Two methods to approximate the superposition of imperfect failure processes

Shaomin Wu
Kent Business School, University of Kent, Canterbury, Kent CT2 7FS, United Kingdom
email: s.m.wu@kent.ac.uk

Suggested citation: Wu, S. (2020) Two methods to approximate the superposition of imperfect failure processes, Reliability Engineering and System Safetty, DOI: 10.1016/j.ress.2020.107332

Abstract

Suppose a series system is composed of a number of repairable components. If a component fails, it is repaired immediately and the effectiveness of the repair may be imperfect. Then the failure process of the component can be modelled by an imperfect failure process and the failure process of the system is the superposition of the failure processes of all components. In the literature, there is a bulk of research on the superimposed renewal process (SRP) for the case where the repair on each component is assumed perfect. For the case that the component causing the system to fail is unknown and that repair on a failed component is imperfect, however, there is little research on modelling the failure process of the system. Typically, the likelihood functions for the superposition of imperfect failure processes cannot be given explicitly. Approximation-based models have to be sought. This paper proposes two methods to model the failure process of a series system in which the failure process of each component is assumed an arithmetic reduction of intensity and an arithmetic reduction of age model, respectively. The likelihood method of parameter estimation is given. Numerical examples and real-world data are used to illustrate the applicability of the proposed models.

Key words: Arithmetic reduction of intensity (ARI) model; arithmetic reduction of age (ARA) model; superimposed ARI (SARI) model; superimposed ARA (SARA) model.

November 23, 2020

1. Introduction

In the reliability literature, repair effectiveness can be categorised into perfect, imperfect and minimal. Suppose an item failed. A perfect repair on the item is equivalent to replacing the item with a new identical item, that is, it brings the failed item to the good-as-new status; a minimal repair restores the item to the status just before the item failed, namely, it brings the failed item to the bad-as-old status; and an imperfect repair brings the item to a status between the good-as-new and bad-as-old statuses. Usually, the renewal process is used to model the interfailure times of perfect repairs; the non-homogeneous Poisson process is for those of minimal repairs; and models such as the arithmetic reduction of intensity (ARI) model and the arithmetic reduction of age (ARA) model are for those of imperfect repairs [1, 2, 3]. There is a bulk of research discussing different types of stochastic processes for modelling failure processes, or simply put, modelling interfailure times, see [1, 2, 3, 4], for example. These models are also applied in maintenance policy optimisation, see [5, 6, 7, 8, 9, 10, 11, 12], for example.

Consider a system that is composed of multiple components in series. Suppose that the failures of the components are statistically independent. Repair is immediately performed upon a component failure and the repair time is negligible. Suppose the effectiveness of the repair is not minimal.

- If the repair is perfect, then the interfailure times of the system is a superimposed renewal process (SRP). The SRP has received plenty of attention from authors (see [13, 14, 15], for example). The reader is referred to [3] for a recently published paper of SRPs in reliability.
- If the repair effectiveness on the failure of each component is imperfect, the superposition of the imperfect failure processes has not been well investigated in the literature yet.

We refer to the case that the components that cause the system to fail are known and that interfailure time data are available as *unmasked failure data*. With unmasked failure data, if the number of failures of each component is large enough, one may develop a model for the failure process of each component and then aggregate those models to describe the failure process of the system. In the real world, nevertheless, maintenance data may be available in an aggregate form. That is, the interfailure time data are available, but which component causes the system to fail may be unknown. Such data are often referred to as *masked failure data*. In the case of modelling on masked failure data, one is unable to build a model for the failure process of each component and then aggregate them as there is no failure data on each individual component. That is, the superposition of imperfect failure processes (SIRP) cannot be explicitly given. In this case, methods that can approximate the SIRP have to be sought.

Examples of such systems can be found from the real-world. For example, one can regard that a section of pavement is composed of a grid of cells. The section may be regarded failed if there is a large defect such as fatigue cracking on a cell. Maintenance should then be carried out to repair the defect and it is usually imperfect. But when the failure data are analysed, data on which cell causes the section to fail may be unavailable due to various reasons such as a lack of precise location. For the pavement owner, it is important to have a model of good performance that can estimate the long term costs of maintaining the pavement. See [16] for other real-world examples. These examples motivate this work.

Given a series system composed of multiple components, this paper assumes that the failure process of each component can be modelled with either the arithmetic reduction of intensity (ARI) model or the arithmetic reduction of age (ARA) model. The reason that this paper uses the ARI and ARA models is due to their wide coverage. Some widely studied models, including the model proposed by [17] and the virtual models proposed by [18, 19], are the special cases of the ARI model and the ARA model, respectively. This paper proposes methods to approximate the superposition of ARI (SARI) model and the superposition of ARA (SARA) model, respectively. Probabilistic properties of the proposed methods are discussed. Artificially generated numerical examples and real-world examples are used to illustrate the proposed methods. This paper extends the work of [4]. Its managerial implication is that practitioners may use the proposed methods in their work such as development of maintenance policies and life cycle costing.

The remainder of this paper is structured as follows. Section 2 gives assumptions and notations that are used in the paper. Section 3 investigates the superposition of imperfect failure processes (SIRP) for the situations when unmasked failure data are available and then gives a method of simulating such an SIRP. Section 4 proposes methods to approximate the superposition of the ARI process and that of the ARA process for the case when only masked failure data are available, respectively. Section 5 gives the likelihood functions of the SARI and the SARA, respectively, and verifies the proposed methods on an artificially generated dataset and then on a real-world dataset. Section 6 discusses an alternative method to approximate the SIRP for the case when only truncated failure data are available and also gives the failure intensity function of the SRP. Section 7 concludes the paper and proposes future research suggestions.

2. Assumptions and notations

This section sets notations and assumptions.

2.1. Notations

The notations in Table 1 will be used in the paper.

2.2. Assumptions

- Suppose a series system is composed of n components, whose interfailure times are statistically independent.
- The failure intensity function of component k is $\frac{1}{n}\lambda_k(t)$ before its first failure.
- Repair is immediately performed upon the failure of a component (or the system). The effectiveness of repair may be perfect, imperfect, or minimal.

Table 1: Notation Table

n	number of components
k	index of a component in the system, $k = 1, 2, \cdots, n$
t	time since the system starts, where $t \geq 0$
N_t	number of failures of the system up to time t
$N_{k,t}$	number of failures of component k up to time t
N	total number of failures of the system
i	index of failures, $i = 1, 2, \cdots, N$
m	order of memory, where $m \geq 1$ in the ARI_m model or the ARA_m model
j	index of the order of memory, $j = 1, 2, \cdots, \infty$
τ	time since the completion of the latest repair, where $\tau > 0$
$\lfloor x \rfloor$	the largest integer that is smaller than or equal to x
b_{N_t}	$b_{N_t} = N_t - n \times \lfloor \frac{N_t}{n} \rfloor$ if $N_t \neq n \times \lfloor \frac{N_t}{n} \rfloor$, $b_{N_t} = n$ otherwise.
T_{N_t}	successive failure times of a repairable system; T_{N_t} is a random variable
T_{k,N_t}	time to the ith failure of component k after the N_tth repair; T_{k,N_t} is a random variable
t_{N_t}	observation of T_{N_t}
t_{k,N_t}	observation of T_{k,N_t}
$\mathscr{H}_{t-}$	history of the failure process up to t (exclusive of t)
$\lambda_k(t)$	failure intensity function of component k if minimal repair is conducted upon failures
$\Lambda_k(t)$	$= \int_0^t \lambda_k(u) du$
$\lambda_{s,k}(t)$	failure intensity function of component k if imperfect repair may be conducted upon failures
$\lambda_s(t)$	failure intensity function of the system at time t
$\lambda_s^a(t)$	approximated failure intensity function of the system at time t
$h_c(t)$	hazard function of a virtual component at time t

- The failure process of a component can be modelled by either the ARI model or the ARA model.
- The failure process of the system can be defined equivalently by the stochastic processes $\{T_j\}_{j\geq 1}$ or $\{N_t\}_{t\geq 0}$ and is characterised by the intensity function.
- Although the failure intensity function of an item (which may be a system or a component) should be denoted with the memory of $\mathscr{H}_{t-}$ such as $\lambda_{s,k}(t|\mathscr{H}_{t-})$ and $\lambda_s(t|\mathscr{H}_{t-})$. For the sake of notational compactness, this paper will omit the symbol $\mathscr{H}_{t-}$ and use $\lambda_{s,k}(t)$ and $\lambda_s(t)$, respectively.
- Repair time is so short that it can be neglected.
- Only the observations of $\{T_j\}_{j\geq 1}$ or $\{N_t\}_{t\geq 0}$ are available, but the source (or component) that causes the system to fail is unavailable. Such failure data is hereinafter referred to as masked failure data.

3. Modelling the failure process with unmasked failure data

In this section, we investigate some properties of SARI and SARA, respectively, assuming that the components that cause the system to fail are known. That is, the failure data are unmasked.

3.1. Related literature on failure process models for multi-component systems

In the literature, there are several papers discussing modelling methods for multi-component systems. Below we give a brief review on the work published in the last two years. More references in this area can be found in [2, 4] and [3], respectively.

There are many publications methods proposed to approximate the SRP (see [13, 14, 15] for example). The two following models, Model I and Model II, which were recently proposed in [2].

Model I is given by,

$$\lambda_s(t) = h_c(t - T_{N_t}) + \lambda_0(t), \tag{1}$$

and Model II is given by,

$$\lambda_s(t) = h_c(t - T_{N_t}) + \frac{1}{n}\left(\sum_{j=0}^{\min\{N_t-1,n-1\}} \lambda_0(t - T_{N_t-j}) + \max\{n - N_t, 0\}\lambda_0(t)\right), \tag{2}$$

where $h_c(t)$ is a hazard function and $\lambda_0(t)$ is a failure intensity function. Model I in (1) and Model II in (2) incorporate both time trends (ageing, reliability growth), which is modelled by the first element $h_c(t-T_{N_t})$, and renewal type behaviour, which is modelled by the second element in the models, respectively.

In essence, Model I and Model II integrate two stochastic processes, which requires more parameters than a single stochastic process. In reality, due to technological advances, today's technical systems have a small number of failures in their service life. It is therefore difficult to collect a good number of time-to-failure data (or interfailure time data), based on which the estimated parameters in a failure process model may have large uncertainty. To reduce the number of parameters, [4] proposes the following model, which is referred to as *Exponential Smoothing of Intensity* model (ESI), to approximate the SRP.

$$\lambda_s(t) = \frac{1}{n}\sum_{j=0}^{\min\{N_t-1,n-1\}} \rho^{n-j-1}\lambda_0(t - T_{N_t-j}) + \frac{\chi\{1 \le N_t < n\}}{n}\sum_{j=N_t}^{n-1} \rho^{n-j-1}\lambda_0(t). \tag{3}$$

where ρ is a parameter and $\rho \in [0,1]$ and $\chi\{A\} = 1$ if A is true, $\chi\{A\} = 0$ if A is false. When $\rho = 1$, the above model reduces to a model, which is referred to as the MAI (*Moving Average of Intensity*) model. According to [4], based on the comparison among ESI and MAI, and nine other existing models on fifteen real-world datasets, the MAI outperforms the ten other models on eleven datasets (out of the fifteen datasets).

Models (1), (2), and (3) are the sum of two intensity functions, which were discussed in the reliability literature for a different purpose, namely, for modelling bathtub shaped non-monotonic intensities. For example, [20] assumes the sum of two nonhomogenous Poisson processes with one intensity function being the power law and the other being the log linear law, or both being the power laws [21], or both being the log linear laws [22].

Models (1), (2), and (3) approximate the SRP (superposition of renewal processes) generated by a multi-component system, in which the repair on each component is assumed perfect. In reality, imperfect repair occurs from time to time and may be a more realistic measure of maintenance effectiveness. However, in the literature, as far as the author's best knowledge, there is little research investigating the superposition of imperfect failure processes (SIRP), which motivates the work of the current paper.

To model the failure process of a single component, reference [1] investigates several models and categorised them into two main classes: ARI$_m$ (Arithmetic Reduction of Intensity model with memory m) and ARA$_m$ (Arithmetic Reduction of Age model with memory m with $m \geq 1$).

The ARI$_m$ model for component k is given by

$$\lambda_{s,k}(t) = \frac{1}{n}\lambda_k(t) - \frac{1}{n}\rho_k \sum_{j=0}^{\min\{m-1,N_t-1\}} (1-\rho_k)^j \lambda_k(T_{N_t-j}), \tag{4}$$

and the ARA$_m$ model for component k is given by

$$\lambda_{s,k}(t) = \frac{1}{n}\lambda_k\left(t - \rho_k \sum_{j=0}^{\min\{m-1,N_t-1\}} (1-\rho_k)^j T_{N_t-j}\right). \tag{5}$$

where ρ_k is a parameter representing the repair effectiveness of component k and m is the order of the memory. That is, Eq. (4) and Eq. (5) assume that the components have different repair effectiveness (i.e., ρ_k) and the same memory m.

Reference [1] also discuss the cases when $m = 1$ and $m = \infty$ for the ARI$_m$ and the ARA$_m$ models as special cases, respectively.

Similar to the methods to approximate the SRP proposed in [2] and [4], we may explore methods to approximate the failure process of a system with the failure process of each component modelled by either ARI$_m$ or ARA$_m$, respectively, as shown in the following section.

3.2. Superposition of the ARI$_m$ and ARA$_m$ processes, respectively

Let's first look at the failure process of a typical system, as shown in Example 1.

Example 1. *Suppose a series system composed of four components, which fail at time points shown in the top four horizontal lines in Figure 1. The superposition of the four imperfect failure processes is shown at the last horizontal line. In this example, we assume that unmasked failure data are available. If the failure*

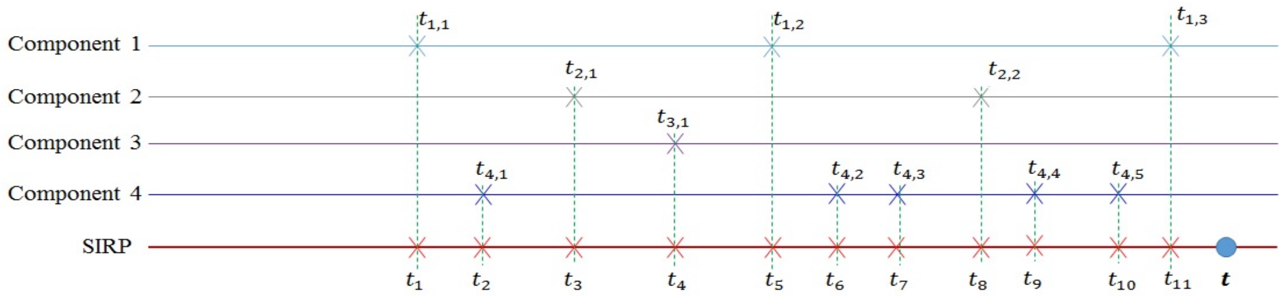

Figure 1: Failure data of a system with four components until time t, where $N_{1,t} = 3, N_{2,t} = 2, N_{3,t} = 1$, and $N_{4,t} = 5$

.

process of each component is modelled by ARI_3, then the superposition of the failure processes is given by

$$\begin{aligned}\lambda_s(t) =& \frac{1}{4}\left[\lambda_1(t) - \rho_1\lambda_1(t_{1,3}) - \rho_1(1-\rho_1)\lambda_1(t_{1,2}) - \rho_1(1-\rho_1)^2\lambda_1(t_{1,1})\right.\\ &+ \lambda_2(t) - \rho_2\lambda_2(t_{2,2}) - \rho_2(1-\rho_2)\lambda_2(t_{2,1})\\ &+ \lambda_3(t) - \rho_3\lambda_3(t_{3,1})\\ &\left.+ \lambda_4(t) - \rho_4\lambda_4(t_{4,5}) - \rho_4(1-\rho_4)\lambda_4(t_{4,4}) - \rho_4(1-\rho_4)^2\lambda_4(t_{4,3})\right]. \end{aligned} \tag{6}$$

Similarly, the superposition of the failure processes for the case when the failure process of each component is modelled by ARA_3 can be easily provided.

Now suppose that component k has $N_{k,t}$ failures that have occurred within $(0, t)$ and the latest failure occurred at time $T_{k,N_{k,t}}$. Then the superposition of the ARI_m processes is given by

$$\lambda_s(t) = \frac{1}{n}\sum_{k=1}^{n}\left(\lambda_k(t) - \sum_{j=0}^{\min\{N_{k,t}-1,m-1\}}\rho_k(1-\rho_k)^j\lambda_k(T_{k,N_{k,t}-j})\right). \tag{7}$$

The above model is referred to as $\text{SARI}_{n,m}$ (superimposed ARI) in this paper.

Similarly, the superposition of failure processes that models the failure process of each component by the ARA models, or the $\text{SARA}_{n,m}$ model can be given by

$$\lambda_s(t) = \frac{1}{n}\sum_{k=1}^{n}\lambda_k\left(t - \sum_{j=0}^{\min\{N_{k,t}-1,m-1\}}\rho_k(1-\rho_k)^j T_{k,N_{k,t}-j}\right). \tag{8}$$

According to [23], the ARI_m model and the ARA_m model have the asymptotic intensities $\lambda_{s,k}(t) = \frac{1}{n}(1-\rho_k)^m\lambda_k(t)$ and $\lambda_{s,k}(t) = \frac{1}{n}\lambda_k((1-\rho_k)^m t)$, respectively. As such, we can obtain the following Lemma.

Lemma 1. *$\lambda_s(t)$ in Eq. (7) has the asymptotic intensity $\frac{1}{n}\sum_{k=1}^{n}(1-\rho_k)^m\lambda_k(t)$ and $\lambda_s(t)$ in Eq. (8) has the asymptotic intensity $\frac{1}{n}\sum_{k=1}^{n}\lambda_k((1-\rho_k)^m t)$.*

Lemma 1 implies: $\lambda_s(t)$ in Eq. (7) (or in Eq. (8)) becomes infinite for $t \to \infty$ if $\lambda_k(t)$ is increasing in t. This result differs from the result of the SRP, on which [24] showed that the SRP tends toward (statistical) equilibrium as the time of operation becomes very large.

3.3. Simulation

In the SRP, each failed component in a series system is replaced with a new identical one. As reiterated in the preceding paragraph, [24] showed that the SRP tends toward (statistical) equilibrium as the time of operation becomes very large, which can be witnessed by viewing numerical examples shown in [3]. It will be interesting to see what trends $\text{SARI}_{n,m}$ and $\text{SARA}_{n,m}$ possess as the time of operation becomes very large. To this end, this subsection aims to use the Monte Carlo simulation to show their trends.

It is noted that $P(T_{i+1} \le t_{i+1} | T_i = t_i) = \dfrac{F(t_{i+1}) - F(t_i)}{1 - F(t_i)} = 1 - \exp(-\Lambda(t_{i+1}) + \Lambda(t_i))$.

Suppose a series system is composed of n components, which are identical when the system start at $t = 0$. Without loss of generality, let $\lambda_1(t)$ be the failure intensity function of a component in the system. Then $\Lambda_1(t) = \int_0^t \lambda_1(u)du$.

The probability of the working time of a given component, component 1, for example, after the i-th repair is given by

- $P(T_{1,1} \le t_{1,1} | T_{1,0} = 0) = 1 - \exp\left(-\Lambda_1(t_{1,1})\right)$,
- when the ARI_m model is used, for $i \ge 1$, we have

$$
\begin{aligned}
&P(T_{1,i+1} \le t_{1,i+1} | T_{1,i} = t_{1,i}, \cdots, T_{1,i-\min\{m-1,i-1\}} = t_{1,i-\min\{m-1,i-1\}}) \\
&= 1 - \exp\Bigg(-\Lambda_1(t_{1,i+1}) + \rho_1 \sum_{j=0}^{\min\{m-1,i-1\}} (1-\rho_1)^j \lambda_1(t_{1,i-j}) t_{1,i+1} \\
&\quad + \Lambda_1(t_{1,i}) - \rho_1 \sum_{j=0}^{\min\{m-1,i-1\}} (1-\rho_1)^j \lambda_1(t_{1,i-j}) t_i \Bigg)
\end{aligned}
\tag{9}
$$

- when the ARA_m model is used, for $i \ge 1$, we have

$$
\begin{aligned}
&P(T_{1,i+1} \le t_{1,i+1} | T_{1,i} = t_{1,i}, \cdots, T_{1,i-\min\{m-1,i-1\}} = t_{1,i-\min\{m-1,i-1\}}) \\
&= 1 - \exp\Bigg[-\Lambda_1 \Bigg(t_{1,i+1} - \rho_1 \sum_{j=0}^{\min\{m-1,i-1\}} (1-\rho_1)^j t_{1,i-j} \Bigg) \\
&\quad + \Lambda_1 \Bigg(t_{1,i} - \rho_1 \sum_{j=0}^{\min\{m-1,i-1\}} (1-\rho_1)^j t_{1,i-j} \Bigg) \Bigg].
\end{aligned}
\tag{10}
$$

Based on the above discussion, with the Monte Carlo simulation, we can simulate the failure process of a system based on a given failure intensity function, $\lambda_1(t)$. For example, if we let $\Lambda_1(t) = \int_0^t \lambda_1(u)du = (\frac{t}{10})^{2.0}$

in the SARI model and $\Lambda_1(t) = \int_0^t \lambda_1(u)du = (\frac{t}{10})^\beta$ in the SARA model, $\rho_1 = 0.5$, $m = 2$. Set the number n of components in a series system to be 5, 50, 100, and 200, respectively, and their numbers of failures are assumed to be $100\times n$, then the failure process according to the failure intensity function given in Eq. (7) and that given in Eq. (8) are shown in Fig. 2 for the SARI$_{n,m}$ model and Fig 3 for the SARA$_{n,m}$ model, respectively. To gain a better understanding, in Fig 3, for the different settings of n's and the numbers of failures, we also show the cases for $\beta = 1.5$, 2, and 2.5, respectively, which are displayed in each message box. In the figures, we have divided the entire failure period into 101 units. For example, if the time to the 20,000th failure is x, then we calculate the number of failures in intervals $(\frac{kx}{101}, \frac{(k+1)x}{101})$ with $k = 0, 1, ..., 101$, and show the number of failures on the Y-axis and the X-axis shows the 101 units.

Fig. 2 and Fig. 3 show that the systems do not develop toward (statistical) equilibrium as the time of operation becomes very large. Instead, they become infinity, which agrees with Lemma 1.

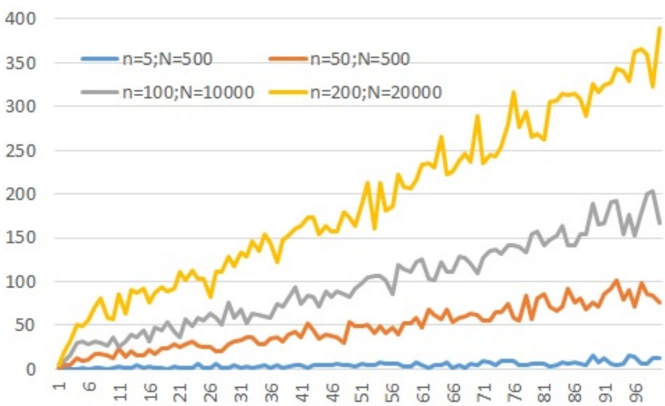

Figure 2: $\rho_1 = 0.5, m = 2, \Lambda_1(t) = (t/10)^2$, n and N are shown for different curves in the figure, for the SARI$_{n,m}$ model.

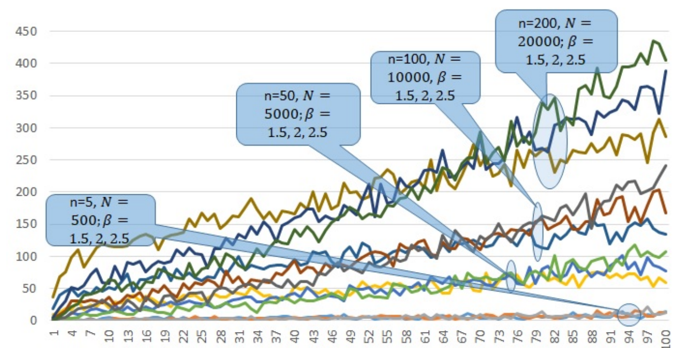

Figure 3: $\rho_1 = 0.5, m = 2, \Lambda_1(t) = (t/10)^\beta$, n, N and β are shown in the figure message boxes, respectively, for the SARA$_{n,m}$ model.

4. Modelling the failure process with masked failure data

If $T_{k,N_{k,t}}$ (for $k = 1, 2, \dots$) are known and we assume that $\lambda_k(t) = \lambda(t)$ and $\rho_k = \rho$, from Eq. (7) and Eq. (8), we obtain

$$
\begin{aligned}
\lambda_s(t) =& \frac{1}{n}\sum_{k=1}^{n}\left(\lambda_k(t) - \sum_{j=0}^{\min\{N_{k,t}-1, m_k-1\}} \rho_k(1-\rho_k)^j \lambda_k(T_{k,N_{k,t}-j})\right) \\
=& \lambda(t) - \frac{1}{n}\sum_{j=0}^{\min\{N_{k,t}-1, m_k-1\}}\left(\rho(1-\rho)^j \sum_{k=1}^{n}\lambda(T_{k,N_{k,t}-j})\right),
\end{aligned}
\tag{11}
$$

and

$$\begin{aligned}\lambda_s(t) =& \frac{1}{n}\sum_{k=1}^{n}\lambda_k\left(t-\sum_{j=0}^{\min\{N_{k,t}-1,m_k-1\}}\rho_k(1-\rho_k)^jT_{k,N_{k,t}-j}\right)\\ =&\frac{1}{n}\sum_{k=1}^{n}\lambda\left(t-\sum_{j=0}^{\min\{N_{k,t}-1,m_k-1\}}\rho(1-\rho)^jT_{k,N_{k,t}-j}\right),\end{aligned} \tag{12}$$

respectively.

Under the assumption that only masked failure data are available, unlike the SRP in which a component after a renewal can be regarded as starting from time 0, which implies the SRP model does not need to remember the component's previous maintenance/repair history. The $\text{SARI}_{n,m}$ or the $\text{SARA}_{n,m}$ processes, however, must remember all of its maintenance history. If we compare the SRP with the SIRP, we can find

- that the age of a component in the SRP is unknown as we do not know when it is installed;
- that the number of failures of a component in the SRP is always 1 as a failed component is renewed;
- that the operating/calendar age of a component in the SIRP is known as it is installed and started at time 0; and
- that the number of failures of a component in the SIRP is unknown.

In what follows, we assume the n components in the series system are identical. If the failure process of a component follows ARI_m (or ARA_m), then the failure processes of the others should follow the same model.

Under the assumption that only masked failure data are available, the value k's in $T_{k,N_{k,t}}$ in Eq. (11) or in Eq. (4) are not observable. As such, it is not possible to use these two models shown in (11) and (4).

Under the assumption that only masked failure data are available, we may have two approaches to approximating the $\text{SARI}_{n,m}$ process or the $\text{SARA}_{n,m}$ process. These two approaches are

Approach 1 to regard the system as one single item and approximate $\text{SARI}_{n,m}$ and $\text{SARA}_{n,m}$ with ARI_m and ARA_m, respectively, or

Approach 2 to take a further development of $\text{SARI}_{n,m}$ and $\text{SARA}_{n,m}$, respectively, and propose new models to approximate these two models, respectively.

Approach 1 uses the ARI_m model and the ARA_m model to approximate the series system if the failure data are masked. That is, in this approach, one regards a multi-component system to be a one-component system, then use the ARI_m model or the ARA_m model to model the failure process.

When the number of failures is small, using the ARI_m or the ARA_m to approximate an SIRP makes the implicit assumption that the kth failure depends on the $(k-1)$th, which may not be true for the case of

$N < n$ (where N is the total number of observed failures), under which each failed component may have only experienced one failure, and the failures of different components are statistically independent. After all, the probability of the occurrences of the first failures within a short time is greater than that of the second failures because: within a given time period $(0, t)$, denote $P(N(t) = k)$ as the probability that the number of failures is k, then $P(N(t) = 2) < P(N(t) = 1)$.

Although the numbers of component failures in a typical system may be different, as shown in Fig 4, the expected numbers of failures for identical components in a given time period are the same. As such, a naive but appealing approach to approximating the SIRP model is to assume that the failures of the system are caused by each component one after another. That is, suppose a system that is composed of identical components in series, we assume the first n failures are due to the first failures of the n components, the failures from the $(n+1)$-th to the $2n$-th failures are due to the second failures of the n failures, and so on. Based on this assumption, we have the following discussion.

Denote

$$b_{N_t} = \begin{cases} N_t - n\lfloor \frac{N_t}{n} \rfloor & N_t \neq n\lfloor \frac{N_t}{n} \rfloor \\ \\ n & N_t = n\lfloor \frac{N_t}{n} \rfloor, \end{cases} \tag{13}$$

where $\lfloor x \rfloor$ is the largest integer that is smaller than or equal to x.

Then we have the following analyses.

(i) **The case of the ARI model.** If we assume that the failure process of each component follows ARI_m shown in Eq. (4), we have the following analyses.

- If $N_t = 0$, the failure intensity of the system is $\lambda(t)$.
- If $0 < N_t \leq n$, there are N_t components whose first failures occur. Each of these components has failure intensity function $\frac{1}{n}\lambda(t) - \frac{1}{n}\rho\lambda(T_k)$ and each of the rest $n - N_t$ components has failure intensity $\frac{1}{n}\lambda(t)$. As such, the failure intensity function of the system after the N_t-th failure is given by

$$\begin{aligned} \lambda_s^a(t) &= \frac{1}{n}\sum_{k=1}^{N_t}(\lambda(t) - \rho\lambda(T_k)) + \frac{n - N_t}{n}\lambda(t) \\ &= \lambda(t) - \frac{1}{n}\sum_{k=1}^{N_t}\rho\lambda(T_k) \end{aligned} \tag{14}$$

- If $n < N_t \leq mn$, there are the two following scenarios.
  - If $N_t = n\lfloor \frac{N_t}{n} \rfloor$, then the failure intensity function of the system is given by

$$\lambda_s^a(t) = \lambda(t) - \frac{1}{n}\sum_{k=1}^{n}\sum_{j=0}^{\lfloor \frac{N_t}{n} \rfloor - 1}\rho(1-\rho)^j\lambda(T_{N_t - nj - k + 1}). \tag{15}$$

- If $N_t \neq n\lfloor \frac{N_t}{n} \rfloor$, then b_{N_t} components have experienced one more failure than the $n - b_{N_t}$ other components. The sum of the failure intensity functions of these b_{N_t} components is $\frac{b_{N_t}}{n}\lambda(t) - \frac{1}{n}\sum_{k=1}^{b_{N_t}}\sum_{j=0}^{\lfloor \frac{N_t}{n} \rfloor} \rho(1-\rho)^j \lambda(T_{N_t-nj-k+1})$, and the sum of the failure intensity function of the $n - b_{N_t}$ other components is $\frac{n-b_{N_t}}{n}\lambda(t) - \frac{1}{n}\sum_{k=b_{N_t}+1}^{n}\sum_{j=0}^{\lfloor \frac{N_t}{n} \rfloor - 1} \rho(1-\rho)^j \lambda(T_{N_t-nj-k+1})$. As such, the failure intensity function of the system is given by

$$\lambda_s^a(t) = \lambda(t) - \frac{1}{n}\sum_{k=1}^{b_{N_t}}\sum_{j=0}^{\lfloor \frac{N_t}{n} \rfloor} \rho(1-\rho)^j \lambda(T_{N_t-nj-k+1}) - \frac{1}{n}\sum_{k=b_{N_t}+1}^{n}\sum_{j=0}^{\lfloor \frac{N_t}{n} \rfloor - 1} \rho(1-\rho)^j \lambda(T_{N_t-nj-k+1}). \tag{16}$$

- Similar to the case of $n < N_t \leq mn$, if $N_t > mn$, then we have
  - If $N_t = n\lfloor \frac{N_t}{n} \rfloor$, then the failure intensity function of the system is given by

$$\lambda_s^a(t) = \lambda(t) - \frac{1}{n}\sum_{k=1}^{n}\sum_{j=0}^{m-1} \rho(1-\rho)^j \lambda(T_{N_t-nj-k+1}), \tag{17}$$

  - If $N_t \neq n\lfloor \frac{N_t}{n} \rfloor$, the failure intensity function of the system is given by

$$\lambda_s^a(t) = \lambda(t) - \frac{1}{n}\sum_{k=1}^{b_{N_t}}\sum_{j=0}^{m-1} \rho(1-\rho)^j \lambda(T_{N_t-nj-k+1}) - \frac{1}{n}\sum_{k=b_{N_t}+1}^{n}\sum_{j=0}^{m-1} \rho(1-\rho)^j \lambda(T_{N_t-nj-k+1}) \tag{18}$$

(ii) **The case of the ARA model.** If we assume that the failure process of each component follows ARA_m shown in Eq. (5), we have the following analyses.

- If $N_t = 0$, the failure intensity of the system is $\lambda(t)$.
- If $0 < N_t \leq n$, there are N_t components whose first failures occur. Each of these components has failure intensity function $\frac{1}{n}\lambda(t - \rho T_k)$ and each of the rest $n - N_t$ components has failure intensity $\frac{1}{n}\lambda(t)$. As such, the failure intensity function of the system after the N_t-th failure is given by

$$\lambda_s^a(t) = \frac{1}{n}\sum_{k=1}^{N_t}\lambda(t - \rho T_k) + \frac{n - N_t}{n}\lambda(t) \tag{19}$$

- If $N_t > n$, a similar discussion as the ARI case can be made.

To sum up, we obtain the following definition, i.e., Definition 1.

Definition 1. *A new $SARI_{n,m}$ model, denoted as $SARI^a_{n,m}$, and a new $SARA_{n,m}$, denoted as $SARA^a_{n,m}$ are defined, respectively, in the following. A $SARI^a_{n,m}$ is defined by*

$$\lambda^a_s(t) = \begin{cases} \lambda(t), & \text{if } N_t < 1; \\ \lambda(t) - \frac{1}{n}\sum_{k=1}^{N_t} \rho\lambda(T_k), & \text{if } 1 \le N_t < n; \\ \lambda(t) - \frac{1}{n}\sum_{k=1}^{b_{N_t}} \rho\lambda(T_{N_t-k+1}) - \frac{1}{n}\sum_{k=1}^{n} \sum_{j=1}^{\min\{\lfloor \frac{N_t}{n} \rfloor -1, m-1\}} \rho(1-\rho)^j \lambda(T_{N_t-n(j-1)-b_{N_t}-k+1}) & \text{if } N_t \ge n \end{cases} \tag{20}$$

and a $SARA^a_{n,m}$ is defined by

$$\lambda^a_s(t) = \begin{cases} \lambda(t), & \text{if } N_t < 1; \\ \frac{1}{n}\sum_{k=1}^{N_t} \lambda\left(t - \rho T_k\right) + \frac{n-N_t}{n}\lambda(t), & \text{if } 1 \le N_t < n; \\ \frac{1}{n}\sum_{k=1}^{b_{N_t}} \lambda\left(t - \rho T_{N_t-k+1}\right) + \frac{1}{n}\sum_{k=1}^{n} \lambda\left(t - \sum_{j=1}^{\min\{\lfloor \frac{N_t}{n} \rfloor -1, m-1\}} \rho(1-\rho)^j T_{N_t-n(j-1)-b_{N_t}-k+1}\right) & \text{if } N_t \ge n \end{cases} \tag{21}$$

On Definition 1, there are the following special cases.

(i) If $n = 1$, then $\text{SARI}^a_{n,m}$ in Eq. (20) and $\text{SARA}^a_{n,m}$ in Eq. (21) reduce to ARI_m and ARA_m in (4) and (5), respectively.

(ii) If $m = 1$ and $N_t > n$, then $\text{SARI}^a_{n,m}$ reduces to $\text{SARI}_{n,1}$,

$$\lambda^a_s(t) = \lambda(t) - \frac{1}{n}\sum_{k=1}^{n} \rho\lambda(T_{N_t-k+1}), \tag{22}$$

and, $\text{SARA}^a_{n,m}$ reduces to $\text{SARA}_{n,1}$,

$$\lambda^a_s(t) = \frac{1}{n}\sum_{k=1}^{n} \lambda\left(t - \rho T_{N_t-k+1}\right). \tag{23}$$

(iii) If $\rho = 0$, then the repair on each component is minimal and both $\text{SARI}^a_{n,m}$ in Eq. (20) and $\text{SARA}^a_{n,m}$ in Eq. (21) reduce the NHPP (non-homogenous Poisson process).

(iv) If $\rho = 1$ and $n = 1$, then

- The failure intensity $\lambda^a_s(t)$ after the N_tth failure in $\text{SARI}_{1,m}$ (see Eq. (20)) is $\lambda(t) - \lambda(T_{N_t})$. At time T_{N_t}, $\lambda(t) - \lambda(T_{N_t}) = 0$ and the system starts from the status with failure intensity 0. But it is important to note that it does not mean that the item is repaired as good as new.
- The failure intensity $\lambda^a_s(t)$ after the N_tth failure in the $\text{SARA}_{1,m}$ model (see Eq. (21)) is $\lambda(t-T_{N_t})$. At time T_{N_t}, $\lambda(t - T_{N_t}) = 0$, which implies that the system is repaired as good as new.

(v) If $\rho = 1$ and $n > 1$, model Eq. (20) and model Eq. (21) reduce to

$$\lambda_s^a(t) = \lambda(t) - \frac{1}{n}\sum_{k=1}^{n} \lambda(T_{N_t-k+1}), \tag{24}$$

and

$$\lambda_s^a(t) = \frac{1}{n}\sum_{k=1}^{n} \lambda\left(t - T_{N_t-k+1}\right), \tag{25}$$

respectively.

The model shown in Eq. (25) is the MAI model, which is a special case of the model shown in Eq. (3).

Remark 1. *The above bullet (v) shows that the MAI model is a special case of the $SARA_{n,m}^a$ model. Numerical data experiments and case studies show that the MAI model has a clear advantage over ten other models on 11 out of 15 real world datasets [4]. As such, the model $SARA_{n,m}^a$ can be regarded as an extension of the MAI model.*

Remark 2. *The existing failure process models can roughly be categorised into three classes, as discussed below.*

- *Models that have one parameter depicting the repair effectiveness of each individual repair. For example, the parameter A_n in the virtual age models $V_n = V_{n-1} + A_n(T_n - T_{n-1})$ and $V_n = A_n(V_{n-1} + T_n - T_{n-1})$ is the parameter depicting the effectiveness of the nth repair [25] (where V_n is the virtual age). Technically, A_n may estimate the repair effectiveness of different components in a system. However, in reality, it is not suitable for modelling the failure process of a multi-component system due to two reasons: on the one hand, the size of masked failure data may not be sufficiently large for estimating a large number of parameters A_n; on the other hand, A_n may be assumed a stochastic process, on which there is little research that has been conducted.*

- *Models that have only one parameter depicting the repair effectiveness for different components in a system. For example, models shown in Eq. (4) and (5) fall in this category. Similarly, if A_n in the virtual age models are set to $A_n = A$ (i.e., A_n are the same over different n's), then the above-mentioned virtual age models have one parameter as well. The shortcoming of such models for modelling the failure process of a multi-component system is discussed in Approach 1 in the above discussion.*

- *Models that approximates the SRP. For example, models shown in Eqs. (1), (2) and (3) fall in this category, which assumes that the repair on each failed component is perfect and is not suitable for a system in which repair on failed component is imperfect.*

Models (20) *and* (21) *are derived for depicting the failure process of a multi-component series system when the failure process of a component follows the ARI model and the ARA model, respectively. They can therefore be used to model the failure process of the pavement system discussed in Section 1, for example.*

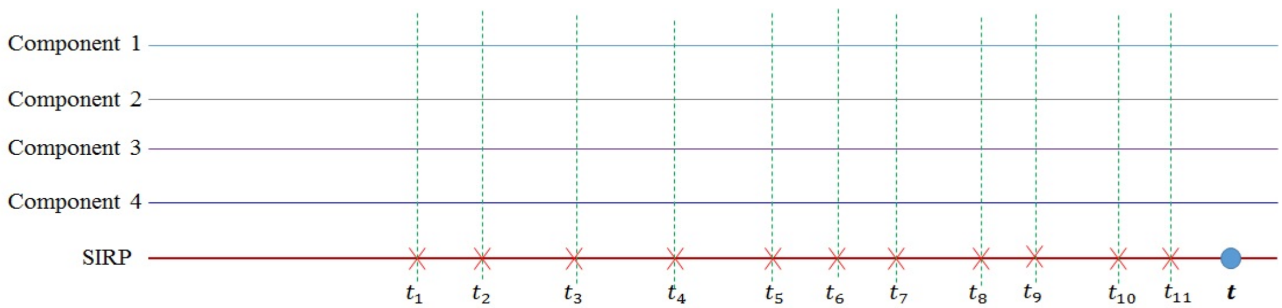

Figure 4: Masked failure data of a system with four components until time t, where $N_{1,t}, N_{2,t}, N_{3,t}$, and $N_{4,t}$ are unknown; $N_t = 11$. where $t_{i,0} = 0$, $j = 1, 2, ..., N_{i,t}$, and $t_j = 0$ if $j \leq 0$.

Example 2. *Suppose that the value k's in $T_{k,j}$ in Example 1 are unknown. But t_j ($j = 1, 2, ..., 11$) are available, as shown in Figure 4. We assume that the four components are identical and that the failure process of each component is ARI_3. The 11 failures are assumed to be caused by three failures of each of three components and two failures of the other component ($3\times3+2 = 11$), respectively, which can be modelled by Model* (20) *in Definition 1. The model, $SARI^a_{4,3}$, is given by*

$$\begin{aligned}\lambda_s^a(t) =& \lambda(t) - \frac{1}{4}\rho(\lambda(t_{11}) + \lambda(t_{10}) + \lambda(t_9)) \\ & - \frac{1}{4}\rho(1-\rho)(\lambda(t_8) + \lambda(t_7) + \lambda(t_6) + \lambda(t_5)) \\ & - \frac{1}{4}\rho(1-\rho)^2(\lambda(t_4) + \lambda(t_3) + \lambda(t_2) + \lambda(t_1)). \end{aligned} \tag{26}$$

In case the causes of the system failures are known, then by plugging $t_{k,i}$ in Figure 3 into Eq. (26), *we obtain*

$$\begin{aligned}\lambda_s^a(t) =& \lambda(t) - \frac{1}{4}\rho(\lambda(t_{1,3}) + \lambda(t_{4,5}) + \lambda(t_{4,4})) \\ & - \frac{1}{4}\rho(1-\rho)(\lambda(t_{2,2}) + \lambda(t_{4,3}) + \lambda(t_{4,2}) + \lambda(t_{1,2})) \\ & - \frac{1}{4}\rho(1-\rho)^2(\lambda(t_{3,1}) + \lambda(t_{2,1}) + \lambda(t_{4,1}) + \lambda(t_{1,1})). \end{aligned} \tag{27}$$

As can be seen, Model (27) differs from Model (6).

In the following, we compare $\lambda_s(t)$ in Eq. (7) with $\lambda_s^a(t)$ in Eq. (20), and $\lambda_s(t)$ in (8) with $\lambda_s^a(t)$ in (21). To do it, we need to introduce an important definition on stochastic ordering.

Definition 2. *Stochastic order (p. 404 in [26]). Assume that X and Y are two random variables. If for every real number r, the inequality*

$$P(X \geq r) \geq P(Y \geq r)$$

holds, then X is stochastically greater than or equal to Y, or $X \geq_{st} Y$. Equivalently, Y is stochastically less than or equal to X, or $Y \leq_{st} X$, or, $E(X) \geq E(Y)$.

Suppose a system is composed of n identical components, each of which follows the same ARI$_m$, then for $N_t > n$ we have

$$\begin{aligned}\lambda_s(t) =&\lambda(t) - \frac{1}{n}\sum_{k=1}^{n}\sum_{j=0}^{\min\{m-1,N_{k,t}-1\}}\rho(1-\rho)^j\lambda(T_{k,N_{k,t}-j})\\ =&\lambda(t) - \frac{1}{n}\sum_{k=1}^{b_{N_t}}\rho\lambda(T_{N_t-k+1}) - \frac{1}{n}\sum_{k=1}^{n}\sum_{j=1}^{\min\{\lfloor\frac{N_t}{n}\rfloor-1,m-1\}}\rho(1-\rho)^j\lambda(T_{N_t-n(j-1)-b_{N_t}-k+1}) + \epsilon_t, \qquad (28)\end{aligned}$$

where

$$\begin{aligned}\epsilon_t =& -\frac{1}{n}\sum_{k=1}^{n}\sum_{j=0}^{\min\{m-1,N_{k,t}-1\}}\rho(1-\rho)^j\lambda(T_{k,N_{k,t}-j})\\ &+\frac{1}{n}\sum_{k=1}^{b_{N_t}}\rho\lambda(T_{N_t-k+1}) + \frac{1}{n}\sum_{k=1}^{n}\sum_{j=1}^{\min\{\lfloor\frac{N_t}{n}\rfloor-1,m-1\}}\rho(1-\rho)^j\lambda(T_{N_t-n(j-1)-b_{N_t}-k+1}).\end{aligned}$$

Lemma 2. *The expectation of ϵ_t has the following bounds:*

$$-\frac{1-\rho^m}{1-\rho}E(\lambda(T_1)) \leq E(\epsilon_t) \leq \rho E(\lambda(T_1)) + \frac{1-\rho^m}{1-\rho}E(\lambda(T_1)) \qquad (29)$$

The proof of Lemma 2 can be found in Appendix.

ϵ_t measures the difference between $\lambda_s(t)$ and $\lambda_s^a(t)$. It should be noted: ϵ_t has a practical implication if the values of ρ's in Eq. (11) and in Eq. (20) are the same and the $\lambda(t)$ in model (20), which is obtained from the masked failure data, and the $\lambda(t)$ in model (11), which is obtained from the unmasked failure data are the same.

5. Parameter estimation and numerical examples

In this section, we derive the maximum likelihood functions for the models proposed in Section 4 and then apply them on a real dataset.

Given a series of successive failure times $t_1, t_2, \cdots, t_N$, on which the system failed; where N is the number of failures and $N > n$. That is, the available data are up to the time at which the last failure occurs.

5.1. Maximum likelihood functions

Below we consider the likelihood for the failure process of a single system before a specified number of failures is occurred. If several independent processes are observed, the log-likelihood can be easily obtained based on the likelihood functions given in this section.

Below we will give the likelihood function of the $\text{SARI}^a_{n,m}$ model and the $\text{SARA}_{n,m}$ models in Section 3. The derivation of the likelihood functions follows from Andersen et al. (1993, sec. II.7) that under our stated conditions, the likelihood function for the observations from a single system is derived below.

Following the definition of b_{N_t}, we define $b_i = i - n\lfloor\frac{i}{n}\rfloor$ if $i \neq n\lfloor\frac{i}{n}\rfloor$, where i is a positive integer. We also define $\mathbb{N}_1 = \{i | 1 \leq i \leq n-1\}$, $\mathbb{N}_2 = \{i | n < i \leq N-1, i \neq \nu n, 1 \leq \nu < \lfloor\frac{N}{n}\rfloor\}$, and $\mathbb{N}_3 = \{i | i = \nu n, \nu < \lfloor\frac{N}{n}\rfloor\}$.

Hence, given a dataset of N successive failure times $t_1, \cdots, t_N$, the likelihood function for the $\text{SARI}^a_{n,m}$ model is

$$
\begin{aligned}
L_{\text{SARI}}(\mathbf{\Theta}) =& \lambda(t_1)\exp(-\Lambda(t_1)) \prod_{i\in\mathbb{N}_1}\left[\left(\lambda(t_{i+1}) - \frac{1}{n}\sum_{k=1}^{i}\rho\lambda(t_k)\right)\exp\left(-\Lambda(t_{i+1}) + \Lambda(t_i) + \frac{(t_{i+1}-t_i)}{n}\sum_{k=1}^{i}\rho\lambda(t_k)\right)\right] \\
&\times \prod_{i\in\mathbb{N}_2}\left[\left(\lambda(t_{i+1}) - \frac{1}{n}\sum_{k=1}^{b_i}\sum_{j=0}^{\min\{\lfloor\frac{i}{n}\rfloor, m-1\}}\rho(1-\rho)^j\lambda(t_{i-nj-k+1}) - \frac{1}{n}\sum_{k=b_i+1}^{n}\sum_{j=0}^{\min\{\lfloor\frac{i}{n}\rfloor-1, m-1\}}\rho(1-\rho)^j\lambda(t_{i-nj-k+1})\right)\right. \\
&\quad \times \exp\left(-\Lambda(t_{i+1}) + \frac{(t_{i+1}-t_i)}{n}\sum_{k=1}^{b_i}\sum_{j=0}^{\min\{\lfloor\frac{i}{n}\rfloor, m-1\}}\rho(1-\rho)^j\lambda(t_{i-nj-k+1})\right. \\
&\quad\quad \left.\left. + \Lambda(t_i) + \frac{(t_{i+1}-t_i)}{n}\sum_{k=b_i+1}^{n}\sum_{j=0}^{\min\{\lfloor\frac{i}{n}\rfloor-1, m-1\}}\rho(1-\rho)^j\lambda(t_{i-nj-k+1})\right)\right] \\
&\times \prod_{i\in\mathbb{N}_3}\left[\left(\lambda(t_{i+1}) - \frac{1}{n}\sum_{k=1}^{n}\sum_{j=0}^{\min\{\lfloor\frac{i}{n}\rfloor-1, m-1\}}\rho(1-\rho)^j\lambda(t_{i-nj-k+1})\right)\right. \\
&\quad \left.\times \exp\left(-\Lambda(t_{i+1}) + \Lambda(t_i) + \frac{(t_{i+1}-t_i)}{n}\sum_{k=1}^{n}\sum_{j=0}^{\min\{\lfloor\frac{i}{n}\rfloor-1, m-1\}}\rho(1-\rho)^j\lambda(t_{i-nj-k+1})\right)\right] \qquad (30)
\end{aligned}
$$

and the likelihood of the $\text{SARA}^a_{n,m}$ model is given by

$$
\begin{aligned}
&L_{\text{SARA}}(\mathbf{\Theta}) = \lambda(t_1)\exp(-\Lambda(t_1)) \\
&\times \prod_{i\in\mathbb{N}_1}\left\{\left[\frac{n-i}{n}\lambda(t_{i+1}) + \frac{1}{n}\sum_{k=1}^{i}\lambda\left(t_{i+1}-\rho t_k\right)\right]\exp\left(-\frac{1}{n}\sum_{k=1}^{i}\Lambda\left(t_{i+1}-\rho t_k\right) + \frac{1}{n}\sum_{k=1}^{i}\Lambda(t_i-\rho t_k) - \frac{n-i}{n}(\Lambda(t_{i+1})-\Lambda(t_i))\right)\right\} \\
&\times \prod_{i\in\mathbb{N}_2}\left\{\left[\frac{1}{n}\sum_{k=1}^{b_i}\lambda\left(t_{i+1} - \sum_{j=0}^{\min\{\lfloor\frac{i}{n}\rfloor, m-1\}}\rho(1-\rho)^j t_{i-nj-k+1}\right) + \frac{1}{n}\sum_{k=b_i+1}^{n}\lambda\left(t_{i+1} - \sum_{j=0}^{\min\{\lfloor\frac{i}{n}\rfloor-1, m-1\}}\rho(1-\rho)^j t_{i-nj-k+1}\right)\right]\right. \\
&\quad \times \exp\left[-\frac{1}{n}\sum_{k=1}^{b_i}\Lambda\left(t_{i+1} - \sum_{j=0}^{\min\{\lfloor\frac{i}{n}\rfloor, m-1\}}\rho(1-\rho)^j t_{i-nj-k+1}\right) - \frac{1}{n}\sum_{k=b_i+1}^{n}\Lambda\left(t_{i+1} - \sum_{j=0}^{\min\{\lfloor\frac{i}{n}\rfloor-1, m-1\}}\rho(1-\rho)^j t_{i-nj-k+1}\right)\right. \\
&\quad\quad \left.\left. + \frac{1}{n}\sum_{k=1}^{b_i}\Lambda\left(t_i - \sum_{j=0}^{\min\{\lfloor\frac{i}{n}\rfloor, m-1\}}\rho(1-\rho)^j t_{i-nj-k+1}\right) + \frac{1}{n}\sum_{k=b_i+1}^{n}\Lambda\left(t_i - \sum_{j=0}^{\min\{\lfloor\frac{i}{n}\rfloor-1, m-1\}}\rho(1-\rho)^j t_{i-nj-k+1}\right)\right]\right\} \\
&\times \prod_{i\in\mathbb{N}_3}\left\{\frac{1}{n}\sum_{k=1}^{n}\lambda\left(t_{i+1} - \sum_{j=0}^{\min\{\lfloor\frac{i}{n}\rfloor-1, m-1\}}\rho(1-\rho)^j t_{i-nj-k+1}\right)\right. \\
&\quad \left.\times \exp\left[-\frac{1}{n}\sum_{k=1}^{n}\Lambda\left(t_{i+1} - \sum_{j=0}^{\min\{\lfloor\frac{i}{n}\rfloor-1, m-1\}}\rho(1-\rho)^j t_{i-nj-k+1}\right) + \frac{1}{n}\sum_{k=1}^{n}\Lambda\left(t_i - \sum_{j=0}^{\min\{\lfloor\frac{i}{n}\rfloor-1, m-1\}}\rho(1-\rho)^j t_{i-nj-k+1}\right)\right]\right\},
\end{aligned}
\qquad (31)
$$

where $t_0 = 0$, $\lambda(t) = 0$, and $\Lambda(t) = \int_0^t \lambda(u)du$.

By maximising $\log(L_{\text{SARI}})$ and $\log(L_{\text{SARA}})$, one can find optimal parameters in $\lambda(t)$ and $\hat{\rho}$, respectively.

5.2. Data examples

In the following, we compare the performance of models NHPP, ARI_m, ARA_m, $\mathrm{SARI}^a_{n,m}$, and $\mathrm{SARA}^a_{n,m}$. We use criteria such AIC (Akaike information criterion), $\mathrm{AIC_c}$ (AIC with a correction), and BIC (Bayesian information criterion) to compare the performance. Those criteria are: $\mathrm{AIC} = -2\log(L) + 2q$, $\mathrm{AIC}_c = -2\log(L) + 2q + \dfrac{2(q+2)(q+3)}{N-q-2}$, and $\mathrm{BIC} = -2\log(L) + q\log(N)$, where L is the maximized value of the likelihood for the model, q is the number of parameters in the model, and N is the total number of failures (observations). The term $2q$, $\dfrac{2(q+2)(q+3)}{N-q-2}$, and $q\log(N)$ in the AIC, $\mathrm{AIC_c}$ and BIC penalise a model with a large number of parameters, respectively. The reader is referred to [27] for details on model performance measures. [28] provides a practical procedure for the selection of time-to-failure models based on the assessment of trends in maintenance data.

We compare the performance of the proposed models $\mathrm{SARI}^a_{n,m}$ and $\mathrm{SARA}^a_{n,m}$ on artificially generated data, which are generated based on the simulation method shown in Section 3.3. We set failure intensity function $\lambda(t) = 0.002869t^{1.5}$ (which has $\Lambda(t) = (\frac{t}{15})^{2.5}$), $n = 3$, $m = 2$, and $\rho = 0.5$, that is, each of the three components has a failure process ARI_2. We compare the models: NHPP, ESI, MAI, ARI, ARA, $\mathrm{SARI}^a_{n,m}$, and $\mathrm{SARA}^a_{n,m}$ on the dataset. Table 2 shows that the $\mathrm{SARI}_{3,7}$ outperforms other models in terms of the $-\log$(likelihood).

Table 2: Model comparison on artificially generated data.

		NHPP	ESI	MAI	ARI	ARA	SARI	SARA
SARI Data	$-$log-likelihood	85.78	85.14	85.13	84.10	84.14	84.09	84.12
$(\alpha = 15, \beta = 2.5$	BIC	179.38	182.01	178.09	179.94	180.01	179.92	179.97
$n = 3, N = 50)$	AIC_c	176.08	177.16	174.79	175.09	175.16	175.07	175.12

We also compare the performance of the proposed models $\mathrm{SARI}^a_{n,m}$ and $\mathrm{SARA}^a_{n,m}$ on the Bus514 dataset shown in [29]. On this dataset, we know neither the number of components nor whether the components are identical. We compare the models: NHPP, ESI, MAI, ARI, ARA, $\mathrm{SARI}^a_{n,m}$, and $\mathrm{SARA}^a_{n,m}$ on the dataset. Table 3 shows that the $\mathrm{SARI}^a_{3,1}$ outperforms other models.

Table 3: Model comparison on the Bus514 dataset.

		NHPP	ESI	MAI	ARI	ARA	SARI	SARA
	-log-likelihood	532.74	530.82	533.18	530.84	532.74	<u>530.38</u>	532.74
Bus514 Data	BIC	1073.46	1073.61	1074.34	1073.65	1077.45	<u>1072.73</u>	1077.45
	AIC_c	1069.96	1068.46	1070.84	1068.50	1072.30	<u>1067.58</u>	1072.30

The above two examples show that $\mathrm{SARI}^a_{n,m}$ results in the smallest $-\log$(likelihood values), but $\mathrm{SARA}^a_{n,m}$ does not perform so well as the $\mathrm{SARI}^a_{n,m}$ model. Nevertheless, since the MAI model shows its outstanding performance and $\text{SARA}^a_{n,m}$ is an extension of MAI, one may set $\rho = 1$ in $\text{SARI}^a_{n,m}$ in case $\text{SARA}^a_{n,m}$ shows poor performance on a dataset. To gain a better view on the comparison of the performance of the models, Figure 5 shows their values of BIC and AIC_c.

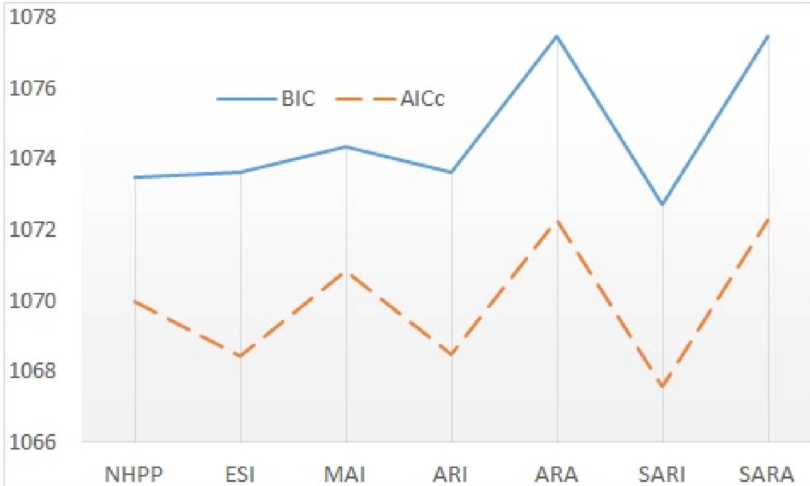

Figure 5: Comparison of BIC and AIC_c.

6. Discussion

6.1. An approximation method for left-truncated masked failure data

Section 4 discusses the scenario where a full history of masked failure data can be collected. Now we consider the case that M failure observations of the earliest occurrences are not available, that is, $T_1, ..., T_M$ are not available, but $T_{M+1}, T_{M+2}, \ldots, T_{N_t}$ are available. Such data are *masked left-truncated failure data*. One can the models $\text{SARI}^a_{n,m}$ and $\text{SARA}^a_{n,m}$ in Section 4 to fit the data, which assumes that the first n failures are due to the n components, respectively. An alternative method is to simply take $T_{N_t}, T_{N_t-1}, ..., T_{N_t-n+1}$ as the last failure times of the n components, and take $T_{N_t-n}, T_{N_t-1}, ..., T_{N_t-2n+1}$ as the 2nd last failure times of the n components, and so on. Under such an assumption, we propose the following models.

Definition 3. *A $SARI^a_{n,m}$ is defined by*

$$\lambda^a_s(t) = \lambda(t) - \frac{\Phi_{N'_t}}{n}\sum_{k=1}^{n}\sum_{j=0}^{\min\{\lfloor \frac{N'_t}{n}\rfloor -1, m-1\}} \rho(1-\rho)^j \lambda(T_{N'_t-nj-k+1}) - \frac{\Psi_{N'_t}}{n}\sum_{k=1}^{r_t}\rho(1-\rho)^{\lfloor \frac{N'_t}{n}\rfloor}\lambda(T_{N'_t-n\lfloor \frac{N'_t}{n}\rfloor-k+1}), \quad (32)$$

and a $SARA^a_{n,m}$ is defined by

$$\lambda^a_s(t) = \frac{\Phi_{N'_t}}{n}\sum_{k=1}^{n}\lambda\left(t - \sum_{j=0}^{\min\{\lfloor \frac{N'_t}{n}\rfloor -1, m-1\}} \rho(1-\rho)^j T_{N'_t-nj-k+1}\right) + \frac{\Psi_{N'_t}}{n}\sum_{k=1}^{n}\lambda\left(t - \sum_{k=1}^{r_t}\rho(1-\rho)^{\lfloor \frac{N'_t}{n}\rfloor}T_{N'_t-n\lfloor \frac{N'_t}{n}\rfloor-k+1}\right), \quad (33)$$

where $m \geq 1$, $N'_t = N_t - M + 1$, $\Phi_{N'_t} = \chi\{N'_t \geq n\}$, $\Psi_{N'_t} = \chi\{\lfloor \frac{N'_t}{n} \rfloor < m \cap b_{N'_t} \neq 0\}$, $r_t = N'_t - n\lfloor \frac{N'_t}{n} \rfloor$ if $N'_t \neq n\lfloor \frac{N'_t}{n} \rfloor$, and $r_t = N'_t$ otherwise.

Definition 3 differs from Definition 1, as shown in Figure 6 and Example 3

- Fig. 6 shows the difference between the two definitions. The notes above the SIRP line shows how Definition 1 defines a cycle, which is a set of T_k with the same power of $(1-\rho)$ in a $\text{SARI}^a_{n,m}$ model or in a $\text{SARA}^a_{n,m}$ model, and the notes under the SIRP line shows how Definition 3 defines a cycle.
- Eq. (34) in Example 3 shows the $\text{SARI}^a_{4,3}$ by Definition 3, in which $\lambda(t_8)$ has a coefficient $\frac{1}{4}\rho$ and $\lambda(t_5)$ has a coefficient $\frac{1}{4}\rho(1-\rho)$ whereas in Example 2, $\lambda(t_8)$ has a coefficient $\frac{1}{4}\rho(1-\rho)$ and $\lambda(t_5)$ has a coefficient $\frac{1}{4}\rho$.

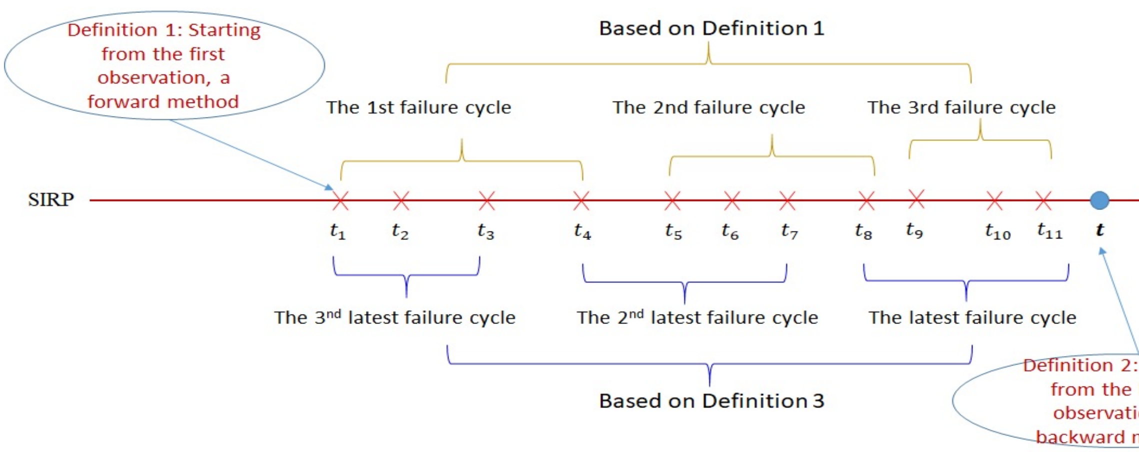

Figure 6: Comparison of definitions of different cycles of Definition 1 and Definition 2.

Example 3. *Further to Example 2, a model $SARI^a_{4,3}$, defined by Definition 3, and is given by*

$$\begin{aligned}\lambda^a_s(t) =& \lambda(t) - \frac{1}{4}\rho(\lambda(t_{11}) + \lambda(t_{10}) + \lambda(t_9) + \lambda(t_8)) \\ & - \frac{1}{4}\rho(1-\rho)(\lambda(t_7) + \lambda(t_6) + \lambda(t_5) + \lambda(t_4)) \\ & - \frac{1}{4}\rho(1-\rho)^2(\lambda(t_3) + \lambda(t_2) + \lambda(t_1)).\end{aligned} \tag{34}$$

6.2. Failure intensity function of the SRP

Although the SRP has been well studied (see [3] for more detailed discussion), to the author's best knowledge, its failure intensity function has not been given in the existing literature and is given in Lemma 3.

Lemma 3. *Given a series system on which a failed component is replaced with an identical new component immediately, the failure intensity function of the system after the N_t-th replacement is given by*

$$\lambda(t|\mathscr{H}_{t-}) = \begin{cases} \frac{1}{n}\sum_{k=1}^{n}\lambda_k(t), & \text{if } N_t = 0, \\ \frac{1}{n}\sum_{k=1}^{n-1}\lambda_{i_k}(t - T_{i_k, N_t - j_k}) + \frac{1}{n}\lambda_{i_n}(t - T_{i_n, N_t}), & \text{if } N_t \geq 1. \end{cases} \tag{35}$$

where $j_k \in \{1, 2, \ldots, N_t\}$, $i_k \in \{1, 2, \ldots, n\}$, $T_0 = 0$, $i_{k_1} \neq i_{k_2}$ for $k_1 \neq k_2$, $j_{k_1} \neq j_{k_2}$ if $k_1 \neq k_2$ and $j_{k_1} j_{k_2} > 0$.

The proof of Lemma 3 can be found in Appendix.

7. Conclusions and further work

In the real world, systems are normally composed of multiple components and the failure data may be masked due to insufficient failure cause data or such data are unattainable because of physical constraints or lack of resources. This needs to develop a method to model the superposition of a number of imperfect failure processes. However, since the failure data are masked, the components that cause the system to fail are unknown. This needs to develop methods to approximate the failure process of the system.

While the superposition of renewal processes has been extensively studied, the superposition of imperfect failure processes (SIRP) has received little attention in the literature. There is a need to conduct research on SIRP, which is the focus of this paper.

The main contributions of this paper include the following.

- This paper developed two methods: one for untruncated masked failure data and one for left-truncated data, to approximate the superposition of the imperfect failure processes of the components in a series system in which the failure process of each component follows two widely used models. The imperfect failure process models are the arithmetic reduction of intensity (ARI) model or the arithmetic reduction of age (ARA) model, respectively.
- The paper showed that unlike the superposition of renewal processes, the superposition of the ARI processes (SARI) (or the ARA processes (SARA)) does not tend toward (statistical) equilibrium as the time of operation becomes very large;
- The MAI (Moving Average of Intensity) model proposed in [4] is a special case of the superposition of the ARA processes; and
- It developed a method to simulate the SARI process and the SARA process, respectively, and gave likelihood functions of the SARI model and the SARA model (for the untruncated data), respectively.

Our future work will be focused on the derivation of statistical properties of the models.

Acknowledgements

The author is indebted to the reviewers and the editor for their helpful comments.

References

[1] L. Doyen, O. Gaudoin, Classes of imperfect repair models based on reduction of failure intensity or virtual age, Reliability Engineering & System Safety 84 (1) (2004) 45–56.

[2] S. Wu, P. Scarf, Two new stochastic models of the failure process of a series system, European Journal of Operational Research 257 (3) (2017) 763–772.

[3] S. Wu, Superimposed renewal processes in reliability, in: F. Ruggeri (Ed.), Wiley StatsRef: Statistics Reference Online, Wiley, 2019, pp. 1–12. doi:10.1002/9781118445112.stat08228.

[4] S. Wu, A failure process model with the exponential smoothing of intensity functions, European Journal of Operational Research 275 (2) (2019) 502–513.

[5] Z.-S. Ye, M. Xie, L.-C. Tang, Reliability evaluation of hard disk drive failures based on counting processes, Reliability Engineering & System Safety 109 (2013) 110–118.

[6] M. L. Gámiz, B. H. Lindqvist, Nonparametric estimation in trend-renewal processes, Reliability Engineering & System Safety 145 (2016) 38–46.

[7] B. Liu, S. Wu, M. Xie, W. Kuo, A condition-based maintenance policy for degrading systems with age-and state-dependent operating cost, European Journal of Operational Research 263 (3) (2017) 879–887.

[8] L. Yang, Y. Zhao, R. Peng, X. Ma, Hybrid preventive maintenance of competing failures under random environment, Reliability Engineering & System Safety 174 (2018) 130–140.

[9] L. Yang, Z.-s. Ye, C.-G. Lee, S.-f. Yang, R. Peng, A two-phase preventive maintenance policy considering imperfect repair and postponed replacement, European Journal of Operational Research 274 (3) (2019) 966–977.

[10] M. Compare, P. Baraldi, I. Bani, E. Zio, D. McDonnell, Industrial equipment reliability estimation: a Bayesian Weibull regression model with covariate selection, Reliability Engineering & System Safety (2020) 106891.

[11] L. Brenière, L. Doyen, C. Bérenguer, Virtual age models with time-dependent covariates: A framework for simulation, parametric inference and quality of estimation, Reliability Engineering & System Safety (2020) 107054.

[12] B. de Jonge, P. A. Scarf, A review on maintenance optimization, European Journal of Operational Research (2019).

[13] P. Kuehn, Approximate analysis of general queuing networks by decomposition, IEEE Transactions on communications 27 (1) (1979) 113–126.

[14] W. Whitt, Approximating a point process by a renewal process, i: Two basic methods, Operations Research 30 (1) (1982) 125–147.

[15] X. Liu, Y. Dijoux, J. Vatn, H. Toftaker, Performance of prognosis indicators for superimposed renewal processes, Probability in the Engineering and Informational Sciences (2020) 1–24.

[16] W. Zhang, Y. Tian, L. A. Escobar, W. Q. Meeker, Estimating a parametric component lifetime distribution from a collection of superimposed renewal processes, Technometrics 59 (2) (2017) 202–214.

[17] J.-K. Chan, L. Shaw, Modeling repairable systems with failure rates that depend on age and maintenance, IEEE Transactions on Reliability 42 (4) (1993) 566–571.

[18] M. A. K. Malik, Reliable preventive maintenance scheduling, AIIE transactions 11 (3) (1979) 221–228.

[19] M. Kijima, H. Morimura, Y. Suzuki, Periodical replacement problem without assuming minimal repair, European Journal of Operational Research 37 (2) (1988) 194–203.

[20] H. Ascher, H. Feingold, Repairable systems reliability: modeling, inference, misconceptions and their causes, M. Dekker New York, 1984.

[21] G. Pulcini, Modeling the failure data of a repairable equipment with bathtub type failure intensity, Reliability Engineering & System Safety 71 (2) (2001) 209–218.

[22] B. M. Mun, S. J. Bae, P. H. Kvam, A superposed log-linear failure intensity model for repairable artillery systems, Journal of Quality Technology 45 (1) (2013) 100–115.

[23] L. Doyen, Asymptotic properties of imperfect repair models and estimation of repair efficiency, Naval Research Logistics (NRL) 57 (3) (2010) 296–307.

[24] R. Drenick, The failure law of complex equipment, Journal of the Society for Industrial and Applied Mathematics 8 (4) (1960) 680–690.

[25] M. Kijima, Some results for repairable systems with general repair, Journal of Applied probability 26 (1) (1989) 89–102.

[26] S. M. Ross, Stochastic processes, 2nd Edition, John Wiley & Sons New York, 1996.

[27] K. P. Burnham, D. Anderson, Model selection and multi-model inference: a practical information-theoretic approach, Springer Science and Business Media, 2002.

[28] D. M. Louit, R. Pascual, A. K. Jardine, A practical procedure for the selection of time-to-failure models based on the assessment of trends in maintenance data, Reliability Engineering & System Safety 94 (10) (2009) 1618–1628.

[29] M. Guida, G. Pulcini, Reliability analysis of mechanical systems with bounded and bathtub shaped intensity function, IEEE Transactions on Reliability 58 (3) (2009) 432–443.

Appendix

Below gives the proof of Lemma 2.

Proof. Apparently, $\epsilon_t \geq -\frac{1}{n}\sum_{k=1}^{n}\sum_{j=0}^{\min\{m-1,N_{k,t}-1\}}\rho(1-\rho)^j\lambda(T_{k,N_{k,t}-j}) \geq -\frac{1}{n}\sum_{k=1}^{n}\sum_{j=0}^{m-1}\rho(1-\rho)^j\lambda(T_{k,N_{k,t}-j})$ and $\epsilon_t \leq \frac{1}{n}\sum_{k=1}^{b_{N_t}}\rho\lambda(T_{N_t-k+1}) + \frac{1}{n}\sum_{k=1}^{n}\sum_{j=1}^{\min\{\lfloor\frac{N_t}{n}\rfloor-1,m-1\}}\rho(1-\rho)^j\lambda(T_{N_t-n(j-1)-b_{N_t}-k+1}) \leq \frac{1}{n}\sum_{k=1}^{n}\rho\lambda(T_{N_t-k+1}) + \frac{1}{n}\sum_{k=1}^{n}\sum_{j=0}^{m-1}\rho(1-\rho)^j\lambda(T_{N_t-n(j-1)-b_{N_t}-k+1})$. Note that $T_{k,N_{k,t}-j} \geq_{\mathrm{st}} T_1$, $T_{N_t-k+1} \geq_{\mathrm{st}} T_1$, and $T_{N_t-n(j-1)-b_{N_t}-k+1} \geq_{\mathrm{st}} T_1$. According to Definition 2, we have $E(T_{k,N_{k,t}-j}) \geq E(T_1)$, $E(T_{N_t-k+1}) \geq E(T_1)$, and $E(T_{N_t-n(j-1)-b_{N_t}-k+1}) \geq E(T_1)$. We can therefore easily obtain that the expectation of ϵ_t has the following bounds:

$$-\frac{1-\rho^m}{1-\rho}E(\lambda(T_1)) \leq E(\epsilon_t) \leq \rho E(\lambda(T_1)) + \frac{1-\rho^m}{1-\rho}E(\lambda(T_1)) \tag{36}$$

■

Below gives the proof of Lemma 3.

Proof. The condition $j_{k_1} \neq j_{k_2}$ if $k_1 \neq k_2$ and $j_{k_1}j_{k_2} > 0$ implies that (1) there is one component renewed, which has failure rate function $\lambda_{i_n}(t - T_{i_n,N_t})$; and (2) within the others, some may have not renewed since installation time $t = 0$ and have the same age, and the others may have failed and then renewed at different failure times.

Before the first failure, none of the components is replaced. Hence, the failure intensity of the system is $\frac{1}{n}\sum_{k=1}^{n}\lambda_k(t)$. During the period from the first replacement until the time when the last component installed at time $t = 0$ is replaced.

$\blacksquare$