Recoverability modelling of power distribution systems using accelerated life models: The case of power cut due to extreme weather events in Norway

Authors: Bjarte Rød[1*], Abbas Barabadi[2], Masoud Naseri[3]

[1] M.Sc., UiT The Arctic University of Norway, Department of Technology and Safety, P.O. 6050 Langnes 9037 Tromsø, Norway, e-mail: bjarte.rod@uit.no

[2] Prof., UiT The Arctic University of Norway, Department of Technology and Safety, P.O. 6050 Langnes 9037 Tromsø, Norway, e-mail: abbas.b.abadi@uit.no

[3] Assoc. Prof., UiT The Arctic University of Norway, Department of Technology and Safety, P.O. 1063, 9480 Harstad, Norway, e-mail: masoud.naseri@uit.no

[*] Corresponding author: E-mail: bjarte.rod@uit.no, Tel: (+47) 98615331

ABSTRACT

Today's societies rely on electrical power distribution systems. Recent weather events have illustrated that the loss of such service can lead to severe consequences for societies and stakeholders. Hence, in order to reduce the impact of such extreme events on infrastructure systems and to limit the associated losses, it is crucial to design infrastructure that can bounce back and recover rapidly after disruptions (i.e. to be resilient). In this regard, it is vital to have knowledge of technical, organizational, internal, and external factors that influence the infrastructure's recovery process. These factors can broadly be categorized into two different groups, namely observed and unobserved risk factors. In most studies on resilience, the effect of unobserved covariates is neglected. This may lead to erroneous model selection for analyzing the time to recovery of the disrupted infrastructure, as well as wrong conclusions and thus decisions. The aim of this paper is to identify the risk factors (observed and unobserved) affecting the recovery process of disrupted infrastructure. To this aim, the paper extends the application of accelerated failure time (AFT) models, to model the recovery time of disrupted critical infrastructures in the presence of unobserved and observed risk factors. This model can be used to analyse how important these factors are from the viewpoint of resource allocation and decision-making. The application and implications of the model are presented in a case study, from both technical and management perspectives. The case study investigated in this paper applies the developed model, analysing recovery times from 73 disruption reports on Norwegian electric power distribution grids after four major extreme weather events. The analysis indicates that failures in the regional grid, natural conditions, area affected, and failures in operational control system have a significant impact on the recovery process.

Keywords: recovery, resilience, electric power distribution systems, critical infrastructure, extreme weather events, accelerated failure time models.

1 INTRODUCTION

Over recent decades, it has been evident that society relies heavily on infrastructure systems to provide and maintain vital societal functions (Rinaldi et al. 2001). Traditionally, in order to ensure the delivery of such functions, the focus of industry has been on the protection of the infrastructure systems from adverse and extreme events, such as hurricanes, tsunamis, floods, and so forth. However, recent events, such as Hurricane Sandy (Comes and Van de Walle 2014) and the tsunami that hit Japan in 2011, leading to a nuclear disaster (Bacon and Hobson 2014), illustrate that it is very difficult, and often not feasible, to protect such systems from all kinds of possible hazards. Hence, there has been a shift from the protection of critical infrastructure to the resilience of critical infrastructure, increasing the focus on preparedness, response and recovery (Pursiainen and Gattinesi 2014; Haimes 2012). In other words, having a resilient infrastructure, with the ability to limit the consequences of an impact through timely and efficient recovery processes, will certainly benefit the infrastructure operators and society as a whole (Choi et al., 2019). To effectively recover infrastructures from extreme events, it is essential for infrastructure operators to have knowledge of the factors (external, technical and organizational) that influence the recovery process. Such knowledge helps the analysts and decision-makers to make realistic estimates of the recovery rate and recovery time of the infrastructures.

Despite the growing number of studies on resilience in engineering systems, there is no common agreement as regards the definition of the concept or, more importantly, of how to assess and measure resilience (Hosseini et al. 2016). However, the most common resilience metric is the well-known resilience triangle, illustrating the loss of performance over time (Bruneau et al. 2003), as shown in Figure 1, adapted from Honfi et al. (2017). The figure illustrates the performance (Q) over time for a system experiencing some kind of incident, occurring at time t_i. The system develops a failure mechanism f. At time t_f, the system gradually starts to recover, through the process which is described by the recovery path r in Figure 1. At time t_r, the system is fully recovered and performs its required function at the same standard as before the incident.
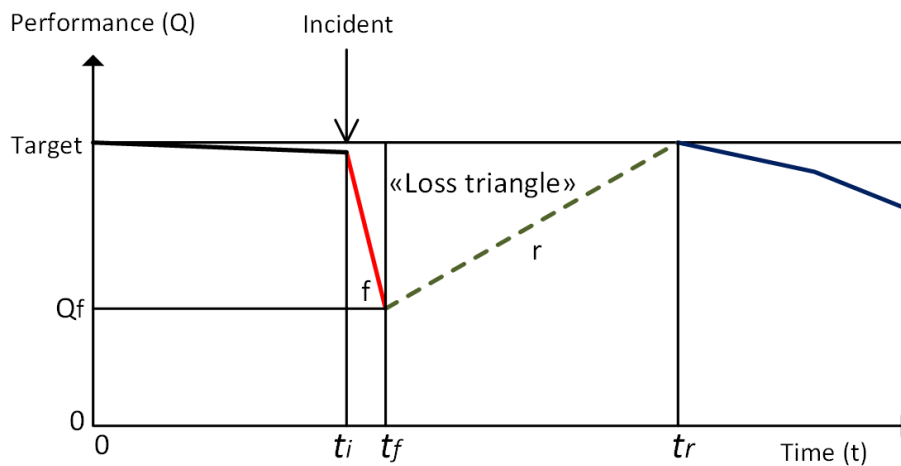

Figure 1. The performance loss function. Adapted from Honfi et al. (2017).

The resilience triangle illustrates the performance of the infrastructure over time, and, the smaller the triangle, the more resilient the infrastructure is. However, considering the trajectory of the recovery path and recovery time without investigating the environmental conditions and other conditions under which the recovery process takes place (i.e. influencing variables), such as number of crew, available resources, environmental conditions, region, technical condition of the system, etc., leads to a great deal of uncertainty and, thus, unreliable analysis results. A possible explanation for such results is that the recovery time as a random variable is, to a great extent, dependent on a set of prevailing operating or environmental conditions, which, through different mechanisms, can affect the length of the recovery time and, thus, recovery rate.

In general, having an effective contingency plan requires extensive knowledge concerning the recovery time of the specific system. Moreover, to have a reliable estimation of the recovery time, the effects of all factors that can influence the recovery process and path should be quantified, using appropriate models. Such models can be used as a basis for developing preparation plans, developing resource allocation strategies, identifying vulnerable recovery scenarios, and learning from the incidents. Influencing factors on the trajectory of the recovery path can be categorized into two groups: *i)* observed risk factors and *ii)* unobserved risk factors. Observed risk factors describe the recovery process characteristics (e.g. type of equipment used during the recovery process, number of maintenance personnel involved, etc.) or the environmental characteristics under which the recovery process took place (e.g. location of the disrupted infrastructures, cause of failures, weather conditions, etc.). Unobserved risk factors are independent variables that may have a significant impact on the recovery time of the infrastructure. However, these are not reported and thus not available in recovery databases. Observed and unobserved risk factors may lead to observed and unobserved heterogeneity. For example, in some situations, local people might help the repair crew to repair the failures and recover the infrastructure. However, their efforts and contribution to a reduced recovery time are not recorded in the corresponding databases. In this regard, their effect on recovery time should be modelled using unobserved risk factors.

Some methods, such as accelerated failure time (AFT) and proportional hazard (PH) models, have been widely used in order to analyse the effects of observed and unobserved risk factors, also known as covariates, on random duration time in survival analysis (e.g. Wei (1992); Bradburn et al. (2003); Orbe et al. (2002); Cox (2018); Fine and Gray (1999)). Although, in survival analysis, some studies have used PH and AFT models to analyse the impact of covariates on the hazard rate and survival time in various fields (e.g. Alvehag and Soder (2011); Alvehag and Soder (2008); Rocchetta et al. (2015); Tian et al. (2005); Peng and Huang (2007)), there is a gap in the literature, in which the application of such methods can be explored in the area of the recovery process of disrupted infrastructures and, in general, in resilience studies.

In this regard, the key novelty of the paper lies in exploring the application of AFT models in analysing the recoverability of disrupted infrastructures, in addition to analysing the impact of observed and unobserved risk factors on the recovery time. This is achieved by considering the operating conditions and other covariates, where the recovery time is selected to be the random variable of interest. Therefore, the results of this study enable managers to make informed decisions regarding resource allocation,

contingency plans, and preparedness plans. From a managerial perspective, the response and recovery process can be optimized by taking these factors into consideration. In so doing, the consequences for the customer and society will also be reduced.

Although the presented approach in the current study can be used in various critical infrastructures, the scope of the present case study is electric power distribution systems. The case study is resolved by analysing the recovery times from outages corresponding to 73 disruptions within the Norwegian system from 2013 to 2016, after four extreme weather events. Therefore, the method is illustrated by a case that consists of major parts of the electric power distribution and transmission grid in Norway. The main driver for choosing electric power distribution as the case study is the fact that it is among the most important critical infrastructures. Considering the high level of interdependency between critical infrastructures in our modern societies (e.g., transportation, health, power distribution, communication, water supply, etc.), any disruption in electric power distribution can trigger the disruption of other mentioned infrastructures. Hence, all electric power distribution companies should have clear understanding regarding the recoverability, i.e. the ability of the organization to recover from disruptions, of their power distribution systems. Moreover, it is crucial for the operator to know how to optimize the recovery process with limited resources in place. Normally, these companies apply relatively simple deterministic models, based on damage assessment in the field, to estimate recovery time, which can only be applied after the event has occurred. Such models are not able to identify the significance of the (observed and unobserved) risk factors and the extent of their impact. Considering the importance of power distribution systems for our society, there is an urgent need to develop some new statistical approaches for modelling the effect of observed and unobserved risk factors on their recoverability. To this aim, the contribution of this paper goes further in employing the AFT model to identify important parameters affecting the recovery of Norwegian electric power distribution systems and in analysing how important these factors are regarding resource allocation and decision-making in future disruptions of the power distribution grid. In addition, this study gives guidance on the use of suitable statistical models for generating accurate and reliable results, which can provide infrastructure operators with valuable information when making important decisions before, during and after a disruption.

The rest of the paper is organized as follows. First, a general discussion about resilience is presented, followed by a literature review about prediction and modelling of resilience and recovery. Thereafter, the Norwegian electric power distribution system is described. The data and methodology are then described, and results from the case study are presented. Finally, some conclusions and recommendations for future works are provided.

2 BACKGROUND AND LITERATURE REVIEW

2.1 Resilience definition and metrics

The definition of resilience is a contested one, and there is no clear definition of the concept, at this time, that could be applied universally (Rochas et al. 2015). The original meaning of the word comes from the

Latin word *resilire*, which can be understood as the "ability to rebound or jump-back" (Dalziell and McManus 2004), highlighting the essence of the concept – the ability to bounce back. In material science, resilience is understood as the ability of materials to recover their shape after being stretched or deformed (Dessavre et al. 2016). In the context of disaster risk reduction, the United Nations (UNISDR n.d.) provides a comprehensive and general description of resilience, as follows: "*the ability of a system, community or society exposed to hazards to resist, absorb, accommodate, adapt to, transform and recover from the effects of a hazard in a timely and efficient manner, including through the preservation and restoration of its essential basic structures and functions through risk management*". This definition emphasises that resilience has its temporal dimensions, including the ability to resist. From a social perspective, Cutter et al. (2008) describe social resilience as "*the ability of a social system to respond and recover from disasters that includes those inherent conditions that allow the system to absorb impacts and cope with an event, as well as adaptive processes that facilitate the ability of the social system to reorganize, change and learn in response to a threat*". It can be argued that, to some extent, adaptive and absorptive capacities are developed prior to the event, implying that the pre-event stage is also included here. This implies that resilience is, as stated by Lange et al. (2017a), "*a process that has to be present and enhanced before, during and after a crisis or disruption of services*". Nevertheless, the exact effect of measures implemented before a crisis is only known after the event.

From an engineering and technical point of view, the key elements related to resilience consist of concepts such as resistance, absorption, adaption and recovery (Francis and Bekera 2014). In many ways, resilience integrates, for better or worse, existing measures of risk, vulnerability, reliability, robustness, survivability, adaptability, maintainability, availability, and so forth, in order to measure resilience (Faturechi and Miller-Hooks 2014). Hence, how one measures and quantifies resilience will, of course, depend on the concept(s) one includes and the metrics and units that are applied to measure these concepts. In essence, this relates to the drop or loss in performance (as described in Figure 1), as a direct or indirect consequence of an abnormal situation. Bollinger and Dijkema (2016) measure this loss in performance in terms of service level, evaluating the resilience level of the Dutch electricity transmission network as a mean fraction of demand served across the range of possible extreme event magnitudes. Cimellaro et al. (2014) measure the infrastructure service level after the 2011 Tohoku earthquake in Japan as the restoration ratio between the number of households without service and the total number of households. Hossain et al. (2019) emphasise that "*Resilience is the ratio of recovery*", measured as recovered production capacity to lost production capacity. Other studies focus more on the functional level of the infrastructure systems. For instance, Ouyang et al. (2012) state that "*The performance level is measured by the number of normally operating components within an infrastructure system*". Similarly, Rochas et al. (2015) use the total length of functioning pipelines as a 'figure of merit' to measure the functional level of an infrastructure. There are also studies that focus on the general quality of the infrastructure systems, such as Mendonça and Wallace (2006), who investigated the number of disruptions for each infrastructure during various weeks of the event, in order to analyse the impacts of the World Trade Centre attack in New York on city critical infrastructures over a three-month period.

For this case study, resilience will be defined in accordance with the time to recovery, where the resilience metric is the service level, measured in terms of end users with power supply. Hence, based on this definition, the AFT model is used to model the time to recovery, considering the effect of observed and unobserved risk factors. It should be mentioned that the recoverability will be modelled by the number of customers affected by the disruption. A limitation of this metric is that the occurrence of disruptions in less populated areas may not reflect the magnitude of the disruption and the level of physical damage. However, in many quantitative resilience assessment methodologies, the recoverability or recovery rate is considered only as a minor part of the resilience definitions (the other parts are prevention, absorption, adaptation) (see e.g. Francis and Bakera (2014) and McEvoy et al. (2012)). However, the weight placed on the recovery phase, compared to other phases of resilience, may vary, based on the selected definition. For example, in the engineering and technical area, where resilience is often divided into several phases and described by several concepts, such as resistance, absorption, adaption and recovery (see e.g. Ouyang & Wang (2015); Kong et al. (2019)), the developed model in this study would then be a part of a more comprehensible definition of resilience. In other words, regardless of the definition of resilience, the recovery is always an important phase, and the model presented in this study is thus applicable in the resilience context.

2.2 Prediction and modelling of resilience and recovery

In general, as stated by Hosseini et al. (2016), quantitative resilience assessment methodologies can be divided into general measures and structurally based models. General measures include probabilistic and deterministic measures, while structurally based models include optimisation, simulation and fuzzy logic approaches. Modelling and simulation of critical infrastructures has become a key field of study, and numerous approaches have been developed over recent years (see studies such as Ouyang (2014) and Hosseini et al. (2016)). A common feature of such models is that they investigate how the structure of the system impacts the resilience level. This is done by observing the system behaviour and modelling and simulating the characteristics of the system. Many of these models represent a real-life restoration process, including a high level of detail (e.g. Çağnan et al. (2006)), which requires a huge amount of data to be being collected and processed.

Probabilistic approaches, categorised as general measures, account for uncertainty, and the stochastic behaviour of the disruptive events, as well as the stochastic behaviour and randomness of duration (i.e. recovery time), are, to a large extent, captured. For instance, Youn et al. (2011) describe resilience by using two traditional concepts, namely, reliability and restoration, where restoration is described as the joint probability of a system failure event, a correct diagnosis event, and a mitigation/recovery action success event. Restoration and recoverability is often referred to as maintainability in conventional reliability engineering, defined as "*the ability of an item under a given condition of use, to be retained in, or restored to, a state in which it can perform a required function, when maintenance is performed under given conditions and using stated procedures and resources*" (International Electrotechnical Vocabulary (IEV) 191 2007). In maintainability analysis methods, the repair or restoration time is considered a random variable (Blanchard et al. 1995, Dhillon 1999). The aim of such

analysis methods is to model the probability that a successful repair process takes place within a stated time interval under procedures and resources (Barabadi et al. 2011), also known as survival analysis. Qiao et al. (2019) classify survival models as non-parametric, semiparametric, or fully parametric. The nonparametric can be easily implemented and does not require any assumptions. However, as stated by the definition of maintainability, the time required for restoration or repair depends on a range of conditions under which the restoration process occurs. Such conditions may include technical features, organisational aspects, and environmental conditions. The nonparametric models do not have the ability to relate these external factors to the restoration function. In order to capture the impact of these conditions and elements, also known as influencing variables or covariates, fully parametric models can be used. AFT and PH models are often used (e.g. Barabadi et al. (2011); Kayrbekova et al. (2011); Naseri (2017)) to study the extent to which the repair time or maintainability depends on the underlying conditions. In an analogy with the maintainability analysis, one may focus on the application of AFT and PH models in the recoverability of an infrastructure unit after a disruptive event. Such models provide the analysts with an opportunity to analyse the impact of different influencing parameters on recovery time or, in general, on recoverability. In this regard, recoverability can be defined as the ability of an organisation to restore an infrastructure unit to a level that is able to deliver required functions as before the occurrence of the disruptive event.

The study by Liu et al. (2007) was one of the first to implement survival analysis to model power outage restoration times during hurricanes and ice storms, using AFT and Cox proportional hazard (Cox PH) models. The authors conclude that AFT is better than Cox, mainly because the results from AFT are easier to interpret. Nateghi et al. (2011) compare five statistical models for estimating power outage duration times: AFT, Cox PH, and data mining techniques (regression trees, Bayesian additive regression tree (BART), and multivariate additive regression splines). They state that BART yields the best predication accuracy but emphasise that the AFT model "*provides a further basis for examining the influence of each covariate on the restoration periods*". Similar statistical methods have been applied in a variety of fields and disciplines, such as health science (e.g. Bakhshi et al. (2017)), accident investigations (e.g. Saeed et al. (2019a and b)), project management (e.g. Qiao et al. (2019)), and the oil and gas industry (e.g. Ilbeigi & Dilkina (2017)) – underlining the broad application area of such methods. However, as mentioned in the introductory section, these studies do not consider the effect of unobserved risk factors. In general, due to the nature of the recovery process, recovery procedures, location of the accident, type of accident, culture of the people affected by the disrupted infrastructures and so on, it is very difficult to capture and record all risk factors in the recovery database. Moreover, our experience with the Norwegian electric power distribution systems and oil and gas industries can confirm this fact: that most of the available recovery data are not very well collected and they do not reflect the actual environmental conditions of the recovery site of the infrastructure. Considering the fact that the results of the recovery analysis will be used later for learning processes in contingency planning, neglecting the impact of unobserved risk factors would lead to biased results and thus unrealistic resource distribution and planning.

2.3 The Norwegian Electrical Power Distribution System

The Norwegian electric power distribution system is divided into three different levels, namely, the distribution grid, the regional grid, and the transmission grid. Consistent with international terminology, we often use 'distribution grid' as an umbrella term for both the distribution and the regional grid in Norway. The transmission grid, has the highest voltage level, ranging from 132 kV to 400 kV; it acts as a link between the producers and the customers in a nationwide system. The transmission grid is about 11,000 km. It is mainly operated by Statnett SF, which is the only Transmission System Operator (TSA) owned by the state; licensed by the Norwegian Water Resources and Energy Directorate (NVE), it is regulated by the Norwegian Energy Act of 1990. The regional grid is the link between the transmission grid and the distribution grid. However, some parts of the grid also consist of production and consumption radials. In total, the regional grid is 19,000 km, of which 8% comprises sea and underground cables. The distribution grid serves the end user, such as households, public services, and industry, with power. The voltage level ranges from 22 kV to 230 V. In total, the distribution grid consists of 100,000 km of lines with a voltage level above 1 kV, of which 40% comprises sea and underground cables (Hatlen and Knudsen Aarrestad 2015).

The Norwegian Water Resources and Energy Directorate (NVE), organised under the Norwegian Ministry of Petroleum and Energy, has the overall responsibility for maintaining the national power supply. One of the directorate's tasks is to issue regulations on system responsibility and to ensure the quality of the power supply. All Norwegian grid companies are obliged to report interruptions to NVE. In 2015, a total of 159 companies operated in the Norwegian electric power grid on one or several levels. These 159 companies cover different geographical areas in Norway, and there is a large deviation among the companies in terms of the number of customers served, size of the service area, geographical characteristics, and so forth. Each company is regulated under the 'compensation for non-delivery of energy' (KILE). This gives distribution companies reduced income in the event of an interruption. As stated by the Norwegian government, "*The KILE scheme is a means for distribution companies to be confronted with customer interruptions cost and take into consideration these costs when making decisions*". This KILE scheme thus ensures that reliability is taken into account when the companies make important decisions, both during operation and with respect to future investments.

In the case of interruptions in the power supply, and to ensure the quality of the supply, each company is obliged to report failure data to the regulator, which is NVE. This is done through the Fault and Supply Interruption information Tool (FASIT), developed in the 1990s. Since 1995, all Norwegian grid operator companies are required to use this tool for the collection and reporting of component fault and delivery point interruption data (Heggset et al. 2009). In addition, when extreme weather events occur, such as major autumn and winter storms, each company that is affected by the storm must prepare and submit extensive reports to NVE. Such a report includes a range of qualitative and quantitative data. The qualitative data concerns the operator's subjective opinions on how the organisation managed to prevent or recover from disruption and power cut. Such data is a valuable source of information that gives a much clearer picture of the recovery process, integrating the organisational and technical resilience domains.

3 DATA AND METHODOLOGY

In this case study, considering the available data, the AFT model is applied. As emphasised in the literature review, there exists a wide range simulation and modelling approaches applicable for infrastructure systems. However, without detailed information about the system characteristics, such methods might produce inaccurate results.

3.1 Model

In risk and reliability analysis fields, the time to failure of a system or the time to repair a failed component is considered a random variable (Rausand and Høyland 2004). This can also be applied to analysing the resilience of infrastructures, including power distribution grids, where the time that it takes to have the grid in the new equilibrium state or back to its normal operating level can also be considered a random variable (Francis and Bekera 2014, Hosseini et al. 2016). The randomness of the time to recover a power distribution grid thus requires the application of probabilistic models.

More specifically, in the current modelling setting, the variable of interest is the duration or the length of time that the recovery process takes. This parameter, which is inherently a random variable, is often referred to as recovery time, as shown by the length $t_f - t_r$ in Figure 1. Such a time interval begins with the initiation of recovery efforts, which is usually upon noticing the power outage, until the recovery process is finished and electricity is again provided for customers.

Let T be a positive random variable, denoting the recovery time. Also, let $f(t)$ be the corresponding probability density function (pdf) of random variable T. Thus, the cumulative distribution function (cdf), $F(t)$ of random variable T (Rausand and Høyland 2004) – which, in the current modelling framework is recoverability denoted by $R(t)$ – expresses the probability that the recovery process is completed at time $T < t$. Therefore, the recoverability can be defined by Equation (1):

$$R(t) = \Pr(T < t) = \int_0^t f(u)du \tag{1}$$

Using such terminology, the recovery rate, denoted by $r(t)$, is defined as the probability that the recovery is completed in the time interval $(t, t + \Delta t]$ when it is known that the recovery has not been completed until time t (i.e., it is known that electricity is still down at time t and customers experience a power cut at time t):

$$r(t) = \Pr(t < T \leq t + \Delta t | T > t) = \frac{\Pr(t < T \leq t + \Delta t)}{\Pr(T > t)} = \frac{f(t)}{1 - R(t)} \tag{2}$$

By combining Equations (1) and (2), the recoverability function, $R(t)$, can be expressed as:

$$R(t) = 1 - \exp\left[-\int_0^t r(u)du\right] \tag{3}$$

Survival function, $S(t)$, is another important concept in duration analysis, given as (Rausand and Høyland 2004):

$$S(t) = \Pr(T \geq t) = 1 - \int_0^t f(u)du$$

In the context of the present study, $S(t)$ states the probability that the recovery cannot be completed before some specified time t:

$$S(t) = 1 - \int_0^t f(u)du = 1 - F(t) = \exp\left[-\int_0^t r(u)du\right] \tag{4}$$

However, Equations (1) to (4) do not include the impact of any covariate or operating condition on the recoverability or the recovery time of the power grid. In survival analysis, various models including accelerated failure time (AFT) and proportional hazard (PH) models have been widely used, in order to analyse the effects of explanatory variables (also known as covariates) on the random duration time (e.g. Wei (1992); Bradburn et al. (2003); Orbe et al. (2002); Cox (2018); Fine and Gray (1999)). The main difference between the AFT and PH models lies in modelling the impact of covariates on the random dependent variable, i.e. duration of recovery time. While, in AFT models, covariates have multiplicative effects on time, in PH models, covariates have multiplicative effects on hazard rate (Kumar and Klefsjö 1994, Nelson 2009). Such models have also been widely used in reliability, availability, and maintainability analyses, in order to capture the impact of covariates on failure and repair times (see e.g. Bagdonavicius and Nikulin (2001); Ghodrati and Kumar (2005b); Crowder (2017); Naseri and Barabady (2016); Naseri et al. (2016)). Different types of covariates are used in such studies including environmental conditions (Barabadi 2014), weather conditions (Naseri et al. 2016), and skill level of operation crew (Ghodrati and Kumar 2005), as well as location of the plant and batch of the production, as discussed in studies by Ansell and Philipps (1997); Dale (1985), Jardine et al. (1987) and Kumar et al. (1992).

Given the above-mentioned discussion, and due to the fact that the the recoverability of an infrastructure unit and its recovery rate after the occurrence of a disruption depend on a number of parameters and conditions under which the recovery process takes place, the current study employs the AFT model to investigate the impact of the influencing parameters (i.e., operating and environmental conditions) on the recovery time of power grids after disruption.

As mentioned earlier, in AFT models, the effects of covariates or explanatory variables on the random variable time are expressed as multiplicative factors to the time (Bagdonavicius and Nikulin 2001, Kumar and Klefsjö 1994, Nelson 2009). In other words, according to the general log-linear relationship between time T and a vector of covariates, the natural logarithm of recovery time is expressed as a linear model of the covariates (Nelson 2009), as given by Equation (5):

$$\ln T = \alpha_0 + \sum_{k=1}^n \alpha_k x_k \tag{5}$$

where n is the total number of covariates, $x_k, k = 1, \dots n$ is the kth covariate, $\alpha_k, k = 1, \dots n$ is the regression coefficient, and α_0 is a constant error term. The distributional form of the error term determines the regression model. Various distributions can be used to develop the recovery time model, including Weibull, exponential, and lognormal (Lee and Wang 2003, Rausand and Høyland 2004).

Due to the flexibility of the Weibull distribution in modelling different patterns of hazard rates, this study uses the Weibull distribution as the underlying distribution model, which has a probability density function given by:

$$f(t) = \frac{\beta}{\eta^\beta} t^{\beta-1} 1 - e^{-\left(\frac{t}{\eta}\right)^\beta} \tag{6}$$

where η and β are the scale and shape parameters. The recoverability function can then be obtained by substituting Equation (6) into Equation (1):

$$R(t) = 1 - e^{-\left(\frac{t}{\eta}\right)^{\beta}} \tag{7}$$

According to the approach suggested by the Department of Defence (1991), in AFT models, the independent random time variable is modelled by multiplying the baseline time, say t_0, by a functional form, $\exp(\alpha_k x_k + \alpha_k x_k + \cdots + \alpha_k x_k), k = 1, \dots, n$, which represents the impact of covariates on the independent variable, time. Thus

$$t = \exp(\alpha_k x_k + \alpha_k x_k + \cdots + \alpha_k x_k) \, t_0 \tag{8}$$

where t_0 is the recovery time under base conditions. By substituting Equation (8) into Equation (7) and according to the equivalent age concept (Naseri et al. (2016), Department of Defence (1991), the recoverability function under the impact of covariates can be rewritten as:

$$R(t|x_k) = 1 - e^{-\left(\frac{t}{\eta \exp(\alpha_k x_k + \alpha_k x_k + \cdots + \alpha_k x_k)}\right)^{\beta}} \tag{9}$$

or

$$R(t|x_k) = 1 - e^{-\left(\frac{t}{\exp(AX)}\right)^{\beta}} \rightarrow R(t|x_k) = 1 - \exp\left[-t^{\beta} \exp[-\beta AX]\right] \tag{10}$$

where β is the shape parameter of the Weibull distribution, A is the regression coefficient row vector including the constant error term, $\alpha_k, k = 0, \dots n$, where $\alpha_0 = \ln \eta_0$ and X is the covariate column vector with $x_k, k = 0, \dots n$, where $x_0 = 1$ and η_0 is the scale parameter under base conditions.

Equation (10) can be used to express the impact of covariates or environmental conditions on the recoverability of the power grid. By substituting Equation (10) into Equation (3), the recovery rate under the influence of covariates can be obtained:

$$r(t|x_k) = \beta t^{\beta} \exp[-\beta AX] \tag{11}$$

In Equations (10) and (11), covariates $x_k, k = 1, \dots n$, can be dependent or independent of time. In other words, the corresponding values of these covariates either change with time or can be assumed to be constant. In the present study, it can be assumed that these covariates do not change within the time frame of recovery. In other words, the covariates $x_k, k = 1, \dots n$ are assumed to be time-independent. Regression coefficients are estimated using maximum likelihood estimation methods (Lee and Wang 2003, Neath and Cavanaugh 2012, Pan 2001, Volinsky and Raftery 2000).

Traditionally, AFT and PH models are used with the assumption of homogeneity of the cumulative distribution function across the individuals (i.e. observations). However, this assumption leads to a great deal of uncertainty – if not wrong results – if some heterogeneity is present among the observations. Moreover, traditional analyses assume that the observations are independent (Hougaard 2016, Mohammadian and Doherty 2006, Yashin et al. 1995). However, in the context of the current study, it can be argued that, in certain cases, some failed components of the system are repaired, and the electricity power grid is brought back to operation so that a group of customers receives electricity. This indicates a group recovery for some power cut scenarios, i.e. electricity is provided for a group of customers, by repairing certain failed components.

Given the above discussion, one should account for the unobserved heterogeneity in the observation; different approaches are used for this in the literature. Some researchers have used random parameter models (Seraneeprkarn et al. 2017; Rahman Shaon et al. 2018; Afghari et al. 2019; Saeed et al. 2019) in estimating car crashes and the impact of some explanatory variables on the number of crashes in a road segment. In such studies, the coefficients of the covariates are assumed to be random variables. This implies that the coefficients have different effects on different observations. In other words, the heterogeneity of the explanatory variables is estimated through the randomness of coefficients. Some justifications for choosing random regression coefficients and thus using random parameter models is provided in a study by Mannering et al. (2016), where the random effect of different variables, including human elements, vehicle characteristics, safety-feature indicators, as well as roadway and traffic characteristics, on the number of road car crashes is discussed. Another approach, which is used in duration analysis and is employed in the current study, relies on shared frailty models, where the effect of heterogeneity is modelled by introducing a multiplicative parameter, known as shared frailty, to the hazard function (Yue and Chan 1997; Hougaard 1995; Matsuoka 2010; Hanagal 2017; Nath et al. 2016; Hesam et al. 2018; Fagbamigme et al. 2019). This also accounts for the presence of unobserved covariates that affect the recovery rate and recoverability. Shared frailty is, in fact, a group-specific unobserved or latent random effect, which is multiplied by the recovery rate function. Another role of shared frailty in the recoverability model is to generate some dependency among the observations that can be grouped together (Gutierrez 2002).

In order to account for the shared frailty, let the data consist of M groups, with one of them consisting of N_i individuals. The frailty of the ith group is then denoted by ε_i, which is a positive random number with mean equal to 1, variance θ, and the probability density function $g(\varepsilon_i)$. Those individuals or observations with $\varepsilon_i > 1$ are said to be frailer, for reasons left unexplained by the observed covariates, and will experience a higher recovery rate. Conversely, those individuals or observations with $\varepsilon_i < 1$ are less frail and will tend to have a lower recovery rate. Observations with higher and lower recovery rates tend to be associated with lower and higher recovery times, respectively. By introducing the frailty parameter, the conditional recovery rate and recoverability function for individual j in the ith group can be written as in Equations (12) and (13), respectively (Gutierrez 2002, Hougaard 1995, Wienke 2010):

$$r_{ij}(t|\varepsilon_i) = \varepsilon_i h_{ij}(t) \tag{12}$$

$$R_{ij}(t|\varepsilon_i) = 1 - \exp\left[-\int_0^t \varepsilon_i r_{ij}(u)du\right] = 1 - \left[S_{ij}(t)\right]^{\varepsilon_i} \tag{13}$$

where $j = 1, \dots, N_i$, $i = 1, \dots M$, $h_{ij}(u)$ and $F_{ij}(t)$ are individual non-frailty recovery rate and recoverability functions, respectively. The unconditional survival function and unconditional recoverability function for individual j in the ith group, when the frailty is present, are then obtained using Equation (13) and are given by Equations (14) and (15), respectively (Gutierrez 2002, Hougaard 1995, Wienke 2010):

$$R'_{ij}(t) = \int_0^\infty g(\varepsilon_i)R_{ij}(t|\varepsilon_i)d\varepsilon_i = \int_0^\infty g(\varepsilon_i)\left[1 - \left[S_{ij}(t)\right]^{\varepsilon_i}\right]d\varepsilon_i \tag{14}$$

$$S'_{ij}(t) = 1 - \int_0^\infty g(\varepsilon_i)R_{ij}(t|\varepsilon_i)d\varepsilon_i = \int_0^\infty g(\varepsilon_i)\left[S_{ij}(t)\right]^{\varepsilon_i}d\varepsilon_i \tag{15}$$

Gamma distribution is a common distribution model for handling the heterogeneity of the data. By assuming a gamma-distributed shared frailty, given by Equation (16) (Gutierrez 2002, Hougaard 1995, Wienke 2010):

$$g(\varepsilon_i) = \frac{\varepsilon_i^{1/\theta-1} \exp(-\varepsilon_i/\theta)}{\Gamma(1/\theta)\theta^{1/\theta}} \tag{16}$$

The unconditional survival function and unconditional recoverability function for individual j in the ith group, when the frailty is present, can be written as (Gutierrez 2002, Hougaard 1995, Wienke 2010):

$$R'_{ij}(t) = 1 - \left[1 - \theta \ln\left[1 - R_{ij}(t)\right]\right]^{-\frac{1}{\theta}} \tag{17}$$

By also introducing the observed covariates, Equation (13) can be rewritten as:

$$R_{ij}\left(t|\varepsilon_i, x_k^{ij}\right) = 1 - \exp\left[-\int_0^t \varepsilon_i r_{ij}\left(u|x_k^{ij}\right)du\right] = 1 - \left[S_{ij}\left(t|x_k^{ij}\right)\right]^{\varepsilon_i} \tag{18}$$

where $x_k^{ij}, k = 1, \dots, n$ is the kth covariate of individual j in the ith group. By assuming a gamma-distributed frailty, the unconditional form of Equation (18) can be written as:

$$F'_{ij}\left(t|x_k^{ij}\right) = 1 - \left[1 - \theta \ln\left[1 - F_{ij}\left(t|x_k^{ij}\right)\right]\right]^{-\frac{1}{\theta}} \tag{19}$$

where $R_{ij}\left(t|x_k^{ij}\right)$ can be obtained using Equation (10) for a Weibull distribution model.

3.2 Data collection and extraction

In this case study, data are extracted and analysed from 73 interruption reports from electric power distribution companies, reported from 2013 to 2016, after four extreme weather events, namely "Hilde", "Ivar", "Tor", and "Nina". This data is partly sensitive, and the reports are not publicly available. Through an agreement with the regulator, the authors of this study were granted access to data from six extreme weather events. However, due to inconsistency in the reporting procedure, only four of the events were selected for further analysis. Moreover, the four weather events selected have quite similar characteristics, which is believed to be an advantage when comparing the recovery processes. The four events are described below.

- *Hilde:* 'Hilde' took place on January 16-17, 2013, with wind speed corresponding to violent storm, and with hurricane force for shorter periods, affecting the area between Trondheim and Bodø. The strength of the weather peaked in the evening and, at 3 am, on January 17, the extreme weather situation was considered over. Approximately 83,000 end users experienced interruptions during the event, while 27,674 customers had their power supply recovered within one hour. The total economic consequence of the event, including KILE costs, was estimated at NOK 51 million. In total, around 400 persons were involved in the short-term recovery process. Only four grid operator companies were affected by this event. However, it should be noted that these companies cover large areas of Norway.
- *Ivar:* The extreme weather 'Ivar' struck middle parts of Norway in the afternoon of December 12, 2013. A low pressure moved in from Great Britain and hit Trøndelag County and Møre og Romsdal County, with wind speed corresponding to violent storm and hurricane. The extreme period of the

weather lasted for a relatively short period, ending after six hours at 9 pm in the evening. Approximately 110,000 end users were affected by the weather, of which 81,000 experienced interruptions of over one hour, and 29,000 had an outage lasting for more than 12 hours. The total economic consequence was estimated at NOK 93 million, and around 630 persons were involved in the short-term recovery process.

- *Tor:* This weather event took place on January 29, 2016. It moved in from the North Sea and first hit the southern parts of Norway and then moved northward to Nordland County. A wind of hurricane strength was measured in several places, with a maximum speed of 48.9 m/s. The severity of the weather declined during the night and, from the morning of January 30, the wind strength was no longer characterised as extreme. Approximately 180,000 outages were registered, of which only 1000 were longer than 24 hours. In total, 150,000 customers were affected by interruptions over the course of the event, some of which experienced several outages. The total damage caused by the event was estimated at NOK 41 million, and more than 800 persons were involved in the short-term recovery. A total of 37 grid companies were affected by this event.

- *Nina:* The extreme weather event 'Nina' struck south-western parts of Norway on January 10, 2015. According to The Norwegian Meteorological Institute (MET) (2015), Nina was one of the five strongest storms registered in Norway during the last 60-70 years. The storm affected large parts of southern Norway, including the urban areas around Oslo. The extreme period of the weather lasted for almost 12 hours. In total, 250,000 end users experienced interruptions during the event, of which 40% had their power supply recovered within one hour, while over 100,000 end users were without power for more than 12 hours. The total damage caused by the event was estimated at NOK 175 million, while 927 persons were involved in the short-term recovery process.
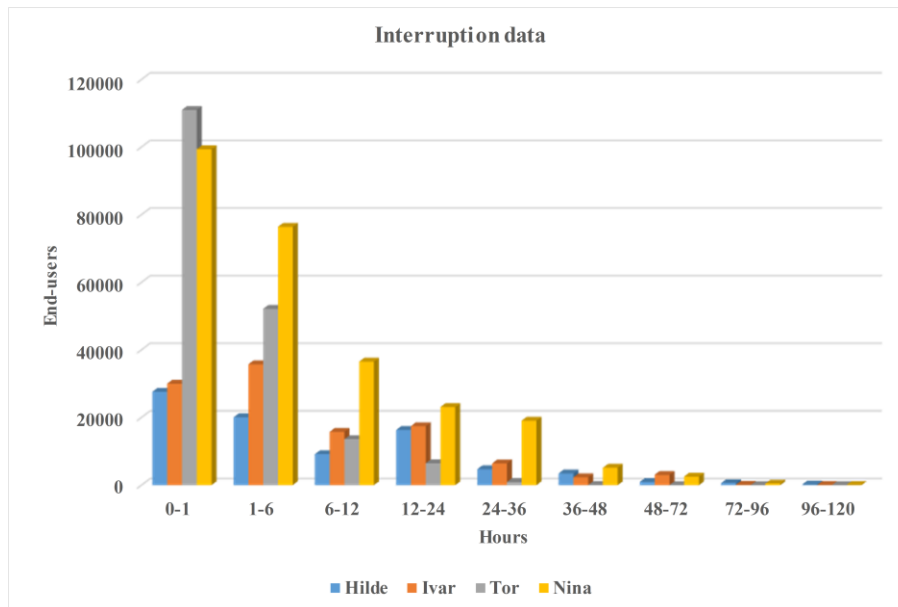

Figure 2. Chart showing number of outages in given time intervals for four different extreme weather events**.**

As discussed, in the case of interruption in the distribution companies' grids, they are obliged to report this in the FASIT-system. Moreover, each grid company affected by an extreme event, such as the events described above, is committed to deliver written reports to NVE. The data set used in this case study is based on such reports, and a brief description of the content of these reports is provided in the next paragraph. A summary of the reported data is found in Table 1.

Table 1. Summary of reported data.

Report metric	Sub-categories/metric	Description
County		19 Norwegian counties
Time and date		Time and date of impact
Place	City, urban and rural	Description of place (more than one is possible)
Natural conditions causing failures	Lightning, precipitation/flooding, Vegetation/trees, wind, salting, avalanche, pollution, fire, birds/animals	Qualitative description of the natural conditions that caused failures
Technical failures	Wear, mechanical failure, heat, electrical failure, fatigue, corrosion	Qualitative description of the types of technical failures
Number of persons involved in the recovery process	Internal employees, external entrepreneurs, landowners, other resources	Operators points out the number of persons involved in the recovery process, divided in four categories.
Costs	Production loss, material costs, KILE-costs, labor costs, compensation costs, other.	Operator estimate the cost out the outages caused by the storm, divided in six categories.
Damage in the grid	Transmission grid , regional grid, distribution grid	Operator states which objects in the grid that is affected, and at which voltage level.
Stations affected	Transmission grid , regional grid, distribution grid	Operator states stations that are damaged (transformation station or connecting station),
Customers without power supply	0-1 hrs., 1-6 hrs., 6-12 hrs., 12-24 hrs., 24-36 hrs., 36- 48 hrs., 2-3 days, 3-4 days, …7-8 days.	Number of customers without power supply reported in time intervals

The first section of the report includes a general description of the event. This includes information about the time of impact, which could have an effect on the length of the recovery time, e.g. if it is at night, during holidays, on weekdays or weekends. In addition, the companies describe the areas affected, whether city, urban area, and/or countryside. Furthermore, the operator specifies the cause of the disruption, distinguishing between natural causes and technical causes/failures. Next, the total number of employees (not man-hours) involved in the recovery process is reported. The same information is given with regard to external personnel. Moreover, the companies can report what other resources they have had access to, such as boats, helicopters and excavators.

The next section in the report deals with the economic consequences of the event. The companies provide estimates for associated costs, divided into several categories. Then, the companies highlight the affected objects in the system, differentiating between different parts of the system, such as voltage level. The companies are also required to state the number of failures that have occurred in each subsystem. However, the exact number of failures is often missing and, if provided, is usually limited to the total number of failures in the system as a whole. In the last part of the report, the companies report whether they have experienced failures relating to other objects, such as their operational control system and operational centres.

At the end of the quantitative part of the report, the companies provide detailed interruption data in terms of the number of outages in certain intervals, where one outage equates to one end user without power for a given time. The outages are not reported in chronological order, and it should also be noted that the same customers may experience more than one outage during one event, i.e. the sum of the number of outages does not necessarily represent the total number of customers affected. Figure 2 compares these four events by illustrating the interruption data for each event in terms of number of outages.

The last part of the report consists of a set of questions, wherein the operators can carry out some qualitative evaluations. Some key aspects here are how they experienced the communication process during the event (both internal and external), the role of exercises, their assessment of the recovery process, condition monitoring and forest clearance, and the effect of the operational control system in the recovery process.

3.3 Variables for analysis

Table 2 shows the list of covariates and their values used in this study. Due to the quality of the reported data and the limited number of data points (n=73), a few key variables are selected to be included in the analysis. Many of the reports contained incomplete data, and, hence, some report metrics in the reports was excluded for the analysis. The selection of variables was based on a literature review and recommendations from the regulator (NVE).

The variable *Event* is a categorical variable that denotes the extreme weather event that has caused the outages. *Location* is a categorical variable, denoting the location of the affected area. *County* is a categorical variable, which lists the counties of Norway and represents the county wherein the extreme event and, thus, the outage has occurred and been reported. *NaturalCondition* is a categorical variable that represents the natural conditions causing the failures and interruption in the power supply. *FailureRegNet* is a logical variable, describing whether the companies have experienced failures in higher voltage levels or in the regional grid or not (i.e. Yes/No). *FailureStation* is a logical variable, stating whether the companies had failures in stations, which could be both transforming stations and connecting stations, independent of voltage level. *CommunicationQuality* is a categorical variable that represents the quality of communication among the actors and personnel during the recovery process, categorised on three levels: poor, sufficient and good. *FailureControlSystem* is a logical variable, stating whether the companies have experienced any complications in their operation control system, which is an essential system used to localise failure and to

reroute power supply. *Exercises* is a logical variable that refers to whether the companies have performed exercises based on similar scenarios. *TreeFallPercent* is a continuous numerical variable, assigned by the operator, that denotes what percentage of the failures is due to trees falling over or hitting the power lines.

Table 2. Model covariates, selected for further analysis, and their possible values.

Covariate (variable)	Value
Event	1: Tor, 2: Hilde, 3: Ivar, 4: Nina
Location	1: City, 2: Urban, 3: Countryside, 4: City and urban, 5: Urban and countryside, 6: City and countryside, 7: City, countryside, urban
County	1: Finnmark, 2: Troms, 3: Nordland, 4: Nord-Trøndelag, 5: Sør-Trøndelag, 6: Møre og Romsdal, 7: Sogn og Fjordane, 8: Hordaland, 9: Rogaland, 10: Vest-Agder, 11: Øst-Agder, 12: Telemark, 13: Vestfold, 14: Buskerud, 15: Akershus, 16: Oslo, 17: Østfold, 18: Oppland, 19: Hedmark, 20: Oppland and Hedmark, 21: Hordaland and Rogaland, 22: Vest-Agder og Øst-Agder
Natural condition	1: Wind, 2: Trees/vegetation, 3: Salt, 4: Snow/ice, 5: Wet soil/ground, 6: Lightning, 7: Precipitation, 8: Avalanche, 9: Wind, trees/vegetation and lightning, 10: Wind and trees/vegetation, 11: Salt and lightning, 12: Wind and salt, 13: Wind and snow/ice, 14: Snow/ice and precipitation, 15: Wind, trees/vegetation, salt, snow/ice, lightning, 16: Wind, trees/vegetation, salt, snow/ice, 17: Wind, trees/vegetation, salt, snow/ice, lightning, avalanche, 18: Wind, trees/vegetation, lightning, 19: Wind, snow/ice, precipitation, 20: Wind, salt, avalanche, 21: Wind, trees/vegetation, salt, 22: Wind, trees/vegetation, wet soil/ground, 23: Lightning, precipitation, trees/vegetation, wind, salt, snow/ice, 24: Wind, trees/vegetation, lightning, snow/ice, 25: Wind, lightning, precipitation, salt, 26: Wind, snow/ice, trees/vegetation
FailureRegNet	0: No, 1: Yes
FailureStation	0: No, 1: Yes
CommunicationQuality	1: Poor, 2: Sufficient, 3: Good
FailureControlSystem	0: No, 1: Yes
Exercises	0: No, 1: Yes
TreeFallPercent	0% - 100%

4 RESULTS AND DISCUSSION

In order to identify the impact of different covariates on the recovery rate and recoverability of the power grid, the recoverability function was developed using AFT models, as described in the Data and Methodology section. It should be noted that the accuracy of the developed models and the range of model parameters depend, among other factors, on the number of available observations or data points and, thus, degree of freedom (Nisbet et al. 2009). According to several runs of different combinations of covariates, Table 2 presents the final model covariates that are selected for further analysis in this study.

Stata software was used to estimate the coefficients. For this purpose, the Weibull distribution was used as the underlying distribution, due to its flexibility in representing different recovery rates, including constant, increasing and decreasing. Using the list of covariates presented in Table 2, and by assuming a Weibull distribution and a Gamma-distributed shared frailty, the model was run. The results are shown in Table 3. Stata uses a maximum likelihood estimation approach to estimate the model coefficients. The

statistical significance of the coefficients can be evaluated by comparing the reported p-values (see Table 3) for each coefficient against a pre-defined threshold, which is usually taken as 0.05. By considering a threshold of 0.05 for p-value, one can analyse which parameter has a significant effect on power distribution system recoverability and its recovery rate. In general, if the p-value is less than 0.05, the null hypothesis, which says that the covariate has no significant effect, will be rejected in favour of the alternative hypothesis, which says the identified covariates have a significant effect on the recoverability. For example, as presented in Table 3, the covariate *Event* (i.e., 1: Tor, 2: Hilde, 3: Ivar, 4: Nina, with 2 being the base value) has p-values equal to 0.22, 0.7 and 0.369 for Tor, Ivar, and Nina respectively. Hence, it can be concluded that the covariate *Event* has no significant effect on the grid's recoverability. In other words, there are no significant differences between these events, and all of them have more and less the same effect on the grid's recoverability. The insignificant effect of the variable *Event* indicates that the recovery rate, recoverability and, thus, the expected recovery time are statistically independent of the type of the event, which is a valid point, as these storms took place during December and January, two months associated with very similar atmospheric and oceanographic conditions in Norway.

However, one should note that the significance level and the extent of the effects of covariates on grid recoverability, which are estimated in this study, to a great extent depend on the number of data points, which is 73. In general, the collected data should represent the real conditions. Here, according to the expert, the polar nights in northern areas could affect visibility for the recovery crew. However, there are only two incidents associated with a northern area county (Troms), one of which took place in the city area, where accessibility time could have been shorter. Hence, any interpretation of the results should be carried out with caution. To obtain more precise results, more accidents in the areas should be included in the database.

Regarding the impact of natural conditions ($NaturalCondition$) on recoverability, as presented in Table 3, only some of the conditions have a statistically significant effect, including wind and tree/vegetation, wind/snow-ice/wet-soil, lightning. However, these conditions are related to the cause of power cuts and, thus, might vary during the recovery phase. Moreover, the fact of whether the companies experienced failures in higher voltage levels and/or in the regional grid or not (i.e., $FailureRegNet$ = 0:No or 1:Yes), as well as the fact that the companies had failures in stations, which could be both transforming stations and connecting stations, independent of voltage level (i.e., $FailureStation$ = 0:No or 1:Yes) has a noticeable and significant impact on recoverability. The same argument holds for the covariate *Exercises*. However, although one expects to notice significant differences in recoverability, in terms of the quality of communication, the analysis results in this study using available data do not suggest any significant correlation. This could be due to either the lack of field data or to the recovery process in general not being very sensitive to the quality of communication among actors.

After identifying the covariates which have significate effect on the recoverability of the grids, the important question which should be answered is: how much these covariates will affect the recoverability of the grids? By having the magnitude of covariates, the future planning will be much effective. In Table 3, the column "coef." shows the regression coefficient of identified covariates. It shows the change in recovery

rate due to the identified covariates. These numbers provide essential input for improving the future recovery process. For example, for location, we will find that location no. 2, which represents *urban area,* with p-value equal to 0.037, has a significant effect grid recoverability.

Regarding the shared frailty and the presence and impact of unobserved covariates, as presented in Table 3, it can be seen that the p-value for the likelihood-ratio test of the hypothesis $\theta = 0$ is 0.283, indicating that the unobserved heterogeneity is negligible. This means that the collected covariates fully reflect the real conditions under which the recovery process is taking place.

However, to illustrate the importance of always testing the impact of unobserved heterogeneity, the model was run, but this time the covariate *Exercises* was excluded from the analysis. The results are presented in Table 4. As shown, the p-value for likelihood-ratio test of the hypothesis $\theta = 0$ is 0.002, indicating the presence of unobserved heterogeneity in the model. Or it tells that there is one or more unobserved covariate (here, *Exercises*), which needs to be considered during the future planning. Moreover, a comparison between Table 3 and Table 4 shows that the regression coefficients are changed significantly, for example in no. 22 *Natural condition* (wind, trees/vegetation). The regression coefficient is changed from -2.5 to -2.7. When this situation arises, the analyst needs to review the recovery process carefully, to identify all possible missing covariates for consideration in future analysis.

The developed model has a high potential to quantify the effect of observed and unobserved covariates. However, the most available data are not collected for this type of analysis, which make its application a challenging task. For example, in this case study, the original interruption reports that the companies must complete and report to the FASIT-system contain more information than that listed in Table 2 as covariates. However, plugging all the provided information into the model, using only 73 data points, led to a high degree of freedom and, thus, to a non-converging solution. This computational issue could have been fixed by collecting a sufficiently large amount of data, which is one of the limitations in the current study. For this purpose, the model was constructed using only a number of important covariates that are expected to have significant effects on recoverability. However, given the amount of collected data and the number of model covariates, it is expected to have significant effect on any unobserved covariate. Nonetheless, the unobserved covariate effect was shown to have a significant impact on recoverability, once a covariate was deliberately removed from the list of model covariates.

Another important factor to keep in mind while analysing the results provided in Table 3 is significance level, which is indeed dependent on the number of covariates and amount of available data. In other words, the statistical interpretation of the model and identification of the range of influencing parameters, as well as the extent of their effects, depends, to a great extent, on the number of covariates used in the model and the amount of available data.

Table 3. Model coefficient with covariates listed in Table 2.

_t	Coef.	Std. Err.	z	P>|z|	[95% Conf. Interval]	
Event						
1	1.028046	.83781	1.23	0.220	-.6140316	2.670123
3	.2861738	.7430107	0.39	0.700	-1.1701	1.742448
4	.83674	.9128939	0.92	0.359	-.9524992	2.625979
Location						
2	-1.53093	.7321808	-2.09	0.037	-2.965978	-.0958822
3	-.7558757	.6309408	-1.20	0.231	-1.992497	.4807455
5	-.8686251	.6343052	-1.37	0.171	-2.11184	.3745902
6	-1.648443	.9556707	-1.72	0.085	-3.521524	.2246366
7	-.6629752	.759607	-0.87	0.383	-2.151778	.8258271
County						
4	-2.215043	.9404198	-2.36	0.019	-4.058232	-.371854
5	-3.647929	.7092393	-5.14	0.000	-5.038013	-2.257845
6	-3.04726	.869635	-3.50	0.000	-4.751713	-1.342806
7	-3.270747	.859844	-3.80	0.000	-4.95601	-1.585484
8	-2.861533	.9030273	-3.17	0.002	-4.631434	-1.091632
9	-2.414826	.9768403	-2.47	0.013	-4.329398	-.5002547
12	-2.613753	1.013233	-2.58	0.010	-4.599653	-.6278533
14	-4.209499	.8687938	-4.85	0.000	-5.912303	-2.506694
18	-4.392309	.9380548	-4.68	0.000	-6.230862	-2.553755
20	-4.360538	.923181	-4.72	0.000	-6.16994	-2.551137
21	-2.555203	.9320204	-2.74	0.006	-4.381929	-.7284766
22	-3.281713	1.164456	-2.82	0.005	-5.564004	-.9994222
NaturalCondition						
2	-1.497308	.543355	-2.76	0.006	-2.562265	-.4323521
6	.0029784	.5076176	0.01	0.995	-.9919338	.9978907
9	-.6075864	.3337785	-1.82	0.069	-1.26178	.0466075
10	-.0465679	.2305315	-0.20	0.840	-.4984013	.4052656
11	1.77619	.5499074	3.23	0.001	.6983917	2.853989
12	-.6970775	.6481872	-1.08	0.282	-1.967501	.573346
13	-1.231593	.3807633	-3.23	0.001	-1.977875	-.4853104
14	-.0729835	.5094491	-0.14	0.886	-1.071486	.9255184
15	.245784	.5220478	0.47	0.638	-.7774109	1.268979
16	.6299441	.5508856	1.14	0.253	-.4497718	1.70966
17	.1059363	.6819122	0.16	0.877	-1.230587	1.44246
18	-.2882191	.674657	-0.43	0.669	-1.610522	1.034084
19	-1.02575	.8175962	-1.25	0.210	-2.628209	.5767095
21	.0220111	.2890767	0.08	0.939	-.5445688	.588591
22	-2.753224	.7549168	-3.65	0.000	-4.232834	-1.273614
23	0	(omitted)				
24	.6753564	.6864794	0.98	0.325	-.6701184	2.020831
25	0	(omitted)				
26	-.2486999	.5550894	-0.45	0.654	-1.336655	.8392552
1.FailureRegNet	-.5960308	.2394162	-2.49	0.013	-1.065278	-.1267836
1.FailureStation	-.4786911	.1862439	-2.57	0.010	-.8437223	-.1136598
CommunicationQuality						
2	-1.20731	1.167934	-1.03	0.301	-3.496418	1.081798
3	-1.12274	1.093496	-1.03	0.305	-3.265953	1.020474
1.FailureControlSystem	.5230183	.2092241	2.50	0.012	.1129466	.9330899
1.Exercises	.8454695	.1921927	4.40	0.000	.4687789	1.22216
TreeFallPercent	.0023817	.0030071	0.79	0.428	-.0035122	.0082755
_cons	4.675124	1.35377	3.45	0.001	2.021783	7.328465
/ln_p	1.299625	.2037711	6.38	0.000	.9002406	1.699009
/lntheta	-.2661752	.5901616	-0.45	0.652	-1.422871	.8905203
p	3.66792	.7474162			2.460195	5.468524
1/p	.2726341	.055555			.1828647	.4064718
theta	.7663048	.4522437			.2410211	2.436397

LR test of theta=0: chibar2(01) = 0.33   Prob >= chibar2 = 0.283

Table 4. Model coefficient with covariates listed in Table 2 and excluding the covariate Exercises.

_t	Coef.	Std. Err.	z	P>|z|	[95% Conf. Interval]	
Event						
1	1.408279	.7584623	1.86	0.063	-.0782794	2.894838
3	.1467681	.7037525	0.21	0.835	-1.232561	1.526098
4	1.004677	.8484019	1.18	0.236	-.6581601	2.667514
Location						
2	-1.67496	.818511	-2.05	0.041	-3.279212	-.0707079
3	-1.083507	.7077462	-1.53	0.126	-2.470664	.3036495
5	-1.355051	.6894708	-1.97	0.049	-2.706389	-.0037134
6	-1.913905	1.021022	-1.87	0.061	-3.915072	.0872618
7	-1.067815	.7663694	-1.39	0.164	-2.569871	.4342416
County						
4	-2.456754	1.010048	-2.43	0.015	-4.436412	-.4770969
5	-3.638092	.8713631	-4.18	0.000	-5.345933	-1.930252
6	-3.771002	.9496189	-3.97	0.000	-5.632221	-1.909783
7	-3.617453	.93747	-3.86	0.000	-5.454861	-1.780046
8	-3.46762	1.004057	-3.45	0.001	-5.435537	-1.499704
9	-2.764427	1.088997	-2.54	0.011	-4.898823	-.6300315
12	-3.729887	1.184112	-3.15	0.002	-6.050703	-1.40907
14	-4.95731	.9731571	-5.09	0.000	-6.864663	-3.049957
18	-5.163519	1.057567	-4.88	0.000	-7.236312	-3.090726
20	-4.821263	1.049337	-4.59	0.000	-6.877926	-2.7646
21	-2.712613	1.051376	-2.58	0.010	-4.773272	-.6519536
22	-3.565478	1.18324	-3.01	0.003	-5.884586	-1.246371
NaturalCondition						
2	-1.360661	.6946758	-1.96	0.050	-2.7222	.0008788
6	-.5779705	.5091871	-1.14	0.256	-1.575959	.4200179
9	-.2821037	.3797377	-0.74	0.458	-1.026376	.4621685
10	-.1403828	.2715039	-0.52	0.605	-.6725207	.3917552
11	1.468806	.6152645	2.39	0.017	.2629099	2.674703
12	-.2514787	.6694731	-0.38	0.707	-1.563622	1.060664
13	-.6702752	.4281	-1.57	0.117	-1.509336	.1687853
14	-.946001	.5745301	-1.65	0.100	-2.072059	.1800572
15	.5126769	.5667224	0.90	0.366	-.5980786	1.623432
16	.4691369	.623665	0.75	0.452	-.7532241	1.691498
17	.3588291	.6742763	0.53	0.595	-.9627281	1.680386
18	-.1029516	.6797526	-0.15	0.880	-1.435242	1.229339
19	-1.585196	.8663967	-1.83	0.067	-3.283302	.1129105
21	.1677997	.3616901	0.46	0.643	-.5410998	.8766992
22	-2.53772	.7561812	-3.36	0.001	-4.019808	-1.055632
23	0	(omitted)				
24	.7953579	.7741048	1.03	0.304	-.7218596	2.312575
25	0	(omitted)				
26	.5187389	.5932845	0.87	0.382	-.6440772	1.681555
1.FailureRegNet	-.5319109	.2872808	-1.85	0.064	-1.094971	.0311492
1.FailureStation	-.3745567	.2311232	-1.62	0.105	-.8275498	.0784365
CommunicationQuality						
2	-1.536637	1.091337	-1.41	0.159	-3.675617	.6023438
3	-1.380925	.9910387	-1.39	0.163	-3.323325	.5614749
1.FailureControlSystem	.6595764	.2465005	2.68	0.007	.1764442	1.142709
TreeFallPercent	.0072869	.0035805	2.04	0.042	.0002693	.0143046
_cons	5.718609	1.508803	3.79	0.000	2.761409	8.67581
/ln_p	1.310773	.2065564	6.35	0.000	.90593	1.715616
/lntheta	.101794	.4954553	0.21	0.837	-.8692805	1.072869
p	3.70904	.7661257			2.474232	5.5601
1/p	.2696116	.05569			.1798529	.4041658
theta	1.107155	.548546			.4192531	2.923755

LR test of theta=0: chibar2(01) = 8.69   Prob >= chibar2 = 0.002

In order to analyse the type of effects that each covariate and its corresponding values have on recoverability or recovery rate, one can analyse the recovery rate model which is given by Equation (11). To this aim, a recovery time for each outage, as the dependent random variable, T_R, should be known prior to running the model and estimating the coefficients. In this study, the recovery time for each outage is a weighted averaging of the duration of outages and their corresponding number of end users, reported for each incident. In other words, Let $C = \{C_1, C_2, \dots, C_m\}$ be the number of end users, corresponding to each incident, that experienced a power cut for a specific period of time, denoted by I_k, $k = 1,2,\dots,8$, where I, a vector of time intervals in hours, is given by Equation (20):

$$I = \{I_1, I_2, \dots, I_8\} = \{[1 - 6], [6 - 12], [12 - 24], [24 - 36], [36 - 48], [48 - 72], [72 - 96], [96,120]\} \text{ h} \tag{20}$$

Then, the recovery time (i.e., the duration of outage) for each reported incident can be calculated using Equation (21):

$$T_R = \frac{\sum_{k=1}^{8} C_k i_k}{\sum_{k=1}^{8} C_k} \tag{21}$$

where T_R is the recovery time of the reported incident, and i_k is the average of the outage time interval I_k.

The estimated coefficients presented in Table 3 can be used to analyse the effect of covariates on recoverability. To this aim, one can expand Equation (11), as:

$$r(t|x_k) = \beta t^{\beta-1} \exp[-\beta(AX)] \tag{22}$$

$$r(t|x_k) = \beta t^{\beta-1} \exp[-\beta(x_0 + x_1 a_1 + x_2 a_2 + \cdots)] \tag{23}$$

While parameter β in Equation (23) is the Weibull shape factor, given as $p = 3.66792$ in Table 3, the term x_0 is the constant term given in Table 3, $x_0 = 3.55204$, x_k is the covariate value, and a_k is the corresponding coefficient given in Table 3. Given the form of Equation (23), if the coefficient of a covariate is negative, it will increase the recovery rate. In other words, a covariate with a negative sign improves the recoverability of the system. This means, at a given time, the probability that the system has recovered will increase. Similarly, a positive sign for a covariate means that the system will recover with a lower probability, i.e., the system recovery rate and recoverability are reduced for those covariates with a positive sign.

For example, let us consider the covariate $Location$, which can have seven different values, as given in Table 2, i.e., $Location$: 1: City, 2: Urban, 3: Countryside, 4: City and urban, 5: Urban and countryside, 6: City and countryside, 7: City, countryside, urban. As presented in Table 3, all values of $Location$ have an insignificant coefficient, except "2: Urban". Note that, in the data set, there is no incident reported relating to "4: City and urban", and the base value for $Location$ is set "1: City" by default by Stata, as it is the minimum value of the covariate $Location$. According to Equation (23), the coefficient -1.53093, corresponding to $Location$ 2: Urban, states that the recovery rate increases if the recovery process is taking places in urban areas compared to the base case, which is City.

As another example, as given in Table 3, the coefficient of $FailureRegNet$ 1: Yes is -0.5960308, which states that the recovery rate is higher for those cases where the companies experienced failures in

higher voltage levels or in the regional grid. The negative coefficient of *FailureStation* 1: Yes is -0.4786911, which states that the recovery rate is higher for those incidents where the failure has occurred in stations (transforming stations and/or connecting stations).

According to Equation (23), a positive coefficient will reduce the recovery rate and thus recoverability. For instance, let us consider the covariate *FailureControlSystem* 1: Yes, which states that companies have experienced complications in their operation control system, which is an essential system used to localise failure and to reroute power supply. According to Table 3, the positive sign of the coefficient of *FailureControlSystem* 1: Yes indicates that the recovery rate will be lower, compared to the base case where no complications in companies' operation control systems is experienced.

From a management perspective, the findings in this study can help the operator to take the necessary measures to minimize the impact of future events. For instance, trees/vegetation had a significant influence on recoverability, and thus the implementation of a better forest clearance programme could be an effective preventive measure. Where the level of accessibility is low, the most critical parts of the grid should be prioritized, in order to avoid future disturbances due to trees and vegetation hitting the power lines. Moreover, having a robust operational control system is essential in a recovery process. Many of the companies have experienced failure in those systems. Improving the reliability of such systems, for instance by adding redundant fibre lines, might increase the possibility of rerouting the power supply and, hence, increasing recoverability.

As illustrated by the above-mentioned discussion, the significance level and the estimated coefficients can be used to compare the recovery rate and, thus, recoverability for certain scenarios that involve the listed covariates. A quantitative value for the extent of the effects of the covariates on recoverability can be estimated, using Equation (7), which is beyond the scope of this study. Once such a quantitative evaluation is performed, a probabilistic risk assessment can be performed, in order to find the bottlenecks of the recovery process for budget allocations or possible improvements. Some studies (e.g., see Hasan et al. (2013)) validate the developed AFT models, by dividing the data into several groups and running the model for each group of the data. Some statistical tests are further performed to see if the estimated coefficients for each group of data are statistically equal. However, such an approach requires a huge amount of data, which was a limitation in this study. Nevertheless, the present study's probabilistic model has been developed based on principles of AFT models which are widely used and acceptable in duration analysis. In this regard, although the developed mathematical framework is acceptable, the results should be used with caution until additional data is collected in order to statistically validate the model.

5 CONCLUSION

Prediction of the recovery time of disrupted infrastructures provides us with essential inputs for developing an effective contingency plan when making important decisions in the recovery phase, in order to minimize the impact of the disruptive events. Recovery processes are complex tasks, and there are many factors that can affect such processes, including operational, environmental, organizational, as well as human, factors.

In addition to these observed risk factors, there are always other factors which the analysts may not identify or about which there might not be sufficient information at the time of the analysis. The common practice in most studies is to neglect the effects of such factors on the recovery process. Such a practice biases the analysis results and, consequently, increases the uncertainties associated with the effectiveness of any contingency plan. Here, in this paper, we extend the application of AFT models, which are used frequently in reliability engineering, in order to model the effects of observed and unobserved risk factors on the recoverability of disrupted infrastructures. The model is applied to Norwegian electric power distribution systems facing extreme weather events. In such a modelling framework, the infrastructure's recoverability is modelled as a function of time, observed and unobserved risk factors.

The developed statistical model is applied to 73 reports on power outage in the electrical power distribution grid in Norway. The model is used to model the recoverability as a function of time and some influencing parameters. The results from the case study indicate how the impact of covariates on recovery rate and recoverability can be analysed. Certain covariates increase the recovery rate and improve recoverability, or, in other words, the recovery rate is higher under certain scenarios than others. It is indicated that the recovery rate is higher if the failure has occurred in the regional grids, which is not intuitive. The reason for the increase in recovery rate is most likely due to the fact that failures in the regional grid affects more customers. Hence, if failures have occurred in the regional grid, it is likely that the repair rate will be higher, since a successful repair will have influence on a high amount of customers. As another example, the recovery rate is lower if the companies have experienced complications in their operation control system. It is crucial to have operational control system that is working when localizing failures. In addition, such systems also gives the operator the opportunity for rerouting and isolation of failures, decreasing the impact of the failure. It is also seen that the covariates $County$, $Exercises$, $FailureStation$, have a significant effect on recovery rate. The geographical areas where these four events took place have quite different characteristics, which is believed to influence the trajectory of the recovery process.

While some covariates have a significant effect on recovery rate, some other covariates, including $Event$ and $CommunicationQuality$, do not significantly influence the recovery rate. An explanation of the former could be that these four events have quite similar weather characteristics. It is a bit surprising that communication quality does influence the recovery process significantly, but it should be noted that almost all companies reported that the quality of the communication was good.

In this regard, the results of the model analysis can be used to identify the parameters affecting the recovery process of infrastructure systems, providing the operator and regulator of the infrastructure with valuable information to improve both the technical systems and the organisational aspects of the infrastructure, in order to enhance the resilience level of the sociotechnical systems as a whole. In that way, they are better prepared for future events. The studied data indicates that not all companies utilise the possibility to learn from previous disruptions. Moreover, it is clear that missing information is an ongoing challenge for the regulator and the operator of the Norwegian electric power distribution systems. Although stakeholders in Norway use the same reporting form, the level of detail provided varies. It is evident that,

in order to fully utilise the presented modelling approach and have statistically significant results, a comprehensive amount of data will be required.

6 LIMITATIONS AND FUTURE RESEARCH

The authors acknowledge that validation is a critical part of every modelling effort. The probabilistic model in this study is developed based on AFT models which are widely used approaches in duration analysis for analysing the impact of some influencing covariates on independent random variables. Validating the developed model will require many observations. One possibility could be to divide the data into two groups, and the model is then run for either of those groups, in order to estimate the model coefficients. A specification test can further be performed, to check whether or not the estimated coefficients corresponding to these two groups of data are statistically equal. However, one of the main limitations in the current study was lack of a large database. In this regard, the results of the current study should be used with caution. By using more data in the future, the model can be run again and validated. However, the findings in the study are consistent with the evaluations performed by the regulator (see Norwegian Water Resources and Energy Directorate (2013a, 2013b, 2015, 2017)).

In addition to the scarcity of data, there are also some limitations in the data collection procedure, which influence the quality of the data and the analysis results. The collected data are based on reports from the grid operators. The structure of the reports and the way the questions are formulated leave considerable room for subjective evaluations by the operator. Moreover, as some of the raw data are qualitative and descriptive, the results will, to some extent, depend on the author's interpretation in the data extraction process. The vagueness in the data will, of course, influence the reliability of the analysis.

It should be mentioned that the reports are sent from the operator to the regulating authority. Operators might try to protect their own reputation and business. Hence, there might be circumstances where the operator paints a different picture from the reality. This discussion highlights the need for a better data collection and power cut reporting procedure, which is a research gap that needs to be filled in the future. A national digital fault registration system already exists, but the data concerning influencing factors are not fully captured at the moment. In addition, to make the data more reliable, one option could be to anonymize some of the data.

ACKNOWLEDGMENTS

NOTATION LIST

The following symbols are used in this paper:

AFT	Accelerated failure time
FASIT	Fault and Supply Interruption information Tool
NVE	Norwegian Water Resources and Energy Directorate
PH	Proportional hazard
TSA	Transmission System Operators
A	Coefficient vector
α_k	Covariate coefficient
$f(t)$	Probability density function
$F(t)$	Cumulative distribution function
$r(t)$	Recovery rate, which is the probability that the recovery is completed in the time interval $(t, t + \Delta t]$, when it is known that it has not been finished at time t
$R(t)$	Recoverability of a power distribution grid, described as the probability that the system can be recovered from a disruption event before time t
$S(t)$	Survival function. In the current study's scope, $S(t)$ is the probability that the recovery cannot be completed before some specified time t
x_k	Covariate
X	Covariate vector
β	Shape factor of a Weibull distribution
ε_i	The frailty of the ith group, which is a positive random number with mean equal to 1, variance θ, and the probability density function $g(\varepsilon_i)$.

DATA AVAILABILITY

Some data, models, or code generated or used during the study are available from the corresponding author by request:

- Extracted and coded data from the 73 interruptions reports
- Stata software analysis code

Some data, models, or code generated or used during the study are proprietary or confidential in nature and may only be provided with restrictions (e.g. anonymised data):

- The interruption reports completed by the distribution companies, as they include sensitive information. Data from these reports can, to a large extent, be provided in coded/anonymised form (as indicated above).

REFERENCES

Afghari, A. P., S. Washington, C. Prato, and M.M. Haque. 2019. Contrasting case-wise deletion with multiple imputation and latent variable approaches to dealing with missing observations in count regression models. Analytic Methods in Accident Research, 24, 100104.

Alvehag, K., and L. Soder. 2008. "A stochastic weather dependent reliability model for distribution systems." In *Proc., 10th International Conference on Probabilistic Methods Applied to Power Systems (PMAPS08), 25-29 May 2008, Rincon, Puerto Rico*. IEEE, 1-8.

Alvehag, K., and L. Soder. 2010. "A reliability model for distribution systems incorporating seasonal variations in severe weather." *IEEE Transactions on Power Delivery*, 26, 910–919.

Ansell, J., and M. Philipps. 1997. "Practical aspects of modelling of repairable systems data using proportional hazards models." *Reliability Engineering & System Safety*, 58 (2), 165–171.

Bacon, P., and C. Hobson. 2014. *Human security and Japan's triple disaster: Responding to the 2011 earthquake, tsunami and Fukushima nuclear crisis*. London and New York: Routledge.

Bagdonavicius, V., and M. Nikulin. 2001. *Accelerated life models: modeling and statistical analysis*. Chapman and Hall/CRC. Bakhshi, E., R.A.A. Khoei, A. Azarkeivan, M. Kooshesh, and A. Biglarian. 2017. "Survival analysis of thalassemia major patients using Cox, Gompertz proportional hazard and Weibull accelerated failure time models." *Medical journal of the Islamic Republic of Iran,* 31, 97.

Bakhshi, E., R.A.A. Khoei, A. Azarkeivan, M. Kooshesh, and A. Biglarian. 2017. "Survival analysis of thalassemia major patients using Cox, Gompertz proportional hazard and Weibull accelerated failure time models." Medical journal of the Islamic Republic of Iran, 31, 97.

Barabadi, A. 2014. "Reliability analysis of offshore production facilities under Arctic conditions using reliability data from other areas." *Journal of Offshore Mechanics and Arctic Engineering*, 136 (2), 021601.

Barabadi, A., J. Barabady, and T. Markeset. 2011. "Maintainability analysis considering time-dependent and time-independent covariates." *Reliability Engineering & System Safety*, 96 (1), 210–217.

Blanchard, B. S., D. C. Verma, and E. L. Peterson. 1995. *Maintainability: a key to effective serviceability and maintenance management*. John Wiley & Sons.

Boin, A., and A. McConnell. 2007. "Preparing for critical infrastructure breakdowns: the limits of crisis management and the need for resilience." *Journal of Contingencies and Crisis Management*, 15 (1), 50–59.

Bollinger, L. A., and G. P. Dijkema. 2016. "Evaluating infrastructure resilience to extreme weather–the case of the Dutch electricity transmission network." *European Journal of Transport and Infrastructure Research*, 16 (1).

Bradburn, M. J., T. G. Clark, S. Love, and D. Altman. 2003. "Survival analysis part II: multivariate data analysis–an introduction to concepts and methods." *British Journal of Cancer*, 89 (3), 431.

Bruneau, M., S. E. Chang, R. T. Eguchi, G. C. Lee, T. D. O'Rourke, A. M. Reinhorn, M. Shinozuka, K. Tierney, W. A. Wallace, and D. Von Winterfeldt. 2003. "A framework to quantitatively assess and enhance the seismic resilience of communities." *Earthquake Spectra*, 19 (4), 733–752.

Burnard, K., and R. Bhamra. 2011. "Organisational resilience: development of a conceptual framework for organisational responses." *International Journal of Production Research*, 49 (18), 5581–5599.

Çağnan, Z., R. A. Davidson, and S. D. Guikema. 2006. "Post-earthquake restoration planning for Los Angeles electric power." *Earthquake Spectra*, 22 (3), 589–608.

Choi, J., N. Naderpajouh, D.J. Yu, and M. Hastak. 2019. "Capacity Building for an Infrastructure System in Case of Disaster Using the System's Associated Social and Technical Components." *Journal of Management in Engineering,* 35(4), 04019013.

Cimellaro, G. P., D. Solari, and M. Bruneau. 2014. "Physical infrastructure interdependency and regional resilience index after the 2011 Tohoku Earthquake in Japan." *Earthquake Engineering & Structural Dynamics*, 43 (12), 1763–1784.

Comes, T., and B. Van de Walle. 2014. "Measuring disaster resilience: The impact of hurricane Sandy on critical infrastructure systems." *ISCRAM*, 11, 195–204.

Cox, D. R. 2018. *Analysis of survival data*. Routledge.

Crowder, M. J. 2017. *Statistical analysis of reliability data*. Routledge.

Cutter, S. L., L. Barnes, M. Berry, C. Burton, E. Evans, E. Tate, and J. Webb. 2008. "A place-based model for understanding community resilience to natural disasters." *Global Environmental Change*, 18 (4), 598–606.

Dale, C. 1985. "Application of the proportional hazards model in the reliability field." *Reliability Engineering*, 10 (1), 1–14.

Dalziell, E. P., and S. T. McManus. 2004. "Resilience, vulnerability, and adaptive capacity: implications for system performance." International Forum for Engineering Decision Making (IFED), University of Canterbury, Christchurch.

Department of Defense, 1991. Military Handbook MIL-HDBK-217F - Reliability Predition of Electronic Equipment. Department of Defense, Washington D.C

Dessavre, D. G., J. E. Ramirez-Marquez, and K. Barker. 2016. "Multidimensional approach to complex system resilience analysis." *Reliability Engineering & System Safety*, 149, 34–43.

Dhillon, B. S. 1999. *Engineering maintainability: How to design for reliability and easy maintenance.* Gulf Professional Publishing.

Fagbamigbe, A. F., R. F. Afolabi, K. Y. Alade, A.S. Adebowale, and B. O. Yusuf. 2019. "Unobserved Heterogeneity in Determinants of Under-five Mortality in Nigeria: Frailty Modeling in Survival Analysis." *African Journal of Applied Statistics*, 6(1), 565-587.

Faturechi, R., and E. Miller-Hooks. 2014. "Measuring the performance of transportation infrastructure systems in disasters: A comprehensive review." *Journal of Infrastructure Systems*, 21 (1), 04014025.

Fine, J. P., and R. J. Gray. 1999. "A proportional hazards model for the subdistribution of a competing risk." *Journal of the American Statistical Association*, 94 (446), 496–509.

Francis, R., and B. Bekera. 2014. "A metric and frameworks for resilience analysis of engineered and infrastructure systems." *Reliability Engineering & System Safety*, 121, 90–103.

Ghodrati, B., and U. Kumar. 2005a. "Operating environment-based spare parts forecasting and logistics: a case study." *International Journal of Logistics: Research and Applications*, 8 (2), 95–105.

Ghodrati, B., and U. Kumar. 2005b. "Reliability and operating environment-based spare parts estimation approach: A case study in Kiruna Mine, Sweden." *Journal of Quality in Maintenance Engineering.* 11(2), 169-184.

Gutierrez, R. G. 2002. "Parametric frailty and shared frailty survival models." *The Stata Journal*, 2 (1), 22–44.

Haimes, Y. Y. 2012. "Systems-based approach to preparedness for, response to, and recovery from natural and human-made disasters." *Leadership and Management in Engineering*, 12(4), 288-298.

Hanagal, D. D. 2017. Frailty Models in Public Health. In Handbook of Statistics Vol. 37, pp. 209-247. Elsevier.

Hasan, S., R. Mesa-Arango, and S. Ukkusuri. 2013. "A random-parameter hazard-based model to understand household evacuation timing behavior." *Transportation research part C: emerging technologies*, 27, 108-116.

Hatlen, L. M., and K. Knudsen Aarrestad. 2015. "FAKTA Energi- og vannressurser i Norge 2015," Norwegian Ministry of Petroleum and Energy.

Heggset, J., G. Kjolle, and K. Sagen. 2009. "FASIT-A tool for collection, calculation and reporting of reliability data." In *Proc., CIRED 2009-20th International Conference and Exhibition on Electricity Distribution-Part 1*, IET, 1-4.

Hesam, S., M. Mahmoudi, A.R. Foroushani, M. Yaseri, and M. A. Mansournia. 2018." A cause-specific hazard spatial frailty model for competing risks data." *Spatial Statistics,* 26, 101-124.

Honfi, D., D. Lange, A. Malm, P. Mindykowski, M. Alheib, C. Bouffier, L. Cauvin, A. Willot, I. Ioannou, and B. Rød. 2017. "Technological resilience concepts applied to critical infrastructure." IMPROVER project.

Hossain, N. U. I., R. Jaradat, S. Hosseini, M. Marufuzzaman, and R. K. Buchanan. 2019. "A framework for modeling and assessing system resilience using a Bayesian network: A case study of an interdependent electrical infrastructure system." *International Journal of Critical Infrastructure Protection*, 25, 62–83.

Hosseini, S., K. Barker, and J. E. Ramirez-Marquez. 2016. "A review of definitions and measures of system resilience." *Reliability Engineering & System Safety*, 145, 47–61.

Hougaard, P. 1995. "Frailty models for survival data." *Lifetime Data Analysis*, 1 (3), 255–273.

Hougaard, P. 2016. "Modelling heterogeneity in survival data." *Journal of Applied Probability*, 28 (3), 695–701.

Ilbeigi, M. and B. Dilkina. 2017. "Statistical approach to quantifying the destructive impact of natural disasters on petroleum infrastructures." *Journal of Management in Engineering*, 34(1), 04017042.

International Electrotechnical Vocabulary (IEV) 191. 2007. "Chapter 191: dependability and quality of service." Accessed December 3, 2019. <http://std.iec.ch/iec60050 >.

Jardine, A., P. Anderson, and D. Mann. 1987. "Application of the Weibull proportional hazards model to aircraft and marine engine failure data." *Quality and Reliability Engineering International*, 3 (2), 77–82.

Kayrbekova, D., A. Barabadi, and T. Markeset. 2011. "Maintenance cost evaluation of a system to be used in Arctic conditions: a case study." *Journal of Quality in Maintenance Engineering*, 17 (4), 320–336.

Kong, J., S. P. Simonovic, and C. Zhang, C. (2019). "Sequential hazards resilience of interdependent infrastructure system: A case study of Greater Toronto Area energy infrastructure system." *Risk Analysis*, 39(5), 1141-1168.

Kumar, D., and B. Klefsjö. 1994. "Proportional hazards model: a review." *Reliability Engineering & System Safety*, 44 (2), 177–188.

Kumar, D., B. Klefsjö, and U. Kumar. 1992. "Reliability analysis of power transmission cables of electric mine loaders using the proportional hazards model." *Reliability Engineering & System Safety*, 37 (3), 217–222.

Labaka, L., J. Hernantes, and J. M. Sarriegi. 2015. "Resilience framework for critical infrastructures: An empirical study in a nuclear plant." *Reliability Engineering & System Safety*, 141, 92–105.

Lange, D., D. Honfi, J. Sjöström, M. Theocharidou, G. Giannopoulos, N. K. Reitan, K. Storesund, L. Melkunaite, H. Rosenquist, L. Peterson, R. Almeida, B. Rød, C. Bouffier, E. Serafinelli, and M. Lexin Lin. 2017a. "IMPROVER Deliverable 5.1 Framework for implementation of resilience concepts to Critical Infrastructure."

Lange, D., D. Honfi, M. Theocharidou, G. Giannopoulos, N. K. Reitan, and K. Storesund. 2017b. "Incorporation of resilience assessment in critical infrastructure risk assessment frameworks." In *Proc., 27th European Safety and Reliability Conference, ESREL 2017, 18 June 2017 through 22 June 2017*, CRC Press/Balkema, 1031-1038.

Lee, E. T., and J. W. Wang. 2003. *Statistical methods for survival data analysis.* Hoboken, NJ: John Wiley & Sons.

Liu, H., R. A. Davidson, and T. V. Apanasovich. 2007. "Statistical forecasting of electric power restoration times in hurricanes and ice storms." *IEEE Transactions on Power Systems*, 22 (4), 2270–2279.

Mannering, F. L., V. Shankar, and C.R. Bhat. 2016. "Unobserved heterogeneity and the statistical analysis of highway accident data." *Analytic methods in accident research*, 11, 1-16.

Matsuoka, T. 2015. "Unobserved heterogeneity in price-setting behavior: A duration analysis approach." *Japan and the World Economy.* 22(1), 13-20.

McEvoy, D., I. Ahmed, and J. Mullet. 2012. "The impact of the 2009 heat wave on Melbourne's critical infrastructure." *Local Environment,* 17 (8), 783–796.

Mendonça, D., and W. A. Wallace. 2006. "Impacts of the 2001 World Trade Center attack on New York City critical infrastructures." *Journal of Infrastructure Systems*, 12 (4), 260–270. https://doi.org/10.1016/j.cegh.2015.11.008

Ministry of Petroleum and Energy. n.d. "Strømnettet". Accessed September 12, 2018. <https://energifaktanorge.no/norsk-energiforsyning/kraftnett/>.

Mohammadian, A., and S. T. Doherty. 2006. "Modeling activity scheduling time horizon: Duration of time between planning and execution of pre-planned activities." *Transportation Research Part A: Policy and Practice*, 40 (6), 475–490.

Naseri, M. 2017. "On maintainability of winterised plants operating in Arctic regions." In *Proc., ASME 2017 36th International Conference on Ocean, Offshore and Arctic Engineering*, American Society of Mechanical Engineers, V03BT02A018-V003BT002A018.

Naseri, M., and J. Barabady. 2016. "An expert-based approach to production performance analysis of oil and gas facilities considering time-independent Arctic operating conditions." *International Journal of System Assurance Engineering and Management*, 7 (1), 99–113.

Naseri, M., P. Baraldi, M. Compare, and E. Zio. 2016. "Availability assessment of oil and gas processing plants operating under dynamic Arctic weather conditions." *Reliability Engineering & System Safety*, 152, 66–82.

Nateghi, R., S. D. Guikema, and S. M. Quiring. 2011. "Comparison and validation of statistical methods for predicting power outage durations in the event of hurricanes." *Risk Analysis*, 31 (12), 1897–1906.

Nath, D. C., A. Bhattacharjee, and R. K. Vishwakarma. 2016. "Risk assessment in liver transplantation patients: A shared frailty parametric approach." *Clinical Epidemiology and Global Health*, 4(1), 1-15.

Neath, A. A., and J. E. Cavanaugh. 2012. "The Bayesian information criterion: background, derivation, and applications." *Wiley Interdisciplinary Reviews: Computational Statistics*, 4 (2), 199–203.

Nelson, W. B. 2009. *Accelerated testing: Statistical models, test plans, and data analysis*. John Wiley & Sons.

Nisbet, R., J. Elder, and G. Miner. 2009. *Handbook of statistical analysis and data mining applications.*, Academic Press.O'Rourke, T. D. 2007. "Critical infrastructure, interdependencies, and resilience." *BRIDGE-Washington-National Academy of Engineering*, 37 (1), 22.

Orbe, J., E. Ferreira, and V. Núñez-Antón. 2002. "Comparing proportional hazards and accelerated failure time models for survival analysis." *Statistics in Medicine*, 21 (22), 3493–-3510.

Ouyang, M. 2014. "Review on modeling and simulation of interdependent critical infrastructure systems." *Reliability Engineering & System Safety*, 121, 43–60.

Ouyang, M., L. Dueñas-Osorio, and X. Min. 2012. "A three-stage resilience analysis framework for urban infrastructure systems." *Structural Safety*, 36, 23–31.

Pan, W. 2001. "Akaike's information criterion in generalized estimating equations." *Biometrics*, 57 (1), 120–125.

Peng, L., and Y. Huang. (2007). "Survival analysis with temporal covariate effects." *Biometrika*, 94 (3), 719–733.

Petersen, L., L. Fallou, P. Reilly, and E. Serafinelli. 2017. "European expectations of disaster information provided by critical infrastructure operators: Lessons from Portugal, France, Norway and Sweden." *International Journal of Information Systems for Crisis Response and Management (IJISCRAM)*, 9 (4), 23–48.

Pursiainen, C., and P. Gattinesi. 2014. "Towards testing critical infrastructure resilience." *JRC Scientific and Policy Reports.*

Qiao, Y., S. Labi, and J. D. Fricker. 2019. "Hazard-based duration models for predicting actual duration of highway projects using nonparametric and parametric survival analysis." *Journal of Management in Engineering*, 35 (6), 04019024.

Rausand, M., and A. Høyland. 2004. *System reliability theory: Models, statistical methods, and applications.* Hoboken, NJ: John Wiley & Sons.

Rinaldi, S. M., J. P. Peerenboom, and T. K. Kelly. 2001. "Identifying, understanding, and analyzing critical infrastructure interdependencies." *IEEE Control Systems*, 21 (6), 11–25.

Rocchetta, R., Y. F. Li, and E. Zio. 2015. "Risk assessment and risk-cost optimization of distributed power generation systems considering extreme weather conditions." *Reliability Engineering & System Safety*, 136, 47–61.

Rochas, C., T. Kuzņecova, and F. Romagnoli. 2015. "The concept of the system resilience within the infrastructure dimension: application to a Latvian case." *Journal of Cleaner Production*, 88, 358–368.

Rose, A. 2004. "Defining and measuring economic resilience to disasters." *Disaster Prevention and Management: An International Journal*, 13 (4), 307–314.

Rose, A., and S. Y. Liao. 2005. "Modeling regional economic resilience to disasters: A computable general equilibrium analysis of water service disruptions." *Journal of Regional Science*, 45 (1), 75–112.

Saeed, T. U., R. Nateghi, T. Hall, and B. S. Waldorf, B. S. 2019a. "Statistical Analysis of Area-wide Alcohol-related Driving Crashes: A Spatial Econometric Approach." *Geographical Analysis.*

Saeed, T. U., T. Hall, H. Baroud, and M. J. Volovski. 2019b. "Analyzing road crash frequencies with uncorrelated and correlated random-parameters count models: An empirical assessment of multilane highways." *Analytic Methods in Accident Research*, 100101.

Seraneeprakarn, P., S. Huang, V. Shankar, F. Mannering, N. Venkataraman, and J. Milton. 2017. "Occupant injury severities in hybrid-vehicle involved crashes: A random parameters approach with heterogeneity in means and variances." *Analytic Methods in Accident Research,* 15, 41-55.

Shaon, M. R. R., X. Qin, M. Shirazi, D. Lord, and S. R. Geedipally. 2018. Developing a Random Parameters Negative Binomial-Lindley Model to analyze highly over-dispersed crash count data. Analytic methods in accident research, 18, 33-44.

The Norwegian Meteorological Institute (MET). 2015. "Ekstremverrapport ", Norwegian Meteorological Institute <https://www.met.no/publikasjoner/met-info/ekstremvaer>. Accessed 10 March 2019.

The Norwegian Water Resources and Energy Directorate (NVE). 2014a. "Erfaringer etter ekstremværet Hilde." Rapport 2014:8.

The Norwegian Water Resources and Energy Directorate (NVE). 2014b. "Erfaringar fra ekstremværet Ivar." Rapport 2014:9.

The Norwegian Water Resources and Energy Directorate (NVE). 2015. "Erfaringar frå ektremveret Nina." Rapport 2015:55.

The Norwegian Water Resources and Energy Directorate (NVE). 2017. "Erfaringer fra ekstremværet Tor. Sammenlignet med erfaringer fra Dagmar." Rapport 2017:41.

Tian, L., D. Zucker, and L.J. Wei. 2005. "On the Cox model with time-varying regression coefficients." *Journal of the American Statistical Association*, 100, 172–183.

UNISDR. n.d. "Terminology on disaster risk reduction ", <https://www.unisdr.org/we/inform/terminology>. Accessed 15 September 2019.

Volinsky, C. T., and A. E. Raftery. 2000. "Bayesian information criterion for censored survival models." *Biometrics*, 56 (1), 256–262.

Wei, L.-J. 1992. "The accelerated failure time model: a useful alternative to the Cox regression model in survival analysis." *Statistics in Medicine*, 11 (14-15), 1871–1879.

Wienke, A. 2010. *Frailty models in survival analysis*. Chapman and Hall/CRC.

Yashin, A. I., J. W. Vaupel, and I. A. Iachine. 1995. "Correlated individual frailty: An advantageous approach to survival analysis of bivariate data." *Mathematical Population Studies*, 5 (2), 145–159.

Youn, B. D., C. Hu, and P. Wang. 2011. "Resilience-driven system design of complex engineered systems." *Journal of Mechanical Design*, 133 (10), 101011.

Yue, H. and K. S. Chan. "A Dynamic Frailty Model for Multivariate Survival Data". *Biometrics*. 53(3), 785 - 793.