GEOACTA **42(1)** (2017) ISSN 1852-7744
© 2017 Asociación Argentina de Geofísicos y Geodestas

RELLENO DE SERIES DE PRECIPITACIÓN DIARIA PARA LARGOS PERIODOS DE TIEMPO EN ZONAS DE LLANURA. CASO DE ESTUDIO CUENCA SUPERIOR DEL ARROYO DEL AZUL

FILLING SERIES OF DAILY PRECIPITATION FOR LONG PERIODS OF TIME IN PLAIN AREAS. CASE STUDY SUPERIOR BASIN OF STREAM DEL AZUL

Cristian Guevara Ochoa[1,2], Ninoska Briceño[1,4], Erik Zimmermann[2,3], Luis Vives[1], Martin Blanco[1,4], Georgina Cazenave[1,4], Guadalupe Ares[1,2]

[1]Instituto de Hidrología de Llanuras "Dr. Eduardo Jorge Usunoff", IHLLA
[2]Consejo Nacional de Investigaciones Científicas y Técnicas, CONICET
[3]Centro Universitario Rosario de Investigaciones Hidroambientales, UNR
[4]Comisión de Investigaciones Científicas de la Provincia de Buenos Aires, CIC
E-mail: ingguevara8a@gmail.com

RESUMEN

El relleno de datos faltantes de precipitación diaria es un problema común en los estudios hidrológicos. El objetivo de este artículo es realizar una comparación y una evaluación de diferentes métodos que permiten rellenar los datos de precipitación diaria faltante para largos periodos en zona de llanura. Este estudio se realiza en la cuenca superior del arroyo del Azul para periodo de nueve años (2006-2014) y se emplean tres estaciones que tienen 3287 datos completos de precipitación diaria y seis estaciones con datos faltantes. Se implementaron siete métodos para el relleno de datos diarios de precipitación: el método de regresión lineal (MRL), el método de razones de distancia (MRD), el método de coeficientes de correlación con estaciones vecinas (MRC), el método de la razón promedio (MRP), método del inverso de la distancia al cuadrado (MIDW), seguido del método de cadenas de Márkov (MKV) y por último el método redes neuronales (MRN). Para la comparación y análisis de las diferentes metodologías se aplicaron diferentes estadísticos y gráficas temporales las cuales miden el ajuste de los datos calculados. Las redes probabilísticas y neuronales son los métodos más adecuados para rellenar datos en zonas de llanura. Los métodos que se aplicaron en el estudio obtuvieron un mejor ajuste en la época de otoño-invierno con menores precipitaciones, en comparación con el periodo primavera-verano en donde se obtuvieron ajustes más bajos debido a que en estas épocas se presentan tormentas convectivas con intensidades muy altas de precipitación.

Palabras clave: Relleno serie de datos, Precipitación diaria, Llanura

ABSTRACT

The filling of missing daily precipitation data is a common problem in hydrological studies. The aim of this article is to make a comparison and evaluation of different methods to fill in missing daily rainfall data for long periods in plain areas. This study is carried out in the upper basin of the Azul stream for a period of nine years (2006-2014) and three stations that have 3287 complete data of daily precipitation and six stations with incomplete data are used. Seven methods were implemented for the filling of daily rainfall data: the linear regression method (MRL), the distance reasons method (MRD), the coefficients correlation method with neighboring stations (MRC) average reason method (MRP), the reasons distance method (MRD), the inverse distance weighted method (MIDW), following Markov chain method (MKV) and finally neural networks method (MRN). For the comparison and analysis of the different methodologies, different statistics and temporal graphs were applied, which measure the adjustment of the calculated data. Probabilistic and neural networks are the most suitable methods to fill data in plain areas. The methods applied in the study obtained a better adjustment in the autumn-

winter season with lower rainfall, compared to the spring-summer period where lower adjustments were obtained because to that in these times there are convective storms with very high rainfall intensities.

Key words: Filling data series, Daily Precipitation, Plain

INTRODUCCIÓN

La precipitación origina la fase terrestre del ciclo hidrológico, la cual se manifiesta como una variable aleatoria en función del espacio y del tiempo (Orsolini et al., 2008). Es también un factor de control principal del ciclo hidrológico de una región, así como de la ecología, el paisaje y los usos del suelo.

Cuando se lleva a cabo un análisis de los datos de precipitación, especialmente si se utilizan estaciones de medición diaria, se suele encontrar días donde no se registran observaciones (Dingman, 2002; Teegavarapu y Nayak, 2017). Esos datos faltantes pueden ser periodos cortos o días aislados, o pueden extenderse durante largos períodos. Lo cual es un problema más complejo debido al relleno de una mayor cantidad de datos, este último caso es desarrollado en este artículo.

Contar con series continuas de precipitación es importante para el análisis de: tendencias temporales climáticas (Bonsal et al., 2017; Li et al., 2017; Serrano et al., 2017), pronósticos hidrometeorológicos (Davolio et al., 2017; Nguyen et al., 2017), cambio climático (Almazroui et al., 2017; Díaz y Vera, 2017; Dye et al., 2017; Hu et al., 2017), inundaciones y sequías (Scarpati y Capriolo, 2013; Guevara, 2015; Correa et al., 2017; Rudd et al., 2017; Wongchuig et al., 2017), ecosistemas (Schuur, 2003; Bai et al., 2008; Del Grosso et al., 2008; Yang et al 2008;), disponibilidad de agua (Putnam y Broecker, 2017; Yehia et al., 2017), agricultura (De Silva et al., 2007), planificación y gestión del recurso hídrico (Kolokytha et al., 2017), etc.

La falta de datos a partir de una serie de tiempo puede presentar varios obstáculos cuando se realizan estudios de modelación hidrológica, hidrogeológica, hidráulica (Kuligowski y Barros, 1998; Su et al., 2008; Chen et al, 2017), debido a que uno de los inconvenientes en el ajuste de este tipo de modelos es que se debe contar con series de precipitación continuas y con alta resolución temporal.

A su vez, llenar los vacíos generados por datos faltantes es esencial para el análisis de los cambios potenciales actuales y futuros de la precipitación ya que para Argentina según Giorgi (2002) y Barros et al. (2015), la precipitación en la mayor parte del país ha aumentado durante el siglo XX, sobre todo a partir de la década de 1960. Esto también se ha evidenciado para la provincia de Buenos Aires, con tendencias significativas hacia condiciones más húmedas (Scarpati y Capriolo, 2013; Scian y Pierini, 2013; Antico y Sabbione, 2010; Canziani, 2008; Haylock et al., 2006), así como un incremento notable en la frecuencia de fuertes lluvias sobre umbrales que van de 50 a 150 mm (Barros et al., 2008). Según los escenarios de cambio climático del Proyecto del Intercomparación de Modelos Acoplados (CMIP5), para Argentina en las próximas tres décadas se esperan tendencias proyectadas positivas de precipitación en las regiones norte y centro (Vera et al., 2006; Stocker et al., 2013 Diaz y Vera, 2017; Maenza et al., 2017), mientras que para las regiones centro-occidental y para la Patagonia se esperan tendencias negativas de precipitación (Vera et al., 2006; Barros et al., 2015).

En la actualidad los cambios potenciales en la precipitación han recibido una creciente atención (Beck et al., 2017), ya que afectan la dinámica hídrica (inundaciones o sequías), comprometiendo la disponibilidad espacio-temporal de los recursos hídricos, y la sustentabilidad de diversas actividades humanas como es la agricultura, industria y el desarrollo urbano en zonas de llanura (Huang et al., 2017).

Se han realizado varios estudios para evaluar los diferentes métodos para el relleno de datos de precipitación. Los métodos que han obtenido los mejores rendimientos son las regresiones lineales, regresiones múltiples y razones promedio, los cuales han dado buenos resultados en los estudios de Tapia et al. (2009), Romero y Casimiro (2015), Campozano et al. (2015). Uno de los métodos más utilizados para la estimación de datos faltantes en hidrología y geografía es la interpolación espacial por medio de la ponderación de la distancia inversa (De Silva et al., 2007; Vicente et al., 2010). El éxito de este método depende principalmente de la existencia de autocorrelación espacial positiva (Teegavarapu y Chandramouli, 2005). Los métodos probabilísticos, como las cadenas de Markov, han dado buenos resultados para la generación de series sintéticas de precipitación para largos periodos de tiempo (Baigorria y Jones, 2010; Mehan et al., 2017). El método de las redes neuronales es un enfoque basado en relaciones de aprendizaje entre variables dependientes e independientes y ha tenido buena aceptación para el pronóstico, ocurrencia e intensidad de la precipitación

(Teegavarapu y Chandramouli, 2005). Más adelante, en la sección de materiales y métodos, se mencionan distintos estudios que han utilizado las metodologías aquí aplicadas y que han sido comparadas para distintas zonas.

La finalidad de este trabajo consiste en determinar la confiabilidad de algunos métodos de relleno de datos diarios de precipitación por medio de estadísticos de ajuste como es el coeficiente de determinación (R^2), raíz cuadrada media del error (RMSE) y el coeficiente de eficiencia del modelado Nash Sutcliffe (NS).

Zona de estudio

El estudio se realiza en la cuenca superior del arroyo del Azul ubicada en el centro de Provincia de Buenos Aires, entre los 59º 46' y 60º 09' de longitud oeste y 36º 49' y 37º 21' de latitud sur (Figura 1). Abarca una superficie 1024 km², aguas arriba de la ciudad de Azul. Esta cuenca nace en la localidad de Chillar (60 km al sur de la ciudad de Azul) y sus afluentes más importantes son los arroyos Videla y Santa Catalina. La altitud de la cuenca varía entre 142 m a 367 m, posee una cabecera serrana que se transforma en su parte media-baja en una llanura de transición.

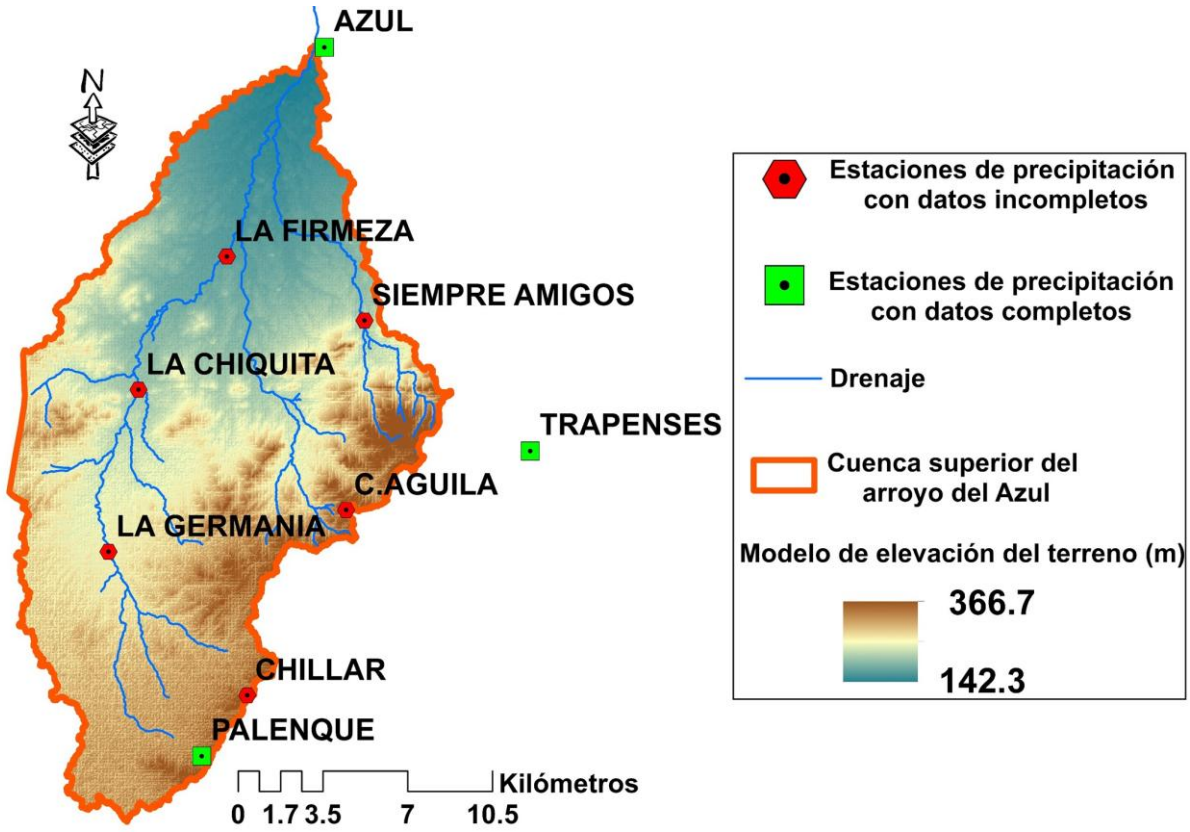

Figura 1. Cuenca superior del arroyo del Azul y distribución espacial de las estaciones de medición de la precipitación.
Figure 1. Superior basin of the Azul stream and spatial distribution of precipitation measuring stations.

Según la estación meteorológica Azul perteneciente al Servicio Meteorológico Nacional (SMN) la precipitación media anual es de 902 mm. Se observa que la distribución de la precipitación media mensual en la serie 2006-2014 de la estación Azul (Figura 2), evidencia un régimen isohigro. Se destaca el mes de enero por ser el más lluvioso (con un promedio 114 mm) y el mes de junio como el menos lluvioso (con un promedio de 38 mm). La distribución de la precipitación anual dentro de la cuenca no es uniforme, existe una tendencia general de lluvias más abundantes en el norte y sur, mientras que en la zona intermedia de la cuenca se presentan menores precipitaciones (IHLLA, 2003).

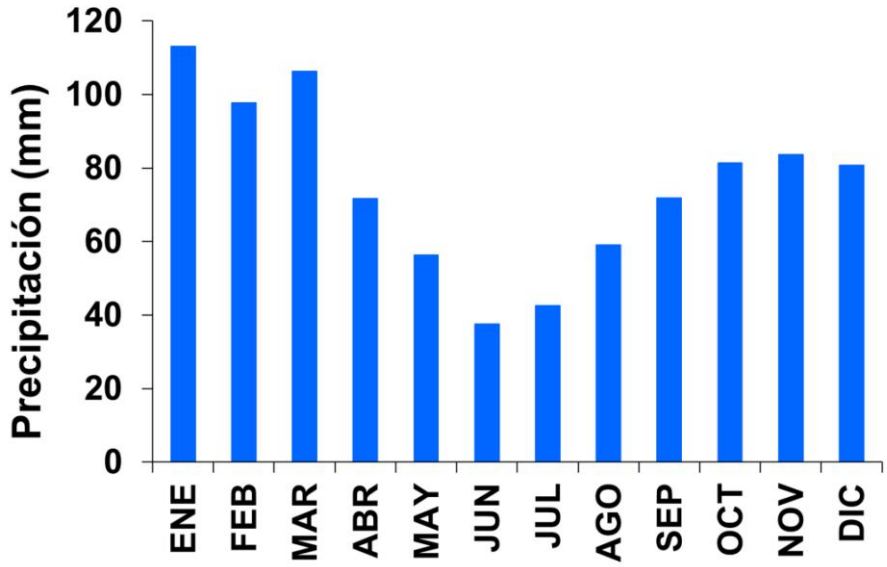

Figura 2. Precipitación media mensual para la estación Azul (SMN) periodo de 2006-2014.
Figure 2. Average monthly precipitation for the Azul station (SMN) from period 2006-2014.

MATERIALES Y MÉTODOS

Para este estudio se contó con tres estaciones convencionales con datos completos en un periodo de análisis desde 2006 al 2014. La estación de Azul perteneciente al Servicio Meteorológico Nacional (SMN) y las estaciones de Palenque y Trapenses son privadas. Se utilizaron también seis estaciones con datos faltantes; Chillar, La Germania, Cerro del Águila, La Chiquita, Siempre Amigos y La Firmeza, que pertenecen a la red del sistema alerta temprana para inundaciones de la ciudad de Azul, gestionada por el Instituto de Hidrología de Llanuras (IHLLA). Son estaciones automáticas y disponen de telemetría por radio (Cazenave y Vives, 2014).

Los datos de precipitación diaria se toman a las 9 am. para todas las estaciones convencionales, mientras que las estaciones automáticas registran cada minuto y posteriormente se calcula el acumulado diario. En la Tabla 1 se observa la cantidad de datos y el período de datos faltantes para cada estación.

Estación	Datos observados		Datos faltantes		Período de datos faltantes	
	N° de datos	Porcentaje	N° de datos	Porcentaje	Desde	Hasta
Azul	3287	100	0	0	-	-
Palenque	3287	100	0	0	-	-
Trapenses	3287	100	0	0	-	-
Chillar	2819	85,8	468	14,2	1/8/2008	12/11/2009
La Germania	2505	76,2	782	23,8	1/8/2008	22/09/2010
Cerro Del Águila	2505	76,2	782	23,8	1/8/2008	22/09/2010
La Chiquita	2505	76,2	782	23,8	1/8/2008	22/09/2010
Siempre Amigos	2505	76,2	782	23,8	1/8/2008	22/09/2010
La Firmeza	2819	85,8	468	14,2	1/8/2008	12/11/2009

Tabla 1. Cantidad de datos precipitación diaria y período de datos faltantes para cada estación.
Table 1. Amount of daily precipitation data and missing data period for each station.

Con la finalidad de determinar las inconsistencias en los datos de las estaciones de referencia, las cuales pueden ser originadas por un cambio en la ubicación del instrumental, una variación en las condiciones del lugar de medición o un cambio del observador, se contrastaron con los datos registrados por las estaciones de la red del IHLLA a partir del método de doble acumulación y gráficas de regresión lineal. Se determinó también la confiabilidad de las estaciones referencia por medio del estudio de Varni y Custodio (2013).

Es importante mencionar que los datos de precipitación de la red del IHLLA se utilizan para el sistema de alerta temprana de inundaciones en la ciudad de Azul (Cazenave y Vives, 2014) y han sido evaluados para diferentes estudios hidrometeorológicos como Vénere et al. (2004) y Guevara (2015).

Según Orsolini et al. (2008), el método de doble acumulación se basa en el análisis de la consistencia de datos de una estación *Y* por comparación con los datos confiables de otra estación de referencia *X*, si los valores acumulados de la estación *Y* son proporcionales a los valores acumulados *X* se obtiene una recta, pero si la recta cambia su pendiente es porque los datos son inconsistentes.

A continuación, se describen los distintos métodos de relleno de datos faltantes y se discuten las ventajas y desventajas de las metodologías utilizadas para rellenar datos diarios de precipitación.

Método Regresión Lineal (MRL)

Este método se basa en la correlación cruzada entre dos estaciones, una con datos faltantes y otra con datos confiables (Romero y Casimiro 2015; Camarillo et al., 2010; Gutiérrez et al., 2004). Para que este método de resultados aceptables, se analiza un periodo de datos completos entre las dos estaciones, se obtiene la línea de tendencia con su ecuación. El resultado es una ecuación tipo $Y = ax + b$ y se obtiene un coeficiente de determinación R^2, el cual nos indica la similitud entre los datos de la estación que se va a rellenar con los datos de referencia.

Luego se sustituye la ecuación de la siguiente forma.

$$Px = (a \times Pc) + b \tag{1}$$

Donde Px es la precipitación estimada para el día j, Pc el dato registrado por la estación de confianza (mm), a la pendiente de la recta y b el punto de intersección con el eje y.

No puede ser aplicado indiscriminadamente, dado que es necesario saber si la calidad del ajuste es buena, por eso se aplica anteriormente gráficas de regresión lineal y dobles acumulaciones para analizar el ajuste de la estación de referencia. La regresión lineal, es sugerida por WMO (1983). Según Alfaro y Pacheco (2000), una desventaja del método es que raras veces se observa una relación lineal perfecta debido a que los fenómenos que se estudian en climatología son usualmente multivariables.

Método de razones de distancia (MRD)

Este método se basa en la relación que tiene la distancia en planta en kilómetros con la precipitación, entre la estación que vamos a rellenar con dos estaciones de confianza alineadas, y solo es confiable para zonas llanas, no siendo adecuado para zonas montañosas (Romero y Casimiro 2015; Pizarro, 1993).

La ecuación para determinar el dato faltante es la siguiente:

$$Px = Pa + a \times \frac{Pb - Pa}{a + b} \tag{2}$$

Donde Px es el dato faltante de precipitación, Pa el dato registrado de precipitación estación de confianza a (mm), Pb el dato registrado de precipitación estación de confianza b (mm), a la

distancia entre la estación con dato faltante y la estación de confianza a y b la distancia entre la estación con dato faltante y la estación de confianza b .

Método coeficientes de correlación con estaciones vecinas (MRC)

Este método calcula el dato faltante de precipitación diaria a partir de correlación de Pearson entre la estación que se va a rellenar, con *n* estaciones de confianza (Tapia et al., 2009; Unesco, 1982).

$$Px = \frac{Pxa \times rxa + Pxb \times rxb + Pxn \times rxn}{rxa + rxb + rxn} \quad (1)$$

Donde Px es el dato faltante de precipitación, Pxa el dato de precipitación registrado estación de confianza a (mm), Pxb el dato de precipitación registrado estación de confianza b (mm), Pxn el dato de precipitación registrado estación de confianza n (mm), rxa el coeficiente de correlación de Pearson para toda la serie entre la estación de confianza a y la estación con dato faltante, rxb el coeficiente de correlación de Pearson para toda la serie entre la estación de confianza b y la estación con dato faltante y rxn el coeficiente de correlación de Pearson para toda la serie entre la estación de confianza n y la estación con dato faltante.

Método de la razón promedio (MRP)

Este método implementado por Tapia et al. (2009), Pizarro y Ramírez (2003), Ponce (1989), Unesco (1982) y Linsley et al. (1977), estima el dato de precipitación faltante, como el promedio de la precipitación ocurrida en n estaciones diferentes confiables, este método requiere que la precipitación anual difiera en menos del 10% con la estación que tiene datos faltantes. El dato faltante se determina con la siguiente expresión:

$$Px = \frac{Ppx}{n} \times \left(\frac{Pa}{Ppa} + \frac{Pb}{Ppb} + \ldots \frac{Pn}{Ppn} \right) \quad (4)$$

Donde Px es el dato faltante de precipitación, Pa el dato de precipitación registrado estación de confianza a (mm), Pb el dato de precipitación registrado estación de confianza b (mm), Pn el dato de precipitación registrado estación de confianza n (mm), Ppa el promedio de precipitación anual para toda la serie estación de confianza a , Ppb el promedio de precipitación anual para toda la serie estación de confianza b , Ppn el promedio de precipitación anual para toda la serie estación de confianza n , Ppx el promedio de precipitación anual para toda la serie estación con datos faltante y n el número de estaciones.

Método inverso de la distancia al cuadrado (MIDW)

Según Diaz et al. (2008), este método utiliza una asignación de peso más grande al punto más cercano y este peso disminuye a medida que aumenta la distancia, dependiendo del coeficiente de potencia. Para este caso se utilizó el inverso de la distancia al cuadrado, es decir β =2.

$$wi = \frac{N\left(\frac{1}{Di}\right)^{\beta}}{\sum_{i=1}^{N}\left(\frac{1}{Di}\right)^{\beta}} \quad (5)$$

Donde N es el número de total de estaciones, Di la distancia entre el sitio a estimar y la estación i y el coeficiente de potencia

El valor del dato faltante se estima en base a un promedio ponderado de los datos medidos en cada estación donde se le asigna un peso Wi dependiendo de su localización, la ecuación resultante es:

$$Px = \sum_{i=1}^{N} WiPi \quad (6)$$

Donde Px es el dato faltante de precipitación, Pi el dato medido en la estación i (mm), Wi el peso de la estación (ec.5) y i la representa el punto estimado en el sitio e

Este método es aplicado en diferentes estudios por Kim y Ryu (2016), Campozano et al. (2015), Vicente et al. (2010), De Silva et al. (2007) y Ahrens (2006), es clasificado como un método determinístico, y ha demostrado su eficiencia y credibilidad para zonas llanas y de montaña. Según Cheng y Liu (2012), este método obtiene mejores resultados para periodos secos que húmedos.

Método de cadenas de Márkov por distribución exponencial (MKV)

Es un método estocástico en el que la probabilidad de que ocurra un evento de precipitación depende del estado húmedo o seco del día anterior. Cuando se genera un día húmedo se utiliza una distribución exponencial para generar la cantidad de precipitación.

La metodología de rellenar datos de precipitación por medio de las cadenas de Márkov de primer orden se ha aplicado en diferentes estudios como Telesca et al. (2017); Mehan et al. (2017); Baigorria y Jones (2010); Cowden et al. (2008); Schuol y Abbaspour (2007). El MKV es un modelo dinámico dependiente del tiempo y que no tiene memoria, utiliza el último evento para condicionar las posibilidades de los eventos futuros. Con esta información se puede predecir el comportamiento del sistema a través del tiempo.

Uno de los inconvenientes que tienen las cadenas de Márkov es que no son apropiados para algunas zonas porque tienden a generar series con pocos periodos secos. Para rellenar datos por medio de la cadena de Márkov de primer orden se requiere calcular, la ocurrencia de un día húmedo o seco a través de las siguientes dos ecuaciones:

$$Pi\left(\frac{D}{W}\right) = 1 - pi\left(\frac{W}{W}\right) \quad (7)$$

Donde $Pi\left(\frac{D}{W}\right)$ es la probabilidad de un día seco en el día i dado un día húmedo en el día i - 1 y $Pi\left(\frac{W}{W}\right)$ la probabilidad de un día húmedo en un día i dado un día húmedo en el día i – 1.

$$Pi\left(\frac{D}{D}\right) = 1 - Pi\left(\frac{W}{D}\right) \quad (8)$$

Donde $Pi\left(\frac{D}{D}\right)$ es la probabilidad de un día seco en un día i dado un día seco en el día i – 1 y $Pi\left(\frac{W}{D}\right)$ la probabilidad de un día húmedo en el día i dado un día seco en el día i - 1.

Para calcular la precipitación diaria en un día faltante, se aplica una función de distribución de probabilidad.

En este caso se aplicó la distribución exponencial. Esta distribución requiere de menos insumos y es más recomendada en áreas donde hay datos limitados sobre eventos de precipitación disponibles. La precipitación diaria es calculada por medio de la siguiente ecuación:

$$Px = \mu mon.\left(-\ln\left(rnd_1\right)\right)^{r\exp} \quad (9)$$

Donde Px es la cantidad de lluvia en un día dado i (mm), μmon el promedio de la precipitación diaria (mm) para el mes i, rnd_1 un número aleatorio entre 0 y 1, $r\exp$ un exponente que se debe establecer entre 1 y 2.

Un valor alto incrementa el número de eventos de precipitaciones extremas durante el año.

$$\mu mon = \frac{PCPMM}{PCPD} \quad (10)$$

Donde $PCPMM$ es la cantidad promedio de precipitación que cae en mes i (mm) y $PCPD$ el promedio de días con precipitación en el mes i.

Para definir un día como húmedo o seco, la cadena genera un número aleatorio rnd_1 entre 0 y 1. Este número aleatorio se compara con la correspondiente probabilidad, $Pi\left(\frac{W}{W}\right)$ o $Pi\left(\frac{W}{D}\right)$. Si el número aleatorio es igual o menor que la probabilidad de seco-húmedo, el día se define como húmedo. Si el número aleatorio es mayor que la probabilidad de seco-húmedo, el día se define como seco.

Método redes neuronales (MRN)

Una red neuronal es un método que reproduce el funcionamiento simplificado del cerebro humano, a través de un procesador distribuido en paralelo con una disposición natural a almacenar conocimiento experimental y convertirlo en disponible para su uso (Haykin, 1998).

Las redes neuronales generalmente usadas son: perceptrón multicapa y función de base radial. Su arquitectura se caracteriza porque las neuronas se agrupan en 3 capas (entrada, oculta, salida).

Recientemente se han utilizado redes neuronales como herramienta para reconstruir y pronosticar datos hidrometeorológicos como: relleno de series históricas de precipitación (Teegavarapu et al.,2017; Sattari et al., 2017; Nkuna y Odiyo, 2011; Kim y Pachepsky, 2010; Saba et al., 2008), predicción de precipitación, caudal y nivel de agua entre otras aplicaciones (Kashiwao et al., 2017; Ruigar y Golian, 2016; Rezaeianzadeh et al.,2015; Coulibaly et al.,2000).

En la Figura 3 se observa la red neuronal implementada en este estudio, la cual está formada por un conjunto interconectado de elementos simples de procesamiento de información denominados nodos.

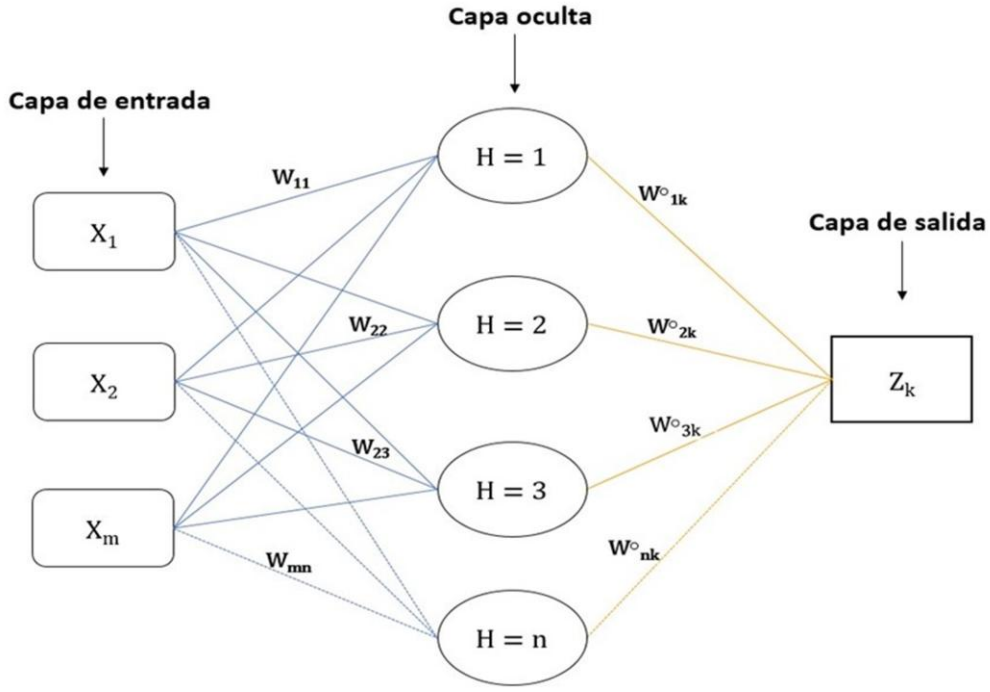

Figura 3. Estructura de la red neuronal con su capa oculta.
Figure 3. Structure of the neural network with its hidden layer.

En esta estructura se muestran *m* variables de entrada $x = \{x_1, x_2, ...x_m\}$; éstas son las variables independientes, es decir las estaciones con serie de datos completos. W corresponde a los pesos de conexión de la capa de entrada con cada nodo H de la capa oculta. La función de activación para la capa oculta es de base radial (exponencial normalizada).

W^o son los pesos de conexión entre los nodos de la capa oculta y los nodos de la capa de salida. Se presentan k variables de salida $z = \{z_1, z_2, ...z_k\}$, las cuales representan a las variables dependientes o estaciones con serie de datos faltantes. La función de activación de la capa de salida es la identidad.

El proceso de ajuste de los parámetros de la red se hace en dos fases. La primera fase o llamada fase de entrenamiento tiene el objetivo de estimar los pesos y la segunda fase la de validación consiste en comprobar si la red predice bien a partir de otros datos los cuales no fueron utilizados en la fase de entrenamiento. Para la fase uno se usó el 70 % y para la fase dos se utilizó el 30% de los datos.

Estadísticos de ajuste

Para determinar la confianza del método con el que se van a rellenar los datos faltantes se utilizan los siguientes estadísticos:

El coeficiente de determinación (R^2) es un estadístico que describe la proporción de la variabilidad total en los datos observados, explicada por la recta de regresión. Es el cuadrado del coeficiente de correlación de Pearson, que varía entre 0 y 1. Cuando los valores de R^2 son altos indican menos varianza del error y en general, los valores superiores a 0,7 se consideran aceptables. El R^2 ha sido ampliamente utilizado para la evaluación hidrológica, aunque es más sensible a los valores extremos.

$$R^2 = \frac{\sigma^2 xy}{\sigma^2 x \sigma^2 y} \quad (11)$$

Donde σxy es la covarianza de x, y, σx la desviación típica variable x y σy la desviación típica variable y

La raíz cuadrada media del error (RMSE) representa la desviación estándar de la muestra de las diferencias entre los valores calculados y los valores observados. Estas diferencias individuales se denominan residuos, cuando los cálculos se realizan sobre la muestra de datos que se utilizó para la estimación se denominan errores de predicción. El RMSE es una buena medida de precisión, pero sólo para comparar los errores de predicción para una variable particular y no entre las variables, ya que es dependiente de la escala. El valor deseado para RMSE es 0, lo cual nos indica que el método no produjo errores.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (\hat{y}t - y)^2}{n}} \tag{12}$$

Donde $\hat{y}t$ son los alores observados, y los alores calculados y n el número de datos analizados

El coeficiente de eficiencia Nash Sutcliffe (NS) es un estadístico normalizado que determina la magnitud relativa de la varianza residual ("ruido") en comparación con la variación de datos medidos (Nash y Sutcliffe, 1970). El NS indica qué tan bien se correlacionan los datos observados frente a los datos rellenados en un ajuste de línea 1:1. Este estadístico es computado por medio de la siguiente ecuación:

$$NS = 1 - \frac{\sum_{i=1}^{n} (\hat{y}t - y)^2}{\sum_{i=1}^{n} (\hat{y}t - ym)^2} \tag{2}$$

Donde $\hat{y}t$ son los valores observados, y los valores calculados y Ym el promedio de los valores observados.

El NS oscila entre los valores entre -∞ y 1, los valores > 0 generalmente son vistos como niveles aceptables de rendimiento, mientras que los valores < 0 indican que el valor medio observado es un mejor predictor que el valor simulado, lo que indica un rendimiento inaceptable.

RESULTADOS Y DISCUSIÓN

Como primer paso se construyó un análisis de curva de doble acumulación (Figura 4) y gráficos de dispersión (Figura 5) con el fin de analizar la confiabilidad de las estaciones de referencia. Este proceso se realiza para un periodo común sin datos faltantes. Se utilizaron 943 datos para el período comprendido entre 01/01/2006 al 01/08/2008, para las 3 estaciones de referencia (Azul, Palenque, Trapenses) y las 6 estaciones que hacen parte de la red de estaciones hidrometeorológicas del IHLLA (Chillar, La Germania, Cerro del Águila, La Chiquita, Siempre Amigos y La Firmeza).

Se encontró, como era de esperar por su distancia, que la estación Chillar presenta una mejor consistencia con los datos de la estación Palenque, mientras que para las estaciones La Germania y Cerro del Águila se obtuvo un buen ajuste con las tres estaciones de referencia. Sin embargo, se presenta una mejor correlación con los datos de la estación de Trapenses. Las estaciones La Chiquita, Siempre Amigos y La Firmeza, que están ubicadas en la parte media y baja de la cuenca, presentan un mejor ajuste con la estación de Azul, como se muestra en las Figuras 3 y 4.

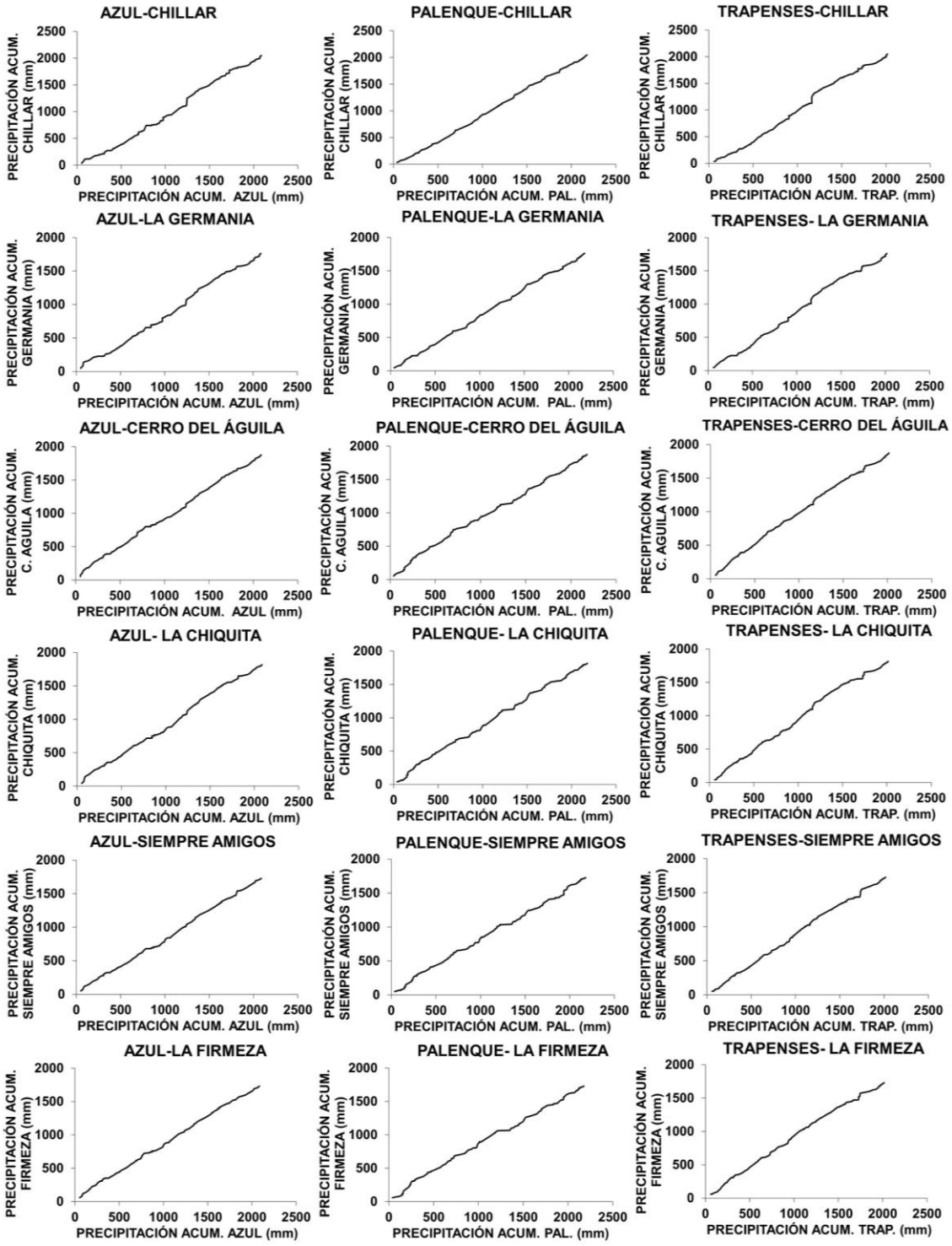

Figura 4. Curva de doble acumulación entre estaciones con datos completos y faltantes para el período comprendido entre 01/01/2006 al 01/08/2008.

Figure 4. Double accumulation curve between stations with complete and missing data for the period from 01/01/2006 to 08/01/2008.

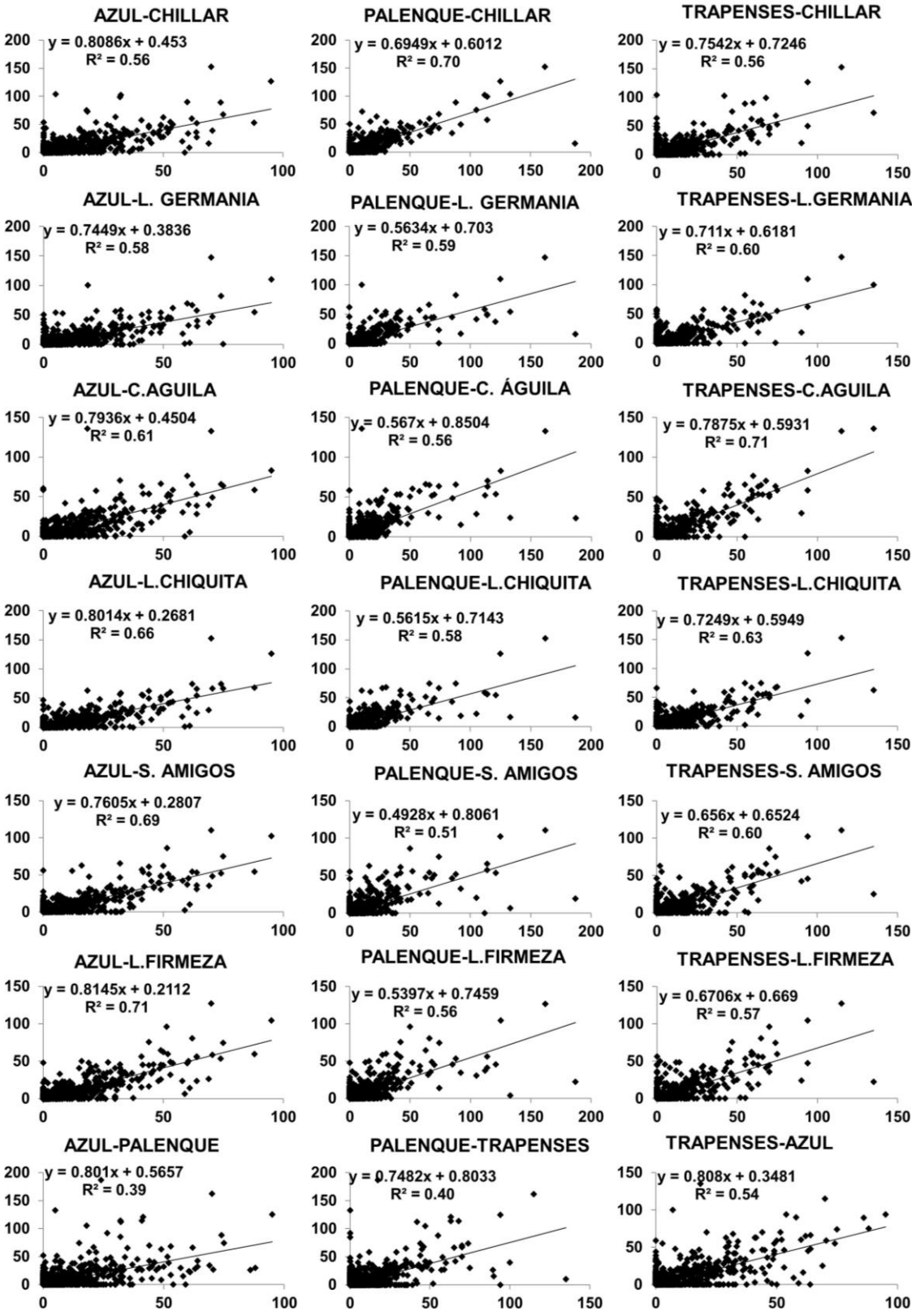

Figura 5. Regresiones lineales para el período comprendido entre 01/01/2006 al 01/08/2008 para 3 estaciones con datos completos (Azul, Palenque, Trapenses) y 6 estaciones con datos faltantes diarios (Chillar, La Germania, Cerro del águila, La Chiquita, Siempre amigos y La Firmeza.

Figure 5. Linear regressions for the period from 01/01/2006 to 08/01/2008 for 3 stations with complete data (Azul, Palenque, Trapenses) and 6 stations with missing daily data (Chillar, La Germania, Cerro del Águila, La Chiquita, Siempre Amigos y La Firmeza.

Como se observa en la Figura 3, al analizar los acumulados de los datos de precipitación diarios se encontró que son bastante confiables debido a que no han sufrido variaciones o cambios de pendiente importantes que indiquen valores erróneos.

En la Figura 4 se muestran los gráficos de dispersión con sus rectas de tendencia lineal entre las estaciones con datos faltantes (ubicados en el eje Y) y las estaciones de referencia (ubicados en el eje X). Se observa un mejor ajuste de las estaciones La Chiquita, Siempre Amigos y La Firmeza con la estación de referencia Azul R^2 (0.66-0.71), mientras que La Germania y Cerro del Águila presentaron mejores ajustes con la estación Trapenses R^2 (0.6-0.71). La parte alta de la cuenca donde se encuentra ubicada la estación Chillar presentó un mejor ajuste con la estación de referencia Palenque R^2 (0.70).

En la Figura 4 se observa que hay menor dispersión de datos con las estaciones Azul y Trapenses, lo cual confirma que estas dos estaciones serían las más confiables para realizar un relleno de datos para la cuenca del arroyo del Azul. Esto también se observa con el método de doble acumulación (Figura 3).

Para la evaluación de los resultados se analizan los datos diarios y posteriormente los acumulados mensuales por medio de estadísticos descritos anteriormente los cuales miden el ajuste de los métodos implementados.

Análisis de datos diarios

Para evaluar el rendimiento de los métodos para el relleno de datos faltantes diarios se utilizaron diferentes estadísticos como la raíz cuadrada media del error, el coeficiente Nash Sutcliffe y los coeficientes de determinación, los cuales se presentan en la Figura 6 y en la Tabla 2. Estos estadísticos se calcularon para el periodo del 2006 al 2014 sin considerar el periodo de datos faltantes.

Los métodos que obtuvieron los mejores rendimientos estadísticos en orden consecutivo fueron el MKV R^2 (0.81-0.98), NS (0.81-0.9), RMSE (0.7 mm-4.1 mm) y el MRN R^2 (0.77-0.83), NS (0.77-0.83), RMSE (3.61 mm-4.46 mm), estos dos métodos ajustaron bastante bien para todas las estaciones. Los métodos MRP Y MRC dieron resultados muy similares con El MRP se obtuvo R^2 (0.74-0.79), NS (0.74-0.79), RMSE (3.99 mm-4.38 mm); y el MRC alcanzo ajustes R^2 (0.74-0.79), NS (0.72-0.79), RMSE (4.06 mm-4.52 mm). El MRL logró un ajuste a escala diaria con R^2 (0.62-0.85), NS (0.62-0.84), RMSE (3.48 mm-5.70 mm). Los métodos MRP, MRC Y MRL lograron buenos rendimientos en la parte baja de cuenca en las estaciones de La Chiquita, La Firmeza y Siempre Amigos, mientras que las estaciones Chillar, La Germania y Cerro del Águila ubicadas en la parte alta de cuenca presentaron rendimientos más bajos.

Los métodos que obtuvieron los ajustes estadísticos más bajos y los errores más altos para el relleno de datos diarios fueron MIDW y MRD. El MIDW obtuvo R^2 (0.64-0.8), NS (0.54-0.79) y RMSE (3.9 mm-6.27 mm) produjo ajustes bajos en casi la mayoría de estaciones a diferencia de la estación de Cerro del Águila, la cual presento un buen ajuste. El MRD R^2 (0.62-0.77), NS (0.53-0.76) y RMSE (4.27 mm-5.98 mm) produjo un rendimiento inferior para todas las estaciones en comparación con todos los demás métodos probados en este estudio.

Al evaluar el desempeño del relleno de datos diarios se encontró que para las estaciones Chillar, La Germania y Cerro del Águila las cuales se encuentran ubicadas en la parte alta de la cuenca en el sistema de sierras de Tandilla, presentan errores más altos para todas la metodologías aplicadas en el estudio obteniendo un rango de error de 2.21 mm a 6.27 mm a diferencia de las estaciones La Chiquita, La Firmeza y Siempre Amigos las cuales se encuentran ubicadas en la parte intermedia-baja de la cuenca las cuales alcanzaron errores de 1 mm a 4.51 mm. Esta diferencia se da porque las estaciones de referencia correlacionan mejor con las estaciones ubicadas en la parte baja de la cuenca, donde las cantidades de lluvia entre estaciones disminuye rápidamente al aumentar la distancia.

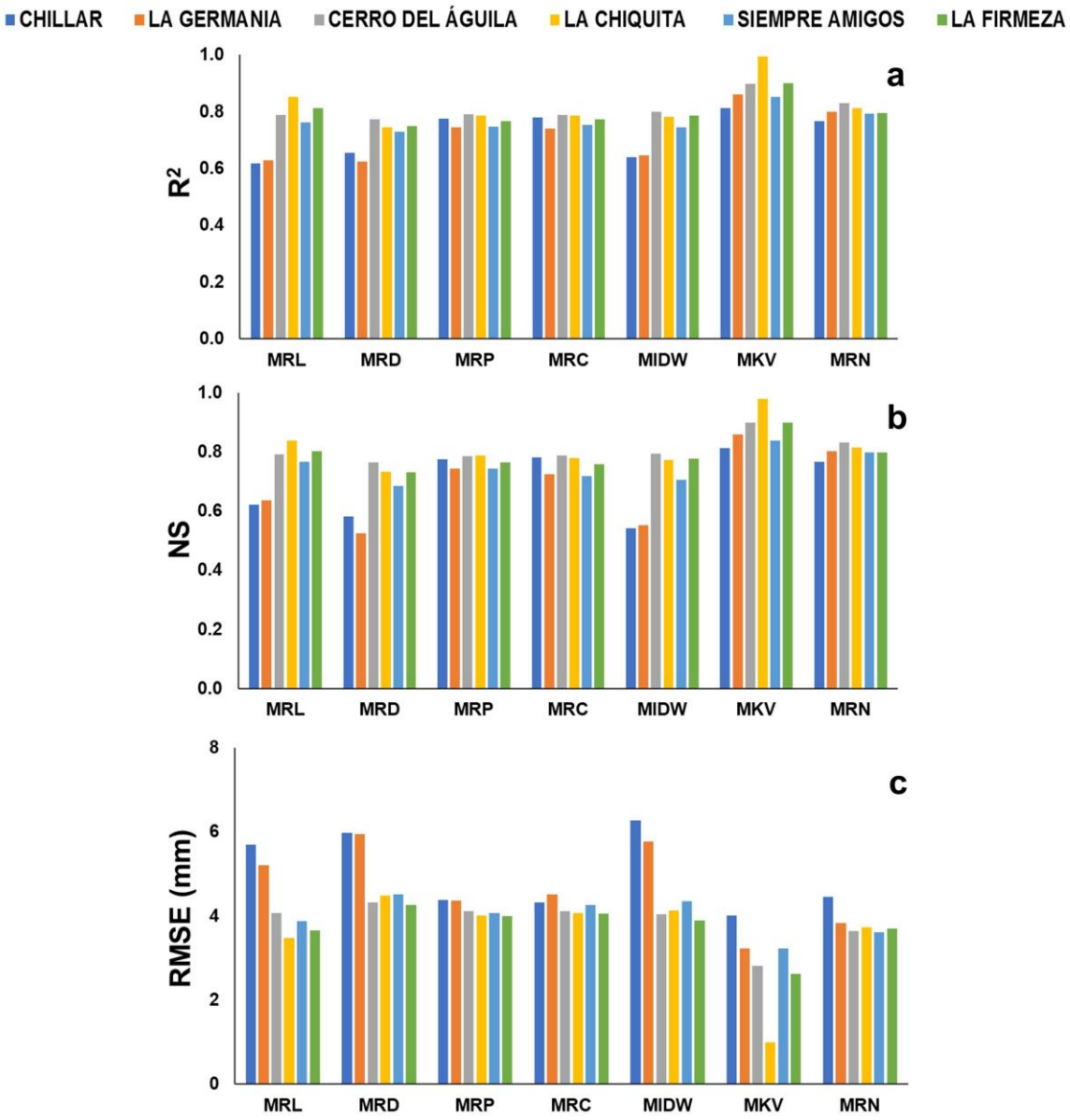

Figura 6. Estadísticos de medición del ajuste entre los datos rellenados por diferentes métodos, con los datos observados de las estaciones con datos faltantes. a. Coeficientes de determinación (R^2). b. Coeficiente Nash-Sutcliffe (NS). c. Raíz cuadrada media del error diario (RMSE).
Figure 6. Statistical measuring of the fit between the data filled by different methods, with the observed data of the stations with missing data. to. Coefficients of determination (R^2). b. Nash-Sutcliffe Coefficient (NS). c. Average square root of the daily error (RMSE)

Análisis de datos mensuales

Al analizar los acumulados mensuales a partir del relleno de datos diarios para el periodo 2006 al 2014 sin considerar el periodo de datos faltantes por medio de las diferentes metodologías (Figura 7), se encontró que los métodos que obtuvieron los coeficientes de determinación más altos con todas las estaciones fueron el MKV, con un promedio (R^2>0.89), el MRN (R^2>0.88), MRP (R^2>0.82), MRC (R^2>0.82), MRD (R^2>0.82). Los coeficientes de determinación más bajos lo presentaron el MIDW (R^2>0.80) y MRL (R^2>0.77).

Método		MRL	MRD	MRP	MRC	MIDW	MKV	MRN
CHILLAR	R^2	0,62	0,66	0,77	0,78	0,64	0,81	0,77
	NS	0,62	0,58	0,78	0,78	0,54	0,81	0,77
	RMSE	5,70	5,98	4,38	4,32	6,27	4,01	4,46
LA GERMANIA	R^2	0,63	0,62	0,74	0,74	0,65	0,86	0,80
	NS	0,64	0,53	0,75	0,73	0,55	0,86	0,80
	RMSE	5,21	5,95	4,36	4,52	5,77	3,22	3,84
CERRO DEL ÁGUILA	R^2	0,79	0,77	0,79	0,79	0,80	0,90	0,83
	NS	0,79	0,76	0,79	0,79	0,79	0,90	0,83
	RMSE	4,07	4,33	4,12	4,11	4,04	2,81	3,64
LA CHIQUITA	R^2	0,85	0,75	0,79	0,79	0,78	0,98	0,81
	NS	0,84	0,73	0,79	0,78	0,77	0,98	0,82
	RMSE	3,48	4,48	4,00	4,07	4,14	1,00	3,73
SIEMPRE AMIGOS	R^2	0,76	0,73	0,75	0,75	0,75	0,85	0,79
	NS	0,77	0,68	0,74	0,72	0,71	0,84	0,80
	RMSE	3,88	4,51	4,07	4,26	4,35	3,23	3,61
LA FIRMEZA	R^2	0,81	0,75	0,77	0,77	0,79	0,90	0,80
	NS	0,80	0,73	0,77	0,76	0,78	0,90	0,80
	RMSE	3,66	4,27	3,99	4,06	3,90	2,62	3,71

Tabla 2. Estadísticos de medición del ajuste entre los datos rellenados por diferentes métodos, con los datos observados de las estaciones de la red del IHLLA. a. Coeficientes de determinación (R^2). b. Coeficiente Nash-Sutcliffe (NS). c. Raíz cuadrada media del error diario (RMSE).
Table 2. Statistics of measurement of the fit between the filled data by different methods, with the observed data of the stations of the IHLLA network. to. Determination coefficients (R^2). b. Nash-Sutcliffe Coefficient (NS). c. Average square root of the daily error (RMSE).

Las correlaciones tanto diarias como mensuales fueron significativas con un valor de $P<0.01$. Al analizar los acumulados mensuales a través de la serie con relleno de datos diarios por diferentes métodos, se encontró que estos acumulados presentaban valores más altos de R^2, debido a que hay una menor variabilidad de los datos observados.

Cuando se analiza el error por medio de RMSE (Figura 8), se encontró que los métodos que tienen los residuos más bajos son el MKV con un rango (1.47-38.86 mm), el MRN (6.07-37.72 mm), (MRP 3.16-44.91 mm), MRC (3.08-47.05 mm), MIDW (2.71-90.48 mm). Los métodos que obtuvieron los errores más altos fueron el MRD (4.04-85.47 mm) y MRL con un RMSE (4.73-87.98 mm).

Al analizar los acumulados mensuales a partir del relleno de datos diarios de precipitación se encontró que los métodos MIDW, MRL, MRD son métodos que sobrestiman la precipitación mensual observada presentando errores que pueden llegar a ser mayores de 80 mm.

Al evaluar el rendimiento de los acumulados mensuales a partir del relleno de datos diarios se encontró que las diferentes metodologías dan mejores resultados en invierno y otoño con las menores precipitaciones a comparación de verano y primavera donde obtuvieron ajustes más bajos.

En la región Pampeana para los meses de verano la mayoría de las precipitaciones son de origen convectivo, debido a que estas se forman al ascender el aire que ha sido fuertemente calentado in situ, lo cual es propio de áreas de gran insolación y humedad. Las tormentas convectivas provienen de nubes cumuliformes y tienden a ser lluvias localizadas, de corta duración y alta intensidad, por lo que ciertamente es difícil obtener buenos ajustes en los meses más cálidos. Mientras que para los meses invernales la escasa precipitación fundamentalmente es de origen estratiforme (Sarochar et al., 2005) y ocupan enormes extensiones, generalmente duran varias horas, y son de baja intensidad.

En la Figura 9 muestra el diagrama de caja de la precipitación promedio mensual calculada a partir del relleno de datos diarios sin considerar el periodo de datos faltantes. De esta forma se comparó la distribución de los datos observados con respecto a los datos calculados.

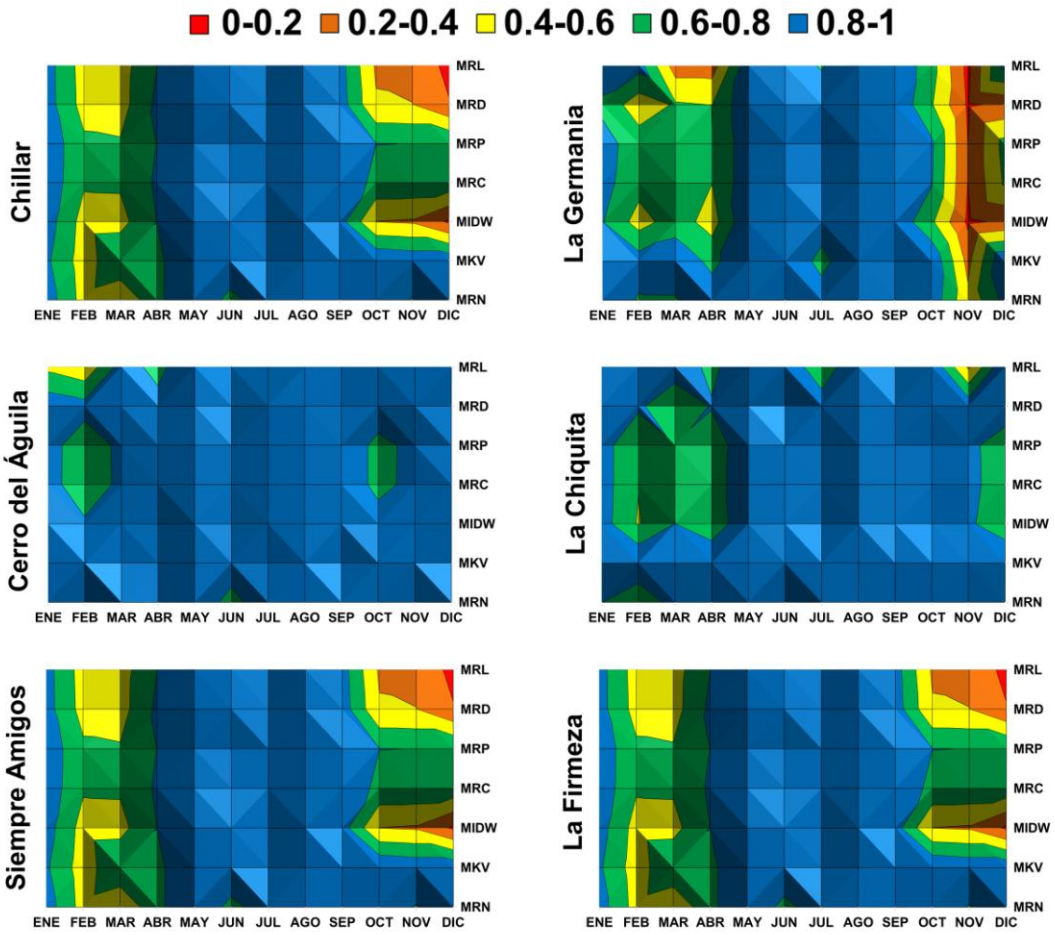

Figura 7. Coeficientes de determinación mensual entre los datos calculados por los diferentes métodos para cada estación con datos faltantes.
Figure 7. Coefficients of monthly determination between the data calculated by the different methods for each station with missing data.

Para las estaciones Chillar, La Germania y La Firmeza los datos observados son bastantes cercanos a la simetría ya que la mediana está cerca del centro de la caja, mientras que para las estaciones Cerro del Águila, La Chiquita y Siempre Amigos la locación de la mediana en las partes superiores de las cajas indica una tendencia hacia la asimetría negativa.

Para las estaciones Chillar, La Germania, La Firmeza, La Chiquita y Siempre Amigos, los datos medidos se agrupan alrededor de 70 mm aproximadamente. En todos los casos la distribución de los datos observados es similar a la obtenida por la mayoría de los métodos calculados, exceptuando el MRL para el cual los datos tienen una menor distribución. En el caso particular de la estación La Germania todos los métodos presentaron una distribución similar a la observada con valores cercanos a 76 mm, Sin embargo, todos los métodos sobrestiman los valores extremos.

Análisis del relleno de datos de precipitación

En esta sección se presenta un análisis del relleno de datos faltantes por medio del método MKV el cual obtuvo el mejor rendimiento. En la Figura 10, se observa el periodo de relleno de datos faltantes para todas las estaciones de la red del IHLLA comparado con la estación más confiable con datos completos en este estudio: la estación perteneciente al Servicio Meteorológico Nacional (SMN) en Azul. Se encontró que el relleno de datos representa de manera adecuada la variabilidad climática observada. Cabe aclarar que el periodo de relleno se da en una época de sequias pluviométricas.

En la Figura 11 se vislumbra la serie de precipitación diaria completada para cada estación de la red del IHLLA a partir del relleno del MKV para el periodo 2006-2014. En las estaciones (La

Germania, Cerro del Águila, La Chiquita y Siempre amigos), el periodo de relleno fue desde el 1/8/2008 al 22/9/2010, mientras que para las estaciones de Chillar y La Firmeza el periodo de relleno es del 8/1/2008 al 12/11/2009.

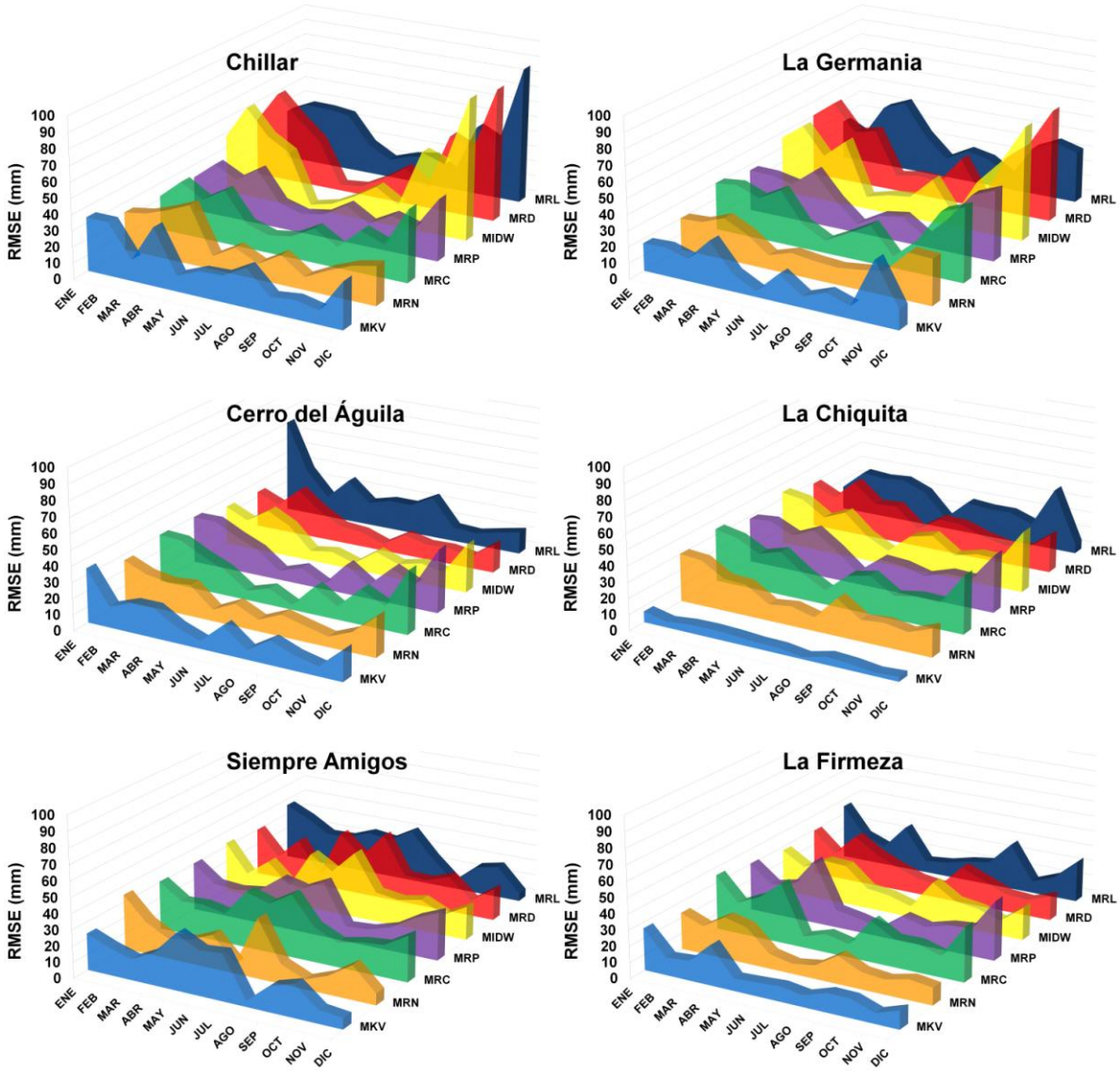

Figura 8. Raíz cuadrada media del error mensual entre los datos calculados por los diferentes métodos para cada estación con datos faltantes.
Figure 8. Average square root of the monthly error between the data calculated by the different methods for each station with missing data.

A partir del relleno de datos faltantes por medio del MKV para el periodo 2006-2014, se realizó una interpolación espacial por medio del método Anudem (Hutchinson, 1989) para todas las estaciones a escala promedio anual, a fin de analizar el patrón espacial de la precipitación en la cuenca superior del arroyo del Azul (Figura 11). Se encontró que se presentan mayores precipitaciones en parte norte donde se ubica el sistema de sierras de Tandilia, debido a que en estas zonas se generan precipitaciones de tipo convectivas-orográficas. En la zona intermedia-baja de la cuenca presentan precipitaciones más bajas con diferencias de alrededor de 60 mm, lo cual puede deberse a que en esta zona hay una mayor influencia precitaciones de tipo convectivas.

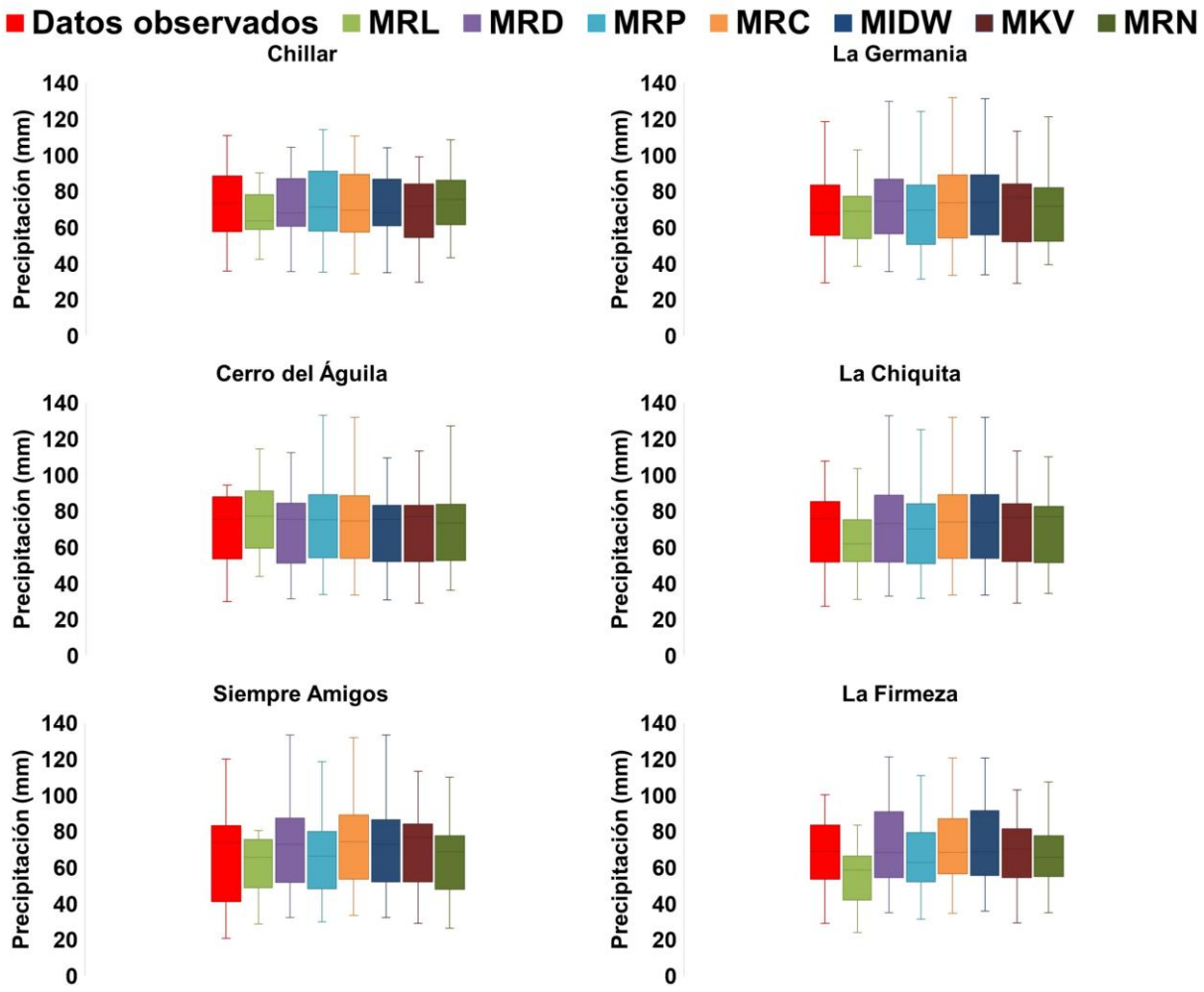

Figura 9. Comparación de la distribución entre la precipitación promedio mensual observada y calculada con los diferentes métodos de relleno para las estaciones de la red del IHLLA. Periodo 2006-2014, sin considerar los datos faltantes.

Figure 9. Comparison of the distribution between the average monthly precipitation observed and calculated with the different filling methods for the stations of the IHLLA network. Period 2006-2014, without considering the missing data.

CONCLUSIONES

Para la cuenca superior del arroyo del Azul, ajustó mejor la parte alta de la cuenca con los datos de Palenque, la parte media de la cuenca con los datos de estación Trapenses, mientras que en cuenca baja el ajuste más alto fue con los datos de la estación Azul, la cual pertenece al SMN.

Al analizar las diferentes metodologías para el relleno de datos faltantes diarios, se encontró que los modelos basados en redes probabilísticas (MKV) y neuronales (MRN) son modelos apropiados para rellenar datos de precipitación en zonas de llanura. Estos métodos son conceptualmente superiores a otros enfoques probados en el presente estudio debido a que utilizan técnicas estadísticas modernas de inferencia y estimación las cuales produjeron estimaciones relativamente más precisas tanto a escala diaria como a escala mensual.

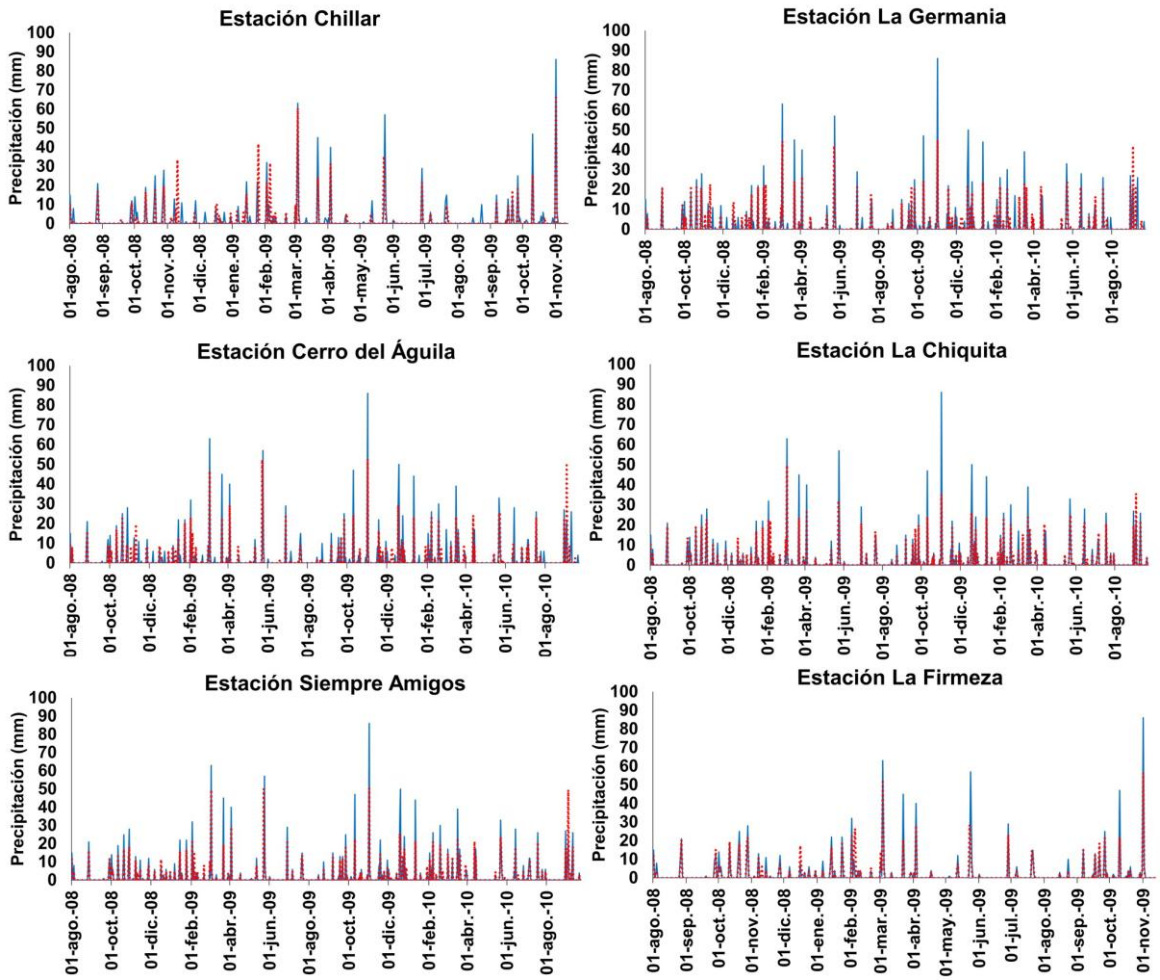

Figura 10. Comparación del periodo rellenado de datos faltantes para todas las estaciones de la red del IHLLA (línea punteada color rojo) con la estación hidrometeorológica Azul SMN (línea color azul)
Figure 10. Comparison of the period filled with missing data for all the stations of the IHLLA network (dotted line red color) with the hydrometeorological station Azul SMN (blue line).

El método que obtuvo los ajustes más bajos fue razones de distancia (MRD), debido a que necesita dos estaciones de referencia, las cuales tienen que estar ubicadas en línea recta con la estación con datos faltantes, situación que no se da con las estaciones de referencia en la cuenca.

Aunque muchos estudios recomiendan el relleno de datos de precipitación por medio de MIDW, MRL y MRD, al analizar los acumulados mensuales a partir del relleno de datos diarios de precipitación no se recomienda utilizarlos debido a que sobrestiman la precipitación mensual observada.

Los métodos que se aplicaron en el estudio tienen un mejor ajuste en la época de otoño-invierno con menores precipitaciones en comparación con el periodo primavera-verano en donde se obtuvieron ajustes más bajos debido a que en estas épocas se presentan tormentas convectivas, las cuales son localizadas y con intensidades muy altas de precipitación.

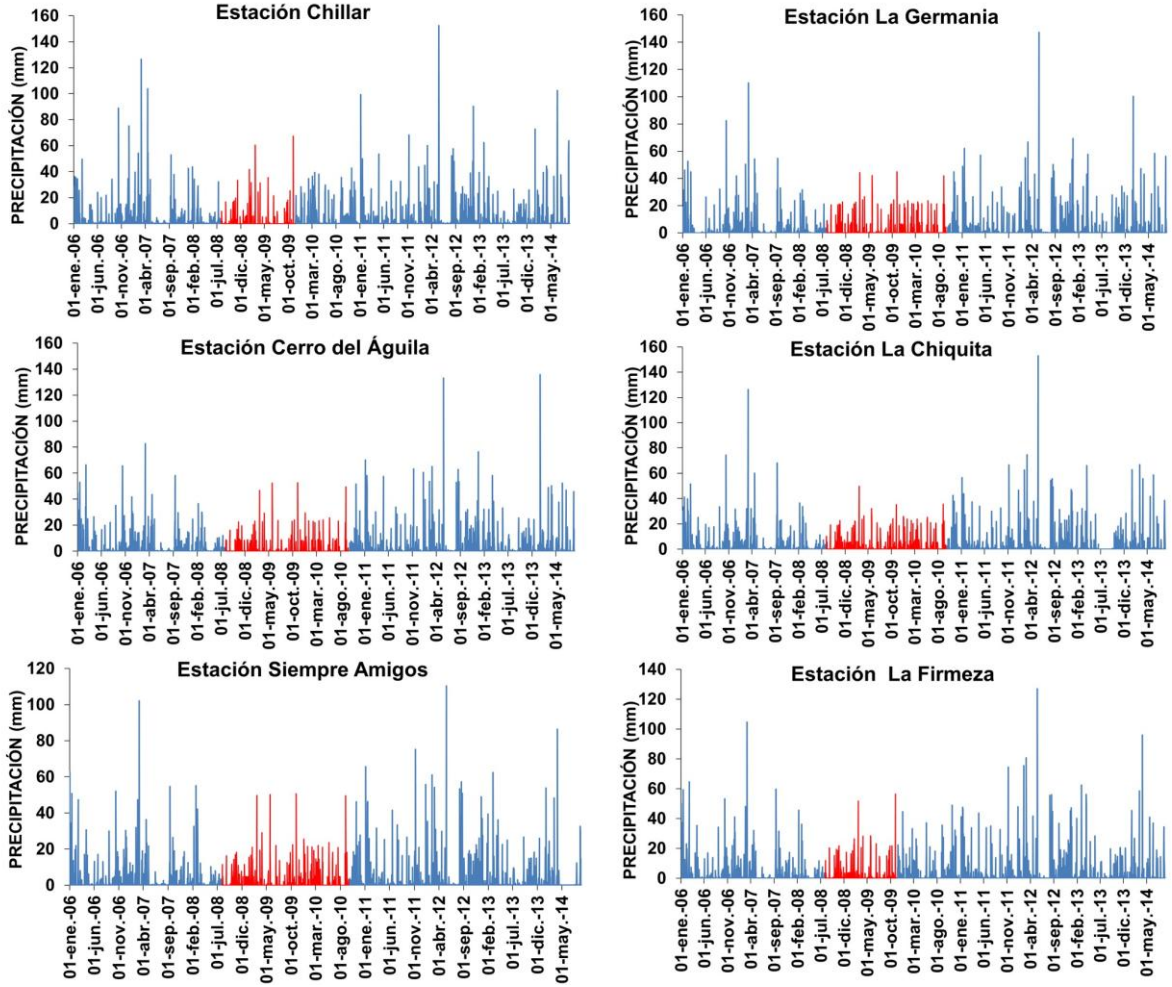

Figura 11. Serie de precipitación completada a escala diaria para el periodo 2006-2014 por el MKV para las estaciones de la red del IHLLA. Datos observados diarios para cada estación (color azul) y datos rellenados para cada estación (color rojo) periodo 2006-2014.
Figure 11. Precipitation series completed on a daily scale for the period 2006-2014 by the MKV for the stations of the IHLLA network. Data observed daily for each station (blue color) and data filled for each station (red color) period 2006-2014.

En este estudio se enfatiza la necesidad de contar con buenos métodos para el relleno de datos diarios de precipitación, a fin de representar los procesos del ciclo hidrológico de la manera más real posible. Se recomienda que las validaciones de los diferentes métodos de relleno de datos diarios de precipitación se realicen a nivel local para tomar mejores decisiones en cuanto a qué método es mejor utilizar. Este estudio es de gran utilidad para la toma de decisiones sobre las metodologías a aplicar cuando se requiere contar con series de datos sin faltantes, al menos para estaciones que están ubicadas en regiones con características similares a la presentada.

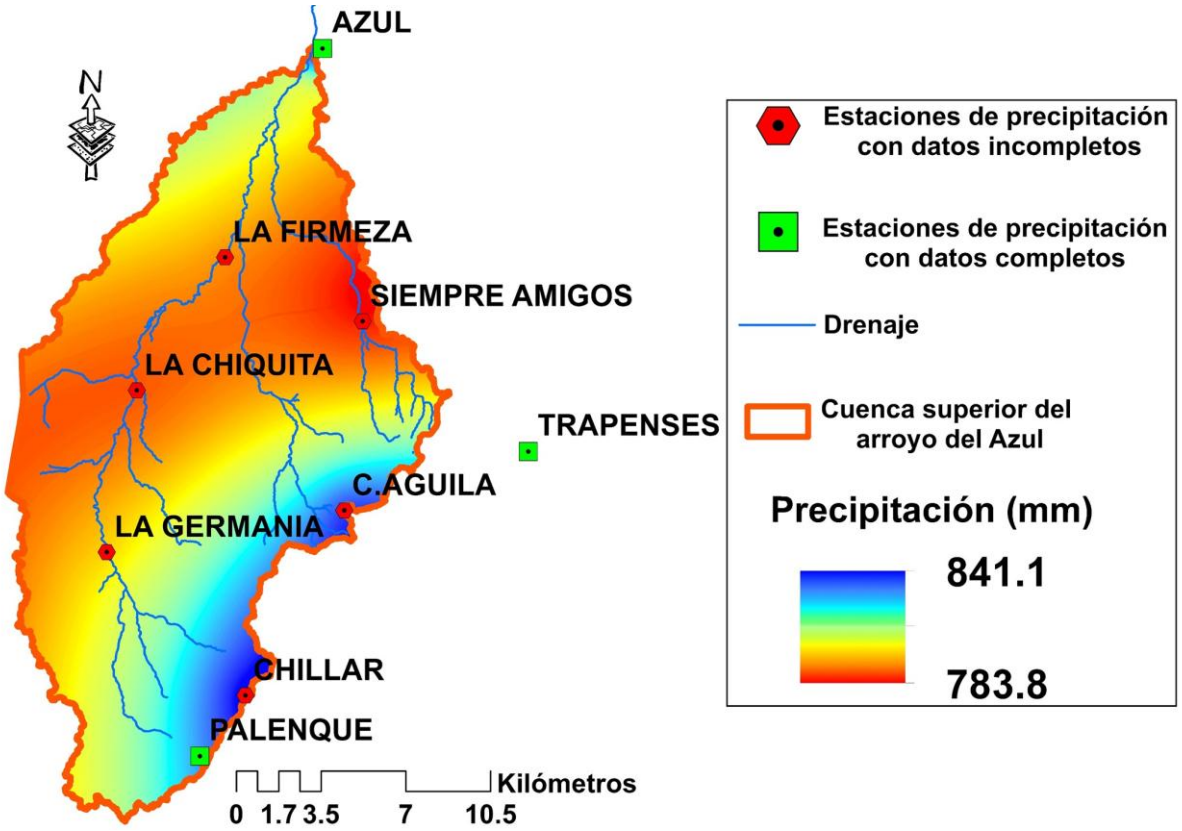

Figura 12. Distribución espacial de la precipitación promedio anual con el periodo rellenado por medio del MKV para la cuenca superior del arroyo del Azul.
Figure 12. Spatial distribution of average annual rainfall with the period filled in by means of the MKV for the upper basin of the Azul stream.

Agradecimientos: Los autores agradecen al Instituto de Hidrología de Llanuras (IHLLA) por proporcionar los datos pluviométricos y a los revisores anónimos por sus importantes aportes.

REFERENCIAS

Ahrens, B., (2006). Distance in spatial interpolation of daily rain gauge data. Hydrology and Earth System Sciences Discussions, 10(2): 197-208.

Alfaro, R. y R. Pacheco, (2000). Aplicación de algunos métodos de relleno a series anuales de lluvia de diferentes regiones de Costa Rica. Tópicos Meteorológicos y Oceanográficos, 7(1): 19-26.

Almazroui, M., M.N. Islam, F. Saeed, A.K. Alkhalaf, R. Dambul, (2017). Assessing the robustness and uncertainties of projected changes in temperature and precipitation in AR5 Global Climate Models over the Arabian Peninsula. Atmospheric Research 194: 202-213.

Antico, P. y N. Sabbione, (2010). Variabilidad temporal de la precipitación en la ciudad de La Plata durante el periodo 1909-2007: tendencias y fluctuaciones cuasiperiódicas. GEOACTA, 35: 44-53.

Bai, Y., J. Wu, Q. Xing, Q. Pan, J. Huang, D. Yang, X. Han, (2008). Primary production and rain use efficiency across a precipitation gradient on the Mongolia plateau. Ecology, 89(8): 2140-2153.

Baigorria, G.A. and J.W. Jones, (2010). GIST: A stochastic model for generating spatially and temporally correlated daily rainfall data. Journal of Climate, 23(22): 5990-6008.

Barros, V.R., J.A. Boninsegna, I.A. Camilloni, M. Chidiak, G.O. Magrín, M. Rusticucci, (2015). Climate change in Argentina: trends, projections, impacts and adaptation. Climate Change, 6(2): 151-169.

Barros, V.R., M.E. Doyle, I.A. Camilloni, (2008). Precipitation trends in southeastern South America: relationship with ENSO phases and with low-level circulation. Theoretical and Applied Climatology, 93(1): 19-33.

Beck, H.E., A.I. Van Dijk, V. Levizzani, J. Schellkens, D.J. Miralles, B. Martens, A. De Roo, (2017). MSWEP: 3-hourly 0.25 global gridded precipitation (1979-2015) by merging gauge, satellite, and reanalysis data. Hydrology and Earth System Sciences, 21(1): 589-615.

Bonsal, B.R., C. Cuell, E. Wheaton, D.J. Sauchyn. E. Barrow, (2017). An assessment of historical and projected future hydro-climatic variability and extremes over southern watersheds in the Canadian Prairies. International Journal of Climatology, 37: 3934-3948.

Camarillo, J., F. Álvarez, A. Aguilar, M. López, (2010). Metodología de cálculo de valores de precipitación diaria en estaciones meteorológicas automáticas para la obtención de indicadores climáticos mensuales. Congreso Nacional de Tecnologías de la Información Geográfica, pp 679-690.

Canziani, O., (2008). Cuenca del Salado: 100 años de lluvias. Revista Hydria 18.

Campozano, L., E. Sánchez, A. Aviles, E. Samaniego, (2015). Evaluation of infilling methods for time series of daily precipitation and temperature: The case of the Ecuadorian Andes. Maskana, 5(1): 99-115.

Cazenave, G. y L. Vives, (2014). Predicción de inundaciones y sistemas de alerta: Avances usando datos a tiempo real en la cuenca del arroyo del Azul. Revista de Geología Aplicada a la Ingeniería y al Ambiente, 33: 83-91.

Chen, C.J., S.U. Senarath, I.M. Dima-West, M.P. Marcella, (2017). Evaluation and restructuring of gridded precipitation data over the Greater Mekong Subregion. International Journal of Climatology, 37(1): 180-196.

Chen, F. and C. Liu, (2012). Estimation of the spatial rainfall distribution using inverse distance weighting (IDW) in the middle of Taiwan. Paddy and Water Environment, 10(3). 209-222.

Correa, S.W., R.C. Dias, J.C. Espinoza W. Collischonn, (2017). Multi-decadal Hydrological Retrospective: Case study of Amazon floods and droughts. Journal of Hydrology, 549: 667-684.

Coulibaly, P., F. Anctil, B. Bobee, (2000). Daily reservoir in flow forecasting using artificial neural networks with stopped training approach. Journal of Hydrology, 230 (3-4): 244-257.

Cowden, J.R., D.W. Watkins, J.R. Mihelcic, (2008). Stochastic rainfall modeling in West Africa: parsimonious approaches for domestic rainwater harvesting assessment. Journal of Hydrology, 361(1): 64-77.

Davolio, S., F. Silvestro, T. Gastaldo, (2017). Impact of rainfall assimilation on high-resolution hydro-meteorological forecasts over Liguria (Italy). Journal of Hydrometeorology, 18: 2659-2680.

De Silva, R.P., N.D.K. Dayawansa, M.D. Ratnasiri, (2007). A comparison of methods used in estimating missing rainfall data. Journal of agricultural sciences, 3(2): 101-108.

Del Grosso, S., W. Parton, T. Stohlgren, D. Zheng, D. Bachelet, S. Prince, R. Olson, (2008). Global potential net primary production predicted from vegetation class, precipitation, and temperature. Ecology, 89(8): 2117-2126.

Díaz, L.B. and C.S. Vera, (2017). Austral summer precipitation interannual variability and trends over Southeastern South America in CMIP5 models. International Journal of Climatology, 37: 681-695.

Díaz, G., I. Cohen, R. Quiroz, J. Payán, C. Thorpe, I. Cruz, (2008). Interpolación espacial de la precipitación pluvial en la zona de barlovento y sotavento del Golfo de México. Agricultura Técnica en México, 34(3): 279-287.

Dingman, S.L., (2002). Physical Hydrology, Prentice Hall, 2nd ed. Prentice Hall.

Dye, S., P. Buckley, J. Pinnegar, (2017). Impacts of Climate Change on the Coastal and Marine Physical Environments of Caribbean Small Island Developing States (SIDS). En Report Card: Science Review, pp 1-9.

Giorgi, F., (2002). Variability and trends of sub-continental scale surface climate in the twentieth century. Part I: observations. Climate Dynamics, 18(8): 675-691.

Guevara, C., (2015). Una metodología para el manejo integral de extremos hídricos en una cuenca rural en zona de llanura. Tesis de Maestría en ecohidrología, Universidad Nacional de la Plata, Facultad de Ingeniería y Facultad de Ciencias Naturales y Museo, pp 154.

Gutiérrez, J.M., R. Cano, A.S. Cofiño, C.M. Sordo, (2004). Redes probabilísticas y neuronales aplicadas a las ciencias atmosféricas. INM, Ministerio de Medio Ambiente, Madrid.

Haykin, S., (1998). Neural Networks: A Comprehensive Foundation, 2nd ed. Macmillan College Publishing. Nueva York.

Haylock, M.R., T.C. Peterson, L.M. Alves, T. Ambrizzi, Y.M.T. Anunciação, J. Baez, V. Corradi, (2006). Trends in total and extreme South American rainfall in 1960 2000 and links with sea surface temperature. Journal of climate, 19(8): 1490-1512.

Hu, W., A. Duan, B. He, (2017). Evaluation of intra-seasonal oscillation simulations in IPCC AR5 coupled GCMs associated with the Asian summer monsoon. International Journal of Climatology, 37:476-496.
Huang, J., A.T. Islam, F. Zhang, Z. Hu, (2017). Spatiotemporal analysis the precipitation extremes affecting rice yield in Jiangsu province, southeast China. International Journal of Biometeorology, 61: 1863-1872.
Hutchinson, M.F., (1989). A new procedure for gridding elevation and stream line data with automatic removal of spurious pits. Journal of Hydrology, 106(3-4): 211-232.
IHLLA, (2003). Sistema de soporte para la gestión eficiente de los recursos hídricos en la llanura bonaerense, pp 1-267.
Kashiwao, T., K. Nakayama, S. Ando, K. Ikeda, M. Lee, A. Bahadori, (2017). A neural network-based local rainfall prediction system using meteorological data on the Internet: A case study using data from the Japan Meteorological Agency. Applied Soft Computing, 56: 317-330.
Kim, J.W. and Y.A. Pachepsky, (2010). Reconstructing missing daily precipitation data using regression trees and artificial neural networks for SWAT streamflow simulation. Journal of Hydrology, 394 (3–4): 305-314.
Kim, J.W. and J.H. Ryu, (2016). A heuristic gap filling method for daily precipitation series. Water resources management, 30(7): 2275-2294.
Kolokytha, E., S. Oishi, R.S. Teegavarapu, (2017). Sustainable Water Resources Planning and Management Under Climate Change. Singapore: Springer.
Kuligowski, R. and A. Barros, (1998). Using artificial neural networks to estimate missing rainfall data. Jawra Journal of the American Water Resources Association, 34: 1437-1447.
Li, J., J. Dodson, H. Yan, D.D. Zhang, X. Zhang, Q. Xu, J. Ni, (2017). Quantifying climatic variability in monsoonal northern China over the last 2200 years and its role in driving Chinese dynastic changes. Quaternary Science Reviews, 159: 35-46.
Linsley, R.K., M.A. Kohler, J.L. Paulhus, (1977). Hidrología para ingenieros, 2nd ed. McGraw-Hill Latinoamericana, S.A.
Maenza, R.A., E.A. Agosta, M.L. Bettolli, (2017). Climate change and precipitation variability over the western 'Pampas' in Argentina. International Journal of Climatology, 37: 445-463.
Mehan, S., T. Guo, M.W. Gitau, D.C. Flanagan, (2017). Comparative Study of Different Stochastic Weather Generators for Long-Term Climate Data Simulation. Climate, 5(2): 26.
Nash, J.E. and J.V. Sutcliffe, (1970). River flow forecasting through conceptual models part I-A discussion of principles. Journal of hydrology, 10 (3): 282-290.
Nguyen - Le, D., T.J. Yamada D. Tran-Anh, (2017). Classification and forecast of heavy rainfall in northern Kyushu during Baiu season using weather pattern recognition. Atmospheric Science Letter, 18: 324-329.
Nkuna, T.R. and J.O. Odiyo, (2011). Filling of missing rainfall data in Luvuvhu River Catchment using artificial neural networks. Physics and Chemistry of the Earth, 36 (14-15): 830-835.
Orsolini, H., E. Zimmermann, P. Basile, (2008). Hidrología procesos y métodos, segunda ed. UNR editora, Rosario, Argentina.
Pizarro, R., (1993). Elementos Técnicos de Hidrología III. En: Proyecto Regional Mayor sobre Uso y Conservación de Recursos Hídricos en Áreas Rurales América Latina y el Caribe Universidad de Talca, pp 135.
Pizarro, R. y C. Ramírez, (2003). Análisis comparativo de cinco métodos para la estimación de precipitaciones areales anuales en períodos extremos. Bosque (Valdivia), 24(3): 31-38.
Ponce, V.M., (1989). Engineering Hydrology: Princples and Practices. 1st ed. Editorial Prentice Hall.
Putnam, A.E. and W.S. Broecker, (2017). Human-induced changes in the distribution of rainfall. Science Advances, 3(5): 1-14.
Rezaeianzadeh, M., L. Kalin C.J. Anderson, (2015). Wetland Water-Level Prediction Using ANN in Conjunction with Base-Flow Recession Analysis. Journal of Hydrologic Engineering, 22(1): 1-11.
Romero, E.L. y W.L. Casimiro, (2015). Evaluación de métodos hidrológicos para la completación de datos faltantes de precipitación en estaciones de la cuenta Jetepeque, Perú. Tecnológica-ESPOL-RTE, 28(3): 42-52.
Rudd, A., V. Bell, A. Kay, (2017). National-scale analysis of simulated hydrological droughts (1891–2015). Journal of Hydrology, 550: 368-385.
Ruigar, H. and S. Golian, (2016). Prediction of precipitation in Golestan dam watershed using climate signals. Theoretical and Applied Climatology, 123(3-4): 671-682.
Saba, I., J. Ortega, F. Cedeño, (2008). Estimación de datos faltantes en estaciones meteorológicas de Venezuela vía un modelo de redes neuronales. Revista de Climatología, 8: 51-70.

Sattari, M.T., A. Rezazadeh-Joudi, A. Kusiak, (2017). Assessment of different methods for estimation of missing data in precipitation studies. Hydrology Research, 48 (4):1032-1044.
Sarochar, R.H., H.H. Ciappesoni, N.E. Ruiz, (2005). Precipitaciones convectivas y estratiformes en la Pampa Húmeda: una aproximación a su separación y aspectos climatológicos de ambas. Meteorológica, 30(1-2): 77-88.
Scarpati, E. y D. Capriolo, (2013). Sequías e inundaciones en la provincia de Buenos Aires (Argentina) y su distribución espacio-temporal. Investigaciones Geográficas, Boletín del Instituto de Geografía UNAM, 82: 38-51.
Scian, B. and J. Pierini, (2013). Variability and trends of extreme dry and wet seasonal precipitation in Argentina. A retrospective analysis. Journal Atmosfera, 26(1): 3-26.
Schuol, J. and K.C. Abbaspour, (2007). Using monthly weather statistics to generate daily data in a SWAT model application to West Africa. Ecological Modelling, 201(3): 301-311.
Schuur, E.A., (2003). Productivity and global climate revisited: the sensitivity of tropical forest growth to precipitation. Ecology, 84(5): 1165-1170.
Serrano-Notivoli, R., M. De Luis, M. Beguería, (2017). An R package for daily precipitation climate series reconstruction. Environmental Modelling & Software, 89: 190-195.
Stocker, T.F., D. Qin, G.K. Plattner, M. Tignor, S.K. Allen, J. Boschung, B.M. Midgley, (2013). IPCC, 2013: climate change 2013: the physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change.
Su, F., Y. Hong, D.P. Lettenmaier, (2008). Evaluation of TRMM Multisatellite Precipitation Analysis (TMPA) and its utility in hydrologic prediction in the La Plata Basin. Journal of Hydrometeorology, 9(4): 622-640.
Tapia, R.P., P.A. Tapia, D.A. Garrido, C.S. Pool, (2009). Evaluación de métodos hidrológicos para la completación de datos faltantes de precipitación en estaciones de la Región del Maule, Chile. Aqua- LAC, 1(2): 172-185.
Teegavarapu, R.S. and V. Chandramouli, (2005). Improved weighting methods, deterministic and stochastic data-driven models for estimation of missing precipitation records. Journal of Hydrology, 312(1): 191-206.
Teegavarapu, R.S., A., Aly, C.S. Pathak, J. Ahlquist, H. Fuelbergd, J. Hoode, (2017). Infilling missing precipitation records using variants of spatial interpolation and data-driven methods: use of optimal weighting parameters and nearest neighbour based corrections. International Journal of Climatology. International Journal of Climatology.
Teegavarapu, R.S. and A. Nayak, (2017). Evaluation of Long-term Trends in Extreme Precipitation: Implications of In-filled Historical Data Use for Analysis. Journal of Hydrology, 550: 616- 634.
Telesca, V., D. Caniani, S. Calace, L. Marotta, I.M. Mancini, (2017). Daily Temperature and Precipitation Prediction Using Neuro-Fuzzy Networks and Weather Generators. In: Gervasi O. et al. (eds) Computational Science and Its Applications – ICCSA 2017. Lecture Notes in Computer Science, 10409: 441-455.
UNESCO, (1982). Guía Metodológica para la Elaboración del Balance Hídrico de América del Sur. Oficina Regional Ciencia y Tecnología, Unesco para América Latina y el Caribe (ROSTLAC) Montevideo - Uruguay, pp 10-36.
Varni, M. y E. Custodio, 2013. Evaluación de la recarga al acuífero del Azul: 1. Análisis de las características climáticas. VIII Congreso Hidrogeológico Argentino y VI Seminario Hispano Latinoamericano sobre temas actuales de hidrología subterránea, La Plata - Argentina, pp 31-37.
Vénere, M.J., A. Clausse, D. Dalponte, P. Rinaldi, G., Cazenave, M. Varni, E. Usunoff, (2004). Simulación de Inundaciones en Llanuras. Aplicación a la Cuenca del Arroyo Santa Catalina-Azul. Mecánica Computacional, 23: 1135-1149.
Vera, C., G. Silvestri, B. Liebmann, P. González, (2006). Climate change scenarios for seasonal precipitation in South America from IPCC-AR4 models. Geophysical Research Letters, 33(13): 1-4.
Vicente, S.M., S. Beguería, J.I. López-Moreno, M.A. García-Vera, P. Stepanek, (2010). A complete daily precipitation database for northeast Spain: reconstruction, quality control, and homogeneity. International Journal of Climatolog, 30(8):1146-1163.
WMO, (1983). Guide to Climatological Practices. WMO, pp 20-100.
Wongchuig, S., R.C. Dias, J.C. Espinoza, W. Collischonn, (2017). Hydrological Retrospective of floods and droughts: Case study in the Amazon. En: EGU General Assembly Conference Abstracts, pp 891.
Yang, Y., J. Fang, W. Ma, W. Wang, (2008). Relationship between variability in aboveground net primary production and precipitation in global grasslands. Geophysical Research Letters, 35 (23): 1-4.

Yehia, A.G., K.M. Fahmy, M.A Mehany, G.G. Mohamed, (2017). Impact of extreme climate events on water supply sustainability in Egypt: case studies in Alexandria region and Upper Egypt. Journal of Water and Climate Change, 8(3): 484-494.

Recibido: Septiembre, 2017
Aceptado: Febrero, 2018