



COLEGIO DE POSTGRADUADOS

INSTITUCION DE ENSEÑANZA E INVESTIGACION EN CIENCIAS AGRÍCOLAS

CAMPUS MONTECILLO

POSTGRADO DE HIDROCIENCIAS

“Modelación Espacial de la Precipitación mediante Regresión Lineal Multivariada en el Norte de México”

Hugo Alberto Fuentes Hernández

TESIS

PRESENTADO COMO REQUISITO PARCIAL
PARA OBTENER EL GRADO DE:

MAESTRO EN CIENCIAS

MONTECILLO, TEXCOCO, EDO. DE MEXICO

2014

La presente tesis titulada: **Modelación Espacial de la Precipitación mediante Regresión Lineal Multivariada en el Norte de México**, realizada por el alumno: **Hugo Alberto Fuentes Hernández**, bajo la dirección del Consejo Particular indicado, ha sido aprobada por el mismo y aceptada como requisito parcial para obtener el grado de:

MAESTRO EN CIENCIAS

HIDROCIENCIAS

CONSEJO PARTICULAR

CONSEJERO



Dr. Abel Quevedo Nolasco

ASESOR



Dr. Daniel Núñez López

ASESOR



Dr. Roberto Ascencio Hernández

Montecillo, Texcoco, Estado de México, Noviembre de 2014.

DEDICATORIA

Para ti madre mía, porque me diste todo tu amor incondicional, me dedicaste tu tiempo, tu paciencia, tu sabiduría, me enseñaste que el camino es largo y que los sueños se logra con esfuerzo, dedicación y amor, me diste el ejemplo, este pequeño detalle es para ti y donde quiera que estés, siempre sabrás que mis logros y éxitos son gracias a tus grandes consejos y a esa enorme fuerza que siempre mostraste, gracias por todo, aquí sigues estando en mi corazón.

AGRADECIMIENTOS

A mi futura esposa Priscila Tenorio que amo tanto, por estar presente en todo momento, en mis proyectos profesionales y en la vida diaria, por darme todo su amor y comprensión, y a toda su familia por acogerme como un miembro más de la familia y apoyarme incondicionalmente.

A mis hermanos y sobrina por estar siempre en las buenas y en las malas, por darme su apoyo, comprensión y cariño.

Al Dr. Abel Quevedo Nolasco por todos sus consejos tan valiosos como asesor y amigo, por enseñarme uno de los grandes valores de la vida la humildad.

Al Dr. Daniel Núñez López, por brindarme sus conocimientos en el desarrollo del presente trabajo, por toda su labor como asesor y por ser un buen amigo.

Al Dr. Roberto Ascencio Hernández, por brindarme todo su apoyo incondicional, muchas cosas no se hubieran podido llevar a cabo sin su apoyo y asesoramiento.

Índice de contenido	pags.
Índice de cuadros.....	7
Índice de figuras.....	7

ÍNDICE DE CONTENIDO	PAGS.....	5
----------------------------	------------------	----------

ÍNDICE DE CUADROS	7
--------------------------------	----------

ABSTRACT	8
-----------------------	----------

RESUMEN.....	9
---------------------	----------

1. INTRODUCCIÓN.....	10
-----------------------------	-----------

2. OBJETIVO	13
--------------------------	-----------

2.1 Objetivos Específicos.....	14
---------------------------------------	-----------

3. HIPÓTESIS	15
---------------------------	-----------

4.- REVISIÓN DE LITERATURA (ANTECEDENTES).....	15
---	-----------

4.1 Modelos de Estimación de la Precipitación	15
--	-----------

4.2 Escalas de Referencia en Estudios de Modelación de la Precipitación.....	19
---	-----------

4.3. Influencia de la Topografía en la Precipitación.....	20
--	-----------

5.- MATERIALES Y METODOS.....	28
--------------------------------------	-----------

5.1 Localización Geográfica	28
--	-----------

5.2 Información Ambiental y Ecológica.....	29
5.3 Materiales	32
5.4 Metodología.....	32
5.4.1. Red Histórica Global de Climatología (GHCN).....	32
5.4.2. Descripción de variables independientes.....	34
5.4.2.1 Variables Geográficas	35
5.4.2.2 Variables del Relieve.....	35
5.4.2.3 Variables de proximidad	36
5.4.3. Construcción de Modelos de Regresión	37
5.4.4. Parámetros Estadísticos de validación	38
5.4.4. Procesos de interpolación Espacial y Calibración	41
6.- RESULTADOS	45
6.1 Proceso de escalamiento.....	45
6.2 Matriz de Correlación	46
6.3 Modelos de Regresión	49
6.4 Interpolación Espacial	51
6.5 Calibración Espacial	53
6.3.1 Pruebas de validación estadística.....	56
7.- DISCUSIÓN.....	58
8.- CONCLUSIONES.....	62
9.- RECOMENDACIONES.....	63
10.- BIBLIOGRAFÍA	64

Índice de cuadros

Cuadro 1. Clasificación de la Precipitación	46
Cuadro 2. Matriz de Correlación	48
Cuadro 3. Variables del Modelo de Regresión.....	50
Cuadro 4. Parámetros estadísticos simple y calibrados	57

Índice de figuras

Figura 1. Ubicación geográfica del área de estudio.....	28
Figura 2. Malla de puntos con series de precipitación, ubicada en el Norte de México.	34
Figura 3. Calculo de la distancia euclidiana.....	36
Figura 4. Función de Algebra de Mapas	42
Figura 5. Remuestreo y promedios focalizados en diferentes resoluciones	45
Figura 6. Precipitación calculada periodo húmedo 1980-2010 mm.	51
Figura 7. Interpolación con Spline aplicando Tensión	52
Figura 8. Errores Residuales en porcentaje.....	52
Figura 9. Calibración periodo húmedo 1980-2010.	53
Figura 10. Comparación modelo calculado vs modelo calibrado ..	54
Figura 11. Periodo Seco 1950 – 2010 mm.....	55
Figura 12. Periodo Húmedo 1950 – 2010 mm.	56

SPATIAL MODELING OF PRECIPITATION BY MULTIVARIATE LINEAR REGRESSION IN NORTHERN MÉXICO.

Fuentes Hernández Hugo Alberto M.C.

Colegio de Postgraduados 2014.

ABSTRACT

The objective of this study was the spatial modeling of precipitation through multivariate statistical models of linear regression of precipitation in the Northern part of México (in the states of Chihuahua, Durango, Sinaloa, Coahuila, Zacatecas, Nuevo León, Tamaulipas, part of de Sonora, San Luis Potosí, Aguascalientes, Nayarit y Jalisco). Where reconstructed precipitation data through the Global Historical Climatology Network (GHCN version 2) and data from the Monitoring System Climate Anomaly (CAMS) were used (Chen *et al.*, 2001). Independent variables were latitude, longitude, topographic factors, distance to the coastal areas, the standard deviation of the digital elevation models of the territory (over a period of 60 years from 1950 to 2010). A correlation matrix and analysis of statistical significance, for the purpose of identify the most influential variables were performed and were not the result of chance correlation. In the dry season, the elevation (focalized at 5 km) and latitude; for annual and wet season, distance or proximity to the coast, are the most representative for precipitation. The calibration of the models, enhanced results by 30 percent on average.

Key words: North of Mexico; precipitation; linear regression; spatial regression; geoestadística.

MODELACIÓN ESPACIAL DE LA PRECIPITACIÓN MEDIANTE REGRESIÓN LINEAL MUTIVARIADA EN EL NORTE DE MÉXICO

Fuentes Hernández Hugo Alberto M.C.

Colegio de Postgraduados 2014.

RESUMEN

El objetivo del presente trabajo fue la modelación espacial de la precipitación a través de modelos estadísticos de regresión lineal multivariada de la precipitación en la parte Norte de la República Mexicana (en los Estados de Chihuahua, Durango, Sinaloa, Coahuila, Zacatecas, Nuevo León, Tamaulipas, parte de Sonora, San Luis Potosí, Aguascalientes, Nayarit y Jalisco). Donde se utilizaron los datos de precipitación reconstruidos a través de la Red Histórica de Climatología Global (GHCN, versión 2) y los datos del Sistema de Monitoreo de Anomalías Climáticas (CAMS) (Chen *et. al.*, 2001). Como variables independientes fueron la latitud, longitud, factores topográficos, distancia a las zonas costeras, la desviación estándar de los modelos de elevación digital del territorio (en un periodo de 60 años de 1950 a 2010). Se realizó una matriz de correlación y un análisis de significancia estadística, con el fin de identificar las variables con mayor influencia y que no fuera la correlación producto del azar. En el periodo seco, la elevación (focalizado a 5 kilómetros) y la latitud; para el periodo anual y húmedo, la distancia o cercanía a las costas, son las más representativas para la precipitación. El proceso de calibración de los modelos permitió mejorar los resultados en un 30 por ciento en promedio.

Palabras clave: Norte de México; precipitación; regresión lineal; regresión espacial; geoestadística;

1. INTRODUCCIÓN

La descripción espacial y temporal de variables climáticas es de gran utilidad para comprender el funcionamiento de procesos físicos, biológicos y ambientales. La precipitación es una de las principales variables climáticas requeridas para la estimación de balances hídricos, recargas de frontera en la modelación de flujos de agua subterránea, evaluar procesos de erosión así como para definir las condiciones climáticas actuales. Sin embargo, la fiable representación espacial de la precipitación, es particularmente compleja particularmente en zonas montañosas con escasa disponibilidad de estaciones climáticas en donde el efecto orográfico es grande. En el caso de la precipitación, diversos métodos estadísticos han sido desarrollados para predecir su distribución espacial que difieren en su concepto y formulación matemática. Métodos de interpolación locales como polígonos de Thiessen e inverso de distancia ponderada (IDW) han sido de los más frecuentemente utilizados en la predicción climática y desarrollo cartográfico (Vicente-Serrano et al., 2003), al igual que procedimientos geo-estadísticos como Kriging, co-Kriging y Spline (Vicente-Serrano et al., 2003; Apaydın et al., 2004;). Diversos estudios comparan la eficiencia de procedimientos geoestadísticos y de métodos locales de interpolación espacial, con el fin de determinar el método más apropiado para representar de forma espacial la precipitación (VicenteSerrano et al., 2003; Apaydın et al., 2004). Desafortunadamente, estos métodos de interpolación consideran sólo las relaciones espaciales de proximidad entre las estaciones climatológicas sin tomar en cuenta características del paisaje que influyen en la distribución de la precipitación, por lo que tienden a efectuar predicciones sesgadas.

Una de las principales limitantes para la determinación de la distribución espacial de la precipitación es la escasa e irregular distribución de estaciones meteorológicas distribuidas en el norte de México además de intervalos de tiempo intermitentes. Dado la variabilidad en espacio y tiempo, de la precipitación, junto con la escasez de datos en países en vías de desarrollo se limita la modelación de los patrones de precipitación.

Muchos son los factores que determinan la variabilidad de la distribución espacial y temporal de la precipitación a diversas escalas espaciales y temporales. Por ejemplo a escala global, existen procesos océano-atmosféricos que se desarrollan a consecuencia de las variaciones en las temperaturas oceánicas tanto en el Atlántico como el Pacífico (El Niño/Niña, “Atlantic Multi-Decadal Oscillation” AMO, “Pacific Decadal Oscillation” PDO). Otro es la zona de convergencia intertropical (ZCIT) que tiende a desplazarse hacia el Norte durante el verano, que es la responsable de las lluvias de verano en México (Nicholson, 1996).

A escala sinóptica, el sistema atmosférico denominado Monzón Norteamérica que se desarrolla en el Golfo de California y los frentes fríos; en tanto que localmente, factores como el relieve y la proximidad de las zonas marítimas, pueden incidir de manera importante en la distribución de la precipitación (Ahrens, 2007). La influencia de la elevación del relieve sobre la distribución de la precipitación muestra en diversos estudios relaciones significativas que indican que la precipitación tiende a incrementarse con la elevación del relieve debido al efecto orográfico de las montañas (Daly et al., 2002); sin embargo, esta relación es más complicada debido a que existen otros factores, como la exposición y complejidad del relieve, que juegan un papel determinante en la

distribución de la precipitación en zonas montañosas (Daly et al., 2002; Hession y Moore, 2011).

En México existen series de datos de precipitación registrados en estaciones meteorológicas del Servicio Meteorológico Nacional, administradas por la Comisión Nacional del Agua, que son almacenados en el sistema Climate Computing Project (CLICOM) desarrollado por las Naciones Unidas. Este contiene series de datos desde 1920 hasta el 2010, de aproximadamente 3,817 estaciones climáticas distribuidas en el territorio nacional.

Desafortunadamente, la información registrada presentan una cantidad importante de inconsistencias entre las cuales puede mencionarse: escasa e irregular distribución espacial de estaciones meteorológicas, así como ausencia de registros por periodos prolongados, difíciles de completar. Estas estaciones miden las variables climatológicas e hidrométricas que se registran en el país. 1,064 son estaciones de referencia de donde se obtienen datos que determinan el comportamiento normal del clima nacional, de estas se obtiene información para calcular la precipitación media de 30 años y cuentan con más de 80% de los datos. Las estaciones climatológicas miden temperatura, precipitación pluvial, evaporación, velocidad y dirección del viento. Actualmente se tiene una carencia de estaciones climatológicas de referencia en el noroeste, norte, noreste y sureste del país, principalmente entre los estados de Chihuahua y Coahuila (Atlas digital del Agua 2012). Al Norte de México las características fisiográficas, los sistemas montañosos y el macizo forestal tienen un clima predominante de templado -frio y en ciertas zonas debido a las condiciones del relieve existen áreas cálidas - húmedas donde predomina la selva baja caducifolia, también existen planicies donde predominan el clima semiárido, aquí se localizan los principales desiertos del país,

por la diversidad geográfica del Norte de México y debido a las inconsistencias de los datos soportados por el CLICOM se utilizó la Red Histórica Global de Climatologías (GHCN) ya que tan solo para el área de estudio en el Norte de México se cuentan con 445 puntos con aproximadamente 28, 000 registros continuos de precipitación mensual desde 1948 hasta el 2010.

La calidad del modelo y la eficiencia en su respuesta a la hora de aplicarlo en zonas donde no se disponga con registros históricos de precipitación, pero se tengan condiciones similares permiten su modelación espacial (Spreen, 1947). Por lo anterior se planteó la modelación de la precipitación espacial por medio de regresión lineal múltiple en el Norte de México.

2. OBJETIVO

El objetivo del presente trabajo fue representar espacialmente y de manera fiable la distribución de la precipitación anual y estacional que ocurre en una región representativa del Norte México mediante el uso de modelos de regresión lineal multivariados, considerando como variables predictoras rasgos geográficos (latitud, longitud), de relieve (elevación, y complejidad del relieve) y de proximidad a las zonas marítimas.

2.1 Objetivos Específicos

- Determinar las normales climatológicas de la precipitación anual y estacional de una fracción representativa del Norte de México para los periodos de 1950-1980, 1960-1990, 1970-2000 y 1980-2010 utilizando series de datos provenientes de la Red Histórica Global de Climatología (GHCN).
- Determinar la influencia de los factores geográficos, topográficos y de proximidad a las zonas marítimas sobre la distribución espacial de la precipitación media anual y estacional en la región.
- Llevar a cabo procesos de interpolación espacial de la precipitación media anual y estacional de las normales climatológicas, mediante el uso de modelos de regresión lineal múltiple.
- Efectuar procesos de calibración espacial de la precipitación anual y estacional, con el uso de los residuales de los modelos de regresión lineal múltiple.
- Evaluar la fiabilidad de las interpolaciones espaciales de la precipitación anual y estacional de cada periodo, previa y posterior a los procesos de calibración espacial de la precipitación.

3. HIPÓTESIS

La distribución de la precipitación anual y estacional que ocurre en el Norte de la República Mexicana está determinada por la presencia de procesos océano-atmosféricos que ocurren a escala global, sin embargo rasgos de escala regional tales como la posición geográfica (latitud, longitud), las características del relieve (elevación, complejidad del relieve) y la proximidad a las zonas marítimas son factores que influyen en su distribución espacial, por lo cual puede ser representada espacialmente mediante el uso de métodos estadísticos multivariados

4.- REVISIÓN DE LITERATURA (ANTECEDENTES)

Existen diferentes trabajos y estudios donde se han llevado a cabo modelaciones de la precipitación, a partir de diferentes metodologías, por lo que sus resultados han sido tema de discusión y reflexión sobre cuál es más idóneo a utilizar con respecto a esta variable climática.

4.1 Modelos de Estimación de la Precipitación

En México se han desarrollado diferentes estudios para estimar la precipitación uno de ellos fue desarrollado en el Instituto de Ingeniería de la Universidad Autónoma de México UNAM donde se presenta un método de calibración de un modelo de estimación de lluvia denominado Hidroestimador el cual estima la precipitación en la Región Hidrológica número 30 de México llamada Cuenca del

Grijalva – Usumacinta, utilizando imágenes satelitales y estaciones climatológicas instaladas en campo que envían la información de manera constante.

Para la calibración del modelo se hizo un análisis de la información que ofrecen las estaciones climatológicas instaladas en la región, una regionalización de la zona así como una selección de tormentas representativas divididas en tres periodos, ciclones, frentes fríos y estiaje. Con los datos obtenidos se realizó una calibración del modelo utilizando un análisis de sensibilidad de los parámetros tomando como referencia coeficientes de determinación entre lo estimado y lo observado, se determinó el coeficiente de eficiencia del modelo y se aplicaron los valores sensibilizados con mayor relación de R^2 (Velázquez 2014).

En el valle de México se desarrolló un modelo de correogionalización lineal para la estimación espacial de la precipitación, en este trabajo se presenta una modificación del procedimiento geoestadístico de estimación espacial de la precipitación, el cual aplica el método de cokriging ordinario, combinando imágenes de radar meteorológico con datos de pluviómetros. Aquí, a diferencia de este procedimiento, se incluye en el método de cokriging ordinario la dependencia espacial conjunta de radar-pluviómetros mediante un modelo de correogionalización lineal (Díaz *et.al.* 2009).

Estudios como los de (Kurtzman y Kadmon, 1999; Oliver y Webster, 1990; Mitás y Mitasova, 1988) modelan la distribución espacial de una variable climática, al utilizar diferentes métodos de interpolación. De estas técnicas se pueden obtener resultados satisfactorios a partir de datos limitados, con base principalmente en la geografía de los puntos de muestreo o estaciones climáticas, la relación topológica entre estos puntos y el valor de la variable a medir.

Sin embargo los métodos de interpolación solo tienen en cuenta las relaciones espaciales entre los puntos de muestreo y dejan de lado otras propiedades del paisaje.

La precipitación generalmente aumenta con la altitud (Spreen, 1947; Smith, 1979), por lo general muchos autores han incorporado la elevación con enfoques geoestadísticos (Martínez - Cob, 1996; Prudhomme y Duncan, 1999; Goovaerts, 2000). Otros han desarrollado relaciones entre precipitación y variables topográficas como la altitud, latitud, continentalidad, la pendiente, la orientación o la exposición, al usar regresión (Basist *et al*, 1994; Goodale *et al.*, 1998; Ninyerola *et al.*, 2000; Wotling *et al*, 2000. Weisse y Bois, 2001). Sin embargo, la precisión de los resultados obtenido por estos métodos en las regiones montañosas es todavía muy limitado (Basist *et al*, 1994; Daly *et al*, 1994).

En un estudio en las regiones autónomas de Austria en el Norte de España se realizó una modelación de la precipitación en las regiones montañosas mediante el uso de sistemas de información geográfica y un análisis multivariado este enfoque de regresión permitió describir entre el 58 – 68 % de la variabilidad espacial de la precipitación anual y la precipitación media estacional se utilizaron cinco variable topográficas (Modelo de elevación digital, la elevación media de la sub-cuenca, la pendiente media de la sub-cuenca, la distancia más corta de una línea al oeste, proximidad a las costas) los resultados mostraron una mejora sobre los modelos desarrollados usando sólo variables relacionadas con la altitud.

También se utilizó las sub-cuencas como una variable la cual hizo una gran contribución al modelo obteniendo una mejora del 10% en relación con el uso de variables topológicas puntuales.

Según los resultados, ese enfoque permitió identificar peculiaridades de los patrones de precipitación con las variables de distancia más corta al oeste y de proximidad a las zonas costeras, que están relacionadas con la posición de frentes dominantes, mostrando una alta precipitación en el Noreste de la región y una baja precipitación en las cuencas internas. Las variables de elevación y altura media de la sub-cuenca muestran un aumento en la precipitación más rápido, a más de 1000 msnm., mientras que la pendiente media de la sub-cuenca se convierte en una variable de factor importante en la función con pendientes de más de 12 grados.

Con este método también se estimó la precipitación en zonas donde no hay estaciones cercanas, estos datos no se deducen directamente de las estaciones más cercanas, sino más bien de las particularidades topográficas del terreno, siempre y cuando las características topográficas estén dentro de los intervalos utilizados en los ajustes de la ecuación de regresión.

Una de las ventajas de los métodos de interpolación es que la calidad de los datos no depende de la densidad de estaciones de medición en las áreas vecinas, es suficiente tener estaciones en zonas de similares características topográficas, a pesar de que estén relativamente separadas. Por la misma razón, el área de estimación también se puede extender más allá de la última estación de medición en la zona, siempre y cuando las características topográficas y la escala de trabajo sean similares

4.2 Escalas de Referencia en Estudios de Modelación de la Precipitación

Muchos factores juegan un papel en la formación de la precipitación a través de diversas escalas espaciales y temporales, el conocimiento de estos factores se puede utilizar para para realizar esfuerzos de modelado, por ejemplo en la superficie terrestre o el océano, elevación orográfica, y otros procesos de mesoescala, este se refiere al proceso que se produce en regiones que van desde unos pocos kilómetros a unos cien kilómetros de diámetro (Ahrens, 2007), incluyendo la superficie terrestre, el océano y la orografía de terreno montañoso. En escala sinóptica son procesos de impacto sobre la distribución espacial de la precipitación en áreas que van de 100 a 1000 km².

Estos procesos incluyen las zonas alta y baja presión y los frentes meteorológicos asociados. En África Oriental, una región de variable topografía, con características de escala sinóptica como la presencia de baja presión más al Sur en enero trae del Noreste vientos [al Norte de la zona de convergencia intertropical (ZCIT)] que entra desde el Océano Índico (Nicholson, 1996). En julio, la celda de baja presión se mueve a el Noreste, a lo largo de Asia y la India, lo que resulta en el verano se estacionen ahí los monzones, y al Sur hasta el Suroeste vientos del Este de África.

Los procesos a escala global, como el dipolo del Océano Índico y la ZCIT también juegan un papel en los patrones asociados con la lluvia. En los trópicos, la posición de la ZCIT y sus bandas asociadas de las precipitaciones se mueven hacia el Norte y hacia el sur durante el año, dominando los patrones espaciales de precipitación.

En la práctica, los efectos, de la escala sinóptica, de mesoescala y los procesos de escala global se pueden modelar a través del uso de variables predictoras o independientes en un análisis estadístico de correlación y regresión, teniendo en cuenta las variables que son altamente correlacionado con la precipitación y para que los datos sean más fácilmente disponibles, tales como la elevación y sus derivados (medidas de procesos de mesoescala), medidas geográficas de ubicación (procesos de escala sinóptica y global), y el análisis de estación específica (teniendo en cuenta la variación que hay entre los procesos de escala sinóptica y global), pueden ser útiles en la predicción de la precipitación en espacios no muestreados.

4.3. Influencia de la Topografía en la Precipitación

Estimar el gasto instantáneo de un escurrimiento superficial, es un aspecto básico para la planeación, diseño y construcción de obras para su manejo y control, además del manejo del agua. Para su estimación existen varios métodos que se describen a continuación.

Las técnicas que utilizan la covarianza pueden hacer uso de la relación que hay entre las variables, para estimar la precipitación con mayor precisión, sus patrones en el espacio y el tiempo. Muchos investigadores se han basado en el establecimiento de una correlación entre la precipitación y la elevación para mejorar nuestra comprensión de los patrones de precipitación (Spreen, 1947; Daly *et al.*, 1994; Pardo-Iguzquiza, 1998; Goovaerts, 1999a, 1999b, 2000; KiefferWeisse y Bois, 2001; Marquínez *et al.*, 2003.; Kyriakidis *et al.*, 2004; Arora *et al.*, 2006). Los niveles relativamente altos de correlación han sido identificados; la relación es en general un aumento a uno (Spreen 1947; Arora *et al.*, 2006.):

A medida que aumenta la elevación, la precipitación aumenta. Esto se debe principalmente a los efectos de la orografía en los terrenos montañosos. Sin embargo, la precipitación puede tener una relación muy compleja con la elevación (Arora *et al.*, 2006.); el análisis de esta relación puede ser complicado por la distribución de las estaciones y de otros factores (Hulme y Nueva, 1997). Algunos investigadores se han centrado en el uso de la elevación como una única variable predictiva, pero hay muchos otros factores que se han identificado que desempeñan un papel importante en la distribución de las precipitaciones. Otros investigadores han ampliado la lista de posibles variables predictoras para incluir los derivados de la elevación y otras variables que resume la ubicación geográfica (Spreen, 1947; Hutchinson, 1998b; Kieffer Weisse y Bois, 2001; Marquinez *et al.*, 2003; Diodato, 2005; Oettli y Camberlin, 2005). Por ejemplo, (Spreen 1947) encontró que la distribución de la precipitación también depende de variables tales como la pendiente, exposición y orientación.

Del mismo modo, Keiffer Weisse y Bois (2001) encontraron que las variables topográficas se correlacionan con eventos de lluvias fuertes (v.g., de 10 y 100 años-eventos), en particular cuando se refiere a medidas de corta duración (i.e., menos de 3 hrs), las variables topográficas regionales (v.g., la distancia hasta el Mediterráneo, la caracterización de la forma general de los Alpes, la distancia a características de los Alpes) se ha encontrado que influyen en las lluvias intensas, mientras que las medidas locales de topografía (v.g., la altitud, la pendiente, o azimut) eran menos influyentes. Cuando se evaluaron períodos de datos más largos (v.g., varios meses o años), con análisis de regresión utilizando la precipitación mensual o estacional como la variable dependiente para tener en consideración los patrones temporales (Daly *et al.*, 1994; Marquinez *et al.*, 2003; Diodato, 2005; Oettli y Camberlin, 2005). En este sentido los datos fueron estratificados temporalmente y los análisis de regresión independientes de una duración de cada período de tiempo. (Daly *et al.* 1994). También se estratificaron datos de lluvia en espacios particulares a los que se denominaron “facetas topográficas” para dar importancia a una topografía variada sobre grandes regiones (v.g., la totalidad oeste de los Estados Unidos), en lugar de desarrollar un modelo de regresión multivariada para representar la variación espacial.

Se realizaron análisis de regresión independiente de precipitación contra la elevación por cada faceta. Muchos de las referencias de estudios anteriores con base en mínimos cuadrados ordinarios (OLS) de regresión, sin embargo, no se incluyó de manera explícita la autocorrelación espacial.

Los métodos estadísticos que se han utilizado recientemente que relacionan la precipitación con la elevación incluyen otras variables predictoras, estos métodos estadísticos varían ampliamente de OLS regresión simple, hasta métodos explícitos para análisis geoestadísticos espaciales. La mayoría de los estudios espacialmente explícitos incorporan tan sólo uno o dos factores de predicción de las precipitaciones, incluyendo en general la elevación, y comparan las técnicas de interpolación que van desde los polígonos de Theissen, distancia inversa, métodos de ponderación hasta geoestadística más sofisticada, tales como kriging con una derivada externa y cokriging (ambos incorporan variables independientes o datos de covarianza como la elevación para mejorar las predicciones de precipitación). El enfoque de los estudios espacialmente explícitos es a menudo para desarrollar las estimaciones más precisas de precipitaciones sin necesariamente entender el papel de los diversos factores predictivos, ni probar hipótesis acerca de ellos formalmente. Schabenberger y Gotway (2005) observaron que el uso de las variables de predicción no es el enfoque principal de muchos de estos tipos de estudios: las variables predictoras fueron a menudo usados para “estimar una variación espacial media y evitar el sesgo”. En muchos casos, no hubo intención de interpretar las relaciones entre las variables predictoras y la variable dependiente, o su significado. Los estudios diseñados para probar hipótesis acerca de múltiples predictores y su influencia en las precipitaciones dependen en gran medida de regresión OLS. Sin embargo, la regresión OLS no tiene en cuenta la autocorrelación espacial y por lo tanto, las estimaciones son sesgadas cuando se utiliza para evaluar los datos autocorrelacionados espacialmente.

Otros métodos que se han utilizado para estimar la distribución espacial de la precipitación son: método suavizado de splines laplacian (en el caso de dos dimensiones, estos también se conocen como splines suavizado de placa delgada) se utiliza para suavizar e interpolar datos de precipitación en el espacio utilizando coordenadas de localización de la información (Hutchinson y Bischof, 1983; Hutchinson, 1998a, 1998b). En lugar de incorporar explícitamente la autocorrelación espacial en los datos de estimación de los parámetros, sin embargo, estos estudios utilizan un procedimiento de adelgazamiento sistemático de datos para evitar problemas con la correlación de corto alcance en el espacio. Hutchinson (1998a, 1998b), evaluó los patrones espaciales en 1 día de datos de precipitación., Hutchinson y Bischof (1983) estudiaron a largo plazo medias estacionales (v.g., los datos de precipitación fueron estratificados por periodos estacionales) y la precipitación anual. Además, los procesos de modelos físicos básicos se han utilizado para predecir la distribución espacial de la precipitación en zonas montañosas.

Al respecto, Barros y Lettenmaier (1993) modelaron la advección de humedad durante barreras topográficas utilizando un modelo 4D de Lagrange. Este modelo simula orografía, se indujo la precipitación a una escala suficiente para resolver los rasgos topográficos dominantes. Aunque se emplearon métodos que utilizan splines o modelos físicos basados en procesos para interpolar los datos de lluvia sobre el espacio, no se utilizaron para la prueba de hipótesis formal de relaciones entre la precipitación y las posibles variables explicativas.

Sin embargo, Hutchinson (1998b) y Sharples *et al.* (2005), de manera informal evaluaron la influencia de las variables explicativas sobre los resultados del modelo. En este sentido, Sharples *et al.* (2005) demostraron que la adición de un término de elevación al modelo mejorado, los resultados redujeron significativamente los errores de interpolación.

Cuando el objetivo de un análisis es la prueba de hipótesis formal, la regresión espacial proporciona una solución que permite realizar pruebas estadísticas formales de hipótesis relacionadas con múltiples variables independientes, a la vez que representa la autocorrelación espacial. Schabenberger y Gotway (2005), identificaron la regresión espacial como una forma de análisis de datos, donde la atención se centra en el modelado y la comprensión de la función de media.

La comprensión se logra cuando se identifican las variables importantes que juegan un papel en los patrones de precipitación. Schabenberger y Gotway (2005) previeron que en una inversión de métodos de tipo kriging, los parámetros de covarianza pueden, tener a veces, parámetros de perturbación.

En un estudio en África Oriental se retomó el trabajo de Oettliand Camberlin (2005) en el cual se hace una incorporación de regresión espacial como un enfoque más innovador para la prueba de hipótesis que incorpora explícitamente autocorrelación espacial en los datos. El objetivo fue mejorar el modelado de las precipitaciones en el este de África por lo que se representa para la dependencia espacial en los datos, lo que evita las trampas comunes de análisis de regresión OLS de los procesos espacialmente autocorrelacionados: subestimación de la incertidumbre en los parámetros del modelo y la sobrevaloración de la fiabilidad de las predicciones del modelo.

Otros objetivos fueron diferenciar entre los efectos regionales y localizados a través de la consideración de múltiples escalas espaciales, y para evaluar los efectos de la estacionalidad de las precipitaciones. Este análisis además permitió la evaluación independiente de los mecanismos dominantes de la generación de precipitaciones durante todo el año.

Este estudio incorpora espacialmente métodos explícitos para modelar las precipitaciones en el este de África usando factores predictivos en diversas geografías (9×9 km a 213×213 km) y escalas climáticas (mesoescala, escala sinóptica y la escala mundial). El uso de modelos espacialmente explícitos y múltiples variables predictoras en muchas escalas resultó en mejoras sobre la investigación previa al capturar la variabilidad espacial de la precipitación regional para proporcionar estimadores más eficientes y precisos. Por otra parte, los resultados de este estudio indicaron que los procesos de mesoescala y de escala global pueden estar bien diferenciados utilizando técnicas de regresión espacial.

Diferenciar estos procesos, y ser capaz de utilizar modelos de precipitación en áreas con pocos datos, es útil para la investigación meteorológica y climatológica aplicada, por lo tanto, justifica este enfoque.

Oettliand Camberlin (2005) no reportaron los coeficientes de sus modelos, por lo que no es posible una comparación directa de los resultados; sin embargo, existen algunas similitudes. Hay una correspondencia directa en los signos de los coeficientes de regresión para la latitud, la longitud, la pendiente y la desviación estándar.

En este estudio, la pendiente fue significativa para una mayor gama de escalas espaciales (i.e., hasta 213 km y 71 km en el estudio), al parecer los mecanismos de mayor escala de la precipitación fueron capturados por los principales componentes derivados, para representar la topografía. La capacidad predictiva general de los modelos MCO (que se resumen por el coeficiente de determinación o R^2) fue mayor para Oettliand Camberlin (2005), sin embargo, el estudio mostró que la regresión OLS era menos apropiada que la regresión espacial debido a la presencia de autocorrelación espacial en los datos y puede resultar en valores inflados de R^2 . El resultado de este estudio mejora el rigor estadístico incorporando específicamente dependencia espacial en el modelado de precipitación, dando estimaciones más precisas de la incertidumbre y fiabilidad en coeficientes del modelo y predicciones. Este enfoque con lleva el supuesto de que la precipitación media en cada estación es constante en el tiempo (i.e. no hay tendencias significativas en el tiempo), en consistencia con las suposiciones hechas Oettliand Camberlin (2005). Será necesario en trabajos futuros evaluar mejor este supuesto y para contrastar las tendencias temporales que se identifican.

5.- MATERIALES Y METODOS

5.1 Localización Geográfica

El área de estudio se ubicó en una fracción representativa del Norte de los Estados Unidos Mexicanos, entre las coordenadas extremas $31^{\circ} 20' 02.01''$ N a $22^{\circ} 02' 37.73''$ O y $109^{\circ} 30' 30.21''$ N a $97^{\circ} 44' 38.24''$ O, que contiene una amplia variedad de condiciones geográficas y topográficas, abarcando los estados de Chihuahua, Durango, Sinaloa, Coahuila, Zacatecas, Nuevo León, Tamaulipas, parte de Sonora, San Luis Potosí, Aguascalientes, Nayarit y Jalisco.



Figura 1. Ubicación geográfica del área de estudio.

5.2 Información Ambiental y Ecológica

La ubicación geográfica (posición latitudinal cercana al Trópico de Cáncer), el gradiente relativo de la influencia continental y marítima, y las condiciones orográficas, con altitudes que varían desde el nivel del mar hasta por arriba de los 5 000 msnm, generan un mosaico de condiciones climáticas y ambientales en la región Norte de México.

El Trópico de Cáncer, además de su influencia en el régimen térmico, define también en forma aproximada la transición entre el clima árido y semiárido de la zona anticiclónica al Norte y el clima húmedo y semi-húmedo al Sur, influenciado por los vientos alisios y los ciclones que provienen del trópico.

Por otro lado, el efecto de los litorales y la alineación de los principales sistemas montañosos en el Norte del país, junto con el gradiente altitudinal, se manifiestan de manera decisiva en la distribución de la humedad y de la temperatura (J. Vargas 2003).

A pesar de la gran diversidad de condiciones climáticas y ambientales ocasionada por los factores anteriores, es posible establecer algunas generalidades particularmente en esta región. Por ejemplo, predominan ampliamente los climas secos sobre los húmedos. Las zonas en donde la vegetación dispone de humedad suficiente en todas las épocas del año no son mayores del 5% de la superficie total de la región. También es notorio que la vertiente del Atlántico, en el Este, es un poco más húmeda que la vertiente del Pacífico en el Oeste.

En la mayor parte de la región Norte predomina el régimen de lluvias de verano que prevalece en casi todo el país, excepto en algunas partes de las montañas del Noroeste, donde existe un cambio gradual hacia el régimen de lluvias en invierno.

En la región Norte del país se presentan desde los climas de tipo cálido y húmedo, con largas temporadas de sequía, en las planicies costeras y tierras bajas de ambos litorales, hasta los climas de tipo frío o polar en las partes más altas de las montañas. En la zona del altiplano central predominan los climas secos con diferente grado de aridez, incluyendo a los de tipo desértico. En cambio, en las zonas montañosas de la región predominan los climas templados y húmedos. Sin embargo, debido al mosaico de condiciones topográficas, es común observar que sitios muy cercanos pero situados en lados opuestos de una sierra presenten condiciones climáticas completamente diferentes entre sí, en términos de precipitación y temperatura.

Con estas condiciones se presenta una gran diversidad de condiciones térmicas, desde temperaturas media anual por debajo de 0° C en las cimas de las montañas más elevadas, con nieves perpetuas y glaciares, hasta los 25-28 °C en algunas zonas costeras. Sin embargo, la mayor parte del altiplano central y de las serranías en esta región tiene una temperatura media anual entre 10 y 20°C. Un aspecto más importante, sin embargo, es la oscilación térmica anual que se presenta en esta región del país, con diferencias entre las temperaturas medias en diferentes épocas del año desde 12 hasta más de 20°C, indicando una estacionalidad térmica bien definida. Las temperaturas extremas registradas en el año varían desde los -15 °C en algunas partes de la Sierra Madre Occidental en Chihuahua, hasta los 55 – 60 °C en los desiertos (J. Vargas 2003).

En cuanto a la precipitación, el promedio anual en la región Norte varía desde menos de 50 mm en algunas zonas del desierto, hasta cerca de los 2 000 mm en algunos sitios de las vertientes oceánicas de las Sierras Madre Occidental y Oriental.

La temporada lluviosa del año varía desde 1-2 meses hasta 6-7 meses en la región norte del país, aunque entre el 60 y 80 % de la precipitación anual generalmente se presenta en no más de tres meses. Por lo tanto, si se considera tanto la distribución anual de la precipitación, como de la temperatura y de la humedad relativa ambiental, puede decirse que en la mayor parte de la región Norte existe un fuerte déficit de humedad en la mayor parte del año.

La cubierta vegetal de México es una de las más variadas del planeta. La gran diversidad biológica se debe, por un lado, a las condiciones fisiográficas, geológicas y climáticas, que presentan una amplia gama de variación y combinaciones, generando un mosaico de nichos ecológicos (Rzedowski, 1994). Sin embargo, también es importante el hecho de que la mayor parte del territorio mexicano es considerado por los biogeógrafos como la transición entre la región Neotropical, con elementos meridionales sudamericanos, y la región Neártica, con elementos boreales norteamericanos (Benítez y Neyra, 1997). De esta manera, en la región norte están representados la mayor parte de los grandes biomas existentes en el planeta, exceptuando los bosques tropicales húmedos; desde los tipos de vegetación de las zonas bajas y cálidas, hasta los desiertos, los bosques templados de coníferas y los páramos de alta montaña (J.Vargas 2003).

5.3 Materiales

- Los materiales utilizados fueron:
 - Un equipo de cómputo:
 - PC Windows 7 u 8
 - Requerimiento mínimo de memoria RAM 2 GB
 - Requerimiento mínimo de Procesador Intel Core I3
 - Sistemas de Información Geográfica y un paquete estadístico computacional:
 - ERDAS, Arc Gis, Arc Info, Statistica.
- Información espacial y cartográfica del área de estudio y las estadísticas de precipitación de la Red Histórica de Climatología Global.

5.4 Metodología

5.4.1. Red Histórica Global de Climatología (GHCN)

Las series de precipitación utilizadas en el presente estudio provienen de la Red Histórica Global de Climatología, en su versión 2.0, (GHCN) las cuales se derivaron a partir de un análisis de la precipitación mensual registrada en más de 17,000 estaciones colectadas por la Agencia NOAA, NCDC y el Sistema de Monitoreo de Anomalías Climáticas NOAA / CPC. La información se registra en una malla equidistante de puntos distribuida a cada 2.5° de latitud/longitud en las zonas terrestres globales para el período 1948-2010 (Chen *et. al.* 2001). La malla del GHNC forma parte integral de los esfuerzos para construir un análisis de la precipitación mensual en todo el mundo durante un largo periodo, desde 1948 hasta el presente.

Conocida como la reconstrucción de precipitación, este análisis mensual de precipitación consiste en el análisis de observaciones calibradas sobre la tierra y los campos reconstruidos sobre el océano.

La construcción de la malla se llevó a cabo mediante los algoritmos de interpolación ampliamente descritos por Cressman (1959); Barnes (1964), Shepard (1968), y el método de interpolación óptima de Gandín (1965) para examinar su desempeño en la interpolación de las observaciones calibradas de todos los días, de un grupo de cinco, y la precipitación mensual. Los algoritmos se desempeñan mejor en la interpolación de anomalía de precipitación que los valores totales y que el esquema de interpolación óptima, produce estadísticas superiores y estables en la mayoría de los casos en comparación con los otros tres algoritmos. La correlación entre los valores de análisis y las observaciones de la estación independientes retirados, es de aproximadamente 0.8 y el sesgo es casi 0 para la interpolación de precipitación mensual utilizando el algoritmo de Gandín. Con base en estos resultados, se seleccionó el algoritmo interpolación óptima para construir el análisis de la precipitación mensual en áreas globales de tierras por un período de 62 años 1948-2010.

Para el desarrollo del presente estudio se descargó el historial de datos de precipitación mensual que conforman la malla regular de GHCN. Esta malla de datos que tiene una distribución mundial, fue proyectada en una cobertura de puntos utilizando un Sistema de Información Geográfica. De esta cobertura mundial, se obtuvieron los puntos de la malla que cubre la zona de estudio en el Norte de México (Figura 2).

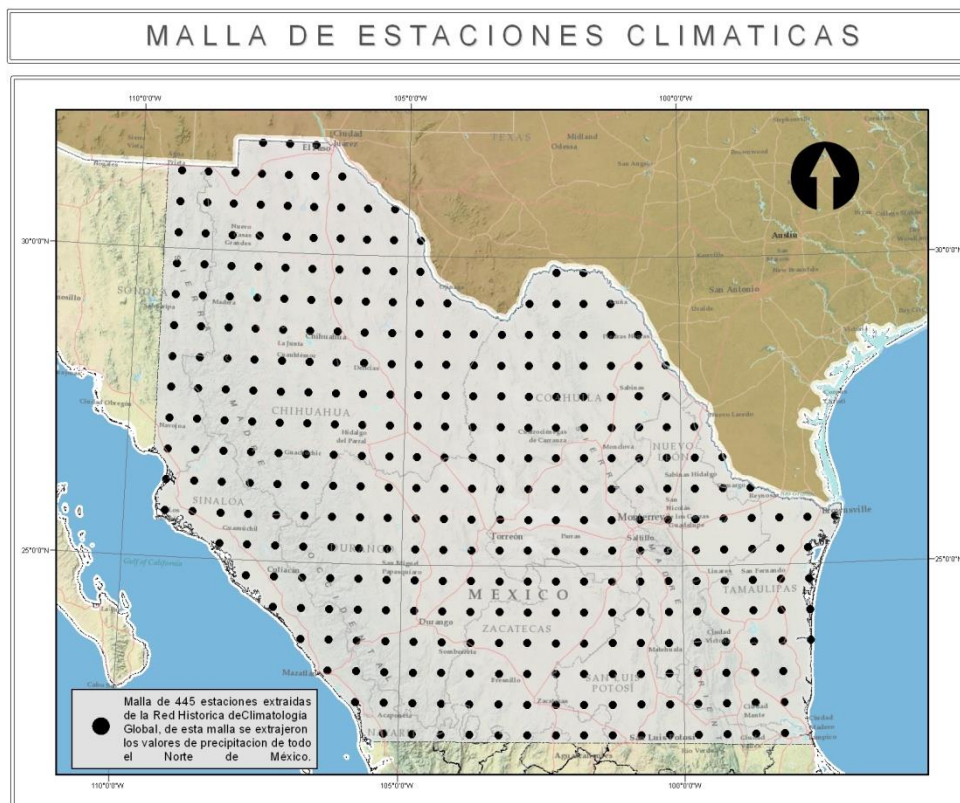


Figura 2. Malla de puntos con series de precipitación, ubicada en el Norte de México.

La zona de estudio fue cubierta con un total de 445 puntos de la malla GHCN. Para cada punto de la malla se calcularon las normales climatológicas de la precipitación anual y estacional (periodos húmedo y seco) de los periodos 1950 – 1980, 1960 – 1990, 1970 – 2000 y 1980 – 2010.

5.4.2. Descripción de variables independientes

En el estudio se identificaron las siguientes variables independientes o predictoras: a).- variables geográficas (latitud, longitud), b).- variables de relieve (elevación, y complejidad del relieve y c).- variables de proximidad (distancia a las zonas de costas) las cuales se describen con mayor detalle a continuación.

5.4.2.1 Variables Geográficas

Para cada punto de la malla se obtuvieron los valores de Latitud y Longitud.

5.4.2.2 Variables del Relieve

Para evaluar la influencia de las variables topográficas desde diversas escalas sobre la distribución espacial de la precipitación, espacial se utilizó un modelo digital de elevaciones (MDE) con resolución espacial de 30 metros proveniente de INEGI (2012). El MDE fue remuestreado a una resolución de 0.25 x 0.25 km, con base en esta resolución espacial y utilizando función “focal mean” del software Arc Info. Se derivaron modelos que representan el promedio de la elevación a escalas de 0.5 x 0.5, 1 x 1, 2.5 x 2.5, 5 x 5 km.

Esta función permite determinar el valor de altitud media observada desde un MDE de 250 mts de resolución, en ventanas de 500 x 500 mts, 1 x 1 km, 2.5 x 2.5 km y 5 x 5 km. La función que se realiza en la entrada es una estadística, en este caso el promedio de todos los valores que se encuentran al interior de estas ventanas. (ESRI 2012). De manera similar y con la finalidad de identificar el efecto de la variabilidad de la altitud del relieve se derivaron también los valores de desviación estándar de la elevación a las escalas de 0.5 x 0.5, 1 x 1, 2.5 x 2.5, 5 x 5 km.

5.4.2.3 Variables de proximidad

Esta variable se determinó con el método de distancia euclidiana, la herramienta describe la relación de cada celda con la malla de puntos que contiene la serie de datos de precipitación, basándose en la distancia de la línea recta expresada en metros.

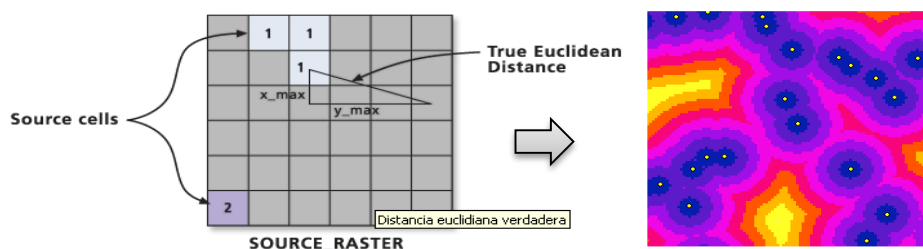


Figura 3. Calculo de la distancia euclidiana.

La distancia euclidiana se calcula desde el centro de la celda de origen hasta el centro de cada una de las celdas circundantes. Conceptualmente, el algoritmo euclidiano funciona del siguiente modo: para cada celda, la distancia a cada celda de origen se determina al calcular la hipotenusa con x_{max} y y_{max} como los otros dos lados del triángulo.

Este cálculo deriva la verdadera distancia euclidiana, en vez de la distancia de la celda. Se determina la distancia más corta a un origen, y si es menor que la distancia máxima especificada, el valor se asigna a la ubicación de la celda en el ráster de salida (ESRI 2012).

Una vez que se realizaron los cálculos de cada una de las variables independientes o predictoras para las diferentes resoluciones del modelo y para cada periodo estacional, se sobrepuso la malla de puntos de los datos de precipitación y con la herramienta “Sample” (ESRI 2012) de Arc GIS 10.1, se extrajeron los valores de cada una de las variables y se construyó una base de datos, los valores de precipitación se convirtieron a logaritmos. A esta base de datos se obtuvo una matriz de correlación para evitar que las variables estuvieran correlacionadas entre si y provocar redundancias en el modelo, de igual manera se identificaron las variables con mayor significancia estadística.

5.4.3. Construcción de Modelos de Regresión

Para explorar el efecto combinado de las variables independientes sobre la precipitación media anual y estacional, se llevaron a cabo análisis de regresión lineal múltiple mediante el procedimiento por pasos, con selección de variables hacia delante “Forward stepwise”.

Con este procedimiento, las variables significativas fueron seleccionándose en función del nivel de significancia expresado por el valor de $p < 0.05$. El método ofrece la ventaja de que una vez que las variables independientes al ser seleccionadas por su significancia estadística, se determina también el porcentaje de variabilidad que explica en cada paso, con lo cual se hace posible identificar el aporte que cada variable independiente presenta sobre el modelo de regresión.

Del total de la malla de puntos, se seleccionaron aleatoriamente el 60% para la obtención y ajuste de modelos de regresión lineal múltiple:

$$y = \beta_0 + \beta_0 \sum_{i=1}^n (X_i) \quad (1)$$

Donde:

Y = Variable dependiente o explicada.

X_1, X_2, \dots, X_n : variables explicativas o independientes.

$\beta_0, \beta_1, \beta_2, \dots, \beta_n$ = Parámetros, miden la influencia que las variables explicativas tienen sobre la variable dependiente donde β_0 es la intersección o término "constante", las β_i ($i > 0$) son los parámetros respectivos a cada variable independiente, y n es el número de parámetros independientes a tener en cuenta en la regresión.

5.4.4. Parámetros Estadísticos de validación

La validación ha sido definida como una comparación de las predicciones del modelo con el mundo real para determinar si el modelo es adecuado para el fin previsto (McKinion y Baker, 1982). La validación del modelo es un paso obligatorio en la compleja tarea de simulación. Descripciones generales de este proceso se pueden encontrar en la mayoría de los textos de simulación (Jorgensen, 1986; Bratley *et al*, 1987; Ripley, 1987), con puntos de vista más amplios algo más raro (Dent y Blackie, 1979).

La capacidad predictiva de cada modelo de regresión fue evaluada empleando los valores de precipitación media mensual observados en el restante 40% de los puntos de la malla que no fueron incluidas en los procesos de ajuste de los modelos.

Los parámetros estadísticos obtenidos para determinar la fiabilidad de cada modelo de regresión previo y posterior al proceso de calibración fueron: el error medio absoluto MAE, el valor porcentual del error medio absoluto MA%E, y el coeficiente de eficiencia (COE). El COE es definido como uno menos, de la suma del absoluto, de las diferencias cuadradas entre los valores normalizados de lo predicho y lo observado, por la varianza de los valores observados, durante el periodo de investigación.

$$COE = 1 - \frac{\sum_{i=1}^n (O_i - P_i)^2}{\sum_{i=1}^n (O_i - \bar{O})^2} \quad (2)$$

Cuando el rango del COE se encuentra entre $-\infty$ y 1 se dice que es un ajuste perfecto. Una eficiencia menor que cero indica que el valor medio en la serie de tiempo observada habría sido un mejor predictor que el modelo (Velázquez 2014).

$$MAE = (\sum |y_i - \hat{y}_i|) / n \quad (3)$$

$$MA\%E = 100[\sum |y_i - \hat{y}_i| / |y_i|] / n \quad (4)$$

Donde y_i representa los valores observados, \hat{y}_i los valores calculados, y n el número de total de datos.

MA%E se utiliza ampliamente, aunque con diferentes nombres, en los estudios demográficos (Siegal, 1972; Cohen, 1986; Smith, 1987) y otros campos (Hameed, 1974). Como MAE se encuentra en las mismas unidades que los datos y MA%E es relativo, ambos pueden ser medidas informativas.

Se debe tener en cuenta que MAE también se puede utilizar para el error medio algebraico (Smith, 1987), en cuyo caso las señales absolutas se omiten de la ecuación. (3). En este caso, no es una verdadera medida de la desviación, como desviaciones positivas y negativas se anulan entre sí. Más bien, su promedio general mide cualquier sesgo del modelo mediante la comparación de la distribución de las diferencias contra cero.

Como sesgo se mide mejor por otras pruebas estadísticas, no se recomienda este uso, y adoptamos MAE como su (absoluta) definición original. Una alternativa al uso de las diferencias absolutas es utilizar un segundo momentos (Picard y Cook, 1984), y utilizando el error medio cuadrático RMSE:

$$\text{RMSE} = \{[\sum(y_i - \hat{y}_i)^2]/n\}^{0.5} \quad (5)$$

Algebraicamente, $\text{RMSE} > \text{MAE}$ (debido a la influencia de la cuadratura de los valores más grandes), con estas medidas son aproximadamente iguales si las diferencias absolutas son de magnitud similar. Con la desviación al cuadrado, RMSE puede ser útil en la obtención de propiedades estadísticas. En resumen estas medidas del grado relativo de desviaciones, ya sea MAE o RMSE pueden ser utilizadas.

Kleijnen (1987) recomienda MA%E, y sugiere un 10% como límite superior en la aceptabilidad. Dado que el modelo de validez depende en gran medida de tanto el tipo de modelo y sus usos previstos (Bratley *et al.*, 1987), no es práctico para establecer un único límite absoluto. En muchos casos estas medidas se utilizan simplemente para comparar los diferentes modelos o técnicas.

Existe un problema potencial en la ecuación. (3) con la división por y_i . Obviamente, MA%E no está definido si cualquier valor observado (y_i) es igual a cero, pero esta restricción no se aplica a los valores simulados (\hat{y}_i). También existen problemas con bajos valores de y_i , como MA% E tiende hacia el infinito como cualquiera y_i tiende hacia cero.

Dado que MA%E se ve fuertemente influenciado por los valores individuales bajos de y_i , que debe ser visto con precaución a menos que todos y_i son de magnitud similar. Dada esta limitación, una mejor medida la desviación relativa puede ser el error medio absoluto con respecto a la media observada = MAE/\bar{y} , que también se puede utilizar si MAE no está definido.

5.4.4. Procesos de interpolación Espacial y Calibración

La interpolación espacial de la precipitación anual y estacional de cada periodo se llevó a cabo utilizando las ecuaciones resultantes de los modelos de regresión lineal concebidos con el 60% de los puntos de la malla de datos. Dichas ecuaciones fueron resueltas utilizando como datos de entrada las variables independientes representadas espacialmente mediante procesos de algebra de mapas en Arc Gis 10.1 (Figura 4).

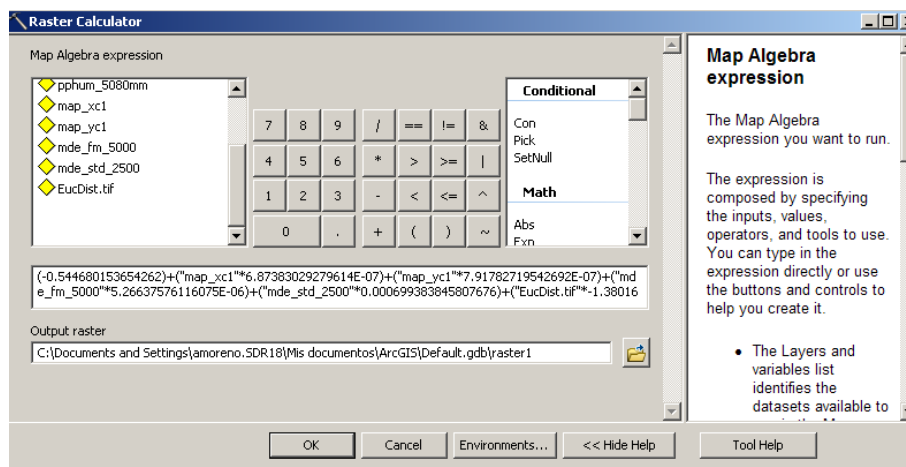


Figura 4. Función de Algebra de Mapas

El modelo se corrió al utilizar los raster de cada variable independiente y los valores predichos de cada variable según el modelo.

Los errores residuales calculados en una tabla, se vectorizaron con base en las coordenadas creando una malla de puntos estos valores se interpolaron para cada periodo estacional y anual. Para interpolar se utilizó el método Spline, el cual ajusta una función matemática a una cantidad especificada de puntos de entrada más cercanos mientras pasa a través de los puntos de muestra. Este método es mejor para generar superficies que varían levemente, por lo que la precipitación es un buen candidato.

La técnica de curvatura mínima básica también se conoce como interpolación por lámina delgada. Asegura una superficie suave (continua y diferenciable), junto con superficies continuas de la primera derivada. En las inmediaciones de los puntos de datos pueden ocurrir cambios rápidos en la gradiente o pendiente (primera derivada), por lo que este modelo no es adecuado para estimar la segunda derivada (curvatura).

Por esta razón se aplicó la opción de Tensión que modifica el criterio de minimización para que los términos de la primera derivada se incorporen a los criterios de minimización. El parámetro de peso especifica el peso adjunto a los términos de la primera derivada durante la minimización, que denomina Φ (phi) en la tabla. Un peso de cero resulta en la interpolación por spline por lámina delgada básica. Al utilizar un valor de peso más grande, se reduce la rigidez de la lámina y, a medida que phi se acerca al infinito en el límite, la superficie se aproxima a la forma de una membrana, u hoja de goma, que pasa por los puntos. La superficie interpolada es suave. Las primeras derivadas son continuas pero no suaves.

Para el método de spline de tensión, el parámetro de peso define el peso de la tensión. Cuanto más alto es el peso, más gruesa es la superficie de salida. Los valores introducidos deben ser iguales o mayores que cero. La cantidad de puntos identifica el número de puntos que se utilizan en el cálculo de cada celda interpolada. Cuantos más puntos de entrada especifique, mayor será la influencia de los puntos distantes sobre cada celda y más suave será la superficie de salida. Cuanto más grande sea la cantidad de puntos, más tiempo llevará procesar el ráster de salida (v.g. Periodo Húmedo 1980-2010). Los valores predichos están normalizados y no hay mucha dispersión en los errores de la interpolación.

La calibración del modelo consistió en compensar el error, el raster de valores calculados se le sumo al raster de errores residuales, obteniéndose un raster de valores calibrados.

Una vez realizada la calibración se determinaron los parámetros estadísticos para el 40 % restante, se obtuvieron las coordenadas de cada estación del 40 % y se convirtieron en un archivo vectorial de puntos el cual se sobrepuso al raster calibrado de cada periodo estacional y anual, para los 60 años de registros y se le extrajo el valor con la función Sample de Arc Gis 10.1, de esta manera se construyó una base de datos con valores calibrados y se calcularon los parámetros estadísticos.

6.- RESULTADOS

El proceso de modelar la precipitación espacial, requiere de ciertos procesos tales como: escalamiento, correlación, parámetros estadísticos, modelo de regresión lineal multivariado y calibración, que a continuación se describen.

6.1 Proceso de escalamiento

Remuestreo del modelo de 30 metros aun modelo de 250 metros, promedios focalizados 0.5, 1, 2.5 y 5 kilómetros (Figura 5).

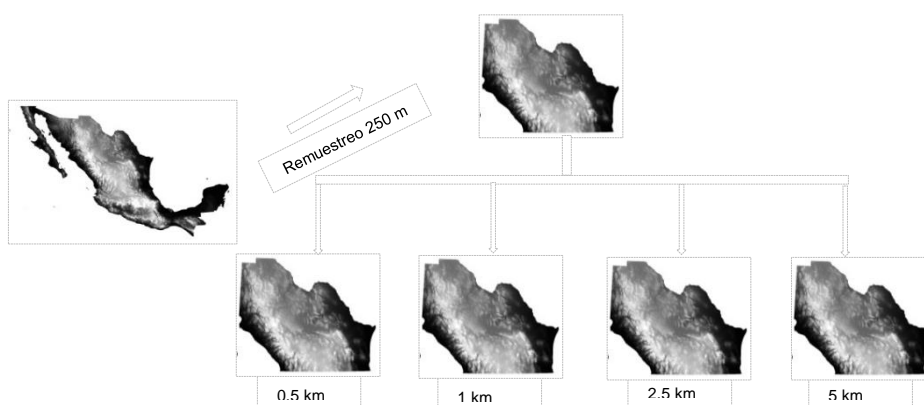


Figura 5. Remuestreo y promedios focalizados en diferentes resoluciones

Para representar la complejidad de la elevación del relieve se obtuvo la desviación estándar de (z) con la función de algebra de mapas, para obtener los promedios de las desviaciones individuales de cada valor de elevación con respecto a la media de la distribución, lo cual nos da el grado de dispersión o variabilidad se eligió el raster de 2.5 kilómetros, También se determinó la proximidad de las costa con el cálculo de distancia euclidiana en todo el país.

La clasificación de la precipitación se realizó con normales Climatológicas en periodos de estacionales y anual.

Cuadro 1. Clasificación de la Precipitación

Periodo Estacional	Normal Climatológica
Periodo Seco, Periodo Húmedo y Anual.	1950 - 1980
	1960 - 1990
	1970 - 2000
	1980 - 2010

6.2 Matriz de Correlación

Las variables independientes que se determinaron de acuerdo con la matriz de correlación fueron la latitud, la longitud, el modelo de elevación digital (z), la desviación estándar de la elevación y la distancia o la cercanía con las costas

En el periodo seco se observa que la longitud tiene una correlación positiva de 0.43 en promedio, en los otros periodos (húmedo y anual) no hay una correlación apreciable (Cuadro 2).

Para la latitud tiene una correlación negativa en promedio del 0.52 en el periodo húmedo, así mismo en el período anual y no hay una correlación durante el período seco.

En la elevación (z) focalizado de 5 km y a 2.5 km se presenta una correlación negativa y es similar en ambas escalas. Donde es mayor para el período seco 0.52, para el período anual de 0.215 y para el período húmedo del 0.065 en promedio, respectivamente.

En términos de la desviación estándar de la elevación a diferentes escalas, el comportamiento es el siguiente, se presenta un correlación positiva en el período húmedo que varía entre 0.26-0.35, en el período anual hay una correlación entre 0.23-0.30, y en el período seco no es apreciable.

En el caso de la distancia euclidiana hay una correlación negativa para todos los periodos. Para los períodos húmedo y anual la correlación es de -0.8 en promedio y en el período seco la correlación disminuye a -0.4, pero también negativa.

Cuadro 2. Matriz de Correlación

Matriz de Correlación (Variable Independientes)													
		PS-1950-1980	PS-1960-1990	PS-1970-2000	PS-1980-2010	PH-1950-1980	PH-1960-1990	PH-1970-2000	PH-1980-2010	A-1950-1980	A-1960-1990	A-1970-2000	A-1980-2010
Longitud	x	0.41	0.43	0.43	0.45	-0.03	-0.04	-0.04	-0.03	0.11	0.10	0.10	0.13
Latitud	y	-0.05	-0.08	-0.03	-0.01	-0.51	-0.51	-0.53	-0.53	-0.46	-0.46	-0.47	-0.46
Modelo (z)	2.5 km	-0.51	-0.50	-0.51	-0.52	-0.07	-0.06	-0.05	-0.06	-0.21	-0.20	-0.20	-0.22
	5 km	-0.51	-0.50	-0.51	-0.53	-0.08	-0.07	-0.05	-0.06	-0.22	-0.21	-0.20	-0.23
Desviación Estándar de (z)	0.5 km	0.05	0.05	0.03	0.01	0.26	0.28	0.28	0.28	0.23	0.24	0.25	0.24
	1 km	0.02	0.02	0.00	-0.02	0.30	0.31	0.31	0.32	0.26	0.27	0.27	0.27
	2.5 km	0.05	0.05	0.02	0.00	0.33	0.34	0.34	0.35	0.29	0.30	0.30	0.29
	5 km	0.01	0.01	-0.02	-0.04	0.28	0.29	0.29	0.30	0.24	0.24	0.25	0.24
Proximidad	Distancia Euclidiana	-0.44	-0.44	-0.40	-0.37	-0.83	-0.82	-0.81	-0.80	-0.83	-0.82	-0.82	-0.81

*PS= Periodo Seco, PH = Periodo Húmedo, A = Anual

6.3 Modelos de Regresión

En el cuadro 3, se indican los coeficientes de regresión de los modelos lineales para cada uno de los periodos citados y para cada década. El coeficiente de determinación para los dos periodos estacionales seco y húmedo está en los rangos de 0.6 – 0.7 (R^2).

Para el periodo seco las variables que aportaron entre el 0.20 y el 0.25 fueron la Latitud y el modelo de elevación (z) focalizado a 5 kilómetros en promedio con respecto al coeficiente de determinación (R^2) que es de 0.65. Para el periodo húmedo y anual la variables con mayor aporte es cercanía a las costas con más del 0.60 aproximadamente con respecto al coeficiente de determinación (R^2) que es mayor de 0.70 (Cuadro 3).

Cuadro 3. Variables del Modelo de Regresión

Periodo	Variable	1950-1980	R ² change	1960-1990	R ² change	1970-2000	R ² change	1980-2010	R ² change
Seco	B ₀	0.54468015	-	0.10733355	-	0.12870649	-	0.24783126	-
	Longitud	0.00000069	0.05947005	0.00000060	0.06847534	0.00000059	0.06489278	0.00000062	0.06691557
	Latitud	0.00000079	0.21820088	0.00000067	0.20240755	0.00000068	0.23087387	0.00000070	0.23833600
	Modelo de Elevación (z)	0.00000527	0.25820190	0.00000487	0.25408551	0.00001854	0.25763358	0.00000332	0.27584406
	Desviacion Estandar de (z)	0.00069938	0.01446305	0.00060963	0.01441998	0.00054554	0.01264876	0.00053933	0.01208797
	Proximidad a las Costas	0.00000138	0.10103071	0.00000119	0.10515591	0.00000111	0.07442156	0.00000109	0.06546853
	R ²		0.65136659		0.64454429		0.64047054		0.65865213
Humedo	B ₀	3.0858501	-	3.1081696	-	3.1984665	-	3.18481201	-
	Longitud	-0.0000001	0.0023523	-0.0000001	0.0037461	-0.0000001	0.0067138	0.00000009	0.01346392
	Latitud	-0.0000001	0.0023110	-0.0000001	0.0022878	-0.0000001	0.0056884	0.00000016	0.01106961
	Modelo de Elevación (z)	0.0000170	0.0203892	0.0000152	0.0237373	0.0000103	0.0265329	NS	NS
	Desviacion Estandar de (z)	0.0004111	0.0075066	0.0004296	0.0093698	0.0004420	0.0100349	0.0004892	0.02523996
	Proximidad a las Costas	-0.0000010	0.6919154	-0.0000010	0.6710328	-0.0000009	0.6561228	0.00000081	0.64651963
	R ²		0.7244746		0.7101738		0.7050929		0.6962931
Anual	B ₀	2.6469856	-	2.6828445	-	2.7956190	-	0.00000009	NS
	Longitud	0.0000001	0.0082785	0.0000001	0.0078167	0.0000001	0.0079552	0.00000005	0.01309210
	Latitud	0.0000001	0.0032998	0.0000001	0.0028207	NS	NS	NS	0.00278032
	Modelo de Elevación (z)	NS	NS	NS	NS	NS	NS	0.00051540	NS
	Desviacion Estandar de (z)	0.0005249	0.0089272	0.0005170	0.0107361	0.0004866	0.0118152	0.00000090	0.01105319
	Proximidad a las Costas	-0.0000011	0.6937990	-0.0000010	0.6788423	-0.0000009	0.6669159	NS	0.65082728
	R ²		0.7143045		0.7002158		0.6866863		0.6777529

6.4 Interpolación Espacial

La interpolación espacial de la precipitación anual y estacional de cada periodo se llevó a cabo utilizando las ecuaciones resultantes de los modelos de regresión lineal concebidos con el 60% de los puntos de la malla de datos.

La precipitación calculada del periodo seco oscila desde 50 hasta 500 mm, para el periodo húmedo desde 140 hasta 930 y para el periodo anual desde 200 hasta 1130 mm, como ejemplo tomamos el periodo húmedo de 1980 – 2010 la precipitación oscila desde 141 hasta 846 mm (Figura 6).

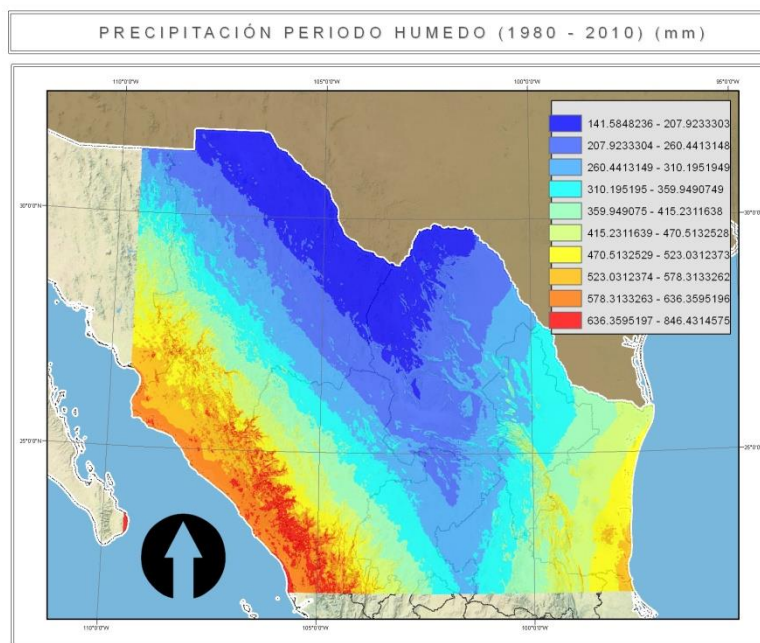


Figura 6. Precipitación calculada periodo húmedo 1980-2010 mm.

La Figura (7) muestra la distribución espacial del 60% de los puntos de la malla con datos de errores residuales aplicando el método de interpolación Spline con tensión.

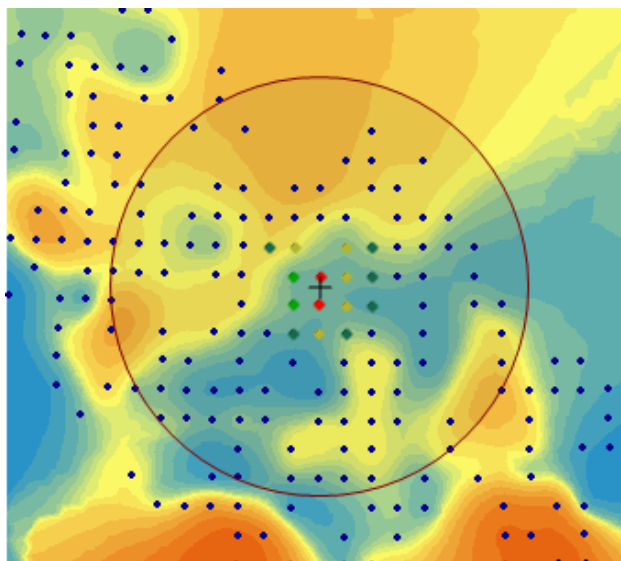


Figura 7. Interpolación con Spline aplicando Tensión

Los errores residuales oscilan de -50 a 50% aproximadamente en los periodos temporales (seco - húmedo) y anual, como ejemplo se muestra los errores residuales del periodo húmedo que van del -64% hasta el 54% (Figura 8).

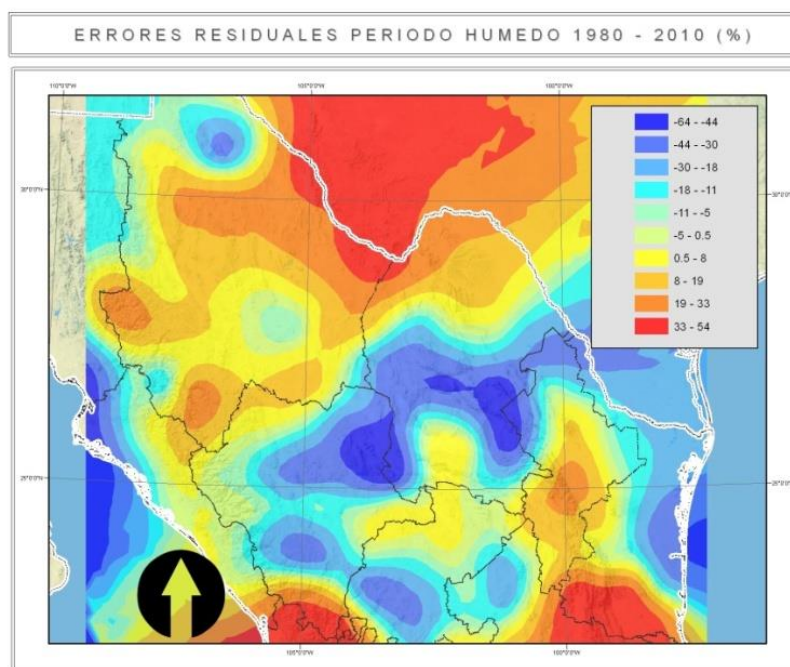


Figura 8. Errores Residuales en porcentaje.

6.5 Calibración Espacial

La calibración consiste en procedimiento para determinar si compensando espacialmente el error residual de los modelos, se mejora o al menos se mantiene constantes las pruebas de validación estadística.

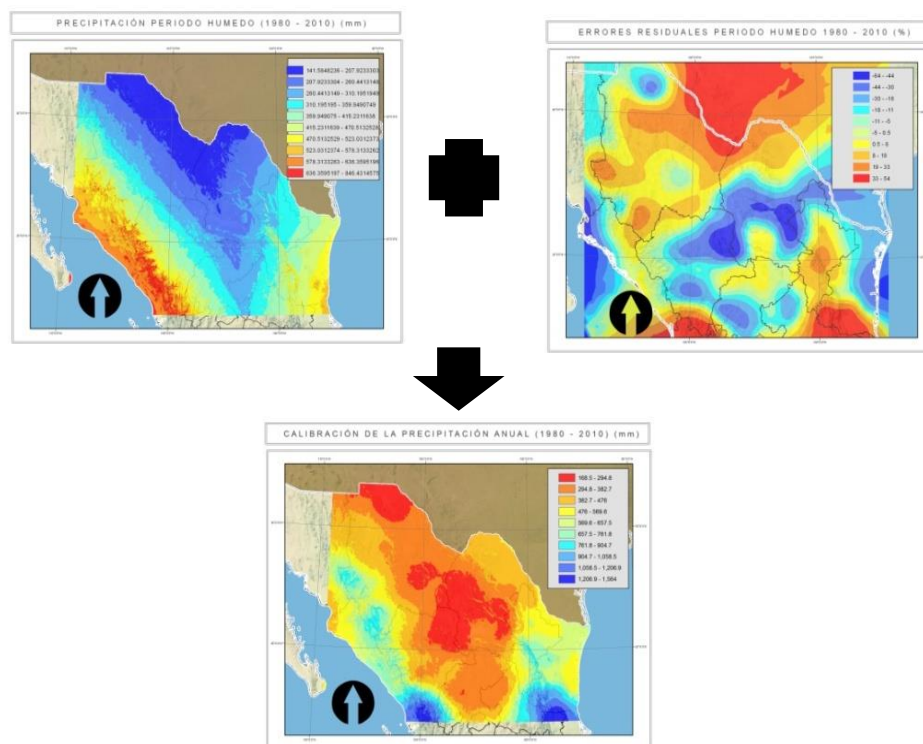


Figura 9. Calibración periodo húmedo 1980-2010.

Tal y como puede constatararse en la cuadro 4 el modelo al compensar el error mostró una mejoría sustancial, por ejemplo si analizamos la precipitación del periodo húmedo de 1980 – 2010 el mapa de precipitación representa los valores congruentes en espacio, sin embargo no se aprecia una uniformidad, simplemente la representación de los valores, mientras que el modelo calibrado muestra áreas más homogéneas y responde a una mayor representatividad en espacio (Figura 10), ya que hay una relación

estrecha entre la precipitación y la orografía del sitio, además de las congruencias con el medio físico y la climatología del Norte de México.

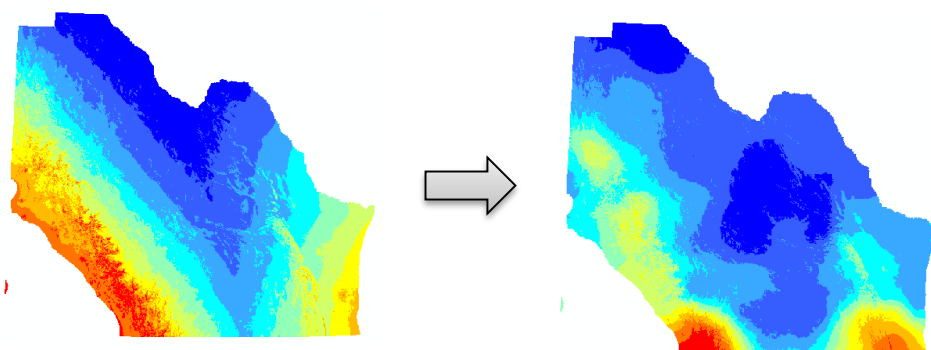


Figura 10. Comparación modelo calculado vs modelo calibrado

El periodo Calibrado Seco para 1950 -1980 oscila entre menos de 30 – 420 mm, de 1960 – 1990 entre 60 – 450 mm, de 1970 – 2000 entre los 37 – 430 mm y finalmente para 1980 – 2010 va de los 17 – 415 mm, en los periodos del 60 al 2000 hubo un aumento en la precipitación mientras que del 2000 al 2010 la precipitación fue menor, al igual que en el periodo de 1950 – 1980, por lo que se puede decir que en la década de los 50 y la década del 2000 fueron más secas (Figura 11).

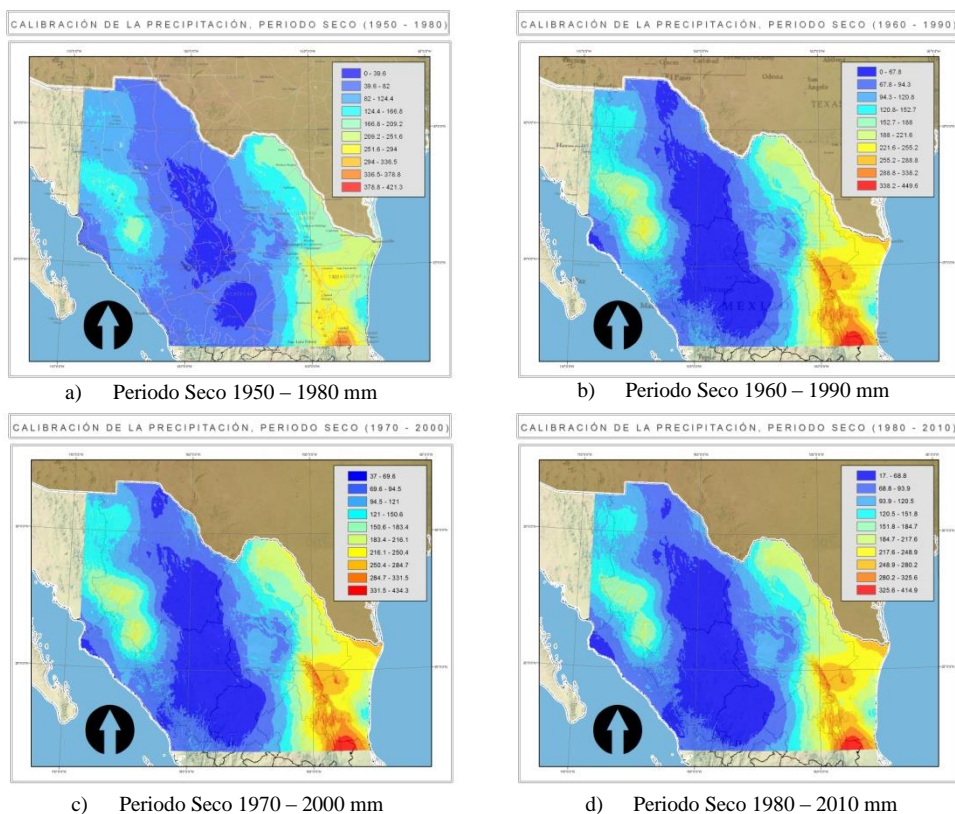


Figura 11. Periodo Seco 1950 – 2010 mm.

Para el periodo húmedo de 1950 – 1980, 1960 – 1990, 1970 – 2000, la precipitación oscila entre 120 – 1500 mm, exceptuando el periodo de 1980 – 2010 donde la máxima precipitación se registró de 1300 mm (Figura 12).

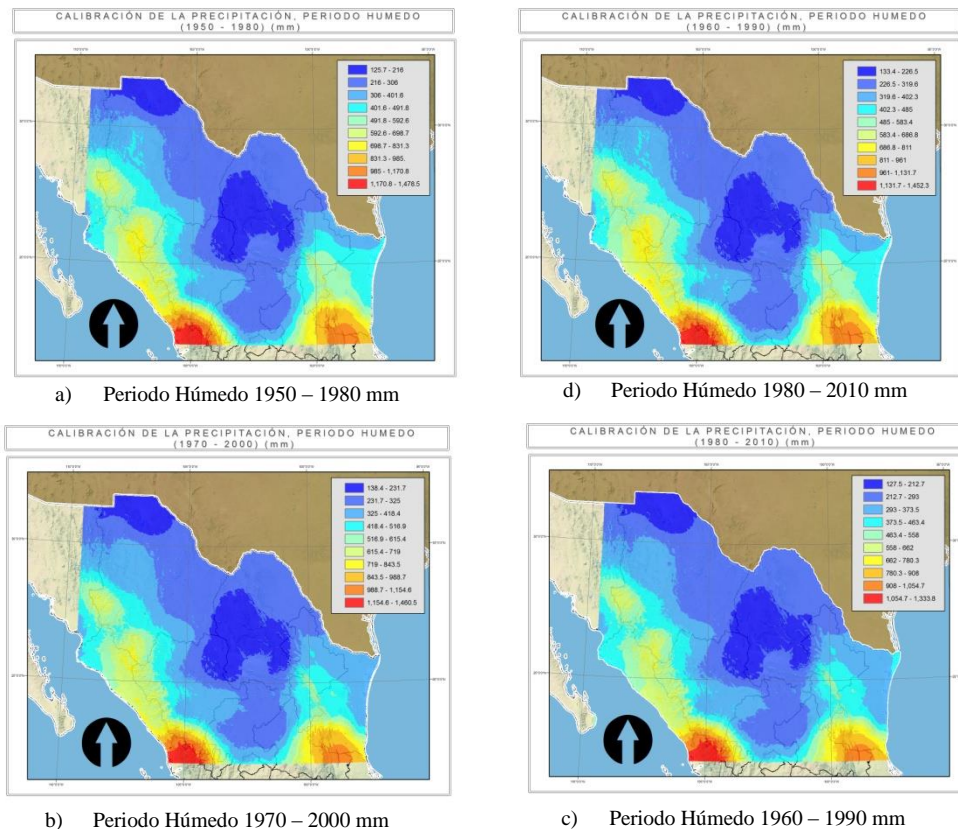


Figura 12. Período Húmedo 1950 – 2010 mm.

6.3.1 Pruebas de validación estadística

La fiabilidad de cada modelo de regresión se llevó a cabo utilizando el 40 por ciento de los datos elegidos aleatoriamente. Los resultados de las pruebas estadísticas obtenidas se describen en el (Cuadro 4) en donde se puede observar rangos de eficiencia comprendidos entre el 0.5 y 0.6 con errores medios absolutos del 18 al 28 por ciento y con coeficientes de eficiencia positivos.

Cuadro 4. Parámetros estadísticos simple y calibrados

	SIN CALIBRAR			CALIBRADO		
	COE	MAE%	MAE	COE	MAE%	MAE
PS-1950-1980	0.53	27.93	29.86	0.92	10.62	11.52
PS-1960-1990	0.57	23.51	29.92	0.92	9.51	12.05
PS-1970-2000	0.60	22.99	28.06	0.92	9.36	11.58
PS-1980-2010	0.60	21.86	26.47	0.93	9.25	10.75
PH-1950-1980	0.46	21.25	101.12	0.92	9.34	38.04
PH-1960-1990	0.54	21.58	96.46	0.91	9.67	39.40
PH-1970-2000	0.54	21.31	93.08	0.90	10.06	39.87
PH-1980-2010	0.53	20.58	80.76	0.90	10.49	36.88
A-1950-1980	0.53	19.99	113.61	0.90	9.78	51.03
A-1960-1990	0.52	19.24	114.24	0.89	9.67	52.09
A-1970-2000	0.52	18.87	109.31	0.88	9.75	50.83
A-1980-2010	0.50	18.37	97.76	0.87	10.16	47.50

*COE= Coeficiente de Eficiencia, RSME = Erro medio cuadrático, MA%E = Valor porcentual del error medio absoluto, MAE = Erro medio absoluto.

En la validación estadística del modelo de regresión sin calibra el coeficiente de eficiencia para el periodo seco oscila entre 0.5 -0.6, para el periodo húmedo de 0.4-0.5 y para el anual de 0.5 a 0.53, el error medio cuadrático para el periodo seco es de 40 a 45, para el periodo húmedo en el rango de 130 – 170, para el anual dentro del rango de 150 – 180, el valor porcentual del error medio absoluto es oscila de 18 a 28 por ciento en todos los periodos, en el periodo seco los errores medios absolutos oscilan de 26 a 30y en el periodo húmedo y anual están por arriba de 90.

En la validación estadística del modelo de regresión calibrado el coeficiente de eficiencia en todos los periodos mejoró en un 40 %, pasando del 50 al 90 %, La R^2 en el modelo calibrado, oscila por arriba de los 0.93 por lo que se considera una buena respuesta del modelo (Cuadro 4).

7.- DISCUSIÓN

A partir de las correlaciones entre variables (Cuadro 2) se observa que la longitud tiene una correlación positiva de 0.43 en promedio en el periodo seco, en los otros periodos (periodo húmedo y el anual) no hay una correlación apreciable. En el período seco, que incluye el invierno, la zona intertropical de convergencia (ZITC) se desplaza hacia el Sur, los vientos alisios actúan solo en esta parte, mientras que en el Norte la influencia proviene de los vientos del Oeste de las latitudes medias. Las depresiones ciclónicas asociadas con estos vientos que atraviesan la parte Norte de México, entonces los Nortes son vientos producidos por las masas de aire frío que se desplazan en invierno, desde Canadá y Estados Unidos, hacia el Sur. La influencia de los ciclones que se originan en relación con las ondas del Este, producen gran parte de la precipitación del verano y principios de otoño.

Según la climatología dinámica, señala la existencia de una “vaguada” de aire húmedo que se presenta comúnmente en verano, sobre la costa del Pacífico. Sands (1960) afirma que esta vaguada de aire húmedo del verano penetra por el Pacífico hacia el Norte, sobre la Altiplanicie Mexicana y la Sierra Madre Occidental, para internarse en los Estados Unidos. Donde es posible que la precipitación que pudiera ocurrir durante este periodo dependa de estas depresiones ciclónicas que se asocian con la longitud, para esta parte del país.

La latitud presentó un aporte negativa únicamente durante el periodo húmedo y en el anual. Por lo que la precipitación en el periodo húmedo y anual es inversamente proporcional la cantidad de lluvia con la latitud, así a mayor latitud menor lluvia y a la inversa, por ello en el periodo seco hay una relación entre la posición latitudinal y la elevación (z) de entre el 0.20 y 0.28 según el aporte con respecto al coeficiente de determinación (R^2) que está en el rango de 0.64 y 0.66. Esto se puede explicar por el desplazamiento en latitud, hacia el Norte de la zona intertropical de convergencia que introduce variaciones anuales de la precipitación (Cuadro 3).

En el periodo húmedo tanto como el anual la elevación tuvo un aporte de 0.02 y distancia a la proximidad de las costas tuvo un aporte de más de 0.64 con respecto al coeficiente de determinación (R^2) que oscila entre 0.69-0.73 se explica porque cuando los vientos que se desplazan hacia el norte u oeste a medida que chocan con las montañas de la sierra madre occidental va descargando la humedad, de igual manera los vientos o corrientes de aire que entran por el golfo de México descargan en las montañas de la sierra madre oriental, esto indica una relación directa de la distribución de la precipitación con respecto a las zonas costeras y la elevación (Cuadro 3).

Para el caso de la elevación (z) focalizado de 5 km y 2.5 km se encontró una correlación negativa y es similar en ambas escalas, siendo mayor para el periodo seco (0.52), para el periodo anual de 0.215, mientras que para el periodo húmedo del 0.065 en promedio, respectivamente.

Esto indica que la elevación promedio a dos escalas presentan una correlación similar (a 5 km y 2.5 km), con respecto a la precipitación y mantiene una relación directa, es decir, llueve más cuando la elevación promedio es mayor, y esta correlación es mayor en el período seco, y menor para el período anual y húmedo, como se indica.

En términos de la desviación estándar de la elevación a diferentes escalas, se observó una aporte promedio del 0.01 aproximadamente con respecto a la R^2 en el período húmedo y anual, es decir cuando hay una mayor variabilidad de la elevación en un área se presenta una mayor precipitación; de igual manera en el periodo seco (Cuadro 3).

En el caso de la distancia euclidiana se encontró una correlación negativa para todos los periodos, es decir hay una mayor precipitación cuando la distancia euclidiana es menor. Para los períodos húmedo y anual la correlación es de -0.8 en promedio y con un aporte del 85 por ciento en cuanto al coeficiente de determinación (R^2) que está en los rangos de 0.65 y 0.73, lo que significa que, entre la precipitación y la distancia euclidiana es alta la correlación pero invertida (mayor cantidad de precipitación a menor distancia euclidiana).

Lo anterior refleja las relaciones en términos de correlaciones (positivas o negativas) y de aportes con respecto al coeficiente de eficiencia entre la precipitación y las variables independientes (latitud, longitud, elevación digital (z) a 5 km, la desviación estándar de (z) a 2.5 km y la proximidad a las costas.

El coeficiente de determinación para los periodos estacionales seco y húmedo está en los rangos de 0.6 – 0.7 (R^2) (Cuadro 3) algunos autores consideran que para fenómenos naturales es aceptable incluso hasta un 0.3 R^2 , se considera que hay una buena calidad en el modelo para replicar sus resultados.

La aplicación del modelo de regresión en el 40 % aleatorio de las estaciones climáticas determinó rangos de eficiencia entre el 0.5 y 0.6 y con errores medios absolutos del 18 al 28 por ciento (Cuadro 4). R^2 se interpreta como la proporción de la variación explicada por la línea de regresión ajustada. El grado de ajuste disminuye a medida que estas estadísticas se alejan de uno. Para la regresión, la línea de mejor ajuste no puede ser peor que $y = \hat{y}$, por lo que R^2 tiene un límite inferior de cero.

La eficiencia tiene un “teórica” límite inferior de infinito negativo, un valor de cero indica que el ajuste $y = \hat{y}$ es igual a la forma de $y = \bar{y}$, con valores de eficiencia menor que cero resultante de un ajuste peor. El uso de esta validación es una extensión de la estadística de R^2 para la clase de modelos de regresión no lineales que no son $y = \bar{y}$, como un posible caso. En estas situaciones, los valores negativos son también posibles.

El coeficiente de eficiencia es de este modo una indicación general de la bondad del ajuste. Cualquier modelo que da un valor negativo no se puede recomendar, con valores deseados cercanos a uno que indica un modelo "casi perfecto". Los valores calculados, aunque muestran errores porcentuales significativos de hasta casi el 30 % se consideran dentro de los límites aceptables, el coeficiente de determinación es de más del 60 % por lo que se considera un buen modelo, además de que la eficiencia no tiene valores negativos, sin embargo el error medio cuadrático si es alto por lo que se calibró el modelo.

La calibración permitió ajustar los parámetros estadístico del 40 % de las estaciones seleccionadas al azar. Los resultados mostraron una mejoría notable indicado por el coeficiente de eficiencia, que mejoró en un 40 %, pasando del 50 al 90 %, de igual manera la R^2 oscila por arriba de los 0.93 por lo que se considera una buena respuesta del modelo (Cuadro 4).

8.- CONCLUSIONES

La modelación de la precipitación por medio de la regresión lineal espacial multivariada resulto una herramienta que permitió entender la distribución espacial y temporal del dicha variable en el Norte de México.

Entre las variables analizadas la elevación y la variabilidad con respecto a su media responde mejor con promedios focalizados a escalas más pequeñas dentro de los rangos de mesoescala. Donde los rasgos geográficos, orográficos y de proximidad a las fuentes de humedad, son factores que tienen influencia directa con la precipitación media anual y estacional.

Dentro del contexto estacional las variables con mayor influencia estadística en la predicción de la precipitación en el periodo seco son: la elevación (focalizado a 5 kilómetros) y la latitud principalmente.

En caso del periodo húmedo que fue similar el comportamiento con respecto al periodo anual, la variable estadísticamente significativa fue la distancia o cercanía a las costas de los puntos de la malla de valores de precipitación.

El proceso de calibración permitió hacer las correcciones de los modelos para mejorar los resultados en un 30 por ciento aproximadamente, dado que al conocer los residuales (entre datos estimados y observados de los puntos de referencia) y al calibrarlos con el resto, se logró observar que está mejor.

La mejora de la capacidad predictiva de los modelos que se resume por el coeficiente de determinación R^2 mayores de 0.93, lo que demuestra que la regresión espacial es más apropiada, al mejorar su eficiencia estadística.

9.- RECOMENDACIONES

Se recomienda realizar el estudio focalizando promedios en escala sinóptica y ver si sigue comportándose del mismo modo. Aplicar la presente metodología en todo el territorio del país utilizando la malla de estaciones que lo cubra por completo. Así mismo realizar una validación de los modelos.

10.- BIBLIOGRAFÍA

- Ahrens CD. (2007). *Meteorology Today: An Introduction to Weather, Climate, and the Environment*, 8th edn. Brooks/Cole-Thompson Learning: Pacific Grove, CA.
- Atlas Digital del Agua México (2012). Sistema Nacional de Información del Agua.
(<http://www.conagua.gob.mx/atlas/ciclo10.html>)
- Apaydin , F. Kemal Sonmez, Y. Ersoy Yildirim (2004). Spatial interpolation techniques for climate data in the GAP region in Turkey. Vol. 28: 31–40.
- Arkin, P. A., and P. Xie. (1994): The Global Precipitation Climatology Project: First Algorithm Intercomparison Project. *Bull. Amer. Meteor. Soc.*, 75, 401–419.
- Arora M, Singh P, Goel NK, Singh RD. (2006). Spatial distribution and seasonal variability of rainfall in a mountainous basin in the Himalayan Region. *Water Resources Management* 20: 489–508.
- Barnes, S. L., 1964: A technique for maximizing details in numerical weather map analysis. *J. Appl. Meteor.*, 3: 396–409.
- Barros AP, Lettenmaier DP. (1993). Dynamic modeling of the spatial distribution of precipitation in remote mountainous areas. *Monthly Weather Review* 121(4): 1195–1214.
- Basist, A., Bell, G.D., Meentemeyer, V. (1994). Statistical relationships between topography and precipitation patterns. *J. Clim.* 7 (9), 1305–1315.
- Benítez, D.H., y Neyra, G. (1997). La biodiversidad de México y su potencial económico. In: *Economía Ambiental; lecciones de América Latina*. Instituto Nacional de Ecología, SEMARNAP: 195-204.

- Bradley, S.G., Dirks, K.N., Stow, C.D. (1998). High resolution studies of rainfall on Norfolk Island, Part III: a model for rainfall redistribution. *J. Hydrol.* 208, 194–203.
- Brammer, R.F. (1989). Unified image computing based on fractals and chaos model techniques. *Opt. Eng.*, 28: 726-734.
- Bratley, P., Fox, B.L. and Schrage, L.E. (1987). *A Guide to Simulation*. Springer-Verlag, New York, NY.
- Camberlin P, Planchon O. 1997. Coastal precipitation regimes in Kenya. *Geografiska Annaler Series A: Physical Geography* 79A: (1–2): 109–119.
- Camberlin P, Okoola RE. 2003. The onset and cessation of the “long rains” in eastern Africa and their interannual variability. *Theoretical and Applied Climatology* 75: 43–54.
- Chen Mingyue, Xie Pingping, Janowiak John E. and Arkin. Phillip A. (2001) Global Land Precipitation: A 50-yr Monthly Analysis Based on Gauge Observations. *Journal of Hydrometeorology*, 3: 246-266.
- Cressman, G. P., 1959: An operational objective analysis system. *Mon. Wea. Rev.*, 87, 367–374.
- Díaz, H. F., R. S. Bradley, and J. K. Eischeid, 1989: Precipitation fluctuations over global land areas since the late 1800s. *J. Geophys. Res.*, 94, 1195–1210.
- Daly, C, Neilson, RP, Phillips, DL. (1994). A statistical topographic model for mapping climatological precipitation over mountainous terrain. *Journal of Applied Meteorology* 33: 140–158.
- Daly C, Gibson WP, Taylor GH, Johnson GL, Pasteris P. 2002. Knowledge-based approach to the statistical mapping of climate. *Climate Research* 22: 99–113.
- Datos climáticos diarios del CLICOM del SMN con gráficas del CICESE (<http://clicom-mex.cicese.mx>)

- Dent J.B., Blackie M.J. (1979) Systems simulation in agriculture, Applied Science Publishers Ltd., London, United Kingdom.
- Diodato N. (2005). The influence of topographic co-variables on the spatial variability of the precipitation over small regions of complex terrain. *International Journal of Climatology* 25: 351–363.
- Eischeid, J. K., H. F. Diaz, R. S. Bradley, and P. D. Jones, 1991: “A comprehensive precipitation dataset for global land areas. U.S. Department of Energy Tech”. Rep. DOE/ER-6901T-H1, 81 pp. [Available from National Technical Information Service, U.S. Dept. of Commerce, Springfield, VA 22161.]
- ESRI. 2012. ESRI® ArcMap™ 10 Copyright© 1995-2012 ESRI Inc. All Rights Reserved.
- Franke, R. 1982. Smooth Interpolation of Scattered Data by Local Thin Plate Splines. *Computer and Mathematics with Applications*. Vol. 8. N.º 4. pág. 273–281. Gran Bretaña.
- Gandin, L. S., 1965: Objective Analysis of Meteorological Fields. Israel Program for Scientific Translations, 242 pp.
- Goodale, C.L., Alber, J.D., Ollinger, S.V. (1998). Mapping monthly precipitation, temperature and solar radiation for Ireland with polynomial regression and digital elevation model. *Clim. Res.* 10, 35–49.
- Goovaerts, P., 2000. Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. *J. Hydrol.* 228, 113–129.
- Hameed, S. (1974). Modelling urban air pollution. *Atmos. Environ.*, 8: 555-561.
- Hession S. L. and Moore N. (2011). A spatial regression analysis of the influence of topography on monthly rainfall in East Africa. MI 48823,

- Hulme M, New M. (1997). Dependence of large-scale precipitation climatologies on temporal and spatial gauge sampling. *Journal of Climate* 10: 1099–1113.
- Hutchinson MF. 1995. Interpolating mean rainfall using thin plate smoothing splines. *International Journal of Geographical Information Systems* 9: 385–403.
- Hutchinson, MF 1998a. Interpolation of rainfall data with thin-plate smoothing splines I: Two dimensional smoothing of data with short range correlation. *Journal of Geographic Information and Decision Analysis* 2: 152–167.
- INEGI. (2012). Derechos Reservados © INEGI. <http://www.inegi.org.mx/>.
- Jorgensen, S.E., Kamp-Nielsen, L., Christensen, T., Windolf-Nielsen, J. and Westergaard, B., (1986). Validation of a prognosis based upon a eutrophication model. *Ecol. Modelling*, 32: 165-182.
- Kieffer Weisse A, Bois P. (2001). Topographic effects on statistical characteristics of heavy rainfall and mapping in the French Alps. *Journal of Applied Meteorology* 40: 720–740.
- Kleijnen, J.P.C. (1987). *Statistical Tools for Simulation Practitioners*. Marcel Dekker, New York, NY.
- Kurtzman, D., Kadmon, R., 1999. Mapping of temperature variables in Israel: a comparison of different interpolation methods. *Clim. Res.* 13, 33–43.
- Kyriakidis PC, Miller NL, Kim J. (2004). A spatial time series framework for simulating daily precipitation at regional scales. *Journal of Hydrology* 297: 236–255.
- McKinion, J.M. and Baker, D.N. (1982). Modeling, experimentation, verification and validation: Closing the feedback loop. *Trans. Am. Soc. Agric. Eng.*, 25: 647-653.

- Marquínez Jorge, Lastra Javier, García Pilar. (2003). Estimation models for precipitation in mountainous regions: the use of GIS and multivariate analysis. *Journal of Hydrology* 270:1-11
- Marquinez J, Lastra J, Garcia P. (2003). Estimation models for precipitation in mountainous regions: the use of GIS and multivariate analysis. *Journal of Hydrology* 270: 1-11.
- Martínez-Cob, A. (1996). Multivariate geostatistical analysis of evapotranspiration and precipitation in mountainous terrain. *J. Hydrol.* 174 (1-2), 19-35.
- Miller, D.R., (1974). Sensitivity analysis and validation of simulation models. *J. Theor. Biol.*,48: 345-360.
- Miller, D.R., Butler, G. and Bramall, L. (1976). Validation of ecological system models. *J. Environ. Manage.*, 4: 383-401.
- Mitas, L. y H. Mitsova. (1988). General Variational Approach to the Interpolation Problem. *Computer and Mathematics with Applications*.16(12): 983-992.
- Nicholson SE. (1996). A review of climate dynamics and climate variability in East Africa. In *The Limnology, Climatology and Paleoclimatology of the East African Lakes*, Johnson TC, Odada, EO (eds). Gordon and Breach Publishers: Amsterdam.
- Ninyerola, M., Pons, X., Roure, J.M. (2000). A methodological approach of climatological modelling of air temperature and precipitation through GIS techniques. *Int. J. Climatol.* 20 (14),1823-1841.
- Oettli P, Camberlin P. 2005. Influence of topography on monthly rainfall distribution over East Africa. *Climate Research* 283: 199-212.
- Oliver, M.A., Webster, R. (1990). Kriging: a method of interpolation for geographical information systems. *Int J. Geogr. Inform.Syst.* 4 (3), 313-332.

- Pardo-Iguzquiza E. (1998). Comparison of geostatistical methods for estimating the areal average climatological rainfall mean using data on precipitation and topography. *International Journal of Climatology* 18: 1031–1047.
- Picard, R.R. and Cook, R.D. (1984). Cross-validation of regression models. *J. Am. Stat. Assoc.*, 79: 575-583.
- Prudhomme, C., Duncan, W.R., 1999. Mapping extreme rainfall in a mountainous region using geostatistical techniques: a case study in Scotland. *Int. J. Climatol.* 19 (12), 1337–1356.
- Rzedowski, J. (1994). *Vegetación de México*. México. Editorial LIMUSA.
- Sands, R. S. (1960). A study in the regional dynamic climatology of México with precipitation as the correlative factor. Adissertation for the Ph.D. Degree. Clark Univ. Woncester, Mass. Trewartha. G.T. 1961. *The earth's problems climates*. The University of Wisconsin Press. Madison
- Santiago de la Fuentes Fernández (2011). “*Regresión Múltiple*”. Madrid España. Facultad de Ciencias Económicas y Empresariales. Universidad Autónoma de Madrid. [pag. 1-10].
- Schaeffer, D.L. (1980). A model evaluation methodology applicable to environmental assessment models. *Ecol. Modelling*, 8: 275-295.
- Sharples JJ, Hutchinson MF, Jellett DR. (2005). On the horizontal scale of elevation dependence of Australian monthly precipitation. *Journal of Applied Meteorology* 44: 1850–1865.
- Siegal J.S. (1972). Development and accuracy of projections of population and households in the United States. *Demography*, 9: 51-68.

- Smith, S.K., (1987). Tests of forecast accuracy and bias for county population projections. *J. Am. Stat. Assoc.*, 82: 991-1003.
- Smith, R.B. (1979). The influence of mountains on the atmosphere. *Adv. Geophys. Academic Press.* 21, 87–230.
- Spreen WC. 1947. A determination of the effect of topography upon precipitation. *Transactions – American Geophysical Union* 28: 285–290.
- Schabenberger O, Gotway CA. 2005. *Statistical Methods for Spatial Data Analysis*. Chapman & Hall: Boca Raton, FL.
- Vargas Hernández, J.J. (2003). Estado de la diversidad biológica de los árboles y bosques en el Norte de México. Documentos de Trabajo: Recursos Genéticos Forestales. FGR/60S. Servicio de Desarrollo de Recursos Forestales, Dirección de Recursos Forestales, FAO, Roma.
- Velázquez José Eduardo. (2014) Calibración de un modelo de estimador de lluvia con imagines de satélite, utilizando datos de estaciones climatológicas, para la region hidrológica número 30 de México. Tesis de ingeniería, Universidad Autónoma de México.
- Vicente-Serrano SM, Saz MA, Cuadrat JM (2003) Comparative analysis of interpolation methods in the middle Ebro valley (Spain): application to annual precipitation and temperature. *Clim Res* 24:161–180.
- Weisse, A.K., Bois, P., 2001. Topographic effects on statistical characteristics of heavy rainfall and mapping in the French Alps. *J. Appl. Meteorol.* 40 (4), 720–740.
- Wotling, G., Bouvier, Ch., Danloux, J., Fritsch, J.-M. (2000). Regionalization of extreme precipitation distribution using the principal components of the topographical environment. *J. Hydrol.* 233: 86–101.

