NTNU – Trondheim
Norwegian University of
Science and Technology

Markov process applied to degradation modelling

Different modelling alternatives and their
properties

Zhiqi Feng

Markov process applied to degradation modelling:
Different modelling alternatives and their properties

Zhiqi Feng

Master of Science in Mathematics

Submission date: 1st of June 2013

Supervisor: Bo Henry Lindqvist, MATH

Co-supervisor: Thomas Welte, SINTEF

Norwegian University of Science and Technology

Department of Mathematical Sciences

Abstract

New scientific methods are required in industry to avoid loss of money and human lives and provide more profit. The purpose of this paper is to study stochastic models and statistical methods for analysis of lifetimes. Different alternatives of sequential continuous-time Markov-processes and semi-Markov process applied to modeling of technical degradation are considered, such as Markov processes similar to non-homogeneous Poisson process and semi-Markov process with sojourn time belonging to the Weibull distribution. A case study from an electrical network is performed. The object is an analysis of wood poles of power lines where the data from the company are panel data which contain observations over multiple time periods for the same individual pole. The wood poles are monitored by means of visual inspection techniques and their conditions are classified into five states, representing five different technical conditions. This paper describes how to predict the remaining lifetime given how long the poles have been in use and in which state they are observed. These corresponding Markov models are implemented in R-studio, parameters of the models are obtained based on the maximum likelihood methods. In order to compare and evaluate the models, likelihood ratio test statistics and hypothesis tests are applied. The simulations, based on the Markov models, are used to assess the quality of data and evaluate the goodness of models. Results from application of Markov models, discussion about advantages and disadvantages of the models, and suggestion are presented.

Preface

The present Master Thesis was carried out as the last requirement for the MSc in Department of Mathematical Sciences (MATH) in the autumn 2012 and spring 2013. The work was performed at the Norwegian University of Science and Technology (NTNU), Department of Mathematical Sciences; in collaboration with SINTEF Energy Research.

Acknowledgements

I would like to thank Professor Bo Henry Lindqvist and Thomas Welte for their committed and enthusiastic guidance during the process of developing this master project.

Table of Contents

Chapter 1: Introduction

In the last decades, several accidents, like collapse of bridges, breakdown of machines and power blackouts, have already become the point of social attention. They have resulted in the loss of human lives and extremely high cost. That is why more and more companies pay more attention to asset management. However, many kinds of factors make the asset management to be a difficult problem. For example, limited cost of maintenance, quality of maintenance and different methods of inspection will lead to different results. In recent decades, a large number of technical assets are inspected visually, such as pipelines, sewers, bridges and wood poles and the visual inspection method is the subjective interpretation of the level of deterioration of assets [1]. Such visual inspections are carried out periodically and result in subjective judgments of the asset condition. In this proposed research, conditions of components are described by discrete states where a transition from one state to the following state represents a degradation of the asset condition. The example is shown in Table 1.

Table 1: Description of state

State code	Description
0	No problems of the asset indicated (as good as new)
1	Some minor problems indicated (worse than "as good as new")
2	Serious problems indicated (poor condition)
3	Critical problems indicated (imminent failure)
4	Fault state.

For the sake of using the information gathered by visual inspections for estimating the remaining lifetime of an asset, a mathematical model should be applied to represent the deterioration process over time. A number of mathematical methods have been applied in asset management in many researches. In the literature, many people would like to use Markov process to simulate the process. Chan and Asgarpoor used the Markov process to find the optimum maintenance policy for an electric component and its failures due to deterioration are taken into consideration [2]. Xueqing and Hui used a discrete-time semi-Markov process to do the road maintenance optimization [3], and Kallen and van Noortwijk used a Markov process with sequential phases to simulate a deterioration process, which is an application of periodic inspection of objects [4]. Thus, the continuous-time Markov-process is often used as stochastic

model for mathematical representation of a degradation process, and the uncertainty and possible variation related to this process, this is also done in the following sections.

Chapter 2: Characteristics of the data

This chapter presents the characteristics of data gathered by visual inspection. Assuming no replacement and maintenance occur after installation of assets, visual inspections are assumed to be carried out with fixed time interval such as every 2nd year or every 5th year. In practice, the inspection time interval is not fixed, because it is not practical to carry out inspections of all assets at the same time, if we have more than hundreds of assets. Then the process of inspection would last for a long period. Thus the dataset would be formed with time series and each time series contains the technical condition states with an increasing number as presented in Table 2 which presents the degradation processes and the condition of assets.

Table 2: Example of the several time series with condition monitoring data

Object no.	Time of first inspection	State	Time of second inspection	State	Time of third inspection	state	……
1	1985	0	1992	1	2005	1	……
2	1988	0	1995	0	2008	1	……
3	2000	0	2003	2	2010	3	……
4	1992	0	1998	0	2008	1	……
5	1990	0	2001	0	2001	1	……
6	2000	0	2005	0	2010	0	……
7	1987	0	2000	1	2009	1	……
……	……	……	……	……	……	……	……

In statistics, the data in Table 2 refer to multi-dimensional data. We call it panel data which contain observations on multiple conditions observed over multiple time periods for the same individuals [5]. In the literature, Kallen used Markov process to analyze panel data of the bridge in Netherlands [1] and Jackson said that "Panel data are observations of a continuous-time process at arbitrary times and Multi-state models for such data are generally based on the Markov assumption" [5].

However, the data we have in this proposed research is incomplete. In application, the dataset will include lots of noise. This may be due to typing errors or other kinds of errors made during the registration of inspection results. Moreover, the problem we face is that the datasets with monitoring data cover so short time period that would hardly contain inspections representing the whole lifetime of assets. For example, the construction year (2000) is close to

the current date, when the periodic inspections are carried out every 5 years. Then only 2 observations are available. This means that only a few data are available. Furthermore, the assessments of condition are also subject to personal judgments. The visual inspection is carried out periodically by different technicians and it would result in subjective judgments of the component condition.

Another problem related to the dataset is censoring. We have to be satisfied with incomplete dataset. It is impractical or too expensive to wait until the process moves to the failure state, and the inspection data of individual asset was not registered for some reasons (missing entries or type errors). That is why it is hard to get all information through the lifetime of assets, and then censoring occurs. For example, some of the assets do not reach the final state and the correct time of failure is not known at the inspection, because the companies will replace or repair their asset before failures and the failure will happen between two inspections. For most of the assets, the degradation process is censored, because the inspections are not carried out continuously and some of the assets will not reach the end of the lifetime. When a transition happened between two inspections, the inspector only knows there must be a transition between two inspections, but he does not know the right time when the transitions occurred. Thus, the interval-censoring will happen between two observations.

In summary, the dataset we have is incomplete and heavily censored. What we know is how many failures between two consecutive inspections and the time of each inspection. The Markov process would be our candidate to describe the deterioration process of assets because of the multiple states and continuous time.

Chapter 3: Model description

3.1 General introduction of Markov processes

In the literature, Markov processes have been applied quite frequently in the engineering [6]-[7]. It is suitable to model the random progress through discrete states. A Markov process is a mathematical system that undergoes transitions from one state to another, among a finite numbers of possible states. With the Markov property, future states are independent of past states given the present state. In this section, the theory of Markov chain is based on Ross's theory [8],

3.1.1 Markov Chain

A stochastic process $\{X_n, n = 0,1,2,3, \cdots\}$ with a finite number of possible values will be discussed. These possible values in the stochastic process are denoted by the corresponding non-negative values $\{0,1,2,3, \cdots\}$. If $X_{n+1} = j$, it indicates the stochastic process is said to stay in state j given the time is $n + 1$. That is

$$\Pr(X_{n+1} = j | X_n = i, X_{n-1} = i_n, \dots , X_2 = i_2, X_1 = i_1)$$
$$= \Pr(X_{n+1} = j | X_n = i) = P_{ij} \tag{1}$$

for all states $j, i, i_n, \dots, i_0$ and all $n \geq 0$. This kind of stochastic process is called as a Markov chain [8]. "Equation (1) may be interpreted as stating that, for a Markov chain, the conditional distribution of any future state X_{n+1}, given the past states $X_{n-1}, \cdots, X_0$ and the present state X_n, is independent of the past states and depends only on the present state." [8]. The value P_{ij} represents the probability that the stochastic process will make a transition to the next state j given that the process is in state i. As is noted in Ross's theory [8], since all of the probabilities are larger or equal to 0 and also the stochastic process must make a transition to the other states, so what we have is that

$$P_{ij} \geq 0 \text{ and } \sum_{j=0}^{\infty} P_{ij} = 1, i = 0,1,2, \cdots \tag{2}$$

The $n + m$-step transition probabilities are denoted by $\Pr(X_{n+m} = j | X_0 = i) = P_{ij}^{n+m}$ and it may be computed by the Chapman-Kolmogorov equation which is show below:

$$P_{ij}^{n+m} = \sum_{k=0}^{\infty} P_{ik}^{n} P_{kj}^{m} \tag{3}$$

for all $n, m \geq 0$. "It is easily understood by noting that $P_{ik}^{n} P_{kj}^{m}$ represents the probability that starting in i the process will go to state j in $n + m$ transitions through a path which takes it into state k at the nth transition" [8].

3.1.2 Markov process

Generally speaking, the Markov process is the extension of the Markov chain. Let's denote the state space as $E = \{0,1,2,3,...\}$ and let $(X(t)|t \geq 0)$ with index t denote a Markov process. According to the definition of Markov chain, the process $(X(t)|t \geq 0)$ is said to be a continuous-time Markov chain, if all states $i_n, i, j \in E$ and $0 \leq h_1 < h_2 < \cdots < h_n < s < s + t$

$$\begin{aligned}
&P_{ij}(s, s + t) \\
&= \Pr(X(s + t) = j | X(s) = i, X(h_n) = i_n, X(h_{n-1}) = i_{n-1}, X(h_{n-2}) = i_{n-2} \dots X(h_1) = i_1) \; (4) \\
&= \Pr(X(s + t) = j | X(s) = i)
\end{aligned}$$

The probability $P_{ij}(s, s + t)$ of the process, which is called the transition probability of Markov process, represents that the process will make a transition into state j at time $s + t$, provided that the process is in state i at time s. Such a Markov process has the Markovian property that the conditional distribution of future state is not depending the past state $X(h_k) = i_k$, but only depending on present state $X(s) = i$ [8].

If the equation (4) can be interpreted as the equation below

$$P_{ij}(t) = \Pr(X(s + t) = j | X(s) = i) \tag{5}$$

we call this Markov process as continuous-time homogeneous Markov process and $P_{ij}(t)$ denotes the transition probability of Markov process in following sections [9]. Since the transition probabilities do not depend on the actual time s but only depends on the length of time interval t.

If the equation (4) can't be interpreted as the equation (5), it indicates the transition probability depends on the starting time t and actual time interval $(s, s + t)$. According the definition of stationary or homogeneous transition probabilities [8], such transition probability of the Markov process is always changing with time, which is called as continuous-time

inhomogeneous Markov process in following sections.

The Chapman-Kolmogorov equation of the continuous Markov process is the solution of transition probabilities within one time period. The equation can be expressed as:

$$P_{ij}(s+t) = \Pr(X(s+t) = j|X(0) = i) = \sum_{k \in E} P_{ik}(s)P_{kj}(t) \quad for\ i,j,k \in E \qquad (6)$$

This equation represents that the process will go to state j within time period $s+t$ starting from i with a path which would lead the process into state k at time point s [8].

Let the transition probability function equation given by equation (4) denote the elements of the corresponding matrix $\boldsymbol{P}(s, s+t) = \|P_{ij}(s, s+t)\|$ of Markov process. Let $\boldsymbol{Q} = \|Q(s+t)\|$ denotes the intensity matrix of Markov process. The transition probability matrix and intensity matrix are the solution of differential equation [10]:

$$\frac{d}{d(s+t)}\boldsymbol{P}(s, s+t) = \boldsymbol{P}(s, s+t)\boldsymbol{Q}(s+t) \qquad (7)$$

3.1.3 Semi-Markov process

A more general Markov process is semi-Markov process. "Suppose that a process can be in any one of N states $1, 2, \ldots, N$, and that each time it enters state i it remains there for a random amount of time having mean μ_i and then makes a transition into state j with probability P_{ij}. Such a process is called a semi-Markov process. Note that if the amount of time that the process spends in each state before making a transition is identically 1, then the semi-Markov process is just a Markov chain." [8] .

For semi-Markov process $(X_t|t \geq 0)$, it is an extension of the Markov chain in which a random time is added between each transition, and then the transition probability of semi-Markov process transiting to state j within a time interval less than or equal to t, provided starting from state i, is expressed as

$$Q_{ij}(t) = Pr(T_{n+1} - T_n \leq t, X(T_{n+1}) = j|X(T_n) = i) \qquad (8)$$

where $T_0 < T_1 < \cdots < T_n < T_{n+1}$ and they are the times when the process changes its state [1]. The random time between every transition can be interpreted in terms of the distribution function:

$$F_{ij}(t) = \Pr(T_{n+1} - T_n \leq t|X(T_{n+1}) = j, X(T_n) = i) \qquad (9)$$

where $T_{n+1} - T_n$ is the waiting time before the process moves to state j given the process has already moved to state i and n is the number of observations [1].

So the probability of the process transiting into the state j with a fixed time, which is less than or equal to t, given that the process already moved into the state i, is shown below [1]:

$$Q_{ij}(t) = Pr(T_{n+1} - T_n \leq t, X(T_n) = j | X(T_{n-1}) = P_{ij}F_{ij}(t) \tag{10}$$

Equation (10) indicates that the transition of semi-Markov model has two steps, see [1]:

1. When the process moves into state i, the semi-Markov process will choose the next state j depending on the transition probability P_{ij}.

2. Before moving into the next state j, the semi-Markov process will wait for an random time which belongs to the distribution $F_{ij}(t)$

In fact, deterioration procedure can be interpreted as a Markov process, since the property of a Markov process, which is a stochastic process, is that the values of $X(w)$ with $w > t$ are not dependent on the values of $X(u), u < t$, given the value of $X(t)$. That is why the Markov processes are useful for modeling the stochastic deterioration with independent increments. In order to model the deterioration, two kinds of Markov processes would be useful. They are shown in Figure 1 and Figure 2.
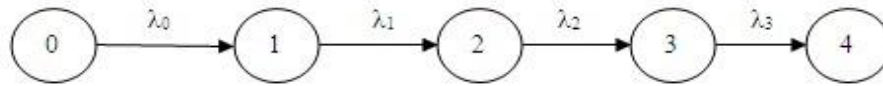

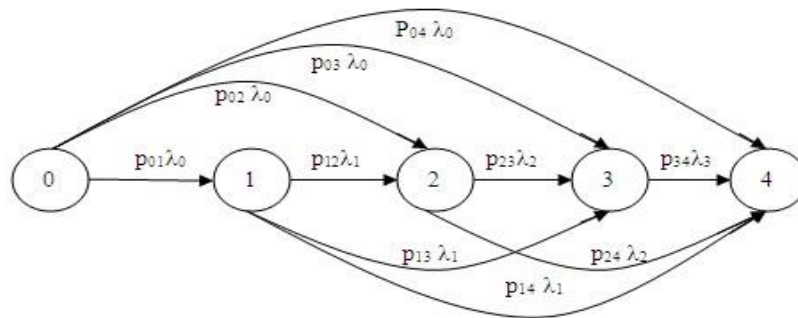Figure 1: Markov process with sequential structure

Figure 2: Markov process with progressive structure

For both kinds of Markov process shown in Figure 1 and Figure 2, the total lifetime of object is the sum of sojourn times in each state. Thus, in the following sections, we focus on the sequential structure of Markov process. Therefore, Markov process should satisfy at least two characteristics, see [1].

1. Different states have different technical conditions, which should be ordered depending on the correct level of the deterioration.
2. The stochastic process must go through all of the condition states, see Figure 1.

3.2 Nonhomogeneous Poisson Process

Esary and Marshall found out lifetime distribution of the instrument and a sequence of shocks were considered as the events occurring randomly in time which is governed by a homogeneous Poisson process (HPP) [11]. They calculated the probability of damage which results from the shock accumulates until it excesses the threshold results in failure. Abdel-Hameed and Proschan extended the shock model to a new shock model governed by nonhomogeneous Poisson process (NHPP) rather than HPP [12]. Compared with homogeneous Poisson process, NHPP is more flexible and its rate of occurrences of failures change**s** with time. Here, the theory of NHPP is based on the Rausand's book [13]. The assumption under NHPP is that the system is under the minimal maintenance which means that the system continues as if no events had happened after a minimal maintenance. In this section, the application of NHPP modulated by a continuous Markov process will be discussed.

The purpose of this section is to review some of the most important mathematical properties of the homogeneous and non-homogeneous Poisson process. It includes the definition of the HPP and NHPP, as well as methods for parameter estimation and simulation.

First, let's have a look at the definition of HPP and NHPP. For homogeneous Poisson process, the conditions should be assumed as follows:

1. $N(0) = 0$, where $N(t)$ denotes the number of failures at time t
2. The probability of the event occurring in the interval $(t, t + \Delta t]$ is independent of current time and may be written as $\lambda \cdot \Delta t + o(\Delta t)$, where λ is a positive constant
3. $P(X(t + \Delta t) - X(t) \geq 2) = o(\Delta t)$ means that more than one event will not happen with a very small time interval. So value of $o(\Delta t)$ is close to zero

15

4. The mean number of events happened in the time interval $[s, s + t]$ is

$$W(t + \Delta t) - W(t) = E\big(N(t + \Delta t) - N(t)\big) = \lambda \Delta t \qquad (11)$$

The stochastic process $\{N(t), t \geq 0\}$ is an HPP with intensities λ which is constant [13]. According to the definition of HPP, it is not used to model the process where the failure rate is not constant and changing with time such as continuous-time inhomogeneous Markov process. Thus, it can be applied to model time homogeneous Markov process.

For NHPP with the rate $\lambda(t)$ for $t \geq 0$, the main assumptions of NHPP are shown below,

1. $X(0) = 0$
2. $\{X(t), t \geq 0\}$ has independent increments
3. $P(X(t + \Delta t) - X(t) \geq 2) = o(\Delta t)$ means that more than one failures will not happen at the same time. So value of $o(\Delta t)$ is close to zero
4. $P(X(t + \Delta t) - X(t) = 1) = \lambda(t)\Delta t + o(\Delta t)$

This kind of counting process $X(t), t \geq 0$ is called NHPP [13]. For the characteristics of Poisson process, since the HPP is special case of non-homogenous Poisson process, let's start to discuss the properties of NHPP. According to the definition of NHPP, the parameter is rate function $\lambda(t)$. This function is also called Rate of Occurrence of Failures (ROCOF) function of NHPP [13]. The cumulative intensity of the process is

$$\Lambda(t) = \int_0^t \lambda(u)du \qquad (12)$$

The number of failures within interval $(0, t)$ is Poisson distributed according to the reliability theory [13] and the distribution of $X(t) = n$ and mean number of failures are shown below

$$P(X(t) = n) = \frac{[\Lambda(t)]^n}{n!} e^{-\Lambda(t)} \quad \text{for } n = 0, 1, 2, \dots \qquad (13)$$

$$E\big(X(t)\big) = \int_0^t \lambda(u)du = \Lambda(t) \qquad (14)$$

According to the equation (14) and (12), the cumulative function $\Lambda(t)$ is the mean number of failures and its variance within interval $(0, t]$ for NHPP. According to the equation (13), it follows

$$P(X(t + v) - X(t) = n) = \frac{[\Lambda(t+v)-\Lambda(t)]^n}{n!} e^{-[\Lambda(t+v)-\Lambda(t)]} \qquad (15)$$

for $n = 0, 1, 2 \dots$. The mean number of failures in the interval $(t, t + v]$ is given by Rausand and Høyland, and it is shown below

$$E(X(t + v) - X(t) = n) = \Lambda(t + v) - \Lambda(t) = \int_t^{t+v} \lambda(u)du \qquad (16)$$

Let T_n denote the time until nth failure for $n = 1, 2, \ldots$, where $T_0 = 0$. The distribution of T_n is given by [13]:

$$P(T_n > t) = P(X(t) \le n - 1) = \sum_{k=0}^{n-1} \frac{\Lambda(t)^k}{k!} e^{-\Lambda(t)} \tag{17}$$

Let T_1 denote the time of first failure from $t_0 = 0$ until the first failure. Then, the survival function of T_1 is shown below

$$R(T_1) = e^{-\Lambda(t)} = e^{-\int_0^{T_1} \lambda(u)du} \tag{18}$$

Assuming that there is no failure happen within time interval $(T_0, T_0 + t)$ given that process is observed at time T_0 and t denote the time until the next failure. The equation (18) can be expressed as

$$P(X(T_0 + t) - X(T_0) = 0) = e^{-[\Lambda(T_0+t)-\Lambda(T_0)]} = e^{-\int_{T_0}^{T_0+t} \lambda(u)du} \tag{19}$$

In following sections, the power law model will be used to describe the ROCOF of a NHPP, it is defined as below [13]:

$$\lambda(t) = abt^{b-1} \quad \text{for } a > 0, b > 0 \text{ and } t \ge 0 \tag{20}$$

The failure rate function of Weibull distribution is:

$$p(t) = \frac{\beta}{\eta^\beta} t^{\beta-1} \tag{21}$$

From equation (20), we know the mean number of failures within interval $(0, t]$. If $\alpha = \frac{1}{\eta^\beta}$ and $b = \beta$, the failure rate function $p(t)$ of Weibull distribution can be defined as $\lambda(t) = abt^{b-1}$.

Because of the polynomial nature of the ROCOF, this model is very flexible and it can model both increasing $(b > 1)$ and decreasing $(0 < b < 1)$ failure rates. When $b = 1$, the model reduces to the HPP model with constant intensity.

Probabilities of a given number of failures for the NHPP model are calculated by an extension of the formulas for the HPP. Thus, for any NHPP

$$P(X(t) - X(0) = k) = \frac{\Lambda(t)^k}{k!} e^{-\Lambda(t)} \tag{22}$$

With the Power Law model:

$$P(X(t) - X(0) = k) = \frac{\alpha^k t^{bk}}{k!} e^{-\alpha t^b} \tag{23}$$

In practice, it has four types of the Markov process with different kinds of intensities according to the properties of equation (20), when the NHPP is applied to simulate the continuous Markov

process. The four types of Markov processes are called Model A, Model B, Model C, and Model D.

3.3 Model description

3.3.1 Model A and Model B

One special case in NHPP is that $\lambda(t)$ is constant all the time, when $b_i = 1 \; and \; a_i = a = \lambda(t)$. Thus the cumulative intensity of the Markov process is only depending on the length of time, so this kind of Markov process is called Model A in following sections. The Markov process is called Model B in following sections, when $b_i = b = 1$ and $\lambda_i(t) = a_i$, which is state-dependent and time-independent. Model A and Model B belong to the time homogeneous Markov process, since they are all time-independent and the sojourn times in all state are random variables with exponential distribution with constant intensities. Their intensity matrices are shown in Table 3.

Table 3: Sample of structure of Time Homogeneous Markov Process

Model A						Model B				
$-\lambda$	λ	0	0	0		$-\lambda_1$	λ_1	0	0	0
0	$-\lambda$	λ	0	0	Time Homogeneous	0	$-\lambda_2$	λ_2	0	0
0	0	$-\lambda$	λ	0	Markov Process	0	0	$-\lambda_3$	λ_3	0
0	0	0	$-\lambda$	λ		0	0	0	$-\lambda_4$	λ_4
0	0	0	0	0		0	0	0	0	0

In practice, the time homogeneous Markov process with sequential structure will not always stay on the original state and it will move to the different state. The further waiting time is not depending on the times spent in the other states because of the properties of exponential distribution (memoryless). The structure of Markov process is shown in Figure 1.

Let the state space as $\boldsymbol{E} = \{0,1,2,3,4\}$. The CDF and probability density function (PDF) for Model A and B, which are illustrated by M. Rausand and A. Høyland [13], are given below respectively:

$$F_i(t) = 1 - e^{-\lambda_i t} \tag{24}$$

$$f_i(t) = \lambda_i e^{-\lambda_i t} \tag{25}$$

where the intensity $\lambda_i > 0$, and for all $i \in \boldsymbol{E}$.

For Model A, its mean number of events and variance of every state within time interval $(0, t]$ is the same.

$$E(X(t)) = \int_0^t \lambda(u)du = \lambda t \tag{26}$$

According to the equation (13), the number of events happens within time interval $(0, t]$ is:

$$P(X(t) = N) = \frac{[\lambda t]^N}{N!}e^{-\lambda t} \quad \text{for } N = 0, 1, 2, \ldots\ldots \tag{27}$$

Based on equation (16), the mean number of failures in the interval $(t, t + v]$ is

$$E(X(t + v) - X(t)) = \Lambda(t + v) - \Lambda(t) = \lambda v \tag{28}$$

which indicates the mean number of events in the time interval $(t, t + v]$ is just dependent on the length of time for Model A. Thus the CDF given that no event happens in the time interval $(t, t + v]$, is

$$P(X(t + v) - X(t) = 0) = e^{-[\Lambda(t+v)-\Lambda(t)]} = e^{-\lambda v} \tag{29}$$

Assuming T_n denote the time until nth failure for $N \in \mathbf{E}$, where $T_0 = 0$. According to the equation (17), The distribution of T_n for Model A is given by [13]:

$$P(T_n > t) = P(X(t) \le n - 1) = \sum_{k=0}^{N-1}\frac{\lambda t^k}{k!}e^{-\lambda t} \tag{30}$$

For Model A, let $\boldsymbol{\theta} = \{\lambda_0, \lambda_1, \cdots\}$ denote the parameter set. If those parameters are known, the remaining lifetime of reaching failure state is belonging to Erlang distribution which is shown below:

$$f(t; \lambda_i, N) = \lambda_i^N t^{N-1}e^{-\lambda_i t}/(N - 1)! \tag{31}$$

where N is shape parameter and $\lambda_i = \lambda$ with $\lambda > 0$ [8].

For Model B, the remain lifetime of reaching failure state is belonging to Hypoexponential distribution which is

$$f(t; \lambda_i) = \sum_{i \in \mathbf{E}} \lambda_i e^{-\lambda_i t}(\prod_{i,j \in \mathbf{E}, i \ne j} \lambda_i/\lambda_j - \lambda_i) \tag{32}$$

with $\lambda_i \ne \lambda_j$ and $i, j \in \boldsymbol{E}$ [1].

Because of the property of exponential distribution, the lifetime of Model A and Model B is the sum of the sojourn time of each state.

$$E(T_{\text{lifetime}}) = E(S_0) + E(S_1) + \cdots \tag{33}$$

with $E(S_i) = \frac{1}{\lambda_i}$ and $E(S_i)$ denote the sojourn time of each state. And we can also compute the

mean time to failure from state 1, 2 and 3 to state 4 in similar way.

According to the memoryless of exponential distribution, the remaining lifetime given the Markov process stays at state i is:

$$E\left(T_{\text{remaining lifetime}}\right) = \sum_i^N E(S_i) \tag{34}$$

where $N, i \in E$ and $T_{\text{remaining lifetime}}$ denotes the remaining lifetime of Markov process.

3.3.2 Model C and Model D

Compared with characteristics of intensity of Model A and B, if $b_i \neq 1$, $b_i > 0$ and $a_i > 0$, there are two cases for time inhomogeneous Markov process. According the value of b_i and a_i, the first case is that the first time inhomogeneous Markov process is called as Model C with $\lambda(t) = abt^{b-1}$ in following sections, when $b_i = b \neq 1$ and $a_i = a > 0$. The second case is that the time-inhomogeneous Markov process is called Model D with intensity $\lambda_i(t) = a_i b_i t^{b_i-1}$ in the following sections, when the values of b_i and a_i are not constant. Their intensity matrices are shown in Table 4.

For time inhomogeneous Markov process, the intensities of two time inhomogeneous Markov processes are not constant and they change with time. The Hypoexponential distribution and Erlang distribution are suitable for finding the joint CDF of time inhomogeneous Markov processes.

For Model C and D, Model C is a standard NHPP, and the equations from equation (12) to equation (23) are suitable for Model C and Model D.

Table 4: Sample of structure of time inhomogeneous Markov Process

Model C						Model D				
$-\lambda(t)$	$\lambda(t)$	0	0	0	Time	$-\lambda_1(t)$	$\lambda_1(t)$	0	0	0
0	$-\lambda(t)$	$\lambda(t)$	0	0	Inhomogeneous	0	$-\lambda_2(t)$	$\lambda_2(t)$	0	0
0	0	$-\lambda(t)$	$\lambda(t)$	0	Markov Process	0	0	$-\lambda_3(t)$	$\lambda_3(t)$	0
0	0	0	$-\lambda(t)$	$\lambda(t)$		0	0	0	$-\lambda_4(t)$	$\lambda_4(t)$
0	0	0	0	0		0	0	0	0	0

Let P denote the probability transition matrix of the Markov process. By computing the transition probability matrix P of two time inhomogeneous Markov process depending on the equation (7), the cumulative density function of time from state i to final state j is shown below [10]:

$$P_j(t) = \sum_{i=0}^{4} P(X(0) = i) \, P_{ij}(0, t) \tag{35}$$

The CDF of time from state 0, 1, 2, 3 to final state 4, namely absorbing state, can be computed based on equation (35), the uncertainty can be expressed as the lifetime distribution, which is shown below:

$$F(t) = \sum_{i=0}^{4} P(X(0) = i) \, P_{ij}(0, t) \tag{36}$$

The mean lifetime $T_{\text{Mean lifetime}}$ of the object can be computed as follow:

$$T_{\text{Mean lifetime}} = E(T) = \int_{0}^{\infty} 1 - F(t) dt \tag{37}$$

In summary, if the residence time of each state belongs to the exponential distribution with constant intensities, it is called time homogeneous Markov process. If the intensities of Markov process are changing with time, and then it is called time inhomogeneous Markov process in this proposed research.

For model A, its intensities do not depend on the state and time. The intensities of model B do not depend on time but depend on time. The intensity of model C is time-dependent and state-independent. And finally, the intensity of model D is time-dependent and state-dependent.

3.4 Semi-Markov process with sojourn time belonging to Weibull distribution

In order to take the problem of censoring data into consideration, survival analysis will be applied to data analysis. Assuming that the semi-Markov process with Weibull distributed sojourn time, which is called as the Model E in the following sections, has five states given no maintenance and replacement carried out, the technical condition of the asset will run all of the states and reach at the fifth state finally. The semi-Markov process can be described as the deterioration process curve as shown in

Figure 1. For the ongoing research, let the residence time S_0 S_1, S_2 and S_3 to be Weibull distributed and the lifetime of the object is also sum of sojourn time of each state.

For analysis, it was assumed that the residence time t in state i is Weibull distributed with scale parameter $\beta_i > 0$ and shape parameter $\eta_i > 0$. The probability density function of Weibull distribution is

$$f(t; \beta_i, \eta_i) = \begin{cases} \frac{\eta_i}{\beta_i} (\frac{t}{\beta_i})^{\eta_i - 1} \exp(-(\frac{t}{\beta_i})^{\eta_i}), & t \geq 0) \\ 0, & t < 0 \end{cases} \tag{38}$$

Assume that its parameter set is $\boldsymbol{\theta} = (\beta_i, \eta_i)$, and its CDF and expectation time in each state are

$$F(t; \beta_i, \eta_i) = 1 - \exp(-(\frac{t}{\beta_i})^{\eta_i}) \tag{39}$$

$$E(S_i) = \beta_i \Gamma(1 + 1/\eta_i) \tag{40}$$

$$\lambda(t) = \frac{\eta_i}{\beta_i} (\frac{t}{\beta_i})^{\eta_i - 1} \tag{41}$$

where Equation $F(t; \beta_i, \eta_i)$ is the CDF, $E(S_i)$ is the mean sojourn time in state i and $\lambda(t)$ is the hazard function, which are all given in the book of Rausand and Høyland [13], In order to estimate the remaining lifetime and the transition probabilities, the research will focus on the estimating parameters.

Chapter 4: Model application

In this section, the models described in chapter 3 will be applied to the dataset gathered by visual inspections. For the infrastructures in the industrial fields, the remaining lifetime and the duration of stay in each state are of interest in this chapter.

In practice, the most important aspect of inspection is that the technical condition should be registered pole by pole and the technical condition will be used as the basics to find the state of the wood pole.

4.1 Description of information

It is difficult to extract the data used for likelihood estimation from the dataset and use it to do parameter estimation. There are many reasons why some information in the dataset may mislead us or it may not be included.

1. In practice, the constructing year of the asset may not have been registered in the dataset.

2. The incorrect date of inspection may have been assigned to the dataset due to personal factors (typing errors and misclassifications)

Besides the identifiable faulty data in the dataset, the missing information will happen and personal errors are present in the dataset. That is why it is difficult for us to extract the right information from dataset. For the original data in this project, the data include 479 poles, the original data description is:

1. 410 poles are still in state 0;

2. 56 poles moved into state 1;

3. 9 poles moved into state 2;

4. 3 poles moved into state 3;

5. Only one pole failed;

In order to get more information of the dataset, one of the best ways of summarizing multi-state data is to build a frequency table of all states. For example, the number of observations and the number of transitions from state i to state j should be summarized. Such table is shown in Table 5. It calculates a frequency table counting the number of observations. For example, the

first row shows that the states of poles start from the state 0 when the inspections were carried out. The number of the first row indicates that the poles were found in state 0 for 860 times, in state 1 for 59 times , in state 2 for 7 times, in state 3 for 2 times and in state 4 for only one time respectively. The second row indicates that the state of pole start from state 1 at the first inspection and the numbers of second row show that the poles were found in state 1 for 19 times, in state 2 for 2 times and in state 3 for only one time respectively. The same holds the other rows. According to the Table 5, most of the observations indicate that poles were found in state 0, and only a small proportion of the observations show that the poles were found in state 1. For state 2, 3 and 4, only a few observations show that the poles were found in state 2, 3 and 4.

Table 5: Table of observations

From\to	0	1	2	3	4
0	860	59	7	2	1
1	0	19	2	1	0
2	0	0	1	0	0
3	0	0	0	1	0
4	0	0	0	0	1

Algorithms will be used for extracting data. The original data is not convenient for parameter estimation in R computer software. According to the requirement of R computer software, the data extracted from the database will be stored in text file as below.

Table 6: sample of data in the text file

1	0	0	6	0	NA	NA
⋮	⋮	⋮	⋮	⋮	⋮	⋮
100	0	0	9	1	16	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮
200	0	0	44	2	51	2
⋮	⋮	⋮	⋮	⋮	⋮	⋮
300	0	0	45	3	52	3
⋮	⋮	⋮	⋮	⋮	⋮	⋮
479	0	0	36	4	43	4

The first column includes the identifier of different object, the second column and third column show the installation date and the beginning state respectively. The fourth column is the date of the first inspection, and the fifth column is the state of first inspection, the sixth column is

date of second inspection and the last column is the state of the second inspection. The dataset in the Table 6 is suitable for the Model A, B, C and D.

For Model E, the original dataset should also be transferred for every state. The transferring process is shown in the following. We take the No.3 of pole in Table 2 as one example and it is shown as one timeline below:

Figure 3: Timeline of pole of No. 3

Assuming that the first time is the installation time (2000) of pole, the timeline above indicates the range of sojourn time in state 0, 1, 2 and 3. They are shown below:

$$0 < S_0 < 3, \qquad 0 < S_1 < 3, \qquad 0 < S_2 < 10, \qquad 0 < S_3 \tag{42}$$

The right form in R should be like the form for first state below:

Table 7: Sample of dataset for Model E

L	R
6	NA
6	NA
NA	12
6	13
29	NA
29	NA
⋮	⋮

L indicates the left side of the time interval and R indicates the right side of time interval for each state. For instance, the first row represents the time interval is $S_i > 6$ (right censoring), and third row indicates the time interval is $S_i < 12$ (left censoring) and the fourth row represents the time interval is $6 < S_i < 13$ (interval censoring). The rest form of data for the other state could be deduced for the example above.

4.2 Parameter estimation

4.2.1 Maximum likelihood method for Model A, B, C and D

For this method, the process in Figure 1 will be discussed. Let $\boldsymbol{\theta}$ denote the vector with all parameters, $T_0 < T_1 < T_2, \dots, < T_n$ with $T_0 = 0$ and $x_n \epsilon E$ with $x_{n-1} \leq x_n$, so the probability function given in Kallen's paper [10] of an object is:

$$f(\boldsymbol{x}; \boldsymbol{\theta}) = \Pr(X(T_n) = x_n | X(T_{n-1}) = x_{n-1}, X(T_{n-2}) = x_{n-2}, \dots, X(T_1) = x_1, X(T_0) = 0)) \quad (43)$$

with $x_n \in E$, the equation above can be written as:

$$f(\boldsymbol{x}; \boldsymbol{\theta}) =$$
$$\Pr(X(T_n) = x_n | X(T_{n-1}) = x_{n-1}) \dots \Pr(X(T_1) = x_1 | X(T_0) = 0) \quad (44)$$
$$f(\boldsymbol{x}; \boldsymbol{\theta}) = \prod_{i=0}^{n} \Pr(X(T_i) = x_i | (T_{i-1}) = x_{i-1}) \quad (45)$$

As illustrated by Kallen, if there are m objects in the dataset, so the likelihood function is the product of the probability density function of each asset:

$$L(\boldsymbol{\theta}; x) = \prod_{j=1}^{m} \prod_{i=0}^{n_j} \Pr_j(X(T_i) = x_i | (T_{i-1}) = x_{i-1}) \quad (46)$$

with $j = (1,2,3, \dots, m)$ denoting index of objects and n_j denoting the number of inspections of the jth wood pole. So the likelihood function for non-homogeneous Poisson process is:

$$L(\boldsymbol{\theta}; x) = \prod_{j=1}^{m} \prod_{i=0}^{n_j} \Pr_j(X(T_i) = x_i | (T_{i-1}) = x_{i-1})$$

$$= \prod_{j=1}^{m} \prod_{i=0}^{n_j} \frac{[\Lambda(T_{i,j}) - \Lambda(T_{i-1,j})]^{(x_{i,j} - x_{i-1,j})}}{(x_{i,j} - x_{i-1,j})!} e^{-[\Lambda(T_{i,j}) - \Lambda(T_{i-1,j})]} \quad (47)$$

with $\Lambda(T_{i,j}) = \int_0^{T_{i,j}} abt^{b-1} dt$. It is the solution of time inhomogeneous Markov process. The procedure of the maximum likelihood method in this paper can be shown in Table 8.

In this project, it is not easy to do estimation by applying the formula of non-homogeneous Poisson process into the maximum likelihood function. Thus the transition probability matrix is the solution of equation (46) and the transition probability matrices are shown below [14]:

$$P(s, s + t) = \exp(tQ) = \sum_{k=0}^{\infty} \frac{t^k}{k!} Q^k \quad (48)$$

where t and s denote the two time points and $t, s \geq 0$. This is used to estimate the parameters of time homogeneous Markov process with sojourn time belonging to the exponential distribution such as Model A and Model B.

Table 8: The Procedure of maximum likelihood method

$T_1 = 0$	$X(T_1) = 1$	$P(1)$	$f(x; \boldsymbol{\theta}) = \displaystyle\prod_{i=0}^{n_j} \Pr(i)$
$T_2 = 0$	$X(T_2) = 1$	$P(2)$	
$T_3 = 0$	$X(T_3) = 2$	$P(3)$	
$T_4 = 0$	$X(T_4) = 3$	$P(4)$	$L(\boldsymbol{\theta}; x) = \displaystyle\prod_{j=1}^{m}\prod_{i=0}^{n_j} f_j(x; \boldsymbol{\theta})$
$T_5 = 0$	$X(T_5) = 3$	$P(5)$	
$T_6 = 0$	$X(T_6) = 4$		
…	…	…	where j denotes the number of poles and n_j denotes the number of inspections of No. j pole.

For time inhomogeneous Markov process, the transition probability matrix is computed by using the equation (7) and Euler method, and it can be written as:

$$\boldsymbol{P}(s, t + \Delta t) = \boldsymbol{P}(s, t)(\boldsymbol{Q}(t) + \boldsymbol{I}) \tag{49}$$

when $s = t$, $P(s, t) = \boldsymbol{I}$ identity matrix [1]. So the equation (49), which is suitable for parameter estimation of Model C and Model D, is also one of the solutions of equation (46).

Now the main challenge is maximizing the likelihood equation and maximizing methods, like Quasi-Newton methods, BFGS (Broyden–Fletcher–Goldfarb–Shanno) method, Fisher's method of scoring and Newton's method, will be applied into the maximization.

4.2.2 Maximum likelihood method for Model E

For Model E, the maximum likelihood method is also the most suitable parameter estimation method. Depending on the characteristics of data from chapter 2, the data set contains three kinds of censored information.

The likelihood function of right-censoring, left-censoring and interval-censoring data is shown as below [15].

$$L(\boldsymbol{\theta}; L, I, R) = \prod_{L} F(T_L; \boldsymbol{\theta}) \prod_{I} [F(T_{I,\max}; \boldsymbol{\theta}) - F(T_{I,\min}; \boldsymbol{\theta})] \prod_{R} (1 - F(T_R; \boldsymbol{\theta})) \tag{50}$$

with $F(T; \boldsymbol{\theta})$ is CDF of the residence time distribution, $\boldsymbol{\theta}$ is a vector with all parameters. L denotes the left-censoring data, I denotes the interval-censoring data and R denotes the right-

censoring data. $T_{I,\max}$ and $T_{I,\min}$ are the upper and lower limits of interval-censoring information for each state.

In the literature, this suggested method will fail when most of the censored information overlaps with each other [16]. If a few inspections are available, this will happen frequently. By looking at Figure 3 and the equation (42), the ranges of sojourn time of state 0 and state 1 are the same. If most of the ranges of sojourn are the same with each other, the estimate of parameter will be not applicable. So this method is suitable for parameter estimation with the enough information. This Maximum Likelihood Method in this project is used to estimate the parameters of semi-Markov process with sojourn time belonging to the Weibull distribution. By applying this method, the statistics software "R" and the package "fitdistrplus" in CRAN (Comprehensive R Archive Network) are necessary for estimating the parameters. The results of semi-Markov process with sojourn time belonging to the Weibull distribution are shown in Table 13. Meanwhile, this maximum likelihood method is also suitable for Model A and Model B, since Model A and B are special case of the semi-Markov process.

4.3 Results of parameter estimation

4.3.1 Results of Model A and Model B

By applying the maximum likelihood method explained in Table 8, the values of the parameter λ_i in time homogeneous Markov process with sojourn time belonging to the exponential distribution are shown in Table 9. The intensities of Model A are all the same and it equals to 0.0061, since the Markov process is time independent and state-independent. By comparing the values of parameter and standard error, the value standard error is larger than the parameter which shows that the estimate of parameter is not good. Because the dataset is not large enough, that is why this condition happened in Model A. Using $E(S_i) = \frac{1}{\lambda_i}$, the total lifetime of the model A is 655.74, which is given below:

$$E(T_{\text{lifetime}}) = E(S_0) + E(S_1) + E(S_2) + E(S_3) = 655.74 \qquad (51)$$

Table 9: Results of Model A and B

State	Model A		Model B	
i	λ_i	Standard error	λ_i	Standard error
0	0.0061	0.04106959	0.00508	0.00061683
1	0.0061	0.04106959	0.02077	0.00630612
2	0.0061	0.04106959	0.07573	0.05144133
3	0.0061	0.04106959	0.03440	0.06703010

By looking at the result of model B, the value of parameter for the state 0 to state 2 is increasing from 0.00508 to 0.07351, and the standard errors are all smaller than the values of parameters. However, the parameter of the state 3 is 0.0324, and its standard error is larger than the value of parameters, which indicates the estimate of the parameter of state 4 is not applicable. The intensities are fixed and they do not change with time, but they change with states, since time homogeneous Markov process (Model B) is time-independent and state-dependent. The trend of the values of the intensities can be explained by the fact that the process will spend more time on the good states and the sojourn time for each state will decrease as the condition of poles goes from good to bad.

Using $E(S_i) = \frac{1}{\lambda_i}$ and result of model B and in Table 9, which is computed on R computer software by loading the package "msm", the sojourn time for each state is shown in Table 10. The total lifetime (290) of pole is sum of sojourn time in every state and the sojourn times are decreasing with the increasing states except the state 3. For state 3, the range of sojourn between the low and upper bound is so relatively large compared with real value of estimate of sojourn time (29.1) for state 3, which means the estimate of parameter of state 3 may be misfit.

With analysis in depth, checking the PDF would be necessary. By using the equation (31), the plot of the probability density function of Model A below, see Figure 4, indicates that the failure probability increases so slow and arrives at mode around 500 years, and then it decreases slowly after 500 years. The straight line indicates the mean lifetime of poles that is mean lifetime of pole by using result of equation (51). For Figure 5, the median time, when survival function is equal to 0.5, is approximately equal to the mean sojourn time for wood poles. The survival function of time from initial state (state 0, 1, 2 and 3) to state 4 shows that there is almost difficult to see any difference between the cross-over points, and it indicates the mean sojourn time for every state is equal to each other.

Table 10: The sojourn time of Model B

	Estimates	SE	Low	Upper
State 0	197.0	23.73	155.60	249.5
State 1	48.2	13.97	27.42	85.2
State 2	13.2	7.78	4.43	41.8
State 3	29.1	33.16	3.75	253.7
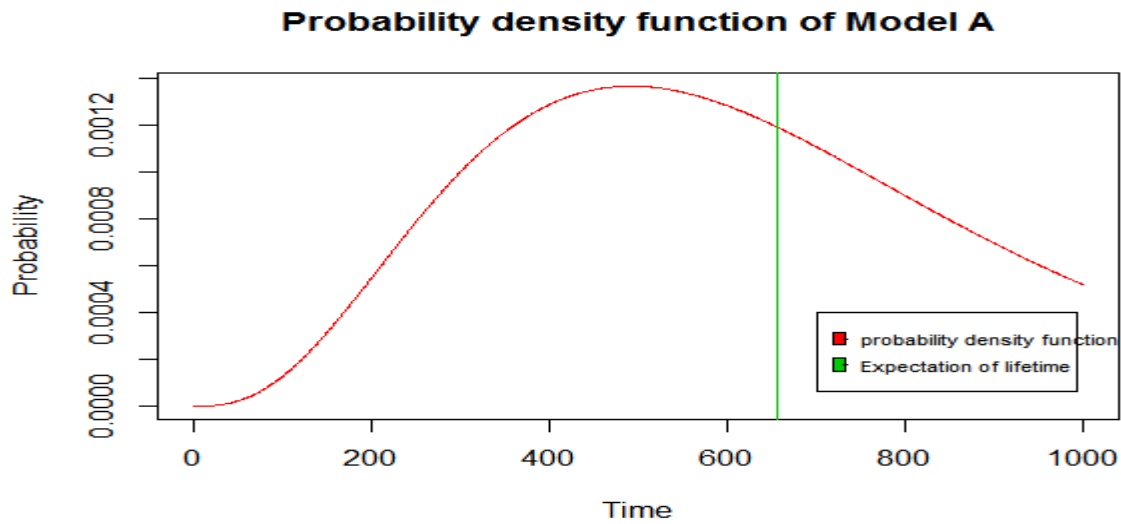
Figure 4: Probability density function of Model A
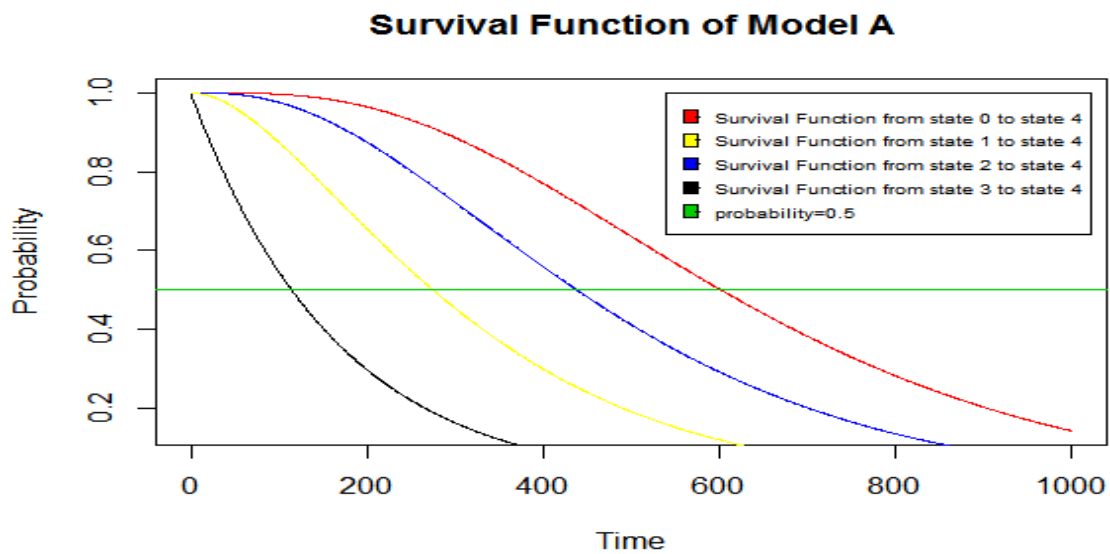
Figure 5: Survival plot of Model A

For model B, the probability density function is shown in the Figure 6 by using the equation (32). The Figure 6 also indicates that the mean lifetime of pole is 290 and its probability density function increases very fast from 0 to 150 years and it decreases slowly after 150 years.

In order to compare the mean sojourn time spent in different states visually, it is necessary to plot the survival function from initial state, such as state 0, 1, 2, and 3, to absorbing state, namely state 4, and it is shown in Figure 7. Since the time point, when survival function is equal to 0.5, is approximately equal to the mean sojourn time for wood poles. According to the crossover points of the Figure 7, it is easy to find that the Markov process spends around 50% of life time in state 0, 20% of lifetime in state 1 and the rest in state 2 and 3.
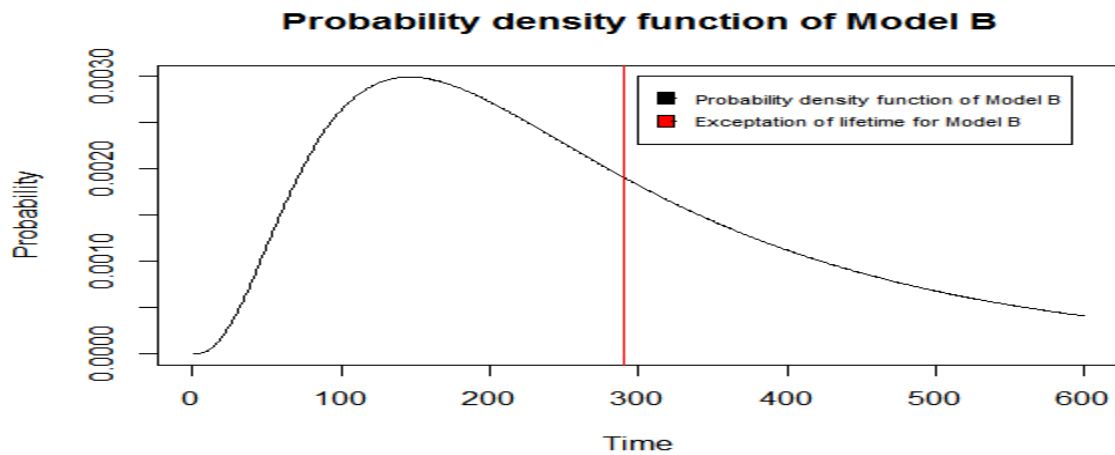
Figure 6: Probability density function of Model B
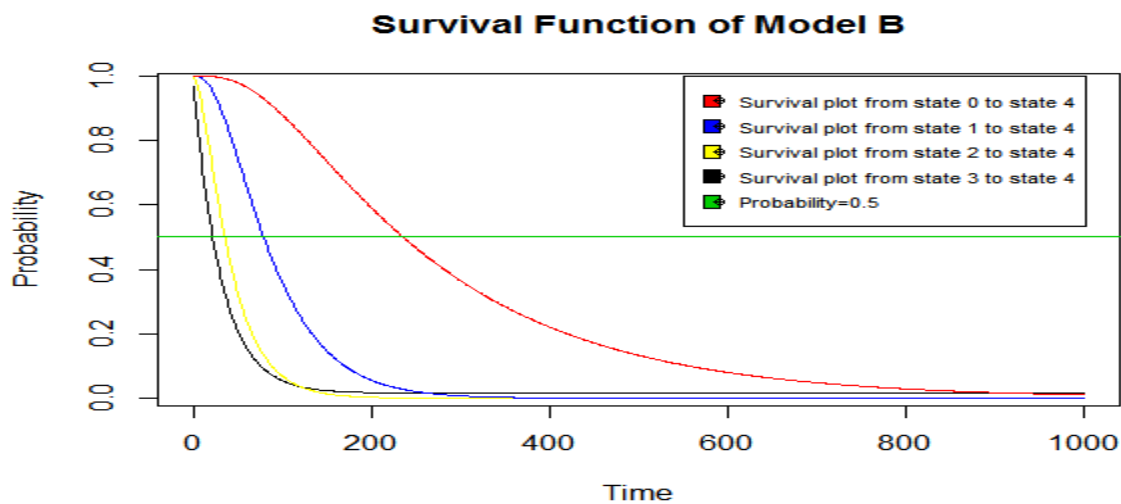
Figure 7: Survival Function of Model B

4.3.2 Result of Model C and Model D

By applying equation (47), the result of time inhomogeneous continuous Markov process is shown in Table 11. According to the table above, it indicates that there is a positive relationship between intensity and age of poles, since b_i is larger than 1. As is known to us, the time homogeneous Markov process with sojourn time belonging to exponential distribution is one of special case of the time inhomogeneous Markov Process when $b_i = 1$. So the comparison of Model A and C would give us more information about the characteristics of poles and the difference between time-independent Markov process and time-dependent Markov process. By applying the equation (14) and result of Table 9, the cumulative intensities of model A and C are shown in Figure 8.

In Figure 8, the red and black lines represents cumulative intensity $\Lambda(T_{i,j})$ of Model A and Model C respectively. Since the intensity of Model A is constant, the cumulative intensity would be linear, but the intensity of Model C is change with time, so the cumulative intensity is non-linear. The conclusion is that the cumulative failure intensity of Model A is larger than Model C before 25 years, but the cumulative failure intensity of Model C increase so fast after 25 years and it is larger than Model A. This can be explained by the age of poles since the pole would spend more time in the initial state and the failure probability will increase more and much faster alongside with the lifetime. By applying the equation (7) and (49), the probability transition matrix is a function of time. It is possible to find the remain lifetime given how long it started from initial states, such as state 0, 1, 2 and 3, to final state 4. Its uncertainty of time can be explained as lifetime distribution through equation (36), which is displayed in Figure 9. From the cross-over point of survival function plot, it indicates that the wood pole would spend approximately 20% of lifetime in state 0, 20% of lifetime in state 1, 25% of lifetime in state 2 and 35% of lifetime in state 3. In application, the Model C will spend less time in good conditions compared with the time spent in bad conditions such as the state 3. The result of time inhomogeneous Markov process model D is shown in Table 12.

Table 11: Result of Model C

State	Time inhomogeneous continuous Markov process (Model C)	
i	a_i	b_i
	0.001617743	1.415699726

Figure 8: Comparison between Model A and Model C
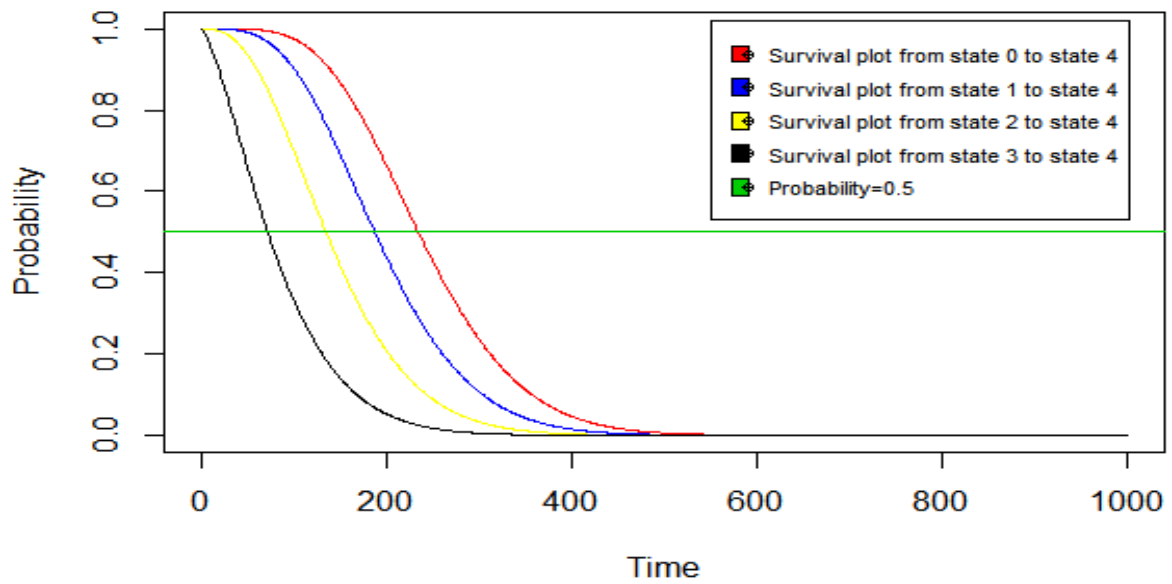
Figure 9: Survival function plot of Model C

Table 12: Parameters of Model D

State	Time inhomogeneous continuous Markov process with different $\lambda_i(t) = a_i b_i t^{b_i-1}$ with $i \in E$ Model D	
i	a_i	b_i
0	0.00537	1.00005
1	0.01519	1.04365
2	0.08541	1.03794
3	0.17817	1.00000

According to the Table 12, we can get that the values of parameter a_i are increasing from state 0 to state 3 and the values of parameter b_i of state 0 and state 3 are equal to 1, which indicates the intensities of state 0 and 3 may be independent of the time. Generally speaking, the values of parameter b_i are almost the same with each other. Then the values of a_i have main effects on the intensities of states. It means that the process will stay in state 0 for a longer time than the other states and the sojourn time in state 0 is the longest time within all of the sojourn times for all states, since the $\lambda_1(t)$ is the smallest.
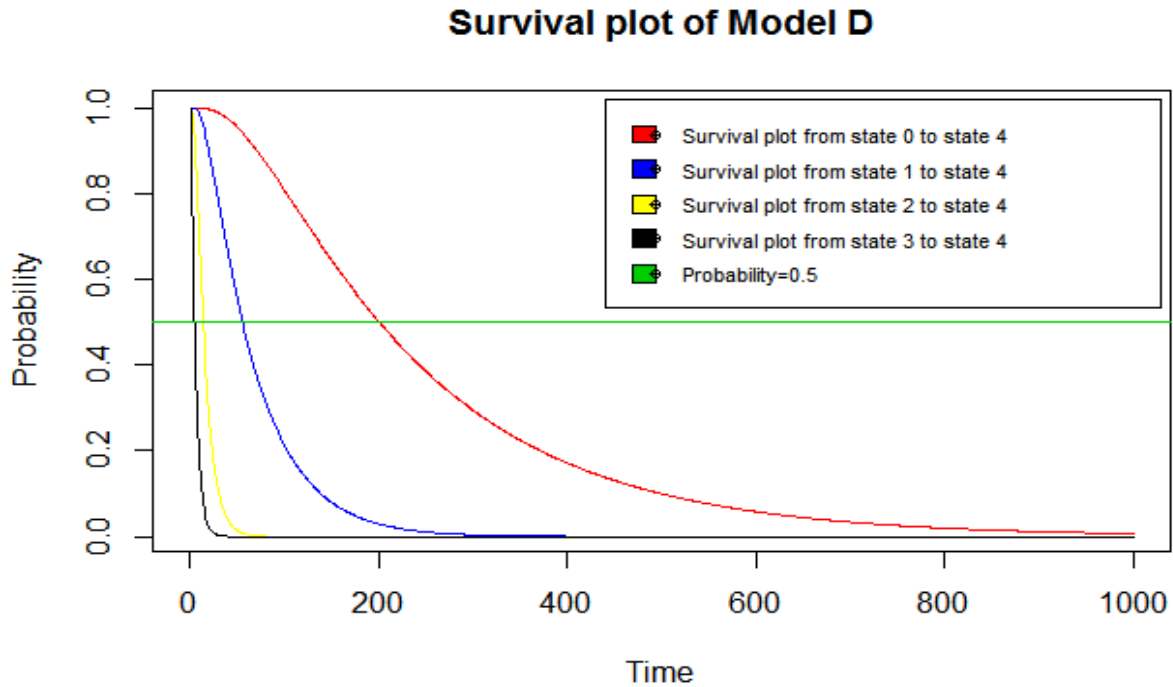
Figure 10: Survival Plot of Model D

By applying the results in Table 12, it is possible for us to determine the time to reach the absorbing state, namely state 4, when the process starts from the function states, namely state 0, 1, 2, and 3. So this kind of uncertainty in the time to failure state is shown through the survival function of time in Figure 10.

As is shown in the Figure 10, it indicates the survival function from function states to the failure state. The survival plot indicates that the wood pole will spend around 80% of lifetime in state 0 and 20% of lifetime in the other states.

4.3.3 Result of semi-Markov process with sojourn time belonging to Weibull distribution

For semi-Markov process with sojourn time belonging to Weibull distribution, the first maximum likelihood estimation is used to do the parameter estimation. According to the property of Weibull distribution, the mean time of each state is computed by applying the equation (40). The result is shown in Table 13.

The table above indicates that the Markov process stays in state 0 for a long period compared with the state 1 and state 2. As it mentioned above, the total lifetime of the poles is the sum of mean time in every state. So the lifetime in semi-Markov process with sojourn time belonging to Weibull distribution is 98.07 years. By having a look at the standard error of parameters, the standard errors of parameters for state 1, 2 and state 3 are too large which indicates few information of latter states results in bad parameter estimates. Because of little observations, the estimates of state 1, 2 and 3 are not reasonable according to large standard errors.

Table 13: Result of semi-Markov process with sojourn time belonging to Weibull distribution

State	Coefficient		Standard error	Mean sojourn time
0	β_0	61.8	3.5302595	55.37
	η_0	3.23	0.3377198	
1	β_1	7.76	12.26334	7.3
	η_1	7.96	122.57866	
2	β_2	6.93	12.84374	6.56
	η_2	9.09	1626.36959	
3	β_3	29.62	4938149	28.84
	η_3	20.35	10116552	

4.4 Summary

Table 14 indicates all of lifetime for 5 models. For time homogeneous Markov processes, the difference between Model A and B is large, it may show that the property of state-dependent has a large influence on the lifetime of the poles. By comparing the model A and C, the difference between 50% of lifetimes of Model A and C may show that the property of time-dependent also has large influence on the lifetime of poles. The 50% quantiles of lifetime of Model B, C and D are similar with the result of GompitZ model [17], which shows the lifetime of poles is around 220 years. However, the quantiles of lifetime of Model B, C and D are so close to each other. Thus it is so difficult for us to find the best model from Model B, C and D. By comparing the value of quantiles of lifetime of Model A with the other models, its value is too large. On the contrary, the lifetime of semi-Markov process with sojourn time belonging Weibull distribution is close to 100 years, which is so short compared with the 50% quantile of lifetime of other models.

In order to carry on the depth analysis, the comparisons of the plots are necessary. Figure 11 shows the comparison of survival plots for Model A, B, C, D and E. The plot indicates the lifetime curves of Model B, C and D are similar with each other. The Model A and Model E are different with the other models. The approximate lifetime of Model B, C and D are close to each other and it is equal to around 230 years according the crossover points, the lifetime of Model E is the smallest and the mean lifetime of Model A is the largest compared with the mean lifetime of the other models.

In conclusion, Model B, C and D may be the suitable models according to the comparison [17]. Thus, the comparison of models is necessary, which is shown in following sections. The time homogeneous Markov models and time inhomogeneous Markov models are nested, and the properties of state-dependent and time-independent have a large influence on the lifetime of poles. The Model A is not suitable model for the original dataset, since the standard error is larger the value of the parameters. The Model E is time independent model and its value is so different with the other model. In fact, we can't ignore the influence of continuous time, which is why Model E may be not a good choice. In this paper, Model D is a general model, since it is one general model for the Model A, B and C, and it includes the influence of time-dependent and state-dependent, which makes it to be more applicable compared with the other models in the report.

Table 14: Approximate quantiles of lifetime of Markov processes

Model	Approx. quantiles		
	25%	50%	75%
A	415	602	837
B	143	234	373
C	181	235	296
D	125	200	331
E	84	98	110
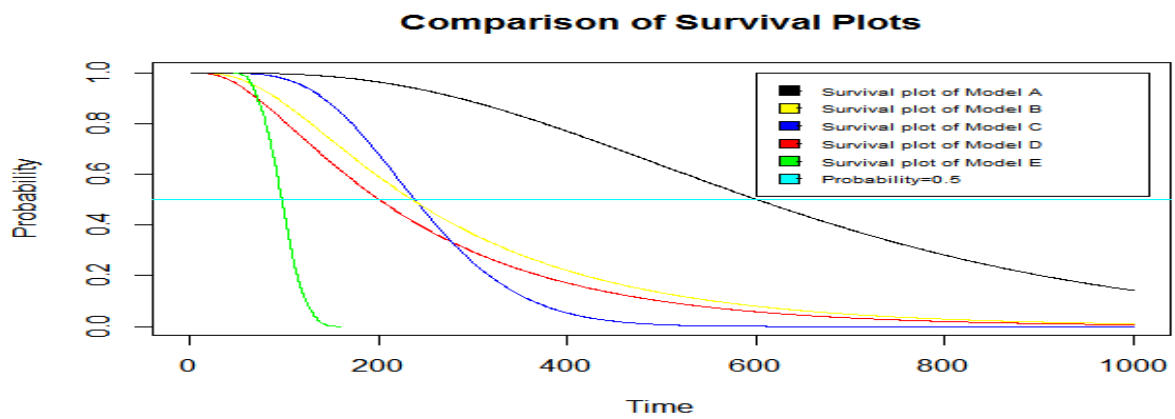
Figure 11: Comparison of Survival Plots

Chapter 5: Evaluation of Markov models

5.1 Comparison of models

One way of checking the model is to compare it with the general model maximal number of parameters, which is saturated model. As mentioned before, Model A, B, C and D are nested model, that is Model A, B and C are one of special case of Model D, so we consider Model D as a general model in this paper. Thus, we can use the likelihood ratio test to compare the models which are nested. In statistics, the likelihood ratio test is a statistical test, it is frequently used to compare the fit of two models. The test is based on the likelihood ratio statistics, and it can be used to compute a test statistics and compared to a critical value to decide whether to reject the null model in favor of the general model. The likelihood ratio statistics, which is belonging to the chi-square distribution, is shown below [18]:

$$D = 2\ln\left(\frac{\text{Likelihood for general model}}{\text{Likelihood for model under H}_0}\right)$$

$$= 2[\ln(\text{Likelihood for full model}) - \ln(\text{Likelihood for model under H}_0)]$$

(52)

The likelihood ratio statistics has a Chi-Square distribution with $K_2 - K_1$ degrees of freedom, where K_2 and K_1 denote the number of parameters in the general model and the model under H_0 respectively. If $D < \chi^2_{0.05,K_2-K_1}$, where $\chi^2_{0.05,K_2-K_1}$ is the 0.05-quantile of Chi-Square distribution with $K_2 - K_1$ degrees of freedom, H_0 should be rejected. By taking model A and model D as an example, the hypothesis is shown below:

$$H_0: \text{Model A is true}$$
$$H_1: \text{Model D is true}$$

By insert the negative likelihood value of Table 15 into the likelihood ratio statistics, likelihood ratio test statistics is shown below.

Table 15: Negative log likelihood results of models

Model	State-dependent	Time-dependent	Number of parameters	Negative log-likelihood
Model A	No	No	1	-328.3538
Model B	Yes	No	4	-313.7357
Model C	No	Yes	2	-297.3978
Model D	Yes	Yes	8	-297.0001

$$D_1 = 2(-297.0001 + 328.3538) = 62.7074 > 14.067 = \chi^2_{0.05,7} \qquad (53)$$

which indicate that null hypothesis is rejected, that is Model A has a poor description of dataset compared with Model D. It is the same for the comparison of Model B and D and comparison of Model C and D, the procedures of the test are shown respectively in equation (54) and (55).

$$D_2 = 2(-297.0001 + 313.7357) = 33.4712 > 9.488 = \chi^2_{0.05,4} \qquad (54)$$

$$D_3 = 2(-297.0001 + 297.3978) = 0.7954 < 12.59 = \chi^2_{0.05,6} \qquad (55)$$

Finally, Model B has a poor description of dataset compared with the Model D, and Model C is better than the model D according to (55), although model D is the general model and has more parameters compared with the other models. For model C, it has two parameters in the model, maybe this indicates that model D is an overfitting model. Thus, model C is the best model compared with Model A, B and D.

By checking whether Model C is good model or not, it necessary to check the standard error and p-value of both parameters. Some of the results are shown in Table 16. First, the standard errors of parameters are very small, and they are all far less than the values of the parameters. By checking the p-value of parameters, both parameters are significant at the 5% level, which indicates Model C has a good the description of dataset.

Table 16: Evaluation of Model C

Model C	Coefficient	Standard Error	Z	p-value
a	0.001617743	0.0003413179	4.739695	2.140404e-06
b	1.415699726	0.0507671294	27.886149	0.000000e+00

5.2 Simulation

In order to evaluate the models, the tools of simulation are useful in practice. There are several advantages to carry out a simulation on software rather than actually making the designs of project and testing model. As is known to us, building the design, testing and redesigning for everything can be a large and expensive project. Simulations can remove the building and rebuilding model out of the process by using the model already used in the design process. So it would help us save enough time and money.

According the result of comparison of models by likelihood ratio test, the Model C is the

best model compared with the other models in proposed research, the simulation based on the Model C will be more useful compared with the simulation based on the other models. In this section, the Model C is chosen to do simulation based on the parameters $\mathbf{a = 0.00162}$ and $\mathbf{b = 1.4157}$. The mean lifetime of Model C is around 235 years. For simulation process, 300 values are sampled from Possion distribution with intensity of different time intervals,

$$\text{Number of failures in } (s,t) \sim \text{Poisson}(\int_s^t \lambda(u)du) \tag{56}$$

where $\int_s^t \lambda(u)du$ represents the intensity of Poisson distribution from s to t. The sampling intensities of each state are shown in Table 17.

Since most of the poles still stay in the first state after 40 years and the lifetime of poles is around 250 years, 4 periodic inspections carried out every 40 years starting from the installation date are chosen in this simulation. The sample values are shown in Table 18 by using the code "rpois" in R software. In order to use the sample values in R, it is necessary to transfer the sample values into the form which can be analyzed by R computer software. By combining properties of Poisson distribution, the Markov process will stay in the first state, if the sum of value in row is equal to 0. The Markov process will be in final state (state 4), if the sum of sample values is no less than 4. Some of the sample values are larger than 2, it indicates that at least two failures happened within 40 years. However, the total failure is no larger than 4. The same holds the other cases.

Since the simulation will be carried out on the R computer software, the data should be transferred to the correct form which is shown in the Table 19. In order to compare the simulation data with the original data, the summarized data frame is shown in Table 20, which indicates that more and more transitions happened between four states. By applying the methods in chapter 4 and the results of the models are shown in Table 21 and Table 22.

Table 17: Sampling mean number of events of different time intervals

Time interval	Intensity
$(0, 40)$	$\lambda_0 = \int_0^{40} abt^{b-1}\, dt = 0.3002972$
$(40, 80)$	$\lambda_1 = \int_{40}^{80} abt^{b-1}\, dt = 0.5008632$
$(80,120)$	$\lambda_2 = \int_{80}^{120} abt^{b-1}\, dt = 0.6212073$
$(120,160)$	$\lambda_3 = \int_{120}^{160} abt^{b-1}\, dt = 0.7150414$

Table 18: Sample value from Poisson distribution with different intensity

Sample value with λ_0	Sample value with λ_1	Sample value with λ_2	Sample value with λ_3	Sum of sample values
0	0	0	0	0
0	0	0	0	0
⋮	⋮	⋮	⋮	⋮
0	0	0	0	1
0	0	0	0	1
⋮	⋮	⋮	⋮	⋮
0	1	1	0	2
0	1	1	0	2
⋮	⋮	⋮	⋮	⋮
0	2	0	1	3
0	0	3	0	3
⋮	⋮	⋮	⋮	⋮
0	2	3	1	6
0	2	1	1	4

Table 19: Input data of R

Installation	state	First inspection time (year)	state	Second inspection time (year)	state	Third inspection time (year)	state	Fourth inspection time (year)	state
0	0	40	0	80	0	120	0	160	0
0	0	40	0	80	0	120	0	160	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
0	0	40	0	80	0	120	0	160	1
0	0	40	0	80	0	120	0	160	1

⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
0	0	40	0	80	1	120	2	160	2
0	0	40	0	80	1	120	2	160	2
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
0	0	40	0	80	2	120	2	160	3
0	0	40	0	80	0	120	3	160	3
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
0	0	40	0	80	2	120	4	160	4
0	0	40	0	80	2	120	3	160	4

Table 20: Table of transitions of the simulation data

From\to	0	1	2	3	4
0	483	198	61	6	0
1	0	139	76	33	8
2	0	0	69	37	19
3	0	0	0	27	19
4	0	0	0	0	25

Table 21: Simulation result of Markov models

State	Model A	Model B	Model C		Model D		Model E	
i	λ	λ_i	a	b	a_i	b_i	β_i	η_i
0	0.013052	0.010968	0.007447	1.10305	0.007198	1.065937	97.228631	1.596601
1	0.013052	0.015129	0.007447	1.10305	0.015272	1.000000	57.06482	4.2253002
2	0.013052	0.015978	0.007447	1.10305	0.009065	1.062756	52.240345	1.295354
3	0.013052	0.014991	0.007447	1.10305	0.011972	1.075039	49.351623	1.542848

Table 22: Approx. quantiles of the simulation for models

Models of simulation	Approx. quantiles		
	25%	50%	75%
A	194	281	391
B	181	262	366
C	198	275	373
D	189	269	370
E	177	221	277

By applying the results above, all of the mean sojourn times increase a bit except the Model A. For model A, the intensity of simulation is higher than the original intensity of Model A and approximate quantiles of the lifetime are all smaller than the original results. For Model B, almost all intensities of simulation are smaller than the original results. However, the

interquartile range is $(181, 366)$, which is smaller and more reasonable compared with the interquartile range of original result $(143, 373)$. It indicates the result of Model B becomes more and more applicable, since more and more data are included in the dataset. For Model C, the simulation values of parameters are a little different form the original values of parameter, since the simulating result of a is larger than the original result and the simulating result of b is smaller than the original result. Meanwhile, the 25% and 50% quantiles of the simulating lifetime are close to the original result of Model C, and the reasons of difference between the original results and simulation results will be discussed in chapter 6. For Model D, the results of the parameter are almost different and all of the quantiles of lifetime are a little larger than the results of Model D based on the simulation dataset. For Model E, the simulation of mean lifetime and the scale parameters are much larger than the original model (Semi-Markov process with sojourn time belonging the Weibull distribution). The mean lifetime of pole is 234.3207 years and 50% quantile of lifetime is 221, which are almost the same with simulation model (Model C).

In conclusion, the results of models become more and more precise, it is all because of the improvement of dataset. By comparing the simulation results of models with the results of original models, the differences of the lifetime between the models are becoming smaller and smaller, since more and more data are added into the dataset. For example, the results of quantiles of lifetime are so different from each other, the 50% quantile of lifetime is ranging from 98 to 602 years. However, the 50% quantile of lifetime is ranging from 221 to 281 based on the simulation dataset.

Since Model A, B, C and D are nested models, it is possible to compare the simulation results of models by applying the methods in comparison of models. The negative log-likelihood values are shown in Table 23. Using equation (52), Model D is one of the best models compared with the Model A, B and C.

Table 23: Simulation results of negative log-likelihood

Model	State-dependent	Time-dependent	Number of parameters	Negative log-likelihood
Model A	No	No	1	1099.904
Model B	Yes	No	4	1091.586
Model C	No	Yes	2	819.3681
Model D	Yes	Yes	8	509.3265

5.3 Verification of Model E

For Model E, the simulating results of model E are totally different with original Model E but its mean lifetime and 50% quantile of lifetime are so close to the simulation model (Model C).

In fact, so many kinds of factors result in this condition that the mean lifetime of simulating results of model E are close to the mean lifetime and 50% quantile of lifetime of the original Model C. For example, maybe more and more data are included in the data sets, which result in the more and more applicable estimate of parameter of model E. On the other hand, the method of model E may be not correct and it has the appropriate results by coincidence or otherwise. In order to check the validity of the method of model E, the second simulation based on the simulating results of model E is carried out in following sections. Because the results of model E based on the original dataset is not the appropriate choice since the original results of state 2 and state 3 are the same which indicates the data is not enough and the estimates of parameter is not good. So the results of the model based on the dataset generated from the Model C, which is shown in Table 21, are chosen as the simulation model of semi-Markov process.

First, 300 random samples are generated on the R studio and the table is shown in Table 24. The first column shows the random times generated from the first pair of parameter of the simulation model of semi-Markov process for the state 0, the second column shows the random time generated from the second pair of parameter of the simulation model of semi-Markov process for the state 1. The rest can be deduced from the first column.

Table 24: 300 random sampled from simulation model of semi-Markov process

T_1	T_2	T_3	T_4
85.042798	60.857911	10.362073	49.295499
26.197953	27.234497	60.190359	65.396184
199.08462	54.474931	101.8884	73.885971
23.207897	175.06266	27.683402	44.913149
$\vdots$	$\vdots$	$\vdots$	$\vdots$
127.27707	20.361248	44.930916	81.466653
149.04694	37.97433	56.880658	63.345582
112.24784	41.456171	9.91369	80.257991
106.42081	197.66188	20.558458	23.210813

Second, 40 years is chosen as the fixed inspection of time interval by considering the whole row as one sample of the pole. By taking the first pole as an example, the pole stays in the state 0 at the installation time and the first inspection (40 years) and second inspection (80 years), it stays in the state 1 at the third inspection time (120 years), and it stays in the state 2 at the fourth inspection time (160 years). And the rest can be deduced from the first row. The rearranged data sets are shown in Table 25. Third, the dataset in the Table 25 should be also transferred for each state. The process is shown in the following. We take the second pole as one example and it is shown as one timeline Figure 12.

Table 25: Data template

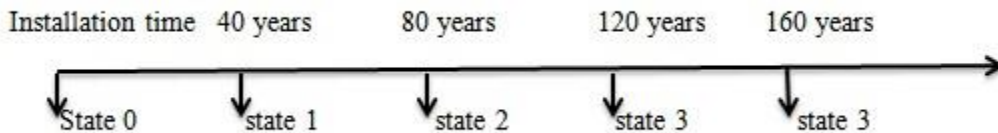| 0	0	40	0	80	0	120	1	160	2
0	0	40	1	80	2	120	3	160	3
0	0	40	0	80	0	120	0	160	0
0	0	40	1	80	1	120	1	160	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
0	0	40	0	80	0	120	0	160	2
0	0	40	0	80	0	120	0	160	1
0	0	40	0	80	0	120	1	160	2
0	0	40	0	80	0	120	1	160	1

Figure 12: Timeline of pole for simulation

The timeline above indicates the range of sojourn time in state 0, 1, 2 and 3. They are shown below:

$$0 < S_0 < 40,$$
$$0 < S_1 < 80,$$
$$0 < S_2 < 80,$$
$$40 < S_3$$

(57)

We can deduce the rest from this example. The right form in R should be like the form for state 0 in Table 26:

Table 26: Data sample for the state 0

L	R
80	120
0	40
160	Na
0	40
⋮	⋮
120	160
120	160
80	120
80	120

L indicates the left side of the time interval and R indicates the right side of time interval for each state. For instance, the third row represents the time interval is $S_i > 160$ (right censoring) and the first row represents the time interval is $80 < S_i < 120$ (interval censoring). The rest can be deduced as the first state (state 0).

After estimation of parameter on R studio, the simulating results were compared with first simulating result of model E which is shown in Table 27. According to the Table 27, the mean lifetimes of two simulating results are so close to each other. The most interesting point is that the parameters of first state for estimation model are almost the same with each other and the rest of second simulating parameters of model E are all smaller than the parameters of simulation model but the values of parameters for estimation model are so close to the parameters of simulation model. However, the parameters of final state in the estimation model are so strange since the standard errors of both parameters are larger than the value of parameters, which indicates the incorrect of parameter estimation.

Table 27: Comparison of simulation model and estimation model

Simulation model					Estimation model				
β_i	*Std. Error*	η_i	Std. Error	Mean sojourn time	β_i	Std. Error	η_i	Std. Error	Mean sojourn time
97.2286	3.8790	1.5966	0.0909	87.1898	97.7260	3.6291	1.7074	0.0976	87.1698
57.0648	4.2253	4.2253	0.1262	54.4393	52.2377	3.4115	1.4057	0.1604	47.5803
52.2403	4.4401	1.2953	0.2064	48.2830	49.5706	3.9166	1.4466	0.2343	44.9613
49.3516	5.0952	1.5428	0.3873	44.4084	42.2391	47.8518	4.6134	95.9246	38.6019

According to the Table 27, although the parameters of final state in the estimation model are so strange, the other parameters are very close to the parameters of simulation. Meanwhile, the 40 years is chosen as the fixed inspection time interval and 4 inspections are assumed to be carried out, so only a little information about final state is useful to do the parameter estimation. Maybe that is why the standard errors of final state are so large. The conclusion is that the results of estimation model are almost the same with the results of simulation model, so the method behind the model E may be correct. Thus, the assumption that the method of model E is not correct but it has the appropriate results by coincidence or otherwise does not exist. More and more data included in the data set is the reason why the mean lifetime of first simulation of model E is so close to the original result of model C.

5.4 Summary

In conclusion, the Model D maybe one of the best models based on the simulation. By simulation of the models, the quantiles of the lifetime are in all models are a little larger than in the original models, except the model A. In detail, the simulation results of 50% quantiles of lifetimes are similar to the original results of Model B, C and D, but for the other simulation quantiles of lifetimes are a little different from the original results of Model B, C and D. The most important point is that the results of lifetime of each model are close to each other alongside with the improvement of the dataset. This indicates that the models in this project fit data well and that is correct to use the Markov process to model the deterioration of wood poles. By applying likelihood ratio statistics, Model D is the choice compared with the model A, B and C based on the simulation. For model E, the results of the simulation are totally different with the original result. The mean lifetime of the simulation is more or less close to the original results of model C, since the dataset is improved compared with the heavy censored dataset. However, the model E ignores the dependency of the time, which is not correct according to the mathematical theory. That is why Model D maybe the best model compared with the Model A, B, C and E based on the simulation.

By applying the intensity matrix of Model D and equation (7), it is possible to compute the probability transition matrix from installation date to the end of lifetime. Then the cumulative probability function of the remaining lifetime of a pole is easy to compute given the state and the time counted from installation date by using equation (15).

Chapter 6: Discussion

For the results using Markov processes above, it is possible to fit all the models to the dataset, except Model E, which is illustrated in section 5.4. The model we choose will have a large influence on the results of parameters and mean lifetimes, since the dataset in this research is heavily censored and new and the models have different conclusions based on two likelihood ratio tests based on the different datasets.

According to the results of 50% quantiles of lifetime, it seems that the estimated lifetime of the pole is too large, since it is larger than the actual lifetime of the pole line which consists several wood poles. According to the suggestion from the technicians in the power company, the actual lifetime of pole line is ranging from 30 to 50 years. According to the theory on the book of Statistical methods for reliability data [15], the lifetime of a single wood pole is reasonable compared with the lifetime of the pole line. Assume that the lifetime of every pole is belonging to the exponential distribution and all of wood poles have the same intensity in a pole line. The structure of a pole line is assumed to be series with several wood poles.

$$\text{MTTF}_\text{p} = \frac{1}{\lambda_p} \text{ and MTTF}_\text{L} = \frac{1}{\lambda_L} \tag{58}$$

where MTTF_p and λ_p denote the mean lifetime to failure (MTTF) and intensity of wood pole respectively, MTTF_L and λ_L represents mean lifetime to failure and intensity of pole line respectively.

According the definition of survival probability of series system, the formula is

$$R_L = \prod R_{P_i} \tag{59}$$

where P_i is the number of pole, R_L and R_{P_i} denote the survival probability of pole line and single pole respectively, and $R_{P_i} = e^{-\lambda_{P_i} t}$. Thus we can get

$$R_L = \prod_1^N R_{P_i} = e^{-\Sigma_1^N \lambda_{P_i}} \tag{60}$$

$$\text{MTTF}_\text{L} = \frac{1}{\Sigma_1^N \lambda_p} = \text{MTTF}_\text{p}/N \tag{61}$$

Table 28: Relationship between $MTTF_L$ and $MTTF_p$

$MTTF_L$\No. Poles	10	25	50	100
5	50	125	250	500
10	100	250	500	1000
25	250	625	1250	2500
50	500	1250	2500	5000

If the lifetime of pole line is from the 5 to 50, Table 28 can be produced. The grey row represents the number of wood poles, the grey column represents the different lifetimes from 5 to 50 years and the rest represent the lifetime of the single pole. For example, when having a line with 10 poles and requiring MTTF of 5 years for the line of poles, the pole must have minimum MTTF of 50 years when the lifetime of a single pole is exponentially distributed. From the tables, we can see that we must require a quite high MTTF of each pole when the line consists of several wood poles. Generally speaking, the MTTF of a single pole is larger than the MTTF of a pole line.

In addition, the original result and simulation result of Model C are different from each other, which are shown in chapter 5. Generally speaking, it is impossible to get the same result with original results based on the simulation in this present research, since the dataset generated form R software is changed. For example, the last row but one in Table 17 indicates the total failures of a pole is larger than 4, but it is transferred into the row in Table 18 manually, which indicates the pole stays in state 4 from third and fourth inspection, according to the procedure of simulation in chapter 5. The first change is that the maximum number of failures of pole is assumed to be equal to 4, no matter how many failures of a pole is generated within the time interval (0, 160), and the second change is that the pole will stay in state 4 all the time when more than 4 failures happened before the fourth inspection. However, the total number of data is 300, maybe it is not enough and this kind of situation that final state is observed frequently does not exist in the original dataset, since only two inspections are available and few poles reached the final state. Maybe that is why the original result and the simulation result are different. Meanwhile, the values of parameters are so small, the difference will be so obviously compared with the value of parameters, if the dataset is changed a little.

By comparing the properties of degradation processes and Markov processes, it is good to use Markov process to analyze the deterioration of a pole. Although the methods in this report

seem to be suitable for the original data, they all have some advantages and disadvantages in different applications.

For Markov process with NHPP, the advantage of using NHPP is that it can be applied to the inspection with fixed time interval and inspection of non-fixed time interval. Three Markov processes by applying NHPP provide the comparison between models and a good way to find the best model. The failures, which are considered as the independent increments, are independent and the number of failures in a time interval is independent of the number of failures in any earlier time intervals. Meanwhile, it also has some disadvantages. According to the properties of Poisson distribution and assuming that the waiting time is exponentially distributed, the number of transitions belongs to the Poisson distribution. If more than one transition occurred between two inspections, it is difficult to find the right time of each transition and only the number of transitions is known, in that case the intensity of NHPP is consistent (HPP). Although NHPP is a continuous probability distribution that expresses the probability of a given number of events occurring in a time interval, NHPP just counts the number of the events in the time interval and it does not count the different types of the event in any time interval.

For Model A, it is a special case of NHPP with constant intensity, which is called the time homogeneous Markov process. It records the number of transitions in the inspection time intervals. Model A is independent of time and state which indicates that Model A does not include the information when the transition happened and which kind of event happened. In fact, the speed of the deterioration process is changing with time, i.e. the intensity of deterioration is time varying. Thus, Model A is time-independent and state-independent. Meanwhile, the transition intensity of this model is so small and it ignores the dependency of state and time, which indicates that Model A is not the perfect model. Meanwhile, the results of likelihood ratio tests also indicated the Model A is not the choice in this presented research.

For Model B, it is also a special case of a non-homogeneous Poisson process with different intensities for each state which is belonging to the time homogeneous Markov process. This model includes information such as the number of transition within the inspection time interval and the types of events. However, its intensities are independent of time which indicates that this model does not include the information when the transition happened. Thus, Model B is state-dependent and time-independent and the results of likelihood ratio tests also indicated the Model A is not the choice in this presented research in present research.

For Model C, it is the standard non-homogeneous Poisson process which is belonging to the time inhomogeneous continuous Markov process with time-varying intensities. It can help us to find the number of transitions which happened in the inspection time interval and the time point when the transition happened between two inspections. However, the assumption behind this model is that all of events are the same, which indicates that Model C is time-dependent and state-independent and it does not contain the different types of event in this model. The mean lifetime of a pole for this model is around 235 years which is similar with the result of Gompitz model [17]. According to results of the hypothesis tests, Model C may be one of the best models in this proposed research.

For model D, it is based on non-homogeneous Poisson process which is belonging to the time inhomogeneous Markov process with time-varying and state-varying intensities. It includes the information such as the number of transitions happened, the time point when the transition happened and the type of the events in the inspection time interval. It seems to be the best model according to the mathematical theory since it overcomes all of the disadvantages of Model A, B and C. Its 50% quantile of lifetime is 200 years which is also close to the result of the Gompitz model [17]. However, Model D may be an overfitting model according to the likelihood ratio statistics based on the original dataset. Meanwhile, the Model D may be the best model compared with the Model A, B and C according to the likelihood ratio test based on the dataset of simulation. Different kinds of results based on the original and simulation datasets do not indicate that Model D is not one of best models in this research.

For Model E (semi-Markov process with Weibull distribution), Weibull distribution is a continuous probability distribution and it is also one of the lifetime distributions which is most widely used in reliability and survival analysis. Weibull distribution can display the characteristics of other types of distributions depending on the value of the shape parameter. In this project, Weibull distributions with different parameters are used to simulate the states of a pole. The assumption behind this application is that sojourn times in all states are independent. For instance, the lifetime of a pole can be divided as five states and how long time the pole stays in the first state does not influence the time staying in the other states. Actually, this may be not true. It can be explained by quantitative analysis.

For example, we take the No.3 of pole in Table 2 as one example and it is shown as one timeline in Figure 3. Assuming that the first time is the installation time (2000) of pole, the

timeline above indicates the range of sojourn time in state 0, 1, 2 and 3. They are shown below:

$$
\begin{aligned}
&0 < S_3, \\
&0 < S_0 < 3, \\
&0 < S_1 + S_0 < 3, \\
&S_3 + S_2 + S_1 + S_0 < 10, \\
&10 < S_3 + S_2 + S_1 + S_0
\end{aligned}
\tag{62}
$$

Assuming that S_3, S_2, S_1 and S_0 are independent and they are all Weibull distributed, the contribution to the likelihood of poles is shown below:

$$
\begin{aligned}
&Likelihood\ contribution\ of\ a\ pole \\
&= \int_0^3 \int_0^{3-S_0} \int_{3-S_0-S_1}^{10-S_0-S_1} \int_{10-S_0-S_1-S_2}^{\propto} f_0(S_0)f_1(S_1)f_2(S_2)f_3(S_3)dS_0dS_1dS_2dS_3
\end{aligned}
\tag{63}
$$

where $f_i(S_i)$ is the probability density function of sojourn time in the states. According to the CDF above and equation (62), we know sojourn time staying in one state is depending on the sojourn time staying in the other state. So the Markov process with sojourn time belonging to the Weibull distribution ignores the dependency of the time spent in each state.

Meanwhile, overlap of time intervals will happen such as the range of S_1 and S_0 in equation (62). The transitions from the good states to state 3 and state 4 are rare, that is the data of states 3 and 4 is not enough, which would result incorrect estimates of parameters. Although semi-Markov process contains the censored and non-censored information, the dataset does not include the complete and enough non-censored information in the dataset which makes the parameter estimations very difficult.

Chapter 7: Conclusion

By comparing the different results of the models, the conclusion is that it is possible to fit all models to the dataset except the Model E, and different models have different results in this presented research, since the dataset is so new and heavily censored. The lifetime of the poles is ranging from 200 to 300 years according to the results of five models.

The study of estimating parameters based on Markov processes has been discussed in the presented research. The information of the original data of wood poles has been displayed in chapter 2. It is feasible to use Markov process to model the deterioration process.

However, the information for the more accurate and more detailed analysis is not available in the original data. In order to support the statistical analysis, the dataset should be designed depending on the requirements of statistics. Many factors, such as incomplete and wrong data registration, personal factors and too few observations for poles, are reasons that why the more objective analyses should be conducted.

According to the discussion above, all of three Markov processes also have some disadvantages. There are some evidences that the models in this project have some disadvantages and a more applicable time inhomogeneous Markov process should be constructed which can provide a better description of the dataset.

As is known for us, if few observations are available, the maximum likelihood method will be not suitable for parameter estimation of all models. In this project, we have sufficient number of poles in the data set, but we do not have enough observations for each pole. In detail, the dataset has enough information for state 0 and state 1, and it does not include enough information of state 2, state 3 and state 4. Thus the models above can describe the beginning of a deterioration process such as transitions from state 0 to state 1. The models have a poor description of the middle and end of the deterioration process of poles such as transitions from state 2 to state 3 and from state 3 to state 4. Thus, it is necessary to update the dataset we have.

If the data source is not sufficient, we should find alternative information. In the literature, some people would like to combine data and expert judgments together to analyze practical problems. Welte and Eggen used the Bayesian approach for estimation of parameters of sojourn time distribution based on expert opinion and monitoring data, when the data set contains a little

information [19]. Mohammadi, Longinow and Williams described the process of refinement of monitoring data and how to do statistical evaluation of models compared with data collection from experts opinion [20]. Thus, the expert judgment from related fields will be useful and precious. For example, we can get information about expectation of the sojourn time in each state, we can measure the value of parameters compared with the expert judgment and evaluate the models. Many methods about taking advantage of the expert judgment could be used in this research.
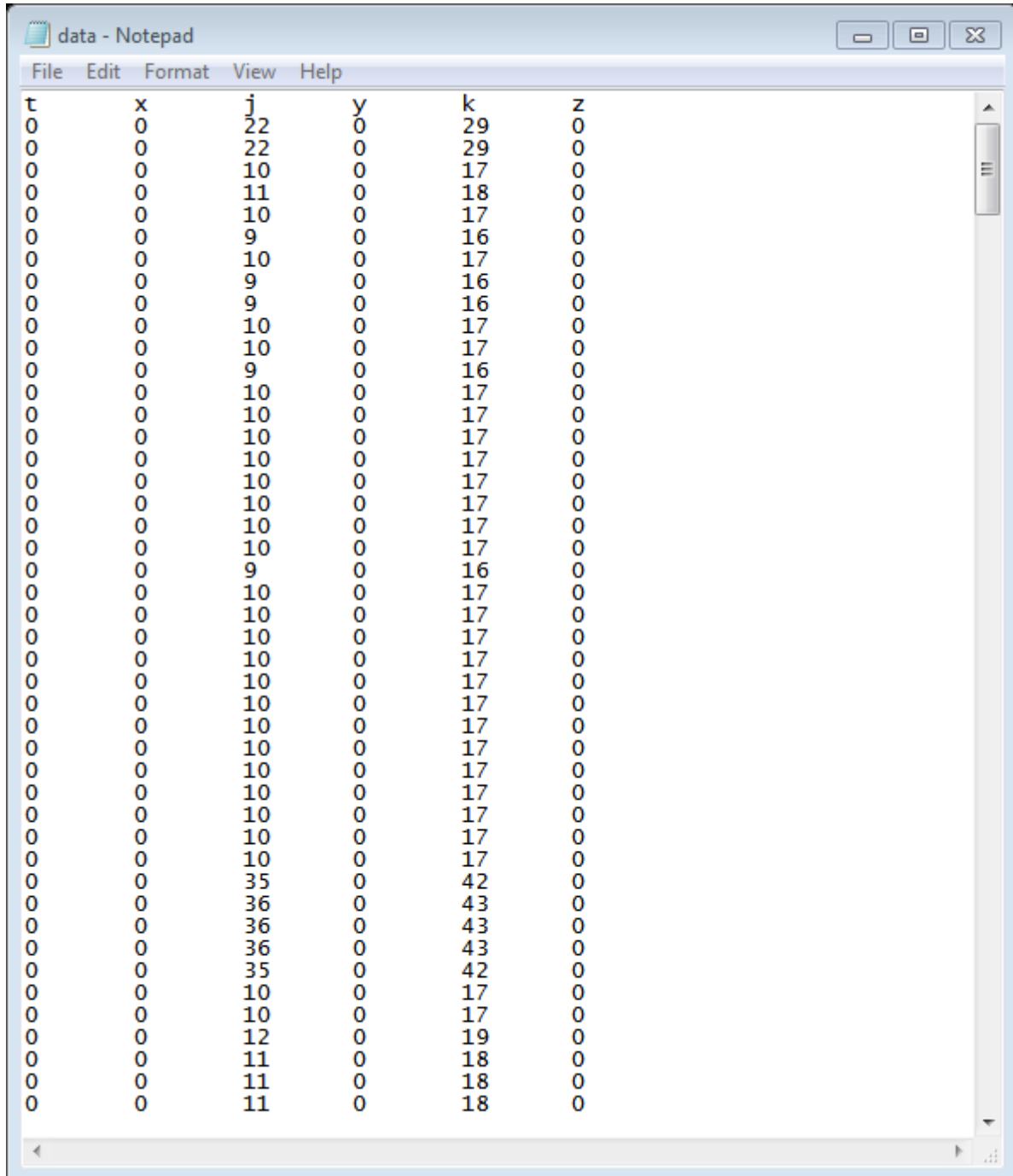
References

1. Kallen, M.J. Markov processes for maintenance optimization of civil infrastructure in the Netherlands. Ph.D. Dissertation.Electrical Engineering, Mathematics and Computer Science Delft University of Technology. 2007

2. Chan, G.K. and S. Asgarpoor, Optimum maintenance policy with Markov processes. *Electric Power Systems Research*. **76**(6–7): p. 452-456. 2006

3. Zhang, X. and H. Gao, Road maintenance optimization through a discrete-time semi-Markov decision process. *Reliability Engineering & System Safety*. **103**(0): p. 110-119. 2012

4. Kallen, M.J. and J.M. van Noortwijk, Optimal periodic inspection of a deterioration process with sequential condition states. *International Journal of Pressure Vessels and Piping*. **83**(4): p. 249-255. 2006

5. Jackson, C., Multi-State Models for Panel Data: The msm Package for R. *J Stat Softw*. **38**: p. 1-28. 2011

6. Chen, D. and K.S. Trivedi, Optimization for condition-based maintenance with semi-Markov decision process. *Reliability Engineering & System Safety*. **90**(1): p. 25-29. 2005

7. Tomasevicz, C.L. and S. Asgarpoor, Optimum maintenance policy using semi-Markov decision processes. *Electric Power Systems Research*. **79**(9): p. 1286-1291. 2009

8. Ross, S.M., *Introduction to Probability Models (Tenth Edition)*, Boston: Academic Press. 2010

9. Nielsen, S.F., Continuous-time homogeneous Markov chains. 2009

10. Kallen, M.J. and J.M.V. Noortwijk. Statistical inference for Markov deterioration models of bridge conditions in the Netherlands. in *Proceedings of the Third International Conference on Bridge Maintenance, Safety and Management (IABMAS*. Netherlands: Taylor & Francis Group. 2006

11. Esary, J.D. and A.W. Marshall, Shock Models and Wear Processes. *The Annals of Probability*. **1**(4): p. 627-649. 1973

12. A-Hameed, M.S. and F. Proschan, Nonstationary shock models. *Stochastic Processes and their Applications*. **1**(4): p. 383-404. 1973

13. Høyland, M.R.a.A., *System Reliability Theory; Models, Statistical Methods and Applications*, Hoboken, New Jersey: Wiley. 2004

14. Moore, G., Orthogonal polynomial expansions for the matrix exponential. *Linear Algebra and its Applications*. **435**(3): p. 537-559. 2011

15. Escobar, W.Q.M.a.L.A., *Statistical methods for reliability data*, New York: Wiley-Interscience. 1998

16. Welte, T.M. and H. Kile. Parameter estimation in degradation modelling: A case study using condition monitoring data from wood pole inspections. in *PowerTech, 2011 IEEE Trondheim*. Year

17. Welte, T., GompitZ applied to power line poles, SINTEF. 2010

18. Dobson, A.J., *An introduction to generalized linear models* Boca Raton Chapman & Hall/CRC. 2002

19. Welte, T.M. and A.O. Eggen. Estimation of Sojourn Time Distribution Parameters Based on Expert Opinion and Condition Monitoring Data. in *Proceedings of the 10th International Conference on Probabilistic Methods Applied to Power Systems* Rincon. Year

20. Mohammadi, J., A. Longinow, and T.A. Williams, Evaluation of system reliability using expert opinions. *Structural Safety*. **9**(3): p. 227-241. 1991

Appendix 1

Input for Model A, B, C and D

```
data - Notepad
File   Edit   Format   View   Help
t        x        j        y        k        z
0        0        22       0        29       0
0        0        22       0        29       0
0        0        10       0        17       0
0        0        11       0        18       0
0        0        10       0        17       0
0        0        9        0        16       0
0        0        10       0        17       0
0        0        9        0        16       0
0        0        9        0        16       0
0        0        10       0        17       0
0        0        10       0        17       0
0        0        9        0        16       0
0        0        10       0        17       0
0        0        10       0        17       0
0        0        10       0        17       0
0        0        10       0        17       0
0        0        10       0        17       0
0        0        10       0        17       0
0        0        10       0        17       0
0        0        10       0        17       0
0        0        9        0        16       0
0        0        10       0        17       0
0        0        10       0        17       0
0        0        10       0        17       0
0        0        10       0        17       0
0        0        10       0        17       0
0        0        10       0        17       0
0        0        10       0        17       0
0        0        10       0        17       0
0        0        10       0        17       0
0        0        10       0        17       0
0        0        10       0        17       0
0        0        10       0        17       0
0        0        35       0        42       0
0        0        36       0        43       0
0        0        36       0        43       0
0        0        36       0        43       0
0        0        35       0        42       0
0        0        10       0        17       0
0        0        10       0        17       0
0        0        12       0        19       0
0        0        11       0        18       0
0        0        11       0        18       0
0        0        11       0        18       0
```

Figure 13

56

where t, j and k are installation time, first inspection time and second inspection time respectively. Then x, y and z are state of installation, state of first inspection and state of second inspection respectively.

Estimation of Model A:

R program of Model A:

```
ModelA=function(p,t,j,k,x,y,z){
  m=matrix(c(-p[1],p[1],0,0,0,
             0,-p[1],p[1],0,0,
             0,0,-p[1],p[1],0,
             0,0,0,-p[1],p[1],
             0,0,0,0,0),
    nrow=5,ncol=5,byrow=TRUE)
    mle=0
    for (l in c(1:475)){
    m1=expm(m*j[l])
    m2=expm(m*(k[l]-j[l]))
    mle=mle+log(m1[x[l],y[l]])+log(m2[y[l],z[l]])}
  print(p)
  return(-mle)}
parmeter1= optimize(ModelA,c(0.001,0,1),
                t=data$t,
                j=data$j,
                k=data$k,
                x=data$x+1,
                y=data$y+1,
                z=data$z+1)
```

Output

```
$minimum
[1] 0.006091755 #### Value of parameter

$objective
[1] 328.3538 #### log likelihood statistics
```

Plot of Model A:

```
R1=0
R2=0
R3=0
R4=0
FF1=function(t){lambda^n*t^(n-1)*exp(-lambda*t)/factorial(n-1)}
FF2=function(t){lambda^3*t^(3-1)*exp(-lambda*t)/factorial(3-1)}
FF3=function(t){lambda^2*t^(2-1)*exp(-lambda*t)/factorial(2-1)}
FF4=function(t){lambda^1*t^(1-1)*exp(-lambda*t)/factorial(1-1)}
for (i in c(1:1000)){
  R1[i]=FF1(i)
  R2[i]=FF2(i)
  R3[i]=FF3(i)
  R4[i]=FF4(i)}
plot(1-cumsum(R1),col="red",,type="l",main="Survival Function of Model
A",ylab="Probability",xlab="Time")
lines(1-cumsum(R2),col="blue",type="l")
lines(1-cumsum(R3),col="yellow",type="l")
lines(1-cumsum(R4),col="black")
legend(430,1,c("Survival Function from state 0 to state 4",
               "Survival Function from state 1 to state 4",
               "Survival Function from state 2 to state 4",
               "Survival Function from state 3 to state 4"),
       cex=0.7,c("red","yellow","blue","black"),pch=20,lty=0)
```

Estimation of Model B:

R program of Model B:

```
ModelB=function(p,t,j,k,x,y,z){
    m=matrix(c(-p[1],p[1],0,0,0,
               0,-p[2],p[2],0,0,
               0,0,-p[3],p[3],0,
               0,0,0,-p[4],p[4],
               0,0,0,0,0),nrow=5,ncol=5,byrow=TRUE)
    mle=0
    for (l in c(1:475)){
      m1=expm(m*j[l])
      m2=expm(m*(k[l]-j[l]))
      mle=mle+log(m1[x[l],y[l]])+log(m2[y[l],z[l]])}
    print(p)
    return(-mle)}
par1= optim(c(0.0061,0.0061,0.0061,0.0061),ModelB,
                t=data$t,
                j=data$j,
                k=data$k,
                x=data$x+1,
                y=data$y+1,
                z=data$z+1)
```

Output of Model B:

```
$par
[1] 0.005076535 0.020766575 0.075734637 0.034398854 #### Value of parameter
$value
[1] 313.7357 #### log likelihood statistics
$counts
function gradient
301        NA  #### Number of iterations
$convergence
[1] 0 #### It indicates the convergence
$message
NULL
```

Plot of Model B:

```
a1=0.005076
a2=0.020766
a3=0.075734637
a4=0.034398854
FF4=function(t){ffff=a1*exp(-a1*t)*(a2/(a2-a1))*(a3/(a3-a1))*(a4/(a4-a1))+
  a2*exp(-a2*t)*(a1/(a1-a2))*(a3/(a3-a2))*(a4/(a4-a2))+
  a3*exp(-a3*t)*(a2/(a2-a3))*(a1/(a1-a3))*(a4/(a4-a3))+
  a4*exp(-a4*t)*(a2/(a2-a4))*(a1/(a1-a4))*(a3/(a3-a4))}
FF3=function(t){ffff=
  a2*exp(-a2*t)*(a3/(a3-a2))*(a4/(a4-a2))+
  a3*exp(-a3*t)*(a2/(a2-a3))*(a4/(a4-a3))+
  a4*exp(-a4*t)*(a2/(a2-a4))*(a3/(a3-a4))}
FF2=function(t){ffff=
  a3*exp(-a3*t)*(a4/(a4-a3))+
  a4*exp(-a4*t)*(a3/(a3-a4))}
FF1=function(t){ffff=
  a4*exp(-a4*t)}
RR1=0
RR2=0
```

```
RR3=0
RR4=0
for (i in c(1:1000)){
  RR1[i]=cumsum(FF1(i))
  RR2[i]=cumsum(FF2(i))
  RR3[i]=cumsum(FF3(i))
  RR4[i]=cumsum(FF4(i))}
plot(1-cumsum(RR1),col="red",,type="l",main="Survival    Function    of    Model
B",ylab="Probability",xlab="Time")  ###survival  function  of  firth  time  to
failure
lines(1-cumsum(RR2),col="blue",type="l")
lines(1-cumsum(RR3),col="yellow",type="l")
lines(1-cumsum(RR4),col="black")
legend(260,1,c("Survival Function from state 0 to state 4",
                "Survival Function from state 1 to state 4",
                "Survival Function from state 2 to state 4",
                "Survival Function from state 3 to state 4"),
       cex=0.7,
       c("red","yellow","blue","black"),pch=20,lty=0)
```

Estimation of ModelC:

R program of ModelC:
```
ModelC=function(p,t1,t2,t3,x1,x2,x3){
  logl=sum((x3-x2)*log(p[1]*(t3^p[2])
                       -p[1]*(t2^p[2]))
          +(x2-x1)*log(p[1]*(t2^p[2]))
          -p[1]*(t3^p[2])
          +p[1]*(t1^p[2]))
  print(p)
  result=c(-logl,sd(p))
  return(-logl)}

  parameter3=optim(c(0.003,1),ModelC,method="BFGS",
                   t1=data$t,
                   t2=data$j,
                   t3=data$k,
                   x1=data$x+1,
                   x2=data$y+1,
                   x3=data$z+1)
```
Output:
```
$par
[1] 0.00162 1.41576  #### Value of parameters
$value
[1] 297.3978 #### log likelihood statistics
$counts
function gradient
52       6    #### Number of iterations
$convergence
[1] 0       #### It indicates the convergence
$message
NULL
```

Plot of Model C:

```
p=c(0.00162,1.4157)
L = list(diag(5))
o3=0
k3=0
l3=0
v3=0
for (i in c(1:1000)){
  t=i-1
  m=matrix(c(-p[1]*p[2]*t^(p[2]-1),p[1]*p[2]*t^(p[2]-1),0,0,0,
            0,-p[1]*p[2]*t^(p[2]-1),p[1]*p[2]*t^(p[2]-1),0,0,
            0,0,-p[1]*p[2]*t^(p[2]-1),p[1]*p[2]*t^(p[2]-1),0,
            0,0,0,-p[1]*p[2]*t^(p[2]-1),p[1]*p[2]*t^(p[2]-1),
            0,0,0,0,0),nrow=5,ncol=5,byrow=TRUE)
  result = matrix(unlist(L[i]),ncol=5)%*%(m+diag(5))
  o3[i]=result[1,5]
  k3[i]=result[2,5]
  l3[i]=result[3,5]
  v3[i]=result[4,5]
  L[[length(L)+1]] = result}
lines(1-o3,col="blue")
plot(1-o3,col="red",type="l",main="Survival      Function      of      Model
C",ylab="Probability",xlab="Time")
lines(1-k3,col="blue",type="l")
lines(1-l3,col="yellow",type="l")
lines(1-v3,col="black")
legend(260,1,c("Survival Function from state 0 to state 4",
              "Survival Function from state 1 to state 4",
              "Survival Function from state 2 to state 4",
              "Survival Function from state 3 to state 4"),
      cex=0.7,
      c("red","blue","yellow","black"),pch=20,lty=0)
```

Estimation of Model D:

R program of Model D:

```
ModelD=function(p,j,k,x,y,z){
  L = list(diag(5))
  for (i in c(1:160)){
    t=i-1
    m=matrix(c(-p[1]*p[2]*t^(p[2]-1),p[1]*p[2]*t^(p[2]-1),0,0,0,
              0,-p[3]*p[4]*t^(p[4]-1),p[3]*p[4]*t^(p[4]-1),0,0,
              0,0,-p[5]*p[6]*t^(p[6]-1),p[5]*p[6]*t^(p[6]-1),0,
              0,0,0,-p[7]*p[8]*t^(p[8]-1),p[7]*p[8]*t^(p[8]-1),
              0,0,0,0,0),nrow=5,ncol=5,byrow=TRUE)
    result = matrix(unlist(L[i]),ncol=5)%*%(m+diag(5))
    L[[length(L)+1]] = result
  }
  mle = 0
  for (l in c(1:475)){
    o=matrix(unlist(L[j[l]]),ncol=5)
    h=matrix(unlist(L[k[l]]),ncol=5)%*%solve(o)
```

```
      mle=mle+log(o[x[l],y[l]])+log(h[y[l],z[l]])}
   print(p)
   return(-mle)}
paramter4=optim(c(0.1,1,0.1,1,0.1,1,0.1,1),ModelD,control=list(maxit=200000),
      j=data$j,
      k=data$k,
      x=data$x+1,
      y=data$y+1,
      z=data$z+1
```

Output:
```
$par
[1]  0.005367249  1.000000100  0.015193468  1.043649102  0.085412186  1.037940474
0.178171350
[8] 1.000004421 #### value of parameters
$value
[1] 297.0001 #### log likelihood statistics
$counts
function gradient
2603       NA    #### number of iterations
$convergence
[1] 0   #### it indicates the convergence
$message
NULL
```

Plot of Model D:
```
p=c(0.005367249,   1.000040100,   0.015193468,   1.043649102,   0.085412186,
1.037940474, 0.178171350,1.000004421)
L = list(diag(5))
k4=0
o4=0
v4=0
l4=0
for (i in c(1:1000)){
  t=i-1
  m=matrix(c(-p[1]*p[2]*t^(p[2]-1),p[1]*p[2]*t^(p[2]-1),0,0,0,
             0,-p[3]*p[4]*t^(p[4]-1),p[3]*p[4]*t^(p[4]-1),0,0,
             0,0,-p[5]*p[6]*t^(p[6]-1),p[5]*p[6]*t^(p[6]-1),0,
             0,0,0,-p[7]*p[8]*t^(p[8]-1),p[7]*p[8]*t^(p[8]-1),
             0,0,0,0,0),nrow=5,ncol=5,byrow=TRUE)
  result = matrix(unlist(L[i]),ncol=5)%*%(m+diag(5))
  print(result)
  o4[i]=result[1,5]
  k4[i]=result[2,5]
  l4[i]=result[3,5]
  v4[i]=result[4,5]
  L[[length(L)+1]] = result}
plot(1-o4,col="red",,type="l",main="Survival        plot        of       Model
D",ylab="Probability",xlab="Time")  ###survival  function  of  firth  time  to
failure
lines(1-k4,col="blue",type="l")
lines(1-l4,col="yellow",type="l")
lines(1-v4,col="black")
abline(h=0.5,col="3")
```

```
legend(440,1.02,c("Survival plot from state 0 to state 4",
                  "Survival plot from state 1 to state 4",
                  "Survival plot from state 2 to state 4",
                  "Survival plot from state 3 to state 4",
                  "Probability=0.5"),
       cex=0.7,
       c("red","blue","yellow","black","3"),pch=10,lty=0)
```

Estimation of Model E:
```
library(splines)
library(survival)
library(fitdistrplus)
####weibull distribution for state 2
d1=data.frame(
  left=state.1$v1,
  right=state.1$V2)
weibull1=fitdistcens(d1,"weibull")
summary(weibull1)
####weibull distribution for state 2
d2=data.frame(
      left=state.2$v1,
     right=state.2$V2)
 weibull2=fitdistcens(d2,"weibull")
 summary(weibull2)
####weibull distribution for state 3
d3=data.frame(
  left=state.3$v1,
  right=state.3$V2)
weibull3=fitdistcens(d3,"weibull")
summary(weibull3)
####weibull distribution for state 3
d4=data.frame(
  left=state.4$v1,
  right=state.4$V2)
weibull4=fitdistcens(d4,"weibull")
summary(weibull4)
```

Appendix 2

Code of Simulation

Model A:

```
model1=function(p,t1,t2,t3,t4,t5,s1,s2,s3,s4,s5){
  m=matrix(c(-p[1],p[1],0,0,0,
             0,-p[1],p[1],0,0,
             0,0,-p[1],p[1],0,
             0,0,0,-p[1],p[1],
             0,0,0,0,0),nrow=5,ncol=5,byrow=TRUE)
  mle=0
  for (l in c(1:300)){
    m1=expm(m*t2[l])
    m2=expm(m*(t3[l]-t2[l]))
    m3=expm(m*(t4[l]-t3[l]))
    m4=expm(m*(t5[l]-t4[l]))

mle=mle+log(m1[s1[l],s2[l]])+log(m2[s2[l],s3[l]])+log(m3[s3[l],s4[l]])+log(m4
[s4[l],s5[l]])}
  return(-mle)}
parameter11= optimize(model1,c(0.000001,1),
             t1=simulation.data$V1,
             t2=simulation.data$V3,
             t3=simulation.data$V5,
             t4=simulation.data$V7,
             t5=simulation.data$V9,
             s1=simulation.data$V2+1,
             s2=simulation.data$V4+1,
             s3=simulation.data$V6+1,
             s4=simulation.data$V8+1,
             s5=simulation.data$V10+1)
###output
$minimum
[1] 0.01305441

$objective
[1] 1099.904
```

Model B:

```
model2=function(p,t1,t2,t3,t4,t5,s1,s2,s3,s4,s5){
  m=matrix(c(-p[1],p[1],0,0,0,
             0,-p[2],p[2],0,0,
             0,0,-p[3],p[3],0,
             0,0,0,-p[4],p[4],
             0,0,0,0,0),nrow=5,ncol=5,byrow=TRUE)
  mle=0
  for (l in c(1:300)){
    m1=expm(m*t2[l])
    m2=expm(m*(t3[l]-t2[l]))
    m3=expm(m*(t4[l]-t3[l]))
    m4=expm(m*(t5[l]-t4[l]))

mle=mle+log(m1[s1[l],s2[l]])+log(m2[s2[l],s3[l]])+log(m3[s3[l],s4[l]])+log(m4
[s4[l],s5[l]])}
  return(-mle)}
parameter22= optim(c(0.013052,0.0130521,0.013052,0.013052),model2,
```

```
            t1=simulation.data$V1,
            t2=simulation.data$V3,
            t3=simulation.data$V5,
            t4=simulation.data$V7,
            t5=simulation.data$V9,
            s1=simulation.data$V2+1,
            s2=simulation.data$V4+1,
            s3=simulation.data$V6+1,
            s4=simulation.data$V8+1,
            s5=simulation.data$V10+1)
### output
$par
[1] 0.01096761 0.01512947 0.01597799 0.01499122

$value
[1] 1091.586

$counts
function gradient
131        NA

$convergence
[1] 0

$message
NULL
```

Model C

```
model3=function(p,t1,t2,t3,t4,t5,s1,s2,s3,s4,s5)
{logl=sum((s5-s4)*log(p[1]*(t5^p[2])-p[1]*(t2^p[2]))
        +(s4-s3)*log(p[1]*(t4^p[2])-p[1]*(t3^p[2]))
        +(s3-s2)*log(p[1]*(t3^p[2])-p[1]*(t2^p[2]))
        +(s2-s1)*log(p[1]*(t2^p[2])-p[1]*(t1^p[2]))
        -p[1]*(t5^p[2])
        +p[1]*(t1^p[2]))
  return(-logl)}
  parameter33=optim(c(0.003,0.1),model3,method="BFGS",
            t1=simulation.data$V1,
            t2=simulation.data$V3,
            t3=simulation.data$V5,
            t4=simulation.data$V7,
            t5=simulation.data$V9,
            s1=simulation.data$V2,
            s2=simulation.data$V4,
            s3=simulation.data$V6,
            s4=simulation.data$V8,
            s5=simulation.data$V10)
### output

$par
[1] 0.007447071 1.103050950

$value
[1] 819.3681

$counts
function gradient
189       35

$convergence
[1] 0

$message
```

```
NULL
```

Model D

```
model4=function(p,s1,s2,s3,s4,s5){
    L = list(diag(5))
    for (i in c(1:160)){
      t=i-1
      m=matrix(c(-p[1]*p[2]*t^(p[2]-1),p[1]*p[2]*t^(p[2]-1),0,0,0,
                0,-p[3]*p[4]*t^(p[4]-1),p[3]*p[4]*t^(p[4]-1),0,0,
                0,0,-p[5]*p[6]*t^(p[6]-1),p[5]*p[6]*t^(p[6]-1),0,
                0,0,0,-p[7]*p[8]*t^(p[8]-1),p[7]*p[8]*t^(p[8]-1),
                0,0,0,0,0),nrow=5,ncol=5,byrow=TRUE)
      result = matrix(unlist(L[i]),ncol=5)%*%(m+diag(5))
      L[[length(L)+1]] = result
    }
    mle = 0
    for (l in c(1:300)){
      o1=matrix(unlist(L[40]),ncol=5)
      o2=matrix(unlist(L[80]),ncol=5)%*%solve(o1)
      o3=matrix(unlist(L[120]),ncol=5)%*%solve(o2)
      o4=matrix(unlist(L[160]),ncol=5)%*%solve(o3)
        mle=mle+log(o1[s1[l],s2[l]])+log(o2[s2[l],s3[l]])
        +log(o3[s3[l],s4[l]])
        +log(o4[s4[l],s5[l]])}
    print(p)
    return(-mle)}
paramter44=optim(c(0.1,1,0.1,1,0.1,1,0.1,1),model4,control=list(maxit=20000),
                s1=simulation.data$V2+1,
                s2=simulation.data$V4+1,
                s3=simulation.data$V6+1,
                s4=simulation.data$V8+1,
                s5=simulation.data$V10+1)

### output

$par
[1] 0.007197686 1.065937308 0.015271995 1.000000596 0.009064900 1.062756488
0.011972459
[8] 1.075039386
$value
[1] 506.5859
$counts
function gradient
1503      NA

$convergence
[1] 0

$message
NULL
```