

Modelación de la distribución de probabilidad de las precipitaciones máximas en la Estación Meteorológica Salache

Modeling of the probability distribution of maximum rainfall in the Salache Meteorological Station

Patricio Germánico Chávez-Zapata¹

RESUMEN:

El presente artículo establece la mejor distribución probabilística de acuerdo al análisis de frecuencias de lluvias máximas mensuales, utilizando para ello el registro de la Estación Meteorológica Salache, ubicado en el cantón Latacunga, provincia de Cotopaxi. El análisis considera las distribuciones Normal, Log-Normal, Gumbel, Log-Gumbel, Pearson Tipo III y Log-Pearson Tipo III; y para determinar la más apropiada se utilizó las pruebas de bondad de ajuste Chi-Cuadrado y la prueba de verificación de Kolmogorov-Smirnov. Los resultados determinaron que la distribución Log-Normal se ajusta más a los datos de precipitación máxima proporcionados por la Estación Meteorológica Salache correspondientes a los años 2020 y 2021.

Palabras claves: *Distribución probabilística, análisis de frecuencias, precipitaciones máximas.*

Recibido 12 de mayo de 2022; revisión aceptada 28 de octubre de 2022

ABSTRACT:

This article establishes the best probabilistic distribution according to the analysis of maximum monthly rainfall frequencies, using the data record of the Salache Meteorological Station, located in the Latacunga canton, Cotopaxi province. The analysis considers the Normal, Log-Normal, Gumbel, Log-Gumbel, Pearson Type III and Log-Pearson Type III distributions; and to determine the most appropriate, the Chi-Square goodness-of-fit tests and the Kolmogorov Smirnov verification test were used. The results determined that the Log-Normal distribution is more in line with the maximum precipitation data provided by the Salache Meteorological Station corresponding to the years 2020 and 2021.

Keywords: *Probabilistic distribution, frequency analysis, maximum rainfall.*

¹ Agencia Digital MindWareNet, Latacunga, Cotopaxi, Ecuador, patriciochavez@mindwarenet.com

1. INTRODUCCIÓN

Una de las variables meteorológicas más importantes que se estudia en hidrometeorología es la precipitación máxima a través de un régimen pluviométrico que interviene sobre el clima tanto en el aspecto espacial como temporal, para ello se analiza las frecuencias de este tipo de eventos extremos para determinar la probabilidad de ocurrencia de este tipo de eventos.

Con frecuencia, el periodo de retorno de los eventos excede el periodo de las observaciones y consecuentemente se deben efectuar extrapolaciones a partir de los valores registrados. Una técnica más objetiva es encontrar la distribución de probabilidad teórica que se ajuste mejor a los datos medidos y usar esta función para la extrapolación [1].

De aquí surge la necesidad de seleccionar el modelo de distribución probabilística que describa adecuadamente el comportamiento de los datos de precipitación máxima. El análisis estadístico de series pluviométricas aun cuando por su naturaleza tienen un alto grado de irregularidad a través del tiempo permite identificar periodos de mayor o menor precipitación.

La modelación de eventos extremos está relacionada con la necesidad básica del científico empírico, dado que, a través de esas construcciones teóricas, los modelos, podía experimentar sobre aquello que la realidad no le permitía. Por otra parte, un modelo resulta extremadamente útil, siempre que se corresponda con la realidad que pretende representar o predecir, de manera que ponga de relieve las propiedades más importantes del mundo que nos rodea, aunque sea a costa de la abstracción y la consecuente simplificación que implica todo modelo [2].

Para estimar este tipo de eventos a menudo se recurre al análisis de frecuencias, el cual relaciona la magnitud de los eventos extremos con la probabilidad de ocurrencia de los mismos, por medio de distribuciones de probabilidad que en comportamientos hidrológicos aquellos que describen de mejor manera los datos observados son: Normal, Log-Normal, Pearson, Gumbel, Log-Gumbel, Log-Pearson tipo III.

2. METODOLOGÍA

Procedimiento para la selección de una distribución teórica de probabilidades

En general el procedimiento se basa en seleccionar una distribución teórica, estimar parámetros de acuerdo al registro de datos observados, efectuar un test de bondad de ajuste y de acuerdo a esta prueba aceptar la distribución teórica analizada o caso contrario rechazarla y probar una nueva distribución [3].

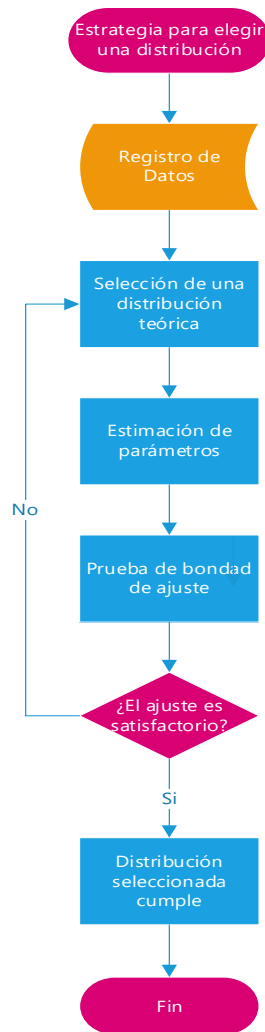


Figura 1. Flujograma de procedimiento para seleccionar la mejor distribución teórica.

El grado de magnitud de un fenómeno extremo está relacionado de forma inversa con su frecuencia de ocurrencia (periodicidad), las precipitaciones muy intensas ocurren con una frecuencia menor que las moderadas o débiles [4].

Para el análisis de frecuencias, se arregla la serie en orden decreciente y cada valor tendrá el número de orden m (m va de 1 a n) siendo n el tamaño de la muestra, el número de meses como una serie natural, se calcula la frecuencia observada a través de la relación de Weibull:

$$P = \frac{m}{n + 1} \quad T = \frac{n + 1}{m} \quad (1)$$

En donde:

P: probabilidad empírica o frecuencia relativa acumulada. **m**: es el número de orden.

n: es el número de observaciones. **t**: es el periodo de retorno.

Distribuciones Normal y Log-Normal

Entre las principales aplicaciones en hidrología de estas distribuciones se tiene:

- Ajuste de distribuciones empíricas de variables hidrológicas
- Análisis de errores aleatorios en las observaciones o mediciones hidrológicas
- Análisis comparativo de varias distribuciones teóricas de ajuste en una distribución empírica

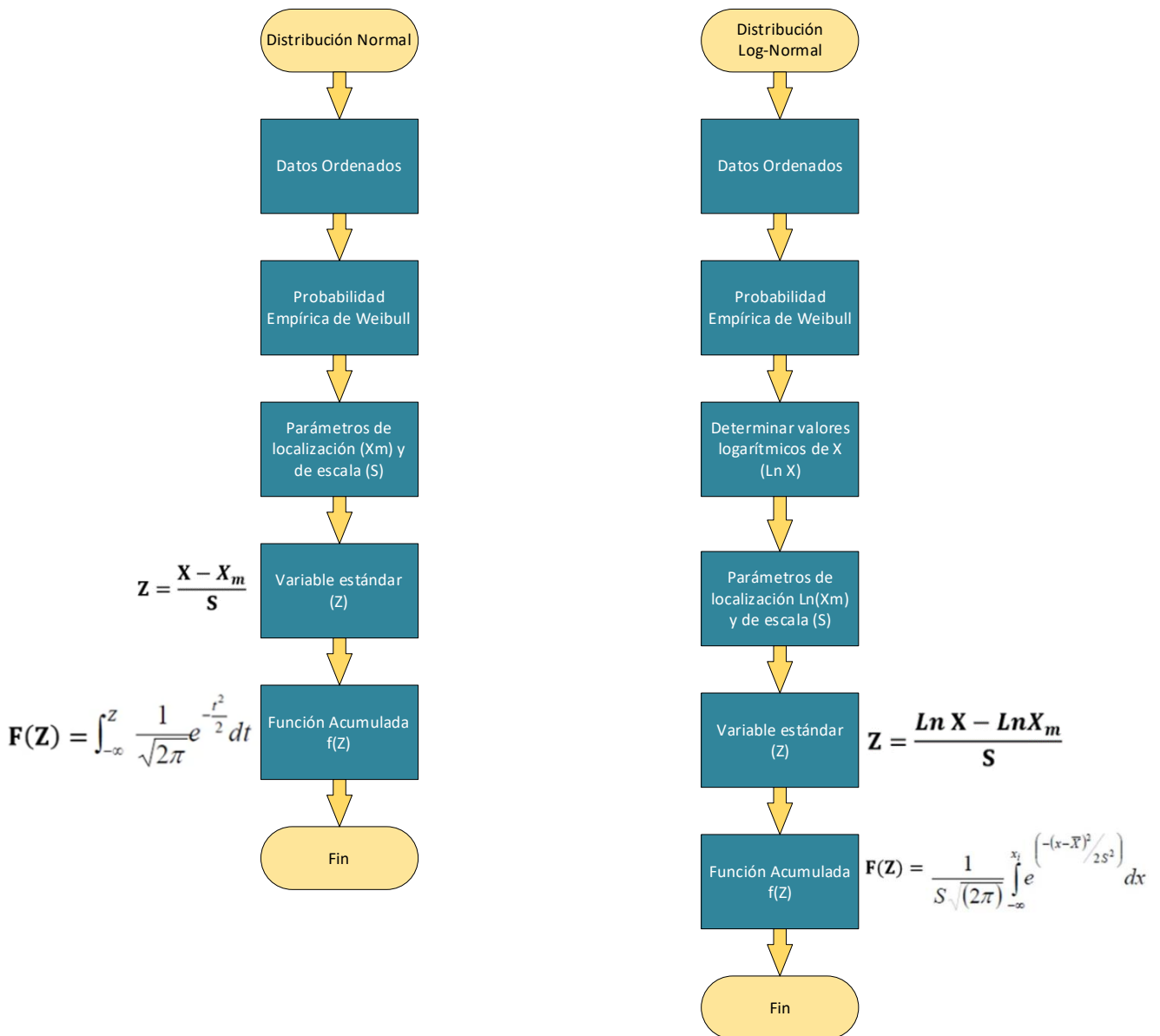


Figura 2. Flujogramas de Distribución Normal y Log-Normal.

Distribuciones Pearson y Log-Pearson Tipo III [5]

Distribución Pearson tipo III: También se le denomina distribución gamma de 3 parámetros (media, desviación estándar y un parámetro de posición, permite describir la distribución de probabilidad de picos decrecientes de máximos. Distribución Log Pearson tipo III: Esta distribución se utiliza para un análisis de probabilidad de eventos extremos, se basa en transformar los valores extremos en logaritmo de base 10, con el objetivo de disminuir la asimetría de los datos. En hidrología se limita para el tratamiento de datos de picos de inundación. En los flujogramas se puede apreciar que g es el coeficiente de asimetría.

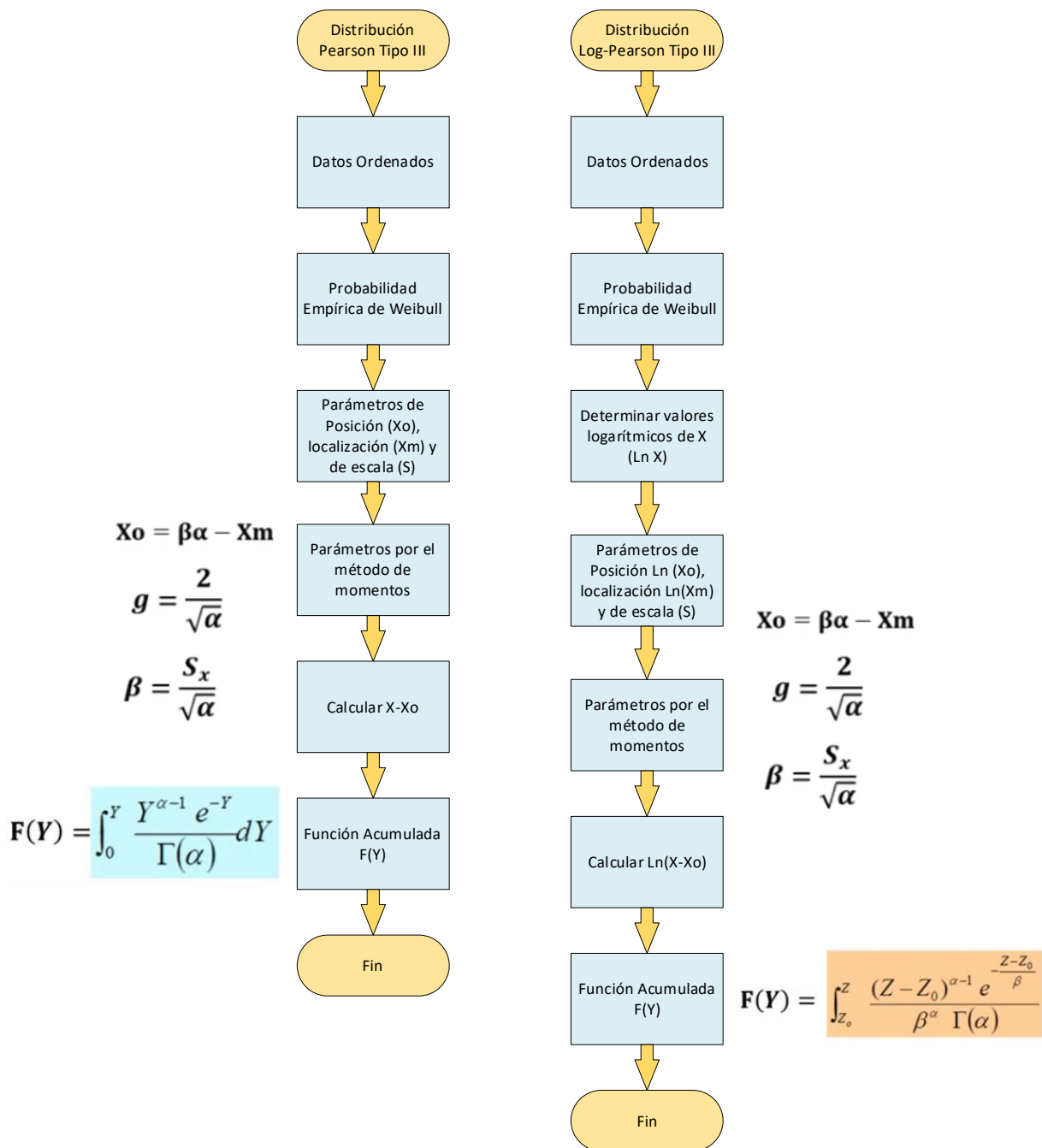


Figura 3. Flujogramas de Distribución Pearson y Log-Pearson Tipo III

Distribuciones Gumbel y Log-Gumbel [6]

Su principal aplicación es para modelar la distribución del máximo o el mínimo, es decir es utilizada para calcular valores extremos.

Generalmente es usada para ajustar a una expresión matemática, las distribuciones empíricas de frecuencia de caudales máximos anuales, precipitaciones máximas anuales.

Las distribuciones de valores extremos seleccionados de conjuntos de muestras de cualquier distribución de probabilidad convergen en una de las tres formas de distribución de valor extremo, como son: Tipo I: Gumbel; Tipo II: Frechet; Tipo III: Weibull

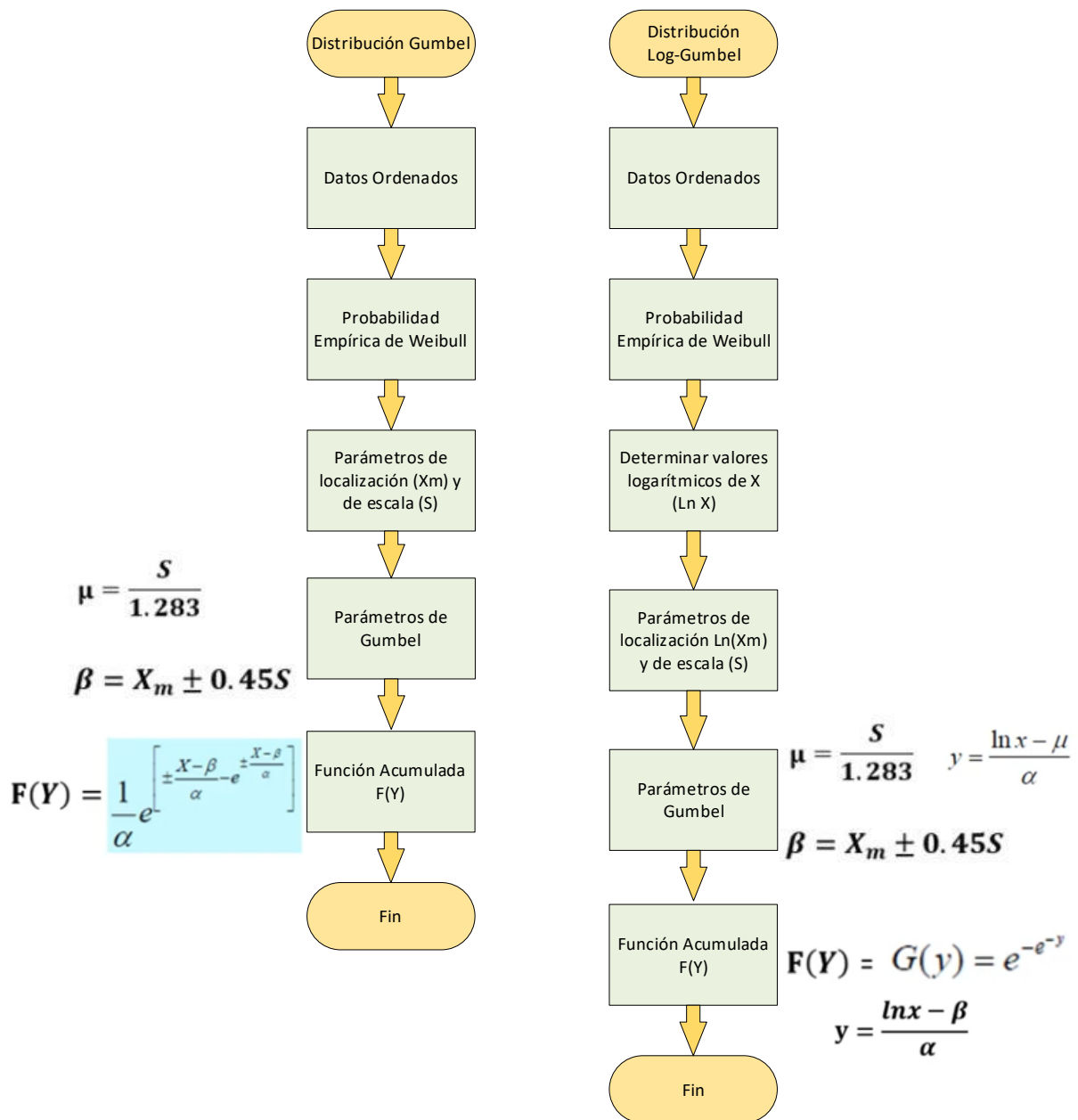


Figura 4. Flujogramas de Distribución Gumbel y Log-Gumbel

En donde: μ y β son los parámetros de Gumbel.

Pruebas de bondad de ajuste

Son pruebas de hipótesis para verificar si los datos observados en una muestra aleatoria se ajustan a un nivel de significancia de acuerdo con una distribución de probabilidad. [7]

$$\text{Hipótesis Nula} \rightarrow H_0: f(x) = f_0(x)$$

$$\text{Hipótesis Alternativa} \rightarrow H_1: f(x) \neq f_0(x)$$

La hipótesis nula indica la distribución propuesta, mientras que la hipótesis alternativa nos indica que la variable en estudio tiene una distribución que no se ajusta a la distribución propuesta.

Para seleccionar la mejor distribución probabilística teórica se puede utilizar tres tipos de estrategias:

- Ajuste gráfico
- Prueba de ajuste Chi-Cuadrado [8]
- Prueba de verificación de Kolmogorov-Smirnov [9]

3. ANÁLISIS DE RESULTADOS

Los resultados obtenidos de acuerdo al análisis efectuado, se presentan en las siguientes tablas y gráficas:

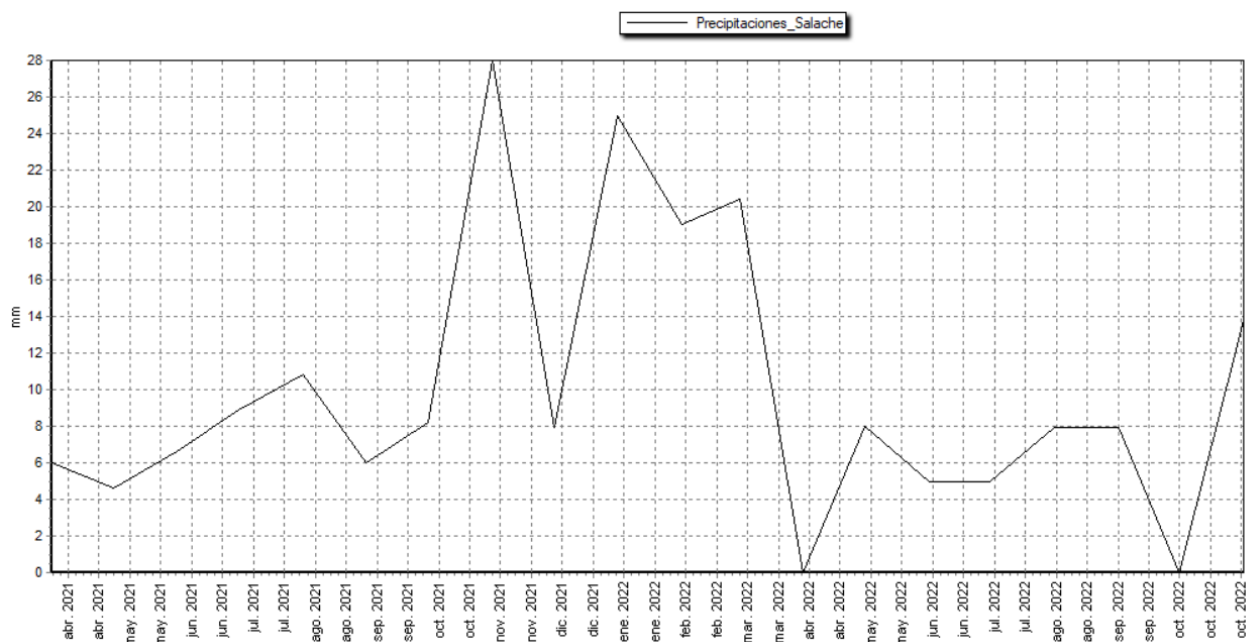


Figura 5. Precipitaciones Máximas Mensuales de la Estación Meteorológica Salache

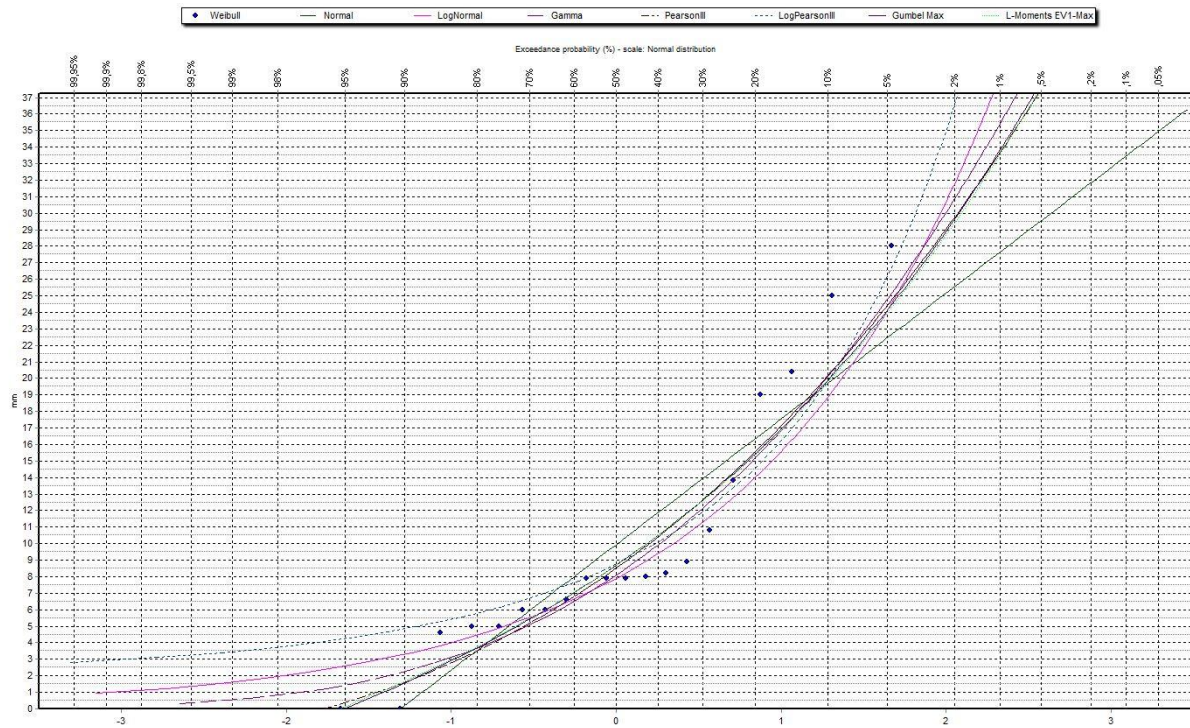


Figura 5. Distribuciones teóricas analizadas.

Tabla 1. Prueba de bondad Chi-Cuadrado.

Prueba X-Square	a=1%	a=5%	a=10%	Alcanzado	Pearson Param.
Normal	REJECT	REJECT	REJECT	0,19%	12,5
LogNormal	ACCEPT	ACCEPT	ACCEPT	13,53%	4
Gamma	ACCEPT	ACCEPT	ACCEPT	10,54%	4,5
Pearson III	ACCEPT	REJECT	REJECT	4,55%	4
Log Pearson III	ACCEPT	ACCEPT	ACCEPT	11,38%	2,5
EV1-Max (Gumbel)	ACCEPT	ACCEPT	ACCEPT	13,53%	4
EV1-Max (Gumbel, L-Moments)	ACCEPT	ACCEPT	ACCEPT	13,53%	4

Tabla 2. Prueba de bondad Kolmogorov-Smirnov.

Prueba Kolmogorov-Smirnov	a=1%	a=5%	a=10%	Alcanzado	DMax
Normal	ACCEPT	ACCEPT	ACCEPT	21,92%	0,22161
LogNormal	ACCEPT	ACCEPT	ACCEPT	96,60%	0,09765
Gamma	ACCEPT	ACCEPT	ACCEPT	85,10%	0,12295
Pearson III	ACCEPT	ACCEPT	ACCEPT	69,27%	0,14559
Log Pearson III	ACCEPT	ACCEPT	ACCEPT	52,71%	0,16782
EV1-Max (Gumbel)	ACCEPT	ACCEPT	ACCEPT	62,11%	0,15509
EV1-Max (Gumbel, L-Moments)	ACCEPT	ACCEPT	ACCEPT	61,61%	0,15576

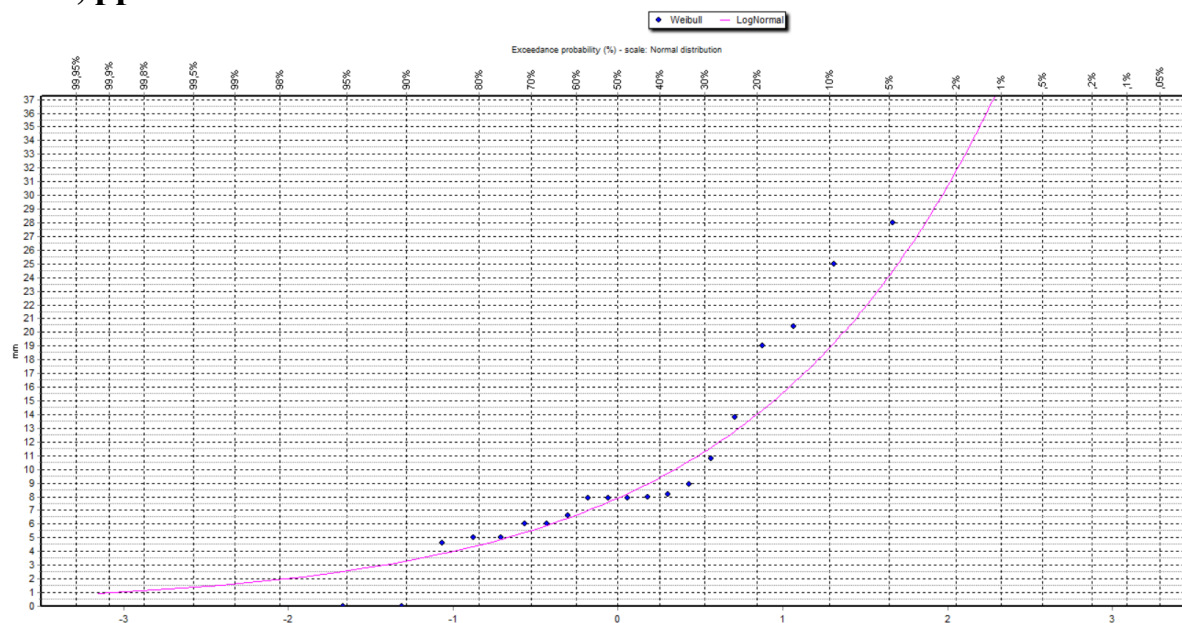


Figura 6. Distribución seleccionada que tiene mejor ajuste a los datos observados.

Los resultados muestran que la distribución de probabilidad Log-Normal tiene el mejor ajuste con los datos observados y proporcionados por la Estación Meteorológica Salache, de acuerdo con la prueba de Kolmogorov-Smirnov en la que se obtiene el valor menor de D_{max} que es la desviación máxima entre la probabilidad (fórmula de Weibull) y la función acumulada de la distribución, para nuestro caso $D_{max} = 0.09765$ alcanzando un 96,60 %, con un nivel de significación de 0.05 que es el más utilizado en estudios hidrológicos.

En la prueba Chi-Cuadrado los resultados lo comparten por igual las distribuciones Log-Normal y Gumbel.

4. CONCLUSIONES

- La distribución de probabilidad Log-Normal tuvo un mejor ajuste a los datos de precipitación máxima mensual de la Estación Meteorológica Salache.
- Es importante considerar factores de nivel de ajuste e incertidumbre de las estimaciones al momento de seleccionar una distribución de probabilidad.
- Cuando varias distribuciones se ajustan satisfactoriamente a los datos, es conveniente utilizar un análisis gráfico para realizar la selección más adecuada.

5. BIBLIOGRAFÍA

- [1] R. Flowers-Cano. y F. Rivera-Trejo, "Evaluación de criterios de selección de modelos probabilísticos", SCIELO, 12 02 2014. [En línea]. Available: http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S2007-24222014000500012. [Último acceso: 22 01 2024]
- [2] Epidat 4, "Ayuda de Distribuciones de Probabilidad", SERGAS, 2 10 2014. [En línea]. Available: https://www.sergas.es/Saude-publica/Documents/1899/Ayuda_Epidat_4_Distribuciones_de_probabilidad_Octubre2014.pdf. [Último acceso: 22 01 2024].
- [3] A. Mayoral y J. Morales, "Distribuciones Continuas", 02 03 2022. [En línea]. Available: <https://bookdown.org/content/944ffa0f-050e-47cb-afaa-4dff15a9ed00/distribuciones-continuas.html>. [Último acceso: 22 01 2024].
- [4] J. Gutiérrez, V. Vargas y R. Rodríguez, "Periodos de retorno de lluvias torrenciales para el estado de Tamaulipas, México", Investigaciones Geográficas, vol. 1, n° 76, pp. 20-33, 2011.
- [5] J. Rivera y O. Penalba, "Distribución de probabilidades de los caudales mensuales en las regiones de cuyo y patagonia (Argentina): Aplicación al monitoreo de sequías hidrológicas", 14 07 2017. [En línea]. Available: http://www.scielo.org.ar/scielo.php?script=sci_arttext&pid=S1850-468X2018000200002&lng=es&nrm=iso. [Último acceso: 22 01 2024].
- [6] N. Johnson y S. Kotz, "Continuous Univariate Distributions", New York: Wiley, 1994.
- [7] C. Pérez, Y. Palacios y J. Herrera, "ANÁLISIS DE PRECIPITACIONES PARA MODELO PREDICTIVO DE LLUVIAS EN ZIPAQUIRÁ, CUNDINAMARCA", SCIELO, 01 03 2022. [En línea]. Available: <http://scielo.sld.cu/pdf/rus/v14n2/2218-3620-rus-14-02-266.pdf>. [Último acceso: 22 01 2024].
- [8] S. Ross, "Introducción a la estadística", Palo Alto: Reverté, 2018.
- [9] A. O'Connor, M. Modarres y A. Mosleh, "Probability Distributions Used in Reliability Engineering", Maryland: DML International, 2019.