Discrimination between parametric survival models for removal times of bird carcasses in scavenger removal trials at wind turbines sites

Regina Bispo, Joana Bernardino, Tiago A. Marques and Dinis Pestana

Abstract Wind power is one of the most promising energy sources found in nature. Despite being considered a clean energy source, the existence of potential environmental impacts, namely, on flying vertebrates, is broadly recognized. In monitoring studies, estimation of avian (or bats) mortality caused by collision has particular interest and must take into account carcass removal by scavengers. For this purpose, scavenger removal trials are conducted at wind turbines sites. Data from scavenger removal trials refer to time until removal of the carcass and are "classical" examples of survival times.
Parametric survival models based on the exponential, Weibull, log-logistic and log-normal distributions are among the most repeatedly used throughout literature. In this study we aim to discriminate between these four competing parametric models to analyze removal data from trials conducted in ten Portuguese wind farms. Both goodness of fit measures and plotting procedures are used and discussed.

Regina Bispo
Departamento de Estatística, ISPA - Instituto Universitário, Lisboa, Portugal and CEAUL - Centro de Aplicações e Estatística da Universidade de Lisboa, e-mail: rbispo@ispa.pt

Joana Bernardino
Bio3 - Estudos e Projectos em Biologia e Valorização de Recursos Naturais, Portugal, e-mail: joana.bernardino@bio3.pt

Tiago Marques
Centre for Research into Ecological and Environmental Modeling, The Observatory, Buchanan Gardens, St Andrews, Scotland UK and CEAUL - Centro de Aplicações e Estatística da Universidade de Lisboa, e-mail: tiago@mcs.st-and.ac.uk

Dinis Pestana
Departamento de Estatística e Investigação Operacional, Faculdade de Ciências da Universidade de Lisboa,Lisboa, Portugal and CEAUL - Centro de Aplicações e Estatística da Universidade de Lisboa, e-mail: ddpestana@fc.ul.pt

1 Introduction

Nowadays, wind is considered as one of the most promising energy sources found in nature. Despite being considered a clean energy source, the existence of potential environmental impacts, namely, on flying vertebrates, must be recognized [1]. Although the nature of the impacts can be diverse, including habitat modification, barrier effect or disturbance in nesting areas, there is a major concern with the direct mortality caused by the collision with the rotating turbine rotor blades and other parts of the wind plant structures [2]. The estimation of collision-caused mortality is, therefore, of crucial importance. It provides information about direct impacts by particular projects, allows comparisons between research studies, enables impact trend studies, provides a basis for legislation and enables the comparison with the impacts caused by other forms of energy generation or other human activities [3].

The estimation of mortality must take into account carcass removal by scavengers and/or by decomposition. Data from scavenger removal trials refer to time until removal of the carcass and are classical examples of "survival" times. Survival data can be modeled either by using non-parametric or parametric approaches. Parametric methods, by assuming a specific form to the underlying data distribution, have the advantage to allow more precise inferences [4]. Naturally, the drawback is that then the used model is assumed correct and estimates, and correspondent variances, depend on the validity of the assumptions. Hence, one of the most important aspects when using parametric survival methods is the selection of the lifetime distribution [5]. Methods to discriminate between parametric models depend on whether populations are homogeneous or not and on the existence of censored observations.

Plotting procedures are common methods used to explore and visually inspect the adequacy of a specific lifetime distribution. In particular, plots based on the linearization of the survivor function are often used. When models are adequate, these plots are expected to be roughly linear. Also, empirical and parametrical estimated functions can be draw together to visually check the model adjustment. For non-homogeneous populations, both types of plots may be constructed in strata defined by the components of the regression vector [6].

The focus of this paper is to discriminate between parametric survival models suitable for modeling the removal times of bird carcasses in scavenger removal trials that were conducted at ten Portuguese wind farms. Methods included both graphical techniques and the calculation of adjustment measures. The agreement between the different approaches is discussed.

2 Motivating data

Carcass removal trials were conducted in ten wind farms located in the north and center of Portugal (for confidentiality reasons sites names are coded from WF1 to WF10). Trials were spread over two seasons (May/June and September/October or January/February and July/August) and three bird size classes were considered (small: $\leq$ 15 cm, medium: between 15 and 25 cm, large: $>$ 25 cm). Carcasses were placed in randomly chosen locations beneath the wind turbines, independently of size class, at a minimum distance of 500 m from each other. The carcasses were checked daily for a maximum period of 20 days and time until removal was recorded.

3 Discrimination between parametric survival models

The methods of assessing the adequacy of data to a specific probability distribution or of selecting the best parametric survival model among several competitors can generically be grouped into plotting procedures and formal methods for goodness of fit evaluation. In the next two sections, we describe some frequently used methods that fall under these two categories and that were used in this study.

3.1 Plotting procedures

Plotting procedures based on the estimated survivor function can be used to visually check the adequacy of the parametric model. In particular, plots based on the linearization of the survivor function give information on the underlying lifetime distribution [7]. These plots should be roughly linear if the assumed model is reasonable. For the mentioned probability distributions, expected approximated linear relationships are summarized in Table 1. The linear agreement can then be appreciated by eye (which can be misleading) or be measured using the standard coefficient of determination.

Table 1 Required linear transformations of survival probability and time scales for different lifetime distributions for graphical inspection of the parametric survival models adequacy. t and $S(t)$ denote, respectively, time and the survivor function.

Lifetime distribution	Time scale	Probability scale
Exponential	t	$-\log S(t)$
Weibull	$\log t$	$\log(-\log S(t))$
Log-logistic	$\log t$	$\log\left(\frac{S(t)}{1-S(t)}\right)$
Log-normal	$\log t$	$\Phi^{-1}(1-S(t))$

Another graphical procedure can be achieved by superimposing graphically the empirical (Kaplan-Meier) and the parametrical estimated survivor functions to check visually the adjustment between the observed and the fitted functions.

For censored data, described plotting procedures are probably the most widely useful graphical approaches for comparing competing parametric models [8].

Whenever covariates were found to be significant in modeling lifetime data, both types of plots were constructed for each covariate level (or combination of levels).

3.2 Goodness of fit measures

There are several well known formal goodness of fit tests and tests for discriminating among models when checking the adequacy of parametric survival models upon which inferences are based [4]. In this study we formally measured model adjustment using the Akaike's and the Bayesian's Information Criteria (AIC and BIC, respectively). These measures enable comparisons between not nested models. Smaller values of both measures indicate better adjusted models.

4 Results

Data were analyzed using the accelerated failure time models general class to model the effect of season and body size carcasses as explanatory variables for removal times. The data analysis showed that removal times were not affected significantly by season and body size factors in 6 out of the 10 wind farms (WF1 to WF6). In WF7 and WF8 wind farms, season proved to have a significant effect ($p < 0.001$) and in WF9 and WF10 wind farms, both covariates had a significant effect on the removal times ($p < 0.001$). Subsequent sections detail results in each of the 10 wind farms, according to the type of fitted model. Although the described plotting procedures were used for all the 10 analyzed data sets, plots based on the linearization of the survivor function are shown only for homogeneous populations and plots superimposing the empirical and the adjusted models are used to ilustraste the adequacy of the models accounting for dependency on explanatory variables.

4.1 Homogeneous populations

After fitting the four models to the data sets of each wind farm, adjustment measures were calculated to compare their respective adequacy (Table 2). For WF1, WF3, WF4 and WF6, Akaike's and Bayesian's criterions were found to be the lowest for the log-normal model. At WF2 and WF5 wind farms log-logistic models presented lower values. However, differences of AIC and BIC values between log-logistic and log-normal models were minimal, suggesting similar goodness of fit.

Plots based on the linearization of the survivor function (Figure 1) show that the exponential model has the poorest fit in all six wind farms (smaller coefficients of determination), which reflects the relative inadequacy of the exponential distribution to model removal times. The remaining parametric models give fairly good approx-

Table 2 Akaike's and Bayesian's criteria measures for exponential, Weibull, log-logistic and log-normal fitted models.

Wind farms	Measure	Exponential	Weibull	Log-logistic	Log-normal
WF1	AIC	243.086	244.704	234.988	234.516
	BIC	244.914	248.361	238.645	238.174
WF2	AIC	305.819	307.818	286.966	288.420
	BIC	307.913	312.007	291.155	292.609
WF3	AIC	100.087	101.613	99.540	98.841
	BIC	101.803	103.604	101.532	100.833
WF4	AIC	265.271	255.003	251.872	247.138
	BIC	267.365	259.192	256.061	251.327
WF5	AIC	91.700	93.568	85.062	85.854
	BIC	92.695	95.560	87.054	87.845
WF6	AIC	321.679	323.106	314.369	312.525
	BIC	323.773	327.295	318.558	316.714
WF7	AIC	408.365	409.753	401.941	399.218
	BIC	423.129	416.899	409.087	406.364
WF8	AIC	186.134	175.516	179.168	178.530
	BIC	189.461	180.507	184.159	183.520
WF9	AIC	278.585	262.217	263.874	262.395
	BIC	286.962	272.689	274.345	272.867
WF10	AIC	152.066	147.576	146.880	147.010
	BIC	156.463	153.439	152.743	152.873

imated linear relationships, with slight differences between them. The coefficients of determination, point to log-logistic and log-normal models as the most suitable ones, matching results from AIC and BIC values. However, in WF4 wind farm the best linear relationship was found for the Weibull model.

4.2 Non-homogeneous populations

Figure 2 shows the plots representing both Kaplan-Meier curves and adjusted parametric models. Based on these plots, comparisons between the four adjusted models seems hard as differences between the models are almost eye imperceptible. Hence, model selection based on these type of plots is risky and can be misleading. The calculated goodness of fit measures (table 2) assume, therefore, a special importante role in this context.

Regarding WF7 and WF8 wind farms, AIC and BIC values were smaller for the log-normal and the Weibull models, respectively, suggesting these models as the most suitable ones to model carcass removal times in these wind farms. For WF9 and WF10 data, these measures point, respectively, to Weibull and log-logistic based models as the most suitable. Note, however, that these measures present very close values for Weibull, log-logistic and log-normal regression models, which, in fact,

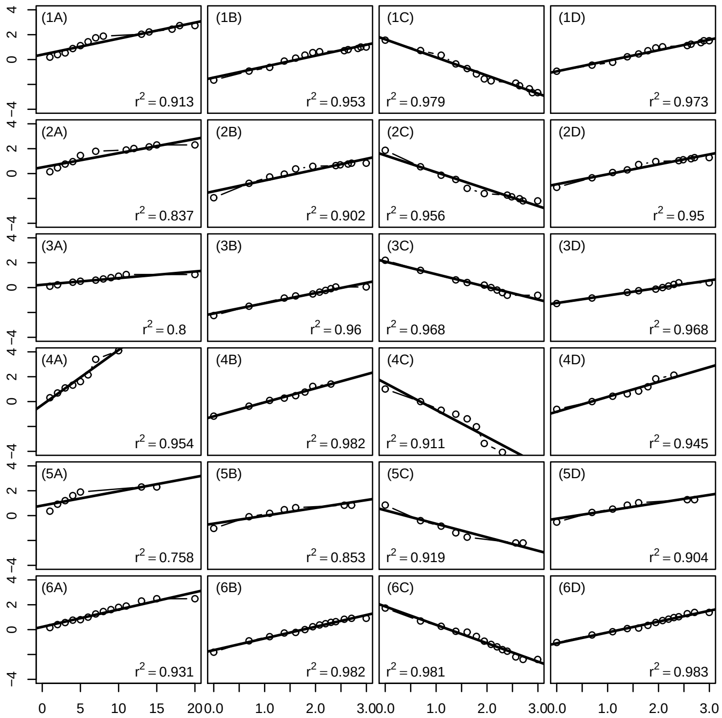

Fig. 1 Plots based on the linearization of the survivor function for the inspection of the fitted parametric survival models adequacy in 1 - WF1, 2 - WF2, 3 - WF3, 4 - WF4, 5 - WF5 and 6 - Pinhal interior wind turbines sites, regarding A - exponential, B - Weibull, C - log-logistic and D - log-normal fitted models.

was expected given the minor differences between correspondent plots displayed in Figure 2.

5 Conclusion

In monitoring studies at wind turbines sites, the final goal is to estimate the number of avian and/or bat fatalities attributable to wind turbine collisions for the entire project on an annual basis [9]. To correctly estimate mortality it is important to consider carcass scavenger removal. For that reason it is a standard procedure to conduct carcass removal trials in wind farms, collecting data regarding carcass removal times. As the final goal of the analysis is to model data and use the fitted parametric models to estimate carcass persistence probabilities, and these can de-

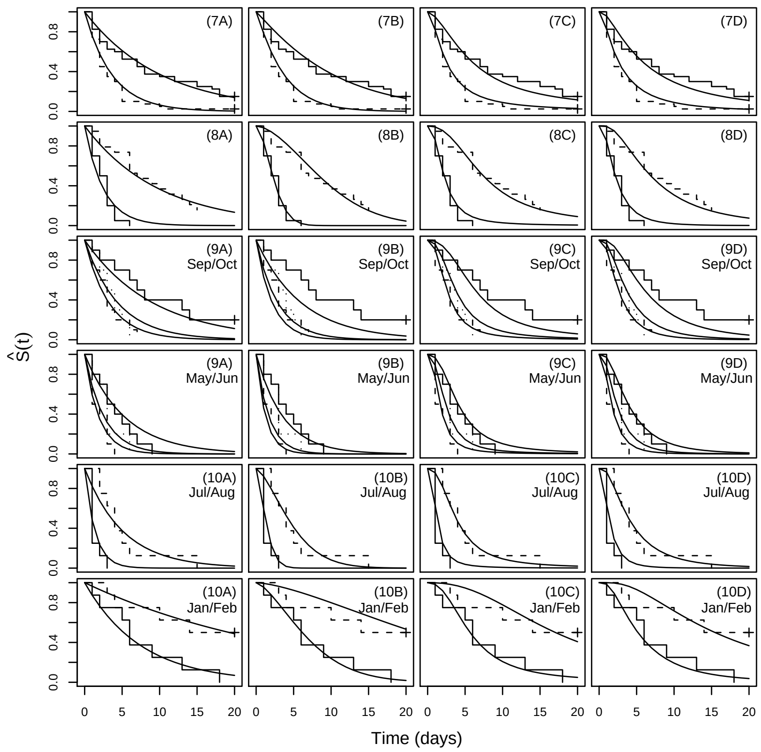

Fig. 2 Empiric (step functions) and fitted survivor functions for non-homogeneous populations. Plots represent estimated functions at 7 - WF7 (step solid line: Jan/Feb and step dashed line: Jul/Aug), 8 - WF8 (step solid line: May/Jun and step dashed line: Sep/Oct), 9 - WF9 (step solid line: small size carcasses, step dashed line: medium size carcasses and step dotted line: large size carcasses) and 10 - WF10 (step dashed line: medium size carcasses and step dotted line: large size carcasses) wind turbines sites, regarding A - exponential, B - Weibull, C - log-logistic and D - log-normal models.

pend heavily on the model selected, procedures used to check model adequacy are particulary important in this context. Lawless (2003) underlines that

> Often data are analyzed under a particular model simply because (1) the model has been used before in similar situations, or (2) it fits the data on hand. This does not imply any absolute validity of the model, and we should ask whether inferences change much if another similar "plausible" model is used instead.

So, recognizing that the carcass persistence probabilities can, in fact, depend heavily on the model selected, this study aimed to discuss some common methods used to discriminate between survival models and, simultaneously, to select the most suitable model for the data sets from the 10 Portuguese wind farms.

Regarding used methods, we found plotting procedures to be clearly not enough for perform model selection. In this study we used plots based on the linearization of the survivor function and plots that presented simultaneously empirical and parametrical estimated survivor functions. In both cases, the diferences between statistical models were of difficult eye perception. The analysis of the plots based on the linearization of the survivor function have the advantage of being coadjuvante by coefficients of determination, making easier the models evaluation.

The calculated goodness of fit measures — AIC and BIC — showed a very strong agreement between them, concerning the model choice. In all the 10 analyzed data sets both measures pointed to the same model selection. These measures allowed to choose the best fitted model, even when values were very close between fitted models, i.e., even when a similar model appropriateness was found. In 8 wind farms, log-normal and log-logistic regression models fitted better the data. In the two remaining wind farms, the Weibull model adjusted better.

We conclude that plotting procedures do not discriminate sufficiently enough the fitted models and that AIC and BIC allow to select the best model and point consistently to the same choice. Plots are nonetheless interesting to ilustraste model adjustment after model choice.

Acknowledgements This research was conducted in collaboration with BIO3 - Estudos e Projectos em Biologia e Valorização de Recursos Naturais.

References

1. G. Johnson, W.P. Erickson, M.D. Strickland, M.F. Shepherd, D.A. Shapherd, S.A. Sarappo, Am. Midl. Nat. **150**, 332 (2003)
2. A.L. Drewitt, R.H.W. Langston, Ann. N. Y. Acad. Sci. **1134**, 233 (2008)
3. K.S. Smalwood, The J. Wildl. Manag. **71(8)**, 2781 (2007)
4. D. Collet, *Modelling Survival Data in Medical Research* (Chapman & Hall/CRC, 2003)
5. A.D. Block, L.M. Leemis, IEEE Trans. on Reliab. **57**, 248 (2008)
6. J.D. Kalbfleisch, R.L. Prentice, *The Statistical Analysis of Failure Time Data* (John Wiley and Sons, 2002)
7. J.F. Lawless, *Statistical Models and Methods for Lifetime Data* (Wiley, 2003)
8. D.R. Cox, D. Oakes, *Analysis of Survival Data* (Chapman & Hall, 1998)
9. W.P. Erickson, J. Jeffrey, K. Kronner, K. Bay., Stateline wind project wildlife monitoring final report. Tech. rep., Western EcoSystems Technology, Inc. and Northwest Wildlife Consultants, Inc. (2004)