



Research article

The accelerated failure time regression model under the extended-exponential distribution with survival analysis

Veronica Kariuki^{1,*}, Anthony Wanjoya², Oscar Ngesa³, Mahmoud M. Mansour^{4,5}, Enayat M. Abd Elrazik^{4,5} and Ahmed Z. Afify⁵

¹ Department of Mathematics, Pan African University Institute for Basic Sciences, Technology and Innovation, Nairobi 00200, Kenya

² Department of Statistics and Actuarial Sciences, Jomo Kenyatta University of Agriculture and Technology, Nairobi 00200, Kenya

³ Department of Mathematics, Statistics and Physical Science, Taita Taveta University, Voi 80300, Kenya

⁴ Department of Management Information Systems, Yanbu, Taibah University, Yanbu 46422, Saudi Arabia

⁵ Department of Statistics, Mathematics, and Insurance, Benha University, Benha 13511, Egypt

* **Correspondence:** Email: veronica.kariuki@students.jkuat.ac.ke; Tel: +254729719495.

Abstract: In this paper, we propose a parametric accelerated failure time (AFT) hazard-based regression model with the extended alpha-power exponential (EAPE) baseline distribution. The proposed model is called the extended alpha-power exponential-AFT (EAPE-AFT) regression model. We show that the EAPE distribution is closed under the AFT model. The parameters of the proposed EAPE-AFT model have been estimated by using the method of maximum likelihood estimation. An extensive simulation study was also conducted to examine the performance of the estimates under several scenarios based on the shapes of the baseline hazard function. Finally, real-life censored survival data has been used to illustrate the applicability of the proposed model.

Keywords: censored survival data; exponential distribution; survival regression model; hazard rate; maximum likelihood estimation; simulation

Mathematics Subject Classification: 60E05, 62E99

1. Introduction

Historically, the exponential distribution has been widely applied to model lifetime data due to its analytical simplicity and the availability of its simple statistical methods such as the existence of

its closed-form solutions of the cumulative distribution function (CDF) [1]. Notably, the exponential distribution is the only known continuous distribution characterized by a constant hazard (failure) rate as well as a memoryless property [2]. Due to the limitations of the hazard rate function (HRF) of the exponential distribution, its applicability to modeling survival data is somewhat limited, as the available data exhibit varying hazard shapes. As a result, in recent years, the Weibull, the log-normal, the log-logistic, and the gamma distributions have become the most popular distributions applied in survival analysis, particularly for time-to-event data analysis. The popularity of the wide range of applications has been motivated by the nature of their HRFs which range from monotonic to non-monotonic shapes; see, e.g., [1, 3, 4].

Lately, researchers have become extremely interested in extending distributions to realize more flexible distributions that can accommodate different hazard shapes; see e.g., [5–7]. Due to the availability of various methods of distribution modification, the exponential distribution has been modified by several researchers to achieve a better fit for real data. Some of these extensions include the exponentiated exponential [8] as an alternative to the gamma distribution and the generalized exponential distribution [9], which outperformed the Weibull and gamma distributions. As well as the odd exponentiated half-logistic exponential distribution [10] whose hazard rate can assume increasing, decreasing, or bathtub shapes, The exponentiated exponential [11], beta-exponential [12], generalized gull alpha-power exponential [13], and exponentiated generalized alpha-power exponential [14], among others.

Furthermore, many researchers have been focused on finding the association between the survival time or HRF and the covariates for the analysis of uncensored time-to-event data. From this, this relationship can be quantified and these covariates can be employed in fields such as medicine for disease management and prevention see [15–18]. These covariates can be classified into either time-dependence covariates or covariates that do not change with time. Thus, they accommodate sophisticated lifetime models that are acceptable in various industrial and medical settings. For instance, in a medical setup, gender and the type of treatment can be classified as time-independent covariates while calorie intake, age, and the level of dosage are all examples of covariates that may depend and even change with time [19].

This idea of finding and quantifying the association between the survival time and the covariates has led to the development of hazard-based regression models, with the main aim of estimating the regression coefficients for the corresponding covariate components. Moreover, the extended distributions have actively been used as baseline distributions in the development of hazard-based regression models to realize more accurate survival predictions. For instance, [20] extended the Weibull distribution by using the generalized Topp-Leone generator and used it as a baseline distribution in the formulation of an accelerated failure time (AFT) model. The log-logistic distribution has been generalized by Muse et al. [7] who developed a hazard-based regression model with the extended model as the baseline distribution. A multi-parametric hazard-based regression model called the extended exponential-Weibull AFT model has been proposed in [21]. This model performed effectively modeled right-censored time-to-event data.

In recent years, a tremendous shift from non-parametric to parametric regression models for survival analysis has resulted in the provision of attractive and flexible methods for the analysis of time-to-event data from different disciplines such as population health, medicine, and disease modeling [22]. There are several parametric regression models among which, the proportional hazard (PH) model and the

AFT model are the most commonly used in the analysis of censored time-to-event data; see [7, 23]. Others include the proportional odds model and the accelerated hazard model.

The AFT models are a type of statistical model that is used to analyze time-to-event data [24, 25]. These models assume that the time to an event (such as death, failure, or disease onset) follows a particular probability distribution and that certain factors (such as age, sex, and treatment group) may influence the rate at which the event occurs [26]. When the distribution of the model's error component is well known, the AFT model is said to be parametric; otherwise, it is called a semi-parametric model; see [27, 28]. The parametric AFT models have different advantages over the semi-parametric AFT models in survival analysis. Some of the significant advantages are as follows:

- (1) **Efficiency:** Parametric AFT models estimate fewer parameters than semi-parametric models. This frequently results in more efficient calculations, particularly when the underlying distributional assumptions are roughly true. Parametric models can provide more accurate estimates of the survival distribution since are required for estimation [29].
- (2) **Flexibility:** Parametric AFT models provide additional flexibility in the simulation of the structure of the survival distribution. Researchers can capture a broader range of survival patterns and perhaps increase model fit by specifying a parametric form for the survival function (SF) (such as Weibull, exponential, log-normal, etc.) as opposed to semi-parametric models, which assume fewer distributional features [29].
- (3) **Interpretability:** The parameters of parametric AFT models are frequently interpreted in terms of the underlying survival timescale. For example, in the Weibull distribution, the shape parameter controls whether the hazard rises, falls, or stays constant over time. This interpretability is useful for gaining insight into the elements that determine survival [30].

Furthermore, parametric AFT models, such as those using proportional hazard metrics, allow variables to accurately predict survival outcomes [28]. Parametric AFT models are used to estimate the effects of these factors on the event time and to make predictions about future events. They are commonly used in fields such as biostatistics, reliability engineering, and survival analysis.

The AFT model has been considered to be more applicable and realistic than the PH model. Additionally, more distributions together with their generalizations have been proven to be closed under the AFT but not under the PH model [24]. Furthermore, according to [23, 31], the covariates of the AFT model directly influence the time to event while those of the PH model only affect the HRF. Also, the constant acceleration assumption made by both the PH and the AFT models is open for relaxation in the AFT allowing for time dependency related to modeling non-PH systems [32].

The AFT model is said to be a parametric model if the error component in the model follows a well-known distribution. The model considers the regression of survival time on one or more covariates; hence, the model is mainly applied to elucidate the effects of the covariates that either decelerate or accelerate the survival time; see e.g., [20, 33].

Based on the above discussion, we propose a fully parametric AFT regression model under the baseline extended alpha-power exponential (EAPE) distribution. The proposed regression model aims to expand the scope of the exponential distribution in the analysis of time-to-event data and also to introduce an alternative model that can compete with the most commonly used classical Weibull, log-logistic, and gamma regression models. The primary motivation for this research was to broaden the application of the exponential distribution in the analysis of survival data. The proposed AFT regression model is quite versatile and can be used for a variety of survival analysis applications.

The proposed EAPE-AFT model is useful in survival analysis because of its adaptability, interpretability, and efficiency. The model explicitly describes the survival time distribution, allowing for a broader range of hazard functions and more flexible modeling of covariate effects. This direct parameterization allows for easy reading of the model's parameters, resulting in a better understanding of how covariates influence survival time. Overall, the parametric EAPE-AFT model developed in this study is a useful and accessible tool for studying survival data, providing insights that may be used to make decisions in a variety of situations.

The rest of this article is organised as follows: Section 2 gives the formulation of the proposed model. The maximum likelihood (ML) estimation methodology for the parameters of the proposed model is described in Section 3. In Section 4, a simulation study to assess the performance of the estimates is given. Section 5 gives a practical illustration of the model through the use of real data while a discussion is given in Section 6. Finally, conclusions and recommendations are given in Section 7.

2. Model formulation

AFT models are typically formulated by using a parametric approach. The basic idea is to model the logarithm of the event time as a linear function of the predictor variables, with the assumption that the error term follows a probability distribution. The most commonly used probability distributions in AFT models are the Weibull and log-normal distributions.

For instance, the Weibull AFT model can be formulated as follows:

$$\log(T) = \mathbf{X}\boldsymbol{\beta} + e,$$

where T is the event time, \mathbf{X} is a matrix of the predictor variables, $\boldsymbol{\beta}$ is a vector comprising the coefficients to be estimated, and e is the error term, which follows an extreme value distribution if the survival times follow a Weibull distribution.

Once the model is formulated, it can be fit to the data by using ML estimation or other methods. The coefficients estimated from the model can be used to make predictions about future event times and assess the effects of predictor variables on the event time.

2.1. The EAPE distribution

In this study, we consider the EAPE distribution [34] as a baseline model. The CDF of the EAPE distribution can be expressed as follows

$$F(x) = \left(\frac{\alpha^{1-e^{-\lambda x}} - 1}{\alpha - 1} \right)^b, \quad \alpha, b, \lambda > 0, \alpha \neq 1, \quad (2.1)$$

where α and $b > 0$ are shape parameters and λ is a scale parameter. By manipulating these parameters, one can customize the EAPE distribution to accurately represent diverse patterns and characteristics observed in lifetime data.

Figure 1 gives a visual display of the HRF shapes for the EAPE model indicating that the EAPE HRF can take four shapes including constant, unimodal, monotonically increasing, and monotonically decreasing. The HRF is a crucial term in survival analysis which is a field of statistics that examines the time until a specific event occurs, such as the death or failure of an item. The shape of the HRF

might reveal vital information about the survival of the item or population under consideration. In the case of an unimodal HRF, the hazard rate is initially low, then rises to a peak, and then falls again. This form is frequently observed in biological systems, such as the aging of living organisms. The inflection point or mode is the highest point in the HRF and it reflects the moment when the HRF is at its peak. In practice, this shape facilitates identification of the most susceptible phase for an item or population. For example, in a medical study, the inflection point may represent the age at which a person is most likely to encounter a specific health event, such as a heart attack or stroke. Understanding the shape of the HRF can assist researchers and practitioners in making sound decisions regarding preventative actions or treatments to reduce the risk of such incidents [35].

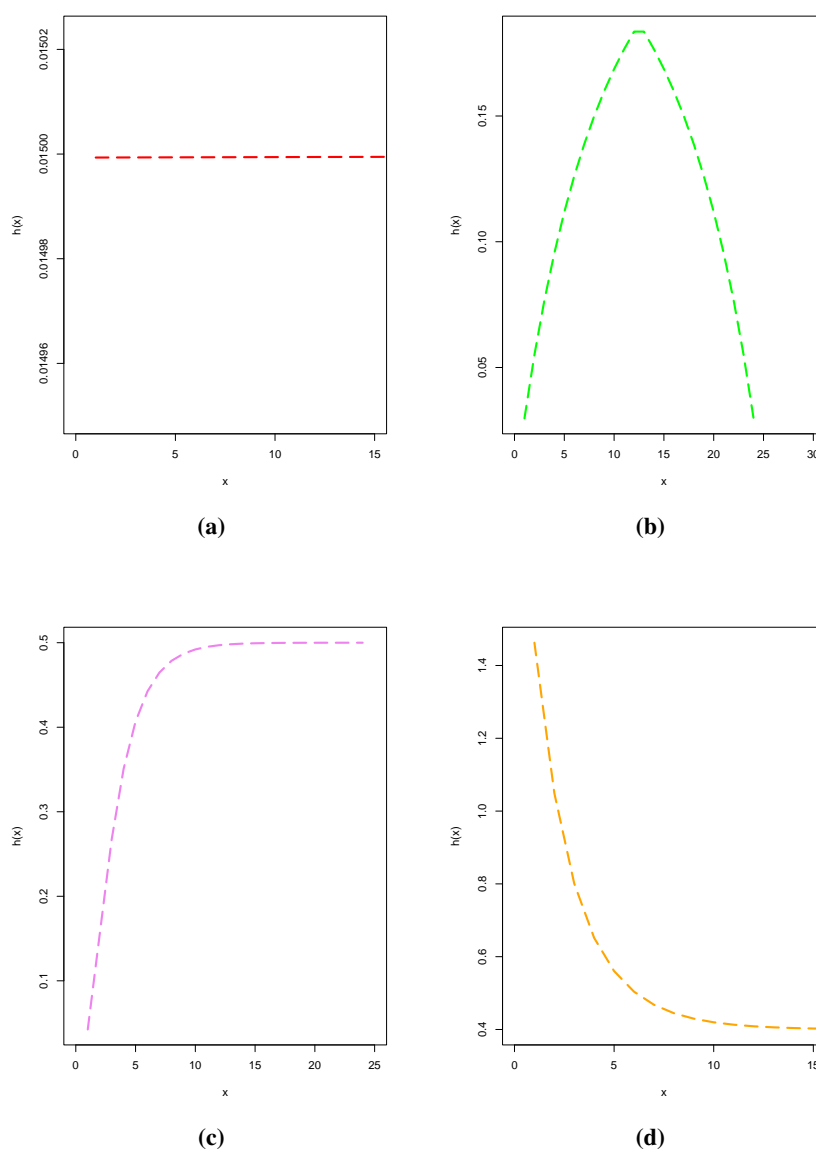


Figure 1. The HRF shapes of EAPE distribution: (a) constant; (b) unimodal; (c) increasing; (d) decreasing.

The probability density function (PDF) of the EAPE model reduces to

$$f(x) = \frac{e^{-\lambda x}}{(\alpha - 1)^b} (\log(\alpha)) \alpha^{1-e^{-\lambda x}} (\alpha^{1-e^{-\lambda x}} - 1)^{b-1}. \quad (2.2)$$

The S of the EAPE distribution is given by

$$S(x) = \frac{(\alpha - 1)^b - (\alpha^{1-e^{-\lambda x}} - 1)^b}{(\alpha - 1)^b}. \quad (2.3)$$

Hence, the HRF and the cumulative-HRF (CHRF) have the following respective forms

$$h(x) = \frac{b\lambda e^{-\lambda x} (\log(\alpha)) \alpha^{1-e^{-\lambda x}} (\alpha^{1-e^{-\lambda x}} - 1)^{b-1}}{(\alpha - 1)^b - (\alpha^{1-e^{-\lambda x}} - 1)^b} \quad (2.4)$$

and

$$H(x) = -\log \left[\frac{(\alpha - 1)^b - (\alpha^{1-e^{-\lambda x}} - 1)^b}{(\alpha - 1)^b} \right]. \quad (2.5)$$

2.2. The AFT model and its assumptions

If the random variable T is the event time, \mathbf{x} is a vector of covariates (also known as regressor variables), $\boldsymbol{\beta}$ is a vector of regressor coefficients, and $\psi(\mathbf{x}'\boldsymbol{\beta})$ is a link function for the regressor variables, then we have the following assumptions:

- (i) $\psi(0) = 1$;
- (ii) $\psi(\mathbf{x}'\boldsymbol{\beta}) > 0$;
- (iii) $\psi(\mathbf{x}'\boldsymbol{\beta})$ is a one-to-one monotonic function.

Then, the HRF of the AFT model is described as follows:

$$h(t) = h_0(t \psi(\mathbf{x}'\boldsymbol{\beta})) \psi(\mathbf{x}'\boldsymbol{\beta}). \quad (2.6)$$

Using the assumption that

$$\psi(\mathbf{x}'\boldsymbol{\beta}) = e^{\mathbf{x}'\boldsymbol{\beta}}$$

the HRF of the AFT model is given by

$$h(t) = h_0(t e^{\mathbf{x}'\boldsymbol{\beta}}) e^{\mathbf{x}'\boldsymbol{\beta}}. \quad (2.7)$$

The corresponding SF reduces to

$$S(t) = S_0(t e^{\mathbf{x}'\boldsymbol{\beta}}). \quad (2.8)$$

Hence, the CHRF can be derived as follows:

$$H(t) = H_0(t e^{\mathbf{x}'\boldsymbol{\beta}}). \quad (2.9)$$

Further, the PDF and CDF of the AFT model are respectively derived as follows:

$$f(t) = f_0(t e^{\mathbf{x}'\boldsymbol{\beta}}) e^{\mathbf{x}'\boldsymbol{\beta}} \quad (2.10)$$

and

$$F(t) = 1 - S(t) = 1 - S_0(t e^{\mathbf{x}'\boldsymbol{\beta}}). \quad (2.11)$$

2.3. The proposed EAPE-AFT model

Considering the EAPE distribution and the AFT models, the HRF and the SF of the proposed EAPE-AFT model respectively have the following forms

$$h(t, \Theta, \beta) = \frac{b\lambda(\log(\alpha))e^{x'\beta}\alpha^{1-e^{-\lambda t e^{x'\beta}}}\left(\alpha^{1-e^{-\lambda t e^{x'\beta}}}-1\right)^{b-1}}{(\alpha-1)^b-\left(\alpha^{1-e^{-\lambda t e^{x'\beta}}}-1\right)^b} \quad (2.12)$$

and

$$S(t; \Theta, \beta) = \frac{(\alpha-1)^b-\left(\alpha^{1-e^{-\lambda t e^{x'\beta}}}-1\right)^b}{(\alpha-1)^b}. \quad (2.13)$$

Clearly, by taking

$$\lambda e^{x'\beta} = \lambda^*,$$

the HRF and SF in Eqs (2.12) and (2.13) can respectively be rewritten as

$$h(t, \Theta, \beta) = \frac{b\lambda^*(\log(\alpha))\alpha^{1-e^{-\lambda^*t}}\left(\alpha^{1-e^{-\lambda^*t}}-1\right)^{b-1}}{(\alpha-1)^b-\left(\alpha^{1-e^{-\lambda^*t}}-1\right)^b} \quad (2.14)$$

and

$$S(t; \Theta, \beta) = \frac{(\alpha-1)^b-\left(\alpha^{1-e^{-\lambda^*t}}-1\right)^b}{(\alpha-1)^b}. \quad (2.15)$$

Equations (2.12) and (2.13) represent the HRF and SF of the EAPE distribution, respectively. This means that the EAPE distribution is closed under the AFT model.

The corresponding CHRF, PDF, and CDF of the EAPE-AFT model are given, respectively, as follows:

$$H(t; \Theta, \beta) = b \log(\alpha - 1) - \log \left[(\alpha - 1)^b - \left(\alpha^{1-e^{-\lambda t e^{x'\beta}}} - 1 \right)^b \right], \quad (2.16)$$

$$f(t, \Theta, \beta) = \frac{b\lambda e^{x'\beta}}{(\alpha-1)^b}(\log(\alpha))\alpha^{1-e^{-\lambda t e^{x'\beta}}}\left(\alpha^{1-e^{-\lambda t e^{x'\beta}}}-1\right)^{b-1} \quad (2.17)$$

and

$$F(t, \Theta, \beta) = \left(\frac{\alpha^{1-e^{-\lambda t e^{x'\beta}}}-1}{\alpha-1} \right)^b. \quad (2.18)$$

3. The ML estimation

In this section, the parameters of the EAPE-AFT regression model are estimated by using the ML estimation approach under a right-censoring scheme. The ML method is known to produce unique and asymptotically efficient estimators [36]. Other properties include asymptotic bias, asymptotic normality, and consistency.

Consider the lifetimes, T_1, T_2, \dots, T_n , of n individuals. If the data are subject to right censoring and the component $C_j > 0$ is the possible censoring time for the j individual, then

$$t_j = \min(T_j, C_j).$$

Further, let

$$\begin{cases} \sigma_j = I(T_j \leq C_j) = 1, & T_j \leq C_j, \\ \sigma_j = 0, & \text{otherwise.} \end{cases}$$

Let $(t_j, \sigma_j$ and $\mathbf{x}_j)$ be a set of right-censored sample data with

$$j = 1, 2, \dots, n, \quad \mathbf{x}_j = (x_1, x_2, \dots, x_n)^T$$

be an $n \times 1$ vector of regressor variables for the j individual depending on the value of σ_j , and t_j be the censoring time for the j individual. Furthermore, the assumption of non-informative censoring, which means that the survival time distribution provides no information on the censoring time distribution and vice versa, is considered. More details on non-informative censoring are explored in [1, 37]. Under this assumption, σ_j and t_j are considered to be random variables with a joint probability distribution function given by

$$f(t_j)^{\sigma_j} S(t_j)^{1-\sigma_j}. \quad (3.1)$$

Therefore, if the event of interest occurs at time t_j , the j individual contributes $f(t_j)$ to the likelihood function and $S(t_j)$ if censored at time t_j . Hence, we obtain the likelihood function by incorporating the information obtained from the censored observations, as follows:

$$L(\Theta) = \prod_{j=1}^n (f(t_j))^{\sigma_j} (S(t_j))^{1-\sigma_j},$$

where Θ is the baseline vector of parameters. In our case, $\Theta = (\alpha, \lambda, b, \beta)$.

We have that

$$f(t) = h(t)S(t);$$

thus, the likelihood function takes the following form:

$$L(\Theta) = \prod_{j=1}^n [h(t_j)S(t_j)]^{\sigma_j} (S(t_j))^{1-\sigma_j} = \prod_{j=1}^n [h(t_j)]^{\sigma_j} S(t_j).$$

Hence, the log-likelihood function reduces to

$$\ell(\Theta) = \sum_{j=1}^n \sigma_j \log h(t_j) + \sum_{j=1}^n \log S(t_j). \quad (3.2)$$

Therefore, the log-likelihood of the EAPE-AFT regression model becomes

$$\ell(\Theta) = \sum_{j=1}^n \sigma_j \log [h_0(t_j e^{\mathbf{x}'_j \beta}) e^{\mathbf{x}'_j \beta}] + \sum_{j=1}^n \log [S_0(t_j e^{\mathbf{x}'_j \beta})].$$

Hence

$$\begin{aligned} \ell(\Theta) = & \sum_{j=1}^n \sigma_j \log e^{x_j \beta} + \sum_{j=1}^n \sigma_j \log b + \sum_{j=1}^n \sigma_j \log \lambda - \sum_{j=1}^n \sigma_j \lambda t_j e^{x_j \beta} \\ & + \sum_{j=1}^n \sigma_j \omega_j \log(\alpha) + (b-1) \sum_{j=1}^n \sigma_j \log(\alpha^{\omega_j} - 1) \\ & - \sum_{j=1}^n \sigma_j \log [(\alpha-1)^b - (\alpha^{\omega_j} - 1)^b] + \sum_{j=1}^n \log [(\alpha-1)^b - (\alpha^{\omega_j} - 1)^b] - n b \log(\alpha-1), \end{aligned} \quad (3.3)$$

where

$$\omega_j = 1 - e^{-\lambda t_j e^{x_j \beta}}.$$

The ML estimates can be obtained directly by optimizing the log-likelihood function or by solving a system of equations that has been derived by equating the first partial derivatives of the log-likelihood function with respect to the parameters to zero. In this study, we employed the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method in the optimization of the log-likelihood function given in Eq (3.3) to obtain the ML estimators of the EAPE-AFT parameters. The BFGS algorithm was independently introduced by the authors of [38–41]. The BFGS algorithm is one of the most efficient algorithms for solving unconstrained optimization problems. It is an iterative technique that begins with an initial guess value Θ_0 and an initial Hessian matrix H_0 to provide a solution to a given function. The following steps are followed such that for $k = (0, 1, \dots)$ an approximation point Θ_k and an $m \times m$ matrix H_k are obtained on the k iteration.

Given parameters: $\Theta_0, H_0 > 0$ and unconstrained optimization problem $\ell(\Theta)$.

(1) First, the quasi-Newton direction is obtained as follows:

$$\Delta \Theta_k = -H_k^{-1} \nabla \ell(\Theta_k).$$

(2) The step length d_k is obtained such that it meets certain line search conditions.

(3) The next iteration is obtained as follows:

$$\Theta_{k+1} = \Theta_k + d_k \Delta \Theta_k.$$

(4) An important feature of the algorithm is the choice of H_k . That is, it must be positive definite and satisfy the quasi-Newton formula given by

$$H_{k+1} \alpha_k = \gamma_k,$$

where

$$\alpha_k = d_k \Delta \Theta_k \quad \text{and} \quad \gamma_k = \nabla \ell(\Theta_k + \alpha_k) - \nabla \ell(\Theta_k).$$

(5) Finally, the matrices denoted by H_k are updated in accordance with the following BFGS formula

$$H_{k+1} = H_k - \frac{H_k \alpha_k \alpha_k^T H_k}{\alpha_k^T H_k \alpha_k} + \frac{\gamma_k \gamma_k^T}{\alpha_k^T \gamma_k},$$

where $\Theta = (\alpha, \lambda, b, \beta)$ is a vector of parameters of the EAPE distribution.

4. Simulation study

In this section, we show the ability of the proposed model to make inferences by presenting simulation results. Specifically, we demonstrate how the model can estimate parameters accurately by using the absolute bias (AB), mean square error (MSE), and relative bias (RB) measures. We check the ability of the model to recover the baseline HRF shapes. Moreover, we obtain the Akaike information criteria (AIC) to choose models that accurately describe the underlying HR shape. We also investigate how the model's inferential capabilities are affected by different proportions of censoring.

The AFT regression framework given in Eq (2.6) was assumed for data generation and simulation design. We generated samples of sizes $n = 2000$ and $n = 5000$. The covariate vector,

$$\mathbf{x} = (x_1, x_2, x_3, x_4)',$$

was considered in the simulation study where x_1 and x_2 are binary variables derived from a Bernoulli(0.5) distribution whereas x_3 and x_4 are continuous variables derived from a standard Gaussian distribution. The vector corresponds to the following coefficients of the AFT regression

$$\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_3, \beta_4)'$$

whose initial values were chosen to be $(-2.5, 0.5, -0.75, 0.5, -0.75)$.

In simulating lifetime censored data from the AFT framework, we used the technique of the inverse transform as explained in [19]. Further, because the main aim of the study was to test the robustness of the proposed model when applied to model different types of time-to-event data, we considered the inverse transform of the exponentiated Weibull (EW) distribution under the assumption of the presence of a regression intercept and covariate effects [3, 5, 21, 42]. Also, the assumption of non-informative censoring has been incorporated for the generation of censored data.

If a random variable T comes from the EW distribution with parameters α , b and λ , the PDF of T is given by

$$f(t; \alpha, b, \lambda) = \alpha b \lambda (bt)^{\alpha-1} \left\{ 1 - e^{-(bt)^\alpha} \right\}^{\lambda-1} e^{-(bt)^\alpha}. \quad (4.1)$$

This distribution accommodates all of the most basic HRF shapes which include constant, monotonic (increasing and decreasing), and non-monotonic (unimodal and bathtub), and it is also closed under the AFT regression models. It follows that the CHRF for the EW distribution is defined as follows:

$$H_0(t; \alpha, b, \lambda) = -\log \left(1 - \left\{ 1 - e^{-(bt)^\alpha} \right\}^\lambda \right). \quad (4.2)$$

The inverse CHRF is given by

$$H_0^{-1}(p; \alpha, b, \lambda) = -\frac{\log \left[(e^{-p} - 1)^{1/\lambda} - 1 \right]^{\frac{1}{\alpha}}}{b}. \quad (4.3)$$

4.1. Simulated scenarios

This study was conducted for three simulation scenarios based on monotonically increasing, monotonically decreasing, and unimodal hazard shapes to assess how well the proposed EAPE-AFT

model performs in comparison with some selected competing distributions, i.e., the log-logistic-AFT (LL-AFT) and Weibull-AFT (W-AFT). The simulations were also conducted to explore the effects of the baseline HRF on the inferential properties of the hazard-based regression models.

First scenario: monotonically increasing HRF

In this scenario, we generated lifetime data from the EW distribution with $(\alpha = 1.5, b = 1.4, \lambda = 1.1)$ and two censoring percentages, i.e., $14.45\% \approx 15\%$ and $22.15\% \approx 22\%$ which were determined from the censoring times derived from an exponential distribution by using the rate parameter $\theta = 0.26$ and $\theta = 0.15$, respectively.

Second scenario: monotonically decreasing HRF

In this scenario, we generated lifetime data from the EW distribution with $\alpha = 1.2, b = 1.2, \lambda = 0.8$ and two censoring percentages, i.e., 15% and 22% which were determined from the censoring times derived from an exponential distribution by using the rate parameter $\theta = 0.56$ and $\theta = 0.31$, respectively.

Third scenario: unimodal HRF

In this scenario, we generated lifetime data from the EW distribution with $\alpha = 0.9, b = 3.5, \lambda = 1.2$ and two censoring percentages, i.e., $14.75\% \approx 15\%$ and 22% which were determined from the censoring times derived from an exponential distribution using the rate parameter $\theta = 0.27$ and $\theta = 0.19$, respectively.

Using the different parameter combinations, the plots of hazard shapes of the EW distribution were constructed as reported in Figure 2 for the cases of monotonically increasing, monotonically decreasing, and unimodal. Figure 3 gives the total time on test (TTT) plots for the lifetime data generated for each of the three scenarios.

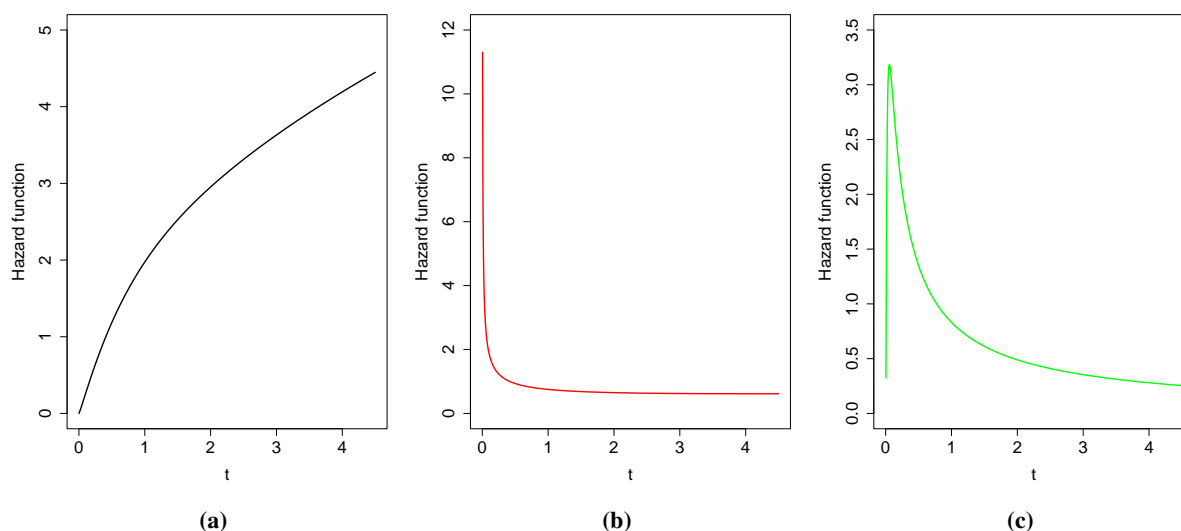


Figure 2. The EW hazard shapes for the considered simulation scenarios: (a) increasing; (b) decreasing; (c) unimodal.

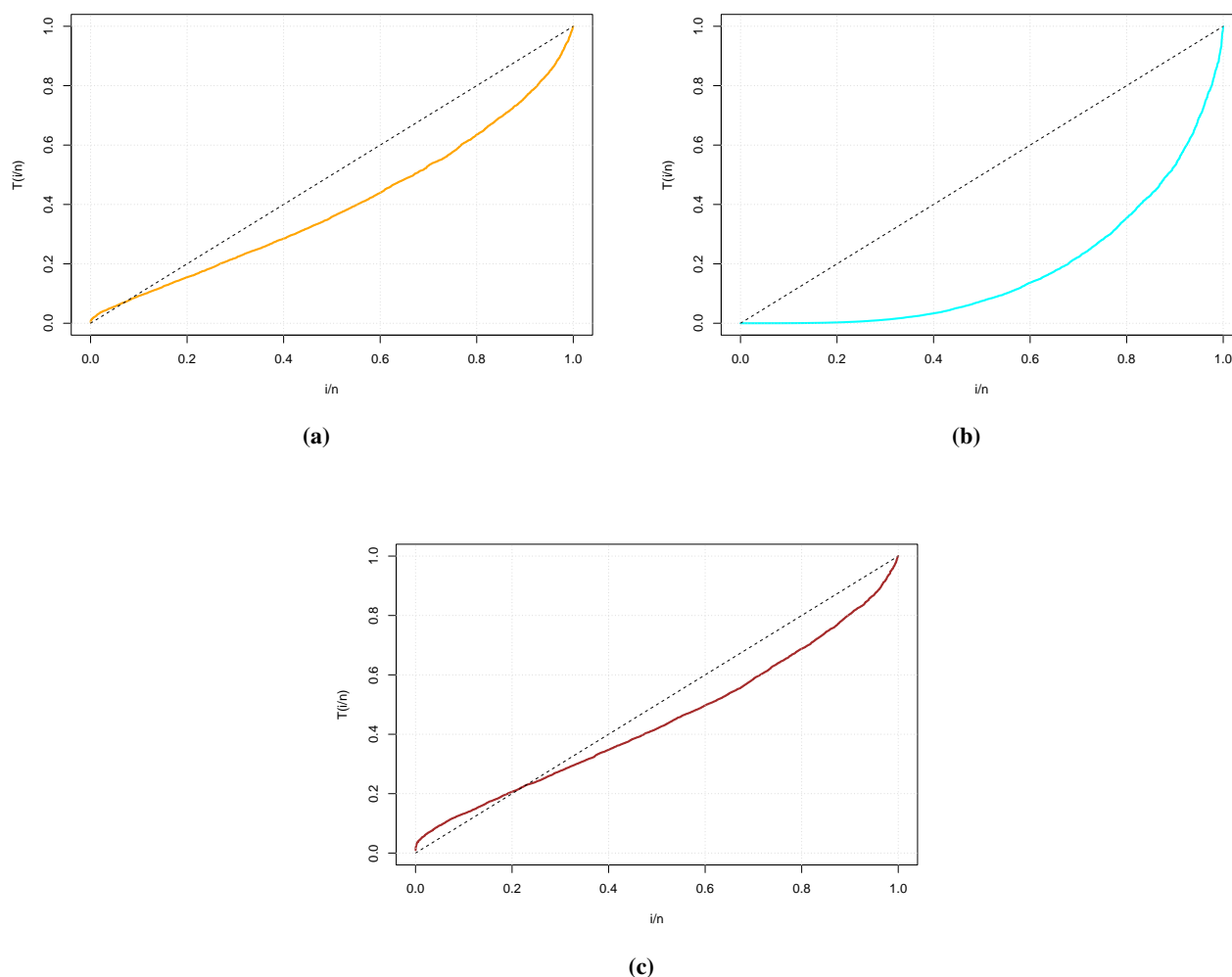


Figure 3. The TTT plots for the generated data for the three scenarios: (a) increasing; (b) decreasing; (c) unimodal.

4.2. Analysis of simulated data

To evaluate the predictive and analytical capabilities of the proposed model in each scenario, the EAPE-AFT model is fit to the lifetime data that were generated based on the EW distribution as was the case for the W-AFT and LL-AFT models. Additionally, for all the fitted models, the AB, RB, and MSE were calculated to evaluate the stability of the regression coefficient estimators. Finally, the AIC was computed for model comparison.

Given that our main objective was to evaluate the accuracy of the estimates (or stability of the estimators) of the covariate coefficients, rather than analyze the characteristics of the optimization process, we utilized the parameter values from the generating model to initiate the optimization procedure in every scenario. We conducted the analysis by using the R programming language, and

the optimization step was executed by using the “nlminb” function in the R software.

4.3. Simulation results

Simulation results for the first scenario are presented in Tables 1–4. Based on these results, it can be inferred that the proposed EAPE-AFT model performs better than the EW-AFT, the W-AFT, and the LL-AFT models, as it achieved the lowest AIC. Additionally, the AIC values for the EW-AFT, EAPE-AFT and W-AFT distributions can be seen as close to each other but they are superior to those for the LL-AFT model. When the censoring percentage was increased to 22% and the sample size was increased, the EAPE-AFT model consistently outperformed all of the considered models with the lowest AIC values. For the covariates, x_1, x_2, x_3, x_4 , and the intercept, all models indicated low values of the AB, MSE, and RB. Increasing the sample size from 2000 to 5000, the values of AB, MSE, and RB decreased indicating consistency of the estimates. Moreover, the AB, MSE, and RB values for the covariates of the EAPE-AFT, EW-AFT, and W-AFT models dominated those of the LL-AFT model. However, increasing the censoring percentage resulted in increased estimation error.

In Scenario 2 as presented in Tables 5–8, all competing models could account for the diminishing HRF shape, but the EAPE-AFT model was found to be the best in terms of the value of the AIC. The EAPE-AFT, EW-AFT, W-AFT, and LL-AFT had similar AIC.

The results for Scenario 3 are presented in Tables 9–12. These results show that the EAPE-AFT model outperformed all of the considered models with the lowest AIC values, and that the EW-AFT, and LL-AFT models had the lowest values for the AIC. On the other hand, the W-AFT model generates the least accurate estimates in terms of the AB, MSE, and RB.

Table 1. Simulation results for the first scenario with 15% censoring and $n = 2000$.

Model	Parameter	True value	MLE	AB	MSE	RB	
EAPE-AFT (AIC = 2572.532)	α	1.8	2.053	0.253	0.064	0.141	
	λ	1.1	0.089	1.011	1.022	0.919	
	b	2.2	0.779	1.421	2.018	0.646	
	β_0	-2	-3.041	1.041	1.083	-0.520	
	β_1	0.65	1.104	0.454	0.206	0.698	
	β_2	-0.7	-1.710	1.010	1.021	-1.443	
	β_3	0.65	0.914	0.264	0.070	0.406	
	β_4	-0.7	-1.663	0.963	0.927	-1.375	
	EW-AFT(True model) (AIC = 2574.291)	α	1.5	1.533	0.033	0.001	0.022
		λ	1.4	1.522	0.122	0.015	0.087
b		1.1	1.151	0.051	0.003	0.046	
β_0		-3	-3.037	0.037	0.001	-0.012	
β_1		0.45	0.449	0.001	0.000	0.002	
β_2		-0.65	-0.697	0.047	0.002	-0.072	
β_3		0.45	0.439	0.011	0.000	0.024	
β_4		-0.65	-0.652	0.002	0.000	-0.003	
LL-AFT (AIC = 2591.397)		α	3	3.261	0.261	0.068	0.087
		b	1.5	1.837	0.337	0.114	0.225
	β_0	-3.5	-3.394	0.106	0.011	-0.030	
	β_1	0.75	0.912	0.162	0.026	0.216	
	β_2	-0.85	-0.873	0.023	0.001	-0.027	
	β_3	0.75	0.793	0.043	0.002	0.057	
	β_4	-0.85	-0.868	0.018	0.000	-0.021	
	W-AFT (AIC = 2576.459)	α	2.9	2.833	0.067	0.005	0.023
		b	1.8	1.900	0.100	0.010	0.056
		β_0	-1.5	-1.593	0.093	0.009	-0.062
β_1		0.35	0.331	0.019	0.000	0.054	
β_2		-1.05	-1.047	0.003	0.000	-0.003	
β_3		0.35	0.395	0.045	0.002	0.127	
β_4		-1.05	-1.069	0.019	0.000	-0.018	

Table 2. Simulation results for the first scenario with 15% censoring and $n = 5000$.

Model	Parameter	True value	MLE	AB	MSE	RB	
EAPE-AFT (AIC = 2513.041)	α	1.8	1.790	0.010	0.000	0.006	
	λ	1.1	1.106	0.006	0.000	0.005	
	b	2.2	2.279	0.079	0.006	0.036	
	β_0	-2	-2.189	0.189	0.036	-0.094	
	β_1	0.65	0.678	0.028	0.001	0.043	
	β_2	-0.7	-0.622	0.078	0.006	-0.111	
	β_3	0.65	0.639	0.011	0.000	0.017	
	β_4	-0.7	-0.711	0.011	0.000	-0.015	
	EW-AFT(True model) (AIC = 2513.31)	α	1.5	1.506	0.006	0.000	0.004
		λ	1.4	1.512	0.112	0.012	0.080
b		1.1	1.108	0.008	0.000	0.007	
β_0		-3	-3.051	0.051	0.003	-0.017	
β_1		0.45	0.458	0.008	0.000	0.018	
β_2		-0.65	-0.671	0.021	0.000	-0.033	
β_3		0.45	0.429	0.021	0.000	0.046	
β_4		-0.65	-0.648	0.002	0.000	-0.003	
LL-AFT (AIC = 2525.861)		α	3	3.295	0.295	0.087	0.098
		b	1.5	1.787	0.287	0.082	0.191
	β_0	-3.5	-4.017	0.517	0.267	-0.148	
	β_1	0.75	1.066	0.316	0.100	0.421	
	β_2	-0.85	-0.692	0.158	0.025	-0.186	
	β_3	0.75	1.246	0.496	0.246	0.661	
	β_4	-0.85	-1.007	0.157	0.025	-0.185	
	W-AFT (AIC = 2513.426)	α	2.9	2.449	0.451	0.203	0.156
		b	1.8	1.814	0.014	0.000	0.008
		β_0	-1.5	-1.605	0.105	0.011	-0.070
β_1		0.35	0.408	0.058	0.003	0.166	
β_2		-1.05	-1.172	0.122	0.015	-0.116	
β_3		0.35	0.316	0.034	0.001	0.097	
β_4		-1.05	-1.076	0.026	0.001	-0.024	

Table 3. Simulation results for the first scenario with 22% censoring and $n = 2000$.

Model	Parameter	True value	MLE	AB	MSE	RB	
EAPE-AFT (AIC = 6401.272)	α	1.8	1.721	0.079	0.006	0.044	
	λ	1.1	1.130	0.030	0.001	0.027	
	b	2.2	2.133	0.067	0.004	0.030	
	β_0	-2	-2.233	0.233	0.054	-0.116	
	β_1	0.65	0.694	0.044	0.002	0.068	
	β_2	-0.7	-0.702	0.002	0.000	-0.004	
	β_3	0.65	0.653	0.003	0.000	0.005	
	β_4	-0.7	-0.605	0.095	0.009	-0.136	
	EW-AFT(True model) (AIC = 6406.657)	α	1.5	1.531	0.031	0.001	0.021
		λ	1.4	1.294	0.106	0.011	0.076
b		1.1	1.125	0.025	0.001	0.022	
β_0		-3	-2.705	0.295	0.087	-0.098	
β_1		0.45	0.468	0.018	0.000	0.039	
β_2		-0.65	-0.643	0.007	0.000	-0.011	
β_3		0.45	0.636	0.186	0.035	0.413	
β_4		-0.65	-0.702	0.052	0.003	-0.080	
LL-AFT (AIC = 6417.564)		α	2.9	2.145	0.755	0.570	0.260
		b	1.8	1.265	0.535	0.286	0.297
	β_0	-1.5	-1.745	0.245	0.060	-0.163	
	β_1	0.35	0.686	0.336	0.113	0.960	
	β_2	-1.05	-1.012	0.038	0.001	-0.036	
	β_3	0.35	0.644	0.294	0.086	0.840	
	β_4	-1.05	-1.032	0.018	0.000	-0.017	
	W-AFT (AIC = 6407.219)	α	3	2.851	0.149	0.022	0.050
		b	1.5	1.208	0.292	0.085	0.195
		β_0	-3.5	-3.080	0.420	0.176	-0.120
β_1		0.75	0.674	0.076	0.006	0.101	
β_2		-0.85	-1.004	0.154	0.024	-0.181	
β_3		0.75	0.634	0.116	0.013	0.155	
β_4		-0.85	-1.018	0.168	0.028	-0.198	

Table 4. Simulation results for the first scenario with 22% censoring and $n = 5000$.

Model	Parameter	True value	MLE	AB	MSE	RB	
EAPE-AFT (AIC = 6142.081)	α	1.8	1.901	0.101	0.010	0.056	
	λ	1.1	1.132	0.032	0.001	0.029	
	b	2.2	2.503	0.303	0.092	0.138	
	β_0	-2	-2.016	0.016	0.000	-0.008	
	β_1	0.65	0.669	0.019	0.000	0.029	
	β_2	-0.7	-0.732	0.032	0.001	-0.045	
	β_3	0.65	0.708	0.058	0.003	0.089	
	β_4	-0.7	-0.800	0.100	0.010	-0.143	
	EW-AFT(True model) (AIC = 6145.041)	α	1.5	1.525	0.025	0.001	0.016
		λ	1.4	1.402	0.002	0.000	0.001
b		1.1	1.109	0.009	0.000	0.008	
β_0		-3	-2.962	0.038	0.001	-0.013	
β_1		0.45	0.460	0.010	0.000	0.021	
β_2		-0.65	-0.683	0.033	0.001	-0.051	
β_3		0.45	0.469	0.019	0.000	0.043	
β_4		-0.65	-0.699	0.049	0.002	-0.075	
LL-AFT (AIC = 6156.652)		α	2.9	2.164	0.736	0.542	0.254
		b	1.8	1.727	0.073	0.005	0.041
	β_0	-1.5	-1.764	0.264	0.070	-0.176	
	β_1	0.35	0.367	0.017	0.000	0.048	
	β_2	-1.05	-1.011	0.039	0.002	-0.037	
	β_3	0.35	0.370	0.020	0.000	0.058	
	β_4	-1.05	-1.099	0.049	0.002	-0.047	
	W-AFT (AIC = 6144.867)	α	3	2.853	0.147	0.022	0.049
		b	1.5	1.515	0.015	0.000	0.010
		β_0	-3.5	-3.106	0.394	0.155	-0.113
β_1		0.75	0.648	0.102	0.010	0.136	
β_2		-0.85	-0.899	0.049	0.002	-0.057	
β_3		0.75	0.776	0.026	0.001	0.035	
β_4		-0.85	-0.973	0.123	0.015	-0.145	

Table 5. Simulation results for the second scenario with 15% censoring and $n = 2000$.

Model	Parameter	True value	MLE	AB	MSE	RB	
EAPE-AFT (AIC = 4709.928)	α	1.3	1.409	0.109	0.012	0.084	
	λ	1.1	1.100	0.000	0.000	0.000	
	b	0.72	0.735	0.015	0.000	0.021	
	β_0	-2	-2.897	0.897	0.805	-0.449	
	β_1	0.65	0.775	0.125	0.016	0.192	
	β_2	-0.7	-0.632	0.068	0.005	-0.097	
	β_3	0.65	0.689	0.039	0.002	0.060	
	β_4	-0.7	-0.653	0.047	0.002	-0.067	
	EW-AFT(True model) (AIC = 4710.583)	α	1.5	1.500	0.000	0.000	0.000
		λ	1.1	1.173	0.073	0.005	0.067
b		0.85	0.853	0.003	0.000	0.004	
β_0		-3	-2.709	0.291	0.085	-0.097	
β_1		0.45	0.431	0.019	0.000	0.042	
β_2		-0.65	-0.630	0.020	0.000	-0.030	
β_3		0.45	0.429	0.021	0.000	0.047	
β_4		-0.65	-0.653	0.003	0.000	-0.005	
LL-AFT (AIC = 4712.974)		α	1.3	1.913	0.613	0.376	0.472
		b	1.5	1.599	0.099	0.010	0.066
	β_0	-3.5	-4.162	0.662	0.438	-0.189	
	β_1	0.75	1.119	0.369	0.136	0.492	
	β_2	-0.85	-0.632	0.218	0.048	-0.256	
	β_3	0.75	0.874	0.124	0.015	0.165	
	β_4	-0.85	-0.854	0.004	0.000	-0.004	
	W-AFT (AIC = 4710.68)	α	1.3	1.202	0.098	0.522	0.075
		b	1.8	1.760	0.040	0.105	0.022
		β_0	-1.5	-1.222	0.278	0.522	-0.185
β_1		0.35	0.321	0.029	0.548	0.083	
β_2		-1.05	-1.057	0.007	0.274	-0.007	
β_3		0.35	0.382	0.032	0.225	0.093	
β_4		-1.05	-1.051	0.001	0.208	-0.001	

Table 6. Simulation results for the second scenario with 15% censoring and $n = 5000$.

Model	Parameter	True value	MLE	AB	MSE	RB	
EAPE-AFT (AIC = 3974.238)	α	1.3	1.310	0.010	0.000	0.008	
	λ	1.1	1.112	0.012	0.000	0.011	
	b	0.72	0.753	0.033	0.001	0.046	
	β_0	-2	-2.209	0.209	0.044	-0.105	
	β_1	0.65	0.600	0.050	0.003	0.077	
	β_2	-0.7	-0.757	0.057	0.003	-0.082	
	β_3	0.65	0.640	0.010	0.000	0.015	
	β_4	-0.7	-0.755	0.055	0.003	-0.079	
	EW-AFT(True model) (AIC = 3974.291)	α	1.5	1.517	0.017	0.000	0.011
λ		1.1	1.133	0.033	0.001	0.030	
b		0.85	0.851	0.001	0.000	0.001	
β_0		-3	-3.084	0.084	0.007	-0.028	
β_1		0.45	0.449	0.001	0.000	0.003	
β_2		-0.65	-0.662	0.012	0.000	-0.019	
β_3		0.45	0.452	0.002	0.000	0.005	
β_4		-0.65	-0.658	0.008	0.000	-0.012	
LL-AFT (AIC = 3978.462)		α	1.3	1.394	0.094	0.009	0.072
	b	1.5	1.573	0.073	0.005	0.049	
	β_0	-3.5	-3.389	0.111	0.012	-0.032	
	β_1	0.75	0.740	0.010	0.000	0.013	
	β_2	-0.85	-0.866	0.016	0.000	-0.018	
	β_3	0.75	0.690	0.060	0.004	0.080	
	β_4	-0.85	-0.863	0.013	0.000	-0.015	
	W-AFT (AIC = 3975.069)	α	1.3	1.381	0.081	0.007	0.063
		b	1.8	1.953	0.153	0.023	0.085
β_0		-1.5	-1.340	0.160	0.026	-0.107	
β_1		0.35	0.388	0.038	0.001	0.109	
β_2		-1.05	-1.652	0.602	0.362	-0.573	
β_3		0.35	0.215	0.135	0.018	0.386	
β_4		-1.05	-1.065	0.015	0.000	-0.014	

Table 7. Simulation results for the second scenario with 22% censoring and $n = 2000$.

Model	Parameter	True value	MLE	AB	MSE	RB	
EAPE-AFT (AIC = 11701.29)	α	1.3	1.306	0.006	0.000	0.005	
	λ	1.1	1.115	0.015	0.000	0.013	
	b	0.72	0.757	0.037	0.001	0.052	
	β_0	-2	-2.090	0.090	0.008	-0.045	
	β_1	0.65	0.642	0.008	0.000	0.012	
	β_2	-0.7	-0.753	0.053	0.003	-0.075	
	β_3	0.65	0.648	0.002	0.000	0.003	
	β_4	-0.7	-0.745	0.045	0.002	-0.064	
	EW-AFT(True model) (AIC = 11714.72)	α	1.5	1.532	0.032	0.001	0.021
λ		1.1	1.109	0.009	0.000	0.008	
b		0.85	0.821	0.029	0.001	0.034	
β_0		-3	-3.068	0.068	0.005	-0.023	
β_1		0.45	0.410	0.040	0.002	0.089	
β_2		-0.65	-0.653	0.003	0.000	-0.005	
β_3		0.45	0.483	0.033	0.001	0.073	
β_4		-0.65	-0.652	0.002	0.000	-0.003	
LL-AFT (AIC = 11715.66)		α	1.3	1.265	0.035	0.001	0.027
	b	1.5	1.638	0.138	0.019	0.092	
	β_0	-3.5	-3.484	0.016	0.000	-0.005	
	β_1	0.75	0.805	0.055	0.003	0.073	
	β_2	-0.85	-0.830	0.020	0.000	-0.023	
	β_3	0.75	0.823	0.073	0.005	0.097	
	β_4	-0.85	-0.846	0.004	0.000	-0.005	
	W-AFT (AIC = 11717.28)	α	1.3	1.318	0.018	0.000	0.014
		b	1.8	1.836	0.036	0.001	0.020
β_0		-1.5	-1.498	0.002	0.000	-0.001	
β_1		0.35	0.401	0.051	0.003	0.145	
β_2		-1.05	-1.463	0.413	0.171	-0.393	
β_3		0.35	0.378	0.028	0.001	0.081	
β_4		-1.05	-0.163	0.887	0.786	-0.844	

Table 8. Simulation results for the second scenario with 22% censoring and $n = 5000$.

Model	Parameter	True value	MLE	AB	MSE	RB	
EAPE-AFT (AIC = 9495.63)	α	1.3	1.320	0.020	0.000	0.015	
	λ	1.1	1.107	0.007	0.000	0.007	
	b	0.72	0.756	0.036	0.001	0.050	
	β_0	-2	-2.066	0.066	0.004	-0.033	
	β_1	0.65	0.595	0.055	0.003	0.084	
	β_2	-0.7	-0.744	0.044	0.002	-0.062	
	β_3	0.65	0.594	0.056	0.003	0.086	
	β_4	-0.7	-0.837	0.137	0.019	-0.196	
	EW-AFT(True model) (AIC = 9504.849)	α	1.5	1.547	0.047	0.002	0.031
λ		1.1	1.103	0.003	0.000	0.003	
b		0.85	0.865	0.015	0.000	0.018	
β_0		-3	-3.052	0.052	0.003	-0.017	
β_1		0.45	0.440	0.010	0.000	0.022	
β_2		-0.65	-0.749	0.099	0.010	-0.152	
β_3		0.45	0.439	0.011	0.000	0.025	
β_4		-0.65	-0.649	0.001	0.000	-0.002	
LL-AFT (AIC = 9506.04)		α	1.3	1.371	0.071	0.005	0.054
	b	1.5	1.611	0.111	0.012	0.074	
	β_0	-3.5	-3.939	0.439	0.193	-0.125	
	β_1	0.75	0.899	0.149	0.022	0.199	
	β_2	-0.85	-1.048	0.198	0.039	-0.233	
	β_3	0.75	0.706	0.044	0.002	0.059	
	β_4	-0.85	-1.005	0.155	0.024	-0.182	
	W-AFT (AIC = 9518.466)	α	1.3	1.385	0.085	0.007	0.065
		b	1.8	1.940	0.140	0.020	0.078
β_0		-1.5	-1.605	0.105	0.011	-0.070	
β_1		0.35	0.300	0.050	0.003	0.143	
β_2		-1.05	-1.047	0.003	0.000	-0.002	
β_3		0.35	0.341	0.009	0.000	0.026	
β_4		-1.05	-1.048	0.002	0.000	-0.002	

Table 9. Simulation results for the third scenario with 15% censoring and $n = 2000$.

Model	Parameter	True value	MLE	AB	MSE	RB	
EAPE-AFT (AIC = 2572.532)	α	2.9	2.983	0.083	0.007	0.029	
	λ	1.5	1.379	0.121	0.015	0.080	
	b	0.2	0.457	0.257	0.066	1.283	
	β_0	-2	-2.077	0.077	0.006	-0.038	
	β_1	0.65	0.716	0.066	0.004	0.101	
	β_2	-0.7	-0.755	0.055	0.003	-0.079	
	β_3	0.65	0.671	0.021	0.000	0.032	
	β_4	-0.7	-0.656	0.044	0.002	-0.062	
	EW-AFT(True model) (AIC = 2574.291)	α	1.9	1.859	0.041	0.002	0.022
λ		2.5	2.511	0.011	0.000	0.004	
b		2.2	2.101	0.099	0.010	0.045	
β_0		-3	-3.056	0.056	0.003	-0.019	
β_1		0.45	0.470	0.020	0.000	0.044	
β_2		-0.65	-0.661	0.011	0.000	-0.016	
β_3		0.45	0.466	0.016	0.000	0.036	
β_4		-0.65	-0.783	0.133	0.018	-0.204	
LL-AFT (AIC = 2589.642)		α	1.9	1.731	0.169	0.029	0.089
	b	3.8	3.728	0.072	0.005	0.019	
	β_0	-1.5	-1.331	0.169	0.029	-0.113	
	β_1	0.35	0.271	0.079	0.006	0.226	
	β_2	-1.05	-1.065	0.015	0.000	-0.014	
	β_3	0.35	0.366	0.016	0.000	0.047	
	β_4	-1.05	-1.095	0.045	0.002	-0.043	
	W-AFT (AIC = 2591.397)	α	3.9	3.468	0.432	0.187	0.111
		b	0.5	0.411	0.089	0.008	0.178
β_0		-3.5	-2.928	0.572	0.327	-0.163	
β_1		0.75	0.697	0.053	0.003	0.071	
β_2		-0.85	-1.057	0.207	0.043	-0.244	
β_3		0.75	0.662	0.088	0.008	0.117	
β_4		-0.85	-1.078	0.228	0.052	-0.268	

Table 10. Simulation results for the third scenario with 15% censoring and $n = 5000$.

Model	Parameter	True value	MLE	AB	MSE	RB	
EAPE-AFT (AIC = 1871.013)	α	2.9	2.813	0.087	0.008	0.030	
	λ	1.5	1.611	0.111	0.012	0.074	
	b	0.2	0.083	0.117	0.014	0.584	
	β_0	-2	-2.039	0.039	0.001	-0.019	
	β_1	0.65	0.702	0.052	0.003	0.081	
	β_2	-0.7	-1.053	0.353	0.125	-0.505	
	β_3	0.65	0.764	0.114	0.013	0.175	
	β_4	-0.7	-1.054	0.354	0.125	-0.506	
	EW-AFT(True model) (AIC = 1873.013)	α	1.9	1.923	0.023	0.001	0.012
		λ	2.5	2.689	0.189	0.036	0.075
b		2.2	2.264	0.064	0.004	0.029	
β_0		-3	-3.172	0.172	0.030	-0.057	
β_1		0.45	0.370	0.080	0.006	0.178	
β_2		-0.65	-0.605	0.045	0.002	-0.069	
β_3		0.45	0.558	0.108	0.012	0.239	
β_4		-0.65	-0.649	0.001	0.000	-0.002	
LL-AFT (AIC = 1884.112)		α	1.9	1.770	0.130	0.017	0.068
		b	3.8	4.286	0.486	0.236	0.128
	β_0	-1.5	-1.370	0.130	0.017	-0.087	
	β_1	0.35	0.704	0.354	0.125	1.011	
	β_2	-1.05	-1.054	0.004	0.000	-0.004	
	β_3	0.35	0.764	0.414	0.171	1.183	
	β_4	-1.05	-1.054	0.004	0.000	-0.004	
	W-AFT (AIC = 1896.38)	α	3.9	3.476	0.424	0.180	0.109
		b	0.5	0.501	0.001	0.000	0.002
		β_0	-3.5	-4.995	1.495	2.235	-0.427
β_1		0.75	0.682	0.068	0.005	0.091	
β_2		-0.85	-1.031	0.181	0.033	-0.213	
β_3		0.75	0.733	0.017	0.000	0.023	
β_4		-0.85	-1.027	0.177	0.031	-0.208	

Table 11. Simulation results for the third scenario with 22% censoring and $n = 2000$.

Model	Parameter	True value	MLE	AB	MSE	RB	
EAPE-AFT (AIC = 6401.272)	α	2.9	2.088102	0.812	0.659	0.280	
	λ	1.5	1.257451	0.243	0.059	0.162	
	b	0.2	0.40004	0.200	0.040	1.000	
	β_0	-2	-1.8396	0.160	0.026	-0.080	
	β_1	0.65	0.693149	0.043	0.002	0.066	
	β_2	-0.7	-0.67023	0.030	0.001	-0.043	
	β_3	0.65	0.652168	0.002	0.000	0.003	
	β_4	-0.7	-0.7452	0.045	0.002	-0.065	
	EW-AFT(True model) (AIC = 6404.219)	α	1.9	1.877667	0.022	0.000	0.012
		λ	2.5	2.393933	0.106	0.011	0.042
b		2.2	1.124529	1.075	1.157	0.489	
β_0		-3	-3.39882	0.399	0.159	-0.133	
β_1		0.45	0.467593	0.018	0.000	0.039	
β_2		-0.65	-0.50064	0.149	0.022	-0.230	
β_3		0.45	0.535904	0.086	0.007	0.191	
β_4		-0.65	-0.80204	0.152	0.023	-0.234	
LL-AFT (AIC = 6407.494)		α	1.9	1.645	0.255	0.065	0.134
		b	3.8	0.265	3.535	12.496	0.930
	β_0	-1.5	-1.245	0.255	0.065	-0.170	
	β_1	0.35	0.686	0.336	0.113	0.960	
	β_2	-1.05	-1.012	0.038	0.001	-0.036	
	β_3	0.35	0.644	0.294	0.086	0.840	
	β_4	-1.05	-1.032	0.018	0.000	-0.017	
	W-AFT (AIC = 6417.564)	α	3.9	3.408	0.492	0.242	0.126
		b	0.5	0.6208	0.121	0.015	0.242
		β_0	-3.5	-3.775	0.275	0.076	-0.079
β_1		0.75	0.674	0.076	0.006	0.101	
β_2		-0.85	-1.004	0.154	0.024	-0.181	
β_3		0.75	0.634	0.116	0.013	0.155	
β_4		-0.85	-1.018	0.168	0.028	-0.198	

Table 12. Simulation results for the third scenario with 22% censoring and $n = 5000$.

Model	Parameter	True value	MLE	AB	MSE	RB	
EAPE-AFT (AIC = 4543.506)	α	2.9	2.801	0.099	0.010	0.034	
	λ	1.5	1.556	0.056	0.003	0.038	
	b	0.2	0.204	0.004	0.000	0.019	
	β_0	-2	-2.090	0.090	0.008	-0.045	
	β_1	0.65	0.670	0.020	0.000	0.030	
	β_2	-0.7	-0.790	0.090	0.008	-0.129	
	β_3	0.65	0.709	0.059	0.003	0.090	
	β_4	-0.7	-0.800	0.100	0.010	-0.143	
	EW-AFT(True model) (AIC = 4545.103)	α	1.9	1.867	0.033	0.001	0.018
		λ	2.5	2.462	0.038	0.001	0.015
b		2.2	2.189	0.011	0.000	0.005	
β_0		-3	-3.334	0.334	0.112	-0.111	
β_1		0.45	0.460	0.010	0.000	0.021	
β_2		-0.65	-0.600	0.050	0.002	-0.076	
β_3		0.45	0.392	0.058	0.003	0.128	
β_4		-0.65	-0.687	0.037	0.001	-0.057	
LL-AFT (AIC = 4547.09)		α	1.9	1.664	0.236	0.056	0.124
		b	3.8	3.827	0.027	0.001	0.007
	β_0	-1.5	-1.640	0.140	0.020	-0.093	
	β_1	0.35	0.367	0.017	0.000	0.048	
	β_2	-1.05	-1.011	0.039	0.002	-0.037	
	β_3	0.35	0.270	0.080	0.006	0.227	
	β_4	-1.05	-0.994	0.056	0.003	-0.053	
	W-AFT (AIC = 4567.423)	α	3.9	3.431	0.469	0.220	0.120
		b	0.5	0.145	0.355	0.126	0.710
		β_0	-3.5	-3.801	0.301	0.091	-0.086
β_1		0.75	0.648	0.102	0.010	0.136	
β_2		-0.85	-0.988	0.138	0.019	-0.162	
β_3		0.75	0.676	0.074	0.005	0.099	
β_4		-0.85	-0.973	0.123	0.015	-0.145	

In conclusion, the study has revealed that both sample size and censoring percentage impact the performance of the considered models. When the censoring and sample size were increased, the EAPE-AFT model consistently outperformed the EW-AFT, W-AFT and LL-AFT models.

5. Survival analysis

In survival analysis, the right censored data are the most common type of censored data available. The time to event in these analyses is commonly taken to be the time between survival and death. In this section, we focus on the use of multi-parametric (more than one covariate) hazard-based regression models. Hence, the EAPE-AFT regression model has been adopted to analyze real-world right-censored survival data. The main goal of this part of the study was to compare the EAPE-AFT model with its sub-models which include the alpha-power exponential AFT (APE-AFT), exponentiated exponential AFT (EE-AFT), and exponential AFT (E-AFT) distributions.

5.1. Data description

Breast cancer is a disease characterized by the abnormal growth of cells in breast tissue, affecting both men and women and constituting the most common type of cancer among women worldwide. According to the World Health Organization, it is responsible for over two million new cases and more than 600,000 deaths annually across the globe. This study utilized data previously analyzed by the authors of [43] on breast cancer patients from the Rotterdam tumor bank which contained records of 2982 primary breast cancer patients, and which formed the original data set. A full description of

the data can be found in [43].

The following covariates were taken into account for the illustrations of this study. We considered the following for each patient:

- (1) t : days to death or last follow-up; death: used for censoring.
- (2) x_1 : hormon (hormonal treatment).
- (3) x_2 : meno (menopausal status).
- (4) x_3 : age (age at surgery).

We begin with the exploration of the data and Table 13 gives the summary statistics for the data. In the sample of the study, approximately 57.383% of the data was censored, indicating that the exact observational lifetimes for those individuals are not known. The explanatory variable “age” represents the age of patients at the time of surgery and it was recorded as a continuous variable. The average age in the breast cancer patients’ data set from the Rotterdam bank was 44.507 years, with a standard deviation of 29.620.

Table 13. Summary statistics for breast cancer data.

Covariates		Percentage observations	Mean (standard deviation)
Days until death or the last follow-up	-	-	1.346(0.628)
Death	0	1710(57.383%)	
	1	1272(42.617%)	
Age at surgery (years)	-	-	44.507(29.620)
Meno	0	1312(44.027%)	
	1	1670(55.973%)	
Hormon	0	2643(88.691%)	
	1	339(11.309%)	

The variable “meno” describes the status of menopause referring to the classification of women based on their hormonal status as related to menopause. It is a categorical variable with two levels:

Premenopausal (0): This level indicates that women in the data set are still in their reproductive phase and have not yet experienced menopause.

Postmenopausal (1): This level indicates that women in the data set have gone through menopause, which is the permanent cessation of menstruation. Postmenopausal women can no longer conceive naturally as their ovaries have stopped releasing eggs. The data set consisted of 1312(44.027%) premenopausal patients and 1670(55.973%) postmenopausal patients. The variable “hormon” indicates whether the patients received hormonal treatment.

It is a categorical variable with two levels:

No (0): This level signifies that individuals in the data set have not undergone hormonal treatment.

Yes (1): This level indicates that individuals in the data set have received hormonal treatment. They have undergone medical intervention involving hormonal therapy or medication aimed at altering or regulating their hormone levels. There were 339 (11.309%) patients who received hormonal treatment.

The Cox PH plot is given in Figure 4, where panel (a) illustrates that the survival probabilities deviate from linearity at later time points, indicating an increasing HRF.

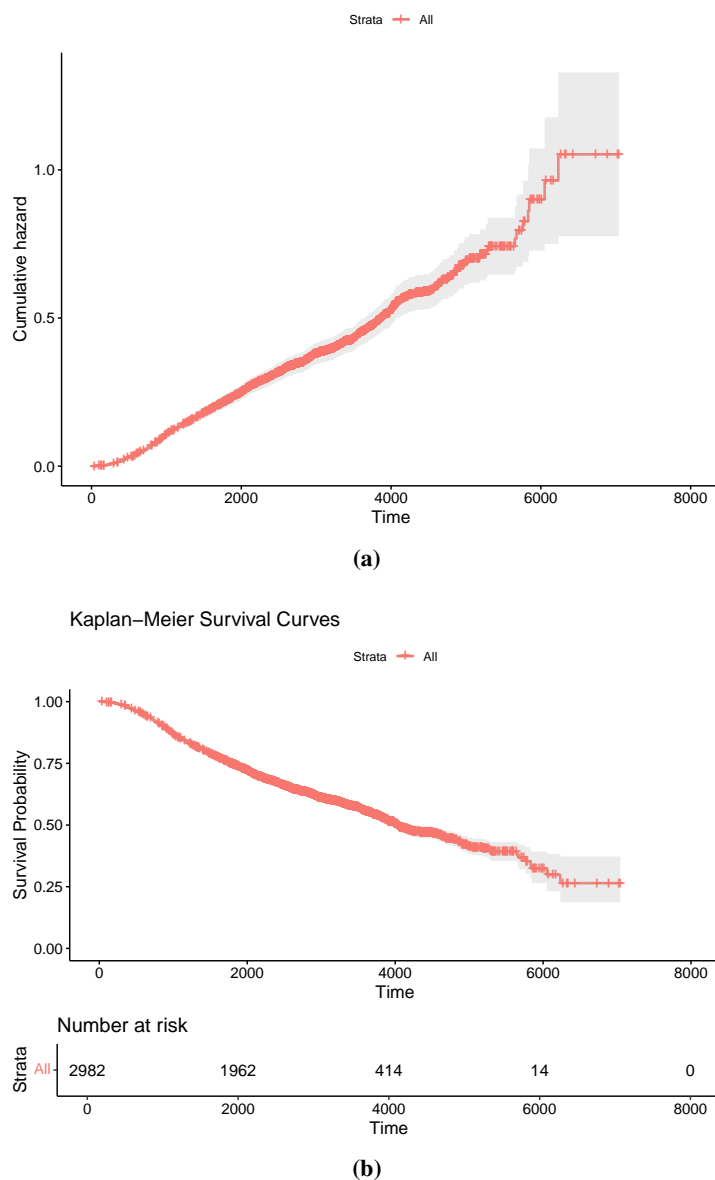


Figure 4. The Cox PH plot (a) and the Kaplan-Meier (b) for the breast cancer data.

The shape of the HRF is affirmed by the shape of the Kaplan-Meier plot in Figure 4. This plot shows a steep initial decrease in survival probability followed by a more steady decline over time reflecting an increasing HRF [31].

As depicted by the histogram and the non-parametric kernel density estimation results in Figure 5. The distribution of the data was asymmetrical and positively skewed, which is a common characteristic of survival data.

Furthermore, to identify extreme observations, we constructed box plots and violin plots as in Figure 5 which revealed the presence of some extreme observations.

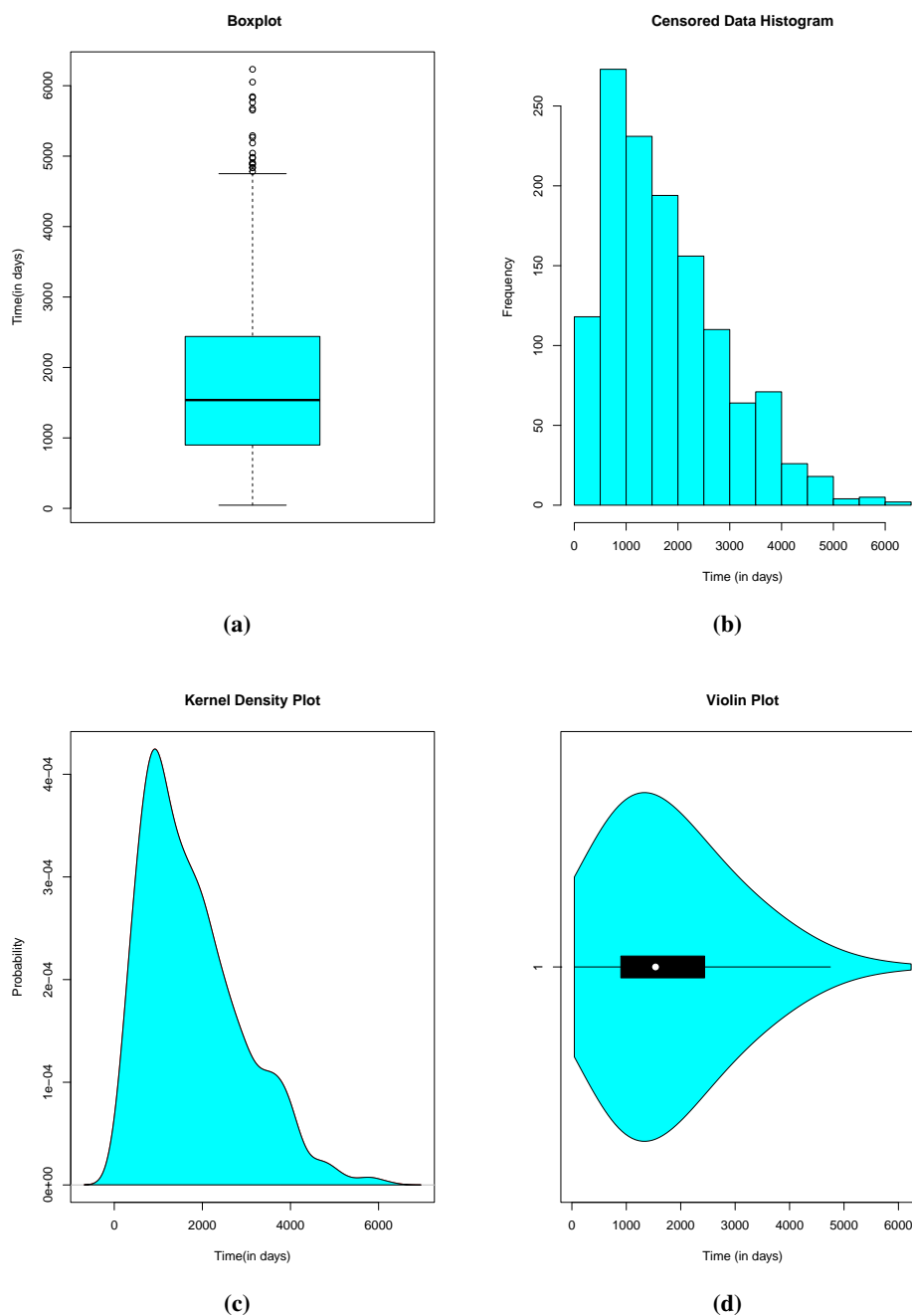


Figure 5. (a) Boxplot, (b) histogram, (c) kernel density, and (c) violin plots for breast cancer data.

5.2. The Cox PH model

To investigate the association between survival time and the factors believed to influence it, we simulated a Cox PH model. The Cox PH model was used to estimate the parameters. Table 14 presents the results of the regression analysis for the Cox PH model, displaying the regression coefficients, standard errors (SEs), and p-values. Notably, all covariates except the menopausal status yielded a

significant impact on the days to the death for breast cancer patients at a 5% level of significance.

Table 14. Results of the regression analysis for the Cox PH model.

Covariate	Coefficient	SE	p-value
Age	0.0146	0.0037	< 0.0001
Meno	0.0989	0.0967	0.3064
Hormon	0.2840	0.0870	0.0011

Schoenfeld residuals are useful for assessing the PH assumption in the Cox model. In this study, we applied the global test, which is a statistical test that is used to evaluate the overall PH assumption in a Cox PH model. It assesses whether there is a significant departure from proportionality across all covariates in the model. Table 15 provides evidence that the PH models do not adequately fit the breast cancer data. Additionally, Figure 6 clearly demonstrates the rejection of the assumption of PH for all covariates included. In other words, the PH models do not adequately capture the patterns observed in the breast cancer patient data set.

Table 15. Chi-squared result and p-value for Schoenfeld residual test at 5% level of significance.

Covariate	Chi square	p-value
Age	12.5480	0.0004
Meno	4.4860	0.0342
Hormon	0.4580	0.4986
GLOBAL	14.0560	0.0028

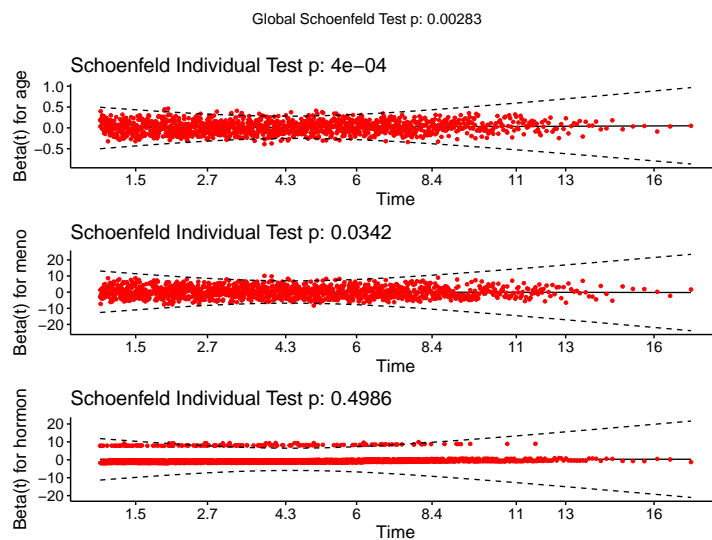


Figure 6. The conventional Schoenfeld residuals for breast cancer patient data, taking the test p-value for the meno, hormon, and age covariates into account.

5.3. Analysis of AFT model

Table 16 gives the values of the ML estimates and their corresponding Ses for the parameters of all considered regression models. Note that all of the coefficients were significant except for $\hat{\beta}_2$. This is

consistent with the results for the Cox PH model which indicated that the covariate hormon does not influence the remaining days to the death for breast cancer patients.

Table 16. The ML estimations and their corresponding SEs for AFT regression models.

MODEL	$\hat{\alpha}$	$\hat{\lambda}$	\hat{b}	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$
EAPE-AFT	1.024 (0.718)	0.089 (0.004)	1.486 (0.040)	0.122 (0.037)	0.074 (0.074)	1.197 (0.067)
APE-AFT	5.248 (0.974)	0.099 (0.007)	- -	0.165 (0.039)	0.065 (0.077)	0.259 (0.068)
EE-AFT	1.444 (0.051)	0.083 (0.005)	- -	0.145 (0.038)	0.085 (0.076)	0.250 (0.069)
E-AFT	0.054 (0.003)	- -	- -	0.180 (0.048)	0.100 (0.096)	0.253 (0.087)

Moreover, Table 17 provides the p-values and confidence intervals for each parameter, as they are measures that aid in determining the statistical significance of a parameter, and, together with information criterion measures, they were used for model comparison. The measures include the Akaike information criterion (AIC), the consistent Akaike's information criterion (CAIC), and the Bayesian information criterion (BIC). A p-value less than the level of significance indicates that the parameter is statistically significant and has a considerable influence on the outcome. The confidence interval defines a range within which we can be confident that the parameter's true value lies. If this interval excludes 0, it indicates that the parameter is statistically significant. The results given in Table 17 confirm that at 5% level of significance, all parameters were significant except for β_2 for all for the considered models.

Table 17. Z-values, 95% confidence interval (CI) and information measures for the EAPE-AFT model and other AFT models.

Model	Estimates	z-value	p-value	95%CI	AIC	CAIC	BIC
EAPE-AFT	$\hat{\alpha}$	3.674	0.004	(0.326 ,1.722)	9549.884	9591.886	9585.886
	$\hat{\lambda}$	26.221	0.000	(0.070 ,0.107)			
	\hat{b}	31.302	0.000	(1.324, 1.647)			
	$\hat{\beta}_1$	4.241	0.002	(0.046 ,0.197)			
	$\hat{\beta}_2$	1.260	0.342	(-0.079 , 0.228)			
APE-AFT	$\hat{\beta}_3$	3.673	0.004	(0.061 , 0.332)	9601.634	9636.636	9631.636
	$\hat{\alpha}$	5.390	0.000	(3.340 , 7.156)			
	$\hat{\lambda}$	13.518	0.000	(0.085 , 0.113)			
	$\hat{\beta}_1$	4.217	0.000	(0.088 , 0.242)			
	$\hat{\beta}_2$	0.837	0.403	(-0.087 , 0.217)			
EE-AFT	$\hat{\beta}_3$	3.786	0.000	(0.125 , 0.393)	9552.995	9597.997	9592.997
	$\hat{\alpha}$	28.392	0.000	(1.344 , 1.544)			
	$\hat{\lambda}$	16.449	0.000	(0.073 , 0.093)			
	$\hat{\beta}_1$	3.781	0.000	(0.070 , 0.220)			
	$\hat{\beta}_2$	1.106	0.269	(-0.065 , 0.235)			
E-AFT	$\hat{\beta}_3$	3.625	0.000	(0.115 , 0.385)	9644.280	9672.281	9668.281
	$\hat{\alpha}$	16.212	0.000	(0.047 , 0.061)			
	$\hat{\beta}_1$	3.749	0.000	(0.086 , 0.274)			
	$\hat{\beta}_3$	2.918	0.004	(0.083 , 0.423)			
	$\hat{\beta}_2$	1.035	0.301	(-0.089 , 0.289)			

The Kaplan-Meier survival curve is depicted in Figure 7, and it demonstrates a statistically significant differentiation in survival durations between the menopausal status group and the hormonal treatment, as evidenced by a p-value below 0.0001. Additionally, the Kaplan-Meier plot

confirms the violation of the Cox PH assumption when considering the covariate hormon; it also affirms that under the AFT model, the covariate meno does not influence the outcome variable.

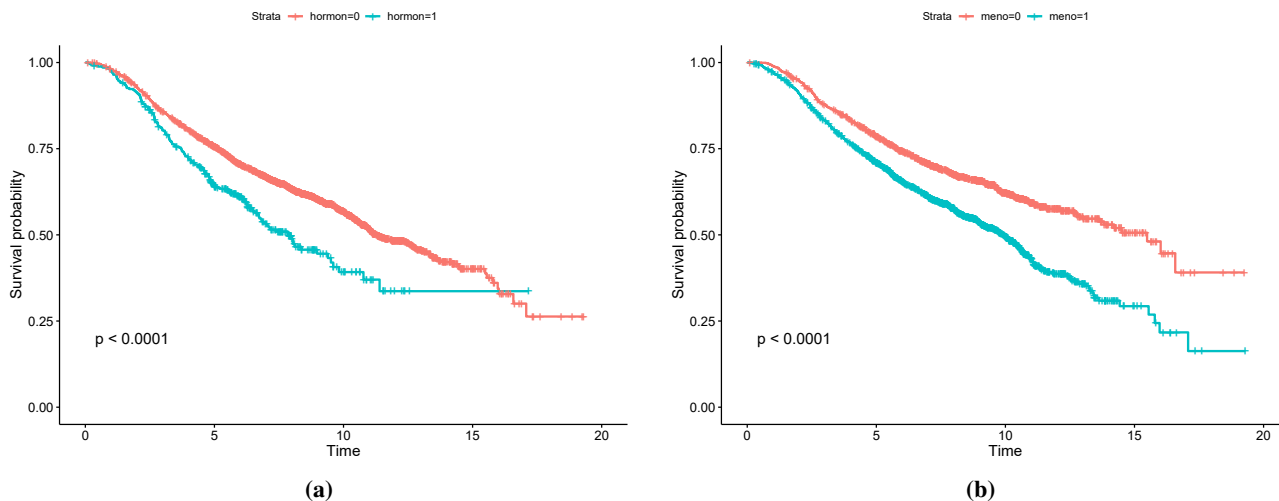


Figure 7. Kaplan Meier curve (a) for covariate hormon (b) for covariate meno

In the analysis of the proposed model, we have compared the fit of the EAPE-AFT model with the APE-AFT, EE-AFT, and E-AFT sub-models whose PDFs are respectively given by

$$f_{APE-AFT}(t, \alpha, \lambda, \beta) = \frac{\lambda e^{x'\beta}}{(\alpha - 1)} (\log(\alpha)) e^{-\lambda t e^{x'\beta}} \left(\alpha^{1 - e^{-\lambda t e^{x'\beta}}} \right), \tag{5.1}$$

$$f_{EE-AFT}(t, b, \lambda, \beta) = b \lambda e^{x'\beta} e^{-\lambda t e^{x'\beta}} \left(1 - e^{-\lambda t e^{x'\beta}} \right)^{b-1}, \tag{5.2}$$

$$f_{E-AFT}(t, b, \lambda, \beta) = \lambda e^{x'\beta} e^{-\lambda t e^{x'\beta}}. \tag{5.3}$$

To compare the fit of our model with its sub-models, three information criteria have been utilized. These include the AIC, CAIC, and BIC. The values of the information criteria are provided in Table 17 which shows that the EAPE-AFT model outperformed all AFT regression models. Additionally, the HRFs for all considered competing AFT models were fit to the data and are displayed in Figure 8.

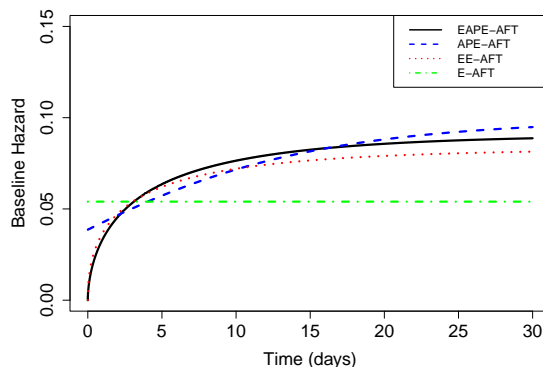


Figure 8. Estimated HRFs for the considered distributions.

6. Discussion

The main goal of this study was to develop a more adaptable model than the Cox model to improve the ability of the exponential distribution to model real-life data. An extensive Monte Carlo simulation was utilized to evaluate the estimators of the EAPE-AFT parameters. Finally, the proposed model was validated on breast cancer data and the choice of the AFT model was justified by the simulation results confirming a direct relationship between covariates and time to event, which improved interpretability.

Regarding the three hazard rate shapes, i.e., the increasing, decreasing, and unimodal shapes, the simulation results revealed that the extended EAPE-AFT model outperformed other existing AFT models, i.e., the W-AFT and LL-AFT models, in terms of ability to different HRFs. The values of the performance metrics such as the AB, RB, and MSE indicate the usefulness of the proposed EAPE-AFT model.

While the proposed model facilitates modeling and the interpretation of survival data, it is limited to data that exhibit bathtub hazard shapes. Additionally, the model is unable to handle crossing survival curves.

7. Conclusions and recommendation

In this paper, we have proposed a new survival regression model for analyzing and modeling censored survival data, called the EAPE-AFT model. The new model outperformed other existing AFT models. We have developed methods for estimating the EAPE-AFT parameters by using the ML estimation technique. We have conducted simulation studies to compare the performance of the proposed model with that of other models, such as the W-AFT and the LL-AFT models. Based on our study, we have found out that the new AFT model provides a more accurate fit for various shapes of hazard rates and different levels of censoring. We have also applied the model to breast cancer patients data and found that it provides a better fit than its sub-models. In conclusion, the EAPE-AFT regression model has the potential to be a useful tool for analyzing various forms of survival data, as it can be applied to several widely known AFT models as special cases.

Future research should investigate residual analysis methodologies, diagnostic criteria for model fit, and model extension to improve the handling of diverse censoring processes.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

The authors extend their appreciation to the Deputyship for Research and Innovation, Ministry of Education in Saudi Arabia for funding this research work through project number 445-9-876. The authors would also like to thank the Editor and the reviewers for their constructive comments and suggestions which have greatly improved the paper.

Conflict of interest

The authors of this paper declare no conflicts of interest that might have interfered with the findings or development of this study.

References

1. J. F. Lawless, *Statistical models and methods for lifetime data*, John Wiley & Sons, Inc., 2011. <https://doi.org/10.1002/9781118033005>
2. H. Piriaei, G. Yari, R. Farnoosh, E-Bayesian estimations for the cumulative hazard rate and mean residual life based on exponential distribution and record data, *J. Stat. Comput. Simul.*, **90** (2020), 271–290. <https://doi.org/10.1080/00949655.2019.1678623>
3. A. H. Muse, S. Mwalili, O. Ngesa, H. M. Alshanbari, S. K. Khosa, E. Hussam, Bayesian and frequentist approach for the generalized log-logistic accelerated failure time model with applications to larynx-cancer patients, *Alex. Eng. J.*, **10** (2022), 7953–7978. <https://doi.org/10.1016/j.aej.2022.01.033>
4. S. A. B. Mastor, O. Ngesa, J. Mung’atu, N. M. Alfaer, A. Z. Afify, The extended exponential Weibull distribution: Properties, inference, and applications to real-life data, *Complexity*, **2022** (2022), 4068842. <https://doi.org/10.1155/2022/4068842>
5. S. A. Khan, Exponentiated Weibull regression for time-to-event data, *Lifetime Data Anal.*, **24** (2018), 328–354. <https://doi.org/10.1007/s10985-017-9394-3>
6. C. D. Lai, *Generalized Weibull distributions*, Springer, 2014, https://doi.org/10.1007/978-3-642-39106-4_2
7. A. H. Muse, S. Mwalili, O. Ngesa, C. Chesneau, A. Al-Bossly, M. El-Morshedy, Bayesian and frequentist approaches for a tractable parametric general class of Hazard-based regression models: an application to oncology data, *Mathematics*, **10** (2022), 3813. <https://doi.org/10.3390/math10203813>
8. R. D. Gupta, D. Kundu, Exponentiated exponential family: an alternative to gamma and Weibull distributions, *Biomet. J.*, **43** (2001), 117–130. [https://doi.org/10.1002/1521-4036\(200102\)43:1%3C117::AID-BIMJ117%3E3.0.CO;2-R](https://doi.org/10.1002/1521-4036(200102)43:1%3C117::AID-BIMJ117%3E3.0.CO;2-R)
9. S. Nadarajah, F. Haghghi, An extension of the exponential distribution, *Statistics*, **45** (2011), 543–558. <https://doi.org/10.1080/02331881003678678>
10. A. Z. Afify, M. Zayed, M. Ahsanullah, The extended exponential distribution and its applications, *J. Stat. Theory Appl.*, **17** (2018), 213–229. <https://doi.org/10.2991/jsta.2018.17.2.3>
11. S. Nadarajah, S. Kotz, The exponentiated type distributions, *Acta Appl. Math.*, **92** (2006), 97–111. <https://doi.org/10.1007/s10440-006-9055-0>
12. S. Nadarajah, S. Kotz, The beta exponential distribution, *Reliab. Eng. Syst. Saf.*, **91** (2006), 689–697. <https://doi.org/10.1016/j.ress.2005.05.008>
13. M. Kilai, G. A. Waititu, W. A. Kibira, M. M. A. El-Raouf, T. A. Abushal, A new versatile modification of the Rayleigh distribution for modeling COVID-19 mortality rates, *Results Phys.*, **35** (2022), 105–260. <https://doi.org/10.1016/j.rinp.2022.105260>

14. E. A. ElSherpieny, E. M. Almetwally, The exponentiated generalized Alpha power family of distribution: properties and applications, *Pak. J. Stat. Oper. Res.*, **18** (2022), 349–367. <https://doi.org/10.18187/pjsor.v18i2.3515>
15. D. Alvares, E. Lázaro, V. Gómez-Rubio, C. Armero, Carmen Bayesian survival analysis with BUGS, *Stat. Med.*, **40** (2021), 2975–3020. <https://doi.org/10.1002/sim.8933>
16. F. N. Demarqui, V. D. Mayrink, Yang and Prentice model with piecewise exponential baseline distribution for modeling lifetime data with crossing survival curve, *Braz. J. Probab. Stat.*, **35** (2020), 172–186. <https://doi.org/10.1214/20-BJPS471>
17. F. J. Rubio, L. Remontet, N. P. Jewell, A. Belot, On a general structure for hazard-based regression models: an application to population-based cancer research, *Stat. Methods Med. Res.*, **28** (2019), 2404–2417. <https://doi.org/10.1177/09622802187822>
18. H. Zhou, T. Hanson, *Bayesian spatial survival models*, Springer, 2015. https://doi.org/10.1007/978-3-319-19518-6_11
19. L. M. Leemis, L. H. Shih, K. Reynertson, Variate generation for accelerated life and proportional hazards models with time-dependent covariates, *Stat. Probab. Lett.*, **10** (1990), 335–339. [https://doi.org/10.1016/0167-7152\(90\)90052-9](https://doi.org/10.1016/0167-7152(90)90052-9)
20. M. Ashraf-Ul-Alam, A. A. Khan, Generalized Topp-Leone-Weibull AFT modelling: a Bayesian analysis with MCMC tools using R and Stan, *Aust. J. Stat.*, **50** (2021), 52–76. <https://doi.org/10.17713/ajs.v50i5.1166>
21. A. B. Mastor, O. Ngesa, J. Mung’atu, A. Z. Afify, The extended exponential-Weibull accelerated failure time model with applications to cancer data set, *International Conference on Mathematics and Its Applications in Science and Engineering*, 2022.
22. K. Burke, M. C. Jones, A. Noufaily, A flexible parametric modelling framework for survival analysis, *J. R. Stat. Soc. Ser. C Appl. Stat.*, **69** (2020), 429–457. <https://doi.org/10.1111/rssc.12398>
23. X. Wang, Y. Yue, J. J. Faraway, *Bayesian regression modeling with INLA*, 1 Ed., Chapman and Hall/CRC, 2018. <https://doi.org/10.1201/9781351165761>
24. S. A. Khan, N. Basharat, Accelerated failure time models for recurrent event data analysis and joint modeling, *Comput. Stat.*, **37** (2022), 1569–1597. <https://doi.org/10.1007/s00180-021-01171-7>
25. J. Le-Rademacher, X. Wang, Time-to-event data: an overview and analysis considerations, *J. Thorac. Oncol.*, **16** (2021), 1067–1074. <https://doi.org/10.1016/j.jtho.2021.04.004>
26. S. Choi, H. Cho, Accelerated failure time models for the analysis of competing risks, *J. Korean Stat. Soc.*, **48** (2019), 315–326. <https://doi.org/10.1016/j.jkss.2018.10.003>
27. R. Mokarram, M. Emadi, A. H. Rad, M. J. Nooghabi, A comparison of parametric and semi-parametric survival models with artificial neural networks, *Commun. Stat.*, **47** (2018), 738–746. <https://doi.org/10.1080/03610918.2017.1291961>
28. I. Selingerova, S. Katina, I. Horova, Comparison of parametric and semiparametric survival regression models with kernel estimation, *J. Stat. Comput. Simul.*, **91** (2021), 2717–2739. <https://doi.org/10.1080/00949655.2021.1906875>
29. J. P. Klein, M. L. Moeschberger, *Survival analysis: techniques for censored and truncated data*, 2 Eds., Springer, 2003. <https://doi.org/10.1007/b97377>

30. T. M. Therneau, P. M. Grambsch, T. M. Therneau, P. M. Grambsch, The Cox model, In: *Modeling survival data: extending the Cox model*, Springer, 2000. https://doi.org/10.1007/978-1-4757-3294-8_3
31. D. G. Kleinbaum, M. Klein, *Survival analysis: a self-learning text*, 3 Eds., Springer, 2012. <https://doi.org/10.1007/978-1-4419-6646-9>
32. M. J. Crowther, P. Royston, M. Clements, A flexible parametric accelerated failure time model, *ArXiv*, 2020. <https://doi.org/10.48550/arXiv.2006.06807>
33. J. D. Kalbfleisch, R. L. Prentice, *The statistical analysis of failure time data*, 2 Eds., John Wiley & Sons, Inc., 2001. <https://doi.org/10.1002/9781118032985>
34. V. Kariuki, A. Wanjoya, O. Ngesa, M. Alqawba, A flexible family of distributions based on the alpha power family of distributions and its application to survival data, *Pak. J. Statist.*, **39** (2023), 237–258.
35. D. G. Kleinbaum, M. Klein, *Survival analysis a self-learning text*, 3 Eds., Springer, 1996. <https://doi.org/10.1007/978-1-4419-6646-9>
36. S. Wang, W. Chen, M. Chen, Y. Zhou, Maximum likelihood estimation of the parameters of the inverse Gaussian distribution using maximum rank set sampling with unequal samples, *Math. Popul. Stud.*, **30** (2023), 1–21. <https://doi.org/10.1080/08898480.2021.1996822>
37. D. G. Kleinbaum, *Evaluating the proportional hazards assumption*, 3 Eds., Springer, 2012. https://doi.org/10.1007/978-1-4757-2555-1_4
38. C. G. Broyden, A new method of solving nonlinear simultaneous equations, *Comput. J.*, **12** (1969), 94–99. <https://doi.org/10.1093/comjnl/12.1.94>
39. R. Fletcher, *A class of methods for nonlinear programming with termination and convergence properties*, North-Holland, 1970.
40. D. Goldfarb, A family of variable-metric methods derived by variational means, *Math. Comput.*, **24** (1970), 23–26. <https://doi.org/10.2307/2004873>
41. D. F. Shanno, Conditioning of quasi-Newton methods for function minimization, *Math. Comput.*, **24** (1970), 647–656.
42. S. A. Khan, S. K. Khosa, Generalized log-logistic proportional hazard model with applications in survival analysis, *J. Stat. Distrib. Appl.*, **3** (2016), 16. <https://doi.org/10.1186/s40488-016-0054-z>
43. P. Royston, D. G. Altman, External validation of a Cox prognostic model: principles and methods, *BMC Med. Res. Methodol.*, **13** (2013), 33. <https://doi.org/10.1186/1471-2288-13-33>



©2024 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)