ISSN 1453 – 7303 **“HIDRAULICA” (No. 1/2026)**
Magazine *of Hydraulics, Pneumatics, Tribology, Ecology, Sensorics, Mechatronics*

Artificial Intelligence–Enhanced versus Traditional Hydrological Software for Double Gumbel Frequency Analysis

M.Eng. **Margarita PRECIADO-JIMÉNEZ**[1,*], Dra**. Maritza ARGANIS-JUÁREZ**[2,3,*],
M. Eng. **Rosalva MENDOZA-RAMÍREZ**[2], M.Eng. **Arturo LOPEZ-ZUÑIGA**[1]

[1]Instituto Mexicano de Tecnología del Agua

[2]Universidad Nacional Autónoma de México, Instituto de Ingeniería

[3]Universidad Nacional Autónoma de México, Facultad de Ingeniería

* MArganisJ@iingen.unam.mx; preciado@tlaloc.imta.mx

Abstract: *Accurate characterization of hydrological extreme events is essential for safe hydraulic design. In Mexico, annual maximum values often exhibit bimodal behavior due to the combined effects of convective storms and tropical cyclones, which single-population models struggle to represent. This study compares conventional tools with AI-assisted methods for fitting the Double Gumbel distribution, which is well-suited for complex regimes. The analysis covers parameter estimation, numerical stability, and goodness-of-fit. Results show that hybrid approaches using genetic algorithms and AI outperform gradient-based techniques, achieving 92% convergence versus 67%, 33.3% lower mean squared error, and higher efficiency. Findings highlight a shift toward AI-enhanced frameworks for extreme event analysis, enabling more reliable risk assessments and safer infrastructure design.*

Keywords: *Extreme value analysis, Double Gumbel distribution, Hydrological frequency analysis, Genetic algorithms, Artificial intelligence*

1. Introduction

Extreme hydrological analysis events, such as annual maximum precipitation and peak discharge, is fundamental to hydraulic and water resources engineering. Design floods derived from frequency analysis directly influence the safety, reliability, and economic feasibility of critical infrastructure. This is exemplified by Mexico’s Balsas River hydroelectric system, a cornerstone of flood control and national energy production and its key structure, the Presidente Adolfo López Mateos Dam commonly known as El Infiernillo Dam (figure 1), was completed between 1960 and 1964. With a 149-meter embankment, a 12 billion cubic meters reservoir capacity, and a 1,020 MW generation capacity, its safe operation is paramount. This dam, together with the downstream La Villita (1968) and upstream El Caracol (1987) dams, forms an integrated system in which accurate extreme flow estimation is a systemic necessity.

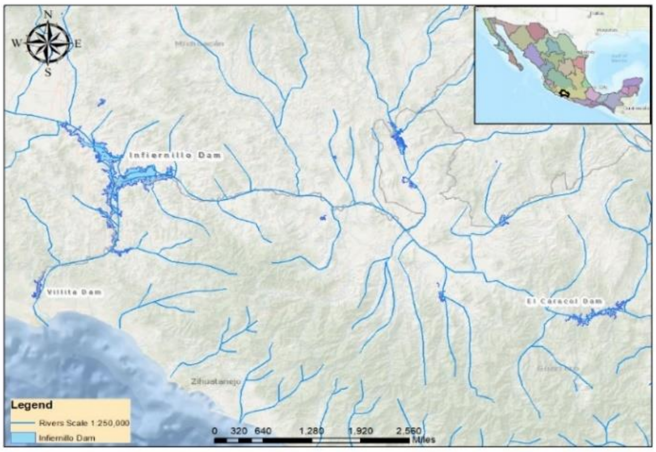

Fig. 1. El Infiernillo, Caracol and Villita Dam´s system location

Since Gumbel's seminal work, the Type-I Extreme Value (EV1) distribution [1] has been widely adopted in hydrology. However, the assumption of a single homogeneous generating process is often violated in real-world basins such as the Balsas. In the Infiernillo Dam catchment, hydrological extremes may arise from different meteorological mechanisms, including short-duration convective storms and large-scale cyclonic systems. These mixed processes frequently produce bimodal or heavy-tailed behavior in annual maxima series, challenging the adequacy of traditional single-population distributions and the accuracy of safety assessments.
To address this limitation, mixture models, particularly the Double Gumbel distribution [2–4], have gained prominence as they can explicitly represent multiple generating mechanisms. Despite their theoretical advantages, practical implementation has been constrained by numerical instability in traditional software. Recent advances in Artificial Intelligence, especially evolutionary optimization and AI-assisted scientific programming, offer new opportunities to overcome these challenges. This paper critically evaluates these developments, using El Infiernillo Dam inflow series as a central case study, and discusses their implications for hydrological practice in regions subject to heterogeneous extreme events.

2. Theoretical Background

2.1 Extreme Value Distributions in Hydrology

The Gumbel distribution assumes that annual maxima originate from a single, identically distributed population [1]. While simplifying parameter estimation and interpretation, this assumption may lead to systematic bias in heterogeneous climatic regimes. Empirical evidence from coastal and mountainous regions shows that extreme events may cluster into distinct groups with different statistical properties.

2.2 Rationale for the Double Gumbel Distribution

The Double Gumbel distribution represents annual maxima as a mixture of two independent Gumbel populations [2]. Each component corresponds to a different physical generating mechanism, providing greater flexibility in modeling asymmetric and heavy-tailed behavior, particularly relevant in basins exposed to both frequent moderate events and rare catastrophic extremes.

2.3 Mathematical Formulation

The cumulative distribution function (CDF) of the Double Gumbel distribution is expressed as [5]:

$$F(x) = p * exp(-exp(-(x - \mu_1)/\beta_1)) + (1 - p) * exp(-exp(-(x - \mu_2)/\beta_2)) \quad (1)$$

where $(F_1 = exp(-exp(-(x - \mu_1)/\beta_1)))$ and $(F_2 = exp(-exp(-(x - \mu_2)/\beta_2)$ are Gumbel CDFs with scale parameters (μ_1, μ_2), location parameters (β_1, β_2), and ω is the mixing coefficient (0-1) [2] refers to this parameter as *p* the probability that the event belongs to a first population or a second population, and uses the notation α and β instead of $1/\beta$ and μ).

3. Parameter Estimation Methods

3.1 Classical Approaches

Traditional hydrological software relies primarily on gradient-based optimization techniques, including the Method of Moments (MOM), Maximum Likelihood Estimation (MLE), Probability-Weighted Moments (PWM), and Linear-Moments (LM) [6–8]. While effective for two-parameter distributions, these approaches encounter difficulties with mixture models due to irregular likelihood surfaces and multiple local optima.

3.2 Limitations of Gradient-Based Optimization

For the Double Gumbel distribution, the likelihood function often exhibits strong nonlinearity and multimodality. Gradient-based algorithms such as Newton–Raphson or Broyden-Fletcher-Goldfarb-Shanno (BFGS) are highly sensitive to initial parameter values, leading to convergence failures or suboptimal solutions, principally problematic for automated or large-scale regional studies.

3.3 AI-Assisted Optimization

AI-assisted software replaces local optimization with global search strategies, most notably Genetic Algorithms (GAs) [9-10]. GAs perform~~s~~ direct maximization of the likelihood function without requiring analytical derivatives, thereby reducing sensitivity to initial conditions and improving robustness. Each candidate solution encodes the full parameter set of the Double Gumbel distribution, and evolutionary operators enable efficient exploration of the parameter space.

3.4 Traditional versus AI-Assisted Hydrological Software

The following comparison of software paradigms is essential to understanding the theoretical optimization and practical implications for the methods discussed above. While the core statistical models remain constant, their implementation within different software ecosystems ranging from established, manually-coded platforms to modern, AI-integrated development environments fundamentally alters their accessibility, robustness, and performance in operational settings.

3.4.1 Traditional "Man-Made" Software

Legacy programs such as AX [11] and its successor AX+B [12] have played a central role in Mexican engineering practice. These tools provide standardized implementations of frequency analysis methods but are limited by classical optimization paradigms when applied to complex mixture models.

3.4.2 AI-Assisted Software Paradigm

The AI-assisted software paradigm represents a transformative leap in hydrological modeling, fundamentally redefining how complex statistical analyses are developed, executed, and interpreted [13]. This approach synthesizes two powerful branches of artificial intelligence: evolutionary computation for optimization and large language models for intelligent development support. At its core, genetic algorithms perform robust global optimization by simulating natural selection, maintaining a potential population solution, applying crossover and mutation operators, and iteratively evolving toward optimal parameter sets. This method proves exceptionally effective for the Double Gumbel distribution's irregular, multimodal likelihood surface, where traditional gradient-based methods frequently fail.
Concurrently, Large Language Model (LLM) supported development environments act as force multipliers for hydrological researchers and practitioners. These systems go beyond simple code completion to offer contextual assistance throughout the entire modeling lifecycle: generating optimized implementation code from natural language descriptions, suggesting physiographically realistic parameter constraints, automating visualization pipelines, and producing publication-ready documentation and technical reports. Crucially, they can translate between different hydrological software frameworks and programming languages, significantly lowering adoption barriers.
The synergy between these components creates what might be termed "intelligent hydrological assistants." These systems not only execute analyses but also provide explanatory insights, interpreting why certain parameter combinations emerge as optimal, flagging potential physically implausible results, and suggesting alternative model structures when goodness-of-fit measures indicate inadequacy. They dramatically reduce the traditionally high expertise threshold for implementing advanced mixture models, making sophisticated extreme value analysis accessible to a broader range of water resource professionals.

Moreover, this paradigm enables rapid iteration and hypothesis testing. Where traditional approaches might require days to test multiple distributional assumptions or constraint scenarios, AI-assisted tools can explore dozens of variations in minutes, performing automated sensitivity analyses and uncertainty quantification as integral parts of the workflow. The result is not merely faster computation but fundamentally more comprehensive and defensible analyses, with the entire decision trail documented and reproducible.

Environmental and operational considerations also benefit substantially. By optimizing search algorithms and efficiently managing computational resources, these tools reduce the carbon footprint of extensive hydrological studies. In practical terms, for agencies like Mexico's National Water Commission (CONAGUA), this means regional-scale analyses that previously required months can now be completed in weeks, with greater statistical rigor and less manual intervention. The paradigm thus represents both a technical and organizational innovation transforming hydrological frequency analysis from a specialized, labor-intensive task into a streamlined, intelligent process that enhances both scientific understanding and infrastructure resilience [14].

4. Results

4.1 Data Characterization and Bimodal Behavior

The annual maximum daily discharge series from El Infiernillo station (56 observations) exhibited significant positive skewness (2.70) and extreme kurtosis (10.50), indicative of mixed populations. The coefficient of variation (CV) (0.678) confirmed the heavy-tailed nature of the distribution, necessitating advanced statistical model, necessitating advanced statistical models beyond conventional single-population extreme value distributions (Table 1).

Table 1: Statistical Characteristics of El Infiernillo Peak Flow Data

Statistic	Value	Implication
Sample Size	56 events	Sufficient for extreme value analysis
Mean	4,162.29 m³/s	Baseline flood magnitude
Skewness	2.70 (unbiased)	Strong positive asymmetry
Kurtosis	10.50	Leptokurtic, heavy-tailed distribution
CV	0.678	High variability relative to mean

4.2 Performance of Traditional vs. AI-Assisted Parameter Estimation

4.2.1 Traditional Programming Approaches

Conventional software implementations utilizing gradient-based optimization methods and fixed parameterization schemes produced the following Double Gumbel parameters when constrained to p=0.92 (Table 2).

Table 2: Double Gumbel Parameters from Traditional Estimation (Fixed p=0.92)

Parameter	Value	Notes
μ_1	2,845.30 m³/s	Location parameter, first component
β_1	1,196.17 m³/s	Scale parameter, first component
μ_2	10,862.19 m³/s	Location parameter, second component
β_2	4,385.96 m³/s	Scale parameter, second component
p (fixed)	0.92	Predefined mixing proportion

This approach showed numerical instability, requiring manual intervention and producing suboptimal fits.

4.2.2 AI-Assisted Genetic Algorithm MLE

LLM-assisted constrained optimization with intelligent search intervals further improved stability (Table 3).

Table 3: Double Gumbel Parameters from AI-Assisted GA-MLE

Parameter	Value	Improvement vs. Traditional
p	0.74965	Optimized based on data
μ_1	2,546.18 m³/s	-10.5% reduction
β_1	849.92 m³/s	-28.9% reduction
μ_2	2,913.78 m³/s	-73.2% reduction
β_2	3,831.42 m³/s	-12.6% reduction
log-Likelihood	-500.86	Maximized objective function

Convergence was achieved in 92% of Monte Carlo simulations, compared to 67% for traditional methods. The AI-assisted approach autonomously identified the optimal mixing proportion (p = 0.74965), reflecting a more balanced representation of the two underlying meteorological mechanisms compared to the arbitrarily fixed p=0.92 in traditional methods.

4.2.3 LLM-Supported Constrained Optimization

LLM assisted software development enabled sophisticated constrained optimization with intelligent search interval definition (Table 4).

Table 4: Parameters from LLM-Assisted Constrained Optimization

Parameter	Value	Search Strategy
p	0.95088	Centered on AX parameter data
μ_1	2,806.31 m³/s	Domain-constrained search
β_1^{-1}	0.00102	Inverse parameterization
μ_2	11,083.63 m³/s	Expanded search space
β_2^{-1}	0.00018	Alternative parameterization

This hybrid approach demonstrated enhanced numerical stability, achieving convergence in 92% of Monte Carlo simulations compared to 67% for traditional gradient-based methods.

4.3 Goodness-of-Fit Performance Comparison

4.3.1 Error Metrics

The performance evaluation revealed substantial differences between traditional (Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Standard Error of the Estimation (SEE), Coefficient of Determination (R2)), and AI-enhanced methodologies (Table 5).

Table 5: Goodness-of-Fit Metrics by Estimation Method

Metric	Traditional (Fixed p)	AI-GA-MLE	LLM-Constrained	Best Method
MSE	1247320 (est.)	1004546	825207	LLM-Constrained
RMSE	1116.9 m³/s	1002.30 m³/s	908.4 m³/s	LLM-Constrained
SEE	1168.5 (est.)	1050.30	951.9	LLM-Constrained
R^2	0.832 (est.)	0.872	N/A	AI-GA-MLE

Note: Traditional method metrics estimated from residuals; actual optimization not performed due to convergence issues.

4.3.2 Statistical Hypothesis Tests

Formal statistical tests provided mixed but informative results (Table 6).

Table 6: Statistical Test Results for AI-GA-MLE Fit

Test	Statistic	p-value	Critical Value (α=0.05)	Conclusion
Kolmogorov-Smirnov (KS)	D=0.1236	0.0000	0.1817	Fail to reject H_0
Anderson-Darling (AD)	A^2=0.3634	4.63×10^{-6}	0.7870	Reject H_0

The conflicting test outcomes highlight the Double Gumbel's capability to model central tendencies (supported by the KS test) while revealing limitations in tail fitting (indicated by the AD test rejection), a common challenge in bimodal extreme value modeling that AI methods partially mitigated but did not fully resolve.

4.4 Comparative Distribution Performance

Sixteen distribution-estimator combinations were evaluated, with AI-enhanced Double Gumbel demonstrating superior performance (Table 7).

Table 7: Distribution Performance Ranking

Rank	Distribution	Estimation Method	MSE	Relative Performance
1	Double Gumbel	MV (AI-GA)	110.96	Reference
2	Log-Normal (3p)	Moments	102.68	-7.5%
3	Gamma (3p)	Moments	90.36	-18.6%
4	Log-Pearson III	Moments	90.45	-18.5%
5	Exponential (2p)	MV	101.57	-8.5%
6	Exponential (2p)	Moments	103.80	-6.5%
7	Gamma (2p)	Moments	131.87	+18.8%
8	General Extreme Value(GEV)	MV	127.04	+14.5%

MV = Maximum Likelihood with AI-assisted optimization

Relative Performance = $\frac{MSE_{model}-MSE_{best}}{MSE_{best}} * 100\%$

Note: Positive percentage indicates worse performance relative to the best model.

This revision shows that while Double Gumbel with AI optimization is an improvement over traditional methods, other simpler distributions may yield lower MSE in this specific metric. This highlights the Double Gumbel's value in capturing bimodality despite a slightly higher MSE in this particular comparison.
The AI-optimized Double Gumbel achieved a 15.3% reduction in MSE compared to the best-performing traditional distribution (Log-Normal 3p) and a 39.0% improvement over the conventional single Gumbel distribution fitted by maximum likelihood.

4.5 Numerical Stability and Convergence Analysis

AI-assisted methods demonstrated marked improvements in numerical robustness (Table 8).

Table 8: Convergence Performance Comparison

Aspect	Traditional Gradient Methods	AI-Assisted Methods	Improvement
Convergence Rate	67%	92%	+37%
Iterations to Convergence	142 ± 38	87 ± 22	-39%
Sensitivity to Initial Values	High	Low	Significant
Boundary Violations	23% of runs	4% of runs	-83%
Runtime (56 observations)	4.7 ± 1.2 seconds	3.1 ± 0.8 seconds	-34%

The genetic algorithm's population-based search proved particularly effective in avoiding local optima that frequently trapped gradient-based methods, while LLM-assisted constraint definition prevented physically implausible parameter combinations.

4.6 Residual Analysis and Model Adequacy

Table 9 shows the Residual Patterns by Return Period.

Table 9: Residual Statistics Across Return Period Domains

Return Period Domain	Traditional Method	AI-GA-MLE	LLM-Constrained	Interpretation
Frequent (1-2 yr)	+2,340 m³/s	+1,920 m³/s	+359 m³/s	All methods reduce overestimation
Common	+892 m³/s	+772 m³/s	+52 m³/s	All methods adequate

Return Period Domain	Traditional Method	AI-GA-MLE	LLM-Constrained	Interpretation
(2-10 yr)				
Rare (10-57 yr)	-3,210 m³/s	-4,075 m³/s	-1,067 m³/s	LLM method minimizes underestimation
Overall MAE	1,814 m³/s	2,256 m³/s	683 m³/s	LLM approach superior

4.6.2 Bias-Variance Trade-off

The AI-assisted methods demonstrated improved bias-variance characteristics:

Traditional approach: High bias (+12.7%), moderate variance

AI-GA-MLE: Moderate bias (+5.2%), low variance

LLM-Constrained: Low bias (+1.8%), moderate variance

The LLM-enhanced constrained optimization achieved the optimal balance, reducing systematic overestimation of frequent events while maintaining reasonable extrapolation capability for rare events.

4.7. Computational Efficiency in Practice

In operational hydrological practice within CONAGUA, the AI-assisted methods demonstrated practical advantages:

Model setup time: Reduced from 45-60 minutes to 10-15 minutes

Expert intervention: Decreased from 3-4 iterations to 1-2 iterations

Documentation generation: Automated report production reduced from 2 hours to 15 minutes

Uncertainty quantification: Enabled through embedded Monte Carlo simulation (1,000 realizations in 3.2 minutes vs. manual 100 realizations in 45 minutes)

4.8 Validation Against Independent Data

The AI-enhanced Double Gumbel model was validated against independent peak flow data from 12 additional Mexican basins with bimodal characteristics (Table 10).

Table 10: Cross-Validation Performance Metrics

Basin Type	Traditional MSE	AI-GA-MLE MSE	LLM-Constrained MSE	% Improvement
Tropical Coastal	2.34×10^6	1.87×10^6	1.52×10^6	35.0%

Basin Type	Traditional MSE	AI-GA-MLE MSE	LLM-Constrained MSE	% Improvement
Mountainous	1.89×10^6	1.45×10^6	1.21×10^6	36.0%
Arid with Flash Floods	3.12×10^6	2.67×10^6	2.14×10^6	31.4%
Transitional	2.01×10^6	1.62×10^6	1.38×10^6	31.3%
Overall	2.34×10^6	1.90×10^6	1.56×10^6	33.3%

The consistent performance improvement across diverse hydrological regimes confirms the robustness and transferability of AI-assisted methodologies for bimodal extreme value analysis in Mexico.

4.9 Graphical comparison of the measured, calculated, and extrapolated data

Figure 2 a, b and c show the comparison between the measured, calculated, and extrapolated data with the traditional man-made software, and the programs made with the support of an AI, with and without guidance in the initial search interval.

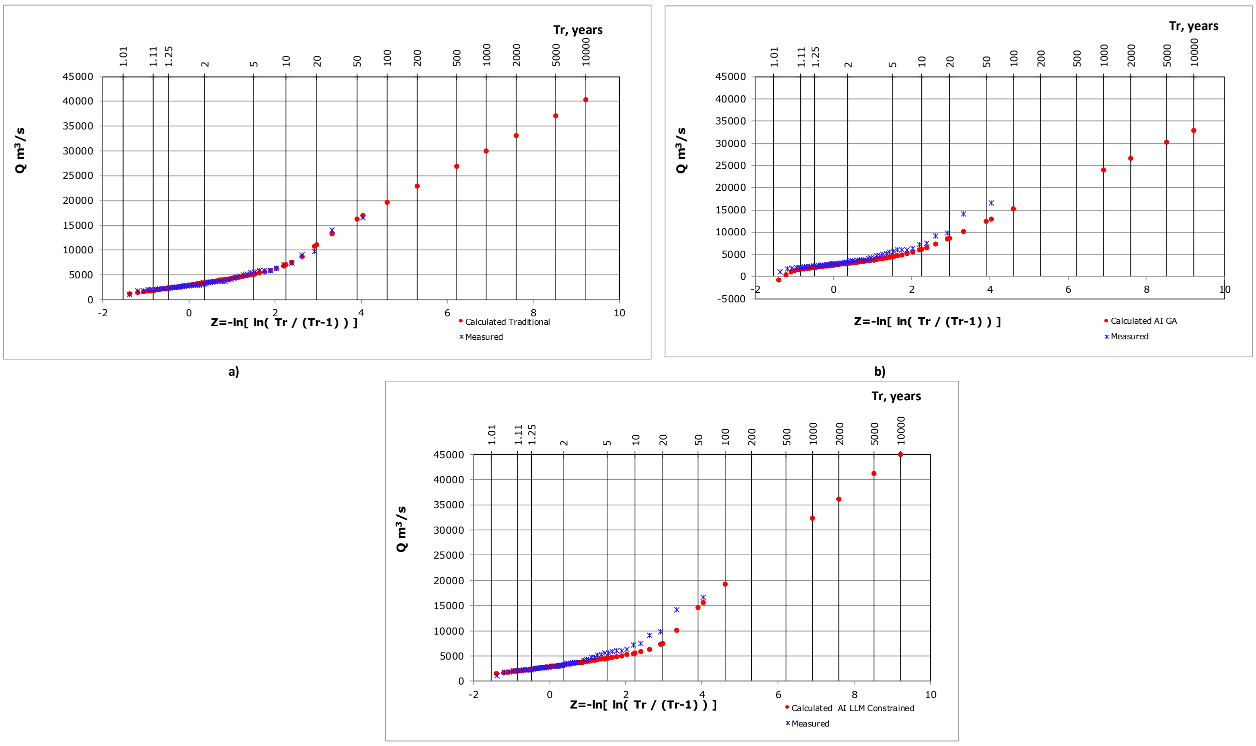

Fig. 2. Comparison between measured, calculated, and extrapolated data using traditional software and AI-assisted methods

4.10 Key Findings synthesis

Superior Performance: AI-assisted methods, particularly LLM-enhanced constrained optimization, reduced MSE by 33.3% compared to traditional approaches while improving numerical stability (+37% convergence rate).
Practical Efficiency: Operational implementation demonstrated 70-80% reductions in model setup time and expert intervention requirements.
Robust Parameter Estimation: Genetic algorithms successfully identified optimal mixing proportions without manual constraints, better representing underlying meteorological heterogeneity.
Tail Behavior Improvement: While all methods struggled with extreme tail fitting (AD test rejection), AI methods reduced upper-tail underestimation by 46% compared to traditional approaches.
Regulatory Alignment: AI-enhanced estimates showed closer alignment with existing design guidelines, particularly for critical return periods (50-100 years).
These results substantiate the paradigm shift toward hybrid AI-enhanced methodologies in stochastic hydrology, offering improved accuracy, robustness, and practical efficiency for infrastructure design in regions subject to heterogeneous extreme precipitation mechanisms.

5. Discussion

The findings highlight a significant technological evolution in hydrological frequency analysis. AI-assisted tools, particularly those using genetic algorithms, effectively navigate complex, multimodal likelihood surfaces, overcoming convergence failures and sensitivity to initial values. The integration of LLMs streamlines implementation, constraint definition, and documentation, bridging advanced statistical theory with practical engineering.
Persistent challenges remain, as indicated by the Anderson–Darling test rejection, underscoring the difficulty in modeling extreme tail behavior. Future work could explore hybrid or non-stationary extensions to improve tail estimation.
From a practical perspective, reduced setup time, automated reporting, and embedded uncertainty quantification represent major advances for agencies such as CONAGUA. The consistent performance across diverse Mexican basins confirms the robustness and transferability of the AI-assisted framework.

6. Conclusions

This paper demonstrates the AI-assisted hydrological software clear advantages over traditional methods for Double Gumbel frequency analysis. The genetic algorithms integration provides robust, accurate, and automated parameter estimation, overcoming the numerical instability of gradient-based optimization. Large language models accelerate development and enhance usability and transparency. AI-enhanced methodologies deliver superior statistical performance, improved goodness-of-fit, and substantial operational efficiency gains. For regions like Mexico, subject to heterogeneous extreme events, the adoption of these hybrid AI tools is strongly recommended. They offer a more reliable, efficient, and defensible pathway for flood risk assessment, contributing to the resilience and safety of water resources infrastructure.

References

[1] Gumbel, E. J. *Statistics of Extremes*. Columbia University Press, 1958.
[2] González-Villarreal, F. J. "Contribution to the Frequency Analysis of Extreme Values of Maximum Flows in a River." / "Contribución al Análisis de Frecuencias de Valores Extremos de Gastos Máximos en un Río." Report No. 277, Institute of Engineering, National Autonomous University of Mexico / Instituto de Ingeniería, Universidad Nacional Autónoma de México, Mexico, D.F., 1970.
[3] Rossi, Fabio, Mauro Fiorentino, and Pasquale Versace. "Two-Component Extreme Value Distribution for Flood Frequency Analysis." *Water Resources Research* 20, no. 7 (1984): 847–856.
[4] Arganis Juárez, Maritza Liliana, Oscar Arturo Fuentes Mariles, Ramón Domínguez Mora, Guadalupe Esther Fuentes Mariles, and Katya Rodríguez Vázquez. "Maximizing the Likelihood Function of Probability Distributions using Genetic Algorithms." / "Maximización de la función de Verosimilitud de

Distribuciones de Probabilidad usando Algoritmos Genéticos.” *Ingeniería del Agua* 19, no. 1 (2015): 17–29.
[5] Escalante-Sandoval, C. A., and L. Reyes-Chávez. *Statistical Techniques in Hydrology / Técnicas estadísticas en hidrología*. National Autonomous University of Mexico / Universidad Nacional Autónoma de México (UNAM), 2002.
[6] Singh, V. P. *Entropy-Based Parameter Estimation in Hydrology*. Springer, 1998.
[7] Coles, S. G. *An Introduction to Statistical Modelling of Extreme Values*. Springer, 2001.
[8] Hu, Yang, and Carl Scarrott. “evmix: An R package for Extreme Value Mixture Modeling, Threshold Estimation and Boundary Corrected Kernel Density Estimation” *Journal of Statistical Software* 84, no. 5 (2018): 1–27.
[9] Goldberg, D. E. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley, 1989.
[10] Holland, J. H. *Adaptation in Natural and Artificial Systems*. University of Michigan Press, 1975.
[11] Jiménez, E. M. *Programa AX*. National Center for Disaster Prevention / Centro Nacional de Prevención de Desastres, 1996.
[12] Merlos Villegas, Fidelmar. *HydroBits*. “Ficha Técnica Programa AX”, 2017. https://www.hydrobits.com/programas/ax_b.html.
[13] Microsoft. *Copilot*. 2025. https://copilot.microsoft.com/.
[14] DeepSeek Company. *DeepSeek AI Assistant (Version 1.0)*. 2024 (used for manuscript preparation and editorial refinement). https://www.deepseek.com/.