

Article

The Extended Exponential-Weibull Accelerated Failure Time Model with Application to Sudan COVID-19 Data

Adam Braima S. Mastor ^{1,*} , Abdulaziz S. Alghamdi ², Oscar Ngesa ³, Joseph Mung'atu ⁴,
Christophe Chesneau ^{5,*}  and Ahmed Z. Afify ⁶ 

¹ Department of Mathematics (Statistics Option) Programme, Pan African University, Institute for Basic Sciences, Technology and Innovation (PAUSTI), Nairobi, 62000-00200, Kenya

² Department of Mathematics, College of Science & Arts, King Abdulaziz University, P.O. Box 344, Rabigh 21911, Saudi Arabia

³ Department of Mathematics and Physical Sciences, Taita Taveta University, Voi 635-80300, Kenya

⁴ Department of Mathematics, Jomo Kenyatta University of Agriculture and Technology (JKUAT), P.O. box 62000-00200, Kenya

⁵ Department of Mathematics, LMNO, CNRS-Université de Caen, Campus II, Science 3, 14032 Caen, France

⁶ Department of Statistics, Mathematics and Insurance, Benha University, Benha, 13511, Egypt

* Correspondence: soliman.adam1@students.jkuat.ac.ke (A.B.S.M.); christophe.chesneau@unicaen.fr (C.C.)

Abstract: A fully parametric accelerated failure time (AFT) model with a flexible, novel modified exponential Weibull baseline distribution called the extended exponential Weibull accelerated failure time (ExEW-AFT) model is proposed. The model is presented using the multi-parameter survival regression model, where more than one distributional parameter is linked to the covariates. The model formulation, probabilistic functions, and some of its sub-models were derived. The parameters of the introduced model are estimated using the maximum likelihood approach. An extensive simulation study is used to assess the estimates' performance using different scenarios based on the baseline hazard shape. The proposed model is applied to a real-life right-censored COVID-19 data set from Sudan to illustrate the practical applicability of the proposed AFT model.

Keywords: baseline hazard; survival regression model; maximum likelihood; Monte Carlo simulation; COVID-19 data

MSC: 65C05; 62N01; 62N02; 62P10



Citation: Mastor, A.B.S.; Alghamdi, A.S.; Ngesa, O.; Mung'atu, J.; Chesneau, C.; Afify, A.Z. The Extended Exponential-Weibull Accelerated Failure Time Model with Application to Sudan COVID-19 Data. *Mathematics* **2023**, *11*, 460. <https://doi.org/10.3390/math11020460>

Academic Editors: Marina Alexandra Pedro Andrade and Maria Alves Teodoro

Received: 20 December 2022

Revised: 7 January 2023

Accepted: 10 January 2023

Published: 15 January 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Over the past few decades, the semi-parametric Cox model has been extensively adopted in the analysis of survival data. Modifications to remove the assumption of “proportional hazards (PH)” are discussed in Cox’s original paper [1]. Many efforts have been made to increase the adaptability of hazard-based regression models using flexible functions for both the baseline hazard and the inclusion of time-dependent parameters, primarily using modified probability distributions [2–5].

The two most popular techniques for parametric hazard-based regression models of survival data are PH models and accelerated failure time (AFT) models [6,7]. Specifically, under the parametric PH assumption, only a few probability models are closed, and none of them are flexible enough to explain a large range of survival data [8]. In some cases, and under certain probability distributions, the AFT model is a more valuable and realistic option than the PH model [9]. However, the basic structure of such models remains the same: the PH model is written as a baseline hazard rate function (HRF), multiplied by the exponential function of the covariates [10], in the following form:

$$h(t; \mathbf{x}) = h_0(t) \exp(\mathbf{x}'\boldsymbol{\beta}),$$

where $h(t; \mathbf{x})$ is the HRF for a subject $\mathbf{x} = (x_1, x_2, \dots, x_n)'$ is the vector of covariates, $h_0(t)$ is the baseline HRF obtained with $\mathbf{x} = 0$, and $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_p)'$ represents the regression coefficients. The intercept β_0 can be included in this vector, but we leave it out for clarity.

Other alternative PH models have been proposed, and they can directly account for the time-dependent effects of covariates. The AFT model is a notable example of these earliest alternatives to the PH [11]. According to the AFT model, the covariates have a direct effect on the time to event, as opposed to the PH model, where the covariates affect the HRF only [12].

Some medical studies have shown that AFT models are frequently used to analyze survival data [13]. In comparison to the interpretation of a hazard rate, which denotes a relative rise or decrease in the event rate, the interpretation of an acceleration factor can be thought of as being more obvious because it directly affects the survival time, either by increasing it or decreasing it [14].

Additionally, parametric survival models are essential for assessing survival data [15]. These models can be applied to various applications [14,16,17]. For instance, (i) when the baseline hazard is theoretically expected in a healthcare data set, a survival analysis can be applied to produce a relatively better estimation, (ii) the survival models are applicable to the spatial models that predict disease prevalence, (iii) the models can provide better estimates for mixed effects in the clustered survival data-sets, and (vi) the survival rates can be estimated using the random effects-frailty models, which are part of the parametric survival models.

Furthermore, to formulate the parametric AFT model framework, the Weibull, log-logistic, and log-normal distributions are commonly utilized as baseline HRFs [8]. The Weibull family can accommodate monotone HRFs (i.e., increasing and decreasing), while log-logistic and log-normal can accommodate non-monotone HRFs [18,19]. These distributions cannot accommodate both monotone and non-monotone HRFs [20]. To address this problem, we looked for a baseline distribution that can accommodate different HRFs as well as be closed under the AFT model framework.

The foundation for the PH models is the idea that the hazard ratio will remain constant throughout time [21]. However, in some cases, this assumption is flawed [22].

Results that are skewed and deceptive may emerge from violating this premise [23]. The authors in [24–27] provide inspiring examples of comorbidity studies with COVID-19 patients. This study aims to determine whether covariates within a patient affect how long the patient stays in the hospital. The standardized Schoenfeld residuals, along with the proportional hazards assumption of the Cox model for covariates (comorbidity and sex), are used. Figure 1 shows that all covariates reject the null hypothesis of the test of PH, so the Cox PH models can't be used to study the influence of COVID-19's covariates on discharge time.

Based on the above argument, in this research study, we proposed a more flexible parametric survival regression model called the extended-exponential-Weibull (ExEW) AFT model.

The main contribution of this study is to offer a useful addition to the toolkit for analyzing survival data that can be used with different hazard regression models in a more general way. There are several reasons why this work focuses on the extension of the AFT model under the ExEW HRF baseline. Firstly, although some classical distributions closed under the AFT model framework exist, none of them are versatile enough to include both monotone and non-monotone HRFs. Secondly, the parametric AFT model may produce more accurate estimates than the semi-parametric PH model. Thirdly, the use of the flexible distribution, which captures both monotone and non-monotone HRFs, is what makes our work distinctive and more appealing.

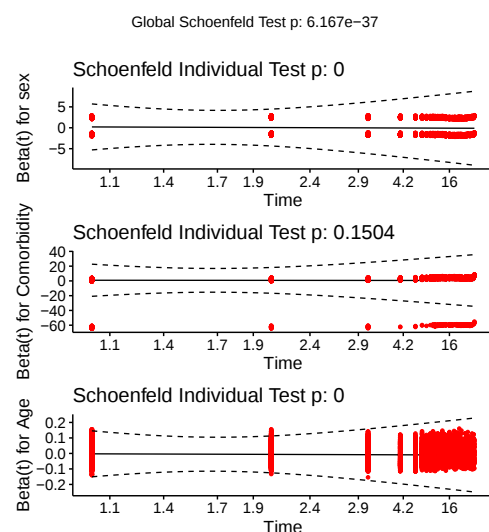


Figure 1. The conventional Schoenfeld residuals from the application of the coronavirus disease 2019 (COVID-19) data set while taking the test p -value for the sex and comorbidity covariates into account.

The motivations for the study can be briefly summarized as follows:

- To propose an ExEW-AFT model that is quite adaptable and can easily accommodate a variety of applications in reliability and survival analysis.
- The ExEW-AFT model could be viewed as a multiple-parameter survival regression, which is perhaps more adaptable than the common single-parameter survival regression model, including the Weibull and exponential AFT models.
- The results obtained with the application are considered the main contributions of this work.
- COVID-19 data from Sudan are used in the proposed model to determine how the disease risk factors affect the length of hospital stays.

The subsequent sections of the paper are structured as follows: The model formulation is presented in Section 2, and baseline HRFs are discussed in Section 3. The proposed model is examined in Section 4. Section 5 discusses the estimation of parameters. Section 6 presents the results of the simulation study. Application of the proposed model to COVID-19 data is offered in Section 7. Discussion is given in Section 8. In Section 9, some conclusions are cited.

2. AFT Model Formulation

For regression analysis of time-to-event data, there are two prevalent classes: odds-based and hazard-based regression models [2]. While the formulation in hazard-based models is dependent on the tractability of the baseline distribution's HRF and cumulative HRF (CHRF), the formulation in odds-based models is dependent on the tractability of the odds function and its derivative [28].

The three common regression models in the context of hazard-based regression models are: PH [2], AFT [18], and accelerated hazard (AH) [20] models. On the other hand, the three most popular regression models in the context of odds-based regression models are proportional odds (PO) [29], accelerated odds (AO) [28], and AFT models. Hence, the AFT model is the only survival regression model that is closed under both odds-based and hazards-based regression models. The formulation of the AFT model is defined as follows:

Let $\mathbf{x}$ be a vector of explanatory variables, and $\psi(\mathbf{x}'\boldsymbol{\beta})$ be the link function for the explanatory variables, where $\boldsymbol{\beta}$ is a vector of regression coefficients. The HRF and survival function (SF) of the AFT model are expressed as follows:

$$h(t; \mathbf{x}) = h_0(t\psi(\mathbf{x}'\boldsymbol{\beta}))\psi(\mathbf{x}'\boldsymbol{\beta}), \quad t \geq 0 \quad (1)$$

and

$$S(t; \mathbf{x}) = S_0(t\psi(\mathbf{x}'\boldsymbol{\beta})). \quad (2)$$

The link function has the following properties:

- (i) $\psi(x) > 0$ for all $x \neq 0$.
- (ii) $\psi(x)$ is a one-to-one monotone function.
- (iii) $\psi(0) = 1$.

In this study, we employed the link function as a standard exponential function. Hence, the HRF and SF of the AFT model can be re-written as follows:

$$h(t; \mathbf{x}) = h_0(te^{\mathbf{x}'\boldsymbol{\beta}})e^{\mathbf{x}'\boldsymbol{\beta}} \quad (3)$$

and

$$S(t; \mathbf{x}) = S_0(te^{\mathbf{x}'\boldsymbol{\beta}}). \quad (4)$$

3. The Baseline ExEW Distribution

In this study, we consider a very flexible lifetime distribution as the baseline distribution: the four-parameter ExEW distribution, also denoted as $\text{ExEW}(a, b, c, \alpha)$ distribution, as presented in Mastor et al. [30] (with some corrections). Some background on it is given below. To begin, the cumulative distribution function (CDF) of the ExEW distribution takes the form

$$F(z) = 1 - \exp[-\alpha(az + bz^c)], \quad z > 0, \quad (5)$$

where $a > 0$, $c > 0$ and $\alpha > 0$ are shape parameters, and $b > 0$ is a scale parameter. The probability density function (PDF) corresponding to Equation (4) reduces to

$$f(z) = \alpha(a + bcz^{c-1}) \exp[-\alpha(az + bz^c)], \quad z > 0. \quad (6)$$

The SF of the ExEW distribution is

$$S(z) = \exp[-\alpha(az + bz^c)], \quad z > 0. \quad (7)$$

The HRF of the ExEW distribution is expressed as

$$h(z) = \alpha(a + bcz^{c-1}), \quad z > 0. \quad (8)$$

The corresponding CHRF is obtained as

$$H(z) = \alpha(az + bz^c), \quad z > 0. \quad (9)$$

By following the CDF in Equation (5), the HRF and CHRF for the baseline distribution have been corrected from the original equations in the original paper [30].

4. The Proposed Model

The proposed AFT model is developed by extending the $\text{ExEW}(a, b, c, \alpha)$ distribution to incorporate covariates. The corresponding SF with covariate vector $\mathbf{x}$ is given by

$$S(t; \mathbf{x}) = \exp\left\{-\alpha\left[a\left(te^{\mathbf{x}'\boldsymbol{\beta}}\right) + b\left(te^{\mathbf{x}'\boldsymbol{\beta}}\right)^c\right]\right\}, \quad t > 0, \quad (10)$$

which corresponds to the SF of the ExEW distribution under the following configuration: $a^* = ae^{\mathbf{x}'\boldsymbol{\beta}}$ and $b^* = b\left(e^{\mathbf{x}'\boldsymbol{\beta}}\right)^c$, that is

$$S(t; \mathbf{x}) = \exp[-\alpha(a^*t + b^*t^c)], \quad t > 0.$$

This demonstrates that, under the AFT model framework, the ExEW distribution is closed in the distributional sense.

Furthermore, the HRF with covariates $\mathbf{x}$ is written as follows:

$$h(t; \mathbf{x}) = \alpha \left[a + b c \left(t e^{\mathbf{x}'\boldsymbol{\beta}} \right)^{c-1} \right] e^{\mathbf{x}'\boldsymbol{\beta}}, \quad t > 0. \quad (11)$$

The PDF with covariates vector $\mathbf{x}$ is given by

$$f(t; \mathbf{x}) = \alpha \left[a + b c \left(t e^{\mathbf{x}'\boldsymbol{\beta}} \right)^{c-1} \right] \exp \left\{ -\alpha \left[a \left(t e^{\mathbf{x}'\boldsymbol{\beta}} \right) + b \left(t e^{\mathbf{x}'\boldsymbol{\beta}} \right)^c \right] \right\} e^{\mathbf{x}'\boldsymbol{\beta}}, \quad t > 0. \quad (12)$$

Single-parameter hazard-based regression (SPHBR) models are commonly used to relate covariates to one parameter of specific interest. In these SPHBR models, the role of the other (explanatory independent variables) parameters are often little more than to give the model sufficient generality to adapt to the data. A more tractable method is to relate these other parameters to covariates; this method is known as multi-parameter hazard-based regression (MPHBR) models [31–34]. The primary focus of this paper is the development of MPHBR models in the context of time-to-event analysis.

In our case, this is a multi-parameter AFT model, which is different from a single-parameter AFT model, like the Weibull AFT model or the log-logistic AFT model, among others, where only the scale parameter is changed. As a result, except for the α parameter, the covariates influence the majority of the baseline distribution parameters. This is what makes our work unique and different from the common classical AFT models.

5. Estimation of the ExEW-AFT Parameters

To estimate the model parameters, maximum likelihood estimation (MLE) is used. Let $T_1, T_2, \dots, T_n$ be the lifetimes of n individuals. If the data are subject to right censoring, then $t_i = \min(T_i, C_i)$, where $C_i > 0$ corresponding to a potential censoring time for individual i . Suppose that $\delta_i = I(T_i \leq C_i) = 1$ for $T_i \leq C_i$ and $\delta_i = 0$ otherwise. Hence, the observed data for an individual i consists of $\{t_i, \delta_i\}$, for $i = 1, 2, \dots, n$, where t_i is a censoring time or lifetime according to whether $\delta_i = 0$ or 1, respectively and $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{in})$ is a column vector of n external covariates for the i th individual.

In this scenario, non-informative censoring is assumed to be in place, meaning that neither the distribution of survival times nor the distribution of censoring times can be inferred from one another.

It is important to note that the assumption of non-informative censoring is justifiable when censoring is random (it is assumed that the failure rates for observations that are censored, uncensored, and remain in the risk set are equal) and/or independent (in other words, censorship is supposed to be random within any interested subgroup); for additional information on the non-informative censoring, see [35].

In this setting, the censored likelihood function is therefore defined as follows:

$$L(\boldsymbol{\theta}) = \prod_{i=1}^n [f(t; \mathbf{x}_i)]^{\delta_i} [S(t; \mathbf{x}_i)]^{1-\delta_i}, \quad (13)$$

where $\boldsymbol{\theta}$ is the vector of the involved parameters.

Based on Equation (13), the log-likelihood function for a parametric AFT model is defined as follows:

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^n \delta_i \log[f(t; \mathbf{x}_i)] + \sum_{i=1}^n (1 - \delta_i) \log[S(t; \mathbf{x}_i)], \quad (14)$$

The Newton-Raphson optimization procedure can be used to directly optimize this, and interval estimates of the model parameters and hypothesis testing are both possible under the approximative normally distributed MLE estimates [8].

The likelihood function in the ExEW-AFT model is given by

$$L(\theta, D) = \prod_{i=1}^n [f(t_i; a, b, c, \alpha, \boldsymbol{\beta}, \mathbf{x}_i)]^{\delta_i} [S(t_i; a, b, c, \alpha, \boldsymbol{\beta}, \mathbf{x}_i)]^{1-\delta_i}, \quad (15)$$

where $S(t; a, b, c, \alpha, \boldsymbol{\beta}, \mathbf{x})$ and $f(t; a, b, c, \alpha, \boldsymbol{\beta}, \mathbf{x})$ of the ExEW-AFT model are derived from Equations (10) and (12), $\theta = (a, b, c, \alpha, \boldsymbol{\beta})$, $D = ((t_1, \delta_1, \mathbf{x}_1), (t_2, \delta_2, \mathbf{x}_2), \dots, (t_n, \delta_n, \mathbf{x}_n))$. We recall that the censoring indicator satisfies $\delta = 0$ if the observation is censored and $\delta = 1$ if the observation is failed, and $\mathbf{x}_i$ is the matrix of covariates, which is known as the design matrix or model matrix. After expressing the PDF in terms of HRF and SF and taking the logarithm of both sides of the likelihood function, the log-likelihood can be written as follows:

$$\ell(\theta; D) = \sum_{i=1}^n [\delta_i \log h(t_i; a, b, c, \alpha, \boldsymbol{\beta}, \mathbf{x}_i) + \log S(t_i; a, b, c, \alpha, \boldsymbol{\beta}, \mathbf{x}_i)]. \quad (16)$$

As a result, the full log-likelihood function for the ExEW-AFT model can be written as follows:

$$\begin{aligned} \ell(\theta; D) = \sum_{i=1}^n \delta_i \log \left[\alpha \left(a + b c \left(t e^{\mathbf{x}_i' \boldsymbol{\beta}} \right)^{c-1} \right) e^{\mathbf{x}_i' \boldsymbol{\beta}} \right] \\ + \sum_{i=1}^n \left(-\alpha \left[a \left(t e^{\mathbf{x}_i' \boldsymbol{\beta}} \right) + b \left(t e^{\mathbf{x}_i' \boldsymbol{\beta}} \right)^c \right] \right). \end{aligned} \quad (17)$$

The MLE vector estimate of θ can be obtained by maximizing Equation (17) directly, with respect to the parameter vector θ . It is denoted by $\hat{\theta} = (\hat{a}, \hat{b}, \hat{c}, \hat{\alpha}, \hat{\boldsymbol{\beta}})$, so that $\hat{a}$, $\hat{b}$, $\hat{c}$, $\hat{\alpha}$, and $\hat{\boldsymbol{\beta}}$ are the MLEs of a , b , c , α and $\boldsymbol{\beta}$, respectively. By directly maximizing the total log-likelihood function with the aid of the tools R, MATHEMATICA, and MATLAB, the parameter estimations can be derived. R software is the program that is utilized in this work.

6. Simulation Study

In this section, we demonstrate the inferential capabilities of the proposed model using simulation results. Here, we demonstrate parameter estimation, the inclination to recover baseline HRF shapes using standard error (SE), average bias (AB), mean square error (MSE), and relative bias (RB) to pick models that accurately reflect the underlying HRF shape, and the effect of censoring proportions on the model's inferential features.

6.1. Simulation Designs and Data Generation

Assuming the AFT regression model framework presented in Equation (11), we specifically simulated $n = 1000$ and 5000 data sets. Four variables were taken into account in the simulation study when we considered covariates. Two binary covariates, x_1 and x_2 , are produced using the Bernoulli (0.5) distribution, while two continuous covariates, x_3 and x_4 , were produced using the standard normal distribution. The covariate vector $\mathbf{x} = (x_1, x_2, x_3, x_4)'$ corresponds to the values for the AFT regression coefficients, which are selected to be $(-2, 0.75, -0.75, 0.5, -0.5)$. Using the inverse transform technique, the exponentiated Weibull (EW) distribution is used to simulate lifetime data from the AFT model framework [36].

In the regression equation, the effects of the covariates and the intercept are presumptive.

The PDF of the EW distribution takes the form

$$f(t; a, b, c) = abc(bt)^{a-1} (1 - e^{-(bt)^a})^{c-1} e^{-(bt)^a}, \quad t > 0, \quad (18)$$

where $a, b, c > 0$ are the distribution's parameters.

6.2. Simulation Algorithm for the Proposed AFT Model

The following are the steps for executing the proposed AFT model:

- (i) Set the parameters of the model's initial values,
- (ii) Utilize the inverse transform technique, create the lifetime data by inverting the CHRF of the proposed model,
- (iii) Utilize the various estimates to evaluate the estimations' values,
- (iv) Analyze the inferential properties of the estimates, taking into account the SE, AB, MSE, and RB.

The Inverse Transform Technique

A popular probabilistic approach for creating datasets using regression survival models is the inverse transform technique [37–40]. This technique is based on the association between the CHRF of a lifetime random variable and a standard uniform random variable. Whenever the CHRF of the baseline distribution has an explicit form solution, it may be used, reversed, and easily used in R.

The CDF is derived from the SF as follows:

$$F(t; x) = 1 - S(t; x). \quad (19)$$

Given this, when generating data, if Y is a random variable that has this CDF, then $U = F(Y)$ follows a uniform distribution throughout the range $[0; 1]$ and $[1 - U]$ also follows a uniform distribution $U[0, 1]$. At the end, for a realization u of U , by using the baseline CHRF $H_0(t; x)$, we get

$$1 - u = \exp\{-H_0(t; x)\}. \quad (20)$$

The inverse of the CHRF must only be calculated if the baseline HRF is strictly positive for every t to simulate lifetime data. An expression of the random live corresponding to the AFT model is as follows:

$$T = \frac{H_0^{-1}(-\log(1 - U))}{e^{x'\beta}}. \quad (21)$$

In this study, we used the EW baseline distribution to generate survival times that can accommodate all of the basic HRF shapes, including decreasing, constant, increasing, unimodal, and bathtub shapes. The EW distribution is likewise closed in the context of the AFT regression model [18].

As a last comment, we recall that the CHRF of the EW model is

$$H_0(t; a, b, c) = -\log\left[1 - \left(1 - e^{-(bt)^a}\right)^c\right]. \quad (22)$$

and thus the inverse of the CHRF is written as follows:

$$H_0^{-1}(t; a, b, c) = \frac{\left(-\log\left[1 - \left(1 - e^{-t}\right)^{1/c}\right]\right)^{1/a}}{b}. \quad (23)$$

6.3. Simulated Scenarios

Based on the shapes in Figure 2, we provide the results of four simulation scenarios based on non-monotone HRF (bathtub or unimodal), and monotone HRF (decreasing or increasing) to evaluate the performance of the ExEW-AFT model in comparison with the Weibull AFT (W-AFT), log-logistic AFT (LL-AFT), and EW-AFT models and to investigate the impact of the baseline HRF shape specification on the AFT model's inferential qualities.

Scenario 1: monotone (increasing) HRF :

The lifetime data for this scenario are created using the EW model, and the parameter for ($a = 1.0, b = 1.5$, and $c = 1.5$). The censoring times are generated from the exponential

distribution with rate parameter λ :

For $n = 1000$, $\lambda = 0.33$ and 0.2 .

For $n = 5000$, $\lambda = 0.4$ and 0.22

Scenario 2: monotone (decreasing) HRF:

The EW model was used to create the lifetime data for this scenario, and the parameter values for ($a = 0.80$, $b = 0.80$, and $c = 1.9$). The censoring times are generated from the exponential distribution with rate parameter λ :

For $n = 1000$, $\lambda = 0.38$ and 0.25

For $n = 5000$, $\lambda = 0.47$ and 0.25

Scenario 3: non-monotone (bathtub) HRF:

The EW model was used to create the lifetime data for this scenario, and the parameter values for ($a = 0.75$, $b = 1.50$, and $c = 2.0$). The censoring times are generated from the exponential distribution with rate parameter λ :

For $n = 1000$, $\lambda = 0.33$ and 0.2

For $n = 5000$, $\lambda = 0.65$ and 0.22

Scenario 4: non-monotone (unimodal) HRF:

The EW model was used to create the lifetime data for this scenario, and the parameter values for ($a = 1.65$, $b = 1.50$, and $c = 0.95$). The censoring times are generated from the exponential distribution with rate parameter λ :

For $n = 1000$, $\lambda = 0.29$ and 0.18

For $n = 5000$, $\lambda = 0.47$ and 0.21

Figure 2 shows the four Scenarios 1–4 (increasing (blue line), decreasing (red line), bathtub (green line), and unimodal (purple line) respectively), depending on the parameter values we chose, there were, on average, 20 and 30 percent censored observations.

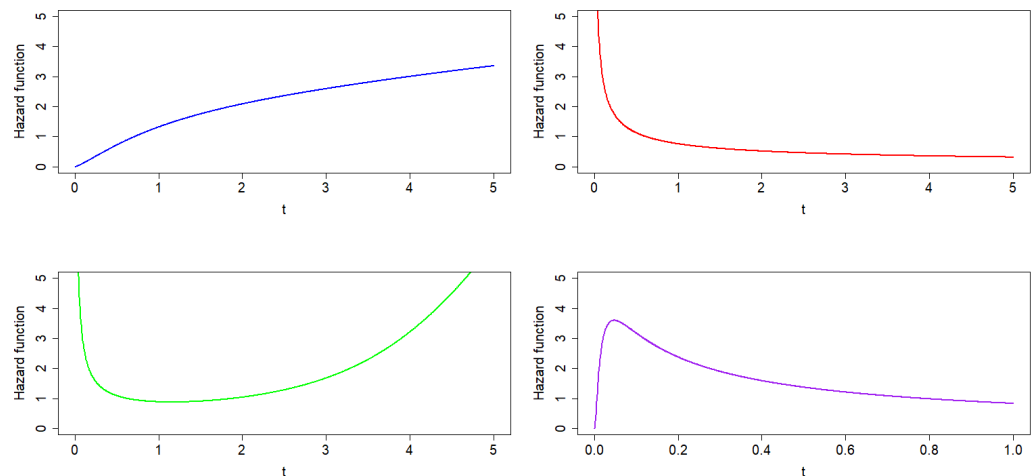


Figure 2. HR shapes for generated distribution for all scenarios.

6.4. Analyses of Simulated Data

To evaluate the inferential properties of the proposed models in all simulated scenarios, the ExEW-AFT model is fitted to the appropriate true generating model from the EW-AFT model. We also fitted the sub-models into each scenario. Furthermore, the estimates of the regression coefficients for each model are evaluated for stability based on the SE, RB, MSE, and the AB. These quantities are given by

$$AB = \frac{1}{n} \sum_{i=1}^n (\hat{v}_i - v), \quad (24)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{v}_i - v)^2 \quad (25)$$

and

$$RB = \frac{\frac{1}{n} \sum_{i=1}^n (\hat{v}_i - v)}{v}, \quad (26)$$

where v denotes each of the considered parameters.

Instead of examining the properties of the optimization process, our purpose was to examine the qualities of the estimates.

In all circumstances, we used the parameter values from the generating model as our optimization step starting points. The R programming language is used to do the analysis. With the aid of the R software “nlminb()”, the optimization stage is complete.

6.5. Performance Measures

The flexibility of the models for the covariates is evaluated in this study using measures such as the mean (estimated), AB, MSE, RB, and SE.

6.6. Simulation Results

According to the findings of scenario 1 in Tables 1–4 the SE, AB, MSE, and RB indicate that the proposed model performed better than others. Furthermore, it appears that sample size and censoring percentage have an effect on how well models match data. When censoring and sample size are increased, our proposed ExEW-AFT model often outperforms the W-AFT and LL-AFT models. As anticipated, all models equally integrated the growing HRF. However, our proposed model performs better in the case of heavy censoring. Theoretically, the findings of scenario 2 in Tables 5–8 show that all of the competing models can take into account the decreasing HRF shape. Our proposed ExEW-AFT model outperformed the W-AFT and LL-AFT models, and even the genuine produced model in terms of SE, AB, MSE, and RB. Moreover, when the censoring and sample size increase, our proposed model is once again the best-suited one and makes a wise choice of heavy censoring. The results of scenario 3 in Tables 9–12 reveal that the only model that has the lowest value in terms of SE, AB, MSE, and RB is our proposed ExEW-AFT model. Generally, the W-AFT and LL-AFT models generated the least accurate estimates for AB, MSE, and RB according to Scenario 3, as expected (i.e., bathtub hazard). The findings of scenario 4 in Tables 13–16 show that the proposed ExEW-AFT model produced estimates that had the lowest SE, AB, MSE, and RB values for all the regression coefficients while producing estimates that are equivalent to the genuine. Finally, the proposed AFT model outperforms the other competing models in all circumstances, including heavy censoring.

Table 1. Simulation study for scenario 1 ($n = 1000$) with about 20% censored observations is used to compare model performance.

Model	Parameter	(True Value)	Mean 20% Censoring	SE	AB	MSE	RB
ExEW-AFT	β_0	−2.0	−1.118	10.444	0.882	−2.749	−0.441
	β_1	0.75	1.144	0.057	0.394	0.746	0.525
	β_2	−0.75	−1.233	0.062	−0.483	0.958	0.644
	β_3	0.50	0.924	0.116	0.424	0.604	0.848
	β_4	−0.50	−0.789	0.116	−0.288	0.371	0.576
	a	1.0	−8.288				
	b	1.50	8.463				
	c	1.50	1.010				
	α	0.033	0.165				
EW	β_0	−2.0	−3.381	85.370	−0.880	4.293	0.440
	β_1	0.75	1.179	0.066	0.429	0.828	0.572
	β_2	−0.75	−1.265	0.069	−0.515	1.037	0.687
	β_3	0.50	0.952	0.127	0.452	0.657	0.904
	β_4	−0.50	−0.813	0.122	−0.313	0.411	0.626
	a	1.0	1.027				
	b	1.50	0.827				
	c	1.50	1.784				
W-AFT	β_0	−2.0	−2.288	44.235	−0.288	1.236	0.144
	β_1	0.75	1.206	0.060	0.456	0.891	0.608
	β_2	−0.75	−1.268	0.066	−0.518	1.046	0.691
	β_3	0.50	0.974	0.122	0.474	0.698	0.948
	β_4	−0.50	−0.815	0.119	−0.315	0.414	0.630
	a	1.0	1.288				
	b	1.50	0.685				
LL-AFT	β_0	−2.0	−2.288	2.785	−0.288	1.236	0.144
	β_1	0.75	1.206	0.066	0.456	0.891	0.608
	β_2	−0.75	−1.268	0.068	−0.518	1.046	0.691
	β_3	0.50	0.974	0.013	0.474	0.698	0.948
	β_4	−0.50	0.012	0.815	−0.315	0.414	0.630
	a	1.0	1.288				
	b	1.50	0.685				

Table 2. Simulation study for scenario 1 ($n = 5000$) with about 20% censored observations is used to compare model performance.

Model	Parameter	(True Value)	Mean 20% Censoring	SE	AB	MSE	RB
ExEW-AFT	β_0	−2.0	−1.356	1.507	0.644	−2.161	−0.322
	β_1	0.75	1.255	0.028	0.505	1.013	0.673
	β_2	−0.75	−1.297	0.029	−0.547	1.120	0.729
	β_3	0.50	0.948	0.054	0.448	0.649	0.896
	β_4	−0.50	−0.830	0.053	−0.330	0.439	0.660
	a	1.0	−6.156				
	b	1.50	6.649				
	c	1.50	1.026				
	α	0.033	0.083				
EW-AFT	β_0	−2.0	−3.288	11.628	−0.811	3.900	0.406
	β_1	0.75	1.283	0.031	0.533	1.083	0.711
	β_2	−0.75	−1.321	0.032	−0.571	1.181	0.761
	β_3	0.50	0.973	0.057	0.473	0.697	0.946
	β_4	−0.50	−0.852	0.056	−0.352	0.476	0.704
	a	1.0	1.034				
	b	1.50	0.744				
	c	1.50	2.078				
W-AFT	β_0	−2.0	−1.931	12.884	−1.224	6.394	0.612
	β_1	0.75	1.253	0.029	0.503	1.008	0.671
	β_2	−0.75	−1.299	0.030	−0.549	1.126	0.732
	β_3	0.50	0.937	0.055	0.437	0.628	0.874
	β_4	−0.50	−0.823	0.054	−0.323	0.428	0.646
	a	1.0	1.961				
	b	1.50	1.241				
LL-AFT	β_0	−2.0	−2.310	4.515	0.141	−0.543	−0.070
	β_1	0.75	1.292	0.029	0.503	1.009	0.671
	β_2	−0.75	−1.322	0.032	−0.557	1.146	0.743
	β_3	0.50	0.985	0.057	0.479	0.709	0.958
	β_4	−0.50	−0.866	0.057	−0.359	0.487	0.718
	a	1.0	1.310				
	b	1.50	0.688				

Table 3. Simulation study for scenario 1 ($n = 1000$) with about 30% censored observations is used to compare model performance.

Model	Parameter	(True Value)	Mean 30% Censoring	SE	AB	MSE	RB
ExEW-AFT	β_0	−2.0	−0.694	6.442	1.306	−3.518	−0.653
	β_1	0.75	1.053	0.043	0.303	0.546	0.404
	β_2	−0.75	−1.087	0.045	−0.337	0.619	0.449
	β_3	0.50	0.756	0.087	0.256	0.322	0.512
	β_4	−0.50	−0.821	0.089	−0.321	0.424	0.642
	a	1.0	−8.194				
	b	1.50	8.315				
	c	1.50	1.013				
	α	0.033	0.195				
EW-AFT	β_0	−2.0	−2.329	13.650	−0.296	1.271	0.148
	β_1	0.75	1.067	0.047	0.317	0.576	0.423
	β_2	−0.75	−1.099	0.048	−0.349	0.645	0.465
	β_3	0.50	0.746	0.090	0.246	0.306	0.492
	β_4	−0.50	−0.837	0.092	−0.337	0.451	0.644
	a	0.50	1.572				
	b	1.50	0.920				
	c	1.50	1.821				
W-AFT	β_0	−2.0	−1.121	11.686	0.879	−2.743	−0.440
	β_1	0.75	1.043	0.044	0.293	0.526	0.391
	β_2	−0.75	−1.087	0.045	−0.337	0.619	0.449
	β_3	0.50	0.730	0.088	0.230	0.283	0.460
	β_4	−0.50	−0.822	0.089	−0.322	0.426	0.644
	a	1.0	7.959				
	b	1.50	1.385				
LL-AFT	β_0	−2.0	−1.963	8.908	0.037	−0.148	−0.018
	β_1	0.75	1.083	0.047	0.333	0.610	0.444
	β_2	−0.75	−1.104	0.048	−0.354	0.656	0.472
	β_3	0.50	0.745	0.089	0.245	0.304	0.490
	β_4	−0.50	−0.841	0.090	−0.341	0.458	0.682
	a	1.0	0.963				
	b	1.50	0.601				

Table 4. Simulation study for scenario 1 ($n = 5000$) with about 30% censored observations is used to compare model performance.

Model	Parameter	(True Value)	Mean 30% Censoring	SE	AB	MSE	RB
ExEW-AFT	β_0	−2.0	−1.839	2.163	0.161	−0.618	−0.080
	β_1	0.75	1.130	0.018	.380	0.714	0.507
	β_2	−0.75	−1.149	0.017	−0.399	0.758	0.532
	β_3	0.50	0.828	0.039	0.328	0.436	0.656
	β_4	−0.50	−0.766	0.039	−0.266	0.337	0.532
	a	1.0	−4.313				
	b	1.50	5.871				
	c	1.50	1.105				
	α	0.033	0.093				
EW-AFT	β_0	−2.0	−2.495	6.116	−0.495	2.227	0.248
	β_1	0.75	1.142	0.022	0.392	0.742	0.523
	β_2	−0.75	−1.157	0.022	−0.407	0.776	0.543
	β_3	0.50	0.828	0.040	0.328	0.435	0.656
	β_4	−0.50	−0.771	0.039	−0.271	0.344	0.542
	a	1.0	1.277				
	b	1.50	0.862				
	c	1.50	2.032				
W-AFT	β_0	−2.0	−1.194	102.989	0.806	−2.575	−0.403
	β_1	0.75	1.119	0.023	0.369	0.690	0.492
	β_2	−0.75	−1.143	0.022	−0.393	0.745	0.524
	β_3	0.50	0.804	0.039	0.304	0.396	0.608
	β_4	−0.50	−0.749	0.041	−0.249	0.311	0.498
	a	1.0	7.893				
	b	1.50	1.377				
LL-AFT	β_0	−2.0	−2.698	2.097	−0.698	3.281	0.349
	β_1	0.75	1.144	0.022	0.394	0.746	0.525
	β_2	−0.75	−1.170	0.021	−0.420	0.807	0.560
	β_3	0.50	0.850	0.041	0.350	0.473	0.700
	β_4	−0.50	−0.750	0.042	−0.250	0.312	0.500
	a	1.0	0.297				
	b	1.50	0.574				

Table 5. Simulation study for scenario 2 ($n = 1000$) with about 20% censored observations is used to compare model performance.

Model	Parameter	(True Value)	Mean 20% Censoring	SE	AB	MSE	RB
ExEW-AFT	β_0	−2.0	−2.190	0.287	−0.190	0.798	0.095
	β_1	0.75	1.431	0.197	0.681	1.486	0.908
	β_2	−0.75	−1.447	0.171	−0.697	1.531	0.929
	β_3	0.50	1.104	0.169	0.605	0.970	1.210
	β_4	−0.50	−1.103	0.168	−0.603	0.968	1.206
	a	0.80	−0.977				
	b	0.80	1.427				
	c	1.90	0.701				
	α	0.09	0.114				
EW-AFT	β_0	−2.0	−2.529	0.275	−0.529	2.394	0.264
	β_1	0.75	1.433	0.285	0.683	1.491	0.911
	β_2	−0.75	−1.445	0.286	−0.695	1.526	0.927
	β_3	0.50	1.126	0.207	0.626	1.018	1.252
	β_4	−0.50	−1.114	0.180	−0.614	0.990	1.228
	a	0.80	0.816				
	b	0.80	0.263				
	c	1.90	3.846				
W-AFT	β_0	−2.0	−2.864	0.288	−0.864	4.202	0.432
	β_1	0.75	1.450	0.323	0.700	1.540	0.933
	β_2	−0.75	−1.454	0.319	−0.704	1.552	0.939
	β_3	0.50	1.135	0.222	0.635	1.039	1.270
	β_4	−0.50	−1.110	0.181	−0.610	0.982	1.220
	a	0.80	1.664				
	b	0.80	1.368				
LL-AFT	β_0	−2.0	−2.864	0.278	−0.864	4.201	0.432
	β_1	0.75	1.450	0.288	0.700	1.539	0.933
	β_2	−0.75	−1.454	0.290	−0.704	1.551	0.939
	β_3	0.50	1.135	0.210	0.635	1.038	1.270
	β_4	−0.50	−1.110	0.180	−0.610	0.982	1.220
	a	0.80	1.663				
	b	0.80	1.368				

Table 6. Simulation study for scenario 2 ($n = 5000$) with about 20% censored observations is used to compare model performance.

Model	Parameter	(True Value)	Mean 20% Censoring	SE	AB	MSE	RB
ExEW-AFT	β_0	−2.0	−2.190	0.279	−0.190	0.798	0.095
	β_1	0.75	1.431	0.111	0.681	1.486	0.908
	β_2	−0.75	−1.447	0.109	−0.697	1.531	0.929
	β_3	0.50	1.105	0.167	0.605	0.970	1.210
	β_4	−0.50	−1.103	0.170	−0.603	0.968	1.206
	a	0.80	−0.087				
	b	0.80	0.719				
	c	1.90	0.743				
	α	0.09	0.074				
EW-AFT	β_0	−2.0	−2.529	4.844	−0.529	2.394	0.264
	β_1	0.75	1.433	0.123	0.683	1.491	0.911
	β_2	−0.75	−1.445	0.125	−0.695	1.526	0.927
	β_3	0.50	1.126	0.235	0.626	1.018	1.252
	β_4	−0.50	−1.114	0.235	−0.614	0.990	1.228
	a	0.80	0.816				
	b	0.80	0.263				
	c	1.90	3.846				
W-AFT	β_0	−2.0	−5.213	49.582	−3.213	23.173	1.606
	β_1	0.75	1.417	0.114	0.667	1.446	0.889
	β_2	−0.75	−1.449	0.117	−0.699	1.537	0.932
	β_3	0.50	1.077	0.227	0.577	0.911	1.154
	β_4	−0.50	−1.101	0.228	−0.601	0.963	1.202
	a	0.80	0.823				
	b	0.80	0.640				
LL-AFT	β_0	−2.0	−2.864	0.124	−0.864	4.201	0.432
	β_1	0.75	1.450	0.129	0.700	1.539	0.933
	β_2	−0.75	−1.454	0.130	−0.704	1.551	0.939
	β_3	0.50	1.135	0.178	0.635	1.038	1.270
	β_4	−0.50	−1.110	0.175	−0.610	0.982	1.220
	a	0.80	1.664				
	b	0.80	1.368				

Table 7. Simulation study for scenario 2 ($n = 1000$) with about 30% censored observations is used to compare model performance.

Model	Parameter	(True Value)	Mean 30% Censoring	SE	AB	MSE	RB
ExEW-AFT	β_0	−2.0	−2.029	0.522	−0.029	0.118	0.015
	β_1	0.75	1.212	0.076	0.462	0.907	0.616
	β_2	−0.75	−1.187	0.052	−0.437	0.848	0.583
	β_3	0.50	0.822	0.156	0.322	0.426	0.644
	β_4	−0.50	−0.930	0.158	−0.430	0.614	0.860
	a	0.80	0.869				
	b	0.80	−0.556				
	c	1.90	1.082				
	α	0.09	0.583				
EW-AFT	β_0	−2.0	−0.422	17.156	1.577	−3.821	−0.788
	β_1	0.75	1.254	0.098	0.504	1.011	0.672
	β_2	−0.75	−1.375	0.101	−0.625	1.328	0.833
	β_3	0.50	1.056	0.189	0.556	0.866	1.112
	β_4	−0.50	−1.185	0.192	−0.685	1.154	1.370
	a	0.80	1.301				
	b	0.80	0.244				
	c	1.90	5.801				
W-AFT	β_0	−2.0	−4.290	75.287	−2.290	14.406	1.145
	β_1	0.75	1.247	0.089	0.497	0.992	0.663
	β_2	−0.75	−1.389	0.094	−0.639	1.366	0.852
	β_3	0.50	1.038	0.183	0.538	0.828	1.076
	β_4	−0.50	−1.176	0.186	−0.676	1.132	1.352
	a	0.80	1.099				
	b	0.80	0.699				
LL-AFT	β_0	−2.0	−2.524	59.316	−0.524	2.371	0.262
	β_1	0.75	1.278	0.095	0.528	1.071	0.704
	β_2	−0.75	−1.386	0.053	−0.636	1.359	0.848
	β_3	0.50	1.056	0.185	0.556	0.866	1.112
	β_4	−0.50	−1.184	0.184	−0.684	1.152	1.368
	a	0.80	1.324				
	b	0.80	1.210				

Table 8. Simulation study for scenario 2 ($n = 5000$) with about 30% censored observations is used to compare model performance.

Model	Parameter	(True Value)	Mean 30% Censoring	SE	AB	MSE	RB
ExEW-AFT	β_0	−2.0	−3.402	0.169	−1.402	7.573	0.701
	β_1	0.75	1.106	0.011	0.356	0.661	0.475
	β_2	−0.75	−1.159	0.021	−0.409	0.781	0.545
	β_3	0.50	0.888	0.040	0.388	0.539	0.776
	β_4	−0.50	−0.768	0.037	−0.268	0.340	0.536
	a	0.80	−0.391				
	b	0.80	1.464				
	c	1.90	1.234				
	α	0.09	1.067				
EW-AFT	β_0	−2.0	−1.541	7.520	0.459	−1.624	−0.2304
	β_1	0.75	1.142	0.022	0.392	0.742	0.523
	β_2	−0.75	−1.157	0.022	−0.407	0.776	0.543
	β_3	0.50	0.828	0.040	0.328	0.435	0.656
	β_4	−0.50	−0.771	0.040	−0.271	0.344	0.542
	a	0.80	3.314				
	b	0.80	0.862				
	c	1.90	2.032				
W-AFT	β_0	−2.0	−1.795	117.420	0.205	−0.776	−0.102
	β_1	0.75	1.119	0.020	0.369	0.690	0.492
	β_2	−0.75	−1.143	0.020	−0.393	0.745	0.524
	β_3	0.50	0.804	0.038	0.304	0.396	0.608
	β_4	−0.50	−0.749	0.038	−0.249	0.311	0.498
	a	0.80	4.324				
	b	0.80	1.377				
LL-AFT	β_0	−2.0	2.086	21.810	−0.086	0.353	0.043
	β_1	0.75	1.153	0.022	0.403	0.767	0.537
	β_2	−0.75	−1.158	0.021	−0.408	0.777	0.544
	β_3	0.50	0.836	0.040	0.336	0.449	0.672
	β_4	−0.50	−0.783	0.039	−0.283	0.363	0.566
	a	0.80	0.886				
	b	0.80	0.594				

Table 9. Simulation study for scenario 3 ($n = 1000$) with about 20% censored observations is used to compare model performance.

Model	Parameter	(True Value)	Mean 20% Censoring	SE	AB	MSE	RB
ExEW-AFT	β_0	−2.0	−2.204	0.762	−0.204	0.857	0.102
	β_1	0.75	1.114	0.046	0.364	0.678	0.485
	β_2	−0.75	−1.170	0.045	−0.420	0.807	0.560
	β_3	0.50	0.786	0.092	0.286	0.368	0.572
	β_4	−0.50	−0.707	0.091	−0.207	0.249	0.414
	a	0.50	−0.925				
	b	1.50	1.351				
	c	2.0	1.171				
	α	1.0	0.449				
EW-AFT	β_0	−2.0	−1.580	8.413	0.420	−1.503	−0.210
	β_1	0.75	1.116	0.051	0.366	0.683	0.488
	β_2	−0.75	−1.150	0.052	−0.400	0.760	0.533
	β_3	0.50	0.821	0.095	0.321	0.423	0.642
	β_4	−0.50	−0.770	0.093	−0.270	0.343	0.540
	a	0.50	2.788				
	b	1.50	0.849				
	c	2.00	2.572				
W-AFT	β_0	−2.0	−1.769	38.554	0.231	−0.871	−0.116
	β_1	0.75	1.071	0.046	0.321	0.584	0.428
	β_2	−0.75	−1.134	0.049	−0.384	0.723	0.512
	β_3	0.50	0.777	0.094	0.277	0.353	0.554
	β_4	−0.50	−0.750	0.093	−0.250	0.313	0.500
	a	0.50	4.3492				
	b	1.50	1.588				
LL-AFT	β_0	−2.0	−2.263	69.060	−0.263	1.123	0.132
	β_1	0.75	1.133	0.094	0.383	0.722	0.511
	β_2	−0.75	−1.151	0.048	−0.401	0.763	0.535
	β_3	0.50	0.830	0.093	0.330	0.440	0.660
	β_4	−0.50	−0.777	0.092	−0.277	0.354	0.554
	a	0.50	0.763				
	b	1.50	0.524				

Table 10. Simulation study for scenario 3 ($n = 5000$) with about 20% censored observations is used to compare model performance.

Model	Parameter	(True Value)	Mean 20% Censoring	SE	AB	MSE	RB
ExEW-AFT	β_0	−2.0	−2.510	0.172	−0.510	2.300	0.255
	β_1	0.75	1.137	0.023	0.387	0.731	0.516
	β_2	−0.75	−1.231	0.022	−0.481	0.953	0.641
	β_3	0.50	0.863	0.041	0.363	0.495	0.726
	β_4	−0.50	−0.733	0.040	−0.233	0.288	0.466
	a	0.50	−0.485				
	b	1.50	2.241				
	c	2.00	1.414				
	α	1.00	0.249				
EW-AFT	β_0	−2.0	−1.550	9.767	0.450	−1.597	−0.225
	β_1	0.75	1.205	0.026	0.455	0.888	0.607
	β_2	−0.75	−1.229	0.026	−0.479	0.949	0.639
	β_3	0.50	0.906	0.046	0.406	0.571	0.812
	β_4	−0.50	−0.793	0.045	−0.293	0.378	0.586
	a	0.50	2.105				
	b	1.50	0.788				
	c	2.00	2.763				
W-AFT	β_0	−2.0	−0.755	9.161	1.245	−3.429	−0.623
	β_1	0.75	1.172	0.023	0.422	0.810	0.563
	β_2	−0.75	−1.207	0.024	−0.457	0.894	0.609
	β_3	0.50	0.868	0.044	0.368	0.503	0.736
	β_4	−0.50	−0.765	0.043	−0.265	0.336	0.530
	a	0.50	9.709				
	b	1.50	1.541				
LL-AFT	β_0	−2.0	−2.145	29.658	−0.145	0.601	0.072
	β_1	0.75	1.210	0.026	0.460	0.901	0.613
	β_2	−0.75	−1.229	0.029	−0.479	0.947	0.639
	β_3	0.50	0.909	0.047	0.409	0.577	0.818
	β_4	−0.50	−0.800	0.045	−0.300	0.390	0.600
	a	0.50	0.645				
	b	1.50	0.542				

Table 11. Simulation study for scenario 3 ($n = 1000$) with about 30% censored observations is used to compare model performance.

Model	Parameter	(True Value)	Mean 30% Censoring	SE	AB	MSE	RB
ExEW-AFT	β_0	−2.0	−2.273	0.381	−0.273	1.165	0.136
	β_1	0.75	0.906	0.019	0.156	0.258	0.208
	β_2	−0.75	−0.970	0.023	−0.220	0.378	0.293
	β_3	0.50	0.645	0.035	0.145	0.167	0.290
	β_4	−0.50	−0.621	0.036	−0.121	0.135	0.242
	a	0.50	−0.165				
	b	1.50	0.953				
	c	2.00	2.311				
	α	1.00	0.851				
EW-AFT	β_0	−2.0	−2.074	4.228	−0.074	0.303	0.037
	β_1	0.75	0.902	0.020	0.152	0.252	0.203
	β_2	−0.75	−0.917	0.0207	−0.167	0.279	0.223
	β_3	0.50	0.607	0.038	0.107	0.119	0.214
	β_4	−0.50	−0.636	0.038	−0.136	0.154	0.272
	a	0.50	0.993				
	b	1.50	1.886				
	c	2.00	2.253				
W-AFT	β_0	−2.0	−1.182	7.313	0.818	−2.603	−0.409
	β_1	0.75	0.884	0.018	0.134	0.220	0.179
	β_2	−0.75	−0.906	0.019	−0.156	0.258	0.208
	β_3	0.50	0.590	0.037	0.090	0.099	0.180
	β_4	−0.50	−0.629	0.037	−0.129	0.146	0.258
	a	0.50	3.165				
	b	1.50	3.148				
LL-AFT	β_0	−2.0	−1.857	14.83	0.1437	−0.551	−0.072
	β_1	0.75	0.911	0.020	0.161	0.268	0.215
	β_2	−0.75	−0.923	0.021	−0.173	0.289	0.231
	β_3	0.50	0.611	0.038	0.111	0.124	0.222
	β_4	−0.50	−0.634	0.039	−0.134	0.152	0.268
	a	0.50	0.357				
	b	1.50	0.252				

Table 12. Simulation study for scenario 3 ($n = 5000$) with about 30% censored observations is used to compare model performance.

Model	Parameter	(True Value)	Mean 30% Censoring	SE	AB	MSE	RB
ExEW-AFT	β_0	−2.0	−2.233	0.628	−0.233	0.987	0.116
	β_1	0.75	0.999	0.013	0.249	0.436	0.332
	β_2	−0.75	−0.987	0.012	−0.237	0.412	0.316
	β_3	0.50	0.729	0.019	0.228	0.280	0.456
	a	0.50	−0.051				
	b	1.50	0.919				
	c	2.00	2.192				
	α	1.00	0.302				
EW-AFT	β_0	−2.0	−1.285	7.195	0.715	−2.350	−0.358
	β_1	0.75	0.994	0.011	0.244	0.426	0.325
	β_2	−0.75	−1.006	0.013	−0.256	0.450	0.341
	β_3	0.50	0.690	0.024	0.190	0.226	0.380
	β_4	−0.50	−0.664	0.024	−0.164	0.191	0.328
	a	0.50	3.459				
	b	1.50	1.625				
	c	2.00	1.703				
W-AFT	β_0	−2.0	−1.191	26.590	0.809	−2.581	−0.404
	β_1	0.75	0.980	0.012	0.230	0.397	0.307
	β_2	−0.75	−0.997	0.012	−0.247	0.431	0.329
	β_3	0.50	0.676	0.023	0.176	0.207	0.352
	β_4	−0.50	−0.652	0.023	−0.152	0.175	0.304
	a	0.50	4.701				
	b	1.50	2.325				
LL-AFT	β_0	−2.0	−2.038	13.067	−0.038	0.153	0.019
	β_1	0.75	1.007	13.202	0.257	0.452	0.343
	β_2	−0.75	−1.012	0.015	−0.262	0.461	0.349
	β_3	0.50	0.704	0.024	0.204	0.246	0.408
	β_4	−0.50	−0.678	0.024	−0.177	0.208	0.354
	a	0.50	0.539				
	b	1.50	0.350				

Table 13. Simulation study for scenario 4 ($n = 1000$) with about 20% censored observations is used to compare model performance.

Model	Parameter	(True Value)	Mean 20% Censoring	SE	AB	MSE	RB
ExEW-AFT	β_0	−2.0	−1.704	2.801	0.296	−1.097	−0.148
	β_1	0.75	1.310	0.092	0.560	1.154	0.747
	β_2	−0.75	−1.410	0.093	−0.660	1.425	0.880
	β_3	0.50	1.115	0.174	0.615	0.993	1.230
	β_4	−0.50	−1.019	0.176	−0.519	0.789	1.038
	a	1.650	3.176				
	b	1.500	−2.714				
	c	0.95	1.021				
	α	0.90	0.178				
EW-AFT	β_0	−2.0	−3.877	33.412	−1.877	11.033	0.939
	β_1	0.75	1.303	0.093	0.553	1.136	0.737
	β_2	−0.75	−1.413	0.096	−0.663	1.433	0.884
	β_3	0.50	1.114	0.181	0.614	0.992	1.228
	β_4	−0.50	−1.013	0.178	−0.513	0.776	1.026
	a	1.65	1.851				
	b	1.50	0.730				
	c	0.95	1.205				
W-AFT	β_0	−2.0	−5.793	14.192	−3.793	29.562	1.897
	β_1	0.75	1.289	0.097	0.539	1.100	0.719
	β_2	−0.75	−1.405	0.094	−0.655	1.412	0.873
	β_3	0.50	1.101	0.178	0.601	0.962	1.202
	β_4	−0.50	−1.004	0.176	−0.504	0.758	1.008
	a	1.65	0.318				
	b	1.50	0.844				
LL-AFT	β_0	−2.0	−2.328	38.052	−0.328	1.419	0.164
	β_1	0.75	1.343	0.095	0.593	1.242	0.791
	β_2	−0.75	−1.419	0.96	−0.669	1.452	0.892
	β_3	0.50	1.144	0.181	0.644	1.059	1.288
	β_4	−0.50	−1.022	0.183	−0.522	0.794	1.044
	a	1.65	1.978				
	b	1.50	1.046				

Table 14. Simulation study for scenario 4 ($n = 5000$) with about 20% censored observations is used to compare model performance.

Model	Parameter	(True Value)	Mean 20% Censoring	SE	AB	MSE	RB
ExEW-AFT	β_0	−2.0	−3.163	3.356	−1.163	6.002	0.582
	β_1	0.75	1.576	0.054	0.826	1.920	1.101
	β_2	−0.75	−1.633	0.055	−0.883	2.103	1.177
	β_3	0.50	1.208	0.102	0.708	1.210	1.416
	β_4	−0.50	−1.095	0.101	−0.595	0.950	1.190
	a	1.650	−0.715				
	b	1.500	3.136				
	c	0.95	0.744				
	α	0.90	0.107				
EW-AFT	β_0	−2.0	−0.047	6.920	1.953	−3.998	−0.977
	β_1	0.75	1.591	0.057	0.841	1.968	1.121
	β_2	−0.75	−1.644	0.058	−0.894	2.142	1.192
	β_3	0.50	1.270	0.105	0.770	1.363	1.540
	β_4	−0.50	−1.139	0.105	−0.639	1.047	1.278
	a	1.65	7.015				
	b	1.50	0.245				
	c	0.95	4.711				
W-AFT	β_0	−2.0	−3.650	16.810	−1.650	9.324	0.825
	β_1	0.75	1.572	0.058	0.822	1.909	1.096
	β_2	−0.75	−1.639	0.057	−0.889	2.124	1.185
	β_3	0.50	1.191	0.104	0.691	1.169	1.382
	β_4	−0.50	−1.084	0.103	−0.584	0.925	1.168
	a	1.65	5.055				
	b	1.50	0.655				
LL-AFT	β_0	−2.0	−2.540	96.820	−0.540	2.451	0.270
	β_1	0.75	1.600	0.056	0.850	1.998	1.133
	β_2	−0.75	−1.646	0.057	−0.896	2.147	1.195
	β_3	0.50	1.257	0.106	0.757	1.329	1.514
	β_4	−0.50	−1.133	0.105	−0.633	1.034	1.266
	a	1.65	2.190				
	b	1.50	1.316				

Table 15. Simulation study for Scenario 4 ($n = 1000$) with about 30% censored observations is used to compare model performance.

Model	Parameter	(True Value)	Mean 30% Censoring	SE	AB	MSE	RB
ExEW-AFT	β_0	−2.0	−2.197	3.087	−0.197	0.828	0.098
	β_1	0.75	1.144	0.066	0.394	0.747	0.525
	β_2	−0.75	−1.235	0.069	−0.485	0.963	0.647
	β_3	0.50	0.887	0.132	0.387	0.538	0.774
	β_4	−0.50	−0.931	0.133	−0.431	0.618	0.862
	a	1.650	−5.965				
	b	1.500	6.806				
	c	0.95	0.989				
	α	0.9	0.260				
EW-AFT	β_0	−2.0	−3.049	25.049	−1.049	5.294	0.524
	β_1	0.75	1.151	0.067	0.401	0.763	0.535
	β_2	−0.75	−1.237	0.069	−0.487	0.967	0.649
	β_3	0.50	0.894	0.133	0.394	0.549	0.788
	β_4	−0.50	−0.940	0.134	−0.440	0.634	0.880
	a	1.65	1.648				
	b	1.50	0.724				
	c	0.95	1.397				
W-AFT	β_0	−2.0	−1.975	33.399	0.025	−0.101	−0.013
	β_1	0.75	1.137	0.069	0.387	0.731	0.516
	β_2	−0.75	−1.229	0.070	−0.479	0.947	0.639
	β_3	0.50	0.880	0.138	0.380	0.524	0.760
	β_4	−0.50	−0.924	0.140	−0.424	0.603	0.848
	a	1.65	6.545				
	b	1.50	0.922				
LL-AFT	β_0	−2.0	−1.909	18.106	0.091	−0.356	−0.046
	β_1	0.75	1.169	0.071	0.419	0.803	0.559
	β_2	−0.75	−1.234	0.072	−0.484	0.959	0.645
	β_3	0.50	0.900	0.141	0.400	0.561	0.800
	β_4	−0.50	−0.961	0.142	−0.461	0.674	0.922
	a	1.65	1.559				
	b	1.50	0.928				

Table 16. Simulation study for scenario 4 ($n = 5000$) with about 30% censored observations is used to compare model performance.

Model	Parameter	(True Value)	Mean 30% Censoring	SE	AB	MSE	RB
ExEW-AFT	β_0	−2.0	−3.159	0.409	−1.159	5.982	0.580
	β_1	0.75	1.279	0.029	0.529	1.073	0.705
	β_2	−0.75	−1.279	0.030	−0.529	1.074	0.705
	β_3	0.50	0.925	0.056	0.425	0.605	0.850
	β_4	−0.50	−0.873	0.056	−0.373	0.513	0.746
	a	1.650	2.843				
	b	1.500	1.372				
	c	0.95	0.835				
	α	0.9	0.114				
EW-AFT	β_0	−2.0	−3.107	10.417	−1.107	5.654	0.553
	β_1	0.75	1.297	0.032	0.547	1.119	0.729
	β_2	−0.75	−1.313	0.032	−0.563	1.161	0.751
	β_3	0.50	0.950	0.060	0.450	0.652	0.900
	β_4	−0.50	−0.870	0.059	−0.370	0.506	0.740
	a	1.65	1.716				
	b	1.50	0.715				
	c	0.95	1.432				
W-AFT	β_0	−2.0	−2.209	0.059	−0.209	0.878	0.104
	β_1	0.75	1.280	0.030	0.530	1.076	0.707
	β_2	−0.75	−1.303	0.031	−0.553	1.134	0.737
	β_3	0.50	0.931	0.058	0.431	0.618	0.862
	β_4	−0.50	−0.855	0.057	−0.355	0.481	0.710
	a	1.65	5.863				
	b	1.50	0.924				
LL-AFT	β_0	−2.0	−1.970	26.224	0.030	−0.119	−0.015
	β_1	0.75	1.313	0.032	0.563	1.162	0.751
	β_2	−0.75	−1.316	0.033	−0.566	1.170	0.755
	β_3	0.50	0.971	0.061	0.471	0.693	0.942
	β_4	−0.50	−0.888	0.061	−0.388	0.539	0.776
	a	1.65	1.620				
	b	1.50	0.915				

7. Application to COVID-19 Data

In this Section, we demonstrate the adaptability and utility of the ExEW-AFT model, considering real-world right-censored COVID-19 data from Sudan.

7.1. Sudan COVID-19 Data

COVID-19 is an infectious illness. Many studies have been conducted since it was declared a global health emergency to better understand the disease's clinical, epidemiological, and prognostic aspects [41–45].

In Sudan, the epidemiological data are disclosed by the epidemiology department of the federal ministry of health (FMH) www.fmoh.gov.sd (accessed on 25 May 2022). Therefore, according to the investigation of FMH, there are 35,321 patients who are infected with the virus. Moreover, every positive case between 13 March 2020 and 31 December 2021 is included in the study sample. The period from the date of admission until the date of the sample result is considered the length of the hospital stay.

Overview of Covariates of COVID-19 Data

For each patient ($i = 1, \dots, 35,321$), the following covariates are taken into account.

- y : The length of stay in the hospital (by days).
- Status: For censoring.
- x_1 : Age.
- x_2 : Sex group.
- x_3 : The comorbidity group.

Table 17 shows some descriptive statistics of these covariates. The average hospital stay lasts three days. The scaled total time on test (TTT) plot for the COVID-19 hazard rate shape is displayed in Figure 3, which shows that the hazard rate shape of COVID-19 is unimodal. The initial density shape of the length of stay in the hospital is reported using the non-parametric kernel density estimation (KDE) approach in Figure 4, beside the histogram. It is noted that the density is asymmetrical and positively skewed, which is a common feature for survival data and makes the normal distribution inappropriate to analyze [46]. This is one of the points that motivated us to use the ExEW distribution as a baseline hazard in the AFT model to fit this data. The box plot and violin plot in Figure 4 are used to identify the extremes, and they reveal that some of these extreme observations are recorded.

Table 17. Statistical summary of the covariates for COVID-19 data.

Covariate		Number of Observations (%)	Mean (Standard Deviation (sd))
Length of stay in the hospital (in days)		-	3 (6.227)
Age (in years)		-	45 (18.750)
Status	1	32,731 (92.7%)	-
	0	2590 (7.3%)	-
Comorbidity	yes	845 (2.4%)	-
	no	34,476 (97.6%)	-
Sex	male	20,654 (58.5%)	-
	female	14,667 (41.5%)	-

Regarding the censoring status (0 is censored, 1 is observational lifetime), there is 7.3% censored in the sample of the study. Also, age as an explanatory variable refers to the patient's age at diagnosis. Age is recorded as a continuous variable. Table 17 shows that the average age in the COVID-19 data set is 45 years, and the standard deviation is 6.226. Moreover, sex describes the gender of patients, which is a categorical variable with only two levels. Males are assigned the value 1, while females are assigned the value 2. There are 20,654 (58%) males and 14,667 (42%) females in the data.

It is recorded that COVID-19 is more common in men as compared to women in Sudan. Furthermore, the comorbidity status is a two-level categorical variable, with value 1 referring to a patient who has a chronic disease and value 2 referring to a patient who does not have a chronic disease. There are 845 (2.4%) patients who have a chronic disease in Sudan.

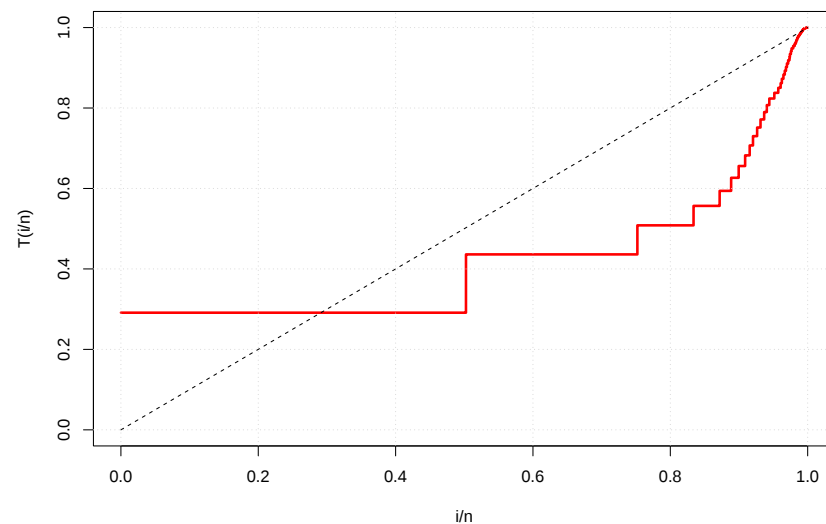


Figure 3. TTT plot for the Sudan COVID-19 data.

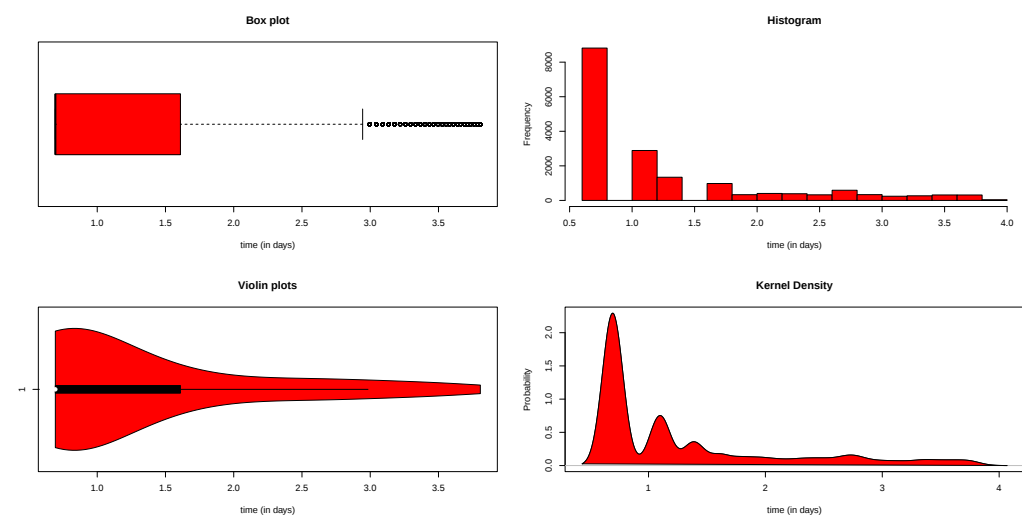


Figure 4. Some non-parametric plots for the Sudan COVID-19 data.

We analyze and compare the fitting of the ExEW-AFT model with that of sub-models such as the W-AFT, the exponentiated exponential (EE) AFT (EE-AFT), and the exponential Weibull AFT (ExW-AFT) models.

The AFT models for the competing models are as follows:

1. The W-AFT model:

$$f_{W-AFT}(t; \beta, \mathbf{x}) = ac \left(te^{\mathbf{x}'\beta} \right)^{a-1} \exp \left\{ -c \left(te^{\mathbf{x}'\beta} \right)^a \right\} e^{\mathbf{x}'\beta}, t > 0. \quad (27)$$

2. The EE-AFT model:

$$f_{EE-AFT}(t; \beta, \mathbf{x}) = ac \left[\exp \left\{ -c \left(te^{\mathbf{x}'\beta} \right) \right\} \right] \left[1 - \exp \left\{ -c \left(te^{\mathbf{x}'\beta} \right) \right\} \right]^{a-1} e^{\mathbf{x}'\beta}, t > 0. \quad (28)$$

3. The ExW-AFT model:

$$f_{\text{ExW-AFT}}(t; \boldsymbol{\beta}, \mathbf{x}) = \left[a + bc \left(te^{\mathbf{x}'\boldsymbol{\beta}} \right)^{c-1} \right] \exp \left\{ -a \left(te^{\mathbf{x}'\boldsymbol{\beta}} \right) + b \left(te^{\mathbf{x}'\boldsymbol{\beta}} \right)^c \right\} e^{\mathbf{x}'\boldsymbol{\beta}}, t > 0. \quad (29)$$

The aforementioned Equations (27)–(29) demonstrate how the covariates act multiplicatively on time, causing an acceleration or a deceleration of time. The analytical measurements, such as the Bayesian information criterion (BIC), and the consistent Akaike information criterion (CAIC) were used to decide which AFT model matches the COVID-19 data the best. Additionally, goodness-of-fit metrics like the log-likelihood ratio test are used. The BIC is given by

$$BIC = k \log(n) - 2\ell \quad (30)$$

and the CAIC is provided by

$$CAIC = \frac{2nk}{n - k - 1} - 2\ell, \quad (31)$$

where ℓ refers to the log-likelihood function calculated at the MLEs, k for the number of model parameters, and n for the sample size.

7.2. Cox PH Model

To ascertain the relationship between survival time and the covariates thought to affect survival time, the Cox PH model is conducted. The Cox PH model parameters are estimated.

Table 18 shows the regression analysis of the Cox-PH model, including regression coefficients, SE, p -value, likelihood ratio test (LRT), and BIC values. Furthermore, all of the covariates (age, sex, and comorbidity) significantly affect the length of stay in the hospital at a 5% level of significance.

Table 18. Results of Cox-PH model including the coefficients, SE, p -value, LRT, BIC, and CAIC.

Covariates	Coefficients	SE	p -Value
Age	−0.005	0.000	<0.022
Sex	−0.103	0.011	<0.022
Comorbidity	0.792	0.044	<0.022
LRT	909.600		<0.000
BIC	621,599		

Testing PH Assumption

The Schoenfeld residual is used in this study to test the PH assumption as follows: Based on the test, the Schoenfeld residuals are obtained for age and sex as covariates. The results in Table 19 provide evidence of the rejection of the assumption of PH for all covariates considered in the COVID-19 data. In other words, the PH models present an inadequate fit for this COVID-19 data.

Table 19. Chi-square (χ^2) test and p -value for Schoenfeld residual test at level of significance 1%.

Covariate	χ^2	p -Value
Age	65.75	<−0.000
Sex	84.88	<−0.000
Comorbidity	2.07	<−0.15
Global	171.46	<0.000

7.3. AFT Model Analysis

In this subsection, we present the analysis of the ExEW-AFT, W-AFT, EE-AFT, and ExW-AFT models using the Sudan COVID-19 data.

We calculated the LRT statistics for the three sub-models. According to the LRT statistics in Table 20, the ExEW-AFT model fits the COVID-19 data the best. Table 21 showed that the SE of $\hat{\beta}$ s for the ExEW-AFT, W-AFT, EE-AFT, and ExW-AFT models is small enough. Moreover, at 5% significance level, the parameters of all AFT models are significant, as shown in Table 22. Furthermore, the analytical measures of competing AFT models are shown in Table 23 which reveals that the ExEW-AFT model has the lowest BIC, and CAIC. In conclusion, the ExEW-AFT model is the best fit for the data among the W-AFT, EE-AFT, and ExW-AFT models under consideration. Moreover, at 5% level of significance, all the covariates (age, sex, and comorbidity) significantly influence the length of stay at the hospital.

Figure 5 shows the Kaplan–Meier (KM) survival curve, which indicates that the difference in survival times between the comorbidity group and the sexual group is statistically significant (p -value < 0.0001). Figure 6 depicts estimated HRFs for the competitive baseline hazards, and Figure 7 represents the average KM estimator and population SF. All of these figures show that our proposed ExEW-AFT model fits the data better than its competitors, including the Cox-PH and its sub-models.

Table 20. The LRT statistics for COVID-19 data at significance level 1%.

Model	Hypothesis	LRT	p -Value
W-AFT	$H_0 : a = 0$, and $\alpha = 1$ vs $H_1 : H_0$ is false	1987.469	<0.000
EE-AFT	$H_0 : b = 0$ vs $H_1 : H_0$ is false	2197.355	<0.000
ExW-AFT	$H_0 : \alpha = 1$ vs $H_1 : H_0$ is false	1544.978	<0.000

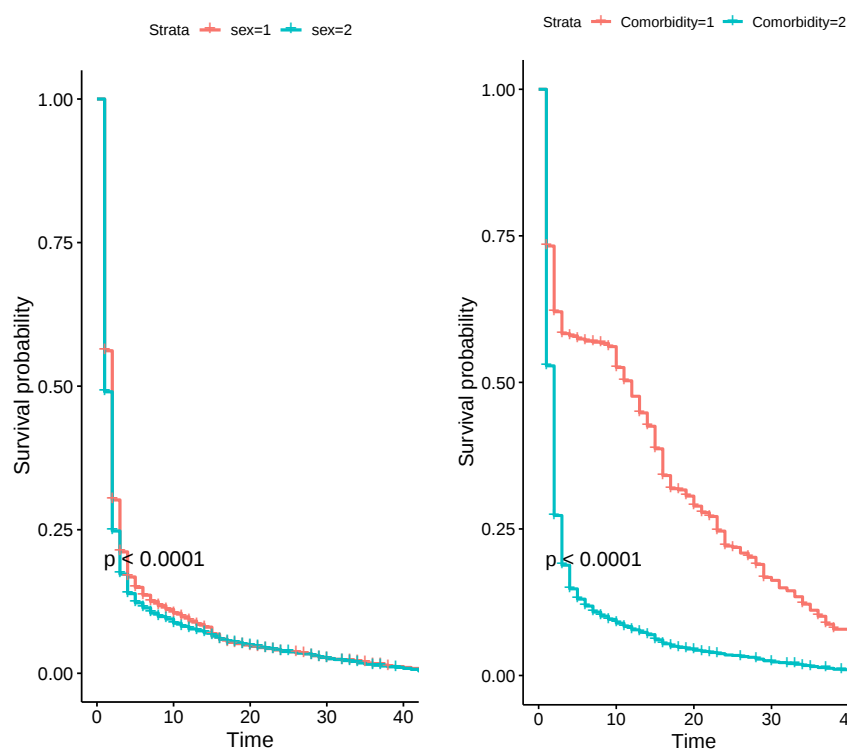


Figure 5. The KM curves for time.

Table 21. MLE fits of the ExEW, W, EW, and ExW AFT models with SE (in parentheses) for COVID-19 data.

Model	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\hat{\alpha}$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$
ExEW-AFT	0.469 (0.029)	0.568 (0.029)	0.977 (0.001)	1.516 (0.037)	−0.201 (0.007)	0.017 (0.014)	0.506 (0.021)
W-AFT	3.772 (0.615)	0.869 (0.003)	- -	- -	−0.147 (0.005)	−0.001 (0.009)	0.050 (0.075)
EE-AFT	1.107 (0.009)	0.027 (0.002)	- -	- -	−0.105 (0.005)	0.050 (0.011)	1.175 (0.032)
ExW-AFT	0.007 (0.016)	0.141 (0.002)	0.867 (0.008)	- -	−0.108 (0.006)	0.027 (0.011)	0.466 (0.038)

Table 22. z-value, p-value, and confidence interval (CI) for the AFT estimates for each model at the level of significance 5%.

AFT-Model	z-Value	p-Value	CI 95%
$\hat{a}_{ExEW}$	16.378	<0.000	(0.412, 0.525)
$\hat{b}_{ExEW}$	19.461	<0.000	(0.510, 0.625)
$\hat{c}_{ExEW}$	805.912	<0.000	(0.975, 0.979)
$\hat{\alpha}_{ExEW}$	40.551	<0.000	(1.443, 1.589)
$\hat{\beta}_{1ExEW}$	27.457	<0.000	(−0.215, −0.187)
$\hat{\beta}_{2ExEW}$	7.482	<0.000	(0.010, 0.045)
$\hat{\beta}_{3ExEW}$	23.692	<0.000	(0.464, 0.548)
$\hat{a}_W$	6.136	<0.000	(2.567, 4.977)
$\hat{b}_W$	340.541	<0.000	(0.864, 0.874)
$\hat{\beta}_{1W}$	−27.114	<0.000	(−0.158, −0.137)
$\hat{\beta}_{2W}$	−0.098	<0.018	(−0.100, −0.080)
$\hat{\beta}_{3W}$	12.734	<0.000	(0.035, 0.065)
$\hat{a}_{EE}$	126.464	<0.000	(1.090, 1.124)
$\hat{b}_{EE}$	15.592	<0.000	(0.024, 0.030)
$\hat{\beta}_{1EE}$	−19.384	<0.018	(−0.116, −0.094)
$\hat{\beta}_{2EE}$	4.724	<0.922	(0.029, 0.071)
$\hat{\beta}_{3EE}$	36.634	<0.000	(1.112, 1.238)
$\hat{a}_{ExW}$	0.425	0.671	(−0.024, 0.037)
$\hat{b}_{ExW}$	56.618	<0.000	(0.136, 0.146)
$\hat{c}_{ExW}$	108.2542	<0.000	(0.851, 0.882)
$\hat{\beta}_{1ExW}$	−17.067	<0.000	(−0.121, −0.096)
$\hat{\beta}_{2ExW}$	2.493	<0.013	(0.006, 0.049)
$\hat{\beta}_{3ExW}$	39.738	<12.266	(0.392, 0.541)

Table 23. The analytical performance measures for comparing AFT models for COVID-19 data.

Model	BIC	CAIC
ExEW-AFT	147,089	147,096
W-AFT	149,056	149,061
EE-AFT	149,266	149,271
ExW-AFT	148,624	148,630

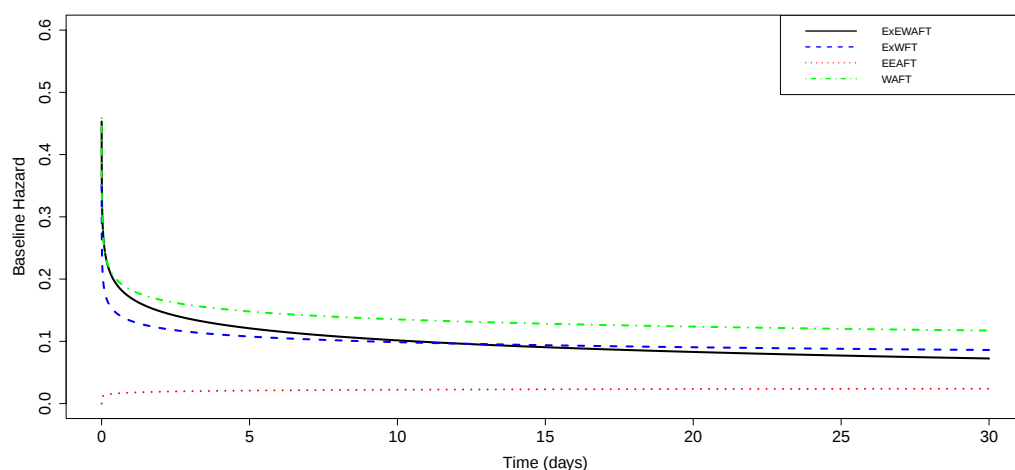


Figure 6. Estimated HRFs for the competing baseline hazards.

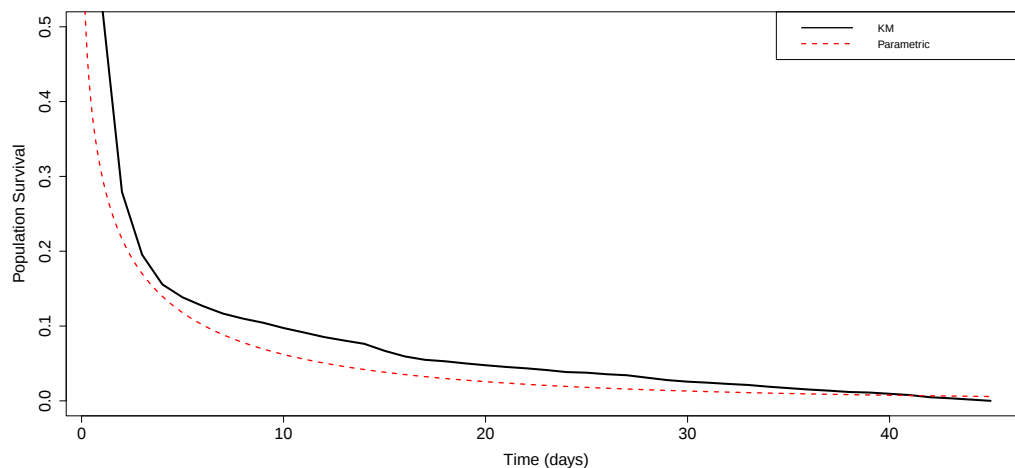


Figure 7. Average KM estimator and population SF.

8. Discussion

The study was designed to produce a more versatile and general model than the Cox model. We assessed the performance of the proposed model's estimators through a comprehensive Monte Carlo simulation study. The proposed model was also applied to Sudan COVID-19 Data. The choice of the AFT model is, therefore, sound as covariates directly relate to the time to event, which eases interpretability.

Considering the four different hazard rate (HR) shapes, including (increasing, decreasing, unimodal, and bathtub HR shapes), the simulation results showed that the ExEW model is capable of representing monotone decreasing, monotone increasing, unimodal, and bathtub HR functions more accurately than the existing AFT models such as: exponentiated-Weibull, Weibull, and log-logistic AFT models. Additionally, the SE, AB, RB, and MSE values showed that the proposed ExEW-AFT model performed well.

A real right-censored COVID-19 data set from Sudan reveals that it is misleading to trust the analysis based on the usual PH model, especially when the data exhibits characteristics such as the proportionality assumption [47]. This choice induces wrong conclusions, which, in turn, may lead to inappropriate clinical practices in terms of the best model that fits the data [11,48]. As demonstrated in our analysis, using different information criteria and some goodness-of-fit tests, including the likelihood ratio test, our proposed AFT model fits the data very well as compared to the common Cox-PH model and some other AFT single-parameter regression models (EW-AFT, W-AFT, and LL-AFT models). The three covariates (age, sex, and comorbidity) are significantly associated

with the length of stay in the hospital. Previous studies have also shown the obtained results [49–51].

The developed model provided an important contribution to the toolset for assessing survival data and can be used with overall hazard regression models. However, our model has some limitations. The model is unable to handle survival data with crossing survival curves. Also, the future will see the development of residual analysis techniques and diagnostic measures for assessing the goodness of fit of the proposed model. Additionally, this work can be extended by proposing an AFT multi-parameter regression model for other types of censored survival data sets, including left censoring, interval censoring, middle censoring, and double censoring mechanisms. Furthermore, we can extend it to more complex survival models, including competing risk models, cure models, frailty models, and mixed effects AFT models, to apply to spatial and clustered survival data sets. In the analysis, only age, sex, and comorbidity covariates were used. Vaccination and place of residence (urban and rural) as variables to analyze the length of stay in the hospital could be very useful, but no information from those variables was collected.

9. Conclusions

In conclusion, the Ex-EW-AFT multi-parameter regression model's flexibility could be highly helpful in representing many forms of time-to-event data that are difficult to predict accurately. Our model can be viewed as an attractive alternative for upcoming studies that assess censored lifetimes. As a result, the study includes a large number of simulation scenarios for evaluating the performance of our proposed model. The results demonstrated that the ExEW-AFT model performed effectively.

Future studies are advised to test the proposed model using other countries' data sets as well as validate if the same factors would be important in a lifetime analysis. Finally, we recommend that in the future we look for a hazard-based regression model that can handle survival data with crossing survival curves.

Author Contributions: Conceptualization, A.B.S.M., A.S.A., O.N., J.M., C.C. and A.Z.A.; Data curation, O.N., J.M. and A.Z.A.; Formal analysis, A.B.S.M., A.S.A., O.N., J.M., C.C. and A.Z.A.; Methodology, A.B.S.M., A.S.A., O.N., J.M., C.C. and A.Z.A.; Investigation, A.B.S.M., A.S.A., O.N., J.M., C.C. and A.Z.A.; Software, A.B.S.M., O.N., J.M., C.C. and A.Z.A.; Project administration, A.S.A., O.N., J.M. and A.Z.A.; Supervision, O.N., J.M. and A.Z.A.; Writing—review and editing, A.B.S.M., A.S.A., O.N., J.M., C.C. and A.Z.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The corresponding author may be able to provide the real data that were used to illustrate the proposed approaches upon reasonable request. Due to ethical restrictions, the data are not publicly available.

Acknowledgments: The first author would like to thank the Pan African University, Institute for Basic Sciences, Technology, and Innovation (PAUSTI), Nairobi, Kenya. Additionally, We would like to thank the Federal Ministry of Health, Sudan for supporting this study with the required data.

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

References

1. Cox, D.R. Regression models and life-tables. *J. R. Stat. Soc. Ser. B* **1972**, *34*, 187–202.
2. Khan, S.A.; Khosa, S.K. Generalized log-logistic proportional hazard model with applications in survival analysis. *J. Stat. Distrib. Appl.* **2016**, *3*, 1–18.
3. Rubio, F.J.; Remontet, L.; Jewell, N.P.; Belot, A. On a general structure for hazard-based regression models: An application to population-based cancer research. *Stat. Methods Med. Res.* **2019**, *28*, 2404–2417.
4. Ashraf-Ul-Alam, M.; Khan, A.A. Generalized Topp-Leone-Weibull AFT modelling: A Bayesian analysis with MCMC tools using R and stan. *Austrian J. Stat.* **2021**, *50*, 52–76.

5. Muse, A.H.; Mwalili, S.; Ngesa, O.; Alshanbari, H.M.; Khosa, S.K.; Hussam, E. Bayesian and frequentist approach for the generalized log-logistic accelerated failure time model with applications to larynx-cancer patients. *Alex. Eng. J.* **2022**, *61*, 7953–7978.
6. Withana Gamage, P.W.; McMahan, C.S.; Wang, L. A flexible parametric approach for analyzing arbitrarily censored data that are potentially subject to left truncation under the proportional hazards model. *Lifetime Data Anal.* **2022**, 1–25. <https://doi.org/10.1007/s10985-022-09579-z>.
7. Khan, S.A.; Basharat, N. Accelerated failure time models for recurrent event data analysis and joint modeling. *Comput. Stat.* **2022**, *37*, 1569–1597.
8. Lawless, J.F. *Statistical Models and Methods for Lifetime Data*; John Wiley & Sons: Hoboken, NJ, USA, 2011.
9. Odell, P.M.; Anderson, K.M.; D’Agostino, R.B. Maximum likelihood estimation for interval-censored data using a Weibull-based accelerated failure time model. *Biometrics* **1992**, *48*, 951–959.
10. Huber, C.; Limnios, N.; Mesbah, M.; Nikulin, M.S. *Mathematical Methods in Survival Analysis, Reliability and Quality of Life*; John Wiley & Sons: Hoboken, NJ, USA, 2013.
11. Aida, H.; Hayashi, K.; Takeuchi, A.; Sugiyama, D.; Okamura, T. An accelerated failure time cure model with shifted gamma frailty and its application to epidemiological research. *Healthcare* **2022**, *10*, 1383.
12. Kleinbaum, D.G.; Klein, M. *Survival Analysis: A Self-Learning Text*; Springer: New York, NY, USA, 2012; Volume 3.
13. Collett, D. *Modelling Survival Data in Medical Research*; CRC Press: Boca Raton, FL, USA, 2015.
14. Crowther, M.J.; Royston, P.; Clements, M. A flexible parametric accelerated failure time model and the extension to time-dependent acceleration factors. *Biostatistics* **2022**. <https://doi.org/10.1093/biostatistics/kxac009>.
15. Sinha, S.K. Robust estimation in accelerated failure time models. *Lifetime Data Anal.* **2019**, *25*, 52–78.
16. Zhang, Z.; Sinha, S.; Maiti, T.; Shipp, E. Bayesian variable selection in the accelerated failure time model with an application to the surveillance, epidemiology, and end results breast cancer data. *Stat. Methods Med. Res.* **2018**, *27*, 971–990.
17. Legrand, C. *Advanced Survival Models*; Chapman and Hall/CRC: Boca Raton, FL, USA, 2021.
18. Khan, S.A. Exponentiated Weibull regression for time-to-event data. *Lifetime Data Anal.* **2018**, *24*, 328–354.
19. Santana, T.V.; Ortega, E.M.; Cordeiro, G.M. Generalized beta Weibull linear model: Estimation, diagnostic tools and residual analysis. *J. Stat. Theory Pract.* **2019**, *13*, 1–23.
20. Muse, A.H.; Ngesa, O.; Mwalili, S.; Alshanbari, H.M.; El-Bagoury, A.A.H. A flexible Bayesian parametric proportional hazard model: Simulation and applications to right-censored healthcare data. *J. Healthc. Eng.* **2022**, 2022, 2051642.
21. Rinne, H. *The Hazard Rate: Theory and Inference (with Supplementary MATLAB-Programs)*; Deutsche Nationalbibliothek: Leipzig, Germany, 2014.
22. Ciampi, A.; Etezadi-Amoli, J. A general model for testing the proportional hazards and the accelerated failure time hypotheses in the analysis of censored survival data with covariates. *Commun. Stat.-Theory Methods* **1985**, *14*, 651–667.
23. Kalbfleisch, J.D.; Prentice, R.L. *The Statistical Analysis of Failure Time Data*; John Wiley & Sons: Hoboken, NJ, USA, 2011.
24. Biazatti, E.C.; Cordeiro, G.M.; Rodrigues, G.M.; Ortega, E.M.; de Santana, L.H. A Weibull-beta prime distribution to model COVID-19 data with the presence of covariates and censored data. *Stats* **2022**, *5*, 1159–1173.
25. Cordeiro, G.M.; Rodrigues, G.M.; Ortega, E.M.; de Santana, L.H.; Vila, R. An extended Rayleigh model: Properties, regression and COVID-19 application. *arXiv* **2022**, arXiv:2204.05214.
26. Biazatti, E.C.; Cordeiro, G.M.; de Lima, M.d.C.S. The dual-Dagum family of distributions: Properties, regression and applications to COVID-19 data. *Model Assist. Stat. Appl.* **2022**, *17*, 199–210.
27. Rodrigues, G.M.; Ortega, E.M.; Cordeiro, G.M.; Vila, R. An extended Weibull regression for censored data: Application for COVID-19 in campinas, Brazil. *Mathematics* **2022**, *10*, 3644.
28. Muse, A.H.; Mwalili, S.; Ngesa, O.; Chesneau, C.; Alshanbari, H.M.; El-Bagoury, A.A.H. Amoud class for hazard-based and odds-based regression models: Application to oncology studies. *Axioms* **2022**, *11*, 606.
29. Economou, P.; Caroni, C. Parametric proportional odds frailty models. *Commun. Stat. Comput.* **2007**, *36*, 1295–1307.
30. Mastor, A.B.S.; Alfaer, N.M.; Ngesa, O.; Mung’atu, J.; Afify, A.Z. The extended exponential Weibull distribution: Properties, inference, and applications to real-life data. *Complexity* **2022**, 2022, 4068842.
31. Jaouimaa, F.Z.; Ha, I.D.; Burke, K. Multi-parameter regression survival modelling with random effects. *arXiv* **2021**, arXiv:2111.08573.
32. Peng, D.; MacKenzie, G.; Burke, K. A multiparameter regression model for interval-censored survival data. *Stat. Med.* **2020**, *39*, 1903–1918.
33. Burke, K.; Jones, M.; Noufaily, A. A flexible parametric modelling framework for survival analysis. *J. R. Stat. Soc. Ser. C Appl. Stat.* **2020**, *69*, 429–457.
34. Burke, K.; Eriksson, F.; Pipper, C. Semiparametric multiparameter regression survival modeling. *Scand. J. Stat.* **2020**, *47*, 555–571.
35. Kleinbaum, D.G.; Klein, M. Evaluating the proportional hazards assumption. In *Survival Analysis*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 161–200.
36. Leemis, L.M.; Shih, L.H.; Reynertson, K. Variate generation for accelerated life and proportional hazards models with time dependent covariates. *Stat. Probab. Lett.* **1990**, *10*, 335–339.
37. Leemis, L.M. Variate generation for accelerated life and proportional hazards models. *Oper. Res.* **1987**, *35*, 892–894.

38. Bender, R.; Augustin, T.; Blettner, M. Generating survival times to simulate Cox proportional hazards models. *Stat. Med.* **2005**, *24*, 1713–1723.
39. Austin, P.C. Generating survival times to simulate Cox proportional hazards models with time-varying covariates. *Stat. Med.* **2012**, *31*, 3946–3958.
40. Muse, A.H.; Mwalili, S.; Ngesa, O.; Chesneau, C.; Al-Bossly, A.; El-Morshedy, M. Bayesian and frequentist approaches for a tractable parametric general class of hazard-based regression models: An application to oncology data. *Mathematics* **2022**, *10*, 3813.
41. Liu, H.; Tian, X. Data-driven optimal control of a SEIR model for COVID-19. *arXiv* **2020**, arXiv:2012.00698.
42. Cordeiro, G.M.; Figueiredo, D.; Silva, L.; Ortega, E.M.; Prata, F. Explaining COVID-19 mortality rates in the first wave in Europe. *Model Assist. Stat. Appl.* **2021**, *16*, 211–221.
43. Marinho, P.R.D.; Cordeiro, G.M.; Coelho, H.F.; Brandão, S.C.S. Covid-19 in Brazil: A sad scenario. *Cytokine Growth Factor Rev.* **2021**, *58*, 51–54.
44. Cabore, J.W.; Karamagi, H.C.; Kipruto, H.K.; Mungatu, J.K.; Asamani, J.A.; Droti, B.; Titi-Ofei, R.; Seydi, A.B.W.; Kidane, S.N.; Balde, T.; et al. COVID-19 in the 47 countries of the WHO African region: A modelling analysis of past trends and future patterns. *Lancet Glob. Health* **2022**, *10*, e1099–e1114.
45. Kiarie, J.W.; Mwalili, S.M.; Mbogo, R.W. COVID-19 pandemic situation in Kenya: A data driven SEIR model. *Med. Res. Arch.* **2022**, *10*. <https://doi.org/10.18103/mra.v10i2.2658>.
46. Alvares, D.; Lázaro, E.; Gómez-Rubio, V.; Armero, C. Bayesian survival analysis with BUGS. *Stat. Med.* **2021**, *40*, 2975–3020.
47. Patel, K.; Kay, R.; Rowell, L. Comparing proportional hazards and accelerated failure time models: an application in influenza. *Pharm. Stat. J. Appl. Stat. Pharm. Ind.* **2006**, *5*, 213–224.
48. Yang, J.; Zheng, Y.; Gou, X.; Pu, K.; Chen, Z.; Guo, Q.; Ji, R.; Wang, H.; Wang, Y.; Zhou, Y. Prevalence of comorbidities and its effects in patients infected with SARS-CoV-2: A systematic review and meta-analysis. *Int. J. Infect. Dis.* **2020**, *94*, 91–95.
49. Thiruvengadam, G.; Lakshmi, M.; Ramanujam, R. A study of factors affecting the length of hospital stay of COVID-19 patients by cox-proportional hazard model in a South Indian tertiary care hospital. *J. Prim. Care Community Health* **2021**, *12*, 21501327211000231.
50. Giacomelli, A.; Ridolfo, A.L.; Milazzo, L.; Oreni, L.; Bernacchia, D.; Siano, M.; Bonazzetti, C.; Covizzi, A.; Schiuma, M.; Passerini, M.; et al. 30-day mortality in patients hospitalized with COVID-19 during the first wave of the Italian epidemic: A prospective cohort study. *Pharmacol. Res.* **2020**, *158*, 104931.
51. Wu, S.; Xue, L.; Legido-Quigley, H.; Khan, M.; Wu, H.; Peng, X.; Li, X.; Li, P. Understanding factors influencing the length of hospital stay among non-severe COVID-19 patients: A retrospective cohort study in a Fangcang shelter hospital. *PLoS ONE* **2020**, *15*, e0240959.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.