Predictive performance of Mixed-Effects Cox regression and Learning Neural Network model with application in agriculture

Judicaël LALY *
Unité de Biostatistique et de Modélisation (UBM), ENSBBA
Université Nationale des Sciences, Technologies, Ingénierie et Mathématiques (UNSTIM)
Dassa-Zoumé, Benin
judicaelmaurilelaly@gmail.com

Charlemagne D.S.J. Gbèmavo
Unité de Biostatistique et de Modélisation (UBM), ENSBBA, UNSTIM
Dassa-Zoumé, Benin
cgbemavo@yahoo.fr

Hubert A. Dossa
Unité de Biostatistique et de Modélisation (UBM), ENSBBA, UNSTIM
Dassa-Zoumé, Benin
hubertadonha@gmail.com

Romain L. Glèlè Kakaï
Laboratoire de Biomathématiques et d'Estimations Forestières (LABEF), UAC
Abomey-Calavi, Benin
glele.romain@gmail.com

Abstract

The Mixed effects Cox model and survival neural networks are commonly used to predict the hazard rate of an interesting event. This paper compared the predictive ability of the two techniques using simulation tools based on sample size and the censoring rate of mixed effects data from agriculture. Thus, varying into three sample sizes and five censoring rates, fifteen different datasets with 1000 replications were simulated based on the characteristics of a real data from *Jatropha Curcas* L. seeds germination trial. Based on three performance metrics (Brier Score, Integrate Brier Score, and Concordance index), the mixed effects Cox model better predicted the survival outcomes than the Partial Logistic Artificial Neural Network original on relatively low sample size (768 or 1536) and middle censoring rate (40 % or 50.65 %). Therefore, the dimension and censoring rate of the relevant dataset must be considered when selecting one of the two models for analysis.

1 Introduction

In agriculture, the developmental and growth events in a plant's life are connected to specific points in time, be it seed germination, seedling emergence, the appearance of the first leaf, flowering, fruit

*Phone number: +22961798611

5th Deep Learning Indaba Conference (DLI 2023).

ripening, wilting, or death [5]. Indeed, techniques of survival analysis are used on agricultural data to predict the risk and survival probability of those event mainly seed germination [6]. Besides, in order to make a good explanation of the developmental and growth events related to the plant's lifetime, the survival data collected usually involves as many experimental factors as possible that could affect the event of interest in a given study. Depending on data type, most survival analysis techniques used are Cox proportional hazards models [6, 12], and mixed effects Cox models but also Survival neural networks from machine learning technique [2].

The Cox proportional hazards (Cox PH) models are introduced and applied in agriculture mainly to estimate the effects of biotic and/or abiotic factor(s) on the instantaneous risk of an interesting event in a hypothetical study [6, 12]. But in applied statistical practice, clustering, hierarchical designs, longitudinal studies, or repeated measurements can all lead to grouped data structures defined by one or more grouping factors [15]. Hence the development of mixed effects Cox models used to estimate the instantaneous risk of an event using data containing fixed and random factors [7, 9, 12]. As an example, a mixed effects Cox model was used by Hamann et al [7] to evaluate the longevity costs of first-year fitness and failed reproduction of *Boechera stricta* using R coxme package v. 2.2-16 [16]. Concerning the machine learning tools, neural network models are useful alternatives in modeling survival data when the assumptions of a classical parametric or semiparametric survival model are seriously violated [3]. Nowadays, several learning neural network models have been developed and adapted to deal with Cox PH models for right-censored survival data depending on study objectives [2], and are either single-layer or multi-layer architectures [1, 3, 8, 14]. As illustrations, we have the original and extended versions of the partial logistic artificial neural network (PLANN) [8] as well as the kernel emerging extreme learning machine Cox model regularized by a based broken adaptive ridge (ELMCoxBAR) [3]. But, in presence of a dataset with mixed factors, which of the tools (the classical mixed effects Cox model or the survival neural network model) can better predict the hazards of the interesting event ?
In order to overcome that challenge, some researches have already been done to compare the prediction performance of survival analysis to that of machine learning tools, especially the comparison of Cox proportional hazard models to artificial neural networks using simulation approaches[8]. The more frequent predictive performance measures for those models are the Brier score (BS), the Integrated Brier score (IBS), and Harrell's concordance index (C-index) [8]. Unfortunately, any of the previous research did not identify the best model between the mixed effects Cox model and the learning neural network for hazard rate prediction, in spite of grouping factors frequently encountered in practice, mainly on agricultural data. The current study proposes to bring out the predictive ability of those two different statistical tools through a mixed effects Cox model (model with random effect) and a survival neural network (PLANN original) using mixed censoring data by answering two (02) research questions: (1) Which of those two statistical tools has the best performance for hazard rate prediction? And (2) what happens to the predictive performance of those two different techniques on real data from the agriculture field? The results of these questions will help to improve the analytics quality of censored data with mixed effects.

2 Data and Methodology

2.1 Data simulation

Data simulation has been done based on a real data from a crop production experiment, especially *Jatropha Curcas* L. seeds germination follow-up containing three (3) categorical variables (Locality, Accession, and Block) and two (2) numerical variables (Time and Status). The experimental design used is a Randomized Complete Block Design with three treatments (Accessions) and 4 repetitions (Blocks). In the real dataset, the factor "Accession" is nested to "Block" and "Block" is nested to "Locality". So the size of each simulated dataset was the product of number of localities, accessions, blocks and block units. Then, three different sample sizes (768, 1536, and 2304) were defined. The Time was generated from Weibull distribution (shape = 1.7067 and scale = 18.2140) and Status with binomial distribution with five scenarios of 20 %, 40 %, 50.65 %, 60 %, and 80 % censoring rates.

2.2 Implementation of models and performance metrics used

The implementation of both MECox model and PLANN has been done on each dataset splitted into two sub-datasets, 75 % for model training and 25 % for model test. For the training of MECox model, the R packages *survival* and *coxme* were used. The hazards were predicted using the equation 1 with baseline hazards λ_0 estimated using *muhaz* R function from the package *muhaz*. To implement the PLANN approach, we followed the steps of data transformation and data training as Kantidakis et al [8] refered to the equation 2. PLANN provides the discrete conditional probability (0 or 1) used to compute estimated hazard in each interval and after the survival probabilities.

$$\lambda(t_{ij}, X_{ij}, b_j) = \lambda_0(t)\exp(X_{ij}\beta + b_j) \tag{1}$$

with t time of event, X fixed factors, b random effect, β vector of the regression parameters, i observations, and j Blocks.

$$\lambda(\tau_I, X, W) = f\left[W'_{0k} + \sum_{h=1}^{H} W'_{hk} g_h\left(W_{0h} + W_{1h}\tau_I + \sum_{j=1}^{j=p} W_{(j+1)h} x_{ij}\right)\right] \tag{2}$$

With j nodes in input layer, h nodes in hidden layer $k = 1$ node in output layer, x_{ij} p elements of covariate vector x_i, w_{jh} weights from input to hidden layer, w'_{hk} weights from the hidden to the output layer, w_{0k} and w_{hk} weights of the bias nodes for the input-hidden and the hidden-output layers respectively. The activation function of both the hidden and output layers was sigmoid function.

Three (03) performance measures were used (Table 1) and two independent sample t-tests to test the significance level of the differences between models for each scenario mention them on barplots.

Table 1: The three (03) performance measures used

Performance metrics	Formular	R function
Brier Score (BS)	$(y - \hat{S}(t_0\|x))^2$	*Brier*
Integrated Brier Score (IBS)	$\frac{1}{t2-t1}\int_{t1}^{t2} BS(u)\,du$	*IBS*
Concordance Index (C-index)	$\frac{Conc\ pairs}{Conc\ pairs + Disconc\ pairs}$	*Cindex*

With $\hat{S}(tx)$ model-based survival probability, x predictor and y actual observation-ignoring censoring.

3 Results

3.1 Comparison of predictive performance over sample size

Figue 1) showed the variation of the performance metrics used over the sample size for both models studied. Thus, IBS (Figure 1A) showed that the MECox model better performed survival outcomes than the PLANN on all three sample sizes studied. The C-index revealed the same comparison when using the first two sample sizes (768 and 1536), but the contrary with the largest sample size of the study (2304) (Figure 1B). For both metrics, the MECox model better predicted survival outcomes on a dataset with relatively low observations (768 or 1536) and 50.65 % censored data than PLANN.

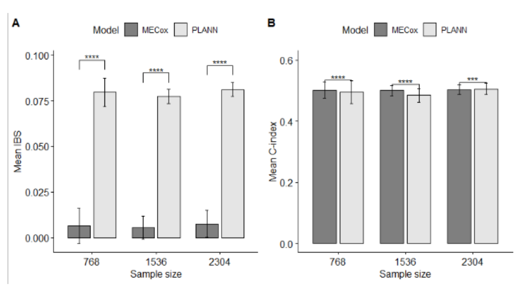
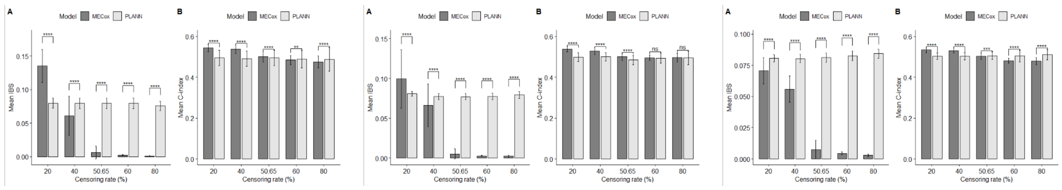


Figure 1: Variation of Integrated Brier Score (A) and C-index (B) over sample size.

3.2 Comparison of predictive performance over censoring rate knowing the sample size

Figure 2 showed the variation of IBS and C-index following the censoring rates for each of the sample sizes. Regarding the sample sizes 768 (Figure 2a) and 1536 (Figure 2b) , both IBS and C-index values decreased when the censoring rate increased for both two models, but this was more pronounced with the MECox model than PLANN. For the sample size 2304 (Figure 2c), IBS and C-index values decreased when the censoring rate increased for the MECox model but kept relatively constant for PLANN. In terms of comparison, IBS and C-index showed that the MECox model better predicted the survival outcomes than the PLANN with a dataset of 768 or 1536 observations and censoring rates of 40 %, and 50.65 %, and also 20 % when the dataset contains until 2304 observations.



(a) Sample size equal to 768 (b) Sample size equal to 1536. (c) Sample size equal to 2304.

Figure 2: Variation of Integrated Brier Score (A) and C-index (B) over censoring rate.

3.3 Application on real data results

During the application section, the data of baobab (*Adansonia digitata L.*) seeds germination was used, extracted from a split plot design (completely randomized) with three repetitions (Block). Two fixed factors (climatic zone and water stress) were tested, with each three modalities. As results, the C-index value was greater for the MECox model than that for PLANN contrary to IBS obtained for the mixed effects Cox model lower than that for PLANN. Since the size of the real data used was low (540) compared to all the sample sizes used in the simulation method, and the censoring rate (65.74 %) closed to 50.65%, the results of simulation tools were confirmed.

4 Discussion

This study assessed the prediction performance of the MECox model and PLANN original on mixed data using Monte Carlo simulation with 1000 replications and its comparison with real data's results. The results varied over sample size and censoring rate, and performance metrics perfectively were reasonables refering to some previous researches [4, 11, 13]. Thus, the comparison of both models studied over the data sample size regarding each of the performance metrics used revealed mainly three different situations of prediction behaviors. There was the case where the MECox model could better perform on survival mixed data than PLANN regardless of the sample size, the case where both models could have a similar performance on a dataset of 1536 objects, and the case where PLANN could better perform with a sample size larger than the largest (2304) used in this study. Refering to the previous studies [4, 13], all those three situations could be possible under some conditions. First, a critical review work of Kantidakis et al [13] saw among 19 studies comparing SNN and Cox model's performance, 9 (47.4 %) claimed better predictive performance of the SNN, while 5 reported a similar or better performance (26.3 %) of the SNN compared to the Cox model, and 5 (26.3 %) showed similar performance. Beside, Wang u.a. [11] showed that the prediction performance of classical survival models was slightly high than the survival neural network, regardless of the data sample size. Regarding the variation of data sample size, the estimated mean c-index obtained by Wang u.a. [11] in the case of small data sample size was high for the cox model and lowest for SNN when the mean IBS showed the contrary with a low difference. Likewise, Kantidakis et al [8] found that the Cox model performed slightly better than the SNNs with a larger sample size (N = 1000), and the predictive performance of both models improved when the sample size increased (smaller Brier scores, higher C-indexes). Concerning the variation of IBS and C-index measured over the censoring rate for each single sample size used, the decreasing of C-index values over the censoring rate observed on all three sample sizes (768, 1536, and 2304) for both two models could be explained by the results of Alabdallah et al [10] where the C-index of CPH and all ML models studied decreased over the censoring rate. The results obtained with the real data of *Adansonia digitata L.* seeds germination also proved how well the simulation properties had been respected.

5 Conclusion

In summary, the predictive performance of the mixed effects Cox model and the partial logistical artificial neural network on agriculture data depends on the sample size and censoring rate. The mixed effects Cox model was the best in most of the data scenarios, especifically for dataset of relatively low observations and a censoring rate less or equal to 50 %. Then, a good choice between both models for hazard prediction should base on those two data components.

Acknowledgements: This study was carried out as part of the Master's Degree in Statistics, major Biostatistics obtained atLaboratoire de Biomathématiques et d'Estimations Forestières (LABEF) of the University of Abomey-Calavi (UAC) through the thesis "Predictive performance of Mixed-Effects Cox regression and Learning Neural Network model with application in agriculture". We express our sincere thanks to all people those contributed to study realization mainly the staffs of LABEF, Green Tech Innovation (GTI) SARL and Biometrics unit of IITA Ibadan.

References

[1] Katzman, Jared L / Shaham, Uri / Cloninger, Alexander / Bates, Jonathan / Jiang, Tingting / Kluger, Yuval(2018): *DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network*, 1: 1–12.

[2] Bayat, Mahmoud / Ghorbanpour, Mansour / Zare, Rozita / Jaafari, Abolfazl / Pham, Binh Thai(2019): *Application of artificial neural networks for predicting tree survival and mortality in the Hyrcanian forest of Iran*104929.

[3] Wang, Hong / Li, Gang(2019): *Extreme learning machine Cox model for high-dimensional survival analysis*, 12: 2139–2156.

[4] Wang, Ping / Li, Yan / Reddy, Chandan K(2019): *Machine learning for survival analysis: A survey*, 6: 1–36.

[5] Humplík, Jan F / Dostál, Jakub / Ugena, Lydia / Spíchal, Lukáš / De Diego, Nuria / Vencálek, Ondřej / Fürst, Tomáš(2020): *Bayesian approach for analysis of time-to-event data in plant biology*1–7.

[6] Romano, Alessandro / Stevanato, Piergiorgio(2020): *Germination Data Analysis by Time-to-Event Approaches*, 5: 617.

[7] Hamann, Elena / Wadgymar, Susana M / Anderson, Jill T(2021): *Costs of reproduction under experimental climate change across elevations in the perennial forb Boechera stricta*, 1948: 20203134.

[8] Kantidakis, Georgios / Biganzoli, Elia / Putter, Hein / Fiocco, Marta(2021): *A Simulation Study to Compare the Predictive Performance of Survival Neural Networks with Cox Models for Clinical Trial Data*.

[9] Aguado, William D / Rogers, Haldre S / Pruetz, Jill D(2022): *Chimpanzees as ecosystem service providers: Seed dispersal of an economically important plant resource*.

[10] Alabdallah, Abdallah / Ohlsson, Mattias / Pashami, Sepideh / Rögnvaldsson, Thorsteinn(2022): *The Concordance Index decomposition: a measure for a deeper understanding of survival prediction models*.

[11] Wang, Meng / Greenberg, Matthew / Forkert, Nils D / Chekouo, Thierry / Afriyie, Gabriel / Ismail, Zahinoor / Smith, Eric E / Sajobi, Tolulope T(2022): *Dementia risk prediction in individuals with mild cognitive impairment: a comparison of Cox regression and machine learning models*, 1: 284.

[12] Cirino, Lauren A / Moore, Patricia J / Miller, Christine W(2022): *High-quality host plant diets partially rescue female fecundity from a poor early start*, 2: 211748.

[13] Kantidakis, Georgios / Hazewinkel, Audinga Dea / Fiocco, Marta / others u.a.(2022): *Neural Networks for Survival Prediction in Medicine Using Prognostic Factors: A Review and Critical Appraisal*.

[14] Kim, Bomee / Jang, Yun Ji / Cho, Hae Ram / Kim, So Yeon / Jeong, Ji Eun / Shim, Mi Kyoung / Kim, Myeong Gyu(2022): *Predicting completion of clinical trials in pregnant women: Cox proportional hazard and neural network models*, 3: 691–699.

[15] Tamási, Bálint / Hothorn, Torsten(2022): *tramME: Mixed-effects transformation models using Template Model Builder*Epub–ahead.

[16] Therneau, T (2012): *Coxme: mixed effects Cox models. R package*
.