DOI: 10.48011/asba.v2i1.1662

Time to Event Analysis for Failure Causes in Electrical Power Distribution Systems ⋆

Matheus S. S. Fogliatto* Luiz Desuó N.* Rafael R. M. Ribeiro* Talysson M. O. Santos* João B. A. London Jr.* Michel Bessani Carlos D. Maciel***

* *Department of Electrical and Computing Engineering, São Carlos School of Engineering, University of São Paulo,São Carlos, SP, Brazil, (e-mail: matheusfogliatto@usp.br, carlos.maciel@usp.br)*
** *Department of Electrical Engineering,Federal University of Minas Gerais,Belo Horizonte, MG, Brazil*

Abstract: Electricity is a fundamental resource for modern society. However, some threats are faced by Electrical Power Distribution Systems, which are responsible for delivering electricity to end consumers. Analyzing how much time these hazards will threaten these systems, causing failure events, is an essential area of study. Through statistical methods, it is possible to study this behaviour from time until failure, as well as to observe the influence of variables at this time, providing models to predict when a failure event will occur. In this study, Reliability Analysis Regression techniques are used on real data, constructing a model for all failures and for different groups of failures, using non-parametric and parametric methods to estimate the reliability and cumulative hazard curves. An analysis of the failure causes directly linked to weather events, using six weather variables, is also made.

Keywords: Causes of Failure, Distribution System, Failure event, Reliability/Survival Regression.

1. INTRODUCTION

Electrical power Distribution Systems (DSs) are the part of the power systems responsible for delivering electricity to final customers, such as homes, hospitals and industries. Therefore, maintaining the operation of them in normal condition is a concern of modern society (of Economic Advisers et al., 2014)(Brem, 2015). However, as DSs generally occupy a large area they are exposed to external environment, and end up being subject to various threats that can cause failures, that is, interruption in the supply of electricity. As a consequence, the study of DS failures is a current concern of researches, public agents, and society, in order to enhance the reliability and resilience of DSs (Zio, 2009).

The leading causes of failures in DSs (Złotecka and Sroka, 2018)(Sroka and Złotecka, 2019) are accidents with animals (Sahai and Pahwa, 2006) or vegetation (Radmer et al., 2002), weather events (Konal et al., 2018)(Pahwa, 2007), load transfer (Rodriguez-Garcia et al., 2019), device failures (Li et al., 2019) and rare events like terrorism, vandalism and cyber-attacks (Ni and Li, 2019). Therefore, several factors may be associated with these failures, from equipment failure to vehicle accidents colliding with network elements, and for a better understanding of them, a separation of these failures into groups (bringing together similar threats or causes) may be interesting.

Two factors are often recurrent in failure studies: how long until it happens (time to failure or lifetime) and how long until the system recover itself (time to repair) (Bessani et al., 2016). For these types of studies, Reliability Analysis Regression (also know as Survival Analysis Regression) (Colosimo and Giolo, 2006)(Cox and Oakes, 2018) provides models to estimate the reliability and hazard using time as variable, making possible to determine the group of failure that is more recurrent, unlike an analysis considering only the number of failure events for each group. This approach is used in areas such as Medicine (Moolgavkar et al., 2018) and Engineering, in general problems (Dantas et al., 2010) or more specifically, as in Reliability Engineering (Murthy et al., 2004). For these models (Mishra et al., 2019), three major techniques can be found in the literature: the parametric (Zhang, 2016), the non-parametric (Rink et al., 2013),and the semi-parametric (Shauly et al., 2011). Parametric Statistics need a family of probability distribution and can be more accurate than non-parametric techniques (Zhang, 2016). From previous works, the Weibull distribution proved a good choice for problems involving DS failures (Fogliatto et al., 2019).

For some failures groups, weather events are directly linked with the occurrence of a failure event. Regression including weather events as covariates can be used to evaluate the impact of a particular weather event in a failure group, by analyzing the value of the coefficients for these variables with the values for the model constructed considering all failure events. A understanding of the characteristics of the different groups of failure events that threaten DSs is

⋆ This work was partially support FAPESP: 2014/50851-0, CNPq: 465755/2014-3; COPEL PD 2866-0504/2018 and BPE Fapesp 2018/19150-6.

useful for improvements in the confiability of the system, and real data is useful because the results achieved can be used in locations with similar characteristics.

In this paper, the different causes of failure events in a real DS from a brazilian city, from a period of almost two years, are separated in five groups, according to the characteristics of these events. A group considering all failures events (all groups together) are used. The main outputs of Reliability Analysis Regression, that are the reliability and cumulative hazard functions, are plotted by non-parametric (Kaplan-Meier and Nelson-Aalen) and parametric (Weibull) techniques for all groups, to estimate the dangerous of each failure group in terms of the lifetime of the system (time intervals between failures). For all groups, the non-parametric estimators and the Weibull Univariate was used considering only the lifetime values, and to include weather events as covariates the Accelerated Failure Time (AFT) model was used. For the failure groups that have some relation with weather events and for the "All Failures" group, an analysis using Weibull AFT model was used to include six covariates (from data of maximum daily values of the number of atmospheric discharges, wind speed, maximum and minimum temperature, precipitation and relative humidity of the analyzed city) to measure the impact of weather in the lifetime of the system. From the models including weather covariates, the reliability function graphics varying the weather values was shown to demonstrate how these covariates modify the lifetime of the DS. Also, tables of predictions in terms of a median and an expected time for the survival of the system from the three groups that weather covariates were included are presented.

The paper is organized in four sections. Section II presents: a brief description of the analyzed real DS; the failure dataset, with examples of the failure groups that were used; an example table of the weather data; an analysis of the issue count of failures for each group, showing the threat order of each group in percentage terms; the statistical theory for the univariate and AFT regression models, for the Kaplan-Meier / Nelson-Aalen methods for the non-parametric and the Weibull distribution for the parametric analyses. Section III summarizes the results for the constructed models, with equations and tables showing the coefficients for the regressions, and graphics of the reliability and cumulative hazard functions, the main outputs for these methods. Section IV presents the conclusions and final remarks.

2. MATERIAL AND METHODS

The real Brazilian DS illustrated in Fig. 1 was used in this research. Associated with this system, data from failure events and data from weather events were used to provide a model to predict the lifetime of the system, using reliability analysis regression techniques.

A Failure event dataset provides the information used in the development of the models of this text. The lifetime and the event information are the information necessary for the Reliability Analysis models. Lifetime, which was calculated using the starting date and hour of the each failure from the original failure dataset, represents the amount of time the system has been operating under

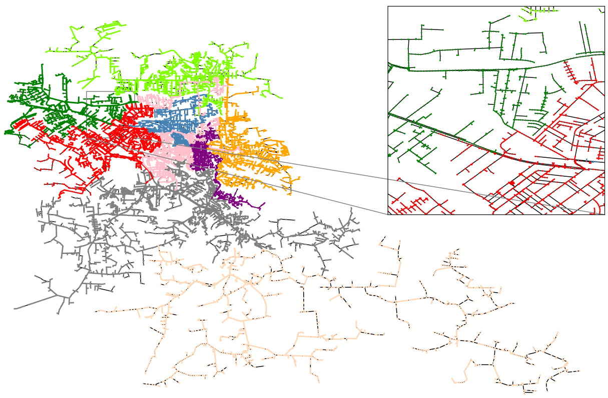


Figure 1. DS of a midsize Brazilian city with 500.000 people, and about 85 years old, composed of 8 substations, 65 feeders, 40143 buses (distribution transformers) and a network with length of 1660.17 kilometers. A subfigure is plotted for more details. Each color represents the feeders of the same substation.

normal conditions until a failure has occurred (time to event/failure) and is presented in minutes. It is considered that the system fully recovers after a failure, thus starting a new survival period. From the analysed period, 12028 failure events were registered (failures of any duration). The "Event" information is to represent the abstention of censored data, that is, all of the lifetimes are complete observations (information most useful in time to failure of equipments or in medicine reasearches, where the study has a time limit. If a patient or the equipment did not fail or died until this deadline, the lifetime of this sample is considered censored).

The weather dataset is composed of the maximum daily values for six weather events in the period from February/2013 to December/2014 (689 days): the number of atmospheric discharges, wind speed in kilometre/hour, amount of precipitation in millimetres, the maximum and minimum temperature in Celsius and relative humidity in percentage.

From all failure events, the maximum lifetime was 6419 minutes. The maximum number of atmospheric discharges was 2267, the maximum wind speed was 97.2 km/h, the maximum amount of precipitation was 80.8 mm, the maximum relative humidity was 98% and the maximum and minimum temperature was 38.3 and 10.2 C, respectively, for the analysed period. A histogram of the lifetime is presented in Fig. 2 and a concentration of the number of occurrences from 0 to 150 minutes is observed. The mean lifetime value was 87.5 minutes.

From the failure dataset, detailed information about the type and the cause of the failures could be found. Type of failure has three different groups: Accidental, Volunteer and Scheduled failures. Examples of failure for each type are presented in Tab. 1. From the used data, 21.82% of failures are scheduled, 23.00% are volunteers and 55.17% are accidental failures. Therefore, more than 75% of the failures events analysed in this research are not planned by the company that manages the distribution system.

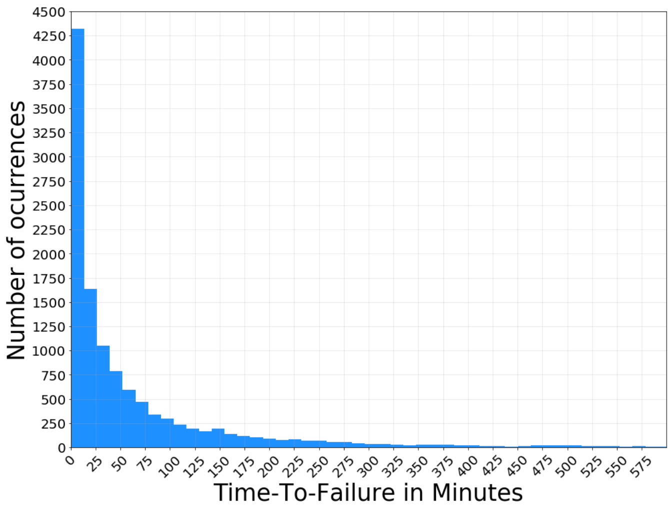


Figure 2. Histogram for the most concentration of survival time of the DS analysed.

Table 1. Table exemplifying the "Type" of failures observed in the failure dataset.

	Ex.1	Ex.2
Accidental	Animals/Insects/Birds	Branches touching the network (pruning)
Volunteer	Maneuvers	Operational opening of another switch
Scheduled	Improvements or/and Expansion	Load transfer

A more specific separation was made establishing six groups for the cause description of the failure events: Equipment Failure, Urban events, Operational, Environmental events, Atmospheric events and a group for failures that did not have the cause identified. These groups are exemplified in Tab. 2. The percentages of the number of failure events are 6.46% from Atmospheric causes, 9.22% for Environmental, 13.77% for Equipment Failure, 45.90% for Operational causes, 7.71% for Urban events and 16.94% have non-identified causes.

For Cause Description, Improvements/Expansion of the network represent 12.98% of the failure events, accidents with animals with 5.45% and atmospheric discharges with 4.55% can be highlighted. About 16.9% of the causes of the fails are not identified.

Table 2. Table exemplifying the Groups of failures observed in the failure dataset.

	Ex.1	Ex.2
Not Identified	-	-
Equipment Failure	Capacitor Bank	Oil Switch
Urban	Vandalism/Theft	Bump
Operational	Third party request	Energy Rationing
Environmental	Burnt/Fire	Branches touching the network (pruning)
Atmospheric	Erosion	Atmospheric Discharge

2.1 Statistics

Reliability Analysis was used with non-parametric and parametric techniques for each one of the proposed groups, and for considering all failures, to analyse the lifetime of a distribution system. Besides that, climatic events are associated with Atmospheric and Environmental failure groups, to observe the influence of different values for these variables in the survival of the analysed DS.

Reliability Analysis have two main outputs (Davidson-Pilon et al., 2019)(Cox and Oakes, 2018): Reliability (Survival) Function, $R(t)$, and Cumulative Hazard Function, $H(t)$. Reliability Function is the probability that a system survives longer that time t, and Cumulative Hazard Function is the accumulation of the hazard (hazard is the event rate at time t conditional on survival until time t or later $(R(t) = P(T \geq t))$ over time. First, for Reliability Function, the non-parametric technique (Chowdhury et al., 2015) is called Kaplan-Meier (Kaplan and Meier, 1958), and for Cumulative Hazard, Nelson-Aaalen (Aalen, 1978) . In Eq. 1 and 2, the general form of Kaplan-Meier and Nelson-Aalen estimators, respectively, are presented. The variable n_j is the number of samples at risk in the time t_j, and d_j is the number of ocurred events at time t_j.

$$\hat{S}(t) = \prod_{j:t_j<t} \left(\frac{n_j - d_j}{n_j} \right) \tag{1}$$

$$\tilde{\Lambda}(t) = \sum_{j:t_j<t} \left(\frac{d_j}{n_j} \right) \tag{2}$$

Weibull model, in the form of the univariate model and the Accelerated Failure Time (AFT), was used as the parametric technique. The univariate model does not include covariates, and was used in conjunction with non-parametric techniques to analyze all failure groups. Eq. 3 and 4 present the reliability and cumulative hazard functions, respectively. The shape parameter and the scale parameter are respectively ρ and λ, and t is the variable for lifetime.

$$R(t) = exp\left(- \left(\frac{t}{\lambda} \right)^{\rho} \right), \lambda > 0, \rho > 0 \tag{3}$$

$$H(t) = \left(\frac{t}{\lambda} \right)^{\rho} \tag{4}$$

The AFT model were applied for three groups, all failures, Atmospheric and Environmental due the used covariates were weather events. Eq. 5, 6 and 7 present the equations for the Weibull AFT model output functions. The coefficients of the regression model are represented by β_i variables, and the real values for the weather events are the x_i variables.The y variable in equations 6 and 7 represent that values for ρ are "independent", in the meaning that it can include covariates or not (in this research, not included).

$$\lambda(x) = exp(\beta_0 + \beta_1 x_1 + ... + \beta_n x_n) \tag{5}$$

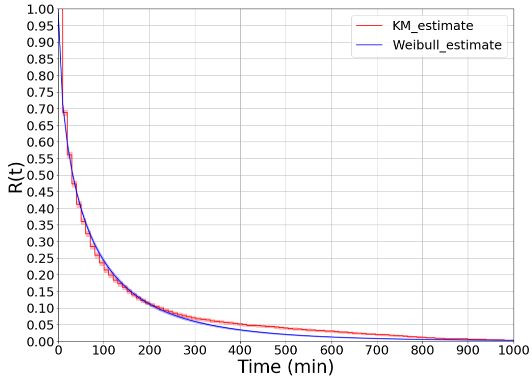
$$R(t; x, y) = exp\left(- \left(\frac{t}{\lambda(x)} \right)^{\rho(y)} \right) \tag{6}$$

$$H(t; x, y) = \left(\frac{t}{\lambda(x)} \right)^{\rho(y)} \tag{7}$$

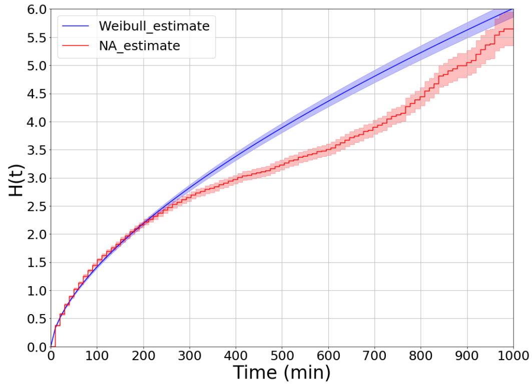
AFT models do the addition of covariates to regression models. The analyse of the values and some statistical criteria can make the relevance for the model of these covariates. The classical approach when evaluating covariates coefficients significance from regression models is the $p - value$ (Wang et al., 2019). $P - values$ equal or lower than 0.05 are considered statistically significant, but this value has been discussed (Amrhein et al., 2019), so there is no need to discard covariates which have a higher value. Another statistical criteria generally presented are the standard error, that is the standard deviation of its sampling distribution, and the confidence interval, which is a range of values that can contain the real value of an unknown parameter. In this research, the 95% confidence interval is presented.

3. RESULTS

Considering all failures, Fig. 3 presents the $R(t)$ and $H(t)$ functions for both the Weibull univariate (blue color) model and Kaplan-Meier and Nelson-Aalen non-parametric techniques (red color). As the proximity between the curves can be observed, we can conclude that the Weibull model presents a good fit for the data used. The non-parametric techniques represent the real data, as some "downhill" can be observed, and the Weibull Model has a "smoothed" curve, being useful for any desired value, and being a more reliable model for forecasting.



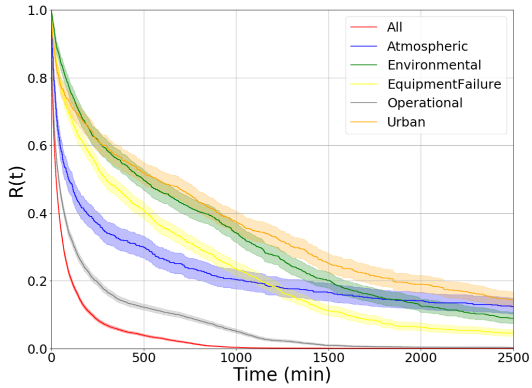
(a)



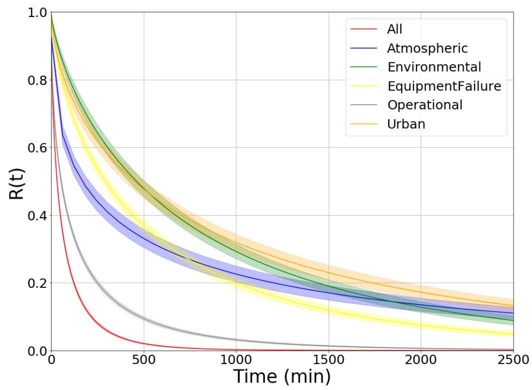
(b)

Figure 3. (a) Reliability Function for the Weibull Fitter, considering all causes of failures. (b) Cumulative Hazard Function for the Weibull Fitter, considering all causes of failures.

The same approach was used for the five groups of failure events, and Fig. 4 and Fig. 5 present the $R(t)$ and $H(t)$ functions for each of them. Considering the period of 0 to 500 minutes (which represents most of the survival periods of the system used), in the order of most dangerous to the lifetime of the DS to the less dangerous, Operational failures was the most impactful, followed by Atmospheric, Equipment Failures, Environmental and Urban. The order of the threat groups differs from the order by the number of failure events occurring, which was Operational, Equipment Failure, Environmental, Urban and Atmospheric.



(a) R(t) for Kaplan-Meier Fitter.

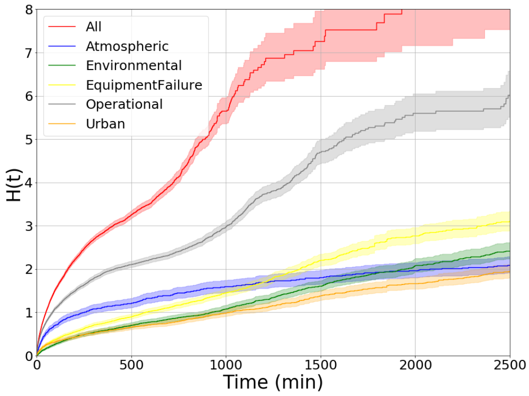


(b) R(t) for Weibull Fitter.

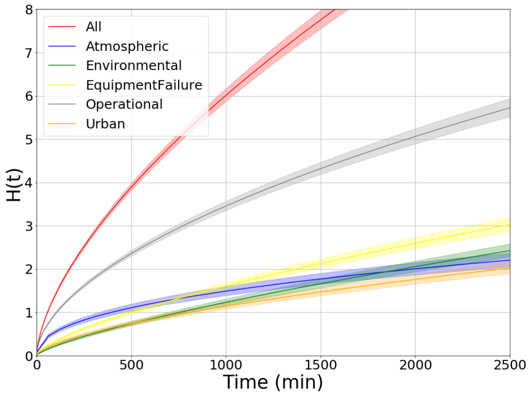
Figure 4. (a) Reliability Function graphic using non-parametric technique of Kaplan-Meier Fitter and (b) using parametric technique of Weibull Fitter, for the groups of failure causes.

Disregarding Operational failures, that are the most dominant cause, the other four groups have a similar percentage in the question of the number of failures. Analysis by the reliability and cumulative risk curves provide a different view. The external factors represented for Atmospheric, Environmental and Urban groups can be evaluated, and the study of weather events that are directly linked with Atmospheric and Environmental failure groups is justified.

The Weibull AFT Model Equations considering all failures are presented in Eq. 8 and 9. "AD" is the number of atmospheric discharges, "W" is the wind speed, "Tx" is the Maximum temperature, "Tn" the minimum temperature, "P" the amount of precipitation and "HR" is the relative humidity. In Fig. 6, the $R(t)$ function is plotted, considering different values (chosen from random days of the used dataset) for the three main covariates: Wind Speed, number of atmospheric discharges and relative humidity. For these curves, all other covariates were considered at their mean values. The baseline curve is considering all covariates at their mean values.



(a) H(t) for Nelson-Aalen Fitter for each failure group.



(b) H (t) for Weibull Fitter for each failure group.

Figure 5. (a) Cumulative Hazard Function graphic using non-parametric technique of Nelson-Aalen Fitter and (b) using parametric technique of Weibull Fitter, for the groups of failure causes.

$$\lambda = exp(5.18783 - 0.00049AD - 0.02995W + 0.00831Tx - 0.00302Tn - 0.00174P - 0.00308HR) \tag{8}$$

$$R(t) = exp\left(-\left(\frac{t}{\lambda}\right)^{-0.42509}\right) \tag{9}$$

In Eq. 8 a positive value, as in the intrinsic and maximum temperature coefficients, represent a positive contribution to the lifetime of the system. Therefore, a negative value, as for atmospheric discharges, wind speed, minimum temperature, amount of precipitation and relative humidity, decrease the survival time (they are a threat for the system).

From the Weibull AFT Model, it is possible to predict in terms of the median (percentile) of survival and the expectation ($E[T|x]$). First, percentiles are measurements that divide the sample in ascending order of data into 100 parts, each with an approximately equal percentage of data, and the median is the 50 percentile. Second, the expectation of a random variable is the sum of the product of each probability of leaving the analysis by its respective value, representing the "expected" average value of an experiment if repeated many times. In Tab. 3 is presented these predictions for real values of the weather data.

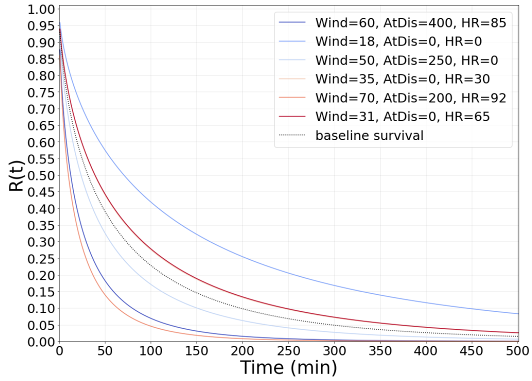


Figure 6. Reliability Function plot for Weibull AFT Regression model, varying the wind speed, the number of atmospheric discharges and the value for relative humidity, and keeping the mean value for the other variables. The baseline survival is considering all covariates as theirs mean value.

Table 3. Prediction of the Percentile and the Expectation of the lifetime of the analysed distribution system for real values of the weather events, considering all failure events.

Percentile	Expectation	AD	W	Tx	Tn	P	HR
46.2876	110.0950	0.0	24.0	28.8	21.5	6.4	77.0
29.7015	70.6451	195.0	37.0	32.6	21.2	0.7	77.0
7.9735	18.9651	108.0	78.0	23.2	19.0	59.2	95.0
44.3196	105.4141	0.0	25.0	29.2	16.0	41.2	68.0
33.57057	79.8474	0.0	37.0	29.8	17.4	0.0	65.0

The behaviour of climate covariates was made fitting the Weibull AFT Regression model for the atmospheric and environmental failure groups, separately. The coefficients for the covariates are presented in Tab. 5 and 7. The values of the coefficients of the covariates can be interpreted multiplying them by the maximum and average values that the variables (weather events) can assume. In Tab. 4 this operation was made using 2000 and 90 for Atmospheric discharges (AD), 100 and 40 for wind speed (W), 40 and 25 for maximum temperature (Tx), 25 and 15 for minimum temperature (Tn), 80 and 10 for precipitation (P) and 100 and 60 for relative humidity (HR). These values are near the maximum and the mean values for the weather events considering all analysed days. This operation was done for the coefficients of the three presented models (all, atmospheric and environmental failures groups), and the resulting values are presented in "module".

Table 4. Product between the coefficient values for the three presented AFT Models and the maximum and mean values of the weather events.

	All		Atmospheric		Environmental	
	Max	Mean	Max	Mean	Max	Mean
AD	0.9800	0.0441	2.2800	0.1026	0.0800	0.0036
W	2.9950	1.1980	3.3560	1.3424	1.4820	0.5928
Tx	0.3324	0.20775	1.5528	0.9705	0.3504	0.2190
Tn	0.0755	0.04530	0.06170	0.0371	1.6323	0.9794
P	0.1392	0.0174	1.2496	0.1562	0.7256	0.0907
HR	0.3080	0.1848	0.4140	0.2484	0.4040	0.2424

The percentile and the expectation for these two groups are shown in Tab. 10 and 9, with real values for the weather events (the weather values differ from each other because they are from events associated with the failure group. So, the values of weather events presented for each table are from days that occurred a failure of the analysed group, either an atmospheric failure or an environmental failure). To maintain the same prediction logic, the weather events that presented a higher $p-value$ in the atmospheric and failure groups regression models were kept.

In Tab. 6 and 8, some statistical criteria for the covariates coefficients are presented. For the atmospheric group, atmospheric discharges, wind speed and precipitation were the covariates with a $p-value$ lower than 0.05. For the environmental failure events, minimum temperature, precipitation and relative humidity are the statistically significant covariates according to their $p-values$. It is possible to observe that with the confidence interval cross zero, the $p-value$ ends up returning a value above 0.05, that is not desirable. However, if the largest portion of the confidence interval focuses on a positive or negative value, it turns out to be odd to disregard the covariate just because of the value of $p-value$.

Table 5. Coefficient values for an Weibull AFT Regression for Atmospheric failure events in the analysed distribution system.

	Intercept	AD	W	Tx	Tn	P	HR
λ	7.18672	-0.00114	-0.03356	0.03882	-0.00247	-0.01562	-0.00414
ρ	-0.79611						

Table 6. P-values, standard errors (se) and confidence interval of the covariates of the Weibull AFT Regression for the Atmospheric failure events.

	p	se	lower 0.95	upper 0.95
AD	$< 5e^{-6}$	0.00025	-0.00163	-0.00065
W	$< 5e^{-6}$	0.00552	-0.04439	-0.02274
Tx	0.13892	0.02623	-0.01259	0.09023
Tn	0.93892	0.03228	-0.06574	0.06079
P	0.01499	0.00642	-0.02821	-0.00303
HR	0.15167	0.00289	-0.00981	0.00152

Table 7. Coefficient values for an Weibull AFT Regression for Environmental failure events in the analysed distribution system.

	Intercept	AD	W	Tx	Tn	P	HR
λ	8.36078	-0.00004	-0.01482	0.00876	-0.06529	-0.00907	-0.00404
ρ	-0.27167						

Table 8. P-values, standard errors (se) and confidence interval of the covariates of the Weibull AFT Regression for the Environmental failure events.

	p	se	lower 0.95	upper 0.95
AD	0.86452	0.00025	-0.00054	0.00045
W	0.00001	0.00341	-0.02150	-0.00814
Tx	0.51777	0.01355	-0.01779	0.03531
Tn	0.00005	0.01604	-0.09674	-0.03385
P	0.01541	0.00374	-0.01641	-0.00173
HR	0.01646	0.00168	-0.00733	-0.00074

The modelling of "ρ" for all covariates is not done due to the characteristics of the weather data, that are independent events (one climactic event value is not directly related with the value of another weather event).

Table 9. Prediction of the Percentile and the Expectation of the lifetime of the analysed distribution system for real values of the climatic events, considering Atmospheric failure events.

Percentile	Expectation	AD	W	Tx	Tn	P	HR
328.0820	1822.7876	195.0	37.0	32.6	21.2	0.7	77.0
20.1203	111.7864	108.0	78.0	23.2	19.0	59.2	95.0
299.9608	1666.5494	0.0	25.0	29.2	16.0	41.2	68.0
234.9093	1305.1307	165.0	40.0	26.9	19.4	0.0	88.0
126.1716	700.9959	138.0	44.0	25.4	19.0	27.8	93.0

Table 10. Prediction of the Percentile and the Expectation of the lifetime of the analysed distribution system for real values of the weather events, considering Environmental failure events.

Percentile	Expectation	AD	W	Tx	Tn	P	HR
162.4786	308.8923	108.0	78.0	23.2	19.0	59.2	95.0
433.7354	824.5855	0.0	25.0	29.2	16.0	41.2	68.0
489.8989	931.3593	0.0	37.0	29.8	17.4	0.0	65.0
407.7476	775.1792	6.0	44.0	23.7	17.8	0.0	65.0
374.5182	712.0061	165.0	40.0	26.9	19.4	0.0	88.0

The intercept value increased for both atmospheric and environmental groups, as expected. In terms of the covariates, in the atmospheric group, for all threats (coefficients with a negative value) the value increased considerably, but for the environmental group, some weather variables decreased (as wind speed and atmospheric discharges) and others presented a drastic increase (precipitation and minimum temperature). The prediction of the lifetimes has higher values when compared with the failure model for all groups, with extreme weather events impacting more in the lifetime considering the atmospheric group than in the environmental group.

From Tab. 4, considering that for the three models only Maximum Temperature (Tx) presented a positive coefficient, that has the meaning of a positive contribution for the lifetime of the system, it can be observed that Wind (W) is the most dangerous threat, due a product bigger than 1 for almost all situations. Wind also presented a higher contribution for the decrease of the lifetime, considering the "maximum" value on the atmospheric failure group model. Assuming a criterion that considering the product of the coefficient with the maximum value of the variable must be greater than 0.1, the minimum temperature for all failures and atmospheric failures group and atmospheric discharges for the environmental failure group presented a "null" impact in these situations.

4. CONCLUSION

Through Reliability Analysis Regression models, an evaluation of the lifetime of a real Brazilian DS due to different factors was made in this research. First, utilizing non-parametric (Kaplan-Meier and Nelson-Aalen) and parametric (Weibull) techniques, the curves for reliability and cumulative hazard function were presented, considering all registered failures from a period of almost two years. As both curves presented a similar behaviour, especially in the time period where most of the survival times of the analyzed DS are concentrated, the parametric model can be said to represent the data well.

The two techniques were applied for each of the different failure groups: Atmospheric, Environmental, Equipment Failure, Operational and Urban. It was observed that the threat level has a different order than the number of occurrences order for the failure groups, which ends up highlighting the group of atmospheric failures. However, for the context analyzed, Operational failures are the principal threat both in terms of the number of failures and the time frame in which they occur.

Using the Weibull AFT Regression Model, an analysis of the Atmospheric and Environmental failure groups, which are most linked with weather events were made using six weather events with their maximum daily values. From the model of the all failures groups, reliability function was plotted using different real values of the significant weather events, demonstrating the impact on survival time given extreme weather events. A discussion was made about the different coefficients values of the weather events covariates, from the All, Atmospheric and Environmental failure groups. For the atmospheric group, all coefficients of negative variables (threats) increased. For the Environmental failure group, some coefficients decreased, but other presented an increase, and specifically, Precipitation presented a great increase.

Predictions for the median (percentile) and the expectation of the lifetime of the system were made for all failures, atmospheric and environmental groups, considering some real values of the weather data. The principal difference was observed in the atmospheric group predictions, where the expected values can be much higher than the percentile values.

REFERENCES

Aalen, O. (1978). Nonparametric inference for a family of counting processes. *Ann. Statist.*, 6(4), 701–726. doi:10.1214/aos/1176344247. URL `https://doi.org/10.1214/aos/1176344247`.

Amrhein, V., Greenland, S., and McShane, B. (2019). Scientists rise up against statistical significance. *Nature*, 567(7748), 305–307. doi:10.1038/d41586-019-00857-9.

Bessani, M., Fanucchi, R., Achcar, J., and Maciel, C. (2016). A statistical analysis and modeling of repair data from a brazilian power distribution system. *Proceedings of International Conference on Harmonics and Quality of Power, ICHQP*, 2016-December, 473–477. doi:10.1109/ICHQP.2016.7783446.

Brem, S. (2015). Critical infrastructure protection from a national perspective. *European Journal of Risk Regulation*, 6(2).

Chowdhury, F., Gulshan, J., and Hossain, S. (2015). A comparison of semi-parametric and nonparametric methods for estimating mean time to event for randomly left censored data. *Journal of Modern Applied Statistical Methods*, 14(1), 196–207. doi:10.22237/jmasm/1430453760.

Colosimo, E. and Giolo, S. (2006). *Análise de sobrevivência aplicada.* Edgard Blücher.

Cox, D. and Oakes, D. (2018). *Analysis of survival data.* doi:10.1201/9781315137438.

Dantas, M., Valença, D., Platiny da Silva Freire, M., Medeiros, P., Da Silva, D., and Aloise, D. (2010). Weibull-regression models to study failure data in oil pumps. *Produção*, 20, 127–134.

Davidson-Pilon, C., Kalderstam, J., Zivich, P., Kuhn, B., Fiore-Gartland, A., Moneda, L., Gabriel, WIlson, D., Parij, A., Stark, K., Anton, S., Besson, L., Jona, Gadgil, H., Golland, D., Hussey, S., Kumar, R., Noorbakhsh, J., Klintberg, A., Ochoa, E., Albrecht, D., dhuynh, Medvinsky, D., Zgonjanin, D., Katz, D.S., Chen, D., Ahern, C., Fournier, C., Arturo, and Rendeiro, A.F. (2019). Camdavidsonpilon/lifelines: v0.22.3 (late). doi:10.5281/zenodo.3364087. URL `https://doi.org/10.5281/zenodo.3364087`.

Fogliatto, M.S.S., Santos, T.M.O., Bessani, M., and Maciel, C.D. (2019). Survival analysis of electrical power distribution systems using weibull regression. *Simpósio Brasileiro de Automação Inteligente.*

Kaplan, E. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53(282), 457–481. doi:10.1080/01621459.1958.10501452.

Konal, M., Öz, I., Uzunoğlu, C., and Kaçar, F. (2018). Electrical distribution network's failure analysis based on weather conditions. *2018 5th International Conference on Electrical and Electronics Engineering, ICEEE 2018*, 269–272. doi:10.1109/ICEEE2.2018.8391344.

Li, Q., Gao, J., and Flowers, G. (2019). Analysis of electromagnetic behaviors induced by contact failure in electrical connectors. *Microwave and Optical Technology Letters*, 61(11), 2579–2585. doi:10.1002/mop.31925.

Mishra, P., Pandey, C., Singh, U., Keshri, A., and Sabaretnam, M. (2019). Selection of appropriate statistical methods for data analysis. *Annals of Cardiac Anaesthesia*, 22(3), 297–301. doi:10.4103/aca.ACA_248_18.

Moolgavkar, S., Chang, E., Watson, H., and Lau, E. (2018). An assessment of the cox proportional hazards regression model for epidemiologic studies. *Risk Analysis*, 38(4), 777–794.

Murthy, D.P., Bulmer, M., and Eccleston, J.A. (2004). Weibull model selection for reliability modelling. *Reliability Engineering System Safety*, 86(3).

Ni, M. and Li, M. (2019). Reliability assessment of cyber physical power system considering communication failure in monitoring function. *2018 International Conference on Power System Technology, POWERCON 2018 - Proceedings*, 3010–3015. doi:10.1109/POWERCON.2018.8601964.

of Economic Advisers, P.C., the U.S. Department of Energy's Office of Electricity Delivery, and Reliability, E. (2014). *Economic benefits of increasing electric grid resilience to weather outages*, volume 2.

Pahwa, A. (2007). Modeling weather-related failures of overhead distribution lines. *2007 IEEE Power Engineering Society General Meeting, PES.* doi:10.1109/PES.2007.386167.

Radmer, D., Kuntz, P., Christie, R., Venkata, S., and Fletcher, R. (2002). Predicting vegetation-related failure rates for overhead distribution feeders. *IEEE Transactions on Power Delivery*, 17(4), 1170–1175. doi:10.1109/TPWRD.2002.804006.

Rink, M., Kluth, L., Shariat, S., Fisch, M., Dahlem, R., and Dahm, P. (2013). Kaplan-meier analysis in urological practice [kaplan-meier-analysen in der urologischen praxis]. *Urologe - Ausgabe A*, 52(6), 838–841. doi:

10.1007/s00120-013-3150-4.

Rodriguez-Garcia, L., Perez-Londono, S., and Mora-Florez, J. (2019). An optimization-based approach for load modelling dependent voltage stability analysis. *Electric Power Systems Research*, 177. doi:10.1016/j.epsr.2019.105960.

Sahai, S. and Pahwa, A. (2006). A probabilistic approach for animal-caused outages in overhead distribution systems. *2006 9th International Conference on Probabilistic Methods Applied to Power Systems, PMAPS*. doi:10.1109/PMAPS.2006.360321.

Shauly, M., Rabinowitz, G., Gilutz, H., and Parmet, Y. (2011). Combined survival analysis of cardiac patients by a cox ph model and a markov chain. *Lifetime Data Analysis*, 17(4), 496–513. doi:10.1007/s10985-011-9196-y.

Sroka, K. and Złotecka, D. (2019). The risk of large blackout failures in power systems. *Archives of Electrical Engineering*, 68(2), 411–426. doi:10.24425/aee.2019.128277.

Wang, B., Zhou, Z., Wang, H., Tu, X., and Feng, C. (2019). The p-value and model specification in statistics. *General Psychiatry*, 32(3). doi:10.1136/gpsych-2019-100081.

Zhang, Z. (2016). Parametric regression model for survival data: Weibull regression model as an example. *Annals of Translational Medicine*, 4(24).

Zio, E. (2009). Reliability engineering: Old problems and new challenges. *Reliability Engineering and System Safety*, 94(2), 125–141. doi:10.1016/j.ress.2008.06.002.

Złotecka, D. and Sroka, K. (2018). The characteristics and main causes of power system failures basing on the analysis of previous blackouts in the world. *2018 International Interdisciplinary PhD Workshop, IIPhDW 2018*, 257–262. doi:10.1109/IIPHDW.2018.8388369.