Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 October 2018 doi:10.20944/preprints201810.0438.v1

Peer-reviewed version available at *Energies* **2018**, *11*, 3034; doi:10.3390/en11113034

Assessing factors impacting on reliability of wind turbines by use of Survival Analysis – A case study

Samet Ozturk[1], Vasilis Fthenakis[1,*], Stefan Faulstich[2]
[1] Columbia University, 116th St & Broadway, New York, NY 10027
[2] Fraunhofer Institute for Energy Economics and Energy System Technology, Kassel, Germany
*corresponding author; email: vmf5@columbia.edu; tel. 212-854-8885

Abstract: Failure of wind turbines is a multi-faceted problem and its monetary impact is often unpredicted. In this study, we present a novel application of survival analysis on wind turbine reliability performance that includes accounting of previous failures and history of scheduled maintenance. We investigate the operational, climatic and geographical factors which affect wind turbine failures and model the risk rate of wind turbine failures based on data from 109 turbines in Germany operating during a period of 19 years. Our analysis showed that adequately scheduled maintenance can increase the survivorship of wind turbine systems and electric subsystems up to 2.8 and 3.8 times, respectively compared to the ones without scheduled maintenance. Geared-drive wind turbines and their electrical systems were observed to have 1.2- and 1.4-times higher survivorship, respectively, compared to direct-drive turbines and their electrical systems. It is also found that survivorship of frequently-failed wind turbine components, such as switches, is worse in geared-drive than in direct-drive wind turbines. We show that survival analysis is a useful tool for guiding the reduction of operating and maintenance costs of wind turbines.

1. Introduction

Wind energy is being deployed increasingly and its global capacity has doubled during the last six years [1]. Although availability of wind turbines reached 98% level in European wind farms [2], improvements on reliability are still needed. The problem is two-fold, one is high operations and maintenance (O&M) cost, the other is lost energy production. Wind turbines are monitored during scheduled maintenances and/or by condition monitoring systems to sustain uninterrupted energy production and avoid high O&M cost. Scheduled maintenances include visual inspection, non-destructive testing methods such as ultrasound and acoustic emissions, oil level test and vibration analysis while condition monitoring systems comprise of pressure, heat and vibration sensors [3, 4]. Wind farm operators need either to develop new techniques or decision support tools for their O&M strategies to reach the goal of maximizing energy production while minimizing O&M cost.

Failure of wind turbines is a multi-factor and often unpredicted problem that deserves further analysis [5]. Our aim is to investigate factors which have impact on wind turbine survivorship and inform actions that need to be taken to reduce the potential for failures. It is expected that proactive maintenance would decrease failure rates and the associated magnitude of consequences in wind turbines, but this was not quantified before. Furthermore, being prepared to respond to failures with readily available spare parts decreases downtime thus increases availability of wind turbines. Hence, spare parts management can be done in an optimum way and downtime can be reduced significantly by means of reliability predictions of wind turbines as it is applied to other industries [6].

In this study, we investigate the potential factors effecting wind turbine failures and model the hazard rate of wind turbine failures by use of survival analysis considering operational, climatic and geographical factors.

2. Literature review

There are several studies determining potential causes and failure mechanisms for wind turbine failures. Tavner et al. [7] investigated wind speed impact on wind turbine reliability and subassemblies of wind

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 October 2018 doi:10.20944/preprints201810.0438.v1

Peer-reviewed version available at Energies 2018, 11, 3034; doi:10.3390/en11113034

turbines in Denmark and concluded, by using cross-correlation analysis, that generator, yaw control, mechanical brake and hydraulic system are more prone to be affected by weather rather than other subassemblies Leite et al. [8] used Markov process to model wind speed characteristics, turbine failure and repair rates and types of turbine and evaluated the availability factor for Brazil. Hau et al. [9] discussed the main causes and mostly affected components in wind turbines and mentioned that proximity to the sea increases the possibility of corrosion and eventually a failure in a wind turbine. Faulstich et al. [10] investigated different factors considering their impacts on failure rates of wind turbines using a reliability ranking method. It was shown that wind turbines located close to the seawater and highlands, and with high wind speeds suffer high failure rates [10]. Fischer et al. [11] proposed reliability-centered maintenance for wind turbine components which were the main drivers of unavailability when components such as gearbox, generator, electrical system, and hydraulic system were failed. They concluded that vibration is the main cause for mechanical failures of those wind turbine components. Tavner [12] investigated the impact of weather conditions on off-shore wind turbine operations. He concluded that high wind speed, turbulence and wind gusts lower reliability of wind turbine blade, pitch and mechanical drive train, whereas temperature and humidity affected electrical components rather than mechanical. Wilson and McMillan [13] produced failure probabilities for offshore wind turbines by using onshore reliability data and offshore weather data and applying Markov chains and Monte Carlo simulation. It is concluded that temperature and humidity have lower impact than wind speed on offshore wind turbine failures [13]. Stafell and Green [14] used actual and theoretical load data for 282 wind farms in the UK and examined if the turbine age has an impact on failure rates. They concluded that aging increases either failure rates or downtime or both since there is a significant power reduction with age. Perez et al. [15] compared the failure rates and downtime values based on different turbine types and aspects and reported that direct-drive turbines have the highest sum of failure rates than geared-drive wind turbines. Reder et al. [16] proposed a framework to analyze Supervisory Control and Data Acquisition (SCADA) data using apriori rule mining and k-means clustering techniques and determined the effects of weather conditions on wind turbine failures. They found that winter is the season in which failure frequencies are increased whereas wind speed did not impact failure occurrences.

Slimacek and Lindqvist [17] analyzed reliability of wind turbines using a Poisson process and survival analysis which consider different factors such as type of turbine, size of turbine, harshness of environment, installation date and seasonal effects applying on WMEP database. They concluded that the turbine reliability is improving over the years and external factors such as lighting, icing and high wind increased the failure rate by 1.7 times. Mazidi et al. [18] proposed a hybrid methodology based on neural networks and a proportional hazard model (PHM) for maintenance management for wind turbines. They used SCADA data to develop a model using PHM to determine stress condition of wind turbines. However, they did not consider external factors due the data constraint. Carlos et al. [19] applied Monte Carlo simulations for maintenance optimization purposes using generic failure database and wind speed data from a Spanish database. They concluded that the optimum scheduled maintenance interval should be 113 days instead of a general industrial application of 180 days. Andrawus [20] proposed an optimal scheduled maintenance interval as 30 days for 26x600-kW wind farm, whereas Kerres et al. [21] found that corrective maintenance which is done upon a failure is a better option for a V44 – 600 kW turbine.

The cited studies investigate reliability of wind turbines and generally use average failure rate per turbine as a metric for reliability of wind turbines. However, in all these studies the survey period ended before wind turbines came to their end-of-life [7,8,10,11]; this type of data is called censored data. Therefore, survival analysis which account for the censorship is a better fit to evaluate reliability of wind turbines and assess the factors which affect the failures of wind turbines. Also, to the authors' knowledge, no study so far applied survival analysis to model the hazard rates to determine the variable impacts of factors such as wind speed, turbine age, distance to seawater and elevational location on wind turbine reliability although these variables are mentioned in different studies which are discussed previously.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 October 2018 doi:10.20944/preprints201810.0438.v1

Peer-reviewed version available at Energies 2018, 11, 3034; doi:10.3390/en11113034

Furthermore, we did not find any published studies quantifying hazard rates by use of survival analysis, based on scheduled maintenance and history of failures in wind turbines.

Survival analysis has been successfully used to determine the factors impacting on reliability of mechanical and civil infrastructural systems [6, 22]. The International Energy Agency (IEA) recommends applying survival analysis to investigate wind turbine reliability in order to develop optimum maintenance strategies [23]. Our study shows a first-time application of survival analysis including new potential-risk related variables such as number of previous failures and history of scheduled maintenance. The results can guide maintenance optimization and spare parts management.

3. Methodology

We investigate the survivorship of wind turbines, subsystems and components of a subsystem by a combination of three methodologies. Firstly, we apply survival analysis for wind turbines from a systems perspective using selected factors (e.g., climatic regions, elevational location, distance to coast, mean annual wind speed, turbine age, turbine type, number of previous failures and scheduled maintenance history) , secondly we investigate factors effecting only a critical subsystem, and lastly we apply survival analysis on frequently failed non-repairable components, to determine factors which impact survivorship of wind turbine components.

3.1. Survival Analysis

Survival analysis is a statistical technique to analyze time-to-event data. In this study, survival analysis is utilized to investigate time-to-failure of wind turbines as a system, a critical subsystem of wind turbines and components of a critical subsystem.

Survival analysis has the advantage of dealing with censorship over regular regression methods when there is no information regarding the exact time that failure occurred.

The survival function demonstrates the probability of a turbine survives beyond time t. The basic equations to define survival analysis are given in Equations 1, 2, 3, 4 and 5 [22]:

$$F(T) = \int_0^T f(x)dx \tag{1}$$

$$S(t) = \int_T^\infty f(x)dx = 1 - F(T) \tag{2}$$

$$S(t) = \exp[-\int_0^T h(x)dx] = \exp[-H(T)] \tag{3}$$

$$h(T) = f(T)/S(T) \tag{4}$$

$$H(T) = \int_T^\infty h(x)dx = -\ln[S(T)] \tag{5}$$

where T is the time to failure, f(x) is the probability density function of having a failure at time x, F (T) is the cumulative distribution function showing that a turbine survives until time T. Also, S(t) is the survival function which denotes the probability of survivorship beyond time T, h(t) is the hazard rate which represents the probability that a turbine at time T would have to fail during the next time interval, and H(T) is the cumulative hazard function.

We identify probability of failure of a wind turbine at a certain time by use of Kaplan-Meier estimator [24] and estimate cumulative hazard by a Nelson-Aalen estimator [25], while comparing survivorship of separate groups of wind turbines by applying statistical tests such as log-rank test [26], which will be explained in the next sections.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 October 2018 doi:10.20944/preprints201810.0438.v1

Peer-reviewed version available at Energies 2018, 11, 3034; doi:10.3390/en11113034

3.1.1. *Non-parametric Survival Analyses: Kaplan-Meier and Nelson-Aalen Estimators*

Kaplan-Meier estimator is a non-parametric method which does not make assumptions for any distribution. Equation 6 defines the Kaplan-Meier estimator [24]

$$\hat{S}(t) = \prod_{j:t_j \le t} \frac{n_j - d_j}{n_j} \tag{6}$$

where d_j is the number of individuals who has an event at time t_j where j=1, …, k and m_j is the number of individuals censored in the interval $[t_j, t_{j+1})$. Also, $n_j = (m_j+d_j) + \ldots + (m_k+d_k)$ is the number of individuals at risk just prior to t_j [24].

On the other hand, Nelson-Aalen estimator is a non-parametric method to estimate and plot cumulative hazard function [25]

$$\hat{H}_{NA}(t) = \sum_{t_i \le t} \frac{d_i}{n_i} \tag{7}$$

where d_i is the number of individuals who has an event at time t_i and n_i is the total individuals at risk at time t_i.

3.1.2. *Log-Rank Test for significance of survivorship*

The log-rank test is used to test the null hypothesis that there is no statistically significant difference between two groups in the probability of an event. The test statistic is the sum of $(O - E)^2/E$ for each group where O is observed, and E is expected number of events. Obtained test statistic value is checked in Chi-distribution table and the corresponding p-value represents the probability of having the event by chance [23].

3.1.3. *Semi-parametric Survival Analysis: Cox Proportional Hazards Model*

Cox Proportional Hazards Model (PHM) which is also known as Cox model include a parametric baseline hazard function along with a non-parametric hazard ratio. Cox PHM is [22]

$$h(t, z, \alpha) = h_0(t)e^{zB} \tag{8}$$

where $h_0(t)$ is the baseline function, z is the variable and B is the hazard coefficient for the variable. The hazard ratio between two groups (z_1 and z_2) in a factor can be estimated using Equation (9).

$$HR(t, z_1, z_0) = e^{B(z_1 - z_2)} \tag{9}$$

In this study the main interests are to determine the differences of survivorship of wind turbines based on selected factors by using Kaplan-Meier and Nelson-Aalen estimators, and to estimate hazard ratios of the factors that impact wind turbine failures by applying Cox regression. Calculations and plots for Kaplan-Meier, Nelson-Aalen estimations and Cox Regression are obtained by using statistical software SPSS V.25 [27]. Proportional hazard assumptions are checked graphically by log-minus-log (survival) against survival times graphs. Cox PHM is applied until only significant factors remain with p-values less than 0.05. Results from Cox regression will be given in tables consisting following parameters of Cox regression:

- Standard Error (SE): SE of the estimate shows the accuracy of estimation for the observed value.
- Wald statistic: Wald statistic is the ratio of regression coefficient B to SE. It is used to evaluate significance of B coefficients of factors.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 October 2018 doi:10.20944/preprints201810.0438.v1

Peer-reviewed version available at Energies 2018, 11, 3034; doi:10.3390/en11113034

- Degrees of freedom: df represents number of sub-factors which are compared against in a factor. For example, design type has two sub factors such as direct and geared and df is 2-1 = 1.
- Significance level (sig.): The probability of having the coefficient by chance for a specific factor.
- Exp (B): Hazard ratio from Cox regression is given as Exp (B).
- 95% Confidence Intervals (CI): 95% Upper and lower levels of coefficients resulting from the regression.

4. Case study based on WMEP data

In this study the application of survival analysis to determine factors for wind turbine reliability is demonstrated by a case study on wind turbines in Germany. The survival analysis involves the investigation of wind turbine failures recorded in the WMEP data-base which covers wind turbines operated in Germany between 1989 and 2008. The events in the WMEP database include scheduled maintenance, scheduled maintenance with replacement or repair, and unscheduled maintenance with a replacement or repair. The WMEP survey collected O&M data from more than 1500 wind turbines, in this study data from 575 of them are ready to be utilized with 6188 turbine years of operation and including 19,242 events considering a repair or replacement.

A participant turbine in this study is defined according to the chosen methodological approach; these are shown in Figure 1. According to the systems and subsystems approach, a participant is a wind turbine with a time interval from either the commissioning date or a start date of a failure to either another start of a failure or end of a survey. In some cases, in the failure data of systems and subsystems, there were unscheduled maintenances which started before the previous failure was resolved. Therefore, for application of survival analysis on non-repairable components, a participant is defined as a wind turbine with time interval between either the commissioning or a replacement of a component and either start of a failure in a component or end of a survey. Wind turbine participant types 1, 2, 3 and 4 shown in Fig. 1 are considered for systems and subsystems approach and participant types 1, 3, 5 and 6 are used for the component approach in which censored values are included in participant types 3, 4 and 6. It must be noted that unscheduled maintenance and scheduled maintenance with a repair or replacement, which are regarded as failure in the WMEP database, are combined for survival analysis.. Also, it is noted that no distinction is been made between initial and repeated participants in the analysis.

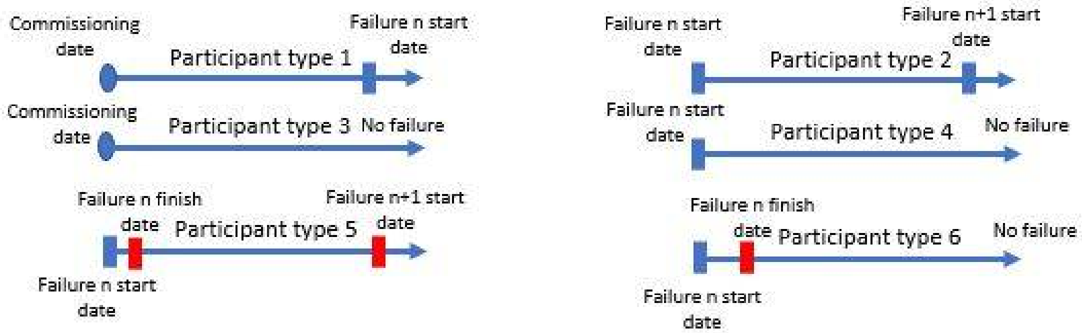


Fig. 1. Types of participants considered in this study for survival analysis

4.1. Survival Analysis Factors

Table 1 lists factors which are considered to have a potential to affect the reliability of wind turbines. The details of the selected factors are given in the subsections that follow. Some of the factors which may for impact the survivorship of wind turbines such as production year, turbulence intensity, presence of an

Peer-reviewed version available at Energies 2018, 11, 3034; doi:10.3390/en11113034

online CMS, distance to service station, grid quality, energy yield and number of wind turbines considered in the analysis, were not taken into account due to data limitations.

Table 1. Factors considered in survival analysis

Geographical and environmental factors	Operational factors
1. Koppen-Geiger Climatic Regions: Cfb, Dfb, Dfc	5. Turbine age (years): 0-3, 4-14
2. Elevational location: High land (>100 m), Low land (<=100 m)	6. Turbine type: Geared-drive, Direct-drive
3. Distance to coast: Coastal (0-20 km), Inland (>20 km)	7. Number of previous failures (NOPF) *Varies*
4. Mean annual wind speed (MAWS) High (>6.25 m/s), Low (<=6.25 m/s)	8. Scheduled maintenance history: Yes, No

4.2.1. Koppen-Geiger climatic regions

Koppen-Geiger is a climate classification that is being cited by about 5,000 studies in variety of disciplines [28]. The Koppen-Geiger climatic regions are determined based on annual precipitation and temperature records along with seasonal temperature records; they are based on 12,396 precipitation and 4,844 temperature data stations world-wide and employ various temperature and precipitation criteria [28]. In Germany, there are four Koppen-Geiger climatic regions as it is seen in Figure 2; these are:

- Cfa: Temperate-without dry season-hot summer
- Cfb: Temperate-without dry season-warm summer
- Dfb: Cold-without dry season-warm summer
- Dfc: Cold-without dry season-cold summer

Criteria for the classification of the climatic regions of interest are given in Table 2 [28]. First criterion which is denoted by a capital letter e.g. C or D, is a climate classification which is evaluated based on the average temperature of the hottest and coldest months. Second criterion is a classification based on annual precipitation level. Last criterion is based on the summer or winter temperature records.

Table 2. Criteria for the climatic region classifications for Germany

1st	2nd	3rd	Description	Criteria*
C			Temperate	$T_{hot} \geq 10$ & $0 < T_{cold} < 18$
	s		- Dry Summer	$P_{sdry} < 40$ & $P_{sdry} < P_{wwet}/3$
	w		- Dry Winter	$P_{wdry} < P_{swet}/10$
	f		- Without dry season	Not (Cs) or (Cw)
		a	- Hot Summer	$T_{hot} \geq 22$
		b	- Warm Summer	Not (a) & $T_{mon10} \geq 4$
		c	- Cold Summer	Not (a or b) & $1 \leq T_{mon10} < 4$
D			Cold	$T_{hot} \geq 10$ & $T_{cold} \leq 0$
	s		- Dry Summer	$P_{sdry} < 40$ & $P_{sdry} < P_{wwet}/3$
	w		- Dry Winter	$P_{wdry} < P_{swet}/10$
	f		- Without dry season	Not (Ds) or (Dw)
		a	- Hot Summer	$T_{hot} \geq 22$

Peer-reviewed version available at Energies 2018, 11, 3034; doi:10.3390/en11113034

b	- Warm Summer	Not (a) & $T_{mon10} \geq 4$
c	- Cold Summer	Not (a, b or d)
d	- Very Cold Winter	Not (a or b) & $T_{cold} < -38$

*T_{hot} = temperature of the hottest month, T_{cold} = temperature of the coldest month, T_{mon10} = number of months where the temperature is above 10, P_{dry} = precipitation of the driest month, P_{sdry} = precipitation of the driest month in summer, P_{wdry} = precipitation of the driest month in winter, P_{swet} = precipitation of the wettest month in summer, P_{wwet} = precipitation of the wettest month in winter.

Figure 2 shows the wind turbine locations in Germany in the WMEP database that we use in this study. There are 427 wind turbines – 4526 turbine years in Cfb region, 122 wind turbines – 1346 turbine years in Dfb region, 25 wind turbines – 306 turbine years in Dfc region in the WMEP database.

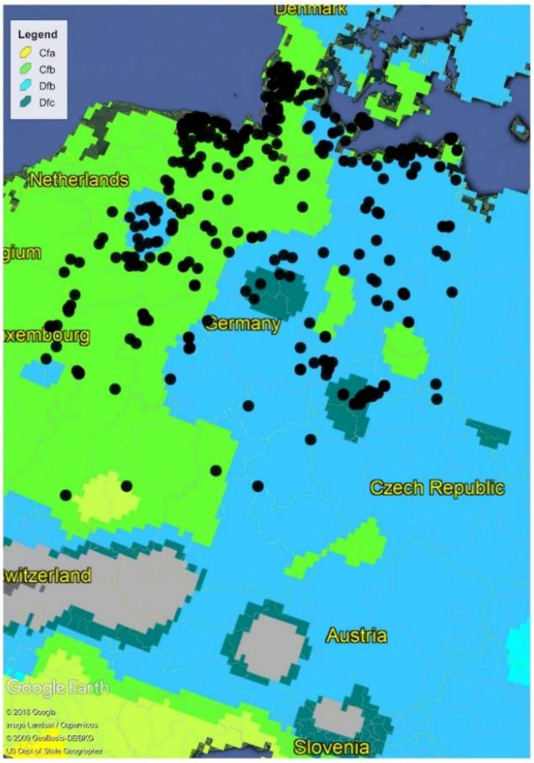


Fig. 2. Map of 575 wind turbines (black dots) in different climatic regions in Germany

4.2.2. Elevational location

The elevational locations where wind turbines are operated are divided in two categories, as low land (<=100 m) and high land (>100 m).

4.2.3. Distance to coast

Wind turbines are also divided in two categories based on their proximity to seawater. Turbines which have distance to coast lesser than or equal to 20 kms are called as coastal, the rest of the turbines are called as inland.

4.2.4. Mean annual wind speed (MAWS)

Mean annual wind speeds at 50 m height for wind turbine locations are gathered from Global Wind Atlas for every event [29]. MAWS values are divided in two categories as "Low" (lower than 6.25 m/s) and "High" (greater than or equal to 6.25 m/s).

Peer-reviewed version available at *Energies* **2018**, *11*, 3034; doi:10.3390/en11113034

4.2.5. Turbine age

Turbine age categorization is based on the operational years at the initial date of the participant. For example, if the turbine commissioning date is the participant start date, then the age of that turbine is considered as 0, if a participant start date is 370 days after the commissioning date of the turbine then the age of the turbine is considered as 1. 0-3 years age category is considered as infant and 4-14 years age category is considered as mature.

4.2.6. Turbine type

Participants are categorized based on their associated wind turbine design types as geared and direct-drive. Table 3 lists the number of participants associated with the design types.

4.2.7. Number of previous failures (NOPF)

NOPF for a participant considers the number of previous failures occurred for a turbine or a subsystem. depending on the methodological approach. In a wind turbine full system approach any error is accounted for NOPF, while in a subsystem approach only failures in regarding subsystem are considered for NOPF. Furthermore, the category intervals vary with the approach. For wind turbine systems approach, categorization is done into five as 0-10, 11-20, 21-30, 31-40, 41+. For electrical subsystem it is done into four as 0-2, 3-5, 6-8, 9+, since less data points are involved in the subsystem investigation than complete systems approach whereas for component approach, categorization is done into two as 0-1 and 2-more.

4.2.8. Scheduled maintenance history

The history of scheduled maintenance categorization in Germany is formed considering the presence of any reported scheduled maintenance without a repair or replacement during the survey period of a participant. Turbines in the WMEP database have routinely scheduled maintenance as certain measures in wind turbines must follow specific industry standards, such as the IEC 61400-1. However, due to lack of information in the WMEP database about completing a scheduled maintenance, the scheduled maintenance historical record must be taken as reporting of scheduled maintenances rather than carrying out them. The IEC 61400-1 standards lead the designers through the whole life-cycle of a turbine, process from design via operation and maintenance (O&M) to decommissioning. Regarding O&M the designer has to establish all requirements on how to handle special wear parts, safety related components, greasing, and so on, and also how and when to provide service. A typical period for recurrent inspections has up to now been three to six months, while currently designers tend to prolong the period to twelve months at least for offshore wind turbines. The operator's manual has to state all these requirements and will be part of the documents given to a certification body for proving the design assumptions and calculations and all accompanying documents. All requirements in technical documents and certificates will then become part of the mandatory preconditions when the government in charge issues the building and operation permit. Furthermore, scheduled maintenance history only considered as a factor for system and subsystem approaches. Due to overwhelming number of participants which have scheduled maintenance history comparing the ones with no scheduled maintenance history between two failures of components, scheduled maintenance is not considered as a factor in the component approach of our study.

Table 3 shows sample participants with their associated factors used for component approach in this study.

Table 3. Sample data used for survival analysis of switch component

Peer-reviewed version available at *Energies* **2018**, *11*, 3034; doi:10.3390/en11113034

Turbine Model	Time to failure (days)	Status	Design type	Climatic regions	Turbine Age	Distance to coast	Elevational location	MAWS
Model A	675	Failed	Geared	Cfb	0-3	Coastal	Low	H
Model A	2978	Censored	Geared	Cfb	0-3	Coastal	Low	H
Model A	1572	Failed	Geared	Cfb	0-3	Inland	Low	L
Model B	3849	Censored	Direct	Dfb	0-3	Coastal	Low	L

4.3. Selected turbine aspects

Table 4 shows number of participants based on the methodological approaches discussed previously. The most represented direct-drive and geared-drive turbine models which have 500 kW power production capacity is selected. For the subsystem approach electrical subsystem is investigated since it is stated to be the most frequently failed subsystem in different studies [2, 10, 30]. There are 39 geared-drive turbines adding up to 432 operational-years and 70 direct-drive turbines with a total of 733 operational-years. Furthermore, fuses and switches, which are the two highest frequently failed components in electrical subsystem in the WMEP database, are considered for the component survivorship investigation in this study.

Table 4. Wind turbine participant characteristics considered in this study

Characteristics	Wind turbine system study		Electrical subsystem study		Component study			
					Fuses		Switches	
Turbine Type	Geared	Direct	Geared	Direct	Geared	Direct	Geared	Direct
Number of participants	1477	3334	269	704	47	123	157	126

5. Results

5.1. Wind turbine system approach

We used Nelson-Aalen cumulative hazard plots to visualize the differences in factors since the Kaplan-Meier survival probability graphs do not distinguish differences well when many data points are involved. The most significant distinctions in the cumulative hazard functions in Fig. 3 are noted in the graphs of design type and history of scheduled maintenance in Fig. 3.a and 3.d, respectively, where there is no significant time-dependency in the distinction. On the other hand, impacts of climatic regions, turbine age, distance to coast, elevational location, NOPF and MAWS, seem to be time-dependent (as shown in Fig. 3.b, 3.c, 3.e, 3.f, 3.g and 3.h) and thus no definitive conclusions can be drawn related to these factors due to violation of Cox regression proportionality assumption. Table 5 summarizes the results from log-rank tests where one can see that numerical results support the graphical representation of design type and history of scheduled maintenance impact on survivorship of wind turbines. From these results, it is inferred that direct-drive wind turbines are significantly more prone to failures than geared-drive wind turbines. A wind turbine which has a history of scheduled maintenance has significantly higher survivorship than a wind turbine with no scheduled maintenance history as can be seen in Fig. 3.d. Moreover, Cox Regression modelling shows that direct-drive turbines are found to be 25% riskier than geared-drive wind turbines while having scheduled maintenance makes a wind turbine 2.8 times safer as it can be seen in Table 6. Lower survivorship of the direct-drive wind turbine can be attributed to its less maturity comparing to the geared-drive wind turbine. Higher risk for survivorship of wind turbines with no scheduled maintenance history is expected since scheduled maintenance are done to improve the reliability of wind turbines. On the other hand, climatic regions seem to not follow any pattern in Fig. 3.b for cumulative hazards whereas early turbine ages seem to have higher hazard rate which is diverging in time in Fig. 3.c. It can be seen in Fig. 3.e that coastal turbines have lower hazard rate than inland turbines

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 October 2018 doi:10.20944/preprints201810.0438.v1

Peer-reviewed version available at Energies 2018, 11, 3034; doi:10.3390/en11113034

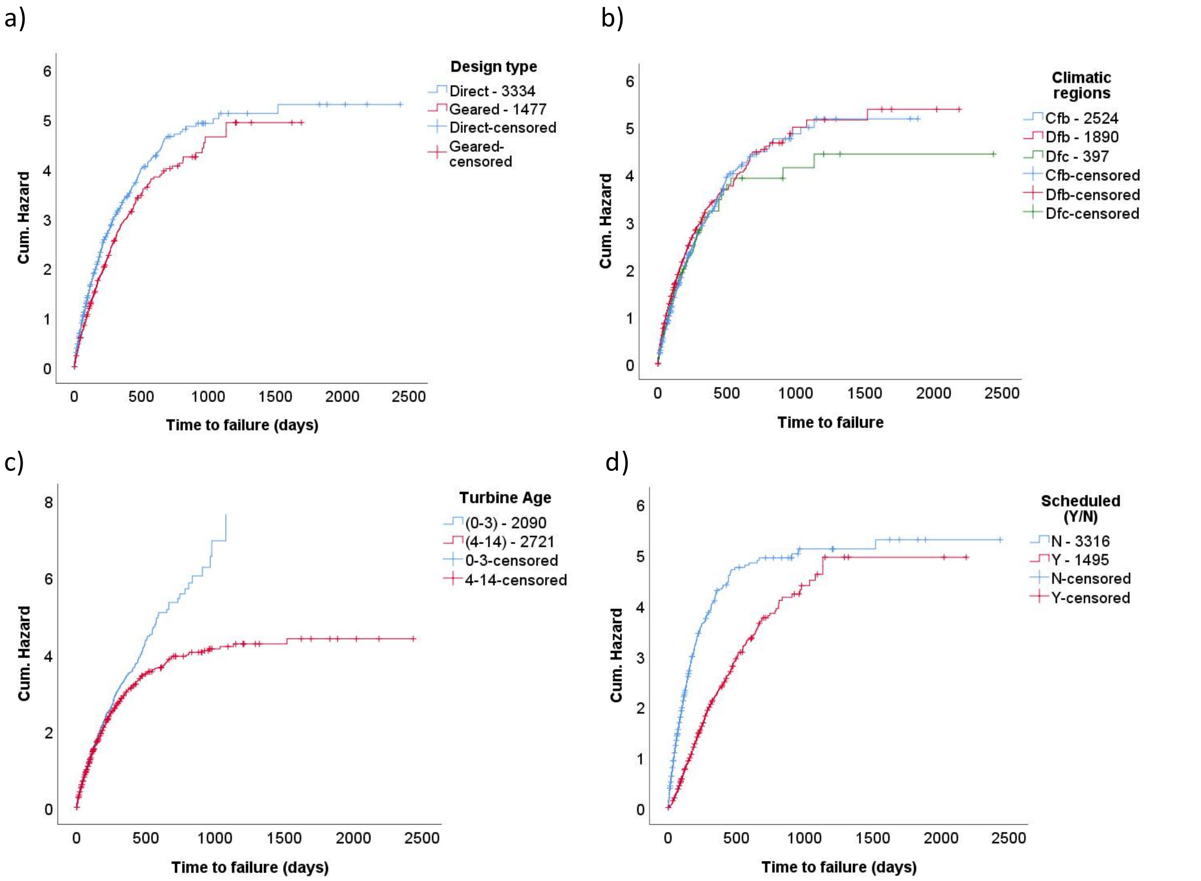
around the first 500 days of operation and it is reversed after that time while wind turbines in high elevations have higher hazard rate than lower elevations around the first 700 days as it is shown in Fig. 3.f and it is reversed after that time. It can be inferred from Figure 3.g that the smaller number of previous failures the high likely that a turbine to fail again after around 500 days. Also, as observed from Figure 3.h, MAWS is not a definitive parameter to conclude that wind speed affects the survivorship of a wind turbine as a system at least for the data considered.

Table 5. Log-rank test results for comparison of factors on wind turbine system failures

Factors	Groups	Test statistics	
		Chi-Square	Sig.
Design type	Direct vs. Geared	53.01	0.000
Scheduled maintenance history	No vs. Yes	991.01	0.000

Table 6. Cox regression results for the factors which satisfy proportionality assumption

							95.0% CI for Exp(B)	
Factors	B	SE	Wald	df	Sig.	Exp(B)	Lower	Upper
Design type	0.22	0.035	39.3	1	0.000	1.25	1.16	1.34
Scheduled (Y/N)	1.02	0.033	935.2	1	0.000	2.77	2.59	2.95



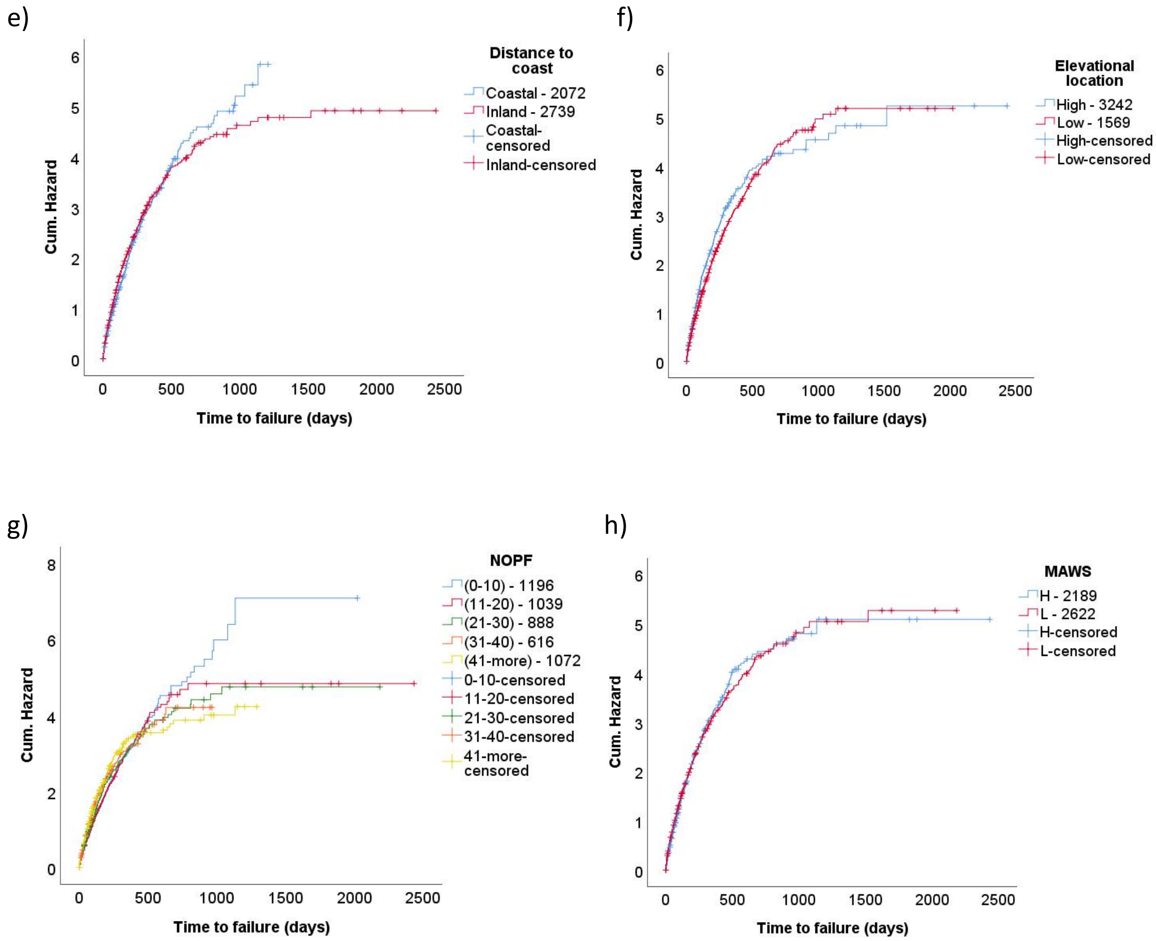


Fig. 3. Nelson-Aalen cumulative hazard functions of wind turbines based on the operational, climatic and geographical factors

5.2. Survival analysis for electrical subsystem

Survivorship of electrical subsystems based on the operational, climatic and geographical factors are depicted in Figure 4; notable differences are observed in the history of scheduled maintenance, wind turbine design type, distance to coast and elevational location of the wind turbines. Design type and history of scheduled maintenance have consistent distinction independent of time as shown in Fig. 4.a and 4.d, respectively, while climatic region, turbine age, distance to coast, elevational location and MAWS are time-dependent factors as it is shown in Fig 4.b, 4c, 4.e, 4.f, 4.g and 4.h, respectively. Table 7 summarizes the log-rank test results which indicate that there is a significant difference between participants with geared-drive and direct-drive wind turbines as well as participants with no prior scheduled maintenance and turbines with having prior scheduled maintenance. Furthermore, Table 8 gives the Cox regression results which show that electrical systems in direct-drive wind turbines have 42% higher risk of having failure than the ones in geared-drive wind turbines while electrical systems with no scheduled maintenance history have the risk of failure 3.8 times more than the electrical systems with a scheduled maintenance history in a wind turbine. It can be inferred from Figure 4.c that the impact of turbine age on survivorship of electrical systems varies with time. Electrical systems in (0-3) years-old wind turbines have higher survivorship around their first 1460 days of operation, counting from the participant entry date) than the turbines at (4-14) years old, then it is reversed. Although electrical systems in wind turbines at coastal and high elevational locations seem to have lower survivorship as can be seen in Figure

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 October 2018 doi:10.20944/preprints201810.0438.v1

Peer-reviewed version available at *Energies* **2018**, *11*, 3034; doi:10.3390/en11113034

4.e and 4.f; the Cox regression results tabulated in Table 8 indicate that these two parameters are not significant. NOPF and MAWS did not satisfy proportionality criterion for Cox regression since the survivorship of electrical systems based on these factors vary with time. Nevertheless, one can see from Figure 4.g and 4.h that a smaller number of previous failure and high MAWS increase the survivorship of electrical systems.

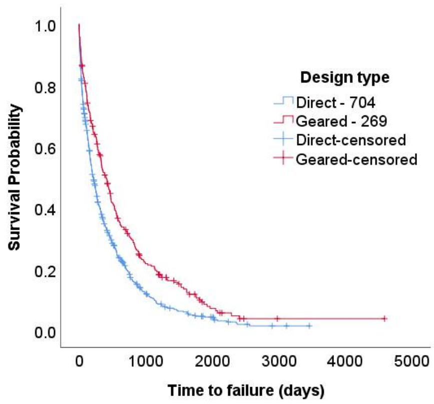
Table 7. Log-rank test results for comparison of factors on electrical subsystem failures

Factors	Groups	Test statistics	
		Chi-Square	Sig.
Design type	Direct vs. Geared	22.77	0.000
Scheduled maintenance history	No vs. Yes	351.76	0.000

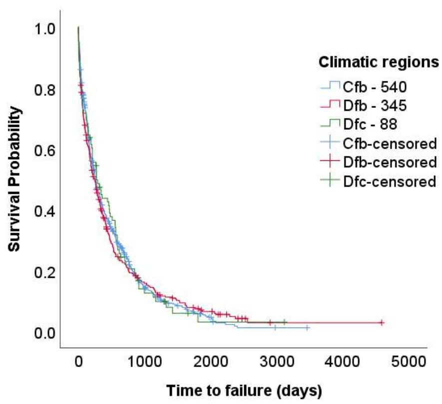
Table 8. Cox regression results for electrical system failures

	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B)	
							Lower	Upper
Design type	0.35	0.080	18.99	1	0.000	1.42	1.21	1.66
Scheduled (Y/N)	1.34	0.075	319.29	1	0.000	3.81	3.30	4.42
Elevational location	0.06	0.077	0.68	1	**0.408**	1.07	0.916	1.24
Distance to coast	0.04	0.087	0.20	1	**0.652**	1.04	0.88	1.23

a)



b)



c)

d)

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 October 2018 doi:10.20944/preprints201810.0438.v1

Peer-reviewed version available at Energies 2018, 11, 3034; doi:10.3390/en11113034

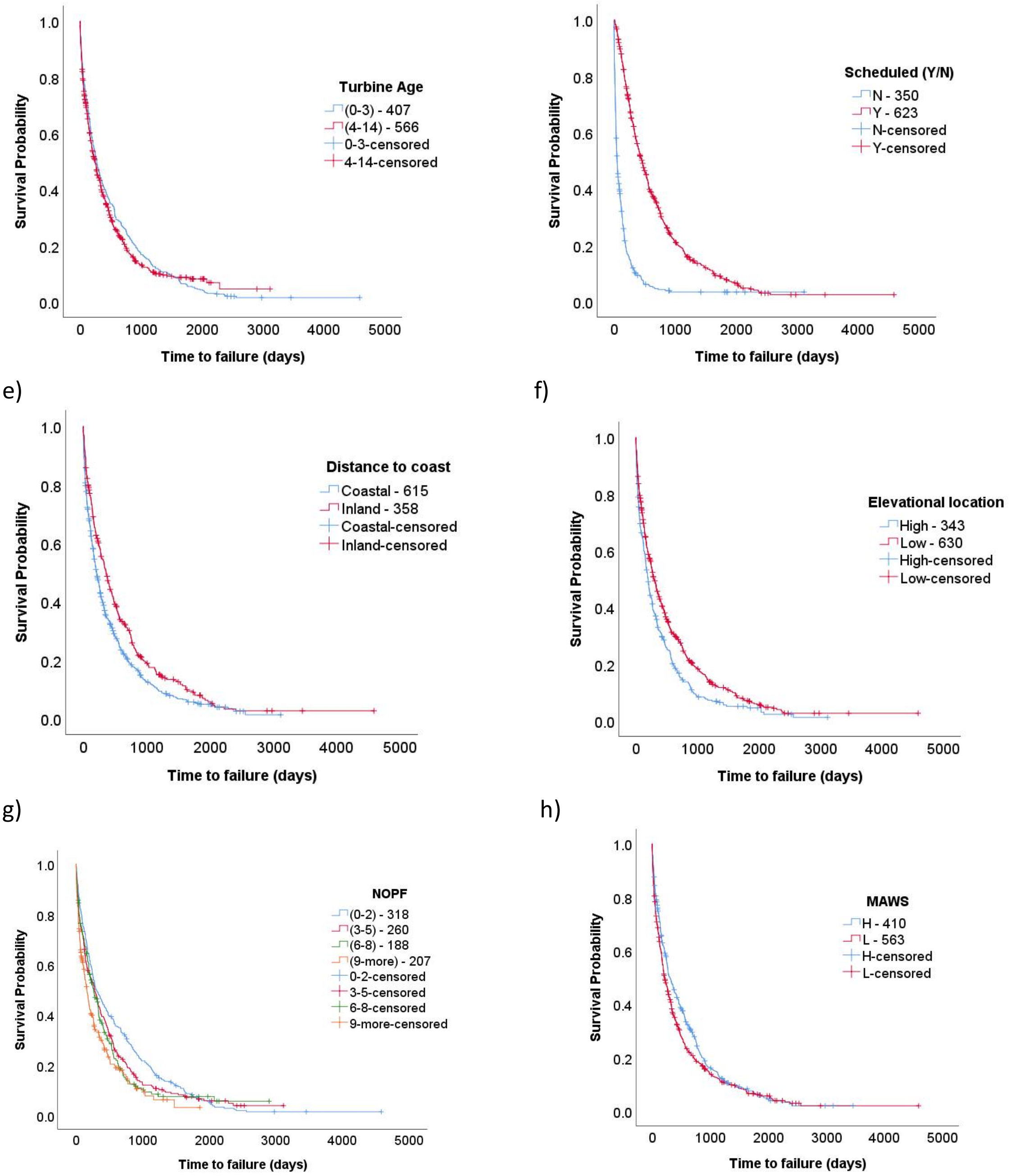


Fig. 4. Kaplan-Meier survival functions of electric subsystem based on the operational, climatic and geographical factors

5.3. Survival analysis for components of electrical subsystems

Fuses and switches are selected for survival analysis in electrical subsystems of wind turbines since both components have the highest frequency in number of failures as well as being non-repairable components. Fuses provide overcurrent protection, while switches start and stop electrical circuits in a wind turbine [31, 32]. They are both discarded after a malfunction since they lose their functionality. Therefore, it is assumed that NOPF and scheduled maintenance would not have an impact on survivorship of fuses and switches for which these factors are not included in the survival analysis.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 October 2018 doi:10.20944/preprints201810.0438.v1

Peer-reviewed version available at Energies 2018, 11, 3034; doi:10.3390/en11113034

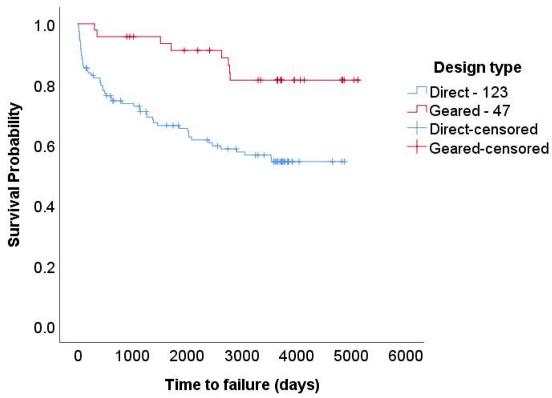
5.3.1. Survival analysis for fuses

Figure 5.a and 5.c depict the dependence of survivorship on the design type and age of wind turbines. Cox regression results show in Table 9 that design type and turbine age are significant factors impacting fuse failures in wind turbines. Fuses in direct-drive wind turbines have 3 times higher risk comparing to fuses in geared-drive wind turbines. Since fuse failures occur because of overcurrent, the direct-drive wind turbines might be claimed to have overcurrent problems more often than geared-drive design ones. Furthermore, fuses in wind turbines at 4-14 years of operation prone to 60% more failure risk than the turbines in their first three years of operation as it can be inferred from Table 9. On the other hand, as shown in Figure 5.b, climatic regions do not seem to affect fuse survivorship while distance to coast, elevational location and MAWS do show some difference, in 5.d, 5.e and 5.f, respectively. However, p-values which are for the representation of significance in Cox regression showed that climatic regions, distance to coast, MAWS, elevational location and NOPF are not significant factors with p-values of 0.883, 0.820, 0.802, 0.479 and 0.173 respectively. Thus, these factors are not included in our hazard rate modeling. The proportionality assumption is verified with the tests for design type and turbine age factors.

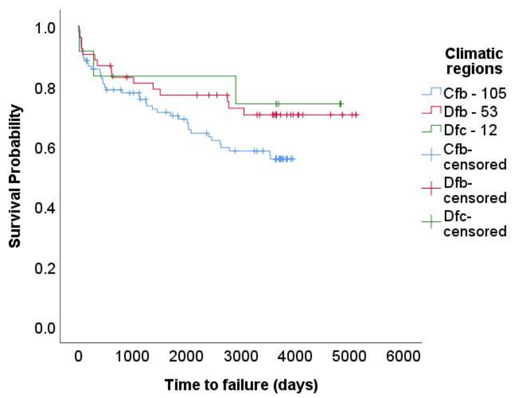
Table 9. Cox regression results for fuses

							95.0% CI for Exp(B)	
	B	SE	Wald	df	Sig.	Exp(B)	Lower	Upper
Design type	1.13	0.381	8.73	1	0.003	3.09	1.46	6.52
Turbine Age	-0.95	0.329	8.35	1	0.004	0.39	0.203	0.74

a)



b)



c)

d)

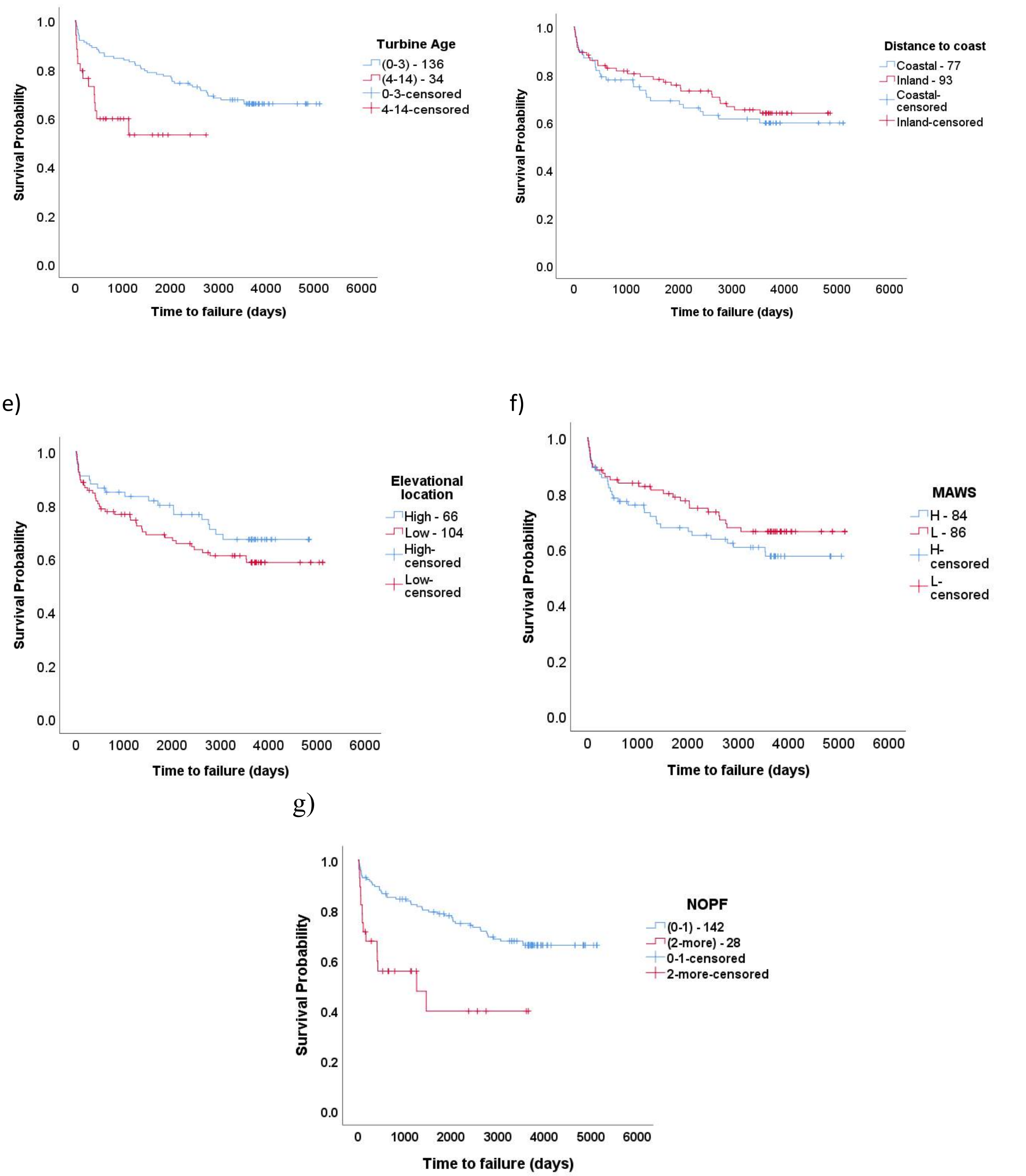


Fig. 5. Kaplan-Meier survival functions for fuses based on the operational, climatic and geographical factors

The hazard rate for survivorship of fuses can be written as in the following:
HR = exp[(1.13 *direct*) + (-0.95 *early age*)]

5.3.2. Survival analysis for switches

The Kaplan-Meier survival functions of switches in wind turbines are depicted in Fig 6. Obvious distinctions are observed in design type, distance to coast, elevational location and MAWS factors as shown in Fig.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 October 2018 doi:10.20944/preprints201810.0438.v1

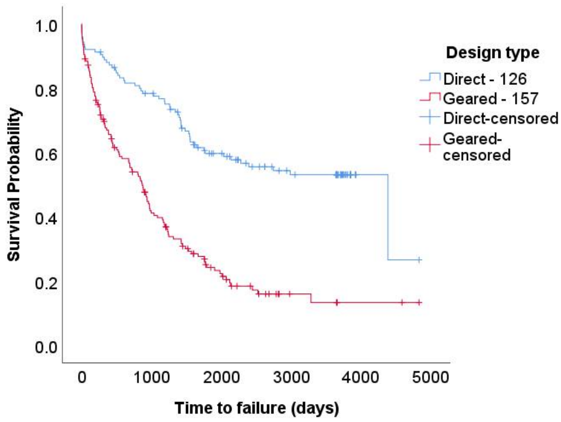
Peer-reviewed version available at *Energies* **2018**, *11*, 3034; doi:10.3390/en11113034

6.a, 6.d, 6.e and 6.f, respectively. However, Cox regression results show that design type, distance to coast and MAWS are the factors which have significant impacts on failure of switches in wind turbines in Table 10. Switches in direct-drive wind turbines have 66% increase of survivorship comparing to the ones in geared-drive wind turbines. This may be explained by differences in material quality or drive design between two different turbine models. The survival functions of switches in wind turbines based on their distance to coast is demonstrated in Fig 6.d. Switches in wind turbines which are within 20 km from the coast have a 39% increase of survivorship comparing to the ones located farther from the coast. This can be attributed to more consistent wind speed regimes in coastal regions compared to the inland regions. In the coastal regions switches deal with less start-stop cycles which improve the survivorship of them. Similarly, turbines located in the regions with high MAWS have 35% increased survivorship as it is seen in Table 10. The reason can be attributed to more consistent spin of wind turbine that reduces the failure rates of switches. Conversely, turbine age, elevational location, NOPF and climatic regions are determined as not significant with 0.883, 0.554, 0.520 and 0.088 p-values, respectively and are not included in our hazard rate model. Proportionality assumption is tested, and it is observed that it is valid with significant factors of p-values less than 0.05.

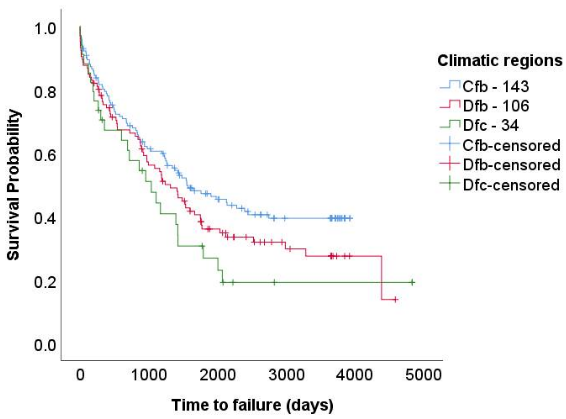
Table 10. Cox regression results for switches

	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B)	
							Lower	Upper
Design type	-1.08	0.167	41.48	1	0.000	0.34	0.25	0.47
Distance to coast	-0.50	0.177	8.04	1	0.005	0.61	0.43	0.86
MAWS	-.43	0.178	5.89	1	0.015	0.65	0.46	0.92

a)



b)

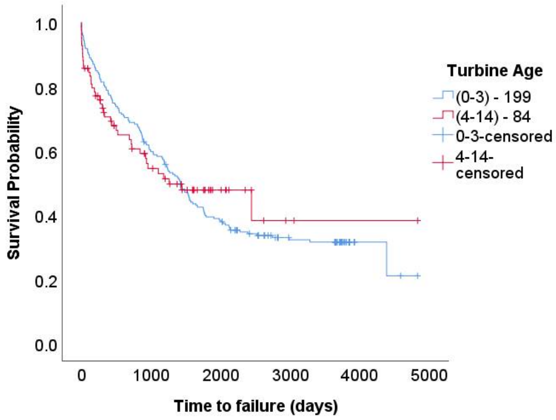


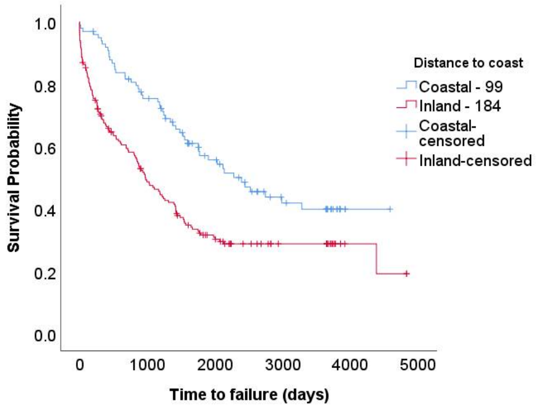
c)

d)

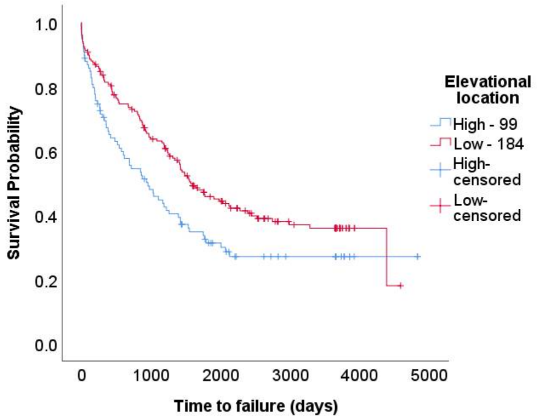
Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 19 October 2018 doi:10.20944/preprints201810.0438.v1

Peer-reviewed version available at Energies 2018, 11, 3034; doi:10.3390/en11113034

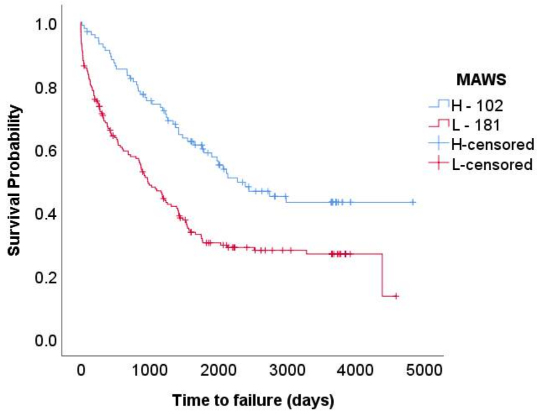




e)



f)



g)

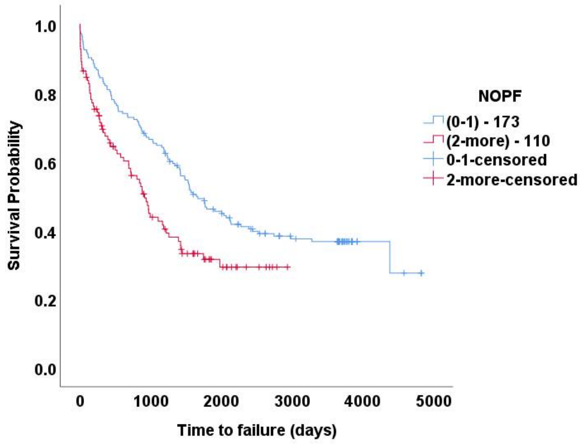


Fig. 6. Kaplan-Meier survival functions of switches based on the operational, climatic and geographical factors

The hazard rate for survivorship of switches can be written as in the following:
HR = exp[(-1.08 *direct*) + (-0.50 *coastal*)+ (-0.43 *high MAWS*)]

6. Conclusions

Peer-reviewed version available at Energies 2018, 11, 3034; doi:10.3390/en11113034

Survival analysis is used considering two novel indicators of survivorship of wind turbines; these are the number of previous failures and the history of scheduled maintenance. We identified several risk factors having definitive impacts on the reliability of wind turbines, their electrical subsystems and components of the electrical subsystem by use of Survival Analysis. Our results are summarized below:

- Geared-drive wind turbines and their electrical systems are observed to have 1.3- and 1.4-times higher survivorship, respectively, comparing to direct-drive wind turbines and their electrical systems. This distinction in survivorship is true for fuses as well while switches show the exact opposite trend where switches in direct-drive turbines are less likely to fail comparing to switches in geared-drive wind turbines. Gearing-type of wind turbines might improve the survivorship of a certain component while reducing the survivorship of another component. For example, the survivorship of fuses was two times higher in direct-drive turbines than in geared-drive turbines, whereas the survivorship of switches was reduced by 66%.
- Although the survival probability graphs show some differences between climatic regions, those are not significant to the survivorship of wind turbines, electrical subsystem and components of electrical systems. However, such significance is related to number of data points and relative number of data points among the factors as well as investigated subsystem and components. The lack of expected significance on climatic regions for this study may be attributed to data scarcity, specifically for the Dfc region with cold climate in the summer.
- Impact of turbine age on survivorship of turbine systems and electrical subsystems vary with time. However, fuses have 60% lower survivorship in their “mature” 4-14 ages than in their early years.
- Scheduled maintenance reporting significantly improves the survivorship of wind turbines; our data and analysis show 2.8- and 3.8-times improvement in survival for wind turbine as a system and electrical subsystems, respectively. In other words, at an instantaneous time there is 2.8 times probability of survivorship for a wind turbine and 3.8 times probability of survivorship for an electrical subsystem which have history of scheduled maintenance than the one which does not have.
- Distance to coast is not found to be a significant reliability factor for wind turbine systems and electrical subsystems. However, shorter distance to coast increases the survivorship of switches by 39%. A potential explanation is that the wind patterns in coastal regions fluctuate less than the ones in land.
- Elevational location is found to be not a significant factor for survivorship of turbine systems, electrical subsystems and fuse and switch components. It must be noted that maximum elevation for considered turbines in this study is 800 m.
- Although hazard rate cannot be quantified due to proportionality violation, it is found that high number of previous failures (NOPF)reduce the survivorship of wind turbine as a system and electrical systems comparing to low NOPF. It is also found that high NOPF shows a lower survivorship for wind turbine components, however in order to determine a significance more data are required.
- MAWS is not shown to be a significant reliability factor for wind turbine systems and electrical subsystems. However, higher MAWS increase the survivorship of switches by 35% which can be attributed to more consistent wind patterns.

It must be noted that the WMEP data on which this study is based on, only cover a part of a turbine lives, the longest being 14 years from start of recorded operation. In the future studies, a complete life of turbines should be considered to obtain more concrete findings. Also, survival analysis can be applied to

Peer-reviewed version available at *Energies* **2018**, *11*, 3034; doi:10.3390/en11113034

improve the survival of expensive wind turbine components such as blades, gearboxes and generators subject to data availability.

References

1- GWEC (2017). Global Wind Report 2016 – Annual market update. [online] Gwec.net. Available at: http://gwec.net/publications/global-wind-report-2/global-wind-report-2016// [Accessed 3 March 2018].
2- Faulstich, S.; B. Hahn; P. J. Tavner. "Wind Turbine Downtime and Its Importance for Offshore Deployment." *Wind Energy* 14, no. 3 (April 1, 2011): 327–37. https://doi.org/10.1002/we.421.
3- Tchakoua, P.; Wamkeue, R.; Ouhrouche, M.; Slaoui-Hasnaoui, F.; Tameghe, T.A.; Ekemb, G. Wind Turbine Condition Monitoring: State-of-the-Art Review, New Trends, and Future Challenges. *Energies* **2014**, *7*, 2595-2630.
4- Amirat, Y.; Benbouzid, M. E. H.; Al-Ahmar, E.; Bensaker, B.; Turri, S. "A Brief Status on Condition Monitoring and Fault Diagnosis in Wind Energy Conversion Systems." *Renewable and Sustainable Energy Reviews* 13, no. 9 (December 1, 2009): 2629–36. https://doi.org/10.1016/j.rser.2009.06.031.
5- Sheng, S.; O'Connor, R. "Chapter 15 - Reliability of Wind Turbines." In *Wind Energy Engineering*, edited by Trevor M. Letcher, 299–327. Academic Press, 2017. https://doi.org/10.1016/B978-0-12-809451-8.00015-1.
6- Ghodrati, B.; Uday, K. "Reliability and Operating Environment-based Spare Parts Estimation Approach: A Case Study in Kiruna Mine, Sweden." *Journal of Quality in Maintenance Engineering* 11, no. 2 (June 1, 2005): 169–84. https://doi.org/10.1108/13552510510601366.
7- Tavner, P.; Edwards, C.; Brinkman, Andy; Spinato, F. "Influence of Wind Speed on Wind Turbine Reliability." *Wind Engineering* 30 (2006): 55–72.
8- Leite, A. P.; Borges, C. L. T.; Falcao, D. M. "Probabilistic Wind Farms Generation Model for Reliability Studies Applied to Brazilian Sites," in *IEEE Transactions on Power Systems*, vol. 21, no. 4, pp. 1493-1501, Nov. 2006. doi: 10.1109/TPWRS.2006.881160
9- Hau, E. "Wind Turbine Installation and Operation." In *Wind Turbines*, 719–87. Springer, Berlin, Heidelberg, 2013. https://doi.org/10.1007/978-3-642-27151-9_18.
10- Faulstich, S.; Lyding, P.; Hahn, B. *"Integrated Wind Turbine Design" Component Reliability Ranking with Respect to WT Concept and External Environmental Conditions Deliverable WP7.3.3, WP7 Condition Monitoring*, 2010.
11- Fischer, K.; Besnard, F; Bertling, L. "Reliability-Centered Maintenance for Wind Turbines Based on Statistical Analysis and Practical Experience." *IEEE Transactions on Energy Conversion* 27, no. 1 (March 2012): 184–95. https://doi.org/10.1109/TEC.2011.2176129.
12- Tavner, P. *Offshore Wind Turbines: Reliability, Availability and Maintenance*. IET Digital Library, 2012. https://doi.org/10.1049/PBRN013E, 10.1049/PBRN013E.
13- Wilson, G.; D. McMillan. "Assessing Wind Farm Reliability Using Weather Dependent Failure Rates." *Journal of Physics: Conference Series* 524, no. 1 (2014): 012181. https://doi.org/10.1088/1742-6596/524/1/012181.
14- Staffell, I.; Green, R. "How Does Wind Farm Performance Decline with Age?" *Renewable Energy* 66 (June 1, 2014): 775–86. https://doi.org/10.1016/j.renene.2013.10.041.
15- Pérez, P.; María, J.; Márquez, F. P. G.; Tobias, A.; Papaelias, M. "Wind Turbine Reliability Analysis." *Renewable and Sustainable Energy Reviews* 23 (July 1, 2013): 463–72. https://doi.org/10.1016/j.rser.2013.03.018.
16- Reder, M.; Yürüşen, N. Y.; Melero. J. J. "Data-Driven Learning Framework for Associating Weather Conditions and Wind Turbine Failures." *Reliability Engineering & System Safety* 169 (January 1, 2018): 554–69. https://doi.org/10.1016/j.ress.2017.10.004.

17- Slimacek, V.; B.H. Lindqvist. “Reliability of Wind Turbines Modeled by a Poisson Process with Covariates, Unobserved Heterogeneity and Seasonality.” *Wind Energy* 19, no. 11 (2016): 1991–2002. https://doi.org/10.1002/we.1964.
18- Mazidi, P.; Bertling Tjernberg, L.; Bobi, M. A. S. “Wind Turbine Prognostics and Maintenance Management Based on a Hybrid Approach of Neural Networks and a Proportional Hazards Model.” *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability* 231, no. 2 (April 1, 2017): 121–29. https://doi.org/10.1177/1748006X16686899.
19- Carlos, S.; Sánchez, A.; Martorell, S.; Marton, I. “Onshore Wind Farms Maintenance Optimization Using a Stochastic Model.” *Mathematical and Computer Modelling*, Public Key Services and Infrastructures EUROPKI-2010-Mathematical Modelling in Engineering & Human Behaviour 2011, 57, no. 7 (April 1, 2013): 1884–90. https://doi.org/10.1016/j.mcm.2011.12.025.
20- Andrawus, J. Maintenance optimisation for wind turbines. PhD thesis, The Robert Gordon University, Aberdeen, UK, 2008
21- Kerres, B.; Fischer, K.; Madlener, R. “Economic Evaluation of Maintenance Strategies for Wind Turbines: A Stochastic Analysis.” SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, March 1, 2014. https://papers.ssrn.com/abstract=2469795.
22- Christodoulou, S. E. “Water Network Assessment and Reliability Analysis by Use of Survival Analysis.” *Water Resources Management* 25, no. 4 (March 1, 2011): 1229–38. https://doi.org/10.1007/s11269-010-9679-8.
23- IEA WIND TASK 33. Reliability Data Standardization of Data Collection for Wind Turbine Reliability and Operations & Maintenance Analyses: Initiatives Concerning Reliability Data (2nd Release); 2013.
24- Kaplan, E. L.; Meier, P. (1958). "Nonparametric estimation from incomplete observations". *J. Amer. Statist. Assoc.* **53** (282): 457–481. doi:10.2307/2281868. JSTOR 2281868.
25- Aalen, O.O. (1978). Nonparametric inference for a family of counting processes, Annals of Statistics 6, 701–726.
26- Bland, J. M.; Altman, D. G. “The Logrank Test.” *BMJ : British Medical Journal* 328, no. 7447 (May 1, 2004): 1073.
27- IBM Corp. Released 2017. IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp.
28- Peel, M.; Finlayson, B.; Mcmahon, T. “Updated World Map of the Koppen-Geiger Climate Classification.” Hydrology and Earth System Sciences Discussions 4 (2007)
29- Global Wind Atlas. Available online: https://globalwindatlas.info/ (June 26, 2018)
30- Ribrant, J.; Bertling, L. M. “Survey of Failures in Wind Power Systems With Focus on Swedish Wind Power Plants During 1997 Ndash;2005.” *IEEE Transactions on Energy Conversion* 22, no. 1 (March 2007): 167–73. https://doi.org/10.1109/TEC.2006.889614.
31- Miller, C. R. *Illustrated Guide to the National Electrical Code*. Albany, N.Y.: Delmar Publishers, 1999. https://catalog.hathitrust.org/Record/004044318.
32- Patrick, D. R. *Electricity and Electronics Fundamentals*. 2. ed. Boca Raton, Fla.: CRC Press, 2008.