axioms

MDPI

Article

Amoud Class for Hazard-Based and Odds-Based Regression Models: Application to Oncology Studies

Abdisalam Hassan Muse [1,2,*], Samuel Mwalili [3], Oscar Ngesa [4], Christophe Chesneau [5,*], Huda M. Alshanbari [6] and Abdal-Aziz H. El-Bagoury [7]

1 Institute for Basic Sciences, Pan African University, Technology and Innovation (PAUSTI), Nairobi 62000-00200, Kenya
2 Faculty of Science and Humanities, School of Postgraduate Studies and Research, Amoud University, Borama 25263, Somalia
3 Department of Statistics and Actuarial Science, Jomo Kenyatta University of Agriculture and Technology (JKUAT), Nairobi 62000-00200, Kenya
4 Department of Mathematics and Physical Sciences, Taita Taveta University, Voi 635-80300, Kenya
5 Department of Mathematics, LMNO, CNRS-Université de Caen, Campus II, Science 3, 14032 Caen, France
6 Department of Mathematical Sciences, College of Science, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia
7 Basic Science Department, Higher Institute of Engineering and Technology, El-Mahala El-Kobra 6734723, Egypt
* Correspondence: abdisalam.hassan@amoud.edu.so (A.H.M.); christophe.chesneau@unicaen.fr (C.C.)

Abstract: The purpose of this study is to propose a novel, general, tractable, fully parametric class for hazard-based and odds-based models of survival regression for the analysis of censored lifetime data, named as the "Amoud class (AM)" of models. This generality was attained using a structure resembling the general class of hazard-based regression models, with the addition that the baseline odds function is multiplied by a link function. The class is broad enough to cover a number of widely used models, including the proportional hazard model, the general hazard model, the proportional odds model, the general odds model, the accelerated hazards model, the accelerated odds model, and the accelerated failure time model, as well as combinations of these. The proposed class incorporates the analysis of crossing survival curves. Based on a versatile parametric distribution (generalized log-logistic) for the baseline hazard, we introduced a technique for applying these various hazard-based and odds-based regression models. This distribution allows us to cover the most common hazard rate shapes in practice (decreasing, constant, increasing, unimodal, and reversible unimodal), and various common survival distributions (Weibull, Burr-XII, log-logistic, exponential) are its special cases. The proposed model has good inferential features, and it performs well when different information criteria and likelihood ratio tests are used to select hazard-based and odds-based regression models. The proposed model's utility is demonstrated by an application to a right-censored lifetime dataset with crossing survival curves.

Keywords: Bayesian analysis; survival models; general odds model; general hazard model; accelerated odds model; accelerated failure time model; Amoud class; proportional odds model; accelerated hazard model; survival analysis; generalized log-logistic distribution; proportional hazard model; censored data

MSC: 62N01; 62N02; 62F15; 65C60; 62P10

Citation: 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. https://doi.org/10.3390/axioms11110606

Academic Editors: Jiajuan Liang and Kaitai Fang

Received: 5 October 2022
Accepted: 27 October 2022
Published: 1 November 2022

1. Introduction

During the last few decades, the semi-parametric Cox proportional hazard (PH) model has dominated survival data analysis. While Cox's original research paper discussed extensions to remove the assumption of PH [1], much work has been carried out to improve the flexibility of survival regression frameworks by using tractable functions for both the

baseline and the inclusion of covariates, primarily using probability distributions, splines, or fractional polynomials [2].

As a matter of fact, the hazard rate and odds functions are two probabilistic functions with significant practical value in survival analysis. They both take into account the hazard rate or odds for a reference level associated with a link function of the covariates, which is often represented by a log-linear or a multiplicative term $\exp(x_i'\beta)$. Each covariate's associated parameters are represented by the vector β. Given a design matrix X and a subject $i, i \in \{1, \ldots, n\}$, the vector x_i represents covariate values. The subject i with all of its covariate values equal to zero $(x_i = 0)$ represents the reference level.

So, based upon the type of probabilistic function utilized as a baseline function, the survival regression model classes can be divided into two primary groups: hazard-based regression models and odds-based regression models. However, the general design of those models did not change; the hazard or odds function was expressed as a baseline function multiplied by the link function of the covariates to either the baseline function, the time scale, or both of them.

Because of the most well-known Cox PH model, hazard-based regression models are the most prevalent survival regression model classes in the field of survival analysis. Consequently, there are four widely used hazard-based regression models: proportional hazard (PH) [1,3], accelerated hazard (AH) [4], accelerated failure time (AFT) [5,6], and general hazard (GH) [7–10]. The odds-based regression models, which are created using a probabilistic function that has recently received more attention and is known as the odds function, are another family of survival regression model classes. Although the odds function is used in epidemiological case-control research, the proportional odds (PO) model class that was presented by Bennet [11] is the first to apply it in survival models. AFT model is another odds-based regression model [3]. As a result, just like hazard-based regression models, odds-based models are divided into four primary categories: PO [11], accelerated odds (AO), AFT [5], and general odds (GO) models.

There are other survival regression models as well, which combine hazard-based and odds-based regression models and are built by taking into account both hazard rate and odds functions. For instance, Yang and Prentice [12] developed the Yang-Prentice model, a semi-parametric survival regression model that can include crossover survival curves. In order to describe survival data with crossed survival curves, Demarqui and Mayrink [13] modified the Yang-Prentice (YP) model using a piecewise exponential baseline distribution. Both the PH and PO models are included as sub-models in the YP model. A generalized odds-rate hazards model was developed by Banerjee et al. [14] and includes the PH, PO, and AFT models as special cases.

For censored lifetime data, Royston and Parmar [15] presented a flexible parametric model based on the PH and PO models. On the other hand, Huang et al. [16] also introduced a general class of regression model called the PH-PO model, which includes PH and PO models as sub-models. Huang and Jiang [17] proposed an extension of the PH-PO model into a more generalized model that takes into account time scale changing effects and time varying coefficient effects. A semi-parametric super model containing six popular survival regression models, including the PH, PO, AFT, AH, YP, and GH models, was recently proposed by Zhang et al. [18]. Davis [19] has recommended the development of further new families that combine hazard-based and odds-based regression models. For additional details, please see [20].

The absence of a general class of odds-based and hazard-based regression models that encompasses all hazard-based and odds-based regression frameworks is an issue that needs to be addressed. Each of the hazard-based and odds-based regression model classes mentioned above can capture different aspects of survival data. On the other hand, choosing which hazard-based or odds-based regression model is the most suitable and precise in reflecting the link between baseline (hazard or odds) and covariates, is an issue and an important research problem that must be addressed. To explicitly nest simpler models and to address the issue, we propose a novel, general, flexible, fully parametric class of hazard-based and odds-based regression framework named "Amoud class (AM)".

In contrast, there are three categories of survival regression model classes: non-parametric, semi-parametric, and parametric models. Compared to non-parametric and semi-parametric methods, parametric models are more informative. They can be used to forecast survival times, hazard rates, as well as mean and median survival times in addition to computing relative effect estimates [21]. They can also be used to plot covariate-adjusted survival curves and forecast absolute risk over time. Semi-parametric models lack the power of parametric models when the parametric form is incorrectly stated. Additionally, they are more effective, resulting in estimates with reduced standard errors and greater accuracy [22,23]. Furthermore, parametric techniques use full maximum likelihood to estimate parameters. Parametric model residuals often take the form of the discrepancy between what was observed and what was expected [24].

Considering the discussion above, the current study proposes a fully parametric class of regression models that comprises formally nested special cases of the PH, PO, AH, AO, AFT, GH, and GO survival regression models. As a result, model selection among these models can be accomplished by conducting approximate likelihood ratio tests using the frequentist approach. To describe baseline hazard or baseline odds, a generalized log-logistic (GLL) distribution containing some of the most frequent parametric baseline distributions used in survival analysis, such as the log-logistic (LL), Bur-XII, exponential, and Weibull distributions, is employed. A right-censoring mechanism is considered, and the proposed model's parameters are evaluated using maximum likelihood estimation and Bayesian estimation techniques. A real-world right-censored survival dataset with a crossing survival curve is utilized to demonstrate how the proposed AM class can be employed.

Hence, the novelty of this research paper is to introduce and investigate a novel, general, tractable, fully-parametric class of hazard-based and odds-based regression model for dealing with right-censored survival data with or without crossing survival curves. This is accomplished by assuming the GLL distribution in the proposed class to cope with the baseline distribution. To the author's best knowledge, no one has ever contemplated employing the parametric AM class of parametric hazard-based and odds-based regression models in general, and with GLL baseline hazard in particular. This class is an extension of the most common hazard-based and odds-based regression frameworks in the literature. On the other hand, another area of interest that has yet to be addressed in the context of the AM class is the use of both the inferential procedures, Bayesian and frequentist approaches. As a result, the strategies are investigated utilizing the frequentist approach using the maximum likelihood estimation (MLE) method and the Bayesian approach using non-informative priors.

The structure of the article is as follows: Section 2 presents a review of the hazard-based, and odds-based regression models in the context of survival, duration, and reliability analysis. The formulation of the AM class, its associated probabilistic functions, and sub-models of the class are discussed in Section 3. Section 4 describes the baseline distribution under examination in this study, as well as some of its special circumstances. Section 5 presents the estimation of the proposed class parameters using both classical MLE and Bayesian estimation approaches. Section 6 focuses on the model selection situations for both nested and non-nested models. Section 7 shows a real-world, right-censored cancer dataset with crossing survival curves. Section 8 finishes the study with a farewell address and recommendations for future research.

2. Recent Literature Review and State of Art

In this section, we review the studies completed in the framework of the hazard-based and odds-based regression models that are closely related to the proposed class in order to illustrate the state of scientific development in the context of current survival, duration, and reliability models.

2.1. Hazard-Based Regression Models

In general, survival datasets are highly skewed and can be censored for some subjects, possibly even the most. Standard linear regression models cannot fit them, and they also only allow for the interpretation of regression coefficients in terms of the mean of time. However, different models can be applied to survival data to generate different interpretations. Observed times' functions rather than the observed times themselves are used for this. The hazard rate and the odds functions, in particular, are two probabilistic functions that are extremely important practically in survival analysis.

There are four major types of hazard-based regression models proposed in the literature to fit survival time data in medical investigations, namely, PH, AH, AFT, and GH models. These models can be used to analyze real-world data in domains other than medicine, such as economics, marketing, engineering, social science, criminology, and education. The modeling approach differs depending on the researcher's event of interest; the general notion is to watch time until the event occurs; however, for some subjects, the event never occurs.

The formulation and construction of four hazard-based regression models are reviewed and discussed in this section. We define the alternative structures below using the hazard rate function (hrf), odds function, survival (complementary distribution) function (sf), and cumulative (or integrated) hazard function (chf) in relation to time t and a vector of covariates x. We suppose that the vector of covariates lacks an intercept to avoid concerns about identifiability. The unknown regression coefficients are represented by a vector β.

2.1.1. PH Model

The semi-parametric PH model introduced by Cox [1] is one of the most well-known hazard-based regression models in survival analysis. The hrf is multiplicatively affected by the impact of the covariates in this model. Different researchers have examined and analyzed studies relating to the parametric PH model utilizing various baseline distributions and inferential techniques. A parametric PH model, with an extended exponential geometric baseline distribution was developed and evaluated by Rezaei et al. [25]. A parametric PH with GLL baseline distribution was also proposed by Khan and Khosa [23]. A modified PH model and a reversed PH model employing the Marshall-Olkin baseline distribution were examined by Balakrishnan et al. [26]. Muse et al. [27] have investigated the Bayesian analysis of the PH model with a GLL baseline distribution.

The PH model's hrf, odds, sf, and the chf can be stated as follows:

$$h_{PH}(t;\beta,x) = h_0(t)e^{x'\beta}, \tag{1}$$

$$R_{PH}(t;\beta,x) = R_0(t)^{e^{x'\beta}}, \tag{2}$$

$$S_{PH}(t;\beta,x) = S_0(t)^{e^{x'\beta}}, \tag{3}$$

$$H_{PH}(t;\beta,x) = H_0(t)e^{x'\beta}, \tag{4}$$

where H_0, R_0, S_0, and H_0 are the baseline hazard rate, odds, survival and cumulative hazard functions.

2.1.2. AFT Model

The PH model is the most popular hazard-based regression model in survival analysis, but it can only be used in situations in which the PH assumption holds. An alternative to the PH model is the AFT model [3,5]. The AFT model is analogous to a hazard-based regression model in which covariates measured on an individual are assumed to act multiplicatively on the time-scale, influencing the rate at which the individual advances along the time axis. Numerous scholars have studied and discussed studies involving the parametric AFT model using various baseline hazards and statistical inference techniques. A parametric AFT model with an exponentiated Weibull baseline distribution was presented and analyzed by Khan [22]. A parametric AFT with a log-exponential power baseline

distribution was also proposed by Olosunde et al. [28]. Ashraf-Ul-alam and Khan [29] used a generalized Top-leone-Weibull baseline distribution to study a parametric AFT model. A parametric AFT model with a GLL baseline distribution was recently proposed by Muse et al. [30].

The hrf, odds, sf, and chf of the AFT model are defined by:

$$h_{AFT}(t;\beta,x) = h_0\left(te^{x'\beta}\right)e^{x'\beta}, \tag{5}$$

$$R_{AFT}(t;\beta,x) = R_0\left(te^{x'\beta}\right), \tag{6}$$

$$S_{AFT}(t;\beta,x) = S_0\left(te^{x'\beta}\right), \tag{7}$$

$$H_{AFT}(t;\beta,x) = H_0\left(te^{x'\beta}\right). \tag{8}$$

2.1.3. AH Model

AFT and PH models have been widely applied to deal with lifetime data in different disciplines of knowledge. Despite being widely used, such hazard-based regression models are not suitable to handle survival data with crossing survival curves. Chen and Wang [4] proposed a semi-parametric hazard-based regression model, named the AH model, allowing the analysis of crossing survival curves. In the context of a parametric AH model, different baseline hazards are available in the AHSurv package [31].

The hrf, odds, sf, and chf of the AH model are defined by:

$$h_{AH}(t;\beta,x) = h_0\left(te^{x'\beta}\right), \tag{9}$$

$$R_{AH}(t;\beta,x) = R_0\left(te^{x'\beta}\right)^{e^{-x'\beta}}, \tag{10}$$

$$S_{AH}(t;\beta,x) = S_0\left(te^{x'\beta}\right)^{e^{-x'\beta}}, \tag{11}$$

$$H_{AH}(t;\beta,x) = H_0\left(te^{x'\beta}\right)e^{-x'\beta}. \tag{12}$$

2.1.4. GH Model

Ciampi and Etazadi-Amoli [7] introduced a general hazard (GH) regression model for testing the PH and the AFT hypothesis in the analysis of censored lifetime data with the presence of covariates. Then, Etazadi-Amoli and Ciampi [8] extended their work by the application of the splines as a baseline function. It is worth to mention that the EHR model requires a careful selection of the knots and could easily lead to an overparametrized or non-identifiable model. On the other hand, Chen and Jewell [10] introduced a general class of semi-parametric hazard-based regression models by following the same procedure as [7] but just by adding the AH framework. The case of parametric GH structure in the context of relative survival framework was introduced by Rubio et al. [2] and developed an R Package called GHSurv available in https://github.com/FJRubio67/GHSurv (accessed on 5 October 2022) and HazReg package that contain several choices of the baseline hazards (https://github.com/FJRubio67/HazReg, accessed on 5 October 2022). Recently, Muse et al. [32] proposed the over survival framework for the GH model using both Bayesian and classical inferences. Consequently, Alvares and Rubio [33] discussed the GH structure in the context of joint models for longitudinal and survival datasets. Finally, in the context of the a spatial survival models, Li et al. [34] extended the GH model to spatial GH model. Recently, Rubio and Drikvandi [35] modified the GH structure into a mixed-effect general hazard (MEGH) model to account for cluster survival datasets.

The hrf, odds, sf, and chf of the GH model are expressed as follows:

$$h_{GH}(t;\beta_1,\beta_2,x) = h_0\left(te^{x'\beta_1}\right)e^{x'\beta_2}, \tag{13}$$

$$R_{GH}(t;\beta_1,\beta_2,x) = R_0\left(te^{x'\beta_1}\right)^{e^{x'(\beta_2-\beta_1)}}, \tag{14}$$

$$S_{GH}(t;\beta_1,\beta_2,x) = S_0\left(te^{x'\beta_1}\right)^{e^{x'(\beta_2-\beta_1)}}, \tag{15}$$

$$H_{GH}(t;\beta_1,\beta_2,x) = H_0\left(te^{x'\beta_1}\right)e^{x'(\beta_2-\beta_1)}, \tag{16}$$

where β_1 and β_2 denote the unknown regression parameters.

2.1.5. Special Cases of the GH Model

All of the hazard-based regression models listed above are incorporated into the GH model of hazard-based models as special cases. The GH model can be used to derive the PH, AH, and AFT models, according to the following proposition.

Proposition 1. *Suppose $h(t;\beta,x)$ is given by Equation (13). Then, we have the following results:*

1. *If $\beta_2 = \beta_1$, then $h(t;\beta,x) = h_o\left(te^{x'\beta}\right)e^{x'\beta}$ giving the AFT model.*
2. *If $\beta_1 = 0$, then $h(t;\beta,x) = h_o(t)e^{x'\beta}$ giving the PH model.*
3. *If $\beta_2 = 0$, then $h(t;\beta,x) = h_o\left(te^{x'\beta}\right)$, giving the AH model.*

Proof of Proposition 1. The proof of Proposition 1 is straightforward. □

2.2. Odds-Based Regression Models

To fit survival time data in medical research, two primary types of odds-based regression models have been proposed in the literature: PO and AFT models. Two more innovative odds-based regression models proposed in this study are the accelerated odds and general odds models. In fields other than medicine, such as economics, marketing, engineering, social science, criminology, and education, these models can be utilized to examine actual data.

The odds function indicates how much more likely it is that a particular event will occur for a given period t. As a result, the odds function is denoted by $R(t;\theta)$, and its mathematical expression is given by the relationship between the cumulative distribution function and its complementary (sf):

$$R(t;\theta) = \frac{F(t;\theta)}{S(t;\theta)} = \frac{1-\exp[-H(t;\theta)]}{\exp[-H(x;\theta)]} = \exp[H(t;\theta)] - 1, \tag{17}$$

where $R(t;\theta), F(t;\theta), S(t;\theta)$, and $H(t;\theta)$ are the odds, cumulative distribution, survival and cumulative hazard functions respectively, and θ is the vector of distributional parameters.

The associated derivative of the odds function is expressed as follows:

$$r(t;\theta) = \frac{dR(t;\theta)}{d(t)} = \frac{h(t;\theta)}{S(t;\theta)} = \frac{f(t;\theta)}{S(t;\theta)^2}, \tag{18}$$

where $r(t;\theta), h(t;\theta)$, and $f(t;\theta)$ represent the odds, hrf, and probability density function (pdf), respectively.

In this section, we review two odds-based regression models that have been explored in the literature along with their formulation. On the other hand, based on the author's knowledge, we present two novel odds-based regression models that have never been used before in the literature. We define the alternative structures below with respect to time t and a vector of covariates x using the odds function $R(.)$, derivative of odds function $r(.)$,

hrf $h(.)$, and sf $S(.)$. We assume that the covariate vector is free of an intercept to ease issues about identifiability. The vector β is used to represent the unknown regression coefficients.

2.2.1. PO Model

The proportional odds (PO) model, originally introduced by Bennett [11], is an odds-based regression model. According to Bennett [11], the PO model is structurally similar to the PH model of Cox and may be used in similar situations. Although the PO model represents an attractive alternative to the PH model.

The odds function of this model is expressed as follows:

$$R_{PO}(t;\beta,x) = R_0(t)e^{x'\beta}, \tag{19}$$

where $R_0(t)$ is the baseline odds function. The associated derivative of the odds function of the PO model is computed as follows:

$$r_{PO}(t;\beta,x) = r_0(t)e^{x'\beta}, \tag{20}$$

where $r_0(t)$ is the baseline derivative odds function.

The hrf of the PO model is computed as follows:

$$h_{PO}(t;\beta,x) = \frac{r_0(t)e^{x'\beta}}{1+R_0(t)e^{x'\beta}}. \tag{21}$$

In terms of the baseline hazard, the hrf, and sf can be expressed as follows using Equation (18):

$$h_{PO}(t;\beta,x) = \frac{\left[\frac{h_0(t)}{S_0(t)}\right]e^{x'\beta}}{1+\left[\frac{F_0(t)}{S_0(t)}\right]e^{x'\beta}} = \frac{h_0(t)e^{x'\beta}}{F_0(t)e^{x'\beta}+S_0(t)}, \tag{22}$$

$$S_{PO}(t;\beta,x) = \left[\frac{1}{1+R_0(t)e^{x'\beta}}\right] = \left[\frac{1}{1+\left[\frac{F_0(t)}{S_0(t)}\right]e^{x'\beta}}\right] = \left[1+\left[\frac{F_0(t)}{S_0(t)}\right]e^{x'\beta}\right]^{-1}. \tag{23}$$

Now, we will put forth two new models that employ methods related to the hazard-based regression models. All of the odds-based regression models in this section will be generalized as well. The model formulation put forward by Chen and Wang [4] served as inspiration for the initial proposed approach. Their model includes accelerated hazards, but we propose a model with accelerated odds. This is how our models differ from theirs. The model formulation proposed by Chen and Jewell [10] served as the basis for the second proposed model. The PH, AH, and AFT models are included in their models as sub-models. In contrast to their model, ours includes the PO, AFT, and AO models as sub-models. The general odds model is the name of this model. The odds-based regression models with different baseline distributions are available in the AmoudSurv Package [36].

2.2.2. Accelerated Forms

The second parametric method of taking into account the effect of covariates, known as the accelerated form, presupposes that the covariates directly rescale time. Accelerated effects of covariates come in two varieties: Two examples of this are the:

i. Accelerated failure time (AFT) model; and
ii. Accelerated odds model.

The accelerated types of the odds-based regression models can be formulated in two different ways, the first of which is similar to the AFT model. The AFT model is the only parametric survival regression framework that belongs to both the hazard-based and odds-based regression models, and both the continuous probability distributions that are closed under the hazard-based regression models and those closed under the odds-based regression models are consistent with the AFT model. For instance, the Weibull and LL

distributions. We will explore these distributions in Section 4 of this study. Based on what the authors know, accelerated odds (AO) model is a new survival regression model that has never been used previously.

The formulation one belongs to the AFT framework and can be expressed as follows:

$$R_{AF}(t;\beta,x) = R_0\left(te^{x'\beta}\right). \tag{24}$$

The associated derivative of the odds function of the AFT model is computed as follows:

$$r_{AF}(t;\beta,x) = r_0\left(te^{x'\beta}\right)e^{x'\beta}. \tag{25}$$

The hrf and sf are expressed as follows:

$$h_{AF}(t;\beta,x) = \frac{R_0'\left(te^{x'\beta}\right)}{1+R_0\left(te^{x'\beta}\right)} = \frac{\left[\frac{h_0\left(te^{x'\beta}\right)}{S_0\left(te^{x'\beta}\right)}\right]e^{x'\beta}}{1+\left[\frac{F_0\left(te^{x'\beta}\right)}{S_0\left(te^{x'\beta}\right)}\right]} = \frac{h_0\left(te^{x'\beta}\right)e^{x'\beta}}{F_0\left(te^{x'\beta}\right)+S_0\left(te^{x'\beta}\right)} = h_0\left(te^{x'\beta}\right)e^{x'\beta}, \tag{26}$$

$$S_{AF}(t;\beta,x) = \left[\frac{1}{1+R_0\left(te^{x'\beta}\right)}\right]^{-1} = \left[\frac{1}{1+\left[\frac{F_0\left(te^{x'\beta}\right)}{S_0\left(te^{x'\beta}\right)}\right]}\right]^{-1} = \left[S_0\left(te^{x'\beta}\right)\right]. \tag{27}$$

This model, as one can see after its derivation and simplification, is similar to the AFT model. As a result, we can remark that the AFT model is the only one of the survival regression models that holds true for both hazard-based and odds-based regression models.

2.2.3. Accelerated Odds Model

A novel parametric odds-based regression model that can incorporate censored lifetime datasets with crossing survival curves is introduced here and named the "accelerated odds (AO)" model. This model is formulated using the odds function, and by using the same procedure as for the AH model, we obtained the following parametric odds-based regression model that is a new one and has not been featured in the literature so far:

$$R_{AO}(t;\beta,x) = R_0\left(te^{x'\beta}\right)e^{-x'\beta}. \tag{28}$$

The associated derivative of the odds function of the AO model is computed as follows:

$$r_{AO}(t;\beta,x) = r_0\left(te^{x'\beta}\right). \tag{29}$$

The hrf and sf are expressed as follows:

$$h_{AO}(t;\beta,x) = \frac{R_0'\left(te^{x'\beta}\right)}{1+R_0\left(te^{x'\beta}\right)} = \frac{h_0\left(te^{x'\beta}\right)}{F_0\left(te^{x'\beta}\right)e^{-x'\beta}+S_0\left(te^{x'\beta}\right)}, \tag{30}$$

$$S_{AO}(t;\beta,x) = \left[\frac{1}{1+e^{-x'\beta}\left[\frac{F_0\left(te^{x'\beta}\right)}{S_0\left(te^{x'\beta}\right)}\right]}\right] = \left\{1+e^{-x'\beta}\left[\frac{F_0\left(te^{x'\beta}\right)}{S_0\left(te^{x'\beta}\right)}\right]\right\}^{-1}. \tag{31}$$

2.2.4. General Odds Model

Another novel general survival regression model, termed the "general odds (GO)" model, is introduced here and consists of three odds-based regression models as special

cases, namely: PO, AFT, and AO models. The odds function of this model can be computed as follows:

$$R_{GO}(t;\beta_1,\beta_2,x) = R_0\left(te^{x'\beta_1}\right)e^{x'(\beta_2-\beta_1)}. \tag{32}$$

The associated derivative of the odds function of the GO model corresponding to the odds function in Equation (32) is computed as follows:

$$r_{GO}(t;\beta_1,\beta_2,x) = r_0\left(te^{x'\beta_1}\right)e^{x'\beta_2}. \tag{33}$$

The hrf and sf of the GO model corresponding to Equation (32) are expressed as follows:

$$h_{GO}(t;\beta_1,\beta_2,x) = \frac{h_0\left(te^{x'\beta_1}\right)e^{x'\beta_2}}{e^{x'(\beta_2-\beta_1)}F_0\left(te^{x'\beta_1}\right) + S_0\left(te^{x'\beta_1}\right)}. \tag{34}$$

$$S_{GO}(t;\beta_1,\beta_2,x) = \left[1 + e^{x'(\beta_2-\beta_1)}\frac{F_o\left(te^{x'\beta_1}\right)}{S_0\left(te^{x'\beta_1}\right)}\right]^{-1}. \tag{35}$$

In terms of the odds function, the sf in Equation (35) of the GO model can be computed as follows:

$$S_{GO}(t;\beta_1,\beta_2,x) = \left[1 + R_0(t)e^{x'(\beta_2-\beta_1)}\right]^{-1}. \tag{36}$$

2.2.5. Special Cases of the GO Model

All of the odds-based regression models listed above are incorporated into the GO model of odds-based models as special cases. The GO model can be used to derive the PO, AO, and AFT models, according to the following proposition:

Proposition 2. *Suppose $r(t;\beta,x)$ is given by Equation (33). Then, we have the following results:*

1. *If $\beta_1 = \beta_2$, then $r(t;\beta,x) = r_o\left(te^{x'\beta}\right)e^{\beta x'}$ giving the AFT model.*
2. *If $\beta_2 = 0$, then $r(t;\beta,x) = r_o(t)e^{x'\beta}$ giving the PO model.*
3. *If $\beta_1 = 0$, then $r(t;\beta,x) = r_o\left(te^{x'\beta}\right)$, giving the AO model.*

Proof of Proposition 2. The proof of Proposition 2 is straightforward. □

3. The Proposed Class

3.1. Why AM Class of Hazard-Based and Odds-Based Regression Models?

All of the hazard-based and odds-based regression models discussed in the preceding Section 2 can model different aspects of time-to-event data. However, determining which model is the most accurate and precise in revealing the correlation between explanatory variables and the baseline hazard (or the baseline odds) is a challenging issue and a significant research question that must be addressed.

In real life, we must decide between hazard-based regression models and odds-based regression models when provided with a dataset. A popular technique would be to fit one model to each of them, and then test the model to determine where it falls well short. However, the possibility of verifying the model assumptions may be constrained due to the finite sample size and other data characteristics. Additionally, if the right time-dependent covariates are taken into account, both the hazard-based models, such as the PH, AFT, AH, and GH models, and the odds-based models, such as the PO, AO, AFT, and GO models, may be able to fit the data relatively well.

Another issue with time-to-event data is that lifetimes can be censored in a variety of ways, including left, right, interval, double, and middle censoring, as well as survival data with crossover survival or hazard curves. Furthermore, a general class containing all of the preceding eight hazard-based and odds-based regression models is required. As a result,

it is difficult to address all of the aforementioned open topics using both frequentist and Bayesian methods.

To address the aforementioned problems and to fill the gap, we introduce the AM class of hazard-based and odds-based survival regression models, a unique, novel, tractable, universal, parametric class of survival regression models that encompasses all hazard-based and odds-based regression models to help applied statisticians to decide which model to fit in a given censored survival dataset. We estimate the model parameters using both frequentist and Bayesian approaches, and we evaluate the proposed model's nested structure using a likelihood ratio test.

In this section, we introduce the new survival regression model, its main probabilistic functions, and some special cases.

3.2. Model Formulation

Let T be a non-negative random variable that represents the length of time until an event of interest occurs. As already sketched, a universal class for hazard-based and odds-based regression models called the "Amoud Class (AM)" has the following closed form in order to accommodate survival data with or without the crossover of the hazard and survival curves:

$$R_{AM}(t;\beta_1,\beta_2,\beta_3,x) = e^{x'(\beta_2-\beta_1)}R_o\left(te^{x'\beta_1}\right)^{e^{x'(\beta_3-\beta_1)}}, \tag{37}$$

where $R_o(.)$ is the baseline odds function. This generality is attained using a structure resembling the general class of hazard-based regression models, with the addition that the baseline odds function is multiplied to a link function (i.e., log-linear function) for the covariates.

The sf for the AM model corresponding to the odds function in Equation (37) is expressed as follows:

$$S_{AM}(t;\beta_1,\beta_2,\beta_3,x) = \left[1+e^{x'(\beta_2-\beta_1)}\frac{F_o\left(te^{x'\beta_1}\right)}{S_0\left(te^{x'\beta_1}\right)}\right]^{-e^{x'(\beta_3-\beta_1)}}. \tag{38}$$

The hrf for the AM model corresponding to Equation (37) is computed as follows:

$$h_{AM}(t;\beta_1,\beta_2,\beta_3,x) = \frac{e^{x'(\beta_2+\beta_3-\beta_1)}h_0\left(te^{x'\beta_1}\right)}{e^{x'(\beta_2-\beta_1)}F_0\left(te^{x'\beta_1}\right)+S_0\left(te^{x'\beta_1}\right)}. \tag{39}$$

3.3. Probabilistic Functions for the Amoud Class Model

In terms of odds function, the sf for the AM model in Equation (38) can be expressed as follows:

$$S_{AM}(t;\beta_1,\beta_2,\beta_3,x) = \left[1+e^{x'(\beta_2-\beta_1)}R_o\left(te^{\beta_1 x'}\right)\right]^{-e^{x'(\beta_3-\beta_1)}}. \tag{40}$$

The derivative of the odds function for the AM model is expressed as follows:

$$r_{AM}(t;\beta_1,\beta_2,\beta_3,x) = r_o\left(te^{x'\beta_1}\right)e^{x'(\beta_2+\beta_1)}R_o\left(te^{\beta_1 x'}\right)^{e^{x'(\beta_3-\beta_1)}-1}. \tag{41}$$

The cumulative distribution function (cdf) for the SM model is computed as follows:

$$F_{AM}(t;\beta_1,\beta_2,\beta_3,x) = 1-\left[1+e^{x'(\beta_2-\beta_1)}\frac{F_o\left(te^{x'\beta_1}\right)}{S_0\left(te^{x'\beta_1}\right)}\right]^{-e^{x'(\beta_3-\beta_1)}}, \tag{42}$$

where the baseline hazard, odds, survival, cumulative distribution, and the derivative of the odds functions are $H_0(.)$, $R_0(.)$, $S_0(.)$, $F_0(.)$, and $r_0(.)$, respectively.

3.4. Special Sub-Models of the Proposed Class

All of the hazard-based and odds-based regression models listed above are incorporated into the AM Class of hazard-based and odds-based survival models as special cases. The AM class can be used to derive the PH, PO, AH, AO, AFT, GH, and GO models, according to the following proposition:

Proposition 3. *Suppose $R(t;\beta,x)$ is given by Equation (37). Then, we have the following results:*

1. *If $\beta_2=\beta_1$, then $R(t;\beta_1,\beta_2,x)=R_o\left(te^{x'\beta_1}\right)^{e^{x'(\beta_2-\beta_1)}}$, giving the GH model.*
2. *If $\beta_3=\beta_1$, then $R(t;\beta_1,\beta_2,x)=R_o\left(te^{x'\beta_1}\right)e^{x'(\beta_2-\beta_1)}$, which is the GO model.*
3. *If $\beta_3=\beta_2=\beta_1$, then $R(t;\beta,x)=R_o\left(te^{x'\beta}\right)$, giving the AFT model.*
4. *If $\beta_3=\beta_1=0$, then $R(t;\beta,x)=R_o(t)e^{x'\beta}$, which is the PO model.*
5. *If $\beta_2=\beta_1=0$, then $R(t;\beta,x)=R_o(t)^{e^{x'\beta}}$, giving the PH model.*
6. *If $\beta_3-\beta_1=0,\beta_2=0$, then $R(t;\beta,x)=R_o\left(te^{x'\beta}\right)^{e^{-x'\beta}}$, which is the AO model.*
7. *If $\beta_2-\beta_1=0,\beta_3=0$, then $R(t;\beta,x)=R_o\left(te^{x'\beta}\right)e^{-x'\beta}$, giving the AH model.*

Proof of Proposition 3. The proof of Proposition 3 is straightforward. □

Figure 1 illustrates the relationship between the proposed AM class and its sub-models including the GH, GO, AFT, AO, AH, PO, and PH models.

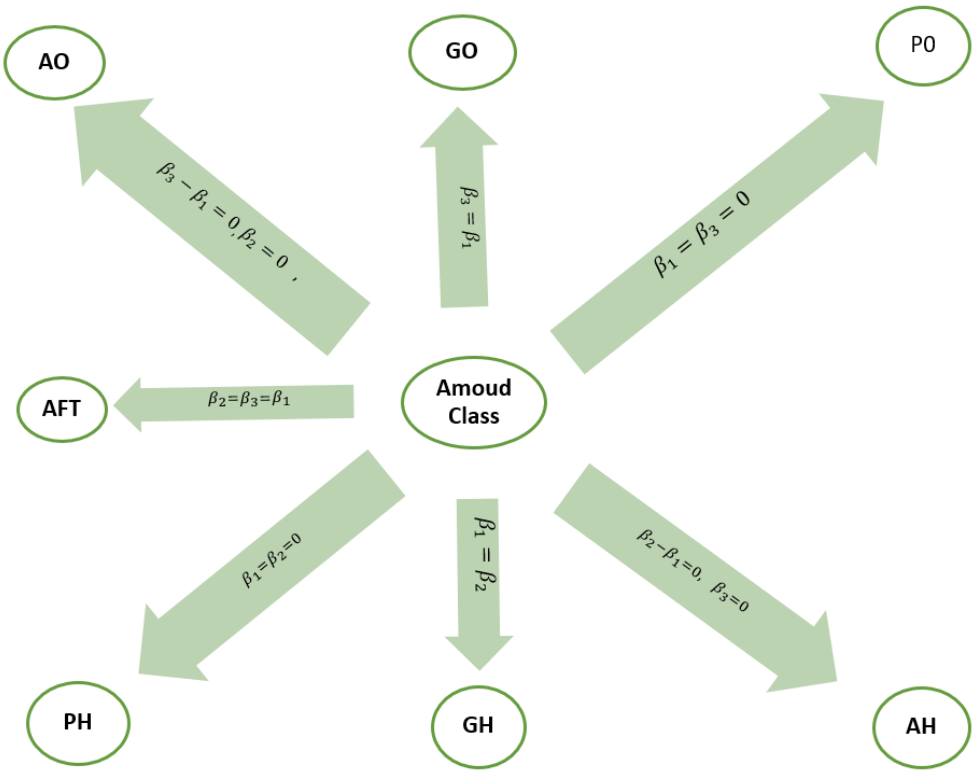

Figure 1. Visual graph illustrating the relationship between the proposed Amoud Class (AM) and its sub-models including the proportional hazard (PH), general odds (GO), general hazard (GH), accelerated failure time (AFT), accelerated odds (AO), accelerated hazard (AH), and proportional odds (PO) models.

4. Baseline Distributions

The Weibull distribution, LL distribution, and a GLL distribution that combines both of them are three different baseline distributions that are presented in this section. The closeness of the Weibull distribution under hazard-based regression models and the closeness of the LL distribution under odds-based regression models were also proven. When applied to censored survival data, the closeness of the distributions is what causes the regression models to produce comparable findings. We propose the use of a modified baseline distribution that demonstrates the differences between the survival regression models taken into consideration in this study because the Weibull and LL distributions have limitations and give the same results under different survival regression models.

4.1. Weibull Baseline for Hazard-Based Regression Models

The Weibull distribution is widely used as a baseline distribution in survival and reliability regression models. Its hrf is monotone. Moreover, hrf and sf can be derived analytically, and as such, censored data can be analyzed easily. Because of the tractability and flexibility of hazard and survival functions, the Weibull model is popular among researchers in survival and reliability analysis. However, the Weibull distribution has the limitation of not capable of accommodating non-monotone unimodal and bathtub-shaped hazard functions [37,38]. Another issue is that Weibull distribution is not a PO model, but this is the only distribution closed under the hazard-based regression models. This means that the PH, AH, and AFT models coincide when the baseline hrf is that of the Weibull distribution. This also means that the GH model is not identifiable.

The hrf and chf of the Weibull distribution are expressed as follows:

$$h_W(t;k,\alpha) = \alpha k(kt)^{\alpha-1}, \quad t \geq 0, \quad k,\alpha > 0, \tag{43}$$

$$H_W(t;k,\alpha) = (kt)^{\alpha}, \quad t \geq 0, \quad k,\alpha > 0, \tag{44}$$

where $k > 0$ and $\alpha > 0$ are the rate and shape parameters, respectively.

The odds function for the Weibull distribution is expressed as follows:

$$R_{LL}(t;k,\alpha) = \exp[H_W(t;k,\alpha)] - 1 = \exp((kt)^{\alpha}) - 1, \quad t \geq 0, \quad k,\alpha > 0. \tag{45}$$

The associated derivative of the odds function of the Weibull distribution is as follows:

$$r_W(t;k,\alpha) = h_W(t;\theta)\exp[H_W(t;\theta)] = \alpha k(kt)^{\alpha-1}\exp((kt)^{\alpha}), \quad t \geq 0, \quad k,\alpha > 0. \tag{46}$$

The Weibull accelerated failure time (W-AFT) model is defined as follows:

$$h_{W-AFT}(t;\beta,x) = H_0\left(te^{x'\beta}\right)e^{x'\beta} = e^{x'\beta}\alpha k\left(kte^{x'\beta}\right)^{\alpha-1} = \left(e^{x'\beta}k\right)^{\alpha}\alpha t^{\alpha-1} = \alpha k^* t^{\alpha-1}. \tag{47}$$

In Equation (47), we can observe the hrf for the Weibull in Equation (43). As mentioned, the scale parameter differs between groups and we can write it as: $T_i \sim Weibull\left(k^* = \left(e^{x'\beta}k\right)^{\alpha}, \alpha\right)$, with scale k^* and shape α.

On the other hand, if a Weibull distribution is assumed for T_i under the PH framework (W-PH) in Equation (1), it then follows that $T_i \sim$ Weibull $\left(k^* = e^{x'\beta}k^{\alpha}, \alpha\right)$, with scale k^* and shape α. The hrf for the W-PH model is rewritten as follows:

$$h_{W-PH}(t;\beta,x) = r_0(t)e^{x'\beta} = \alpha k(kt)^{\alpha-1}e^{x'\beta} = e^{x'\beta}k^{\alpha}\alpha t^{\alpha-1} = \alpha k^* t^{\alpha-1}. \tag{48}$$

This proves that the Weibull baseline is the only baseline distribution that is closed under all hazard-based regression models.

4.2. Log-Logistic Baseline for Odds-Based Regression Models

The LL distribution is a frequently used baseline distribution in survival and reliability regression models. Its hrf is monotone decreasing hazard and non-monotone unimodal. The LL model hazard and probability density shapes are similar to those of the log-normal distribution. but it has explicit algebraic expressions for the hazard rate and survival functions which makes it more suitable for the analysis of censored lifetime data than the log-normal distribution [39,40]. The LL distribution has the limitation of not being capable of accommodating monotone increasing and bathtub shaped hrfs. Another issue is that the LL distribution is not a PH model, but is the only distribution closed under the odds-based regression models. This means that PO, AO, and AFT models coincide when the baseline hazard is LL. This also makes the GO model not identifiable.

The cdf and sf of the LL distribution are expressed as follows:

$$F_{LL}(t;k,\alpha) = \frac{(kt)^{\alpha}}{1+(kt)^{\alpha}}, \quad t \geq 0, \quad k,\alpha > 0, \tag{49}$$

$$S_{LL}(t;k,\alpha) = \frac{1}{1+(kt)^{\alpha}}, \quad t \geq 0, \quad k,\alpha > 0, \tag{50}$$

where $k > 0$, and $\alpha > 0$ are the rate and shape parameters, respectively.

The odds function for the LL distribution is expressed as follows:

$$R_{LL}(t;k,\alpha) = \frac{F_{LL}(t;k,\alpha)}{S_{LL}(t;k,\alpha)} = \frac{\frac{(kt)^{\alpha}}{1+(kt)^{\alpha}}}{\frac{1}{1+(kt)^{\alpha}}} = (kt)^{\alpha}, \quad t \geq 0, \quad k,\alpha > 0. \tag{51}$$

The associated derivative of the odds function of the LL distribution is as follows:

$$r_{LL}(t;k,\alpha) = R'_{LL}(t;k,\alpha) = \alpha k(kt)^{\alpha-1}, \quad t \geq 0, \quad k,\alpha > 0. \tag{52}$$

It is obvious that the odds function for the LL distribution and its derivative are comparable to the chf and hrf functions for the Weibull distribution, respectively. Therefore, it is simple to illustrate that the odds-based regression models simply consider the LL distribution as a closed baseline distribution. the PO and AFT models, as examples.

According to Lawless [21], the LL distribution can be used to support a parametric AFT model, allowing scale parameter to differ between groups. For this, we need to keep the AFT structure that we mentioned above in the odds-based regression model formulation and adopt the derivative odds function of the LL distribution for Equation (52) for the reference group.

The log-logistic AFT (LL-AFT) derivative of the odds function is defined as follows:

$$r_{LL-AFT}(t;\beta,x) = r_0\left(te^{x'\beta}\right)e^{x'\beta} = e^{x'\beta}\alpha k\left(kte^{x'\beta}\right)^{\alpha-1} = \left(e^{x'\beta}k\right)^{\alpha}\alpha t^{\alpha-1} = \alpha k^* t^{\alpha-1}. \tag{53}$$

In Equation (53), we can observe the derivative of the odds structure for the LL distribution in (52). As mentioned, the scale parameter differs between groups and we can write it as $T_i \sim \log-\text{logistic}\left(k^* = \left(e^{x'\beta}k\right)^{\alpha}, \alpha\right)$, with scale k^* and shape α.

On the other hand, if a LL distribution is assumed for T_i under the PO framework (LL-PO) in Equation (20) it then follows that $T_i \sim \log-\text{logistic}\left(k^* = e^{x'\beta}k^{\alpha}, \alpha\right)$, with scale k^* and shape α. The derivative of the odds function for the LL-PO model is thus rewritten as follows:

$$r_{LL-PO}(t;\beta,x) = r_0(t)e^{x'\beta} = \alpha k(kt)^{\alpha-1}e^{x'\beta} = e^{x'\beta}k^{\alpha}\alpha t^{\alpha-1} = \alpha k^* t^{\alpha-1}. \tag{54}$$

This proves that the log-logistic distribution is the only baseline distribution that is closed under all odds-based regression models.

4.3. Generalized Log-Logistic Baseline for All Models

The GLL distribution [23,27,30,41,42] is an example of a baseline distribution that can incorporate both monotone and non-monotone hrfs, as well as be closed under both odds-based and hazard-based regression models, and has the benefit of including both the Weibull and LL models as sub-models [42].

The hrf and the odds function of the GLL distribution are expressed as follows:

$$h_{GLL}(t;k,\alpha,\eta) = \frac{\alpha k(kt)^{\alpha-1}}{[1+(\eta t)^{\alpha}]}, \quad t \geq 0, \quad k,\alpha,\eta > 0, \tag{55}$$

$$R_{GLL}(t;k,\alpha,\eta) = [1+(\eta t)^{\alpha}]^{\frac{k^{\alpha}}{\eta^{\alpha}}} - 1, \quad t \geq 0, \quad k,\alpha,\eta > 0, \tag{56}$$

where $k > 0$, $\alpha > 0$, and $\eta > 0$ are the distributional rate and shape parameters, respectively.

The hrf in Equation (55) consists of different sub-models of the GLL distribution [42].

- Log-logistic (LL) distribution: when $k = \eta$, Equation (55) reduces to the hrf of an LL distribution, which is

$$h_{LL}(t;k,\alpha) = \frac{\alpha k(kt)^{\alpha-1}}{[1+(kt)^{\alpha}]}, \quad t \geq 0, \quad k,\alpha > 0. \tag{57}$$

- Burr-XII (BXII) distribution: when $\eta = 1$, Equation (55) reduces to the hrf of a BXII-2 distribution, which is

$$h_{BXII}(t;k,\alpha) = \frac{\alpha k(kt)^{\alpha-1}}{[1+t^{\alpha}]}, \quad t \geq 0, \quad k,\alpha > 0. \tag{58}$$

- Weibull (W) distribution: when $\eta \to 0$, Equation (55) reduces to the hrf of the W distribution, which is

$$h_{W}(t;k,\alpha) = \alpha k(kt)^{\alpha-1}, \quad t \geq 0, \quad k,\alpha > 0. \tag{59}$$

5. Estimation Based on Frequentist and Bayesian Approaches

In this section, the unknown parameters of the proposed fully parametric AM class with GLL, LL and Weibull baseline distributions are estimated using frequentist MLE and Bayesian approaches.

5.1. MLE for Right-Censored Data

As was earlier indicated, not always an observed time will be a survival time: the subject is observed up to a particular time and is no longer followed up for a reason unrelated to the event occurrence. This is an illustration of a right-censored observed time, which was taken into consideration in this work and is the most common type of censoring in oncology studies. The same survival likelihood functions are reached despite the fact that there are many right-censoring techniques [21]. This ensures the identifiability of the distribution of the observed times under the further assumption that the survival times are independent random variables for all subjects (random censoring) and that the censoring times depend on no parameter associated with the survival function (non-informative censoring) [24].

These presumptions allow for the formulation of a general expression for the survival likelihood function. Assuming that a survival time $T_i = t_i$ or a censored time $C_i = c_i$ are recorded for each subject, $i, 1 \leq i \leq n$. Assume also that survival (censoring) times are independent among all subjects, i.e., $T_1, \ldots, T_n \sim F_T(t;\theta_T)(C_1, \ldots, C_n \sim F_C(c;\theta_C))$. The actual observable time is defined by $Y_i = \min(T_i, C_i)$, whose distribution is indexed by a vector $\theta = (\theta_T, \theta_C)$ of parameters. Then, the information of a subject i is given by the

pair (Y_i, δ_i), where $\delta_i = \mathbb{I}_{T_i<c_i}$ being the censoring indicator random variable. For a pair $(Y_i = t_i, \delta_i = 1\ )$ (a survival observed time), the likelihood contribution is given by:

$$\begin{aligned}\lim_{\varepsilon\to 0^+}\frac{1}{2\varepsilon}P(y_i-\varepsilon<Y_i<y_i+\varepsilon,\delta_i=1;\theta) &= \lim_{\varepsilon\to 0^+}\frac{1}{2\varepsilon}P(y_i-\varepsilon<T_i<y_i+\varepsilon,T_i\le C_i;\theta)\\ &= \lim_{\varepsilon\to 0^+}\frac{1}{2\varepsilon}\int_{y_i-\varepsilon}^{y_i+\varepsilon}\int_t^{\infty} dF_C(c;\theta_C)dF_T(t;\theta_T) \text{ (independence)}\\ &= \lim_{\varepsilon\to 0^+}\frac{1}{2\varepsilon}\int_{y_i-\varepsilon}^{y_i+\varepsilon}[1-F_C(c;\theta_C)]dF_T(t;\theta_T)\\ &= [1-F_C(y_i;\theta_C)]f_T(y_i;\theta_T).\end{aligned} \tag{60}$$

On the other hand, the likelihood contribution for a pair $(Y_i = c_i, \delta_i = 0)$ (right censored observed time), the likelihood contribution is provided by

$$\begin{aligned}\lim_{\varepsilon\to 0^+}\frac{1}{2\varepsilon}P(y_i-\varepsilon<Y_i<y_i+\varepsilon,\delta_i=0;\theta) &= \lim_{\varepsilon\to 0^+}\frac{1}{2\varepsilon}P(y_i-\varepsilon<C_i<y_i+\varepsilon,T_i>C_i;\theta)\\ &= [1-F_T(y_i;\theta_T)]f_C(y_i;\theta_C).\end{aligned} \tag{61}$$

Thus, under a random right censoring, the survival likelihood function for a sample $y = (y_1, ..., y_n)$ of size n has the following expression:

$$L(\theta;y) = \prod_{i=1}^{n}\{[1-F_C(y_i;\theta_C)]f_T(y_i;\theta_T)\}^{\delta_i}\{[1-F_T(y_i\mid\theta_T)]f_C(y_i;\theta_C)\}^{1-\delta_i}. \tag{62}$$

Assuming that censoring is non-informative, i.e., the distribution of the censoring times does not depend on the parameters θ_T from the survival function, the factors $[1-F_C(y_i;\theta_C)]^{\delta_i}$ and $[f_C(y_i;\theta_C)]^{1-\delta_i}$ do not give any information for inference and can be dropped from Equation (62). Thereby, $\theta = \theta_T$ and a simpler survival likelihood function is given by

$$\begin{aligned}L(\theta,\beta;D) &= \prod_{i=1}^{n}[f(t_i;\theta,\beta,x)]^{\delta_i}[S(t_i;\theta,\beta,x)]^{1-\delta_i}\\ &= \prod_{i=1}^{n}\left[\frac{h(t_i;\theta,\beta,x)}{S(t_i;\theta,\beta,x)}\right]^{\delta_i}[S(t_i;\theta,\beta,x)]^{1-\delta_i}\\ &= \prod_{i=1}^{n}[h(t_i;\theta,\beta,x)]^{\delta_i}S(t_i;\theta,\beta,x)\\ &= \prod_{i=1}^{n}[h(t_i;\theta,\beta,x)]^{\delta_i}\exp[-H(t_i;\theta,\beta,x)],\end{aligned} \tag{63}$$

where $D = (t_i, \delta_i, x_i, i = 1, 2, \ldots, n)$ represents the observed data including t_i = survival time, δ_i = censoring time, θ is the vector of baseline distributional parameters, and x_i = covariates. The maximum likelihood estimation can be generated via an iterative optimization method (e.g., the Newton-Raphson algorithm).

The above formulation in Equation (63) is useful for modelling hazard-based regression models, like the PH, AH, and GH models. An alternative version can be obtained only in terms of the odds function and its derivative as follows:

$$\begin{aligned} L(\theta,\beta;D) &= \prod_{i=1}^{n}[f(t_i;\theta,\beta,x)]^{\delta_i}[S(t_i;\theta,\beta,x)]^{1-\delta_i} \\ &= \prod_{i=1}^{n}\Big[r(t_i;\theta,\beta,x)S(t_i;\theta,\beta,x)^2\Big]^{\delta_i}[S(t_i;\theta,\beta,x)]^{1-\delta_i} \\ &= \prod_{i=1}^{n}[r(t_i;\theta,\beta,x)]^{\delta_i}[S(t_i;\theta,\beta,x)]^{1-\delta_i} \\ &= \prod_{i=1}^{n}\left[\frac{r(t_i;\theta,\beta,x)}{1+R(t_i;\theta,\beta,x)}\right]^{\delta_i}[S(t_i;\theta,\beta,x)], \\ &= \prod_{i=1}^{n}\left[\frac{r(t_i;\theta,\beta,x)}{1+R(t_i;\theta,\beta,x)}\right]^{\delta_i}\exp[-H(t_i;\theta,\beta,x)] \\ &= \prod_{i=1}^{n}\left[\frac{r(t_i;\theta,\beta,x)}{1+R(t_i;\theta,\beta,x)}\right]^{\delta_i}\left[\frac{1}{1+R(t_i;\theta,\beta,x)}\right]. \end{aligned} \tag{64}$$

The log-likelihood function corresponding to Equation (63) is written as follows:

$$\ell(\theta,\beta;D) = \sum_{i=1}^{n}\delta_i\log[H_0(t_i;\theta,\beta,x)] - \sum_{i=1}^{n}H_0(t_i;\theta,\beta,x). \tag{65}$$

5.2. The Log-Likelihood Functions

Let $\theta = (\alpha,k)'$, $\Xi = \left(\beta_1',\beta_2',\beta_3'\right)'$, $\Omega = (\theta',\Xi')'$, $a_i = e^{x_i'(\beta_2+\beta_3-\beta_1)}$, $b_i = e^{x_i'(\beta_2-\beta_1)}$, $c_i = e^{x_i'(\beta_3-\beta_1)}$, and $d_i = k.t_i.x_i'\beta_1$ and assume the Weibull baseline distribution, then the log-likelihood function for the Weibull-AM (W-AM) model is

$$\begin{aligned} \ell(\mathbf{\Omega}) = \sum_{i=1}^{n}\delta_i\log\alpha + \sum_{i=1}^{n}\delta_i\log k + (\alpha-1)\sum_{i=1}^{n}\delta_i\log d_i \\ + \sum_{i=1}^{n}\delta_i\log a_i - \sum_{i=1}^{n}\delta_i\log[\{b_i(1-exp(-d_i)^{\alpha})\} + exp(-d_i)^{\alpha}] \\ - \sum_{i=1}^{n}\log\left[c_i\left(1+b_i\left\{\frac{1-exp(-d_i)^{\alpha}}{exp(-d_i)^{\alpha}}\right\}\right)\right]. \end{aligned} \tag{66}$$

The log-likelihood function for the LL baseline distribution under the AM class can be expressed as follows:

$$\begin{aligned} \ell(\mathbf{\Omega}) = \sum_{i=1}^{n}\delta_i\log\alpha + \sum_{i=1}^{n}\delta_i\log k + (\alpha-1)\sum_{i=1}^{n}\delta_i\log(d_i) - \sum_{i=1}^{n}\delta_i\log(1+d_i^{\alpha}) \\ + \sum_{i=1}^{n}\delta_i\log a_i - \sum_{i=1}^{n}\delta_i\log\left[\left\{b_i\left(\frac{d_i^{\alpha}}{1+d_i^{\alpha}}\right)\right\} + \frac{1}{(1+d_i^{\alpha})}\right] \\ - \sum_{i=1}^{n}\log[c_i(1+b_i\{d_i^{\alpha}\})]. \end{aligned} \tag{67}$$

Moreover, assuming $\theta = (\alpha,k,\eta)'$, $m_i = \eta.t_i.x_i'\beta_1$, and regarding the GLL baseline distribution under the AM class, the log-likelihood function can be expressed as follows:

$$\ell(\mathbf{\Omega}) = \sum_{i=1}^{n} \delta_i \log \alpha + \sum_{i=1}^{n} \delta_i \log k + (\alpha - 1) \sum_{i=1}^{n} \delta_i \log(d_i) - \sum_{i=1}^{n} \delta_i \log(1 + m_i^{\alpha}) + \sum_{i=1}^{n} \delta_i \log a_i - \sum_{i=1}^{n} \delta_i \log \left[\left\{ b_i \left(1 - \{1 + m_i^{\alpha}\}^{-\left(\frac{\kappa}{\eta}\right)^{\alpha}} \right) \right\} + \left(\{1 + m_i^{\alpha}\}^{-\left(\frac{\kappa}{\eta}\right)^{\alpha}} \right) \right] - \sum_{i=1}^{n} \log \left[c_i \left\langle 1 + b_i \{1 + m_i^{\alpha}\}^{\left(\frac{\kappa}{\eta}\right)^{\alpha}} - 1 \right\rangle \right]. \tag{68}$$

5.3. Bayesian Inference

In this section, we offer general guidelines for prior selection of the regression coefficients associated with covariates and baseline distribution parameters. We examined a prior independent scenario between the baseline parameters in H_0(t) (baseline hazard) or $R_0(t)$ (baseline odds) and the regression coefficients. Additionally, we determined the prior independence of the regression coefficients in a non-informative scenario with normal distributions of zero mean and a large known variance [43] as

$$\pi(H_0, \beta_1, \beta_2, \beta_3) = \pi(H_0)\pi(\beta_1, \beta_2, \beta_3) = \pi(H_0) \prod_{j=1}^{J} \mathrm{N}\left(\beta_j \mid 0, \sigma_j^2\right), \tag{69}$$

where $\pi(H_0)$ is the prior distribution of all baseline parameters and hyperparameters in H_0(t).

For the baseline hazard parameter θ in baseline distributions, we consider the following priors:

$$\pi(\alpha) \sim G(a_1, b_1) = \frac{b_1^{a_1}}{\Gamma(a_1)} \alpha^{a_1 - 1} e^{-b_1 \alpha}; a_1, b_1, \alpha > 0, \tag{70}$$

$$\pi(\eta) \sim G(a_2, b_2) = \frac{b_2^{a_2}}{\Gamma(a_2)} \eta^{a_2 - 1} e^{-b_2 \eta}; a_2, b_2, \eta > 0, \tag{71}$$

$$\pi(k) \sim G(a_3, b_3) = \frac{b_3^{a_3}}{\Gamma(a_3)} k^{a_3 - 1} e^{-b_3 k}; a_3, b_3, k > 0. \tag{72}$$

The values of the hyper-parameters values of the prior distributions are selected from the historical data of the baseline distribution [42].

For the regression coefficients prior, we have

$$\pi(\beta_1') \sim N(a_4, b_4), \tag{73}$$

$$\pi(\beta_2') \sim N(a_5, b_5), \tag{74}$$

$$\pi(\beta_3') \sim N(a_6, b_6). \tag{75}$$

The joint prior distribution for the distributional parameters and coefficient of regression expressed as follows:

$$\pi(\alpha, k, \eta, \beta_1', \beta_2') = \pi(\alpha)\pi(\eta)\pi(k)\pi(\beta_1')\pi(\beta_2')\pi(\beta_3'). \tag{76}$$

The model must be supplied with data $\mathcal{D} = \{(t_i, \delta_i, \boldsymbol{x}_i), i = 1, \ldots, n\}$, where t_i is the observed lifetime time for the ith individual, δ_i is the censoring status taking 1 if the event of interest has occurred and 0 otherwise, and $\boldsymbol{x}_i$ are the explanotory variables.

Prior knowledge and experimental data are combined in the posterior distribution via the Bayes' theorem, and we get

$$\pi(H_0, \beta_1, \beta_2, \beta_3; \mathcal{D}) \propto \mathcal{L}(H_0, \beta_1, \beta_2, \beta_3)\pi(H_0, \beta_1, \beta_2, \beta_3), \tag{77}$$

where $\mathcal{L}(H_0, \beta_1, \beta_2, \beta_3)$ is the likelihood function of $(H_0, \beta_1, \beta_2, \beta_3)$ given in Equation (63).

This study uses Markov chain Monte Carlo (McMC) techniques for Bayesian inference, and the Metropolis within the Gibbs algorithm is used to sample from the posterior distribution [44]. In our implementation, the independence sampler is used to update each parameter component [45].

6. Model Comparison

6.1. Classical Model Comparison

The comparison between GH, AFT, AH, and PH models based on the GLL baseline hazard was evaluated using different information criteria, and the nested structure of the GH model and its special cases was evaluated using the likelihood ratio test (LRT) as discussed below:

6.1.1. Nested Models

When the models are nested, we can compare them using a LRT. Assume we have two models: $f = f(t \mid \theta \in \Theta)$ and $f_0 = f(t \mid \theta_0 \in \Theta_0 \subset \Theta)$, where $\dim(\Theta) = m$ and $\dim(\Theta_0) = m - r$, respectively. In other words, f is reduced to f_0 by adjusting r of its parameters to constants. Wilks [46] proved that the LRT is expressed as follows:

$$LRT = -2\log\frac{L(\hat{\theta})}{L(\hat{\theta_0})} = 2\left[\ell(\hat{\theta}) - 2\ell_0\left(\widehat{\theta_0}\right)\right] \sim x^2{}_r, \tag{78}$$

where $\hat{\theta}$ is the restricted Maximum likelihood (M L) estimates under the null hypothesis (H_0) and $\hat{\theta}_0$ is the unrestricted ML estimates under the alternative hypothesis (H_1), L is the likelihood function, and ℓ is the log-likelihood function.

In our case, the AM model of hazard-based and odds-based regression models has seven sub-models, namely; GH, GO, PH, AH, PO, AO, and AFT models. In order to assess the following hypotheses, we used the likelihood ratio criterion:

i. H_0: $\beta_2 = \beta_1$, that is the sample is from the GH model.
   H_1 : if H_0 is false, then the sample is from the AM model.
ii. H_0: $\beta_1 = \beta_3$, that is the sample is from the GO model.
   H_1 : if H_0 is false, then the sample is from the AM model.
iii. H_0: $\beta_1 = \beta_2 = \beta_3 = 0$, that is the sample is from the AFT model.
   H_1 : if H_0 is false, then the sample is from the AM model.
iv. H_0: $\beta_1 = \beta_3 = 0$, that is the sample is from the PO model.
   H_1 : if H_0 is false, then the sample is from the AM model.
v. H_0: $\beta_1 = \beta_2 = 0$, that is the sample is from the PH model.
   H_1 : if H_0 is false, then the sample is from the AM model.
vi. H_0: β_3–$\beta_1 = 0$, $\beta_2 = 0$, that is the sample is from the AO model.
   H_1 : if H_0 is false, then the sample is from the AM model.
vii. H_0: β_2–$\beta_1 = 0$, $\beta_3 = 0$, that is the sample is from the AH model.
   H_1 : if H_0 is false, then the sample is from the AL model.

Under the null hypothesis, the LRT follows the Chi-square distribution with degrees of freedom (df) $(df_{\text{alt}} - df_{\text{null}})$. If the p-value is less than 0.05 the null hypothesis is rejected. In other words, if $LRT > X^2_{r,1-\tau}$, we conclude that the fit provided by f is significantly better than f_0 (at the τ level of significance).

6.1.2. Non-Nested Model

More generally, models can be non-nested, which means that there is no parameter configuration that makes the two models' equivalent. As a result, we are unable to use the likelihood ratio test. The Akaike information criterion (AIC) is one of the most extensively used methods for comparing non-nested models. The AIC rewards goodness of fit but penalizes the model for increasing the number of estimated parameters and is expressed as follows:

$$AIC = 2(j + p) - 2l, \tag{79}$$

where l represents the log-likelihood function evaluated as the MLEs, p the number of covariates, and j the number of distributional parameters of the assumed baseline probability distribution (i.e., $j = 3$ for the GLL distribution). Burnham and Anderson [47] provided some basic rules of thumb for the use of AIC as summarized in Table 1.

Table 1. Rule of thumb for AIC differences.

Δ_M	Level of Support of Model M
0–2	Substantial
4–7	Considerably less
>10	Essentially none

Other approaches for model comparison tools for both nested and non-nested models to decide which model best fits the provided data are available. Specifically, Bozdogan's consistent AIC (BCAIC), the Bayesian information criterion (BIC), the CAIC (Consistent AIC), and the Hannan Quin information criterion (HQIC).

In scenarios where the sample size is fairly small when compared to the number of parameters in the model, the CAIC fixes the AIC for overfitting of the data and is calculated as follows:

$$CAIC = AIC + \frac{2(j+p)(j+p+1)}{n-(j+p)-1}. \tag{80}$$

Contrary to the AIC, which is asymptotically efficient, the HQIC is frequently quoted in the literature. It is calculated as follows:

$$HQIC = 2(j+p)\log(\log(n)) - 2l. \tag{81}$$

The BCAIC is another adjusted form of AIC which is consistent and is computed as follows:

$$BCAIC = 2(j+p)\log(\log(n)) - 2l. \tag{82}$$

The BIC also known as the Schwarz information criterion [48], is used in the same way as AIC (we aim to minimize its value) but has a larger penalty for complexity when $n \geq 8$ (which is typically is). The BIC is computed as follows:

$$BIC = (j+p)\log(n) - 2l. \tag{83}$$

6.2. Bayesian Model Comparison

The Watanabe-Akaike information criterion (WAIC) and Leave-one-out cross-validation information criterion (LOOIC) were employed as full Bayesian model selection criteria in this study. They are both techniques for calculating pointwise out-of-sample prediction accuracy using a fitted Bayesian model. Asymptotically, they are equivalent since WAIC is based on the series expansion of leave-one-out cross-validation (LOO). It is helpful to be able to compute both WAIC and cross-validation because they address different prediction questions with finite data. The log-likelihood assessed from the posterior simulations of the parameter values can be used to directly estimate the WAIC and an approximated LOO based on importance sampling. Compared to more basic estimates of prediction error like AIC and DIC, LOOIC and WAIC have a number of advantages, but they are less frequently employed in practice since they require additional computing steps [49,50].

7. Practical Illustrations

A clinical trial right-censored oncology dataset is examined in this section to demonstrate the applicability and tractability of the proposed models, including the fully-parametric AM class, GO, and AO models with three different baseline distributions, including Weibull, LL, and GLL baseline distributions in modelling right-censored survival data with crossing survival curves. We compared the proposed AM class with its sub-models that contain both hazard-based regression models, including PH, AH, AFT, and GH models, and the

odds-based regression models, including PO, AO, AFT, and GO models, using both the MLE frequentist and Bayesian approaches using noninformative priors. The class and its sub-models were compared using different information criteria, including the classical ones (AIC, BIC, BCAIC, CAIC, and HQIC), Bayesian model selection (WAIC, and LOOIC), and checking the nested structure of the AM class using the LRT test.

7.1. IPASS Clinical Trial Data Set

In order to show the applicability of the proposed models, we re-analyzed a large dataset from a randomized clinical trial called IPASS for this study. In a randomized controlled trial, gefitinib vs. carboplatin-paclitaxel was compared for progression-free survival in patients with advanced pulmonary adenocarcinoma. An unadjusted PH model was used to examine the main outcome. Despite the implicit violation of the PH model assumption represented by the crossing of the two survival curves, the study's findings were published using this model [51].

Argyropoulos and Unruh [52] reconstructed and re-published the IPASS dataset, and it is now freely available in an AHSurv R package [31].The features stated in the references are all still there in this reconstructed dataset, which is also accessible to the clinical trial's results. The months of March 2006 through April 2008 are covered by the database. The main objective of the trial is to evaluate the effects of gefitinib versus carboplatin/paclitaxel doublet chemotherapy on progression-free survival (in months) in a subset of patients with non-small-cell lung cancer (NSCLC). According to the trial's design, $n = 1207$ previously untreated individuals in East Asia with advanced lung adenocarcinoma and who were non-smokers or previous light smokers were randomly assigned to either carboplatin + paclitaxel (608 patients) or gefitinib (608 patients) (609 patients). The observations show 965 occurrences of the event of interest (79.3 percent), with 449 (73.7 percent) relating to patients receiving gefitinib and 516 (84.9 percent) related to patients receiving carboplatin+paclitaxel.

The primary goal of this section is to appropriately assess the rebuilt IPASS data and estimate the regression coefficients using the proposed fully-parametric AM class provided in Section 3. For the proposed model, we evaluate both the maximum likelihood and the Bayesian estimating approaches to achieve this goal.

We fit all hazard-based and odds-based regression models as well as the general proposed AM class using three different baseline distributions, namely Weibull, LL, and GLL distributions, letting $x_i = I$(treatment = chemotherapy), which equals 1 if the treatment involves gefitinib and 0 if the treatment involves carboplatin/paclitaxel. Tables 2–4 provide a summary of the numerical results.

Figure 2 displays the total time on test (TTT) plot for the survival time and the survival curves of the two types of drugs where crossing between the curves can be seen, which confirms the efficacy of the proposed novel models in this study, including the AM, GO, and AO models, plus some other existing models in the literature, including the GH, and AH models, and that it is appropriate for the analysis of survival data with crossover survival curves.

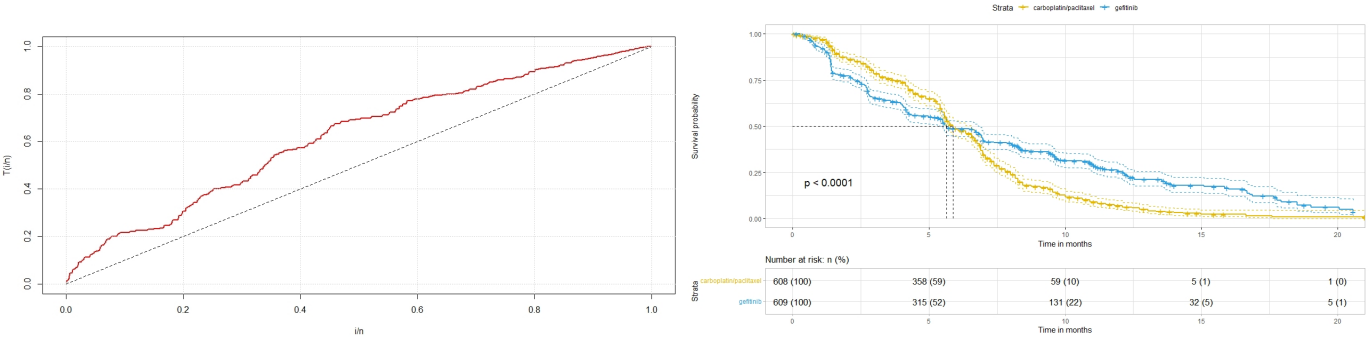

Figure 2. Illustrating the total time on test (TTT) plot, and the crossing survival curves for the two types of drugs determined using the Kaplan-Meier method for the IPASS dataset.

Table 2. Results from the fitted proposed fully-parametric odds-based and hazard-based regression models with W baseline distribution to IPASS dataset.

Models	Parameter(s)	Estimate	SE	AIC	BCAIC	BIC	CAIC	HQIC
W-AM	β_1	0.167	0.008	5552.888	5583.409	5578.409	5552.849	5562.495
	β_2	−1.346	0.109					
	β_3	2.815	0.129					
	α	1.845	0.002					
	κ	7.003	0.009					
W-GH	β_1	1.398	0.007	5686.968	5711.385	5707.385	5686.943	5694.653
	β_2	−0.862	0.129					
	α	1.362	0.129					
	κ	6.613	0.129					
W-GO	β_1	−2.574	0.008	5609.716	5634.133	5630.133	5609.691	5617.402
	β_2	1.921	0.011					
	α	1.596	0.003					
	κ	6.870	0.005					
W-AH	β	−0.875	0.005	5684.474	5702.787	5699.787	5684.460	5690.239
	α	1.365	0.004					
	κ	6.762	0.003					
W-PH	β	−0.320	0.007	5684.474	5702.787	5699.787	5684.460	5690.239
	α	1.365	0.004					
	κ	6.762	0.003					
W-AO	β	−0.567	0.001	5652.746	5671.058	5668.058	5652.731	5658.510
	α	1.419	0.003					
	κ	6.515	0.002					
W-PO	β	−0.003	0.017	5708.640	5726.952	5723.952	5708.625	5714.404
	α	1.341	0.010					
	κ	7.594	0.006					
W-AFT	β	−0.234	0.009	5684.474	5702.787	5699.787	5684.460	5690.239
	α	1.365	0.004					
	κ	6.762	0.003					

Table 3. Results from the fitted proposed fully-parametric odds-based and hazard-based regression models with LL baseline distribution to IPASS dataset.

Models	Parameter(s)	Estimate	SE	AIC	BCAIC	BIC	CAIC	HQIC
LL-AM	β_1	2.185	0.030	5690.051	5720.572	5715.572	5690.014	5699.658
	β_2	−0.361	0.009					
	β_3	−0.171	0.011					
	α	2.291	0.022					
	κ	5.498	0.006					
LL-GH	β_1	1.074	0.018	5688.051	5712.468	5708.468	5688.027	5695.737
	β_2	−0.088	0.019					
	α	2.291	0.029					
	κ	5.498	0.109					
LL-GO	β_1	−0.845	0.038	5757.552	5781.968	5777.968	5757.527	5765.237
	β_2	0.726	0.129					
	α	1.801	0.129					
	κ	5.478	0.129					

Table 3. *Cont.*

Models	Parameter(s)	Estimate	SE	AIC	BCAIC	BIC	CAIC	HQIC
LL-AH	β	1.119	0.014	5687.976	5706.289	5703.289	5687.961	5693.740
	α	2.259	0.008					
	κ	5.614	0.009					
LL-PH	β	−0.128	0.013	5751.598	5769.911	5766.911	5751.584	5757.362
	α	1.832	0.010					
	κ	5.156	0.009					
LL-AO	β	0.061	0.007	5755.552	5773.864	5770.864	5755.537	5761.316
	α	1.801	0.009					
	κ	5.478	0.010					
LL-PO	β	0.049	0.004	5755.552	5773.864	5770.864	5755.537	5761.316
	α	1.801	0.009					
	κ	5.478	0.010					
LL-AFT	β	0.027	0.009	5755.552	5773.864	5770.864	5755.537	5761.316
	α	1.801	0.009					
	κ	5.478	0.010					

Table 4. Results from the fitted proposed fully-parametric odds-based and hazard-based regression models with GLL baseline distribution to IPASS dataset.

Models	Parameter(s)	Estimate	SE	AIC	BCAIC	BIC	CAIC	HQIC
GLL-AM	β_1	0.179	0.003	5554.864	5591.489	5585.489	5554.810	5566.392
	β_2	−1.344	0.001					
	β_3	2.824	0.002					
	α	1.853	0.129					
	κ	0.143	0.129					
	η	0.010	0.129					
GLL-GH	β_1	2.049	0.001	5605.079	5635.600	5630.600	5605.041	5614.686
	β_2	−0.835	0.002					
	α	1.818	0.129					
	κ	0.151	0.129					
	η	0.045	0.129					
GLL-GO	β_1	−0.635	0.002	5645.507	5676.028	5671.028	5645.470	5655.114
	β_2	0.339	0.001					
	α	1.364	0.129					
	κ	0.145	0.129					
	η	0.000	0.129					
GLL-AH	β	1.545	0.001	5656.576	5680.993	5676.993	5656.551	5664.262
	α	1.886	0.129					
	κ	0.154	0.129					
	η	0.101	0.129					
GLL-PH	β	−0.064	0.014	5686.187	5710.604	5706.604	5686.162	5693.872
	α	−0.171	0.129					
	κ	−0.171	0.129					
	η	−0.171	0.129					
GLL-AO	β	−0.708	0.002	5661.339	5685.755	5681.755	5661.314	5669.024
	α	1.482	0.129					
	κ	0.158	0.129					
	η	0.000	0.129					
GLL-PO	β	−0.008	0.021	5708.832	5733.249	5729.249	5708.808	5716.518
	α	1.406	0.129					
	κ	0.139	0.129					
	η	0.031	0.129					
GLL-AFT	β	−0.166	0.014	5688.582	5712.999	5708.999	5688.557	5696.267
	α	1.359	0.129					
	κ	0.143	0.129					
	η	0.000	0.129					

The parameters and related standard errors from the various hazard- and odds-based regression models employing the Weibull baseline distribution and five different information criterion estimations are shown in Table 2. The TTT plot in Figure 2 shows that the data's increasing hazard rate points to the theoretical use of the Weibull baseline distribution. According to the findings in Table 2 and Figure 3, the W-AM class has the lowest values for each information criterion when compared to all other competing regression models, demonstrating its superiority over other hazard-based and odds-based regression models. Another crucial aspect that restricts the use of the Weibull baseline distribution is the fact that all hazard-based regression models yield the same result, which is a weakness of the Weibull baseline distribution.

All hazard-based regression models, including the AH, PH, AFT, and GH models, produce the same findings when compared to the Weibull baseline as illustrated in Table 2. The LL baseline distribution was used to fit and compare all of the regression models after we looked at an alternate baseline distribution. Using the LL baseline distribution and five different information criteria, Table 3 provides estimates of the parameters and related standard errors from the various hazard-based and odds-based regression models. According to the AIC values in Table 3, there is no clear preference for one model over the other. The fact that all odds-based regression models yield the same result is a significant factor that restricts the applicability of the LL baseline distribution.

When the Weibull distribution is used as the baseline distribution, all hazard-based regression models exhibit coincidence as shown in Table 2. On the other hand, when the baseline distribution is an LL distribution, all odds-based regression models exhibit coincidence as illustrated in Table 3, and Figure 4.

These two points recommend looking for and utilizing a modified baseline distribution, which can provide us with various results for all survival regression models, regardless of whether they are hazard-based, odds-based, or a combination of both. We used the GLL baseline distribution, where a sub-model of the Weibull distribution that yields different results for all the regression models taken into consideration in this study, to close the gap and compare the seven different hazard-based and odds-based regression models that are currently in use.

For the proposed AM class and seven different hazard- and odds-based regression models with GLL baseline distribution, the parameter estimates and their associated standard errors are shown in Table 4. We see that for the eight competing models, the estimates of the baseline distribution parameters and their standard errors are quite similar and within a reasonable range. The GLL-AM model appears to be preferred above the other competing models and provided the best-fitting model, according to the values of the five distinct information criteria. The results also showed that the GH and GO models are preferred over their sub-models. Finally, the results indicate that the only basic survival regression models that can be used to model and analyze survival data with crossing survival curves are the AH and AO models.

The GLL-AM regression model is the best model compared to the others, according to the LRT results in Table 5. The previously stated result is supported by the plots of the estimated hazards in Figure 5.

Table 5. LRT values for the AM class and its sub-models using IPASS dataset.

Model	Hypothesis	LRT Statistic	*p*-Value
GH	H_0: $\beta_2 = \beta_1$, H_1: H_0 is false,	52.214	<0.0001
GO	H_0: $\beta_3 = \beta_1$, H_1: H_0 is false,	92.644	<0.0001
AH	H_0: $\beta_2 - \beta_1 = 0, \beta_3 = 0$, H_1: H_0 is false,	105.712	<0.0001
AO	H_0: $\beta_3 - \beta_1 = 0, \beta_2 = 0$, H_1: H_0 is false,	135.322	<0.0001
PH	H_0: $\beta_1 = \beta_2 = 0$, H_1: H_0 is false,	110.474	<0.0001
PO	H_0: $\beta_1 = \beta_3 = 0$, H_1: H_0 is false,	157.968	<0.0001
AFT	H_0: $\beta_1 = \beta_2 = \beta_3$, H_1: H_0 is false,	137.718	<0.0001

According to the LRT results in Table 6, the GLL-GH regression model is the most effective of the alternatives hazard-based regression models.

Table 6. LRT values for the GH model and its sub-models using IPASS dataset.

Models	Hypothesis	LRT Statistic	*p*-Value
GH vs. AH	H_0: $\beta_2 = 0$, H_1: H_0 is false,	53.498	<0.0001
GH vs. PH	H_0: $\beta_1 = 0$, H_1: H_0 is false,	83.108	<0.0001
GH vs. AFT	H_0 $\beta_1 = \beta_2$, H_1: H_0 is false,	85.504	<0.0001

According to the LRT results in Table 7, the GLL-GO regression model is the most effective of the alternatives odds-based regression models.

Table 7. LRT values for the GO Model and its sub-models using IPASS dataset.

Models	Hypothesis	LRT Statistic	*p*-Value
GO vs. AO	H_0: $\beta_2 = 0$, H_1: H_0 is false,	17.830	<0.0001
GO vs. PO	H_0: $\beta_1 = 0$, H_1: H_0 is false,	65.324	<0.0001
GO vs. AFT	H_0: $\beta_1 = \beta_2$, H_1: H_0 is false,	45.074	<0.0001

The W-AM model is the best model among the others, according to the hazard plots of the estimated hazard rates in Figure 3, while other hazard-based models with the Weibull baseline distribution fitted similarly and there was no more difference at all. According to the plots in Figure 4, there is no model that is preferred over the others when the baseline distribution is the LL distribution, and there is no difference between any of the odds-based regression models that were fitted. Finally, as illustrated in Figure 5, the GLL-AM model is the superior compared to the other competitive models.

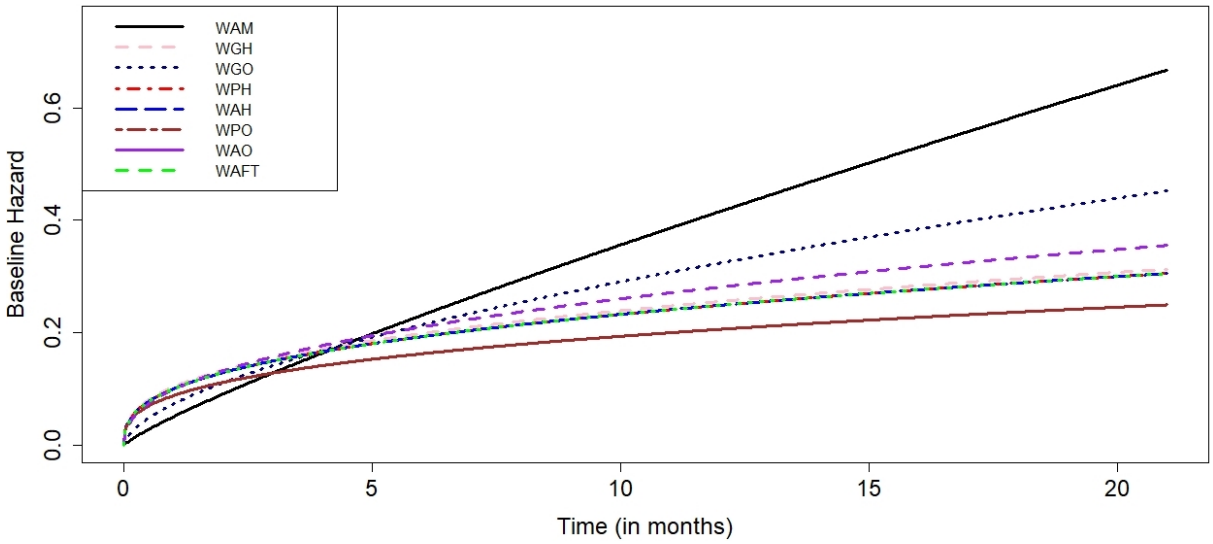

Figure 3. Estimated hrfs for the competitive models of the IPASS dataset.

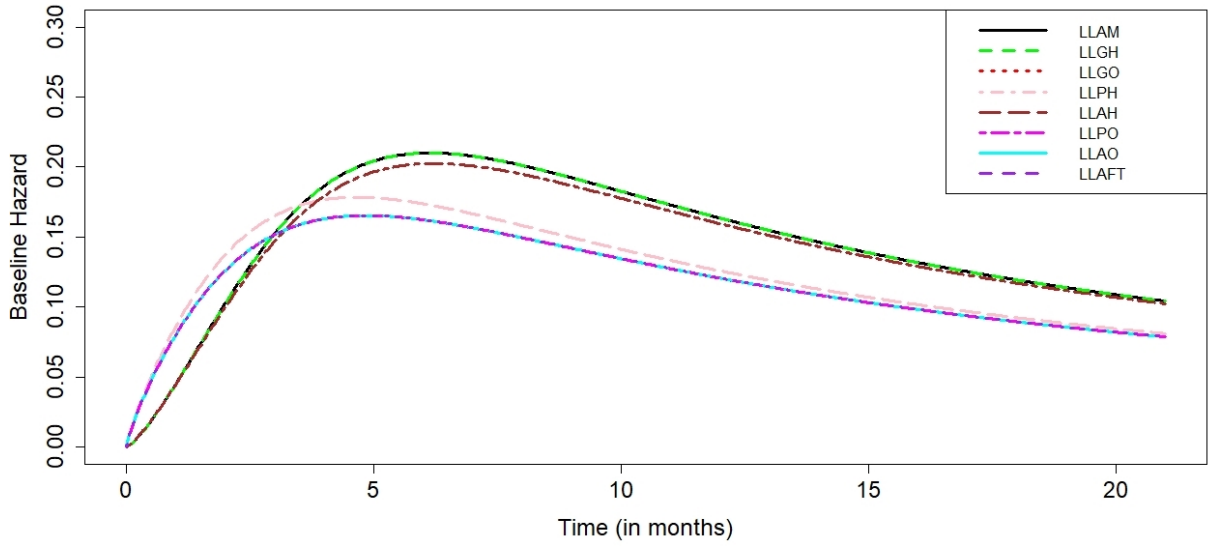

Figure 4. Estimated hrfs for the competitive models of the IPASS dataset.

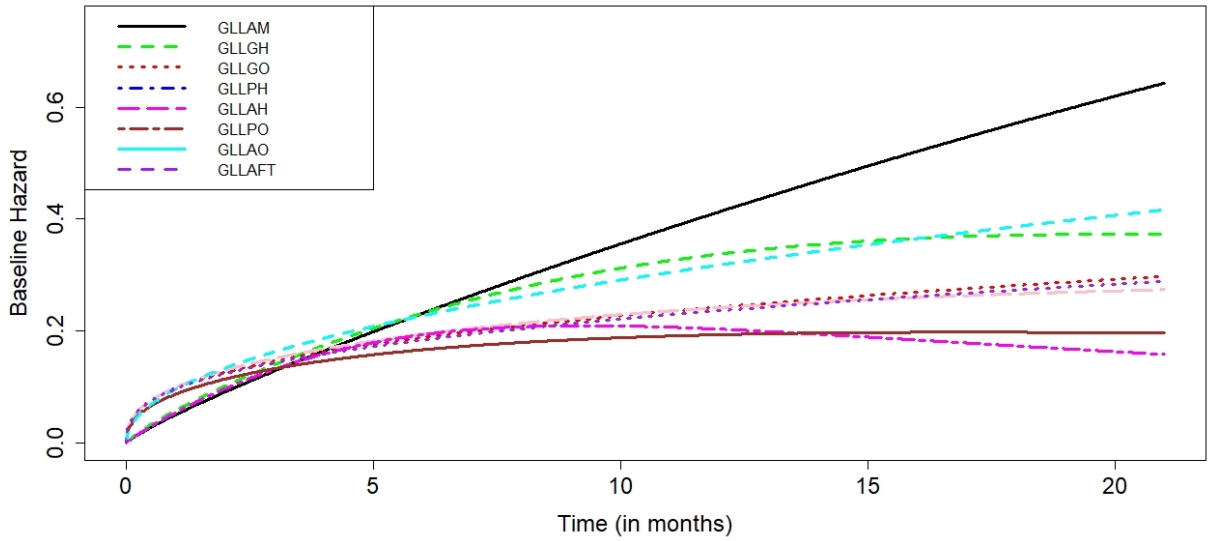

Figure 5. Estimated hrfs for the competitive models of the IPASS dataset.

7.2. Bayesian Analysis

We performed all Bayesian inferential procedures resulting from the combination of the aforementioned general baseline distribution specifications with various prior scenarios in baseline parameters, as well as regression coefficients related to explanatory variables. Using the Rstan package in R [53], the joint posterior distribution for each model was approximated. We performed four parallel chains with 3000 iterations and a burn-in of 1000 for each estimated model. To lessen autocorrelation in the sample, chains were also trimmed by storing after every fifth iteration. With a prospective scale reduction factor close to 1 and an actual number of separate simulation draws of more than 400, convergence to the joint posterior distribution was assured [29,54].

The posterior distribution's numerical summary characteristics are summarized in Table 8. According to the summary results, the McMC algorithm has converged to the joint posterior distribution because the potential scale reduction factor $(\hat{R})$ is 1, the effective sample size $(n - eff)$ is greater than 400, and the Monte Carlo standard error (SE) is less than 5 percent of the posterior standard deviations (SD) for all of the parameters. For visually examining convergence, use trace graphs. The trace plots in Figures 6–13 demonstrate a stationary pattern fluctuating inside a band, demonstrating convergence of

the McMC algorithm. For the proposed AM class, density and autocorrelation graphs are also employed in Figures 14 and 15, respectively, and both show that the McMC algorithm has converged.

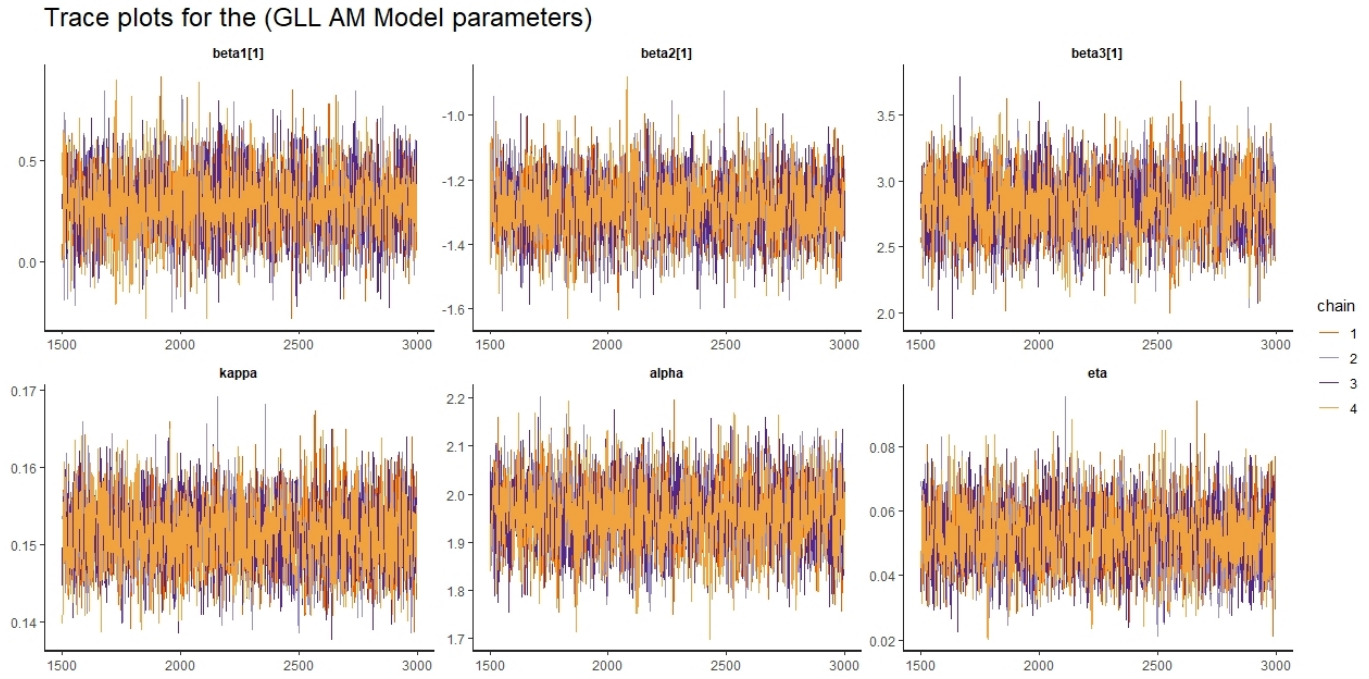

Figure 6. Trace plots for the GLL−AM model parameters.

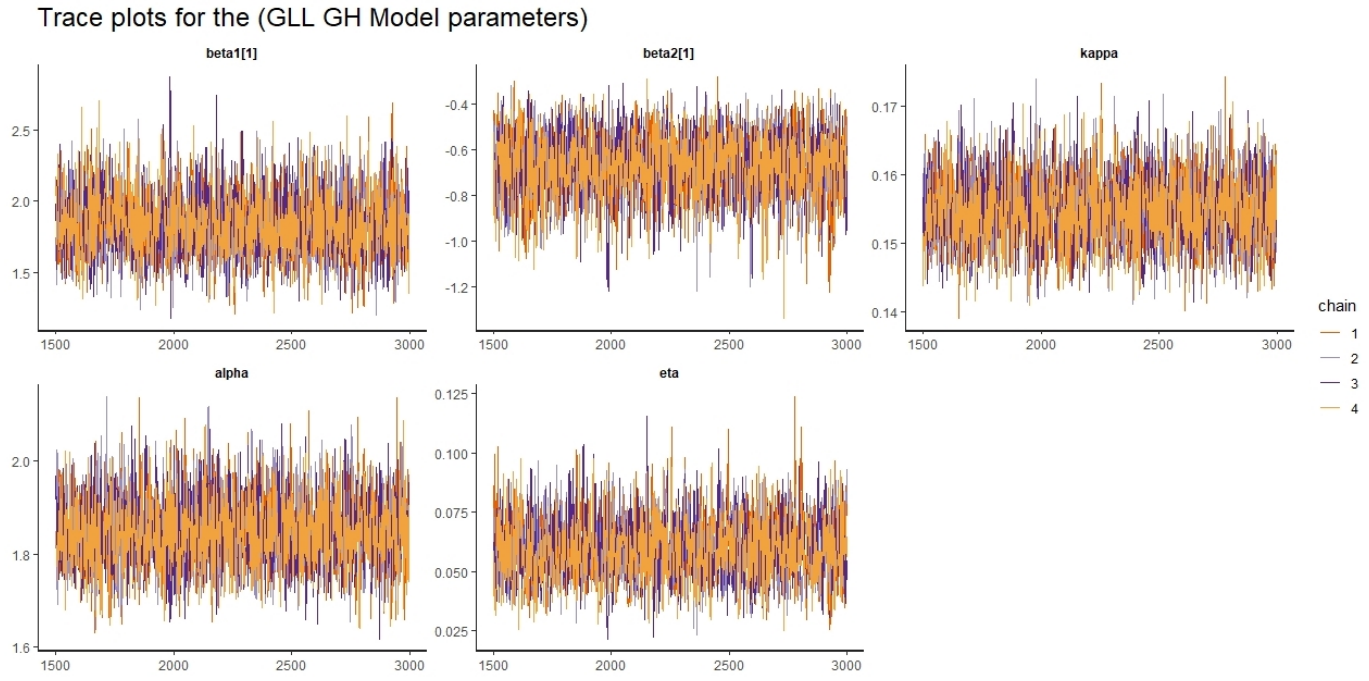

Figure 7. Trace plots for the GLL−GH model parameters.

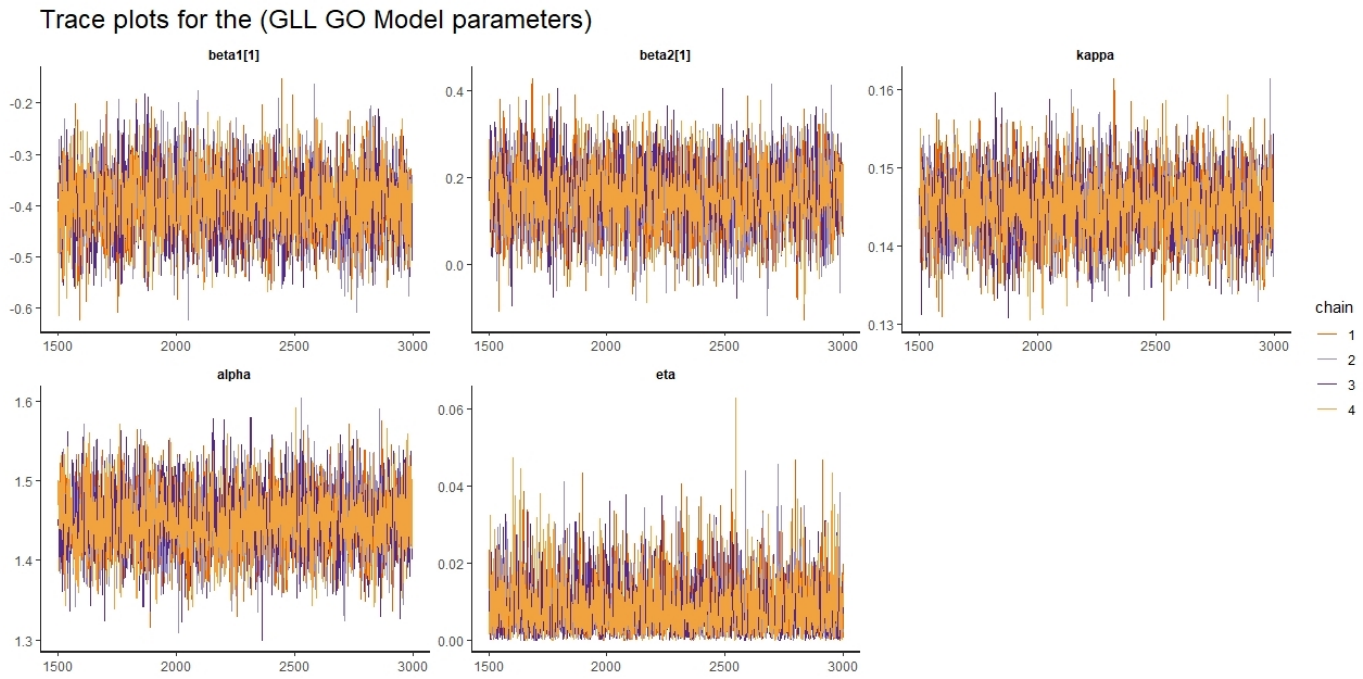

Figure 8. Trace plots for the GLL−GO model parameters.

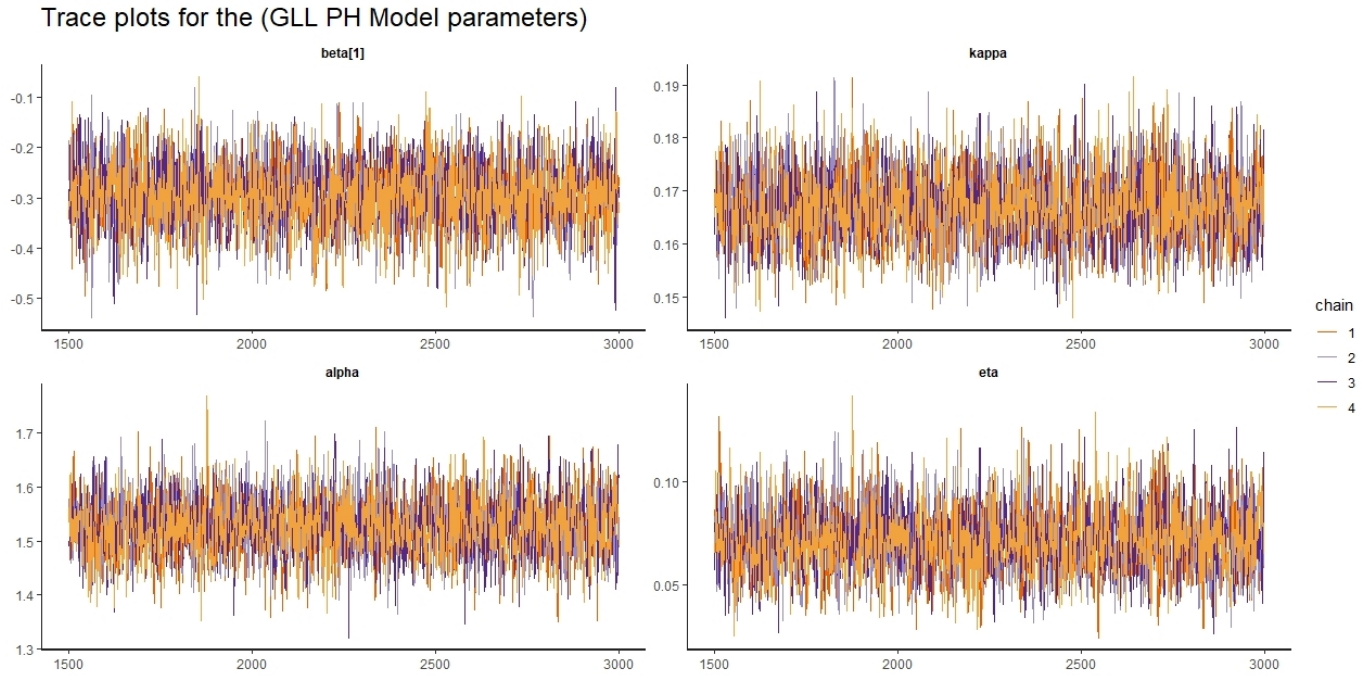

Figure 9. Trace plots for the GLL−PH model parameters.

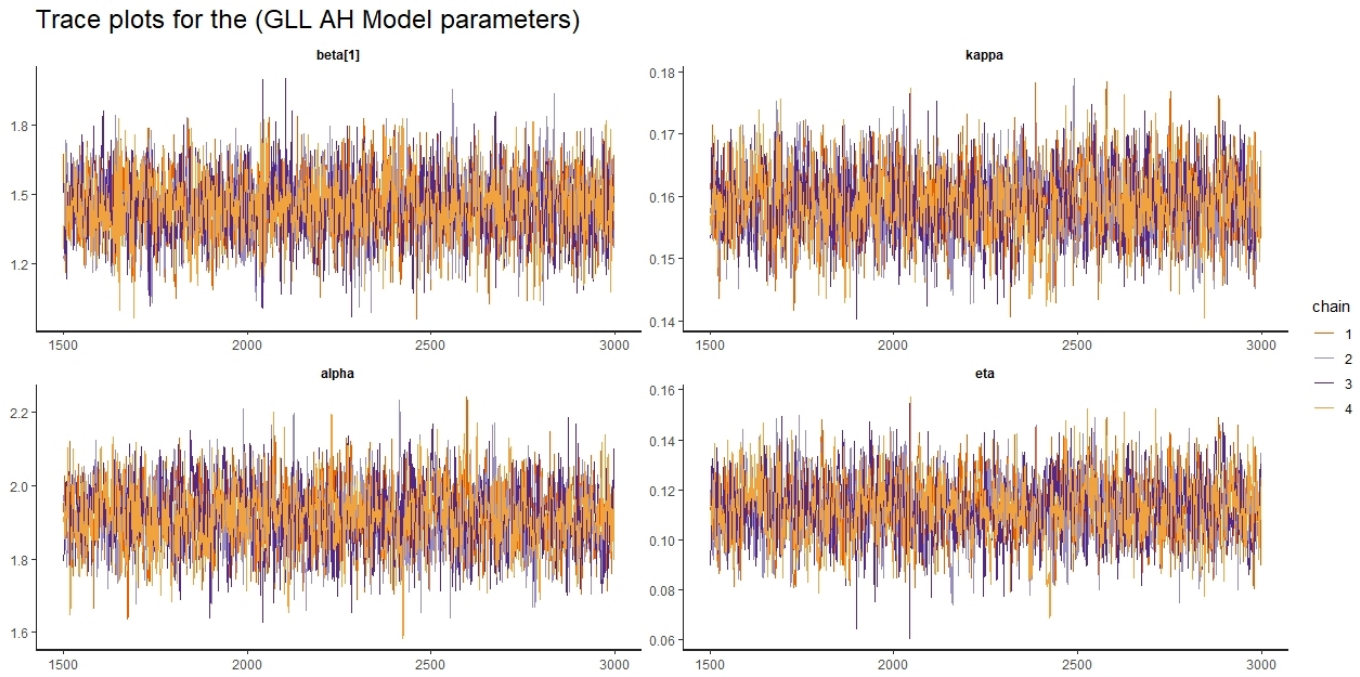

Figure 10. Trace plots for the GLL−AH model parameters.

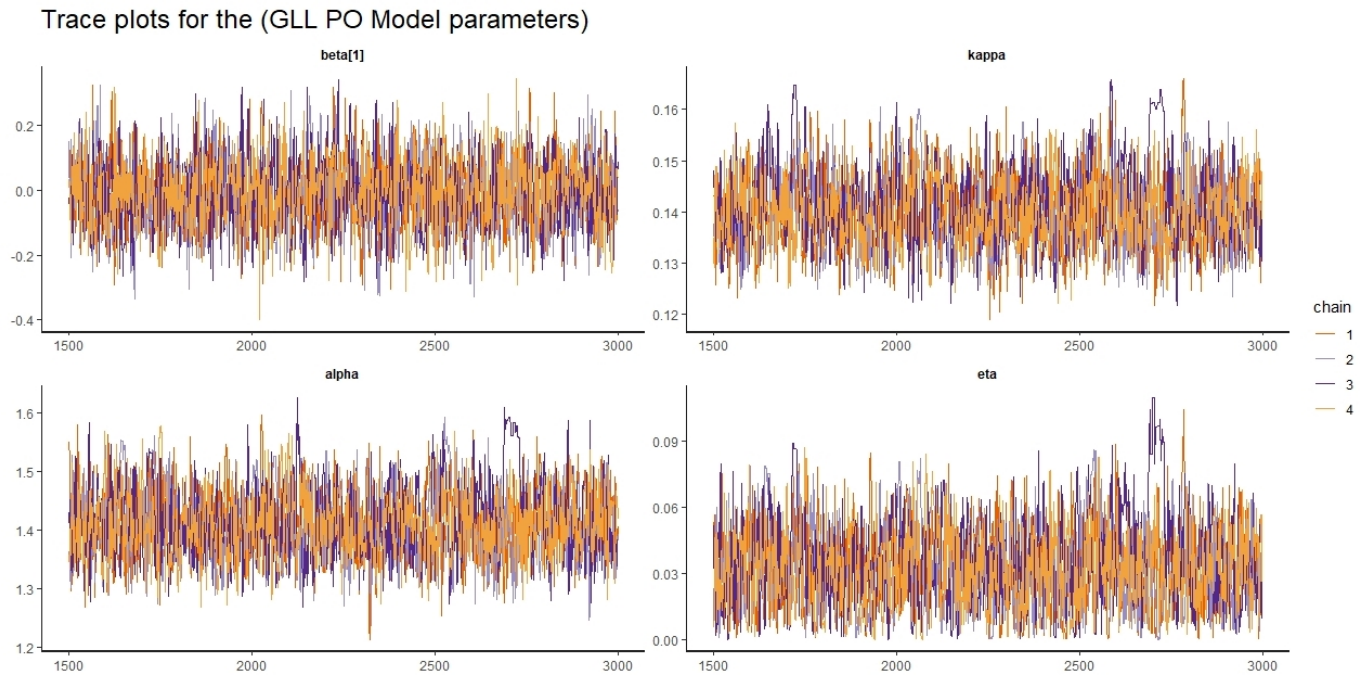

Figure 11. Trace plots for the GLL−PO model parameters.

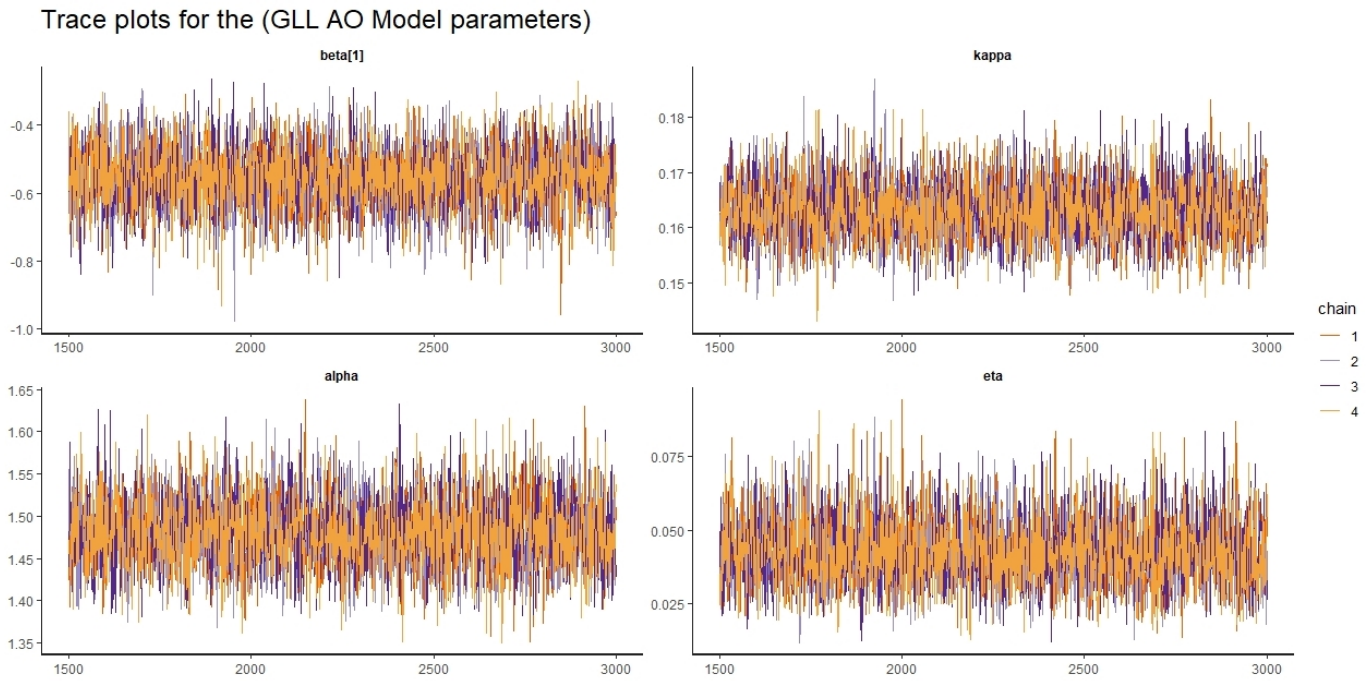

Figure 12. Trace plots for the GLL−AO model parameters.

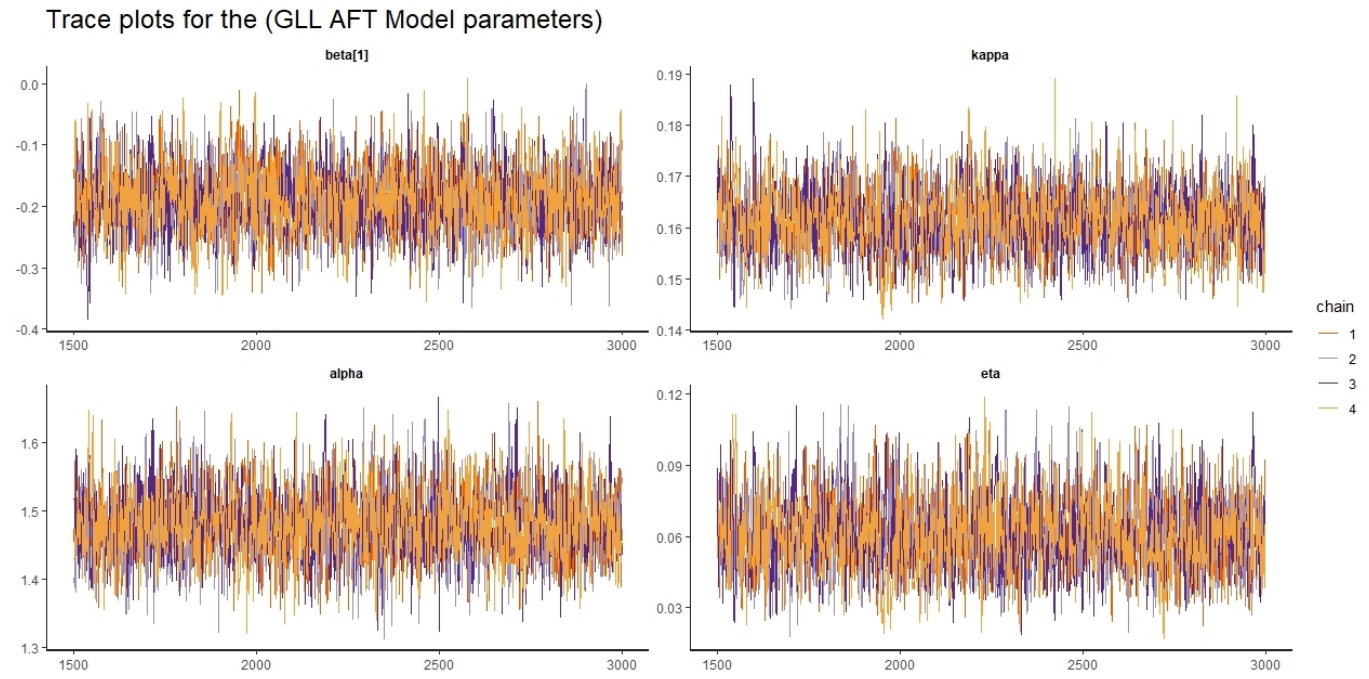

Figure 13. Trace plots for the GLL−AFT model parameters.

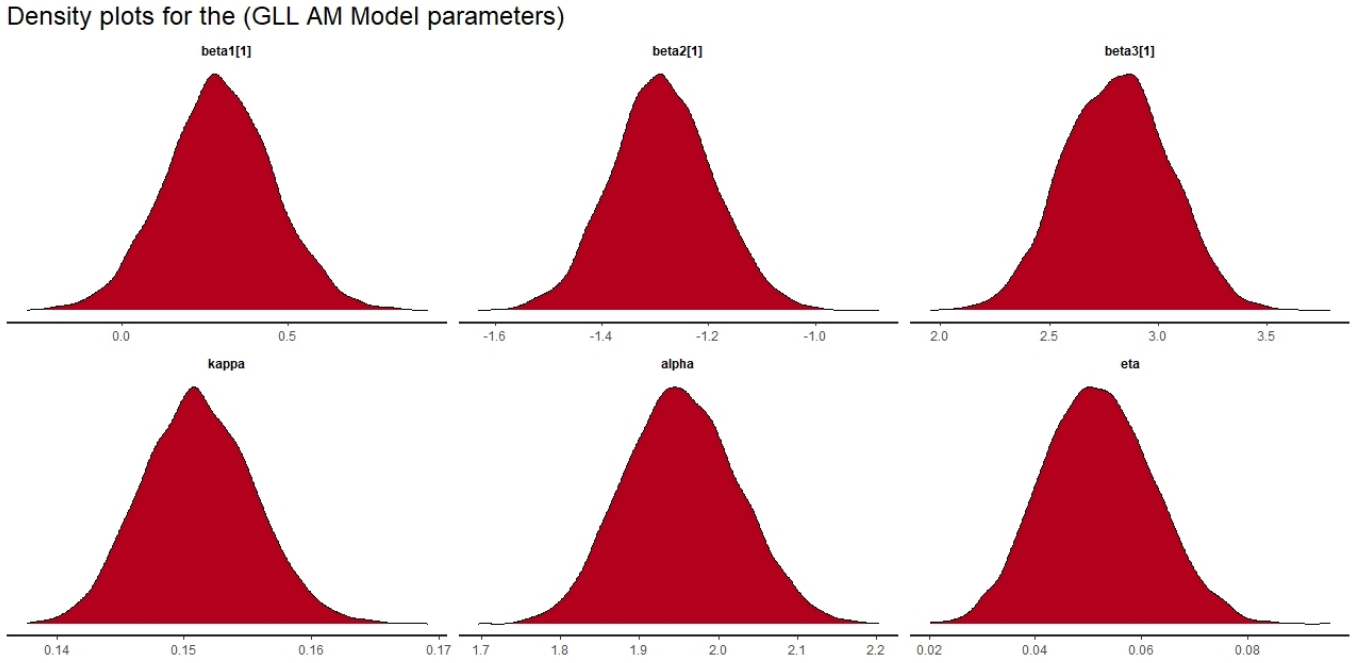

Figure 14. Density plots for the GLL−AM model parameters.

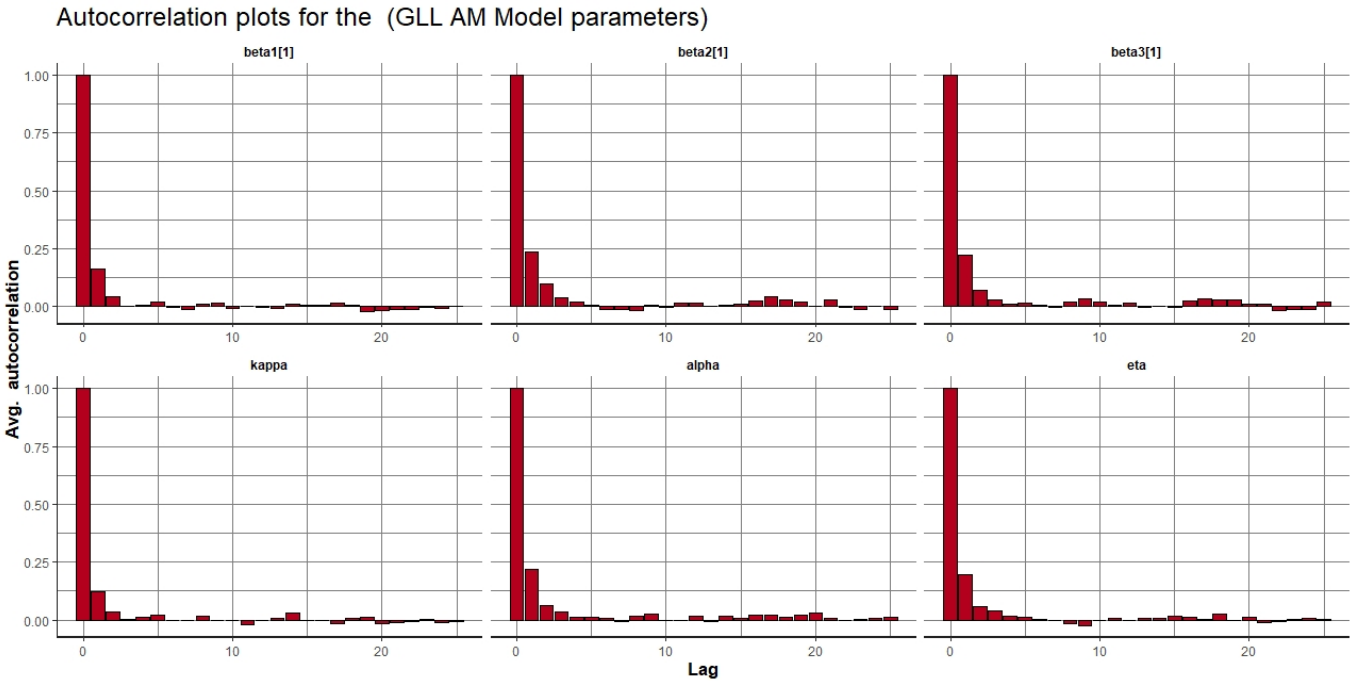

Figure 15. Autocorrelation plots for the GLL−AM model parameters.

Table 8. Results for the posterior properties of the competitive models.

Models	Par (s)	Estimate	SE	SD	2.5%	Medium	97.5%	$n-eff$	$\hat{R}$
GLL-AM	β_1	0.293	0.003	0.166	−0.032	0.291	0.618	4108	1.000
	β_2	−1.287	0.002	0.096	−1.473	−1.288	−1.096	3309	1.000
	β_3	2.810	0.004	0.242	2.349	2.811	3.279	3431	1.000
	α	1.953	0.001	0.072	1.818	1.952	2.095	3477	1.000
	k	0.151	0.000	0.004	0.143	0.151	0.160	4309	1.001
	η	0.052	0.000	0.010	0.032	0.052	0.073	3613	1.001
GLL-GH	β_1	1.839	0.005	0.219	1.444	1.828	2.296	2068	1.001
	β_2	−0.672	0.003	0.139	−0.968	−0.663	−0.425	2232	1.000
	α	1.851	0.001	0.075	1.712	1.849	2.004	2672	1.002
	k	0.155	0.000	0.005	0.145	0.155	0.165	2431	1.001
	η	0.059	0.000	0.013	0.036	0.059	0.087	1739	1.002
GLL-GO	β_1	−0.395	0.001	0.065	−0.524	−0.395	−0.263	3169	1.001
	β_2	0.157	0.001	0.077	0.005	0.158	0.307	3303	1.001
	α	1.450	0.001	0.041	1.371	1.450	1.533	3576	1.000
	k	0.145	0.000	0.004	0.137	0.145	0.154	3523	1.001
	η	0.008	0.000	0.007	0.000	0.007	0.027	3664	1.000
GLL-AH	β	1.448	0.003	0.143	1.163	1.445	1.728	3090	1.000
	α	1.912	0.002	0.089	1.742	1.911	2.087	2446	1.000
	k	0.159	0.000	0.006	0.148	0.159	0.170	2406	1.000
	η	0.112	0.000	0.013	0.088	0.112	0.137	1989	1.000
GLL-PH	β	−0.294	0.001	0.066	−0.427	−0.294	−0.168	3606	1.000
	α	1.526	0.001	0.054	1.425	1.524	1.635	3214	1.001
	k	0.167	0.000	0.007	0.155	0.167	0.180	3082	1.000
	η	0.071	0.000	0.016	0.042	0.071	0.104	2693	1.002
GLL-AO	β	−0.557	0.001	0.092	−0.740	−0.556	−0.378	4181	1.000
	α	1.481	0.001	0.041	1.403	1.480	1.561	3743	1.001
	k	0.163	0.000	0.005	0.153	0.163	0.174	3575	1.001
	η	0.042	0.000	0.011	0.022	0.042	0.067	3654	1.002
GLL-PO	β	−0.011	0.002	0.105	−0.214	−0.011	0.196	2226	1.001
	α	1.412	0.002	0.059	1.308	1.408	1.537	1028	1.003
	k	0.140	0.000	0.007	0.128	0.140	0.156	837	1.003
	η	0.033	0.001	0.019	0.003	0.032	0.075	726	1.004
GLL-AFT	β	−0.188	0.001	0.054	−0.291	−0.190	−0.078	3662	1.000
	α	1.481	0.001	0.051	1.388	1.479	1.584	3041	1.000
	k	0.161	0.000	0.006	0.150	0.161	0.174	2976	1.000
	η	0.061	0.000	0.016	0.032	0.060	0.095	2601	1.000

Table 9 displays the computed models' WAIC and LOOIC values. In comparison to the other fitted models, including the GLL-AH, GLL-AO, GLL-PO, GLL-AFT, GLL-GH, GLL-GO, and GLL-PH models, the GLL-AM model performs better based on the WAIC and LOOIC values. The worst performance is displayed by the most popular survival regression models, such as the PH, PO, and AFT models. This proves that, despite their frequent application, these models are not appropriate for handling survival data with crossing survival curves.

Table 9. Bayesian model selection between the proposed AM class and its sub-models using the GLL baseline distribution.

Model	WAIC	LOOIC
GLL-AM	**5559.50**	**5559.54**
GLL-GH	5608.30	5608.28
GLL-GO	5651.76	5651.69
GLL-AH	5657.40	5697.43
GLL-PH	5692.50	5692.51
GLL-AO	5666.60	5666.61
GLL-PO	5708.60	5708.62
GLL-AFT	5698.00	5698.04

8. Conclusions

We investigated a novel, general, flexible, fully parametric class for hazard-based and odds-based regression models, named the AM class, with a GLL baseline distribution that can incorporate the basic shapes of the failure rate and contains, as specific cases, the main survival regression models of interest in time-to-event analysis: PO, PH, AO, AH, AFT, GO, and GH models. However, the AH, AO, and GO models' restricted utility is mostly due to a lack of reliable and efficient estimating methods. We demonstrated that both classical and Bayesian inference may be performed using existing optimization techniques by adopting a flexible parametric baseline distribution.

The proposed AM class framework is quite adaptable and can easily be applied to a wide range of reliability and survival analysis applications. This framework specifically incorporates and generalizes the practically significant PH, AFT, AH, GH, PO, AO, and GO survival regression models. Additionally, the GLL baseline model, which only requires one additional parameter, accounts for the main hrf shapes (monotone and non-monotone) within some of the most common baseline distributions (Burr type XII, LL, Weibull, and exponential distributions).

The combination of such adaptable parametric odds-based and hazard-based regression models with the AM class structure is a potent tool for modeling survival times. Although we concentrated on overall survival models, the proposed tractable fully parametric AM class is equally useful in excess hazard (relative survival) models. In the AM class, we used the GLL distribution as a baseline distribution; however, other versatile parametric distributions, such as the generalized Weibull, exponentiated Weibull, power generalized Weibull, and generalized gamma distributions, can also accommodate the basic shapes of the hrf including constant, monotone and non-monotone shapes.

We only used the GLL distribution in this case since it allows for a simple implementation, makes parameter interpretation easier, and the accompanying MLEs and Bayesian estimators are consistent and asymptotically normal in the presence of right-censored observations. Finally, an R package called AmoudSurv was developed to fit the odds-based regression models [36].

In the future, we want to develop an R package to fit the most common parametric hazard-based and odds-based regression models, such as the AH, AO, AFT, PH, PO, GO, GH, and AM models, with different baseline distributions that can represent varied hazard rates. This study's technique can also be extended to numerous event scenarios, such as the multi-state model, competing risk model, and to include lifetime data with cure proportion rate and frailty characteristics. It is also possible to adapt it to joint model frameworks, spatial models, mixed effects models, and excess hazard models. Other strategies for censoring observations, such as interval censoring, left censoring, middle-censoring, and double-censoring, could be utilized in future investigations. This is beyond the focus of this study, but it will be covered in many others.

Author Contributions: Conceptualization, A.H.M., S.M., O.N., C.C., H.M.A. and A.-A.H.E.-B.; Data curation, A.H.M., C.C. and H.M.A.; Formal analysis, A.H.M., S.M., O.N., C.C., H.M.A. and A.-A.H.E.-B.; Investigation, A.H.M., S.M., O.N., C.C. and A.-A.H.E.-B.; Methodology, A.H.M., S.M., O.N., C.C., H.M.A. and A.-A.H.E.-B.; Software, A.H.M., S.M., O.N. and C.C.; Supervision, S.M., O.N., C.C., H.M.A. and A.-A.H.E.-B.; Validation, A.H.M., S.M., O.N., C.C. and H.M.A.; Visualization, S.M., O.N., C.C., H.M.A. and A.-A.H.E.-B.; Writing—original draft, A.H.M., C.C. and H.M.A.; Writing—review & editing, A.H.M., S.M., O.N., C.C., H.M.A. and A.-A.H.E.-B. All authors have read and agreed to the published version of the manuscript

Funding: Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2022R299), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Data Availability Statement: Datasets are mentioned along the paper.

Acknowledgments: The authors thank the support from the Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2022R299), Princess Nourah bint Ab-

dulrahman University, Riyadh, Saudi Arabia. We thank the academic editors and referees for their valuable suggestions and comments which improved the paper.

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. [CrossRef]
2. 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. [CrossRef] [PubMed]
3. Kalbfleisch, J.D. Non-parametric Bayesian analysis of survival time data. *J. R. Stat. Soc. Ser. B* **1978**, *40*, 214–221. [CrossRef]
4. Chen, Y.Q.; Wang, M.C. Analysis of accelerated hazards models. *J. Am. Stat. Assoc.* **2000**, *95*, 608–618. [CrossRef]
5. Buckley, J.; James, I. Linear regression with censored data. *Biometrika* **1979**, *66*, 429–436. [CrossRef]
6. Komárek, A.; Lesaffre, E. Bayesian accelerated failure time model with multivariate doubly interval-censored data and flexible distributional assumptions. *J. Am. Stat. Assoc.* **2008**, *103*, 523–533. [CrossRef]
7. 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. [CrossRef]
8. Etezadi-Amoli, J.; Ciampi, A. Extended hazard regression for censored survival data with covariates: A spline approximation for the baseline hazard function. *Biometrics* **1987**, *43*, 181–192. [CrossRef]
9. Louzada-Neto, F. Extended hazard regression model for reliability and survival analysis. *Lifetime Data Anal.* **1997**, *3*, 367–381. [CrossRef]
10. Chen, Y.Q.; Jewell, N.P. On a general class of semiparametric hazards regression models. *Biometrika* **2001**, *88*, 687–702. [CrossRef]
11. Bennett, S. Analysis of survival data by the proportional odds model. *Stat. Med.* **1983**, *2*, 273–277. [CrossRef] [PubMed]
12. Yang, S.; Prentice, R. Semiparametric analysis of short-term and long-term hazard ratios with two-sample survival data. *Biometrika* **2005**, *92*, 1–17. [CrossRef]
13. Demarqui, F.N.; Mayrink, V.D. Yang and Prentice model with piecewise exponential baseline distribution for modeling lifetime data with crossing survival curves. *Braz. J. Probab. Stat.* **2021**, *35*, 172–186. [CrossRef]
14. Banerjee, T.; Chen, M.H.; Dey, D.K.; Kim, S. Bayesian analysis of generalized odds-rate hazards models for survival data. *Lifetime Data Anal.* **2007**, *13*, 241–260. [CrossRef]
15. Royston, P.; Parmar, M.K. Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. *Stat. Med.* **2002**, *21*, 2175–2197. [CrossRef] [PubMed]
16. Huang, T.; Elsayed, E.; Jiang, T. An ALT proportional hazard-proportional odds model. In Proceedings of the 14th ISSAT International Conference on Reliability and Quality in Design, Orlando, FL, USA, 7–9 August 2008.
17. Huang, T.; Jiang, T. An extended proportional hazards-proportional odds model in accelerated life testing. In Proceedings of the 2009 8th International Conference on Reliability, Maintainability and Safety, Chengdu, China, 20–24 July 2009; IEEE: Piscataway, NJ, USA, 2009; pp. 1173–1176.
18. Zhang, J.; Hanson, T.; Zhou, H. Bayes factors for choosing among six common survival models. *Lifetime Data Anal.* **2019**, *25*, 361–379. [CrossRef]
19. Davis, A. Modelling Techniques for Time-To-Event Data Analysis. Ph.D. Thesis, University of Bath, Bath, UK, 2018.
20. Zhou, H.; Hanson, T. Bayesian spatial survival models. In *Nonparametric Bayesian Inference in Biostatistics*; Springer: Cham, Switzerland, 2015; pp. 215–246.
21. Lawless, J.F. *Statistical Models and Methods for Lifetime Data*; John Wiley & Sons: Hoboken, NJ, USA, 2011.
22. Khan, S.A. Exponentiated Weibull regression for time-to-event data. *Lifetime Data Anal.* **2018**, *24*, 328–354. [CrossRef]
23. 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. [CrossRef]
24. Collett, D. *Modelling Survival Data in Medical Research*; CRC Press: Boca Raton, FL, USA, 2015.
25. Rezaei, S.; Hashami, S.; Najjar, L. Extended exponential geometric proportional hazard model. *Ann. Data Sci.* **2014**, *1*, 173–189. [CrossRef]
26. Balakrishnan, N.; Barmalzan, G.; Haidari, A. Modified proportional hazard rates and proportional reversed hazard rates models via Marshall-Olkin distribution and some stochastic comparisons. *J. Korean Stat. Soc.* **2018**, *47*, 127–138. [CrossRef]
27. 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. [CrossRef] [PubMed]
28. Olosunde, A.A.; EJIOFOR, C. Log-Exponential Power Distribution for Accelerated Failure Time Model in Survival Analysis and Its Application. *Afr. Stat.* **2021**, *16*, 2587–2603. [CrossRef]
29. 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. [CrossRef]
30. 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. [CrossRef]

31. Muse, A.H.; Mwalili, S.; Ngesa, O.; Kilai, M. 'AHSurv:An R Package for Flexible Parametric Accelerated Hazards (AH) Regression Models'. 2022. Available online: https://cran.r-project.org/web/packages/AHSurv/index.html (accessed on 18 September 2022).
32. 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. [CrossRef]
33. Alvares, D.; Rubio, F.J. A tractable Bayesian joint model for longitudinal and survival data. *Stat. Med.* **2021**, *40*, 4213–4229. [CrossRef]
34. Li, L.; Hanson, T.; Zhang, J. Spatial extended hazard model with application to prostate cancer survival. *Biometrics* **2015**, *71*, 313–322. [CrossRef]
35. Rubio, F.J.; Drikvandi, R. MEGH: A parametric class of general hazard models for clustered survival data. *Stat. Methods Med. Res.* **2022**, *8*, 1603–1616. [CrossRef]
36. Muse, A.H.; Mwalili, S.; Ngesa, O.; Chesneau, C. 'AmoudSurv:An R Package for Tractable Parametric Odds-Based Regression Models'. 2022. Available online: https://cran.r-project.org/web//packages/AmoudSurv/index.html (accessed on 18 September 2022).
37. S Mastor, A.B.; Ngesa, O.; Mung'atu, J.; Alfaer, N.M.; Afify, A.Z. The Extended Exponential Weibull Distribution: Properties, Inference, and Applications to Real-Life Data. *Complexity* **2022**, *2022*, 4068842. [CrossRef]
38. Alkhairy, I.; Nagy, M.; Muse, A.H.; Hussam, E. The Arctan-X family of distributions: Properties, simulation, and applications to actuarial sciences. *Complexity* **2021**, *2021*, 4689010. [CrossRef]
39. Muse, A.H.; Mwalili, S.M.; Ngesa, O. On the log-logistic distribution and its generalizations: A survey. *Int. J. Stat. Probab.* **2021**, *10*, 93. [CrossRef]
40. Muse, A.H.; Tolba, A.H.; Fayad, E.; Abu Ali, O.A.; Nagy, M.; Yusuf, M. Modelling the COVID-19 mortality rate with a new versatile modification of the log-logistic distribution. *Comput. Intell. Neurosci.* **2021**, *2021*, 8640794. [CrossRef] [PubMed]
41. Al-Aziz, S.N.; Muse, A.H.; Jawad, T.M.; Sayed-Ahmed, N.; Aldallal, R.; Yusuf, M. Bayesian inference in a generalized log-logistic proportional hazards model for the analysis of competing risk data: An application to stem-cell transplanted patients data. *Alex. Eng. J.* **2022**, *61*, 13035–13050. [CrossRef]
42. Muse, A.H.; Mwalili, S.; Ngesa, O.; Almalki, S.J.; Abd-Elmougod, G.A. Bayesian and classical inference for the generalized log-logistic distribution with applications to survival data. *Comput. Intell. Neurosci.* **2021**, *2021*, 5820435. [CrossRef] [PubMed]
43. Lázaro, E.; Armero, C.; Alvares, D. Bayesian regularization for flexible baseline hazard functions in Cox survival models. *Biom. J.* **2021**, *63*, 7–26. [CrossRef]
44. Smith, A.F.; Roberts, G.O. Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods. *J. R. Stat. Soc. Ser. B* **1993**, *55*, 3–23. [CrossRef]
45. Vines, S.; Gilks, W.; Wild, P. Fitting Bayesian multiple random effects models. *Stat. Comput.* **1996**, *6*, 337–346. [CrossRef]
46. Wilks, S.S. The large-sample distribution of the likelihood ratio for testing composite hypotheses. *Ann. Math. Stat.* **1938**, *9*, 60–62. [CrossRef]
47. Burnham, K.P.; Anderson, D.R. Multimodel inference: Understanding AIC and BIC in model selection. *Sociol. Methods Res.* **2004**, *33*, 261–304. [CrossRef]
48. Schwarz, G. Estimating the dimension of a model. *Ann. Stat.* **1978**, *6*, 461–464. [CrossRef]
49. Vehtari, A.; Gelman, A.; Gabry, J. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Stat. Comput.* **2017**, *27*, 1413–1432. [CrossRef]
50. Magnusson, M.; Vehtari, A.; Jonasson, J.; Andersen, M. Leave-one-out cross-validation for Bayesian model comparison in large data. In Proceedings of the International Conference on Artificial Intelligence and Statistics. PMLR, Online, 26–28 August 2020; pp. 341–351.
51. Mok, T.S.; Wu, Y.L.; Thongprasert, S.; Yang, C.H.; Chu, D.T.; Saijo, N.; Sunpaweravong, P.; Han, B.; Margono, B.; Ichinose, Y.; et al. Gefitinib or carboplatin–paclitaxel in pulmonary adenocarcinoma. *New Engl. J. Med.* **2009**, *361*, 947–957. [CrossRef] [PubMed]
52. Argyropoulos, C.; Unruh, M.L. Analysis of time to event outcomes in randomized controlled trials by generalized additive models. *PLoS ONE* **2015**, *10*, e0123784. [CrossRef]
53. Carpenter, B.; Gelman, A.; Hoffman, M.D.; Lee, D.; Goodrich, B.; Betancourt, M.; Brubaker, M.; Guo, J.; Li, P.; Riddell, A. Stan: A probabilistic programming language. *J. Stat. Softw.* **2017**, *76*. [CrossRef]
54. Alvares, D.; Lázaro, E.; Gómez-Rubio, V.; Armero, C. Bayesian survival analysis with BUGS. *Stat. Med.* **2021**, *40*, 2975–3020. [CrossRef] [PubMed]