

Utilización de diferentes metodologías para la construcción de un mapa de precipitación acumulada en la Provincia de Santa Cruz

Using different methodologies for the construction of cumulative rainfall map in the Province of Santa Cruz

Dora Maglione¹, dmaglione@uarg.unpa.edu.ar; Julio Soto¹, sotojulioa@hotmail.com;
José Luis Sáenz¹, jlsaenz@uarg.unpa.edu.ar; Oscar Bonfili², obonfili@gmail.com

¹Instituto de Trabajo, Economía y Territorio - Unidad Académica Río Gallegos - Universidad Nacional de la Patagonia Austral - Av. Gdor. Gregores y Piloto Rivero - Río Gallegos Santa Cruz – Argentina

²Servicio Meteorológico Nacional, Oficina Meteorológica Río Gallegos

Recibido: 26/11/2018. Aceptado: 03/04/2019

RESUMEN

Uno de las aplicaciones más importantes en el área de la Climatología es la caracterización de las variables a través de superficies continuas. Para llevar adelante esta representación gráfica, es necesario recurrir a distintos métodos de interpolación para estimar el valor de la variable a partir de un número discreto de observaciones. En relación a la variable precipitación, en la literatura existe una gran cantidad de propuestas para llevar adelante este proceso. El objetivo de este trabajo fue comparar cuatro métodos para analizar los datos de la precipitación acumulada. Los métodos utilizados en este estudio fueron el de Kriging, Cressman, Barnes y el método de las distancias inversas ponderadas (IDW). El método de Kriging, para la estimación de los pesos, utiliza la estructura de correlación espacial capturada a partir del variograma. En el método de Cressman los pesos dependen de la distancia entre el punto a estimar y los vecinos que se encuentren dentro de un radio de proximidad R. El método de Barnes es una mejora del método de Cressman. El método IDW utiliza como pesos las distancias inversas ponderadas entre el punto a estimar y la de los puntos con información. Para llevar adelante esta comparación se calcularon para cada método, distintas medidas de errores de predicción. Estos errores se determinaron utilizando los valores estimados desde las imágenes generadas en los puntos con información. Estos mapas se obtuvieron a partir de los valores medios observados en puntos pertenecientes a distintos grillados regulares. Una vez seleccionado el mejor método se procedió a la elaboración del mapa de precipitación acumulada de la provincia de Santa Cruz.

Palabras clave: Mapa de precipitación acumulada; Kriging; método de Cressman; método de Barnes; método IDW

ABSTRACT

The characterization of the variables through continuous surfaces is one of the most important activities in the area of Climatology. To carry out this graphic representation, it is necessary to resort to different methods of interpolation to estimate the value of the variable from a discrete number of observations. In the literature there is a great amount of proposals to carry



out this process in relation to the precipitation variable. The objective of this work was to compare four methods to analyze the accumulative rainfall data. The methods used in this study were Kriging, Cressman, Barnes and the inverse distance weighted method (IDW). The Kriging method uses the structure of spatial correlation captured from the variogram for the estimation of weights. In the Cressman method the weights depend on the distance between the point to be estimated and the neighbors that are within a radius of proximity R. The Barnes method is an improvement of the Cressman method. IDW uses as weights the weighted inverse distances between the point to be estimated and that of the points with information. To carry out this comparison, different measures of prediction errors were calculated for each method. To determine these errors the estimated values from the generated images in the information points were taken into account. These maps were obtained from the average values observed in the points belonging to different regular grids. Once the best method was selected, the accumulative rainfall map of the province of Santa Cruz was constructed.

Key words: Accumulative rainfall map; Kriging method; Cressman method; Barnes method; IDW method

INTRODUCCIÓN

En muchos campos de la ciencia de la tierra, en particular en el área de la climatología, es frecuente el generar superficies continuas a partir de un número discreto de observaciones provenientes de distintas ubicaciones espaciales. Para realizar esta tarea, es necesario recurrir a algún método para estimar el valor de una variable en un punto que no fue observado con la ayuda de valores en puntos vecinos que si fueron observados.

En la literatura existe una gran cantidad de propuestas, algunos son métodos interpoladores de un nivel, otros de dos niveles. Ahora, al momento de tener que generar un mapa continuo, el investigador debe optar por alguna de las distintas alternativas. Cada método tiene sus beneficios y sus contrapartes, además, algunos deben cumplir una serie de supuestos que necesitan ser chequeados para su efectiva implementación.

Por otra parte, el mapa generado proporciona estimaciones para los distintos puntos no observados de la superficie. Desde el punto de vista estadístico, este número no tiene demasiada relevancia sino viene acompañado por alguna medida de variabilidad que indique el posible error que se está cometiendo al proveer esa información.

En este artículo, se ofrece una propuesta de trabajo para comparar varios métodos citados en la literatura para la construcción de superficies continuas para la variable precipitación acumulada. La finalidad de esto es el de seleccionar el que mejor ajuste produzca en el sentido de minimizar ciertos errores de predicción, para a partir de éste construir una superficie continua para la provincia de Santa Cruz que caracterice espacialmente el comportamiento de la variable.

El trabajo está organizado de la siguiente manera, en la primera sección se presentan los distintos métodos a ser analizados. En la siguiente se implementan y comparan, y a partir del que mejor ajusta al conjunto de datos disponibles se genera un mapa de precipitación acumulada para la provincia de Santa Cruz. Finalmente, se ofrecen algunas conclusiones generales.

1. MARCO DE REFERENCIA

1.1. Antecedentes

Dentro de los modelos de predicción espacial es posible diferenciar entre dos grandes enfoques, por un lado, el determinístico y por el otro, el probabilístico. Sin embargo, Dobesch, Dumolard y Dryas (2007) y Tveito et al. (2006) clasifican los métodos en tres categorías principales, los dos antes mencionados y un tercero al que denominan otros métodos, los cuales agrupan aquellos especialmente desarrolladas para propósitos meteorológicos usando (una combinación de) métodos determinísticos y probabilísticos. Los métodos deterministas crean una superficie continua utilizando solo las características geométricas de las observaciones puntuales. Los métodos probabilísticos se basan en la teoría probabilística y usan el concepto de aleatoriedad, así el campo interpolado realizado es una de muchas realizaciones lo cual permite incluir la varianza en el proceso de interpolación y calcular la significación estadística de los valores predichos.

Dentro del primer grupo se encuentran modelos muy bien instrumentados como los polígonos de Thiessen (nombrados en honor al meteorólogo estadounidense Alfred H. Thiessen, y más conocidos como diagramas o teselación de Voronoi o teselación de Dirichlet), el método de interpolación lineal, método de interpolación bilineal, la representación polinomial, el método Spline y el de las distancias inversas ponderadas (conocido como IDW por sus siglas en inglés, Inverse Distance Weighting), las redes neuronales (Sluiter, 2009; Hengl, 2009), entre algunas de las muchas propuestas.

Por otro lado, en los métodos probabilísticos la interpolación se realiza en dos etapas, en la primera se determina una función que resume la estructura estadística del campo, la cual servirá de base para la determinación de los pesos que se usarán en la interpolación, y en la segunda etapa los pesos de la aproximación se calculan a través de un proceso de optimización. Como representante de esto último tenemos el método de Kriging descrito por Matheron (1971) donde la estructura estadística del campo se explica a través del semivariograma, y también al de Gandín (Montoya et al., 2000) donde los pesos se estiman a partir de la función de autocorrelación.

Los modelos determinísticos son relativamente sencillos de aplicar y en algunos casos bastantes eficientes (Hengl, 2009; Rauthe et al., 2013), sin embargo, suelen no proporcionar medidas para el error cometido. Es conveniente que los mapas de interpolaciones de la variable estén acompañados con los correspondientes mapas de isolíneas de los errores y de las varianzas de predicción para poder identificar zonas de mayor incertidumbre en las predicciones (Giraldo, 2000).

No existe un método general que sea adecuado para todos los problemas: depende de la naturaleza de la variable y de la escala de tiempo en la que se representa la variable (Sluiter, 2009). En general, hay mucha discusión acerca de cuál método utilizar para generar superficies continuas a partir de un número determinado de ubicaciones con información, como lo demuestran la gran cantidad de estudios que se han realizado para comparar métodos de interpolación para alguna variable climatológica en un contexto determinado (Bustamante, 2003; Alzate Velázquez et al., 2017; Villatoro et al., 2008; Soenario, Sluiter y Plieger, 2010, Hiemstra y Sluiter, 2011; Stepek y Wijnant, 2011; entre otros) por lo cual la elección del mismo debiera realizarse de manera minuciosa (Montoya et al., 2000).

Se puede hacer una segunda distinción entre interpoladores, y diferenciar entre los exactos y aproximados. Los interpoladores exactos reproducen los valores originales en los puntos de datos en los que se basa la interpolación. Los interpoladores aproximados no reproducen los valores originales, asumiendo incertidumbre en estos valores y reduciendo los errores por el efecto de suavizado.

En Tveito (2007) se presenta un marco de recomendación para el desarrollo de un enfoque de interpolación, el cual consta de varios pasos, los cuales contemplan la elección de un método de interpolación correcto, su correcta utilización, la prueba de varios métodos y la validación. Como señala Montoya et al. (2000), se deben analizar las variantes de los métodos contemplados, la distribución espacial de las observaciones, las características del área de estudio, así como la naturaleza y resolución temporal y espacial de la variable climática que se desea interpolar.

Sluiter (2009) presenta una revisión de distintos métodos usados en la bibliografía para datos climáticos. Dentro de los determinísticos presenta a el método Thiessen o Voronoi, las Redes Triangulares Irregulares (TIN), los cuales tienen una aplicación limitada, aunque pueden usarse si se dispone de una red densa de mediciones; el método IDW donde el peso de cada observación está determinado por la distancia inversa, por lo que se da más peso a los puntos cercanos; un modelo de regresión lineal determinística entre dos variables como interpolador global, a partir de relaciones físicas entre las mismas, y las redes neuronales que modelan relaciones complejas no lineales y requieren de alto costo computacional. Dentro de los probabilísticos, los cuales incorporan el concepto de aleatoriedad se incluyen los modelos de regresión lineal, y el de kriging, centrándose especialmente en este último con sus distintas variantes ya que según Burrough y McDonnell (1998) en aplicaciones de geociencia son los que resultan ser los mejores en el caso de que los datos no estén demasiados dispersos. Dentro de otros métodos menciona al método MISH (interpolación meteorológica basada en la base de datos de superficie homogeneizada) que incorpora información de otras variables en el proceso de estimación, y también el método PRISM que es un sistema basado en el conocimiento que utiliza mediciones puntuales de precipitación, temperatura y otros elementos climáticos para producir coberturas digitales continuas.

En Europa, se aplicó PRISM para calcular la climatología de precipitación a largo plazo para la región alpina por Schwarb (Schwarb, 2001). PRISM demostró ser muy poderoso en áreas donde la red de la estación no es representativa de la variación en la topografía.

Alzate Velásquez et al. (2017) modelan las precipitaciones utilizando el modelo regionalizado de lluvia Regionalisierte Niederschlage (Regnie), y en la que se incorporan datos de la pendiente y exposición del terreno. En este caso, los estadísticos de prueba de las superficies Regnie fueron similares a los obtenidos con interpolaciones con los métodos Spline e IDW.

Sluiter (2012) presenta una producción de los mapas en el Atlas Climático 1981-2010 (Klimaatlas 1981-2010) para distintas variables climatológicas: temperatura, precipitación, evaporación, presión atmosférica, humedad y viento, utilizando como métodos a kriging, las distancias inversas ponderadas (IDW) y el equivalente bidimensional de la spline cúbica (Thin Plate Spline, TPS). En el caso de la precipitación, el método utilizado para la construcción de las superficies de precipitación fue el kriging ordinario.

En Eslovenia, el Kriging universal se usó para calcular las precipitaciones normales (Sluiter, 2009).

Schuermans, Bierkens y Pebesma (2007) utilizan cokriging para combinar datos de estación con datos de radar de precipitación (así como Kriging con deriva externa). En general, cokriging da mejores resultados cuando el número de covariables es (mucho) más alto que la variable de interés y cuando la correlación espacial entre variables y covariables es alta.

Soenario, Sluiter y Plieger (2011) comparan los métodos IDW y Kriging ordinario para la construcción de mapas de precipitaciones, llegando a la conclusión de que los últimos resultan ser mejores.

En Colombia se utilizaron distintos métodos interpolar datos de precipitaciones, el de Spline (Saavedra, 1993), Kriging (Izquierdo, 1993), el método de Gandin (Montoya et al., 2000). Montoya et al. (2000) compara el método de Gandin (el cual es una extensión del método

propuesto por Cressman en 1959) con el de Kriging, y concluye que el primero es mucho más eficiente que el segundo.

Así, se puede observar que, en el caso de la precipitación, ha sido modelada en distintas regiones utilizando un espectro variado de métodos. En particular, en este trabajo vamos a considerar cuatro métodos alternativos para la construcción de la superficie continua para la precipitación, ellos son el método IDW, el método de Cressman, el método de Barnes que es una mejora al método de Cressman y el de Kriging.

1.2. Marco teórico

El problema de interpolación se puede plantear de manera general de la siguiente manera: sea z la variable de interés y $z(\mathbf{s}_1), z(\mathbf{s}_2), \dots, z(\mathbf{s}_n)$ los valores de la variable observados en los puntos $\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_n$ y sea $z(\mathbf{s}_o)$ el valor que se desea estimar en el punto \mathbf{s}_o . El interpolado puede expresarse como $\hat{z}(\mathbf{s}_o) = p_1 z(\mathbf{s}_1) + p_2 z(\mathbf{s}_2) + \dots + p_n z(\mathbf{s}_n)$, donde los p_i son los pesos en principio desconocidos. Las distintas formas de elegir o calcular a los pesos son los que dan lugar a los distintos métodos.

Como señaláramos anteriormente, existen una gran diversidad de métodos, muchos de los cuales a su vez ofrecen variantes. A continuación, se van a describir los métodos que van a ser utilizados y comparados para seleccionar el que mejor resulte en el sentido de minimizar algunas medidas para los errores de predicción en puntos con datos existentes.

Método de las distancias inversas ponderadas o más conocido como IDW

Es posiblemente uno de los métodos de predicción espacial más antiguos y simple de aplicar, fue introducido por Shepard (1968). Como muchos otros interpoladores, el valor de la variable en una determinada localización se obtiene como un promedio ponderado, donde los pesos son proporcionales a las distancias inversas:

$$\hat{z}(\mathbf{s}_o) = \sum_{i=1}^n p_i(\mathbf{s}_o) z(\mathbf{s}_i)$$

siendo z la variable de interés, $\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_n$ las localizaciones con información y $p_i(\mathbf{s}_o)$ el peso,

definido como $p_i(\mathbf{s}_o) = \frac{\frac{1}{d(\mathbf{s}_o, \mathbf{s}_i)}}{\sum_{j=1}^n \frac{1}{d(\mathbf{s}_o, \mathbf{s}_j)}}$ donde $d(\mathbf{s}_o, \mathbf{s}_i)$ representa la distancia entre los puntos

\mathbf{s}_o y \mathbf{s}_i . De esta manera se otorga mayor peso a los puntos vecinos y la influencia disminuye a medida que nos alejamos de la ubicación \mathbf{s}_o . También se puede usar en vez $d(\mathbf{s}_o, \mathbf{s}_i)$ una potencia de la distancia $d^\beta(\mathbf{s}_o, \mathbf{s}_i)$, con $\beta > 1$, que es un parámetro que enfatiza la similitud espacial.

Método de Cressman

Los métodos de aplicación polinómica para los pesos en el área de la meteorología según Montoya et al. (2000), fue inicialmente propuesto por Panofsky (1949), para describir los campos de la presión y la temperatura a través de polinomios. Posteriormente el mismo fue desarrollado por Gilchrist y Cressman (1954) y mejorado por Cressman (1959) al introducir la dependencia del inverso de la distancia.

En el caso de este método, los pesos dependen solo de la distancia r entre la ubicación \mathbf{s}_o donde se debe estimar el valor del campo y la ubicación de la observación \mathbf{s}_j : $r = d(\mathbf{s}_o, \mathbf{s}_j)$. En

este caso los pesos se calculan como $p_i(\mathbf{s}_o) = \frac{w_i}{\sum_{j=1}^n w_j}$ donde $w_i = \begin{cases} \frac{R^2 - r^2}{R^2 + r^2} & \text{si } r < R \\ 0 & \text{sino} \end{cases}$

El radio de búsqueda R es un parámetro de control típico y define la escala de longitud sobre la cual se usa una observación. Este parámetro es elegido por el usuario, se puede hacer que esta escala de longitud varíe en el espacio según la cobertura de datos, a la escala física o en función del conocimiento del dominio del problema.

Este método es muy sencillo y fácil de implementar, pero presenta algunos problemas, tales como que no se pueden obtener estimaciones en lugares donde no existen observaciones dentro del radio R y en regiones con pocas observaciones puede devolver campos discontinuos.

Método de Barnes

El método de Barnes es un método muy utilizado para analizar el hidrograma asociado a las precipitaciones.

Como una alternativa a la elección de los pesos propuestos por Cressman, se pueden definir otras funciones de ponderación. En particular, Barnes (1964) propone que los pesos se definen

utilizando una función gaussiana $w_j = e^{-\frac{r^2}{2R^2}}$

Debido a que los pesos de Barnes nunca se anulan, todas las observaciones se utilizan al momento de predecir. Además, permite obtener estimaciones en todos los puntos (las cuales pueden o no ser precisas) y se evitan las discontinuidades artificiales que se producen con el método de Cressman.

Método de Kriging

La interpolación con Kriging, comienza al reconocer que la variación de un atributo continuo es muchas veces demasiado irregular como para ser modelado por una simple función. La variación se puede describir mejor como una superficie estocástica con un atributo conocido como variable regionalizada:

$$Z(s) = \mu(s) + \varepsilon(s)$$

donde $\mu(\cdot)$ es una función determinista que describe la componente estructural del campo (o variación a gran escala) y $\varepsilon(\cdot)$ es un componente aleatorio que captura la correlación espacial del campo (variación a microescala).

El variograma (2γ) o semivariograma (γ) es quien permite capturar la estructura de correlación espacial. El mismo está definido como $\gamma(\mathbf{h}) = \frac{1}{2n} \sum_{i=1}^n (z(s_i) - z(s_i + \mathbf{h}))^2$, donde n es el número de pares de puntos separados por el vector \mathbf{h} . Si el proceso es isotrópico, la función depende sólo de la longitud del vector \mathbf{h} . Un factor importante del variograma es el rango que describe la distancia a partir del cual los datos se vuelven espacialmente independientes.

El variograma es utilizado para el cálculo de los pesos óptimos para la interpolación. Kriging es un interpolador relativamente rápido que puede ser exacto o suavizado según el método, proporciona, además del mapa de predicción, el mapa de errores y probabilidades; aunque requiere la verificación de supuestos y toma de decisiones.

Existen distintos tipos de Kriging. En la predicción por Kriging ordinario, el valor predicho es una combinación lineal de los valores medidos, y la correlación espacial entre los datos capturados por el variograma determinan los pesos. Existen muchas otras variantes de Kriging: Kriging simple, Kriging universal, Cokriging, Kriging residual, Kriging indicador, Kriging disyuntivo, entre otros, los cuales están descriptos en Cressie (1993).

2. METODOLOGIA Y RESULTADOS

Se usaron datos promedios anuales de precipitación acumulada de 159 localizaciones ubicadas entre los 74° y 66° de longitud Oeste y entre los 45° y 53° de latitud Sur¹. En la Fig. 1 se presentan los puntos con información y la representación del valor observado en deciles.

¹ Parte de estos datos han sido recopilados por Diaz et al (2016) y por Monserrat et al (2016)

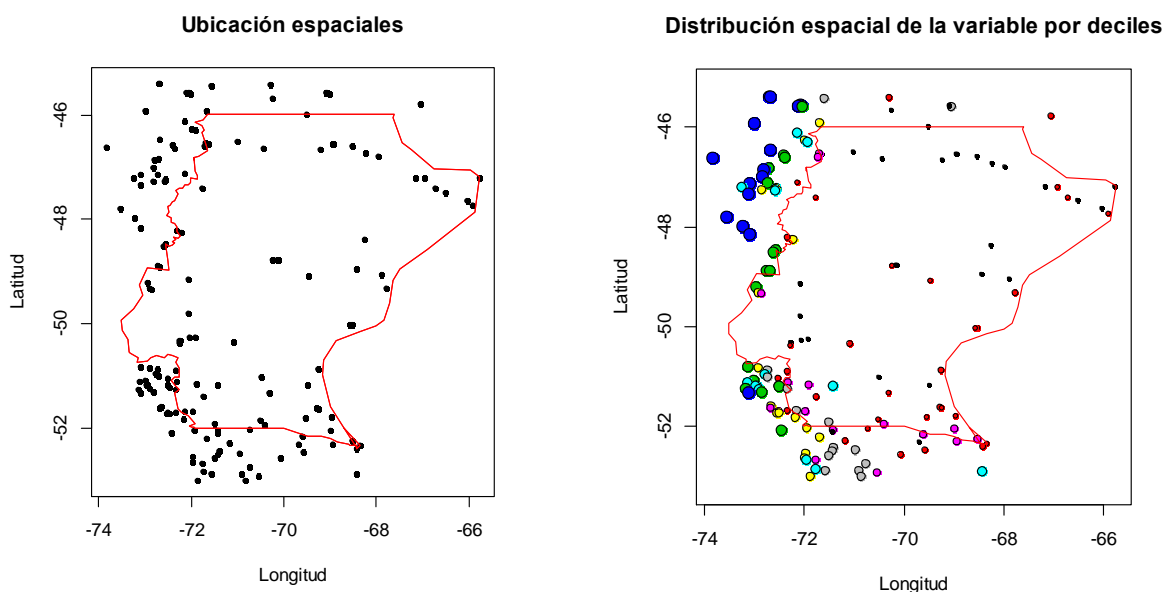


Figura 1: En el gráfico de la izquierda se presentan las ubicaciones espaciales con información. Y en el de la derecha los valores de la precipitación acumulada anual por deciles

Para poder comparar los métodos, se utilizaron distintos grillados regulares, tomando como localización espacial al punto central de cada casilla de la grilla y como valor observado al valor promedio de las observaciones de estaciones pertenecientes a dicha casilla.

Con estos puntos se procedió a realizar las interpolaciones utilizando el método IDW, el método de Cressman, el de Barnes y el de Kriging utilizando el software R. En el caso de Kriging fue necesario realizar una transformación de la variable ya que no se cumplía el supuesto de normalidad y seleccionar el mejor modelo para distintos tipos de semivariogramas y distintos tipos de ajuste para la tendencia a gran escala.

La comparación de las predicciones obtenidas a partir de los distintos interpoladores utilizados fue efectuada mediante la comparación de distintas medidas sobre los errores de predicción en los puntos originales. Esto se repitió para las distintas grillas utilizadas.

Las medidas de errores considerados fueron:

- Promedio del error absoluto (MAE), definido como $MAE = \frac{1}{n} \sum_{i=1}^n |z(s_i) - \hat{z}_m(s_i)|$; el modelo es mejor mientras menor sea MAE
- Raíz del error cuadrático medio (RMSE), definido como $RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (z(s_i) - \hat{z}_m(s_i))^2}$; el modelo es mejor mientras menor sea RMSE
- El error porcentual absoluto medio (MAPE), definido como $MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|z(s_i) - \hat{z}_m(s_i)|}{|z(s_i)|}$; el modelo es mejor mientras menor sea MAPE
- El error cuadrático relativo (RSE), definido como $RSE = \frac{\sum_{i=1}^n (z(s_i) - \hat{z}_m(s_i))^2}{\sum_{i=1}^n (\bar{z} - z(s_i))^2}$; el ajuste perfecto corresponde a RSE=0
- El error absoluto relativo (RAE), definido como $RAE = \frac{\sum_{i=1}^n |z(s_i) - \hat{z}_m(s_i)|}{\sum_{i=1}^n |\bar{z} - z(s_i)|}$; el ajuste perfecto corresponde a RAE=0

donde $z(s_i)$ es el valor observado de la precipitación en el punto s_i , y $\hat{z}_m(s_i)$ es el valor estimado de la precipitación en la localización s_i utilizando el método m (donde m es el método IDW, Cressman, Barnes o Kriging) y \bar{z} es el promedio de los valores observados.

Los tamaños de grillas utilizados fueron (expresados en grados): 0,25x0,25, 0,50x0,50, 0,75x0,75 y 1,00x1,00. En la Fig.2 se presentan los puntos medios de cada una de las celdas definidos para los distintos grillados.

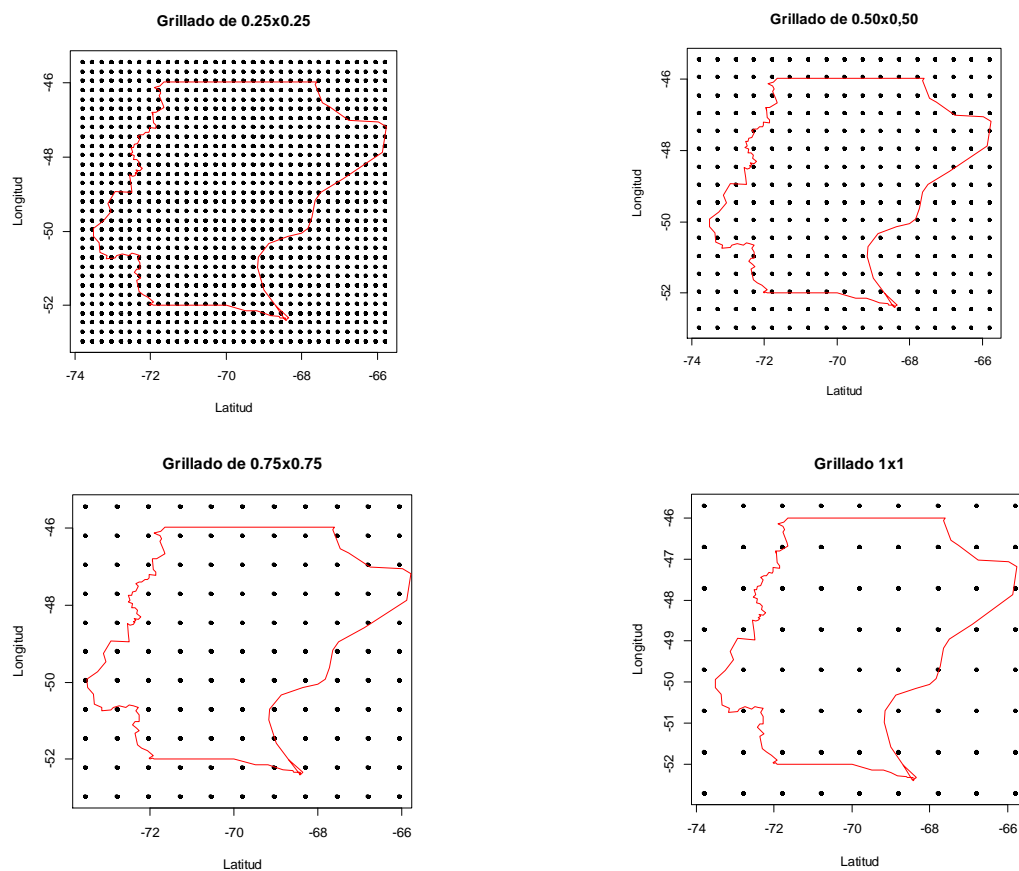


Figura 2: Distintos grillados usados para la generación de superficies continuas

A modo de ejemplo en la Fig. 3 se presentan los valores observados de la precipitación acumulada anual, y los que se obtienen sobre los puntos para la grilla 0,50 x 0,50 que son los valores usados para generar las superficies utilizando los cuatro métodos.

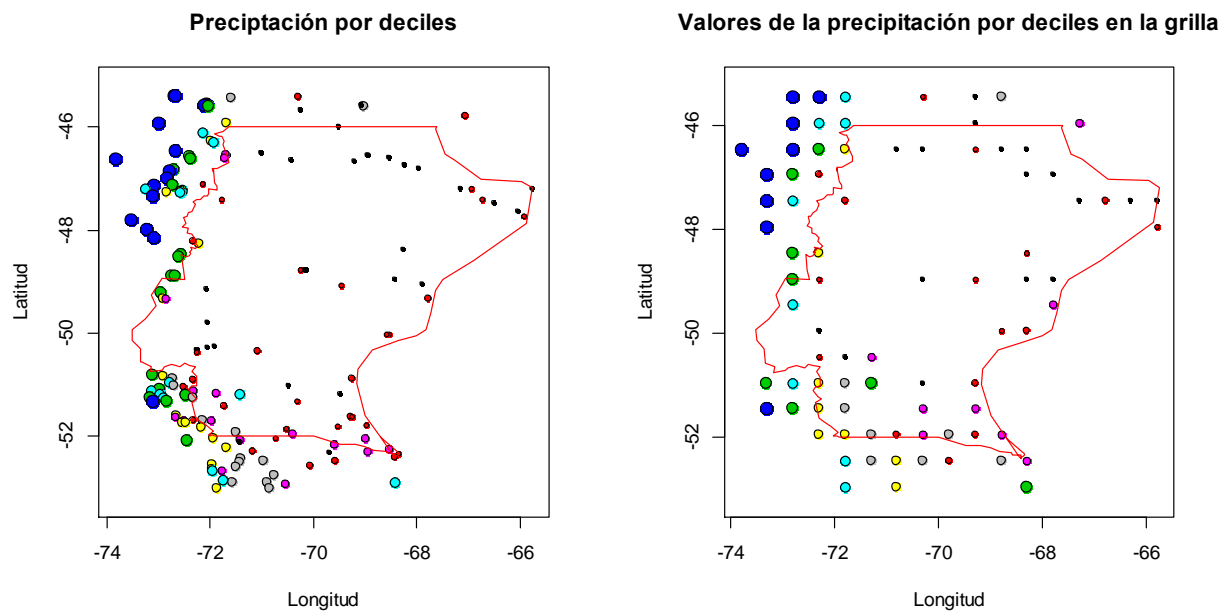


Figura 3: Valores observados de la precipitación acumulada anual (a la izquierda), y sus valores sobre la grilla (a la derecha), mostrados en deciles

A partir de los valores sobre los centros de la grilla (para cada una de ellas) se generan las cuatro imágenes de la precipitación acumulada anual para los cuatro métodos utilizados. En la Fig. 4, se presentan los mapas obtenidos para la grilla 0,50x0,50 para cada uno de los cuatro métodos.

Para el método de Cressman, se usó como radio de influencia a un $R=1^\circ$ en todas las direcciones para este tamaño de grilla (en general, el radio fue del doble del tamaño de la grilla en todos los casos lo que asegura una cobertura de 4 veces el tamaño del grillado en las direcciones de la latitud y la longitud).

Para poder aplicar el método de Kriging, fue necesario realizar una transformación de los datos, trabajando con el logaritmo de las precipitaciones acumuladas anuales desplazadas por una constante que asegurara simetría y curtosis. También fue necesario seleccionar entre varios modelos de variogramas (exponencial y esférico) y entre varios modelos para la tendencia (constante, lineal, cuadrática), usando el criterio de información de Akaike².

² El criterio de información de Akaike es una medida de la calidad relativa de un modelo estadístico, para un conjunto dado de datos. Permite comparar modelos, son mejores los que presentan menor valor.

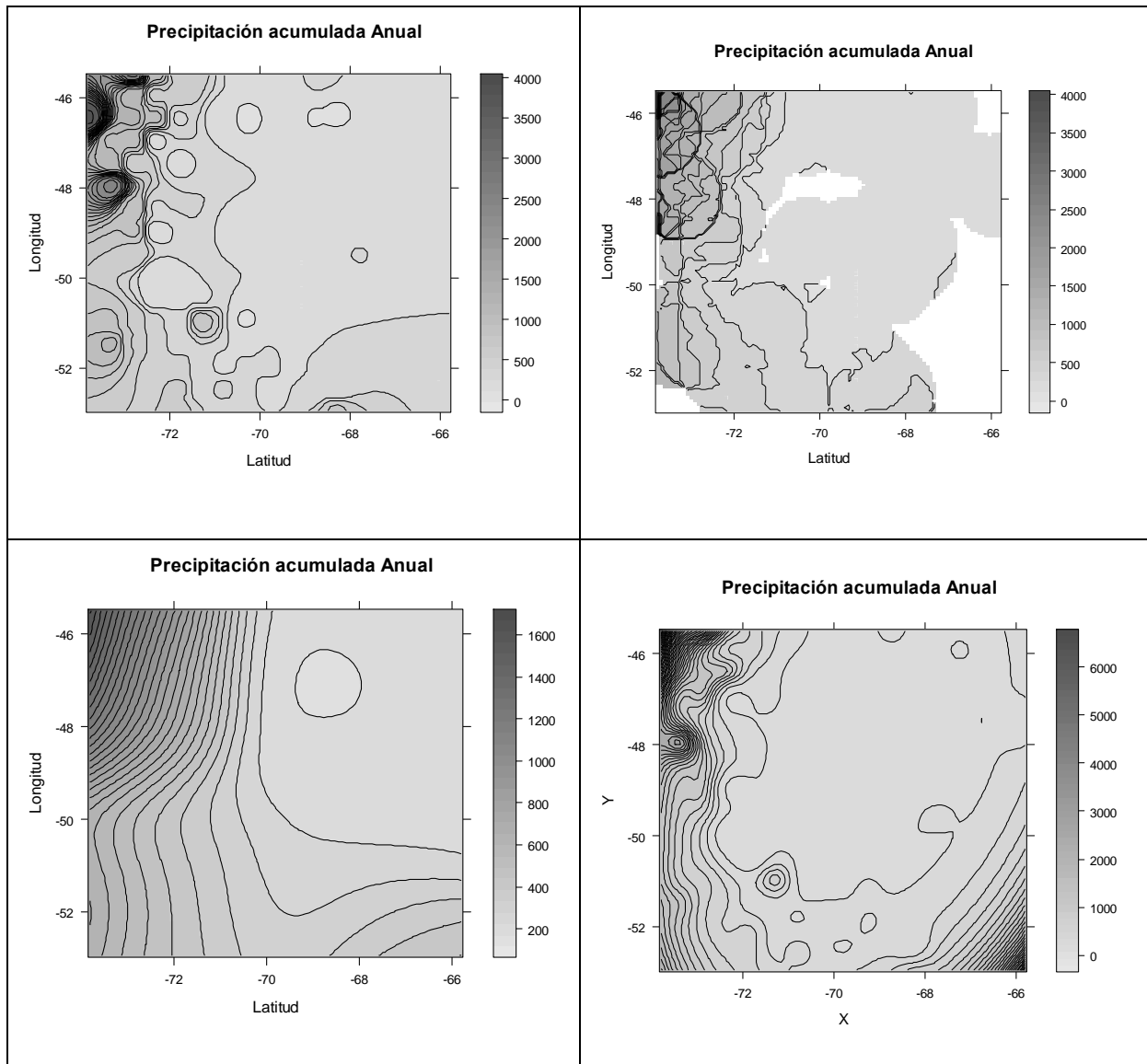


Figura 4: Imágenes obtenidas usando los cuatro métodos a partir de una grilla de 0.50 x 0.50: IDW (superior izquierda), Cressman (superior derecha), Barnes (inferior izquierda) y Kriging (inferior derecha)

Una vez que se generaron las imágenes, los valores predichos en los puntos originales con información fueron calculados. A partir de estos valores, se computaron las medidas de error MAE, RMSE, MAPE, RSE y RAE, para cada uno de los métodos. En la Tabla 1, se presentan los valores obtenidos para los distintos grillados y para cada método.

De la comparación de las medidas de error, se observa que el mejor método resulta ser Kriging. Para éste los valores de MAE, RSME y MAPE son menores que los calculados para los otros métodos para todos los tamaños de grilla considerados. Además, los valores RSE y RAE resultan ser los más cercanos a cero en todos los casos.

Se advierte también que a medida que aumenta el tamaño de la grilla se pierde precisión en las estimaciones.

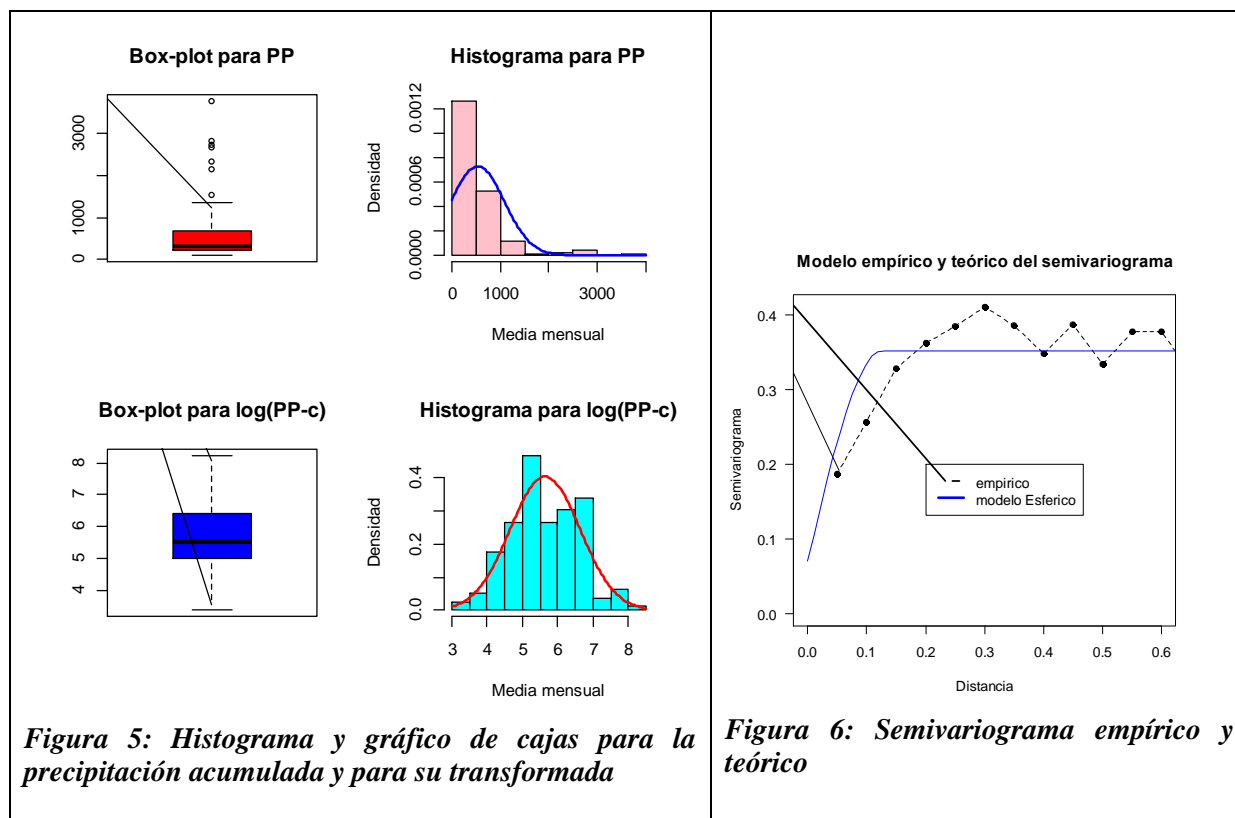
Tabla 1: Valores de los errores calculados para cada método y para cada tamaño de grilla

Tamaño de la Grilla	Errores	IDW	CRESSMAN	BARNES	KRIGING
0.25 x 0.25	MAE	220.4846	102.6879	150.7630	77.6236
	RMSE	369.8339	155.4661	256.9408	126.2741
	MAPE	0.5877	0.2261	0.3132	0.1470
	RSE	0.4621	0.0817	0.2230	0.0539
	RAE	0.6175	0.2876	0.4223	0.2174
0.5 x 0.5	MAE	237.4548	183.4321	207.3230	106.8601
	RMSE	397.3135	337.3855	380.9236	169.9125
	MAPE	0.6186	0.3450	0.4521	0.2039
	RSE	0.5333	0.3846	0.4902	0.0975
	RAE	0.6651	0.5138	0.5807	0.2993
0.75 x 0.75	MAE	241.1841	214.0680	225.7967	108.9038
	RMSE	406.6814	371.8716	404.5587	185.7504
	MAPE	0.6389	0.4505	0.5333	0.2203
	RSE	0.5588	0.4672	0.5529	0.1166
	RAE	0.6755	0.5996	0.6324	0.3050
1 x 1	MAE	244.1382	233.2976	239.0638	149.2381
	RMSE	377.5495	375.1387	410.3175	265.7342
	MAPE	0.6867	0.5075	0.6086	0.2944
	RSE	0.4816	0.4754	0.5688	0.2386
	RAE	0.6838	0.6534	0.6696	0.4180

Luego del método de Kriging, el que mejores predicciones proporciona es el de Cressman, aunque este método tiene la desventaja de no poder estimar siempre todos los valores (ya que sectores que no poseen puntos con información dentro del radio R permanecen sin estimación). Luego le sigue el método de Barnes y finalmente el IDW.

En consecuencia, para la construcción de un mapa predictivo de precipitaciones acumulada para la Provincia de Santa Cruz, se va a utilizar el método de Kriging.

Dado que la precipitación acumulada anual (PP) no sigue una distribución normal, fue necesario aplicar una transformación $f(x)=\log(x-c)$ donde la constante c se seleccionó de manera de asegurar simetría y curtosis para los datos transformados ($c=85\text{ml}$). En la Figura 5, se presentan los gráficos de caja e histograma para la variable en la escala original (parte superior) y para su transformada (en la parte inferior). Para la variable transformada los datos siguen una distribución normal ($p\text{-valor}=0.2371$ para el test de normalidad de Shapiro-Wilks). El índice de Morán indica que existe correlación espacial ($p\text{-valor}=0$), por lo que es una buena estrategia incorporarla en la modelación que es lo que utiliza el método de Kriging al momento de estimar los pesos en las distintas interpolaciones. Además, se puede suponer isotropía, ya que al analizar el comportamiento de la correlación espacial en distintas direcciones (N-S, O-E, SO-NE, NO-SE) se observa que la misma presenta una tendencia similar.



Para la variable transformada, se compararon distintos modelos aditivos de la forma: $Z(s) = \mu(s) + e(s)$, donde s representa un punto del espacio bidimensional, $\mu(\cdot)$ representa la tendencia o variación a gran escala y $\varepsilon(\cdot)$ contiene la estructura de dependencia espacial del proceso bajo estudio o variación a pequeña escala. Para la tendencia se contemplaron como alternativas la superficie constante, la lineal y la de segundo orden; y para modelar la estructura de correlación espacial subyacente en los datos se consideraron distintos modelos de semivariograma pertenecientes a las familias esférica y exponencial. El mejor modelo fue seleccionado a través del criterio de Akaike (Tabla 2), y resultó poseer una tendencia cuadrática a gran escala, dada por $\mu(x, y) = 7.7499 - 9.3081x - 3.0543y + 10.2252x^2 + 4.7365y^2 - 5.1555xy$ y para modelar la estructura de correlación espacial un semivariograma perteneciente a la familia esférica (Fig.6):

$$\gamma(h) = \begin{cases} 0.0706 + 0.2815 \left(\frac{1.5|h|}{0.1261} - \frac{0.5|h|^3}{0.1261^3} \right) & \text{si } |h| \leq 0.1261 \\ 0.3521 & \text{si } |h| > 0.1261 \end{cases}$$

Finalmente, el modelo para caracterizar el comportamiento de la variable precipitación acumulada resulta ser un modelo multiplicativo dado por

$$Z(s) = e^{\mu(s)} e^{\varepsilon(s)} + c$$

donde $\mu(\cdot)$ representa la tendencia o variación a gran escala y $\varepsilon(\cdot)$ contiene la estructura de dependencia espacial del proceso bajo estudio o variación a pequeña escala descritos ambos anteriormente. En las Fig. 7 se presenta el mapa generado y los errores de predicción. En la Fig. 8 se muestran algunas isólinas para la provincia de Santa Cruz.

Tabla 2: Valor Akaike para los modelos comparados para la tendencia general y el semivariograma que modela la estructura de correlación espacial

Modelo	Grados de libertad (gl)	Akaike (AIC)	AIC-2 gl
$\mu(s)=cte$, 2γ de la familia exponencial	4	222.6835	214.6835
$\mu(s)=cte$, 2γ de la familia esférica	4	221.0363	213.0363
$\mu(s)=lineal$, 2γ de la familia exponencial	6	221.2383	209.2383
$\mu(s)=lineal$, 2γ de la familia esférica	6	218.7276	206.7276
$\mu(s)=cuadrática$, 2γ de la familia exponencial	9	203.5956	185.5956
$\mu(s)=cuadrática$, 2γ de la familia esférica	9	203.3388	185.3388

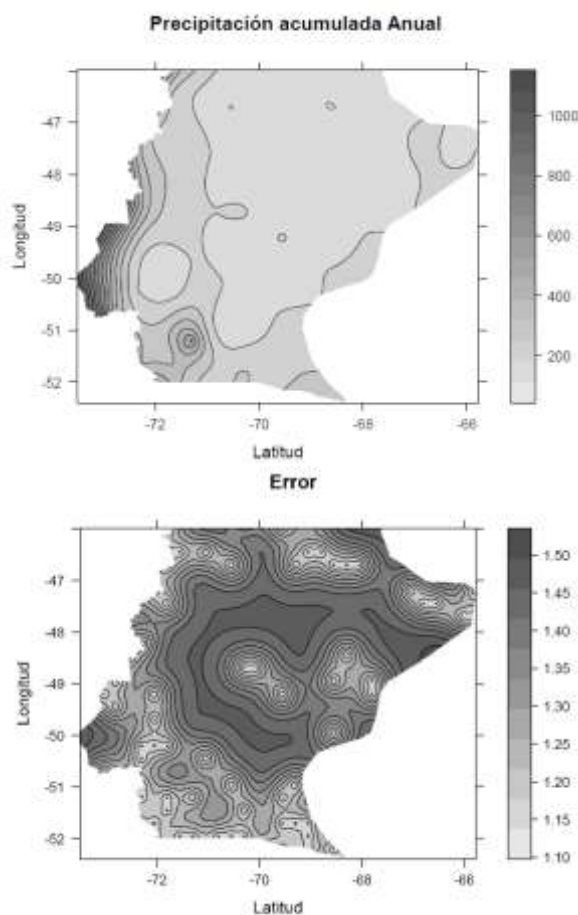


Figura 7: Mapa de precipitación estimado y Mapa de errores de predicción para la precipitación acumulada anual

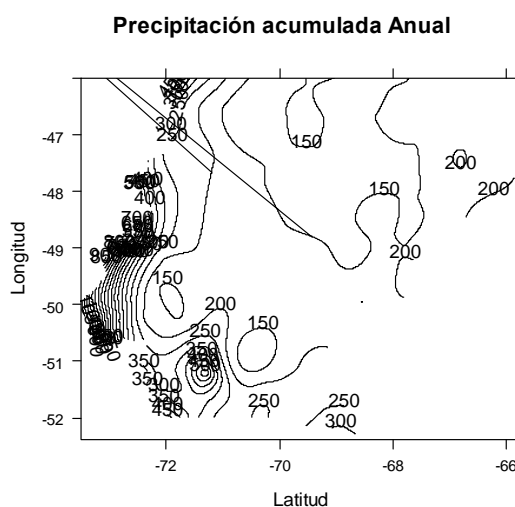


Figura 8: Isolíneas para la precipitación acumulada anual.

3. CONCLUSIONES

En este trabajo, se presenta una metodología para cuando se desea construir un mapa continuo para una variable a partir de un conjunto discreto de valores. Como se observara la aplicación de uno u otro método de interpolación genera distintas superficies para la variable bajo estudio. Así, es importante realizar algún análisis sobre los errores que se cometen al aplicar una u otra técnica, si lo que se desea es cierto grado de precisión en las estimaciones de valores en puntos espaciales no observados.

Particularmente, para esta variable y esta distribución de datos, las mejores predicciones continuas se obtienen utilizando el método estocástico de Kriging, el cual, además, proporciona para cada punto estimado un error de predicción. Luego, le sigue el método de Cressman, aunque tiene la desventaja de no proporcionar valores predictivos para todos los puntos de la región.

En particular, fue posible construir un mapa para las precipitaciones acumuladas anuales a partir del método de Kriging mediante un modelo multiplicativo, ya que el logaritmo de la variable fue modelado por medio de un modelo aditivo con tendencia cuadrática y con estructura de correlación espacial definida a partir de un semivariograma de la familia exponencial.

Los métodos estocásticos tienen la ventaja de generar superficies de incertidumbre o error que describen la calidad de la predicción, sin embargo, para aplicarlos es necesario conocer en profundidad el comportamiento espacial del fenómeno representado, especialmente en aspectos como la estacionariedad, la independencia espacial, la isotropía, la densidad de estaciones, además de las características propias de la variable a interpolar; por lo cual no necesariamente el método probabilístico genera las mejores superficies, lo cual sí se da en este caso.

Una característica a mencionar es que la precisión en el mapa generado, depende en gran medida de la distribución espacial de los datos, mientras mayor sea la cobertura y densidad, mejor serán las predicciones (como se observara al comparar los errores de los distintos grillados para cada método). En el caso de Kriging, cuando mejor sea la calidad del mapeo (Kravchenko, 2003; Burrough y McDonnell, 1998), mejor es posible captar la estructura de correlación espacial y este método es el que proporcionará las mejores interpolaciones; sin

embargo, dependiendo del tipo de datos analizados, su costo y dificultad de obtención determinan que tan valioso es finalmente el uso de la interpolación.

Otra alternativa que podría también aplicarse, que no se ha incluido en este trabajo, es el de incorporar en la modelación variables correlacionadas a la que se desea modelar.

Como conclusión final, podemos decir que no hay un método general único para todos los problemas, por lo tanto, es una buena práctica el probar diferentes métodos para seleccionar aquel que proporcione mayor precisión.

4. REFERENCIAS

- ALZATE VELÁZQUEZ D.F, ARAUJO CARRILLO G.A., ROJAS BARBOSA E. O., GÓMEZ LATORRE D.A., MARTINEZ MALDONADO F.M. (2017). Interpolación Regnie para lluvia y Temperatura en las regiones Andina, Caribe y Pacífica de Colombia. *Colombia Forestal*. 21(1): 102-118.
<https://doi.org/10.14483/2256201x.11601>
- BARNES, S.L. (1964). A technique for maximizing details in numerical weather map analysis. *J.Appl.Meteor.* 3: 396.409.
[https://doi.org/10.1175/1520-0450\(1964\)003<0396:atfmdi>2.0.co;2](https://doi.org/10.1175/1520-0450(1964)003<0396:atfmdi>2.0.co;2)
- BURROUGH, P.A., McDONELLI R.A. (1998). *Principles of Geographical Information Systems*. Oxford University Press. 333p
- BUSTAMANTE, J. (2003). Cartografía predictiva de variables climatológicas: Comparación de distintos modelos de interpolación de la temperatura en España peninsular. *Graellsia*, 2-3, 359-376. <https://doi.org/10.3989/graellsia.2003.v59.i2-3.252>
- CRESSIE, N. (1993). *Statistics for Spatial Data*. Rev.Edn. Wiley. New York. 900 p.
- CRESSMAN, G.P. (1959). An operational objective analysis system. *Month Weather Review*, 87(10).
- DIAZ B., BONFILI O., MONSERRAT M.C., ALMONACID L., STURZEMBAUM V. (2016). Red de cooperación hidrometeorológica en Santa Cruz. XVI Reunión Argentina y VIII Latinoamericana de Agrometeorología.
- DOBESCH, H., P. DUMOLARD P. & DRYAS I., Eds. (2007) *Spatial interpolation for climate data: the use of GIS in climatology and meteorology*. Geographical Information Systems, Series. London: ISTE Ltd.
- GILCHRIST, B., CRESSMAN G.P. (1954). An experiment in objective weather analysis. *Tellus*, 6(4).
- GIRALDO, R. (2002). *Introducción a la geoestadística: teoría y aplicación*. Bogotá, D.C.: Universidad Nacional de Colombia. Facultad de Ciencias. Departamento de Estadística. 94 p. <https://doi.org/10.18273/revint.v35n1-2017006>
- HENGL, T. (2009). *A practical guide to geostatistical mapping*. Luxembourg: Office for Official Publications of the European Communities. 140 p. Recuperado de Hengl, T. (2007). *A practical guide to geostatistical mapping of environmental variables*. Luxembourg: Office for Official Publications of the European Communities. 140 p. <https://doi.org/10.1086/343356>
- HIEMSTRA P., SLUITER R. (2011). *Interpolation of Makkink evaporation in the Netherlands*, Royal Netherlands Meteorological Institute.
- IZQUIERDO, J.C. (1993). Uso de Kriging para construcción de registros de precipitación en Antioquia. *Atmósfera*, 19: 13-24.
- KRAVCHENKO A. (2003). Influence of spatial structure on accuracy of interpolation methods. *Soil Science Society of American Journal*, 67:1564-1571. <https://doi.org/10.2136/sssaj2003.1564>

- MATHERON, G. (1971). The theory of regionalized variables and its applications: Cahierscentre morph. Math. Fontainebleau, V. 5, 211 p.
- MONSERRAT M.C., DIAZ B., BONFILI O., ALMONACID L. (2016). Banco de datos hidrometeorológicos de Santa Cruz. 3r Encuentro de Investigadores en Formación de Recursos Hídricos.
- MONTOYA G., ESLAVA J., CALDERÓN A., VEGA H. (2000). Implementación del método de Gandin para interpolar datos de precipitación en Colombia. Meteorol. Colomb. 1:25-30. ISSN 0124-6984. Santa Fe de Bogotá, D.C. – Colombia.
- PANOFSKY, H.A. (1949). Objective weather analysis Journal of Meteorology, 6(6).
- RAUTHE, M., STEINER, H., U., RIEDIGER, A., MAZURKIEWITCZ, A., GRATZKI, A. (2013). A Central European precipitation climatology–Part I: Generation and validation of a high-resolution gridded daily data set (HYRAS). *Meteorologische Zeitschrift*, 22(3), 235-256. <https://doi.org/10.1127/0941-2948/2013/0436>
- SAAVEDRA, U.H.A. (1993). Utilización de la interpolación cúbica Spline en la complementación de series meteorológicas. *Atmósfera* 19: 38-45. <https://doi.org/10.4206/bosque.1999.v20n2-01>
- SCHUURMANS, J.M., BIERKENS M.F.P., PEBESMA E.J. (2007). Automatic Prediction of High-Resolution Daily Rainfall Fields for Multiple Extents: The Potential of Operational Radar. *Journal of Hydrometeorology* 8, pp. 1204 - 1224.
- SLUITER, R. (2009). *Interpolation methods for climate data–Literature review*. De Bilt: , Royal Netherlands Meteorological Institute. <https://doi.org/10.1175/2007jhm792.1>
- SCHWARB, M. (2001). The Alpine Precipitation Climate. Evaluation of a High-Resolution Analysis Scheme Using Comprehensive Rain-Gauge Data. *Zuercher Klimaschriften* 80,
- SHEPARD, D. (1968). A two-dimensional interpolation function for irregularly-spaced data. *Proceedings of the 1968 ACM National Conference*. pp. 517–524. <https://doi.org/10.1145/800186.810616>
- SOENARIO, I., R. SLUITER R., PLIEGER M. (2010). Optimization of Rainfall Interpolation. De Bilt, Royal Netherlands Meteorological Institute.
- STEPEK A., WIJNANT I.L. (2011). Interpolating wind speed normals from the sparse Dutch network to a high resolution grid using local roughness from land use maps. De Bilt, Royal Netherlands Meteorological Institute.
- TVEITO O.E., (2007). The Developments in Spatialization of Meteorological and Climatological Elements In: H. Dobesch et al (Ed.) *Spatial interpolation for climate data: the use of GIS in climatology and meteorology* pp. 73-86. London: ISTE Ltd. <https://doi.org/10.1002/9780470612262.ch6>
- TVEITO O.E., M. WEGEHENKEL M., VAN DER WEL F., DOBESCH H. (2006). The Use of Geographic Information Systems in Climatology and Meteorology - Final Report COST Action 719.