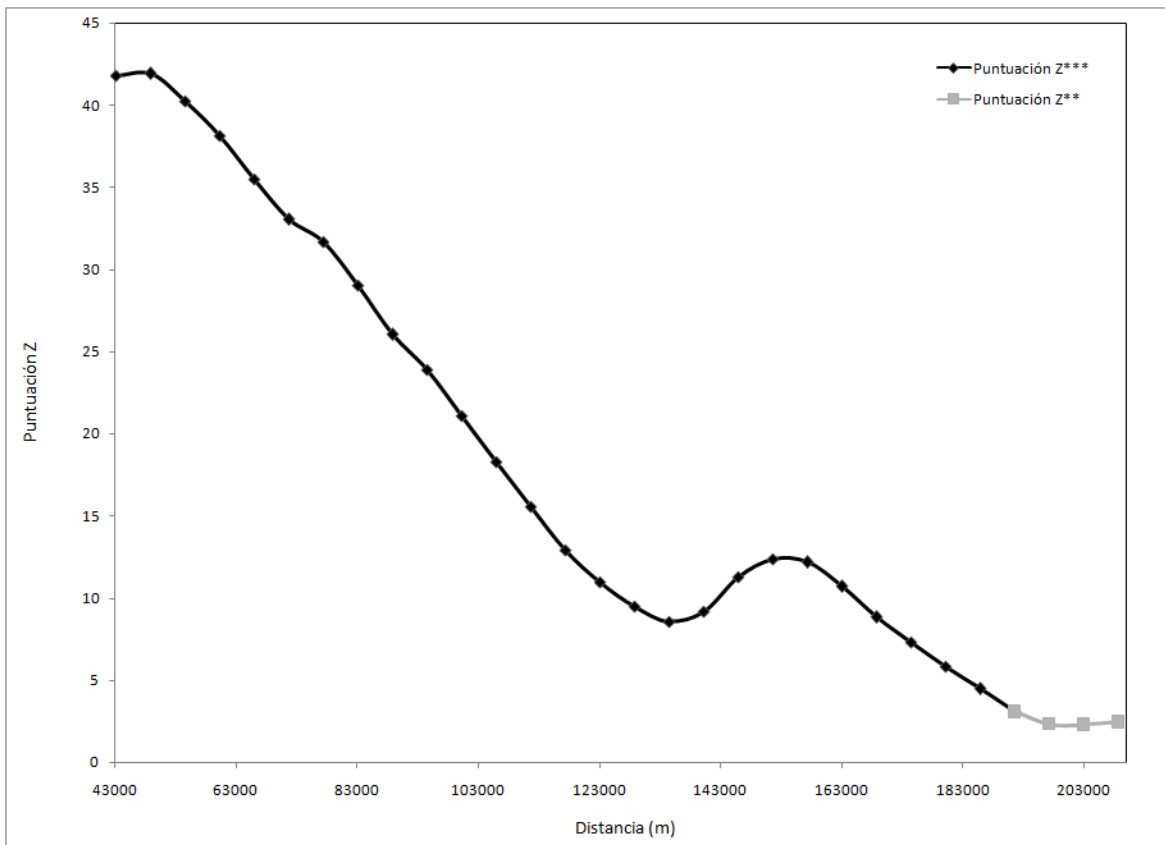
La frecuencia espacial por categoría de precipitación manifiesta que los valores más dominantes son de tipo NS y HH, siendo superiores en las categorías de precipitación Media-Alta y Media-Baja, lo que corrobora que estos grupos de estaciones generan un sobremuestreo de la precipitación mensual (véase tabla 6).

Tabla 6. Frecuencia de ocurrencia de Clúster LISA para la variable densidad por Categoría de Precipitación.

Andrades, J. E. (2015): “Caracterización de los procesos espaciales y temporales y sus interrelaciones en estaciones de precipitación mensual en la zona centro occidental de la República Bolivariana de Venezuela”, *GeoFocus (Artículos)*, n° 16, p. 153-180. ISSN: 1578-5157

Categoría de Precipitación	Clúster de densidad				
	HH	HL	LH	LL	NS
Alta	10	0	3	15	69
Media-Alta	90	0	16	22	249
Media	16	0	7	21	82
Media-Baja	40	0	2	35	202
Baja	21	0	8	18	35
Total	177	0	36	111	637

Los resultados de la AE autoincremental de la densidad de las estaciones de precipitación muestran que la AE es significativa (véase figura 12), ésta alcanza un pico en 48855 m y luego decrece hasta 134531 m donde posteriormente se encuentra un segundo pico en 151666 m, por lo que existen fuertes indicios de que los datos poseen tres escalas, una que coincide con los datos dados por el parámetro $L(d)$ la cual oscila entre 1:40000 a 1:50000, (cercano a los valores HH de intensidad espacial de las estaciones), un segundo pico que se encuentra cercano a 1:100000 y podría alcanzar hasta 1:170000 (cercano a los valores LL de intensidad espacial de las estaciones) y una tercera escala que se encuentra entre 1:50000 y 1:100000 (cercano a los valores NS de intensidad espacial de las estaciones), este fenómeno de diferencia escalar tendrá efectos significativos sobre el modelamiento de los datos ya que se traducirá en varianzas heterogéneas que serán mayores en las zonas de baja densidad de muestreo.



Andrades, J. E. (2015): "Caracterización de los procesos espaciales y temporales y sus interrelaciones en estaciones de precipitación mensual en la zona centro occidental de la República Bolivariana de Venezuela", *GeoFocus (Artículos)*, n° 16, p. 153-180. ISSN: 1578-5157

Figura 12. Valor autoincremental de la AE de la intensidad de las estaciones de precipitación.

Tabla 7. Resumen prueba ANOVA para la Intensidad por categorías de las K Medias de precipitación

	Df	Sum Sq	Mean Sq	F value	p-value
Tipo de Precipitación	4	0,00551	0,001377	10,05	5,69e-08
Residuales	956	0,13098	0,000137		

Con el ANOVA se comprobó la hipótesis de que las medias de las densidades son distintas en al menos una de las categorías de precipitación (véase tabla 7). Para identificar cuáles de las categorías tienen medias diferentes se generó la prueba de Tukey (véase tabla 8), en ella se muestra que la comparación de las medias de las intensidades de precipitación para las categorías asignadas por el análisis de las K medias, Media-Alta con Alta, Media-Alta con Media, Media-Baja con Media-Alta, son distintas. De hecho, son los Grupos Alta, Media y Media-Alta los que poseen una mayor densidad de estaciones que el resto de los grupos. El grupo Media-Baja posee la segunda mayor cantidad de estaciones pero también en promedio una de las distribuciones menos agregadas.

Tabla 8. Resumen de la prueba Tukey de comparación de las medias de las Intensidades por categoría de precipitación.

Categorías de Precipitación	Diff	Lwr	Upr	p-value
Baja/Alta	0,0043	-0,0004	0,0091	0,0957
Media/Alta	0,0015	-0,0027	0,0058	0,8656
Media-Alta/Alta	0,0068	0,0032	0,0105	0,0000
Media-Baja/Alta	0,0032	-0,0004	0,0070	0,1184
Media/Baja	-0,0028	-0,0073	0,0017	0,4385
Media-Alta/Baja	0,0025	-0,0013	0,0064	0,3977
Media-Baja/Baja	-0,0010	-0,0050	0,0029	0,9521
Media-Alta/Media	0,0053	0,0020	0,0086	0,0001
Media-Baja/Media	0,0017	-0,0016	0,0051	0,6298
Media-Baja/Media-Alta	-0,0035	-0,0060	-0,0010	0,0011

De los datos categorizados para cada grupo, un total de 37169 (13%) se categorizaron como de precipitación Baja, 99577 (34%) Media-Baja, 72347 (25%) Media, 46583 (16%) Media-Alta y 34429 (12%) como Alta, esto posiblemente explica parcialmente el comportamiento no normal de los datos (véase tabla 9) ya que no existe una proporción similar en la cantidad de datos por categoría de precipitación ni una distribución espacial homogénea. Esto es refrendado por los histogramas de cada una de las categorías, donde todas las categorías poseen una distribución asimétrica, nótese cómo la frecuencia de valores 0 disminuye en la medida en que aumenta la categoría de precipitación (véase figura 13).

Andrades, J. E. (2015): “Caracterización de los procesos espaciales y temporales y sus interrelaciones en estaciones de precipitación mensual en la zona centro occidental de la República Bolivariana de Venezuela”, *GeoFocus (Artículos)*, n° 16, p. 153-180. ISSN: 1578-5157

Tabla 9. Test de normalidad de Shapiro-Wilk de las categorías de las estaciones de precipitación

Tipo de Precipitación	W	p-value
Baja	0,75	< 2,2e-16
Media-Baja	0,89	< 2,2e-16
Media	0,90	< 2,2e-16
Media-Alta	0,84	< 2,2e-16
Alta	0,93	< 2,2e-16

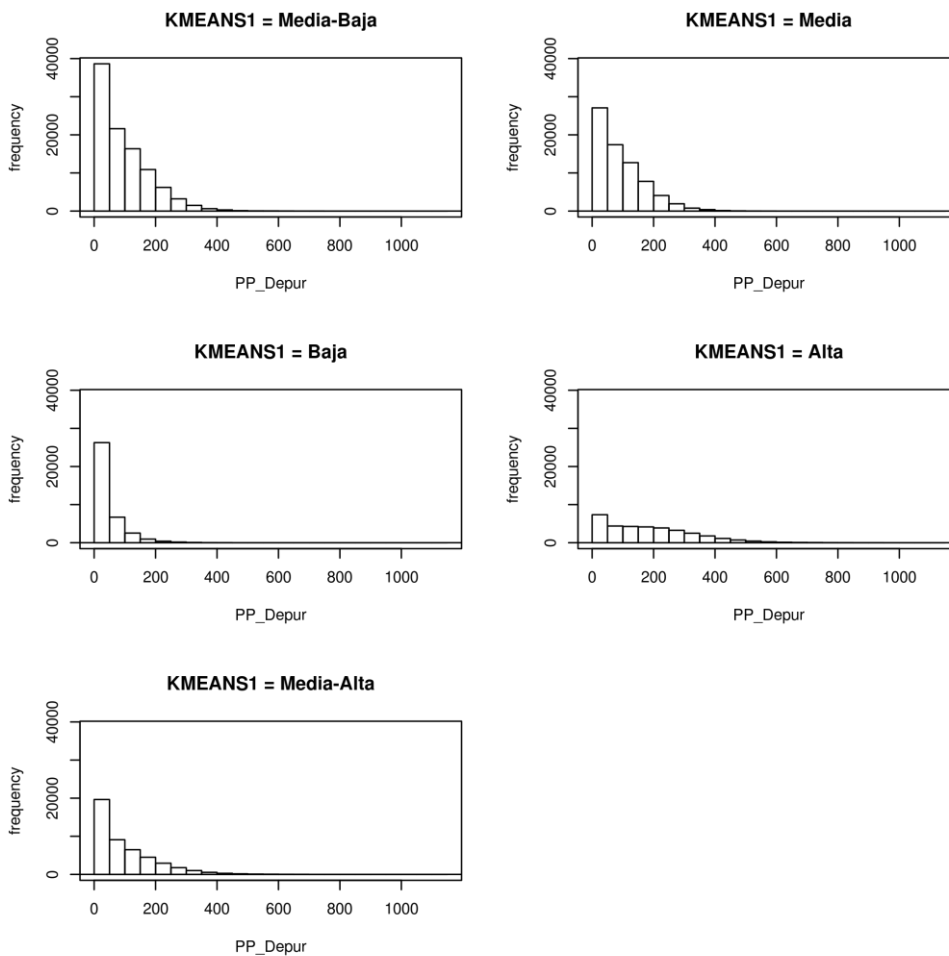


Figura 13. Histogramas por categoría de estaciones de precipitación mensual.

5. Discusión

En este trabajo se utilizó una gran cantidad de datos, si bien esto es una fortaleza para el modelaje espacio temporal, también presenta la debilidad de que no se ha hecho un análisis de un control de calidad de los datos, pudiendo haberse eliminado falsos ceros y valores fuera de rango.

Se prueba un proceso puntual de tipo agregado en las estaciones de precipitación, lo cual tiene una influencia sobre cualquier modelamiento de los datos ya que se está sobreestimando

Andrades, J. E. (2015): "Caracterización de los procesos espaciales y temporales y sus interrelaciones en estaciones de precipitación mensual en la zona centro occidental de la República Bolivariana de Venezuela", *GeoFocus (Artículos)*, n° 16, p. 153-180. ISSN: 1578-5157

información en ciertas zonas y subestimando en otras, es decir, debido a la estructura agregada de las estaciones, es probable que exista una estructura de Heterogeneidad Espacio Temporal. Hasta este momento, la escala de los datos de precipitación es desconocida, esto debido a que la naturaleza de la data es discontinua en el tiempo y que la data es producto de una Geocodificación. Para un proceso de selección futuro, se debe considerar esta situación. Bajo el supuesto de que las estaciones no poseen datos discontinuos en el espacio tiempo, se podría considerar que existen por lo menos tres escalas entre las estaciones, una que oscila entre 1:40000 a 1:50000; otra que oscila entre 1:100000 y 1:170000 y una tercera que oscila entre 1:50000 y 1:100000, sin embargo esto debe ser evaluado de forma más profunda en estudios posteriores con los datos discontinuos.

Se logra determinar que los datos de precipitación mensual poseen AET positiva y ésta es de carácter lineal. En un futuro modelamiento de los datos se debe considerar esta propiedad en el procesamiento de los datos climáticos, ya que si se obvia, se traducirá en sesgos estadísticos y peores ajustes. Debido a que se comprobó la existencia de AET en la precipitación mensual, previamente a cualquier modelado se debe evaluar si los datos son espacio temporalmente separables, es decir, si se puede modelar el componente espacial independiente del temporal o no, de no evaluar esta situación y siendo no separables estas dimensiones, se estaría incluyendo sesgos y/o sobreprocesamiento en el modelado de los datos, ya que el mecanismo estocástico subyacente es de mucha mayor complejidad.

El uso de los Indicadores de intensidad de AET es funcional ya que no requiere que el usuario conozca climáticamente a profundidad el área de estudio, incluso la zona puede contener distintos tipos de clima, pues al caracterizar la AET a través del I de Moran y su variante LISA, con un gran volumen de datos, la gran continuidad espacio temporal de la de la variable (en este caso precipitación mensual) se va a manifestar de forma automática.

Según los criterios planteados por Li *et al.* (2007), la interpretación del I de Moran deben tomarse con cautela pues el valor verdadero de la dependencia espacial puede estarse subestimando cuando el modelo correcto es SAR con parámetros alejados de 0, proceso que en este momento es desconocido, por lo que establecer una medida de la fuerza de dependencia espacial no es posible exclusivamente utilizando el I de Moran.

Debe estudiarse el comportamiento de la AET para posteriormente caracterizar la distribución de frecuencia, ya que según los estudios previos de Moran (1969) y Griffith (2011) es posible que la distribución de los datos esté distorsionada por la AET. Debe considerarse a futuro evaluar el comportamiento de la AET ante una transformación de los datos, con el fin de alcanzar un histograma más aproximado a una distribución normal.

Se logró a través del análisis de Agrupación de las K Medias categorizar las estaciones en cinco categorías Alta, Media-Alta, Media, Media-Baja y Baja según la frecuencia de los clúster HH, LL, HL, LH y NS. En un futuro modelamiento de los datos es posible que deban separarse las estaciones según sus grupos de precipitación mensual ya que en caso contrario podrían violarse los supuestos del modelo, en ese sentido se deben evaluar las implicaciones que generan el número de estaciones sobre los supuestos y optimización del modelo.

Es posible que si se aplican otros métodos de clasificación como los de redes neuronales, máximos verosímiles, clasificadores "suaves", entre otros, puede ser que se logre categorizar de mejor manera las estaciones y su precipitación. Sin embargo, se debe considerar la Ley Geográfica de Tobler (1970) en el proceso de caracterización y simulación de los datos.

Andrades, J. E. (2015): “Caracterización de los procesos espaciales y temporales y sus interrelaciones en estaciones de precipitación mensual en la zona centro occidental de la República Bolivariana de Venezuela”, *GeoFocus (Artículos)*, n° 16, p. 153-180. ISSN: 1578-5157

Se prueba la existencia de diferencias significativas entre la densidad espacial de los grupos de estaciones de precipitación, con ello se cuantifica el grado de afectación que las propiedades de primer orden ejercen sobre las de segundo orden. En un tentativo modelamiento espacio temporal de los datos de precipitación en el que se requiera excluir (de ser necesario) estaciones que están causando efectos de sobre muestreo, se puede pensar en descartar estaciones de los Grupos Alta, Media y Media-Alta que poseen una mayor densidad de estaciones que el resto de los grupos, en caso contrario podría intentarse trabajar con múltiples escalas en el modelamiento de los datos. Es posible que el proceso de primer orden sea uno de los factores para que la data de precipitación mensual no posea distribución normal. Dada la naturaleza no-negativa de la variable precipitación, es posible que ésta genere una distribución asimétrica en los datos analizados, o en su defecto una combinación entre la densidad espacial y la naturaleza del dominio de la variable precipitación. Alternativamente, se podría estar manifestando una mayor cantidad de categorías de estaciones de precipitación que no están siendo caracterizadas en este trabajo.

También debe evaluarse cómo influye el relieve (variable únicamente espacial) matemáticamente sobre los datos espacio temporales de precipitación mensual (variable espacio temporal) y si en esta interrelación se manifiesta correlación espacial.

En cierto sentido, la evaluación de los procesos de primer y segundo orden genera como producto, una categorización de una variable climática sin la necesidad de realizar ningún tipo de modelamiento de los datos. De hecho esta metodología no requiere trabajar con datos continuos en el espacio tiempo (data con valores *Missing* estimados) por lo que puede reemplazar o complementar los mapas climáticos clásicos en los que se necesita completar datos faltantes para realizar luego un proceso de interpolación. La caracterización de los procesos de primer y segundo orden y su interrelación respectiva es aplicable a otras variables espacio temporales.

6. Conclusiones

Estos resultados son similares a estudios previos tales como los presentados por el Servicio de Meteorología de la Fuerza Aerea Venezolana (1955); Fundación Empresas Polar (2004) y Goldbrunner (1984), pero adicionalmente se determinó la existencia de un mecanismo espacio temporal de asociación mucho más complejo de lo que se había trabajado previamente, éste mecanismo condiciona cualquier tratamiento de los datos a futuro.

Referencias bibliográficas

- Alcaldía Mayor de Bogotá D.C. (2012). *Análisis espacial y geoestadístico del comportamiento de los precios de la vivienda nueva ofertada en Bogotá (2009 –2011)*. [Consulta: 06-03-2015]. Disponible en: <http://www.sdp.gov.co/portal/page/portal/PortalSDP/InformacionTomaDecisiones/Estadisticas/Inventarios/Productividad/DICE035-AnalisisPreciosVivienda-31122011.pdf>
- Androva, N. y Boland, S. (2011). “Volume variability diagnostic for 4D datasets”. *Geophysical Research Letters*, 38, pp. 1-5.
- Anselin, L. (1995). “Local indicators of spatial association-LISA”. *Geographical Analysis*, 27, 2, pp. 93–115.

Andrades, J. E. (2015): "Caracterización de los procesos espaciales y temporales y sus interrelaciones en estaciones de precipitación mensual en la zona centro occidental de la República Bolivariana de Venezuela", *GeoFocus (Artículos)*, n° 16, p. 153-180. ISSN: 1578-5157

Anselin, L. (2005). *Exploring spatial Data with geodtm: A Workbook*. Urbana Center for Spatially Integrated Social Science.

Bidegain, M. y Diaz, A. (2011). *Análisis Estadístico de Datos Climáticos Distribuciones de Probabilidad*. [Consulta: 05-03-2015]. Disponible en: http://meteo.fisica.edu.uy/Materias/Analisis_Estadistico_de_Datos_Climaticos/teorico_AEDC/Distribuciones_Probabilidad_2011.pdf

Bivand, R., Pebesma, E. y Gómez-Rubio, V. 2008. *Applied spatial data analysis with R*. New York, Springer.

Bosque, J. (1992). *Sistemas de información geográfica*. Madrid, Ediciones Rialp, S.A.

Chasco, C. (2003). "Métodos gráficos del análisis exploratorio de datos espaciales". *Anales de Economía Aplicada*. [Consulta: 13-01-2015]. Disponible en: <http://www.asepelt.org/ficheros/File/Anales/2003%20-%20Almeria/asepeltPDF/93.PDF>

Chasco, C. (2009). "Análisis exploratorio de datos espaciales al servicio del Geomarketing III Seminario sobre Nuevas Tecnologías de la Investigación, en el Marketing y la Comunicación". [Consulta: 03-03-2015]. Disponible en: https://www.google.co.ve/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0CBwQFjAA&url=http%3A%2F%2Fwww.uam.es%2Fpersonal_pdi%2Feconomicas%2Fcoro%2Fdocencia%2Fgeomark%2FponenciaAedemo_CoroChasco.doc&ei=fer5VP_PM_iPsQS5h4GQBg&usq=AFQjCNHHeZ3lO5u2jZ_C_n2qQCvKlbw-tA&sig2=DMa_Hk9Q6QaN5cWk-JPT0A&bvm=bv.87611401,d.cWc

Clark, P. y Evans, F. (1954). "Distance to nearest neighbor as a measure of spatial relationships in populations". *Ecology*, 35, 4, pp. 445-453.

Cocchi, D. y Bruno, F. (2010). "Considering groups in the statistical modeling of spatio-temporal data". *Statistica*, 70,4, pp. 511-527.

Celemín, J. (2009). "Autocorrelación espacial e indicadores locales de asociación espacial. Importancia, estructura y aplicación". *Revista Universitaria de Geografía*, 18,1, pp. 11-31.

Cortez, A., Rodríguez, M., Rey, J., Ovalles, F., Gonzales, W. y Parra, R. (2011). "Caracterización espacio temporal de la precipitación en el estado Guárico", Venezuela. *XIX Congreso Venezolano de la Ciencia del Suelo*. [Consulta: 06-03-2015] Disponible en: http://www.sian.inia.gob.ve/repositorio/congresos/CVCS19/suelo_espacio_tiempo/SET25.pdf

Cressie, N. (1993). *Statiscs for spatial data*. New York, Wiley-Interscience.

Cressie, N. y Winkle, C. (2011). *Statistics for spatio-temporal data*. New York, Wiley.

Cryer, J. y Chan, K. (2008). *Time series analysis with applications in R*. New York, Springer.

Dixon, P. (2012). Ripley's K function. El-Shaarawi, A., Piegorisch, W. (Ed.): *Encyclopedia of Environmetrics*. 5, Chichester, Wiley, pp 1796-1803.

Fisher, J., Trulio, L., Biging, G. y Chromczak, D. (2007). "An analysis of spatial clustering and implications for wildlife management: A burrowing owl example". *Environmental Management*, 39, pp. 403-411.

Figueroas, S. y Gargallo, M. (2003). *Análisis Exploratorio de Datos*. [Consulta: 06-03-2015]. Disponible en: <http://www.5campus.com/leccion/aed>

Andrades, J. E. (2015): “Caracterización de los procesos espaciales y temporales y sus interrelaciones en estaciones de precipitación mensual en la zona centro occidental de la República Bolivariana de Venezuela”, *GeoFocus (Artículos)*, n° 16, p. 153-180. ISSN: 1578-5157

Fu, W., Jiang, P., Zhou, G. y Zhao, K. (2014). “Using Moran’s I and GIS to study the spatial pattern of forest litter carbon density in a subtropical region of southeastern China”. *Biogeosciences*, 11, 8, pp. 2401–2409.

Fundación Empresas Polar. (2004). *Circulación atmosférica y tipos de clima*.

Garza, J., Morales, B. y González, B. (2013). *Análisis estadístico multivariante*. México, McGraw-Hill.

Getis, A. y Aldstadt, J. 2004. “Constructing the spatial weights matrix using a local statistic”. *Geographical Analysis*, 36, 2, pp. 90–104.

Gilgen, H. 2006. *Univariate time series in geosciences theory and examples*. Berlin, Springer.

Giménez, Y. (2010). *Clasificación no supervisada: El método de k-medias*. Departamento de Matemática Facultad de Ciencias Exactas y Naturales Universidad de Buenos Aires. [Consulta: 15-01-2015] Disponible en: http://cms.dm.uba.ar/academico/carreras/licenciatura/tesis/2010/Gimenez_Yanina.pdf

Gujarati, D. (2004). *Econometría*. México, McGraw-Hill Interamericana.

Goldbrunner, A. (1984). *Atlas climatológico de Venezuela 1951-1970*. Servicio de Meteorología de la Fuerza Aérea Venezolana. Caracas.

Griffith, D. (2011). “Positive spatial autocorrelation, mixture distributions, and geospatial data histograms”. *Spatial Data Mining and Geographical Knowledge Services (ICSDM), 2011 IEEE International Conference on*, 1-6.

Hernández, R. (2008). *Caracterización de la sequía meteorológica en los climas Árido, Semiárido y Subhúmedo seco en los Llanos Centro Orientales de Venezuela, para el manejo de los recursos hídricos*. Centro de Investigación y Postgrado, Universidad Nacional Experimental Politécnica de la Fuerza Armada. [Consulta: 06-03-2015] Disponible en: http://www.inameh.gob.ve/documentos/Caracterizacion_sequia_Llanos_Centro_Orientales_Venezuela.pdf

Herrera, M., Paz, J. y Cid, J. (2012). *Introducción a la Econometría Espacial. Una Aplicación al Estudio de la Fecundidad en la Argentina usando R*. [Consulta: 24-08-2015]. Disponible en: <http://mpra.ub.uni-muenchen.de/41138/>

Huo, X., Li, H., Sun, D., Zhou, L. y Li, B. (2012). “Combining geostatistics with Moran’s I analysis for mapping soil heavy metals in Beijing, China”. *Int. J. Environ. Res. Public Health* 2012, 9, pp. 995-1017.

Jenks, G. (1967). “The data model concept in statistical mapping”. *International Yearbook of Cartography*, 7, pp. 186-190.

Khanhasani, M., Akhavan, R., Sagheb- Talebi, K. y Vardanyan, Z. (2013). “Spatial patterns of oak species in the Zagrosian forests of Iran”. *International Journal of Biosciences*, 3,8, pp. 66-75.

Ketchen, D. y Shook, C. (1996). “The application of cluster analysis in Strategic Management Research: An analysis and critique”. *Strategic Management Journal*, 17, 6, pp. 441–458.

LeSage, J. (1999). *The Theory and Practice of Spatial Econometrics*. Disponible en: <http://www.spatial-econometrics.com/html/wbook.pdf>

Andrades, J. E. (2015): “Caracterización de los procesos espaciales y temporales y sus interrelaciones en estaciones de precipitación mensual en la zona centro occidental de la República Bolivariana de Venezuela”, *GeoFocus (Artículos)*, n° 16, p. 153-180. ISSN: 1578-5157

Li, H., Calder, C. y Cressie, N. (2007). “Beyond Moran’s I: Testing for Spatial Dependence Based on the Spatial Autoregressive Model”. *Geographical Analysis*, 39, pp. 357–375.

Lloyd, C. (2010). *Spatial data analysis an introduction for GIS users*. New York, Oxford University Press.

Lu, Y. (2000). “Spatial cluster analysis for point data: location quotients verses Kernel Density”. *Presentation at Information Science Summer Assembly, Portland, Oregon*. [Consulta: 06-03-2015] Disponible en: <http://dusk.geo.orst.edu/ucgis/web/oregon/papers/lu.htm>

Martelo, M. (2003). *La precipitación en Venezuela y su relación con el sistema climático. Dirección de Hidrología, Meteorología y Oceanología–Dirección General de Cuencas Hidrográficas – MARN*. [Consulta: 06-03-2015]. Disponible en: http://www.inameh.gob.ve/documentos/precipitacion_venezuela_relacion_sistema%20climatico.pdf

Matthew, P. (2011). *R Script for K-Means Cluster Analysis*. [Consulta: 06-03-2015]. Disponible en: <http://www.mattpeoples.net/kmeans.html>

Moran, P. (1950). “Notes on continuous stochastic phenomena”. *Biometrika*, 37, 1, pp. 17–23.

Moran, P. (1969). “Statistical inference with bivariate gamma distributions”. *Biometrika*, 56, 3, pp. 627-634.

Mohd, N. y Bee, Y. (2011). “Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests”. *Journal of Statistical Modeling and Analytics*, 2, 1, pp. 21-33.

Olaya, V. (2011). *Sistemas de información geográfica*. [Consulta: 06-03-2015] Disponible en: http://wiki.osgeo.org/wiki/Libro_SIG

Olivares, B. Parra, R., Guevara, E., Cortez, A. y Rodríguez, M. (2012). “Patrones de homogeneidad pluviométrica en estaciones climáticas del estado Anzoátegui, Venezuela”. *Multiciencias*, 12,1, pp. 11-17.

Othman, F., Akbari, A. y Samah, A. (2011). “Spatial rainfall analysis for an urbanized tropical river basin”. *International Journal of the Physical Sciences*. 6, 20, pp. 4861-4868.

Pérez, M. y Soto, E. (2004). “Régimen de precipitación en Montalbán estado Carabobo, Venezuela”. *Agronomía Tropical*, 54, 4, pp. 359-370.

Pruscha, H. (2013). *Statistical analysis of climate series analyzing, plotting, modeling, and predicting with R*. New York, Springer.

Reynolds, K. y Madden, L. (1988). “Analysis of epidemics using spatio-temporal autocorrelation”. *Phytopatology*, 78, 2, pp. 240-246.

Ripley, B. (1977). “Modelling spatial patterns”. *Journal of the Royal Statistical Society. Series B* 39, pp 172-192.

Rocco, C. (2012). “Análisis exploratorio de datos espaciales en estudios de demanda eléctrica”. *III Congreso Venezolano de Redes y Energía Eléctrica*. Consulta: [02-03-2015] Disponible en: http://www.cnv-cigre.org.ve/congreso_2012/temario/pdf/C1-32.pdf

Servicio de Meteorología de la Fuerza Aerea Venezolana. (1955). *Atlas climatológico provisional periodo 1951/1955*.

Andrades, J. E. (2015): “Caracterización de los procesos espaciales y temporales y sus interrelaciones en estaciones de precipitación mensual en la zona centro occidental de la República Bolivariana de Venezuela”, *GeoFocus (Artículos)*, n° 16, p. 153-180. ISSN: 1578-5157

Silva, G. (2010). *Tipos y subtipos climáticos de Venezuela*. Departamento de Geografía Física Universidad de Los Andes. Consulta: [02-03-2015]. Disponible en: http://www.saber.ula.ve/bitstream/123456789/31158/3/trabajo_ascenso_2010.pdf

Stojanova, D. (2012). *Considering autocorrelation in predictive models*. (Tesis doctoral inédita). Jožef Stefan International Postgraduate School.

Tobler, W., (1970). “A computer movie simulating urban growth in the Detroit region”. *Economic Geography*, 46, 2, pp. 234-240.

Toledo, V. y Hernández-Szczurek, D. (2005). Estudio de la variabilidad temporal y espacial de las lluvias anuales de la Depresión de Carora, estado Lara, Venezuela. Lobo, D., Luján, D. y Soto, G. (Ed.): *Evaluación de Parámetros y Procesos Hidrológicos en el Suelo Compendio de los trabajos presentados en la VII Escuela Latinoamericana de Física de Suelos Tema: Hidrología de Suelos La Serena, Chile, 2-14 de Noviembre de 2003*, París, UNESCO, pp.53-59.

Toral, M. (2001). *El factor espacial en la convergencia de las regiones de la Unión Europea: 1980-1996*. (Tesis doctoral inédita). Centro: Facultad de Ciencias Económicas y Empresariales. Universidad Pontificia Comillas de Madrid.