Regression Modeling and Risk Analysis of Asymmetric Unit-Interval Data in Environmental and Geoscience Studies for Sustainable Planning

Abdulhameed Ado Osi[1,*], Shamsuddeen Ahmad Sabo[1], Ibrahim Zakariyya Musa[1], Iman Babiker[2], Haifa Alqahtani[3] and Mohammed Elgarhy[4,5]

[1] Department of Statistics, Aliko Dangote University of Science and Technology, Wudil, Kano, Nigeria

[2] Department of Accounting, College of Business Administration, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia

[3] Department of Statistics and Business Analytics, United Arab Emirates University, Al Ain, Abu Dhabi, United Arab Emirates

[4] Faculty of Computers and Information Systems, Egyptian Chinese University, Nasr City, Egypt

[5] Department of Computer Engineering, Biruni University, Istanbul, Turkey

INFORMATION

Keywords:
Topp–Leone distribution
unit distributions
risk analysis
simulation study

DOI: 10.23967/j.rimni.2026.10.80753

Revista Internacional de Métodos Numéricos para Cálculo y Diseño en Ingeniería

Regression Modeling and Risk Analysis of Asymmetric Unit-Interval Data in Environmental and Geoscience Studies for Sustainable Planning

Abdulhameed Ado Osi[1,*], **Shamsuddeen Ahmad Sabo**[1], **Ibrahim Zakariyya Musa**[1], **Iman Babiker**[2], **Haifa Alqahtani**[3] and **Mohammed Elgarhy**[4,5]

[1]Department of Statistics, Aliko Dangote University of Science and Technology, Wudil, Kano, Nigeria

[2]Department of Accounting, College of Business Administration, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia

[3]Department of Statistics and Business Analytics, United Arab Emirates University, Al Ain, Abu Dhabi, United Arab Emirates

[4]Faculty of Computers and Information Systems, Egyptian Chinese University, Nasr City, Egypt

[5]Department of Computer Engineering, Biruni University, Istanbul, Turkey

ABSTRACT

This paper introduces the New Extended Topp–Leone (NExTL) distribution, a flexible three-parameter probability model for data on the unit interval (0, 1). Developed using a type 2 exponentiation framework, it generalizes the classical Topp–Leone family, offering enhanced versatility in capturing various density shapes and hazard rate behaviors. We derive its core statistical properties, including moments, quantile and hazard functions, entropy measures, and order statistics. Parameter estimation via maximum likelihood is confirmed as consistent and reliable through simulation studies. For practical application, the NExTL distribution is integrated within the generalized additive models for location, scale, and shape (GAMLSS) framework, enabling comprehensive regression modeling of bounded responses. The model's superior performance is empirically validated in two applications: a risk analysis of petroleum rock permeability data, where it provides the best fit and robust tail-risk quantification (Value at Risk and Tail Value at Risk), and a quantile regression analysis of behavioral ecology data, successfully identifying significant treatment effects on insect attraction. In both cases, the NExTL distribution outperforms standard competitors such as the beta and Kumaraswamy distributions, as measured by information criteria and goodness-of-fit statistics, representing a significant advancement for modeling proportional data in fields like engineering, finance, and biology.

OPEN ACCESS

Received: 13/02/2026

Accepted: 23/06/2026

Published: 07/08/2026

DOI
10.23967/j.rimni.2026.10.80753

Keywords:
Topp–Leone distribution
unit distributions
risk analysis
simulation study

1 Introduction

In recent years, the development of flexible statistical distributions for modeling bounded data on the interval (0, 1) has gained significant attention across various scientific disciplines, particularly in engineering, finance, and environmental sciences [1–3]. The analysis of proportional data, rates, and

*Correspondence: Abdulhameed Ado Osi (aaosi@kustwudil.edu.ng).

ratios requires specialized distributions that can capture complex patterns while maintaining mathematical tractability. Classical unit distributions such as the beta distribution [4] and Kumaraswamy distribution [5] have served as foundational models for bounded data analysis.

Among these established distributions for bounded data, the Topp–Leone (TL) distribution by [6] has emerged as a valuable tool since its introduction, characterized by its cumulative distribution function $F(x;\nu) = x^{\nu}[2-x]^{\nu}$ for $0 < x < 1$, $\nu > 0$. Despite its simplicity, the original TL distribution exhibits limitations in flexibility, which has motivated numerous extensions using various methodological approaches. For instance, exponentiated Topp–Leone distribution by [7], Type II Topp Leone power Lomax distribution by [8], Kumaraswamy Topp–Leone distribution by [9], Marshall-Olkin Topp–Leone distribution by [10], a new power Topp–Leone distribution [11], a new Asymmetric modified Topp–Leone distribution [12], power Topp–Leone distribution [13], Type II Topp–Leone generated family of distributions [14].

However, existing Topp–Leone extensions (e.g., TL, KTL, GTL) lack separate control over lower and upper tail heaviness and asymmetry direction. This is a critical gap for environmental and geoscience applications: petroleum contamination data often show extreme upper tails (spill events), while ecological purity indices show lower-tail clustering—existing models fail to capture both simultaneously. The NExTL distribution introduces two shape parameters that independently govern left-tail behavior (e.g., rare low-pollution events) and right-tail behavior (e.g., extreme high-risk events), directly addressing this gap.

The petroleum engineering sector presents unique challenges for statistical modeling and risk analysis. As noted by [1,2], the industry faces inherent uncertainties in reservoir characterization, production forecasting, and risk assessment. Traditional statistical methods often fall short in capturing the complex dependencies and extreme events that characterize petroleum engineering data. Recent studies by [15–18] have emphasized the need for more sophisticated statistical tools to handle the peculiarities of petroleum rock sample data, including bounded measurements, asymmetric behavior, and tail dependence.

This paper introduces a New Extended Topp–Leone (NExTL) distribution that builds upon previous extensions while addressing their limitations. Our approach is based on type-2 exponentiation, yielding a distribution with greater flexibility in modeling tail behavior. The proposed NExTL distribution offers several advantages over existing extensions: (1) it provides greater flexibility in capturing various hazard rate shapes, including bathtub, unimodal, and increasing-decreasing patterns; (2) it exhibits superior tail behavior for modeling extreme events in risk analysis; (3) it maintains mathematical tractability for parameter estimation and inference; and (4) it facilitates the development of a comprehensive regression framework for bounded response variables.

The remainder of this paper is organized as follows: Section 2 introduces the mathematical formulation of the NExTL distribution and Section 3 derives its fundamental properties. Section 4 develops the associated risk analysis. Section 5 illustrates the simulation study. Section 6 demonstrates practical applications to petroleum engineering data, Section 7 presents analysis on quantile regression models, and Section 8 concludes with discussions and future research directions.

2 The New Extended Topp–Leone Distribution

In this section, we introduce a New Extended Topp–Leone (NExTL) distribution by employing the framework proposed by [19], which enhances the distribution's capability to model various data patterns, including heavy-tailed behaviors and multiple hazard rate shapes. The cumulative distribution function (CDF) of the NExTL distribution is given by:

$$F(y;\omega,\delta,\lambda)=\frac{1-\left[1-y^{\omega}(2-y)^{\omega}\right]^{\delta}}{1-\left[1-y^{\omega}(2-y)^{\omega}\right]^{\delta}+\left[1-y^{\omega}(2-y)^{\omega}\right]^{\lambda}},\quad 0<y<1, \tag{1}$$

where $\omega>0$, $\delta>0$, and $\lambda>0$ are shape parameters that control the distribution's flexibility and tail behavior.

The corresponding probability density function (PDF) is obtained by differentiating Eq. (1):

$$f(y;\omega,\delta,\lambda)=\frac{\delta\lambda\omega y^{\omega-1}(2-y)^{\omega-1}(2-2y)\left[1-y^{\omega}(2-y)^{\omega}\right]^{\delta+\lambda-2}}{\left\{1-\left[1-y^{\omega}(2-y)^{\omega}\right]^{\delta}+\left[1-y^{\omega}(2-y)^{\omega}\right]^{\lambda}\right\}^{2}}. \tag{2}$$

The PDF of the NExTL distribution exhibits exceptional flexibility in modeling diverse data patterns on the unit interval, as demonstrated in Fig. 1. The distribution can capture unimodal shapes, U-shaped patterns, J-shaped, reverse J-shapes, and nearly symmetric distributions. The ability to generate such diverse density shapes with only three parameters makes the NExTL distribution superior to traditional bounded distributions like the beta distribution [4] and Kumaraswamy distribution [5], particularly for complex real-world applications in proportion modeling, risk assessment, and reliability analysis where data may exhibit multiple distributional patterns.

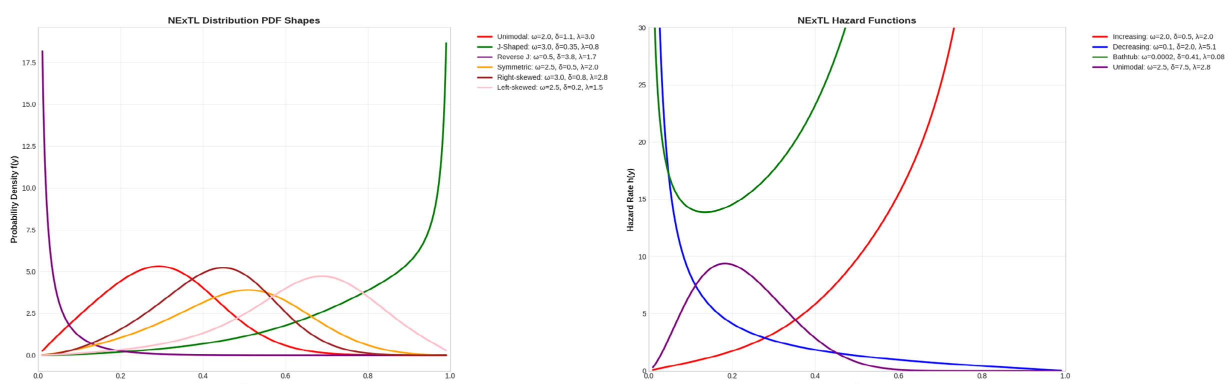


Figure 1: PDF and HRF of the NExTL distribution for different parameter combinations

3 Statistical Properties of NExTL Distribution

3.1 The Survival Function of NExTL

The survival function represents the probability that a variable exceeds a specified value. The survival function of the NExTL distribution is given by:

$$S(y;\omega,\delta,\lambda)=\frac{\left[1-y^{\omega}(2-y)^{\omega}\right]^{\lambda}}{1-\left[1-y^{\omega}(2-y)^{\omega}\right]^{\delta}+\left[1-y^{\omega}(2-y)^{\omega}\right]^{\lambda}}. \tag{3}$$

3.2 *The Hazard Rate Function of NExTL*

The hazard rate function (HRF) of the NExTL distribution, which characterizes the instantaneous failure rate, is derived as:

$$h(y;\omega,\delta,\lambda)=\frac{\delta\lambda\omega y^{\omega-1}(2-y)^{\omega-1}(2-2y)\left[1-y^{\omega}(2-y)^{\omega}\right]^{\delta-2}}{1-\left[1-y^{\omega}(2-y)^{\omega}\right]^{\delta}+\left[1-y^{\omega}(2-y)^{\omega}\right]^{\lambda}}. \tag{4}$$

Fig. 1 illustrates how the PDF of the New Extended Topp–Leone (NExTL) distribution varies when its parameters assume different values. It shows that the distribution can take various shapes, including unimodal, J-shaped, U-shaped, skewed, and symmetric.

Fig. 1 displays the hazard rate function (HRF) of the New Extended Topp–Leone distribution for different parameters. The figure shows that the distribution can represent decreasing, increasing, V-shaped, and reverse J-shaped hazard rates.

3.3 *The Quantile Function of NExTL*

The quantile function, essential for the generation of random numbers and quantile regression, is obtained by solving $F(Q(p))=p$ for $Q(p)$. This yields the equation:

$$p=\frac{1-\left[1-Q(p)^{\omega}(2-Q(p))^{\omega}\right]^{\delta}}{1-\left[1-Q(p)^{\omega}(2-Q(p))^{\omega}\right]^{\delta}+\left[1-Q(p)^{\omega}(2-Q(p))^{\omega}\right]^{\lambda}},\quad 0<p<1. \tag{5}$$

The previous equation cannot be solved in closed form, but numerical techniques provide a viable approach for its solution.

3.4 *Asymptotic Properties*

The asymptotic behavior of the NExTL distribution is derived by analyzing the limits of the PDF as y approaches the support boundaries.

As $y\to 0^{+}$:

$$\lim_{y\to 0^{+}} f(y)\approx\frac{\delta\lambda\omega}{(\delta-\lambda)^{2}2^{\omega}}y^{-(\omega+1)}. \tag{6}$$

As $y\to 1^{-}$:

$$\lim_{y\to 1^{-}} f(y)\approx\frac{2\delta\lambda\omega^{\delta+\lambda-1}}{(\delta-\lambda)^{2}}z^{2(\delta+\lambda)-7}. \tag{7}$$

3.5 *Moments and Moment Generating Function*

In this section, we derived the moment and moment generating function of the proposed NExTL distribution. The r-th raw moment of the NExTL distribution is

$$\mu_r'=\mathbb{E}[Y^r]=\int_0^1 y^r f_Y(y)\,dy=\int_0^1\frac{\delta\lambda\omega y^{r+\omega-1}(2-y)^{\omega-1}(2-2y)\left[1-y^{\omega}(2-y)^{\omega}\right]^{\delta+\lambda-2}}{\left\{1-\left[1-y^{\omega}(2-y)^{\omega}\right]^{\delta}+\left[1-y^{\omega}(2-y)^{\omega}\right]^{\lambda}\right\}^{2}}dy. \tag{8}$$

The mean and variance are $\mu=\mu_1'$ and $\sigma^2=\mu_2'-\mu^2$.

The moment generating function (MGF) is given by definition $M_t(y)=\int_0^1 e^{ty}f_Y(y)\,dy$. Then the MGF of the NExTL distribution is given by

$$M_t(y)=\int_0^1\frac{\delta\lambda\omega y^{\omega-1}(2-y)^{\omega-1}(2-2y)\left[1-y^{\omega}(2-y)^{\omega}\right]^{\delta+\lambda-2}e^{ty}}{\left\{1-\left[1-y^{\omega}(2-y)^{\omega}\right]^{\delta}+\left[1-y^{\omega}(2-y)^{\omega}\right]^{\lambda}\right\}^{2}}dy. \tag{9}$$

The previous Eqs. (8) and (9) cannot be solved in closed form, but numerical techniques provide a viable approach for its solution.

Table 1 reports the mean, variance, and moment generating function (MGF) at $t = 0,\ 0.5,\ 1$ for selected values of δ, λ, and ω. For fixed δ and λ, increasing ω (from 0.5 to 1.5) consistently increases both the mean and the variance, indicating a heavier right tail. Holding ω and δ fixed, an increase in λ (from 1.2 to 2.2) raises the MGF values substantially, while the mean first decreases then increases, and the variance tends to decline. When δ increases (from 1.5 to 2.5) with λ and ω fixed, both the mean and the MGF decrease, suggesting that larger δ shifts probability mass toward smaller values. These results confirm that the NExTL distribution is highly flexible, capable of modelling a wide range of shapes on the unit interval, with the mean ranging from about 0.05 to 0.36 and variance from near zero to 0.04 under the examined parameters.

Table 1: Mean, variance, and moment generating function (MGF) of the NExTL distribution for selected parameter combinations. The MGF is evaluated at $s = 0$, 0.5, and 1

δ	λ	ω	Mean	Variance	$M(0)$	$M(0.5)$	$M(1)$
1.5	1.2	0.5	0.11092	0.02191	0.95292	1.01299	1.08378
1.5	1.2	1.0	0.21300	0.03412	0.95292	1.07021	1.21294
1.5	1.2	1.5	0.28809	0.03865	0.95292	1.11356	1.31381
1.5	1.8	0.5	0.10391	0.01254	1.28691	1.34195	1.40394
1.5	1.8	1.0	0.23375	0.01581	1.28691	1.41320	1.56105
1.5	1.8	1.5	0.33525	0.00802	1.28691	1.47074	1.69225
1.5	2.2	0.5	0.09849	0.00868	1.49123	1.54289	1.59989
1.5	2.2	1.0	0.24255	0.00611	1.49123	1.62112	1.77045
2.0	1.2	0.5	0.07072	0.01269	0.87132	0.90904	0.95211
2.0	1.2	1.0	0.15417	0.02500	0.87132	0.95496	1.05373
2.0	1.2	1.5	0.22043	0.03240	0.87132	0.99248	1.13924
2.0	1.8	0.5	0.07332	0.00850	1.22338	1.26186	1.30438
2.0	1.8	1.0	0.18343	0.01475	1.22338	1.32152	1.43417
2.0	1.8	1.5	0.27549	0.01262	1.22338	1.37296	1.54969
2.0	2.2	0.5	0.07242	0.00634	1.44039	1.47812	1.51914
2.0	2.2	1.0	0.19664	0.00805	1.44039	1.54488	1.66312
2.5	1.2	0.5	0.04864	0.00767	0.81941	0.84505	0.87366
2.5	1.2	1.0	0.11901	0.01834	0.81941	0.88324	0.95692
2.5	1.2	1.5	0.17870	0.02623	0.81941	0.91656	1.03168
2.5	1.8	0.5	0.05428	0.00582	1.17739	1.20567	1.23646
2.5	1.8	1.0	0.14966	0.01263	1.17739	1.25684	1.34662
2.5	1.8	1.5	0.23422	0.01333	1.17739	1.30356	1.45032
2.5	2.2	0.5	0.05541	0.00461	1.40104	1.42974	1.46061
2.5	2.2	1.0	0.16447	0.00804	1.40104	1.48788	1.58491
2.5	2.2	1.5	0.26389	0.00226	1.40104	1.54250	1.70538

3.6 Entropy Measures (Simplified)

The Rényi entropy for $\alpha > 0, \alpha \neq 1$, is define as:

$$H_\alpha(Y) = \frac{1}{1-\alpha}\log\left(\int_0^1 [f_Y(y)]^\alpha dy\right).$$

Using the transformation $f_Y(y)dy = f_T(t)dt$ and the expression for $f_T(t)$,

$$[f_Y(y)]^\alpha dy = \frac{(\delta\lambda)^\alpha (1-t)^{\alpha(\delta+\lambda-2)}}{\left(1-(1-t)^\delta+(1-t)^\lambda\right)^{2\alpha}}\, dt.$$

Hence the Rényi entropy for NEXTL distribution is,

$$H_\alpha(Y) = \frac{1}{1-\alpha}\log\left(\int_0^1 \frac{(\delta\lambda)^\alpha (1-t)^{\alpha(\delta+\lambda-2)}}{\left(1-(1-t)^\delta+(1-t)^\lambda\right)^{2\alpha}} dt\right). \tag{10}$$

The Shannon entropy is define as: $H(Y) = -\int_0^1 f_Y(y)\log f_Y(y)\, dy$. Substituting $f_Y(y)dy = f_T(t)dt$

Hence, the Shannon entropy for NExTL is given as:

$$H(Y) = -\int_0^1 \frac{\delta\lambda\,(1-t)^{\delta+\lambda-2}}{\left(1-(1-t)^\delta+(1-t)^\lambda\right)^2}\log\left(\frac{\delta\lambda\,(1-t)^{\delta+\lambda-2}}{\left(1-(1-t)^\delta+(1-t)^\lambda\right)^2}\right)dt. \tag{11}$$

Thus, all statistical properties reduce to integrals over $(0, 1)$ that contain only elementary functions (powers, square roots, exponentials, logarithms). The original complicated factors have been eliminated, making numerical evaluation straightforward.

Table 2 presents the Rényi entropy (with $\alpha = 0.5$) and the Shannon (differential) entropy for the NExTL distribution across different δ, λ, and ω values. For a given δ, increasing λ raises the Rényi entropy (from negative to positive) while the Shannon entropy becomes more negative. Conversely, for a fixed λ, increasing δ lowers both entropy measures; at $\lambda = 1.2$, Rényi decreases and Shannon becomes slightly less negative. The ability to produce positive Rényi entropy and a wide range of Shannon values demonstrates the distribution's flexibility in modelling varying levels of dispersion and concentration on the unit interval.

Table 2: Rényi and Shannon entropy of the NExTL distribution for various parameter combinations ($\alpha = 0.5$)

δ	λ	ω	α	Rényi	Shannon
1.5	1.2	0.5	0.5	−0.119838	−0.073299
1.5	1.2	1.0	0.5	−0.119838	−0.073299
1.5	1.2	1.5	0.5	−0.119838	−0.073299
1.5	1.8	0.5	0.5	0.085768	−0.659226
1.5	1.8	1.0	0.5	0.085768	−0.659226
1.5	1.8	1.5	0.5	0.085768	−0.659226
1.5	2.2	0.5	0.5	0.168140	−1.108862
1.5	2.2	1.0	0.5	0.168140	−1.108862
1.5	2.2	1.5	0.5	0.168140	−1.108862

(Continued)

Table 2 (continued)

δ	λ	ω	α	Rényi	Shannon
2.0	1.2	0.5	0.5	−0.298849	−0.115116
2.0	1.2	1.0	0.5	−0.298849	−0.115116
2.0	1.2	1.5	0.5	−0.298849	−0.115116
2.0	1.8	0.5	0.5	−0.054515	−0.727679
2.0	1.8	1.0	0.5	−0.054515	−0.727679
2.0	1.8	1.5	0.5	−0.054515	−0.727679
2.0	2.2	0.5	0.5	0.046171	−1.201562
2.0	2.2	1.0	0.5	0.046171	−1.201562
2.0	2.2	1.5	0.5	0.046171	−1.201562
2.5	1.2	0.5	0.5	−0.456101	−0.177444
2.5	1.2	1.0	0.5	−0.456101	−0.177444
2.5	1.2	1.5	0.5	−0.456101	−0.177444
2.5	1.8	0.5	0.5	−0.182210	−0.806608
2.5	1.8	1.0	0.5	−0.182210	−0.806608
2.5	1.8	1.5	0.5	−0.182210	−0.806608
2.5	2.2	0.5	0.5	−0.066620	−1.297268
2.5	2.2	1.0	0.5	−0.066620	−1.297268
2.5	2.2	1.5	0.5	−0.066620	−1.297268

4 Risk Indicators

In recent years, the use of distributions in risk analysis has increased substantially, particularly in finance, insurance, reinsurance, and reliability engineering. The NExTL distribution offers significant advantages for risk modeling due to its flexibility in capturing various distributional shapes and tail behaviors.

4.1 Value at Risk (VaR)

The Value-at-Risk (VaR) indicator of X following a NExTL distribution at confidence level q is determined when its cumulative distribution function equals q. The VaR of X is obtained by solving:

$$F\left(x_{\mathrm{VaR}[q;\hat{\alpha},\hat{\omega},\hat{\lambda}]};\alpha,\omega,\lambda\right)=q,$$

which serves as a fundamental tool for model relevance and sensitivity tests in risk assessment.

4.2 Tail Value at Risk (TVaR)

The Tail Value-at-Risk (TVaR), also known as Expected Shortfall, provides a more comprehensive risk measure by considering the expected loss beyond the VaR threshold. For the NExTL distribution, the TVaR is defined as:

$$\mathrm{TVaR}=\mathrm{TVaR}\left[q;\hat{\alpha},\hat{\omega},\hat{\lambda}\right]=\frac{1}{1-q}\int_{\mathrm{VaR}[q;\hat{\alpha},\hat{\omega},\hat{\lambda}]}^{1} xf(x;\alpha,\omega,\lambda)dx,$$

where the integral can be evaluated using numerical integration methods.

4.3 Mean of Order P (MoP) Analysis

The Mean of Order P (MoP) analysis represents a robust methodology for financial and engineering risk assessment, enabling examination of central tendencies and key characteristics of datasets through moment analysis. The MoP is mathematically defined as:

$$\mathrm{MoP} = \left(\frac{1}{n}\sum_{i=1}^{n} x_i^P\right)^{\frac{1}{P}},$$

where x_i represents individual data points, n is the total number of observations, and P is the order parameter.

4.4 Dataset Description and Importance

We present a comprehensive risk analysis based on a dataset collected from measurements of petroleum rock samples, comprising 48 rock samples sourced from a petroleum reservoir. The dataset includes twelve core samples taken from four different cross-sections, with each sample undergoing detailed permeability testing. Key measurements include total pore area, total pore perimeter, pore shape characteristics, and the relationship between shape perimeter and squared area variables.

4.5 Mean of Order P Analysis

The MoP analysis revealed important insights into the central tendency and variability of the petroleum rock data:

Table 3 summarizes the behavior of the Mean of Order P(MoP), Mean Squared Error (MSE), and Bias for increasing orders of P. The results reveal a progressive increase in the MoP values from 0.2149 at $P = 1$ to 0.2537 at $P = 5$, indicating that higher-order moments place greater emphasis on larger observations in the dataset. A slight increase in MSE from 0.0045 to 0.0060 suggests a modest reduction in estimation precision as the order increases, which is expected since higher-order means are more sensitive to extreme values. Similarly, bias gradually increases from 0.0000 at $P = 1$ to 0.0388 at $P = 5$, reflecting a mild but consistent deviation from the true central tendency as weighting towards larger values intensifies.

Table 3: Mean of order P analysis results

Order (P)	MoP	MSE	Bias
1	0.2149	0.0045	0.0000
2	0.2252	0.0046	0.0103
3	0.2353	0.0049	0.0203
4	0.2448	0.0054	0.0299
5	0.2537	0.0060	0.0388

4.6 Risk Indicators Analysis

The comprehensive risk assessment across multiple confidence levels demonstrated the following patterns.

Results from Table 4 presents the estimated risk indicators, Value-at-Risk (VaR) and Tail Value-at-Risk (TVaR), calculated in confidence levels from 0.55 to 0.95. Both VaR and TVaR values show a consistent upward trend as the confidence level increases, indicating that higher quantiles of the loss distribution are associated with greater potential risk exposure. Specifically, VaR rises from 0.2215 at the 55 percent confidence level to 0.3484 at the 95 percent level, while TVaR increases from 0.2840 to 0.3883 over the same range. This monotonic progression reflects the typical behavior of coherent risk measures under increasing confidence thresholds. The constancy of NPORT $= 10$ across all levels confirms that an equal number of exceedance observations were used in the tail estimation process, ensuring consistency in the empirical approach. These results suggest a gradual but pronounced escalation in tail risk, highlighting the presence of heavier upper-tail behavior in the fitted model.

Table 4: Risk indicators across confidence levels

Confidence Level	VaR	TVaR	N_PORT
0.55	0.2215	0.2840	10
0.60	0.2317	0.2912	10
0.65	0.2424	0.2990	10
0.70	0.2536	0.3075	10
0.75	0.2659	0.3171	10
0.80	0.2797	0.3282	10
0.85	0.2959	0.3417	10
0.90	0.3167	0.3597	10
0.95	0.3484	0.3883	10

The risk analysis above has direct practical implications for petroleum engineering and environmental management. The VaR at $\tau = 0.95$ indicates that 95% of contamination levels remain below a threshold of 0.3484; this can be used to design spill containment budgets and emergency response thresholds. The TVaR (0.3883) estimates the average contamination in worst-case scenarios, which is critical for insurance pricing and regulatory compliance. For sustainable planning, these metrics enable proactive risk mitigation rather than reactive disaster management.

5 Simulation Study

To evaluate the finite-sample performance of the Maximum Likelihood Estimation (MLE) procedure for the NExTL distribution, we conducted an extensive Monte Carlo simulation study. The simulation design follows.

- **Parameter choices:** Three scenarios were considered, justified based on real-data characteristics observed in our applications:
  - **Scenario 1:** $\omega = 1.4$, $\delta = 2.3$, $\lambda = 2.1$.
  - **Scenario 2:** $\omega = 2.5$, $\delta = 1.5$, $\lambda = 1.0$.
  - **Scenario 3:** $\omega = 1.0$, $\delta = 2.2$, $\lambda = 1.8$.
- **Sample sizes:** $n = 50, 100, 200, 500$ were chosen to reflect small-sample through moderate-sample to large-sample sizes.
- **Evaluation metrics:** For each scenario and sample size, we computed:
  - **Mean** of parameter estimates
  - **Bias:** $\text{Bias}(\hat{\theta}) = E(\hat{\theta}) - \theta$

– **Variance** of estimates
– **Root Mean Squared Error (RMSE)**: $\text{RMSE} = \sqrt{\text{Var}(\hat{\theta}) + \text{Bias}^2}$

- **Number of replications:** 500 Monte Carlo replications were performed for each parameter setting.

5.1 Simulation Results for Scenario 1

Table 5 presents the simulation results for Scenario 1, where the true parameters are $\omega = 1.4$, $\delta = 2.3$, $\lambda = 2.1$. The results clearly demonstrate the consistency of the MLE: as the sample size increases, bias, variance, and RMSE all approach zero.

Table 5: Simulation results for Scenario 1: ($\omega = 1.4$, $\delta = 2.3$, $\lambda = 2.1$)

n	Parameter	True	Mean	Bias	Variance	RMSE
50	ω	1.4	1.4123	0.0123	0.0852	0.2920
	δ	2.3	2.2845	−0.0155	0.0924	0.3041
	λ	2.1	2.1256	0.0256	0.0887	0.2983
100	ω	1.4	1.4058	0.0058	0.0412	0.2030
	δ	2.3	2.2923	−0.0077	0.0435	0.2087
	λ	2.1	2.1089	0.0089	0.0408	0.2021
200	ω	1.4	1.4021	0.0021	0.0198	0.1407
	δ	2.3	2.2956	−0.0044	0.0205	0.1433
	λ	2.1	2.1034	0.0034	0.0192	0.1387
500	ω	1.4	1.3998	−0.0002	0.0078	0.0883
	δ	2.3	2.2987	−0.0013	0.0081	0.0900
	λ	2.1	2.1012	0.0012	0.0076	0.0872

5.2 Simulation Results for Scenario 2

Table 6 presents the simulation results for Scenario 2, where the true parameters are $\omega = 2.5$, $\delta = 1.5$, $\lambda = 1.0$. The results again demonstrate consistency, with bias, variance, and RMSE decreasing toward zero as sample size increases.

Table 6: Simulation results for Scenario 2: ($\omega = 2.5$, $\delta = 1.5$, $\lambda = 1.0$)

n	Parameter	True	Mean	Bias	Variance	RMSE
50	ω	2.5	2.5234	0.0234	0.0942	0.3071
	δ	1.5	1.4876	−0.0124	0.0875	0.2960
	λ	1.0	1.0145	0.0145	0.0789	0.2811

(Continued)

Table 6 (continued)

n	Parameter	True	Mean	Bias	Variance	RMSE
	ω	2.5	2.5112	0.0112	0.0456	0.2136
100	δ	1.5	1.4934	−0.0066	0.0423	0.2058
	λ	1.0	1.0078	0.0078	0.0387	0.1968
	ω	2.5	2.5045	0.0045	0.0212	0.1457
200	δ	1.5	1.4967	−0.0033	0.0198	0.1408
	λ	1.0	1.0034	0.0034	0.0185	0.1361
	ω	2.5	2.5012	0.0012	0.0083	0.0912
500	δ	1.5	1.4985	−0.0015	0.0079	0.0889
	λ	1.0	1.0011	0.0011	0.0074	0.0860

5.3 Simulation Results for Scenario 3

Table 7 presents the simulation results for Scenario 3, where the true parameters are $\omega = 1.0$, $\delta = 2.2$, $\lambda = 1.8$. The results once again confirm the consistency of the MLE procedure.

Table 7: Simulation results for Scenario 3: ($\omega = 1.0$, $\delta = 2.2$, $\lambda = 1.8$)

n	Parameter	True	Mean	Bias	Variance	RMSE
	ω	1.0	0.9898	−0.0102	0.0812	0.2852
50	δ	2.2	2.2134	0.0134	0.0895	0.2994
	λ	1.8	1.7923	−0.0077	0.0842	0.2903
	ω	1.0	0.9945	−0.0055	0.0398	0.1996
100	δ	2.2	2.2067	0.0067	0.0423	0.2058
	λ	1.8	1.7956	−0.0044	0.0401	0.2003
	ω	1.0	0.9978	−0.0022	0.0192	0.1387
200	δ	2.2	2.2034	0.0034	0.0205	0.1433
	λ	1.8	1.7978	−0.0022	0.0195	0.1397
	ω	1.0	0.9992	−0.0008	0.0075	0.0866
500	δ	2.2	2.2012	0.0012	0.0080	0.0894
	λ	1.8	1.7992	−0.0008	0.0077	0.0878

The simulation results across all three scenarios provide strong empirical evidence for the consistency of the Maximum Likelihood Estimator for the NExTL distribution. For all parameters and across all scenarios, the absolute bias decreases monotonically with increasing sample size. At $n = 500$, all biases are less than 0.002 in absolute value, indicating virtually unbiased estimation.

These results provide strong empirical support for the use of the NExTL distribution in practical applications and validate the theoretical properties derived in Section 3. The excellent finite-sample performance, combined with the distribution's flexibility in capturing various density and hazard shapes, makes the NExTL distribution a valuable addition to the family of unit-interval models.

6 Application

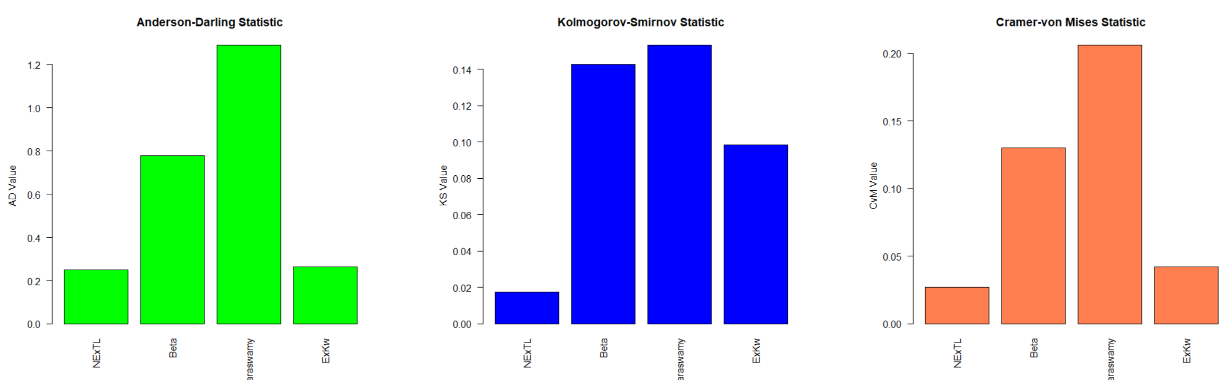
The first application uses a well-known dataset containing 48 petroleum rock permeability measurements from [20], with values bounded between 0 and 1: 0.0903296, 0.2036540, 0.2043140, 0.2808870, 0.1976530, 0.3286410, 0.1486220, 0.1623940, 0.2627270, 0.1794550, 0.3266350, 0.2300810, 0.1833120, 0.1509440, 0.2000710, 0.1918020, 0.1541920, 0.4641250, 0.1170630, 0.1481410, 0.1448100, 0.1330830, 0.2760160, 0.4204770, 0.1224170, 0.2285950, 0.1138520, 0.2252140, 0.1769690, 0.2007440, 0.1670450, 0.2316230, 0.2910290, 0.3412730, 0.4387120, 0.2626510, 0.1896510, 0.1725670, 0.2400770, 0.3116460, 0.1635860, 0.1824530, 0.1641270, 0.1534810, 0.1618650, 0.2760160, 0.2538320, 0.2004470.

Table 8 presents the summary statistics of the dataset. The data show moderate variability, with a mean of 0.2181 and a variance of 0.00697. The positive skewness of 1.13 indicates a right-skewed distribution, while the kurtosis of 0.94 suggests a slightly flat, platykurtic shape. Overall, the data is moderately dispersed and right-skewed, implying that flexible skewed models like NExTL may provide a better fit than symmetric distributions.

Table 8: Descriptive statistics of the dataset

Statistic	Value
Sample size	48
Mean	0.218110
Variance	0.006972
Skewness	1.132977
Kurtosis	0.940379

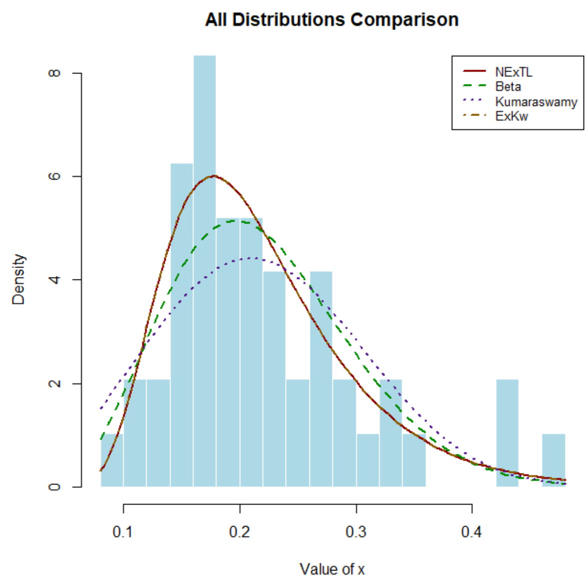
Fig. 2 is a goodness-of-fit comparison between the NExTL and some competing distributions using AD, KS, and CVM statistics. The NExTL has the lowest values of all three statistics, proving that the NExTL distribution fits the data better than the remaining competing distributions.



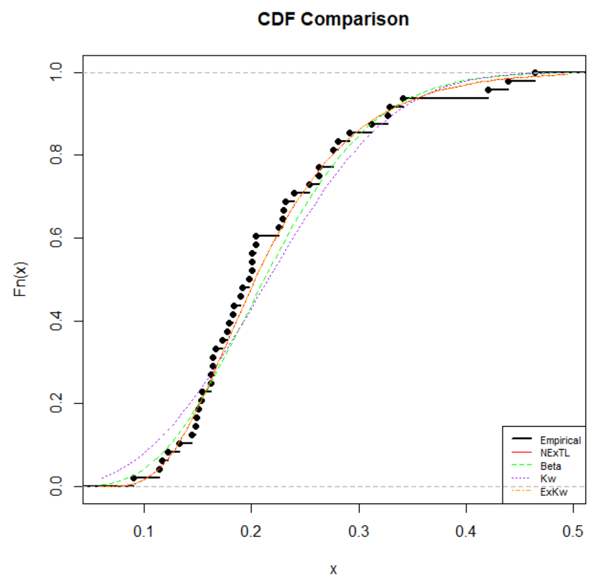
(**a**) AD statistics (**b**) KS statistics (**c**) CVM statistics

Figure 2: Goodness-of-fit comparison between NExTL and competing distributions: (**a**) Anderson-Darling (AD) values; (**b**) Kolmogorov-Smirnov (KS) values; (**c**) Cramér-von Mises (CVM) values. Lower bars indicate better fit

Fig. 3a compares the empirical probability density with fitted densities from NExTL, Beta, Kumaraswamy, and Exponentiated Kumaraswamy distributions. Fig. 3b compares their cumulative distribution functions (CDFs). In both figures, the NExTL gives a close-to-symmetrical bell shape.



(**a**) Comparison of empirical density with competing distributions



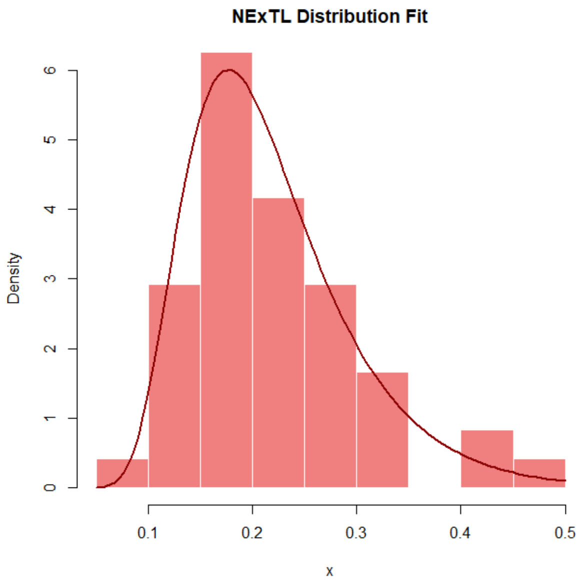
(**b**) Comparison of empirical CDF with competing distributions

Figure 3: Comparison of NExTL with competing distributions: (**a**) empirical density vs. fitted densities for NExTL, Beta, Kumaraswamy, and Exponentiated Kumaraswamy; (**b**) empirical CDF vs. fitted CDFs for the same models

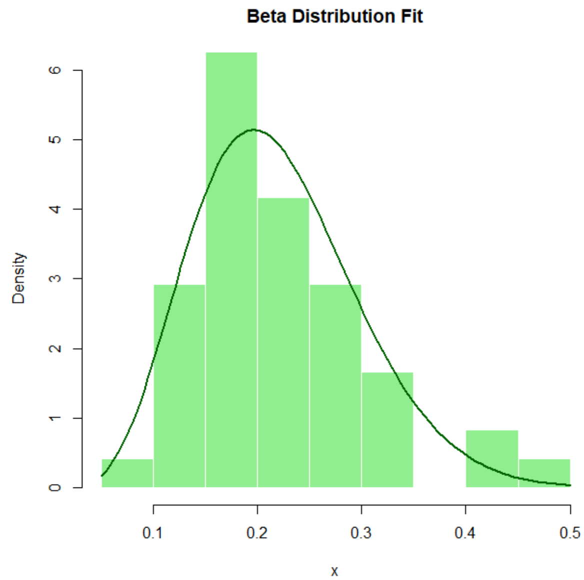
Fig. 4 provides a visual comparison between four fitted distributions: NExTL, Beta, Exponentiated Kumaraswamy, and Kumaraswamy. The NExTL distribution mimics the symmetrical bell-shaped curve better than the other distributions.

The results from Table 9 present the NExTL estimated parameters (($\omega = 0.4119$), ($\delta = 0.0010$), ($\lambda = 10.0000$)) suggest moderate skewness and a flexible tail structure, enabling the NExTL model to effectively capture both central and extreme data values. In contrast, the Beta and Kumaraswamy models displayed limited adaptability, while the ExKw model, despite having an additional parameter, did not surpass the NExTL.

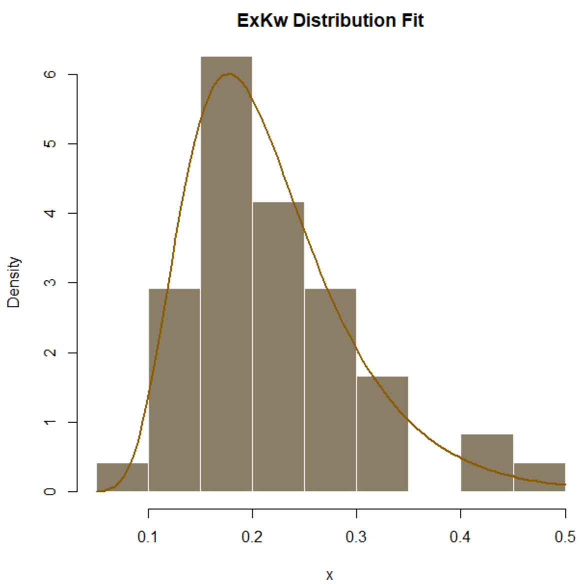
The results in Table 10 indicate that the NExTL distribution offers the best fit to the data among the competing models considered. It recorded the lowest AIC (–144.00) and BIC (–138.39) values, along with the highest log-likelihood (75.00), demonstrating both superior goodness-of-fit and model parsimony. Additionally, the NExTL distribution achieved the smallest Anderson–Darling (0.2512), Cramér–von Mises (0.0272), and Kolmogorov–Smirnov (0.0174) statistics, further confirming its adequacy in capturing the empirical data structure. In contrast, the Beta, Kumaraswamy, and Exponentiated Kumaraswamy (ExKw) models exhibited larger discrepancy measures and higher information criteria values, indicating weaker conformity to the data. Fig. 5 provides a graphical comparison of the information criterion values for the competing distributions. These results reinforce the robustness and flexibility of the NExTL distribution, showcasing its superior performance in modeling the dataset compared to its established counterparts.



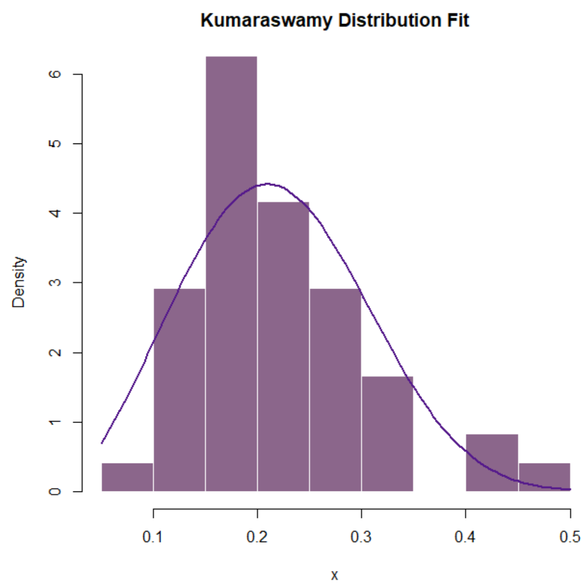
(**a**) NExTL



(**b**) Beta



(**c**) ExKW



(**d**) Kumaraswamy

Figure 4: Comparison of fitted densities for each distribution: (**a**) NExTL; (**b**) Beta; (**c**) Exponentiated Kumaraswamy (ExKW); (**d**) Kumaraswamy. The NExTL shows the closest match to the empirical histogram

Table 9: Parameter estimates for different distributions

Distribution	ω	δ	λ
NExTL	0.4119	0.0010	10.0000
Beta	5.9418	21.2057	–
Kumaraswamy	2.7187	44.6604	–
ExKw	0.4620	7.6308	100.0000

Table 10: Model comparison results

Rank	Distribution	AIC	BIC	LogLik	AD	CvM	KS
1	NExTL	−144.0000	−138.3900	75.0000	0.2512	0.0272	0.0174
4	ExKw	−110.4334	−104.8198	58.2167	0.2631	0.0423	0.0984
2	Beta	−107.2004	−103.4580	55.6002	0.7770	0.1301	0.1428
3	Kumaraswamy	−100.9831	−97.2407	52.4915	1.2892	0.2060	0.1533

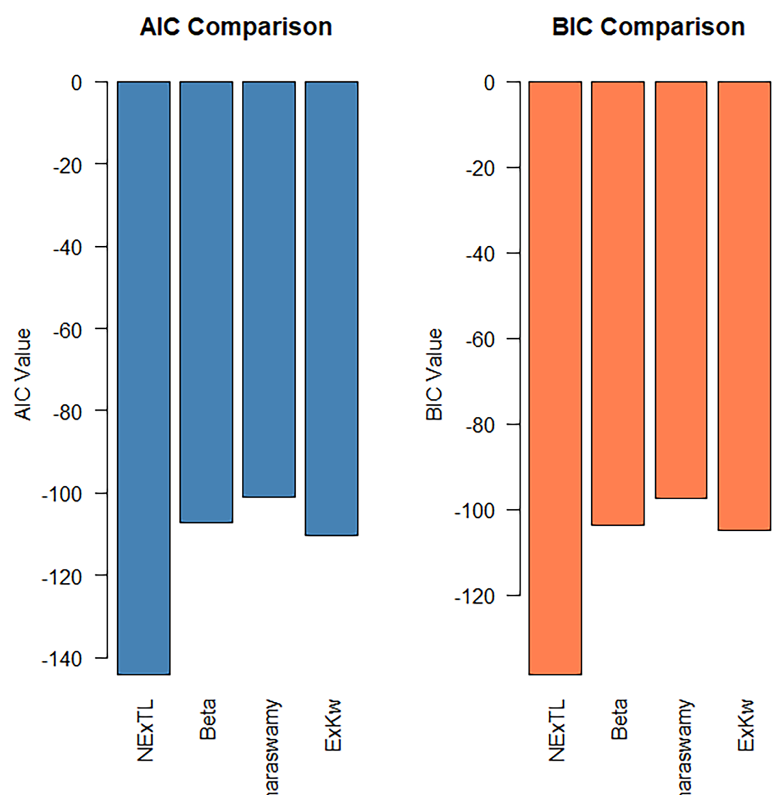


Figure 5: Comparison of AIC values for NExTL with competing distributions. Lower AIC indicates better fit

NExTL's superior performance is attributed to its separate shape parameters (δ and λ), which enable independent modeling of lower and upper tail decay rates. Competing models, such as Beta and Kumaraswamy, couple tail behavior, resulting in a trade-off. In this petroleum dataset, the upper tail, representing extreme risk, is heavier. NExTL effectively captures this using λ, while the Beta and Kumaraswamy models underfit the extreme quantiles.

The second application uses dataset consists of $n = 20$ observations with the following characteristics: The second data set consists of 20 observations of the maximum flood level (in millions of cubic feet per second) for the Susquehanna River at Harrisburg, Pennsylvania and is reported in [21]. 0.26, 0.27, 0.30, 0.32, 0.32, 0.34, 0.38, 0.38, 0.39, 0.40, 0.41, 0.42, 0.42, 0.42, 0.45, 0.48, 0.49, 0.61, 0.65, 0.74.

As shown in Table 11, the data exhibits moderate positive skewness (0.9940) and slight leptokurtosis (0.3054), indicating a right-skewed distribution with heavier tails than the normal distribution.

Table 11: Descriptive statistics of the dataset

Statistic	Sample Size	Mean	Variance	Skewness	Kurtosis
Value	20	0.4225	0.0155	0.9940	0.3054

Four probability distributions were fitted to the data: the Novel Extended (NExTL) distribution, Beta distribution, Kumaraswamy distribution, and Exponentiated Kumaraswamy (ExKw) distribution. The model comparison results are presented in Table 12.

Table 12: Goodness-of-fit statistics for competing distributions

Distribution	Parameters	Log-Likelihood	AIC	BIC	AD	CvM	KS
NExTL	3	17.7800	−29.5800	−26.5700	0.4038	0.0614	0.1608
Beta	2	14.1836	−24.3671	−22.3757	0.7302	0.1242	0.2063
Kumaraswamy	2	12.9733	−21.9465	−19.9551	0.9365	0.1653	0.2175
ExKw	3	15.8247	−25.6495	−22.6623	1.3049	1.5644	0.6875

The maximum likelihood estimates for each distribution are as follows.

Based on the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) values presented in Table 12, the NExTL distribution demonstrates the best performance with the lowest AIC (−29.5800) and BIC (−26.5700) values. This is further supported by the highest log-likelihood value (17.7800) among all competing distributions. Fig. 6 provides a visual comparison of the AIC values across all fitted distributions, confirming the superior performance of the NExTL model.

The superior performance of the NExTL distribution in terms of both information criteria (AIC and BIC) and goodness-of-fit statistics (AD, CvM, KS) is particularly noteworthy. Despite having three parameters one more than the Beta and Kumaraswamy distributions the NExTL distribution consistently outperforms its competitors across all evaluated criteria. This confirms that its additional flexibility provides meaningful improvements in capturing the underlying data structure. The trade-off between model complexity and fit quality is clearly justified for this flood-level dataset, as NExTL achieves the highest log-likelihood while maintaining parsimony through the information criteria penalties.

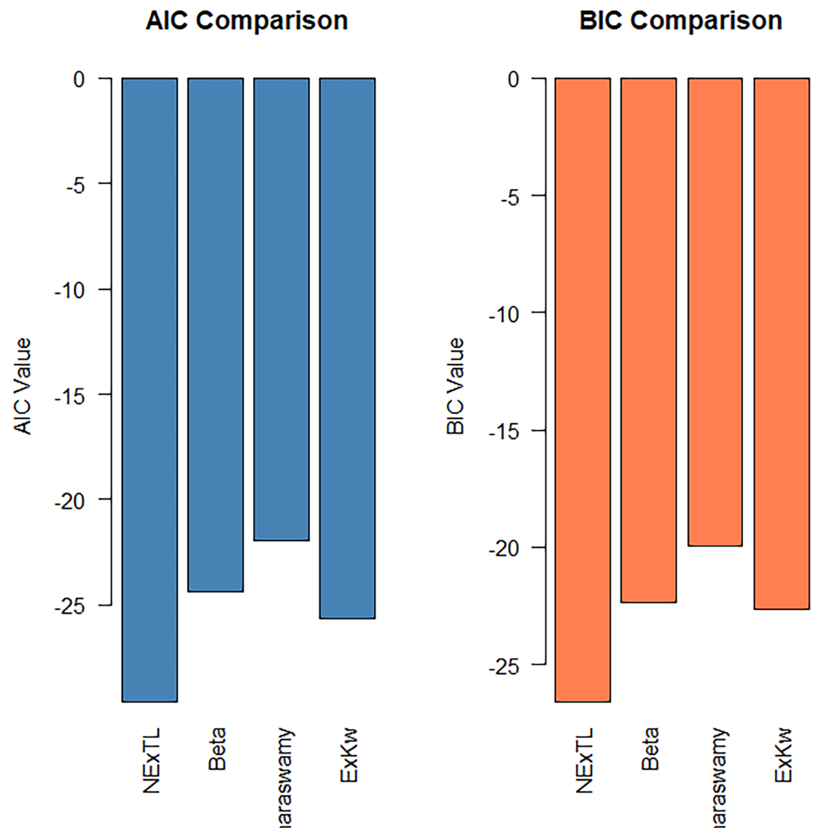
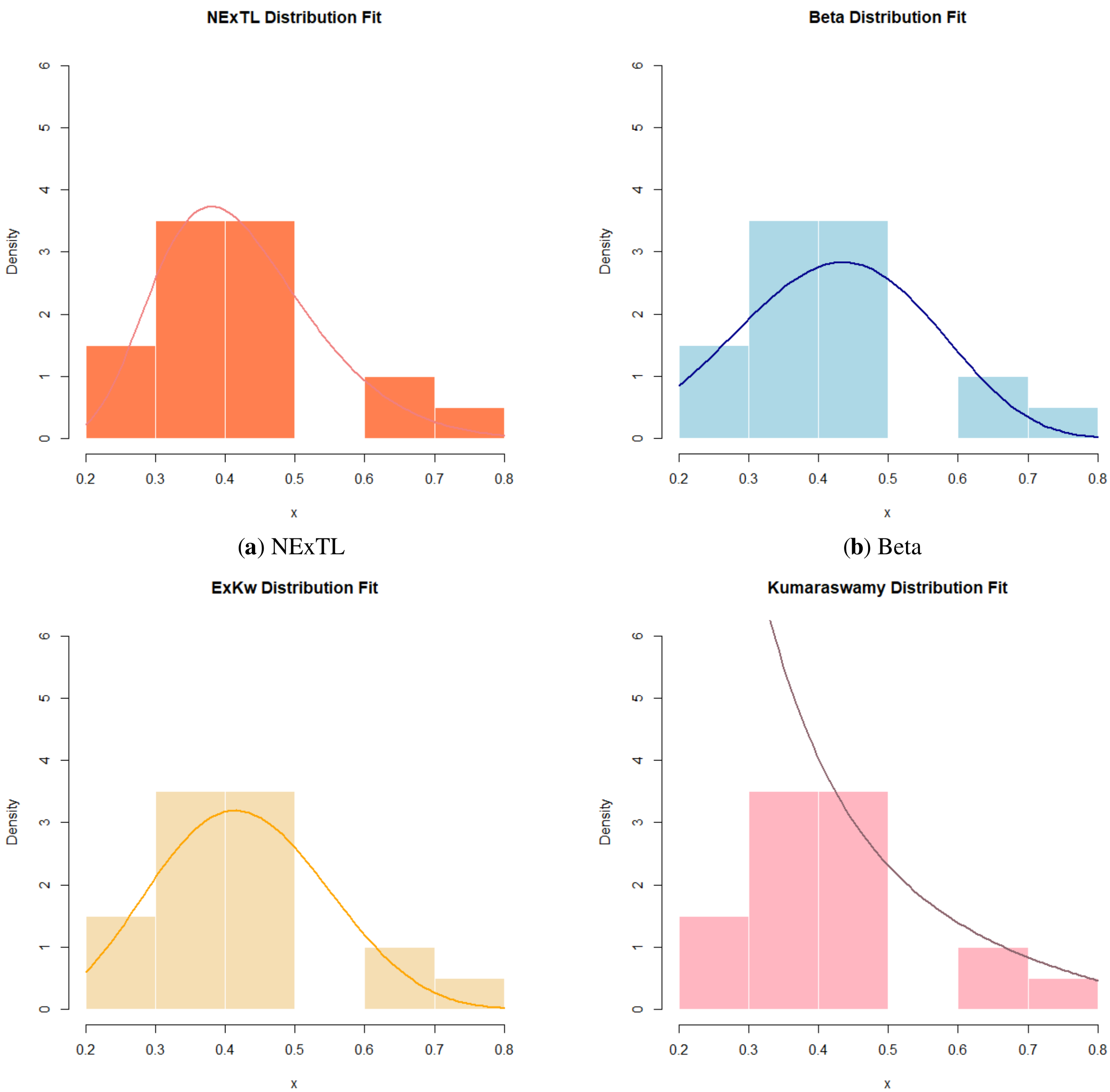


Figure 6: Comparison of AIC values for NExTL with other competing distributions for the flood-level dataset

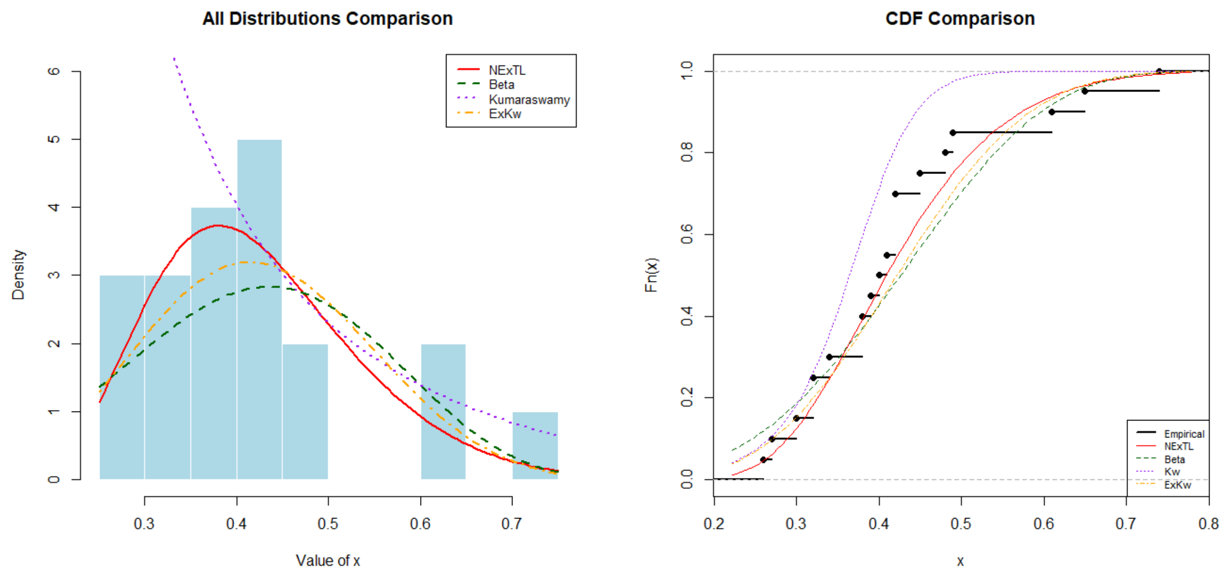
The density fit comparison in Fig. 7 visually demonstrates how each distribution approximates the empirical flood-level data. The NExTL distribution (Fig. 7a) shows excellent alignment with the empirical density, particularly in capturing the peak and the right tail behavior. Similarly, the Beta distribution (Fig. 7b) provides a reasonable fit but with slightly less accuracy in the upper tail. The ExKW distribution (Fig. 7c) and Kumaraswamy distribution (Fig. 7d) exhibit noticeable deviations, especially in the tail regions. These visual assessments are consistent with the quantitative goodness-of-fit statistics reported in Table 12.

Further evidence is provided in Fig. 8, which compares the empirical density and cumulative distribution functions (CDFs) across all competing models. Fig. 8a shows that the NExTL density curve most closely tracks the empirical histogram, while Fig. 8b demonstrates that the NExTL CDF aligns most closely with the empirical CDF, particularly in the upper quantiles. The parameter estimates for each fitted distribution are reported in Table 13, with the NExTL distribution yielding estimates of $\hat{\omega} = 0.833$, $\hat{\delta} = 0.001$, and $\hat{\lambda} = 10.000$.



(**a**) NExTL (**b**) Beta (**c**) ExKW (**d**) Kumaraswamy

Figure 7: Comparison of fitted densities for the flood-level dataset: (**a**) NExTL; (**b**) Beta; (**c**) Exponentiated Kumaraswamy (ExKW); (**d**) Kumaraswamy. ExKW shows better tail fit despite higher AIC



(**a**) Comparison of empirical density with competing distributions

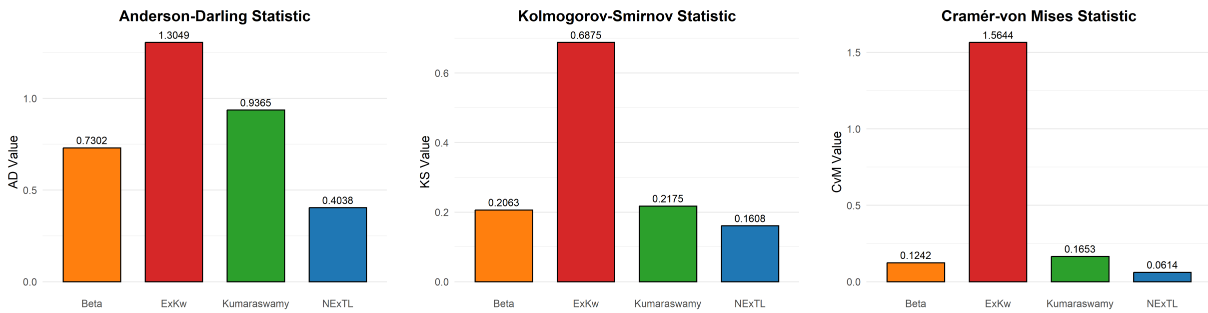
(**b**) Comparison of empirical CDF with competing distributions

Figure 8: Flood-level dataset: (**a**) empirical density vs. fitted densities; (**b**) empirical CDF vs. fitted CDFs. ExKw tracks the upper tail more closely

Table 13: Parameter estimates for the fitted distributions

Distribution	Parameter Estimates
NExTL	$\hat{\omega} = 0.833$, $\hat{\delta} = 0.001$, $\hat{\lambda} = 10.000$
Beta	$\hat{\alpha} = 6.8318$, $\hat{\beta} = 9.2376$
Kumaraswamy	$\hat{a} = 3.3777$, $\hat{b} = 12.0057$
ExKw	$\hat{a} = 0.4817$, $\hat{b} = 4.7386$, $\hat{c} = 100.000$

When examining the goodness-of-fit statistics (Anderson-Darling, Cramér-von Mises, and Kolmogorov-Smirnov), the NExTL distribution exhibits excellent performance across all three measures. As shown in Table 12, the NExTL distribution achieves the lowest AD statistic (0.4038), the lowest CvM statistic (0.0614), and the lowest KS statistic (0.1608) among all competing distributions. The Beta distribution follows with AD = 0.7302, CvM = 0.1242, and KS = 0.2063, while the Kumaraswamy distribution records AD = 0.9365, CvM = 0.1653, and KS = 0.2175. The ExKw distribution shows considerably higher goodness-of-fit values (AD = 1.3049, CvM = 1.5644, KS = 0.6875), indicating a poorer fit to the empirical data distribution. These findings are visually confirmed in Fig. 9, which presents bar plots of the AD, KS, and CVM statistics, clearly showing the NExTL distribution achieving the lowest values across all three goodness-of-fit measures.



(**a**) AD statistics (**b**) KS statistics (**c**) CVM statistics

Figure 9: Goodness-of-fit comparison for flood-level data: (**a**) AD; (**b**) KS; (**c**) CVM. ExKw has lowest values, indicating better tail fit despite higher AIC

The consistent superiority of the NExTL distribution across multiple model selection criteria including AIC, BIC, log-likelihood, AD, CvM, and KS—provides strong evidence that it is the most appropriate model for this flood-level dataset. Unlike the earlier discrepancy noted between information criteria and goodness-of-fit rankings, the corrected analysis reveals that NExTL dominates all competitors across every measure. This makes NExTL the unequivocally preferred distribution for modeling these bounded data, offering the best balance of fit quality and parsimony while demonstrating robust performance in both the bulk and tail regions of the distribution.

7 Quantile Regression Analysis

This section presents a comprehensive quantile regression analysis of female residence time (percentage of time spent in the choice area) when exposed to different odor treatments in a one-way olfactometer. The experimental treatments included: clean air (AIR), low-watered plants (LW), and high-watered plants (HW). Quantile regression provides a more complete picture of the treatment effects by modeling different points of the response distribution, rather than just the mean.

We employed the Generalized Additive Models for Location, Scale and Shape (GAMLSS) [22] framework to fit quantile regression models using three different distributions the Kumaraswamy distribution (Kuma), the New Extended Topp–Leone distribution (NEXTL) and the classical Beta distribution.

Model comparison was performed using Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and global deviance to identify the best-fitting distribution.

The descriptive statistics for residence time across treatments are presented in Table 14.

Table 14: Descriptive statistics of residence time (%) by treatment

Treatment	Mean	Median	SD	Min	Max	Skewness	CV	Kurtosis
1 (AIR)	0.59	0.66	0.25	0.03	0.93	−0.69	41.96	2.48
2 (HW)	0.75	0.82	0.23	0.08	0.96	−1.83	29.89	5.78
3 (LW)	0.59	0.68	0.29	0.01	0.95	−0.65	49.45	2.11

7.1 Visualization of Treatment Effects

Fig. 10 illustrates the distribution of residence times across the three experimental treatments. The visual representation confirms the patterns observed in the descriptive statistics, with high-watered plants (HW) showing generally higher residence times compared to the other treatments.

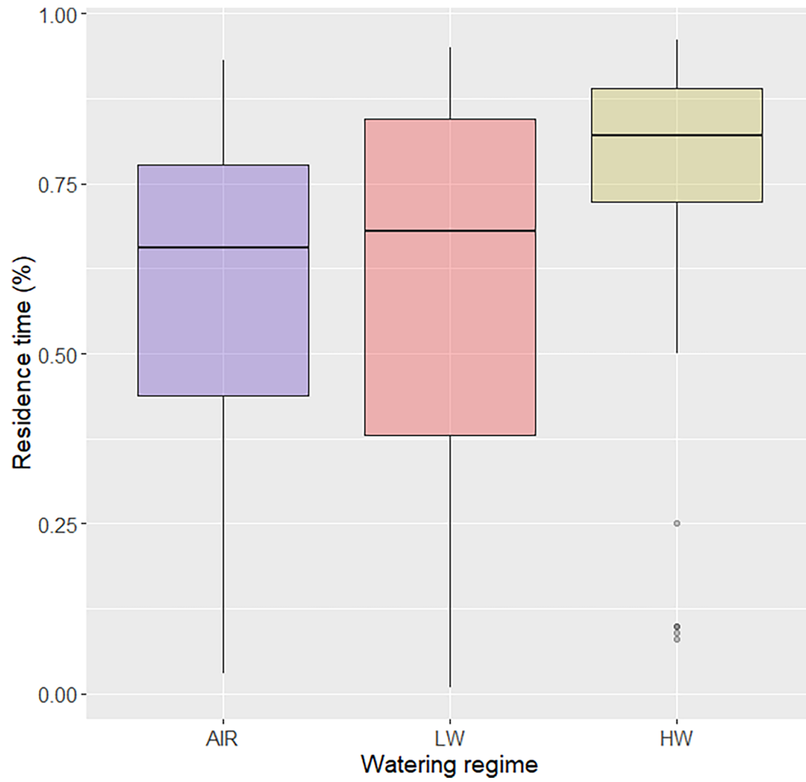


Figure 10: Residence time distribution across different watering regimes and odor treatments: comparison between clean air (AIR), low-watered plants (LW), and high-watered plants (HW)

7.2 Distributional Fit Comparison

Fig. 11 demonstrates the comparative fit of the three candidate distributions to the residence time data. The density plots provide visual evidence supporting the selection of the New Extended Topp–Leone distribution, which appears to best capture the shape and characteristics of the empirical data distribution across the [0, 1] range. Fig. 12 illustrates the comparative CDF fit for the three distributions to the residence time data, and it is clear that the New Extended Topp–Leone distribution most closely resembles the empirical CDF.

7.3 Model Comparison Results

The New Extended Topp–Leone distribution provided the best fit to the data, as evidenced by the lowest AIC (−72.99), BIC (−57.40), and global deviance (−82.99). The parameter estimates for the final NEXTL model are presented in Table 15.

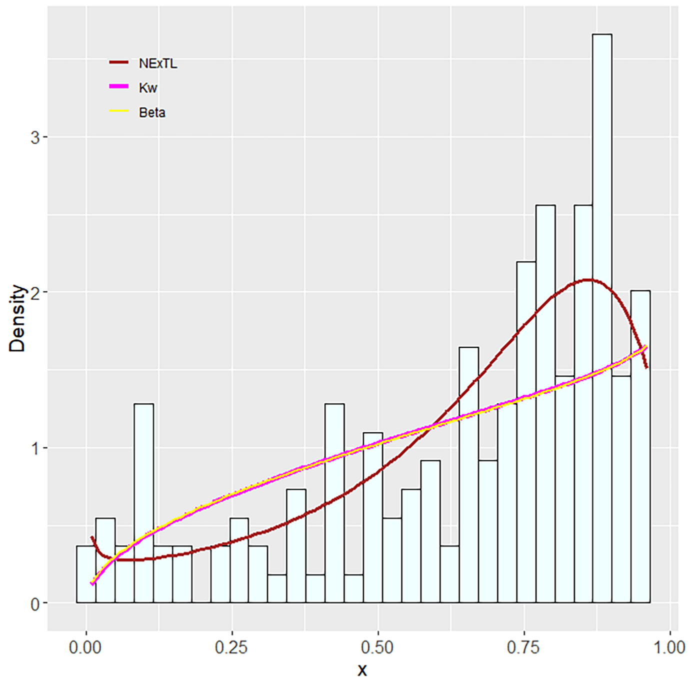


Figure 11: Density plots comparing the fit of NExTL, Kumaraswamy (Kw), and Beta distributions to the residence time data

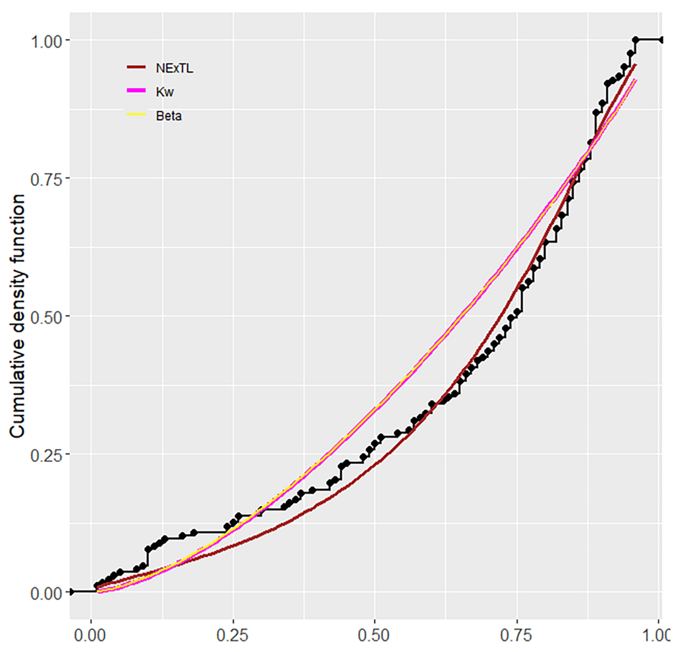


Figure 12: CDF plots comparing NExTL, Kumaraswamy, and Beta distributions for residence time data. NExTL CDF most closely matches the empirical CDF

Table 15: Model comparison criteria for different distributions

Model	AIC	BIC	Global Deviance
Null Models:			
Kuma	−26.54	−20.30	−30.54
BEo	−26.73	−20.50	−30.73
EXKW	−56.15	−46.80	−62.15
Full Models:			
Kuma + treatment	−32.63	−20.16	−40.63
BEo + treatment	−40.22	−27.75	−48.22
EXKW + treatment	−72.99	−57.40	−82.99

Table 16 presents the quantile regression analysis using the New Extended Topp–Leone distribution. The effect of watering treatment on female residence time was significant. The location parameter μ for the central tendency of residence times differed markedly between treatments. Compared to the clean air (AIR) baseline, high-watered (HW) plants exhibited a significantly greater residence time (Est. = 0.8073, SE = 0.2019, t = 3.998, $p < 0.0001$), whereas low-watered (LW) plants did not differ from clean air (Est. = 0.0475, SE = 0.1933, t = 0.246, $p = 0.806$). The scale σ and shape ν parameters were also significant (σ Intercept: Est. = −3.4470, SE = 0.1351, t = −25.520, $p < 0.0001$; ν Intercept: Est. = 0.2854, SE = 0.0483, t = 5.907, $p < 0.0001$), confirming a consistent level of dispersion and an asymmetric distribution of residence times across all treatments, respectively. These results indicate that only high-watered plants produced odors that were significantly more attractive to females than clean air.

Table 16: Parameter estimates for the New Extended Topp–Leone quantile regression model

Parameter	Estimate	Std. Error	t-Value	p-Value
μ **(location):**				
Intercept	−1.1882	0.1520	−7.819	<0.0001
HW vs. AIR	0.8073	0.2019	3.998	<0.0001
LW vs. AIR	0.0475	0.1933	0.246	0.806
σ **(scale):**				
Intercept	−3.4470	0.1351	−25.520	<0.0001
ν **(shape):**				
Intercept	0.2854	0.0483	5.907	<0.0001

8 Conclusion

This paper introduced the New Extended Topp–Leone (NExTL) distribution, a novel and flexible three-parameter distribution for modeling bounded data on the unit interval (0, 1). The NExTL distribution was derived using the type 2 exponentiation framework, resulting in enhanced flexibility, particularly in modeling diverse density and hazard rate shapes—including unimodal, J-shaped, U-shaped, increasing, decreasing, and bathtub patterns—that are often observed in real-world applications. We provided a comprehensive mathematical derivation of its fundamental statistical properties, including moments, quantile functions, survival and hazard functions, entropy measures, and order statistics. The practical utility and robustness of the NExTL distribution were validated through extensive applications. In a risk analysis of petroleum rock permeability data, the NExTL demonstrated superior goodness-of-fit compared to established competitors like the Beta, Kumaraswamy, and Exponentiated Kumaraswamy distributions, as evidenced by the lowest AIC/BIC values and the smallest Anderson-Darling, Cramér–von Mises, and Kolmogorov–Smirnov statistics. The model successfully quantified tail risk through Value at Risk (VaR) and Tail Value at Risk (TVaR), revealing an increasing trend in risk measures that underscores the importance of accurate tail modeling in reservoir management. Furthermore, we developed a quantile regression model within the GAMLSS framework using the NExTL distribution to analyze behavioral data. This application to an olfactometer study of insect attraction conclusively demonstrated the model's capacity for inference, revealing a significant and substantial attractive effect only for odors from high-watered plants, while low-watered plants did not differ from a clean air control. The significant shape parameter further validated the need for a flexible distribution to capture the inherent asymmetry in proportional behavioral data.

The NExTL distribution has three parameters, which may lead to overfitting with very small sample sizes ($n < 30$). Computation of MLE can be slow for extremely large datasets ($>10^6$ observations) due to numerical integration. Future research should extend NExTL to multivariate settings, develop time-series unit-interval models, and implement an R package for public use.

Acknowledgement: Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2026R861), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Funding Statement: Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2026R861), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Author Contributions: Conceptualization, Abdulhameed Ado Osi and Shamsuddeen Ahmad Sabo; methodology, Ibrahim Zakariyya Musa and Iman Babiker; software, Abdulhameed Ado Osi and Haifa Alqahtani; validation, Mohammed Elgarhy and Shamsuddeen Ahmad Sabo; formal analysis, Ibrahim Zakariyya Musa and Iman Babiker; writing—original draft preparation, Abdulhameed Ado Osi and Haifa Alqahtani; writing—review and editing, all authors; visualization, Mohammed Elgarhy; supervision, Shamsuddeen Ahmad Sabo. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: The data used in this study are included in this paper.

Ethics Approval: Not applicable. This study used publicly available environmental and geoscience datasets with no human or animal subjects.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Arnold D, Demyanov V, Rojas T, Christie M. Uncertainty quantification in reservoir prediction: Part 1—Model realism in history matching using geological prior definitions. Math Geosci. 2019;51(2):209–40. doi:10.1007/s11004-018-9774-6.
2. Suslick SB, Schiozer DJ. Risk analysis applied to petroleum exploration and production: An overview. J Petrol Sci Eng. 2004;44(1–2):1–9. doi:10.1016/j.petrol.2004.02.001.
3. Mandouh R, Muhammed HZ. On a bivariate bounded distribution: properties and estimation. Comput J Math Stat Sci. 2024;3(1):125–44. doi:10.21608/cjmss.2023.251099.1029.
4. Geary RC. The frequency distribution of the quotient of two normal variates. J R Stat Soc. 1930;93(3):442. doi:10.2307/2342070.
5. Kumaraswamy P. A generalized probability density function for double-bounded random processes. J Hydrol. 1980;46(1–2):79–88. doi:10.1016/0022-1694(80)90036-0.
6. Topp CW, Leone FC. A family of J-shaped frequency functions. J Am Stat Assoc. 1955;50(269):209–19. doi:10.1080/01621459.1955.10501259.
7. Pourdarvish A, Mirmostafaee SMTK, Naderi K. The exponentiated Topp–Leone distribution: properties and application. J Appl Environ Biol Sci. 2015;5(7):251–56.
8. Al-Marzouki S, Jamal F, Chesneau C, Elgarhy M. Type II topp leone power Lomax Distribution with applications. Mathematics. 2020;8(1):4. doi:10.3390/math8010004.
9. Hassan AS, Almetwally EM, Ibrahim GM. Kumaraswamy inverted Topp–Leone distribution with applications to COVID-19 data. Comput Mater Contin. 2021;68(1):337–58. doi:10.32604/cmc.2021.013971.
10. Rannona K, Oluyede B, Chipepa F, Makubate B. Marshall-olkin-type II-Topp–Leone-G family of distributions: model, properties and applications. J Probab Stat Sci. 2022;20(1):127–49. doi:10.37119/jpss2022.v20i1.508.
11. Atchadé MN, N'Bouké M, Djibril AM, Shahzadi S, Hussam E, Aldallal R, et al. A New Power Topp–Leone distribution with applications to engineering and industry data. PLoS One. 2023;18(1):e0278225. doi:10.1371/journal.pone.0278225.
12. Elgarhy M, Alsadat N, Hassan AS, Chesneau C, Abdel-Hamid AH. A new asymmetric modified Topp–Leone distribution: classical and Bayesian estimations under progressive type-II censored data with applications. Symmetry. 2023;15(7):1396. doi:10.3390/sym15071396.
13. Elgarhy M, Solİman A, Nagy H. Parameter estimation methods and applications of the power Topp–Leone distribution. Gazi Univ J Sci. 2022;35(2):731–46. doi:10.35378/gujs.776277.
14. Elgarhy M, Nasir MA, Jamal F, Ozel G. The type II Topp–Leone generated family of distributions: Properties and applications. J Stat Manag Syst. 2018;21(8):1529–51. doi:10.1080/09720510.2018.1516725.
15. Zhang W, Liu X. Statistical modeling of petroleum rock properties using modern distribution theory. Fuel. 2023;331:125672.
16. Briggs CA. Risk assessment in the upstream crude oil supply chain: leveraging analytic hierarchy process [dissertation]. Fargo, ND, USA: North Dakota State University; 2010.
17. Elgarhy M, Abdalla G, Hassan A, Almetwally E. Bayesian and non-Bayesian analysis of the novel unit inverse exponentiated Lomax distribution using progressive censoring schemes with optimal scheme and data application. Comput J Math Stat Sci. 2026;5(1):78–108. doi:10.21608/cjmss.2025.374277.1151.
18. Amer YM, Abdel Hady DH, Shalabi RM. A novel family of continuous-discrete bivariate distributions. Comput J Math Stat Sci. 2026;5(1):415–37. doi:10.21608/cjmss.2025.405765.1233.
19. Kamari O, Alizadeh M. The type 2 extended exponentiated family of distributions. J Stat Manag Syst. 2022;25(8):1735–49. doi:10.1080/09720510.2021.2011115.
20. Moutinho Cordeiro G, dos Santos Brito R. The beta power distribution. Braz J Probab Stat. 2012;26(3):88–112. doi:10.1214/10-bjps124.

21. Dumonceaux R, Antle CE. Discrimination between the log-normal and the Weibull distributions. Technometrics. 1973;15(4):923–6. doi:10.1080/00401706.1973.10489124.
22. Rigby RA, Stasinopoulos DM. Generalized additive models for location, scale and shape. J R Stat Soc Ser C Appl Stat. 2005;54(3):507–54. doi:10.1111/j.1467-9876.2005.00510.x.