UNIVERSIDAD DE COSTA RICA

SISTEMA DE ESTUDIOS DE POSGRADO

**ANÁLISIS DE INTERPOLACIÓN ESPACIAL PARA DATOS DE PRECIPITACIÓN DEL ESTADO DE TEXAS**

SP-1649 Tópicos de Estadística Espacial Aplicada

Proyecto 3

Maestría Profesional en Estadística

Profesora:

Marcela Alfaro Córdoba

Estudiantes:

Melissa Cordero Diaz

Marco Otoya Chavarría

Ciudad Universitaria Rodrigo Facio, Costa Rica

2019

# RESUMEN

En el presente caso práctico se analizaron datos anuales de precipitación de 21 estaciones climáticas ubicadas en el estado de Texas, tomados de “Introducción a SIG y análisis espacial” con el fin de realizar una interpolación de la precipitación para las áreas en donde no se tiene datos. Las mediciones de precipitación no se tienen para la totalidad de la zona de Texas sino solo para los puntos donde estas ubicadas dichas estaciones climáticas. Para el desarrollo del caso se aplicaran métodos deterministas y estadísticos mediantes la aplicación de Kriging y análisis de semi-variogramas.

# INTRODUCCIÓN

Los estudios orientados a analizar la distribución espacial de las variables climáticas son de gran utilizad, principalmente para comprender el comportamiento de variables que tiene alta incidencia en la vida de los seres humanos, las actividades productivas y los cambios climáticos actuales.

La precipitación es una de las variables climáticas necesarias para comprender el comportamiento hidrológico y la capacidad de captación de una cuenca hidrológica (Santillan Gutierrez, 1980).

Conocer la precipitación también es relevante para los análisis de modelación de flujos de aguas subterráneas. Siendo un tema de mayor interés para las autoridades dada las constantes amenazas de contaminación y los incrementos de los periodos de sequía, pues estos puntos de agua son una alternativa para aumentar la oferta de agua en esta épocas (Castro, 2004) ; por lo que es de gran utilidad para evaluar los procesos de erosión, así como, para definir las condiciones climáticas actuales (López, 2014).

Dentro de las técnicas que se utilizan para analizar este tipo de fenómenos están los métodos de interpolación, dentro de los cuales se tienen dos categorías: los métodos de interpolación deterministas (corte empírico tradicional) y los métodos de interpolación estadísticos. Dentro de los métodos deterministas se puede mencionar técnicas de proximidad ó Thiessen (vecino más cercano o diagramas de Voronói), red de triángulos irregulares (TIN), funciones polinomiales (Spline) o de regresión lineal, redes neuronales y técnicas de distancia inversa ponderada (IDW para abreviar). Finalmente en los métodos de enfoque estadísticos de interpolación se incluyen la tendencia de superficie y Kriging (tipo geo estadístico) (Velásquez, 2017).

El presente trabajo analiza datos anuales de precipitación de 21 estaciones climáticas ubicadas en el estado de Texas, tomados de “Introducción a SIG y análisis espacial”, con el fin de realizar una interpolación una interpolación de la precipitación donde no se tiene datos, aplicando algunas de las técnicas antes mencionados.

## OBJETIVO

Analizar la precipitación anual reportada por 21 estaciones climáticas ubicadas en el estado de Texas, con el fin de realizar una interpolación de la precipitación a las áreas en donde no se tiene datos, mediante la aplicación de los métodos de interpolación espacial determinísticos y probabilísticos.

# METODOLOGIA

## DATOS

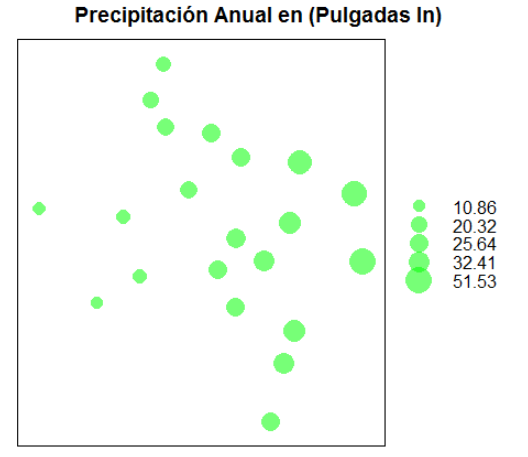
Se cuenta con los datos reportados de precipitación de 21 estaciones climáticas para el estado de Texas, utilizados en el analisis de “Introducción a SIG y análisis espacial” de Manuel Gimond, edición de septiembre de 2019. Los datos pueden descargarse directamente a los siguientes links [Datos de precipitación](http://colby.edu/~mgimond/Spatial/Data/precip.rds) y [Datos del mapa](http://colby.edu/~mgimond/Spatial/Data/texas.rds).

Los datos cuentan con las respectivas coordenadas espaciales y los promedios anuales de precipitación reportados por cada una de las estaciones climáticas. Los datos se encuentran disponibles en formato de puntos espaciales.

## ANALISIS EXPLORATORIO

Para el analisis exploratorio de los datos se inicia explorando el formato de los datos y aplicando un resumen descriptivo para validar que se encuentran en formato de puntos espaciales y que contengan el dato de precipitación por estación. Adicionalmente, se grafican los datos para identificar cuales estaciones reportan mayores niveles de precipitación, tal como se muestra en el siguiente gráfico.

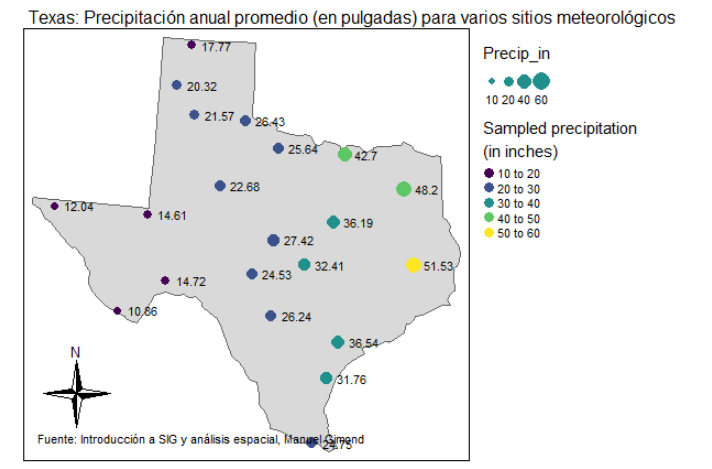
**Gráfico No.1 Precipitación anual medida en pulgadas por estación climática, Texas.**



Fuente: Elaboración propia a partir de los datos de Introducción a SIG y análisis espacial Año: 2019.

Como puede observarse en el gráfico 1, las estaciones climáticas con círculos más grandes reportaron mayores niveles de precipitación y están a la izquierda del gráfico. Adicionalmente, es evidente que en promedio la precipitación anual puede variar de estación en estación lo cual puede estar asociado a la ubicación de las estaciones climáticas; este hallazgo es relevante y debe considerarse al momento de realizar alguna extrapolación con este tipo de datos.

**Mapa No1. Precipitación anual promedio (en pulgadas) de acuerdo a la estaciones, Texas**



Fuente: Elaboración propia a partir de los datos de Introducción a SIG y análisis espacial Año: 2019.

Al analizar el gráfico de los datos analizados sobre el mapa de Texas se pueden identificar que las estaciones que reportan mayor precipitación anual esta ubicadas al este y noreste del Estado, con precipitaciones anuales superiores a las 40 pulgadas.

# RESULTADOS

## MÉTODOS DE INTERPOLACIÓN NO-GEOESTADÍSTICOS

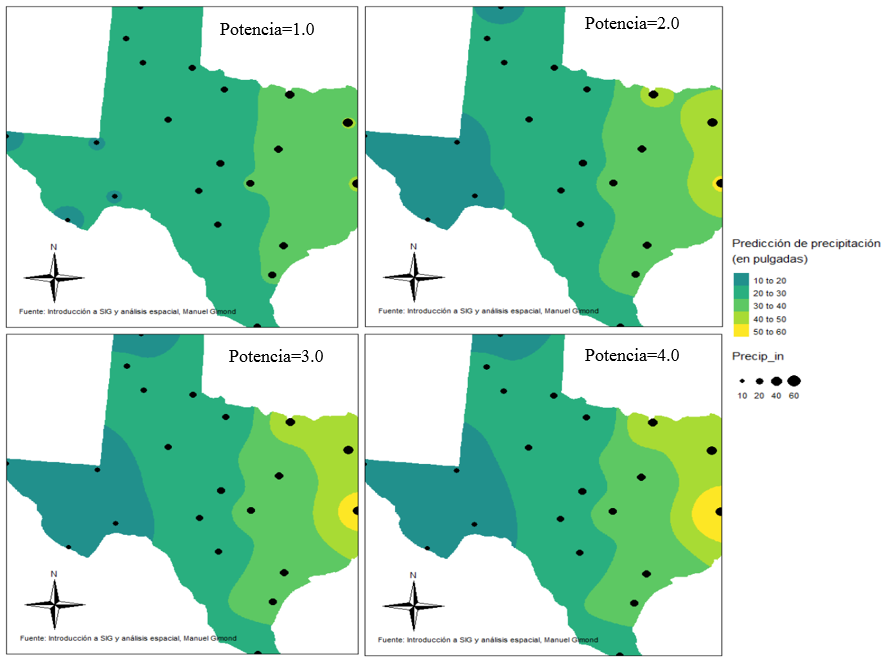
**Interpolación con el método de la distancia inversa ponderada (IDW)**

Este método calcula un promedio ponderado para las ubicaciones donde no hay estaciones identificadas que reporten datos de precipitación, para ello utiliza las ubicaciones ponderadas de las estaciones climáticas más cercanas, con lo cual se logra que se consideren los diferentes promedios de precipitaciones reportados. Los pesos de los puntos donde no hay reporte de estaciones son proporcionales a las estaciones que si reportan precipitaciones más próximas.

**Coeficiente de potencia:** Cuando mayor sea el coeficiente de potencia mayor es el peso de los puntos cercanos (Gimond, 2019).

En los siguientes gráficos se muestra la predicción de la precipitación para el estado de Texas para diferentes niveles de potencia. Con un nivel de potencia bajo (1.0), la interpolación que se puede realizar para las estaciones climáticas que reportaron menos de 20 pulgadas de precipitación y más de 40 pulgadas de precipitación es muy reducida.

**Mapa No.2 Predicción de la precipitación con el método de distancia inversa ponderada (IDW)**



Fuente: Elaboración propia a partir de los datos de Introducción a SIG y análisis espacial Año: 2019.

A medida que se aumenta el nivel de potencia la interpolación que se hace de las estaciones concentradas en estos rangos aumenta. Para efectos de este análisis, dado que interesa tener una interpolación del rango de precipitación de más de 50 pulgadas se seleccionó un nivel de potencia de 3, sin embargo, esta es una decisión que se debe tomar de la mano de un experto en el área.

## ENFOQUE ESTADÍSTICO DE INTERPOLACIÓN

Dentro de los métodos de interpolación estadísticas se puede destacar el método de Kriging. El método tiene varias formas dentro de las que destacan el ordinal, universal y simple. En el presente estudio a los datos analizados aplicó la forma ordinaria.

**Pasos de la aplicación:**

1. Eliminar la tendencia espacial.
2. Definir el modelo del variograma que mejor caracterice la autocorrelación espacial de los datos.
3. Interpolación de la superficie.
4. Unir la superficie interpolada Kriging a la superficie interpolada de tendencia (resultado final).

**Eliminación de la tendencia espacial**

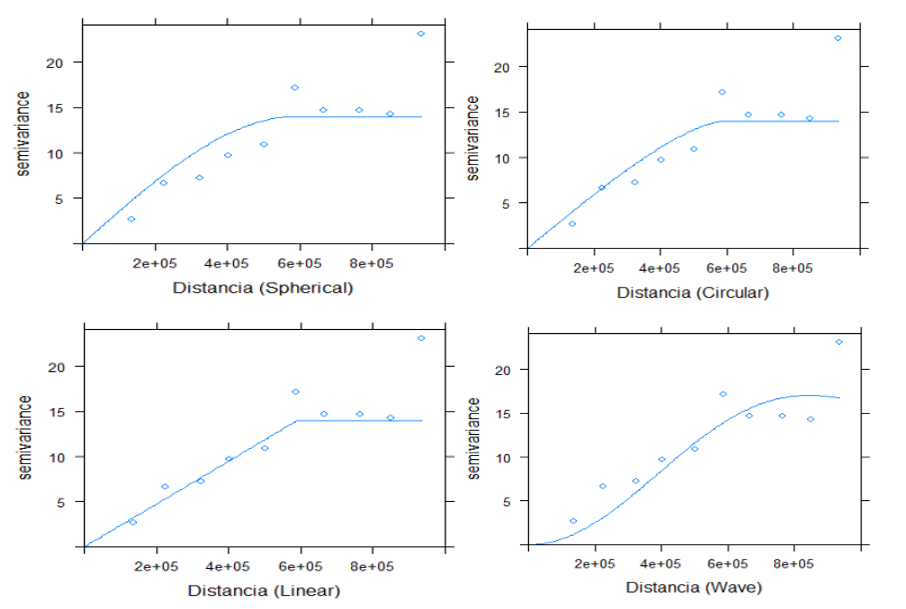
Dentro de los supuesto que deben cumplir para poder aplicar este método se destaca que la media y la variancia de los datos deben ser constante en toda el área de estudios, en los datos analizados se puede observar que las estaciones climáticas del este y noreste del Estado de Texas no tiene una media similar al resto de la estaciones climáticas, con lo que se evidencia el incumplimiento de este supuesto. Para reducir la tendencia de los datos se realiza el ajuste de primer orden (Gimond, 2019).

**Modelo del variograma que mejor caracteriza la autocorrelación espacial**

Seguidamente se debe seleccionar el variograma que mejor representa la autocorrelación espacial de los datos, para este procedimiento se ajustan diferentes modelos de variogramas de los cuales, de acuerdo con el análisis gráfico, se seleccionan aquellos que realizan una mejor representación. En el gráfico 2 se pueden visualizar los mejores modelos seleccionados.

Dentro de los modelos que se probaron los que mostraron un mejor ajuste fueron el modelo esférico, el modelo Circular, el modelo Lineal y el modelo Wave. De estos modelos se selecciona el modelo circular, dado que proporciona un resultado similar al modelo esférico con la diferencia de que se ajusta mejor a los datos iniciales. Para cada uno de los modelos se utilizó un valor de umbral de 14, un valor de nugget de 0 y un rango de 590000.

**Gráfico No.2 Modelos ajustados a variograma residual**

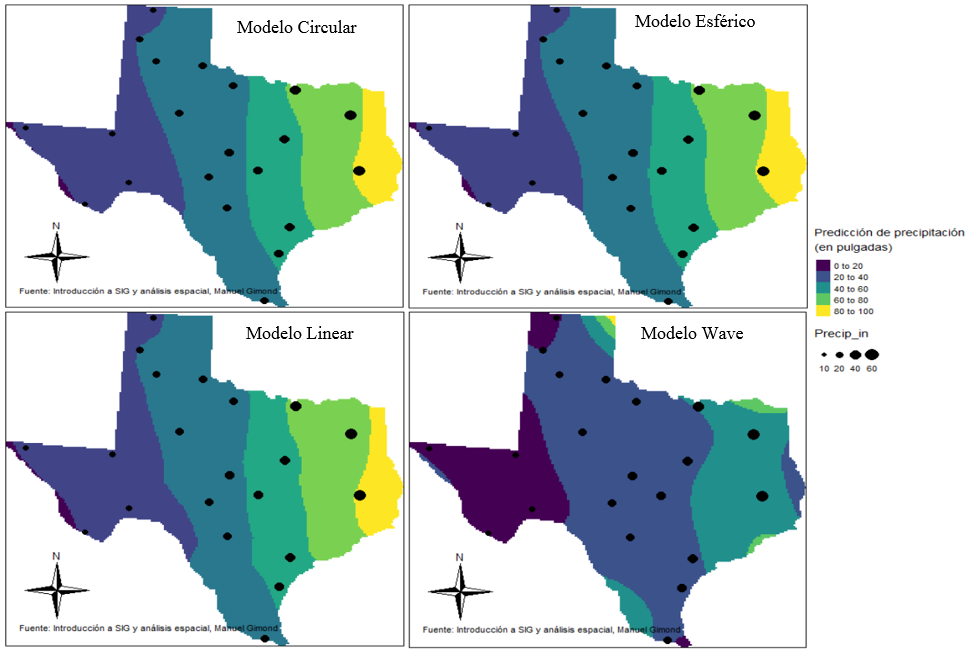


Fuente: Elaboración propia a partir de los datos de Introducción a SIG y análisis espacial Año: 2019.

**Interpolación de la superficie**

Para generar la interpolación de la superficie se utilizan los valores ajustados del modelo circular, para el analisis se utiliza la función “krige” que permite incluir el modelo de tendencia sin tener que realizar el ajuste para reducir la tendencia (Gimond, 2019). El modelo que se seleccionó como el de mejor ajuste de la autocorrelación de los datos es el modelo circular, sin embargo, se incluye la interpolación de la superficie realizada por los demás modelos para ver las diferencias en cada uno. En el siguiente mapa se observa la interpolación de la superficie para los distintos modelos.

**Mapa No.3 Superficie kriged final obtenida con los diferentes modelos ajustados**



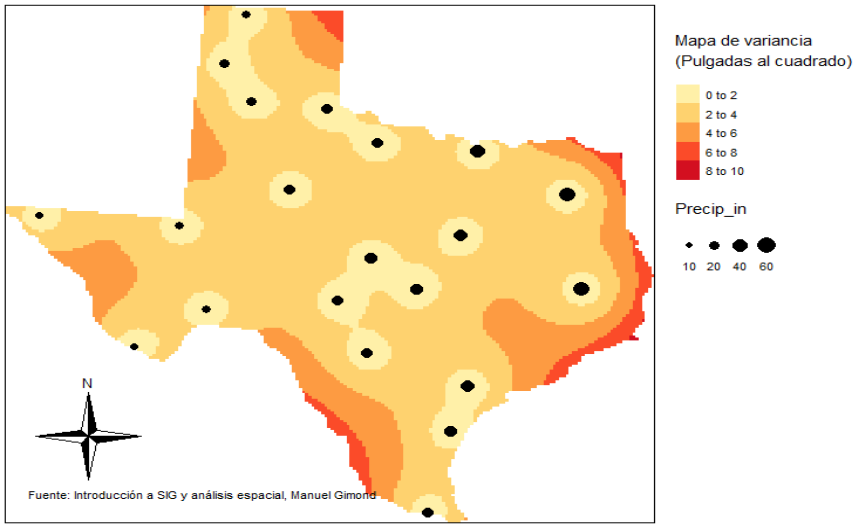
Fuente: Elaboración propia a partir de los datos de Introducción a SIG y análisis espacial Año: 2019.

Como se puede observar el modelo Circular y Esférico realizan interpolaciones de la superficie similares, lo cual no es de sorprender dado que el comportamiento de los variogramas ya mostraba comportamientos similares. El modelo lineal muestra algunas variaciones, principalmente porque amplia la interpolación de la precipitación de 20 a 40 pulgadas, esto reduce levemente la interpolación realizada para la precipitación de 40 a 60 pulgadas.

Finalmente, la interpolación que realiza el modelo Wave presentan un comportamiento distinto, en este caso destaca principalmente porque aparece una interpolación para precipitaciones de 0 a 20 pulgadas y prácticamente desaparece la interpolación para precipitaciones de más de 60 pulgadas, lo cual evidencia que no es el mejor modelo para los datos de precipitación.

Otra de las herramientas que se puede utilizar del método Kriging y que puede aportar mucha información son los mapas de variancia que brinda una medida de variación de los valores que se están utilizando para realizar la interpolación. Lo anterior, con el fin de conocer el grado de incertidumbre de los resultados, en este caso cuando menor sea la variancia de los puntos analizados más consistentes son los resultados obtenidos (Gimond, 2019). Como se puede observar en el mapa la variancia de las estaciones climáticas es baja lo cual favorece los resultados obtenidos de la interpolación.

**Mapa No.4 Variancia obtenida del análisis de Kriging para la precipitación anual promedio, Texas.**



Fuente: Elaboración propia a partir de los datos de Introducción a SIG y análisis espacial Año: 2019.

# CONCLUSIONES

* Para la precipitación, se obtienen resultados satisfactorios para la interpolación tanto con el método de la distancia inversa ponderada (IDW) como con el método de Kriging ordinal.
* En el caso del método de la distancia inversa ponderada (IDW) se debe considerar que las conclusiones de analisis cambian de forma importante ante la aplicación de diferentes niveles de potencia, por lo que es de suma relevancia contar con el criterio del experto en el área para poder asignarlos.
* Para el método de Kriging queda evidenciada la importancia de seleccionar un modelo de variograma adecuado que represente adecuadamente la autocorrelación de los datos, una mala especificación puede alterar las conclusiones del analisis.
* El mejor modelo para realizar la interpolación de la precipitación del estado de Texas es un modelo circular o de igual forma un modelo esférico dado que brinda resultados muy similares al realizar la interpolación.
* Finalmente, una de las conclusiones más relevantes y de importancia práctica es que se deben realizar analisis exhaustivos y robustos, probando distintos métodos y cambiando los parámetros de estos para entender cuál es el efecto de estos cambios sobre los resultados; lo anterior antes de brindar alguna conclusión del analisis cuando no se tenga el suficiente conocimiento del fenómeno analizado.

# REFERENCIAS

Castro, C. C. (2004). Modelado del Flujo de Agua Subterránea en la Región Nordeste de Brasil, Mediante Modelos de Simulación. *SCIELO*.

Estimación del balance hídrico mediante variables climáticas, e. l. (s.f.).

Gimond, M. (2019). *Intro to GIS and Spatial Analysis.* Last edited on 2019-08-09.

Hastie, T., Tibshirani, R., & Friedman, J. (2008). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* Stanford, California: Springer.

López, D. N. (2014). Uso de modelos de regresión para interpolar espacialmente la precipitación media mensual en la cuenca del río Conchos. *SCIELO*.

Santillan Gutierrez. (1980). Estimación del balance hídrico mediante variables climáticas, en la cuenca del río Cazones, Veracruz, México. *Revista Ambiente & Água*.

Velásquez, D. F. (2017). *Regnie interpolation for rain and temperature in the andean, caribbean and pacific regions of Colombia.* Colombia Forestal, vol. 21, núm. 1, 2018.