INVERTER RELIABILITY IN PV SYSTEMS: STATE-SPACE MODELING AND
BAYESIAN ANALYSIS

By

Josue Sánchez

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Electrical and Computer Engineering—Master of Science

2025

ABSTRACT

The push for cleaner energy sources, coupled with declining costs, has facilitated the massive deployment of solar photovoltaic (PV) systems in electric grids worldwide. At the heart of any PV system is the inverter, a device responsible for converting DC power captured by solar cells into AC power suitable for grid use. In recent years, reliability concerns have emerged regarding inverters, with multiple reports identifying central and string inverters as the primary culprits in most forced outages in PV systems. Inverter failures significantly hinder energy production, potentially reducing it to zero. Thus, estimating the reliability of these devices is crucial for forecasting the long-term performance of PV systems.

In this thesis, we develop a state-space reliability model to characterize the failure behavior of string inverters, using a limited and heterogeneous failure dataset from residential and commercial PV systems in the U.S. Despite the data constraints, the proposed model successfully captures both decreasing and increasing failure rate behaviors observed in the data. Additionally, we derive an exponential approximation of the model, enabling system-level reliability evaluation via Markov Reward Models (MRM). To address the uncertainty inherent in limited datasets, we adopt a Bayesian framework, which is better suited for uncertainty quantification under data scarcity. This approach allows us to compute credible intervals on expected energy production by propagating parameter uncertainty through the MRM. Our findings indicate that, although parameter uncertainty is non-negligible, its impact on expected long-term energy yield remains limited—primarily due to the relatively fast replacement of inverters compared to their average time to failure.

Lastly, since the failure rate is an important quantity for reliability optimization and risk assessment, we establish a method for detailed failure rate estimation, providing deeper insights into the failure process. Following this approach, without relying on any major assumptions, the model estimations confirm our assumptions of a bathtub-like failure rate behavior.

To my parents, my brothers, and Ana

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Electric grids are facing two important challenges, an ever-increasing demand for electricity and higher pressure to reduce their reliance on fossil-based generation amid global warming concerns. The combined effect of these two challenges, along with declining costs, is driving the integration of more and more renewable energy resources (RERs) in an attempt to reduce the carbon footprint associated with electricity generation and ensure sufficient supply to meet the load. From the pool of RERs available, solar photovoltaic (PV) is the most widely deployed.

According to the International Energy Association (IEA), in 2023 around 407–446 GW of solar PV were installed globally. In the U.S. alone, 32 GW were installed, closing the year with a cumulative installed capacity of 137.5 GW [3]. Utility-scale installations make up most of the newly installed capacity, but residential installations have been growing steadily [4].

Commercial and residential PV systems give customers the opportunity of reducing their bills, and have more control over their electric supply. However, as more of these systems are installed, concerns about their reliability and how it can affect their expected performance have become critical. Moreover, the inverter, one of the most critical components of the system, has been deemed the most failure-prone [2,5–8]. Since PV systems generate electricity at no cost as long as there is sunlight, the amount of time it is fully functional (i.e, its availability) has a direct impact on the returns of investing in these systems.

Reliability analysis of these systems requires improved models and comprehensive approaches that consider the effects of component failure, along with other significant factors such as solar panel degradation. This thesis develops a state-space reliability model to characterize the failure behavior of string inverters using data and integrates this model into an analytical framework based on Markov Reward Models. A Bayesian approach facilitates accounting for parameter uncertainty, making it possible to establish confidence bounds for the expected energy production. Additionally, the thesis establishes a method for detailed
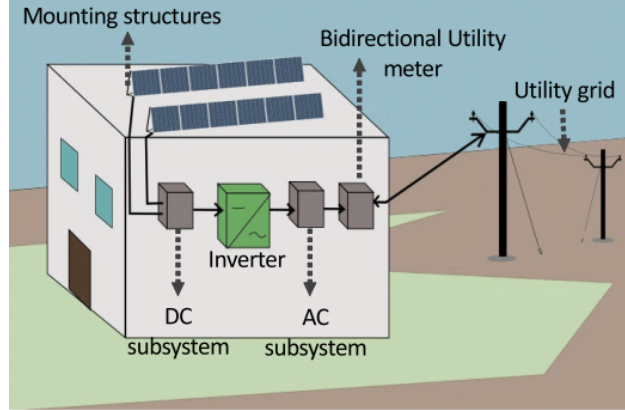
Figure 1.1 Residential grid-tied PV system [2]

failure rate estimation, providing deeper insights into the failure process and confirming a bathtub-like failure behavior.

1.1 Motivation & Challenges

A PV system is made up of several components, including solar panels, mounting structures, watt-meters, inverters, and protective devices. Most small PV systems are grid-tied, meaning they operate connected to the electric grid. The basic configuration of a grid-tied PV system is shown in Figure 1.1. The system design is determined by the type of inverter used. Residential and commercial PV systems can be designed to work with string or microinverters, but they are typically designed to use string inverters because of their lower cost [2].

Several studies have analyzed failure records of PV systems and concluded that inverters are responsible for the majority of forced outages. In [7] failure data were collected by Sandia National Laboratory from PV systems in the U.S. of various sizes and grouped in four portfolios. Portfolio D consists of mostly small systems of up to 500 kW. In all portfolios, inverter failures are the main cause of maintenance tickets, accounting for 50% of all failures in portfolio D.

Gunda et al. [1] analyze a dataset of 55,000 inverter-related maintenance records and apply machine learning techniques to classify each entry into a distinct failure mode. However, because of data constraints, the failure distributions provided are only for time until the

first failure. The parameters of the distributions for some failure modes are shown in Table 1.1. The shape parameter of all distributions is less than one, which indicates a decreasing likelihood of failure. It is worth noting that no information about the rating of inverters is provided, and all systems under study were located in the U.S.

Table 1.1 Weibull parameters for annual failure rates reported in [1]

Inverter Subsystem	Shape (α)	Scale (β)
Communications	0.69	3.29
Ground Faults	0.77	3.60
Heat Mgmt. Systems	0.93	3.35
IGBTs	0.81	6.01

A study conducted by Golnas [5] analyzed energy production in 600 PV systems during a period of 27 months and found that inverter-associated outages accounted for 36% of energy not produced. Furthermore, a comparison between central-, string-, and microinverter-based systems is done in terms of a normalized metric called tickets per inverter-year; showing string systems to be far more reliable than central inverters, but having 3 to 4 times higher tickets per inverter-year than microinverters.

The largest PV system failure database found in the literature is presented in [8]. The dataset consists of 100,000 PV systems in the US and over 7 GW$_{DC}$ of total capacity, collected for five years. Again, inverters make the top of the list in the number of failures for all three system types (residential, commercial, and utility). However, the impact on energy production of inverter failures is significantly higher in residential systems than in commercial and utility systems. The authors argue that this is because of the closer supervision and monitoring in larger systems. Additionally, residential systems may be leased by the property owner, which means that the property and system owners need to coordinate repair schedules, potentially increasing the time to resolve hardware issues.

A possible explanation why inverters fail so often is given in [9] and it is a consequence of three factors. First, inverters have to carry out a multitude of functions with little to no redundancy, and under harsh environmental conditions. Secondly, the pressure to reduce

costs in inverters forces some manufacturers to rely on cheaper materials that may not provide the required resistance to achieve the expected life of the component. Lastly, inverters are operated through software that is sometimes updated remotely, and these updates can cause malfunctions when not enough testing is carried out prior to rollout.

There is evidence in the literature to believe that inverters are the most fragile component in PV systems, and their failures affect significantly energy production; this is true particularly for residential systems. When a string inverter fails, energy production is severely hindered, even dropping to zero [1]. Therefore, accurate modeling of inverter failures is crucial to estimate the reliability of PV systems. This not only will allow a better understanding of how they fail, but also to assess more accurately the impacts of reliability on the economics of PV systems [2].

Reliability studies concerning PV systems have long relied on two kinds of analysis, analytical and simulation-based. The simulation-based approach is conducted by means of a Monte Carlo Simulation (MCS), easily accommodating any effects that could not be modeled analytically. Regardless of this, accurate reliability models are still required to obtain correct results, which has been difficult to achieve in power systems because of the lack of data [10]. However, only steady-state numerical results can be obtained from MCS. Ignoring transient behavior can lead to an under- or overestimation of the risk associated with system failure [11].

Most analytical models used for PV system reliability analysis rely on Continuous-time Markov Chains (CTMCs) to derive expressions for system availability and other performance metrics [12–14]. An inherent assumption of CTMCs is that the failure rate of components remains constant. CTMCs are commonplace in power systems reliability because failure data is usually presented as Mean Time to Failure (MTTF) and Mean Time to Repair (MTTR). However, assuming a constant failure rate limits the capacity of these models to account for the impacts of aging or the occurrence of premature failures, potentially leading to inaccurate estimations of system availability [10].

The constant failure rate assumption may not be appropriate for inverters for two reasons. First, inverters, as any other power converter, include capacitors and power switches as their building blocks, which are known to wear and fail prematurely, causing their failure rate to be non-constant [10]. Secondly, previous research has analyzed inverter failure data and estimated distributions for various failure modes that have time-varying failure rates [1, 7]. As seen in Table 1.1, the shape parameter of these distributions indicates that inverters are less likely to fail later on than just after being put to operation; an indication of decreasing failure rate.

The impacts of non-constant failure rate behavior have been addressed extensively in the literature of power system reliability to model the long-term effects of aging [15–17] and account for component condition [18–20]. These works focus more on the long-term performance rather than the dynamic failure behavior, which also provides a lot of insights into how the component is failing. Nevertheless, very few studies in power systems have examined components where failures are followed by replacements, such as inverters.

In [21], the authors propose using a Semi-Markov Process (SMP) to model the time-varying failure rate due to aging, along with a CTMC for random failures. They calculate the total component availability as the product of the individual availabilities. The SMP employs a Weibull distribution for time to failure, and both models assume an exponential time to repair. While the SMP demonstrates high accuracy, it requires numerically solving a system of integral equations in the time domain, which makes it unsuitable for an analytical whole-system analysis. Lifetime data for the models were derived from Stress-Strength Analysis (SSA) and electrothermal modeling.

Sangwongwanich and Blaabjerg [22] approach the reliability modelling of fault-tolerant power converters using the method of stages. The method consists in approximating the Weibull distribution using an Erlang distribution. The resulting model comprises several stages of constant failure rate states that behave as a Weibull distribution. Even though it is possible to approximate any distribution using the method of stages, difficulties will

arise when the failure data being dealt with exhibits both increasing and decreasing failure rates; only increasing failure rate data is present because the authors generate the lifetime data using electrothermal modeling, which is meant to account only for aging—an increasing failure rate process.

Peyghami et al. [10] compares CTMC, method of stages, and SMP to model the availability of components subject to increasing, constant, and decreasing failure rates. Although the method of stages approximates wear-out failures very well, it behaves similarly to CTMC when modeling premature failures, ignoring the sharp decrease in availability as a result of a decreasing failure rate. The results show that only SMP accurately captures the behavior of decreasing failure rates. However, it has the same limitations as the SMP in [21]. The analyses shown in this work are theoretical, and no actual failure distribution development is undertaken.

1.2 Thesis Contributions

Up until now, reliability modeling of inverters has primarily focused on those used in large-scale installations, which differ from those in smaller systems. The reasons for this focus include the lack of sufficient data to derive a robust model and the perceived lower importance of high reliability in small systems. However, as mentioned in [2, 8], reliability can significantly impact the long-term performance of small systems due to the typical delays in resolving failures, making it imperative to develop accurate models from data that can account for their actual behavior.

Previous research has often assumed that there is no decreasing failure rate stage in the life of power converters, which may not be accurate for inverters. Moreover, critical aspects of reliability modeling, such as parameter uncertainty and distribution fitting, have gone unnoticed, possibly because the focus has been on generating large samples of lifetime data from first principles, which cannot account for hidden defects responsible for premature failures.

In order to bridge the existing gaps in the literature, the contributions of this work can

be summarized as follows:

- We propose a methodology to develop a data-driven state-space reliability model tailored for equipment that exhibits both increasing and decreasing failure rates.

- The methodology is applied to a limited dataset of non-repairable string inverter failures collected from residential and commercial PV systems in the U.S., where data scarcity and variability motivate the use of flexible, probabilistic modeling techniques.

- We derive an exponential approximation for the inverter model, facilitating the analytical comparison of expected energy production between systems designed with string and microinverters.

- A Bayesian approach is adopted to quantify the uncertainty of parameter estimates, seamlessly integrated into a reliability analysis based on Markov Reward Models.

- Leveraging Stan, we propose a detailed failure rate estimation procedure, providing estimates with a granularity of up to 50 days.

1.3 Thesis organization

The remainder of the thesis is organized as follows. Chapter 2 describes the reliability modeling of small PV systems. Chapter 3 explains the concepts concerning failure data analysis and describes the methodology used to develop the inverter reliability model. Chapter 4 introduces the concept of Markov Regeneration Process and Markov Reward Models, and provides the mathematical basis for the exponential approximation of the inverter model. Chapter 5 adopts a Bayesian approach to the failure data analysis. Lastly, Chapter 6 presents the conclusions.

CHAPTER 2

PV SYSTEM RELIABILITY

In this chapter, we analyze small PV systems from a reliability perspective, modeling them as multistate systems. We narrow our focus to residential PV systems for simplicity of design; however, the concepts presented here can be extended to commercial systems. We begin by providing the necessary definitions to formalize the analysis as multistate systems. Next, we examine the components of both types of system designs found in residential PV systems—string inverter and microinverter systems—and discuss their respective reliability models. Finally, the reliability of the system is graphically represented using a reliability block diagram.

2.1   PV system as a multistate system

It is possible to represent any system composed of binary or multistate components as a multistate system (MSS), since the overall performance or efficiency level of the system depends on the availability of its constituent units. If the MSS can occupy K distinct states, and g_i denotes the total system performance in state i ($i \in 1, \dots, K$), then the MSS performance rate at time t is a random variable taking values in the set $\mathbf{g} = g_1, \dots, g_K$ [23].

Different procedures have been employed in the reliability literature to analyze MSS, and they can be classified into state-space and non-state-space methods. The former are typically used to analyze non-repairable MSS, although, under the assumption of independent failures and repairs, they can be used for repairable systems as well. These methods include reliability block diagram, reliability graph, fault-tree analysis, and state-space enumeration. On the other hand, state-space methods have higher modeling power and are not limited by the assumption of independent failures and repairs. Continuous-time Markov Chains, Semi-Markov Process, and Petri nets are among the most widely used state-space methods [24].

To develop a MSS reliability model, it is necessary to understand how the individual components behave and how they affect the performance of the system [23]. Therefore, reliability models for each component must be devised and a analysis method chosen accordingly.

Grid-tied residential PV systems are constituted by [25]:

- Solar panels, an interconnected package of solar cells in charge of converting sunlight into DC electricity.

- Inverter, recognized as the brain of the installation, takes DC electricity coming from the panels and converts it into AC electricity that can be used to feed loads or feed back to the grid.

- Energy meter, keeps track of the energy fed back to the grid and how much is being drawn from it; there could be two energy meters, depending on the compensation scheme used.

- Wiring, electrical boxes and protective devices, provide connectivity between system components and protection against faults.

We will now focus on each component of the PV system, its reliability model, and how system performance is affected once it fails.

2.1.1 Solar panels

In string systems, solar panels are connected in series to form a string. Each string is then connected to one optimizer, whose job is to maximize the energy production of the string. In contrast, in microinverter systems, each solar panel has its own optimizer, and the production of the panels is aggregated on an AC bus. The way the panels are connected determines the impact of their failures on the system performance. If a solar panel in a string fails, the whole string fails unless there is a bypass diode that provides an alternate path for the current. In contrast, if a solar panel in a microinverter system fails, only its output would be lost and the rest of the system continues functioning [26].

Past research has found that solar panels seldom suffer from failures, making them one of the most reliable components in the PV system [1, 5, 7]. In [8], the share of solar panel

failures was about 0.1% of the total hardware failures. However, since there are many solar panels in a system, the impact of their failures needs to be accounted for.

Therefore, a simple two-state model with an exponential time to failure for each solar panel is sufficient to describe their behavior; these models are commonplace in the literature [12, 13, 25–27]. For the string systems in this work, we will assume that a bypass diode is present and the failure of one panel does not cause the failure of the entire string. By doing so, and given that all solar panels in a system have the same rating, the different states associated with solar panel failures can be grouped to reduce the number of system states. For example, a system with two solar panels where each has two states, functional and failed, would have a total of three states instead of four.

The failure rate of solar panels is drawn from [28], where the authors present a mean time to failure of 65,789,474 hours or 7,510.2 years.

2.1.2 Inverters

Inverters are responsible for converting DC into AC, which makes them indispensable for the operation of the PV system. Similar to solar panels, the way the system is designed and wired determines how inverter failures will impact the system. Failure of central, string, and microinverters results in the loss of energy production of all panels connected to them. This behavior suggests a kind of series dependence between solar panels and inverters, regardless of what kind of inverter we are dealing with, and only differs by the aggregated energy production lost, being noticeably lower for microinverter failures.

Microinverters are highly reliable, with Enphase advertising a time between failures of 600 years [29]. This duration is still far from that of solar panels, but is significantly longer than that of string and central inverters. Furthermore, field studies confirm that microinverters fail significantly less frequently than central and string inverters [8]. The reason for why microinverters are more reliable is that their components have significantly lower power handling requirements than string and central inverters [2]. Consequently, as for solar panels, using a simple two-state model with an exponential time to failure is justified.

Unlike microinverters, string and central inverters are prone to failures, and their failures result in the loss of production from several solar panels, rendering crucial having an accurate reliability model. We will address the reliability model of string inverters in the next chapter.

2.1.3 Balance-of-system components

In well-designed and properly installed systems, failures originating from energy meters, wiring, disconnects, and other hardware required to transfer power are unlikely [8]. Lastly, any grid failures would not significantly impact the long-term performance of PV systems because most outages are resolved in a few hours, at most [30]. Furthermore, grid and energy meter failures will affect the reliability of the system in the same way regardless of its architecture, making it possible to compare the reliability of two different inverter configurations without the need of including them in the model.

Therefore, for the reliability diagram of the system, only inverter and solar panel failures will be considered.

2.2 Reliability block diagram

The PV system is decomposed into several subsystems. The way of decomposing the system has to comply with the dependence between components and how repairs are carried out. For string systems, we can decompose them into subsystems comprising one string inverter and the solar panels connected to it. Only the inverter performance is usually monitored in these systems, which leaves solar panel failures to be addressed during maintenance visits. Conversely, when inverter failures happen, the system owner is notified by the monitoring system, prompting its repair. Under these assumptions, we can consider each subsystem as independent.

Microinverter systems are more reliable, and information about their performance is usually available, ensuring that the repair procedure for both, microinverter and solar panel, can be initiated promptly. Therefore, these systems can be decomposed into multiple subsystems of a microinverter and the panels connected to it.

From the way we have defined subsystems, as long as there is one solar panel functioning

and the inverter or microinverter to which it is connected has not failed, the system can produce electricity. Thus, each subsystem is aggregated in parallel to build the complete representation of the PV system. If a two-state model corresponding to DC or AC disconnects, energy meter, or grid failures were to be added, it would be in series with the parallel connection of the subsystems because, if any of these fail, energy production will drop to zero. The reliability diagram of the string inverter or microinverter subsystem is shown in Figure 2.1.
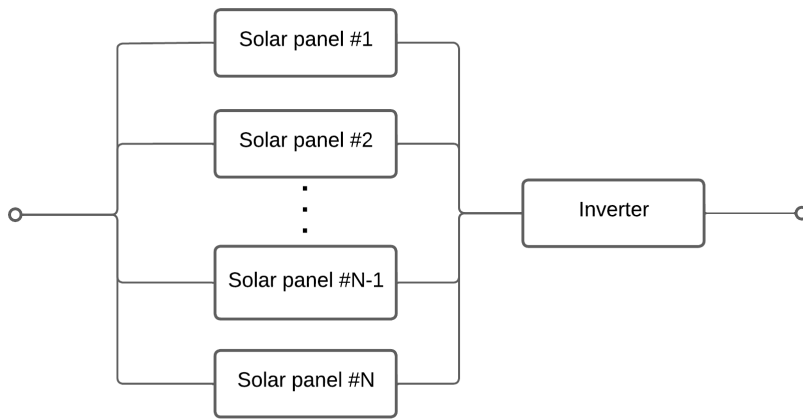
Figure 2.1 Reliability block diagram of subsystems

CHAPTER 3

RELIABILITY MODELING OF INVERTERS

In this chapter, we describe the reliability modeling of string inverters based on failure data. This approach, known as the actuarial approach, assumes that all relevant information about the failure of a component is captured by a time-to-failure distribution $f(t)$. From $f(t)$, reliability characteristics such as the failure rate $\lambda(t)$ can be directly derived.

We analyze the failure data and propose a piecewise exponential distribution to model the observed failures. This distribution also connects with the concept of state-space modeling, which enables the development of a reliability model. Finally, we present numerical results that highlight the importance of accounting for the time-varying failure rate behavior exhibited by inverters.

3.1 Developing reliability models

Reliability modeling of most components, except those present in safety-critical systems, involves the way they fail and how they are brought back to operation upon failure. In our case, the failure of an inverter leads to its subsequent replacement; thus, the component goes back to a good-as-new condition. Under this scenario, the failure-repair process can be considered an alternating renewal process. Therefore, the failure and repair distributions can be established independently [24].

To obtain the failure distribution of a component, it is necessary to carry out life tests, where the time until failure of each component being tested is recorded. However, in many cases, the lifetimes of the components are observed while they are in operation; this kind of measurement campaign is referred to as field data collection. Moreover, when failures are recorded from components in operation, a significant portion of them will not fail before the end of data collection, and instead of a lifetime value, their potential lifetime is obtained. This is called censored data in the reliability literature [31].

The procedure to fit failure models recommended in most reliability textbooks is as follows [32]:

1. Construct a histogram of the failures.

2. Calculate descriptive statistics.

3. Assess the empirical failure rate

4. Make use of any prior knowledge of the failure process

5. Select suitable candidate models

6. Estimate parameters

7. Perform a goodness-of-fit test

Steps 1 through 4 provide insights to narrow down the number of candidate distributions that will be considered in step 5. Step 6 consists of estimating the parameters of each candidate model, usually by means of maximum likelihood estimation, and judging them according to accuracy metrics such as the Akaike Information Criterion. Finally, in step 7, it is determined whether it is reasonable to assume the data came from the hypothesized distribution [31, 32].

A reliability model of the component can be developed from the estimated failure distribution using different approaches. The simplest is to use a Semi-Markov Process where the sojourn time in the up state is distributed according to the failure distribution estimated, and the state space model consists of only two states (up and down) [10]. An alternative approach is to use a CTMC with two or more transient and one absorbing states to approximate the non-exponential failure distribution. The preceding definition corresponds to a family of distributions called phase-type [10, 22, 24]. Some authors have shown that PH approximations are not so accurate, especially in the transient period, when there is a decreasing failure rate behavior present [10].

3.2 Description of the failure database

Data was collected from residential and commercial PV systems located in the Midwest part of the U.S. The earliest installation date of the systems under study was 2013, and

the capacity of the inverters ranges from 5 to 100 kW$_{ac}$. A total of 193 failures were registered. Failures were identified from current and voltage signals, when possible, and from the monitoring system component changelog when there was no communication during the failure.

The histogram of the failures is depicted in Figure 3.1; more than half of the failures occurred within 1000 days, while only 28 failures were recorded beyond 2400 days.
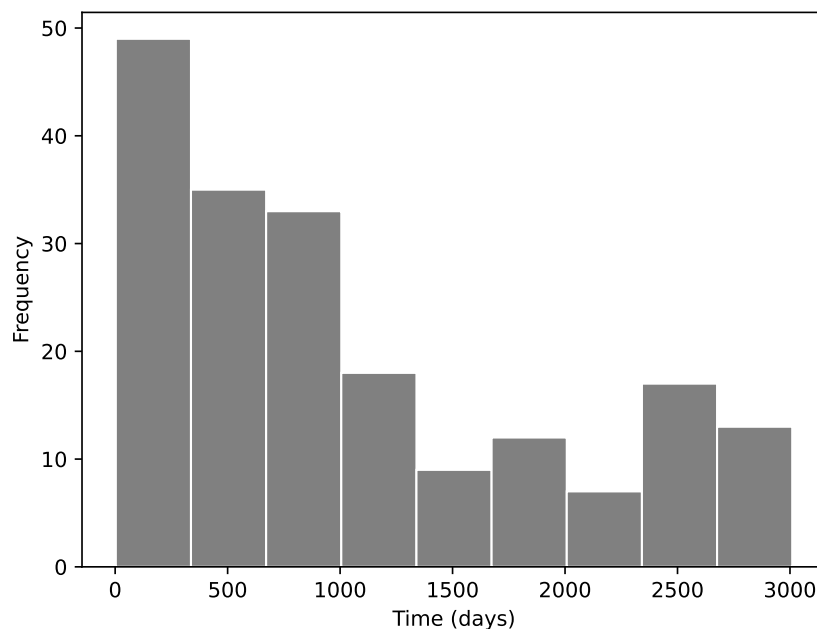
Figure 3.1 Histogram of failures

The failure database is complemented by censored data from 234 inverters, which are currently communicating with the monitoring platform. The censored data in this case is multiply right-censored, that is, each component was placed in operation at different times and only some of them failed while the data was being collected [32].

3.3 Non-parametric models

The procedure of fitting a failure distribution to data can be less of trial and error if there is some indication of the underlying failure behavior embedded in the data. In light of this, non-parametric representations of the failure data provide valuable insights into choosing

better models. The only assumptions made by these representations is that the cumulative distribution function (CDF), $F(t)$, is continuous and monotonically increasing [31].

Two important non-parametric representations will be discussed next, the Kaplan-Meier estimator and the cumulative failure rate.

3.3.1 Kaplan-Meier estimator

The Kaplan-Meier estimator provides an estimate of the survivor function $R(t)$ from censored data [31]. The Kaplan-Meier estimator is defined as:

$$\hat{R}(t) = \prod_{j \in J_t} \frac{n_j - 1}{n_j} \tag{3.1}$$

where $\hat{R}(t)$ is the estimate of the survival function at time t, $t(j)$ represents a failure time, J_t denotes the set of all indices j where $t(j) < t$, and n_j represents the number of items functioning and in observation immediately before time $t(j)$, $j = 1, 2, ..., n$.

When the largest lifetime recorded is censored, the estimator is said to be undefined after the last registered failure.

3.3.2 Cumulative failure rate

The Cumulative failure rate, also Nelson-Aalen estimator, is used to empirically determine the cumulative failure rate, $Z(t)$, in the presence of censored data [31]. The estimator is defined as:

$$\hat{Z}(t) = \sum_{j \in J_t} \frac{1}{n_j} \tag{3.2}$$

$Z(t)$ can be interpreted as the accumulation of the failure rate over time, and its concavity indicates the shape of failure rate. If $Z(t)$ is concave upward, there is an indication of increasing failure rate, and downward concavity indicates decreasing failure rate. A bathtub shaped curve would be seeing as $Z(t)$ being concave downward followed by a transition to a concave upward section.

From the cumulative failure rate, it is possible to ascertain if the data could be well represented by a Weibull distribution if it has a decreasing failure rate. Alternatively, by a

normal distribution, lognormal, or a Weibull if it has an increasing failure rate [32]. However, it is not possible to estimate the numerical value of the failure rate using this estimator.

3.3.3 Non-parametric analysis of the data

Figure 3.2 depicts a Kaplan-Meier estimate that can be thought of as three lines with different slopes. The changes in slope occur at around 1100 and 2500 days. The cumulative failure rate depicted in Figure 3.3 initially has a concave downward section, but suddenly changes to a linear section around 2500 days. The former is an indication of a decreasing failure rate, agreeing with what is shown in Figure 3.1. The latter, on the contrary, is more complicated to analyze because of the lack of sufficient data beyond 2500 days. Although, it is clear that there are two different failure rate behaviors, which is not easy to model using regular continuous distributions. We will show that it is possible to model this kind of behavior using a piecewise exponential distribution.
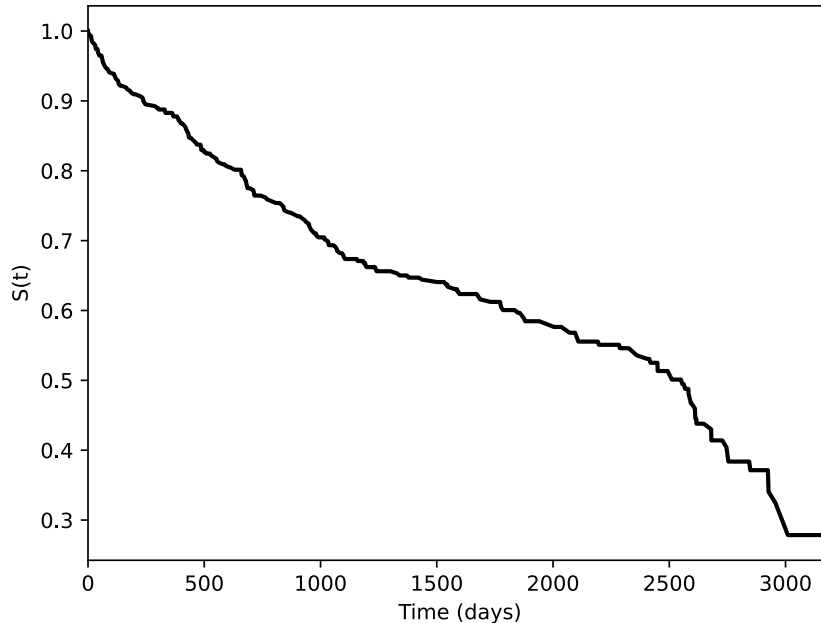
Figure 3.2 Kaplan-Meier estimate
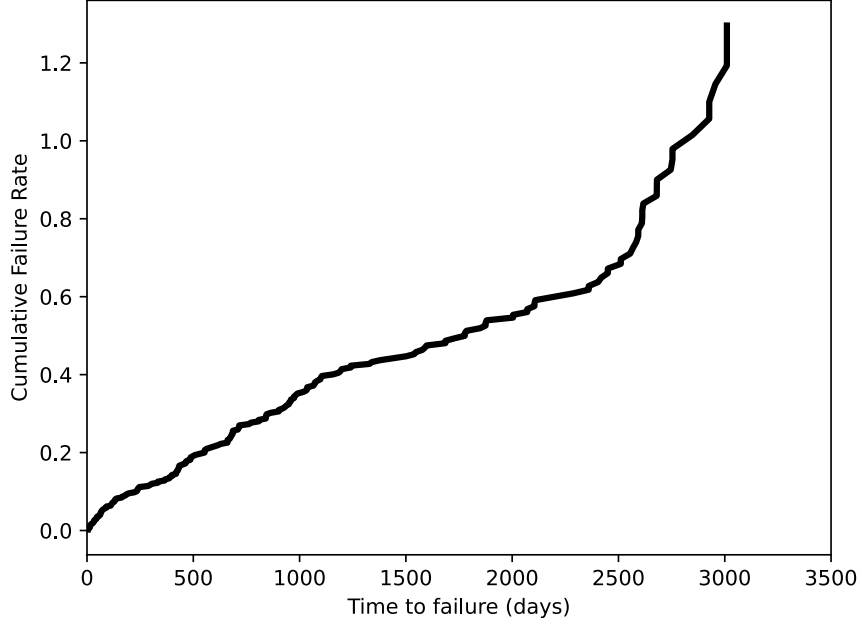
Figure 3.3 Cumulative failure rate

3.4 Piecewise exponential distribution

A random variable T is piecewise exponentially (PWE) distributed if its hazard rate function is piecewise constant with a total of r change points d_k, where $1 \leq k \leq r$ [33]. The hazard and survivor function are given by equations 3.3 and 3.4, respectively.

$$
h(t) = \begin{cases} \lambda_1, & t < d_1 \\ \lambda_2, & d_1 \leq t < d_2 \\ \vdots \\ \lambda_{r+1}, & t \geq d_r \end{cases}
\tag{3.3}
$$

$$
S(t) = \begin{cases} e^{-\lambda_1 t}, & t < d_1 \\ e^{(\lambda_2 - \lambda_1) d_1 - \lambda_2 t}, & d_1 \leq t < d_2 \\ \vdots \\ e^{\left[\sum_{i=1}^{r} (\lambda_{i+1} - \lambda_i) d_i\right] - \lambda_{r+1} t}, & t \geq d_r \end{cases}
\tag{3.4}
$$

18

The parameter estimation process of the PWE distribution is usually done using the maximum likelihood estimator, and it consists of determining the location of the change points and the hazard rates. If the change points are known, the likelihood function becomes smooth and the hazard rates can be found solving the maximization problem. On the other hand, when the change points are unknown, the estimation can be conducted by brute force or by fitting lines to the empiric survivor function plotted in logarithmic scale [33]. A brute force approach involves assuming a number of change points and estimating parameters for numerous combinations of possible locations, which can be tedious when several change points are considered.

In this work, the change points will be determined visually from the survivor function plotted in logarithmic scale.

3.5 Maximum likelihood estimator

The Maximum Likelihood Estimator (MLE) of any parametric distribution $f(\theta)$ is found by determining the parameter values $(\theta_1, \ldots, \theta_k)$ that maximize the likelihood function [32]. The likelihood function is given by equation 3.5.

$$L(\theta_1, \ldots, \theta_k) = \prod_{i=1}^{n} f(t_i \mid \theta_1, \ldots, \theta_k) \tag{3.5}$$

where $i = 1, 2, 3, ..., n$ are the indices corresponding to the number of failure observations.

Due to the multiplicative nature of equation 3.5 it is easier to find the parameters that maximize the logarithm of the likelihood function.

Equation 3.5 is slightly modified to take into account multiply censored observations, resulting in equation 3.6. The term $R(t_i; \theta)$ is added to find the set of parameters that maximize the reliability function for the censored observations.

$$L(\theta) = \prod_{i \in U} f(t_i; \theta) \prod_{i \in C} R(t_i; \theta) \tag{3.6}$$

where U and C are the set of indices corresponding to failure and censored observations, respectively.

3.6 Fitting failure distributions

In this section, we estimate the parameters of different failure distributions and assess how well each fits the data. The parameter identification is done using lifelines library in Python.

For the PWE distribution, it is necessary to identify the change points. In Figure 3.4, the Kaplan-Meier estimate is plotted with the vertical axis in logarithmic scale. Initially, we identified two changes in slope, at 1100 and 2500 days, which are the red dots in Figure 3.4. However, the best fit found was when two more change points are added, at 130 and 2900 days, depicted as green dots in Figure 3.4.
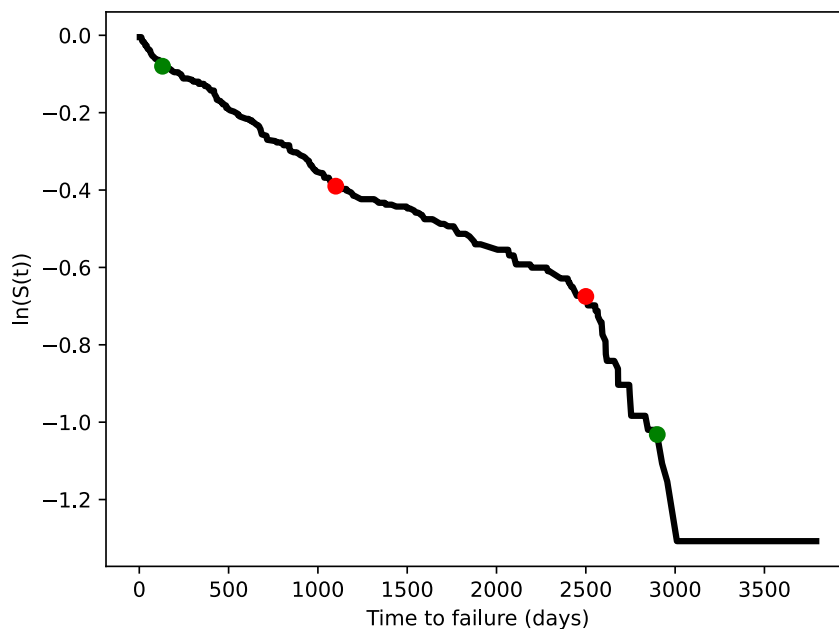
Figure 3.4 Visual identification of change points

Now, we compare distributions to model the data. A qualitative fit assessment is provided by Figure 3.5, where the survivor function of each estimated distribution is compared to the Kaplan-Meier estimate. It can be seen that the PWE distribution is the only whose 95% confidence bounds enclose completely the Kaplan-Meier estimate, a visual indication of a good fit. Quantitatively, the quality of fit can be assessed using the Akaike Information

Criterion (AIC), which is presented in Table 3.1; the PWE distribution achieves the lowest AIC value and is therefore the best model among the six proposed.

Table 3.1 Comparison of distributions according to AIC

Weibull	Exponential	Lognormal
3496	3500	3520
Loglogistic	**PWE**	**Generalized gamma**
3506	3470	3491

The Walt test is used to perform a goodness-of-fit test of the parameters of the PWE distribution. The p-values obtained for each parameter of the distribution are smaller than 0.05, which means that it is possible to reject the null hypothesis that each rate equals the baseline value of 1. In other words, the rates are statistically significant.

3.7 Proposed model

Reliability literature has consistently favored state-space modeling over non-state-space approaches, as it offers greater modeling power and the ability to represent dependencies among various components of a system [34]. Therefore, if the model is meant to be used for a whole-system analysis where there could be dependencies between components, a state-space modeling approach should be adopted.

We propose a state-space reliability model where the PWE distribution becomes a sequence of states, and use a Semi-Markov process as model formalism to describe its dynamic behavior. In the following subsections, we will present the mathematical basis of our approach.

3.7.1 Semi-Markov process

A Semi-Markov process is a generalization of the CTMC where the exponentially distributed sojourn time requirement is relaxed, but the process still follows the Markov property [24].

The SMP is completely characterized by its kernel matrix $Q(t)$ and the initial state of the process. Each element $Q_{ij}(t)$ represents the probability of a one-step transition from state i
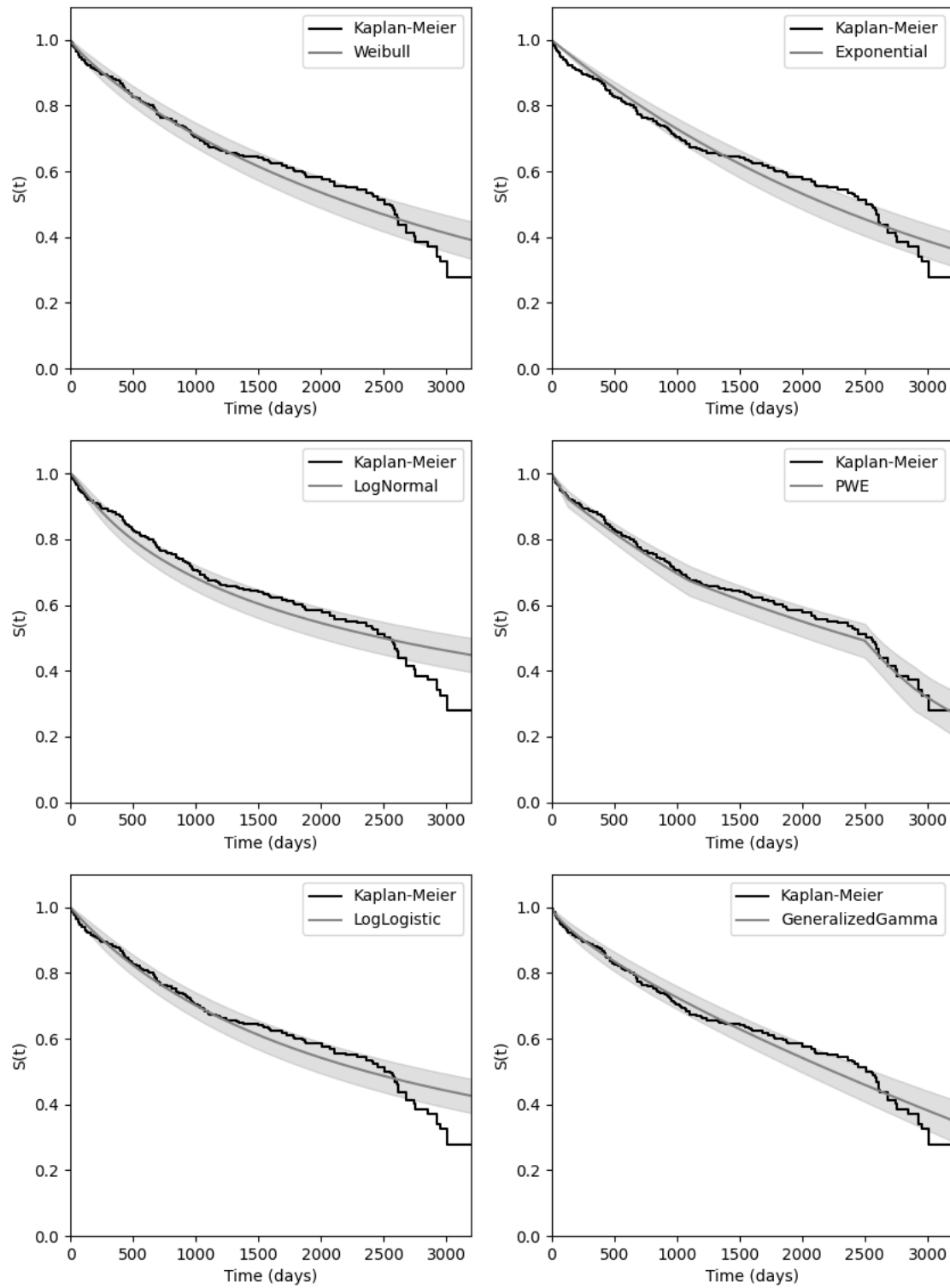
Figure 3.5 Survivor function of each failure distribution estimated

to state j during the interval $[0, t]$. Considering a scenario where a process is in state 0 and can move to two different states (1 and 2), Q_{01} can be found by applying equation 3.7.

$$Q_{01}(t) = \Pr\left\{(T_{0,1} \le t) \cap (T_{0,2} > t)\right\} \tag{3.7}$$

where $\Pr\{\cdot\}$ indicates the probability of the event enclosed in the curly brackets, $T_{0,1}$ and $T_{0,2}$ are random variables that denote the time to transition from state 0 to states 1 and 2, respectively.

The unconditional sojourn time distribution in state i can be obtained from $Q(t)$ using equation 3.8.

$$F_i(t) = \sum_{j=1}^{K} Q_{ij}(t) \tag{3.8}$$

where $j = 1, 2, ..., K$ denote the system states.

The conditional time-dependent probabilities can be found by solving the following system of integral equations [24]:

$$\theta_{ij}(t) = \delta_{ij}\left[1 - F_i(t)\right] + \sum_{k=1}^{K} \int_0^t \frac{dQ_{ik}(\tau)}{d\tau}\theta_{kj}(t - \tau)d\tau \tag{3.9}$$

where δ_{ij} is an indicator variable that evaluates to 0 if $i \ne j$, and 1 if $i = j$. The term $\theta_{ij}(t)$ reads as the probability that the process is in state j at time t, given that the process started from state i.

3.7.2 Deriving the state-space model

The PWE distribution can be viewed as a sequence of J transient states, where $J = r+1$, and one absorbing state F. From each transient state, except the last state in the sequence, it is possible to transition toward the failure state or to the next state in the sequence.

The time until a transition to failure T_{iF}, $i = 1, 2, \ldots, J$, is exponentially distributed with parameter λ_k, corresponding to the k-th interval of the PWE distribution, defined between the change points $d_k - d_{k-1}$. The CDF of T_{iF} is denoted by equation 3.10. Similarly, the

time until a transition to the next state in the sequence T_{ij}, $j = i + 1$, is deterministic with a CDF given by equation 3.11.

$$F_{iF}(t) = \begin{cases} 0, & t < 0, \\ 1 - e^{-\lambda_k t}, & t \geq 0. \end{cases} \tag{3.10}$$

$$F_{ij}(t) = \begin{cases} 0, & t < T_k, \\ 1, & t \geq T_k. \end{cases} \tag{3.11}$$

where $T_k = d_k - d_{k-1}$. When $k = 1$, $T_k = d_1$ Since the process can only start from the first state of the sequence, the probabilities of interest have the form $\theta_{1j}(t)$.

3.7.3   State probabilities

The probability of being in any state at time t can be found by solving equation 3.9 in Laplace domain. These equations can be written in matrix form as equation 3.12 [24].

$$\boldsymbol{\theta}(s) = (\mathbf{I} - \mathbf{Q}(s))^{-1} \left(\frac{1}{s} \mathbf{I} - \mathbf{F}(s) \right). \tag{3.12}$$

where:

$$\theta_{ij}(s) = \int_0^\infty e^{-st} \theta_{ij}(t) \, dt, \tag{3.13}$$

$$Q_{ik}(s) = \int_0^\infty e^{-st} \, dQ_{ik}(t), \tag{3.14}$$

$$F_i(s) = \int_0^\infty e^{-st} F_i(t) \, dt. \tag{3.15}$$

To obtain the time-domain expression of $\theta(s)$, the inverse Laplace operator is applied. It may not be possible to obtain a close form solution for $\theta(t)$; however, the integral that defines the inverse Laplace transform can be numerically evaluated.

Without a transition out of the failure state, the state probabilities for the SMP are nonzero only during the interval where the process could certainly have reached that state, and no transition to the next state could have occurred. Thus, the individual time-dependent

24

probabilities have no meaning in this case. However, adding the probabilities of all transient states yields the survivor function.

The survivor function is derived analytically by solving the system of equations in (3.16) because the singularity of $Q(t)$ when there is no transition out of failure prohibits the use of equation 3.12.

$$\begin{cases} \theta_{jF}(s) = sQ_{j(j+1)}(s)\theta_{(j+1)F}(s) + Q_{(j+1)F}(s) \\ \vdots \\ \theta_{NF}(s) = Q_{NF}(s) \end{cases} \tag{3.16}$$

where $j = 1, \ldots, J - 1$.

The equations can be back-substituted to obtain an expression for $\theta_{1F}(s)$ solely in terms of the complex Laplace-domain variable s. Thereafter, the inverse Laplace operator is applied to obtain the time-domain expression.

3.8   Results and discussion

The PWE distribution fitted in Section 3.6 can be converted into a SMP with five transient states and one absorbing state. The matrix $Q(t)$ has dimensions 6×6 and the initial state is the first state of the sequence. The state-space model is represented by the diagram in Figure 3.6.

Solving equation 3.16 yields a time-domain expression consisting of exponential terms multiplied by Heaviside functions, which describe the same survivor function as the fitted PWE distribution. Thus, the failure process is accurately reproduced by the SMP.

Next, we want to use the model to find the availability of the inverter. To do this, we include a transition out of the failure state. The time to restoration is assumed to follow an exponential distribution with rate μ. In this scenario, $Q(t)$ is full rank, and it is possible to apply equation 3.12. The element $\theta_{16}(s)$, which represents the unavailability of the system, is extracted and its time-domain representation is obtained numerically using Euler's method [35]. The availability of the proposed model, along with that of a two-state
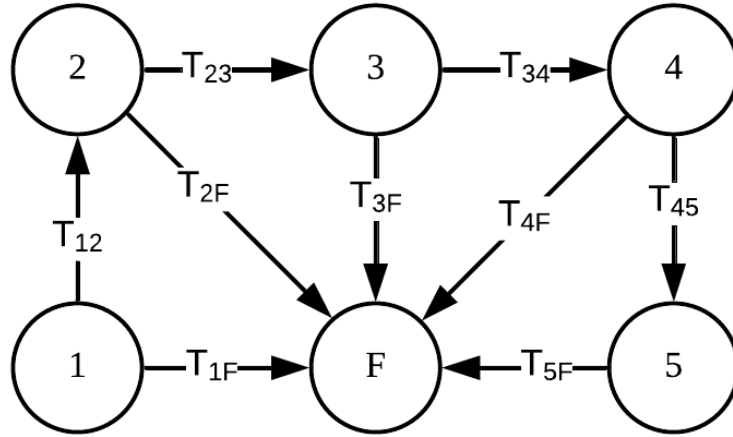
Figure 3.6 Proposed reliability model
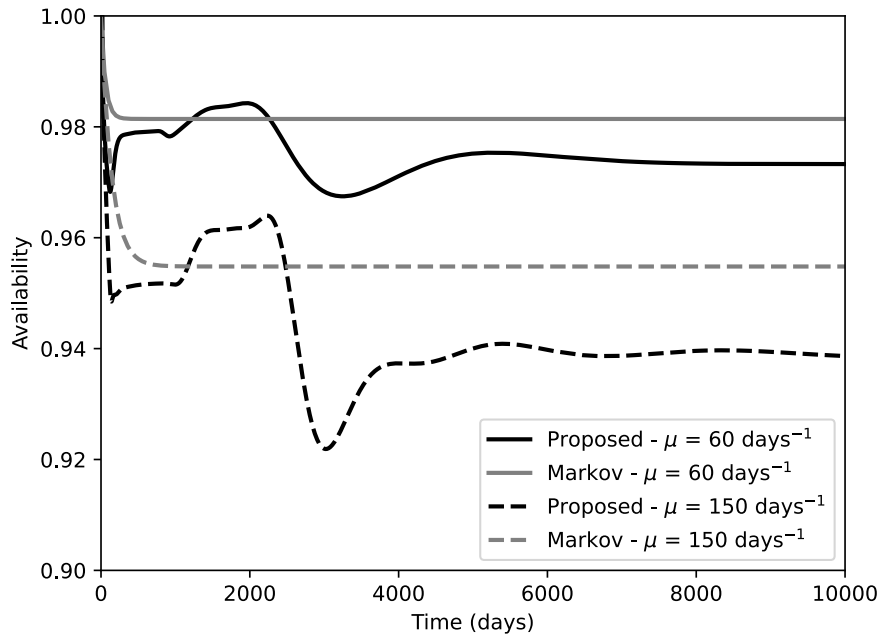
Figure 3.7 System availability

Markov model, is plotted in Figure 3.7. The two-state model failure rate is obtained from the MLE of an exponential distribution from the data.

As shown in Figure 3.7, the availability of the proposed model initially decreases sharply, reflecting a high likelihood of premature failures, which implies a decreasing failure rate. This is followed by an increase in availability and a subsequent dip—suggesting a period of

rising failure rates—before finally stabilizing at a steady-state value. The two decreasing epochs in availability demonstrate the model's capability to capture the underlying patterns of varying failure rates in the data, as confirmed by the cumulative failure rate. In contrast, the two-state CTMC model demonstrates a much simpler behavior, where availability quickly stabilizes at a steady state.

As the mean time to repair increases, the error incurred by using a two-state model becomes more significant. This may not be an issue for large-scale facilities, but it can impact the economic feasibility of systems used in residential and commercial settings, which are not monitored as closely and where repairs can take longer [8].

The accuracy of the inverter model comes with an important caveat: conducting a whole-system analysis with our model requires the use of a modeling formalism capable of maintaining some memory of where the process is in time, because the model is not memoryless. To illustrate this, let us consider a system comprising one inverter and one solar panel; here we assume the solar panel to be a two-state exponential model. The process starts from the state 1 of the inverter and the up state of the solar panel. If the solar panel fails, the clock resets and the progress towards state 2 of the sequence is lost. After the failure of the solar panel, the inverter will restart in state 1 as if no time has elapsed, leading to an incorrect representation of its failure behavior.

In the next chapter, we will introduce a modeling formalism capable of dealing with the non-memoryless nature of our model, called Markov Regenerative Models.

CHAPTER 4

SUBSYSTEM RELIABILITY MODEL

Up until now, we have discussed the reliability of PV systems and reduced it to several subsystems connected in parallel, each consisting of solar panels and the inverter to which they are connected. In the previous chapter, we developed a reliability model for string inverters that captures both premature and wear-out failures, and discussed why the model structure cannot be directly applied to a system analysis under the Semi-Markov framework. In this chapter, we introduce Markov Regenerative Process and elaborate on how to apply it to the solar panel-inverter subsystem. Thereafter, we utilize phase-type distributions to find an exponential representation of the string inverter model. Lastly, we describe Markov Reward Process and its use for estimating energy yield of PV systems.

4.1 Markov Regenerative Process

Markov Regenerative Processes (MRGPs) is a class of stochastic models that can be used to model processes in which there are transitions that do not adhere to the Markov property. MRGPs are a generalization of many stochastic models such as SMP and CTMC. The key concept behind MRGPs is that the Markov property holds for a certain group of states called regeneration states [24].

The embedded time points $\{T_n, n \geq 0\}$ are the regeneration time points (RTP) of the system, that is, the process resets every time it reaches these points. To define a MRGP, two matrices need to be specified: global $K(t)$ and local $E(t)$ kernel matrices. The global kernel describes the occurrence of the next RTP, and can be formulated as equation 4.1. On the other hand, the local kernel expresses the state transitions within the regeneration interval, before the process hits the next RTP; each element of $E(t)$ is computed using the expression in equation 4.2.

$$K_{ij}(t) = P\{Y_1 = j, T_1 \leq t | Y_0 = i\} \tag{4.1}$$

$$E_{ij}(t) = P\{Z(t) = j, T_1 > t | Y_0 = i\} \tag{4.2}$$

where $Z(t)$ is the state of the process at time t, and i, j are states within the state space S.

Using the global and local kernels, the transition probability over $(0, t]$ can be found by solving the generalized Markov Renewal Equation, shown in equation 4.3.

$$V_{ij}(t) = E_{ij}(t) + \sum_{k \in S} \int_0^t dK_{ik}(y)V_{kj}(t - y). \tag{4.3}$$

where $V_{ij}(t)$ is the transition probability.

As for SMPs, equation (4.3) can be converted to the Laplace domain to obtain a matrix equation that can be solved analytically if the kernels can be expressed in closed form. The resulting matrix equation is denoted by equation 4.4.

$$\hat{V}(s) = [I - \hat{K}(s)]^{-1}\hat{E}(s) \tag{4.4}$$

Solving equation 4.4 involves numerous symbolic operations just to find $\hat{V}(s)$, which we will need to numerically integrate to find $V(t)$. This procedure becomes computationally expensive as the state space grows, necessitating an alternative solution approach. A more efficient way of solving equation 4.3 is to use an exponential approximation of the inverter model so that the subsystem model can be expressed as a CTMC, which are easier to analyze than MRGPs [24].

4.2 Exponential approximation

Finding an exponential approximation of a non-exponential stochastic process is called Markovization. This procedure entails replacing the non-exponential distribution with a phase-type distribution, representable as a CTMC with one absorbing state. Furthermore, any discrete state non-exponential stochastic process can be approximated by an equivalent CTMC over an expanded state space using phase-type distributions [24].

4.2.1 Phase-type distributions

Phase-type (PH) distributions are employed to model non-exponential distributions as the time until absorption of a Markov process with one absorbing state. PH distributions of order n can approximate as closely as desired any distribution function [24].

The components of a PH distribution are:

1. State Space: A set of transient states $1, 2, ..., n$ and one absorbing state.

2. Initial Probability Vector: The probability of starting in each transient state is given by α.

3. Sub-Generator Matrix (T): The matrix represents the transition rates between the transient states.

Different techniques have been proposed in the literature for fitting PH distributions: Fitting mixture of Erlangs with moment-matching techniques, MLE, and Expected Maximization with various classes of PH distributions. Computer programs such as PhFit have been developed to fit continuous distributions and experimental data to PH distributions [24].

4.2.2 Deterministic transitions

The inverter model developed in the previous chapter exhibits non-exponential behavior due to deterministic transitions between success states. This deterministic time to transition can be modeled as a PH distribution by representing it as an Erlang random variable, which is the sum of k i.i.d. exponential random variables. Being the sum of k exponential random variables with the same rate parameter means that it is equivalent to a succession of k exponential stages, all with the same rate. The mean and coefficient of variation of an Erlang random variable are given in equations 4.5 and 4.6, respectively.

$$\mu = k/\lambda \tag{4.5}$$

$$\text{CV} = 1/\sqrt{k} \tag{4.6}$$

where k is the number of states of the exponential approximation and λ is the transition rate between states.

Equation 4.6 reveals an interesting property of the exponential approximation by means of an Erlang: The quality of the approximation depends on the number of states. Ideally, a deterministic variable has zero variance or an infinite coefficient of variation. However, for

practical purposes, accurate representations can be achieved without the need for too many states [36].

To find the exponential approximation, both the CV or k, and the mean need to be specified. The mean simply represents the value at which the transition occurs, and the CV or k is set to achieve the desired accuracy. Subsequently, the goodness of the approximation can be assessed by plotting the CDF.

The CDF of the exponential approximation can be found by solving the Kolmogorov differential equation for the transition rate matrix (A) and initial state vectors (p_0), given in equations 4.7 and 4.8, respectively.

$$A = \begin{bmatrix} Q_u & \mathbf{a}^T \\ \mathbf{0} & 0 \end{bmatrix},$$ (4.7)

$$p_0 = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}$$ (4.8)

where Q_u is a $k \times k$ matrix and each of its rows has only two nonzero values, $a_{ii} = -\lambda$ and $a_{i,i+1} = \lambda$. The column vector $\mathbf{a}^T$ is found as $-Q_u \mathbf{e}^T$, where $\mathbf{e}^T$ is a column vector with all entries equal to 1. The initial state vector p_0 is a $1 \times (k+1)$ vector.

The Kolmogorov differential equation is written as [23]:

$$\frac{dp(t)}{dt} = p(t)A$$ (4.9)

Figure 4.1 depicts the exponential approximation of a deterministic transition at 100 days for different values of CV. As denoted by equation 4.6, in Figure 4.1 we can see that the accuracy of the approximation increases with decreasing CV or increasing number of states.

4.2.3 Exponential inverter model

Now, we elucidate the procedure to find the exponential approximation of the inverter model. First, we find the exponential approximation for each deterministic transition. Then, we aggregate these approximations to form the state-space of the inverter model. Figure 4.2 depicts the transition rate diagram of the exponential inverter model. Each state in the same
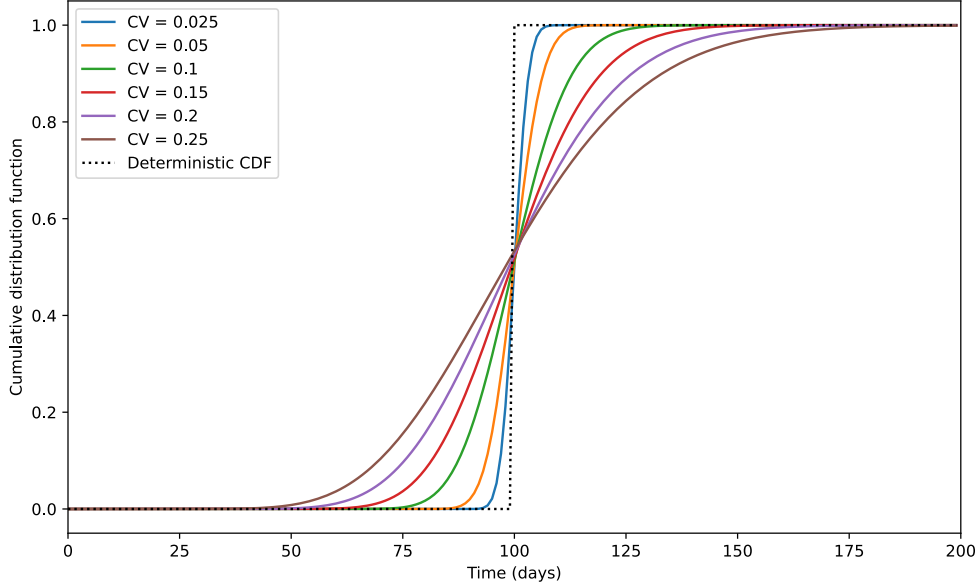
Figure 4.1 Exponential approximation of a deterministic transition for different coefficients of variation

group has a transition to failure given by the PWE distribution, whereas the transition from one transient state to the next is given by the rate determined through the deterministic approximation.

Assuming $k = 25$ for the first and last transition, and $k = 100$ for the rest, we find a good compromise between accuracy and complexity. Figure 4.3 shows the availability of both the exponential model and the SMP. The maximum difference between the two responses is 0.3%. Therefore, we can conclude that the exponential approximation does a very good job at capturing the non-exponential behavior of the inverter model.

4.3 Subsystem model

Now that we have found an exponential approximation of the inverter model, we can create a subsystem model. The transition rate diagram depicted in Figure 4.4 shows the state space of the subsystem model with string inverter and the different failure transitions between states. Horizontal transitions and transitions denote the progression through each state of the inverter model, whereas vertical transitions denote changes in the number of
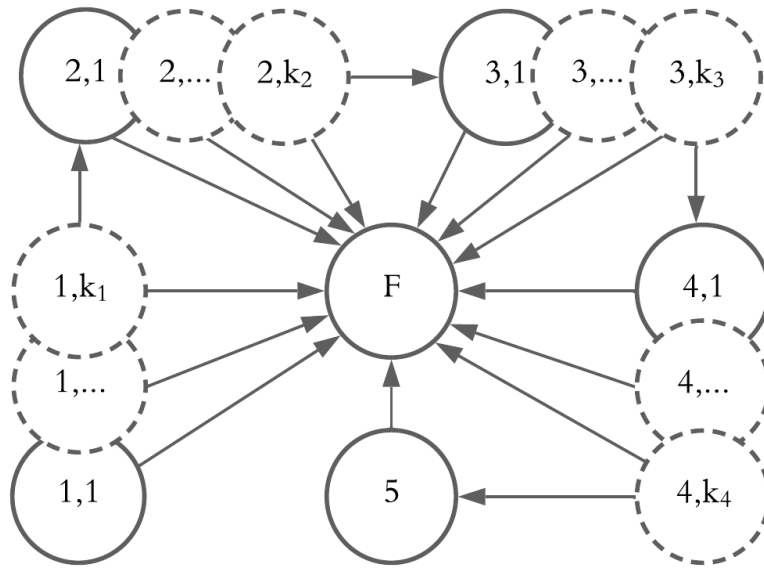
Figure 4.2 Exponential inverter model
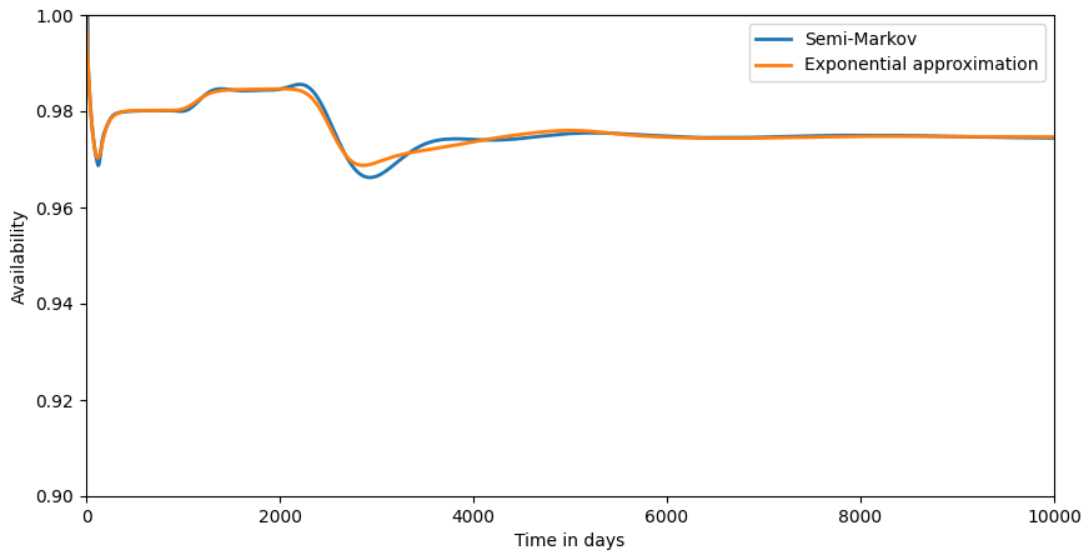
Figure 4.3 Availability of both inverter models

solar panels functioning. From each row to the next, the system moves with a transition rate given in equation 4.10.

$$\lambda_i = n_i \times \lambda_{sp} \quad \text{for } i = 1, \dots, N \tag{4.10}$$

where n_i is number of solar panels functioning in state i and λ_{sp} is the failure rate of a single solar panel.
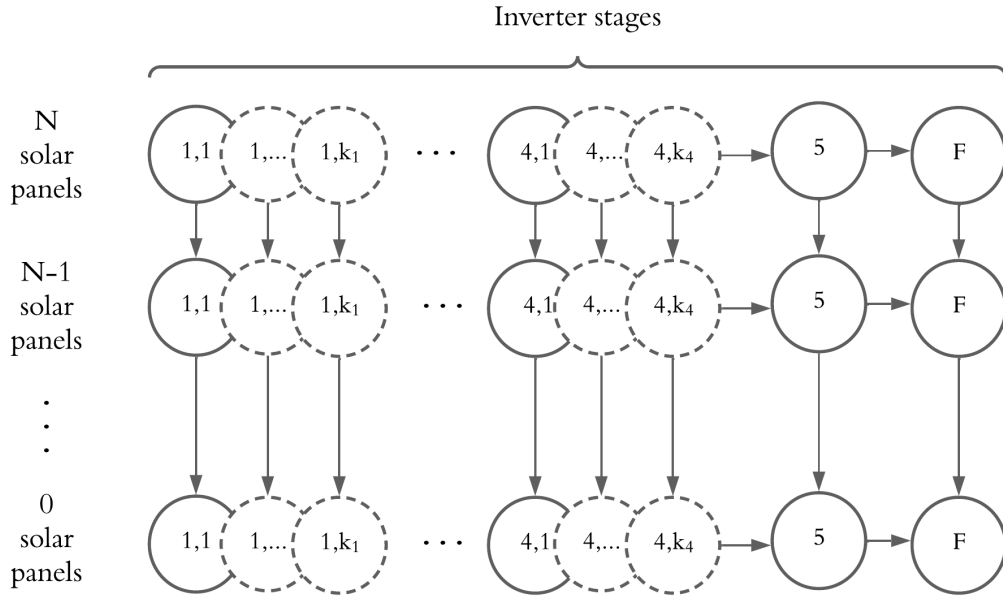
Figure 4.4 Subsystem model with string inverter
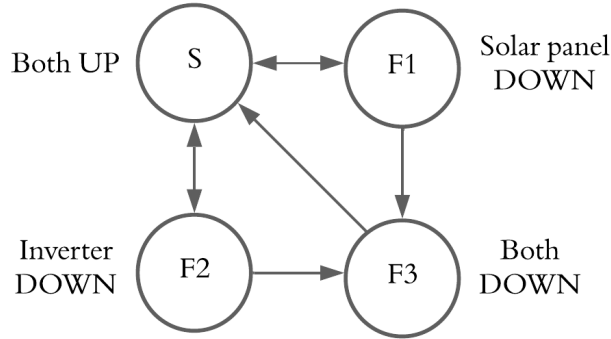
Figure 4.5 Subsystem model with microinverter

The repair transitions are not shown in the diagram. Solar panels repairs take place only during scheduled maintenance. This would be represented by a transition from each state where there is a failed solar panel to the first state of the corresponding column in which is located, with rate given by the mean time to maintenance. In case of an inverter repair, the process moves horizontally to the left, arriving at the first state of the row. The rate at which this transition occurs is determined by how long it takes to solve inverter-related hardware issues.

The subsystem model with microinverter is shown in Figure 4.5. Since information at component level is available, the failure of the solar panels or the microinverter triggers a repair action that will restore the system to the state where all components are functional. All repair transitions are equal and its rate is given by the time it takes to repair the inverter.

4.4   Markov Reward Models

A Markov Reward Model (MRM) or Markov Reward Process is a mathematical framework to analyze stochastic systems that adhere to the Markov property and accumulate some prize or reward by staying in a state or transiting to another [23]. In the context of a PV system, MRMs provide a flexible framework that we can use to estimate the expected energy yield.

An MRM consists of the following components:

1. State space (S): A finite or countable set of states $s_1, s_2, s_3, ..., s_K$.

2. Transition rate matrix (A): A matrix that describes the instantaneous rate of transiting to another state.

3. Reward matrix (R): A matrix that contains the different rewards the process obtains by staying in its current state or transitioning to another. The rewards accrued per unit of time when the system is in the state s_i is r_{ii}. Alternatively, the prize won by transitioning from state i to state j is denoted as r_{ij}.

The total expected accumulated reward until time t when starting from state i is $v_i(t)$. The solution of the system of differential equations with zero as the initial condition provides the accumulated expected reward:

$$\frac{dv_i(t)}{dt} = z_i + \sum_{j=1}^{K} q_{ij} v_j(t), \quad i = 1, \ldots, K \tag{4.11}$$

$$z_i = r_{ii} + \sum_{\substack{j=1 \\ j \neq i}}^{K} q_{ij} r_{ij}, \quad i = 1, \ldots, K \tag{4.12}$$

4.5 Rewards for estimating energy yield

For a PV system, the reward matrix specifies the energy production in each state per unit of time. Energy production in a given state is determined by the number of solar panels functioning and if whether the inverter is operational. Off-diagonal values of the reward matrix are zeros, since there is no reward associated with transitions between states. Equation 4.13 denotes the amount of energy produced in each state.

$$E_i = \delta \times \min(P_{in}, P_{sp} \times (N - M)) \times E_{avg} \tag{4.13}$$

where E_i is the energy production of the i-th state, δ is an indicator variable with a value of 1 if the inverter is operative and 0 if it is not, P_{in} is the inverter rated power, P_{sp} is the solar panel rated power, M is the number of solar panels not working in the i-th state. E_{avg} is the average yearly energy output of a perfectly reliable PV system per kW_{ac} installed.

The rewards matrix can be set up in a way that also accounts for time-dependent changes in the system. It can decrease sequentially to account for the degradation of solar panels, providing a more comprehensive description of the performance of the system.

4.6 Results and discussion

We will now compare the long term performance of two 8 kW systems, one with string inverters and the other with microinverters. The system configurations under study are:

1. One 8 kW string inverter and twenty 400 W solar panels connected to it.

2. Twenty 400 W microinverters, each connected to one 400 W solar panel.

No shading effects are assumed, so the performance of both systems neglecting failures is the same; any differences in performance will be because of reliability. The yearly energy production is 1375 kWh/kW. Solar panel degradation is assumed to be 0.5% yearly.
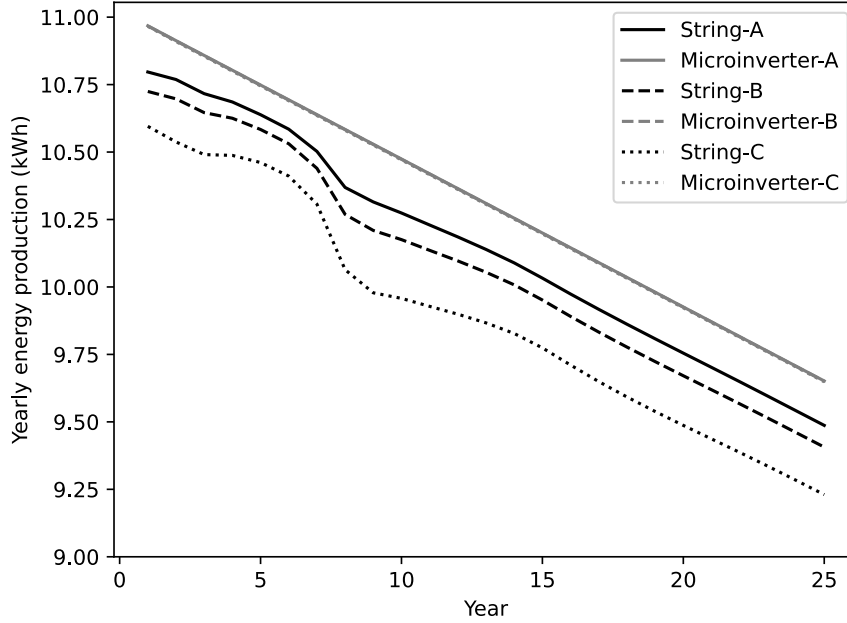
Figure 4.6 Yearly energy production of both system configurations for varying inverter mean time to repair (A=40 days, B=60 days, and C=105 days)

The energy production of both systems is shown in Figure 4.6 for different values of inverter mean time to repair, which for the microinverter system is the time to repair of both solar panel and microinverter. As expected, the higher reliability of microinverters grants them higher energy production overall. It is worth pointing out that the string system production is heavily influenced by the decreasing and increasing failure rate behavior of the inverter; this manifests as a reducing the gap in energy production between the two systems within the first 5 to 6 years, and then increasing it until the difference settles. Another key takeaway from Figure 4.6 is that the string system energy production is very sensitive to the mean time to repair of inverters, whereas microinverter systems are so reliable that even for

higher time to repairs the energy production results unaffected. These repair values were selected from [8] and they and represent the mean, 75th percentile, and max value of the recorded times to solve inverter-related hardware issues.

CHAPTER 5

BAYESIAN INVERTER RELIABILITY MODELING

In Chapter 3, we established time-to-failure distributions using maximum likelihood estimation, as is customary in traditional reliability and lifetime data analysis. In this approach, the notion of probability is frequency-based; that is, it derives from the understanding of probability as the limiting relative frequency of an event in a repeated series of identical trials. An alternative to this notion is Bayesian probability, which differs from the frequency-based paradigm by considering probability as a subjective assessment of the state of knowledge. This philosophical difference enables the use of information that would otherwise remain unutilized, thereby enhancing model building. In Bayesian models, parameters are treated as random variables about which we make probabilistic statements, providing a robust framework to quantify uncertainty—especially vital in the absence of abundant data. In this chapter, we revisit the identification of the PWE distribution from a Bayesian perspective for two reasons. First, we aim to analyze the impact of uncertainty stemming from model parameters on the expected energy production and provide confidence bounds for these estimates. Secondly, we aim to analyze the failure rate as a key indicator of the failure process by employing a PWE distribution with change points evenly placed at distances sufficient to capture how the process varies over time.

5.1 Basics of Bayesian analysis

In Bayesian analysis, probability models are composed of two parts: prior and likelihood. The prior represents our knowledge about the parameters before any data is used. On the other hand, the likelihood function is obtained from the sampling distribution, which essentially describes the probability of observing the data given a set of parameter values. These two elements give rise to the posterior distribution through the application of the Theorem of Bayes, shown in equation 5.1 [37].

$$p(\theta|y) = \frac{f(y|\theta)p(\theta)}{m(y)} \tag{5.1}$$

where $p(\theta|y)$ is the posterior distribution, $p(\theta)$ is the prior density, $m(y)$ is the marginal density of the data, and $f(y|\theta)$ is the likelihood.

5.1.1 Priors

For every parameter θ_i in a Bayesian model, we need to provide a prior. The prior encapsulates an initial plausibility assignment for each possible value the parameter can take. Priors can serve other purposes such as constraining the parameters to be within a reasonable range, taking advantage of any insight we may have about what the correct value should be. Priors can be informative and noninformative; the former is used when it is known that the parameter is more likely to take certain values, whereas the latter is used when very little information is known about the parameter; an example of a noninformative prior would be to use a uniform distribution from zero to one when estimating a population proportion. Since priors can and will affect the posterior, it is necessary to do a prior sensitivity check to gauge how robust the posterior is to changes in the priors [38].

5.1.2 Likelihood

The likelihood specifies the plausibility of observing each sample in the dataset, and it is expressed in terms of the parameters of the model. For instance, if we are interested in the proportion of the number of heads to the total number of tosses, we would choose a binomial probability mass function as the likelihood function. In fact, the likelihood function is what Bayesian analysis has in common with its frequency-based counterpart, and it is the reason why, as the sample size grows, the results of both approaches become very similar [38].

5.1.3 Posterior

Obtaining the posterior consists of updating our beliefs contained in the priors. This updating procedure is carried out by applying equation 5.1. Closed-form expressions for $p(\theta|y)$ are only possible when conjugate priors are used, since finding $m(y)$ is numerically intractable in most cases. Fortunately, Markov Chain Monte Carlo (MCMC) methods can be used to sample from complex posteriors and generate sequences of parameter values upon which inferences can be based. Probabilistic programming tools such as Stan have built-

in samplers that can be used to generate a sample of the posterior for a wide variety of models, so that all the effort can be devoted to finding appropriate priors and writing out the likelihood function, if it is not already implemented in the chosen tool [39].

5.1.4   Parameter uncertainty

The frequency-based way of thinking about parameter uncertainty is based on the sampling distribution of the estimator being used. However, deriving analytical expressions of the sampling distribution for every model is not always possible, and in these cases, inferences about the parameter must rely on asymptotic results—approximating the sample distribution of the parameter when the sample size is "large." The MLE of a parameter θ for a large sample size is approximately normally distributed, with a mean equal to θ and a variance equal to the negative reciprocal of the second derivative of the log-likelihood evaluated at the MLE. Confidence bounds derived following this approach are not probability statements about parameter uncertainty, but rather statements based on repeated sampling that may not even make sense in most scenarios, especially in those where data is limited [37].

Unlike frequency-based confidence bounds, Bayesian probability does provide probability statements about parameter uncertainty, also called credible confidence bounds [38]. Additionally, these probability statements can be easily propagated through complex models, a task that is difficult and sometimes impossible with frequency-based confidence intervals [37].

5.2   Rethinking the PWE distribution under Bayes

Bayesian probability offers a framework to incorporate uncertainty into the reliability and performance analyses conducted so far. The expected energy production is a fixed value once the parameters have been specified; however, if there is uncertainty in the parameters, it becomes a random variable. Using Bayesian credible confidence bounds, we can offer a confidence interval for the expected energy production.

5.2.1   Building the Model

In this section, we will assume that the location of the change points is the same as those in the model of Chapter 3; hence, the only parameters are the failure rates. Despite knowing

the MLE of the parameters, it is desirable to introduce as little bias as possible into the priors because they can heavily influence the posterior. As an initial prior, we have assumed a bathtub-shaped behavior, summarized in Table 5.1. The lowest failure rate is chosen as the inverse of ten years, a reference value used by most manufacturers to establish warranties for string inverters, and the standard deviation is set large enough to allow the sampler to explore the search space.

Table 5.1 Initial priors

Parameter	Prior
λ_1	$\mathcal{N}(4.1096 \times 10^{-4}, 2.25 \times 10^{-8})$
λ_2	$\mathcal{N}(3.4246 \times 10^{-4}, 2.25 \times 10^{-8})$
λ_3	$\mathcal{N}(2.7397 \times 10^{-4}, 2.25 \times 10^{-8})$
λ_4	$\mathcal{N}(3.4246 \times 10^{-4}, 2.25 \times 10^{-8})$
λ_5	$\mathcal{N}(4.1096 \times 10^{-4}, 2.25 \times 10^{-8})$

The likelihood function of the PWE distribution is shown in equation 5.2 [40].

$$
\begin{aligned}
h(t) &= \lambda_i, \quad t \in I_i, \quad i = 1, \ldots, r+1, \\
L(t, \lambda) &= \prod_{i=1}^{N} \lambda_i \cdot d_i \cdot \exp(-\lambda_i \cdot a_i), \\
d_i &= \sum_j x_{ij}, \\
a_i &= \sum_j (t_{ij} - t_{i-1,j})
\end{aligned}
\tag{5.2}
$$

where I_i is the interval $(t_{i-1}, t_i]$, l_i is the likelihood contribution of all data points through λ_i, d_i is the number of failures in interval i, and a_i is the exposure time in interval i.

5.2.2 Estimation results

The model estimation is implemented in RStan, which is an interface in R to Stan, a C++ library for Bayesian inference using the No-U-Turn sampler [39]. The estimation summary is shown in Table 5.2.

The resulting marginal posterior for all parameters has a lower standard deviation than the priors, which indicates that uncertainty was reduced after the Bayesian update. However,

Table 5.2 Posterior summary statistics using normal prior

λ	Mean	SD	2.5%	25%	50%	75%	97.5%
1	5.429E-04	8.056E-05	3.945E-04	4.868E-04	5.393E-04	5.958E-04	7.116E-04
2	3.218E-04	3.210E-05	2.622E-04	2.996E-04	3.206E-04	3.428E-04	3.877E-04
3	1.923E-04	2.887E-05	1.401E-04	1.722E-04	1.908E-04	2.109E-04	2.530E-04
4	5.906E-04	1.003E-04	4.032E-04	5.217E-04	5.870E-04	6.562E-04	7.975E-04
5	4.961E-04	1.180E-04	2.774E-04	4.136E-04	4.934E-04	5.747E-04	7.335E-04

it is important to verify to what extent this reduction was influenced by the choice of priors. To assess the sensitivity of the posterior to changes in the priors, we explore more candidate prior distributions with the same expected value: exponential and gamma. The latter can be considered a small perturbation, whereas the former represents a large perturbation. For the gamma distribution, the parameters α and λ have been set to match the mean and to have a standard deviation closer to that of the normal priors; the exponential distribution is specified by just the mean. The posterior summary statistics using each prior are shown in Tables 5.3 and 5.4.

Table 5.3 Posterior summary statistics using gamma prior

λ	Mean	SD	2.5%	25%	50%	75%	97.5%
1	5.464E-04	8.859E-05	3.882E-04	4.836E-04	5.418E-04	6.032E-04	7.341E-04
2	3.185E-04	3.152E-05	2.598E-04	2.964E-04	3.174E-04	3.393E-04	3.825E-04
3	1.891E-04	2.819E-05	1.384E-04	1.695E-04	1.875E-04	2.072E-04	2.487E-04
4	6.578E-04	1.364E-04	4.187E-04	5.608E-04	6.481E-04	7.445E-04	9.485E-04
5	4.977E-04	1.384E-04	2.660E-04	3.972E-04	4.858E-04	5.834E-04	8.031E-04

Table 5.4 Posterior summary statistics using exponential prior

λ	Mean	SD	2.5%	25%	50%	75%	97.5%
1	5.826E-04	1.020E-04	4.010E-04	5.104E-04	5.769E-04	6.478E-04	7.986E-04
2	3.177E-04	3.242E-05	2.571E-04	2.953E-04	3.166E-04	3.387E-04	3.845E-04
3	1.862E-04	2.824E-05	1.346E-04	1.666E-04	1.848E-04	2.045E-04	2.453E-04
4	8.144E-04	1.857E-04	4.939E-04	6.824E-04	7.995E-04	9.306E-04	1.219E-03
5	6.064E-04	2.305E-04	2.450E-04	4.375E-04	5.779E-04	7.421E-04	1.147E-03

The summary statistics of all three choices of prior are very similar, which indicates that the model setup is not very sensitive to changes in the prior. The only noticeable difference is in λ_5 of the exponential prior—slightly more skewed than for the other priors. This skewness

can be attributed to the likelihood not being as influential as in the rest of the intervals, making the prior set the shape of the posterior.
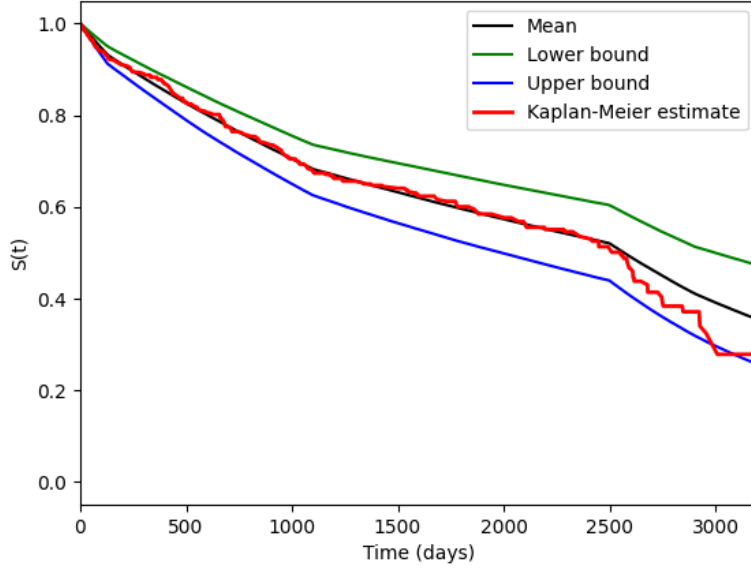

Figure 5.1 Comparison of the survivor function with the Kaplan-Meier estimate for a normal prior

Another way of assessing model quality is by plotting the survival function with its confidence bounds on top of the Kaplan-Meier estimate, as shown in Figures 5.1 to 5.3. For all three priors, the survival functions look identical up until 2500 days, but they differ in how they behave at the tail. Using a normal prior, we get the smoothest tail, even leaving a portion of the Kaplan-Meier estimate outside the 95% confidence bounds. The tail of the gamma prior is slightly less smooth, yet it completely encloses the Kaplan-Meier estimate. Lastly, with the exponential prior, we observe the roughest tail, with higher failure rate estimates in the last two intervals. Tail robustness shown by the distribution with normal priors is desirable, and usually, control actions are taken to ensure this [33].

We will choose the model with the normal prior as the final model, although the gamma prior could also have been chosen.
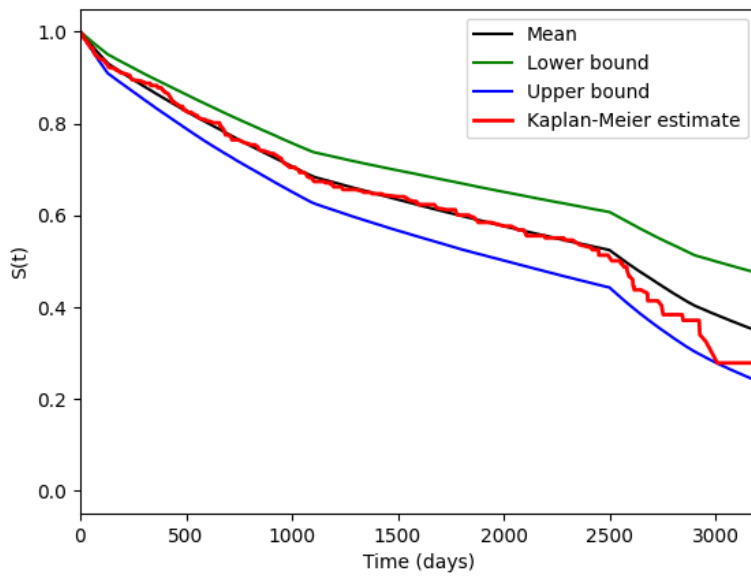
Figure 5.2 Comparison of the survivor function with the Kaplan-Meier estimate for a gamma prior
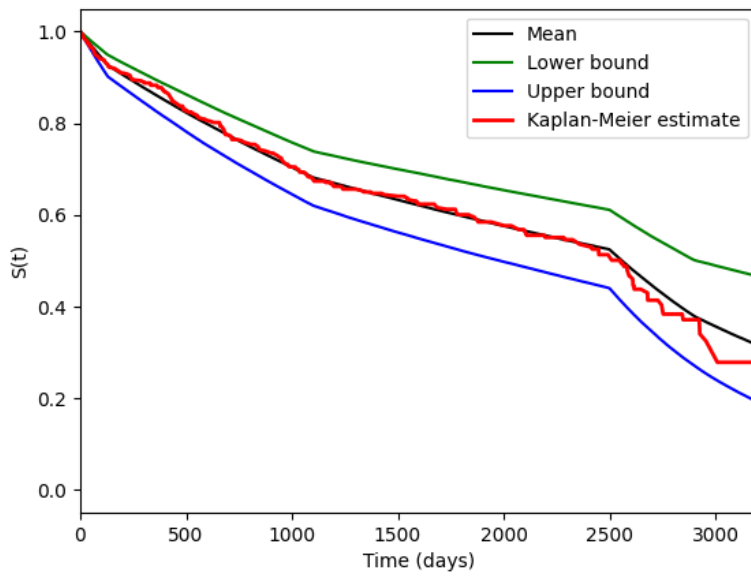

Figure 5.3 Comparison of the survivor function with the Kaplan-Meier estimate for an exponential prior

5.2.3 Expected energy production with parameter uncertainty

Accounting for parameter uncertainty using a Bayesian time-to-failure model is straightforward. From the generated posterior samples, we can select the set of parameters that define the confidence region of interest. For the sake of simplicity, we will assume a two-sided equal-tailed confidence interval of 95%; thus, we will have a set of parameters for the 2.5% bound and another for the 97.5%. Thereafter, we compute the expected energy production for each of these two sets of parameters, following the same procedure outlined in Chapter 4.

The numerical results presented next are based on the case study used in Chapter 4. Figures 5.4 to 5.6 depict the expected energy production for the test system, assuming an inverter mean time to repair of 40, 60, and 105 days.
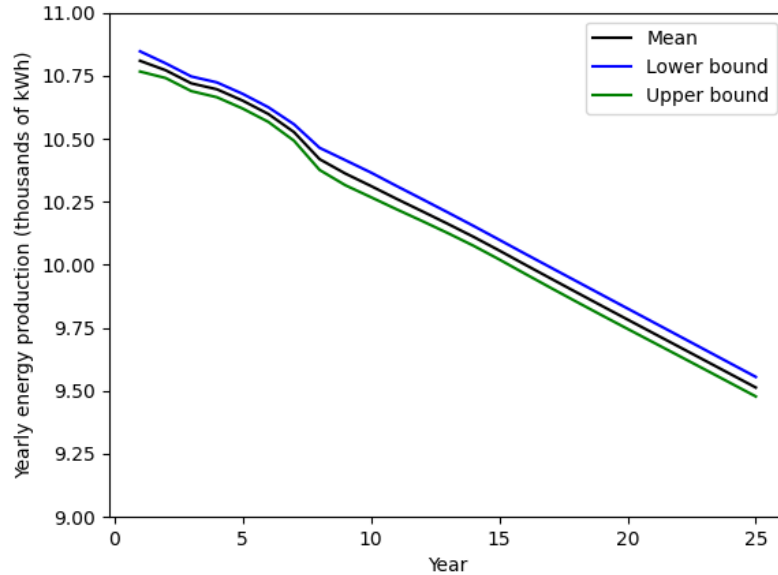

Figure 5.4 Expected energy production with a 40-day mean time to inverter repair.

Figures 5.4 to 5.6 suggest that even though there is significant uncertainty associated with model parameters, its effects on the expected energy production of the system are not significant, due to the rapid execution of repairs. Nevertheless, it is important to acknowledge the ease with which it is possible to account for parameter uncertainty in this Markov-based
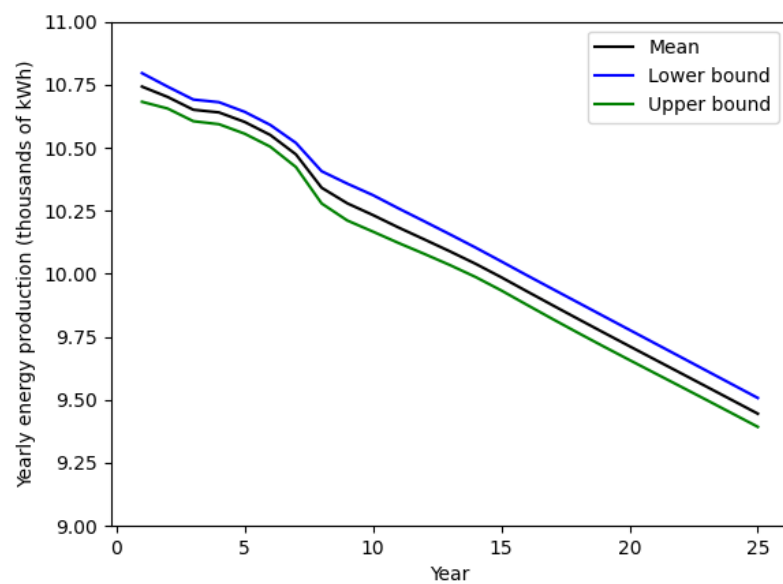
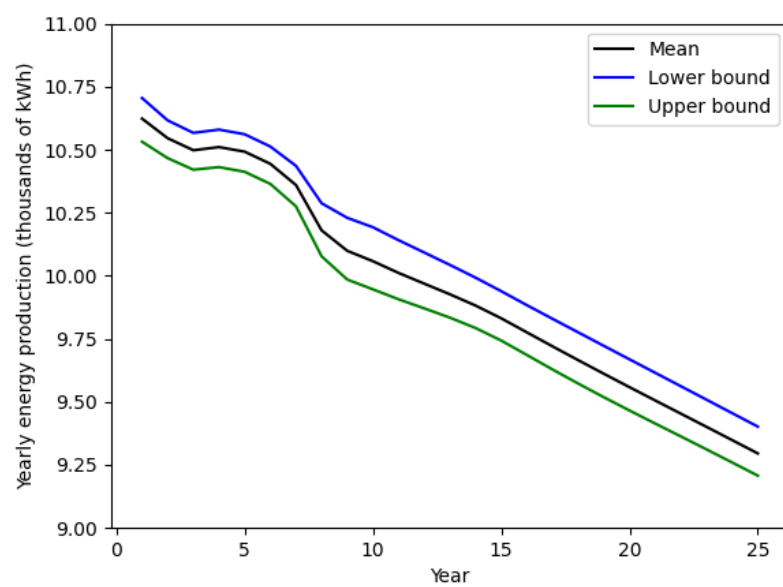Figure 5.5 Expected energy production with a 60-day mean time to inverter repair.

Figure 5.6 Expected energy production with a 105-day mean time to inverter repair.

reliability analysis using a Bayesian approach for time-to-failure modeling.

5.3 Interval-based failure rate analysis

The failure rate is a crucial metric for non-repairable components, not only because it provides a more intuitive interpretation of the failure process but also because it represents the conditional probability of failure given survival up to a certain point in time. Thus, knowing its value can contribute to more effective risk management [41].

The objective of the failure modeling so far has been to find a compact model for integration into a system reliability model, and as such, the behavior of the failure rate has been simplified as much as possible to reduce the complexity of the model. Now, we aim to capture the failure rate behavior with arbitrary granularity. Fitting such a model using MLE is challenging because it uses limited information to estimate the failure rate for each interval. For instance, we will use MLE to find a yearly failure rate model with change points every year from the first to the eighth year. Figure 5.7 illustrates the goodness of fit of this model in terms of the survival function. It is evident that the model overestimates the failure rates, yielding an overly pessimistic outlook. To further illustrate, consider a model with change points placed every 100 days; the fit of this model is depicted in Figure 5.8, and as expected, it performs even worse.

The reason why MLE is not stable is because it uses only local information to estimate the failure rate, and since the data is spread unevenly, it is not possible to provide reasonable estimates in every interval, especially those in which data is very scarce. This instability could be overcome by including a mechanism for exchanging information across intervals or by implementing some form of regularization. Using a simple Bayesian approach, as demonstrated in the previous section, does provide a good model for yearly intervals, shown in Figure 5.9. However, decreasing the interval spacing to 100 days results in a model with wide and noisy confidence bounds, as seen in Figure 5.10. Also, a model with this many change points requires the specification of numerous priors, increasing the modeling burden.
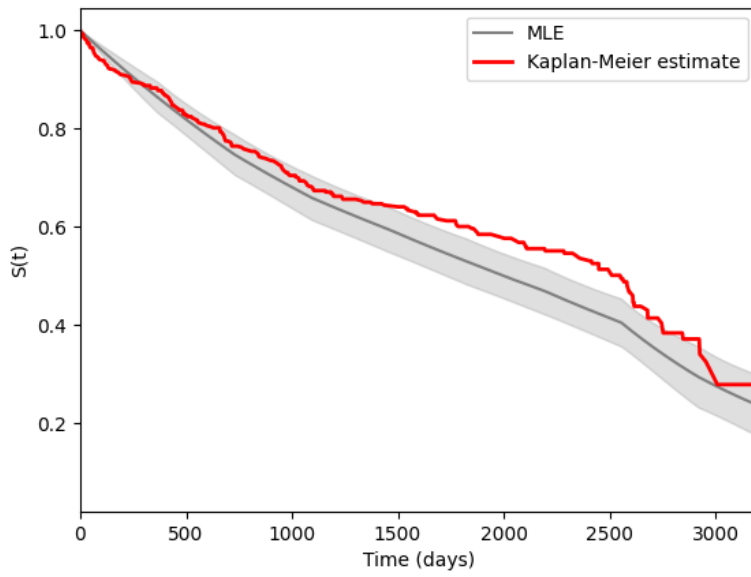
Figure 5.7 MLE estimation of PWE distribution: change points every year
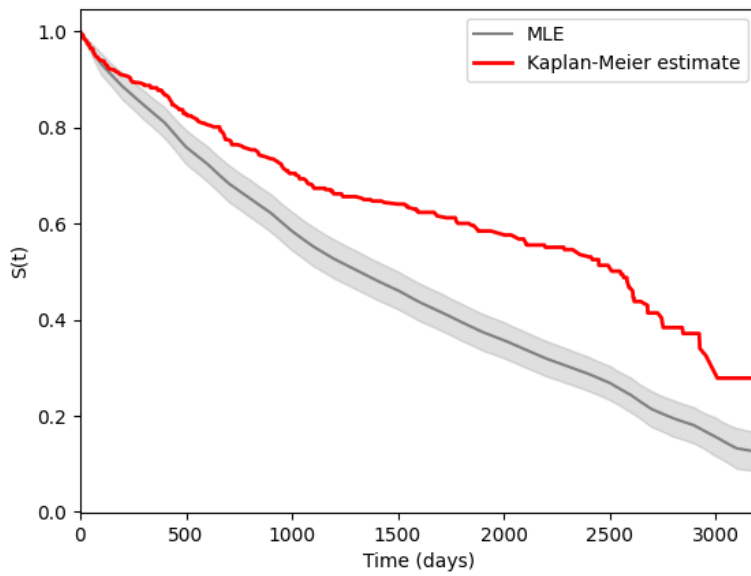

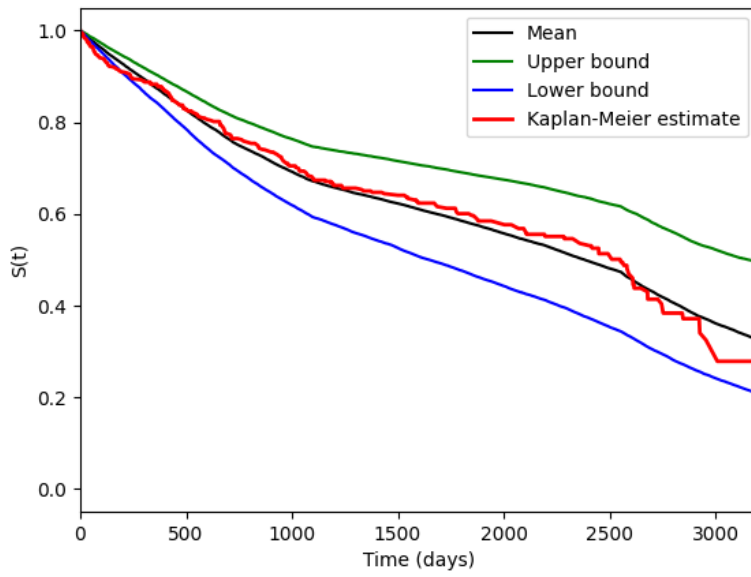Figure 5.8 MLE estimation of PWE distribution: change points every 100 days

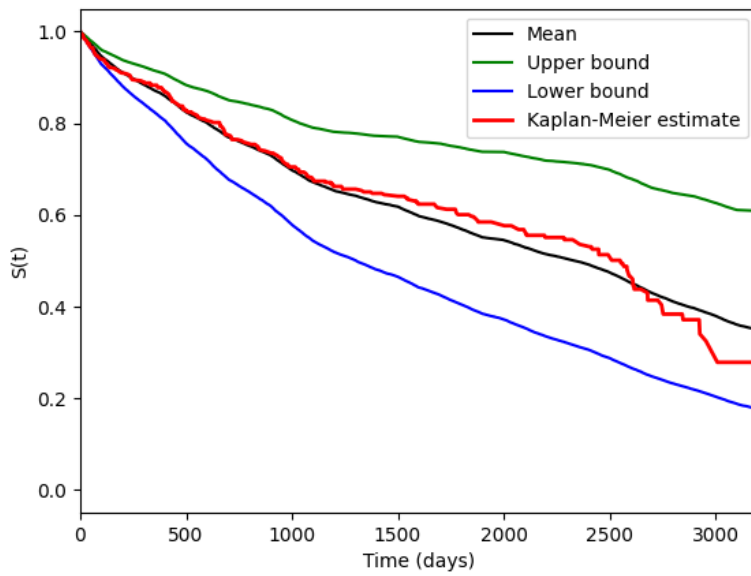Figure 5.9 Bayesian estimation of PWE distribution: change points every year

Figure 5.10 Bayesian estimation of PWE distribution: change points every 100 days

5.3.1 Relating Consecutive Intervals

Using a Bayesian framework and probabilistic programming tools like Stan, we can model complex interactions between model parameters to find stable estimates of the failure rate for any granularity of change point placement. Essentially, we aim to relate failure rates across intervals in a manner that leverages the underlying nature of the failure process—a gradual degradation. We propose relating consecutive rates using equation 5.3; assuming that the next rate is normally distributed with mean given by the previous rate discourages large variations. Now, information flows across intervals, strengthening the failure rate estimation for intervals that lack sufficient data to produce a reliable estimate using local data alone. Another advantage of this approach is that only one prior is needed: the initial failure rate; thereby reducing the complexity of prior sensitivity analysis.

$$\lambda_i \sim \mathcal{N}(\lambda_{i-1}, \sigma^2) \quad i = 2, \ldots, r+1, \tag{5.3}$$

where σ is a tuning parameter.

5.3.2 Model with Change Points Every 100 Days

The change points are now placed every 100 days until 3100 days, thus the model includes a total of 31 change points. One aspect to keep in mind is that there are intervals where no failures occurred. The prior for the initial rate is $\mathcal{N}(4.1096 \times 10^{-4}, 2.25 \times 10^{-8})$—similar to what we assumed for the compact model—and σ is 1×10^{-4}.

The survivor function, plotted alongside the Kaplan-Meier estimate, is depicted in Figure 5.11; the model exhibits good visual agreement with the Kaplan-Meier estimate and tighter, smoother confidence bounds. The failure rate estimates are shown in Fig. 5.12.

5.3.3 Model with change points every 50 days

The granularity of the change points can be further increased without compromising the quality of the resulting distribution. Now, we present results for an even placement every 50 days-a more complex task due to the increased number of intervals with zero failures. The prior for λ_1 and the value of σ are the same as in the previous case.
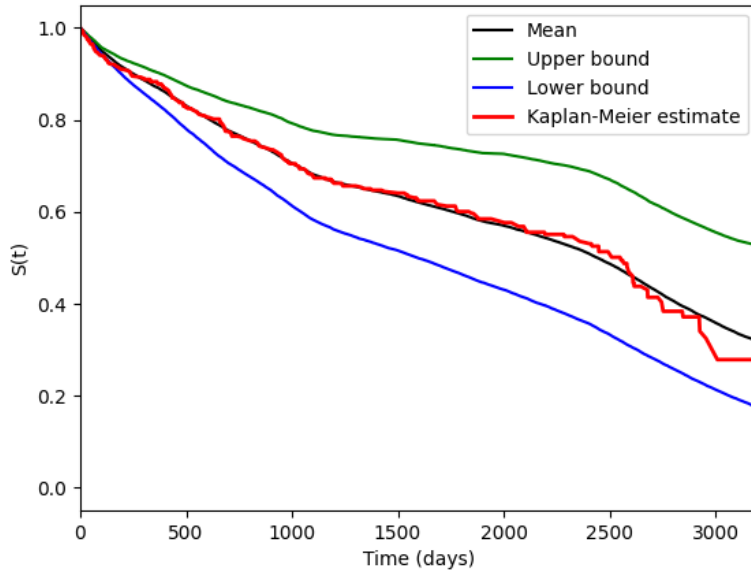
Figure 5.11 Survivor function of PWE distribution with changepoints placed every 100 days
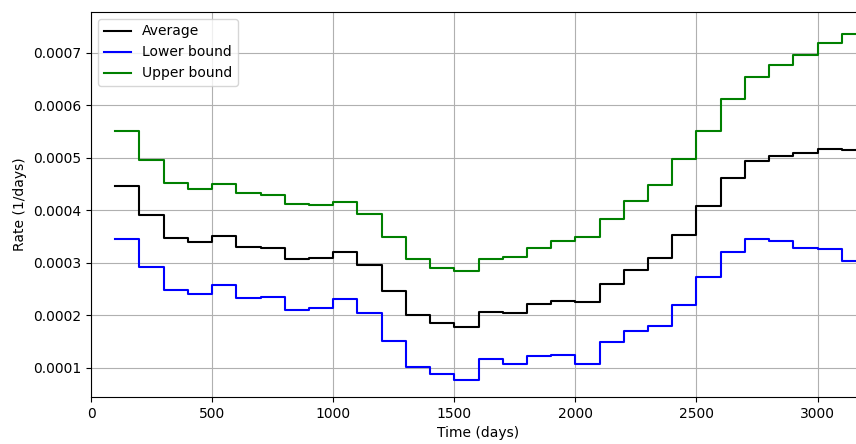
Figure 5.12 Failure rate estimates of PWE distribution with change points placed every 100 days

Figure 5.13 Survivor function of PWE distribution with changepoints placed every 50 days

Figure 5.14 Failure rate estimates of PWE distribution with change points placed every 50 days
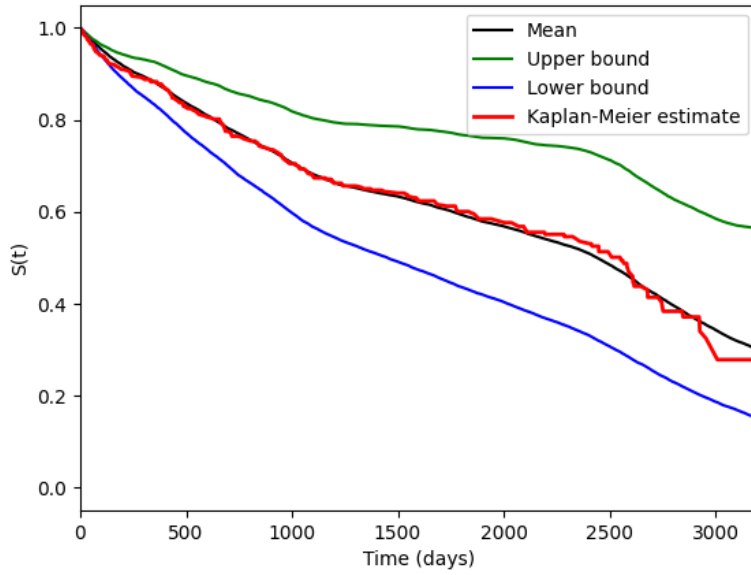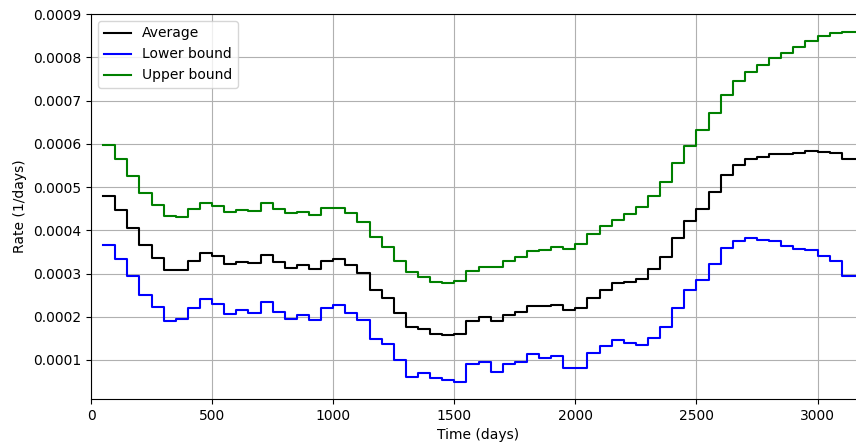
Figures 5.13 and 5.14 depict the survivor function plotted against the Kaplan-Meier estimate and the failure rate estimates, respectively, confirming that robust estimates can also be obtained using this approach for change points every 50 days.

CHAPTER 6

CONCLUSIONS

In this thesis, three main topics have been discussed: developing reliability models from scarce and censored data, converting a non-exponential reliability model into a Markov model amenable to conventional reliability analysis, and leveraging Bayesian probability to account for parameter uncertainty and provide detailed estimates of failure rates.

Using field data instead of test data comes with many challenges, such as insufficient sample sizes and censoring, which complicate the modeling process. Moreover, when the data do not follow typical behavior, employing well-known probability distributions in power systems reliability—such as Weibull, exponential, or log-normal—is not possible. To overcome this, we proposed using a piecewise exponential distribution, more commonly found in settings like clinical trials, which can inherently accommodate all kinds of failure rate behaviors and is more suitable for fitting the data at hand.

We found that not only did it provide a good fit, but it also related to the notion of viewing the failure process as a progression through discrete stages, which is the basis for state-space reliability methods. However, unlike most reliability models in the power systems literature, a model based on the PWE distribution was not memoryless, necessitating a more advanced modeling framework than continuous-time Markov chains. The model was then framed as a Semi-Markov process. By solving the Markov renewal equation, we demonstrated how the availability of the string inverter changes over time when an exponential time to restore is assumed. However, this raised the concern that a whole-system model could not be framed as a Semi-Markov process due to the absence of a global clock to track progress.

The conversion of a non-exponential process into an exponential model is referred to as Markovization, which is based on the use of phase-type distributions. By using these distributions, we approximated each deterministic transition of the Semi-Markov process as a series of exponentially distributed stages, all with the same rate; this sum of exponential stages can be represented by an Erlang distribution. Using the expressions for the mean

and variance of an Erlang distribution, we determined the rate corresponding to a number of stages, which ultimately controls the accuracy of the approximation. Numerical results show that using a model with 252 states provides a sufficiently accurate representation for long-term analysis.

In Chapter 2, we briefly discussed the modeling of small PV systems, focusing on residential systems due to their simplicity in design and the economic rationale behind using string inverters. We proposed decomposing the system model into independent subsystems comprised of one string inverter and the solar panels connected to it, in line with how monitoring systems operate and repairs are carried out. Similarly, for microinverter systems, which are recognized for their higher reliability, we argued that a subsystem can consist of the microinverter and the connected solar panel. This arrangement not only simplified our analysis but also lets us compare the reliability of string inverters with that of microinverters.

Having developed a model for the string inverter, we constructed a completely memoryless subsystem model, suitable for Markov-based reliability analysis. Although reliability is traditionally compared in terms of availability, it is more appropriate to compare the string inverter and microinverter subsystems in terms of expected energy production. For this purpose, we utilized Markov Reward Models, where the rewards represent energy production. Our results indicate that the energy production of microinverters is not sensitive to changes in the mean time to repair, unlike that of string inverters.

The last chapter of this thesis adopted a Bayesian approach to the estimation procedure outlined in Chapter 3. By doing so, we were able to provide confidence bounds for the energy production of the string inverter subsystem without any major changes to the modeling framework itself. Applied to the case study of Chapter 4, we see that, although there is significant uncertainty in the model parameters, most of it is absorbed by the promptness of the repairs. Another strength of Bayesian probability is how it allows for flexible model representation, enabling the modeling of complex interactions between model parameters. Leveraging this approach, we obtained failure rate estimates with minimal prior informa-

tion for arbitrarily placed changepoints, confirming the bathtub-shaped behavior previously suggested by the cumulative failure rate. This analysis is not only valuable for risk assessment and reliability optimization but also serves as an initial method to describe the failure process, facilitating subsequent simplifications based on observed behaviors.

Data on the operation of small PV systems is not widely available, which limits our understanding of their actual reliability. In this work, we attempted to characterize the behavior of string inverters using a small dataset collected from a single region in the U.S. However, this dataset is by no means representative of all string inverters, which constrained the depth of the insights we could draw. To support more conclusive research in this area, a more comprehensive data collection platform must be established.

A robust reliability assessment would require access to additional information such as voltage and current waveforms, environmental parameters like temperature and humidity, and the internal temperature of the device. Moreover, greater effort must be made to accurately categorize failure causes. This would enhance reliability analysis and enable root-cause investigations that improve failure modeling. Such information would not only deepen our understanding of component reliability but also allow us to leverage failure data for prognostic purposes—ultimately enabling the prediction of failures and the implementation of corrective measures to prevent them.

BIBLIOGRAPHY

[1] T. Gunda, S. Hackett, L. Kraus, C. Downs, R. Jones, C. McNalley, M. Bolen, and A. Walker, "A machine learning evaluation of maintenance records for common failure modes in PV inverters," *IEEE Access*, vol. 8, pp. 211 610–211 620, 2020.

[2] T. J. Formica, H. A. Khan, and M. G. Pecht, "The effect of inverter failures on the return on investment of solar photovoltaic systems," *IEEE Access*, vol. 5, pp. 21 336–21 343, 2017.

[3] D. Feldman, J. Zuboy, K. Dummit, D. Stright, M. Heine, S. Grossman, and R. Margolis, "Spring 2024 Solar Industry Update," National Renewable Energy Laboratory, Tech. Rep. NREL/PR-7A40-90042, Jun. 2024.

[4] Solar Energy Industries Association, "Solar Market Insight Report Q1 2024," 2024, accessed: 2024-07-16. [Online]. Available: https://www.seia.org/us-solar-market-insight

[5] A. Golnas, "PV System Reliability: An operator's perspective," *IEEE Journal of Photovoltaics*, vol. 3, no. 1, pp. 416–421, 2013.

[6] T. Gunda and R. Homan, "Evaluation of component reliability in photovoltaic systems using field failure statistics," Sandia National Lab. (SNL-NM), Albuquerque, NM (United States), Tech. Rep., 09 2020. [Online]. Available: https://www.osti.gov/biblio/1660804

[7] G. T. Klise, O. Lavrova, and R. L. Gooding, "PV System Component Fault and Failure Compilation and Analysis," Sandia National Lab. (SNL-NM), Albuquerque, NM (United States), Tech. Rep., 02 2018. [Online]. Available: https://www.osti.gov/biblio/1424887

[8] D. C. Jordan, B. Marion, C. Deline, T. Barnes, and M. Bolinger, "PV field reliability status—Analysis of 100 000 solar systems," *Progress in Photovoltaics: Research and Applications*, vol. 28, no. 8, pp. 739–754, 2020. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/pip.3262

[9] T. Doyle, R. Desharnais, and M. Mills-Price, "2019 PV Inverter Scorecard," PVEL LLC, Technical Report, 2019. [Online]. Available: https://www.pvel.com/inverter-scorecard/

[10] S. Peyghami, M. Fotuhi-Firuzabad, and F. Blaabjerg, "Reliability evaluation in microgrids with non-exponential failure rates of power units," *IEEE Systems Journal*, vol. 14, no. 2, pp. 2861–2872, 2020.

[11] J. Cheng, Y. Tang, and M. Yu, "The reliability of solar energy generating system with inverters in series under common cause failure," *Applied Mathematical Modelling*, vol. 68, pp. 509–522, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0307904X18305687

[12] M. Theristis and I. A. Papazoglou, "Markovian reliability analysis of standalone photovoltaic systems incorporating repairs," *IEEE Journal of Photovoltaics*, vol. 4, no. 1, pp. 414–422, 2014.

[13] S. V. Dhople, A. Davoudi, P. L. Chapman, and A. D. Domínguez-García, "Integrating photovoltaic inverter reliability into energy yield estimation with markov models," in *2010 IEEE 12th Workshop on Control and Modeling for Power Electronics (COMPEL)*, 2010, pp. 1–5.

[14] X. Yu and A. M. Khambadkone, "Reliability analysis and cost optimization of parallel-inverter system," *IEEE Transactions on Industrial Electronics*, vol. 59, no. 10, pp. 3881–3889, 2012.

[15] Wenyuan Li, "Incorporating aging failures in power system reliability evaluation," *IEEE Transactions on Power Systems*, vol. 17, no. 3, pp. 918–923, Aug. 2002. [Online]. Available: http://ieeexplore.ieee.org/document/1033745/

[16] H. Kim and C. Singh, "Reliability modeling and simulation in power systems with aging characteristics," *IEEE Transactions on Power Systems*, vol. 25, no. 1, pp. 21–28, 2010.

[17] P. Jirutitijaroen and C. Singh, "The effect of transformer maintenance parameters on reliability and cost: a probabilistic model," *Electric Power Systems Research*, vol. 72, no. 3, pp. 213–224, 2004. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378779604001129

[18] H. Toftaker and I. B. Sperstad, "Integrating component condition in long-term power system reliability analysis," Jun. 2022. [Online]. Available: https://www.techrxiv.org/doi/full/10.36227/techrxiv.19977497.v1

[19] J. H. Jürgensen, *Condition-based Failure Rate Modelling for Individual Components in the Power System.* Stockholm: KTH Royal Institute of Technology, 2016.

[20] Hagkwen Kim and C. Singh, "Reliability Modeling and Simulation in Power Systems With Aging Characteristics," *IEEE Transactions on Power Systems*, vol. 25, no. 1, pp. 21–28, Feb. 2010. [Online]. Available: http://ieeexplore.ieee.org/document/5247018/

[21] S. Peyghami, F. Blaabjerg, and P. Palensky, "Incorporating power electronic converters reliability into modern power system reliability analysis," *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 9, no. 2, pp. 1668–1681, 2021.

[22] A. Sangwongwanich and F. Blaabjerg, "Reliability assessment of fault-tolerant power converters including wear-out failure," in *2022 IEEE Applied Power Electronics Conference and Exposition (APEC)*, 2022, pp. 300–306.

[23] A. Lisnianski, I. Frenkel, and Y. Ding, *Multi-state System Reliability Analysis and Optimization for Engineers and Industrial Managers.* London: Springer London, 2010. [Online]. Available: http://link.springer.com/10.1007/978-1-84996-320-6

[24] K. S. Trivedi and A. Bobbio, *Reliability and Availability Engineering: Modeling, Analysis, and Applications.* Cambridge University Press, 2017.

[25] A. Garro and F. Barrara, "Reliability Analysis of Residential Photovoltaic Systems," *RE&PQJ*, vol. 9, no. 1, 2011, number: 1. [Online]. Available: https://www.repqj.com

[26] A. M. Mustafa, W. A. Omran, Y. G. Hegazy, and M. Abu-Elnaga, "Reliability assessment of grid connected photovoltaic generation systems," in *2015 International Conference on Renewable Energy Research and Applications (ICRERA)*, 2015, pp. 1543–1549.

[27] S. V. Dhople and A. D. Dominguez-Garcia, "Estimation of photovoltaic system reliability and performance metrics," *IEEE Transactions on Power Systems*, vol. 27, no. 1, pp. 554–563, 2012.

[28] M. Perdue and R. Gottschalg, "Energy yields of small grid connected photovoltaic system: effects of component reliability and maintenance," *IET Renewable Power Generation*, vol. 9, no. 5, pp. 432–437, 2015. [Online]. Available: https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/iet-rpg.2014.0389

[29] Enphase, "Reliability of Enphase Microinverters," Enphase Energy, Inc., Technical Report, 2021. [Online]. Available: https://enphase.com/download/reliability-enphase-microinverters-tech-brief?srsltid=AfmBOorK5ocHPKOgsMKiRcomFhXj62P-MqQ7r3maaueOxk_8SHXM5BL_

[30] U.S. Energy Information Administration, "EIA Form 861: Annual Electric Power Industry Report," U.S. Energy Information Administration, Data Report, 2022. [Online]. Available: https://www.eia.gov/electricity/data/eia861/

[31] M. Rausand and A. Høyland, *System Reliability Theory: Models, Statistical Methods, and Applications*, 2nd ed. Wiley-Interscience, 2004.

[32] C. Ebeling, *An introduction to reliability and maintainability engineering.* McGraw-Hill, 2004.

[33] T. Xu and R. Wen, "PWEXP: An R Package Using Piecewise Exponential Model for Study Design and Event/Timeline Prediction," 2024. [Online]. Available: https://arxiv.org/abs/2404.17772

[34] K. Gaurav, V. Kumar, and B. K. Singh, "Dependability analysis of a system using state-space modeling techniques: A systematic review," *IEEE Transactions on Reliability*, vol. 72, no. 4, pp. 1340–1354, 2023.

[35] A. Cohen, *Numerical Methods for Laplace Transform Inversion.* Springer US, 2007.

[36] M. Colledani, A. Ratti, and C. Senanayake, "An approximate analytical method to evaluate the performance of multi-product assembly manufacturing systems," *Procedia CIRP*, vol. 33, pp. 357–363, 2015, 9th CIRP Conference on Intelligent Computation in Manufacturing Engineering - CIRP ICME '14. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2212827115007258

[37] M. S. Hamada, A. G. Wilson, C. S. Reese, and H. F. Martz, *Bayesian Reliability*, 1st ed. Springer New York, NY, 2008.

[38] R. McElreath, *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*, 2nd ed. Chapman Hall/CRC, 2020.

[39] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin, *Bayesian Data Analysis*, 3rd ed. Chapman and Hall/CRC, 2025.

[40] D. Gamerman, "Bayes estimation of the piece-wise exponential distribution," *IEEE Transactions on Reliability*, vol. 43, no. 1, pp. 128–131, 1994.

[41] M. Finkelstein, *Failure Rate Modelling for Reliability and Risk.* Springer London, 2008.