



Research article

A robust two-component lifetime model with a bathtub-shaped hazard: A Bayesian study using Hamiltonian Monte Carlo simulation

M. G. M. Ghazal^{1,2,*}, Ghadah Alomani³, Khalaf S. Sultan⁴ and Hossam M. M. Radwan¹

¹ Department of Mathematics, Faculty of Science, Minia University, Minia 61519, Egypt

² University of Technology and Applied Sciences, Rustaq College of Education, Sultanate of Oman

³ Department of Mathematical Sciences, College of Science, Princess Nourah Bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia

⁴ Mathematics Department, Faculty of Science, Al-Azhar University, Nasr City 11884, Cairo, Egypt

* **Correspondence:** Email: mohamed.gamal@utas.edu.om.

Abstract: Accurately modeling lifetime data with non-monotonic hazard rates, particularly bathtub-shaped patterns, remains an important problem in lifetime analysis. This work introduced the additive exponential power (AEP) model, a four-parameter distribution obtained by combining two exponential power hazard components within an additive framework. The proposed model is capable of representing increasing and non-monotone hazard functions, including bathtub-shaped patterns, which makes it suitable for a wide range of lifetime behaviors. Several distributional and reliability properties of the AEP model were derived to establish its theoretical basis. Parameter estimation was developed using both maximum likelihood and Bayesian approaches, with Bayesian computation carried out through Hamiltonian Monte Carlo simulation to facilitate posterior inference under parameter dependence. The performance of the proposed estimation procedures was examined through Monte Carlo simulation, and the practical behavior of the model was illustrated using three real datasets with comparisons to existing lifetime models. The results indicated that the AEP model provides a flexible and useful addition to the family of statistical models for analyzing lifetime data.

Keywords: additive exponential power model; bathtub-shaped hazard rate; reliability analysis; Bayesian inference; Hamiltonian Monte Carlo

Mathematics Subject Classification: 60E05, 62E15, 62E20, 62F10, 62F15

1. Introduction

The hazard rate (HR) function, also known as the failure rate, is a fundamental concept in reliability and survival analysis. It provides a mathematical description of the instantaneous risk of failure

at time t , conditional on survival up to that point. As such, the HR function plays an important role in understanding lifetime behavior, assessing system reliability, and supporting maintenance and replacement decisions. Depending on the underlying failure mechanism, the HR may be increasing, decreasing, constant, or non-monotone. Among the non-monotone patterns, the bathtub-shaped hazard curve is particularly important in reliability theory and applications [1].

The bathtub curve typically consists of three phases: (i) An early-failure or burn-in phase, characterized by a decreasing HR due to initial defects or imperfections; (ii) a useful-life phase, during which the HR remains approximately constant; and (iii) a wear-out phase, in which the HR increases because of aging and deterioration. Although this pattern is frequently encountered in engineering and survival data, its statistical modeling remains challenging. Classical lifetime models, such as the Weibull model [2], are often unable to capture all three phases simultaneously, particularly when a relatively flat middle region is present. Consequently, the development of flexible statistical models capable of accommodating a wide variety of HR shapes has remained an active area of research for many years [3–5].

One important strategy for constructing flexible non-monotone lifetime models is the additive hazard approach. In this framework, the lifetime of a system is assumed to be influenced by more than one latent failure mechanism. If X_1 and X_2 denote two component lifetimes, then the system lifetime may be represented as $T = \min(X_1, X_2)$, which is analogous to a series-system structure. Under the assumption of independence, the overall hazard rate is given by the sum of the component hazard rates,

$$h(t) = h_1(t) + h_2(t),$$

while the survival function is the product of the component survival functions,

$$S(t) = S_1(t)S_2(t).$$

This representation is especially appealing when one component describes an early random failure mechanism and the other represents a wear-out mechanism associated with long-term aging. By combining baseline models from the same or different families, the additive hazard framework can generate hybrid models with enhanced flexibility for modeling a broad range of hazard shapes, including bathtub-type behavior.

Over the past two decades, this idea has led to a rich class of additive lifetime models designed to accommodate increasingly flexible hazard-rate patterns. Almalki and Yuan [6] introduced the new modified Weibull distribution to model a variety of hazard shapes, while Thach and Bris [7] developed the additive Chen–Weibull distribution for reliability applications. Abba et al. [8] proposed a flexible additive Chen–Gompertz model for systems involving different failure mechanisms.

Related developments have explored alternative combinations of lifetime components. Abdullahi et al. [9] proposed an additive Chen–exponential power model with Bayesian inference for reliability analysis. Jiang and Murthy [10] studied reliability modeling based on two Weibull distributions, providing an early framework for combining distinct Weibull components, while Wang [11] introduced an additive Burr XII construction capable of representing bathtub-shaped failure rates.

Further extensions broadened the range of distributions used in flexible lifetime modeling. He et al. [12] developed an additive modified Weibull distribution, and Singh [13] proposed an additive Perks–Weibull model with a bathtub-shaped hazard rate. Tarvirdizade and Ahmadpour [14] extended the Chen distribution to provide additional flexibility for lifetime-data applications.

More recent additive constructions have continued this development. Xavier et al. [15] introduced an additive power-transformed half-logistic model for reliability applications. Méndez-González et al. [16] proposed an additive Perks distribution, and later developed an additive Chen distribution for modeling lifetime data [17]. Their subsequent Chen–Perks model further combined these hazard structures and investigated the resulting reliability properties [18].

Recent studies have also focused on complex failure mechanisms and censored reliability data. Jiang et al. [19] developed an improved modified Weibull distribution for predicting the reliability evolution of an aircraft lock mechanism. Abba et al. [20] proposed a robust two-component failure-time model for complete and censored reliability data, whereas Abba and Wang [21] developed a Bayesian failure-time model for one- and two-failure-mode systems using Hamiltonian Monte Carlo (HMC) simulation.

Bayesian computation and multi-risk modeling represent another recent direction in this literature. Abba et al. [22] developed a robust multi-risk reliability model using an HMC-based Bayesian methodology, while Amiru et al. [23] proposed the additive Dhillon–Chen distribution for failure-time data. More recently, Ghazal [24] employed Hamiltonian Monte Carlo for inference in the flexible exponential power–Weibull distribution with reliability applications. Collectively, these developments demonstrate the continuing role of additive constructions in extending hazard-rate flexibility and motivating new computational approaches for lifetime and reliability analysis.

Despite this progress, existing additive models may still face practical limitations. In particular, some models have limited flexibility in representing a smooth transition across the different phases of the bathtub curve, while others may become difficult to estimate reliably when the parameters are strongly dependent or when the sample size is relatively small. These issues motivate the search for alternative constructions that retain flexibility while remaining statistically tractable.

The exponential power (EP) model, introduced by Smith and Bain [25], provides a useful building block in this context because of its ability to generate non-monotone hazard shapes, including bathtub-type behavior for suitable parameter values. Its HR function is given by

$$h_{EP}(x) = \alpha\beta(\alpha x)^{\beta-1}e^{(\alpha x)^\beta}, \quad x > 0, \alpha, \beta > 0, \quad (1.1)$$

where α is a scale parameter and β is a shape parameter. The EP model is well-known for producing bathtub-shaped hazard rates when $0 < \beta < 1$. Several extensions based on the EP family have also been proposed in the literature. For example, Abba et al. [20] introduced the exponential power–Gompertz model, while Abdullahi et al. [9] developed the additive Chen–exponential power model. These contributions highlight the potential of EP-based constructions for flexible lifetime modeling.

Motivated by the above considerations, this paper introduces the additive exponential power (AEP) model, which combines two EP hazard components within an additive framework. The proposed model is intended to provide a flexible four-parameter distribution capable of representing increasing and non-monotone hazard functions, including bathtub-shaped patterns. In this construction, the two components may be interpreted as representing different latent failure mechanisms, such as an early-life failure effect and a later wear-out effect.

In addition to introducing the model, the paper develops its main statistical and reliability properties and studies parameter estimation from both frequentist and Bayesian perspectives. Maximum likelihood estimation is considered for point estimation, while Bayesian inference is implemented using Hamiltonian Monte Carlo (HMC) with the No-U-Turn sampler (NUTS), which is particularly

useful for posterior computation in the presence of parameter dependence. The finite-sample performance of the resulting estimators is examined through Monte Carlo simulation, and the practical behavior of the AEP model is illustrated through real-data applications and comparisons with existing lifetime models.

The main contributions of this paper can be summarized as follows. First, a new additive lifetime model based on two exponential power hazard components is proposed. Second, several distributional and reliability properties of the model are derived. Third, parameter estimation is developed under both maximum likelihood and Bayesian frameworks. Fourth, simulation and real-data analyses are used to examine the practical performance of the proposed model in comparison with existing competitors.

The remainder of this paper is organized as follows. Section 2 introduces the AEP model and presents its main properties. Section 3 discusses parameter estimation under the maximum likelihood and Bayesian frameworks. Section 4 reports the simulation study. Section 5 presents the real-data applications and comparative results. Finally, Section 6 concludes the paper.

2. The additive exponential power model and its properties

This section introduces the AEP model and examines its key statistical and reliability properties. We begin by constructing the model and discussing the shapes of its hazard function. Subsequently, we derive analytical results for its quantiles, mode, and moments. Additional measures, including incomplete moments, Rényi entropy, mean time to failure, mean residual life, and order statistics, are also derived. These results highlight the flexibility and practical utility of the AEP model in reliability and survival analysis.

2.1. AEP model construction

The additive methodology states that the hazard function of a newly constructed model can be expressed as the sum of the hazard functions of its component models. For the AEP model, we combine two exponential power (EP) hazard components with potentially different parameter values, each representing a different failure mechanism. Let $\psi = (\alpha, \beta, \gamma, \lambda)$ denote the parameter vector, where $\alpha, \gamma > 0$ are scale parameters and $\beta, \lambda > 0$ are shape parameters. The HR function of the AEP model is given by

$$h(x; \psi) = h_{EP1}(x) + h_{EP2}(x), \quad (2.1)$$

where

$$h_{EP1}(x) = \alpha\beta(\alpha x)^{\beta-1}e^{(\alpha x)^\beta} \quad \text{and} \quad h_{EP2}(x) = \gamma\lambda(\gamma x)^{\lambda-1}e^{(\gamma x)^\lambda}.$$

Here, $h_{EP1}(x)$ can represent an initially decreasing component associated with early-life failures, while $h_{EP2}(x)$ can represent an increasing component associated with wear-out failures. This additive structure provides the AEP model with exceptional flexibility, enabling it to capture increasing and non-monotone hazard-rate patterns, including bathtub-type and more complex shapes. Such versatility is particularly valuable in reliability engineering, where complex failure patterns often require advanced modeling approaches.

The corresponding HR function for the AEP model is

$$h(x; \psi) = \alpha\beta(\alpha x)^{\beta-1}e^{(\alpha x)^\beta} + \gamma\lambda(\gamma x)^{\lambda-1}e^{(\gamma x)^\lambda}, \quad x > 0. \quad (2.2)$$

The probability density function (PDF) is obtained from the relationship $f(x; \psi) = h(x; \psi) \cdot S(x; \psi)$, where $S(x; \psi)$ is the survival function. Evaluating the cumulative hazard, $H(x; \psi) = \int_0^x h(\tau; \psi) d\tau$, yields

$$H(x; \psi) = e^{(\alpha x)^\beta} + e^{(\gamma x)^\lambda} - 2. \quad (2.3)$$

Consequently, the survival function is

$$S(x; \psi) = e^{(2 - e^{(\alpha x)^\beta} - e^{(\gamma x)^\lambda})} \quad (2.4)$$

and the PDF becomes

$$f(x; \psi) = (\alpha\beta(\alpha x)^{\beta-1} e^{(\alpha x)^\beta} + \gamma\lambda(\gamma x)^{\lambda-1} e^{(\gamma x)^\lambda}) e^{(2 - e^{(\alpha x)^\beta} - e^{(\gamma x)^\lambda})}. \quad (2.5)$$

Finally, the cumulative distribution function (CDF) follows directly as

$$F(x; \psi) = 1 - e^{(2 - e^{(\alpha x)^\beta} - e^{(\gamma x)^\lambda})}. \quad (2.6)$$

From a structural perspective, the AEP model differs from several existing additive lifetime models in an important way. In the additive Weibull (AW) model, the two Weibull components are individually monotone, and the bathtub behavior arises from their additive interaction. Similarly, the flexible additive Chen–Gompertz (FACG) model combines a Chen component with a Gompertz component, where the Gompertz component provides a monotone increasing contribution and can produce a relatively long low-level hazard region and a sharper wear-out transition [8,10]. In contrast, each component of the AEP model is an exponential-power hazard and can itself exhibit a bathtub shape when its corresponding shape parameter is less than one. Thus, the AEP construction allows non-monotone behavior to arise intrinsically at the component level, rather than relying exclusively on the interaction between a decreasing and an increasing component. Moreover, the scale parameters α and γ explicitly control the locations of the component minima, providing direct control over the timing of early-life and wear-out features. This structure gives the AEP model a distinct theoretical interpretation as a combination of two potentially non-monotone failure mechanisms, while retaining only four parameters.

In particular, for the AEP model, the i th component has its minimum at

$$x_i^* = \frac{1}{a_i} \left(\frac{1 - b_i}{b_i} \right)^{1/b_i}, \quad 0 < b_i < 1,$$

showing explicitly how its scale parameter determines the location of the component-level minimum. This provides a direct analytical characterization that is not available from the additive structure alone.

2.2. Identifiability and parameterization

Since the two components of the additive exponential-power hazard function are exchangeable, the parameter vectors $(\alpha, \beta, \gamma, \lambda)$ and $(\gamma, \lambda, \alpha, \beta)$ generate the same density, survival, and hazard functions. Consequently, without an identifying restriction, the model is subject to label-switching non-identifiability.

To resolve this issue, we impose the identifying constraint $\alpha < \gamma$. For both the classical maximum likelihood estimation and the Bayesian inference based on Hamiltonian Monte Carlo (HMC), this constraint is enforced directly through the following reparameterization:

$$\alpha = \exp(u), \quad \beta = \exp(v), \quad \gamma = \exp(u) + \exp(w), \quad \lambda = \exp(z), \quad (2.7)$$

where $u, v, w, z \in \mathbb{R}$. This parameterization guarantees

$$\alpha > 0, \quad \beta > 0, \quad \lambda > 0, \quad \gamma - \alpha = \exp(w) > 0,$$

and hence $\gamma > \alpha$. The identifiable parameterization in Eq (2.7) was used throughout the numerical estimation procedures, including both maximum likelihood optimization and Hamiltonian Monte Carlo (HMC) sampling. Consequently, all optimized and sampled parameter vectors satisfy the identifying constraint $\alpha < \gamma$, thereby avoiding label switching.

2.3. Regularity conditions for maximum likelihood inference

For the identified AEP model, the parameter space is defined as

$$\Theta = \{(\alpha, \beta, \gamma, \lambda) : \alpha > 0, \beta > 0, \gamma > \alpha, \lambda > 0\},$$

where the constraint $\alpha < \gamma$ is imposed to ensure identifiability. The support of the AEP distribution is $\mathcal{X} = (0, \infty)$, which is independent of the parameter vector. Hence, the parameter-independent support condition required for standard maximum likelihood asymptotic theory is satisfied. For every interior point of Θ , the AEP density and log-density are continuously differentiable with respect to the model parameters. The possible divergence of the hazard function as $x \rightarrow 0^+$ when $0 < \beta < 1$ or $0 < \lambda < 1$ does not, by itself, invalidate the regularity conditions at interior parameter values. Indeed, near the origin, the relevant terms in the density and its parameter derivatives are of the form $x^{c-1} |\log x|^k$, $c > 0$, where k is finite. These terms are integrable because

$$\int_0^\varepsilon x^{c-1} |\log x|^k dx < \infty, \quad c > 0.$$

Therefore, the likelihood and the corresponding score and information integrals remain finite for interior parameter values satisfying $\beta > 0$ and $\lambda > 0$.

The values $\beta = 0$ and $\lambda = 0$ are boundary points and are not included in the parameter space. Consequently, any singular behavior associated with these limiting values does not constitute a violation of the regularity conditions at an interior parameter point. For the asymptotic likelihood arguments, the analysis is considered locally on a compact neighborhood contained in the interior of Θ and bounded away from its boundaries.

Under these regularity conditions, together with the usual identifiability and nonsingularity conditions on the Fisher information matrix, standard maximum likelihood theory yields consistency and asymptotic normality of the MLE. In particular,

$$\sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \xrightarrow{d} N_4(\mathbf{0}, I(\boldsymbol{\theta}_0)^{-1}),$$

where $\boldsymbol{\theta} = (\alpha, \beta, \gamma, \lambda)^T$ and $I(\boldsymbol{\theta}_0)$ denotes the Fisher information matrix per observation evaluated at the true parameter vector $\boldsymbol{\theta}_0$. Accordingly, for sufficiently large samples, the covariance matrix of the

MLE is approximated by the inverse of the observed Fisher information matrix. These results follow from standard asymptotic likelihood theory; see, for example, Lehmann and Casella [26] and van der Vaart [27]. To empirically characterize the numerical conditioning of the Fisher information matrix, we computed its standardized (correlation-based) eigenstructure,

$$R(\theta) = D(\theta)^{-1}I(\theta)D(\theta)^{-1}, \quad D(\theta) = \text{diag}(\sqrt{I_{11}}, \sqrt{I_{22}}, \sqrt{I_{33}}, \sqrt{I_{44}}),$$

across a range of parameter configurations, including values near the boundary of Θ and the actual MLEs obtained for the Sylwia, VETEs, and DTPPs datasets (Table 1). This standardization reduces the influence of heterogeneous parameter scales and facilitates examination of the correlation structure among the parameters.

The numerical results reveal persistent strong dependence between the scale and shape parameters of the EP components across the parameter configurations examined, including the MLEs obtained from the three real datasets analyzed in Section 5. The standardized condition numbers remain large and nearly unchanged across these configurations, reflecting the near-collinearity induced by the very strong scale–shape correlation in one of the EP components. These findings suggest that the observed ill-conditioning reflects a persistent feature of the EP parameterization over the configurations considered, rather than a numerical anomaly specific to a particular dataset or boundary parameter value. This behavior is consistent with the known estimation difficulty of the exponential power distribution [25], where the scale and shape parameters can exhibit strong dependence in the likelihood.

Importantly, this high correlation does not invalidate the regularity conditions established above. The Fisher information matrix remains finite and nonsingular at the interior parameter configurations examined. Rather, the observed near-collinearity indicates that precise joint estimation of the scale and shape parameters may require larger samples, which is consistent with the relatively wide confidence intervals observed for some parameter estimates in the real-data applications.

Table 1. Correlation coefficients and standardized condition numbers of the Fisher information matrix at selected parameter configurations and at the MLEs fitted to the three real datasets analyzed in Section 5.

α	β	γ	λ	$\rho(\alpha, \beta)$	$\rho(\gamma, \lambda)$	Standardized condition number
0.002	0.50	0.03	5.00	−0.8777	0.999978	9.1817×10^4
0.002	0.10	0.03	5.00	−0.3841	0.999978	9.1817×10^4
0.002	0.05	0.03	5.00	0.1401	0.999978	9.1817×10^4
0.002	0.01	0.03	5.00	0.9827	0.999978	9.1817×10^4
0.002	0.50	0.03	0.10	0.999978	0.5927	9.1817×10^4
0.002	0.50	0.03	0.05	0.999978	0.5527	9.1817×10^4
0.002	0.50	0.03	0.01	0.999978	0.9846	9.1817×10^4
5.29×10^{-5} *	0.26	0.0808	4.09	−0.9192	0.999978	9.1817×10^4
1.37×10^{-3} †	0.66	0.0026	3.81	−0.6208	0.999978	9.1817×10^4
2.24×10^{-3} ‡	0.54	0.0273	5.39	−0.8765	0.999978	9.1817×10^4

*Note: * MLEs fitted to the Sylwia dataset (Table 9). † MLEs fitted to the VETEs dataset (Table 12). ‡ MLEs fitted to the DTPPs dataset (Table 16).

2.4. Shapes of the HR function

The HR function of the AEP model, defined as the sum of two EP hazard components, exhibits remarkable versatility in capturing various failure patterns observed in reliability applications. To analyze the shape characteristics, we consider the general form of each EP hazard component:

$$h_{EP_i}(x) = a_i b_i (a_i x)^{b_i-1} e^{(a_i x)^{b_i}}, \quad a_i > 0, b_i > 0, i = 1, 2,$$

with $(a_1, b_1) = (\alpha, \beta)$ and $(a_2, b_2) = (\gamma, \lambda)$.

The limiting behavior of each EP component is given by:

$$\lim_{x \rightarrow 0} h_{EP_i}(x) = \begin{cases} 0, & b_i > 1, \\ a_i, & b_i = 1, \\ +\infty, & 0 < b_i < 1, \end{cases} \quad \text{and} \quad \lim_{x \rightarrow \infty} h_{EP_i}(x) = +\infty.$$

Hence, the initial and asymptotic behavior of the AEP HR function is governed by the smaller of the two shape parameters, β and λ . Specifically, the hazard function approaches zero as $x \rightarrow 0$ if both $\beta > 1$ and $\lambda > 1$, while it diverges to infinity if either $0 < \beta < 1$ or $0 < \lambda < 1$.

The derivative of the combined hazard rate is given by

$$h'(x; \psi) = \alpha^2 \beta (\alpha x)^{\beta-2} e^{(\alpha x)^\beta} [(\beta - 1) + \beta (\alpha x)^\beta] + \gamma^2 \lambda (\gamma x)^{\lambda-2} e^{(\gamma x)^\lambda} [(\lambda - 1) + \lambda (\gamma x)^\lambda]. \quad (2.8)$$

For the individual exponential power hazard component, if $b_i \geq 1$, then $h_{EP_i}(x)$ is strictly increasing. In contrast, when $0 < b_i < 1$, $h_{EP_i}(x)$ initially decreases, attains a minimum at $x_i^* = \frac{1}{a_i} \left(\frac{1-b_i}{b_i} \right)^{1/b_i}$, and subsequently increases to infinity. Hence, the scale parameter a_i directly affects the location of the minimum.

For the combined hazard function, its shape cannot, in general, be determined solely from the values of β and λ . Instead, the monotonicity of the combined hazard depends on the sign of $h'(x)$ in Eq (2.8) over the interval of interest. When $\beta > 1$ and $\lambda > 1$, both components are strictly increasing, and consequently the combined hazard is strictly increasing. When one shape parameter is below one and the other is above one, the combined hazard necessarily tends to infinity as $x \rightarrow 0^+$ and as $x \rightarrow \infty$, and therefore has at least one minimum, resulting in a bathtub-type hazard. The precise location and depth of the minimum depend on all four parameters, particularly the scale parameters α and γ .

When $0 < \beta < 1$ and $0 < \lambda < 1$, both EP components are individually bathtub-shaped and each has a unique minimum. Consequently, the combined hazard is not globally decreasing. Its shape is determined by the positive roots of $h'(x) = 0$ and is therefore examined numerically. In particular, a minimum x_c is identified by $h'(x_c) = 0$ with a change in sign from negative to positive, and its depth is given by $h(x_c)$. The resulting critical-point equation is transcendental and generally has no closed-form solution; hence, the locations of the critical points are obtained numerically. The resulting classification should therefore be interpreted over the interval of interest.

Figures 1 and 2 illustrate the flexibility of the AEP model through selected numerical examples of the density and hazard rate functions. Figure 1 shows several representative density shapes, while Figure 2 illustrates increasing and non-monotone hazard patterns, including bathtub-shaped cases.

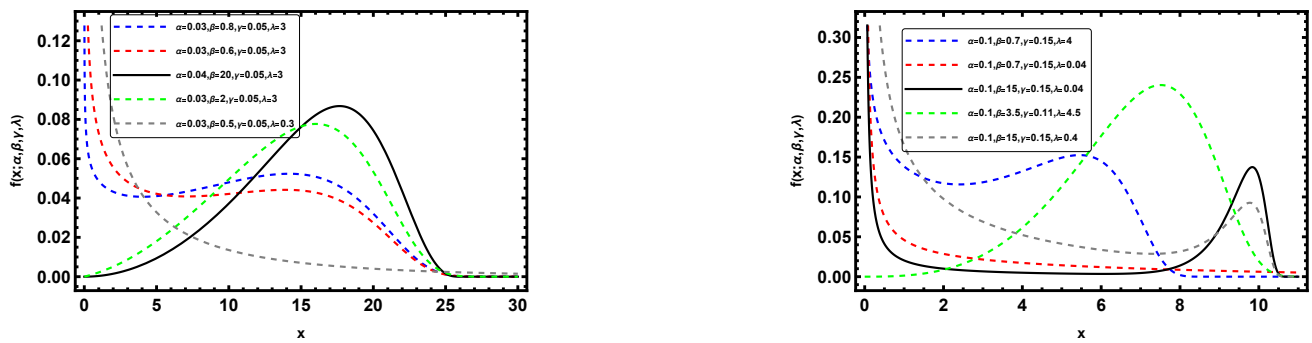
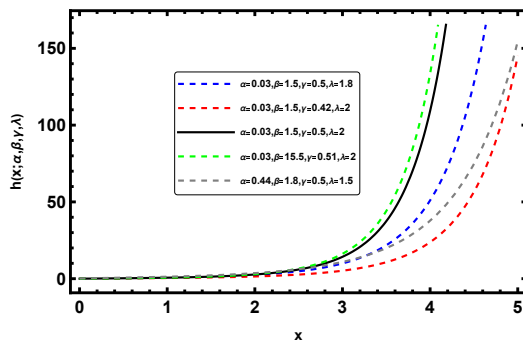
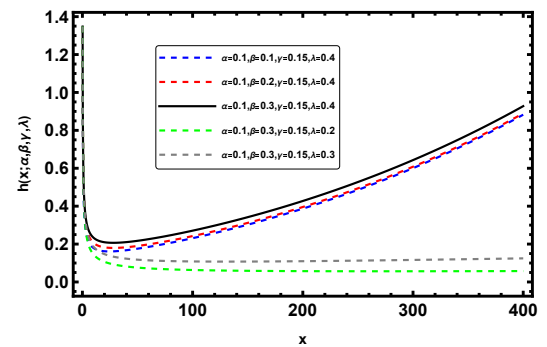


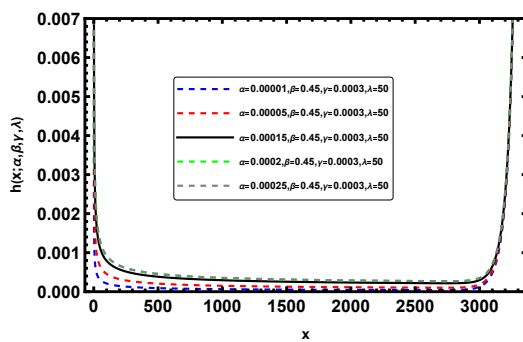
Figure 1. Plots of the AEP density function for selected parameter values.



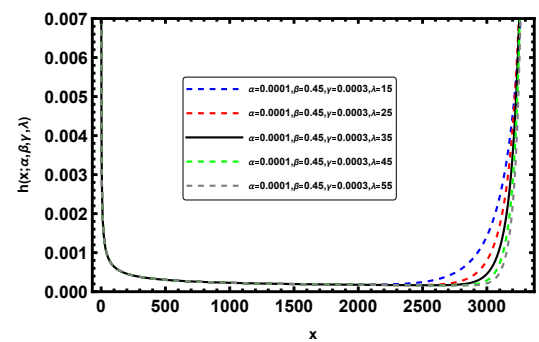
(a) Increasing; $\beta, \lambda > 1$



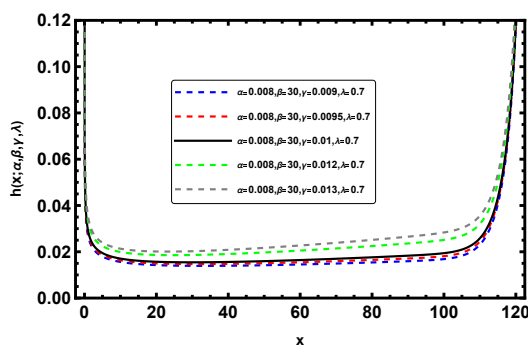
(b) Non-monotone hazard; $\beta, \lambda < 1$



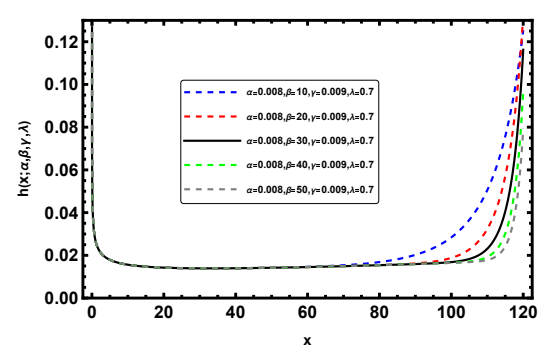
(c) Bathtub-shaped; $\beta < 1 < \lambda$



(d) Bathtub-shaped; $\beta < 1 < \lambda$



(e) Bathtub-shaped; $\beta > 1 > \lambda$



(f) Bathtub-shaped; $\beta > 1 > \lambda$

Figure 2. Plots of the AEP hazard rate function for selected parameter values.

2.5. Quantile function

The quantile function, denoted by $Q(u)$, is essential for statistical characterization, simulation, and reliability assessment. For the AEP model, the u -th quantile x_u is obtained by solving $F(x_u; \psi) = u$ for x_u . Using Eq (2.6), this reduces to:

$$e^{(\alpha x_u)^\beta} + e^{(\gamma x_u)^\lambda} = 2 - \ln(1 - u), \quad 0 < u < 1. \quad (2.9)$$

Due to the nonlinear combination of exponential power terms, Eq (2.9) lacks a closed-form solution and must be evaluated using numerical methods, such as the Newton–Raphson algorithm or the FindRoot function in Wolfram Mathematica. The quartiles (25th, 50th, and 75th percentiles) characterize the model's symmetry and dispersion, serving as key benchmarks for establishing warranty periods, failure thresholds, and maintenance schedules.

2.6. Mode

The mode of a probability distribution is the point where the PDF attains its maximum, representing the most probable failure time. For the AEP model, the mode is obtained by solving $f'(x; \psi) = 0$. Differentiating Eq (2.5) yields:

$$\begin{aligned} & \left[\beta \alpha^\beta x^{\beta-2} (\beta((\alpha x)^\beta + 1) - 1) e^{(\alpha x)^\beta} + \lambda \gamma^\lambda x^{\lambda-2} (\lambda((\gamma x)^\lambda + 1) - 1) e^{(\gamma x)^\lambda} \right. \\ & \quad \left. - \beta^2 \alpha^{2\beta} x^{2\beta-2} e^{2(\alpha x)^\beta} - \lambda^2 \gamma^{2\lambda} x^{2\lambda-2} e^{2(\gamma x)^\lambda} - 2\beta\lambda \alpha^\beta \gamma^\lambda x^{\beta+\lambda-2} e^{(\alpha x)^\beta + (\gamma x)^\lambda} \right] \\ & \quad \times e^{(2 - e^{(\alpha x)^\beta} - e^{(\gamma x)^\lambda})} = 0. \end{aligned} \quad (2.10)$$

Since Eq (2.10) has no closed-form solution, the mode must be determined numerically.

2.7. Moments

Theorem 1. The k -th raw moment about the origin, $\mu'_k = \mathbb{E}[X^k]$, for the AEP model is given by:

$$\mu'_k = \frac{ke^2}{\lambda\gamma^k} \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^n n^m}{n!m!} \left(\frac{\alpha}{\gamma}\right)^{m\beta} G\left(\frac{k + \beta m}{\lambda} - 1\right). \quad (2.11)$$

The integral $G(a)$, defined as $G(a) = \int_1^\infty \frac{(\ln t)^a}{t} e^{-t} dt$, is related to the upper incomplete gamma function $\Gamma(s, 1)$ and can be expressed by differentiating it as follows:

$$G(a) = \frac{d^a}{ds^a} \Gamma(s, 1) \Big|_{s=0}.$$

Proof. See A.1. □

2.8. Incomplete moments

Incomplete moments provide important information about the behavior of a random variable below a specified threshold, making them particularly useful in reliability analysis and risk assessment. Unlike complete moments, which characterize the entire distribution, incomplete moments focus on specific regions of interest, such as the early failure or warranty periods.

Theorem 2. The s -th incomplete moment of the AEP model can be expressed as:

$$\begin{aligned} \mu'_s(z) = & e^2 \beta \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \sum_{i=0}^{\infty} \frac{k^i \alpha^{(m+l+1)\beta} \gamma^{i\lambda} z^{s+(m+l+1)\beta+i\lambda}}{n! m! k! l! i! (s + (m+l+1)\beta + i\lambda)} \\ & + e^2 \lambda \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} \frac{(k+1)^j \alpha^{m\beta} \gamma^{(j+1)\lambda} z^{s+m\beta+(j+1)\lambda}}{n! m! k! j! (s + m\beta + (j+1)\lambda)}. \end{aligned} \quad (2.12)$$

Proof. See A.2. □

2.9. Rényi entropy

Rényi entropy provides a generalized measure of uncertainty in a probability model, extending the classical Shannon entropy. It is particularly valuable in reliability analysis, as it quantifies the unpredictability of failure times—higher entropy values indicate greater uncertainty in a system's lifetime.

Theorem 3. For a continuous random variable X with PDF $f(x; \psi)$, the Rényi entropy of order q ($q > 0, q \neq 1$) for the AEP model is given by:

$$H_q(X) = \frac{1}{1-q} \log \left[e^{2q} \sum_{k=0}^q \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{l=0}^{\infty} \binom{q}{k} \frac{(-q)^{i+j} \alpha^{(q+l-k)\beta} \beta^{q-k} \gamma^{k\lambda} \lambda^k}{i! j! l!} I_{k,j,l} \right],$$

where

$$I_{k,j,l} = \int_0^{\infty} x^{(q+l-k)\beta+k\lambda-q} e^{(k+j)\gamma x^\lambda} dx.$$

The integral $I_{k,j,l}$ generally has no closed-form analytical solution and must be evaluated numerically for specific parameter values.

Proof. See A.3. □

2.10. Mean time to failure

The mean time to failure (MTTF) is a fundamental reliability metric representing the expected lifetime of a non-repairable system.

Theorem 4. The MTTF for the AEP model is given by:

$$MTTF = \frac{e^2}{\gamma \lambda} \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^n n^m}{n! m!} \left(\frac{\alpha}{\gamma} \right)^{m\beta} G\left(\frac{\beta m + 1}{\lambda} - 1 \right), \quad (2.13)$$

where $G(a)$ is as defined in Theorem 1.

Proof. See A.4. □

The MTTF is crucial for assessing product reliability, setting warranty periods, and designing systems.

2.11. Mean residual life

The mean residual life (MRL) function, $m(t) = \mathbb{E}[X - t \mid X > t]$, quantifies the expected remaining lifetime given survival up to time t .

Theorem 5. *The MRL function for the AEP model is given by:*

$$m(t) = e^{e^{(\alpha t)^\beta} + e^{(\gamma t)^\lambda}} \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^n n^m}{n! m!} \left(\frac{\alpha}{\gamma}\right)^{m\beta} G\left(\frac{\beta m + 1}{\lambda} - 1, e^{(\gamma t)^\lambda}\right), \quad (2.14)$$

where $G(a, b) = \int_b^\infty \frac{(\ln t)^a}{t} e^{-t} dt$ is a generalized form of the G function.

Proof. See A.5. □

Table 2 reports numerical estimates of the MRL at selected time points and the MTTF for two parameter settings of the AEP model. Each value is accompanied by its Monte Carlo mean squared error (MSE). The results show that the MRL decreases as t increases, reflecting natural system aging. As expected, at $t = 0$, the MRL equals the MTTF, confirming the theoretical identity $\text{MTTF} = \text{MRL}(t)|_{t=0}$. The MSEs consistently decrease with larger sample sizes n , demonstrating improved precision and estimator stability. MTTF estimates remain stable across all sample sizes.

Figure 3 illustrates the MSE trajectories of the MRL and MTTF across different sample sizes. In both parameter settings, the MSE curves decline steadily as n increases, consistent with asymptotic theory. The most notable improvement occurs when moving from small to moderate samples, after which the reduction rate becomes more gradual, indicating diminishing gains for very large n . This behavior highlights the efficiency of the proposed estimation method and supports its use in practical reliability studies.

Finally, Figure 4 presents the trace plots and posterior density estimates of the MRL and MTTF. The trace plots show stable sampling behavior without evident systematic patterns, while the posterior densities are predominantly unimodal. These graphical results provide additional support for the numerical stability of the simulation procedure.

Table 2. Monte Carlo estimates of the MRL at different time points t and the MTTF for the AEP model under two parameter settings. The MSE for each estimate is reported in parentheses.

n	MRL at time t					MTTF
	$t = 0.1$	$t = 0.35$	$t = 0.8$	$t = 1.3$	$t = 1.7$	
$\alpha = 0.1, \beta = 2, \gamma = 0.5, \lambda = 1.5$						
50	1.23611 (0.09607)	1.05029 (0.08894)	0.80547 (0.08387)	0.58083 (0.08379)	0.44132 (0.08890)	1.31551 (0.09432)
125	1.23080 (0.06000)	1.05366 (0.05894)	0.80043 (0.05474)	0.58136 (0.05260)	0.44052 (0.05600)	1.31617 (0.06096)
225	1.22916 (0.04606)	1.05430 (0.04340)	0.80183 (0.04082)	0.57825 (0.03921)	0.43919 (0.04103)	1.31528 (0.04672)
325	1.23017 (0.03650)	1.05459 (0.03565)	0.80244 (0.03284)	0.58070 (0.03196)	0.43811 (0.03445)	1.31603 (0.03683)
400	1.22987 (0.03313)	1.05327 (0.03246)	0.80324 (0.02914)	0.58084 (0.02903)	0.43994 (0.03079)	1.31534 (0.03344)
$\alpha = 0.05, \beta = 0.8, \gamma = 0.1, \lambda = 2.7$						
50	5.59168 (0.43754)	5.46033 (0.43293)	5.24486 (0.42445)	4.95174 (0.41488)	4.72499 (0.38551)	5.61441 (0.43976)
125	5.56940 (0.28042)	5.46339 (0.26815)	5.22377 (0.27128)	4.94735 (0.24985)	4.72399 (0.25011)	5.58740 (0.28435)
225	5.59191 (0.20741)	5.47153 (0.18788)	5.22737 (0.20153)	4.95043 (0.19577)	4.72377 (0.18953)	5.60879 (0.21063)
325	5.58310 (0.17935)	5.47310 (0.16169)	5.23269 (0.15953)	4.94529 (0.15801)	4.71878 (0.15575)	5.59997 (0.18020)
400	5.58442 (0.15350)	5.47519 (0.15270)	5.23806 (0.15369)	4.95639 (0.14232)	4.72441 (0.15020)	5.60310 (0.15582)

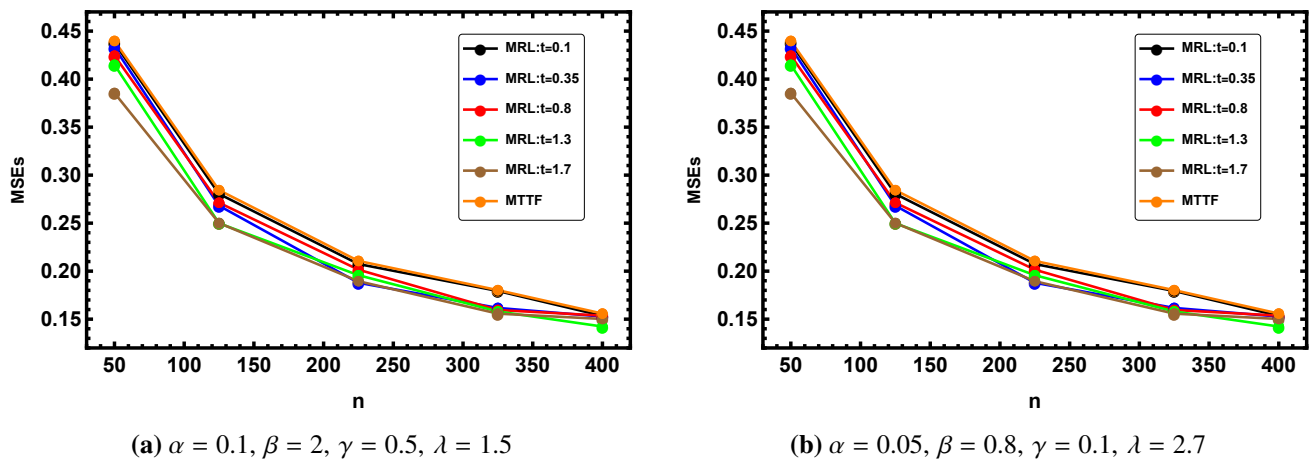


Figure 3. MSE of the MRL and MTTF obtained via Monte Carlo simulation for the AEP model under different sample sizes n and parameter configurations.

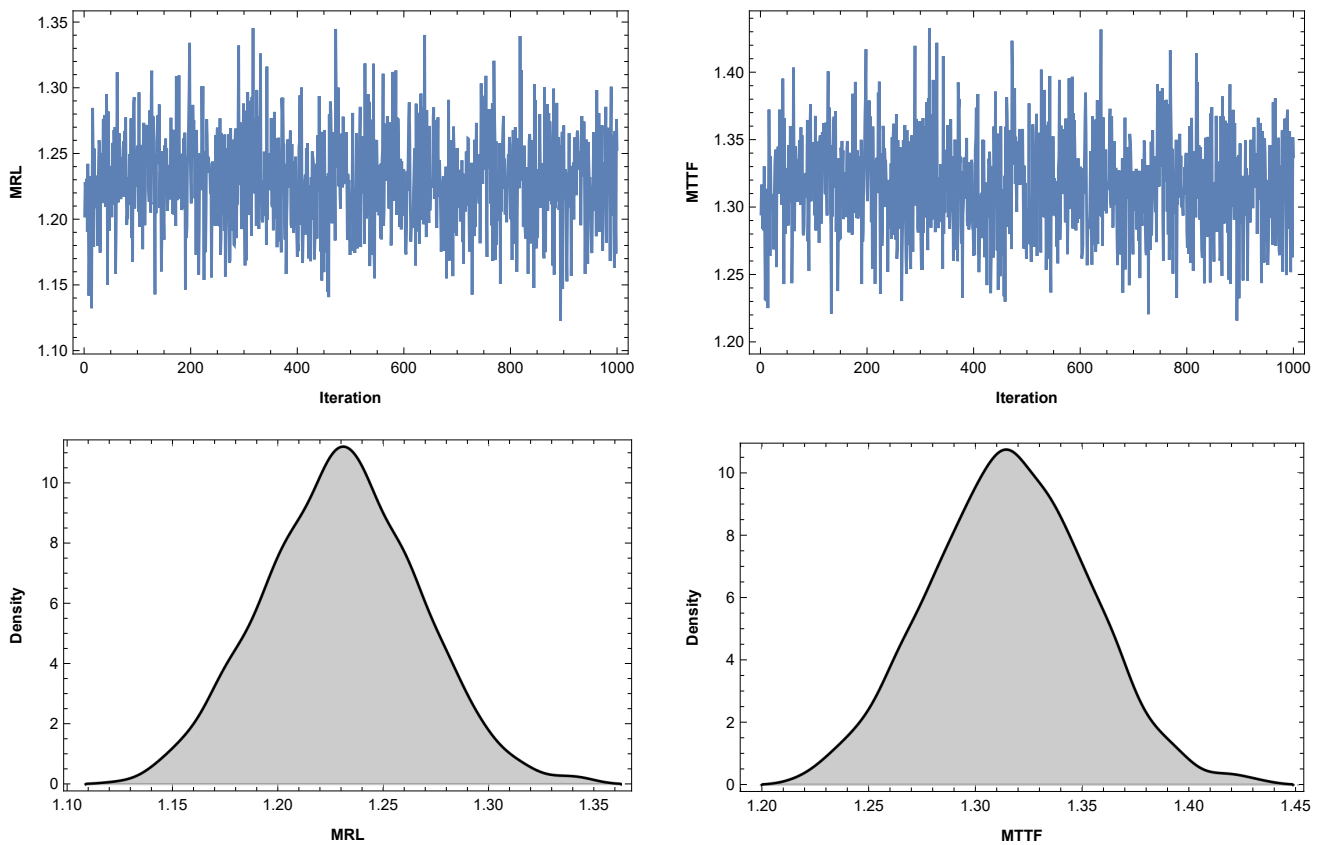


Figure 4. Trace plots and posterior density estimates of the MRL and MTTF obtained from Monte Carlo simulation for the AEP model with $n = 400$ and $t = 0.1$ under parameter setting $\alpha = 0.1, \beta = 2, \gamma = 0.5, \lambda = 1.5$.

2.12. Order statistics

Order statistics are vital for analyzing censored data and modeling system failures.

Theorem 6. Let $X_{1:n} \leq X_{2:n} \leq \dots \leq X_{n:n}$ be the order statistics from a random sample of size n following the AEP model. The PDF of the k -th order statistic $X_{k:n}$ is given by:

$$f_{k:n}(x; \psi) = \frac{n!}{(n-k)!} h(x; \psi) \sum_{j=0}^{k-1} \frac{(-1)^j}{j!(k-j-1)!} e^{-(n+j+1-k)H(x; \psi)}, \quad (2.15)$$

where $h(x; \psi)$ is the HR function defined in Eq (2.2) and $H(x; \psi)$ is the cumulative-hazard function in Eq (2.3).

Proof. See A.6. □

Order statistics facilitate parameter estimation from censored data and help model the failure of series and parallel systems, which is crucial for risk assessment and maintenance planning.

3. Parameter estimation

The parameters of the AEP model are estimated using classical and Bayesian methods. In the classical framework, maximum likelihood estimation (MLE) is used to derive point estimates and asymptotic confidence intervals (CIs). For Bayesian inference, posterior estimation is performed using HMC, specifically NUTS, an adaptive HMC algorithm designed to efficiently explore the posterior distribution.

3.1. Maximum likelihood estimation

The MLE method is a cornerstone of classical statistical inference. It determines the parameter values that maximize the likelihood function, which represents the probability of observing the given data.

Let $\psi = (\alpha, \beta, \gamma, \lambda)$ denote the vector of AEP model parameters, and let $D = (x_1, x_2, \dots, x_n)$ be a set of n independent and identically distributed observations. The likelihood function, $L(\psi|D)$, is given by:

$$L(\psi|D) = \prod_{i=1}^n \left(\alpha \beta (\alpha x_i)^{\beta-1} e^{(\alpha x_i)^\beta} + \gamma \lambda (\gamma x_i)^{\lambda-1} e^{(\gamma x_i)^\lambda} \right) e^{\sum_{i=1}^n \left(2 - e^{(\alpha x_i)^\beta} - e^{(\gamma x_i)^\lambda} \right)}. \quad (3.1)$$

The corresponding log-likelihood function $\ell(\psi|D)$ is:

$$\ell(\psi|D) = \sum_{i=1}^n \log \left(\alpha \beta (\alpha x_i)^{\beta-1} e^{(\alpha x_i)^\beta} + \gamma \lambda (\gamma x_i)^{\lambda-1} e^{(\gamma x_i)^\lambda} \right) + \sum_{i=1}^n \left(2 - e^{(\alpha x_i)^\beta} - e^{(\gamma x_i)^\lambda} \right). \quad (3.2)$$

The MLEs, $\widehat{\psi} = (\widehat{\alpha}, \widehat{\beta}, \widehat{\gamma}, \widehat{\lambda})$, are obtained by solving the system of score equations:

$$\begin{aligned} \frac{\partial \ell(\psi|D)}{\partial \alpha} &= -\beta \alpha^{-1} \sum_{i=1}^n (\alpha x_i)^{\beta-1} e^{(\alpha x_i)^\beta} \\ &\quad + \beta^2 \sum_{i=1}^n \frac{(\alpha x_i)^{\beta-1} (1 + (\alpha x_i)^\beta) e^{(\alpha x_i)^\beta}}{\alpha \beta (\alpha x_i)^{\beta-1} e^{(\alpha x_i)^\beta} + \gamma \lambda (\gamma x_i)^{\lambda-1} e^{(\gamma x_i)^\lambda}} = 0. \end{aligned}$$

$$\begin{aligned}\frac{\partial \ell(\psi|D)}{\partial \beta} &= - \sum_{i=1}^n (\alpha x_i)^\beta \ln(\alpha x_i) e^{(\alpha x_i)^\beta} \\ &\quad + \alpha \sum_{i=1}^n \frac{(\alpha x_i)^{\beta-1} \left(1 + \beta \left(1 + (\alpha x_i)^\beta\right) \log(\alpha x_i)\right) e^{(\alpha x_i)^\beta}}{\alpha \beta (\alpha x_i)^{\beta-1} e^{(\alpha x_i)^\beta} + \gamma \lambda (\gamma x_i)^{\lambda-1} e^{(\gamma x_i)^\lambda}} = 0.\end{aligned}$$

$$\begin{aligned}\frac{\partial \ell(\psi|D)}{\partial \gamma} &= -\lambda \gamma^{-1} \sum_{i=1}^n (\gamma x_i)^{\lambda-1} e^{(\gamma x_i)^\lambda} \\ &\quad + \lambda^2 \sum_{i=1}^n \frac{(\gamma x_i)^{\lambda-1} (1 + (\gamma x_i)^\lambda) e^{(\gamma x_i)^\lambda}}{\alpha \beta (\alpha x_i)^{\beta-1} e^{(\alpha x_i)^\beta} + \gamma \lambda (\gamma x_i)^{\lambda-1} e^{(\gamma x_i)^\lambda}} = 0.\end{aligned}$$

$$\begin{aligned}\frac{\partial \ell(\psi|D)}{\partial \lambda} &= - \sum_{i=1}^n (\gamma x_i)^\lambda \ln(\gamma x_i) e^{(\gamma x_i)^\lambda} \\ &\quad + \gamma \sum_{i=1}^n \frac{(\gamma x_i)^{\lambda-1} \left(1 + \lambda \left(1 + (\gamma x_i)^\lambda\right) \log(\gamma x_i)\right) e^{(\gamma x_i)^\lambda}}{\alpha \beta (\alpha x_i)^{\beta-1} e^{(\alpha x_i)^\beta} + \gamma \lambda (\gamma x_i)^{\lambda-1} e^{(\gamma x_i)^\lambda}} = 0.\end{aligned}$$

Due to the complexity of these equations, closed-form solutions are not feasible. Therefore, numerical optimization methods such as Newton–Raphson or built-in solvers (e.g., FindMaximum or NMaximize in Wolfram Mathematica) are employed.

3.1.1. Asymptotic confidence intervals

Under standard regularity conditions, the MLE vector $\widehat{\psi}$ is asymptotically normal for large sample sizes, with a mean equal to the true parameter vector ψ and variance–covariance matrix given by the inverse of the Fisher information matrix.

In practice, we use the observed Fisher information matrix, $\mathbf{I}(\widehat{\psi})$, calculated by evaluating the negative of the Hessian matrix (the matrix of second partial derivatives) of the log-likelihood function at the MLEs. For this four-parameter model, the 4×4 matrix is:

$$\mathbf{I}(\widehat{\psi}) = - \begin{bmatrix} \frac{\partial^2 \ell(\psi|D)}{\partial \alpha^2} & \frac{\partial^2 \ell(\psi|D)}{\partial \alpha \partial \beta} & \frac{\partial^2 \ell(\psi|D)}{\partial \alpha \partial \gamma} & \frac{\partial^2 \ell(\psi|D)}{\partial \alpha \partial \lambda} \\ \cdot & \frac{\partial^2 \ell(\psi|D)}{\partial \beta^2} & \frac{\partial^2 \ell(\psi|D)}{\partial \beta \partial \gamma} & \frac{\partial^2 \ell(\psi|D)}{\partial \beta \partial \lambda} \\ \cdot & \cdot & \frac{\partial^2 \ell(\psi|D)}{\partial \gamma^2} & \frac{\partial^2 \ell(\psi|D)}{\partial \gamma \partial \lambda} \\ \cdot & \cdot & \cdot & \frac{\partial^2 \ell(\psi|D)}{\partial \lambda^2} \end{bmatrix}_{\psi=\widehat{\psi}}.$$

The estimated asymptotic variance–covariance matrix of the MLEs is obtained by inverting the observed Fisher information matrix:

$$\widehat{\text{Var}}(\widehat{\psi}) = \begin{bmatrix} \widehat{\text{Var}}(\widehat{\alpha}) & \text{Cov}(\widehat{\alpha}, \widehat{\beta}) & \text{Cov}(\widehat{\alpha}, \widehat{\gamma}) & \text{Cov}(\widehat{\alpha}, \widehat{\lambda}) \\ \cdot & \widehat{\text{Var}}(\widehat{\beta}) & \text{Cov}(\widehat{\beta}, \widehat{\gamma}) & \text{Cov}(\widehat{\beta}, \widehat{\lambda}) \\ \cdot & \cdot & \widehat{\text{Var}}(\widehat{\gamma}) & \text{Cov}(\widehat{\gamma}, \widehat{\lambda}) \\ \cdot & \cdot & \cdot & \widehat{\text{Var}}(\widehat{\lambda}) \end{bmatrix}.$$

The diagonal entries of this matrix, $\widehat{\text{Var}}(\widehat{\psi}_j)$, provide the estimated variances of the parameter estimates.

Accordingly, an approximate $100(1 - \eta)\%$ confidence interval for parameter ψ_j is given by:

$$(\widehat{\psi}_j - z_{\eta/2} \sqrt{\widehat{\text{Var}}(\widehat{\psi}_{jj})}, \widehat{\psi}_j + z_{\eta/2} \sqrt{\widehat{\text{Var}}(\widehat{\psi}_{jj})}), \quad j \in \{\alpha, \beta, \gamma, \lambda\},$$

where $z_{\eta/2}$ is the upper $\eta/2$ quantile of the standard normal model.

3.2. Bayesian estimation

Bayesian inference provides a flexible framework for parameter estimation by combining prior information about the model parameters with the information contained in the observed data. For the AEP model, this approach is particularly useful because it allows uncertainty quantification for all parameters and facilitates inference in situations where the posterior distribution may exhibit dependence among parameters.

Let $\psi = (\alpha, \beta, \gamma, \lambda)$ denote the parameter vector, and let $D = (x_1, x_2, \dots, x_n)$ denote the observed sample. The posterior distribution of ψ is proportional to the product of the likelihood function and the prior distribution, that is,

$$\pi(\psi | D) \propto L(D | \psi)\pi(\psi), \quad (3.3)$$

where $L(D | \psi)$ is the likelihood function given in Eq (3.1).

For the real-data applications, gamma priors are assigned to the positive model parameters, subject to the identifying constraint $\alpha < \gamma$ introduced in Subsection 2.2. Specifically,

$$\begin{aligned} \alpha &\sim \text{Ga}(a_1, b_1), \\ \beta &\sim \text{Ga}(a_2, b_2), \\ \gamma &\sim \text{Ga}(a_3, b_3), \\ \lambda &\sim \text{Ga}(a_4, b_4), \end{aligned} \quad (3.4)$$

where $\text{Ga}(a, b)$ denotes the gamma distribution with shape parameter a and rate parameter b , with density proportional to $x^{a-1}e^{-bx}$ for $x > 0$. Under prior independence, the joint prior density is

$$\pi(\psi) \propto \alpha^{a_1-1} \beta^{a_2-1} \gamma^{a_3-1} \lambda^{a_4-1} \exp\{-(b_1\alpha + b_2\beta + b_3\gamma + b_4\lambda)\}. \quad (3.5)$$

Table 3. Hyperparameter values of the gamma priors used in the Bayesian analysis of the three real-data applications.

Parameter	Sylwia		VETEs		DTPPs	
	Shape (a_i)	Rate (b_i)	Shape (a_i)	Rate (b_i)	Shape (a_i)	Rate (b_i)
α	2	20,000	4	1600	1	10,000
β	2	8	5	4	1	4
γ	2	25	4	1500	2	60
λ	4	1	6	2	2	0.5

For the real-data applications, independent gamma priors were assigned to the positive model parameters. We use the shape–rate parameterization, such that

$$\theta_i \sim \text{Gamma}(a_i, b_i), \quad E(\theta_i) = \frac{a_i}{b_i}, \quad \text{Var}(\theta_i) = \frac{a_i}{b_i^2},$$

where a_i and b_i denote the shape and rate hyperparameters, respectively. The hyperparameter values used for the three real-data applications are reported in Table 3.

Combining the likelihood and the prior yields the joint posterior density

$$\begin{aligned} \pi(\psi \mid D) &\propto \alpha^{a_1-1} \beta^{a_2-1} \gamma^{a_3-1} \lambda^{a_4-1} \exp\{-(b_1\alpha + b_2\beta + b_3\gamma + b_4\lambda)\} \\ &\times \exp\left\{-\sum_{i=1}^n \left(e^{(\alpha x_i)^\beta} + e^{(\gamma x_i)^\lambda}\right)\right\} \prod_{i=1}^n h(x_i; \psi), \end{aligned} \quad (3.6)$$

where

$$h(x_i; \psi) = \alpha\beta(\alpha x_i)^{\beta-1} e^{(\alpha x_i)^\beta} + \gamma\lambda(\gamma x_i)^{\lambda-1} e^{(\gamma x_i)^\lambda}.$$

Under the squared error loss function (SELF), the Bayes estimators of α , β , γ , and λ are given by their posterior means:

$$\begin{aligned} \hat{\alpha} &= \mathbb{E}(\alpha \mid D), & \hat{\beta} &= \mathbb{E}(\beta \mid D), \\ \hat{\gamma} &= \mathbb{E}(\gamma \mid D), & \hat{\lambda} &= \mathbb{E}(\lambda \mid D). \end{aligned}$$

The posterior distribution is analytically intractable because of its complex functional form and the absence of a closed-form normalizing constant. Therefore, posterior computation is carried out numerically using Hamiltonian Monte Carlo with the NUTS, as described in Subsection 3.2.1. Since HMC–NUTS operates directly on the joint posterior density, explicit full conditional distributions are not required for implementation.

3.2.1. Hamiltonian Monte Carlo algorithm and the No-U-Turn sampler

While conventional Markov chain Monte Carlo (MCMC) methods, such as Metropolis–Hastings and Gibbs sampling [28–30], remain standard tools for Bayesian estimation, their efficiency often deteriorates in complex settings with correlated parameters. Their random-walk behavior may lead to slow exploration of the parameter space, highly autocorrelated samples, and a reduced effective sample size. These difficulties become more pronounced in flexible lifetime models such as the AEP model, where the posterior distribution may exhibit non-negligible dependence among parameters.

Hamiltonian Monte Carlo (HMC) addresses these limitations by incorporating gradient information from the log-posterior density to guide posterior exploration. The Hamiltonian function is defined as

$$H(\psi, \mathbf{p}) = U(\psi) + K(\mathbf{p}),$$

where

$$U(\psi) = -\log \pi(\psi \mid D)$$

is the potential energy and

$$K(\mathbf{p}) = \frac{1}{2} \mathbf{p}^T \mathbf{M}^{-1} \mathbf{p}$$

is the kinetic energy associated with an auxiliary momentum vector $\mathbf{p}$. Typically, $\mathbf{p}$ is generated from a multivariate normal distribution, $\mathbf{p} \sim N(\mathbf{0}, \mathbf{I})$, and the Hamiltonian dynamics are approximated numerically using the leapfrog integrator, as described in standard treatments of HMC [31,32]. Related implementations of HMC for lifetime and reliability models can be found in [33,34].

To avoid manual tuning of the trajectory length, we employ the NUTS, which adaptively determines the number of leapfrog steps at each iteration and terminates the trajectory when a U-turn criterion is satisfied [35]. This adaptive mechanism improves posterior exploration and reduces the need for manual tuning in HMC implementations [36,37]. More recent studies have further demonstrated the practical use and computational efficiency of the NUTS in complex Bayesian models [38,39]. These features make the NUTS particularly useful for the posterior computation required in the present AEP model.

Recent developments have further extended HMC-based sampling through neural-network-enhanced and adaptive strategies. Levy et al. [40] introduced a neural-network-based generalization of HMC, while Meng et al. [41] proposed an adaptive meta-learning stochastic gradient HMC approach for Bayesian updating. More recently, Meng et al. [42] developed an adaptive principal-component transformation for MCMC sampling to improve sampling efficiency in Bayesian structural system identification. These developments highlight the continuing evolution of gradient-based and adaptive MCMC methods, while the present study employs the standard HMC–NUTS framework for posterior inference.

In this study, Bayesian computation was implemented in two settings. For the real-data applications, the HMC–NUTS algorithm was implemented in Wolfram Mathematica. Four parallel chains were run, with $N_{\text{total}} = 5000$ iterations per chain, of which the first $N_{\text{warm}} = 2000$ iterations were discarded as burn-in. Initial values were selected using the corresponding maximum likelihood estimates and then perturbed slightly to obtain dispersed starting points across chains. Posterior convergence was assessed using the Gelman–Rubin statistic $\hat{R}$ and the effective sample size (ESS).

For the simulation study reported in Section 4, Bayesian estimation was implemented in *R* using *rstan*. In that part of the analysis, four chains were run with 2000 iterations per chain, including 1000 warm-up iterations. The NUTS control settings were `adapt_delta=0.9` and `max_treedepth=12`.

Let $\psi_{l,k}$ denote the k th post-warm-up posterior draw from chain l , for $l = 1, \dots, C$. Under the squared error loss function, the Bayes estimator of ψ is approximated by the posterior mean

$$\hat{\psi} = \frac{1}{C(N_{\text{total}} - N_{\text{warm}})} \sum_{l=1}^C \sum_{k=N_{\text{warm}}+1}^{N_{\text{total}}} \psi_{l,k}.$$

Similarly, credible intervals are obtained from the empirical posterior quantiles of the generated samples.

The resulting posterior draws were also used to construct trace plots, posterior density plots, and pairwise scatter plots for diagnostic purposes. In practice, $\hat{R}$ values close to 1 together with sufficiently large ESS values were taken as indicators of satisfactory convergence and sampling efficiency. The main steps of the HMC–NUTS algorithm used for Bayesian parameter estimation are illustrated in Figure 5.

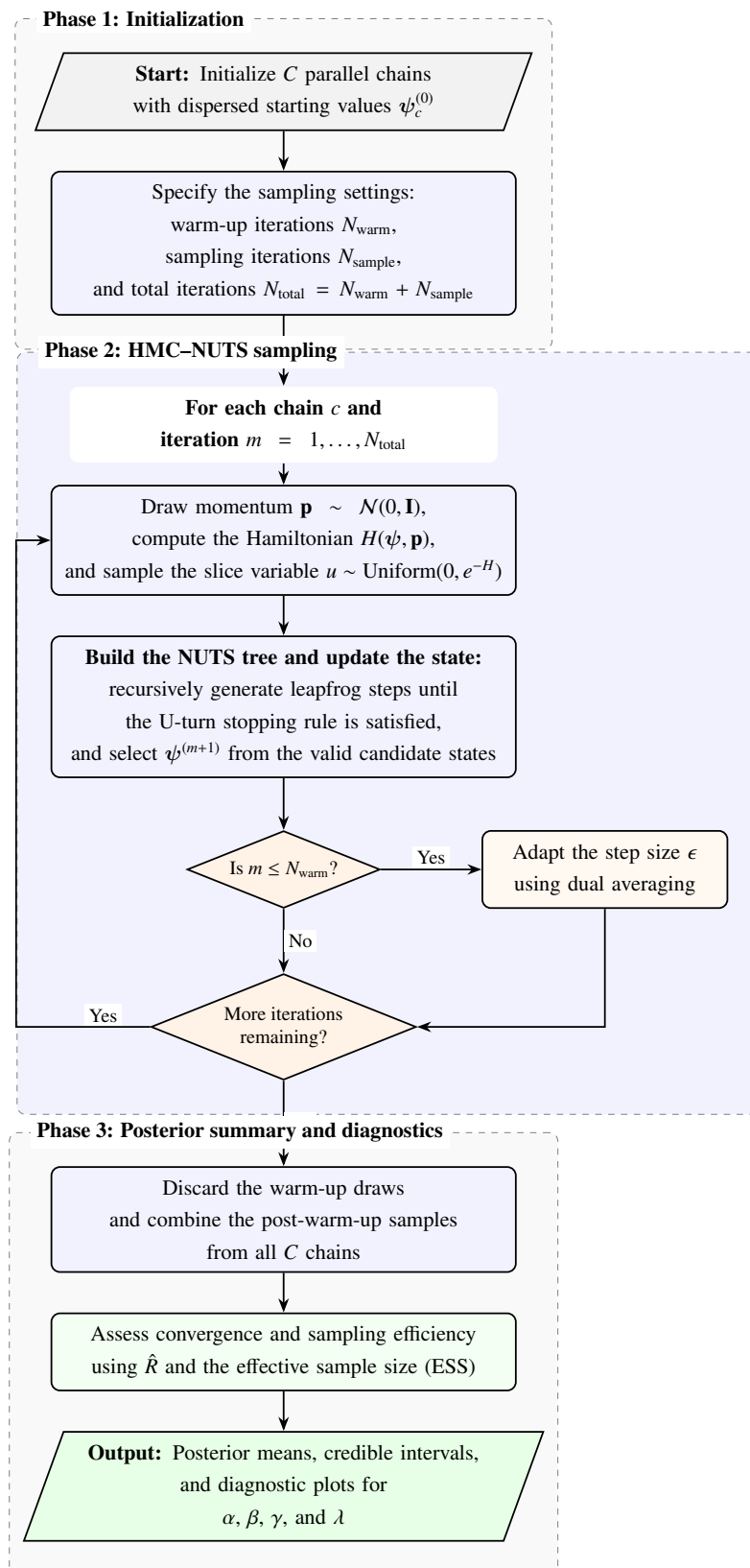


Figure 5. Flowchart of the HMC-NUTS procedure used for Bayesian estimation of the AEP model parameters.

4. Simulation results

In this section, we investigate the finite-sample performance of the proposed estimators for the AEP model under repeated sampling. The simulation study was implemented in *R*, where maximum likelihood estimation was carried out by numerical optimization and Bayesian estimation was performed using *rstan* under the HMC–NUTS framework. For each replication, a random sample was generated from the AEP model, the unknown parameters were estimated by both MLE and Bayesian methods, and the resulting estimates were summarized using bias, mean squared error (MSE), root mean squared error (RMSE), and relative mean squared error (ReMSE). The simulation procedure is summarized in Algorithm 1.

Algorithm 1 Simulation experiment for the AEP model

- 1: Specify the true parameter vector $\psi_0 = (\alpha_0, \beta_0, \gamma_0, \lambda_0)$.
 - 2: Select the sample size $n \in \{30, 60, 120, 180, 300\}$ and the number of replications $N = 1000$.
 - 3: **for** $r = 1, \dots, N$ **do**
 - 4: Generate a random sample of size n from the AEP model by numerically inverting Eq (2.9)
 - 5: Obtain the maximum likelihood estimates by numerically minimizing the negative log-likelihood function.
 - 6: Obtain the Bayes estimates using HMC–NUTS posterior sampling in *rstan*.
 - 7: Store the resulting estimates of α, β, γ , and λ .
 - 8: **end for**
 - 9: For each parameter and estimation method, compute the average estimate, bias, MSE, RMSE, and ReMSE.
 - 10: Compare the results across sample sizes to assess estimation accuracy and stability.
-

For the original simulation scenarios reported in Table 4, the MLEs were computed using numerical optimization with data-dependent initial values chosen in a neighborhood of the true parameter values. In the additional robustness scenarios reported in Tables 5 and 6, a practically feasible data-driven initialization strategy was used instead, without using the true parameter values. For Bayesian estimation, posterior sampling was conducted in *Stan* using four chains, 2000 iterations per chain, and 1000 warm-up iterations, with `adapt_delta`=0.9 and `max_treedepth`=12. Weakly informative prior distributions were specified on the log-scale, as described in Subsection 3.2.1. In each replication, the Bayes estimates were taken as posterior means from the generated posterior samples. Convergence of the Bayesian sampler was monitored using the potential scale reduction statistic $\hat{R}$, and runs failing to satisfy the convergence criterion were excluded from the final summary.

For each parameter $\theta \in \{\alpha, \beta, \gamma, \lambda\}$, let $\hat{\theta}_1, \dots, \hat{\theta}_N$ denote the estimates obtained from the valid replications, and let θ_0 denote the true value. The simulation measures were computed as follows:

$$\begin{aligned} \text{Bias}(\hat{\theta}) &= \frac{1}{N} \sum_{r=1}^N (\hat{\theta}_r - \theta_0), & \text{MSE}(\hat{\theta}) &= \frac{1}{N} \sum_{r=1}^N (\hat{\theta}_r - \theta_0)^2, \\ \text{RMSE}(\hat{\theta}) &= \sqrt{\text{MSE}(\hat{\theta})}, & \text{ReMSE}(\hat{\theta}) &= \frac{\text{MSE}(\hat{\theta})}{\theta_0^2}. \end{aligned}$$

These criteria provide complementary measures of estimation performance in terms of bias, variability, and relative error.

The simulation results for both estimation methods are reported in Table 4 and displayed graphically in Figures 6 and 7. Overall, the results indicate that the performance of both MLE and Bayesian estimation improves as the sample size increases. Across the two parameter settings, the bias, MSE, RMSE, and ReMSE generally decrease with increasing n , which is consistent with the expected large-sample behavior of the estimators. For moderate and large sample sizes, the differences between the two approaches become smaller, and both methods provide more accurate and stable estimates. These findings suggest that both estimation procedures are effective for the AEP model, with improved reliability as the amount of sample information increases.

A more detailed examination of Table 4 shows that the estimation of λ is more challenging than that of the other model parameters, particularly for small sample sizes under the MLE approach. This is reflected in the comparatively larger bias and error measures observed for λ when n is small. The Bayesian estimates are generally more stable in this setting, indicating that posterior-based regularization helps reduce the estimation variability of λ . As the sample size increases, however, this difficulty becomes less pronounced, and the estimates from both methods improve substantially.

Table 4. Simulation results of the MLE and Bayesian estimators for the AEP parameters based on bias, MSE, RMSE, and ReMSE under different sample sizes.

n	Method	Parameter	$\alpha = 0.05, \beta = 0.25, \gamma = 0.08, \lambda = 4$				$\alpha = 0.02, \beta = 0.5, \gamma = 0.03, \lambda = 2$			
			Bias	MSE	RMSE	ReMSE	Bias	MSE	RMSE	ReMSE
30	MLE	$\hat{\alpha}$	0.01113	0.00247	0.04970	0.98798	0.00538	0.00002	0.01506	0.56733
		$\hat{\beta}$	0.00478	0.00424	0.06508	0.06776	0.04681	0.02500	0.15798	0.09983
		$\hat{\gamma}$	0.00238	0.00005	0.00712	0.00792	0.00033	0.00003	0.00502	0.02800
		$\hat{\lambda}$	4.18422	145.10000	12.04781	9.07186	3.83342	82.82000	9.10031	20.70390
	Bayes	$\hat{\alpha}$	0.02290	0.00205	0.04528	0.82016	0.00847	0.00012	0.01089	0.29663
		$\hat{\beta}$	0.00841	0.00374	0.06113	0.05978	0.09448	0.02737	0.16544	0.10948
		$\hat{\gamma}$	-0.00168	0.00005	0.00740	0.00856	-0.00403	0.00003	0.00568	0.03587
		$\hat{\lambda}$	-0.60844	2.25500	1.50171	0.14095	0.37436	0.43620	0.66046	0.10905
60	MLE	$\hat{\alpha}$	0.00503	0.00106	0.03248	0.42189	0.00385	0.00016	0.01275	0.40624
		$\hat{\beta}$	-0.00125	0.00222	0.04710	0.03550	0.03146	0.01481	0.12168	0.05922
		$\hat{\gamma}$	0.00127	0.00002	0.00460	0.00330	-0.00006	0.00002	0.00417	0.01929
		$\hat{\lambda}$	1.27056	15.59600	3.99550	0.99775	1.75200	22.72000	4.7666	5.68011
	Bayes	$\hat{\alpha}$	0.00764	0.00067	0.02579	0.26603	0.00532	0.00007	0.00846	0.17909
		$\hat{\beta}$	-0.00330	0.00178	0.04217	0.02845	0.06196	0.01530	0.12369	0.06120
		$\hat{\gamma}$	0.00081	0.00002	0.00440	0.00302	-0.00255	0.00002	0.00445	0.02198
		$\hat{\lambda}$	-0.18598	1.64700	1.28327	0.10292	0.39267	0.59630	0.77219	0.14907
120	MLE	$\hat{\alpha}$	-0.00101	0.00045	0.02116	0.17905	0.00166	0.00010	0.00992	0.24610
		$\hat{\beta}$	-0.00802	0.00108	0.03280	0.01721	0.01124	0.00780	0.08943	0.03199
		$\hat{\gamma}$	0.00067	0.00001	0.00295	0.00136	-0.00003	0.00000	0.00312	0.01080
		$\hat{\lambda}$	0.29960	2.02272	1.42222	0.12642	0.62940	4.62200	2.14988	1.15550
	Bayes	$\hat{\alpha}$	0.00070	0.00033	0.01814	0.13158	0.00203	0.00004	0.00651	0.10605
		$\hat{\beta}$	-0.00971	0.00096	0.03098	0.01536	0.02358	0.00677	0.08226	0.02706
		$\hat{\gamma}$	0.00072	0.00000	0.00292	0.00133	-0.00079	0.00000	0.00284	0.00898
		$\hat{\lambda}$	-0.17281	0.90170	0.94959	0.05636	0.25846	0.34790	0.58985	0.08698
180	MLE	$\hat{\alpha}$	-0.00238	0.00029	0.01714	0.11756	0.00097	0.00008	0.00883	0.19471
		$\hat{\beta}$	-0.00870	0.00071	0.02669	0.01140	0.00480	0.00559	0.07480	0.02238
		$\hat{\gamma}$	0.00063	0.00000	0.00246	0.00095	0.00007	0.00000	0.00267	0.00789
		$\hat{\lambda}$	0.15338	1.01300	1.00663	0.06333	0.31610	0.88660	0.94158	0.22164
	Bayes	$\hat{\alpha}$	-0.00131	0.00024	0.01552	0.09634	0.00071	0.00004	0.00625	0.09757
		$\hat{\beta}$	-0.01069	0.00069	0.02635	0.01111	0.00898	0.00426	0.06524	0.01703
		$\hat{\gamma}$	0.00072	0.00000	0.00246	0.00094	-0.00002	0.00000	0.00227	0.00574
		$\hat{\lambda}$	-0.14660	0.63820	0.79889	0.03989	0.21350	0.38200	0.61805	0.09550
300	MLE	$\hat{\alpha}$	-0.00252	0.00019	0.01391	0.07737	0.00044	0.00005	0.00724	0.13085
		$\hat{\beta}$	-0.00973	0.00048	0.02197	0.00772	0.00135	0.00336	0.05792	0.01342
		$\hat{\gamma}$	0.00045	0.00000	0.00195	0.00059	0.00008	0.00000	0.00217	0.00525
		$\hat{\lambda}$	0.08930	0.42370	0.65089	0.02648	0.16999	0.35590	0.59654	0.08897
	Bayes	$\hat{\alpha}$	-0.00224	0.00017	0.01299	0.06746	0.00008	0.00003	0.00587	0.08613
		$\hat{\beta}$	-0.01116	0.00047	0.02179	0.00760	0.00075	0.00276	0.05253	0.01104
		$\hat{\gamma}$	0.00057	0.00000	0.00189	0.00056	0.00021	0.00000	0.00193	0.00414
		$\hat{\lambda}$	-0.10536	0.37220	0.61011	0.02326	0.14650	0.23410	0.48385	0.05853

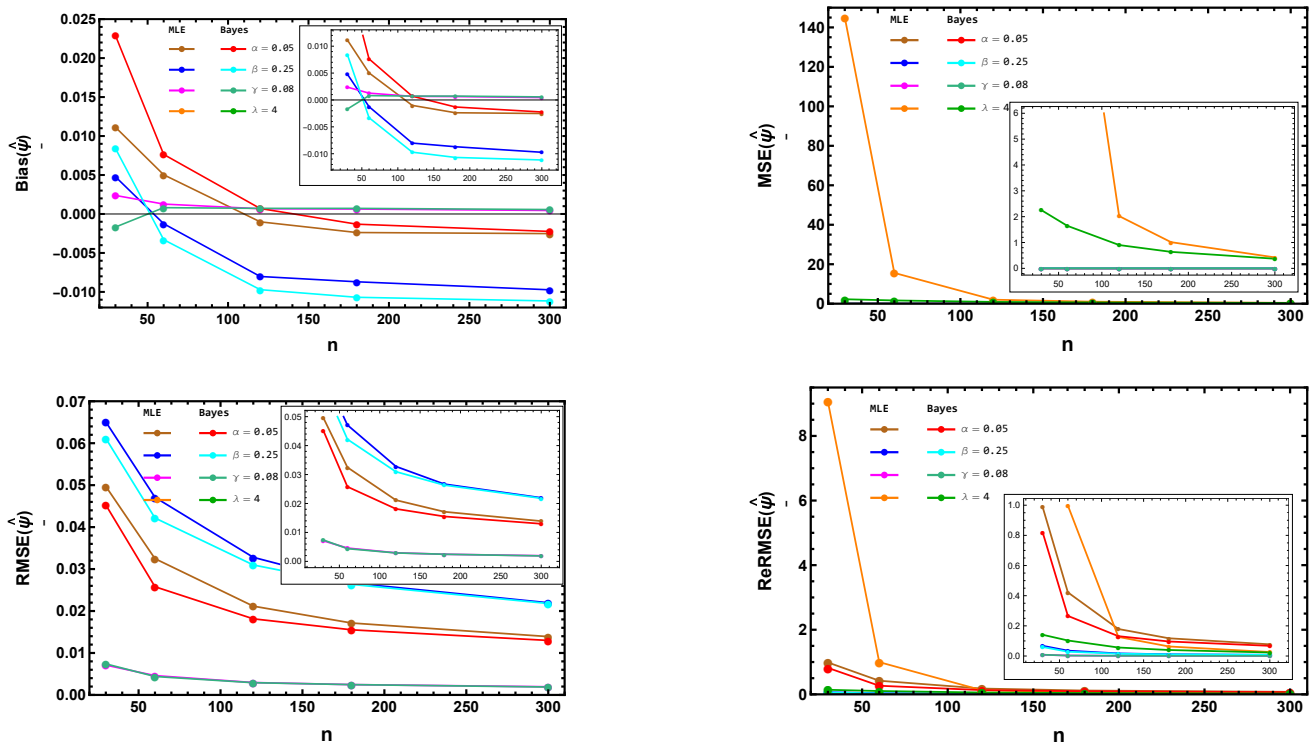


Figure 6. Comparison of (a) bias, (b) MSE, (c) RMSE, and (d) ReMSE for MLE and Bayesian estimators using the parameter set: $\alpha = 0.05, \beta = 0.25, \gamma = 0.08$, and $\lambda = 4$.

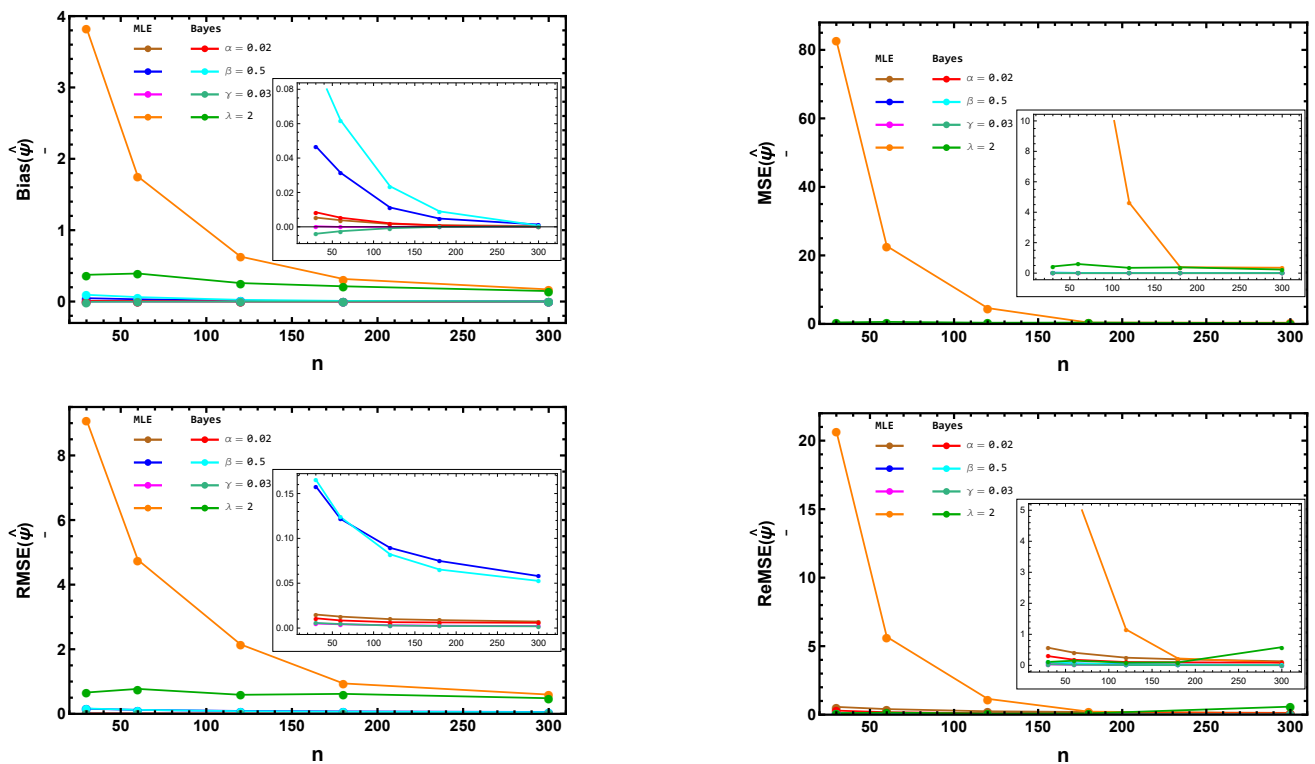


Figure 7. Comparison of (a) bias, (b) MSE, (c) RMSE, and (d) ReMSE for MLE and Bayesian estimators using the parameter set: $\alpha = 0.02, \beta = 0.5, \gamma = 0.03$, and $\lambda = 2$.

Table 5. Simulation results under practically feasible initialization for the unequal component-strength setting $(\alpha_0, \beta_0, \gamma_0, \lambda_0) = (0.005, 0.5, 0.5, 3)$, based on 500 replications.

n	Method	Convergence	Param.	Bias	MSE	RMSE	ReMSE
60	MLE	49/500 (9.8%)	α	0.1764	0.0562	0.2370	2.25×10^3
			β	8.5805	203.80	14.276	8.15×10^2
			γ	-0.0011	0.0007	0.0269	2.90×10^{-3}
			λ	-0.4226	0.2437	0.4936	2.71×10^{-2}
	Bayes	471/500 (94.2%)	α	0.1036	0.0130	0.1138	5.18×10^2
			β	4.1694	45.72	6.7617	1.83×10^2
			γ	-0.0040	0.0002	0.0155	9.60×10^{-4}
			λ	0.2758	0.4760	0.6899	5.29×10^{-2}
180	MLE	55/500 (11.0%)	α	0.3757	0.1577	0.3972	6.31×10^3
			β	5.7005	34.716	5.8920	1.39×10^2
			γ	-0.0495	0.0032	0.0568	1.29×10^{-2}
			λ	-1.0446	1.2020	1.0964	1.34×10^{-1}
	Bayes	308/500 (61.6%)	α	0.0758	0.0116	0.1079	4.66×10^2
			β	1.1151	6.9128	2.6292	2.77×10^1
			γ	-0.0051	0.0001	0.0114	5.20×10^{-4}
			λ	0.0848	0.1669	0.4085	1.85×10^{-2}
300	MLE	18/500 (3.6%)	α	0.3743	0.1572	0.3965	6.29×10^3
			β	5.7782	37.085	6.0897	1.48×10^2
			γ	-0.0485	0.0031	0.0560	1.26×10^{-2}
			λ	-1.1095	1.3319	1.1541	1.48×10^{-1}
	Bayes	169/500 (33.8%)	α	0.1564	0.0542	0.2327	2.17×10^3
			β	2.1604	12.880	3.5896	5.15×10^1
			γ	-0.0135	0.0005	0.0224	2.01×10^{-3}
			λ	-0.3382	0.5839	0.7642	6.49×10^{-2}

Table 6. Simulation results under practically feasible initialization for the boundary and model-misspecification settings, based on 500 replications.

Panel A: Boundary setting ($\alpha_0, \beta_0, \gamma_0, \lambda_0$) = (0.05, 0.95, 0.08, 1.05)							
n	Method	Convergence	Param.	Bias	MSE	RMSE	ReMSE
60	MLE	55/500 (11.0%)	α	0.0144	3.12×10^{-4}	0.0177	0.1248
			β	2.7526	72.02	8.4865	79.801
			γ	0.0019	2.13×10^{-4}	0.0146	0.0333
			λ	-0.1554	0.0385	0.1961	0.0349
	Bayes	496/500 (99.2%)	α	-0.0254	6.54×10^{-4}	0.0256	0.2614
			β	11.164	129.2	11.369	143.21
			γ	0.0260	7.64×10^{-4}	0.0276	0.1193
			λ	-0.1085	0.0255	0.1596	0.0231
180	MLE	122/500 (24.4%)	α	-0.0187	1.12×10^{-3}	0.0334	0.4467
			β	5.0514	197.4	14.049	218.71
			γ	0.0154	4.74×10^{-4}	0.0218	0.0741
			λ	-0.1677	0.0341	0.1846	0.0309
	Bayes	491/500 (98.2%)	α	-0.0267	7.15×10^{-4}	0.0268	0.2862
			β	11.271	133.9	11.570	148.32
			γ	0.0231	5.59×10^{-4}	0.0237	0.0874
			λ	-0.1262	0.0194	0.1391	0.0176
300	MLE	57/500 (11.4%)	α	-0.0154	1.06×10^{-3}	0.0326	0.4251
			β	1.2574	3.127	1.7683	3.4648
			γ	0.0119	3.88×10^{-4}	0.0197	0.0606
			λ	-0.1390	0.0301	0.1736	0.0273
	Bayes	492/500 (98.4%)	α	-0.0270	7.33×10^{-4}	0.0271	0.2933
			β	10.821	125.8	11.215	139.37
			γ	0.0223	5.11×10^{-4}	0.0226	0.0798
			λ	-0.1272	0.0180	0.1343	0.0164
Panel B: Model-misspecification setting							
Data generated from the additive Weibull model and fitted with the AEP model.							
n	MLE Convergence				Mean KS Statistic		
180	41/500 (8.2%)				0.04900		

To further assess the practical robustness of the estimation procedures, the original simulation experiments were retained, and the simulation study was substantially extended to include three additional challenging scenarios: A highly unequal component-strength setting, a boundary setting with shape parameters close to one, and a model-misspecification setting. For these additional experiments, a practically feasible data-driven initialization strategy was adopted, combining the scale estimate $1/\bar{x}$ with a coarse grid search over candidate shape-parameter values. Unlike the initialization used in the original simulation experiments, this strategy does not use the true parameter values. Table 5 reports the results for the highly unequal component-strength setting, whereas Table 6 presents the boundary and model-misspecification scenarios in Panels A and B, respectively. For the unequal setting, $(\alpha_0, \beta_0, \gamma_0, \lambda_0) = (0.005, 0.5, 0.5, 3)$, the practical initialization resulted in relatively low MLE

convergence rates, ranging from 3.6% to 11.0%. The Bayesian procedure achieved higher convergence rates, although its performance also deteriorated as the configuration became more difficult. The estimation errors were particularly pronounced for α and β , indicating substantial difficulty in jointly estimating the parameters of the two components when their scales are highly unequal. For the boundary setting, $(\alpha_0, \beta_0, \gamma_0, \lambda_0) = (0.05, 0.95, 0.08, 1.05)$, the results also indicate increased estimation difficulty, particularly for β . Although the Bayesian procedure maintained high convergence rates (98.2%–99.2%), the relatively large bias and MSE of $\hat{\beta}$ indicate that numerical convergence alone does not necessarily imply accurate parameter recovery near the transition $\beta = 1$ and $\lambda = 1$. Finally, Panel B of Table 6 reports the model-misspecification experiment, in which data were generated from an additive Weibull distribution and the AEP model was fitted. In the considered setting, the MLE convergence rate was 8.2% at $n = 180$, and the mean Kolmogorov–Smirnov statistic among the converged replications was 0.049. This experiment indicates that misspecification may introduce additional numerical and estimation difficulties. Overall, these additional experiments provide a more realistic assessment of the estimation procedures than initialization near the true parameter values. They demonstrate that practical initialization, highly unequal component scales, boundary configurations, and model misspecification can materially affect finite-sample estimation performance.

5. Real data applications

This section evaluates the practical performance of the proposed AEP model by applying it to three real-world datasets commonly used as benchmarks in reliability and survival analysis. The performance of the AEP model is compared against several competing models listed in Table 7.

Table 7. Competing models considered in the analysis.

Model	Survival function	Author
MWE	$e^{\alpha\beta(1-e^{(\frac{x}{\alpha})^\beta})^\gamma}$	Xie et al. [43]
MW	$e^{-\alpha x^\beta} e^{\gamma x}$	Lai et al. [44]
NMW	$e^{1-e^{\alpha x^\beta} \gamma x}$	Ghazal [45]
AW	$e^{-\alpha x^\beta - \gamma x^\lambda}$	Jiang and Murthy [10]
EMWE	$1 - \left[1 - e^{\alpha\lambda(1-e^{(\frac{x}{\alpha})^\beta})^\gamma} \right]^\gamma$	Sarhan and Apaloo [46]
AP	$1 - \frac{(\alpha+1)(\beta e^{-\gamma x} + 1)}{(\beta+1)(\alpha e^{\lambda x} + 1)}$	Méndez et al. [16]
AC	$e^{\alpha(1-e^{x^\beta}) + \gamma(1-e^{x^\lambda})}$	Méndez et al. [17]
GEEW	$1 - \left[1 - e^{-(\alpha x^\beta + \gamma x)^\lambda} \right]^\theta$	Shakhatreh et al. [47]
EAW	$1 - \left[1 - e^{-\alpha x^\beta - \gamma x^\lambda} \right]^\theta$	Ahmad and Ghazal [48]

Estimation procedures follow the methodologies described in Section 3, with Bayesian inference implemented via HMC using the NUTS algorithm. For each dataset, four parallel chains of 5000 samples were generated, discarding the first 2000 samples as burn-in. The remaining 3000 posterior samples per chain were used to obtain Bayesian estimates, complemented by posterior density plots

for visual validation.

To enable a comprehensive model comparison, we employed eight goodness-of-fit metrics: negative log-likelihood ($-\ell$), Akaike information Criterion (AIC), corrected AIC (AICc), Bayesian information Criterion (BIC), Hannan–Quinn information Criterion (HQIC), the Kolmogorov–Smirnov (KS) test, the Cramér–von Mises (CvM) test, and the Anderson–Darling (AD) test. Since the model parameters are estimated from the same datasets, the reported KS, CvM, and AD p -values should be interpreted with caution. Models were ranked according to these statistics, with the rankings visualized using line plots. In addition, comparative performance was assessed through estimated survival, HR, and cumulative-hazard functions. All computations were performed in Wolfram Mathematica.

5.1. Case study I: Lifetime data of 30 devices

The first dataset, referred to as the Sylwia dataset, consists of the failure times of 30 devices reported by Sylwia [49]. This dataset is particularly useful for reliability analysis because its failure pattern is associated with a non-monotone, bathtub-type hazard, providing an appropriate setting for assessing the flexibility of the AEP model.

Figure 8 presents the scaled total-time-on-test (TTT) transform together with the posterior mean TTT curve obtained from the fitted AEP model and its uncertainty band. The empirical TTT curve exhibits the characteristic departure from the diagonal associated with non-monotone hazard behavior. The fitted AEP curve follows the empirical pattern reasonably closely over most of the range, while the uncertainty band reflects the greater posterior uncertainty in the early part of the sample.

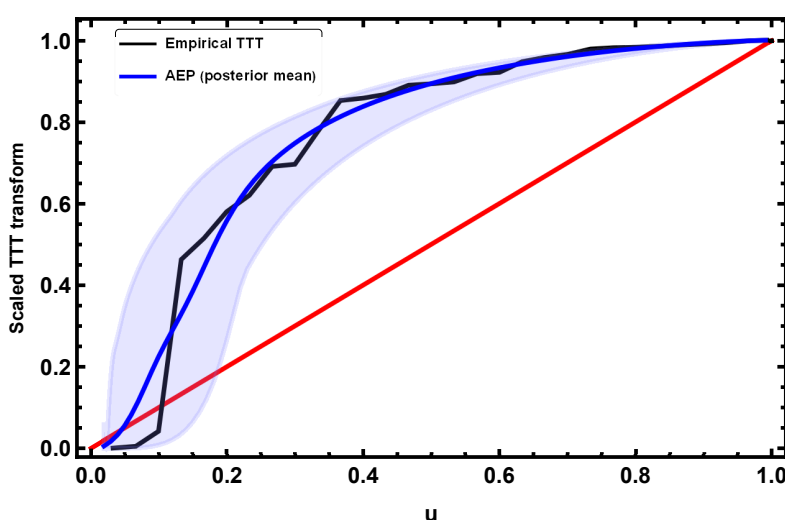


Figure 8. Comparison of empirical and AEP-fitted scaled TTT transform curves for the Sylwia dataset.

Table 8 reports the MLEs and Bayesian estimates together with the corresponding 95% confidence and highest posterior density (HPD) intervals. The Bayesian estimates are generally close to their classical counterparts, although the estimate of α remains very small under both approaches. The HMC convergence diagnostics are satisfactory, with all $\hat{R}$ values very close to one. The ESS values range from 297 for α to 2300 for γ , indicating greater sampling efficiency for some parameters than for others. A total of 23 divergent transitions were observed across the four chains; therefore, the posterior

results are interpreted together with the remaining convergence diagnostics rather than on the basis of $\hat{R}$ alone.

Table 8. MLEs and Bayesian estimates with 95% CIs and HPD intervals, $\hat{R}$, ESS, and HMC diagnostics for the AEP model fitted to the Sylwia dataset.

Parameter	MLEs	95% CI	Bayes	95% HPD	$\hat{R}$	ESS
$\hat{\alpha}$	5.29000×10^{-5}	[-0.00052, 0.00062]	9.41767×10^{-5}	[0.00001, 0.00026]	1.00001	297
$\hat{\beta}$	0.26064	[-0.03447, 0.55576]	0.27171	[0.16964, 0.40441]	1.00048	1090
$\hat{\gamma}$	0.08079	[0.07591, 0.08567]	0.08034	[0.07500, 0.08542]	1.00008	2300
$\hat{\lambda}$	4.08914	[2.55208, 5.62620]	3.99894	[2.74680, 5.53469]	1.00004	1276
Divergent transitions: Chain 1 = 4, Chain 2 = 9, Chain 3 = 3, Chain 4 = 7, Total = 23.						

The graphical HMC diagnostics in Figure 9 provide additional information on posterior sampling. The trace plots show substantial overlap among the four chains without evident persistent between-chain separation, while the posterior density plots are predominantly unimodal and the chain-specific boxplots show broadly comparable posterior distributions. Figure 10 further examines posterior dependence among the AEP parameters. The pairwise scatter plots and reported correlations indicate mostly weak-to-moderate posterior associations, with the most noticeable positive association occurring between α and β . These diagnostics are consistent with the dependence between scale and shape parameters discussed in Subsection 2.3.

Figure 11 complements the numerical comparisons by displaying the fitted survival, hazard-rate, and cumulative-hazard functions. The posterior mean AEP curves follow the corresponding empirical patterns closely over the observed range, while the 95% HPD bands quantify the associated posterior uncertainty. In particular, the estimated hazard function reproduces the non-monotone behavior suggested by the empirical diagnostics. Taken together, the numerical and graphical results indicate that the AEP model provides a close fit to the Sylwia dataset relative to the competing models considered here.

Table 9 reports the MLEs of the AEP model and the nine competing lifetime models. The corresponding goodness-of-fit and model selection results are summarized in Table 10. Among the models considered, the AEP model attains the smallest $-\ell$, AIC, AICc, BIC, and HQIC values. It also yields the smallest KS, CvM, and AD statistics for this dataset. The ranking summary in Figure 12 provides a visual representation of these comparisons across the reported criteria.

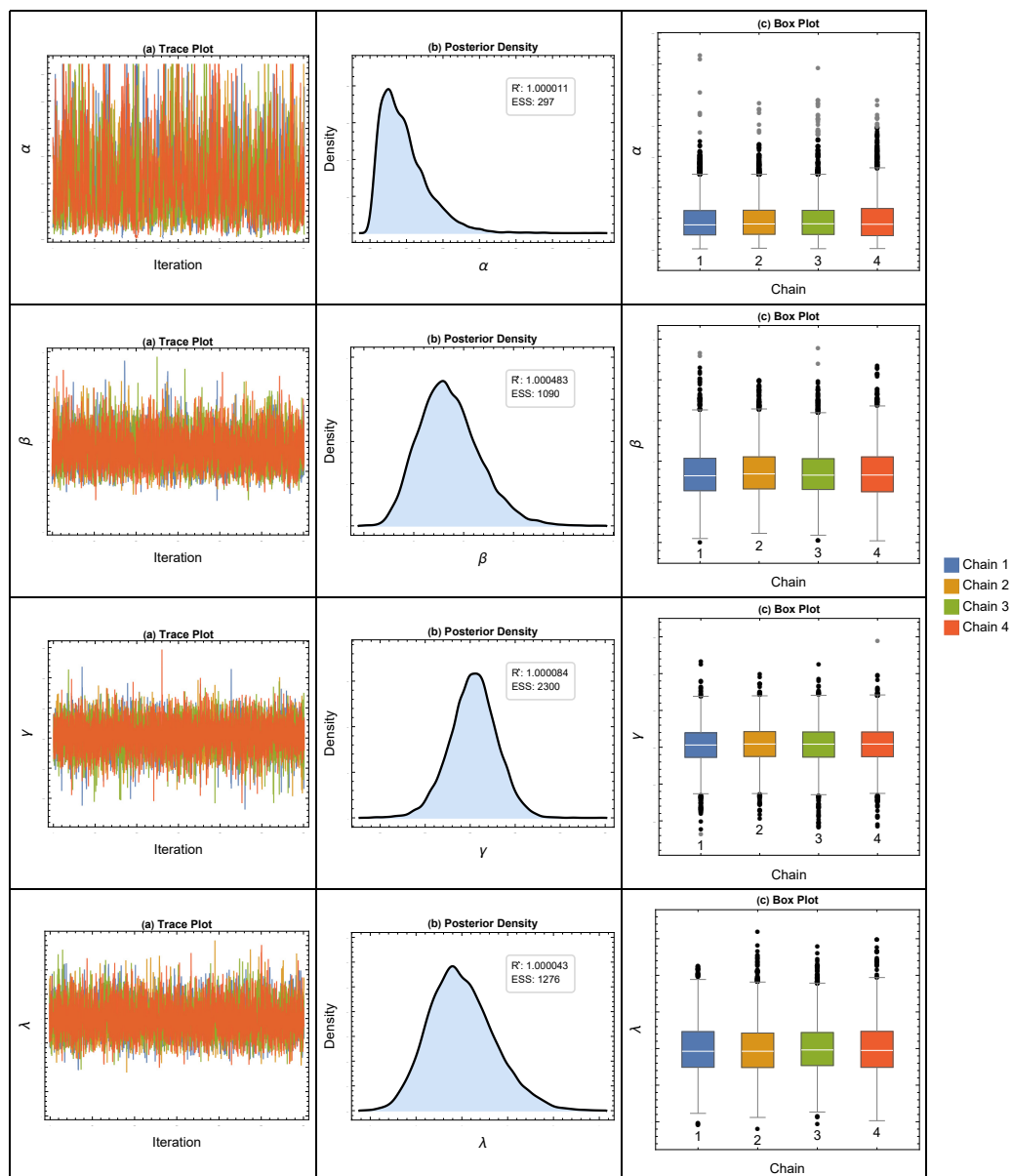


Figure 9. HMC diagnostics for the AEP parameters in the Sylwia dataset: (a) Trace plots, (b) posterior density plots, and (c) chain-specific box plots.

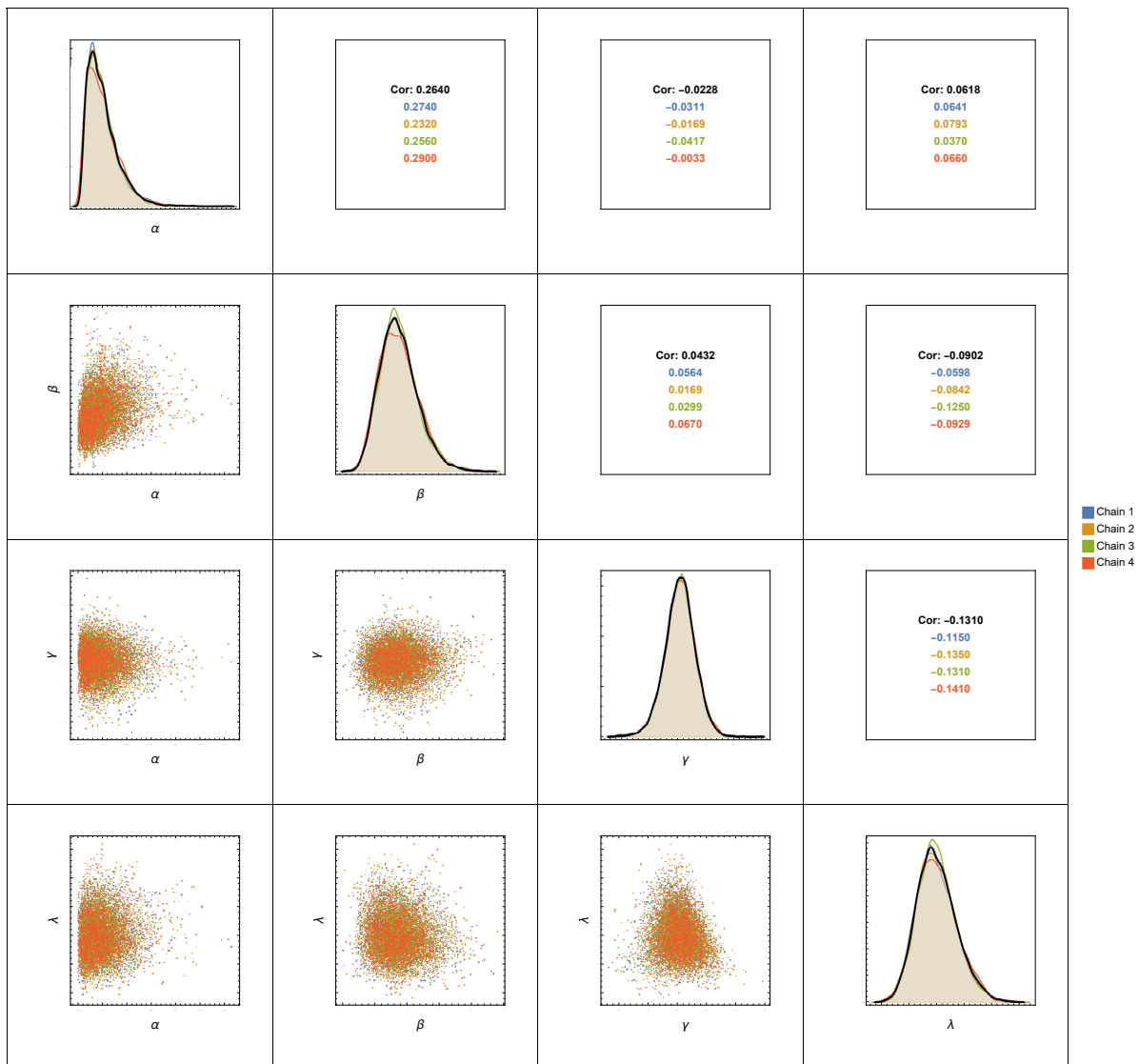


Figure 10. Scatter-plot matrices for pairs of AEP parameters derived from HMC posterior samples using the Sylwia dataset.

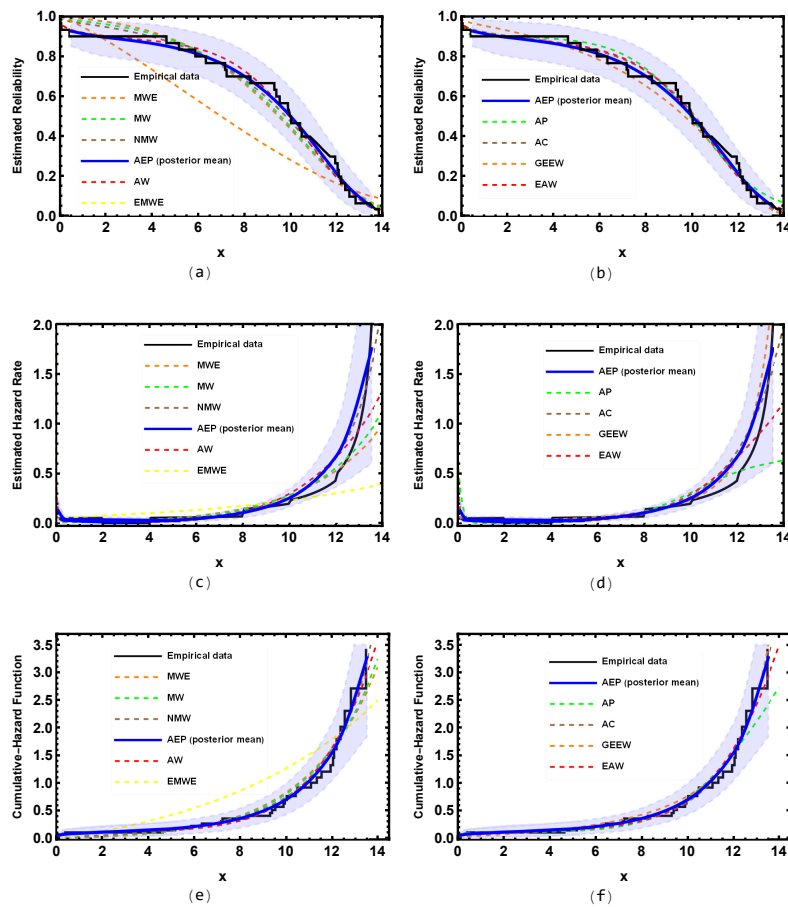


Figure 11. Pairwise scatter plots of the AEP parameters based on HMC posterior draws for the Sylwia dataset.

Table 9. MLEs for the parameters of the AEP model and all competing models fitted to the Sylwia dataset.

Model	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\lambda}$	$\hat{\theta}$
MWE	1.59523	0.00732	0.79176		
MW	0.01811	0.20607	0.33166		
NMW	0.02766	0.21773	0.25007		
AEP	5.29309×10^{-5}	0.26064	0.08079	4.08913	
AW	2.01027×10^{-6}	5.43358	0.08445	0.28963	
EMWE	167.058	3.41015	0.34297	42.3152	
AP	9.72103×10^{-4}	0.11034	6.49423	0.67742	
AC	0.04679	0.17949	6.96796×10^{-4}	0.81699	
GEEW	0.75568	0.02422	0.01493	80.0595	0.14709
EAW	6.06845×10^{-5}	4.20899	2.93817	0.03583	44.8662

Table 10. Goodness-of-fit statistics and model selection criteria for the AEP model and competing models fitted to the Sylwia dataset. Models are ranked (superscript numbers) from 1 (best) to 10 (worst) for each criterion.

Model	$-\ell$	AIC	AICc	BIC	HQIC	KS (p-value)	CVM (p-value)	AD (p-value)
MWE	77.088 ⁹	160.177 ⁹	161.1 ⁹	164.38 ⁹	161.521 ⁹	0.14463 (0.51088) ⁹	0.10494 (0.56408) ⁹	1.25859 (0.24602) ⁹
MW	73.163 ⁸	152.326 ⁸	153.249 ⁷	156.529 ⁶	153.671 ⁷	0.13186 (0.62677) ⁸	0.0825 (0.68105) ⁸	0.62862 (0.61996) ⁸
NMW	71.181 ⁷	148.362 ³	149.285 ²	152.565 ²	149.706 ³	0.1086 (0.83386) ⁶	0.04758 (0.89426) ⁵	0.36647 (0.88033) ⁵
AEP	69.245 ¹	146.489 ¹	148.089 ¹	152.094 ¹	148.282 ¹	0.08045 (0.98146) ¹	0.0261 (0.98842) ¹	0.1693 (0.9967) ¹
AW	70.327 ⁴	148.654 ⁴	150.254 ⁴	154.258 ⁴	150.447 ⁴	0.09639 (0.91809) ⁴	0.04316 (0.91939) ⁴	0.26593 (0.96113) ⁴
EMWE	85.986 ¹⁰	179.972 ¹⁰	181.572 ¹⁰	185.577 ¹⁰	181.765 ¹⁰	0.3366 (0.00156) ¹⁰	0.98294 (0.00242) ¹⁰	5.00406 (0.00292) ¹⁰
AP	70.986 ⁶	149.973 ⁵	151.573 ⁵	155.577 ⁵	151.766 ⁵	0.09751 (0.91152) ⁵	0.05301 (0.86147) ⁶	0.40533 (0.84233) ⁶
AC	69.894 ²	147.787 ²	149.387 ³	153.392 ³	149.58 ²	0.08416 (0.97162) ²	0.02666 (0.98714) ²	0.1731 (0.99616) ²
GEEW	70.768 ⁵	151.536 ⁷	154.036 ⁸	158.542 ⁸	153.777 ⁸	0.12774 (0.66497) ⁷	0.06102 (0.8113677) ⁷	0.40542 (0.84225) ⁷
EAW	70.049 ³	150.098 ⁶	152.598 ⁶	157.104 ⁷	152.339 ⁶	0.09332 (0.93459) ³	0.03609 (0.95471) ³	0.23016 (0.97986) ³

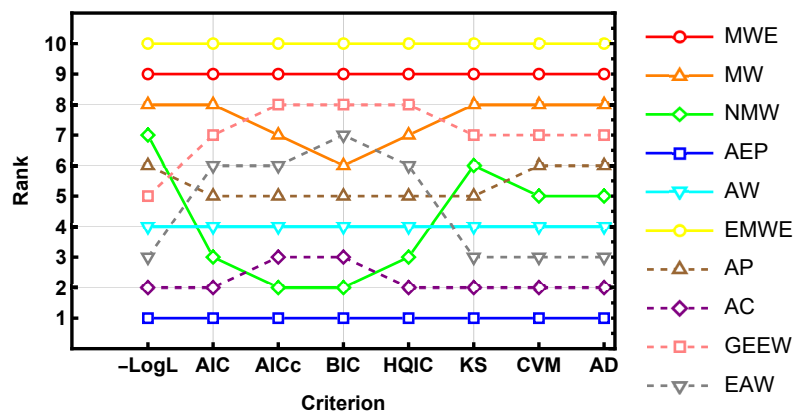


Figure 12. Ranking of fitted models based on performance metrics for the Sylwia dataset.

5.2. Case study II: Voltage endurance life of 58 electrodes

The second dataset is the voltage endurance test electrodes (VETEs) dataset reported by Doganaksoy et al. [50]. It comprises the lifetimes of 58 electrodes subjected to high-voltage stress, including 18 Mode E failures, 27 Mode D failures, and 13 right-censored observations. The presence of distinct failure mechanisms and censoring makes this dataset useful for evaluating the flexibility of the AEP model under a more complex reliability setting.

Figure 13 presents the empirical scaled TTT transform together with the posterior mean curve obtained from the fitted AEP model and its uncertainty band. The empirical curve departs systematically from the diagonal and exhibits a pattern consistent with non-monotone hazard behavior. The posterior mean AEP curve follows the empirical TTT pattern closely over most of the observed range, while the uncertainty band quantifies the corresponding posterior uncertainty.

Table 11 reports the MLEs and Bayesian estimates together with their 95% confidence and HPD intervals. The two estimation approaches give broadly comparable estimates, particularly for γ , while some differences are observed for β and λ . The HMC diagnostics show $\hat{R}$ values very close to one for all four parameters. The ESS values range from 277 for α to 756 for γ , indicating some variation in sampling efficiency across parameters. A total of 25 divergent transitions were observed across the four chains; consequently, the posterior results are assessed jointly with the trace plots, posterior densities,

and between-chain comparisons rather than relying on a single diagnostic measure.

The graphical diagnostics in Figure 14 show substantial overlap among the four chains, predominantly unimodal posterior densities, and broadly comparable chain-specific boxplots. Figure 15 provides further information about posterior dependence among the AEP parameters. The pairwise correlations indicate a moderate positive association between α and β ($r = 0.520$) and a moderate negative association between α and γ ($r = -0.411$). A weaker positive association is observed between α and λ ($r = 0.355$), whereas the remaining pairwise associations are relatively weak. These results indicate some posterior dependence among the scale and shape parameters, consistent with the discussion in Subsection 2.3.

Table 12 reports the MLEs of the AEP model and the nine competing models, while Table 13 summarizes the corresponding goodness-of-fit and model-selection criteria. Among the models considered, the AEP model attains the smallest $-\ell$, AIC, AICc, and HQIC values. For BIC, the NMW model gives the smallest value, whereas the AEP ranks second. The AEP also yields the smallest KS, CvM, and AD statistics among the fitted models.

Figure 16 complements the numerical results by comparing the estimated survival, hazard-rate, and cumulative-hazard functions with their empirical counterparts and those of the competing models. The AEP posterior mean curves generally follow the empirical behavior over the observed range, and the 95% HPD bands display the associated posterior uncertainty. In particular, the fitted hazard function captures the non-monotone behavior suggested by the empirical diagnostics. Taken together, the numerical and graphical results indicate that the AEP model provides a close and competitive fit to the VETEs dataset among the models considered. Figure 17 summarizes these comparisons and shows that the AEP ranks favorably across most of the reported criteria, while also making clear that its relative position depends on the criterion considered.

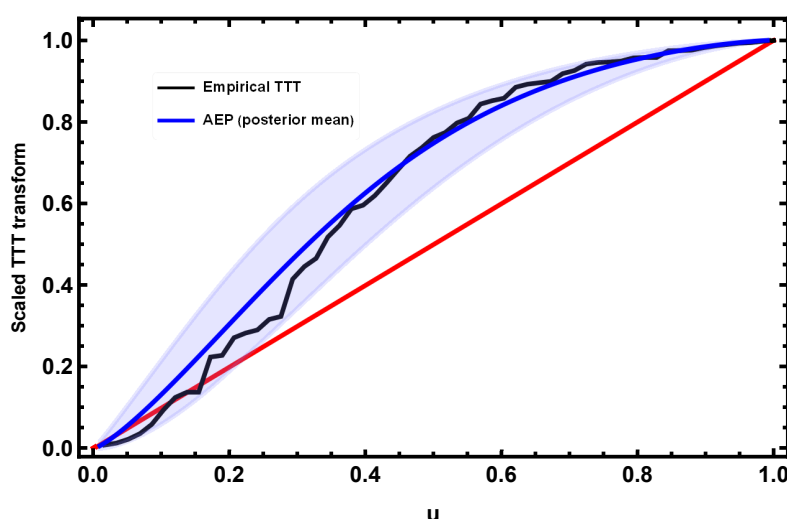


Figure 13. Comparison of empirical and AEP-fitted scaled TTT transform curves for the VETEs dataset.

Table 11. MLEs and Bayesian estimates with 95% CIs and HPD intervals, $\hat{R}$, ESS, and HMC diagnostics for the AEP model fitted to the VETEs dataset.

Parameter	MLEs	95% CI	Bayes	95% HPD	$\hat{R}$	ESS
$\hat{\alpha}$	0.00137	[-0.00001, 0.00275]	0.00157	[0.00076, 0.00231]	1.00025	277
$\hat{\beta}$	0.66208	[0.36202, 0.96213]	0.73340	[0.51658, 0.98599]	1.00004	403
$\hat{\gamma}$	0.00260	[0.00235, 0.00285]	0.00256	[0.00230, 0.00280]	0.99988	756
$\hat{\lambda}$	3.80844	[1.70362, 5.91325]	3.58219	[2.23946, 5.26760]	1.00017	741

Divergent transitions: Chain 1 = 6, Chain 2 = 7, Chain 3 = 6, Chain 4 = 6, Total = 25.

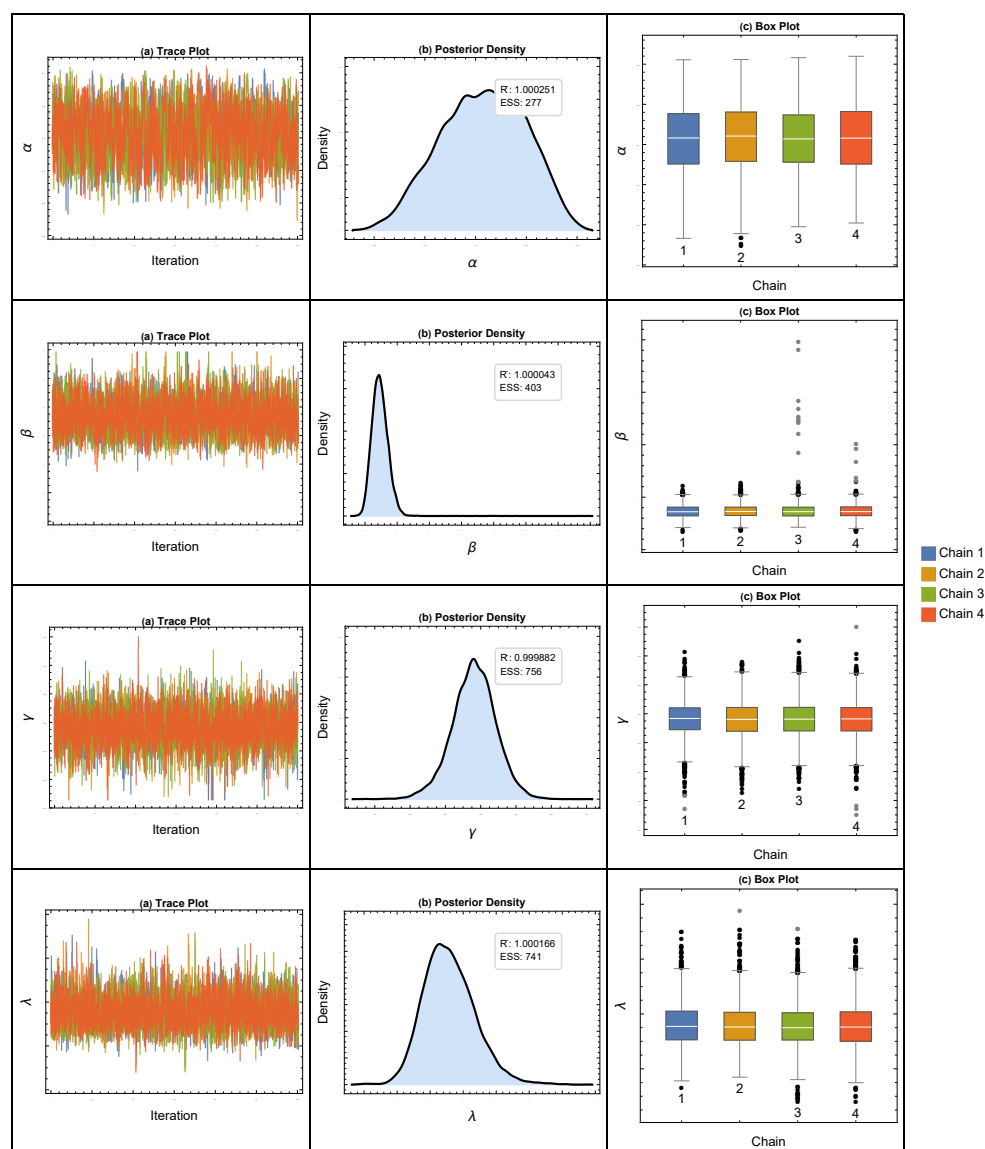


Figure 14. HMC diagnostics for the AEP parameters in the VETEs dataset: (a) Trace plots, (b) posterior density plots, and (c) chain-specific box plots.

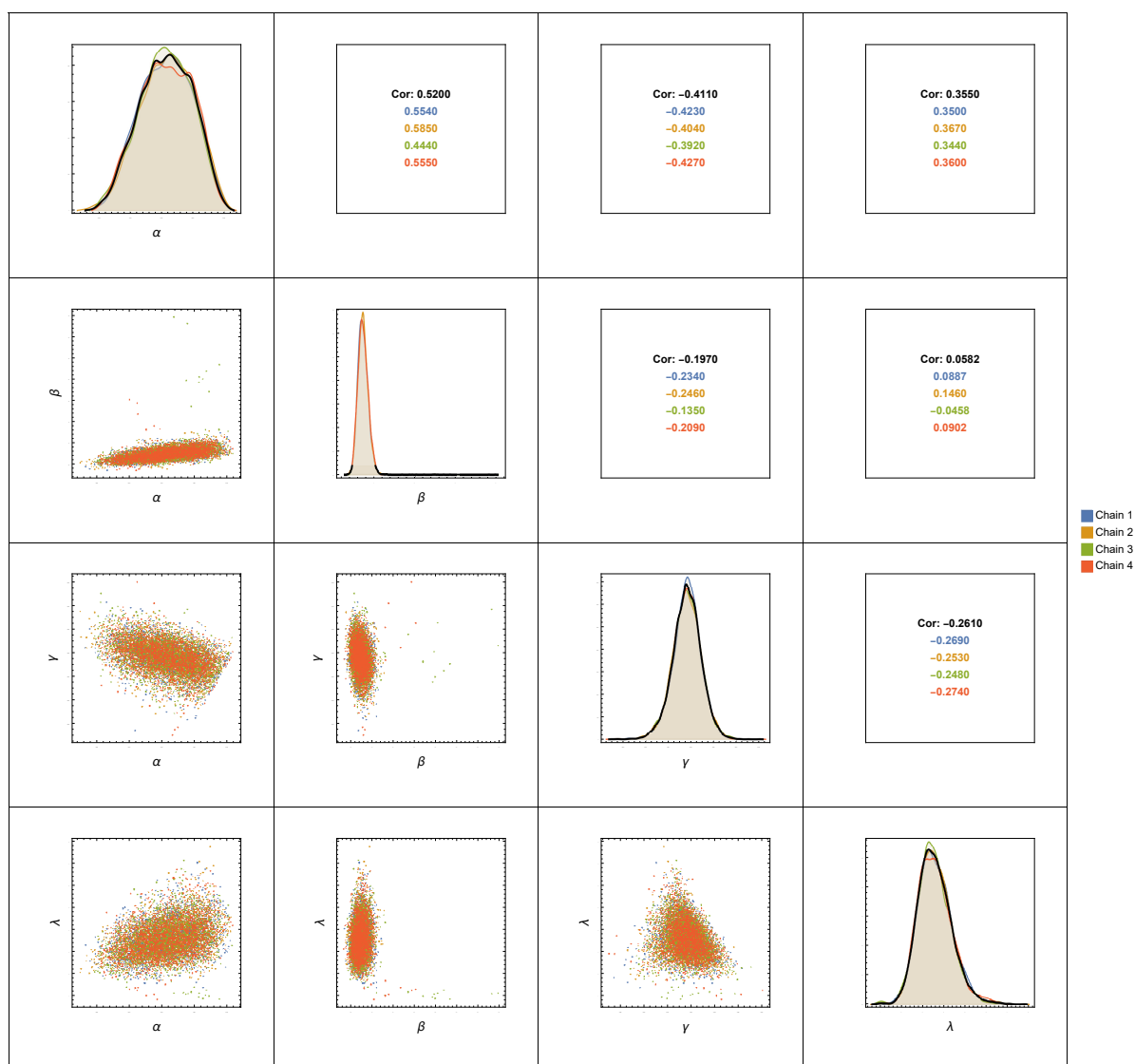


Figure 15. Pairwise scatter plots of the AEP parameters based on HMC posterior draws for the VETEs dataset.

Table 12. MLEs for the parameters of the AEP model and all competing models fitted to the VETEs dataset.

Model	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\lambda}$	$\hat{\theta}$
MWE	86.902	1.43051×10^{-3}	0.74902		
MW	0.01598	0.48172	5.73704×10^{-3}		
NMW	0.01557	0.5152	3.52531×10^{-3}		
AEP	1.36944×10^{-3}	0.66208	2.59582×10^{-3}	3.80843	
AW	2.13036×10^{-10}	3.84687	0.0133	0.67248	
EMWE	217.879	1.96542	0.33407	3.63949×10^{-4}	
AP	1.9221×10^{-5}	4.09437×10^{16}	3.08258×10^{-3}	0.03213	
AC	2.72812×10^{-6}	0.43454	0.01354	0.24991	
GEEW	1.17691	4.21506×10^{-3}	1.89723×10^{-4}	8.04426	181.349
EAW	4.91799×10^{-8}	2.97135	1.19627	0.12827	10.9288

Table 13. Goodness-of-fit statistics and model selection criteria for the AEP model and competing models fitted to the VETEs dataset. Models are ranked (superscript numbers) from 1 (best) to 10 (worst) for each criterion.

Model	$-\ell$	AIC	AICc	BIC	HQIC	KS (p-value)	CVM (p-value)	AD (p-value)
MWE	355.416 ⁹	716.831 ⁹	717.276 ⁹	723.013 ⁸	719.239 ⁹	0.10467 (0.5488) ⁹	0.16675 (0.34304) ⁹	1.17142 (0.27848) ⁹
MW	353.785 ⁸	713.571 ⁷	714.015 ⁷	719.752 ⁷	715.978 ⁷	0.09142 (0.71749) ⁸	0.10633 (0.5562) ⁸	0.6392 (0.61102) ⁸
NMW	352.065 ⁶	710.129 ⁴	710.574 ³	716.31 ¹	712.537 ³	0.0837 (0.81124) ⁷	0.06608 (0.77777) ⁷	0.3963 (0.85185) ⁷
AEP	350.509 ¹	709.018 ¹	709.773 ¹	717.26 ²	712.229 ¹	0.04871 (0.99914) ¹	0.01503 (0.99967) ¹	0.11726 (0.99985) ¹
AW	351.448 ⁵	710.896 ⁶	711.651 ⁶	719.138 ⁶	714.106 ⁶	0.0712 (0.93037) ⁵	0.04288 (0.91943) ⁵	0.25089 (0.96984) ⁴
EMWE	350.622 ²	709.245 ²	709.999 ²	717.486 ³	712.455 ²	0.05287 (0.99691) ³	0.01883 (0.99815) ²	0.13543 (0.99941) ²
AP	350.923 ³	709.846 ³	710.601 ⁴	718.088 ⁴	713.056 ⁴	0.06401 (0.9714) ⁴	0.02545 (0.98924) ⁴	0.29856 (0.939) ⁵
AC	351.197 ⁴	710.395 ⁵	711.15 ⁵	718.637 ⁵	713.605 ⁵	0.05262 (0.99712) ²	0.02301 (0.99367) ³	0.15391 (0.99832) ³
GEEW	359.518 ¹⁰	729.037 ¹⁰	730.191 ¹⁰	735.279 ¹⁰	733.05 ¹⁰	0.12909 (0.28858) ¹⁰	0.22459 (0.22472) ¹⁰	1.33011 (0.22274) ¹⁰
EAW	352.178 ⁷	714.355 ⁸	715.509 ⁸	724.658 ⁹	718.368 ⁸	0.07648 (0.88665) ⁶	0.05351 (0.85657) ⁶	0.35084 (0.89527) ⁶

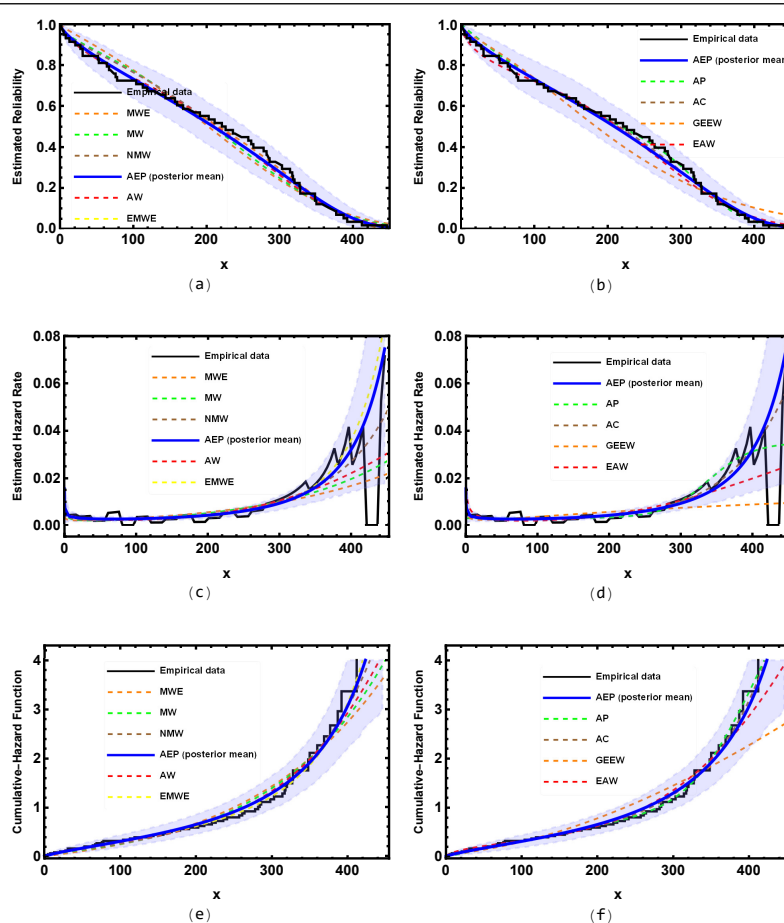


Figure 16. Estimated survival, HR, and cumulative-hazard functions obtained by fitting the AEP model and competing models to the VETEs dataset.

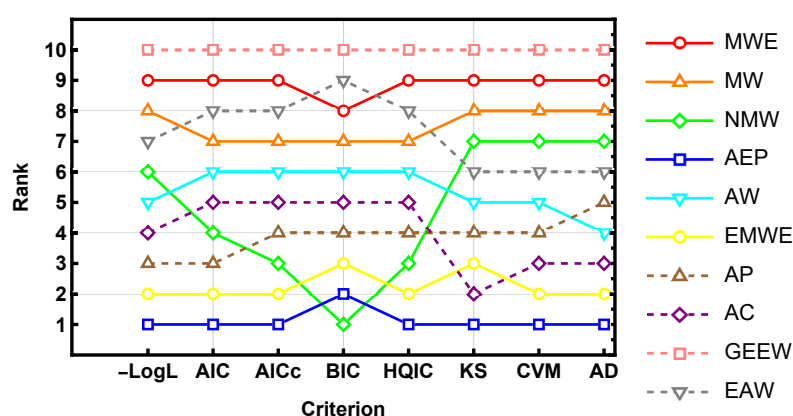


Figure 17. Ranking of fitted models based on performance metrics for the VETEs dataset.

5.3. Case study III: Survival times of 26 psychiatric patients

The third dataset consists of the survival times of 26 psychiatric patients (DTPPs), as reported by Klein and Moeschberger [51]. The dataset contains 12 right-censored observations, corresponding to approximately 46% censoring. The relatively small sample size and substantial censoring make this

application particularly useful for examining both the flexibility of the AEP model and the sensitivity of Bayesian inference to the prior specification.

Figure 18 presents the empirical scaled TTT transform together with the posterior mean curve obtained from the fitted AEP model and its uncertainty band. The empirical curve exhibits a pronounced departure from the diagonal, consistent with non-monotone hazard behavior. The posterior mean AEP curve reproduces the main empirical pattern, while the relatively wide uncertainty band in the early part of the curve reflects the uncertainty associated with the small and substantially censored sample.

Table 14 reports the MLEs and Bayesian estimates together with their 95% confidence and HPD intervals. The HMC diagnostics show $\hat{R}$ values very close to one for all four parameters. The ESS values range from 282 for α to 2195 for γ , indicating noticeably different sampling efficiencies across parameters. A total of 36 divergent transitions were observed across the four chains. Accordingly, the posterior estimates are interpreted jointly with the additional graphical diagnostics rather than relying on $\hat{R}$ alone.

Figure 19 shows substantial overlap among the four chains, predominantly unimodal posterior densities, and broadly comparable chain-specific distributions. The posterior distribution of α is noticeably right-skewed, consistent with its concentration near the lower boundary of the parameter space.

Figure 20 further examines posterior dependence among the AEP parameters. The most noticeable association occurs between α and β , with a moderate positive correlation ($r = 0.437$), whereas the remaining pairwise correlations are relatively weak. In particular, the correlation between γ and λ is -0.153 , while the absolute correlations for the other parameter pairs are below 0.10. These results indicate limited posterior dependence apart from the association between α and β , consistent with the scale–shape dependence discussed in Subsection 2.3.

Because this application has only $n = 26$ observations and a substantial proportion of censoring, a prior sensitivity analysis was conducted to assess the influence of the prior specification on posterior inference. The AEP model was therefore re-fitted using a common vague $\text{Gamma}(0.01, 0.01)$ prior for all four parameters and compared with the primary prior specification reported in Table 3. The results are summarized in Table 15.

The sensitivity analysis reveals a clear difference among the parameters. The posterior means of γ and λ change by only -0.3% and $+4.1\%$, respectively, and their 95% HPD intervals remain broadly similar under the two prior specifications. In contrast, α and β are substantially more prior-sensitive: Their posterior means change by approximately $+1158\%$ and $+134\%$, respectively, and their HPD intervals become considerably wider under the vague priors. Thus, for this small and heavily censored dataset, the posterior inference for γ and λ is considerably less sensitive to the prior specification than that for α and β . These results emphasize the importance of prior sensitivity assessment when applying the AEP model to small or highly censored samples, particularly for parameters that are weakly informed by the likelihood.

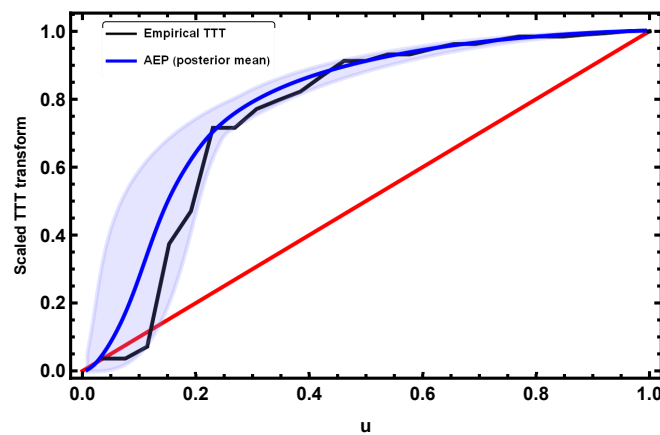


Figure 18. Comparison of empirical and AEP-fitted scaled TTT transform curves for the DTPPs dataset.

Table 14. MLEs and Bayesian estimates with 95% CIs and HPD intervals, $\hat{R}$, ESS, and HMC diagnostics for the AEP model fitted to the DTPPs dataset.

Parameter	MLEs	95% CI	Bayes	95% HPD	$\hat{R}$	ESS
$\hat{\alpha}$	0.00224	[-0.00585, 0.01034]	0.00012	$[4.79402 \times 10^{-6}, 0.00041]$	1.00249	282
$\hat{\beta}$	0.53743	[0.02754, 1.04732]	0.31461	[0.17508, 0.49639]	1.00044	535
$\hat{\gamma}$	0.02730	[0.02577, 0.02884]	0.02741	[0.02552, 0.02916]	0.99993	2195
$\hat{\lambda}$	5.38956	[2.80835, 7.97077]	4.56959	[2.85963, 6.75947]	1.00034	606

Divergent transitions: Chain 1 = 7, Chain 2 = 12, Chain 3 = 5, Chain 4 = 12, Total = 36.

Table 15. Prior sensitivity analysis for the DTPPs dataset ($n = 26$): Posterior means and 95% HPD intervals under the original Gamma priors versus a standard vague gamma(0.01, 0.01) specification.

Parameter	Original priors		Vague priors Ga(0.01, 0.01)		% Change in Mean
	Mean	95% HPD	Mean	95% HPD	
α	0.00012	$[4.79 \times 10^{-6}, 4.1 \times 10^{-4}]$	0.00151	$[1.78 \times 10^{-16}, 1.02 \times 10^{-2}]$	+1158%
β	0.31461	[0.175, 0.4964]	0.73631	[0.0729, 1.0577]	+134%
γ	0.02741	[0.02552, 0.02916]	0.02734	[0.02543, 0.02923]	-0.3%
λ	4.56959	[2.86, 6.7]	4.7553	[2.517, 7.728]	+4.1%

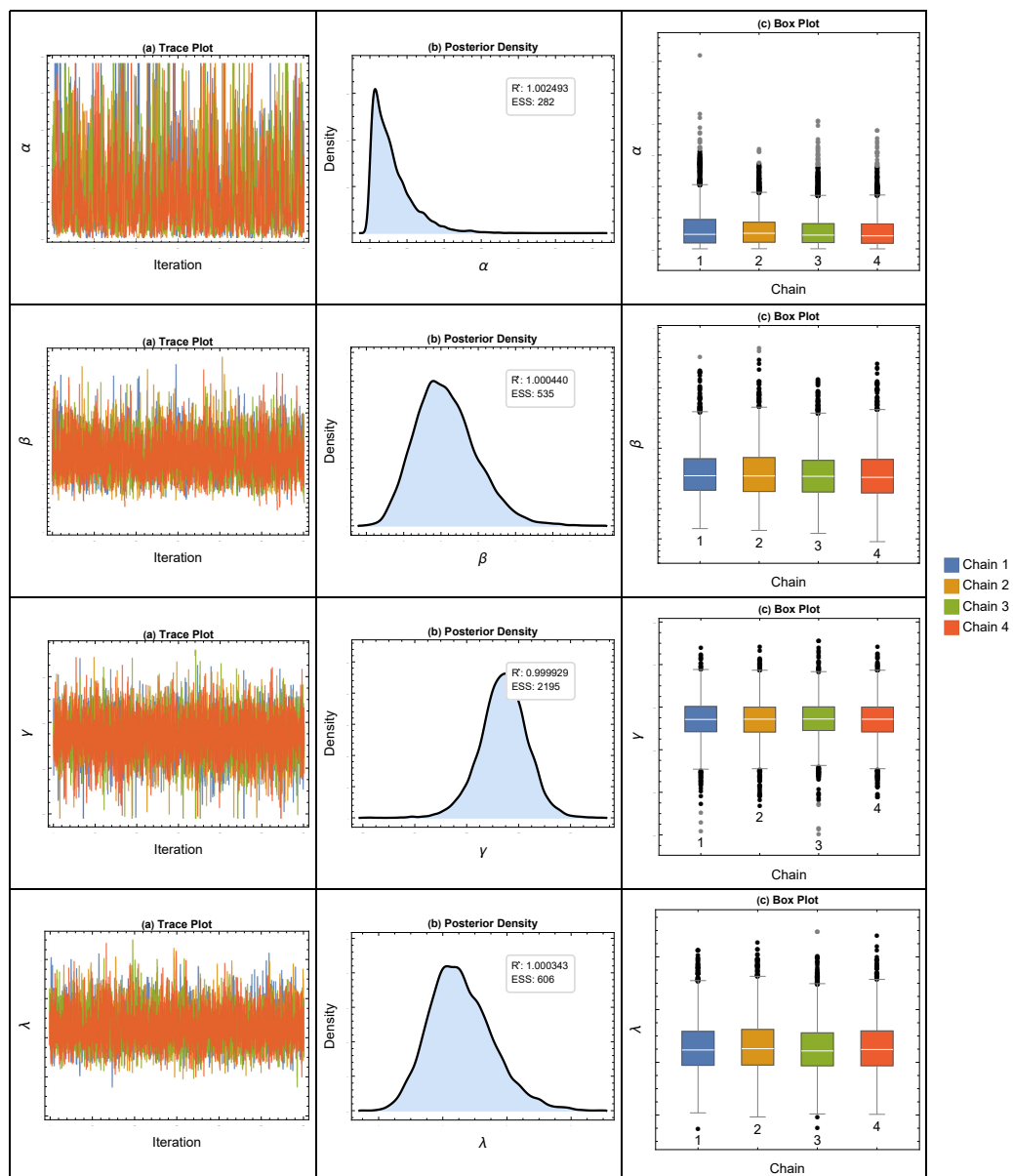


Figure 19. HMC diagnostics for the AEP parameters in the DTPPs dataset: (a) Trace plots, (b) posterior density plots, and (c) chain-specific box plots.

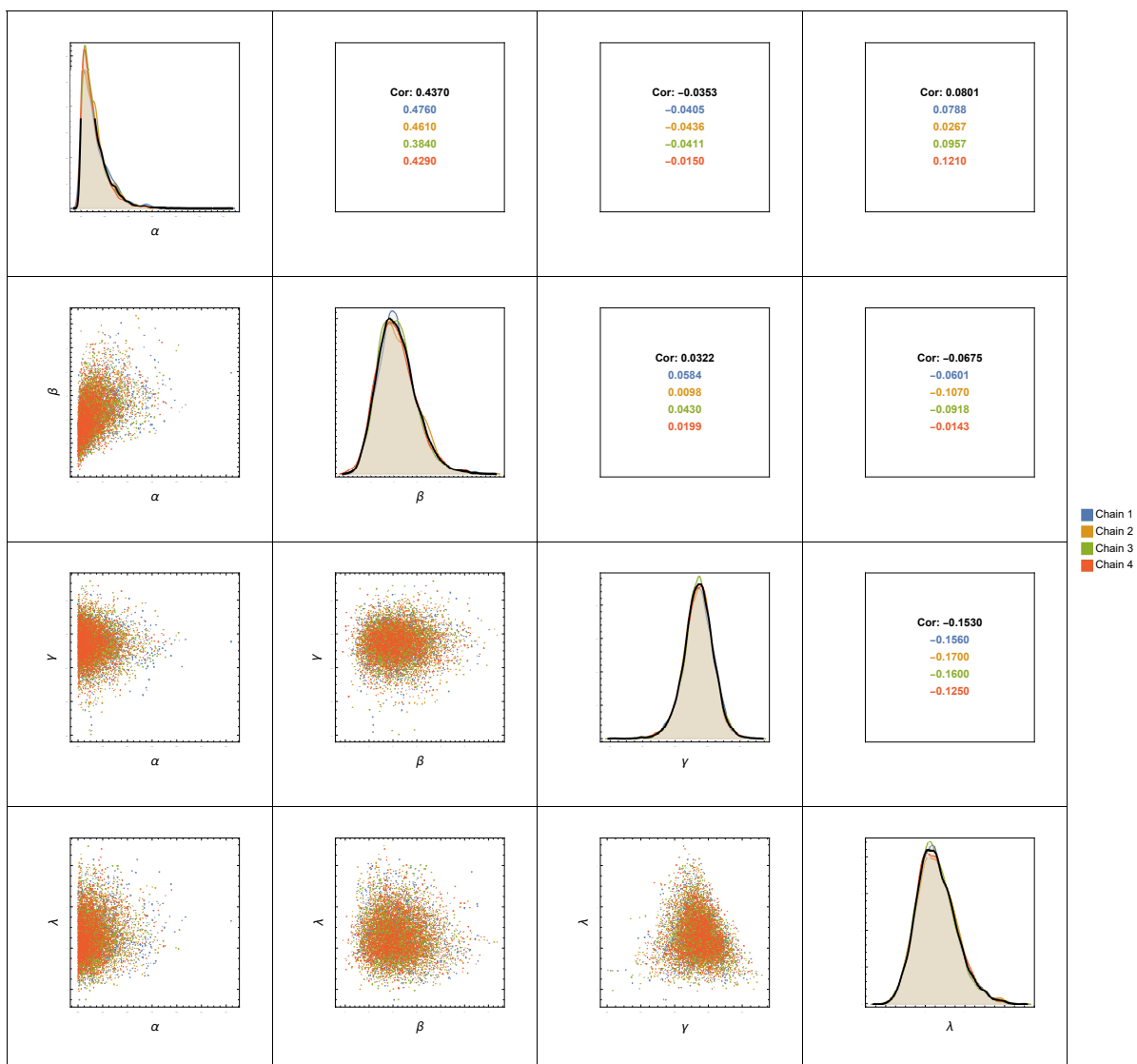


Figure 20. Pairwise scatter plots of the AEP parameters based on HMC posterior draws for the DTPPs dataset.

Table 16 reports the MLEs of the AEP model and the nine competing models, while Table 17 summarizes the corresponding goodness-of-fit and model-selection criteria. Among the models considered, the AEP model attains the smallest $-\ell$, AIC, AICc, BIC, and HQIC values. For the goodness-of-fit statistics, the AC model gives the smallest KS, CvM, and AD values, whereas the AEP model ranks second for these three criteria.

Figure 21 further compares the estimated survival, hazard-rate, and cumulative-hazard functions. The AEP posterior mean curves generally follow the corresponding empirical functions over the observed range, while the 95% HPD bands quantify the posterior uncertainty. The fitted hazard function reproduces the non-monotone behavior suggested by the empirical diagnostics, and the survival and cumulative-hazard curves show close agreement with their empirical counterparts. Taken together, the numerical and graphical results indicate that the AEP model provides a competitive fit to the DTPPs dataset, while the prior sensitivity results also demonstrate the importance of accounting for

posterior uncertainty in small, substantially censored samples. Figure 22 provides a visual summary of this criterion-dependent comparison.

Table 16. MLEs for the parameters of the AEP model and all competing models fitted to the DTPPs dataset.

Model	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\lambda}$	$\hat{\theta}$
MWE	10.1015	0.00309	1.10841		
MW	0.01059	0.30666	0.11236		
NMW	0.01475	0.35209	0.08353		
AEP	2.24358×10^{-3}	0.53743	0.0273	5.38957	
AW	1.77681×10^{-11}	7.03642	0.03842	0.57372	
EMWE	15.9336	2.24483	0.26549	7.60944×10^{-5}	
AP	3.59944×10^{-5}	0.24781	0.14241	0.32356	
AC	0.0265	0.26926	0.00001	0.69619	
GEEW	0.20579	0.19167	0.01546	17.2776	0.17554
EAW	2.82839×10^{-8}	5.06028	3.8132	0.05632	143.791

Table 17. Goodness-of-fit statistics and model selection criteria for the AEP model and competing models fitted to the DTPPs dataset. Models are ranked (superscript numbers) from 1 (best) to 10 (worst) for each criterion.

Model	$-\ell$	AIC	AICc	BIC	HQIC	KS (p-value)	CVM (p-value)	AD (p-value)
MWE	95.021 ⁹	196.042 ⁹	197.133 ⁹	199.817 ⁹	197.129 ⁹	0.15058 (0.5973) ⁹	0.09985 (0.58912) ⁹	1.18584 (0.27258) ⁹
MW	94.352 ⁸	194.703 ⁸	195.794 ⁷	198.477 ⁶	195.79 ⁷	0.15988 (0.51959) ⁸	0.11081 (0.53758) ⁸	0.83758 (0.45349) ⁸
NMW	92.331 ⁷	190.662 ³	191.753 ²	194.436 ²	191.749 ³	0.13634 (0.71922) ⁶	0.07269 (0.73943) ⁵	0.54956 (0.69525) ⁵
AEP	89.671 ¹	187.341 ¹	189.246 ¹	192.374 ¹	188.79 ¹	0.07448 (0.99873) ¹	0.01777 (0.99895) ¹	0.16311 (0.99749) ¹
AW	89.799 ⁴	187.599 ⁴	189.504 ⁴	192.631 ⁴	189.048 ⁴	0.08539 (0.99141) ⁴	0.02299 (0.99423) ⁴	0.17053 (0.99656) ⁴
EMWE	90.118 ¹⁰	188.236 ¹⁰	190.141 ¹⁰	193.268 ¹⁰	189.685 ¹⁰	0.0935 (0.9769) ¹⁰	0.03044 (0.97682) ¹⁰	0.24589 (0.97241) ¹⁰
AP	90.673 ⁶	189.347 ⁵	191.252 ⁵	194.379 ⁵	190.796 ⁵	0.07848 (0.99719) ⁵	0.02968 (0.97922) ⁶	0.24363 (0.97357) ⁶
AC	90.046 ²	188.092 ²	189.996 ³	193.124 ³	189.541 ²	0.06387 (0.99993) ²	0.0162 (0.9995) ²	0.14698 (0.99888) ²
GEEW	93.783 ⁵	197.565 ⁷	200.565 ⁸	200.597 ⁸	199.376 ⁸	0.2031 (0.23376) ⁷	0.23936 (0.2029) ⁷	1.328 (0.22331) ⁷
EAW	89.825 ³	189.65 ⁶	192.65 ⁶	195.94 ⁷	191.461 ⁶	0.11935 (0.85277) ³	0.04877 (0.88775) ³	0.28169 (0.95088) ³

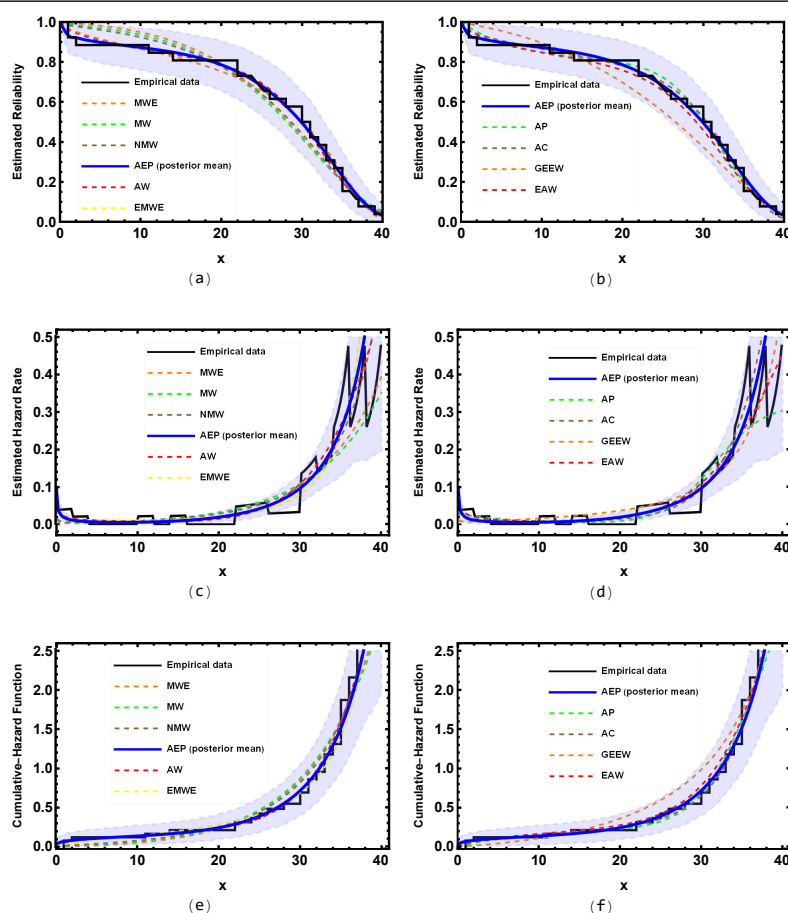


Figure 21. Estimated survival, HR, and cumulative hazard functions obtained by fitting the AEP model and competing models to the DTPPs dataset.

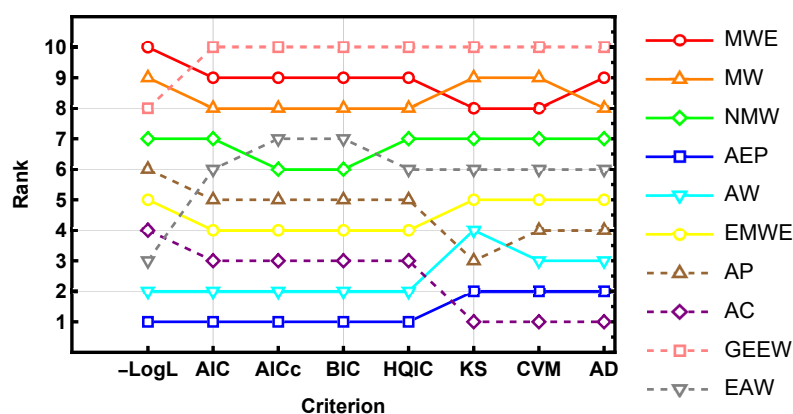


Figure 22. Ranking of fitted models based on performance metrics for the DTPPs dataset.

6. Conclusions

This study introduced the additive exponential power (AEP) model as a flexible four-parameter lifetime model constructed from two exponential power hazard components. The additive structure

provides a unified framework for representing a broad range of hazard-rate behaviors, including increasing, bathtub-type, and more complex non-monotone patterns. Several distributional and reliability properties were derived, and parameter estimation was developed using both maximum likelihood and Bayesian approaches. To address component exchangeability, an identifying parameterization was imposed, while Bayesian inference was implemented using Hamiltonian Monte Carlo with the No-U-Turn sampler (HMC–NUTS).

The main contribution of the AEP model lies in its structural flexibility: Each exponential power component can represent a distinct failure mechanism, while their additive combination allows more complex hazard patterns to be accommodated. The simulation study and three real-data applications illustrate the practical applicability of the model under different lifetime-data settings.

The study also identifies several practical limitations. Estimation can be challenging when component strengths are highly unequal or shape parameters are close to transition values, and the combined hazard may require numerical determination of its critical points. Bayesian inference may also exhibit sensitivity to prior specification when the sample is small or substantially censored, particularly for parameters weakly informed by the likelihood. These findings emphasize the importance of careful initialization, prior sensitivity analysis, and convergence diagnostics when applying the model. Future work may consider regression extensions incorporating covariates, accelerated life-testing and repairable-system applications, and more computationally efficient Bayesian methods for large-scale data.

Author contributions

M. G. M. Ghazal: Conceptualization, methodology, software, validation, formal analysis, investigation, data curation, writing—original draft preparation, writing—review and editing, visualization; Ghadah Alomani: Funding acquisition, resources, writing—review and editing; Khalaf S. Sultan: Supervision, writing—review and editing; Hossam M. M. Radwan: Validation, supervision, writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Use of Generative-AI tools declaration

The authors confirm that no Artificial Intelligence (AI) tools were used in preparing this manuscript.

Funding

This research project was funded by Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2026R226), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Acknowledgments

The authors acknowledge that this research project was funded by Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2026R226), Princess

Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Conflict of interest

The authors declare that they have no conflicts of interest to report regarding the present study.

References

1. W. Q. Meeker, L. A. Escobar, *Statistical methods for reliability data*, John Wiley & Sons, 1998.
2. W. Weibull, A statistical distribution function of wide applicability, *J. Appl. Mech.*, **18** (1951), 293–297. <http://dx.doi.org/10.1115/1.4010337>
3. W. H. Bartley, *Analysis of transformer failures*, WGP 33 (03), Proceedings of the 36th Annual Conference on International Association of Engineering, 2003.
4. M. K. Shakhathreh, A. J. Lemonte, G. M. Arenas, The log-normal modified Weibull distribution and its reliability implications, *Reliab. Eng. Syst. Safe.*, **188** (2019), 6–22. <http://dx.doi.org/10.1016/j.ress.2019.03.014>
5. P. Kosky, R. Balmer, W. Keat, G. Wise, *Industrial engineering*, In: Exploring Engineering, Elsevier, 2021, 229–257. <https://doi.org/10.1016/B978-0-12-815073-3.00011-9>
6. S. J. Almalki, J. Yuan, A new modified Weibull distribution, *Reliab. Eng. Syst. Safe.*, **111** (2013), 164–170. <http://dx.doi.org/10.1016/j.ress.2012.10.018>
7. T. T. Thach, R. Bris, An additive Chen-Weibull distribution and its applications in reliability modeling, *Qual. Reliab. Eng. Int.*, **37** (2021), 352–373. <http://dx.doi.org/10.1002/qre.2740>
8. B. Abba, H. Wang, H. S. Bakouch, A reliability and survival model for one and two failure modes system with applications to complete and censored datasets, *Reliab. Eng. Syst. Safe.*, **223** (2022), 108460. <http://dx.doi.org/10.1016/j.ress.2022.108460>
9. A. U. Abdullahi, H. Wang, Y. Liu, Y. Qin, Robust failure time model with Bayesian inference and its application to reliability analysis, *Qual. Reliab. Eng. Int.*, 2025. <http://dx.doi.org/10.1002/qre.3821>
10. R. Jiang, D. N. Murthy, Reliability modeling involving two Weibull distributions, *Reliab. Eng. Syst. Safe.*, **47** (1995), 187–198. [http://dx.doi.org/10.1016/0951-8320\(94\)00045-P](http://dx.doi.org/10.1016/0951-8320(94)00045-P)
11. F. Wang, A new model with bathtub-shaped failure rate using an additive Burr XII distribution, *Reliab. Eng. Syst. Safe.*, **70** (2000), 305–312. [http://dx.doi.org/10.1016/S0951-8320\(00\)00066-1](http://dx.doi.org/10.1016/S0951-8320(00)00066-1)
12. B. He, W. Cui, X. Du, An additive modified Weibull distribution, *Reliab. Eng. Syst. Safe.*, **145** (2016), 28–37. <http://dx.doi.org/10.1016/j.ress.2015.08.010>
13. B. Singh, An additive perks-Weibull model with bathtub-shaped hazard rate function, *Commun. Math. Stat.*, **4** (2016), 473–493. <http://dx.doi.org/10.1007/s40304-016-0096-z>
14. B. Tarvirdizade, M. Ahmadpour, A new extension of Chen distribution with applications to lifetime data, *Commun. Math. Stat.*, **9** (2021), 23–38. <http://dx.doi.org/10.1007/s40304-019-00185-4>

15. T. Xavier, J. K. Jose, S. Nadarajah, An additive power-transformed half-logistic model and its applications in reliability, *Qual. Reliab. Eng. Int.*, **38** (2022), 3179–3196. <http://dx.doi.org/10.1002/qre.3119>
16. L. C. Méndez-González, L. A. Rodríguez-Picón, I. J. Pérez-Olguin, V. García, D. Luviano-Cruz, The additive Perks distribution and its applications in reliability analysis, *Qual. Technol. Quant. M.*, **20** (2023), 784–808. <http://dx.doi.org/10.1080/16843703.2022.2148884>
17. L. C. Méndez-González, L. A. Rodríguez-Picón, I. J. Pérez-Olguin, R. L. Portilla, An additive Chen distribution with applications to lifetime data, *Axioms*, **12** (2023), 118. <http://dx.doi.org/10.3390/axioms12020118>
18. L. C. Méndez-González, L. A. Rodríguez-Picón, M. I. Rodríguez-Borbón, H. Sohn, The Chen-Perks distribution: Properties and reliability applications, *Mathematics*, **11** (2023), 3001. <http://dx.doi.org/10.3390/math11133001>
19. D. Jiang, Y. Han, W. Cui, F. Wan, T. Yu, B. Song, An improved modified Weibull distribution applied to predict the reliability evolution of an aircraft lock mechanism, *Probabilist. Eng. Mech.*, **72** (2023), 103449. <http://dx.doi.org/10.1016/j.probengmech.2023.103449>
20. B. Abba, H. Wang, M. Muhammad, H. S. Bakouch, A robust bathtub-shaped failure time model for a two-component system with applications to complete and censored reliability data, *Qual. Technol. Quant. M.*, **21** (2024), 309–339. <http://dx.doi.org/10.1080/16843703.2023.2193771>
21. B. Abba, H. Wang, A new failure times model for one and two failure modes system: A Bayesian study with Hamiltonian Monte Carlo simulation, *P. I. Mech. Eng. O-J. Risk Reliab.*, **238** (2024), 304–323. <http://dx.doi.org/10.1177/1748006X221146367>
22. B. Abba, J. Wu, M. Muhammad, A robust multi-risk model and its reliability relevance: A Bayes study with Hamiltonian Monte Carlo methodology, *Reliab. Eng. Syst. Safe.*, **250** (2024), 110310. <http://dx.doi.org/10.1016/j.ress.2024.110310>
23. F. Amiru, U. Usman, S. Shamsuddeen, U. Adamu, B. Abba, The additive Dhillon-Chen distribution: Properties and applications to failure time data, *Int. J. Stat. Distrib. Appl.*, **11** (2025), 1–10. <http://dx.doi.org/10.11648/j.ijdsda.20251101.11>
24. M. G. M. Ghazal, Hamiltonian Monte Carlo-based inference for the flexible exponential power–Weibull distribution with applications in reliability analysis, *AIMS Math.*, **11** (2026), 11659–11705. <https://doi.org/10.3934/math.2026481>
25. R. M. Smith, L. J. Bain, An exponential power life-testing distribution, *Commun. Stat.-Theory M.*, **4** (1975), 469–481. <http://dx.doi.org/10.1080/03610927508827263>
26. E. L. Lehmann, G. Casella, *Theory of point estimation*, 2Eds., Springer, 1998.
27. A. W. van der Vaart, *Asymptotic statistics*, Cambridge University Press, 1998.
28. N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, E. Teller, Equation of state calculations by fast computing machines, *J. Chem. Phys.*, **21** (1953), 1087–1092. <http://dx.doi.org/10.1063/1.1699114>
29. W. K. Hastings, Monte Carlo sampling methods using Markov chains and their applications, *Biometrika*, **57** (1970), 97–109. <http://dx.doi.org/10.1093/biomet/57.1.97>

30. A. E. Gelfand, Gibbs Sampling, *J. Am. Stat. Assoc.*, **95** (2000), 1300–1304. <https://doi.org/10.2307/2669775>
31. R. M. Neal, *MCMC using Hamiltonian dynamics*, In: S. Brooks, A. Gelman, G. Jones, X. L. Meng Eds., *Handbook of Markov Chain Monte Carlo*, CRC Press, Boca Raton, 2011, 113–162.
32. A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, D. B. Rubin, *Bayesian data analysis*, CRC Press, Boca Raton, 2014. <https://doi.org/10.1201/b16018>
33. T. T. Thach, R. Bris, Improved new modified Weibull model: A Bayes study using hamiltonian monte carlo simulation, *P. I. Mech. Eng. O-J. Risk Reliab.*, **234** (2020), 496–511. <https://doi.org/10.1177/1748006X19896740>
34. T. T. Thach, R. Bris, P. Volf, F. P. A. Coolen, Non-linear failure rate: A Bayes study using hamiltonian monte carlo simulation, *Int. J. Approx. Reason.*, **123** (2020), 55–76. <https://doi.org/10.1016/j.ijar.2020.04.007>
35. M. D. Hoffman, A. Gelman, The no-u-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo, *J. Mach. Learn. Res.*, **15** (2014), 1593–1623.
36. M. Betancourt, A conceptual introduction to Hamiltonian Monte Carlo, *arXiv Preprint*, 2017. <http://arxiv.org/abs/1701.02434>
37. M. Nishio, A. Arakawa, Performance of Hamiltonian Monte Carlo and No-U-Turn sampler for estimating genetic parameters and breeding values, *Genet. Sel. Evol.*, **51** (2019). <https://doi.org/10.1186/s12711-019-0515-1>
38. S. Thomas, W. Tu, Learning Hamiltonian Monte Carlo in R, *Am. Stat.*, **75** (2021), 403–413. <https://doi.org/10.1080/00031305.2020.1865198>
39. L. Devlin, M. Carter, P. R. Horridge, P. L. Green, S. Maskell, The No-U-Turn Sampler as a proposal distribution in a sequential Monte Carlo sampler without accept/reject, *IEEE Signal Proc. Let.*, **31** (2024), 1089–1093. <https://doi.org/10.1109/LSP.2024.3386494>
40. D. Levy, M. D. Hoffman, J. Sohl-Dickstein, *Generalizing Hamiltonian Monte Carlo with neural networks*, In: International Conference on Learning Representations (ICLR), 2018. <https://doi.org/10.48550/arXiv.1711.09268>
41. X. Meng, J. L. Beck, Y. Huang, H. Li, Adaptive meta-learning stochastic gradient Hamiltonian Monte Carlo simulation for Bayesian updating of structural dynamic models, *Comput. Method. Appl. M.*, **437** (2025), 117753. <https://doi.org/10.1016/j.cma.2025.117753>
42. X. Meng, Y. Huang, J. L. Beck, K. Jiang, H. Li, MCMC with adaptive principal-component transformation: Rotation-invariant universal samplers for Bayesian structural system identification, *Adv. Eng. Inform.*, **74** (2026), 104758. <https://doi.org/10.1016/j.aei.2026.104758>
43. M. Xie, Y. Tang, T. N. Goh, A modified Weibull extension with bathtub-shaped failure rate function, *Reliab. Eng. Syst. Safe.*, **76** (2002), 279–285. [https://doi.org/10.1016/S0951-8320\(02\)00022-4](https://doi.org/10.1016/S0951-8320(02)00022-4)
44. C. D. Lai, M. Xie, D. N. P. Murthy, A modified Weibull distribution, *IEEE T. Reliab.*, **52** (2003), 33–37. <https://doi.org/10.1109/TR.2002.805788>

45. M. G. M. Ghazal, A new extension of the modified Weibull distribution with applications for engineering data, *Probabilist. Eng. Mech.*, **74** (2023), 103523. <https://doi.org/10.1016/j.probengmech.2023.103523>
46. A. M. Sarhan, J. Apaloo, Exponentiated modified Weibull extension distribution, *Reliab. Eng. Syst. Safe.*, **112** (2013), 137–144. <https://doi.org/10.1016/j.res.2012.10.013>
47. M. K. Shakhathreh, A. J. Lemonte, G. M. Cordeiro, On the generalized extended exponential-Weibull distribution: Properties and different methods of estimation, *Int. J. Comput. Math.*, **97** (2020), 1029–1057. <https://doi.org/10.1080/00207160.2019.1605062>
48. A. E. Ahmad, M. G. Ghazal, Exponentiated additive Weibull distribution, *Reliab. Eng. Syst. Safe.*, **193** (2020), 106663. <http://dx.doi.org/10.1016/j.res.2019.106663>
49. S. Kosznik-Biernacka, Makeham's generalised distribution, *Comput. Method. Sci. Technol.*, **13** (2007), 113–120. <https://doi.org/10.12921/cmst.2007.13.02.113-120>
50. N. Doganaksoy, G. Hahn, W. Meeker, Reliability analysis by failure mode, *Qual. Prog.*, **35** (2002), 47–52.
51. J. P. Klein, M. L. Moeschberger, *Survival analysis: Techniques for censored and truncated data*, Springer, 2003.

Appendix

A. Proofs of the properties of the AEP model

In this appendix, we derive detailed proofs for the theorems on the statistical properties of the AEP model stated in Section 2. Using series expansions and integral transformations, each subsection is dedicated to a specific property, followed by the proof and relevant mathematical arguments.

A.1. Proof of Theorem 1: Moments

Proof. The k -th raw moment about the origin for the AEP model is defined as

$$\mu'_k = \mathbb{E}[X^k] = \int_0^\infty k x^{k-1} S(x; \psi) dx,$$

where the survival function $S(x; \psi)$ is given in Eq (2.4). Substituting, we obtain

$$\mu'_k = k e^2 \int_0^\infty x^{k-1} e^{-e^{(\alpha x)^\beta} - e^{(\gamma x)^\lambda}} dx.$$

To evaluate this integral, we employ a series expansion for the exponential term $e^{-e^{(\alpha x)^\beta}}$:

$$e^{-e^{(\alpha x)^\beta}} = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} e^{n(\alpha x)^\beta} = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^n n^m \alpha^{\beta m}}{n! m!} x^{\beta m}.$$

Substituting this expansion into the integral gives

$$\mu'_k = k e^2 \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^n n^m \alpha^{\beta m}}{n! m!} \int_0^\infty x^{k+\beta m-1} e^{-e^{(\gamma x)^\lambda}} dx.$$

Applying the substitution $u = e^{(\gamma x)^\lambda}$, which implies $x = \frac{1}{\gamma}(\ln u)^{1/\lambda}$ and $dx = \frac{1}{\gamma\lambda}(\ln u)^{1/\lambda-1} \frac{du}{u}$, the integral transforms into

$$\mu'_k = \frac{k e^2}{\lambda \gamma^k} \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^n n^m}{n! m!} \left(\frac{\alpha}{\gamma}\right)^{\beta m} \int_1^{\infty} \frac{(\ln u)^{\frac{k+\beta m}{\lambda}-1}}{u} e^{-u} du.$$

Recognizing the integral as $G\left(\frac{k+\beta m}{\lambda} - 1\right)$ where $G(a) = \int_1^{\infty} \frac{(\ln t)^a}{t} e^{-t} dt$, we arrive at the desired result. $\square$

A.2. Proof of Theorem 2: Incomplete moments

Proof. The s -th incomplete moment of the AEP model is defined as:

$$\mu'_s(z) = E[X^s | X \leq z] = \int_0^z x^s f(x; \psi) dx.$$

Substituting the PDF from Eq (2.4) gives:

$$\begin{aligned} \mu'_s(z) &= e^2 \int_0^z x^s \left(\alpha \beta (\alpha x)^{\beta-1} e^{(\alpha x)^\beta} + \gamma \lambda (\gamma x)^{\lambda-1} e^{(\gamma x)^\lambda} \right) \\ &\quad \times e^{-e^{(\alpha x)^\beta} - e^{(\gamma x)^\lambda}} dx. \end{aligned}$$

Expand the terms $e^{-e^{(\alpha x)^\beta}}$ and $e^{-e^{(\gamma x)^\lambda}}$ using Taylor series gives:

$$e^{-e^{(\alpha x)^\beta}} = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^n n^m \alpha^{\beta m}}{n! m!} x^{\beta m}, \quad e^{-e^{(\gamma x)^\lambda}} = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} e^{k(\gamma x)^\lambda}.$$

Substituting these into the integral yields a sum of two integrals. The first integral, denoted I_1 , arises from the first component of the HR function:

$$I_1 = e^2 \beta \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \sum_{k=0}^{\infty} \frac{(-1)^{n+k} n^m \alpha^{(m+1)\beta}}{n! m! k!} \int_0^z x^{s+(m+1)\beta-1} e^{(\alpha x)^\beta} e^{k(\gamma x)^\lambda} dx.$$

Expanding $e^{(\alpha x)^\beta}$ using the series $\sum_{l=0}^{\infty} \frac{(\alpha x)^{\beta l}}{l!}$ and applying the standard identity:

$$\int_0^t u^{a-1} e^{\alpha u^\beta} du = \sum_{m=0}^{\infty} \frac{\alpha^m t^{a+\beta m}}{m! (a + \beta m)},$$

we obtain:

$$I_1 = e^2 \beta \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \sum_{i=0}^{\infty} \frac{k^i \alpha^{(m+l+1)\beta} \gamma^{i\lambda} z^{s+(m+l+1)\beta+i\lambda}}{n! m! k! l! i! (s + (m+l+1)\beta + i\lambda)}.$$

Similarly, applying the same identity to the second integral gives:

$$I_2 = e^2 \lambda \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} \frac{(k+1)^j \alpha^{m\beta} \gamma^{(j+1)\lambda} z^{s+m\beta+(j+1)\lambda}}{n! m! k! j! (s + m\beta + (j+1)\lambda)}.$$

Combining these results, we arrive at the desired representation of the incomplete moments. This completes the proof. $\square$

A.3. Proof of Theorem 3: Rényi entropy

Proof. The Rényi entropy of order q for a continuous random variable X with density $f(x; \psi)$ is defined as

$$H_q(X) = \frac{1}{1-q} \log \int_0^\infty [f(x; \psi)]^q dx.$$

Substituting the PDF from Eq (2.4) into this definition gives:

$$I_q = e^{2q} \int_0^\infty \left[\alpha \beta (\alpha x)^{\beta-1} e^{(\alpha x)^\beta} + \gamma \lambda (\gamma x)^{\lambda-1} e^{(\gamma x)^\lambda} \right]^q e^{-q e^{(\alpha x)^\beta} - q e^{(\gamma x)^\lambda}} dx,$$

Applying the binomial theorem to the q -th power of the sum yields:

$$\begin{aligned} I_q &= e^{2q} \sum_{k=0}^q \binom{q}{k} \alpha^{(q-k)\beta} \beta^{q-k} \gamma^k \lambda^k \\ &\quad \times \int_0^\infty x^{(q-k)\beta - q + k\lambda} e^{(q-k)(\alpha x)^\beta + k e^{(\gamma x)^\lambda}} e^{-q e^{(\alpha x)^\beta} - q e^{(\gamma x)^\lambda}} dx. \end{aligned}$$

The integrand contains exponential terms that generally do not admit closed-form solutions. Expanding the exponentials $e^{-q e^{(\alpha x)^\beta}}$ and $e^{-q e^{(\gamma x)^\lambda}}$, the entropy becomes

$$\begin{aligned} I_q &= e^{2q} \sum_{k=0}^q \sum_{i=0}^\infty \sum_{j=0}^\infty \binom{q}{k} \frac{(-q)^{i+j} \alpha^{(q-k)\beta} \beta^{q-k} \gamma^k \lambda^k}{i! j!} \\ &\quad \times \int_0^\infty x^{(q-k)\beta + k\lambda - q} e^{(q-k+i)(\alpha x)^\beta + (k+j) e^{(\gamma x)^\lambda}} dx. \end{aligned}$$

Further expanding $e^{(q-k+i)(\alpha x)^\beta}$ and substituting into the integral leads to the final expression after rearranging the summations and integrals. The integral $I_{k,j,l}$ requires numerical evaluation for most parameter values. This completes the proof. $\square$

A.4. Proof of Theorem 4: Mean time to failure

Proof. The MTTF is the first raw moment, $\mathbb{E}[X] = \mu'_1$. The result follows directly from Theorem 1 by setting with $k = 1$:

$$\text{MTTF} = \mu'_1 = \frac{e^2}{\gamma \lambda} \sum_{n=0}^\infty \sum_{m=0}^\infty \frac{(-1)^n n^m}{n! m!} \left(\frac{\alpha}{\gamma} \right)^{m\beta} G\left(\frac{\beta m + 1}{\lambda} - 1 \right).$$

This completes the proof. $\square$

A.5. Proof of Theorem 5: Mean residual life

Proof. The MRL function is defined as

$$m(t) = \mathbb{E}[X - t \mid X > t] = \frac{1}{S(t; \psi)} \int_t^\infty S(x; \psi) dx.$$

Substituting the survival function from Eq (2.3) gives:

$$m(t) = \frac{1}{e^{-e^{(\alpha t)^\beta} - e^{(\gamma t)^\lambda}}} \int_t^\infty e^{-e^{(\alpha x)^\beta} - e^{(\gamma x)^\lambda}} dx.$$

The denominator simplifies to $e^{e^{(\alpha t)^\beta} + e^{(\gamma t)^\lambda}}$. For the numerator, we proceed as in the proof of Theorem 1. We expand $e^{-e^{(\alpha x)^\beta}}$ and change the order of integration and summation. Applying the substitution $u = e^{(\gamma x)^\lambda}$ for the inner integral, the lower limit becomes $u_t = e^{(\gamma t)^\lambda}$. The resulting integral is

$$\int_{u_t}^\infty \frac{(\ln u)^{\frac{\beta m + 1}{\lambda} - 1}}{u} e^{-u} du = G\left(\frac{\beta m + 1}{\lambda} - 1, e^{(\gamma t)^\lambda}\right).$$

Combining these results yields the final expression. This completes the proof. $\square$

A.6. Proof of Theorem 6: Order statistics

For a random sample $\{X_i\}_{i=1}^n$, from the AEP model, the PDF of the k -th order statistic $X_{k:n}$ is given by the standard formula:

$$f_{k:n}(x; \psi) = \frac{n!}{(k-1)!(n-k)!} [F(x; \psi)]^{k-1} f(x; \psi) [1 - F(x; \psi)]^{n-k}.$$

Expanding $[F(x; \psi)]^{k-1}$ using the binomial series gives:

$$[F(x; \psi)]^{k-1} = \sum_{j=0}^{k-1} \binom{k-1}{j} (-1)^j [1 - F(x; \psi)]^j.$$

Substituting this expansion into the order statistics formula yields:

$$f_{k:n}(x; \psi) = \frac{n!}{(n-k)!} \sum_{j=0}^{k-1} \frac{(-1)^j}{j!(k-j-1)!} f(x; \psi) [1 - F(x; \psi)]^{n+j-k}.$$

Substituting the expressions for $f(x; \psi)$ from Eq (2.5) and $1 - F(x; \psi) = S(x; \psi)$ from Eq (2.4):

$$\begin{aligned} f_{k:n}(x; \psi) &= \frac{n!}{(n-k)!} \sum_{j=0}^{k-1} \frac{(-1)^j}{j!(k-j-1)!} \left(\alpha \beta (\alpha x)^{\beta-1} e^{(\alpha x)^\beta} + \gamma \lambda (\gamma x)^{\lambda-1} e^{(\gamma x)^\lambda} \right) \\ &\quad \times e^{-(n+j+1-k)(e^{(\alpha x)^\beta} + e^{(\gamma x)^\lambda} - 2)}. \end{aligned}$$

Recognizing that $h(x; \psi) = \alpha \beta (\alpha x)^{\beta-1} e^{(\alpha x)^\beta} + \gamma \lambda (\gamma x)^{\lambda-1} e^{(\gamma x)^\lambda}$ and $H(x; \psi) = e^{(\alpha x)^\beta} + e^{(\gamma x)^\lambda} - 2$, we arrive at the final expression. This completes the proof.

B. Software codes

The Bayesian simulation study presented in Section 4 was implemented in R using the `rstan` package and the HMC–NUTS algorithm, with 2000 iterations, 1000 warm-up iterations, and four chains. In contrast, the three real-data applications in Section 5 were carried out in Wolfram Mathematica using 5000 iterations, 2000 warm-up iterations, and four chains. In both implementations, the target acceptance probability was set to 0.9 and the maximum tree depth to 12. For the competing models, numerical starting values were selected within the admissible parameter space according to the model-specific constraints. In the additional simulation scenarios, a data-driven initialization strategy based on $1/\bar{x}$ and a coarse grid over the shape parameters was used instead of the true parameter values. Representative commands are provided below to improve the transparency and reproducibility of the numerical analyses.

```
(* Wolfram Mathematica: real-data applications *)
ProbabilityDistribution[pdf,x,{xmin,xmax}]
LogLikelihood[dist,data]
(* Maximum-likelihood estimation *)
FindMaximum[
{loglik,constraints},
{{x,x0},{y,y0},...},
Method -> "PrincipalAxis"]
NMaximize[
{loglik,constraints},
{x,y,...},
MaxIterations -> 1000]
(* Numerical and goodness-of-fit calculations *)
FindRoot[{eqn1,eqn2,...},{x,x0},{y,y0},...]
EmpiricalDistribution[data]
HistogramDistribution[data]
H = DistributionFitTest[
data,dist,"HypothesisTestData"]
H["TestDataTable"]
H["AndersonDarling"]
H["CramerVonMises"]
(* Mathematica HMC-NUTS: real-data Bayesian analysis *)
gradLogPost[theta_List]
leapfrog[theta_List,r_List,eps_]
FindReasonableEps[theta_List]
BuildTree[theta_List,r_List,u_,v_,j_,eps_]
runMCMCChain[
startPointLog_Association,totalSamples_Integer]
# R/rstan: Bayesian simulation study
library(stats)
library(rstan)
```

```

options(mc.cores = parallel::detectCores())
# Data-driven initialization for the additional scenarios
rough_scale <- 1/mean(data)
grid <- expand.grid(
b0 = c(0.5,1,2),
l0 = c(0.5,1,2,4))
# Numerical MLE
optim(
par=init_par,
fn=neg_log_lik_aep,
data=data,
method="L-BFGS-B")
# HMC-NUTS
sampling(
stan_model,
data=stan_data,
iter=2000,
warmup=1000,
chains=4,
control=list(
adapt_delta=0.9,
max_treedepth=12))
# Convergence assessment
summary(fit)$summary
all(summary(fit)$summary[, "Rhat"] < 1.1)

```



AIMS Press

© 2026 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0>)