

DOI: 10.24850/j-tyca-2020-04-06

Artículos

Predicción de lluvias máximas para la república mexicana mediante modelos probabilísticos no estacionarios

Prediction of maximum rainfall in Mexico through non-stationary probabilistic models

Gabriela Álvarez-Olguín¹, <https://orcid.org/0000-0002-7144-2426>

Saúl Martínez-Ramírez², <https://orcid.org/0000-0003-1014-2427>

Brenda I. G. Licona-Morán³, <https://orcid.org/0000-0003-1309-8871>

¹Universidad Tecnológica de la Mixteca, Huajuapán de León, Oaxaca, México, galvarez@mixteco.utm.mx

²Universidad Tecnológica de la Mixteca, Huajuapán de León, Oaxaca, México, saulmr@mixteco.utm.mx

³Universidad Tecnológica de la Mixteca, Huajuapán de León, Oaxaca, México, brenda@mixteco.utm.mx

Autora para correspondencia: Gabriela Álvarez Olguín,
galvarez@mixteco.utm.mx

Resumen

La predicción de eventos de lluvia máxima es la base del diseño de estructuras hidráulicas para la mitigación de inundaciones. Esta predicción se hace tradicionalmente a partir de métodos de análisis de frecuencias, que consisten en estudiar eventos pasados para estimar las probabilidades de ocurrencias futuras. Sin embargo, debido a que la variabilidad climática provoca cambios en la media y la varianza de las series de tiempo de lluvia, los eventos de diseño no son confiables si se estiman a partir de técnicas válidas para condiciones estacionarias. Para México, existe evidencia de que los patrones de lluvia se están modificando; en consecuencia, para que las predicciones sean confiables, se requieren aplicar métodos que contemplen cambios en las características estadísticas de los datos a través del tiempo. Por lo anterior, el objetivo de este trabajo fue estimar eventos de lluvia máxima en 24 horas, a través de modelos probabilísticos no estacionarios. Se analizaron 769 series a las que se aplicaron las pruebas de Pettitt, Mann-Kendall y Descomposición de Modos Empíricos por Conjuntos, para verificar su estacionaridad. Se propusieron diferentes modelos probabilísticos, en los que parámetros de las funciones de distribución lognormal, gamma, gumbel, weibull y general de valores extremos tienen como covariables al tiempo y al índice de oscilación decadal del Pacífico. Los resultados indican que para las series no estacionarias los modelos propuestos representan mejor la variabilidad de los datos que los modelos estacionarios convencionales.

Palabras clave: análisis de frecuencias, eventos hidrometeorológicos extremos, oscilación decadal del Pacífico, puntos de cambio, tendencias y variabilidad climática.

Abstract

The prediction of maximum rainfall is the basis of the design of hydraulic structures for flood mitigation. This prediction is traditionally made from frequency analysis methods that consist of studying past events in order to define the probabilities of future occurrences. However, because climatic variability causes changes in the mean and variance of rainfall time series, design events are unreliable if they are estimated from techniques valid for stationary conditions. For Mexico, there is evidence that rainfall patterns are changing, consequently for predictions to be reliable, it is necessary to apply methods that contemplate changes in the statistical characteristics of the data over time. Therefore, the aim of this study was to estimate annual maximum 24-hour rainfall events associated with different return periods, through non-stationary probabilistic models. We analyzed 769 series to which were applied the Pettitt, Mann-Kendall and Ensemble Empirical Mode Decomposition tests to verify the seasonality. Different probabilistic models were proposed, in which parameters of the Lognormal, Gamma, Gumbel, Weibull and Generalized Extreme Value distribution have as covariates to time and Pacific Decadal Oscillation index. The results indicated that for the non-stationary series the proposed models represent the variability of the data better than conventional stationary models.

Keywords: Frequency analysis, extreme hydro-meteorological events, Pacific Decadal Oscillation, change point and trend and climatic variability.

Recibido: 03/05/2019

Aceptado: 22/10/2019

Introducción

Los patrones de las variables hidrometeorológicas están cambiando (Gleick, 1989; Voss, May, & Roeckner, 2002; Held & Söden, 2006) y se espera que en el futuro los desastres provocados por eventos hidrometeorológicos sean más frecuentes e intensos debido al calentamiento global (IPCC, 2007). Una forma de reducir el riesgo de tales desastres es por medio de una planeación basada en la probabilidad de los eventos extremos. Por ejemplo, para reducir el riesgo de inundaciones se deben construir estructuras hidráulicas para drenaje y control de escurrimientos, diseñadas a partir de eventos de lluvia asociados con un periodo de retorno determinado.

La metodología clásica para predecir eventos de diseño de variables hidrometeorológicas se basa en la teoría de valores extremos de series

de tiempo estacionarias (Khaliq, Ouarda, Ondo, Gachon, & Bobée, 2006; Villarini & Smith, 2010). Bajo condiciones de estacionaridad, la distribución de la variable de interés es invariante en el tiempo, sin tendencias, cambios ni periodicidades (Villarini & Smith, 2010). No obstante, la estacionaridad de las series de tiempo ha dejado de existir debido al cambio en el clima del planeta, que está alterando la media y los extremos de la precipitación, evapotranspiración y gastos de los ríos (Milly *et al.*, 2008). La no estacionaridad de las series hidrometeorológicas debida al cambio climático por lo usual se manifiesta como una tendencia o un salto repentino en las características estadísticas de los datos (Khaliq *et al.*, 2006). La presencia de una tendencia tiene un efecto considerable en el ajuste de una distribución de probabilidad convencional a una muestra de observaciones no estacionarias; por lo tanto, la comprensión y el diagnóstico de un comportamiento no estacionario es muy importante (Khaliq *et al.*, 2006).

En México existe evidencia del incremento en la frecuencia e intensidad de los eventos extremos (Jáuregui, 2000; Peralta-Hernández, Balling, & Barba-Martínez, 2009), por lo que las técnicas tradicionales de predicción son incapaces de estimar con precisión su magnitud. Si la predicción de eventos extremos de diseño no es adecuada, las obras hidráulicas tienen mayor riesgo de falla, y sus efectos pudieran ser socialmente inaceptables aguas abajo. Por el contrario, una sobreestimación de los eventos incrementaría la seguridad de la estructura, pero con elevados costos económicos (Escalante-Sandoval & Reyes-Chávez, 2010). Ante esta situación, se requieren emplear técnicas que consideren la falta de estacionaridad de las series. Sin embargo, para

México, son pocas las investigaciones (Campos-Aranda, 2016; Álvarez-Olguín & Escalante-Sandoval, 2016; Álvarez-Olguín & Escalante-Sandoval, 2017) que incluyen la falta de estacionaridad de las series en los análisis de predicciones hidrometeorológicas. Por lo anterior, el objetivo de este trabajo fue estimar para la república mexicana eventos de lluvia máxima en 24 horas, asociados con periodos de retorno de 2, 5, 10, 50 y 100 años, mediante modelos probabilísticos no estacionarios. Los resultados muestran la influencia de la variabilidad climática en la predicción de la lluvia máxima, lo cual es útil para obtener mejores estimaciones que podrían contribuir a disminuir el riesgo de falla de las obras hidráulicas.

Datos y métodos

Los datos utilizados corresponden a la lluvia diaria medida en las estaciones climatológicas convencionales del Servicio Meteorológico Nacional de la Comisión Nacional del Agua (Conagua). Esta información se extrajo de la base de datos *Clicom* (*Climate Computing Project*) y sólo se consideraron estaciones con al menos 64 años de registros en el periodo de 1950 a 2013, y por lo menos 90% de información disponible. La localización de estas estaciones se muestra en la Figura 1. Para deducir

datos faltantes se aplicó el método de interpolación de la distancia inversa ponderada (Shepard, 1968), considerando dos estaciones de apoyo y un exponente de distancia igual a dos. Se verificó si existe mejora en la estimación de la media y la varianza de las muestras extendidas, mediante el criterio de información relativa (I), definido con la Ecuación (1), descrita por Escalante-Sandoval y Reyes-Chávez (2002).

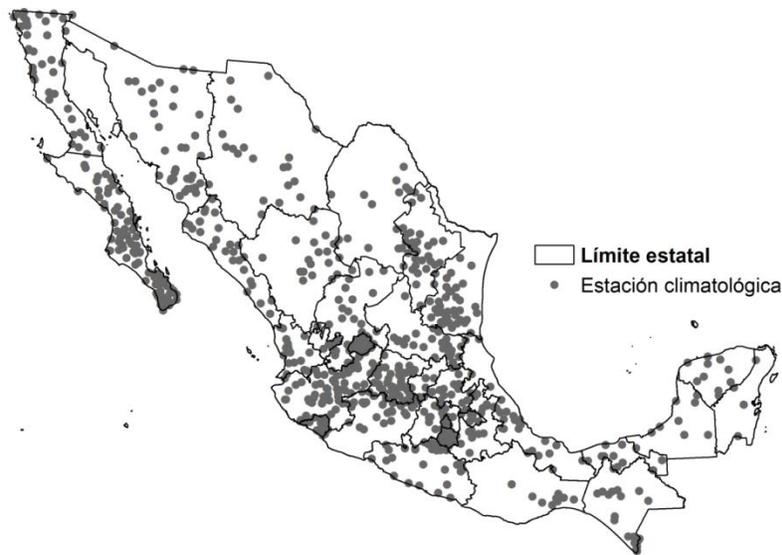


Figura 1. Localización de las 769 estaciones usadas en el estudio.

$$I = \frac{\text{var}(s_{y_1}^2)}{\text{var}(s_y^2)} \quad (1)$$

Donde $Var(S_{y_1}^2)$ es la varianza de la varianza de la serie original y $Var(S_y^2)$ es la varianza de la varianza de la serie extendida.

Si $I > 1$, entonces la varianza de la varianza de la serie extendida no excede a la original; en consecuencia, la extensión de los registros es adecuada. De esta forma, se obtuvieron datos confiables de lluvia diaria de 769 estaciones, con los que se lograron las series anuales de lluvia máxima en 24 horas, para el periodo comprendido entre 1950 y 2013.

El siguiente paso fue verificar la estacionaridad de las series anuales de lluvia máxima en 24 horas. Para comparar de modo espacial los cambios ocurridos en las series de lluvia, se tomó como base la regionalización de la lluvia obtenida por Álvarez-Olguín y Escalante-Sandoval (2017). La homogeneidad de las series se verificó mediante la ausencia de cambios abruptos en la media, a través de la prueba de Pettitt (Pettitt, 1979). La presencia de tendencias en las series se evaluó con la prueba de Mann-Kendall (Kendall, 1975), aplicada a las series anuales de lluvia máxima en 24 horas. Las tendencias también se evaluaron mediante la gráfica de los residuales de las series obtenidas por el método de descomposición de modos empíricos por conjuntos (EEMD, *Ensemble Empirical Mode Decomposition*), desarrollado recientemente (Huang *et al.*, 1998; Wu & Huang, 2009) para descomponer señales en funciones de modos intrínsecos (*Intrinsic Mode Function*). En este método, el cociente de la desviación estándar de la señal de ruido blanco gaussiano adicionada a los datos fue de 0.2.

Es común que la no estacionaridad de una serie hidrometeorológica se deba a la presencia de una tendencia y/o un salto repentino en las

características estadísticas de los datos (Yevjevich & Jeng, 1969; Khaliq *et al.*, 2006). La presencia de una tendencia tiene un efecto considerable en la interpretación de los resultados cuando se ajusta una distribución de probabilidad convencional a una muestra de observaciones no estacionarias (Khaliq *et al.*, 2006). No se ha desarrollado una teoría general para el análisis de frecuencia de procesos no estacionarios, pero una práctica aceptada consiste en modificar las distribuciones de probabilidad estándar para obtener versiones no estacionarias de las mismas (Coles, 2001). Una distribución no estacionaria combina parámetros de la distribución estacionaria con algunas covariables, como el tiempo o índices de circulación atmosférica a gran escala (ENOS, DOP, OMA y NAO), para generar un nuevo tipo de parámetro para la estimación. Por ejemplo, de la distribución general de valores extremos (GVE), con parámetros ν , α y β (ubicación, escala y forma, respectivamente), se puede obtener un modelo para estimar una variable x_t en función del tiempo, como se indica en la Ecuación (2).

$$x_t \sim GEV(\nu(t), \alpha(t), \beta(t)) \quad (2)$$

Donde cada parámetro $\nu(t)$, $\alpha(t)$ y $\beta(t)$ tiene una expresión en términos del tiempo. El parámetro $\nu(t)$ se puede expresar como se muestra en la Ecuación (3).

$$\nu(t) = \nu_0 + \nu_1 t \quad (3)$$

Donde t es un índice que depende del tiempo. Así, las variaciones a través del tiempo en los procesos observados se modelan como una tendencia lineal.

En esta investigación, el análisis de las series se realizó con las funciones para valores máximos descritos en la Tabla 1 y los modelos incluidos en la Tabla 2. Para el caso estacionario, se utilizaron los modelos M_0 ; los modelos M_1 , M_2 y M_3 corresponden al caso no estacionario, y sus parámetros se expresan en términos del tiempo e índice de oscilación decadal del Pacífico (PDO). Este índice se seleccionó por ser una de las fluctuaciones de periodo largo de mayor correlación con la lluvia en México (Álvarez-Olguín & Escalante-Sandoval, 2017). Las series anuales del índice PDO se obtuvieron del promedio de los valores mensuales para el periodo comprendido entre 1950 y 2013 (NOAA, 2014).

Tabla 1. Funciones de densidad de probabilidad para valores máximos utilizadas.

Distribución	Función de distribución	Restricción
Log Normal 3p (LN3)	$f(x) = \frac{1}{(x - x_0)\sigma_z\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln(x-x_0)-\mu_z}{\sigma_z}\right)^2}$	$x > x_0$
Gamma 3p (GA3)	$f(x) = \frac{1}{\alpha\Gamma(\beta)} \left(\frac{x - x_0}{\alpha}\right)^{\beta-1} e^{-\left(\frac{x-x_0}{\alpha}\right)}$	$\alpha > 0$
Gumbel (G)	$f(x) = \frac{1}{\alpha} e^{-e^{-\left(\frac{x-\nu}{\alpha}\right)}} e^{-\left(\frac{x-\nu}{\alpha}\right)}$	$\alpha > 0$

Weibull (W)	$f(x) = \left(\frac{k}{\varphi}\right) \left(\frac{x}{\varphi}\right)^{k-1} e^{-\left(\frac{x}{\varphi}\right)^k}$	$\varphi > 0, \kappa > 0$
General de valores extremos (GVE)	$f(x) = \frac{1}{\alpha} e^{-\left[1+\beta\left(\frac{x-u}{\alpha}\right)\right]^{\frac{1}{\beta}}} \left[1 + \beta \left(\frac{x-u}{\alpha}\right)\right]^{-1-\frac{1}{\beta}}$	$\alpha > 0$

Tabla 2. Modelos propuestos para los parámetros de las funciones de distribución de probabilidad para valores máximos.

Distribución	Estacionario (M_0)	No estacionario		
		M_1	M_2	M_3
Log Normal 3p (LN3)	$\mu_z(t) = \mu_0$ $\sigma_z(t) = \sigma_z$ $x_0(t) = x_0$	$\mu_z(t) = \mu_0 + \mu_1 t$ $\sigma_z(t) = \sigma_z$ $x_0(t) = x_0$	$\mu_z(t) = \mu_0 + \mu_1 \text{PDO}_t$ $\sigma_z(t) = \sigma_z$ $x_0(t) = x_0$	$\mu_z(t) = \mu_0 + \mu_1 t + \mu_2 \text{PDO}_t$ $\sigma_z(t) = \sigma$ $x_0(t) = x_0$
Gamma 3p (GA3)	$x_0(t) = \lambda_0$ $\alpha(t) = \alpha$ $\beta(t) = \beta$	$x_0(t) = \lambda_0 + \lambda_1 t$ $\alpha(t) = \alpha$ $\beta(t) = \beta$	$x_0(t) = \lambda_0 + \lambda_1 \text{PDO}_t$ $\alpha(t) = \alpha$ $\beta(t) = \beta$	$x_0(t) = \lambda_0 + \lambda_1 t + \lambda_2 \text{PDO}_t$ $\alpha(t) = \alpha$ $\beta(t) = \beta$
Gumbel (G)	$v(t) = v_0$ $\alpha(t) = \alpha$	$v(t) = v_0 + v_1 t$ $\alpha(t) = \alpha$	$v(t) = v_0 + v_1 \text{PDO}_t$ $\alpha(t) = \alpha$	$v(t) = v_0 + v_1 t + v_2 \text{PDO}_t$ $\alpha(t) = \alpha$
Weibul (W)	$\varphi(t) = \varphi_0$ $\kappa(t) = \kappa$	$\varphi(t) = \varphi_0 + \varphi_1 t$ $\kappa(t) = \kappa$	$\varphi(t) = \varphi_0 + \varphi_1 \text{PDO}_t$ $\kappa(t) = \kappa$	$\varphi(t) = \varphi_0 + \varphi_1 t + \varphi_2 \text{PDO}_t$ $\kappa(t) = \kappa$
General de valores ex. (GVE)	$v(t) = v_0$ $\alpha(t) = \alpha$ $\beta(t) = \beta$	$v(t) = v_0 + v_1 t$	$v(t) = v_0 + v_1 \text{PDO}_t$	$v(t) = v_0 + v_1 t + v_2 \text{PDO}_t$

		$\alpha(t) = \alpha$	$\alpha(t) = \alpha$	$\alpha(t) = \alpha_0$
		$\beta(t) = \beta$	$\beta(t) = \beta$	$\beta(t) = \beta$

t es un índice que depende del tiempo, para 1950 $t = 1$.

PDO_t es el valor promedio anual del índice PDO, correspondiente a t .

La estimación de los parámetros de los modelos se realizó por medio del método de máxima verosimilitud, que se considera el más eficiente, pues proporciona la menor varianza muestral de los parámetros y de los eventos estimados, en comparación con otros métodos. Además, tiene la ventaja de adaptarse a cambios en la estructura del modelo (Katz, Parlange, & Naveau, 2002).

Para el caso de la función de densidad de probabilidad GVE (general de valores extremos), la función logarítmica de verosimilitud se expresa como:

$$l = -n \ln(\alpha) - \sum_{t=1}^n \left[1 + \beta \left(\frac{x_t - v(t)}{\alpha} \right) \right]^{-\frac{1}{\beta}} - \left(1 + \frac{1}{\beta} \right) \sum_{t=1}^n \ln \left[1 + \beta \left(\frac{x_t - v(t)}{\alpha} \right) \right] \quad (4)$$

Donde $v(t)$, α y β son los parámetros de ubicación, escala y forma, respectivamente.

Para el modelo GVE_ M_3 (modelo M_3 de la función general de valores extremos), se consideró al parámetro de escala en función del tiempo y del índice PDO, por lo tanto, la función logarítmica de verosimilitud propuesta es:

$$l = -n \ln(\alpha) - \sum_{t=1}^n \left[1 + \beta \left(\frac{x_t - (v_0 + v_1 t + v_2 PDO_t)}{\alpha} \right) \right]^{-\frac{1}{\beta}} - \left(1 + \frac{1}{\beta} \right) \sum_{t=1}^n \ln \left[1 + \beta \left(\frac{x_t - (v_0 + v_1 t + v_2 PDO_t)}{\alpha} \right) \right] \quad (5)$$

Donde v_0 , v_1 y v_2 son los parámetros de ubicación; α es el parámetro de escala y β es el parámetro de forma; t es un índice que depende del tiempo ($t = 1$ para el año 1950), y PDO_t es el valor promedio anual del índice PDO correspondiente a t .

Los modelos que mejor describen la variabilidad de los datos se seleccionaron mediante el criterio de información de Akaike (Akaike, 1974), donde se utilizó la Ecuación (6). El mejor modelo es el que tuvo el menor valor de AIC. En el caso de tener valores similares, se seleccionó el modelo con menos parámetros.

$$AIC = -2(l) + 2K \quad (6)$$

Donde l es el máximo valor de la función de verosimilitud y K es el número de parámetros estimados.

Además, se utilizó el método descrito por Coles (2001) para comparar la validez de un modelo M_1 contra otro M_0 , tal que $M_0 \subset M_1$, en el cual se utiliza la medida de discordancia definida por la Ecuación (7). Valores grandes de D indican que el modelo M_1 es más adecuado y explica mejor la variación de los datos, en comparación con el modelo M_0 . El

estadístico D se distribuye de acuerdo con la distribución *chi*-cuadrada (χ^2_v). El parámetro v es la diferencia entre el número de parámetros de los modelos M_1 y M_0 . Valores de D más grandes que las cantidades de la distribución χ^2_v para un nivel de confianza particular se consideran significantes, entonces se rechaza M_0 en favor del modelo M_1 :

$$D = 2\{l_1(M_1) - l_0(M_0)\} \quad (7)$$

Donde $l_i(M_i)$ es la función logarítmica de verosimilitud maximizada del modelo M_i .

Para verificar que los modelos seleccionados se ajustan de forma adecuada a los datos observados, fue necesario transformarlos mediante una estandarización (Krzysztofowicz, 1997). De esta manera, tras ordenar los valores estandarizados y asociarlos con los correspondientes de la distribución empírica de Weibull, se elaboraron las gráficas de residuales (Q-Q) y *worm plots* (Buuren & Fredriks, 2001). En una "worm plot", el eje vertical es la diferencia entre los valores empíricos y los teóricos, y contiene el intervalo de confianza al 95%, estimado como:

$$\pm 1.96g(z)^{-1}\sqrt{(p(1-p)/n)} \quad (8)$$

Donde $g(z)$ es la función de densidad normal; z , un evento asociado con una probabilidad p , y n es el tamaño de la muestra.

Por último, se estimaron eventos de diseño relacionados con diferentes periodos de retorno. Para el modelo GVE_3, los eventos de diseño se estimaron por medio de la Ecuación (9):

$$\hat{X} = \hat{v}_0 + \hat{v}_1 t_0 + \hat{v}_2 \widehat{PDO} - \frac{\hat{\alpha}}{\hat{\beta}} \{1 - [-\ln(p)]^{-\hat{\beta}}\} \quad (9)$$

Donde \hat{v}_0 , \hat{v}_1 y \hat{v}_2 son los estimadores del parámetro de ubicación; $\hat{\alpha}$ es el estimador del parámetro de escala; $\hat{\beta}$ es el estimador del parámetro de forma; t_0 es el valor del índice del tiempo para un año determinado; \widehat{PDO} es un valor promedio del índice PDO y \hat{X} es el evento de diseño asociado con una probabilidad de no excedencia p . El valor \widehat{PDO} se estimó como la media aritmética de los valores PDO correspondientes al mayor periodo de anomalías, es decir, al mayor periodo con persistencia de eventos superiores a cero.

Los datos utilizados en los análisis son del periodo 1950-2013. Por lo tanto, al primer año (1950) le corresponde el valor de $t_0 = 1$; para los años siguientes, el valor de t_0 aumenta de modo consecutivo una unidad cada año. De esta forma, al escenario del año 2013 se asignó un valor de $t_0 = 64$ y al escenario de 2040 un valor de $t_0 = 94$.

Resultados y discusión

A través de la prueba de Pettitt se verificó la presencia de cambios abruptos en las series, mientras que con las pruebas de Mann-Kendall y EEMD se evaluaron tendencias. Las series identificadas como no estacionarias a un nivel de significancia de 0.05, se clasificaron en los siguientes tipos:

1. Series con cambio descendente: presentan un cambio abrupto descendente en la media. La media después del cambio es menor a la del periodo antes del cambio.
2. Series con cambio ascendente: presentan un cambio abrupto ascendente en la media. La media después del cambio es mayor a la del periodo antes del cambio.
3. Series con tendencia decreciente significativa: presentan un decrecimiento gradual a través del tiempo.
4. Series con tendencia creciente: presentan un crecimiento gradual a través del tiempo.
5. Series con cambio descendente y tendencia decreciente: tienen tanto un cambio abrupto descendente como una tendencia decreciente.
6. Series con cambio ascendente y tendencia creciente: tienen tanto un cambio abrupto ascendente como una tendencia creciente.

En total, 22% (167) de las series analizadas se consideró no estacionaria (Figura 2). Corresponde a estaciones ubicadas en las regiones 1 y 2, en el centro y noroeste del país, donde se tienen 67 y 32 estaciones, respectivamente (Tabla 3). De estas series, 89% tiene puntos de cambio ascendentes y/o tendencias monótonas crecientes significativas.

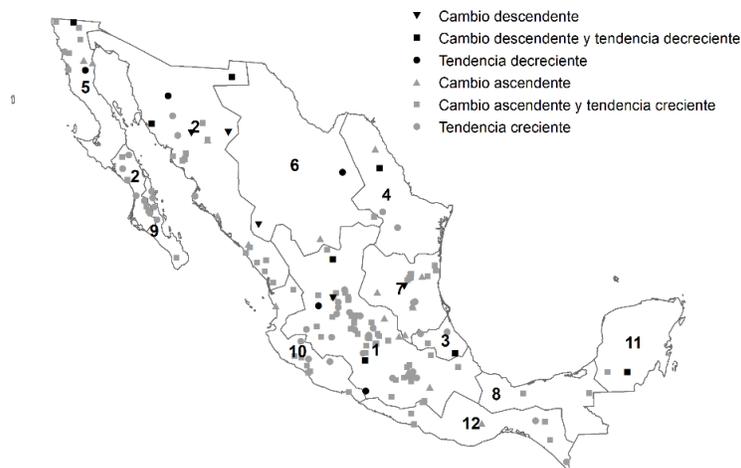


Figura 2. Series anuales no estacionarias de lluvia máxima en 24 horas.

Tabla 3. Cantidad de estaciones analizadas, series de lluvia máxima en 24 horas estacionarias, no estacionarias y tipo de series no estacionarias.

Región	Total	E	NE	Causa de no estacionaridad					
				A	D	TC	TD	A&TC	D&TD
1	270	203	67	4	1	16	2	42	2

2	106	74	32	3	2	8	1	16	2
3	18	14	4	0	0	2	0	1	1
4	45	41	4	1	0	2	0	0	1
5	43	31	12	4	0	0	1	6	1
6	45	41	4	1	1	0	1	1	0
7	72	56	16	5	1	2	0	8	0
8	23	21	2	0	0	0	0	2	0
9	41	36	5	0	0	4	0	1	0
10	51	40	11	1	0	3	0	7	0
11	20	18	2	0	0	0	0	1	1
12	35	27	8	1	0	2	0	5	0
Total	769	602	167	20	5	39	5	90	8

E = estacionaria; NE = no estacionarias; A = punto de cambio ascendente; D = punto de cambio descendente; TC = tendencia creciente; TD = tendencia decreciente; A&TC = punto de cambio ascendente y tendencia creciente; D&TD = punto de cambio descendente y tendencia decreciente.

En la Figura 3 se muestra la variación temporal de la lluvia máxima y los residuales de algunas series no estacionarias; para éstas se determinaron tendencias monótonas crecientes significativas. También se identificaron cambios ascendentes significativos en las series de las estaciones Acámbaro, El Novillo, Presa Rodríguez, La Servilleta y Cerro Ortega; en éstas, tanto la prueba de Pettitt como de Mann-Kendall indicaron incrementos abruptos y graduales, respectivamente. En el caso de Acámbaro, se encontró un punto de cambio en 1987 y una tendencia

creciente (Tabla 4). Además, con la gráfica del residual obtenido por EEMD se aprecia que la tendencia es monótona.

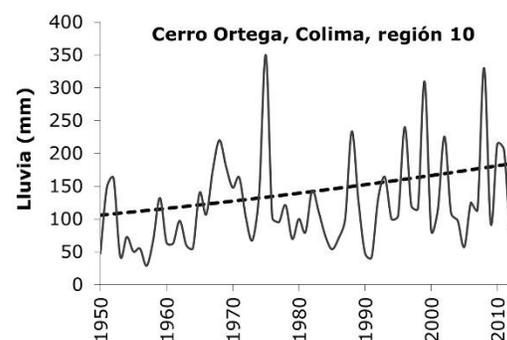
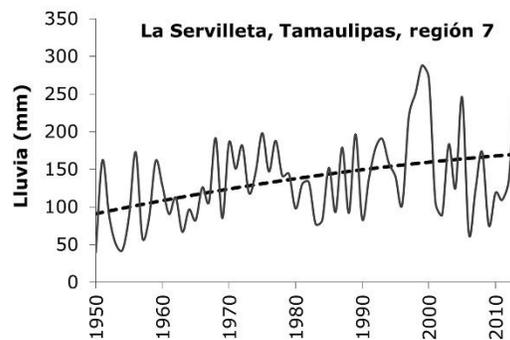
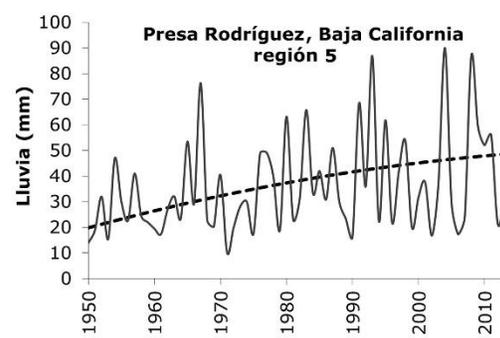
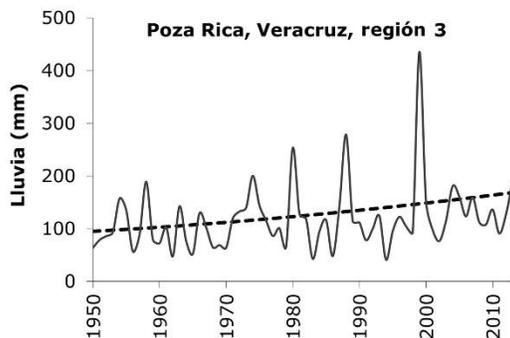
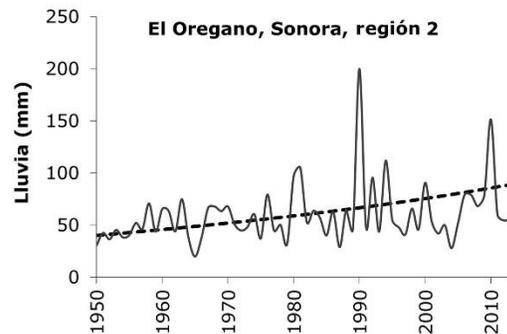
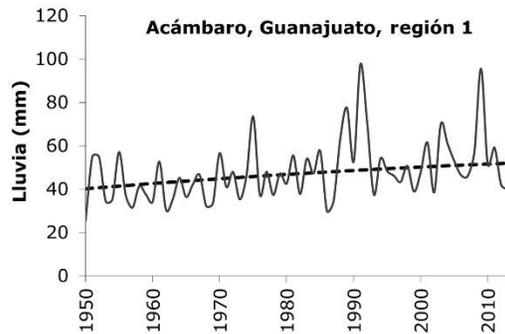


Figura 3. Series de lluvia anual máxima en 24 horas no estacionarias y residuales obtenidos por descomposición de modos empíricos por conjuntos (EEMD).

Tabla 4. Resultados de las pruebas de estacionaridad realizadas a las series de las estaciones San Vicente y Acámbaro.

Prueba	Estadístico/condición	Estación	
		Acámbaro	San Vicente
Pettitt	Punto de cambio	1987	1975
	K_T	529	429
	p -value	0.0036	0.031
	Cambio	Ascendente	Ascendente
	Significativo ($\alpha = 0.05$)	Sí	Sí
Mann-Kendall	τ	0.290	0.157
	S	585	317
	Z	3.38	1.83
	Tendencia	Creciente	Creciente
	p -value	0.0007	0.067
	Significativo ($\alpha = 0.05$)	Sí	No
Estacionaria		No	No
Causa de no estacionaridad		Cambio ascendente y tendencia	Cambio ascendente

Las tendencias crecientes indican que en los próximos años las lluvias extremas serán más intensas y habrá mayor probabilidad de que ocurran desastres relacionados con estos eventos, sobre todo en las regiones del centro y noroeste del país. El aumento en las intensidades de lluvia en Norteamérica está referido por Peterson, Zhang, Brunet-India y Vázquez-Aguirre (2008), y Cavazos *et al.* (2008). Entre las causas

probables de dicho cambio están el aumento del contenido de humedad atmosférica derivado del calentamiento global (IPCC, 2007; Meehl, Arblaster, & Tebaldi, 2005) y la influencia de fenómenos oscilatorios que regulan la variabilidad climática global.

Hay evidencias de la relación entre el comportamiento de la lluvia en México y fenómenos oscilatorios de gran escala como ENSO (El Niño *Southern Oscillation*), PDO (*Pacific Decadal Oscillation*) (Magaña, Vázquez, Pérez, & Pérez, 2003; Méndez-González, Návar-Cháidez, González-Rodríguez, & Treviño-Garza, 2007; Méndez-González, Ramírez-Leyva, Cornejo-Oviedo, Zárate-Lupercio, & Cavazos-Pérez, 2010; Méndez & Magaña, 2009). Por lo tanto, es de esperarse que influyan en las características estadísticas de las series. Ejemplo de la influencia de la PDO sobre las lluvias máximas en el noroeste del país se muestra en la serie de la estación San Vicente, Ensenada, Baja California (Figura 4), que se localiza en la región 5, la única del país con clima mediterráneo, donde 46.7% de la aportación a la precipitación anual se presenta durante el invierno (Álvarez-Olguín & Escalante-Sandoval, 2017). Para la serie, se identificó un punto de cambio ascendente significativo en el año 1975 (Tabla 4), el cual coincide con el inicio de la fase positiva de la PDO, que abarcó el periodo 1976-1998. Se observó que durante esta fase se presentaron tres de los eventos más intensos, en los años 1978, 1993 y 1997. La fase positiva de PDO favorece las lluvias invernales en el norte de México (Mantua, Hare, Zhang, Wallace, & Francis, 1997; Méndez *et al.*, 2010), por lo cual la falta de estacionaridad de la serie se puede atribuir a la influencia de la PDO.

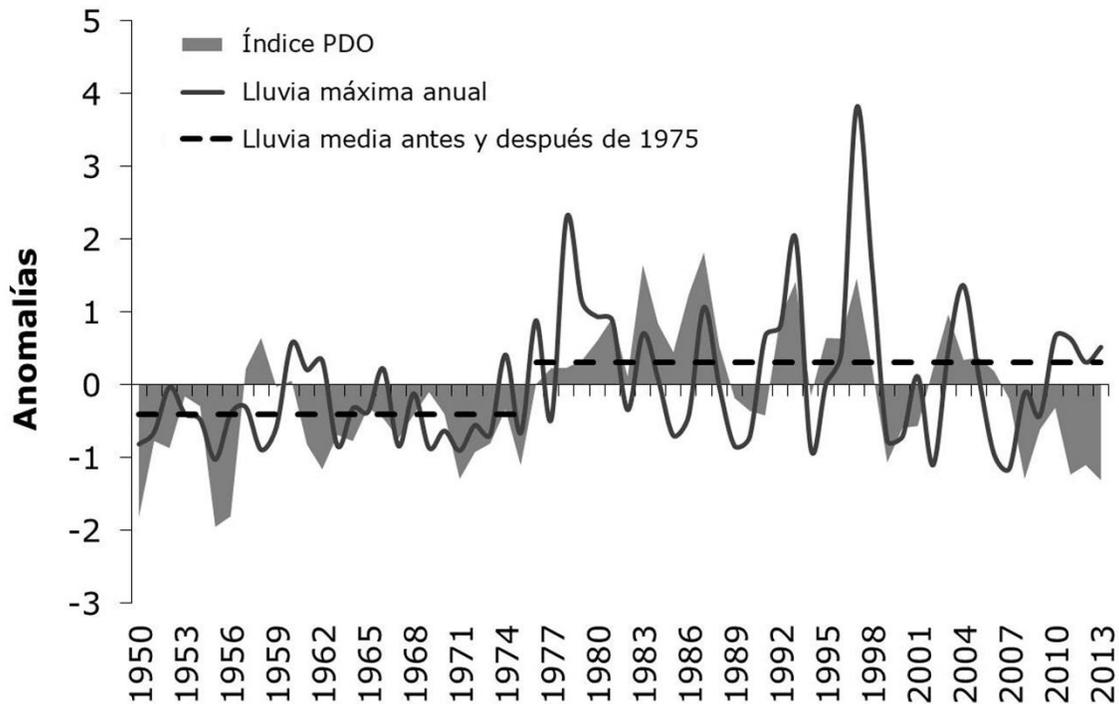


Figura 4. Lluvia anual máxima en 24 horas registrada en la estación San Vicente, Ensenada, Baja California.

Como resultado del análisis de frecuencias, se determinó que para 83% (139) de las series no estacionarias se seleccionó modelos no estacionarios, como los que mejor representan la variabilidad de los datos. Para las 28 series no estacionarias restantes, si bien con el estadístico AIC los modelos no estacionarios aparentemente fueron mejores, el estadístico D fue inferior a 3.84 (valor de χ^2_v a un nivel $\alpha = 0.05$), por lo que no se justificó estadísticamente la utilización de un

modelo M_1 , M_2 o M_3 en lugar de un M_0 . Por otro lado, predominan los modelos de tipo GVE_M_1 elegidos para 91 series (Figura 5). La mayoría de éstas pertenece a estaciones localizadas en las regiones 1 y 7 (centro del país), con un total de 42 y 10, respectivamente (Tabla 5).

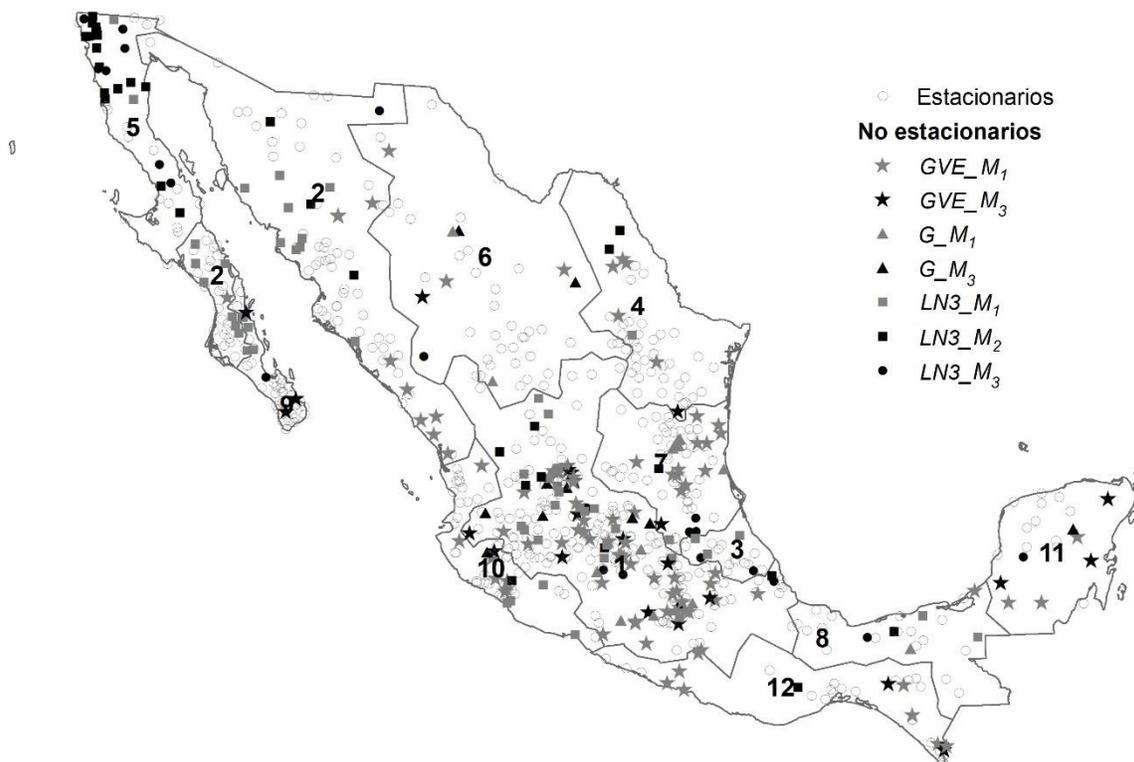


Figura 5. Tipos de modelos seleccionados para la predicción de lluvia máxima en 24 horas.

Tabla 5. Cantidad de series correspondientes a los tipos de modelos probabilísticos seleccionados.

Modelo	Total de estaciones	Región											
		1	2	3	4	5	6	7	8	9	10	11	12
GVE_M ₀	299	94	50	8	23	7	22	29	7	20	20	10	9
GVE_M ₁	91	42	10	0	5	0	3	12	1	0	7	3	8
GVE_M ₃	23	11	0	0	1	0	1	1	0	3	1	3	2
G_M ₀	68	47	3	2	3	0	3	2	1	0	3	1	3
G_M ₁	15	6	0	0	0	0	2	6	1	0	0	0	0
G_M ₃	11	7	0	0	0	0	2	0	0	0	1	1	0
LN3_M ₀	125	28	19	1	8	11	11	9	4	11	14	1	8
LN3_M ₁	55	22	17	3	1	2	0	1	2	3	4	0	0
LN3_M ₂	32	6	3	0	2	16	0	1	1	0	1	0	2
LN3_M ₃	22	4	1	3	0	7	1	3	1	1	0	1	0
GA3_M ₀	20	2	2	1	2	0	0	7	4	0	0	0	2
W_M ₀	8	1	1	0	0	0	0	1	1	3	0	0	1
Total	769	270	106	18	45	43	45	72	23	41	51	20	35

En la región 5 de la península de California se tiene la mayor cantidad de estaciones (23), con modelos LN3_M₂ y LN3_M₃, en los que el parámetro de escala está en función del índice PDO. La selección de modelos del tipo M₂ y M₃ ratifica la asociación entre el índice PDO y las lluvias máximas en la península.

Al comparar la Figura 2 y la Figura 5 se aprecia una correspondencia entre los sitios del país con series no estacionarias y los sitios para los que, a su vez, se han seleccionado modelos no estacionarios, por lo cual

se puede determinar que los resultados bajo el enfoque no estacionario son consistentes con las pruebas de estacionaridad realizadas.

En la Tabla 6 se muestran los resultados de las pruebas de bondad de ajuste y los estimadores de los parámetros de las estaciones Acámbaro y San Vicente, elegidas para mostrar la metodología seguida. Para San Vicente, los estadísticos de bondad de ajuste indicaron que el valor más bajo del AIC es de 517.53 y corresponde al modelo $LN3_M_2$. Además, el estadístico D del modelo M_2 con respecto al M_0 es de 12.68, superior al valor $\chi^2 = 3.84$. Por lo tanto, el modelo M_2 es más adecuado y explica mejor la variación de los datos que el modelo M_0 . En el caso de la estación Acámbaro, tanto el modelo G_M_1 , como el modelo GVE_M_1 tienen valores AIC bajos, iguales a 496.8; sin embargo, el estadístico D para este último es de 14.0 (superior al de la función Gumbel), por lo cual se consideró que el mejor modelo es el GVE_M_1 .

Tabla 6. Resultados de las pruebas de bondad de ajuste y estimadores de los parámetros.

Función	Estación San Vicente							Estación Acámbaro						
	AIC	D	Estimadores de los parámetros					AIC	D	Estimadores de los parámetros				
			$\hat{\mu}_0$	$\hat{\mu}_1$	$\hat{\mu}_2$	\hat{x}_0	$\hat{\sigma}_z$			$\hat{\mu}_0$	$\hat{\mu}_1$	$\hat{\mu}_2$	\hat{x}_0	$\hat{\sigma}_z$
LN3_ M_0	528.2	-	3.111			9.212	0.638	508.8	-	3.360	-		16.57 6	0.427
LN3_ M_1	527.0	3.2	2.942	0.007		7.912	0.581	497.2	13.6	2.963	0.01		17.87	0.404
LN3_ M_2	517.5	12.7	3.275		0.289	6.845	0.513	510.7	0.1	3.356		0.02	16.74	0.429
LN3_ M_3	518.2	10.8	3.182	0.004	0.257	5.889	0.486	498.8	0.4	2.960	0.01	-0.04	17.54	0.397

			$\hat{\lambda}_0$	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\alpha}$	$\hat{\beta}$			$\hat{\lambda}_0$	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\alpha}$	$\hat{\beta}$
GA3_M0	526.5	-	13.513			13.77	1.669	509.4	-	22.716	-		7.403	3.430
GA3_M1	526.4	2.0	16.508	-0.05		16.99	1.263	496.9	14.5	20.753	0.21		7.921	2.578
GA3_M2	526.1	2.4	12.259		3.818	10.61	2.339	509.6	1.8	25.923		1.34	8.857	2.529
GA3_M3	523.6	4.9	19.041	-0.08	1.078	19.15	1.008	497.4	1.5	21.687	0.23	-0.98	9.073	2.073
			$\hat{\nu}_0$	$\hat{\nu}_1$	$\hat{\nu}_2$	$\hat{\alpha}$				$\hat{\nu}_0$	$\hat{\nu}_1$	$\hat{\nu}_2$	$\hat{\alpha}$	
G_M0	531.5	-	28.932	-		12.21		507.4	-	41.935	-		10.33	
G_M1	532.2	1.3	26.123	0.091		12.24		496.8	12.6	34.969	0.23		9.229	
G_M2	525.6	7.9	30.415		5.812	11.8		509.1	0.3	42.061		0.78	10.28	
G_M3	527.5	6.7	29.616	11.83	0.024	5.656		498.6	0.2	34.719	9.21	0.23	-0.570	
			$\hat{\phi}_0$	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\beta}$				$\hat{\phi}_0$	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\beta}$	
W_M0	541.9	-	41.337	-		2.206		528.2	-	53.321	-			3.370
W_M1	534.3	9.6	27.246	0.444		2.432		517.5	12.7	41.074	0.37			3.754
W_M2	523.5	20.4	42.717		10.06	2.680		529.9	0.3	53.113		-1.47		3.389
W_M3	520.7	15.6	36.414	0.188	9.102	2.806		515.4	4.1	37.172	0.47	-4.83		3.950
			$\hat{\nu}_0$	$\hat{\nu}_1$	$\hat{\nu}_2$	$\hat{\alpha}$	$\hat{\beta}$			$\hat{\nu}_0$	$\hat{\nu}_1$	$\hat{\nu}_2$	$\hat{\alpha}$	$\hat{\beta}$
GVE_M0	530.1	-	27.561	-		10.93	0.217	508.8	-	41.525	-		10.04	0.074
GVE_M1	532.0	0.1	27.092	0.020		11.07	0.194	496.8	14.0	34.722	0.22		8.665	0.139
GVE_M2	527.1	5.0	29.754		4.932	11.39	0.077	510.2	0.6	41.576		0.95	9.893	0.092
GVE_M3	531.7	2.3	27.886	0.013	2.932	11.24	0.134	498.2	0.6	34.371	0.23	-1.07	8.636	0.138

Los modelos seleccionados para las series de las estaciones San Vicente y Acámbaro (Figura 6) indican que, para las dos estaciones, los datos de las *worm plot* están dentro de los límites de confianza, y en las gráficas cuantil-cuantil (Q-Q) los datos están cerca de la diagonal unitaria. Por tal razón, los modelos seleccionados se ajustan de forma adecuada con los datos.

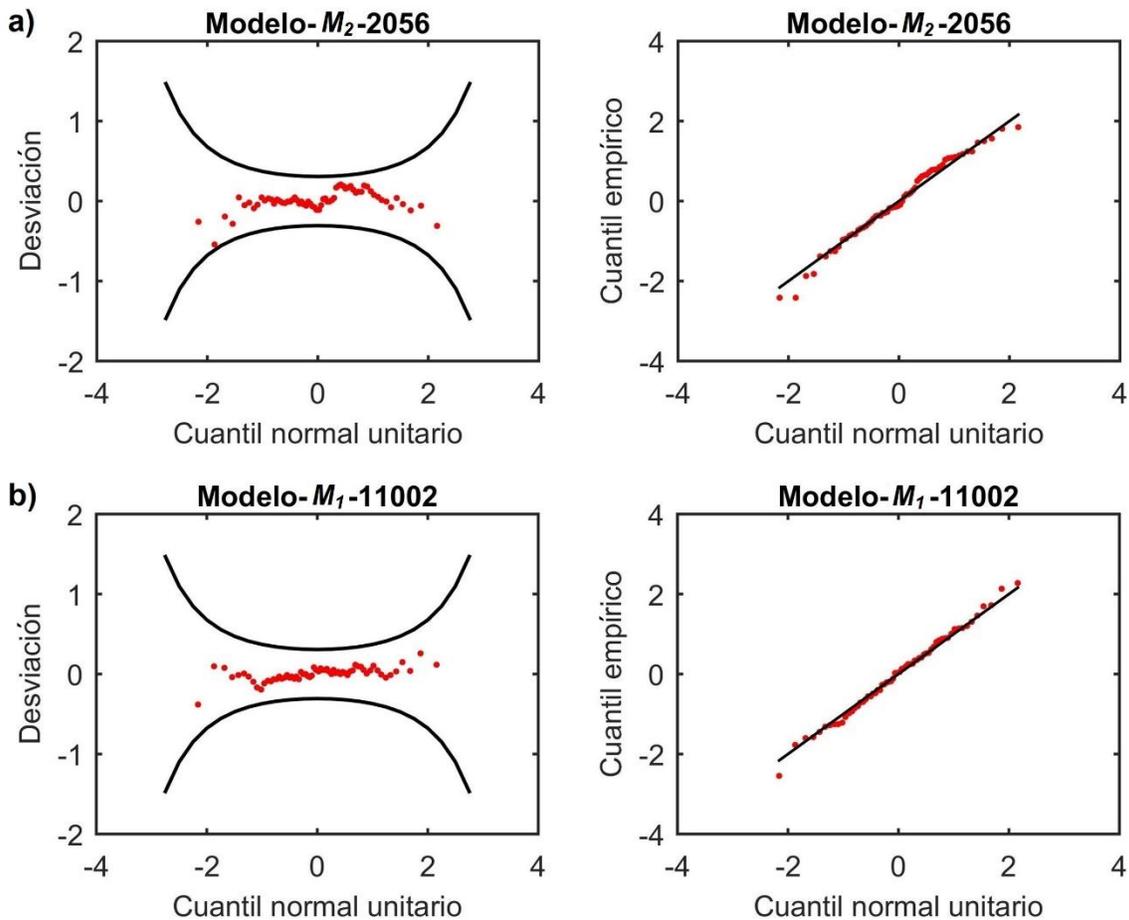


Figura 6. Gráficas *worm plot* (izquierda) y *Q-Q* (derecha) para el análisis visual del ajuste de los modelos seleccionados: a) estación San Vicente (2056), Ensenada, Baja California; b) estación Acámbaro (11002), Acámbaro, Guanajuato.

Los eventos de diseño de la estación San Vicente asociados con diferentes periodos de retorno se estimaron al resolver la Ecuación (10),

Ecuación (11) y Ecuación (12). El valor \widehat{PDO} de la Ecuación (11) se estimó como la media aritmética de los valores PDO mensuales de la fase positiva de esta oscilación, comprendida en el periodo 1976-1998. A pesar de que el modelo M_2 seleccionado no depende del tiempo, se consideró que es un modelo no estacionario. De acuerdo con Villarini y Smith (2010), bajo condiciones de estacionaridad, la distribución de la variable de interés es invariante en el tiempo, sin tendencias, cambios ni periodicidades, por lo que se considera válido asumir que un modelo tipo M_2 es no estacionario debido a la periodicidad intrínseca del PDO incluido como covariable:

$$\hat{X} = \hat{x}_0 + \exp\{F^{-1}(p|\hat{\mu}, \hat{\sigma})\} = 6.845 + \exp\{F^{-1}(p|3.44, 0.513)\} \quad (10)$$

$$\hat{\mu} = \hat{\mu}_0 + \hat{\mu}_2 \widehat{PDO} = 3.275 + 0.289(0.574) = 3.44 \quad (11)$$

$$p = F(x|\hat{\mu}, \hat{\sigma}) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}\left(\frac{z-\hat{\mu}}{\sigma}\right)^2} \quad (12)$$

Donde F^{-1} es la función de distribución de probabilidad log normal inversa, y p es la probabilidad de no excedencia en función de un periodo de retorno T , $p = 1-1/T$.

Los eventos de diseño obtenidos con el modelo convencional son inferiores a los estimados con el modelo no estacionario (Figura 7). La diferencia es mayor en periodos de retorno de 2, 5 y 10 años, en comparación con los periodos de 50 y 100 años. Por ejemplo, los eventos

estimados para un periodo de retorno de 10 años fueron de 60.02 y 67.03 mm, para los casos estacionario y no estacionario, respectivamente. Con lo anterior, se puede determinar que al no considerar la influencia del PDO en la variabilidad de la lluvia en la estación se tendría una subestimación de los eventos de lluvia máxima, principalmente en los de menor periodo de retorno.

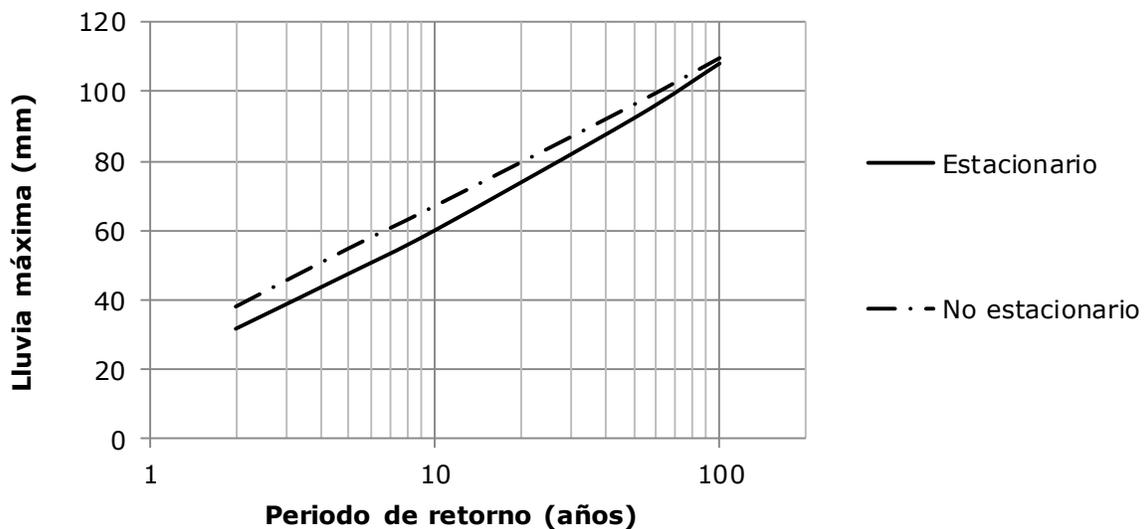


Figura 7. Eventos de diseño de la lluvia anual en 24 horas de la estación San Vicente, Ensenada, Baja California.

Por otro lado, los eventos de la estación Acámbaro se obtuvieron con la Ecuación (13) y la Ecuación (14). Al considerar una dependencia lineal con respecto al tiempo en el parámetro de ubicación de la función GVE, se pudo predecir el incremento de la lluvia máxima de diseño en el futuro. Así, la lluvia máxima de diseño para un periodo de retorno de 10

años será de 73.1 mm en el año 2020; para el año 2050, para el mismo periodo de retorno, se predice una lluvia máxima de 79.6 mm (Figura 8), es decir, la lluvia se incrementará 6.5 unidades en un lapso de 30 años:

$$\hat{X} = \hat{v}(t) - \frac{\hat{\alpha}}{\hat{\beta}} \{1 - [-\ln(p)]^{-\hat{\beta}}\} = \hat{v}(t) - \frac{8.665}{0.139} \{1 - [-\ln(p)]^{-0.139}\} \quad (13)$$

$$\hat{v}(t) = \hat{v}_0 + \hat{v}_1 t_0 = 34.722 + 0.217 t_0 \quad (14)$$

Donde t_0 es el valor del índice del tiempo para un año determinado, $t_0 = 71$ para el año 2020 y $t_0 = 101$ para el año 2050.

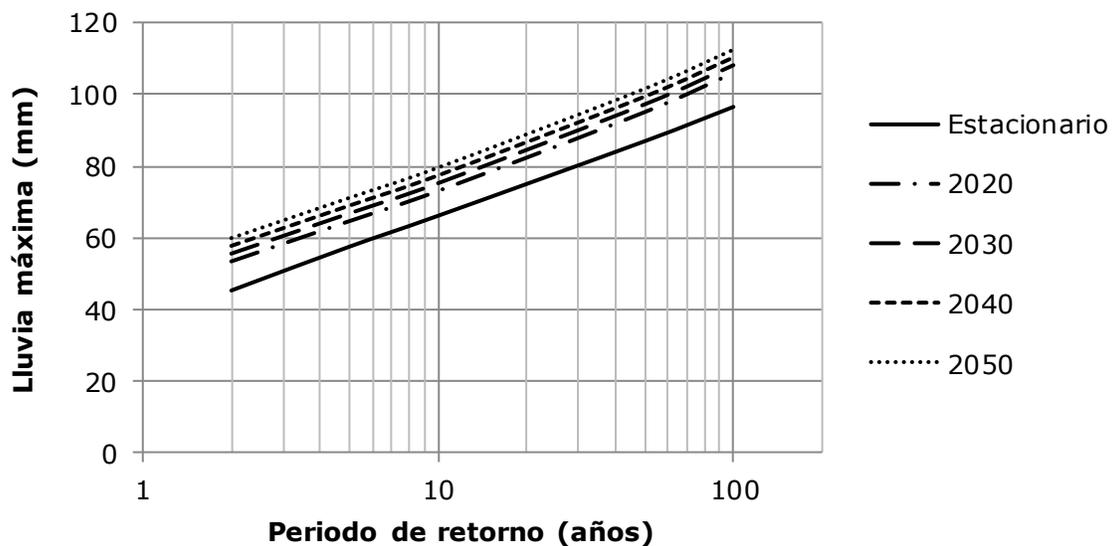


Figura 8. Eventos de diseño de la lluvia anual en 24 horas de la estación Acámbaro, Acámbaro, Guanajuato. Las líneas discontinuas corresponden a los eventos de años futuros.

Conclusiones

Se encontró evidencia de que 22% de las series de lluvia máxima analizadas es no estacionaria. Para 83% de éstas, los modelos no estacionarios seleccionados explican mejor la variabilidad de los datos. Con respecto a las series no estacionarias restantes, el criterio de información de Akaike indicó que los modelos no estacionarios fueron mejores, sin embargo, el estadístico de discordancia no justificó estadísticamente la utilización de un modelo M_1 , M_2 o M_3 en lugar de un M_0 .

Los modelos que incluyen al PDO como covariable son mejores que los estacionarios para el noroeste del país. Por otro lado, la incorporación del tiempo como un índice en los parámetros de las funciones genera mejores ajustes en las series que presentan puntos de cambio y tendencias, de las cuales, en 89% fuer ascendente y creciente, respectivamente. Para estas series, los modelos no estacionarios del tipo

M_1 se ajustan mejor a los datos. Para la estación Acámbaro, los eventos de diseño asociados con diferentes periodos de retorno indican que con el modelo no estacionario GVE_ M_1 se predice un incremento de los valores de la lluvia máxima en el futuro.

El análisis de las series de las estaciones San Vicente y Acámbaro evidenció que, bajo el enfoque clásico estacionario, se subestima la lluvia de diseño y, en consecuencia, su uso en el diseño de alguna obra hidráulica implicaría mayor riesgo de falla.

Referencias

- Akaike, H. (1974). A new look at statistical-model identification. *Automatic Control, IEEE Transactions on*, 19(6), 716-723.
- Álvarez-Olguín, G., & Escalante-Sandoval, C. A. (2017). Modes of variability of annual and seasonal rainfall in Mexico. *JAWRA Journal of the American Water Resources Association*, 53(1), 144-157.
- Álvarez-Olguín, G., & Escalante-Sandoval, C. A. (2016). Análisis de frecuencias no estacionario de series de lluvia anual. *Tecnología y ciencias del agua*, 7(1), 71-88.
- Buuren, S. V., & Fredriks, M. (2001). Worm plot: A simple diagnostic device for modelling growth reference curves. *Statistics in Medicine*, 20(8), 1259-1277.
- Campos-Aranda, D. F. (2016). Modelo probabilístico simple para análisis de frecuencias en registros hidrológicos extremos con tendencia. *Tecnología y ciencias del agua*, 7(3), 171-186.

- Cavazos, T., Turrent, C., & Lettenmaier, D. P. (2008). Extreme precipitation trends associated with tropical cyclones in the core of the North American monsoon. *Geophysical Research Letters*, 35(21).
- Coles, G. S. (2001). *An introduction to statistical modeling of extreme values*. USA: Springer.
- Escalante-Sandoval, C., & Reyes-Chávez, L. (octubre, 2010). *Aplicación de algunas distribuciones mezcladas en el análisis de gastos máximos anuales*. XXI Congreso Nacional de Hidráulica, México.
- Escalante-Sandoval, C., & Reyes-Chávez, L. (2002). *Técnicas estadísticas en hidrología*. México, DF, México: Facultad de Ingeniería, Universidad Nacional Autónoma de México.
- Gleick, P. H. (1989). Climate change, hydrology and water resources. *Reviews of Geophysics*, 27(3), 329-344.
- Held, I. M., & Soden, B. J. (2006). Robust responses of the hydrological cycle to global warming. *Journal of Climate*, 19, 5686-5699.
- Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Tung, C. C., & Liu, H. H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 454(1971), 903-995.
- IPCC, Intergovernmental Panel on Climate Change (2007). *Climate Change 2007: The Physical Science Basis. Contribution of Working*

- Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.)]. Cambridge, United Kingdom and New York, USA: Cambridge University Press.
- Jáuregui, E. (2000). *El clima de la Ciudad de México* (vol. 1). México, DF, México: Universidad Nacional Autónoma de México, Instituto de Geografía y Plaza y Valdés.
- Khaliq, M. N., Ouarda, T. B. M. J., Ondo, J. C., Gachon, P., & Bobée, B. (2006). Frequency analysis of a sequence of dependent and/or non-stationary hydro-meteorological observations: A review. *Journal of Hydrology*, 329, 534-552.
- Katz, R. W., Parlange M. B., & Naveau, P. (2002). Statistics of extremes in hydrology. *Advances in Water Resources*, 25, 1287-1304.
- Kendall, M. G. (1975). *Rank correlation methods* (4^a ed.). London, UK: Charles Griffin.
- Krzyszfowicz, R. (1997). Transformation and normalization of variates with specified distributions. *Journal of Hydrology*, 197, 286-292.
- Magaña, V. O., Vázquez, J. L., Pérez, J. L., & Pérez, J. B. (2003). Impact of El Niño on precipitation in Mexico. *Geofísica Internacional-México*, 42(3), 313-330.
- Mantua, N. J., Hare, S. R., Zhang, Y., Wallace, J. M., & Francis, R. C. (1997). A Pacific interdecadal climate oscillation with impacts on

- salmon production. *Bulletin of the American Meteorological Society*, 78(6), 1069-1079.
- Meehl, G. A., Arblaster, J. M., & Tebaldi, C. (2005). Understanding future patterns of increased precipitation intensity in climate model simulations. *Geophysical Research Letters*, 32(18), 1-4.
- Méndez-González, J., Návar-Cháidez, J. D. J., González-Rodríguez, H., & Treviño-Garza, E. J. (2007). Teleconexiones del fenómeno ENSO a la precipitación mensual en México. *Ciencia UANL*, 10(3), 290-298.
- Méndez-González, J., Ramírez-Leyva, A., Cornejo-Oviedo, E., Zárate-Lupercio, A., & Cavazos-Pérez, T. (2010). Teleconexiones de la Oscilación Decadal del Pacífico (PDO) a la precipitación y temperatura en México. *Investigaciones Geográficas*, 73, 57-70.
- Méndez, M., & Magaña, V. (2009). Regional aspects of prolonged meteorological droughts over Mexico and Central America. *Journal of Climate*, 23(5), 1175-1188.
- Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P., & Stouffer, R. J. (2008). Stationarity is dead: Whiter water management? *Science*, 319, 573-574.
- NOAA, National Oceanic and Atmospheric Administration. (2014). *Climate Indices: Monthly atmospheric and ocean time series*. Recuperado de <http://www.esrl.noaa.gov/psd/data/climateindices/list/index.html>
- Peralta-Hernández, A. R., Balling, R. C., & Barba-Martínez, L. R. (2009). Comparative analysis of indices of extreme rainfall events:

- Variations and tendencies from southern México. *Atmósfera*, 22(2), 219-228.
- Peterson, T. C., Zhang, X., Brunet-India, M., & Vázquez-Aguirre, J. L. (2008). Changes in North American extremes derived from daily weather data. *Journal of Geophysical Research: Atmospheres*, 113(D7), 1-9.
- Pettitt, A. N. (1979). A non-parametric approach to the change-point problem. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 28 (2), 126-135.
- Shepard, D. (1968). A two-dimensional interpolation function for irregularly-spaced data. *ACM '68 Proceedings of the 1968 23rd ACM National Conference*. New York, USA.
- Villarini G., & Smith, J. A. (2010). Flood peak distributions for the Eastern United States. *Water Resources Research*, 46(6), 1-17.
- Voss, R., May W., & Roeckner, E. (2002). Enhanced resolution modeling study on anthropogenic climate change: Changes in extremes of the hydrological cycle. *International Journal of Climatology*, 22, 755-777.
- Wu, Z., & Huang, N. E. (2009). Ensemble empirical mode decomposition: A noise-assisted data analysis method. *Advances in adaptive data analysis*, 1(01), 1-41.
- Yevjevich, V. & Jeng, R.I. (1969). *Properties of non-homogeneous hydrologic time series*. Hydrology papers. Colorado, USA: Colorado State University.