

DOI: 10.24850/j-tyca-2020-03-04

Artículos

**Detección de registros homogéneos en 16 series  
amplias de precipitación anual del Altiplano Potosino,  
México**

**Detection of homogeneous records in 16 large series of  
annual precipitation of the Potosino Plateau, Mexico**

Daniel Francisco Campos-Aranda<sup>1</sup>

<sup>1</sup>Profesor jubilado de la Universidad Autónoma de San Luis Potosí, San Luis Potosí, México, campos\_aranda@hotmail.com

Autor para correspondencia: Daniel Francisco Campos-Aranda,  
campos\_aranda@hotmail.com

**Resumen**

Los métodos de homogenización han sido desarrollados para reducir el impacto de las alteraciones no climáticas en los registros o series de datos procedentes de las estaciones climatológicas. Tales impactos proceden de *cambios* en la ubicación de la estación, en el equipo utilizado y en la técnica de medición, así como por las alteraciones que sufre su entorno. Debido a que la información histórica sobre los cambios ocurridos en la

estación climatológica por lo general no está disponible, se recurre a aplicar pruebas estadísticas para detectar puntos de quiebre que implican cambios en la media y, por lo tanto, pérdida de homogeneidad del registro. Existen dos enfoques de aplicación de las pruebas estadísticas: (1) de manera individual a cada registro por probar y (2) haciendo uso de series de referencia, formadas con base en estaciones circunvecinas al registro que se prueba. En este estudio se aplicó el primer enfoque, en una zona geográfica y haciendo uso de las cuatro pruebas siguientes: Von Neumann, Pettitt, Buishand y SNHT. Se procesaron 16 series de precipitación anual del Altiplano Potosino de México, en el periodo común de 1964 a 2016, cuya amplitud varió de 42 a 53 datos. Se trabajó con datos mensuales, y por ello se expone cómo se dedujeron los valores faltantes, y cómo se detectaron y corrigieron los máximos extremos. También se describe cómo se estimaron los valores anuales faltantes. Se encontró que 11 registros resultaron *confiables* debido a que ninguna prueba los detectó como no homogéneos. Tres registros se clasificaron como *menos confiables*, pues un máximo de dos pruebas los encontró no homogéneos y, por último, dos registros son *no confiables*, ya que un mínimo de tres pruebas los encuentra no homogéneos. El procedimiento adoptado formuló una estrategia práctica para detectar registros homogéneos de precipitación anual, cuyas ventajas se exponen en el estudio.

**Palabras clave:** homogenización, valores máximos dispersos, test de Von Neumann, test del rango de Pettitt, test del rango de Buishand, prueba SNHT, precipitación mensual y anual.

## Abstract

Homogenization methods have been developed to reduce the impact of non-climatic alterations in the records or series of data coming from climatological stations. Such impacts arise from *changes* either in the location of the station, in the equipment used or in the measurement technique, as well as from alterations suffered by the surroundings. Due to the fact that the historic information on the changes that occurred in the climatological station is generally not available, statistical tests are used to detect breaking points that imply changes in the mean and, therefore, loss of homogeneity of the record. There are two approaches to the application of statistical tests: (1) individually to each record to be tested and (2) using reference series based on neighboring stations. In this study, the first approach was applied to a specific geographical area and using the following four tests: Von Neumann, Pettitt, Buishand and SNHT. Sixteen annual precipitation series of the Potosino Plateau of Mexico were processed in the common period from 1964 to 2016, the amount of values varied from 42 to 53 data. Monthly data were used, hence, the procedures to deduce missing values and to detect and correct extreme maximums are exposed. A description of how the missing annual values were estimated is also included. It was found that 11 records were *reliable*, because no test detected them as non-homogeneous. Three records were classified as *less reliable*, because a maximum of two tests found them non-homogeneous and finally, two records are *unreliable*, since a minimum of three tests find them non-homogeneous. The adopted procedure proposes a practical strategy to detect homogeneous records of annual precipitation; its advantages are exposed at the study.

**Keywords:** Homogenization, maximum outliers values, von Neumann test, Pettitt rank test, Buishand rank test, SNHT test, monthly and annual rainfall.

Recibido: 09/10/2018

Aceptado: 22/10/2019

## Introducción

## Generalidades

Al parecer, los eventos extremos del tiempo atmosférico son cada vez más severos y frecuentes, con sus respectivos impactos devastadores; además, quizás son originados por el cambio climático. Estas dos aseveraciones requieren que los análisis estadísticos orientados a su predicción y pronóstico estén basados en datos meteorológicos y climáticos de calidad, es decir, que sean precisos y homogéneos. En la realidad, obtener registros o series amplias de datos climáticos perfectamente homogéneos es casi imposible debido al cambio inevitable

del área que rodea el sitio de la estación climatológica, el cual afecta los datos observados. La pérdida de homogeneidad ocurre cuando hay cambios en el registro generados por causas no climáticas (Guijarro, 2014; Yozgatligil & Yazici, 2016).

En resumen, todos los registros o series de datos que se miden en las estaciones climatológicas, con frecuencia sufren alteraciones que no están relacionadas con las variaciones del clima. Tales modificaciones se originan en los cambios de ubicación de la estación, al reemplazar sus instrumentos o cambiar la técnica de medición, y al alterarse las condiciones físicas que rodean al sitio. La *homogenización* es un procedimiento para detectar y corregir las alteraciones artificiales citadas. Respecto a las técnicas de homogenización, Guijarro (2014) las describe someramente y se enfoca en los paquetes de cómputo disponibles. Una serie de datos climatológicos se considera homogénea cuando las condiciones de medición en la estación no han variado con el tiempo (Beaulieu *et al.*, 2008). Guijarro (2014) destaca la importancia de los servicios meteorológicos en evitar por todos los medios cambios que afecten las mediciones climatológicas, pues ello hace innecesario el uso de los métodos de homogenización.

Las erupciones volcánicas, siendo un fenómeno natural que origina una alteración de los registros climáticos, definen puntos de cambio (*change points*) que no implican pérdida de homogeneidad; por el contrario, los puntos de quiebre (*break points*) son puntos de cambio causados por alteraciones no climáticas y por lo tanto implican pérdida de homogeneidad (Toreti *et al.*, 2011).

La homogenización de las series climáticas es una tarea difícil que debe realizarse con cuidado extremo, en especial cuando los datos adicionales o datos históricos (*metadata*) de la estación climatológica no están disponibles, lo que es un caso común. El objetivo de la homogenización es eliminar, o al menos reducir, las alteraciones no climáticas mientras se preserva la señal climática (Toreti *et al.*, 2011). Se han propuesto muchos métodos para homogeneizar las series climatológicas; una revisión completa que incluye sugerencias para su uso se puede consultar en Peterson *et al.* (1998), así como en Beaulieu, Ouarda y Seidou (2007).

Una primera clasificación de los procedimientos y/o pruebas estadísticas que detectan y corrigen la pérdida de homogeneidad establece dos grupos: el primero se basa en la disponibilidad de datos adicionales sobre la evolución histórica de la estación climatológica, conocidos como *metadatos*, y el segundo en la usencia de tales datos. Rhoades y Salinger (1993) exponen técnicas para el primer grupo bajo dos enfoques: (1) empleando estaciones circunvecinas; (2) en estaciones aisladas, o de registros antiguos o primigenios. En general, los datos históricos de la estación climatológica son esenciales para validar los cambios detectados por las pruebas y, desafortunadamente, casi nunca están disponibles.

Otra clasificación de las técnicas de homogenización las divide en *pruebas absolutas*, las cuales emplean en exclusivo la serie o registro bajo análisis, y en *pruebas relativas*, que emplean registros cercanos o correlacionados para formar la serie de referencia. Ambos enfoques son válidos y útiles, pero también exhiben desventajas. Yozgatligil y Yazici (2016) comparan siete pruebas de cada tipo y encuentran que los test

relativos tienen mejor desempeño; entre ellos, citan en primer lugar la prueba de homogeneidad normal estándar, designada SNHT (*Standard Normal Homogeneity Test*).

Syrakova y Stefanova (2009) indican que debido a la dificultad para establecer *a priori* cuáles estaciones son homogéneas, optan por seguir un esquema iterativo de mejoramiento gradual de la homogeneidad, el cual fue propuesto por Hanssen-Bauer y Forland (1994), y ha sido aplicado también por Moberg y Alexandersson (1997); González-Rouco, Jiménez, Quesada y Valero (2001), y Tuomenvirta (2001).

En cambio, Wijngaard, Klein-Tank y Können (2003) aplican y citan dos ventajas de las pruebas absolutas: (1) fácil aplicación en redes poco densas; (2) superiores a las pruebas relativas en redes con cambios simultáneos, ya que éstas no los detectan. Dhorde y Zarenistanak (2013) utilizan un enfoque mixto en el estudio de homogeneidad de 20 estaciones climatológicas, aplicando las pruebas de Pettitt y SNHT a cada serie, y la prueba SNHT para cambios graduales en la media con serie de referencia.

Ducré-Robitaille, Vincent y Boulet (2003) evaluaron ocho técnicas de homogenización de series de temperatura, empleando series simuladas homogéneas, y conteniendo uno o varios cambios en la media (*steps*). Encontraron que la mayoría tiene buen desempeño, pero dos de ellas tienden a ser más eficientes: la prueba SNHT y la técnica de la regresión lineal múltiple.

Beaulieu *et al.* (2008) compararon también ocho pruebas estadísticas para detectar falta de homogeneidad en series de precipitación anual, utilizando varios miles de series sintéticas tanto homogéneas como no homogéneas; encontraron que ningún método es

eficiente en todos los tipos de pérdida de homogeneidad, pero tres de ellos tienen un desempeño mucho mejor: el test bivariado, el método de Jaruskova y la prueba SNHT.

En México, y en específico para el estado de Veracruz, Guajardo-Panes, Granados-Ramírez, Sánchez-Cohen, Díaz-Padilla y Barbosa-Moreno (2017) han aplicado el enfoque absoluto con las pruebas SNHT, Pettitt y Buishand, bajo un esquema de verificación espacial, para seleccionar estaciones climatológicas circunvecinas a la estación que se prueba, dentro de una misma unidad climática, definida con base en el intervalo de variación del elemento del clima que se analiza. Lo anterior equivale a procesar la información climática por subregiones o zonas climáticas.

## Concepto de control de calidad

González-Rouco *et al.* (2001) han definido que los procesos de *control de calidad* de las series de datos climáticos involucran tres etapas: (1) detección y corrección de valores dispersos (*outliers*); (2) técnicas de interpolación de datos faltantes, y (3) homogenización de la serie. En realidad, la estimación de datos faltantes ocurre de manera mensual, antes de la homogenización y posterior a ésta, cuando se intenta completar series para tener un periodo común de información. Debido a que todas las series de datos climáticos son extremadamente sensibles a

la presencia de valores erróneos y datos dispersos, Eischeid, Baker, Karl y Diaz (1995) abordan la detección y corrección de datos dispersos según un enfoque objetivo de verificación temporal en el registro analizado y espacial con base en seis técnicas de interpolación de datos mensuales.

## Objetivo

Se ha formulado un enfoque práctico o estrategia de verificación de la homogeneidad en registros de precipitación anual, el cual tiene las siguientes características relevantes: (1) se trabaja por regiones o subregiones geográficas, utilizando datos mensuales; (2) se estiman datos mensuales faltantes con una técnica estadística; (3) se detalla cómo se verifican y corrigen los valores mensuales máximos extremos, por truncamiento; (4) se deducen datos anuales faltantes, con una técnica de ponderado regional; (5) se utiliza un enfoque de contraste absoluto, empleando cuatro pruebas estadísticas (Von Neumann, Pettitt, Buishand y SNHT), y (6) se revisa gráficamente cada registro que no fue homogéneo o confiable, y se corrige la pérdida de homogeneidad cuando existe un punto de quiebre. En este estudio se trabajó en la región climática semiárida del Altiplano Potosino de México, procesando 16 registros de periodo amplio, cuyo número de datos anuales varió de 42 a 53.

## Métodos operativos

### Corrección de valores dispersos

Los *valores dispersos (outliers)* son observaciones que distan mucho de la tendencia mostrada por los datos, los cuales corresponden a mediciones erróneas o a eventos meteorológicos extremos. Existen diversos enfoques para contrastar y verificar la variabilidad temporal y/o espacial de la variable climática estudiada, con objeto de identificar los valores dispersos y diagnosticar datos erróneos o valores factibles de ocurrir. Cuando los valores dispersos son sin duda observaciones erróneas, tales datos se eliminan, pero entonces se tiene un problema de datos faltantes. Si son valores factibles, se debe decidir si se corrigen y cómo conviene hacerlo (Eischeid *et al.*, 1995; González-Rouco *et al.*, 2001).

Debido a las dificultades prácticas que implica la verificación de valores dispersos —en los registros originales o de campo—, se adoptó un enfoque de reducción de la dispersión en la cola derecha de la distribución de probabilidades, para buscar mejor desempeño de las técnicas

homogenización que no son resistentes (Lanzante, 1996). Tal enfoque consiste en acotar los *valores dispersos* a una magnitud umbral extrema ( $P_{vd}$ ), que es (Eischeid *et al.*, 1995):

$$P_{vd} = P_{50} + FM \cdot (P_{75} - P_{25}) \quad (1)$$

donde  $P_{50}$  es la mediana o el percentil de 50%;  $P_{75}$  y  $P_{25}$  son los percentiles de 75 y 25%; y  $FM$  es el factor de multiplicación del rango del intercuartil (RIC), con un valor empírico adoptado de 2.75 para los datos de temperatura media y de 4.00 para la precipitación, ambas mensuales. González-Rouco *et al.* (2001) utilizan una expresión similar para sus datos de precipitación mensual:

$$P_{vd} = P_{75} + 3 \cdot (P_{75} - P_{25}) \quad (2)$$

Los valores dispersos superiores a  $P_{vd}$  se sustituyen por tal límite. Este enfoque de *truncamiento* de los valores dispersos reduce el sesgo que éstos originan y también mantiene información relativa a los valores extremos. Eischeid *et al.* (1995) indican que el RIC es utilizado en los procesos de control de calidad debido a que es resistente (Lanzante, 1996) a los valores dispersos. En este estudio se adoptó el criterio definido por la Ecuación (1), por tener mayor justificación teórica y empírica.

## Pruebas del enfoque absoluto

Los test que se aplican a la serie estudiada son los del planteamiento de Wijngaard *et al.* (2003), que utiliza cuatro métodos para probar la pérdida de homogeneidad: el test de Von Neumann, el test del rango de Pettitt, el test del rango de Buishand y la prueba SNHT. Estas cuatro pruebas o test suponen que bajo la hipótesis nula, los valores anuales de la serie ( $Y_i$ ) de la variable  $Y$  que se estudia son independientes y están idénticamente distribuidos. Bajo la hipótesis alternativa —el primer test—, se asume que la serie no está distribuida de forma aleatoria, y los tres restantes consideran que existe un cambio en la media o punto de quiebre. Lógicamente, la *homogeneidad* o consistencia de la serie implica que sus datos pertenecen a una sola población que tiene una media invariante con el tiempo (Machiwal & Jha, 2012).

Las pruebas de Pettitt, Buishand y SNHT permiten localizar el año donde es probable ocurriera el punto de quiebre y, en general, son diferentes en su planteamiento o concepción. Se ha encontrado que las dos primeras son más sensitivas a los puntos de quiebre que suceden hacia la mitad de la serie; por el contrario, la prueba SNHT detecta cambios con más facilidad hacia el inicio y final de serie.

Los test de Von Neumann, Buishand y SNHT asumen que los datos  $Y_i$  tienen distribución normal; en cambio, la prueba de Pettitt no requiere tal consideración, pues se basa en los rangos y no en los valores de la serie, lo cual también la vuelve menos sensible a los valores dispersos

(*outliers*). La prueba de Von Neumann se considera complementaria a las otras tres, debido a que detecta pérdida de homogeneidad por causas diferentes al cambio en la media (Wijngaard *et al.*, 2003).

### Test de Von Neumann

El cociente de Von Neumann ( $N$ ) tiene en el numerador la suma de las diferencias sucesivas al cuadrado y en el denominador la varianza. Su ecuación es (Wijngaard *et al.*, 2003):

$$N = \frac{\sum_{i=1}^{n-1} (Y_i - Y_{i+1})^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \quad (3)$$

en la cual,  $n$  es el número años de la serie o registro procesado. Cuando la serie es homogénea,  $N$  tiene un valor cercano a dos; si existe un cambio en la media,  $N$  disminuye y resulta menor que el valor crítico (Tabla 1); cuando existen variaciones rápidas en la media,  $N$  aumenta y puede ser mayor que dos. Wijngaard *et al.* (2003) exponen sus valores críticos en la Tabla 1, obtenidos por otros autores, los cuales son función de  $n$  y del nivel de significancia ( $\alpha$ ) de la prueba.

**Tabla 1.** Valores críticos de las cuatro pruebas absolutas aplicadas, con nivel de significancia ( $\alpha$ ) de 5 y 1% (Pettitt, 1979; Wijngaard *et al.*, 2003; Khaliq & Ouarda, 2007).

Núm. se datos ( $n$ )	Test de Von Neumann		Test del rango de Pettitt		Test del rango de Buishand		Prueba SNHT	
	5%	1%	5%	1%	5%	1%	5%	1%
20	1.30	1.04	64.8	80.3	1.43	1.60	6.95 (7.089)	9.56 (9.113)
30	1.42	1.20	118.0	146.3	1.50	1.70	7.65 (7.747)	10.45 (10.153)
40	1.49	1.29	181.0	224.4	1.53	1.74	8.10 (8.151)	11.01 (10.771)
50	1.54	1.36	252.3	312.8	1.55	1.78	8.45 (8.432)	11.38 (11.193)
70	1.61	1.45	416.8	516.7	1.59	1.81	8.80 (8.814)	11.89 (11.737)
100	1.67	1.54	710.1	880.5	1.62	1.86	9.15 (9.167)	12.32 (12.228)

### Test del rango de Pettitt

Los rangos ( $R_1, R_2, \dots, R_n$ ) de la serie original ( $Y_1, Y_2, \dots, Y_n$ ) son utilizados para calcular el estadístico de la prueba (Pettitt, 1979; Wijngaard *et al.*, 2003; Yozgatligil & Yazici, 2016):

$$P_k = 2 \cdot \sum_{i=1}^k R_i - k \cdot (n + 1) \quad \text{con } k = 1, 2, \dots, n \quad (4)$$

El cálculo de los rangos es bastante simple. Primero se ordenan los datos en forma creciente y después se compara cada dato ordenado contra todos los originales; cuando coinciden, ese es el rango buscado. Si un punto de quiebre ocurre en el año  $k$ , entonces el estadístico ( $P_k$ ) es máximo o mínimo cerca de tal año, y su valor máximo absoluto excede el valor crítico de la Tabla 1, el cual está definido por la siguiente ecuación en función del nivel de significancia ( $\alpha$ ) y de  $n$  (Pettitt, 1979; Dhorde & Zarenistanak, 2013):

$$VC_{Pettitt} = [-\ln(\alpha) \cdot (n^2 + n^3)/6]^{1/2} \quad (5)$$

La gráfica de  $k$  en las abscisas contra  $P_k$  en las ordenadas muestra los resultados de la prueba. Mallakpour y Villarini (2016) analizan con simulación el desempeño o sensibilidad del test de Pettitt en la detección de cambios súbitos en la media, en muy diversas condiciones y en series de valores extremos.

## Test del rango de Buishand

Se comienza por calcular la curva masa residual, definida como (Buishand, 1982; Wijngaard *et al.*, 2003):

$$S_0^* = 0 \text{ y } S_k^* = \sum_{i=1}^k (Y_i - \bar{Y}) \quad \text{con } k = 1, 2, \dots, n \quad (6)$$

Cuando la serie es homogénea, los valores de  $S_k^*$  fluctúan alrededor de cero debido a que  $Y_i$  no tiene desviaciones sistemáticas con respecto a su media. Si un punto de quiebre ocurre en el año  $k$ , entonces  $S_k^*$  alcanza su máximo (salto negativo) o su mínimo (salto positivo) cerca de tal año. El estadístico de la prueba es el rango ( $R$ ) escalado con la desviación estándar ( $s$ ) de la serie y dividido entre  $\sqrt{n}$ .  $R$  se obtiene con la diferencia entre el  $S_k^*$  máximo menos el mínimo y  $s$  con expresión siguiente:

$$s = \left[ \frac{\sum_{i=1}^n (Y_i - \bar{Y})^2}{n-1} \right]^{1/2} \quad (7)$$

Los valores críticos de  $R/(\sqrt{n} s)$  se tienen en la Tabla 1.

## Prueba SNHT

El estadístico  $T_k$  de la prueba compara la media estandarizada de los primeros  $k$  años del registro contra la relativa a los siguientes  $(n - k)$  años (Alexandersson, 1986; Wijngaard *et al.*, 2003; Syrakova & Stefanova, 2009):

$$T_k = k \cdot \bar{z}_1^2 + (n - k) \cdot \bar{z}_2^2 \quad \text{con } k = 1, 2, \dots, n \quad (8)$$

con:

$$\bar{z}_1 = \left( \frac{1}{k \cdot s} \right) \cdot \sum_{i=1}^k (Y_i - \bar{Y}) \quad (9)$$

$$\bar{z}_2 = \left[ \frac{1}{(n-k) \cdot s} \right] \cdot \sum_{i=k+1}^n (Y_i - \bar{Y}) \quad (10)$$

Este estadístico tiene un comportamiento igual al  $P_k$ , con sus valores críticos definidos en la Tabla 1. Entre paréntesis se han citado los valores críticos desarrollados recientemente por Khaliq y Ouarda (2007), que son más precisos que los originales; abarcan muestras de gran tamaño y han sido expuestos para seis niveles de confiabilidad  $(1 - \alpha)$ , variando de 90 a 99%. Alexandersson y Moberg (1997) exponen la versión de la prueba SNHT para series o registros con tendencia lineal.

Toreti *et al.* (2011) citan resultados de varios autores y verifican con base en 1 000 series sintéticas de 100 años de longitud no homogéneas, que la prueba SNHT es propensa a detectar puntos de quiebre en el inicio o final de la serie cuando tiene pocos datos; por ello se debe evitar aplicarla en segmentos con menos de diez valores.

Respecto al nivel de significancia ( $\alpha$ ) a usar en la prueba SNHT, Syrakova y Stefanova (2009) aceptan el criterio de Hanssen-Bauer y Forland (1994), y de González-Rouco *et al.* (2001) de utilizar un 10% si el punto de quiebre está verificado por los datos históricos de la estación climatológica, y 5% si no existe tal comprobación. Esta selección reduce el riesgo de aceptar la homogeneidad en ausencia de datos sobre los cambios ocurridos en la estación climatológica, lugar, equipo y técnica de medición.

## Corrección de la pérdida de homogeneidad

Cuando un punto de quiebre se ha identificado en un registro de precipitación anual, los datos previos a tal cambio se corrigen al multiplicarlos por el siguiente factor (González-Rouco *et al.*, 2001):

$$f = \frac{\bar{P}_d}{\bar{P}_a} \quad (11)$$

siendo  $\bar{P}_d$  y  $\bar{P}_a$  las medias de los valores anuales de precipitación después y antes del cambio. Syrakova y Stefanova (2009) exponen cómo se realiza la corrección de registros de temperatura anual y estacional que se contrastan con pruebas relativas.

## Clasificación de las series probadas

La clasificación depende del número de pruebas que rechazaron la hipótesis nula, definiéndose tres categorías, cuyas designaciones según Wijngaard *et al.* (2003), y Dhorde y Zarenistanak (2013) son las siguientes: útil, dudosa y sospechosa. De acuerdo con Guajardo-Panes *et al.* (2017) son las siguientes: confiable, medianamente confiable y no confiable. En este estudio se adopta la designación siguiente: en la clase 1 o "confiable" se agrupan las series que fueron detectadas homogéneas por las cuatro pruebas, o bien, sólo una prueba detectó pérdida de homogeneidad; la clase 2 o "menos confiable" engloba series con dos pruebas que detectan falta de homogeneidad, y la clase 3 o "no confiable" son series con tres o cuatro pruebas que rechazan la hipótesis nula.

Con las series clase 1 se pueden realizar análisis de tendencia y estudios de variabilidad climática; con las clase 2, los resultados de los análisis citados deben ser tomados con reservas, pero son factibles si no existen otras series confiables disponibles; las series clase 3 no se deben utilizar para las estimaciones mencionadas.

Existe una diferencia notable al establecer la clasificación anterior, ya que Wijngaard *et al.* (2003) definen un nivel de significancia ( $\alpha$ ) en las pruebas de 1%; en cambio, en los otros dos trabajos citados se emplea un 5%. En este estudio se aceptó  $\alpha = 5\%$ , para coincidir con las recomendaciones de la prueba SNHT (Alexandersson, 1986).

## Datos utilizados y técnicas auxiliares

### Selección de estaciones climatológicas

Este estudio empleó la información climatológica *mensual* del archivo en Excel proporcionado por la Dirección Local San Luis Potosí de la Comisión Nacional del Agua (Conagua), la cual está agrupada en las tres zonas geográficas del estado: Altiplano Potosino, Zona Media y Región Huasteca. Los registros de datos procesados corresponden a la variable climática *precipitación anual* en milímetros, integrado cada dato por la suma de los doce valores de lluvia mensual.

El archivo del Altiplano Potosino abarca más de 60 estaciones climatológicas, pero algunas están repetidas por pertenecer a la extinta Secretaría de Recursos Hidráulicos (SRH) y al Servicio Meteorológico

Nacional (SMN). Así, se eliminaron las estaciones con menos de 40 años de registro y con lapsos de años faltantes. Con tales restricciones, se obtuvo una muestra de 16 estaciones, cuyo año inicial de registros se indica con AI en la Tabla 2 y Tabla 3 de datos por procesar. Todos los registros abarcan hasta el año 2016, excepto la estación Los Pilares, que fue suspendida en 2009.

**Tabla 2.** Valores de la precipitación anual (milímetros) y de sus estadísticos en las ocho estaciones pluviométricas indicadas del Altiplano Potosino, México.

Núm.	1	2	3	4	5	6	7	8
Año	Vanegas	S. M. del Refugio	La Presa	Matehuala	La Maroma	Charcas	Palo Blanco	Reforma
1964	282.0	302.6	-	(614.5)	-	(283.1)	211.0	-
1965	300.8	319.0	-	464.0	312.5	317.0	350.5	374.5
1966	457.6	377.0	-	766.0	408.0	463.7	586.0	598.7
1967	410.9	362.3	-	578.0	462.9	585.5	910.0	580.3
1968	440.4	330.5	-	662.0	490.0	598.0	472.0	393.3
1969	713.7	(127.9)	-	(420.5)	248.0	163.0	(242.2)	247.9
1970	325.0	282.0	-	368.5	205.0	517.0	196.0	300.0
1971	224.5	390.5	-	[327.5]	387.0	684.0	509.7	268.3
1972	253.9	329.0	-	533.5	280.0	(389.7)	461.1	(255.9)
1973	404.0	513.6	-	471.0	434.4	607.6	168.6	*1037.3*
1974	82.4	107.0	-	(328.9)	95.0	(175.0)	191.8	248.6

1975	155.5	163.5	(247.1)	526.0	332.0	[364.3]	329.5	277.0
1976	336.0	501.3	695.3	891.8	586.0	949.5	491.6	650.3
1977	152.5	136.1	470.2	532.6	355.0	199.7	261.3	224.0
1978	249.4	252.7	570.5	698.3	439.0	376.6	335.3	344.1
1979	222.3	223.1	383.2	441.9	303.0	263.9	238.6	225.9
1980	346.0	(249.5)	428.9	474.6	311.0	330.4	(172.4)	293.1
1981	249.9	(178.8)	590.2	[438.5]	346.0	364.5	(211.0)	424.3
1982	288.5	265.1	394.0	(274.3)	360.0	166.6	(297.4)	320.7
1983	457.1	358.6	444.9	568.0	268.0	192.8	194.0	230.3
1984	276.8	370.7	649.2	369.0	440.0	599.5	227.5	342.2
1985	251.4	233.1	636.9	[475.4]	392.0	507.5	122.6	370.4
1986	221.8	240.0	377.4	*668.3*	288.0	(326.7)	184.3	258.4
1987	417.8	(77.0)	591.2	593.9	499.0	591.1	148.2	319.9
1988	329.6	38.7	634.5	557.2	359.8	(443.6)	(155.5)	465.2
1989	122.1	139.0	401.3	(491.7)	212.0	344.0	(31.0)	136.8
1990	341.0	355.9	466.5	(648.2)	327.0	(505.7)	135.0	432.9
1991	336.4	305.0	508.6	695.2	497.1	659.3	[554.5]	497.8
1992	351.6	383.0	597.1	677.3	391.0	773.0	255.0	540.5
1993	185.5	*884.9*	411.0	445.2	256.0	612.3	121.0	283.1
1994	95.0	371.5	464.9	571.1	426.0	227.7	171.0	480.4
1995	52.5	241.0	(461.9)	559.1	351.0	119.0	(264.8)	419.9
1996	144.5	294.1	487.1	310.3	206.1	405.7	222.3	125.1
1997	241.7	561.2	574.7	586.2	367.0	506.2	305.4	269.8
1998	190.8	112.5	316.2	334.8	244.0	393.5	250.6	325.1
1999	182.5	133.0	366.4	347.6	177.0	229.1	(81.8)	111.5

2000	173.5	250.0	(303.8)	313.2	331.0	310.6	175.5	(91.7)
2001	176.5	184.0	(377.2)	431.7	367.0	545.7	200.2	232.2
2002	428.0	410.0	(818.0)	753.1	508.0	834.8	428.8	414.3
2003	307.2	244.0	(497.3)	677.6	502.9	486.0	163.8	374.7
2004	225.0	404.0	(701.6)	648.7	387.8	742.0	245.1	356.1
2005	224.0	179.0	327.1	372.0	276.0	353.8	194.1	214.6
2006	339.0	(149.9)	494.6	624.8	500.0	607.7	315.0	302.4
2007	351.5	72.0	854.2	549.3	474.0	758.2	462.5	513.7
2008	337.0	354.5	664.5	625.5	309.0	552.6	372.0	359.3
2009	352.5	337.6	566.2	406.9	342.0	610.7	373.0	395.2
2010	327.8	371.0	496.6	503.6	593.0	607.2	330.0	(155.8)
2011	137.0	98.0	340.5	292.3	147.0	330.0	182.0	180.6
2012	143.1	150.7	564.0	425.7	386.6	243.2	258.0	323.4
2013	536.8	563.0	973.1	726.4	687.0	889.7	398.5	601.0
2014	258.8	398.5	474.9	525.1	333.0	394.0	326.0	422.1
2015	525.5	581.0	724.1	702.7	576.0	732.3	713.9	614.4
2016	281.0	486.0	630.0	668.5	351.0	356.2	207.0	427.9
AI	1964	1964	1975	1957	1965	1962	1964	1965
<i>n</i>	53	53	42	53	52	53	53	52
Mín	52.5	38.7	247.1	274.3	95.0	119.0	31.0	91.7
Máx	713.7	884.0	973.1	891.8	687.0	949.5	910.0	1037.3
$\bar{x}$	287.1	297.1	523.3	527.5	367.8	464.0	290.7	358.7
<i>Med</i>	281.0	294.1	495.6	532.6	357.4	443.6	245.1	333.6
<i>Cv</i>	0.439	0.531	0.299	0.271	0.324	0.434	0.558	0.461
<i>Cs</i>	0.757	1.039	0.721	0.140	0.255	0.371	1.565	1.433

$Ck$	4.459	5.471	3.728	2.548	3.380	2.623	6.445	7.352
$r_1$	0.241	0.189	-0.025	0.163	-0.08	0.032	0.347	0.044

**Tabla 3.** Valores de la precipitación anual (milímetros) y de sus estadísticos en las ocho estaciones pluviométricas indicadas del Altiplano Potosino, México.

Núm.	9	10	11	12	13	14	15	16
Año	Moctezuma	El Mezquite	El Grito	Los Pilares	Mezquitic	Los Filtros	El Peaje	Villa de Arriaga
1964	{232.5}	197.0	-	281.6	294.5	418.0	456.1	489.0
1965	{459.0}	287.5	-	422.0	233.0	383.6	454.2	426.1
1966	{371.9}	428.0	-	413.0	426.5	565.2	539.0	631.7
1967	{224.0}	495.8	-	672.8	572.1	627.3	557.6	538.0
1968	{417.5}	358.0	-	415.3	540.9	504.1	476.0	(199.9)
1969	{78.0}	183.0	178.6	203.5	295.0	302.7	394.0	174.0
1970	{333.7}	250.0	380.8	323.6	348.5	363.5	369.2	269.7
1971	183.0	297.0	540.3	399.5	471.0	388.6	541.0	593.0
1972	441.6	286.5	394.1	300.5	293.5	350.7	(243.8)	347.0
1973	307.0	760.5	530.3	379.0	545.0	420.1	388.7	328.5
1974	236.6	576.9	268.9	251.0	310.6	235.9	(267.1)	156.9
1975	427.8	700.5	461.4	372.8	557.0	374.4	494.3	280.5
1976	463.9	523.0	607.2	508.6	756.0	503.2	631.1	359.5
1977	343.3	265.8	434.3	319.9	402.2	282.8	307.2	96.0
1978	345.2	419.2	311.9	352.6	429.0	347.2	383.3	507.5
1979	325.5	361.5	270.4	393.5	380.9	354.6	355.7	192.0

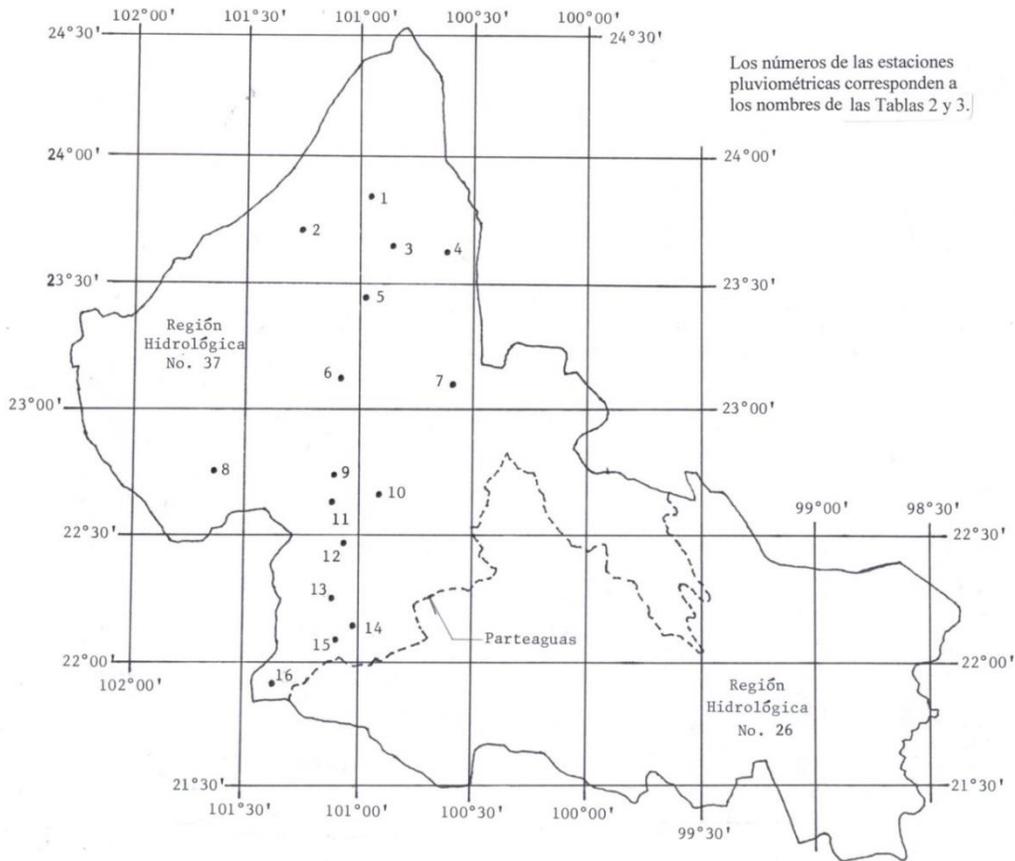
1980	260.7	290.0	280.4	216.5	319.9	253.9	305.8	375.5
1981	374.7	493.8	457.2	474.6	492.9	500.7	504.3	270.9
1982	334.8	321.0	467.8	298.2	305.7	326.7	291.4	292.5
1983	314.3	351.4	507.2	376.8	285.1	280.4	259.3	154.5
1984	563.0	468.7	625.0	377.2	564.9	361.6	507.7	148.0
1985	594.2	(355.7)	534.4	437.6	478.4	489.1	485.3	167.0
1986	296.1	[373.4]	419.2	421.0	439.7	456.7	403.2	504.0
1987	(339.4)	[431.6]	566.6	393.5	427.0	364.8	322.6	504.9
1988	531.4	(337.8)	672.1	536.6	576.3	455.0	399.4	256.0
1989	(219.1)	256.0	262.8	736.2	270.5	239.5	282.9	359.5
1990	315.2	(62.6)	457.0	285.9	420.4	395.5	372.7	728.3
1991	491.9	33.9	539.4	492.5	641.2	528.1	476.0	*876.6*
1992	521.9	90.0	524.9	525.2	480.8	469.6	477.1	837.0
1993	310.6	44.5	407.8	382.9	(85.7)	296.1	404.7	523.0
1994	430.5	358.0	563.0	472.9	(178.1)	332.8	348.0	(562.9)
1995	383.6	(127.8)	343.1	337.3	395.4	285.7	324.7	400.0
1996	305.6	209.6	361.4	281.9	304.5	280.4	331.3	*542.6*
1997	407.5	361.2	551.5	458.1	270.5	432.6	439.0	679.0
1998	177.8	236.0	187.0	238.9	130.7	363.7	339.3	418.0
1999	327.5	208.0	369.6	349.9	163.0	399.8	334.7	249.0
2000	222.5	292.0	1226.4	219.8	300.0	169.6	281.9	122.0
2001	406.6	406.0	305.6	420.9	499.9	273.9	321.8	(111.7)
2002	495.5	693.0	593.7	525.2	642.2	466.8	556.7	(112.1)
2003	418.7	527.0	360.5	367.1	224.7	425.7	319.0	337.0
2004	510.0	825.0	546.4	607.2	474.8	550.8	702.7	688.5

2005	(224.9)	232.0	(210.5)	350.9	(357.4)	295.1	273.8	287.0
2006	258.0	450.0	347.0	342.4	383.0	509.8	556.4	340.0
2007	*832.3*	678.0	611.7	578.3	457.0	516.9	520.2	270.0
2008	353.0	434.0	268.2	437.5	590.5	373.5	*562.7*	546.4
2009	39.0	447.0	360.5	-	448.4	464.5	533.6	201.0
2010	34.0	597.2	542.0	-	508.4	492.7	563.9	(263.8)
2011	15.0	232.0	257.1	-	327.8	218.0	291.8	(75.7)
2012	18.0	471.0	325.5	-	286.4	195.5	244.7	134.9
2013	89.0	810.0	649.0	-	822.4	548.1	626.8	189.2
2014	386.0	675.0	451.0	-	448.5	387.3	(364.0)	165.0
2015	451.0	1040.0	678.5	-	729.7	710.2	680.3	(396.5)
2016	162.0	1040.0	389.0	-	481.1	441.7	346.0	340.1
AI	1971	1962	1969	1963	1943	1950	1963	1962
<i>n</i>	53	53	48	45	53	53	53	53
Mín	15.0	33.9	178.6	203.5	85.7	169.6	243.8	75.7
Máx	832.3	1040.0	678.5	736.2	822.4	710.2	702.7	876.6
$\bar{x}$	332.2	408.5	429.1	398.9	416.4	395.7	419.1	358.9
<i>Med</i>	334.8	361.2	426.7	382.9	426.5	387.3	394.0	337.0
<i>Cv</i>	0.476	0.557	0.320	0.292	0.374	0.285	0.284	0.550
<i>Cs</i>	0.154	0.877	-0.006	0.735	0.317	0.290	0.489	0.720
<i>Ck</i>	4.141	3.932	2.131	3.923	3.264	3.141	2.481	3.037
<i>r</i> <sub>1</sub>	0.237	0.587	-0.056	-0.02	0.156	0.102	-0.108	0.489

Para establecer el *mayor* periodo común de datos se seleccionó el año 1964 para su inicio y, con ello, 13 registros procesados tienen 53 o

52 valores; tres, La Presa, El Grito y Los Pilares, abarcan 42, 48 y 45 años, respectivamente. En la estación Moctezuma, los siete años faltantes de 1964 a 1970, se tomaron del *Boletín Climatológico No. 3* (SARH, 1980), con datos del inicio de las estaciones al año de 1978. Estos valores se han indicado en el inicio de la Tabla 3, con paréntesis de llave. En el boletín citado se verificaron los inicios de las estaciones La Presa y El Grito.

En la Figura 1 se muestra la localización geográfica de las 16 estaciones climatológicas procesadas del Altiplano Potosino, México, destacando que las cuatro primeras forman un subgrupo norte-extremo; las tres siguientes integran un subgrupo norte; las cinco que continúan definen el subgrupo del centro; las cuatro últimas establecen el subgrupo del sur.



**Figura 1.** Ubicación geográfica de las 16 estaciones climatológicas procesadas del Altiplano Potosino, México.

## Estimación de datos mensuales faltantes

En este estudio se adoptó como valor mensual faltante de los años incompletos la *moda* mensual respectiva, estimada con base en el ajuste

de la distribución Gamma Mixta de dos parámetros (Campos-Aranda, 2005), bajo tres condiciones específicas: (1) se estima sólo en el semestre lluvioso, es decir, de mayo a octubre; (2) se completa un máximo de tres meses en cada año incompleto; y (3) cuando la moda no existe, porque la distribución es de tipo J invertida, se emplea la precipitación mensual con probabilidad de no ser excedida de 25% ( $P_{25}$ ), como el caso de octubre en la estación Charcas (ver Tabla 4). En la Tabla 2 y Tabla 3, los valores anuales completados con el procedimiento anterior se han indicado entre paréntesis circulares.

**Tabla 4.** Muestra de resultados del ajuste de la distribución Gamma Mixta en los dos meses indicados de cada estación pluviométrica incompleta del Altiplano Potosino, México.

Estación	Santa Ma. del Refugio		La Presa		Matehuala	
	Mayo	Junio	Abril	Julio	Mayo	Junio
Mes	Mayo	Junio	Abril	Julio	Mayo	Junio
Núm. de datos	50	52	38	41	53	54
Núm. de ceros	7	7	8	0	0	0
<i>EEA</i>	5.1	13.9	6.0	20.4	9.4	18.2
Media	29.0	50.2	30.2	85.7	55.2	84.3
Moda	13.9	20.9	9.0	43.0	22.3	52.4
$P_{25}$	9.7	15.8	4.2	42.4	24.7	47.1

$P_{50}$	23.2	39.1	20.9	72.2	44.9	74.1
$P_{75}$	41.9	72.3	44.8	114.9	74.8	110.9
$P_{vd}$	152.0	265.1	183.3	362.2	245.3	329.3
1º máximo	96.0	(267.0)	102.7	353.0	232.5	(352.0)
2º máximo	89.4	160.0	86.2	217.1	133.0	201.5
<b>Estación</b>	<b>Charcas</b>		<b>Palo Blanco</b>		<b>Reforma</b>	
Mes	Agosto	Oct.	Mayo	Sept.	Junio	Julio
Núm. de datos	51	52	50	48	51	50
Núm. de ceros	1	2	4	1	3	2
$EEA$	6.3	5.0	5.8	12.2	5.5	14.8
Media	60.6	37.1	40.3	59.0	53.9	9.6
Moda	20.7	NE	18.3	23.4	16.1	15.7
$P_{25}$	24.6	9.1	16.1	25.3	19.1	24.1
$P_{50}$	48.0	24.7	32.8	47.7	41.5	52.4
$P_{75}$	83.3	51.6	56.6	80.8	75.4	96.8
$P_{vd}$	282.8	194.7	194.8	269.7	266.7	343.2
1º máximo	192.0	159.6	137.0	204.0	208.4	(347.0)
2º máximo	189.4	117.0	97.0	175.0	155.1	237.3
<b>Estación</b>	<b>Moctezuma</b>		<b>El Mezquite</b>		<b>El Grito</b>	
Mes	Junio	Sept.	Julio	Sept.	Julio	Agosto

Núm. de datos	47	47	52	51	47	47
Núm. de ceros	6	4	7	4	0	1
<i>EEA</i>	28.8	4.6	10.1	5.9	8.1	5.7
Media	62.1	49.1	59.2	61.3	83.6	48.7
Moda	21.0	20.8	4.2	24.4	30.8	19.1
$P_{25}$	18.2	18.7	11.7	23.1	36.2	20.8
$P_{50}$	47.0	39.4	39.6	48.8	67.2	39.3
$P_{75}$	89.4	69.2	84.8	86.3	113.6	66.8
$P_{vd}$	331.8	241.4	332.0	301.6	376.8	223.3
1º máximo	(399.5)	151.0	251.0	205.4	286.0	173.6
2º máximo	152.0	137.0	219.0	200.0	222.9	150.2
<b>Estación</b>	<b>Mexquitic</b>		<b>El Peaje</b>		<b>Villa de Arriaga</b>	
Mes	Mayo	Julio	Julio	Agosto	Julio	Sept.
Núm. de datos	51	51	53	53	54	51
Núm. de ceros	3	1	0	0	4	5
<i>EEA</i>	3.8	9.3	9.4	14.7	38.1	34.6
Media	38.0	77.9	75.7	61.5	74.7	67.0

Moda	6.9	6.7	32.3	29.3	32.3	5.9
$P_{25}$	11.9	24.4	34.8	29.6	29.5	16.4
$P_{50}$	27.9	56.1	62.2	51.4	60.3	46.4
$P_{75}$	53.2	107.9	102.4	82.7	104.7	95.2
$P_{vd}$	193.1	390.0	332.6	263.8	361.1	361.6
1º máximo	137.0	320.5	275.0	(282.2)	(513.0)	(495.0)
2º máximo	127.0	218.2	219.1	138.8	234.0	280.0

*EEA*: error estándar de ajuste, en mm;  $P_{25}$ : precipitación anual con un 25% de probabilidad de ser menor o igual en mm.

En la Tabla 4 se exponen resultados de los ajustes mensuales —sólo dos meses de cada registro incompleto— por limitación de espacio. De la Tabla 4 se deduce que sólo cuatro estaciones pluviométricas tuvieron registros completos: Vanegas, La Maroma, Los Pilares y Los Filtros.

## Corrección de valores mensuales máximos

El ajuste de la distribución Gamma Mixta permitió estimar los elementos de la Ecuación (1) para calcular el valor disperso máximo mensual ( $P_{vd}$ ), citado en la Tabla 4. También se indican entre paréntesis circulares los valores máximos que deberán ser truncados al valor de  $P_{vd}$  por haberlo

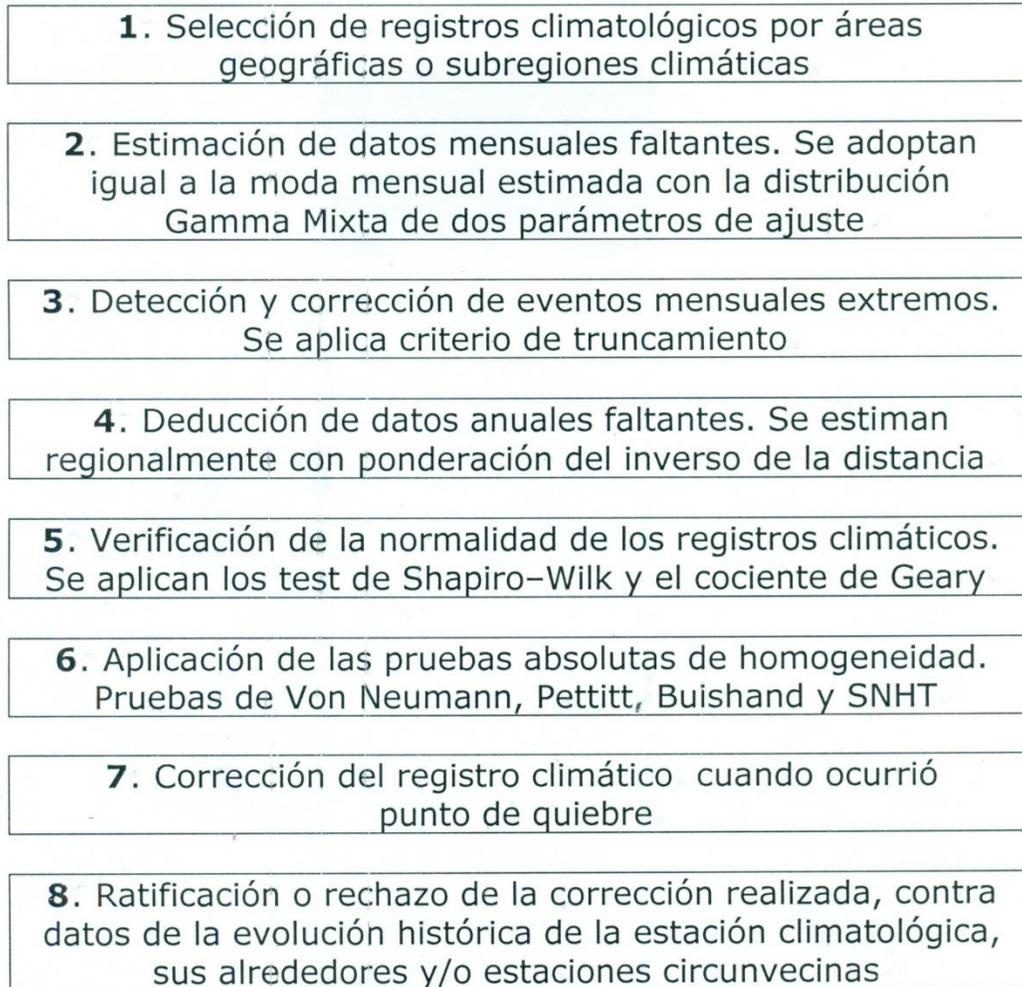
excedido. Los respectivos valores anuales corregidos por el truncamiento descrito se destacan entre asteriscos en la Tabla 2 y Tabla 3. Debido al proceso adoptado de deducción de datos mensuales faltantes, sólo se corrigieron valores máximos extremos en el semestre lluvioso de mayo a octubre, pero se buscaron valores erróneos en el resto de los meses para eliminarlos.

## **Estimación de datos anuales faltantes**

Las estaciones pluviométricas Matehuala, Charcas, Palo Blanco y El Mezquite presentaron datos anuales faltantes; la primera, tres; la última, dos; y el resto, un valor. Al observar en la Tabla 2 y Tabla 3 que los valores anuales disponibles en las estaciones circunvecinas a cada registro incompleto eran bastante similares, se optó por deducir cada valor anual faltante con base en los tres valores de las estaciones más cercanas, según un ponderado con el inverso de la distancia en kilómetros entre cada estación vecina y la incompleta. Este procedimiento fue descrito y aplicado por Eischeid *et al.* (1995), Mishra y Singh (2008), y por Campos-Aranda (2013); es adecuado en áreas de topografía plana y para pocos datos faltantes. En la Tabla 2 y Tabla 3, los siete valores anuales estimados con el enfoque descrito se citan entre paréntesis rectangulares.

Cuando se deduzcan varios datos anuales faltantes en estaciones climatológicas de áreas de montaña se debe usar el método de la relación normalizada (Paulhus & Kohler, 1952), aplicando factores de ponderación definidos con base en los coeficientes de correlación entre los valores anuales de la estación incompleta y las circunvecinas, como los definió Young (1992) y aplicaron Eischeid *et al.* (1995).

En la Figura 2 se citan los conceptos y técnicas numéricas que incluye la estrategia propuesta para la verificación de la homogeneidad en registros de precipitación anual de una determinada área geográfica.



**Figura 2.** Diagrama de bloques de los conceptos y técnicas estadísticas, que incluye la estrategia sugerida para probar y alcanzar homogeneidad de los registros climatológicos.

## Análisis y discusión de los resultados

## Parámetros estadísticos de los registros

En los diez últimos renglones de la Tabla 2 y Tabla 3 se muestran los indicadores estadísticos relativos a cada registro completado. Con AI se indica el año de inicio de cada serie de precipitación anual, cuyo número de datos, dentro del periodo común establecido, es  $n$ . Después de la mediana ( $Med$ ) del registro se citan los coeficientes de variación ( $Cv$ ), asimetría ( $Cs$ ) y curtosis ( $Ck$ ), así como el de correlación serial de orden uno ( $r_1$ ).

## Verificación de la normalidad

Se aplicó el test clásico de Shapiro y Wilk (1965) a los registros de precipitación anual con menos de 50 datos, y el de Shapiro (1998) en aquellos con más 50 valores. Se obtuvo que los seis registros de las estaciones Santa Ma. del Refugio, Palo Blanco, Reforma, El Mezquite, El Peaje y Villa de Arriaga no proceden de una distribución normal.

A estos seis registros se les aplicó la prueba de Cociente de Geary (Machiwal & Jha, 2012; Campos-Aranda, 2015) y con ello se ratificó que los cuatro primeros registros no proceden de la distribución Normal. En la Tabla 2 y Tabla 3 se observa que los coeficientes de asimetría y curtosis de cuatro registros son los que más difieren de los valores normales de cero y tres. Tales registros serán procesados usando el logaritmo natural de sus datos para volverlos normales.

## Valores críticos de las pruebas estadísticas

En la Tabla 1 se observa que los valores críticos de los estadísticos  $N$ ,  $P_k$ ,  $R/(\sqrt{n} s)$  y  $T_k$  de cada prueba son función del tamaño del registro ( $n$ ) y como varía de 42 a 53, en la Tabla 5 se exponen sus magnitudes respectivas interpoladas.

**Tabla 5.** Valores críticos, con nivel de significancia de 5%, de los estadísticos de las pruebas indicadas, según tamaño ( $n$ ) del registro.

Prueba por aplicar	$n = 42$	$n = 45$	$n = 48$	$n = 52$	$n = 53$
Von Neumann	< 1.500	< 1.515	< 1.530	< 1.547	< 1.551
Pettitt	> 194.6	> 215.7	> 237.4	> 267.5	> 275.2

Buishand	> 1.534	> 1.540	> 1.546	> 1.554	> 1.556
SNHT	> 8.21	> 8.29	> 8.38	> 8.47	> 8.49

## Resultados de las pruebas estadísticas

En la Tabla 6 se han concentrado las magnitudes máximas que se obtuvieron en cada prueba para su indicador ( $N$ ,  $P_k$ ,  $R/(\sqrt{n} s)$  y  $T_k$ ), cuyo contraste contra los valores de la Tabla 5 permitieron indicar si cada registro resultó homogéneo (H) o no homogéneo (NH). Se obtuvo que 11 registros son *confiables*, es decir, 68.8% de las series de precipitación anual analizadas; tres son *menos confiables* (Vanegas, Palo Blanco y Villa de Arriaga), y dos son *no confiables* (Moctezuma y El Mezquite). Para estas dos últimas estaciones, la prueba de Pettitt aporta sus valores máximos en los 16 registros procesados: 236 y 234.

**Tabla 6.** Valores máximos de los estadísticos de las pruebas aplicadas y resultados de la homogeneidad de los registros de precipitación anual en las 16 estaciones climatológicas indicadas del Altiplano Potosino, México.

Prueba aplicada	Vanegas	S. Ma. del Refugio	La Presa	Matehuala
Von Neumann	1.52→NH	1.46→NH	1.98→H	1.68→H

Pettitt	114.0→H	206.0→H	164.0→H	148.0→H
Buishand	1.53→H	1.12→H	1.26→H	1.00→H
SNHT	9.20→NH	5.48→H	6.94→H	3.48→H
Resultado	Menos conf.	Confiable	Confiable	Confiable
<b>Prueba aplicada</b>	<b>La Maroma</b>	<b>Charcas</b>	<b>Palo Blanco</b>	<b>Reforma</b>
Von Neumann	2.16→H	1.92→H	1.27→NH	1.67→H
Pettitt	136.0→H	204.0→H	188.0→H	153.0→H
Buishand	1.23→H	1.18→H	1.87→NH	1.29→H
SNHT	4.84→H	5.79→H	6.52→H	4.00→H
Resultado	Confiable	Confiable	Menos conf.	Confiable
<b>Prueba aplicada</b>	<b>Moctezuma</b>	<b>El Mezquite</b>	<b>El Grito</b>	<b>Los Pilares</b>
Von Neumann	1.54→NH	0.78→NH	2.04→H	2.02→H
Pettitt	236.0→H	234.0→H	123.0→H	138.0→H
Buishand	1.59→NH	1.95→NH	1.05→H	1.07→H
SNHT	12.6→NH	10.8 →NH	3.39→H	2.58→H
Resultado	No conf.	No conf.	Confiable	Confiable
<b>Prueba aplicada</b>	<b>Mezquitic</b>	<b>Los Filtros</b>	<b>El Peaje</b>	<b>Villa de Arriaga</b>
Von Neumann	1.68→H	1.80→H	2.22→H	1.02→NH
Pettitt	144.0→H	158.0→H	174.0→H	188.0→H

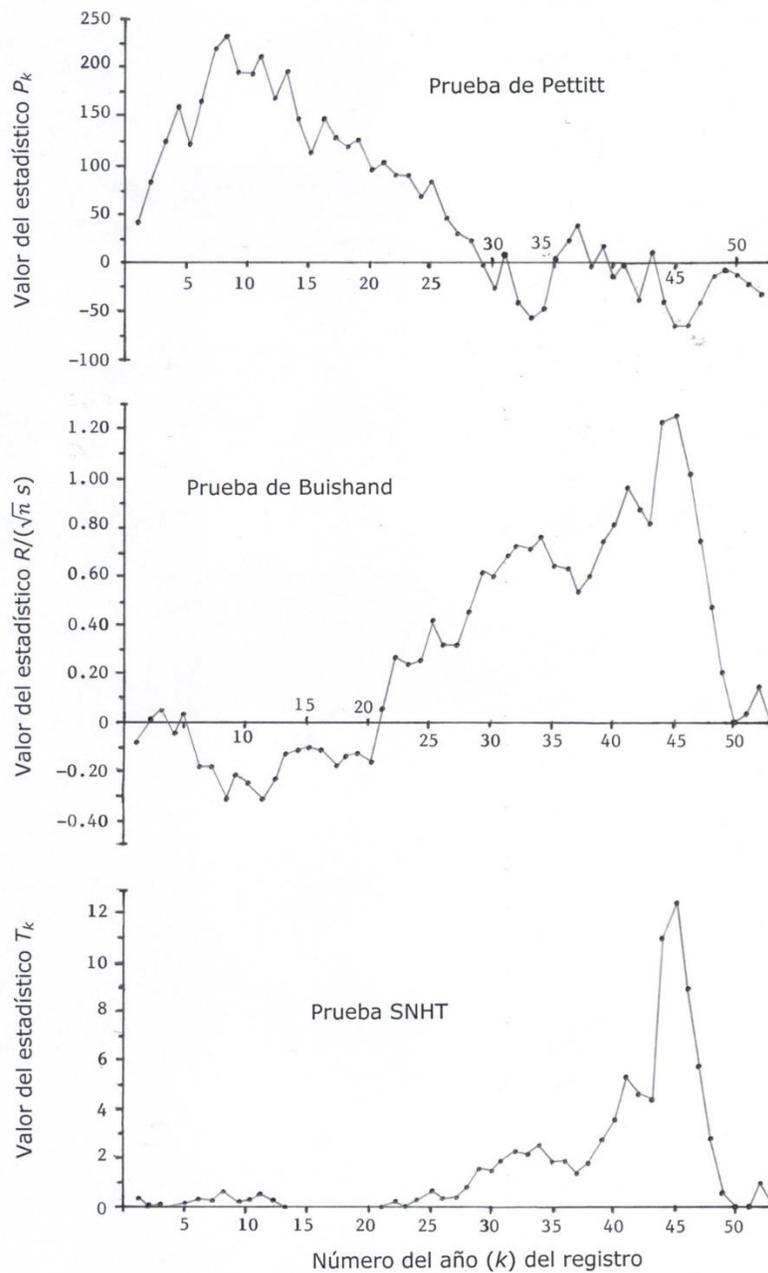
Buishand	1.33→H	1.44→H	1.49→H	1.76→NH
SNHT	7.43→H	5.40→H	4.82→H	5.72→H
Resultado	Confiable	Confiable	Confiable	Menos conf.

## Análisis de los registros no homogéneos

Con respecto a los registros menos confiables de las estaciones Vanegas, Palo Blanco y Villa de Arriaga, es probable que la prueba de Von Neumann detecte la no aleatoriedad debido a la *persistencia* que muestran tales series, cuyos coeficientes de correlación serial de orden uno ( $r_1$ ) fueron 0.241, 0.347 y 0.489; es decir, los más grandes de la Tabla 2 y Tabla 3, junto con el mayor de todos ( $r_1 = 0.587$ ) de la estación El Mezquite. A los registros de las tres estaciones citadas se les aplicó la prueba de Anderson (WMO, 1971; Linsley, Kohler, & Paulhus, 1988), resultando que sus coeficientes  $r_1$  son significativos.

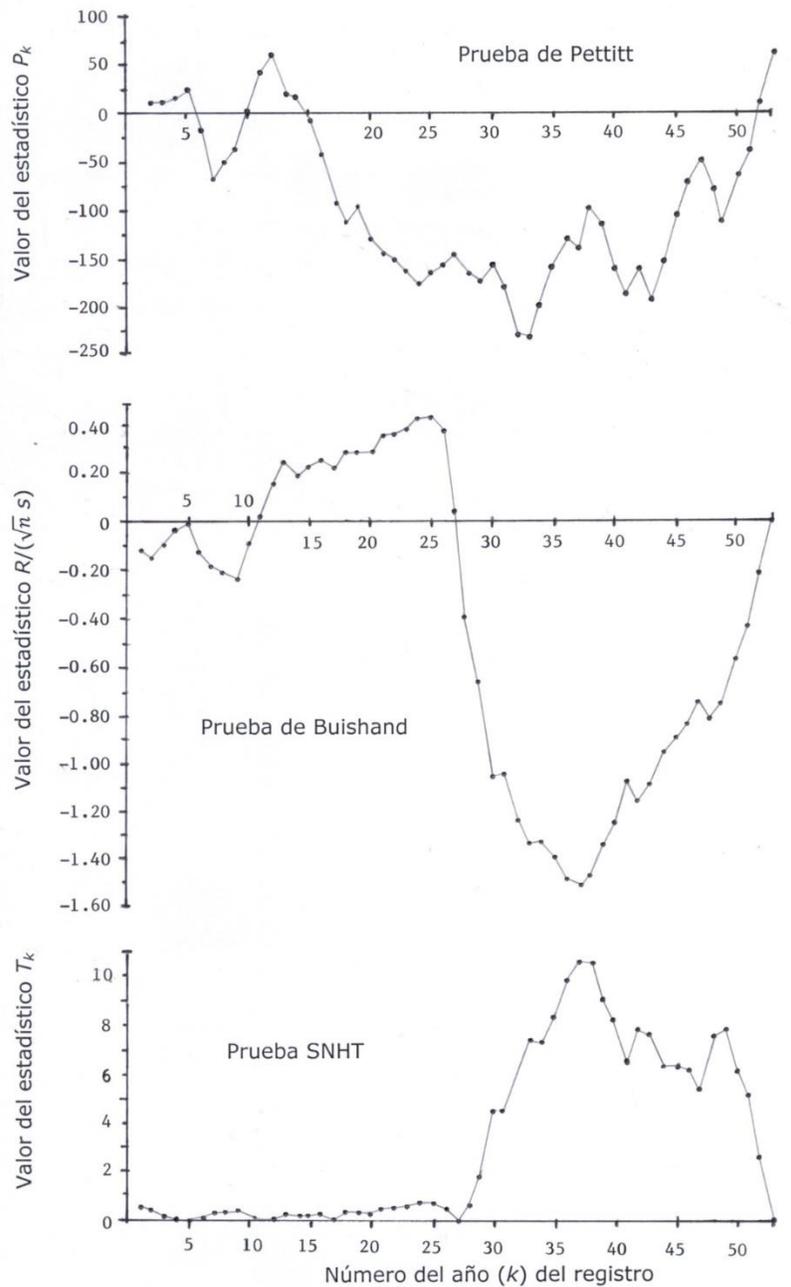
En la Figura 3 se muestran los gráficos de los estadísticos de las pruebas de Pettitt, Buishand y SNHT, aplicadas al registro de la estación Moctezuma. Se observa que las dos primeras definen valores máximos extremos positivos y negativos, y que la prueba SNHT define un máximo positivo en  $k = 45$  (año 2008). Por lo tanto, no existe un único punto de quiebre según las tres pruebas; pero dos de ellas lo definen en  $k = 45$ . En realidad, en 2008 comienza un periodo de siete años extremadamente

secos, cuya presencia fue ratificada en tal zona, para una sequía regional que afectó toda la altiplanicie mexicana. Por lo anterior, no procede la aplicación de la Ecuación (10) en  $k = 45$ , para corregir el registro previo.



**Figura 3.** Gráficos de las pruebas estadísticas indicadas, aplicadas al registro de precipitación anual de la estación Moctezuma del Altiplano Potosino, México.

Por otra parte, en la Figura 4 se exponen los gráficos de las pruebas de Pettitt, Buishand y SNHT, aplicadas al registro de la estación El Mezquite. Las dos primeras definen valores máximos extremos positivos y negativos, y la tercera un único máximo positivo en  $k = 37$  (año 2000). Aceptando tal año como un punto de quiebre, se define un lapso de 2000 a 2016 con 17 valores, y una media de 579.365 mm para después del cambio, y de 36 datos y una media de 327.811 mm para antes del cambio. Con la Ecuación (10) se obtiene un valor de 1.7674 como factor de corrección ( $f$ ).



**Figura 4.** Gráficos de las pruebas estadísticas indicadas, aplicadas al registro de precipitación anual de la estación El Mezquite del Altiplano Potosino, México.

En registro corregido de la estación El Mezquite tiene los siguientes parámetros estadísticos: mínimo = 59.9 mm; máximo = 1344.1 mm;  $\bar{X}$  = 579.4 mm; mediana = 567.3 mm;  $C_v$  = 0.480;  $C_s$  = 0.467;  $C_k$  = 3.486, y  $r_1$  = 0.503. Además, de acuerdo con las pruebas de Shapiro y Geary, resultó que puede proceder de una distribución normal. Los nuevos valores máximos de los estadísticos de las pruebas de Von Neumann, Pettitt, Buishand y SNHT fueron  $N = 1.016$ ;  $P_k = 182.0$ ;  $R/(\sqrt{n} s) = 2.03$ , y  $T_k = 8.02$ . Por lo cual resulta no homogéneo según la primera y tercera prueba, pasando a ser un registro menos confiable con la corrección realizada.

## Conclusiones

El planteamiento expuesto (Figura 2), al procesar la información climática de un área geográfica, permite trabajar con magnitudes que son semejantes, y al establecer un periodo común, tales datos deben mostrar también un comportamiento similar. Lo anterior se puede apreciar en los valores y parámetros estadísticos expuestos en la Tabla 2 y Tabla 3, relativos a los 16 registros amplios de precipitación anual procesados del Altiplano Potosino de México.

Los resultados de este estudio se pueden concretar en lo siguiente: se obtuvo que 11 registros son confiables, es decir, homogéneos; tres son menos confiables (Vanegas, Palo Blanco y Villa de Arriaga), y dos son no confiables (Moctezuma y El Mezquite). Los tres primeros registros no homogéneos deben su inconsistencia a la persistencia, pues tienen coeficientes de correlación serial de orden uno ( $r_1$ ) significativos.

Respecto al registro no confiable de la estación Moctezuma, los gráficos de sus pruebas de Pettitt y Buishand (Figura 3) muestran un comportamiento diferente, sin coincidencia del máximo, dejando claro que no existe punto de quiebre. Por el contrario, los gráficos de la estación El Mezquite (Figura 4), con las pruebas de Buishand y SNHT, tienden a definir un máximo en el año 2000 ( $k = 37$ ); al corregir los datos antes del cambio, se obtiene un registro menos confiable.

La prueba de Pettitt no se considera suficientemente potente o eficiente, ya que no mostró pérdida de homogeneidad en ninguno de los cinco registros no homogéneos, que fueron detectados por las otras tres pruebas aplicadas; lo anterior, en registros de precipitación anual de la zona semiárida del estado de San Luis Potosí, México.

### **Agradecimientos**

Se agradecen las sugerencias y recomendaciones de los dos árbitros anónimos B y C, las cuales permitieron mejorar la descripción del enfoque expuesto para probar pérdida de homogeneidad en registros climatológicos debido a alteraciones no climáticas. La figura 2 se incluyó a sugerencia del árbitro B, y el análisis de los registros no homogéneos,

con base en los resultados de las pruebas de Pettitt, Buishand y SNHT, se formuló por recomendación de ambos árbitros.

## Referencias

- Alexandersson, H. (1986). A homogeneity test applied to precipitation data. *Journal of Climatology*, 6(6), 661-675.
- Alexandersson, H., & Moberg, A. (1997). Homogenization of Swedish temperature data. Part I: Homogeneity test for linear trends. *International Journal of Climatology*, 17(1), 25-34.
- Beaulieu, C., Ouarda, T. B. M. J., & Seidou, O. (2007). Synthèse des techniques d'homogénéisation des séries climatiques et analyse d'applicabilité aux series de precipitations. *Hydrological Sciences Journal*, 52(1), 18-37.
- Beaulieu, C., Seidou, O., Ouarda, T. B. M. J., Zhang, X., Boulet, G., & Yagouti, A. (2008). Intercomparison of homogenization techniques for precipitation data. *Water Resources Research*, 44(2), W02425.
- Buishand, T. A. (1982). Some methods for testing the homogeneity of rainfall records. *Journal of Hydrology*, 58(1-2), 11-27.
- Campos-Aranda, D. F. (2005). FDP Gamma Mixta (Anexo B). En: *Agroclimatología Cuantitativa de Cultivos* (pp. 267-281). México, DF, México: Editorial Trillas.
- Campos-Aranda, D. F. (2013). Caracterización de sequías meteorológicas mediante curvas de severidad-área-frecuencia, en el valle de San Luis Potosí, México. *Tecnología y ciencias del agua*, 4(3), 165-186.

- Campos-Aranda, D. F. (2015). Búsqueda del cambio climático en la temperatura máxima de mayo en 16 estaciones climatológicas del estado de Zacatecas, México. *Tecnología y ciencias del agua*, 6(3), 143-160.
- Dhorde, A. G., & Zarenistanak, M. (2013). Three-way approach to test data homogeneity: An analysis of temperatures and precipitation series over southwestern Islamic Republic of Iran. *Journal of Indian Geophysical Union*, 17(3), 233-242.
- Ducré-Robitaille, J. F., Vincent, L. A., & Boulet, G. (2003). Comparison of techniques for detection of discontinuities in temperature series. *International Journal of Climatology*, 23(9), 1087-1101.
- Eischeid, J. K., Baker, C. B., Karl, T. R., & Diaz, H. F. (1995). The quality control of long-term climatological data using objective data analysis. *Journal of Applied Meteorology*, 34(12), 2787-2795.
- Guajardo-Panes, R. A., Granados-Ramírez, G. R., Sánchez-Cohen, I., Díaz-Padilla, G., & Barbosa-Moreno, F. (2017). Validación espacial de datos climatológicos y pruebas de homogeneidad: caso Veracruz, México. *Tecnología y ciencias del agua*, 8(5), 157-177.
- Guijarro, J. A. (2014). Chapter 24. Quality control and homogenization of climatological series. In: Eslamian, S. (ed.). *Handbook of Engineering Hydrology. Fundamentals and Applications* (pp. 501-513). Boca Raton, USA: CRC Press.
- González-Rouco, J. F., Jiménez, J. L., Quesada, V., & Valero, F. (2001). Quality control and homogeneity of precipitation data in the southwest of Europe. *Journal of Climate*, 14(5), 964-978.

- Hanssen-Bauer, I., & Forland, E. J. (1994). Homogenizing long Norwegian precipitation series. *Journal of Climate*, 7(6), 1001-1013.
- Khaliq, M. N., & Ouarda, T. B. M. J. (2007). On the critical values of the standard normal homogeneity test (SNHT). *International Journal of Climatology*, 27(5), 681-687.
- Lanzante, J. R. (1996). Resistant, robust and non-parametric techniques for the analysis of climate data: Theory and examples, including applications to historical radiosonde station data. *International Journal of Climatology*, 16(11), 1197-1226.
- Linsley, R. K., Kohler, M. A., & Paulhus, J. L. (1988). Chapter 14. Stochastic hydrology. In: *Hydrology for Engineers* (pp. 374-397). London, England: McGraw-Hill Book Co., SI Metric edition.
- Machiwal, D., & Jha, M. K. (2012). Chapter 34. Methods for testing normality of hydrologic time series and methods for time series analysis (pp. 32-84). In: *Hydrologic time series analysis: Theory and practice*. Dordrecht, The Netherlands: Springer.
- Mallakpour, I., & Villarini, G. (2016). A simulation study to examine the sensitivity of the Pettitt test to detect abrupt changes in mean. *Hydrological Sciences Journal*, 61(2), 245-254.
- Mishra, A. K., & Singh, V. P. (2008). Chapter 23. Development of drought SAF curves. In: Singh, V. P. (ed.). *Hydrology and Hydraulics* (pp. 811-833). Highlands Ranch, Colorado, USA: Water Resources Publications.
- Moberg, A., & Alexandersson, H. (1997). Homogenization of Swedish temperature data. Part II: Homogenized gridded air temperature

- compared with a subset of global gridded air temperature since 1861. *International Journal of Climatology*, 17(1), 35-54.
- Paulhus, J. L. M., & Kohler, M. A. (1952). Interpolation of missing precipitation records. *Monthly Weather Review*, 80(8), 129-133.
- Peterson, T. C., Easterling, D. R., Karl, T. R., Groisman, P., Nichols, N., Plummer, N., Torok, S., Auer, I., Boehm, R., Gullett, D., Vincent, L., Heino, R., Tuomenvirta, H., Mestre, O., Szentimrey, T., Salinger, J., Førland, E. J., Hanssen-Bauer, I., Alexandersson H., Jones, P., & Parker, D. (1998). Homogeneity adjustments of *in situ* atmospheric climate data: a review. *International Journal of Climatology*, 18(13), 1493-1517.
- Pettitt, A. N. (1979). A non-parametric approach to the change-point problem. *Applied Statistics*, 28(2), 126-135.
- Rhoades, D. A., & Salinger, M. J. (1993). Adjustment of temperature and rainfall records for site changes. *International Journal of Climatology*, 13(8), 899-913.
- SARH, Secretaría de Agricultura y Recursos Hidráulicos. (1980). *Boletín Climatológico No. 3. Región Hidrológica No. 37 (El Salado)*. México, DF, México: Secretaría de Agricultura y Recursos Hidráulicos. Subsecretaría de Planeación, Dirección General de Estudios, Subdirección de Hidrología.
- Shapiro, S. S., & Wilk, M. B. (1965). An analysis of Variance Test for Normality (Complete Samples). *Biometrika*, 52(3-4), 591-611.
- Shapiro, S. S. (1998). Chapter 6. Selection, fitting and testing statistical models. In: Wadsworth, H. M. (ed.). *Handbook of statistical*

- methods for engineers and scientists*, 2<sup>nd</sup> ed. (pp. 6.1-6.35). New York, USA: McGraw-Hill Inc.
- Syrakova, M. & Stefanova, M. (2009). Homogenization of Bulgarian temperature series. *International Journal of Climatology*, 29(12), 1835-1849.
- Toreti, A., Kuglitsch, F. G., Xoplaki, E., Della-Marta, P. M., Aguilar, E., Prohom, M., & Luterbacher, J. (2011). A note on the use of the standard normal homogeneity test to detect inhomogeneity in climatic time series. *International Journal of Climatology*, 31(4), 630-632.
- Tuomenvirta, H. (2001). Homogeneity adjustments of temperature and precipitation series-Finnish and Nordic data. *International Journal of Climatology*, 21(4), 495-506.
- Wijngaard, J. B., Klein-Tank, A. M. G., & Können, G. P. (2003). Homogeneity of 20<sup>th</sup> century European daily temperature and precipitation series. *International Journal of Climatology*, 23(6), 679-692.
- WMO, World Meteorological Organization. (1971). Annexed III: Standard tests of significance to be recommended in routine analysis of climatic fluctuations. In: *Climatic change* (pp. 58-71) (Technical Note No. 79). Geneva, Switzerland: World Meteorological Organization.
- Young, K. C. (1992). A three-way model for interpolating for monthly precipitation values. *Monthly Weather Review*, 120(11), 2561-2569.

Yozgatligil, C., & Yazici, C. (2016). Comparison of homogeneity tests for temperature using a simulation study. *International Journal of Climatology*, 36(1), 62-81.