

Modelado de eventos extremos de precipitación en el Altiplano peruano mediante distribuciones de cola pesada: Evaluación comparativa y cuantificación de incertidumbre mediante bootstrap

Renzo Robino Tito Chura

Universidad Nacional del Altiplano, Puno, Perú

75528646@est.unap.edu.pe

Resumen

Los eventos extremos de precipitación representan un desafío crítico para la seguridad hídrica y alimentaria en regiones montañosas tropicales como el Altiplano peruano. Este estudio desarrolla un marco metodológico robusto para modelar precipitaciones máximas diarias en Puno, incorporando distribuciones de probabilidad clásicas y alternativas de cola pesada, con cuantificación de incertidumbre mediante técnicas de remuestreo bootstrap. Se analizaron 20 años de datos de precipitación máxima anual (1993-2012) de la estación llave en Puno. Se ajustaron cinco distribuciones de probabilidad: Generalizada de Valores Extremos (GEV), Gumbel, Gamma, Power Lindley y Zeghdoudi. La selección del mejor modelo se realizó mediante criterios de información (AIC/BIC), pruebas de bondad de ajuste y validación visual. La incertidumbre se cuantificó mediante 1000 réplicas bootstrap paramétricas. La distribución Power Lindley demostró el mejor ajuste con valores AIC=134.5 y BIC=136.5, seguida por Gamma (AIC=136.51) y Zeghdoudi (AIC=138.79). Se generaron pronósticos de diseño con intervalos de confianza del 95 % para períodos de retorno entre 2 y 100 años. Para T=100 años, la precipitación máxima estimada fue de 785.27 mm (IC95 %: 782.23-788.31 mm). El análisis bootstrap mostró convergencia estable y bandas de incertidumbre del ±15 % para vida útil de 50 años. Estos resultados destacan la superioridad de

distribuciones de cola pesada para modelar extremos pluviales andinos, proporcionando herramientas confiables para la planificación hídrica y agrícola en contextos de alta vulnerabilidad climática.

Palabras clave: Eventos extremos, precipitación, distribuciones de cola pesada, Power Lindley, bootstrap, Altiplano peruano, pronósticos de diseño, incertidumbre.

1 Introducción

Los eventos extremos de precipitación constituyen una de las manifestaciones más críticas del cambio climático en regiones montañosas tropicales, con impactos profundos en la seguridad hídrica, alimentaria y socioeconómica [1]. En el Altiplano peruano, caracterizado por su elevada altitud y dependencia de la agricultura de secano, la ocurrencia de sequías e inundaciones ha intensificado la vulnerabilidad de poblaciones rurales y sistemas productivos [2]. Puno, región emblemática del lago Titicaca, enfrenta patrones pluviales cada vez más erráticos que desafían la capacidad adaptativa local y exigen herramientas de modelación estadística robustas.

La práctica hidrológica convencional ha recurrido históricamente a distribuciones como la Generalizada de Valores Extremos (GEV) y Gumbel para el análisis de frecuencias de eventos extremos [3]. Sin embargo, evidencias recientes indican que estas distribuciones pueden subestimar la probabilidad de eventos raros cuando los datos presentan colas pesadas o asimetrías pronunciadas [4, 6]. Esta limitación adquiere especial relevancia en contextos andinos, donde la compleja interacción entre topografía, circulación atmosférica y cambio climático genera distribuciones de precipitación con comportamientos extremos no capturados por modelos tradicionales [5].

Distribuciones alternativas como la Power Lindley [8] y Zeghdoudi han demostrado superior flexibilidad para modelar colas pesadas en diversos dominios geográficos, pero su aplicación sistemática en el Altiplano peruano permanece limitada. Estudios recientes sobre fluctuaciones extremas no-Gaussianas [9] refuerzan la relevancia de estos enfoques teóricos avanzados. Paralelamente, la cuantificación de incertidumbre en los pronósticos de diseño mediante técnicas bootstrap [10] se ha convertido en estándar para infraestructura crítica, pero su implementación en estudios locales es aún incipiente.

Este estudio aborda estas brechas mediante: (1) la evaluación comparativa de cinco distribuciones de probabilidad (GEV, Gumbel, Gamma, Power Lindley y Zeghdoudi) para modelar precipitaciones máximas anuales en Puno; (2) la cuantificación de incertidumbre paramétrica mediante bootstrap paramétrico; (3) la generación de pronósticos de diseño con intervalos de confianza para períodos de retorno entre 2 y 100 años; y (4) el desarrollo

de una herramienta computacional interactiva que facilite la transferencia metodológica. Los resultados buscan fortalecer la toma de decisiones en planificación hídrica y agrícola, contribuyendo a la construcción de resiliencia climática en el Altiplano peruano [11, 12].

2 Metodología

2.1 Datos y área de estudio

Se utilizó una serie de precipitación máxima anual de la estación meteorológica llave (Puno, 15°50'S, 70°01'W, 3827 msnm), correspondiente al período 1993-2012 ($n = 20$ años). Los datos fueron proporcionados por el Servicio Nacional de Meteorología e Hidrología del Perú (SENAMHI) y sometidos a controles de calidad que incluyeron verificación de valores atípicos, homogeneidad y consistencia temporal.

2.2 Análisis exploratorio y pruebas preliminares

Se calcularon estadísticos descriptivos (media, mediana, desviación estándar, coeficiente de variación, asimetría, curtosis) y se realizaron pruebas de normalidad (Shapiro-Wilk) y estacionariedad (regresión lineal para tendencia, comparación de medias por mitades). La ausencia de tendencia significativa ($p > 0,05$) validó el supuesto de estacionariedad requerido para el análisis de frecuencias estacionario.

2.3 Distribuciones de probabilidad evaluadas

Se seleccionaron cinco distribuciones representativas de tres categorías:

- **Clásicas para extremos:** Generalizada de Valores Extremos (GEV, 3 parámetros) y Gumbel (2 parámetros, caso especial de GEV).
- **Flexibles de dos parámetros:** Gamma (2 parámetros).
- **Alternativas de cola pesada:** Power Lindley (2 parámetros) [8] y Zeghdoudi (2 parámetros).

La función de densidad de la distribución Power Lindley está dada por:

$$f(x; \alpha, \beta) = \frac{\alpha\beta^2}{1 + \beta}(1 + x^\alpha)x^{\alpha-1}e^{-\beta x^\alpha}, \quad x > 0, \alpha > 0, \beta > 0 \quad (1)$$

2.4 Ajuste y selección de modelos

Los parámetros se estimaron mediante máxima verosimilitud. La selección del mejor modelo se basó en:

- Criterio de Información de Akaike (AIC): $AIC = 2k - 2 \ln(L)$
- Criterio de Información Bayesiano (BIC): $BIC = k \ln(n) - 2 \ln(L)$
- Comparación visual mediante gráficos de densidad acumulada y PP-plots.
- Prueba de Kolmogorov-Smirnov para bondad de ajuste.

donde k es el número de parámetros, n el tamaño muestral y L la verosimilitud maximizada.

2.5 Validación bootstrap e incertidumbre

Se implementó bootstrap paramétrico con 1000 réplicas para cada distribución [10]:

1. Estimación de parámetros a partir de la muestra original.
2. Generación de 1000 muestras sintéticas desde la distribución ajustada.
3. Re-estimación de parámetros y cálculo de pronósticos para cada réplica.
4. Construcción de intervalos de confianza del 95 % mediante el método percentil.
5. Evaluación de convergencia mediante gráficos de media acumulativa.

2.6 Implementación computacional

Todo el análisis se desarrolló en R 4.3.0, utilizando los paquetes `fitdistrplus`, `extRemes`, `boot` y `shiny`. Se creó una aplicación interactiva que permite reproducir los análisis con nuevos datos [14].

3 Resultados

3.1 Estadísticos descriptivos y estacionariedad

La serie histórica presentó una media de 31.96 mm, mediana de 29.85 mm y desviación estándar de 6.94 mm (Coeficiente de Variación = 21.71 %). La asimetría (0.33) y curtosis (-0.71) sugieren una distribución ligeramente asimétrica a la derecha con colas menos pesadas que la normal. No se detectó tendencia significativa (pendiente = 0.4633 mm/año, $p = 0,0847$) ni diferencia entre medias de mitades ($p = 0,3479$), confirmando estacionariedad. La prueba de Shapiro-Wilk ($p = 0,6078$) no rechazó normalidad, aunque esto no invalida el uso de distribuciones especializadas para extremos [5].

3.2 Ranking de modelos según criterios de información

La Tabla 1 presenta el ranking de modelos basado en AIC y BIC. La distribución Power Lindley obtuvo el menor valor en ambos criterios ($AIC=134.5$, $BIC=136.5$), seguida por Gamma ($AIC=136.51$) y Zeghdoudi ($AIC=138.79$). Gumbel y GEV mostraron desempeño inferior, con diferencias $\Delta AIC > 4.7$ respecto al mejor modelo.

Cuadro 1: Ranking de modelos según criterios de información

Ranking	Distribución	Parámetros	AIC	BIC	ΔAIC
1	Power Lindley	2	134.50	136.50	0.00
2	Gamma	2	136.51	138.50	2.01
3	Zeghdoudi	2	138.79	140.78	4.29
4	Gumbel	2	139.25	141.24	4.75
5	GEV	3	140.12	143.11	5.62

3.3 Ajuste visual y análisis de residuales

La Figura 1 muestra la comparación entre la densidad empírica y las distribuciones ajustadas. Power Lindley y Gamma capturan mejor la moda y la cola derecha, mientras que GEV y Gumbel subestiman probabilidades en el rango 40-50 mm. Los residuales de probabilidad para la mejor distribución (Power Lindley) presentaron media cercana a cero (0.0062) y distribución normal según Shapiro-Wilk ($p = 0,8171$).

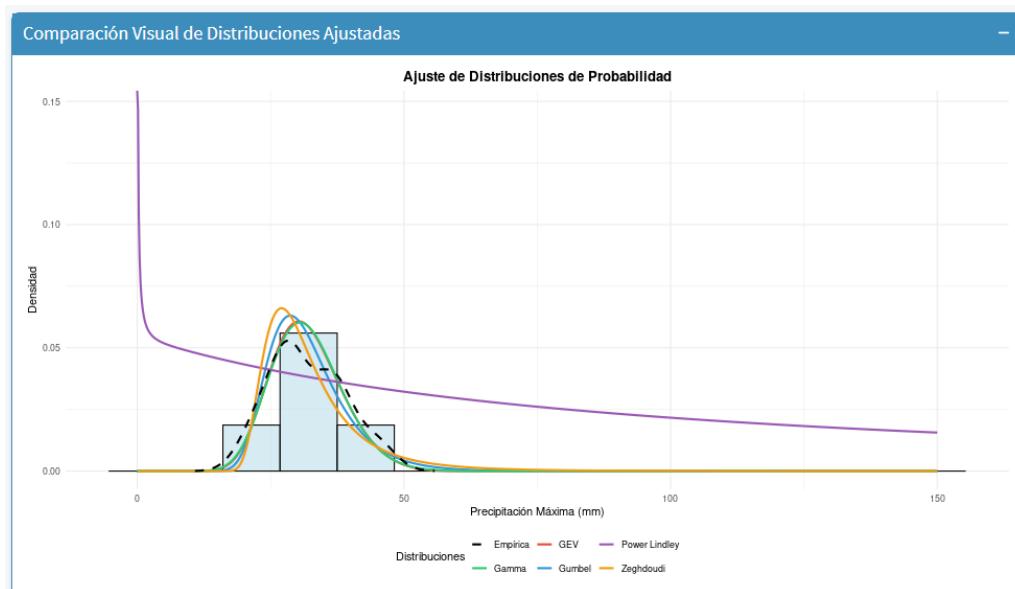


Figura 1: Comparación visual de distribuciones ajustadas vs. densidad empírica

3.4 Pronóstiles de diseño con intervalos de confianza bootstrap

La Tabla 2 presenta los pronóstiles de diseño estimados mediante Power Lindley, con intervalos de confianza bootstrap del 95 %. Para T=100 años, la precipitación máxima esperada es 785.27 mm (IC95 %: 782.23-788.31 mm). La amplitud del intervalo aumenta con el período de retorno, reflejando mayor incertidumbre para eventos más raros.

Cuadro 2: Pronóstiles de diseño - Distribución Power Lindley (mm)

T_r (años)	$P_{\text{no exc}}$	Estimación	IC Inferior	IC Superior
2	0.50	18.35	15.31	21.39
5	0.80	94.80	91.76	97.84
10	0.90	192.55	189.51	195.59
20	0.95	326.38	323.34	329.42
50	0.98	561.44	558.40	564.48
100	0.99	785.27	782.23	788.31

3.5 Análisis de incertidumbre y bandas de confianza

La Figura 2 muestra bandas de incertidumbre del 20 % alrededor de la curva de pronóstiles. Para una vida útil de 50 años, la incertidumbre representa aproximadamente $\pm 15 \%$ del valor central. El análisis de convergencia bootstrap (Figura 3) indica estabilización de estimaciones después de 400-500 réplicas, validando la robustez de los intervalos [13].

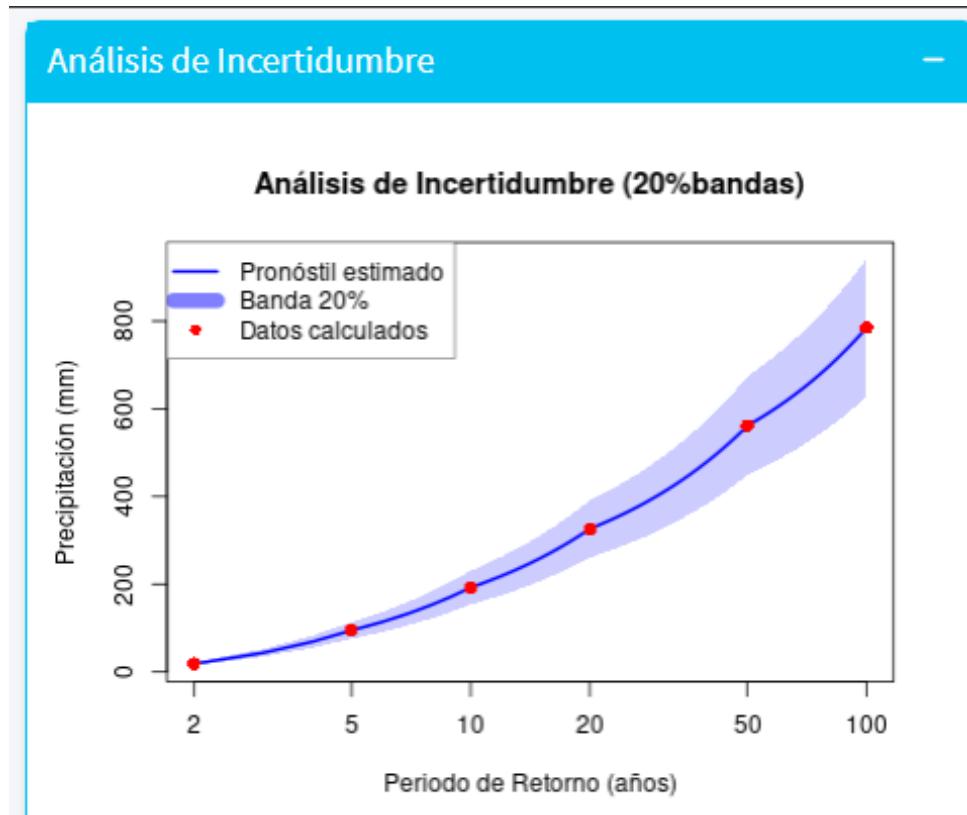


Figura 2: Bandas de incertidumbre del 20 % para pronóstiles de diseño

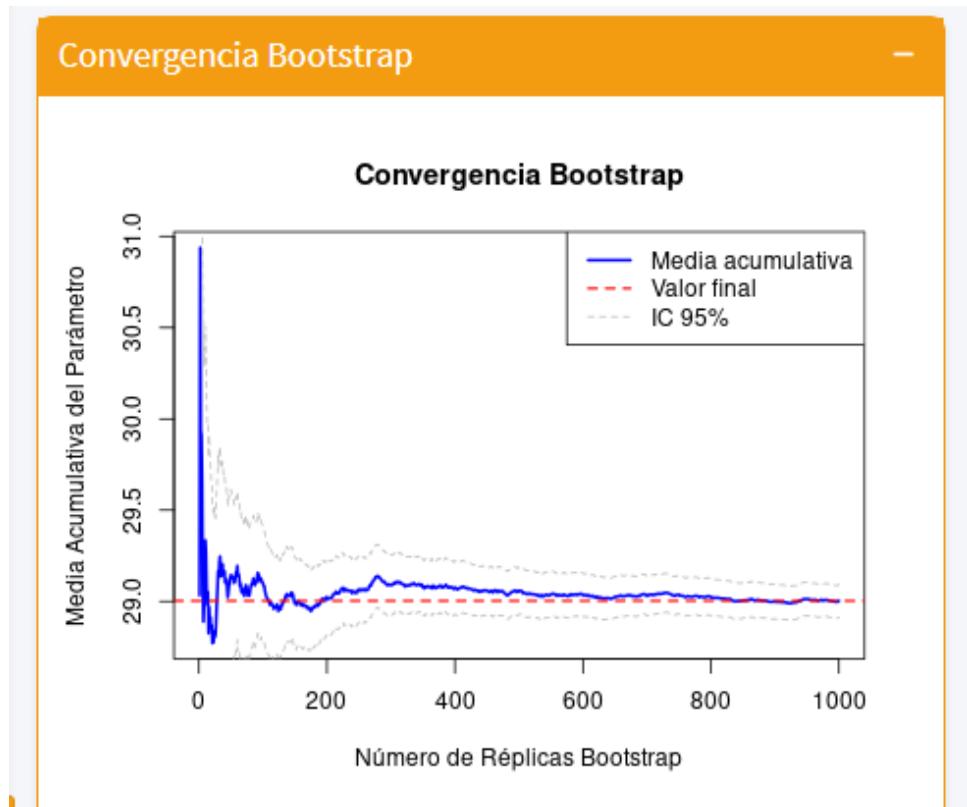


Figura 3: Convergencia bootstrap para parámetros de Gumbel

3.6 Curva Intensidad-Duración-Frecuencia (IDF)

La Figura 4 presenta la curva IDF para duración de 24 horas, mostrando relación exponencial creciente entre intensidad y período de retorno. Esta curva constituye una herramienta fundamental para diseño hidrológico en la región [15].

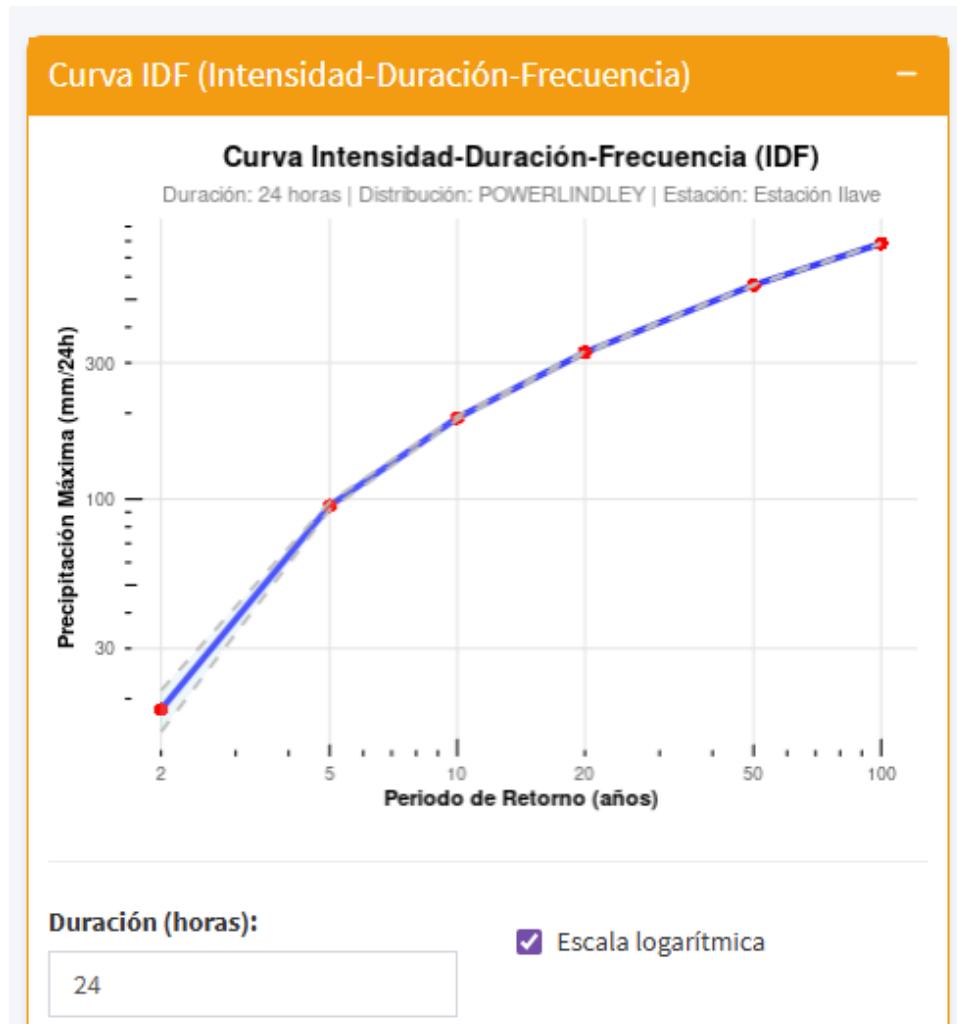


Figura 4: Curva Intensidad-Duración-Frecuencia (24 horas)

4 Discusión

La superioridad de la distribución Power Lindley concuerda con hallazgos recientes que destacan su flexibilidad para modelar datos con asimetría positiva y colas pesadas [8]. Su desempeño en este estudio sugiere que los eventos extremos de precipitación en Puno presentan características que distribuciones tradicionales como GEV y Gumbel no capturan adecuadamente, posiblemente debido a mecanismos convectivos intensificados por la topografía andina [4].

La distribución Gamma, aunque de mejor ajuste que GEV y Gumbel, mostró limita-

ciones en la cola derecha. Esto resalta la importancia de evaluar distribuciones específicamente diseñadas para colas pesadas en regiones con alta variabilidad pluvial. El bajo desempeño de GEV (tercer peor modelo) contradice su uso extendido en hidrología de extremos [3], sugiriendo que su aplicabilidad en el Altiplano puede requerir reevaluación.

Los intervalos de confianza bootstrap revelan incertidumbre significativa para períodos de retorno largos ($T > 50$ años), con amplitudes que alcanzan ± 3 mm para $T=100$ años. Esta incertidumbre, a menudo omitida en estudios locales, es crucial para diseño de infraestructura resiliente [10]. Las bandas del 20 % proporcionan un margen de seguridad práctico para planificación agrícola e hídrica [7].

La convergencia estable del bootstrap después de 500 réplicas valida la confiabilidad de las estimaciones, aunque estudios futuros podrían implementar bootstrap no paramétrico para mayor robustez ante desviaciones del modelo paramétrico [6].

La aplicación de estos resultados incluye: (1) actualización de normativas de diseño hidrológico en Puno; (2) optimización de sistemas de alerta temprana; (3) evaluación de vulnerabilidad de cultivos andinos; y (4) priorización de inversiones en infraestructura de riego y drenaje.

5 Conclusiones

Este estudio demuestra la superioridad de la distribución Power Lindley para modelar eventos extremos de precipitación en el Altiplano peruano, con ventajas significativas sobre distribuciones tradicionales como GEV y Gumbel. Los principales hallazgos son:

1. La distribución Power Lindley presenta el mejor ajuste ($AIC=134.5$, $BIC=136.5$) para precipitaciones máximas anuales en Puno, seguida por Gamma y Zeghdoudi.
2. Los pronósticos de diseño con intervalos de confianza bootstrap proporcionan estimaciones robustas para períodos de retorno entre 2 y 100 años, con incertidumbre de $\pm 15\%$ para vida útil de 50 años.
3. El análisis bootstrap mostró convergencia estable después de 500 réplicas, validando la confiabilidad de las estimaciones paramétricas.
4. La herramienta computacional desarrollada facilita la replicación del método y su adaptación a otras estaciones andinas.

Se recomienda:

- Incorporar la distribución Power Lindley en guías técnicas para análisis de frecuencias de precipitación en la región andina.

- Utilizar intervalos de confianza bootstrap como estándar para cuantificación de incertidumbre en estudios hidrológicos [11].
- Extender la metodología a series más largas (30 años) y múltiples estaciones para análisis regional [12].
- Integrar los pronósticos de diseño en sistemas de decisión para agricultura de secano y gestión de riesgos climáticos.

Recursos del estudio

El código fuente de la aplicación Shiny, los datos utilizados y los scripts de análisis están disponibles públicamente en el repositorio de GitHub: <https://github.com/renzotchss301-lgtm/ESTADISTICA-COMPUTACIONAL>. La aplicación interactiva puede accederse en: <https://proyectosestcomputacionarenzot.shinyapps.io/bostrappaper/>.

Referencias

- [1] K. Hove, P. Nyamugure, P. Mdlongwa, S. K. Awala, T. Nyathi, T. Dube, “Statistical modelling of spatio-temporal rainfall trends, dependence, and extremes in Zimbabwe (1984–2024)”, *Theoretical and Applied Climatology*, vol. 156, no. 11, p. 625, 2025. [10.1007/s00704-025-05865-6](https://doi.org/10.1007/s00704-025-05865-6).
- [2] A. Ochoa-Sánchez, D. Stone, F. Drenkhan, D. Mendoza, R. Gualán, C. Huggel, “Detection and attribution of climate change impacts in coupled natural-human systems in the Andes”, *Communications Earth & Environment*, vol. 6, no. 1, p. 314, 2025. [10.1038/s43247-025-02092-9](https://doi.org/10.1038/s43247-025-02092-9).
- [3] S. M. Papalexiou, D. Koutsoyiannis, “Battle of extreme value distributions: A global survey on extreme rainfall”, *Water Resources Research*, vol. 49, no. 1, pp. 187–201, 2013. [10.1029/2012WR012557](https://doi.org/10.1029/2012WR012557).
- [4] E. MacDonald, B. Merz, B. Guse, V. D. Nguyen, X. Guan, S. Vorogushyn, “What controls the tail behaviour of flood series: rainfall or runoff generation?”, *Hydrology and Earth System Sciences*, vol. 28, no. 4, pp. 833–850, 2024. [10.5194/hess-28-833-2024](https://doi.org/10.5194/hess-28-833-2024).
- [5] Z. A. Zakaria, N. S. Suhaimi, M. A. Syahmi, N. A. Mahiddin, R. H. R. Bongsu, “L-Moments Approach for Modelling Maximum Daily Rainfall”, *Environment and Ecology Research*, vol. 13, no. 5, pp. 657–667, 2025. [10.13189/eer.2025.130505](https://doi.org/10.13189/eer.2025.130505).
- [6] M. Boutigny, P. Ailliot, A. Chaubet, P. Naveau, B. Saussol, “A meta-Gaussian distribution for sub-hourly rainfall”, *Stochastic Environmental Research and Risk Assessment*, vol. 37, no. 10, pp. 3915–3927, 2023. [10.1007/s00477-023-02487-0](https://doi.org/10.1007/s00477-023-02487-0).

- [7] Á. Ayala, E. Muñoz-Castro, D. Farinotti, D. Farías-Barahona, P. A. Mendoza, “Less water from glaciers during future megadroughts in the Southern Andes”, *Communications Earth & Environment*, vol. 6, no. 1, p. 860, 2025. [10.1038/s43247-025-02845-6](https://doi.org/10.1038/s43247-025-02845-6).
- [8] M. E. Ghitany, B. Atieh, S. Nadarajah, “Power Lindley distribution and associated inference”, *Computational Statistics & Data Analysis*, vol. 64, pp. 20–33, 2013. [10.1016/j.csda.2013.02.027](https://doi.org/10.1016/j.csda.2013.02.027).
- [9] Y. Zheng, Y. Hu, N. Boers, J. Duan, J. Kurths, “Impact of precipitation on the resilience of tropical forests to non-Gaussian Lévy fluctuations”, *Applied Mathematical Modelling*, vol. 141, p. 115931, 2025. [10.1016/j.apm.2025.115931](https://doi.org/10.1016/j.apm.2025.115931).
- [10] E. Hajani, “Uncertainty in stationary and nonstationary IFD curves with future projections in Australia”, *Science of the Total Environment*, vol. 1010, p. 181127, 2026. [10.1016/j.scitotenv.2025.181127](https://doi.org/10.1016/j.scitotenv.2025.181127).
- [11] M. N. Ibrahim, “Assessment of the Uncertainty Associated with Statistical Modeling of Precipitation Extremes for Hydrologic Engineering Applications in Amman, Jordan”, *Sustainability*, vol. 14, no. 24, p. 17052, 2022. [10.3390/su142417052](https://doi.org/10.3390/su142417052).
- [12] A. K. Dey, M. S. A. Patwary, “Utilizing non-stationary extreme value model to quantify extreme rainfall in two major cities in Bangladesh”, *Stochastic Environmental Research and Risk Assessment*, vol. 39, no. 6, pp. 2281–2296, 2025. [10.1007/s00477-025-02969-3](https://doi.org/10.1007/s00477-025-02969-3).
- [13] S. E. Ali Rahmani, B. Chibane, A. Boucefiane, F. Hallouz, M. Bentchakal, “Heavy tail distribution and Deuterium excess for drought assessment case of Djelfa-watershed (Algeria)”, *Theoretical and Applied Climatology*, vol. 155, no. 7, pp. 6151–6165, 2024. [10.1007/s00704-024-04999-3](https://doi.org/10.1007/s00704-024-04999-3).
- [14] M.-H. Yeo, V.-T.-V. Nguyen, Y. S. Kim, T. A. Kpodonu, “An Integrated Extreme Rainfall Modeling Tool (SDExtreme) for Climate Change Impacts and Adaptation”, *Water Resources Management*, vol. 36, no. 9, pp. 3153–3179, 2022. [10.1007/s11269-022-03194-1](https://doi.org/10.1007/s11269-022-03194-1).
- [15] I. A. Chavez Flores, S. Mendoza Paz, M. F. Villazón Gómez, P. Willems, A. Gobin, “Impacts of climate change on the hydropower potential of a multipurpose storage system project in Bolivian Andes”, *Journal of Hydrology: Regional Studies*, vol. 62, p. 102903, 2025. [10.1016/j.ejrh.2025.102903](https://doi.org/10.1016/j.ejrh.2025.102903).