

ICCE: Software para el cálculo de indicadores de cambio climático extremo

ICCE: Software to calculate extreme climate change indicators



<https://cu-id.com/2377/v29n2e06>

 Pedro Roura-Pérez¹,  Vivian Sistachs-Vega²,  Dalia Diaz-Sistachs³

¹ Centro del Clima, Instituto de Meteorología, Cuba.

² Facultad de Matemática y Computación, Universidad de La Habana, Cuba.

³ C2S, Consulting 2 Success LLC, Americas Division, Santo Domingo, República Dominicana.

RESUMEN: Debido a las constantes dificultades que trae el cambio climático para la vida, se hace necesaria la creación de herramientas que propicien un estudio más detallado de los efectos que este tiene. Entre las necesidades más apremiantes se encuentra la de establecer una base común para el estudio del cambio climático. El uso de los indicadores de cambio climático es una de las medidas encaminadas a lograr dicha base común. Estos indicadores muestran el comportamiento de lo que los expertos consideran son los mayores efectos del cambio climático. Sin embargo, los indicadores por si solos no brindan suficiente información, ya que a menudo constituyen grandes volúmenes de datos. Para facilitar su estudio y comprensión se desarrolló un software para la automatización de parte del proceso de estudio de los indicadores. Las herramientas que se brindan son: análisis de tendencia y punto de cambio a través de pruebas no paramétricas, como las dócimas de Kendall-Mann y Pettitt y el análisis del período de retorno a través de la teoría de valores extremos, más específicamente la distribución generalizada de valores extremos. Se describe una metodología enfocada en el uso de estas herramientas a fin de sentar las bases para su posterior uso por los expertos del Instituto de Meteorología.

Palabras claves: Indicadores de cambio climático, tendencia, punto de cambio, períodos de retorno, valores extremos, pruebas no paramétricas, software.

ABSTRACT: Due to the constant difficulties that climate change has in daily living, it is necessary to create tools that provide a detailed study of the effects of these changes. The most important need is to establish a common ground to study climate change. The use of climate change indicators is one of the directions to achieve this common ground. These indicators show the behavior of what is considered by the experts as the major effects of climate change. Although, indicators on itself do not provide enough information since more often than not they constitute high volumes of data. To facilitate the study and comprehension of the indicators, a software was developed to automatize the part of the process of study of the indicators. The tools provided are, trend analysis and point of change through non-parametric tests such as Kendall-Mann and Pettitt tests and the analysis of the return periods through the extreme values theory more specifically the generalized extreme values distribution. It is described a methodology focused on the use of these tools that will accommodate the work of the Institute of Meteorology experts.

Key Words: Climate change indicators, trend, change point, return periods, Extreme Value, non-parametric testing, software.

INTRODUCCIÓN

El cambio climático consiste en la modificación a escala global o local del clima con respecto al historial climático. Ejemplo de estas transformaciones son el aumento de las temperaturas a causa de los gases de efecto invernadero, lo que incrementa el riesgo de sequías e incendios; el aumento de la frecuencia y severidad de las inundaciones, dado que las precipitaciones son mucho más intensas aunque sean

menos frecuentes; el incremento de las temperaturas y las variaciones en la humedad puede ocasionar la proliferación de enfermedades en zonas donde antes no existían; el aumento en la severidad y la frecuencia de las olas de calor pone en riesgo la vida de miles de personas; y por último el aumento del nivel del mar debido al derretimiento de los casquetes polares, lo cual da origen a inundaciones más frecuentes y severas (IPCC, 2007, 2014).

*Autor para correspondencia: pedro.roura@insmet.cu

Recibido: 22/02/2023

Aceptado: 02/05/2023

Ante esta situación se hace necesario para el ser humano el adaptarse al clima cambiante, y parte de esta adaptación consiste en la mejora de las capacidades de predicción, para poder asegurar la seguridad tanto de las personas como de la infraestructura. Con este objetivo la comunidad científica ha llevado a cabo numerosos estudios estadísticos con el fin de brindar la información necesaria. Por ejemplo: para la construcción de una presa se necesita saber el caudal de agua que debe ser capaz de soportar, lo cual se ve afectado por las intensas lluvias e inundaciones; los cultivos necesitan de un determinado rango de temperatura y ciertas características de humedad para poder germinar apropiadamente; para la construcción de parques eólicos es necesario un estudio de las características del viento en la zona; y a fin de evitar pérdidas humanas con los eventos meteorológicos extremos como los huracanes, tifones y tornados se hace necesario ser capaz de predecir el comportamiento e intensidad de los mismos (IPCC,2007).

Para poder sentar las bases para proveer una mirada objetiva y científica al cambio climático la Organización de Naciones Unidas (ONU) creó el Panel Intergubernamental del Cambio Climático (IPCC) en 1988. Este panel realiza una exhaustiva evaluación de los eventos extremos como el incremento de días y noches cálidas, la escalada de las olas de calor y los cambios en la frecuencia e intensidad de sequías y huracanes, entre otro [IPCC, 2014]. Para ello un grupo de expertos formuló un conjunto de 27 indicadores para detectar variaciones en el comportamiento del clima. Estos indicadores tienen como propósito la estimación de los cambios en los extremos a partir de una metodología común, con el fin de proveer una base comparativa global.

El Instituto de Meteorología ha realizado algunos estudios sobre el comportamiento de los indicadores de extremo. Sin embargo, aún no se ha aplicado la teoría de valores extremos a los indicadores climáticos, ni se cuenta con una herramienta computacional que brinde a los expertos un manejo simplificado de los indicadores para así facilitar su estudio. Dada la complejidad del cómputo y el análisis de los indicadores de extremo, normalmente los investigadores deben usar más de un programa, cada uno con tareas específicas, para lograr los resultados deseados. Además, estos softwares no son libres y son de difícil adquisición, por lo que es necesaria una versión propia, desarrollada para el Instituto de Meteorología. Este problema motivó el siguiente trabajo, donde se busca crear una herramienta computacional, centrada en la aplicación de la teoría de valores extremos a los indicadores, que brinde facilidades tanto para el cálculo como para el análisis de los indicadores de extremos climáticos a través de métodos estadísticos. En este trabajo se describe un software y las herramientas empleadas para su desarrollo. Se especifica su funcionamiento y la metodología a utilizar en datos de variables climáticas extremas.

Este software es muy importante para el Instituto de Meteorología (INSMET) y en especial para el Centro del Clima (CENCLIM) para brindar servicios sobre la caracterización de las variables extremas en diferentes esferas como la construcción de obras, la planificación urbana, la agricultura, para la hidrología. Además, para El Estado Mayor de la Defensa Civil, debido a que ellos son los encargados de la toma de decisiones ante la ocurrencia de un evento extremo, previniendo a la población evitando así las pérdidas de vidas humanas y la disminución de las pérdidas materiales.

El artículo se estructuró en 3 epígrafes, el primero trata sobre Indicadores de Cambio Climático extremo, el segundo sobre la Teoría de Valores Extremos y las pruebas estadísticas y el último define la metodología a usar en este tipo de problema.

1. INDICADORES DE CAMBIO CLIMÁTICO EXTREMO

En este epígrafe se exponen los distintos métodos estadísticos usados en el desarrollo del software, así como una breve explicación de la base teórica de los mismos. Se tiene en cuenta el comportamiento de los Indicadores de Extremos Climáticos de series diarias de temperatura máxima, temperatura mínima y precipitación. Estos datos se obtienen de la base de datos del Centro del Clima del Instituto de Meteorología, así como de las series diarias de precipitación de la red pluviométrica del Instituto Nacional de Recursos Hidráulicos. Para lograr un mejor análisis se utilizan algunos fundamentos de la teoría de valores extremos, como son: la Distribución del Máximo, el Teorema de Convergencia a Familias y la Distribución Generalizada de Valores Extremos de Gumbel, Fréchet y Weibull. Para un análisis más profundo de los datos se analiza también la existencia de tendencia, puntos de cambio y los períodos de retorno.

El Panel Intergubernamental de Expertos sobre el Cambio Climático (IPCC) en su cuarto informe realiza una exhaustiva evaluación de los eventos extremos desde el incremento de días y noches cálidas, de las olas de calor, cambios en la frecuencia e intensidad de sequías y huracanes, entre otros (IPCC, 2007). Para ello se definieron indicadores mediante los cuales se pueden identificar o encontrar las evidencias del cambio del clima (IPCC, 2014).

Un Indicador es un procedimiento que permite cuantificar o relaciona un fenómeno con una escala. Sirve para “indicar” o sugerir la existencia de ciertas características del fenómeno en estudio. A menudo es utilizable para registrar los cambios en dicho fenómeno (Bautista *et al.*, 2013).

Los indicadores de cambio climático tienen el objetivo de ser elementos matemáticos que se utilizan para identificar, registrar, conocer y hacer evidente el cambio climático (IPCC, 2007). Posteriormente se relacionan dichos cambios con las respuestas de los

organismos, con la producción agrícola, ganadera y forestal, con aspectos de salud y con los riesgos ambientales (Burgos y González, 2012).

Hay indicadores relacionados con las temperaturas extremas (Tabla 1).

2. TEORÍA DE VALORES EXTREMOS Y PRUEBAS ESTADÍSTICAS

2.1. Teoría de Valores Extremos para el cálculo de Períodos de Retorno

La Teoría de Valores Extremos ha emergido en los últimos 50 años como una de disciplinas estadísticas más importantes para la ciencia aplicada. Sus técnicas se han esparcido a lo largo de muchos campos. Por ejemplo: las empresas de seguros, asesoramiento de riesgo en el mercado financiero y en la predicción de tráfico en las telecomunicaciones. El análisis estadístico de valores extremos es comúnmente empleado para predecir los períodos de retorno de dichos valores extremos. La característica distintiva del análisis de valores extremos es el objetivo de cuantificar el comportamiento estocástico de un proceso que alcanza valores inusualmente altos, o bajos. El análisis de valores extremos usualmente requiere estimar la probabilidad de eventos que son más extremos que cualquiera que ya haya sido observado.

La Teoría de Valores Extremos se centra en el comportamiento estadístico de $M_n = \{X_1, \dots, X_n\}$, donde X_1, \dots, X_n es una secuencia de variables aleatorias independientes con una función de distribución común F . En la práctica los X_i usualmente representan mediciones periódicas de un proceso, a fin de que M_n represente el máximo valor del proceso a lo largo de n unidades de tiempo. Si n es el número de observaciones en un año entonces M_n corresponde al máximo anual (Coles, 2001).

Si bien en la teoría la distribución de M_n puede ser derivada exactamente para cada valor de n como:

$$P(M_n \leq x) = P(X_1 \leq x, \dots, X_n \leq x) = P(X_1 \leq x) \dots P(X_n \leq x) = \{F(x)\}^n \quad (1)$$

En la práctica esto no resulta útil ya que la función de distribución F es desconocida.

El Teorema de Tipos Extremales encapsula el rango de posibles distribuciones límite de

$$M_n^* = \frac{M_n - b_n}{a_n} \quad (2)$$

la cual es una normalización lineal de M_n a fin de evitar la degeneración (Coles, 2001).

2.1.1 Teorema de Tipos Extremales:

Si existe una secuencia de constantes $\{a_n > 0\}$ y $\{b_n\}$ tales que:

$$P\left(\frac{M_n - b_n}{a_n} \leq x\right) \rightarrow G(x) \text{ cuando } n \rightarrow \infty \quad (3)$$

Donde G es una función de distribución no degenerada, entonces G pertenece a una de las siguientes familias:

$$1: G(x) = \exp\left\{-\exp\left[-\left(\frac{x-b}{a}\right)\right]\right\}, \quad -\infty < x < \infty \quad (4)$$

$$2: G(x) = \begin{cases} 0, & x \leq b \\ \exp\left[-\left(\frac{x-b}{a}\right)^{-\alpha}\right], & x > b \end{cases} \quad (5)$$

$$3: G(x) = \begin{cases} \exp\left\{-\left[\left(\frac{x-b}{a}\right)^\alpha\right]\right\}, & x < b \\ 1, & x \geq b \end{cases} \quad (6)$$

Para los parámetros $a > 0, b$ y en los casos de las familias 2 y 3, $\alpha > 0$.

Colectivamente a estas tres clases de distribución se les conoce como distribuciones de valores extremos, con los tipos 1, 2 y 3 siendo conocidos como las familias Gumbel, Fréchet y Weibull. Cada familia tiene un parámetro de localización b y otro de escala a respectivamente, adicionalmente los casos Fréchet y Weibull tienen un parámetro de forma α (Coles, 2001) (Medialdea, 2016).

En las primeras aplicaciones de la teoría de extremo era usual adoptar una de estas tres familias, para luego estimar los parámetros relevantes a esta distribución. Este enfoque tiene dos problemas, primeramente, es necesario un amplio conocimiento de los procesos para poder determinar cuál de las familias se ajusta mejor a los datos en cuestión y, además, una vez que se tome la decisión, cualquier inferencia posterior presume que la elección es correcta y no tiene en cuenta las desviaciones que puedan existir debido a esta elección.

2.1.2 Distribución Generalizada de Valores Extremos

Para corregir los problemas antes mencionados se unifican las tres familias originales de la distribución de valores extremos en una reformulación conocida como Distribución Generalizada de Valores Extremos (GEV por sus siglas en inglés):

$$G(x) = \exp\left\{-\left[1 - \varepsilon\left(\frac{x-\mu}{\sigma}\right)\right]^{1/\varepsilon}\right\} \quad (7)$$

Definida sobre el conjunto $\{x | 1 + \varepsilon(x-\mu)/\sigma > 0\}$, donde $-\infty < \mu < \infty, \sigma > 0$ y $-\infty < \varepsilon < \infty$.

Este modelo tiene tres parámetros: un parámetro de localización μ , uno de escala σ y uno de forma ε . Los casos cuando $\varepsilon > 0$ y $\varepsilon < 0$ corresponden a los casos de Fréchet y Weibull, mientras que si $\varepsilon = 0$ se interpreta como el límite de GEV cuando ε tiende a 0, lo cual lleva a la familia Gumbel:

$$G(x) = \exp\left\{-\exp\left[-\left(\frac{x-\mu}{\sigma}\right)\right]\right\} \quad (8)$$

Esta unificación simplifica la implementación estadística, ya que ε se determina a partir de los propios datos con que se cuenta (Coles, 2001) (Medialdea, 2016).

Tabla 1. Indicadores Cambio Climático

Indicador	Significado
FD0	Días con helada (Frost Days). Sea Tn_i la temperatura mínima en el día i . FD es el número de días cuando $Tn_i < 0^\circ C$
SU25	Días de verano (Summer Days). Sea Tx_i la temperatura máxima en el día i . SU es el número de días cuando $Tx_i > 25^\circ C$
ID0	Días con hielo (Icing Days). Sea Tx_i la temperatura máxima en el día i . ID es el número de días cuando $Tx_i < 0^\circ C$
TR20	Noches tropicales (Tropical Nights). Sea Tn_i la temperatura mínima en el día i . TR es el número de días cuando $Tn_i > 20^\circ C$
GSL	Estación de crecimiento (Growing Season Length). Sea T_i la temperatura media en el día i . Cuento el número de días entre la primera ocurrencia de por lo menos 6 días consecutivos con: $T_i > 5^\circ C$. Y la primera ocurrencia después de 1 julio (1° Enero en el Hemisferio Sur) de por los menos 6 días consecutivos con: $T_i < 5^\circ C$
TXx	Temperatura máxima extrema. Sea Tx_i la temperatura máxima en el día i . Entonces: $TXx = \max(Tx_i)$
TNx	Temperatura mínima más alta. Sea Tn_i la temperatura mínima en el día i . Entonces: $TNx = \max(Tn_i)$
TXn	Temperatura máxima más alta, mínimo valor alcanzado por la temperatura máxima. Sea Tx_i la temperatura máxima en el día i . Entonces: $TXn = \min(Tx_i)$
TNn	Temperatura mínima extrema, mínimo valor alcanzado por la temperatura mínima. Sea Tn_i la temperatura mínima en el día i . Entonces: $TNn = \min(Tn_i)$
Tn10p	Frecuencia de noches frías, porciento de noches que se encuentran por debajo del percentil 10 de temperatura mínima. Sea Tn_i la temperatura mínima en el día i y sea $Tnen10$ el día calendario del percentil 10 centrado en una ventana de 5 días. El índice indica el porcentaje de tiempo, donde: $Tn_i < Tnen10$
Tx10p	Frecuencia de días frescos, porciento de días que se encuentran por debajo del percentil 10 de temperatura máxima. Sea Tx_i la temperatura máxima en el día i y sea $Txen10$ el día calendario del percentil 10 centrado en una ventana de 5 días. El índice indica el porcentaje de tiempo, donde: $Tx_i < Txen10$
Tn90p	Frecuencia de noches cálidas, porciento de noches que se encuentran por encima del percentil 90 de temperatura mínima. Sea Tn_i la temperatura mínima en el día i y sea $Tnen90$ el día calendario del percentil 90 centrado en una ventana de 5 días. El índice indica el porcentaje de tiempo, donde: $Tn_i > Tnen90$
Tx90p	Frecuencia de días cálidos, porciento de días que se encuentran por encima del percentil 90 de temperatura máxima. Sea Tx_i la temperatura máxima en el día i y sea $Txen90$ el día calendario del percentil 90 centrado en una ventana de 5 días. El índice indica el porcentaje de tiempo, donde: $Tx_i > Txen90$
WSDI	Duración de los períodos cálidos (Warm Spells Duration Index), suma de los periodos con al menos 6 días con temperatura superior al 90 percentil. Sea Tx_i la temperatura máxima en el día i y sea $Txen10$ el día calendario del percentil 90 centrado en una venta de 5 días. Entonces el número de días por período es sumado donde, en intervalos de por los menos 6 días consecutivos: $Tx_i > Txen90$
CSDI	Duración de los períodos fríos (Cold Spells Duration Index), suma de los periodos con al menos 6 días con temperatura inferior al 10 percentil. Sea Tn_i la temperatura mínima en el día i y sea $Tnen10$ el día calendario del percentil 10 centrado en una venta de 5 días. Entonces el número de días por período es sumado donde, en intervalos de por los menos 6 días consecutivos: $Tn_i < Tnen10$
DTR	Intervalo diario de temperatura (Diurnal Temperature Range), promedio de la diferencia entre la temperatura máxima y la mínima. Sean Tx_i y Tn_i las temperaturas máximas y mínimas respectivamente en el día i . Si l representa el número días, entonces: $DTR = \frac{\sum_{i=1}^l (Tx_i - Tn_i)}{l}$
DC40	Días con temperatura máxima superior o igual a $40^\circ C$. Sea Tx_i la temperatura máxima en el día i , entonces: $DC40 = \text{cant}(Tx_i \geq 40)$
RX1day	Precipitación máxima en un día, máximo valor alcanzado por las precipitaciones en un día. Sea RR_i el total diario de precipitación en el día i . Entonces: $RX1day = \max(RR_i)$
RX5day	Precipitación máxima en 5 días, máximo de la suma de las precipitaciones agrupadas en períodos de 5 días. Sea RR_i la cantidad de precipitación para el intervalo de cinco días terminando en i . Entonces los valores máximos de 5 días son: $RX5day = \max(RR_i)$
SDII	Índice simple de intensidad diaria (Simple Daily Intensity Index), promedio de las precipitaciones superiores a 1 mm. Sea RR_i la cantidad diaria de precipitación en días húmedos (con precipitaciones superiores a 1mm). Sea l la cantidad de días húmedos, entonces: $SDII = \frac{\sum_{k=1}^l RR_k}{l}$
R10	Días con precipitación mayor a 10mm. Sea RR_i la cantidad de precipitación en el día i . Cuento el número de días donde: $RR_i \geq 10mm$
R20	Días con precipitación mayor a 20mm. Sea RR_i la cantidad de precipitación en el día i . Cuento el número de días donde: $RR_i \geq 20mm$
CDD	Días secos consecutivos (Consecutive Dry Days), mayor número de días consecutivos con precipitaciones inferiores a 1 mm. Sea RR_i la cantidad de precipitación en el día i . Cuento el mayor número de días consecutivos donde: $RR_i \leq 1mm$
CWD	Días húmedos consecutivos (Consecutive Wet Days), mayor número de días consecutivos con precipitaciones superiores a 1 mm. Sea RR_i la cantidad de precipitación en el día i . Cuento el mayor número de días consecutivos donde: $RR_i \geq 1mm$
R95pTOT	Días muy húmedos, suma de las precipitaciones por encima del 95 percentil. Sea RR_i la cantidad diaria de precipitación en días húmedos (con precipitaciones superiores a 1mm) y sea $RRin95$ el percentil 95 de precipitación en los días húmedos en el período 1988-2018. Si l representa el número de días húmedos, entonces: $R95p = \sum_{k=1}^l RR_i$ donde $RR_i > RRin95$
R99p	Días extremadamente húmedos, suma de las precipitaciones por encima del 99 percentil. Sea RR_i la cantidad diaria de precipitación en días húmedos (con precipitaciones superiores a 1mm) y sea $RRin99$ el percentil 99 de precipitación en los días húmedos en el período 1988-2018. Si l representa el número de días húmedos, entonces: $R99p = \sum_{k=1}^l RR_i$ donde $RR_i > RRin99$
PRCPTOT	Precipitación total, suma de todas las precipitaciones. Sea RR_i la cantidad de precipitación en el día i , entonces: $PRCPTOT = \sum_{k=1}^l RR_k$

2.1.3 Estimación de los parámetros π , σ y ε empleando momentos LH

Los L momentos fueron introducidos por Jonathan RM Hosking en 1990 para caracterizar las funciones de distribución. La aplicación de L momentos al análisis de extremos ha sido investigado por muchos investigadores desde sus inicios.

Momentos LH, una generalización de L momentos, son introducidos para caracterizar mejor las secciones superiores de una distribución, así como la ocurrencia de eventos mayores a partir de los datos. Los momentos LH propuestos están basados en combinaciones lineales de estadísticas de orden superior (Wang, 1997).

Los momentos LH se definen como:

$$\lambda_1^\eta = \frac{1}{c_{\eta+1}^n} \sum_{i=1}^n c_{\eta+1}^{i-1} x_i \quad (9)$$

$$\lambda_2^\eta = \frac{1}{2c_{\eta+2}^n} \sum_{i=1}^n (c_{\eta+1}^{i-1} - c_{\eta+1}^{i-1} c_1^{n-1}) x_i \quad (10)$$

$$\lambda_3^\eta = \frac{1}{3c_{\eta+3}^n} \quad (11)$$

$$\sum_{i=1}^n (c_{\eta+2}^{i-1} - 2c_{\eta+1}^{i-1} c_1^{n-1} + c_{\eta+1}^{i-1} c_2^{n-1}) x_i$$

$$\lambda_4^\eta = \frac{1}{4c_{\eta+4}^n} \sum_{i=1}^n (c_{\eta+3}^{i-1} - 3c_{\eta+2}^{i-1} c_1^{n-1} + 3c_{\eta+1}^{i-1} c_2^{n-1} - c_{\eta+1}^{i-1} c_3^{n-1}) x_i \quad (12)$$

Donde $C_j^m = \binom{m}{j} = \frac{m!}{j!(m-j)!}$, η es el orden del momento LH, n es el tamaño de la muestra y x_i es el i -ésimo elemento de la muestra ordenada (Busababodhin et al., 2015).

Luego del cálculo de los momentos LH se pasa a calcular los parámetros de la GEV. Primeramente, se halla:

$$\varepsilon = \alpha_0 + \alpha_1 \left(\frac{\lambda_3^\eta}{\lambda_2^\eta} \right) + \alpha_2 \left(\frac{\lambda_3^\eta}{\lambda_2^\eta} \right)^2 + \alpha_3 \left(\frac{\lambda_3^\eta}{\lambda_2^\eta} \right)^3 \quad \text{a partir}$$

de (Tabla 2): (13)

Luego se hallan las estimaciones de σ y μ despejando de λ_2^η y λ_1^η respectivamente según $\varepsilon \neq 0$ o $\varepsilon = 0$:

Para $\varepsilon \neq 0$:

$$\lambda_1^\eta = \mu + \frac{\sigma}{\varepsilon} [1 - \Gamma(1 + \varepsilon)(\eta + 1)^{-\varepsilon}] \quad (14)$$

$$\lambda_2^\eta = \frac{\sigma(\eta + 2)\Gamma(1 + \varepsilon)}{2\varepsilon} [-(\eta + 2)^{-\varepsilon} + (\eta + 1)^{-\varepsilon}] \quad (15)$$

Para $\varepsilon = 0$:

$$\lambda_1^\eta = \mu + \sigma[e + \ln(\eta + 1)] \quad (16)$$

$$\lambda_2^\eta = \frac{(\eta + 2)\sigma}{2} [\ln(\eta + 2) - \ln(\eta + 1)] \quad (17)$$

Donde Γ es la función Gamma y $e \approx 0.5772$ es la constante de Euler (Wang, 1997).

2.1.4 Períodos de Retorno

En los estudios meteorológicos tienen mucha importancia los valores críticos o valores de retorno, ya que permiten prepararse con antelación a eventos de gran magnitud. Son considerados como los cuantiles de la distribución de valores extremos. Si el período de retorno T_p es el cuantil de orden p de una variable, entonces p es la probabilidad de que T_p sea superado una vez al año; y el período de retorno $\frac{1}{p}$, es el número de unidades de tiempo que transcurrirán en media entre dos veces en los que la variable supere el valor de T_p (Cooley, 2011).

Sea X una variable aleatoria continua con función de distribución (x), la probabilidad prefijada $p = p(X > x_0) = 0.01 = 1\%$, equivale a un período de retorno de $\frac{1}{0.01} = 100$ años, o sea, para 100 años podemos esperar que la variable sea igual o mayor que x_0 . Por tanto solo se debe calcular, a partir de la expresión de (x), donde los parámetros han sido estimados previamente, la abscisa crítica x que corresponde a la probabilidad o período de retorno prefijado a través de la función inversa de $F(x)$ (Roura et al., 2019).

$$x(F) = \mu + \frac{\sigma}{\varepsilon} [1 - (-\ln F)^\varepsilon] \quad \text{para } \varepsilon \neq 0 \quad (18)$$

$$x(F) = \mu - \sigma \ln(-\ln F) \quad \text{para } \varepsilon = 0 \quad (19)$$

2.2 Análisis de Tendencia: Dócima de Kendall-Mann

El análisis de tendencia de una variable meteorológica es un procedimiento estadístico que, mediante pruebas de hipótesis, explica el cambio de la variable, arroja la tasa de cambio respecto al tiempo, ayuda a identificar regiones y épocas vulnerables a dichos cambios; y ayuda a conocer si el clima está cambiando o se mantiene estable.

Generalmente se utilizan dos métodos no paramétricos para demostrar la posible existencia de tendencia: uno está basado en el coeficiente de correlación de Spearman y el otro es la dócima de Kendall-Mann.

Kendall-Mann es un método estadístico basado en el ranking (número de orden) de dos variables, amplia-

Tabla 2. Valores de coeficientes

η	a_0	a_1	a_2	a_3
0	0.2849	-1.8213	0.8140	-0.2835
1	0.4823	-2.1494	0.7269	-0.2103
2	0.5914	-2.3351	0.6442	-0.1616
3	0.6618	-2.4548	0.5733	-0.1273
4	0.7113	-2.5383	0.5142	-0.1027

mente usado en estudios de identificación de tendencias en series de variables meteorológicas en cuanto a si son ascendentes o descendentes, pero no cuantifican estas. Además de tendencias detecta saltos. Como antecedente, se puede señalar que ha sido recomendada por la Organización Meteorológica Mundial para su aplicación en series de variables meteorológicas e hidrológicas.

Las observaciones originales x_i , son reemplazadas por el rango m_i que está dado por la magnitud de la posición que ocupan los valores originales en la serie ordenada de forma creciente. Por cada elemento m_i , se calcula el número de elementos n_i que lo preceden ($i > j$) y que cumplan la condición $m_i > m_j$ (Sneyers, 1990) (Roura et al., 2019).

Luego se calculan:

$$t_i = \sum_{j=1}^i n_j \quad (20)$$

Y se calculan su valor esperado y varianza:

$$E(t_i) = \frac{i(i-1)}{4} \text{ y } V(t_i) = \frac{i(i-1)(2i+5)}{72} \quad (21)$$

Con esto se obtiene:

$$KM_i = \begin{cases} \frac{t_i - E(t_i)}{\sqrt{V(t_i)}} & \text{si } i = 2, 3, \dots, n \\ 0 & \text{si } i = 1 \end{cases} \quad (22)$$

La tendencia puede ser creciente o decreciente según KM_i sea positivo o negativo. Para $i = n$ se tiene la tendencia global, con todos los datos de la serie en cuestión.

Determinándose así la región crítica, que sería:

$$\alpha_1 = P(|Z| > |KM|) \quad (23)$$

Si $\alpha_1 > \alpha_0$ se acepta H_0 (No hay tendencia significativa).

Si $\alpha_1 < \alpha_0$ se rechaza H_0 (Hay tendencia significativa).

La representación gráfica simultanea de los valores de KM_i de las series directa y retrograda en la forma de curvas poligonales se utiliza para ubicar aproximadamente un punto de cambio, que se corresponde cronológicamente con el punto de intersección de ambos gráficos, bajo la existencia de una tendencia global significativa.

La serie retrograda y_i' asociada de datos x_i es la serie original invertida, donde. Se procede en forma análoga al caso anterior, pero con ordenación decreciente para el cálculo del rango (Sneyers, 1990) (Roura et al., 2019).

2.3 Dócima de Pettitt para Puntos de Cambio

Dócima de Pettitt es la más común para la detección de puntos de cambio simples en variables con datos continuos. En este caso la hipótesis nula plantea que la serie sigue una o más distribuciones que tienen el mismo parámetro o sea que no existe cambio, contra la hipótesis alternativa que plantea la existencia de un

punto de cambio (Roura et al., 2019) (Pettitt, 1978) (Pohlert, 2020).

El test estadístico no paramétrico se define como:

$$K_T = \max |U_{t,T}| \quad (24)$$

Donde:

$$U_{t,T} = \sum_{i=1}^t \sum_{j=t+1}^T \text{sgn}(x_i - x_j) \quad (25)$$

En caso de existir punto de cambio, este se encuentra en K_T . La probabilidad de significación de K_T está dado por α_1 , que se calcula como:

$$\alpha_1 \cong 2e^{\left(\frac{-6K_T^2}{n^3 + n^2}\right)} \quad (26)$$

Si $\alpha_1 > \alpha_0$ se acepta H_0 (No hay punto de cambio).

Si $\alpha_1 < \alpha_0$ se rechaza H_0 (Hay punto de cambio).

3. METODOLOGÍA

En este capítulo se explica la metodología seguida durante la implementación del software para el cálculo de indicadores de cambio climático extremo (ICCE), la cual consiste en realizar un análisis de tendencia, punto de cambio y período de retorno para el análisis tanto de las variables meteorológicas como de los indicadores de extremos climáticos. Para el análisis de tendencia se emplea la dócima de Kendall-Mann, para la determinación de la existencia de punto de cambio se usa la dócima de Pettitt y para el cálculo de períodos de retorno se utiliza la teoría de valores extremos. Al final del capítulo se discuten los requerimientos computacionales y los detalles de implementación.

Lo novedosos de la metodología implementada en este trabajo yace en que es la primera vez que se aplica la teoría de extremos para el cálculo del período de retorno de los indicadores de extremos climáticos en un estudio del INSMET. Es primera vez también que se emplea la Distribución Generalizada de Valores Extremos (GEV) para realizar dicho análisis, estudios anteriores llevaban a cabo un proceso para seleccionar el modelo que mejor se adapta a los datos, lo cual acarrea posibles errores.

En la concepción de la metodología se consultó a los expertos para facilitarles el uso y comprensión del software tanto en la entrada de los datos recogidos por el INSMET como en el análisis y visualización de los resultados. A fin de ayudar a la comprensión de los resultados obtenidos, se brindan métodos estadísticos orientados al análisis de las propiedades de los mismos. La metodología propuesta se ilustra en la figura 1.

Como se aprecia en la metodología implementada se calculan los indicadores de extremos a partir de los datos suministrados. Además, se da la posibilidad de calcular un grupo de estadígrafos tanto para los datos originales como para los estadígrafos. El análisis de existencia de tendencia y puntos de cambio, así como el cálculo de los períodos de retorno se puede realizar también a los datos originales y a los indicadores.

Finalmente, toda esta información, excepto los estadígrafos, puede ser graficada para facilitar su comprensión.

PASO 1. Entrada de datos

Para el cálculo de los indicadores de extremo el software utiliza datos asociados a la temperatura, como son las temperaturas máximas, mínimas y medias de cada período; y otros asociados a las precipitaciones como son la cantidad de precipitaciones en el período, la lluvia máxima en 24 horas y la cantidad de días con lluvia. También se recogen la fecha de la medición y el identificador de la estación de la cual se toman las muestras. La entrada de los datos al programa ocurre a través de ficheros “.csv” que contienen los datos recogidos por el INSMET a lo largo de los años. Estos ficheros contienen la información en formato diario, mensual y anual (Figura 2).

Aunque el INSMET garantiza la completitud y veracidad de los datos suministrados, se decidió verificar que no existan faltantes en los valores de entrada, pues estos deben ser continuos para poder realizar correctamente el cálculo de los indicadores de extremo. En caso de que sea necesario se da una advertencia al usuario de que un dato debe ser rellenado. Para así garantizar la robustez y buen funcionamiento del software.

PASO 2. Cálculo de los indicadores

Una vez se han cargado los datos se pasa a realizar el cálculo de los indicadores de extremo. Para ello el usuario debe seleccionar los indicadores que desea calcular y el período de tiempo en el cual desea calcularlos, mensual o anual. Cabe destacar que este proceso no requiere que el usuario especifique nada respecto a los datos, pues los ficheros contienen toda la información necesaria y el software la selecciona automáticamente. Por ejemplo, para el indicador de noches tropicales (TR), que son aquellas en que la

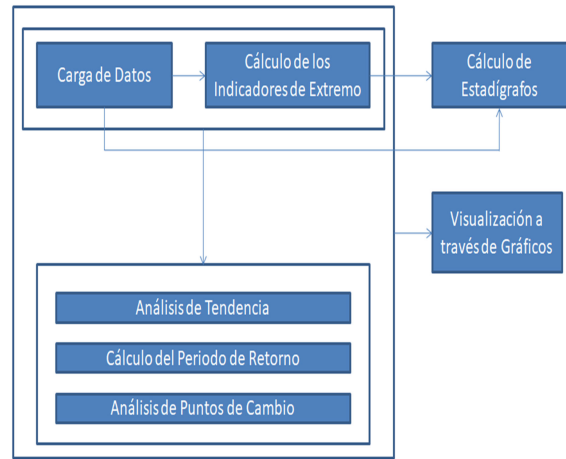


Figura 1. Diagrama de flujo seguida en el software ICCE.exe

temperatura mínima es superior a los 20 grados, el programa extrae de los datos la columna de temperaturas mínimas y agrupa los datos según el período en el que el usuario eligió (Figura 3).

Al terminar el proceso de cálculo los resultados se muestran agrupados en una tabla ordenada cronológicamente, donde cada celda contiene el valor del indicador para el mes o año según se halla especificado (Figura 4).

PASO 3. Cálculo de estadígrafos

Se le da al usuario la posibilidad de calcular diversos estadígrafos, como lo son la media, la mediana, la varianza, la desviación típica y el coeficiente de variación, entre otros. Esto se hace para brindar al usuario una idea del comportamiento que tienen los datos e indicadores, para facilitar el entendimiento del comportamiento de la información. Para calcular dichos estadígrafos el usuario solo debe seleccionar las variables que desea (Figura 5).

Como resultado se muestra una tabla con los estadígrafos calculados (Figura 6).

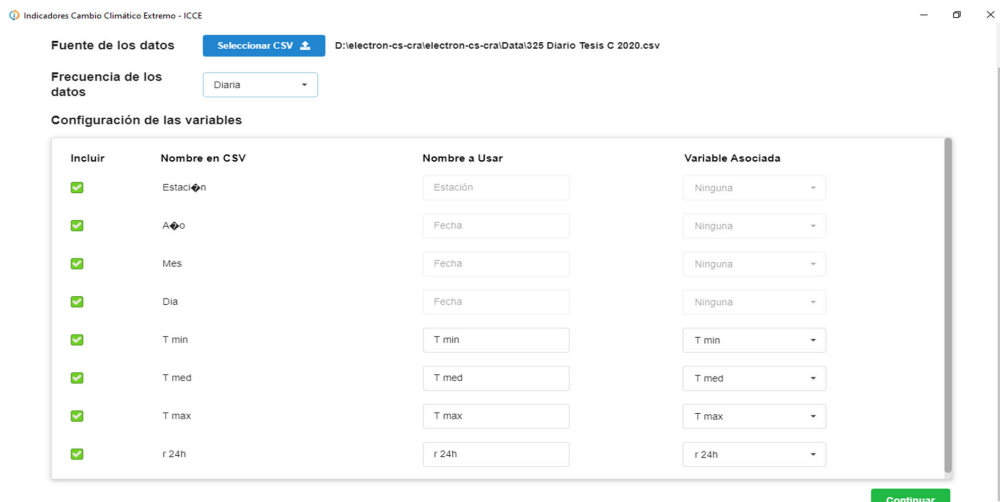


Figura 2. Selección y configuración del fichero

Indicadores de Cambio Climático

Periodo: Indicadores:

Figura 3. Selección de Indicadores

Indicadores Cambio Climático Extremo - ICCE

Herramientas **Indicadores Anuales**

Mostrando 10 filas

Fecha	SU25	TR20
1988	320	167
1989	343	190
1990	359	186
1991	326	231
1992	325	179
1993	325	181
1994	335	198
1995	325	209
1996	323	198
1997	338	234

Filas [1 - 10] / 31

Figura 4. Vista de los indicadores calculados

Estadígrafos

Variables:

Figura 5. Selección de variables

Indicadores Cambio Climático Extremo - ICCE

Estadígrafos

Mostrando 10 filas

Estadígrafos	T min	T max
Mínimo	2.40	13.90
Máximo	27.00	38.20
Media	19.64	29.52
Moda	22.00	32.00
Mediana	20.60	30.20
Varianza Muestral	14.20	9.34
Varianza Poblacional	14.20	9.34
Desviación Típica	3.77	3.06
Desviación Poblacional	3.77	3.06
Coefficiente de Variación Muestral	19.19	10.36

Filas [1 - 10] / 18

Figura 6. Estadígrafos de las variables seleccionadas

PASO 4. Análisis de tendencia

El análisis de tendencia se emplea para determinar si el incremento en los valores de una variable aleatoria es significativo o no. Por ejemplo, si los valores de una variable aleatoria están distribuidos a lo largo de una recta, entonces el análisis de tendencia se basa en si la pendiente de dicha recta se considera pronunciada o no. Por supuesto, para casos no lineales no es tan

sencillo determinar la velocidad de incremento, por lo que se emplean métodos más avanzados, basados en pruebas de hipótesis, para determinar si existe tendencia.

Para realizar un análisis de tendencia el usuario debe especificar qué variable o indicador desea analizar, qué método se debe emplear y qué nivel de significación debe ser usado en las pruebas de hipótesis (Figura 7).

Como resultado se muestran los valores calculados durante la ejecución del programa, para que el usuario pueda verificar la ejecución de los mismos, así como la conclusión final de si el indicador seleccionado presenta tendencia o no (Figura 8).

PASO 5. Análisis de punto de cambio

El análisis de puntos de cambio trata de identificar momentos en los cuales la función de distribución de una variable aleatoria, como los indicadores de extremo, cambia. En general el problema yace en detectar si han ocurrido cambios o no y de ser posible identificar el momento en que ocurrió. El principal uso de este análisis es la detección de comportamientos anómalos en las variables aleatorias. Uno de los métodos más usados para la detección de los puntos de cambio es la Dócima de Pettitt.

Al igual que con el análisis de tendencia, para realizar el análisis de punto de cambio el usuario debe especificar qué variable o indicador desea analizar, qué método se debe emplear y qué nivel de significación debe ser usado en las pruebas de hipótesis (Figura 9).

Como resultado se le muestra al usuario los valores calculados, si existe punto de cambio o no, y en caso de que exista se muestra el año en el que se alcanza. Nótese que este análisis se debe realizar a nivel anual debido a que existen marcadas diferencias en los valores de las temperaturas y precipitaciones entre las distintas estaciones del año (Figura 10).

PASO 6. Cálculo de períodos de retorno

Como se explicó anteriormente, el período de retorno es un valor asociado a la probabilidad de ocurrencia de un suceso. Dado un nivel de retorno, que repre-

Seleccione los parámetros para el análisis de tendencia

Método: Variable: Alpha:

Figura 7. Configuración del análisis de tendencia

Indicadores Cambio Climático Extremo - ICCE

Tend SU25 (KM, 0.05)

Tendencia SU25 (Kendall-Mann, 0.05)

Serie Directa

Fecha	m(i)	n1(i)	n2(i)	t(i)	ValorEsperado(t)	Varianza(t)	KM(i)
1988	3.00	0.00	0.00	0.00	0.00	0.00	0.00
1989	28.00	1.00	0.00	1.00	0.50	0.25	1.00
1990	31.00	2.00	0.00	3.00	1.50	0.92	1.57
1991	11.50	1.00	0.00	4.00	3.00	2.17	0.68
1992	8.00	1.00	0.00	5.00	5.00	4.17	0.00
1993	8.00	1.00	0.50	6.50	7.50	7.08	-0.38
1994	23.50	4.00	0.00	10.50	10.50	11.08	0.00
1995	8.00	1.00	1.00	12.50	14.00	16.33	-0.37
1996	4.50	1.00	0.00	13.50	18.00	23.00	-0.94
1997	25.00	7.00	0.00	20.50	22.50	31.25	-0.36

Filas [1 - 10] / 31

No hay tendencia en los datos (0.27 > 0.05)

Figura 8. Resultado del análisis de tendencia (Serie directa)

Seleccione los parámetros para el análisis de punto de cambio

Método: Variable: Alpha:

Figura 9. Configuración del análisis de punto de cambio

senta el umbral a partir del cual una ocurrencia del suceso se considera relevante, el período de retorno es el tiempo promedio entre dos ocurrencias del suceso que sobrepasan el nivel de retorno.

Al igual que los análisis de tendencia y punto de cambio el usuario debe seleccionar qué variable o indicador desea analizar y especificar si desea realizar un análisis de máximos o mínimos (Figura 11).

Como resultado el usuario obtiene los niveles de retorno para los períodos de 100, 50, 20 y 10 años (Figura 12), así como los parámetros de la Distribución Generalizada, y una tabla en la que se muestran los valores de las funciones de densidad y distribución (Figura 13).

PASO 7. Visualización de la información a través de gráficos

Con excepción del cálculo de estadígrafos, todas las tablas anteriores pueden ser visualizadas en forma de gráficos para facilitar su comprensión. Para ello el

usuario solo debe elegir la ventana de gráfico a partir de un botón en la parte superior derecha (Figura 14).

CONCLUSIONES

Se diseñó e implementó una metodología para el análisis de los indicadores de extremo climático. El cálculo de los indicadores puede mostrar un marcado ascenso o descenso en las variables de estudio. El análisis de tendencia puede explicar el cambio de la variable ayudando a conocer si el clima está cambiando o se mantiene estable, se utilizan dos métodos no paramétricos para demostrar la posible existencia de tendencia: uno está basado en el coeficiente de correlación de Spearman y el otro es la dócima de Kendall-Mann. Para la detección de puntos de cambio simples en variables con datos continuos lo más común es usar la Dócima de Pettitt.

Se confeccionó un software (ICCE) el cual resulta ser una herramienta muy útil para el INSMET para realizar servicios para la caracterización de una serie

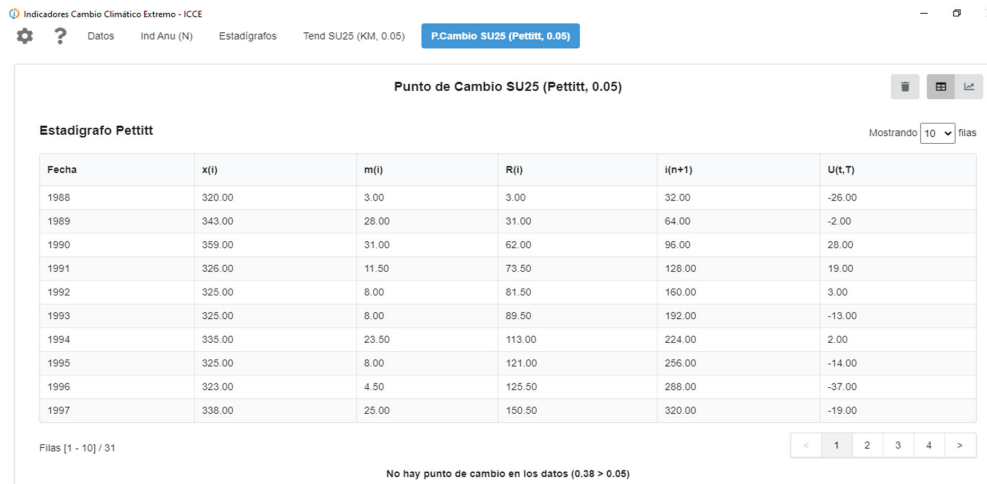


Figura 10. Resultado del análisis de punto de cambio

Seleccione los parámetros para el análisis de período de retorno

Figura 11. Configuración del análisis de periodos de retorno



Figura 12. Resultado del análisis de periodos de retorno (Niveles de Retorno)

Gumbel => Parámetro de localización: 325.70 | Parámetro de escala: 8.38 | Parámetro de forma: 0.03

Función de densidad y Función de distribución

SU25	Función de Densidad	Función de Distribución
295.00	0.00	0.00
319.00	0.03	0.11
320.00	0.03	0.14
323.00	0.04	0.25
323.00	0.04	0.25
324.00	0.04	0.29
325.00	0.04	0.34
325.00	0.04	0.34
325.00	0.04	0.34
326.00	0.04	0.38

Filas [1 - 10] / 31

Figura 13. Resultado del análisis de períodos de retorno (GEV)

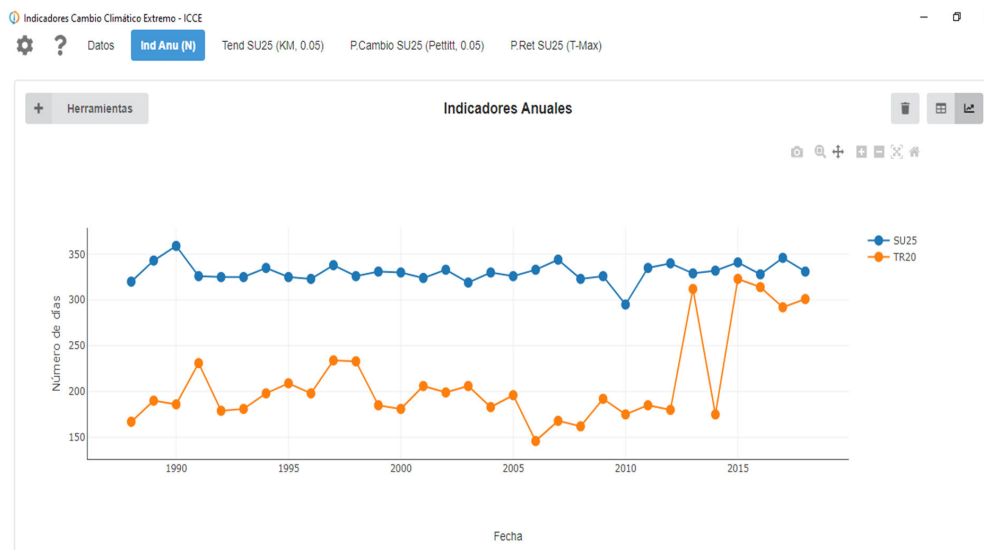


Figura 14. Indicadores SU25 y TR20, frecuencia anual, Casablanca (1988-2018)

de datos de variables asociadas a eventos climatológicos extremos obteniendo beneficios económicos.

Se recomienda además utilizar el software para estudios de tendencia, punto de cambio y períodos de retorno de otras variables climatológicas que puedan ser de interés. También utilizar el software en todos los Centros Meteorológicos Provinciales (CMP) del país y ampliar en la medida de lo posible la longitud temporal de la serie de datos disponibles, para obtener un mejor resultado.

AGRADECIMIENTOS

A Jorge Cristhian Rey Gil, en su momento perteneciente a la Facultad de Matemática y Computación, Universidad de la Habana el cual defendió la Tesis de Diploma de la Especialidad Ciencia de la Computación titulada ICCE: Software para el cálculo de indicadores de cambio climático extremo. Este estudiante se desempeñó con muy buenos resultados en el momento de las investigaciones en los diferentes temas relacionados con la programación e implementación de este software.

REFERENCIAS

- Bautista, F., Pacheco, A., y Ayala, G. (2013). Indicadores de cambio climático con datos diarios (ICC).
- Burgos, Y., y González, I. (2012). Análisis de indicadores de extremos climáticos en la isla de Cuba.
- Busababodhin, P., Seo, Y.A., Park, J.S., y Kumphon, B. (2015). LH-moment estimation of Wakeby distribution with hydrological applications.
- Coles, S. (2001). An Introduction to Statistical Modeling of Extreme Values.
- Cooley, D. (2011). Return Periods and Return Levels Under Climate Change.
- IPCC. (2007). Cambio climático 2007. Contribución de los Grupos de trabajo I, II y III al Cuarto Informe de evaluación del Grupo Intergubernamental de Expertos sobre el Cambio Climático. Informe de síntesis. Ginebra, Suiza. IPCC: ISBN 92-9169-322-7.
- IPCC. (2014). Climate Change 2014. Contribución de los Grupos de trabajo I, II y III al Quinto Informe de evaluación del Grupo Intergubernamental de

- Expertos sobre el Cambio Climático. Informe de síntesis. Ginebra, Suiza.
- Medialdea, A. (2016). Análisis de Valores Extremos - Modelización Espacial.
- Pettitt, A.N. (1978). A Non-parametric Approach to the Change-point Problem.
- Pohlert, T. (2020). Non-Parametric Trend Tests and Change-Point Detection.
- Roura, P., Sistachs V., Rodríguez, J.A., y Vega, R. (2019). SEVECLIM: Software para Analizar Modelos Probabilísticos para Caracterizar Variables Climáticas Extremas. Revista Investigación Operacional.
- Roura, P., Sistachs, V., y Arenas, J.C. (2019). TrendSoft: Software para el análisis de tendencia y puntos de cambios de variables climatológicas.
- Ruíz, O., Espejel, D., Ontiveros, R.E., Enciso, J.M., Galindo, M. A., Quesada, M. L., Grageda, J., Ramos, R., y Ruíz, J.A. (2016). Tendencia de temperaturas máximas y mínimas mensuales en Aguascalientes, México. Revista Mexicana de Ciencias Agrícolas, Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias, Estado de México, México.
- Sneyers, R. (1990). On the statistics analysis of series of observations. LA nota técnica 143.
- Wang, Q. J. (1997). LH moments for statistical analysis of extreme events.

Conflicto de intereses: Los autores declaran no tener conflictos de intereses.

Contribución de autoría: Concepción de la idea: P. Roura. Obtención de datos: P. Roura. Análisis eventos extremos : P. Roura. Análisis estadístico: P. Roura, V. Sistachs. Programación e implementación: D. Sistachs. Creación de la metodología: P. Roura, V. Sistachs, D. Sistachs. Análisis de los resultados obtenidos: P. Roura, V. Sistachs, D. Sistachs. Revisión: P. Roura, V. Sistachs, D. Sistachs. Edición: P. Roura.

Este artículo se encuentra bajo licencia [Creative Commons Reconocimiento-NoComercial 4.0 Internacional \(CC BY-NC 4.0\)](https://creativecommons.org/licenses/by-nc/4.0/)