

**SOCIEDAD CHILENA DE INGENIERÍA HIDRÁULICA**  
**XXVII CONGRESO CHILENO DE INGENIERÍA HIDRÁULICA**

**UN MARCO METODOLÓGICO PARA LA EVALUACIÓN DE MÉTODOS DE  
DESAGREGACIÓN DE PRECIPITACIÓN PARA LA REPRESENTACIÓN DE  
LLUVIA EXTREMA**

**CLAUDIO SANDOVAL V.<sup>1</sup>**  
**JORGE GIRONÁS L.<sup>2</sup>**  
**CRISTIÁN CHADWICK I.<sup>3</sup>**

**RESUMEN**

Los datos de precipitación de alta resolución temporal son esenciales para analizar eventos de lluvia extrema, que se requieren para la modelación hidrológica y el diseño de infraestructura. Frente a la escasez de registros de alta resolución, los métodos de desagregación temporal resultan una herramienta clave. Sin embargo, hasta ahora no existe un marco metodológico que permita comparar métodos de desagregación de manera integral en su capacidad para representar extremos. Este trabajo introduce un marco metodológico innovador para evaluar métodos de desagregación de precipitaciones de escala diaria a horaria, enfocado en su desempeño frente a eventos extremos. El marco integra tanto el análisis de curvas Intensidad-Duración-Frecuencia (IDF) como de índices extremos de precipitación (ERIs), empleando métricas que consideran la exactitud y precisión de las curvas IDF y la variabilidad y el sesgo en los ERIs a través de un índice de eficiencia Kling-Gupta modificado. El marco propuesto abarca múltiples dimensiones del desempeño, facilitando la identificación de fortalezas y limitaciones específicas de cada método según el contexto. El marco se prueba en cinco sitios con tres métodos de desagregación (SOC, k-NN y HUFF). Los resultados muestran que las metodologías de desagregación presentan fortalezas diferenciadas según el tipo de indicador analizado, evidenciando un desempeño altamente dependiente del contexto local y del objetivo del estudio. El marco propuesto permite visualizar de manera integrada estas variaciones y orientar la selección de la metodología de desagregación más adecuada según las condiciones y requerimientos específicos de cada sitio.

---

<sup>1</sup> Ingeniero Civil Hidráulico, M.Sc. - email: [claudio.sandoval@uc.cl](mailto:claudio.sandoval@uc.cl)

<sup>2</sup> Profesor Asociado, Departamento de Ingeniería Hidráulica y Ambiental, Pontificia Universidad Católica de Chile - email: [jgironas@uc.cl](mailto:jgironas@uc.cl)

<sup>3</sup> Profesor Asociado, Facultad de Ingeniería y Ciencias, Universidad Adolfo Ibáñez - email: [cristian.chadwick@uai.cl](mailto:cristian.chadwick@uai.cl)

## 1. INTRODUCCIÓN

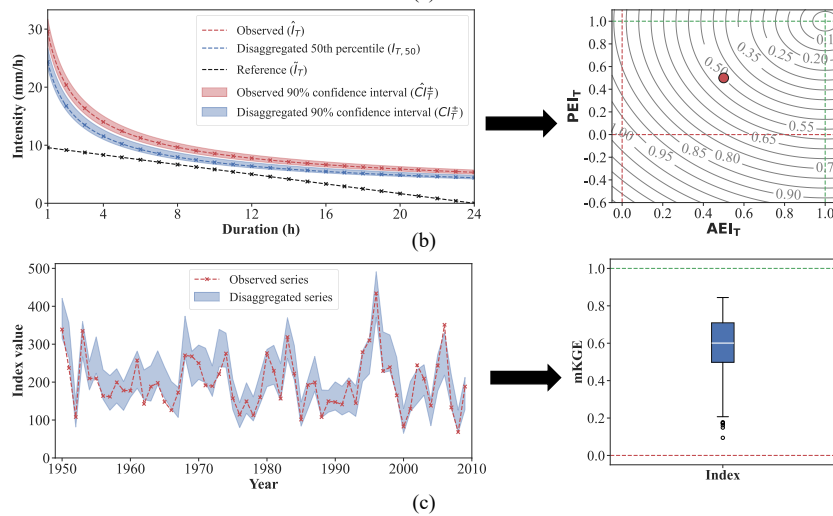
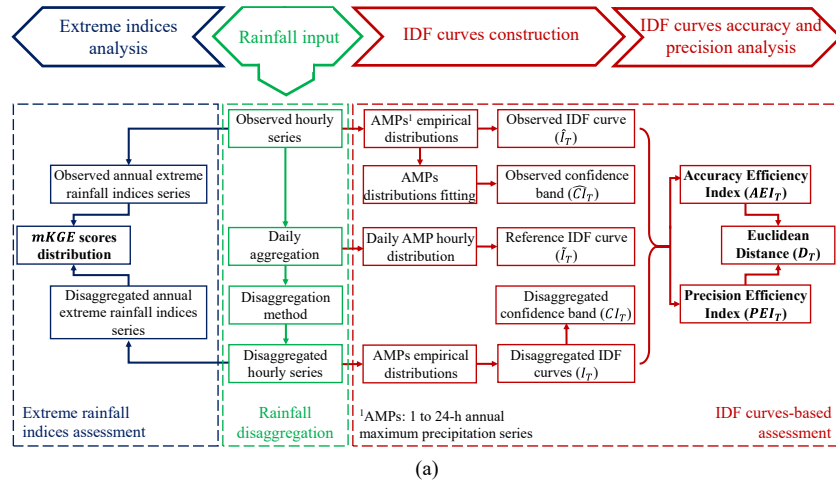
Los datos de precipitación de alta resolución temporal son clave para analizar eventos de lluvia extrema, cruciales para la modelación hidrológica y el diseño de infraestructura (Bruni et al., 2015; Ochoa-Rodríguez et al., 2015). Pese a su importancia, su disponibilidad es limitada, ya que registros subdiarios suelen estar restringidos a pocas estaciones (Woldemeskel et al., 2016). Así, los métodos de desagregación temporal surgen como una alternativa natural que permite generar información en escalas finas a partir de registros a resoluciones más gruesas. Entre las metodologías más usadas destacan los modelos estocásticos basados en clústeres Poisson como los de Bartlett–Lewis y Neyman–Scott (Rodríguez-Iturbe et al., 1987; Onof & Wheater, 1993), los modelos de cascadas aleatorias (Gupta & Waymire, 1993; Molnar & Burlando, 2005), y otros modelos no paramétricos que no requieren suponer distribuciones de probabilidad, como el método de análogos de los  $k$ -vecinos más cercanos o  $k$ -NN (Lall & Sharma, 1996; Sharma & Srikanthan, 2006).

La validación de estos métodos suele basarse según su capacidad de replicar estadísticas generales de las observaciones: media, desviación estándar, autocorrelación y proporción de periodos secos (e.g., Onof & Wheater, 1993; Socolofsky et al., 2001), y se ha evaluado su capacidad de reproducir extremos, como las precipitaciones máximas anuales (AMP) de distintas duraciones y percentiles extremos (e.g., Bhattacharyya & Saha, 2023; Molnar & Burlando, 2005). Aún no hay consenso para elegir métodos de desagregación orientados a extremos: los estudios suelen enfocarse en casos o métricas puntuales y omiten la incertidumbre del proceso (Qin & Dai, 2024); más aún, tampoco existe un marco metodológico común.

El objetivo de este trabajo es proponer y evaluar un marco metodológico para comparar el desempeño de desagregadores de precipitación diaria-horaria con foco en extremos. El marco evalúa (i) la exactitud de las curvas IDF desagregadas, entendida como el sesgo entre la mediana (percentil 50) de las realizaciones y la curva IDF observada, y (ii) la precisión, entendida como la consistencia entre desagregaciones (dispersión intramétodo, de naturaleza estocástica) en relación con la incertidumbre estadística de la curva IDF observada. Complementariamente, se consideran índices extremos de precipitación (ERIs) análogos a los propuestos por el Equipo de Expertos sobre Detección e Índices del Cambio Climático (ETCCDI) y adaptados a resolución horaria (véase Sección 2.2.1): TDH, RX5-h, R95-h, R95p-h. El marco se aplica en cinco sitios usando tres metodologías (véase Sección 3): un modelo de pulsos rectangulares (SOC),  $k$ -NN y un método basado en curvas de Huff (HUFF).

## 2. MARCO METODOLÓGICO

El marco evalúa el desempeño de los métodos de desagregación en la reproducción de curvas IDF, cuatro ERIs y sus respectivas incertidumbres. Consta de tres componentes (Figura 1): (1) desagregación de series diarias a horarias con los métodos evaluados; (2) evaluación de exactitud y precisión en la reproducción de las curvas IDF; (3) evaluación de la capacidad de los métodos evaluados para reproducir las series anuales de cuatro ERIs relevantes para la caracterización hidroclimática.



**Figura 1.** Marco de evaluación. (a) Esquema general. (b) Ejemplo de evaluación de curvas IDF para distintos periodos de retorno ( $T$ ): observadas (rojo), desagregadas (azul) y de referencia (negro), usadas para calcular  $AEI_T$ ,  $PEI_T$  y  $D_T$  (esta última representada como elipses concéntricas). (c) Ejemplo de evaluación de ERIs: los valores de  $mKGE$  entre la serie observada y las desagregadas para cada ERI se muestran como diagramas de caja.

## 2.1. EVALUACIÓN BASADA EN CURVAS IDF

### 2.1.1. CONSTRUCCIÓN DE CURVAS IDF

La curva IDF observada y las desagregadas son construidas a partir de un análisis de frecuencia empírico mediante posición de ploteo de Weibull sobre las series de AMP observada y desagregadas para duraciones  $d$  entre 1 a 24 h. Para cada duración, las AMPs se obtienen mediante ventanas móviles, calculando la precipitación acumulada máxima en intervalos horarios continuos de  $d$  horas y seleccionando el máximo en cada año. Además, se propone una curva IDF de referencia, de naturaleza determinística, construida a partir de la serie observada agregada a escala diaria. Su propósito es servir de *benchmark* en la evaluación del desempeño de las curvas IDF desagregadas. Para su construcción, primero la precipitación máxima diaria de cada año ( $AMP_{24}$ ) se distribuye en 24 pulsos horarios

suponiendo un perfil de intensidad lineal decreciente con el pulso más lluvioso ubicado en la primera hora. Esta elección obedece a un criterio de diseño no paramétrico simple y de fácil implementación, que no requiere estimación de parámetros. Luego, los pulsos horarios se calculan con la Ecuación (1):

$$I(d) = \frac{AMP_{24}}{12} \cdot \left(1 - \frac{d}{24}\right) \quad (1)$$

donde, por construcción, la curva se evalúa en  $d = 0.5, 1.5, \dots, 23.5$  h para obtener las intensidades entre 1 y 24 h. Luego, para cada duración, se selecciona el tramo consecutivo más lluvioso en cada uno de los años de registro. El análisis de frecuencia empírico sobre estos valores permite estimar la intensidad correspondiente a cada periodo de retorno, generando así la curva IDF de referencia al repetir este proceso para cada duración.

### 2.1.1.1. TIPOS DE INCERTIDUMBRE EN CURVAS IDF

Las curvas IDF observadas y desagregadas traen consigo fuentes de incertidumbre de distinta naturaleza. La incertidumbre de las curvas desagregadas es estocástica y refleja la variabilidad intrínseca que un método de desagregación estocástico genera entre distintas realizaciones del proceso. En cambio, la incertidumbre de la curva IDF observada es estadística y surge porque el registro observado usado para la construcción de la curva es una muestra finita del clima, con una extensión temporal e historial de eventos específicos.

En ambos casos, la banda de incertidumbre se cuantifica como un intervalo de confianza (IC) del 90%, pero el modo de estimarlo difiere. Para cada duración  $d$ , cada serie desagregada genera su propia serie de AMPs y, sobre esa serie, se estima la AMP de período de retorno  $T$  años mediante análisis de frecuencia empírico. Al reunir las AMPs de  $T$  años entre todas las series desagregadas, los percentiles 5-95 definen el IC del 90% (con la mediana siendo la curva percentil 50). Repitiendo para todo  $d$  se obtiene la banda de incertidumbre estocástica. Por otra parte, para la IDF observada, para cada  $d$  se ajusta una distribución Gumbel a las AMP observadas y se ejecuta un *bootstrap* paramétrico: se generan  $B$  muestras sintéticas con extensión igual al número de años observados y en cada una se calcula empíricamente la AMP de  $T$  años. Los percentiles 5-95 de esas  $B$  AMPs definen el IC del 90% por duración; al repetirlo para todo  $d$ , se obtiene la banda de incertidumbre estadística.

### 2.1.2. ANÁLISIS EXACTITUD-PRECISIÓN DE CURVAS IDF

Para evaluar el desempeño de las metodologías de desagregación en replicar las curvas IDF, este estudio propone y emplea tres métricas novedosas: el Índice de Exactitud ( $AEI_T$ , por sus siglas en inglés), el Índice de Precisión ( $PEI_T$ , por sus siglas en inglés) y la Distancia Euclidiana ( $D_T$ ).

El  $AEI_T$ , inspirado en el Coeficiente de Eficiencia de Nash-Sutcliffe Modificado (Legates & McCabe, 1999), cuantifica qué tan exacta, en términos de sesgo absoluto, es la curva IDF desagregada percentil 50 respecto a la curva observada y se calcula según la Ecuación (2), donde, para cada duración  $d$  y periodo de retorno  $T$ ,  $\hat{I}_d^T$  es la intensidad de la curva IDF

observada,  $I_{d,50}^T$  es la intensidad desagregada percentil 50 proveniente de la banda de incertidumbre estocástica, e  $\tilde{I}_d^T$  es la intensidad de la curva de referencia. El  $AEI_T$  varía entre  $-\infty$  y 1: valores cercanos a 1 indican mayor coincidencia con la curva observada, mientras que valores menores a 0 reflejan un desempeño inferior a la curva de referencia.

El  $PEI_T$  cuantifica si la incertidumbre estocástica es comparable con la incertidumbre estadística, y se calcula de acuerdo con la Ecuación (3), donde  $\widehat{CI}_{T,d}^\pm$  y  $CI_{T,d}^\pm$  son los límites superior e inferior del IC del 90% para la incertidumbre estadística y estocástica, respectivamente. El  $PEI_T$  varía entre  $-\infty$  y 1: valores entre 0 y 1 son preferibles, ya que un valor de 1 implica ausencia de incertidumbre adicional por la desagregación, mientras que valores negativos indican incertidumbre superior a la variabilidad natural.

Finalmente, la métrica  $D_T$  integra  $AEI_T$  y  $PEI_T$  y corresponde a la distancia euclidiana entre el punto  $(AEI_T, PEI_T)$  y el óptimo  $(1,1)$ , de acuerdo con la Ecuación (4). Un valor de  $D_T = 0$  indica un desempeño perfecto (máxima exactitud y precisión). Los factores de ponderación  $w_{AEI}$  y  $w_{PEI}$  permiten priorizar una métrica sobre la otra según la preferencia de quien modela.

$$AEI_T = 1 - \frac{\sum_{d=1}^{24} \frac{1}{d} |\hat{I}_d^T - I_{d,50}^T|}{\sum_{d=1}^{24} \frac{1}{d} |\hat{I}_d^T - \tilde{I}_d^T|} \quad (2)$$

$$PEI_T = 1 - \frac{\sum_{d=1}^{24} \frac{1}{d} \left( \frac{CI_{T,d}^+ - CI_{T,d}^-}{I_{d,50}^T} \right)}{\sum_{d=1}^{24} \frac{1}{d} \left( \frac{\widehat{CI}_{T,d}^+ - \widehat{CI}_{T,d}^-}{\tilde{I}_d^T} \right)} \quad (3)$$

$$D_T = \sqrt{w_{AEI}(AEI_T - 1)^2 + w_{PEI}(PEI_T - 1)^2} \quad (4)$$

La Figura 1b ilustra un ejemplo del análisis de exactitud y precisión, donde los índices  $AEI_T$  y  $PEI_T$  se representan en los ejes horizontal y vertical, respectivamente.  $D_T$  corresponde a la distancia al punto óptimo  $(1,1)$ , visualizada en forma de elipses concéntricas según los factores de ponderación empleados. Nótese que para el cálculo de  $AEI_T$  y  $PEI_T$  se usan ponderaciones inversamente proporcionales a la duración ( $d$ ), para dar más importancia a las intensidades cortas, relevantes para crecidas de rápida respuesta (Viglione & Blöschl, 2009).

## 2.2. EVALUACIÓN DE ÍNDICES EXTREMOS DE PRECIPITACIÓN

### 2.2.1. ÍNDICES EXTREMOS DE PRECIPITACIÓN (ERIs)

La Tabla 1 muestra los ERIs utilizados en este estudio, basados en los índices hidroclimáticos propuestos por el ETCCDI. Estos han sido ampliamente utilizados en la literatura (e.g. Klein Tank & Können, 2003; Taschetto & England, 2009; Kioutsioukis et al., 2009; Zhang et al., 2011; Vicuña et al., 2013). Aunque originalmente desarrollados para datos diarios, aquí se utilizan sus versiones homólogas a escala horaria. Estos ERIs representan distintos aspectos del comportamiento extremo de la precipitación en un año: (1) Horas totales secas (TDH, *total dry hours*), (2) precipitación máxima en 5 horas consecutivas (RX5-h), (3) número de

horas con precipitación sobre el percentil 95 observado (R95-h) y (4) precipitación acumulada en las horas con lluvia sobre dicho percentil (R95p-h). Los dos últimos reflejan, respectivamente, la duración y magnitud de pulsos extremos.

**Tabla 1.** ERIs usados en este estudio.

Índice	Descripción	Unidad
TDH	Número de horas sin lluvia ( $P < 0.2$ mm)	h
RX5-h	Máxima P en cinco horas consecutivas	mm
R95-h	Número de horas en que $P > P_{95}$ del registro observado	h
R95p-h	P acumulada en horas con P sobre $P_{95}$ del registro observado	mm

### 2.2.2. ÍNDICE DE EFICIENCIA DE KLING-GUPTA MODIFICADO (*mKGE*)

Para evaluar si las series horarias desagregadas generan ERIs similares a las observadas, este estudio propone un coeficiente de eficiencia Kling-Gupta modificado (*mKGE*), basado en Gupta et al. (2009):

$$mKGE = 1 - \sqrt{(\sigma_{sim}/\sigma_{obs} - 1)^2 + (\mu_{sim}/\mu_{obs} - 1)^2} \quad (5)$$

donde  $\sigma_{obs}$  y  $\sigma_{sim}$  son las desviaciones estándar, y  $\mu_{obs}$  y  $\mu_{sim}$  las medias de las series anuales observada y desagregada de cada ERI, respectivamente. De esta forma, por cada ERI se tendrá tantos valores de *mKGE* como series desagregadas haya. A diferencia del *KGE* original, el *mKGE* no incluye el coeficiente de correlación lineal, ya que su incorporación generaría ruido en el desempeño de los métodos, los cuales, por su naturaleza estocástica, no buscan reproducir las series de tiempo exactas, sino los patrones estadísticos de dichas series. En este contexto, donde los índices se estiman a partir de valores anuales agregados, el coeficiente de correlación carece de significado hidrológico. El *mKGE* varía de  $-\infty$  a 1 (óptimo), con un umbral mínimo aceptable de desempeño  $mKGE_{ref} = 0$  ( $\mu_{sim} = \mu_{obs}$  y  $\sigma_{sim} = 0$ ), siguiendo las sugerencias de Knoben et al. (2019). Se busca, por lo tanto, obtener valores altos de *mKGE* y baja dispersión entre desagregaciones. La Figura 1c ilustra este cálculo y los valores de referencia.

## 3. MÉTODOS DE DESAGREGACIÓN

### 3.1. MÉTODO ESTOCÁSTICO DE TIPO PULSO (SOC)

Esta técnica de desagregación, basada en el trabajo de Socolofsky et al. (2001), divide la precipitación diaria en eventos subdiarios mediante pulsos rectangulares, con intensidad, duración y hora de inicio asignadas aleatoriamente a partir de registros horarios mensuales, de manera de preservar la estacionalidad. El método incluye la calibración de un umbral mínimo de evento mensual  $\epsilon$ , para no subestimar la probabilidad de no lluvia (Choi et al., 2008). Se presenta aquí un resumen; para detalles adicionales, véase Socolofsky et al. (2001).

- i. Ajuste de distribución: Para cada mes, los eventos de precipitación se separan a partir de una serie horaria de referencia, considerando al menos 6 horas secas entre ellos como criterio de independencia (Huff, 1967). Luego, se ajustan distribuciones

- Gamma de dos parámetros y Exponencial a las profundidades de los eventos y se selecciona la mejor según la prueba de bondad de ajuste Kolmogorov-Smirnov.
- ii. Asignación del primer pulso: La precipitación diaria a desagregar se evalúa en la función de densidad acumulada de la distribución escogida para el mes respectivo, y se determina su probabilidad de no excedencia. A partir de esta probabilidad, se genera un valor uniforme comprendido entre 0 y dicha probabilidad, cuyo cuantil asociado corresponde a la profundidad del primer pulso rectangular subdiario desagregado.
  - iii. Iteración hasta convergencia: Tras obtener el primer pulso subdiario, se actualiza la precipitación diaria restante y se repite el procedimiento con el nuevo monto hasta que esta sea menor o igual al umbral mensual  $\varepsilon$ , momento en que termina la desagregación.
  - iv. Duración y ubicación: A cada pulso se le asigna duración y hora de inicio, usando duraciones históricas similares (mediante la estrategia de búsqueda de análogos de nombre homónimo al método explicado en la siguiente subsección) y ubicándolos aleatoriamente íntegramente dentro del día; si se solapan, las intensidades se suman.
  - v. Repetición: El proceso se repite para cada día de la serie a desagregar.

### 3.2. MÉTODO k-NN

Este método no paramétrico utiliza un vector de “fragmentos” que representa las proporciones de precipitación horaria derivadas del monto diario original (Pui et al., 2012), seleccionado de un día histórico análogo identificado mediante el algoritmo k-NN (Lall y Sharma, 1996). Así, se mantiene la dependencia temporal diaria, a la vez que se asigna aleatoriedad a escala horaria (Sharma y Srikanthan, 2006). El tamaño de ventana  $l$  para buscar días análogos se determina de acuerdo con Alam & Elshorbagy (2015), con el propósito de preservar la estacionalidad. Para más detalles, véanse las referencias originales.

- i. Preparación de datos: Se conforma una base de días de referencia con datos diarios y horarios. Para cada día, se calculan los fragmentos horarios adimensionales como proporción respecto al total diario.
- ii. Selección de días análogos: Para cada día a desagregar, se buscan potenciales análogos en una ventana estacional ( $l$ ) a lo largo de  $N$  años, formando un vector de tamaño  $n = l \cdot N$ .
- iii. Identificación de los vecinos más cercanos: Se calcula la diferencia entre el valor diario a desagregar y cada análogo; se seleccionan los  $k$  más similares ( $k = \sqrt{n}$ ).
- iv. Selección aleatoria ponderada: Se elige uno de los  $k$  días análogos usando una probabilidad ponderada  $\left( p_j = \frac{\frac{1}{j}}{\sum_{i=1}^k \frac{1}{i}} \right)$ , privilegiando los más similares.
- v. Desagregación: El valor diario objetivo se reparte en 24 horas usando los fragmentos horarios del día análogo seleccionado, asegurando el balance de masa.
- vi. Repetición: El procedimiento se repite para cada día a desagregar.

### 3.3. MÉTODO BASADO EN PATRONES ADIMENSIONALES DE TORMENTA (HUFF)

Se propone y aplica un método sencillo de desagregación diaria a horaria basado en patrones adimensionales de tormenta mediante curvas de Huff (Huff, 1967), que describen la relación entre tiempo transcurrido y precipitación acumulada escalada de 0 a 100% para eventos históricos en cada sitio. A continuación, se resume el método; para más detalle sobre el cálculo de las curvas, véase Huff (1967).

- i. Identificación de eventos: Se identifican los eventos de lluvia desde la serie horaria de referencia, considerando al menos seis horas secas para la independencia de estos (Huff, 1967).
- ii. Clasificación y construcción de curvas: Los eventos se agrupan en cuartiles según el momento de mayor acumulación de precipitación, y se construyen curvas adimensionales (precipitación acumulada vs. tiempo, ambos de 0 a 100%).
- iii. Análisis de frecuencia: Para cada cuartil y tiempo, se generan curvas para distintas probabilidades de excedencia (10–90%).
- iv. Selección de curva y desagregación: Se selecciona aleatoriamente un cuartil y una curva de Huff. Se asignan duración y hora de inicio como en SOC, y la distribución horaria diaria se define según la curva seleccionada.
- v. Repetición: Se aplica el proceso a cada día a desagregar.

#### 4. CASO DE ESTUDIO

El marco metodológico se evaluó empleando registros horarios de cinco sitios con diferentes regímenes de precipitación: Altheim (Alemania), Hikone (Japón), Pathhead (Reino Unido), Oregon (Estados Unidos) y Quinta Normal (Chile). El Anexo A1 muestra la ubicación y el régimen de lluvias de cada sitio. El set de datos proviene del Global Sub-Daily Rainfall Dataset (INTENSE; Lewis et al., 2019, 2021) y de la Dirección Meteorológica de Chile (Quinta Normal), todos con al menos 10 años de registros y menos de 0.8% de datos faltantes. Para un mayor detalle de los registros empleados, se sugiere visitar el Anexo A2.

Los registros horarios observados de cada sitio fueron empleados de forma independiente en la calibración y la posterior validación de los métodos de desagregación, empleando en ambos casos el esquema de validación cruzada “*leave-one-year-out*”. Este enfoque implica agregar a escala diaria los valores de las series observadas horarias de cada año, y luego desagregarlos con los propios datos horarios, pero de los años restantes. De esta forma, para cada sitio y metodología, 300 series fueron desagregadas para el preproceso de calibración de SOC y k-NN, y para la estimación de las curvas de Huff. Posteriormente, se generaron otras 300 series desagregadas para el proceso de validación, permitiendo una estimación robusta de la incertidumbre estocástica.

En la evaluación de curvas IDF, se consideraron periodos de retorno de 2, 5, 10 y 20 años, excluyéndose valores mayores de  $T$  para determinar un estándar común entre sitios con disponibilidad heterogénea de datos y evitar así una fuente de incertidumbre adicional asociada al uso de muestras acotadas en extensión (e.g., Altheim, que tiene sólo 10 años, para el que las intensidades de  $T = 20$  años son obtenida mediante la extrapolación lineal de la función densidad acumulada empírica inversa de cada duración  $d$ ). Por cada  $T$ , se generó una curva IDF por serie desagregada, para un total de 300, permitiendo una estimación fiel de la curva IDF percentil 50 y la banda de incertidumbre estocástica. Para la incertidumbre

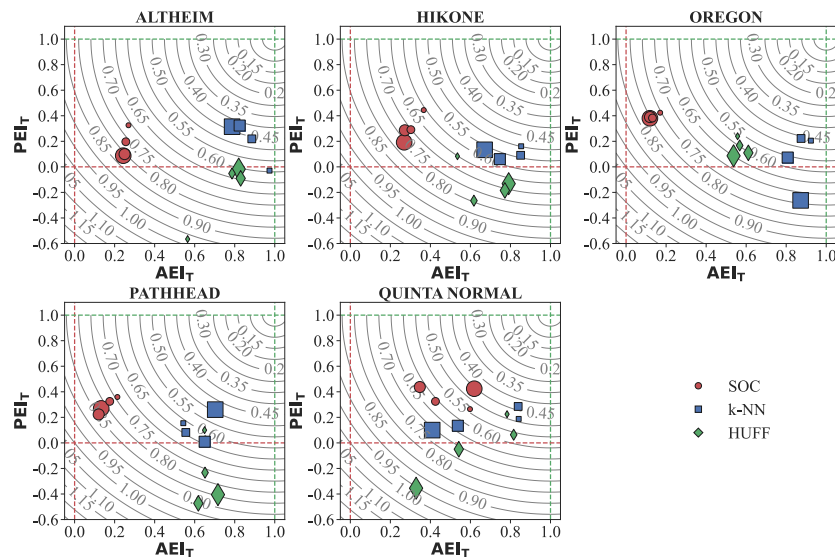
estadística, se generaron 1000 series de AMP por duración, siguiendo el procedimiento de *bootstrapping* paramétrico descrito en la Sección 2.1.1.1. Para la Distancia Euclidiana ( $D_T$ ) se asignó más peso a la componente de exactitud ( $w_{AEI} = 2/3$ ) que a la de precisión ( $w_{PEI} = 1/3$ ), priorizando que la curva IDF percentil 50 represente bien la curva observada.

Finalmente, para los ERIs, se obtuvo la serie anual de cada índice para la serie observada y para cada una de las 300 series desagregadas, permitiendo obtener un valor de  $mKGE$  por desagregación, contabilizando un total de 300 valores de  $mKGE$  por índice. Luego, para cada índice, se obtiene el valor del  $mKGE$  percentil 50 y se analiza la dispersión (Figura 1c).

## 5. RESULTADOS Y DISCUSIÓN

### 5.1. ANÁLISIS DE LAS CURVAS IDF

La Figura 2 muestra los resultados de exactitud ( $AEI_T$ ) y precisión ( $PEI_T$ ) en la representación de las curvas IDF para cada sitio y método. Todas las metodologías presentan buenos desempeños en  $AEI_T$  (siempre  $\geq 0$ ), mientras que  $PEI_T$  exhibe al menos un valor negativo en cada sitio, aunque esto no implica necesariamente mal desempeño. Valores de  $PEI_T$  cercanos a 0 son esperables y señalan que la incertidumbre del proceso estocástico de desagregación es comparable a la variabilidad natural estadística. Las diferencias entre los comportamientos manifestados por  $AEI_T$  y  $PEI_T$  evidencian que el desempeño es altamente dependiente del método estudiado y del sitio analizado, diferencias que son apreciables directamente a partir del marco metodológico propuesto.



**Figura 2.** Análisis de exactitud ( $AEI_T$ ) y precisión ( $PEI_T$ ) de las curvas IDF. Los periodos de retorno de 2, 5, 10 y 20 años se representan con un tamaño de marcador mayor a medida que aumenta su valor.

El marco permite ver rápidamente qué método muestra el mejor desempeño en cada métrica de análisis. La Figura 2 permite ver que SOC es el que tiende a representar de mejor manera el  $PEI_T$ , sugiriendo una menor variabilidad entre desagregaciones. Al observar  $AEI_T$ , el

comportamiento es opuesto, ocurriendo que los pares ordenados que ofrece SOC en los distintos lugares tienden a estar más hacia la izquierda (a excepción de Quinta Normal, donde se observa un desempeño similar entre metodologías), sugiriendo una menor representación de exactitud con respecto a k-NN y HUFF.

Al considerar un análisis integrado de ambas métricas mediante  $D_T$  (elipses concéntricas), se identifica claramente qué metodología tiende a mostrar una menor distancia con respecto al punto (1,1).  $D_T$  funciona entonces como curva de indiferencia: para un cierto nivel de distancia (desempeño) esperable, se puede definir qué combinaciones de  $AEI_T$  y  $PEI_T$  son permisibles de acuerdo con el criterio preestablecido y con ello qué metodologías y periodos de retorno lo satisfacen. Si se quiere priorizar una componente, basta ajustar los pesos y recalcular las combinaciones óptimas. En este estudio, exactitud pesa el doble que precisión: una caída de una unidad en  $AEI_T$  penaliza dos veces más que la misma caída en  $PEI_T$ .

La Figura 2 también muestra diferencias entre combinaciones ( $AEI_T, PEI_T$ ) para un mismo método. En primer lugar, el desempeño es dependiente del  $T$  analizado. Si bien se aprecia una tendencia al agrupamiento de puntos de una misma metodología, esto no siempre ocurre, como es el caso de Quinta Normal con todos los métodos, o en Oregon, donde k-NN ofrece pares más dispersos. En segundo lugar, dicho desempeño no es transversalmente consistente entre sitios para una metodología dada. Un claro ejemplo es el par ( $AEI_T, PEI_T$ ) de  $T = 20$  bajo k-NN, el cual tiene el mejor desempeño en Pathhead y Altheim entre los 4 valores de  $T$ , pero el peor en Oregon y en Quinta Normal. En tercer lugar, no existe una relación monótonica consistente entre un valor de  $T$  y el desempeño logrado. Por ejemplo, SOC sugiere que un mayor  $T$  implica un menor  $PEI_T$  (salvo en Quinta Normal), mientras que k-NN sugiere lo contrario, llegando a apreciarse una relación opuesta para  $T = 20$  en Pathhead respecto a los otros periodos de retorno, y algo similar en Quinta Normal, donde la correlación lineal entre  $T = 2$  y  $T = 5$  es opuesta a la existente entre  $T = 10$  y  $T = 20$ . Algo similar se aprecia con  $AEI_T$ : mientras SOC y k-NN muestra consistentemente que un mayor  $T$  implica un menor  $AEI_T$ , este comportamiento es opuesto al que ofrece HUFF, donde valores más altos de  $T$  tienden a ofrecer un  $AEI_T$  más alto (salvo en Oregon y Quinta Normal).

**Tabla 2.** Valores promedio de exactitud ( $\overline{AEI_T}$ ), precisión ( $\overline{PEI_T}$ ) y distancia euclidiana ( $\overline{D_T}$ ) de las curvas IDF.  $\overline{D_T}$  corresponde al valor promedio de  $D_T$  para los cuatro periodos de retorno.

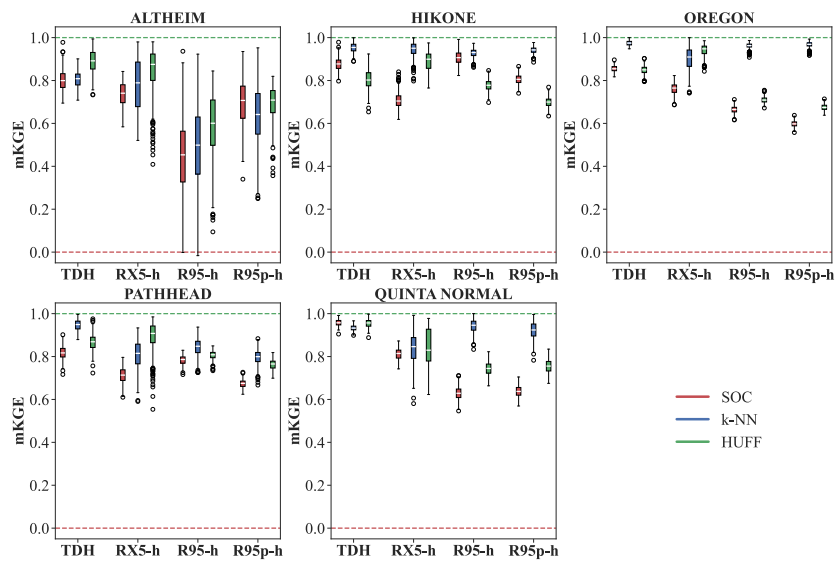
Lugar	$\overline{AEI_T}$			$\overline{PEI_T}$			$\overline{D_T}$		
	SOC	k-NN	HUFF	SOC	k-NN	HUFF	SOC	k-NN	HUFF
Altheim	0.254	0.868	0.750	0.177	0.207	-0.181	0.773	0.476	0.713
Hikone	0.367	0.780	0.679	0.304	0.111	-0.125	0.698	0.547	0.707
Oregon	0.135	0.869	0.568	0.394	0.059	0.151	0.788	0.555	0.605
Pathhead	0.161	0.613	0.659	0.294	0.126	-0.253	0.797	0.597	0.777
Quinta Normal	0.497	0.656	0.617	0.363	0.177	-0.029	0.557	0.564	0.677

Para comparar desempeños entre las distintas metodologías, la Tabla 2 resume el desempeño promedio de cada método por sitio ( $\overline{AEI_T}, \overline{PEI_T}$  y  $\overline{D_T}$ ). En general, k-NN destaca en  $\overline{AEI_T}$  (y en  $\overline{D_T}$  en consecuencia). En tanto, SOC suele obtener mejores resultados en  $\overline{PEI_T}$ . La superioridad de k-NN para representar los AMP y las curvas IDF ha sido reportada por otros

autores (Pui et al., 2012; Poschlod et al., 2018). Por su parte, la tendencia de SOC a subestimar intensidades extremas es coherente con modelos tipo pulso, ya que la asignación aleatoria de duraciones puede generar una subrepresentación de eventos cortos e intensos (Kaczmarska et al., 2014; Cross et al., 2018).

## 5.2. ANÁLISIS DE LOS ERIs

La Figura 3 muestra los resultados de  $mKGE$  para los ERIs según sitio y método de desagregación. En general, todos los métodos presentan buen desempeño y baja dispersión, salvo en Altheim, donde la mayor variabilidad se atribuye a la menor extensión temporal del registro (10 años). Sin embargo, todos los valores de  $mKGE$  superan el umbral mínimo de 0. En los otros sitios, el desempeño es excelente, especialmente para el índice TDH (mayor a 0.7 en la mayoría de los casos), indicando una buena representación de las horas secas.



**Figura 3.** Diagramas de caja con los valores de  $mKGE$  obtenidos para las series de ERIs en los diferentes lugares de estudio.

El marco permite evaluar rápidamente la capacidad de cada método para reproducir los ERIs mediante el  $mKGE$ . La Figura 3 muestra que el marco metodológico presenta al  $mKGE$  como una distribución y no como un único valor. Esto habilita comparar métodos por tendencia central (percentil 50) y dispersión: a igual mediana, un rango intercuartil (IQR) alto indica mayor incertidumbre que IQR bajo. Además, el marco deja ver qué método replica mejor cada ERI. En este estudio, el desempeño aceptable se fija con un umbral  $mKGE \geq mKGE_{ref}$ , el cual es ajustable según los criterios operativos del usuario.

El percentil 50 de los valores del  $mKGE$  para cada ERI permite comparar el desempeño entre las metodologías (Tabla 3). En TDH, k-NN lidera en tres de los cinco lugares ( $mKGE > 0.81$ ); en RX5-h, HUFF es superior en tres de los cinco lugares ( $mKGE > 0.83$ ); y para R95-h y R95p-h, k-NN domina en cuatro de los cinco lugares, aunque presenta su peor desempeño en Altheim. En general, todos los valores de  $mKGE$  indican un buen desempeño. Estos resultados reflejan la naturaleza de cada método: k-NN sobresale en TDH por su capacidad

de preservar la intermitencia natural de la lluvia, mientras que SOC, pese a su enfoque tipo pulso, no logra replicarla bien debido a la asignación de los pulsos no sigue un proceso Poisson, a diferencia de los métodos de procesos puntuales. k-NN también destaca en percentiles extremos por captar la magnitud y distribución histórica de eventos. HUFF, en tanto, es especialmente efectivo para eventos de duración intermedia, al basarse en patrones históricos acumulados.

**Tabla 3.** Percentil 50 de los valores de  $mKGE$  de los ERIs en cada lugar.

Lugar	TDH			RX5-h			R95-h			R95p-h		
	SOC	k-NN	HUFF	SOC	k-NN	HUFF	SOC	k-NN	HUFF	SOC	k-NN	HUFF
Altheim	0.800	0.809	0.892	0.741	0.789	0.875	0.453	0.499	0.601	0.707	0.642	0.709
Hikone	0.877	0.954	0.802	0.704	0.950	0.899	0.907	0.930	0.778	0.806	0.943	0.700
Oregon	0.855	0.974	0.850	0.765	0.910	0.948	0.665	0.963	0.709	0.598	0.969	0.674
Pathhead	0.817	0.949	0.869	0.713	0.815	0.908	0.786	0.848	0.808	0.675	0.799	0.765
Quinta Normal	0.957	0.932	0.950	0.814	0.846	0.830	0.629	0.946	0.744	0.637	0.924	0.755

## 6. CONCLUSIONES

Se desarrolló un marco metodológico para evaluar métodos de desagregación diaria-horaria según su capacidad para representar extremos, aplicado sobre tres enfoques y cinco sitios distintos. Las principales conclusiones son:

- El marco integra dos vistas complementarias de los extremos: (i) curvas IDF, donde se distingue explícitamente exactitud y precisión en la representación de las curvas observadas, y (ii) índices extremos de precipitación (ERIs), evaluados con la métrica  $mKGE$ . En ambos casos, los resultados se reportan como distribuciones (percentil 50 y dispersión) y no como valores únicos.
- En línea con lo anterior, el marco considera en la evaluación dos fuentes intrínsecas de incertidumbre para el análisis de las curvas IDF: la estocástica (propia del proceso de desagregación) y la estadística (asociada al carácter muestral del registro observado). Ambas se expresan mediante intervalos de confianza del 90%, comparables entre métodos, duraciones y sitios. El marco también reconoce la incertidumbre estocástica en la evaluación de los ERIs por medio de la distribución de los  $mKGE$ .
- El marco incorpora criterios de decisión flexibles y adaptables para evaluar el desempeño, como la distancia euclidiana  $D_T$  (por medio de pesos relativos para exactitud y precisión) para las curvas IDF y el  $mKGE_{ref}$  para los ERIs.
- El marco es modular: admite nuevos periodos de retorno, ERIs, resoluciones temporales o métodos de desagregación sin cambiar la lógica de evaluación.

El caso de estudio mostró que el desempeño resulta altamente dependiente del contexto: no emergió un método universalmente superior; más bien, algunos enfoques destacaron en métricas ligadas a percentiles (R95-h/R95p-h) o a intermitencia (TDH), mientras otros lo hicieron en representar máximos anuales de duración intermedia (RX5-h). Esto resalta y ejemplifica la capacidad del marco para evidenciar fortalezas y limitaciones de acuerdo con el objetivo y con las condiciones locales.

## AGRADECIMIENTOS

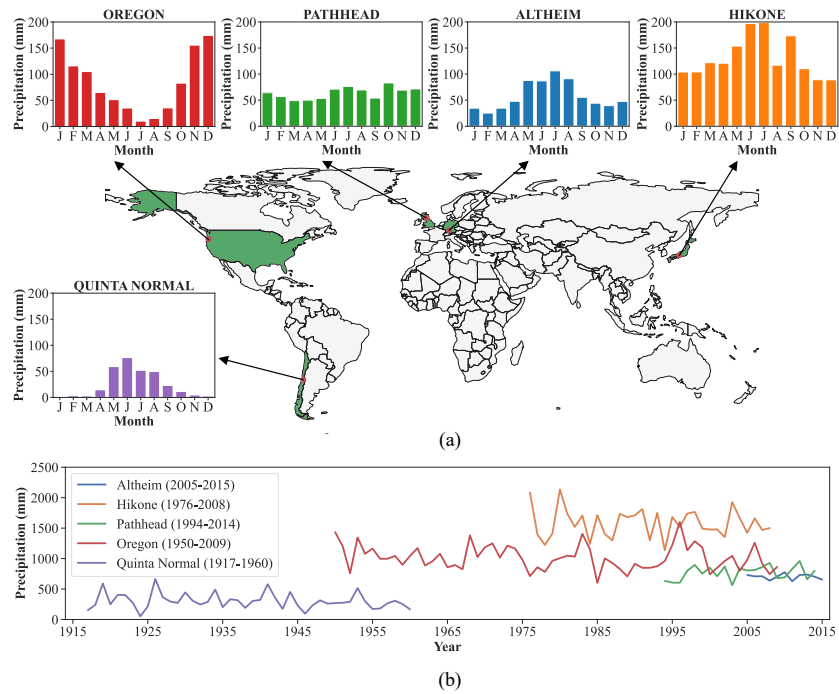
Agradecemos el financiamiento de FONDECYT (1241163), ANID (FONDAP 1523A0004 y 1523A0009), el apoyo de la Universidad Adolfo Ibáñez (FONDECYT Puente 2025) y la Beca Arturo Cousiño Lyon de la Fundación San Carlos de Maipo.

## REFERENCIAS

- Alam, M. S., & Elshorbagy, A. (2015). Quantification of the climate change-induced variations in intensity–duration–frequency curves in the Canadian Prairies. *Journal of Hydrology*, 527, 990–1005. <https://doi.org/10.1016/j.jhydrol.2015.05.059>
- Bhattacharyya, D., & Saha, U. (2023). Deep learning application for disaggregation of rainfall with emphasis on preservation of extreme rainfall characteristics for Indian monsoon conditions. *Stochastic Environmental Research and Risk Assessment*, 37(1021–1038). <https://doi.org/10.1007/s00477-022-02331-x>
- Bruni, G., Reinoso, R., van de Giesen, N. C., Clemens, F. H. L. R., & ten Veldhuis, J. A. E. (2015). On the sensitivity of urban hydrodynamic modelling to rainfall spatial and temporal resolution. *Hydrology and Earth System Sciences*, 19(2), 691–709. <https://doi.org/10.5194/hess-19-691-2015>
- Choi, J., Socolofsky, S. A., & Olivera, F. (2008). Hourly disaggregation of daily rainfall in Texas using measured hourly precipitation at other locations. *Journal of Hydrologic Engineering*, 13(6), 476–487. [https://doi.org/10.1061/\(ASCE\)1084-0699\(2008\)13:6\(476\)](https://doi.org/10.1061/(ASCE)1084-0699(2008)13:6(476))
- Cross, D., Onof, C., Winter, H., & Bernardara, P. (2018). Censored rainfall modelling for estimation of fine-scale extremes. *Hydrology and Earth System Sciences*, 22(1), 727–756. <https://doi.org/10.5194/hess-22-727-2018>
- Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, 377(1–2), 80–91. <https://doi.org/10.1016/j.jhydrol.2009.08.003>
- Gupta, V., & Waymire, E. (1993). A statistical analysis of mesoscale rainfall as a random cascade. *Journal of Applied Meteorology*, 32, 251–267. [https://doi.org/10.1175/1520-0450\(1993\)032<0251:ASAOMR>2.0.CO;2](https://doi.org/10.1175/1520-0450(1993)032<0251:ASAOMR>2.0.CO;2)
- Huff, F. A. (1967). Time distribution of rainfall in heavy storms. *Water Resources Research*, 3(4), 1007–1021. <https://doi.org/10.1029/WR003i004p01007>
- Kaczmarzka, J., Isham, V., & Onof, C. (2014). Point process models for fine-resolution rainfall. *Hydrological Sciences Journal*, 59(11), 1972–1991. <https://doi.org/10.1080/02626667.2014.925558>
- Kioutsioukis, I., Melas, D., & Zerefos, C. (2009). Statistical assessment of changes in climate extremes over Greece (1955–2002). *International Journal of Climatology*, 30(11), 1723–1737. <https://doi.org/10.1002/joc.2030>
- Klein Tank, A. M. G., & Können, G. P. (2003). Trends in indices of daily temperature and precipitation extremes in Europe, 1946–99. *Journal of Climate*, 16(24), 3665–3680. [https://doi.org/10.1175/1520-0442\(2003\)016%3C3665:TIHODT%3E2.0.CO;2](https://doi.org/10.1175/1520-0442(2003)016%3C3665:TIHODT%3E2.0.CO;2)
- Knoben, W. J. M., Freer, J. E., & Woods, R. A. (2019). Technical note: Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores. *Hydrology and Earth System Sciences*, 23(10), 4323–4331. <https://doi.org/10.5194/hess-23-4323-2019>
- Lall, U., & Sharma, A. (1996). A nearest neighbor bootstrap for resampling hydrologic time series. *Water Resources Research*, 32(3), 679–693. <https://doi.org/10.1029/95WR02966>
- Legates, D. R., & G. J. McCabe Jr. (1999). Evaluating the use of “goodness-of-fit” measures in hydrologic and hydroclimatic model validation. *Water Resources Research*, 35(1), 233–241. <https://doi.org/10.1029/1998WR900018>
- Lewis, E., Fowler, H. J., Alexander, L. V., Dunn, R., McClean, F., Barbero, R., Guerreiro, S., Li, X.-F., & Blenkinsop, S. (2019). GSDR: A global sub-daily rainfall dataset. *Journal of Climate*, 32(14), 4715–4729. <https://doi.org/10.1175/JCLI-D-18-0143.1>
- Lewis, E., Pritchard, D., Villalobos-Herrera, R., Blenkinsop, S., McClean, F., Guerreiro, S., Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Rustemeier, E., & Fowler, H. J. (2021). Quality control

- of a global hourly rainfall dataset. *Environmental Modelling & Software*, 144, 105169. <https://doi.org/10.1016/j.envsoft.2021.105169>
- Molnar, P., & Burlando, P. (2005). Preservation of rainfall properties in stochastic disaggregation by a simple random cascade model. *Atmospheric Research*, 77(1-4), 137–151. <https://doi.org/10.1016/j.atmosres.2004.10.024>
- Ochoa-Rodriguez, S., Wang, L.-P., Gires, A., Pina, R. D., Reinoso-Rondinel, R., Bruni, G., Ichiba, A., Gaitan, S., Cristiano, E., van Assel, J., Kroll, S., Murlà-Tuyls, D., Tisserand, B., Schertzer, D., Tchiguirinskaia, I., Onof, C., Willems, P., & ten Veldhuis, M.-C. (2015). Impact of spatial and temporal resolution of rainfall inputs on urban hydrodynamic modelling outputs: A multi-catchment investigation. *Journal of Hydrology*, 531(Part 2), 389–407. <https://doi.org/10.1016/j.jhydrol.2015.05.035>
- Onof, C., & Wheater, H. S. (1993). Modelling of British rainfall using a random parameter Bartlett-Lewis Rectangular Pulse Model. *Journal of Hydrology*, 149(1–4), 67–95. [https://doi.org/10.1016/0022-1694\(93\)90100-N](https://doi.org/10.1016/0022-1694(93)90100-N)
- Poschlod, B., Hodnebrog, Ø., Wood, R.R, Alterskjær, K., Ludwig, R., Myhre, G., & Sillmann, J. (2018). Comparison and Evaluation of Statistical Rainfall Disaggregation and High-Resolution Dynamical Downscaling over Complex Terrain. *Journal of Hydrometeorology*, 19(12), 1973-1982. <https://doi.org/10.1175/JHM-D-18-0132.1>
- Pui, A., Sharma, A., Mehrotra, R., Sivakumar, B., & Jeremiah, E. (2012). A comparison of alternatives for daily to sub-daily rainfall disaggregation. *Journal of Hydrology*, 470-471, 138–157. <https://doi.org/10.1016/j.jhydrol.2012.08.041>
- Qin, X., & Dai, C. (2024). Investigation of rainfall disaggregation with flexible timescales based on point process models. *Journal of Hydrology*, 634, 131101. <https://doi.org/10.1016/j.jhydrol.2024.131101>
- Rodriguez-Iturbe, I., Cox, D. R., & Isham, V. (1987). Some models for rainfall based on stochastic point processes. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 410(1839), 269–288. <https://doi.org/10.1098/rspa.1987.0039>
- Sharma, A., & Mehrotra, R. (2010). Rainfall generation. In M. Gebremichael (Ed.), *Rainfall: state of the science* (pp. 215-246). Washington: American Geophysical Union. <http://dx.doi.org/10.1029/2010GM000973>
- Sharma, A., & Srikanthan, R. (2006). Continuous rainfall simulation: a nonparametric alternative. In 30th Hydrology and Water Resources Symposium. Launceston, Tasmania: Conference Design.
- Socolofsky, S., Adams, E. E., & Entekhabi, D. (2001). Disaggregation of daily rainfall for continuous watershed modeling. *Journal of Hydrologic Engineering*, 6(1), 300-309. [https://doi.org/10.1061/\(ASCE\)1084-0699\(2001\)6:4\(300\)](https://doi.org/10.1061/(ASCE)1084-0699(2001)6:4(300))
- Taschetto, A. S., & England, M. H. (2009). An analysis of late twentieth century trends in Australian rainfall. *International Journal of Climatology*, 29(6), 791–807. <https://doi.org/10.1002/joc.1736>
- Vicuña, S., Gironás, J., Meza, F.J., Cruzat, M.L., Jelinek, M., Bustos, E., Poblete, D., & Bambach, N. (2013). Exploring possible connections between hydrological extreme events and climate change in central south Chile. *Hydrological Sciences Journal*, 58(8), 1598-1619. <https://doi.org/10.1080/02626667.2013.840380>
- Viglione, A., & Blöschl, G. (2009). On the role of storm duration in the mapping of rainfall to flood return periods. *Hydrology and Earth System Sciences*, 13(2), 205–216. <https://doi.org/10.5194/hess-13-205-2009>
- Woldemeskel, F. M., Sharma, A., Mehrotra, R., & Westra, S. (2016). Constraining continuous rainfall simulations for derived design flood estimation. *Journal of Hydrology*, 542, 581–588. <https://doi.org/10.1016/j.jhydrol.2016.09.028>
- Zhang, X., Alexander, L., Hegerl, G. C., Jones, P., Klein Tank, A., Peterson, T. C., Trewin, B., & Zwiers, F. W. (2011). Indices for monitoring changes in extremes based on daily temperature and precipitation data. *WIREs Climate Change*, 2(6), 851–870. <https://doi.org/10.1002/wcc.147>

# ANEXOS



**Anexo A1.** Estaciones seleccionadas. (a) Ubicación y precipitación media mensual, (b) Series de precipitación anual.

## Anexo A2. Estaciones seleccionadas

Estación	Latitud	Longitud	Faltantes (%)	Inicio	Término
Altheim (Alemania)	48.15	9.46	0.001	Ene 2005	Dic 2015
Hikone (Japón)	35.27	136.24	0.006	Ene 1976	Dic 2008
Oregon (Estados Unidos)	44.13	-123.22	0.273	Ene 1950	Dic 2009
Pathhead (Reino Unido)	55.85	-2.99	0.156	Ene 1994	Dic 2014
Quinta Normal (Chile)	-33.44	-70.68	0.801	Ene 1917	Dic 1960