Informative Bayesian Survival Analysis to Handle Heavy Censoring in Lifetime Data

1st Ryan Leadbetter
School of Elec Eng, Comp and Math Sci
Curtin University
Perth, Australia
ryan.leadbetter@postgrad.curtin.edu.au

1st Aloke Phatak
School of Elec Eng, Comp and Math Sci
Curtin University
Perth, Australia
loke.phatak@curtin.edu.au

2nd Adriano Polpo
School of Physics, Maths and Computing
The University Of Western Australia
Perth, Australia
adriano.polpo@uwa.edu.au

2nd Melinda Hodkiewicz
School of Engineering
The University Of Western Australia
Perth, Australia
melinda.hodkiewicz@uwa.edu.au

2021 International Conference on Maintenance and Intelligent Asset Management (ICMIAM) | 978-1-6654-6671-4/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICMIAM54662.2021.9715184

***Abstract*— Developing a fixed time replacement policy requires reliable estimates of the lifetime of an asset. Estimation of the lifetime is conventionally performed by fitting a Weibull distribution to historical lifetime data using maximum likelihood. However, in industries such as mining and mineral processing, data are often heavily censored. This censoring results in biased parameter estimates that can mislead a replacement policy which the analysis informs. In this paper, we demonstrate for practitioners how high levels of censoring in lifetime data affect inference about the Weibull parameters and how a Bayesian approach can be used to constrain the parameter estimates to more sensible values by using domain expert knowledge. Furthermore, we elaborate on a previous method from the literature which elicits domain expert knowledge on the outcome space in order to construct a joint prior for the Weibull distribution; we also show how this method is more effective at reducing bias caused by high levels of censoring compared to other informative Bayesian approaches in the reliability literature. Finally, we present a small simulation study to show that the bias reducing effect of the informative joint prior is reproducible.**

Keywords*— Bayesian, domain expert knowledge, Prior elicitation, and Survival analysis.

I. INTRODUCTION

In the domain of maintenance and reliability it is common for lifetime data to be heavily censored, particularly when the data is observational. The standard approach to analyzing such data is to fit a Weibull distribution using maximum likelihood [1]. However, in the case of heavy censoring, this approach yields biased inference about the Weibull parameters. The bias arises because the censored data do not inform the full distribution. Fitting the model by using a Bayesian approach with strongly informative priors can leverage domain expert knowledge to inform the model where the data cannot, resulting in a reduction in the bias caused by the censoring. With this in mind, this paper is a guide for practitioners that demonstrates how to construct informative prior distributions so that a Bayesian framework can be used to model heavily censored data in a way that still allows us to gain insight from the data.

Section II provides a brief background on the Weibull distribution, a description of right and interval censoring, background information about the motivating problem, and a short overview of Bayesian approaches that have been proposed as a way of encoding domain expert knowledge into the model. Then in Section III, we describe a method to simulate censored lifetime data from a known distribution to demonstrate the effects of censoring and evaluate how effective the different informative priors are at mitigating bias. Section IV demonstrates how using maximum likelihood to fit a model to the simulated data results in biased parameter estimates. Section V begins by providing a comparison of the maximum likelihood results with an objective Bayesian approach to show that both suffer from bias introduced by censoring. Section V then goes on to present two different ways in which informative priors can be elicited and highlights the hidden implications of the two methods before assessing the performance of each informative prior at reducing the effects of the bias. Finally, in Section VI, we implement a small simulation study to compare the modelling choices in a more comprehensive way.

II. BACKGROUND

A. Weibull distribution

The Weibull distribution is commonly used in reliability analysis for modeling lifetime data because of its ability to model an increasing, constant, or decreasing risk of failure [2]. In addition, the Weibull distribution is the limiting distribution for the minimum value in a sample when the sample space is lower-bound, such as lifetime, which must be greater than zero [3]. An asset is made up of multiple components, each of which has several catastrophic failure modes that could cause failure of the asset. Therefore, the failure time of the asset is equal to the time of the earliest occurring catastrophic failure mode. In other words, the time from installation to the earliest time of occurrence out of all the catastrophic modes for all components is the lifetime of the asset. Throughout this paper we will mainly use the coupled parameterization of the two-parameter Weibull distribution [4], which has probability density function (PDF)

$$f(t) = (t/\eta)^{1-\beta}\exp\left[-(t/\eta)^{\beta}\right] \tag{1}$$

and cumulative distribution function (CDF)

$$F(t) = 1 - \exp\left[-\left(\frac{t}{\eta}\right)^{\beta}\right] \tag{2}$$

where β and η are the shape and scale parameters, respectively. However, when eliciting independent marginal priors for the two parameters of the Weibull in Section V. we use the alternative, uncoupled, parameterization of the distribution where $\lambda = 1/\eta^\beta$. The uncoupled parameterization has PDF

$$f(t) = \beta\lambda(t)^{1-\beta}\exp\left[-\lambda t^\beta\right] \quad (3)$$

and CDF

$$F(t) = 1 - \exp\left[-\lambda t^\beta\right] \quad (4)$$

where λ is the uncoupled scale parameter. This reparameterization is made to make elicitation and encoding of domain expert knowledge more convenient.

In both parameterizations the shape parameter β dictates whether the risk of failure increases, decreases, or remains constant with respect to the exposure time t. If $\beta = 1$ the hazard is constant, in which case the Weibull reduces to an exponential distribution; if $\beta > 1$ the hazard will increase with respect to t; and if $\beta < 1$ the hazard will decrease with respect to t. Practically speaking, $\beta > 1$ corresponds to a wear out failure mechanism, whereas $\beta < 1$ corresponds to infant mortality [4].

B. Censoring

In reliability data we often encounter censored data. This censoring commonly occurs in the form of either right or interval censoring. Right censoring occurs when the component is either still in operation at the end of the observation period or has been preventatively replaced. In both situations, the information from the right censored lifetime is that the true life of the component is greater than the censoring time. Interval censoring occurs when the exact lifetime is unknown, but it is known to be between two values. An example of interval censoring is when a component fails between two inspection times, thus the lifetime is known to be greater than the first inspection time but less than the second. Further discussion on right and interval censoring can be found in [5]. The information contained in both types of censoring can be included in the likelihood function, which can be written as

$$L(\theta|t_i) = \prod_{i=1}^{n}[f(t_i)]^{\delta_O}\,[1 - F(t_i)]^{\delta_R}\,[F(t_i + t_{start}) - F(t_i)]^{\delta_I} \quad (5)$$

where δ_O, δ_R, and δ_I are indicator variables denoting whether the lifetime is fully observed, right censored or interval censored respectively.

C. Motivating problem

The motivating problem for the work described in this paper was to justify and inform the design of a fixed time replacement policy for idlers on an overland conveyor belt using historic lifetime data. The historic lifetime data are the recorded failures times for repeatedly replaced idlers which have been converted into lifetimes by calculating the time between failures for each idler. In their textbook, [4] suggests some cost functions for practitioners to design fixed time replacement policies using a Weibull survival model. However, [4] also stresses that before electing to use a fixed time replacement policy it is essential to be confident that the component suffers from wear out. In other words, the lower confidence bound of the shape parameter β should be greater than one. However, this criterion cannot be guaranteed for the idlers because of incompleteness in the data due to censoring. Censoring occurs because of a combination of a limited window of good quality historical data, long-lasting components, and preventative replacement of components. Unfortunately, the incompleteness of the data means that any inference about the Weibull parameters based on the data alone results in biased parameter estimates and wide uncertainty bounds. The consequence of the bias and wide uncertainty in the estimates is that we cannot confidently say whether the shape parameter lies above one using solely the data.

D. Bayesian lifetime analysis

The Bayesian framework provides a formal and principled way of incorporating prior information into a statistical model. In the Bayesian framework, inference is performed by first starting with a prior distribution $P(\theta)$ that encodes our belief about the parameters before seeing the data. This prior is then combined with the likelihood of the data, $L(\theta|y)$. Through an updating process we obtain the posterior distribution $P(\theta|y)$, our updated belief about the parameters after observing the data. The updating process is performed using Bayes rule $P(\theta|y) \propto L(\theta|y) \times P(\theta)$. For a more detailed introduction to Bayesian methods in reliability see [6] and for a comprehensive resource see [2].

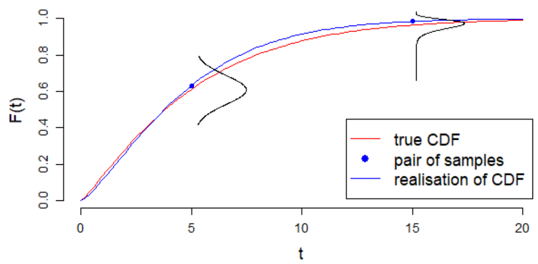

Fig. 1. Example of joint prior elicitation on the outcome space (adapted from [13]).

A prior for a given sampling distribution which results in a posterior that comes from the same family as the prior but with different parameters is known as a conjugate prior [15]. a prior which results in the posterior distribution Conjugate priors are mathematically convenient. However, there is no joint continuous conjugate prior for the two parameter Weibull distribution [7]. When formulating the prior distribution there is no default method and so it is up to the model builder to design a suitable prior. If we have very little prior information we may use an objective prior. An objective prior is a prior which results in the posterior being dominated by the likelihood [15]. When objective priors are used the Bayesian approach is equivalent to non-Bayesian inference methods and therefore suffers from the same bias when data are heavily censored. However, cases where we have no prior belief about the parameters are rare. Information from several different sources can be incorporated to construct an informative prior.

One source of prior information is domain expertise. The process of encoding a domain expert's knowledge to construct an informative prior is known as knowledge elicitation. For a comprehensive text on knowledge elicitation, we point the reader to [8]. Some examples of previous works that have used an informative Bayesian Weibull model to analyze censored lifetime data are [6], [9]–[11]. These approaches construct an informative prior directly on the parameter space. To construct the prior in this way the assumption of prior independence is made between the parameters to simplify the elicitation process. The assumption of independence is convenient because it means that the joint prior can be constructed as

$P(\beta, \eta) = P(\beta) \times P(\eta)$ and so information about the two parameters can be elicited separately. However, the assumption also implicitly states that there is no covariance between the parameters. This is not correct, as we would expect there to be some covariance between the shape and scale parameters. Furthermore, since the Weibull parameters do not have an intuitive interpretation, eliciting information directly on the parameter space results in arbitrary values being chosen for the hyperparameters of the marginal prior distributions that do not accurately reflect the true state of knowledge [11].

A more intuitive method for constructing an informative prior, which also indirectly describes how we would like the parameters to covary with one another, is to elicit information on the outcome space and then use this information to derive the prior for the parameters [12]. Such an approach was proposed by [13] for constructing a joint prior for the two parameter Weibull distribution. The method outlined in [13] elicits information about the Weibull CDF at two exposure times, t_1 and t_2. This elicitation happens in two parts; first the domain expert gives an estimate of value of the CDF at t_1 and t_2, $[E(\hat{F}(t_1)), E(\hat{F}(t_2))]$; second, the expert quantifies the uncertainty about these estimates, $[\sigma(\hat{F}(t_1)), \sigma(\hat{F}(t_2))]$. This information about the value and uncertainty of the CDF at t_1 and t_2 is interpreted as the mean and standard deviation of two $Beta$ distributions. Fig. 1. shows a visual representation of the elicitation process. By taking a sample from the $Beta$ distribution at t_1 and another from the $Beta$ distribution at t_2, conditional on the sample at t_2 being greater than the sample at t_1, we produce a doublet which is a realization of the CDF at t_1 and t_2, $[\hat{F}_i(t_1), \hat{F}_i(t_2)]$. Calculating the parameters of the Weibull CDF which pass through $[(t_1, \hat{F}_i(t_1)), (t_2, \hat{F}_i(t_2))]$ we obtain a draw from the joint distribution of the Weibull parameters, $[\beta_i, \eta_i]$, that encodes the domain expert knowledge. By repeating this sampling process, we can construct an informative joint prior for the Weibull distribution. In the original paper, [13] use exposure times which are early in the lifetime distribution. They elect to do so because the application area is warranty analysis, and this is the region of the CDF for which prior knowledge is available. The joint prior is then updated with binomial survival data in the form of the number of failures out of n units at a given exposure time. In this research we elaborate on the method proposed in [13] so that it can be extended to domains such as reliability. In particular, we show how it can be used for constructing a joint prior and explain how it encodes prior knowledge. We also extend [13] by using the full distribution of the lifetime data to update the prior by implementing the prior sampling method in STAN [14].

III. SIMULATION METHOD

[12] emphasize the value of simulating synthetic data to evaluate a model during the Bayesian model building process. This is a valuable step in any statistical modeling process when dealing with complex data, not just in Bayesian approaches. The idea of using synthetic data is to assess, in a controlled environment, how well the proposed model can reclaim the the known parameter values of the synthetic data. We have followed the recommendations of [12], and simulated synthetic idler data sets of similar size and structure to the true data from the motivating problem. Simulation of the synthetic data was done in two stages; first, we simulate the data generation method, the repeated replacement of idlers; then we simulate the observation process, only observing a window of data due to unreliable replacement records before a given time.

A. Data generating process

To simulate the data generating process we start by taking 9900 samples from a Weibull distribution with parameter values which we believe are reasonable: $\beta = 1.15$ and $\eta = 5.253$, yielding a mean life of 5 years. We then assign 99 lifetimes to theoretical idlers to generate a data set of 100 idlers with 99 lifetimes each. We then take the cumulative sum over the 99 lifetimes for each idler to emulate the repeated replacement of the idler. The outcome of this process is demonstrated in Fig. 2.A for 5 idlers with 12 lifetimes each. The horizontal axis is time in years. $t = 0$ is the equivalent of the commissioning of the conveyor.

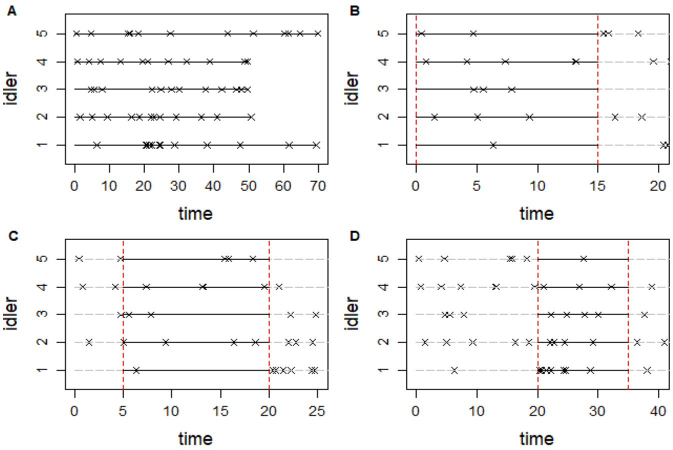

Fig. 2. Example of data simulation process. (A) shows a complete view of the generated data for 5 idlers each with 12 lifetimes. (B) shows the observation scenario where observation begins at $t = 0$ and ends at $t = 15$. (C) shows the case where observation starts at $t = 5$ and ends at $t = 20$. (D) shows the case where observation starts at t = 20 and ends at t = 35.

B. Observation process

To emulate the observational process, we specify a start and end time for the observation period and then discard any information which lies outside this window. We have used three different scenarios which result in increasingly complex censoring. The first is shown in Fig. 2.B, where the start of the observation period, t_{start}, is equal to $t_{start} = 0$ and the end of the observation period is $t_{end} = 15$, three times the mean lifetime. The scenario depicted in Fig. 2.B is equivalent to having a complete data set going back to the commissioning of the conveyor, with t_{end} being the present. Therefore, any idlers whose lifetimes surpass t_{end} are still in operation. For instance, the second lifetime of idler 1 in Fig. 2.B is right censored because it has not failed by $t = 15$ and so we only know that this lifetime is greater that $t = 8.5$. The next observational scenario, shown in Fig. 2.C, is when the observational window no longer starts at $t_{start} = 0$ but at one mean lifetime from the commissioning of the conveyor ($t_{start} = 5$). This could be the case where failure records are not reliable before $t = 5$. Note that the length of the observation period is still $t = 15$. When the observation period no longer starts at $t = 0$ this means that an idler which was installed before t_{start} and fails within the observation period, for example the first lifetime of idler 1 in Fig. 2.C, is interval censored. For cases that are interval censored we know that the true lifetime must be greater than the time for which we observed it, but it cannot be greater than the time that the conveyor has been in operation for. The final observational scenario, shown in Fig. 2.D, is a slightly more complex version of the previous scenario in Fig. 2.C, where t_{start} is equal to four mean lifetimes ($t_{start} = 20$). This scenario is more complex because the interval censoring resulting from the start of the observation period contains much less information about the true idler lifetime and so is more like right censoring.

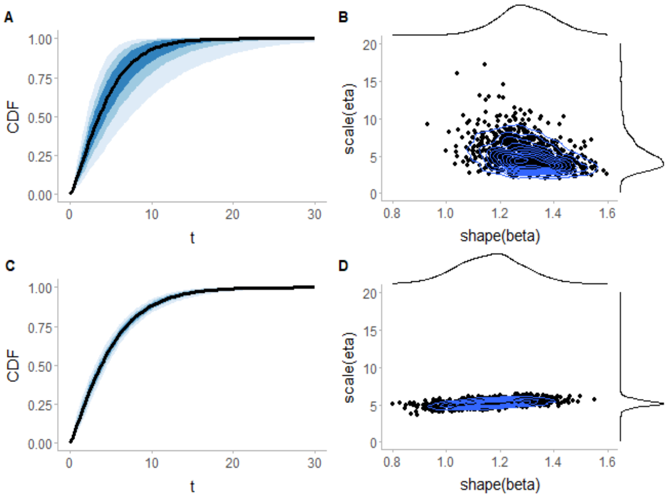

Fig. 3. Prior predictive checks. (A) is a ribbon plot of 1000 realizations of the CDF from the informative prior constructed from two independent marginal priors while (B) is a plot of the parameter values used to generate the realizations in (A). Similarly, (C) shows 1000 realizations of the CDF generated from the joint prior elicited on the output space and (D) plots the parameter values of these realizations.

IV. Maximum likelihood

Estimation of and inference about the Weibull parameters is usually performed by fitting the model using maximum likelihood [1]. The details of maximum likelihood can be found in [4]. Here we demonstrate how, when lifetime data is incomplete, the maximum likelihood estimate yields biased parameter estimates and wide uncertainty bounds. The first two columns in Table I present the results of fitting the likelihood in (5) to the newly simulated data using maximum likelihood. From Table I we can see that the maximum likelihood estimate of the scale moves further away from the true value as t_{start} increases. We also observe that the confidence intervals grow wider as t_{start} is increased, and even for $t_{start} = 5$ the confidence interval no longer contains the true value. This increase in bias and inaccurate quantification of uncertainty suggest that it would be dangerous to design a fixed time replacement policy using a Weibull model that has been fitted using maximum likelihood when data are heavily censored.

V. Bayesian approach

A. Objective prior

Before demonstrating the different ways of constructing an informative prior we first fit a Bayesian model with an objective prior to demonstrate the similarity to fitting a model using maximum likelihood. For an objective prior we use $Gamma$ marginal priors for both the shape and scale parameter with hyper parameters that are close to zero, $Gamma(0.001, 0.001)$, [15]. In the third and fourth columns of Table I we can see that the mean and 95% credible intervals for the shape and scale parameter obtained by fitting this non-informative model. Just as when maximum likelihood is used, the shape parameter is overestimated as t_{start} is increased and the credible interval no longer contains the true value when $t_{start} > 0$. This again suggests that it would be dangerous to design fixed time replacement policies using an objective prior.

B. Informative prior

1) Independent marginal priors

When the assumption of prior independence is made, it is straightforward to define a joint prior for the two parameters. To construct independent marginal priors, we have two pieces of expert knowledge. The first is that the shape parameter of a Weibull distribution for the lifetime of steel rolling element bearings is most likely between 1.1 and 1.5. Since steel rolling element bearings are the main cause of failure in an idler, we would expect that the lifetimes of idlers would have a similar value for the shape parameter. The second piece of information is that the mean lifetime of the idlers is between three and seven years. This information comes from the manufacturer (the Weibull distribution which the data has been simulated from has a mean lifetime $t = 5$ to match this information). Using the feasible ranges of the shape and mean lifetime we can calculate a range for the uncoupled scale parameter, λ, via the mean lifetime equation, $E(T) = \lambda^{-1/\beta}\Gamma(1 + 1/\beta)$. By rearranging for λ and substituting in the upper and lower limits for $E(T)$ and β we get $0.046 \leq \lambda \leq 0.287$. Recall that here we are using the uncoupled parameterisation of the Weibull. This parameterization makes elicitation easier because when the shape parameter β is known the conjugate prior for the uncoupled scale λ is a $Gamma$ distribution. It is much more convenient to work with than the conjugate prior for the coupled scale η, which is

TABLE I. Results of simulation study

	maximum likelihood				objective prior				independent marginal prior				joint prior			
	shape		scale		shape		scale		shape		scale		shape		scale	
simulated data set	est	95% UI	est	95% UI	est	95% UI	est	95% UI	est	95% UI	est	95% UI	est	95% UI	est	95% UI
t_start = 0	1.19	(1.08, 1.3)	5.26	(4.74, 5.77)	1.19	(1.08, 1.3)	5.28	(4.76, 5.81)	1.21	(1.12, 1.3)	5.27	(4.8, 5.8)	1.19	(1.09, 1.28)	5.27	(4.85, 5.72)
t_start = 5	1.21	(1.1, 1.33)	6.00	(5.42, 6.58)	1.21	(1.1, 1.34)	6.02	(5.46, 6.64)	1.22	(1.13, 1.33)	6.00	(5.44, 6.61)	1.21	(1.11, 1.31)	5.78	(5.31, 6.24)
t_start = 20	1.16	(1.06, 1.27)	6.44	(5.75, 7.13)	1.16	(1.06, 1.27)	6.46	(5.79, 7.18)	1.19	(1.1, 1.29)	6.42	(5.77, 7.19)	1.18	(1.08, 1.28)	5.98	(5.49, 6.49)

^UI is the Uncertainty Interval. When maximum likelihood is used to fit the model this is the confidence interval and when Bayesian methods are used this is the credible interval.

an $Inverse\ gamma$. Because it is conjugate, we use a $Gamma$ distribution as the marginal prior for λ; we also use a $Gamma$ distribution for the marginal prior of β because it is a flexible distribution with a support that is constrained to be positive. We construct informative marginal priors by placing 95% of the prior distributions mass between the feasible values of each parameter. This yields $\beta \sim Gamma(156.2,120.2)$ and $\lambda \sim Gamma(6.2,42.6)$ and so the joint informative prior under the assumption of prior independence is $P(\beta,\lambda) \sim Gamma(\beta|156.2,120.2) \times Gamma(\lambda|6.2,42.6)$. Fig. 3.B shows the joint density of 1000 draws from this prior.

2) Joint prior

In contrast to [13], for whom the early part of the lifetime distribution is important, we wish to elicit information in the upper tail of the distribution while still allowing the lower tail to be flexible enough to be updated by the data. To do this we elicit information at the exposure times $t_1 = 3.8$ and $t_2 = 15$. As an estimate and uncertainty for the CDF at t_1 we use $\hat{F}(t_1) = 0.5$ and $sd(\hat{F}(t_1)) = 0.042$ (8% of the estimate) and for the estimate and uncertainty at t_2 we use $\hat{F}(t_2) = 0.96$ and $sd(\hat{F}(t_2)) = 0.013$ (around 1% of the estimate). It is important to note that because we have simulated the data, we know the true value of the CDF at t_1 and t_2. Therefore, the uncertainty around the domain expert's estimates of the CDF have been selected so that the resulting marginal priors of the Weibull parameters contain the true value and are also comparable to the marginal priors from Section V. B. 1). Solving for the beta parameters gives us the two $Beta$ distributions $\hat{F}(t_1) \sim Beta(70.2, 69.6)$ and $\hat{F}(t_2) \sim Beta(188.9, 6.8)$. We then sample pairs $(\hat{F}(t_1)_i, \hat{F}(t_2)_i)$ from the two distributions. If $\hat{F}(t_1)_i \geq \hat{F}(t_2)_i$ the sample is discarded because the CDF must be monotonically increasing. From each pair the parameters of the corresponding CDF can be obtained as

$$\beta = \frac{F^*(t_2) - F^*(t_1)}{\ln\left(\frac{t_2}{t_1}\right)}$$

$$\eta = \exp\left(\ln(t_1) - \frac{F^*(t_1)}{\beta}\right) \qquad (6)$$

where $F^*(t) = \ln(-\ln(1 - F(t)))$. Fig. 3.D shows 1000 draws from the joint prior of the Weibull distribution sampled using this method. From the joint distribution in Fig. 3.D we can see that constructing the prior in this way results in a clear covariance structure.

1) Prior predictive checks

We can use prior predictive checking to compare the information contained in the two different informative prior distributions. Prior predictive checks are performed by sampling realizations of the parameters from the prior and then using them to simulate synthetic data. These checks are a way of gauging what information is contained in each prior distribution and are a good way of refining the prior so that it properly reflects what we believe to be plausible results before seeing the data. However, because we have not modelled the censoring mechanism, out model is not a fully generative model [12], and we therefore cannot generate data from the prior. We can, however, sample realizations of the parameters and plot the resulting CDFs. A and C in Fig. 3. show the realizations of the CDF resulting from the 1000 draws from the joint prior distributions in B and D, respectively. Furthermore, the marginal distribution of the parameters from the samples are show opposite the corresponding axes in B and D of Fig. 3. Fig. 3. highlights the difference in the two methods, even though there is only a small difference in the marginal distributions. The difference arises because the elicitation method proposed by [13] also captures information about how we would like the parameters to covary. The joint prior elicitation method still allows for flexibility in the important range of the distribution, between $t = 0$ and $t = 20$, but is much more restrictive, especially in the upper tail of the CDF (Fig. 3.C) when compared to the prior constructed using independent marginal distributions, which has a significant amount of uncertainty along the entire CDF (Fig. 3.A).

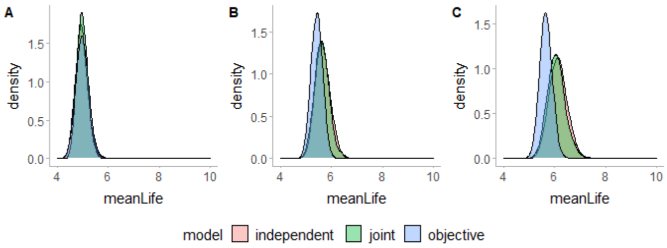

Fig. 4. The posterior distribution of the mean lifetime for the Weibull model with an objective prior, informative independent marginal prior, and joint prior. (A), (B), and (C) are the data sets with observation time starting at t = 0, 5, and 20 respectively.

TABLE II. RESULTS OF SIMULATION STUDY

		average bias of mean life est			
simulated data set	model	mean	2.5% percentile	97.5% percentile	density
t_start = 0	objective	0.056	-0.432	0.533	
	independent	0.018	-0.456	0.470	
	joint	0.006	-0.328	0.304	
t_start = 5	objective	0.449	-0.024	0.843	
	independent	0.417	-0.042	0.796	
	joint	0.294	-0.018	0.556	
t_start = 20	objective	1.548	0.762	2.239	
	independent	1.457	0.698	2.124	
	joint	0.884	0.457	1.245	

C. Posterior comparison

Now that we have two informative joint priors, one constructed using the product of independent marginal priors and the other using the methods of [13], we can update these priors with the likelihood of the simulated data to obtain the posterior distribution. The Stan code which performs this updating is available at [16]. The last four columns in Table I summarise the marginal posterior distributions of the shape and scale using the mean and 95% credible interval for the different combinations of informative prior and simulated data when the two different informative priors are used. The mean and credible intervals of the scale parameter show that the informative prior constructed using the method of [13] is much more effective at reducing the bias introduced by the more complex censoring mechanism when $t_{start} > 0$.

However, it is difficult to interpret the joint posterior of the parameters. Therefore, we repeatedly sample from the joint posterior and calculate the mean lifetime for each draw using the uncoupled parameterization of the mean lifetime equation, $E(T) = \eta\Gamma(1 + 1/\beta)$, to construct a posterior distribution of the mean lifetime. Calculating the distribution of functions in this way is one of the advantages of working in the Bayesian paradigm. Fig. 4. compares the posterior of the mean lifetime when an objective, informative marginal independent priors or the informative joint prior constructed using [13] is used. A, B, and C in Fig. 4 show this comparison for the different simulated data sets where $t_{start} = 0$, $t_{start} = 5$, and $t_{start} = 20$ respectively. In the three plots in Fig. 4. the impact of the joint prior elicited using the method of [13] is much clearer, especially in plot C. We can also see that an informative prior constructed as the product of two independent marginal priors performs little better than a completely objective prior.

VI. Simulation Study

The results from the previous section demonstrate how an informative prior elicited on the outcome space can be used to reduce the bias caused by high proportions of censoring in data. However, the conclusions drawn are based on a single example. To make the results more defensible we have conducted a small-scale simulation study. In the simulation study we repeat the same process performed in Section IV to obtain the posterior distribution of the mean lifetime for each of the nine combinations of synthetic data sets and model priors 100 times, generating new synthetic data for each iteration. We then summarise the bias of each of the posterior distributions using the measure $bias = 5 - mean(\hat{E}(T))$, where $\hat{E}(T)$ is the posterior of the mean lifetime. The resulting 100 bias values for each synthetic data set and model combination are summarized in Table IV by the mean, 0.025 quantile, and 0.975 quantile. The rightmost column in the table shows the density of the 100 bias values from the simulation. Similar to the individual example, we see an increase in the bias as t_{start} increases. Additionally, it is apparent, especially from the density outline, that using the informative prior constructed using the method of [13] is much more effective at reducing the average bias as well as the variation of bias caused by censoring, particularly for the case when $t_{start} = 20$. It is important to note that the bias is still present for the joint model and on average equal to 0.88 mean lifetimes when $t_{start} = 20$. However, the fact that the variance of the bias has been reduced for the extreme case means that a bias adjustment would be more reliable if the informative joint prior is used.

VII. Concluding Remarks

In this work we have demonstrated that biased parameter estimates result from high levels of censoring in lifetime data and explored two different informative Bayesian approach to reduce that bias. The findings show that an informative Bayesian prior can be used to reduce the effect of the bias but only if the prior is constructed in a principled way. If an informative prior is constructed under the assumption of independent marginal priors, the informative model peforms no better than an objective model. However, if information is elicited on the outcome space in a way which captures covariance structure, the bias reducing effect of an informative Bayesian approach is much more significant. The difference in the performance of these two informative priors emphasizes the importance of prior elicitation is when using domain expert knowledge. We should spend just as much mental effort constructing a prior as we do performing exploratory data analysis, selecting a likelihood, or evaluating model fit. Using a correctly constructed prior which accurately reflects beliefs not only reduces the average bias, but also the variability of the bias meaning that some form of bias adjustment may be suitable to correct the inference. The simulation study conducted here is not comprehensive. We have explored a limited number of scenarios and only used a small number of iterations. Furthermore, the priors explored have reflected the case where domain expert knowledge is accurate. In order to formulate advice on how to adjust for bias, a larger simulation study is necessary which looks at a wider range of scenarios and also the cases when the domain expertise used to construct the prior is incorrect and may not contain the true parameter values. However, despite the lack of a comprehensive simulation study, the advice outlined in this work can help practitioners because they now have a way forward in dealing with heavily censored lifetime data; are aware of the possibility of misleading results due to bias; and can understand the way in which these results may be misleading.

Acknowledgment

This research was supported by the Australian Government through the Australian Research Council's Industrial Transformation Training Centres scheme (project IC180100030).

References

[1] U. Genschel and W. Q. Meeker, "A comparison of maximum likelihood and median-rank regression for weibull estimation," *Quality Engineering*, vol. 22, no. 4, pp. 236–255, 2010.

[2] M. S. Hamada, A. G. Wilson, C. S. Reese, and H. Martz, *Bayesian reliability*, vol. 173. 2008, pp. 935–935.

[3] H. Rinne, *The weibull distribution: A handbook*. CRC press, 2008.

[4] A. K. Jardine and A. H. Tsang, *Maintenance, replacement, and reliability: Theory and applications*. CRC press, 2005.

[5] J. F. Lawless, *Statistical models and methods for lifetime data*, vol. 362. John Wiley & Sons, 2011.

[6] M. Li and W. Q. Meeker, "Application of bayesian methods in reliability data analyses," *Journal of Quality Technology*, vol. 46, pp. 1–23, 2014.

[7] R. M. Soland, "Bayesian analysis of the weibull process with unknown scale and shape parameters," *IEEE Transactions On Reliability*, Nov. 1969.

[8] A. O'Hagan, C. E. Buck, A. Daneshkhah, J. R. Eiser, P. H. Garthwaite, and D. J., *Uncertain judgements: Eliciting experts' probabilities*, 1. Aufl. New York: Wiley, 2006.

[9] D. Kundu and D. Mitra, "Bayesian inference of weibull distribution based on left truncated and right censored data," *Computational Statistics and Data Analysis*, vol. 99, pp. 38–50, 2016.

[10] F. Ducros and P. Pamphile, "Bayesian estimation of weibull mixture in heavily censored data setting," *Reliability Engineering & System Safety*, vol. 180, pp. 453–462, Dec. 2018.

[11] C. W. Zhang, R. Pan, and T. N. Goh, "Reliability assessment of high-quality new products with data scarcity," *International Journal of Production Research*, vol. 59, no. 14, pp. 4175–4187, 2021.

[12] A. Gelman *et al.*, "Bayesian workflow," *arXiv preprint arXiv:2011.01808*, 2020.

[13] M. P. Kaminskiy and V. V. Krivtsov, "A simple procedure for bayesian estimation of the weibull distribution," *IEEE Transactions on Reliability*, vol. 54, pp. 612–616, 2005.

[14] Stan Development Team, "The Stan Core Library." 2021.Available: http://mc-stan.org/

[15] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin, *Bayesian data analysis*. Chapman; Hall/CRC, 2013. doi: 10.1201/B16018.

[16] R. K. Leadbetter, *Code to reproduce Leadbetter et.al. (2021)*. 2021, Available: https://github.com/rleadbett/Informative-Bayesian-Survival-Analysis-to-Handle-Heavy-Censoring-in-Lifetime-Data