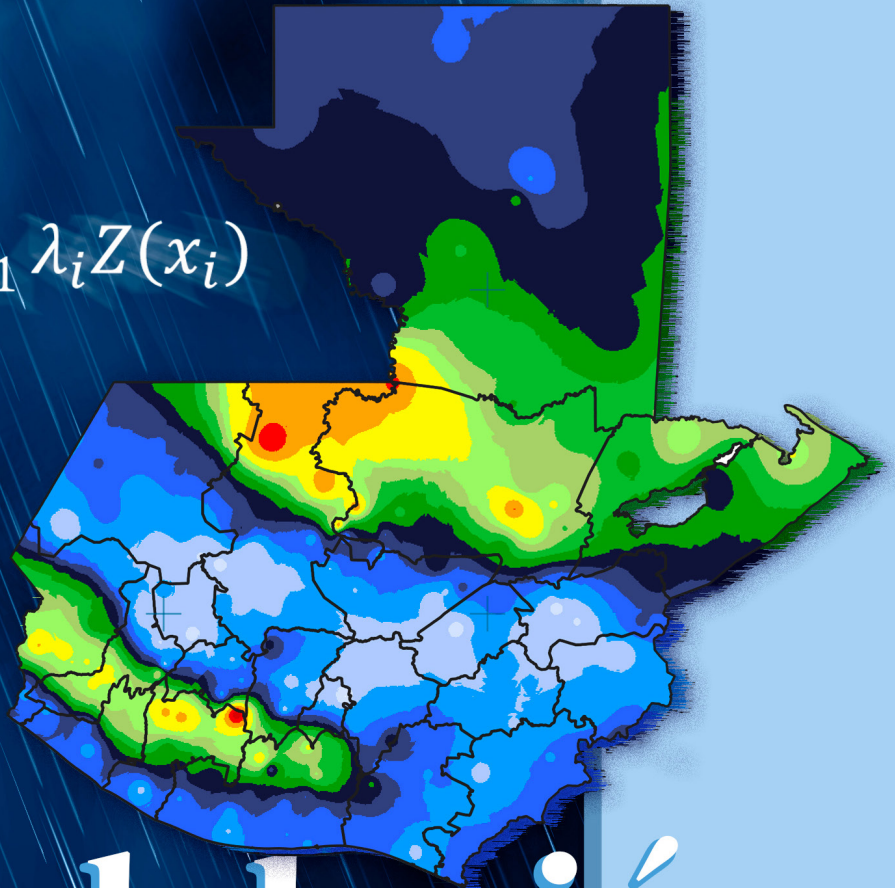


$$Z = \frac{\sum_{i=1}^n \frac{1}{(d_i)^p} Z_i}{\sum_{i=1}^n \frac{1}{(d_i)^p}}$$

$$\check{Z}(x_0) = \sum_{i=1}^n \lambda_i Z(x_i)$$



Modelación Matemática

ANÁLISIS COMPARATIVO DE MODELOS DE INTERPOLACIÓN DE LA PRECIPITACIÓN EN GUATEMALA (1981-2019)

Comparative analysis of precipitation interpolation models in Guatemala (1981-2019)

José Luis Argueta Mayorga

Dr. en Ciudad, Territorio y Sostenibilidad
jlam.argueta@gmail.com

Mayra Virginia Castillo Montes

Dra. en Educación
mayracastillom@yahoo.com

Walter Arnoldo Bardales Espinoza

Mtro. en Recursos Hidráulicos
bardaleswa@gmail.com

William Adolfo Polanco Anzueto

Lic. en Matemática Aplicada
william.polanco.gt@gmail.com

Eugenio Miguel Polanco Sotoj

Estudiante de Ingeniería en Ciencias y Sistemas
eugenio.polanco.ing@gmail.com

Recibido: 30 de noviembre de 2021. | Revisado: 17 de enero de 2022. | Aprobado: 8 de marzo de 2022.

RESUMEN

El estudio del comportamiento de la precipitación es muy importante, pues representa una variable de la cual dependen diferentes fenómenos naturales. A nivel mundial, aún se discute el método de interpolación que mejor representa el fenómeno de la precipitación para cada región geográfica. En particular en Guatemala no se han realizado análisis de los modelos que mejor representan el comportamiento de la precipitación en todo el territorio; por tanto, el presente trabajo consiste en comparar los métodos de interpolación *IDW*, *Co-Kriging*, *Ordinary Kriging*, *Universal Kriging*, *Drifted External Kriging*, *ANUSPLIN* y *Spline*. Como resultados relevantes se establece que los métodos *ANUSPLIN* e *IDW* estiman mejor la precipitación a lo largo del año; sin embargo, se observa que para época con menor precipitación el método *IDW* y *Spline* son los que mejores resultados demuestran, mientras que, para la época con mayor precipitación, los métodos *IDW* y *ANUSPLIN* reportan mejores resultados de estimación.

PALABRAS CLAVE

Modelo de interpolación, precipitación, validación cruzada, Guatemala.

ABSTRACT

The to study the behavior of precipitation since the precipitation is a variable on which different natural phenomena. However, worldwide, the interpolation method that best represents the phenomenon of precipitation for each geographic region is still debated. In Guatemala, there has not been an analyzed or proposal of a method that better represents precipitation throughout the territory; therefore, the present research consists of the comparative analysis between the IDW, Co-Kriging, Ordinary Kriging, Universal Kriging, Drifted External Kriging, ANUSPLIN and spline interpolation methods. The main results obtained, shows that ANUSPLIN and IDW were the methods which better represent the interpolation of precipitation throughout the year; however, it was observed that for the season with less precipitation the IDW and SPLINE Elevation method were the ones that showed the best results and for the season with the highest rainfall the IDW and ANUSPLIN methods showed better results.

KEYWORDS

Interpolation model, precipitation, cross validity, Guatemala

INTRODUCCIÓN

Uno de los fenómenos poco estudiados a nivel global es el de la selección de los métodos de interpolación de la precipitación, así como el método que mejor se adapta al contexto geográfico sobre latitudes y longitudes del globo terráqueo, inclusive existen diferentes factores que aún no se han estudiado a profundidad como: los métodos que actualmente se aplican, las variables que afectan a la precipitación, la topografía, los métodos de validación, los métodos para mejorar las bases de datos, los modelos de elevación digital y su influencia sobre la precipitación.

El fin primordial de los modelos de interpolación radica en estimar la precipitación en lugares donde no se tienen mediciones a partir de lugares donde sí se cuenta con información. Por lo que, el proceso de interpolación parte de una metodología que establece relaciones con distintos fenómenos o comportamientos climáticos por división geográfica, político-social o latitud y longitud.

Prevalece la discusión de qué metodología y método estima mejor la precipitación y su representación para las delimitaciones geográfica, político-social o latitud-longitud que se han planteado como base para delimitar el área de estudio. De esta cuenta, se ha observado que inclusive en una misma región se han identificado métodos de interpolación diferentes que representan mejor la precipitación en dicha región, por lo que, aún se continúa encontrando nuevos hallazgos, nuevos métodos, nuevas bases de datos y nuevas metodologías que como resultado brindan una mejor representación de la precipitación. Por lo que, la aportación al conocimiento sobre el comportamiento de la precipitación es importante, dado que, existen diferentes resquicios y vacíos de conocimiento desde donde se puede aportar; e inclusive se debe particularizar el comportamiento por región o delimitación socio-política, ya que, la precipitación está influida por diferentes factores tanto topográficos como de influencia de otros fenómenos climáticos. También se ve influida por factores como la orografía, elevación, pendiente, distancia hacia el mar, dirección

de los vientos, latitud y altitud, que necesitan de una metodología específica y de una escala temporal adecuada (horaria, diaria, mensual, anual, lustro, década, decalustro u otras).

DESARROLLO DEL ESTUDIO

El estudio es de enfoque cuantitativo, diseño no experimental y carácter comparativo con base en diferentes métricas cuantitativas. La categorización espacial de los estudios para la interpolación de la precipitación se puede realizar con base en la delimitación física-orográfica, política-social, conceptual o de acuerdo con el grado de complejidad física de la formulación de la estructura de análisis de la precipitación (Ly et al., 2013). En el presente estudio se adopta la división política-social, pues se adapta al área de estudio que se enfoca en la República de Guatemala y corresponde a un área de 108,889 kilómetros cuadrados, que se encuentra entre las latitudes de 13°44'19.86" N y 17°48'55.67"N y longitud 88°14'04.55"O y 92°12'38.32"O.

Se combinan aportes de diferentes autores que han utilizado la metodología de la validación cruzada, la cual tiene como base la estadística y la estimación de los errores en los resultados obtenidos por cada método de interpolación que se aplica en la división política-social, puesto que se busca obtener un análisis completo de toda la delimitación espacial que incluyen las estaciones meteorológicas de distintas bases de datos. Por ello, se fundamenta en la literatura de cada método de validación MSE (Price et al., 2000; Taesombat & Sriwongsitanon, 2009; Newlands, et al., 2011; Plouffe et al., 2015; Zhang et al., 2018), MAE (Taesombat & Sriwongsitanon, 2009; Newlands et al., 2011), R2 (Portalés et al., 2008), NSE Error Cuadrático Medio Estandarizado (RMSE) Varianza Equivalente, Kling Gupta E y Nash Sutcliffe.

Se recopila la información mensual de precipitación desde enero de 1981 a diciembre de 2020, para lo cual se integran registros históricos de varias redes de observación que existen en Guatemala: la red climática del Instituto Nacional de Sismología, Vulcanología,

Meteorología e Hidrología -INSIVUMEH-, Instituto Nacional de Electrificación -INDE-, -EMPAGUA-, Instituto Privado de Cambio Climático -ICC- y estaciones fronterizas de los servicios meteorológicos de El Salvador, Honduras, Belice y México (ver figura 1). También, se integran los datos de Rainfall Estimates from Rain Gauge and Satellite Observations (CHIRPS) de la versión 2.0, para completar, extender o rellenar los vacíos de información. El ajuste de los valores ante la ausencia o incongruencia de los datos se realiza a través de la regresión lineal; por tanto, en los espacios carentes de datos o que no poseen alguna consistencia, se reemplazan por valores que provienen de la regresión lineal.

Desde el punto de vista operativo, la variable dependiente es la precipitación medida en milímetros y las variables independientes son la elevación, las distancias de la estación meteorológica hacia el océano Pacífico, océano Atlántico, Golfo de México, la pendiente, longitud y latitud en donde se encuentra la estación meteorológica.

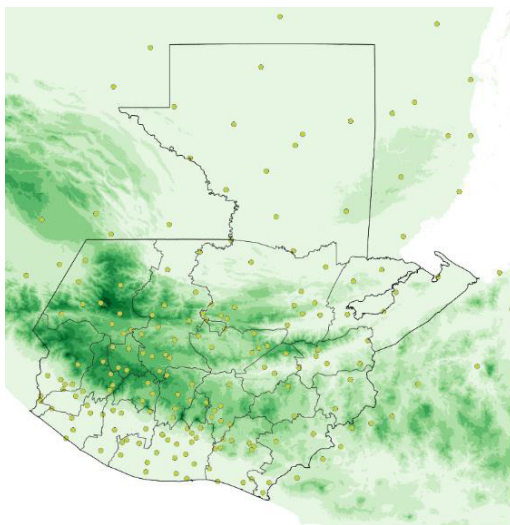


Figura 1. Localización de las estaciones meteorológicas utilizadas en el estudio.

Fuente: elaboración propia.

El set de datos se divide en dos grupos, el 80 % se destina a entrenamiento y el 20 % para validación, se repite el proceso 50 veces de forma aleatoria, de tal

manera que para cada mes se generan 50 grupos de entrenamiento y validación diferentes. Se divide el análisis de la precipitación por cada mes; por lo que, la división temporal permite demarcar la primera escala de tiempo de análisis que se refiere a la precipitación mensual, que hasta el momento no se ha abordado en el contexto de Guatemala.

Los modelos de interpolación que se aplican se describen a continuación.

El método Ponderación Inverso la Distancia (IDW por sus siglas en inglés), es un modelo determinístico de interpolación que se basa en la presunción de que los puntos interpolados se ven más afectados por los puntos cercanos y se afectan menos los puntos más distantes (Wang et al., 2014; Hadi & Tombul, 2018).

$$Z = \frac{\sum_{i=1}^n \frac{1}{(d_i)^p} Z_i}{\sum_{i=1}^n \frac{1}{(d_i)^p}} \quad (1)$$

donde Z es el valor determinado de un punto interpolado; Z_i es el valor de punto recolectado i ($i=1,2, \dots, n$); n es el número de puntos recolectados; d_i es la distancia entre los valores interpolados y recolectados; p representa el parámetro de poder el cual es un número real positivo (Wang et al., 2014). En este estudio se estudió la variación de p de 1 a 4.

Kriging Ordinario (OK en sus siglas en inglés), es un método geoestadístico de interpolación que se basa en la teoría de variograma y análisis estructural, en la cual se utiliza la mejor predicción no sesgada de las variables regionales dadas en un área determinada. El principio del método es que el cambio espacial continuo debe describirse apropiadamente dentro de una superficie determinada, a diferencia de un modelo matemático uniforme porque el cambio espacial es irregular (Xu et al., 2014). Es una estimación lineal no sesgada óptima en los puntos de muestra desconocidos después de considerar la forma, el tamaño, la posición y las relaciones espaciales entre cada par de puntos conocidos y la información de estructura proporcionada por el variograma.

La ecuación de cálculo del método OK se expresa de la siguiente forma:

$$\check{Z}(x_0) = \sum_{i=1}^n \lambda_i Z(x_i) \quad (2)$$

donde $\sum_{i=1}^n \lambda_i Z(x_i)$ es el valor determinado en x_0 y $Z(x_i)$.

El método Co-Kriging toma en cuenta los efectos combinados de múltiples variables que afectan la información espacial y pueden derivarse del Kriging Ordinario. La información se introduce en Co-Kriging como el segundo tipo de factores de influencia (Xu et al., 2014). Si se toma como variable la elevación se establece la ecuación siguiente:

$$Z(x) = \sum_{i=1}^n \lambda_i Z_{ui} + \lambda [y(x) - m_y + m_z] \quad (3)$$

donde $Z(x)$ es el estimado de precipitación por la interpolación en el punto X ; Z_{ui} es la actual medida de precipitación en la estación i ; $y(x)$ es la elevación del punto x ; n es el número de estaciones meteorológicas; λ y λ_i son los coeficientes de ponderación de Co-Kriging; m_y es el promedio de elevaciones y m_z es el promedio de atributos meteorológicos.

El método de Kriging con desviación externa es una variante del Kriging universal para la cual la variable $Z(x)$ a interpolar es estimado como la suma de un término estocástico $y(x)$ y un término determinístico $m(x)$ (Delrieu et al., 2014), también denominado la desviación:

$$Z(x) = Y(x) + m(x) \quad (4)$$

donde x denota una posición dada. En la aplicación del método de Kriging con desviación externa, se establece que la desviación tiende a ser una función lineal de $R(x)$. Un estimador lineal $Z_{KED}(x_0)$, corresponde a la suma ponderada de las medidas de las estaciones $G(x_i)$, que es usado para la estimación en un punto dado o sobre un dominio X_0 dado:

$$Z_{KED}(x_0) = \sum_{i=1}^n \lambda_i^{KED} G(X_i) \quad (5)$$

La ponderación λ_i^{KED} se estima para minimizar la estimación de la desviación estándar:

$$\sigma_{KED}(x_0) = \sqrt{E(Z(x_0) - Z_{KED}(x_0))} \quad (6)$$

$$\text{con } E(Z(x_0)) = E(Z_{KED}(x_0)) \quad (7)$$

El método Spline se aplica para suavizar una superficie continua minimizando la curvatura (Taesombat & Sriwongsitanon, 2009). Para un planteamiento bivariado de Spline para medir los puntos $i=1 \dots N$, se calcula con base en Hadi & Tombul (2018):

$$Z(s_i) = f(s_i) + \epsilon(s_i) \quad (8)$$

donde $\epsilon(s_i)$ representa los errores aleatorios, los cuales pueden ser asumidos a ser errores aleatorios no correlacionados e independiente con un promedio de cero y una varianza σ^2 y $f(s_i)$ representa una función suavizadora determinística, que se estima por minimización:

$$\sum_{i=1}^n [Z(s_i) - f(s_i)] + \lambda J_m^d \quad (9)$$

donde $f(s_i)$ son las cantidades correspondientes a la función que encaja en cada uno de los puntos i ; λ es el también llamado regularizador o parámetro y J_m^d es una medida de lo suavizante de la función f . La forma J_m^d depende de dos parámetros: el número de variables independientes d y el orden de las derivadas m .

El método *Anusplin* se basa en el suavizado del *Spline*, usualmente con al menos dos variables independientes, longitud y latitud, en este orden y en unidades de grados decimales. La tercera variable independiente, la elevación sobre el nivel del mar, es normalmente apropiada cuando se ajustan superficies a temperatura o precipitación.

El rendimiento y eficacia de cada una de las técnicas de interpolación se puede evaluar mediante la comparación del punto observado, que se conoce como validación cruzada (Seaman, 1988). Las técnicas de evaluación y validación de desempeño son ampliamente utilizadas por geoestadísticos para evaluar los métodos de interpolación espacial (Hijmans et al., 2005; Hancock & Hutchinson, 2006). Con base

en se realizaron los análisis de error absoluto medio (MAE), Error Cuadrático Medio Estandarizado (RMSE), MSE, Varianza Equivalente, Kling Gupta E, Nash Sutcliffe, que se incluyen a continuación.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \tilde{Y}_i| \quad (10)$$

donde y_i es un valor observado en la posición i y \tilde{Y}_i es un valor estimado para la posición i , n es el número de puntos de datos.

El Error Cuadrático Medio determina la diferencia cuadrática media entre la verdadera precipitación y su estimado (Di Piazza et al., 2011):

$$MSE = \frac{1}{N_v} \sum_{i=1}^{N_v} |z(x_i) - \check{Z}(x_i)|^2 \quad (11)$$

donde $\check{Z}(x_i)$ es el valor estimado de lluvia y $z(x_i)$ es el valor de lluvia medido en N_v puntos de validación a analizar.

Nash Sutcliffe Efficiency (NSE):

$$NSE = 1 - \frac{\sum_{i=1}^N (O_i - M_i)^2}{\sum_{i=1}^N (O_i - \bar{O})^2} \quad (12)$$

donde O_i es el valor de punto observado i , M_i es el valor estimado en i , y N es el total de número de observaciones y \bar{O} es el promedio del valor observado (Wagner et al., 2012; Chen et al., 2017;). $NSE = 1$ indica correspondencia perfecta entre las simulaciones y observaciones; $NSE=0$ indica que las simulaciones del modelo tienen el mismo poder explicativo como el promedio de las observaciones; y $NSE < 0$ indica que el modelo es un mal predictor de las observaciones (Knoben et al., 2019).

Kling Gupta Efficiency, representa la correlación, el sesgo y similitud en la variabilidad entre los datos observados y cuadrículados (Nashwan et al., 2019).

$$KGE = 1 - \sqrt{(r - 1)^2 + \left(1 - \frac{\mu_{sim}}{\mu_{obs}}\right)^2 + \left(\frac{\frac{\sigma_{sim}}{\mu_{sim}}}{\frac{\sigma_{obs}}{\mu_{obs}}}\right)^2} \quad (13)$$

donde r es la correlación de Pearson; μ y σ representan el promedio y la desviación estándar, respectivamente, de los datos simulados (sim) y observados (obs); n se refiere al número de puntos dentro de la grilla.

El factor R^2 es denominado la varianza esperada, estima el grado de predicción de un modelo, por definición es un número menor o igual a 1, y entre más se acerque a 1, mejor es la forma en cómo encajan los datos calculados y medidos (Caruso & Quarta, 1998).

$$R^2 = 1 - \frac{\sum_{i=1}^{N} (v_i^c - v_i^0)^2}{\sum_{i=1}^{N} (v_i^c - \bar{v}_i^0)^2} \quad (14)$$

donde v_0 son los valores observados y v_c son los valores calculados, y \bar{v} indica el promedio de los valores.

La interpretación de la varianza explicada es una medida de precisión basada en los valores predichos que se derivaron de la validación cruzada. (Li, 2016).

$$VEcv = \left(1 - \frac{SSD}{SST}\right) * 100(\%) = \left(1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}\right) * 100(\%) \quad (15)$$

$$\text{con } SSD = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad \text{y} \quad SST = \sum_{i=1}^n (y_i - \bar{y})^2 \quad (16)$$

RESULTADOS OBTENIDOS

El comportamiento de la precipitación en Guatemala se representa en Figura 5. En el periodo de enero a abril se observa que las precipitaciones prevalecen en el norte del país en dirección hacia el Este, lo que denota una fuerte influencia del océano Atlántico. Asimismo, sobre la región donde prevalece la lluvia en el mes de enero, no supera los 200 milímetros de lluvia y lo interesante es la identificación de las áreas montañosas como factores que demarcan y delimitan las áreas de comportamiento de la precipitación.

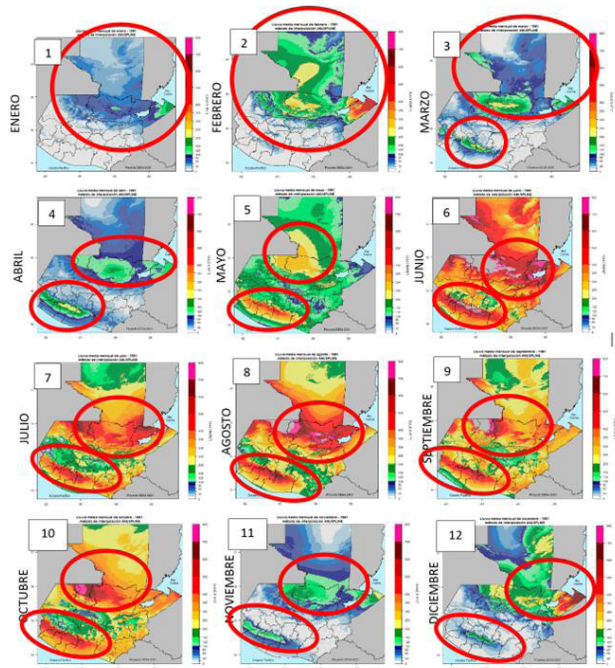


Figura 2. Método de interpolación Anusplin y comportamiento de la precipitación de enero a diciembre de 1981.

Fuente: elaboración propia.

Durante la época de enero a mayo se observa una tendencia de precipitación al Norte que se inclina hacia el Oeste del país, siempre demarcado por la cordillera norte que colinda con la Franja Transversal del Norte. Por otro lado, la precipitación al sur del país se concentra en la franja volcánica y la cordillera que divide el país con el lado del océano Pacífico.

En el caso de la precipitación de junio, julio, agosto, septiembre y octubre se puede observar que a nivel de todo el país se incrementa considerablemente la precipitación; no obstante, se observa cómo la precipitación se centra en dos regiones muy marcadas que en los meses anteriores también registran altas precipitaciones. En el caso de noviembre y diciembre las precipitaciones del país disminuyen, pero se sigue observando en menor medida influencia de la precipitación y la delimitación de la precipitación a partir de las cordilleras y montañas que delimitan el norte y sur del país.

Con relación a la métrica Kling Gupta E y la interpolación de la lluvia mensual que se muestra en la Figura 3, se observa cómo el método ANUSPLIN denota ser el método que mejor ajuste tiene a lo largo del año. El método IDW y el método Spline Simple denotan ser los métodos que le siguen con base esta métrica. Se observa que para los meses de la época con más lluvia denotaron menores variaciones de la métrica Kling Gupta E y para los meses de la época con menos lluvia denotan comportamientos mucho más variables. El método que denota menor variación de la métrica Kling Gupta E a lo largo del año es el método IDW. Los métodos KED y UK demuestran mejor resultado que el método OK, aunque la variación no es considerable en comparación de esos métodos.

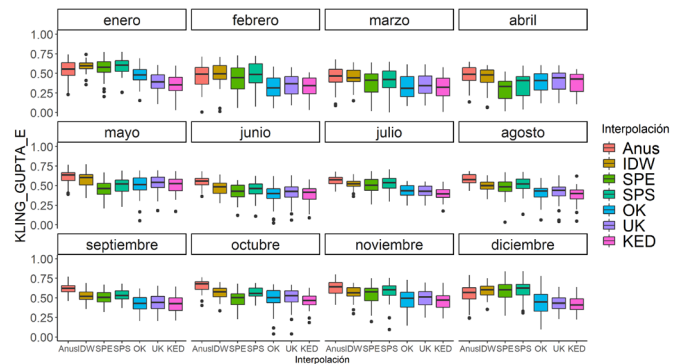


Figura 3. Métrica Kling Gupta E y la interpolación de lluvia mensual de enero a diciembre, métodos Anusplin, IDW, SPE, SPS, OK, UK y KED.

Fuente: elaboración propia.

Para la métrica MAE, con los métodos de interpolación para la precipitación mensual (Figura 4), se observa que en la época con mayor precipitación (de mayo a octubre) los métodos Anusplin e IDW, son los que presentan menor variación y mantienen los valores estimados más bajos. Los métodos Spline simple y Spline elevación se comportan de manera parecida, es decir, varían en magnitudes con valores más altos que Anusplin e IDW. Los valores estimados por los métodos OK, KED y UK se encuentran por arriba de los generados por Anusplin e IDW, pero por debajo de los valores Spline (elevación y simple). Se observa también un patrón que, en cada mes de esta época,

aumentan los valores de los residuos, siendo más bajos al inicio de la época lluviosa-mayo y terminando con valores notables altos al final de la época lluviosa-octubre, que se puede interpretar como un aumento de la variabilidad en la precisión de los métodos empleados.

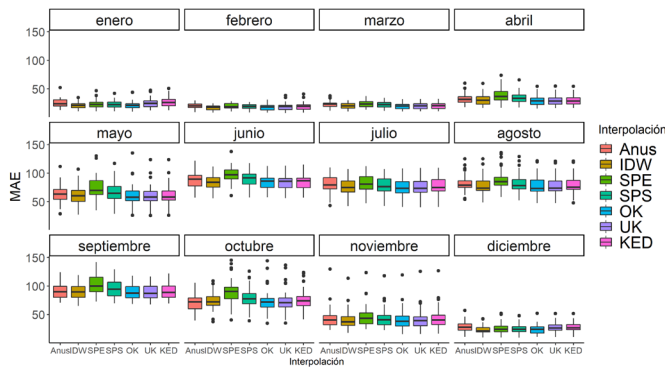


Figura 4. Métrica MAE y la interpolación de lluvia mensual de enero a diciembre, con Anusplin, IDW, SPE, SPS, OK, UK y KED.

Fuente: elaboración propia.

Para la época con menor lluvia (de noviembre a abril) se identifica con base en la métrica MAE, que ocurre lo contrario con relación al comportamiento de los residuos en todos los métodos de interpolación aplicados, las magnitudes disminuyen y presentan valores altos al inicio de la época y termina con valores bajos al finalizar (abril). Las métricas indican que los métodos Anusplin e IDW son los que presentan los valores más bajos en los residuos, aunque con poca diferencia de los demás métodos, el resultado del patrón de cada uno de los métodos denota poca variabilidad en su precisión y se caracterizan por contar con valores muy bajos en los residuos y muy parecidos.

Con relación a la métrica MSE se observa diferente comportamiento para la época con menor precipitación (Figura 5), lo que indica que disminuye el MSE y para la época lluviosa se observa comportamientos diversos en cuanto a la variación de la precipitación. Los métodos de interpolación Anusplin e IDW presentan menor variación para la época lluviosa y se perfilan como los métodos que menor magnitud o valor reportan. Los

métodos Spline Elevación y Spline simple presentan mayores valores en residuos que IDW y Anusplin. La distribución de los residuos en los métodos OK, KED y UK en algunos meses (junio, julio, agosto) presentan una distribución de valores muy parecidos a los del método Spline elevación y Spline simple, aunque la tendencia en los residuos generados por los distintos métodos es aumentar desde el inicio de la época con mayor precipitación (mayo hasta septiembre), con un descenso en el mes de octubre.

Las métricas indican que el método IDW presenta los valores más bajos en los residuos, aunque en los meses de noviembre y diciembre los valores presentan valores más extremos para todos los métodos.

Por último, se observa muy poca diferencia entre los métodos en los meses de febrero y marzo.

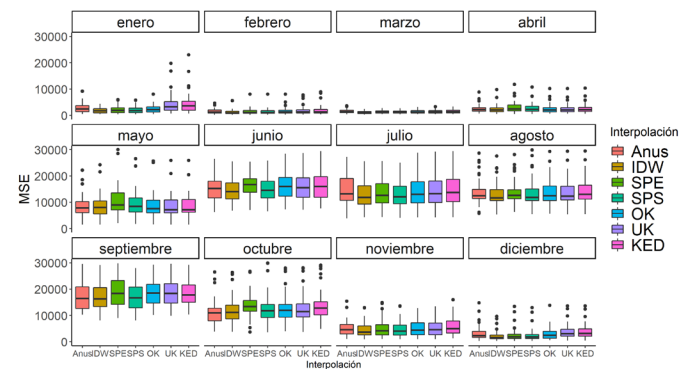


Figura 5. Métrica MSE y la interpolación de lluvia mensual de enero a diciembre de los métodos Anusplin, IDW, SPE, SPS, OK, UK y KED.

Fuente: elaboración propia.

Con respecto a la métrica Máxima Varianza Explicada-VEcv, indica que los valores cercanos al 100 %, sugieren predicciones perfectas (Figura 6); un 0 % indica que las predicciones son tan precisas como usar el promedio de los datos de validación y los valores negativos indican que las predicciones son menos precisas que usar el promedio para predicciones (Li, 2016). Para la época con mayor precipitación (de mayo a octubre) se observa que el método IDW es el que mejor predice la precipitación, seguido por el método

Anusplin, que en el mes de octubre cae en el rango de bueno, mientras IDW no. En la época con considerable precipitación ningún método de interpolación supera el rango de predicciones promedio y los métodos OK, KED y UK presentan valores extremos muy bajos hasta caer al rango de pobre y muy pobre (junio, agosto y octubre).

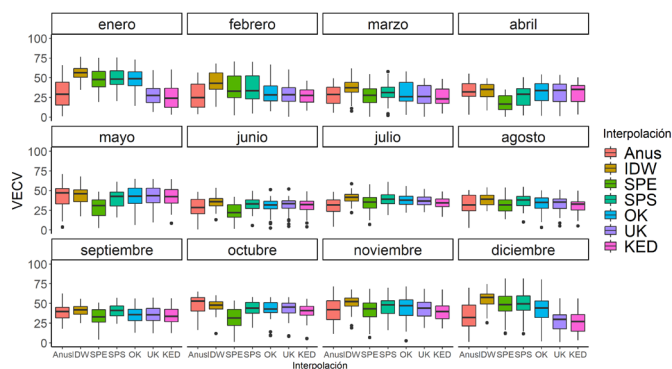


Figura 6. Métrica VECv y la interpolación de lluvia mensual de enero a diciembre de los métodos Anusplin, IDW, SPE, SPS, OK, UK y KED.

Fuente: elaboración propia.

Para la época con reducida precipitación (noviembre a abril) se obtiene que el método IDW es el que posee el porcentaje más alto; luego, se ubica en el rango de bueno para los meses de noviembre, diciembre y enero. Durante los meses de febrero, marzo y abril baja su rango promedio; mientras que el método Anusplin mantiene el rango promedio.

Además, los métodos Spline elevación, Spline simple y OK se comportan de manera similar en el rango inferior del promedio y los métodos UK y KED se ubican en el rango de pobre.

Sobre la distribución del coeficiente de determinación R², para la época lluviosa (de mayo a octubre) se calculan valores entre 0.4 y 0.6 en los métodos Anusplin e IDW (Figura 7), es decir, con estos métodos se podría replicar los resultados con la proporción de variación de los mismos de un 40 % hasta un 60 % o también interpretarse como la que la proporción de la varianza explicada por éstos métodos se ajusta a la distribución de los datos entre un 40 % a 60 %. Los métodos que le

siguen son OK, UK y Spline (elevación y simple) en los meses de junio y julio, con alrededor de un 40 % de la proporción de la varianza explicada. También se observa en esta época una tendencia en la disminución de este coeficiente, empezando por mayo y terminando en septiembre, ya que en el mes de octubre presenta un aumento cercano al 0.6.

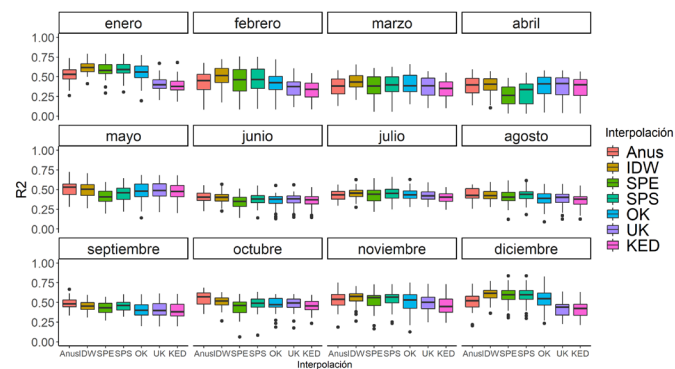


Figura 7. Métrica R² y la interpolación de lluvia mensual de enero a diciembre de los métodos Anusplin, IDW, SPE, SPS, OK, UK y KED.

Fuente: elaboración propia.

El método IDW presenta los valores más altos, mayores de 0.6 en noviembre, también en diciembre y enero se observa mayor presencia en sus valores extremos; para los meses de febrero, marzo y abril varía entre 0.4 y 0.5. Los demás métodos (OK y SP-elevación y simple) tienen coeficientes parecidos para los meses noviembre, diciembre y marzo (entre 0.35 y 0.6), también se observa mayor variabilidad en todos los métodos para los meses de febrero, marzo y abril; es decir, la precisión de los métodos oscila más en estos meses.

Con base en la métrica NSE se puede observar que el método IDW sobresale considerablemente (Figura 8), seguido de los métodos Spline elevación y Spline simple, luego en una categoría de promedio el método Anusplin y entre pobre y muy deficiente estarían los métodos UK, OK y KED. Mientras que para la época con menor precipitación se observa que el método IDW sobresale y tiene un comportamiento similar a lo largo del año; por otro lado, en la época lluviosa el método Spline elevación baja a la categoría de promedio junto con Anusplin, mientras que el Spline simple se

mantiene en la categoría de bueno. Los métodos OK, UK y KED se mantienen entre la calidad de pobre a muy deficiente

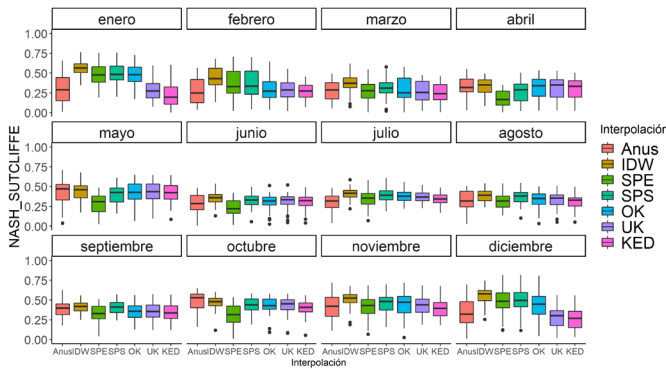


Figura 8. Métrica NSE y la interpolación de lluvia mensual de enero a diciembre de los métodos Anusplin, IDW, SPE, SPS, OK, UK y KED.

Fuente: elaboración propia.

De acuerdo con los métodos de validación se puede observar que las técnicas de interpolación que mejores resultados estiman son los métodos IDW y Anusplin (Tabla 1), seguidos por los métodos de Spline simple y Spline elevación que no generan tan bajos resultados y se encuentran dentro de la categoría de bueno; mientras que los métodos OK, UK y KED demostraron ser los métodos de interpolación con muy deficiente resultado..

Tabla 1 Validación y método de interpolación en los doce meses del año. 1981 a 2019.

Método de validación	Método de interpolación efectivo a lo largo del año				
	Excelente	Bueno	Promedio	Pobre	Muy deficiente
MAE	Anusplin, IDW	Spline Simple Spline Elevación		OK, UK	KED
MSE	Anusplin, IDW	Spline Simple Spline Elevación		KED	OK, UK
NASH SUTCLIFFE	IDW	Spline Simple Spline Elevación	Anusplin	OK	UK, KED
KLING GUPTA	Anusplin	Idw, Spline Simple Elevación	Spline		OK, UK, KED
R ²	IDW, Anusplin	Spline Simple Elevación	Spline	OK	UK, KED
Vecv	Anusplin	Spline Simple Elevación	Spline	OK	UK, KED

Fuente: elaboración propia.

Se observa para la época con menor precipitación que el método que sobresale dentro de una categoría de excelente es IDW (Tabla 2); mientras que los métodos Spline simple, Anusplin, OK y Spline elevación tienen un comportamiento entre bueno y promedio; no obstante, los métodos OK y KED demuestran comportamientos entre pobre y muy deficiente. Es importante resaltar que para la época sin tanta precipitación el método Anusplin, que toma como base la elevación, no sobresale de los demás métodos.

Tabla 2 Validación y método de interpolación en época de precipitación intensa. 1981 a 2019.

Método de validación	Método de interpolación efectivo en época precipitación intensa				
	Excelente	Bueno	Promedio	Pobre	Muy deficiente
MAE	IDW, Anusplin	OK, UK, KED	SPS		SPE
MSE	IDW	Anusplin	Spline simple	Spline elevación, UK	OK, KED
NASH SUTCLIFFE	IDW	Spline simple	Spline elevación, Anusplin	OK, UK	KED
KLING GUPTA	Anusplin	IDW, Spline simple	Spline elevación	OK, UK	KED
R ²	Anusplin	IDW	Spline simple, Spline elevación	OK, UK	OK, UK, KED
Vecv	IDW	Spline simple, Anusplin	OK, UK	KED	SPLINE ELEVACIÓN

Fuente: elaboración propia.

El método IDW es el que sobresale para la época seca y lluviosa, así como para todo el año, por lo que es importante profundizar en el número de estaciones vecinas, el número de potencia y la base de datos de las estaciones que generan los mejores resultados. En ese sentido, para el análisis de la interpolación se establecen diferentes números de estaciones vecinas que van de la 5 a la 20 y diferentes números de potencia que varían desde el 1.2 a 4.0.

En la Figura 9 se observa la frecuencia con los mejores resultados del número de estaciones vecinas por mes de interpolación, a partir de lo cual se establece que la

cantidad de estaciones vecinas debe estar entre 15 a 20, mientras que en la época con más precipitación (mayo a octubre) la cantidad de varía de 8 a 10.

Al seleccionar los valores de IDP (Figura 10), se concluye que para los meses de época con menor precipitación (noviembre a abril) el mejor valor es 1.4, mientras que para los meses de mayo, junio y octubre estos valores oscilan entre 1.4 a 1.8, y para los meses de julio a septiembre los valores con mejores resultados para potencia oscilan de 2 a 2.4.

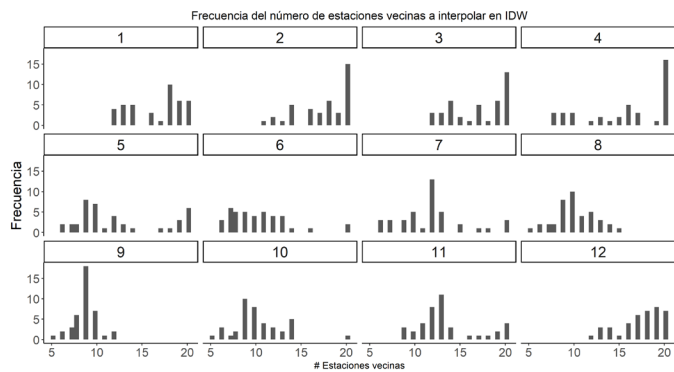


Figura 9. Vinculación de las estaciones vecinas por mes de análisis y la frecuencia en relación de las estaciones vecinas cercanas.

Fuente: elaboración propia.

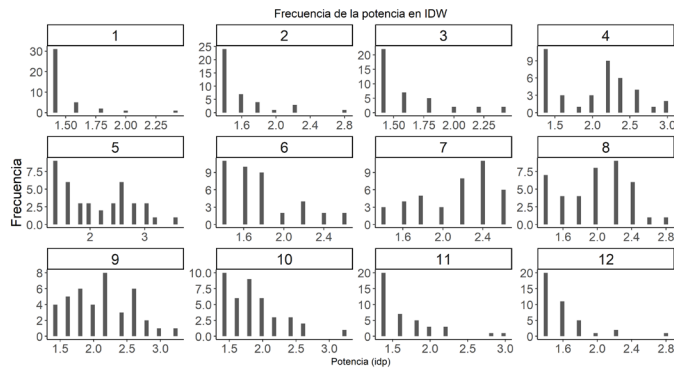


Figura 10. Vinculación de la potencia para cada mes de análisis y la frecuencia de la potencia por cada mes.

Fuente: elaboración propia.

DISCUSIÓN DE RESULTADOS

Se observa en la exploración de variables independientes tales como la distancia hacia el

océano Atlántico, distancia hacia el océano Pacífico, distancia hacia el Golfo de México, elevación, latitud, longitud y pendiente, la relación con la variable dependiente precipitación, por lo que se identifica que existe necesidad de distribuirla en función de una regionalización. Es decir, en el proceso de exploración se observa la influencia en los registros de precipitación en las estaciones cercanas al océano Pacífico, pero dicha influencia varía en lugares de convergencia o donde el océano Atlántico podría influir considerablemente. Por lo que, una de las vertientes a futuro es explorar estadísticamente la base de datos que se depuró, la influencia en las diferentes áreas del país, que puede generar ecuaciones de correlación entre las variables dependientes e independientes. También, en el proceso de regionalización sería importante la consideración de modelos multivariantes que permitan profundizar en el comportamiento estadístico de la precipitación en cada región o de cada distribución.

La metodología que la presente investigación aplica es un aporte al conocimiento, puesto que abarca la metodología seguida para integrar una base de datos, establecer datos para la creación del set de validación cercano al 20 % de la data disponible, mientras que el 80 % se destina como set de entrenamiento que permite evaluar el comportamiento al aplicar cada uno de los modelos de interpolación y los métodos de validación estadística, para definir el método que mejor representa la precipitación en el territorio de la república de Guatemala.

Otro de los aportes importantes de la presente investigación son los mapas de interpolación que se elaboran para los doce meses del año y a partir de los años 1981, 1986, 1991, 1996, 2001, 2006, 2011 y 2016; se interpola aplicando los métodos Anusplin, Spline elevación e IDW. Los tres modelos de interpolación seleccionados para los mapas corresponden a los tres métodos que a lo largo del año demostraron mejores bondades de ajuste. Y la realización de la base de mapas correspondientes a distintos métodos de interpolación constituyen uno de los productos importantes de la investigación, puesto que, con los mapas se pueden

realizar diferentes análisis en investigaciones futuras que para estimar o explicar comportamientos a partir de diferentes escalas temporales.

De acuerdo con los métodos de validación a lo largo del año se puede observar que las técnicas de interpolación que mejores resultados generan son los métodos IDW y Anusplin, seguidos por los métodos de Spline simple y Spline elevación como los que métodos que se encuentran dentro de la categoría de bueno; mientras que los métodos OK, UK y KED demostraron ser los métodos de interpolación con muy deficiente resultado.

Por otro lado, es importante destacar la temporalidad de la precipitación y los métodos que puedan mostrar tendencia en relación con una alta y baja precipitación. Se observa que los métodos que sobresalen dentro de una categoría de excelente para la época con mayor precipitación son: IDW y Anusplin; mientras que para el método Spline simple y para Spline elevación, se caracteriza un comportamiento entre bueno y promedio; no obstante, los métodos OK, UK y KED demuestran comportamientos entre pobre y muy deficiente. Sin embargo, cuando se presenta una menor precipitación el método que sobresale dentro de una categoría de excelente es IDW mientras que los métodos Spline simple, Anusplin, OK y Spline elevación denotan un comportamiento entre bueno y promedio. Es importante resaltar que para la época con precipitación media o baja, el método Anusplin que toma como base la elevación, no sobresale de los demás métodos.

Dado que el método IDW es el de mejor ajuste en distintas épocas de intensidad de lluvia en el territorio de la república de Guatemala, se evidencia que para la época con menor precipitación (noviembre a abril) la cantidad de estaciones vecinas debe estar entre 15 y 20, mientras que en los meses con la época con mayor precipitación (mayo a octubre) la cantidad de estaciones vecinas necesarias oscila entre 8 y 10.

Para el método IDW es necesario considerar valores de potencia (IDP) de acuerdo con la época de mayor o

menor precipitación, se concluye que para los meses de época con menor precipitación (noviembre a abril) el mejor valor es 1.4, mientras que para los meses de mayo, junio y octubre estos valores oscilan entre 1.4 a 1.8, y para los meses de julio a septiembre los valores con mejores resultados para potencia oscilan entre 2 y 2.4.

Se identifica que la variable elevación se puede asociar con un mejor ajuste en modelos como el Anusplin y el Spline elevación; es decir, las variables ambientales agregan un comportamiento importante y vital en el desarrollo de la interpolación de la precipitación.

Por lo anterior, se recomienda explorar los análisis con otras variables ambientales como las distancias a los océanos o cuerpos de agua importantes que inciden en la variabilidad climática en Guatemala, las pendientes, usos de suelo, regionalización por microclimas, latitud, longitud, áreas boscosas, entre otros.

Al comparar los mapas de interpolación se puede observar el suavizado de la interpolación de la precipitación con el método Anusplin; es decir, la base del método Anusplin depende de la elevación y eso influye en la interpolación de la precipitación, ya que se adapta a la superficie existente. Lo mismo ocurre con el método Spline elevación ya que su interpolación depende de la elevación de cada uno de los lugares y se adapta a la superficie del territorio de Guatemala. Mientras que el método IDW se asocia con la localización de las estaciones, por lo que se observa que la interpolación tiene el comportamiento de anillos o círculos que demarcan áreas de precipitación; por lo que, en los anillos de interpolación del método IDW se pueden encontrar vacíos debido a carencia de estaciones en lugares donde puede existir otro comportamiento de la precipitación.

Con relación a los mapas de interpolación del método IDW sería importante observar los puntos de anillos que identifican comportamientos de precipitación aislados y nutrir la base de datos a partir de otras estaciones meteorológicas que permitan brindar continuidad al

mapa IDW o bien, que se pueda adecuar o alinear los datos, previo a utilizar el método IDW.

Finalmente, el análisis comparativo de los modelos de interpolación de registros históricos de la precipitación en el territorio de Guatemala, identifica potencialidades y limitaciones de los métodos más utilizados con base en métricas cuantitativas; lo que genera aportes para futuras investigaciones acerca del tema.

Se agradece el apoyo de la Dirección General de Investigación de la USAC, para la realización del estudio.

CONCLUSIONES

1. Entre los métodos que mejor comportamiento en la predicción de la precipitación se obtiene el método IDW y el Anusplin, en consideraciones del contexto geográfico sobre latitudes y longitudes del globo terráqueo, la topografía y su influencia sobre la precipitación para la base de datos, territorio y consideraciones planteadas en este estudio sobre la precipitación en el territorio de Guatemala. Esta conclusión se aplica tanto para época con precipitación intensa, como la que no posee dichas condiciones.
2. Del método de IDW se recomienda la utilización de valores de potencia 1.4 a 1.8 para los meses de mayo junio y octubre; de 2 a 2.4 para los meses de julio a septiembre y de noviembre a abril de 1.4. Estos valores permiten hacer eficiente el trabajo con los números de estaciones vecinas que entre noviembre y abril debe estar entre 15 y 20 y de mayo a octubre de 8 a 10. Por tanto, el número de vecinos y potencia permiten que el método IDW presente mejor ajuste en la determinación de los vacíos de información de precipitación en el territorio de la República de Guatemala
3. Uno de los ricos aportes de esta investigación es la conjugación de diferentes bases de datos, provenientes de diferentes instituciones y que se procesan para eliminar errores como valores

cero e incongruencias. Por lo que, la metodología empleada para conformar una base de datos congruente es el aporte que esta investigación plantea.

RECOMENDACIONES

1. Una de las vertientes a futuro es explorar estadísticamente la base de datos depurada, la influencia en las diferentes áreas del país, que puede generar ecuaciones de correlación entre las variables dependientes e independientes.
2. También, el proceso de regionalización de las estaciones o valores de precipitación sería importante la consideración de modelos multivariantes que permitan profundizar en el comportamiento estadístico de la precipitación en cada región o de cada distribución.
3. Se recomienda continuar con el trabajo de comparar nuevos métodos de interpolación que han surgido o se han vuelto eficientes, que permitan tomar en cuenta las formas en cómo se validan los resultados y se integran o establece la congruencia de la base de datos.

REFERENCIAS BIBLIOGRÁFICAS

- Caruso, C., & Quarta, F. (1998). *Interpolation methods comparison. Computers and Mathematics with Applications*, 35(12), 109–126. [https://doi.org/10.1016/S0898-1221\(98\)00101-1](https://doi.org/10.1016/S0898-1221(98)00101-1)
- Chen, D., Ou, T., Gong, L., Xu, C. Y., Li, W., Ho, C. H., & Qian, W. (2010). Spatial interpolation of daily precipitation in China: 1951-2005. *Advances in Atmospheric Sciences*, 27(6), 1221–1232. <https://doi.org/10.1007/s00376-010-9151-y>
- Chen, T., Ren, L., Yuan, F., Yang, X., Jiang, S., Tang, T., Liu, Y., Zhao, C., & Zhang, L. (2017). Comparison of spatial interpolation schemes for rainfall data and application in hydrological modeling. *Water (Switzerland)*, 9(5), 1–18. <https://doi.org/10.3390/w9050342>

- Delrieu, G., Wijbrans, A., Boudevillain, B., Faure, D., Bonnifait, L., & Kirstetter, P. E. (2014). Geostatistical radar-raingauge merging: A novel method for the quantification of rain estimation accuracy. *Advances in Water Resources*, 71, 110–124. <https://doi.org/10.1016/j.advwatres.2014.06.005>
- Di Piazza, A., Conti, F. Lo, Noto, L. V., Viola, F., & La Loggia, G. (2011). Comparative analysis of different techniques for spatial interpolation of rainfall data to create a serially complete monthly time series of precipitation for Sicily, Italy. *International Journal of Applied Earth Observation and Geoinformation*, 13(3), 396–408. <https://doi.org/10.1016/j.jag.2011.01.005>
- Hadi, S. J., & Tombul, M. (2018). Comparison of Spatial Interpolation Methods of Precipitation and Temperature Using Multiple Integration Periods. *Journal of the Indian Society of Remote Sensing*, 46(7), 1187–1199. <https://doi.org/10.1007/s12524-018-0783-1>
- Hancock, P. A., & Hutchinson, M. F. (2006). Spatial interpolation of large climate data sets using bivariate thin plate smoothing splines. *Environmental Modelling & Software*, 21(12), 1684–1694. <https://doi.org/10.1016/J.ENVSOFT.2005.08.005>
- Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25(15), 1965–1978. <https://doi.org/10.1002/JOC.1276>
- Knoben, W. J. M., Freer, J. E., & Woods, R. A. (2019). Technical note: Inherent benchmark or not? Comparing Nash-Sutcliffe and Kling-Gupta efficiency scores. *Hydrology and Earth System Sciences*, 23(10), 4323–4331. <https://doi.org/10.5194/hess-23-4323-2019>
- Li, J. (2016). Assessing spatial predictive models in the environmental sciences: Accuracy measures, data variation and variance explained. *Environmental Modelling and Software*, 80, 1–8. <https://doi.org/10.1016/j.envsoft.2016.02.004>
- Ly, S., Degré, A., & Charles, C. (2013). Different methods for spatial interpolation of rainfall data for operational hydrology and hydrological modeling at watershed scale. A review. *Biotechnology, Agronomy and Society and Environment*, 17(2), 392–406. <https://doi.org/10.6084/m9.figshare.1225842.v1>
- Nashwan, M. S., Shahid, S., & Wang, X. (2019). Uncertainty in estimated trends using gridded rainfall data: A case study of Bangladesh. *Water (Switzerland)*, 11(2), 5–8. <https://doi.org/10.3390/w11020349>
- Newlands, N. K., Davidson, A., Howard, A., & Hill, H. (2011). Validation and inter-comparison of three methodologies for interpolating daily precipitation and temperature across Canada. *Environmetrics*, 22(2), 205–223. <https://doi.org/10.1002/env.1044>
- Portalés, C., Boronat-Zarceño, N., Pardo-Pascual, J., & Balaguer-Beser, A. (2008). *Un nuevo método para el cálculo de precipitaciones medias mediante técnicas de interpolación geoestadística considerando las características geográficas y topográficas del terreno* (pp. 1–8). Valencia: Congreso Internacional de Ingeniería Geomática y Topográfica.
- Price, D. T., McKenney, D. W., Nalder, I. A., Hutchinson, M. F., & Kesteven, J. L. (2000). A comparison of two statistical methods for spatial interpolation of Canadian monthly mean climate data. *Agricultural and Forest Meteorology*, 101(2–3), 81–94. [https://doi.org/10.1016/S0168-1923\(99\)00169-0](https://doi.org/10.1016/S0168-1923(99)00169-0)
- Seaman, R. S. (1988). Some real data tests of the interpolation accuracy of Bratseth's successive correction method. *Dynamic Meteorology and Oceanography*, 40(2), 173–176. <https://doi.org/10.3402/tellusa.v40i2.11791>
- Taesombat, W., & Sriwongsitanon, N. (2009). Areal rainfall estimation using spatial interpolation techniques. *ScienceAsia*, 35(3), 268–275. <https://doi.org/10.2306/>

- scienceasia1513-1874.2009.35.268
- Wagner, P. D., Fiener, P., Wilken, F., Kumar, S., & Schneider, K. (2012). Comparison and evaluation of spatial interpolation schemes for daily rainfall in data scarce regions. *Journal of Hydrology*, 464–465, 388–400. <https://doi.org/10.1016/j.jhydrol.2012.07.026>
- Wang, S., Huang, G., Lin, Q., Li, Z., Zhang, H., & Fan, Y. (2014). Comparison of interpolation methods for estimating spatial distribution of precipitation in Ontario, Canada. *International Journal of Climatology*, 34(14), 3745–3751. <https://doi.org/10.1002/joc.3941>
- Xu, W., Zou, Y., Zhang, G., & Linderman, M. (2014). A comparison among spatial interpolation techniques for daily rainfall data in Sichuan Province, China. *International Journal of Climatology*, 35(10), 2898–2907. <https://doi.org/10.1002/joc.4180>
- Zhang, M., Leon, C. de, & Migliaccio, K. (2018). Evaluation and comparison of interpolated gauge rainfall data and gridded rainfall data in Florida, USA. *Hydrological Sciences Journal*, 63(4), 561–582. <https://doi.org/10.1080/02626667.2018.1444767>
- William Adolfo Polanco Anzueto, Licenciado en Matemática Aplicada.
Afilación laboral: Facultad de Ingeniería. USAC.
- Eugenio Miguel Polanco Sotoj, Bachiller en Ciencias y Letras, estudiante de Ingeniería en Ciencias y Sistemas.

INFORMACIÓN DE LOS AUTORES

José Luis Argueta Mayorga, Ingeniero Civil, Maestro en Ingeniería Vial, Doctor en Ciudad Territorio y Sostenibilidad.

Afilación laboral: Facultad de Ingeniería. USAC.

Mayra Virginia Castillo Montes, Licenciada en Matemática Aplicada, Doctora en Educación.

Afilación laboral: Facultad de Ingeniería. USAC.

Walter Arnoldo Bardales Espinoza, Ingeniero Agrónomo, Maestro en Recursos Hidráulicos.

Afilación laboral. Facultad de Ingeniería.