

Kent Academic Repository

Full text document (pdf)

Citation for published version

Wu, Shaomin (2019) A failure process model with the exponential smoothing of intensity functions. *European Journal of Operational Research*, 275 (2). pp. 502-513. ISSN 0377-2217.

DOI

<https://doi.org/10.1016/j.ejor.2018.11.045>

Link to record in KAR

<https://kar.kent.ac.uk/70299/>

Document Version

Author's Accepted Manuscript

Copyright & reuse

Content in the Kent Academic Repository is made available for research purposes. Unless otherwise stated all content is protected by copyright and in the absence of an open licence (eg Creative Commons), permissions for further reuse of content should be sought from the publisher, author or other copyright holder.

Versions of research

The version in the Kent Academic Repository may differ from the final published version.

Users are advised to check <http://kar.kent.ac.uk> for the status of the paper. **Users should always cite the published version of record.**

Enquiries

For any further enquiries regarding the licence status of this document, please contact:

researchsupport@kent.ac.uk

If you believe this document infringes copyright then please contact the KAR admin team with the take-down information provided at <http://kar.kent.ac.uk/contact.html>

A failure process model with the exponential smoothing of intensity functions

Shaomin Wu*

Kent Business School, University of Kent, Canterbury, Kent CT2 7FS, United Kingdom

Abstract

This paper proposes a new model and investigates its special case model, both of which model the failure process of a series system composed of multiple components. We make the following assumption: (1) once the system fails, the failed component can be immediately identified and replaced with a new identical one, and (2) once the system fails, only the time of the failure is recorded; but the component that causes the system to fail is not known. The paper derives a parameter estimation method and compares the performance of the proposed models with nine other models on artificially generated data and fifteen real-world datasets. The results show that the two new models outperform the nine models in terms of the three most commonly used penalised model selection criteria, the Akaike's information criterion (AIC), corrected Akaike's information criterion (AICc) and Bayesian information criterion (BIC), respectively.

Keywords: Maintenance, non-homogeneous Poisson process, superimposed renewal process, Akaike information criterion, higher order Markov process.

1 Introduction

Modelling the failure process of a repairable system is an important research topic since it is needed in various applications such as maintenance policy optimisation and lifecycle costing (see, Huang, Huang, and Ho (2017); Olde Keizer, Teunter, and Veldman (2017); Cha, Finkelstein, and Levitin (2018); Yang, Ye, Lee, Yang, and Peng (2018), for example).

In the last three decades, researchers have developed many failure process models, which include the geometric process (GP) (Lam, 1988), the virtual age models (Kijima, 1989), the arithmetic

*Email: s.m.wu@kent.ac.uk. Telephone: 0044 1227 827 940.

reduction of intensity (ARI) models and the arithmetic reduction of age (ARA) models (Doyen & Gaudoin, 2004), and many other models. The reader is referred to Wu and Scarf (2017) for a classification of the existing failure process models. In addition, Doyen, Gaudoin, and Syamsundar (2017) study the possibility of building imperfect maintenance models with a geometric reduction of age or geometric reduction of Intensity. Wu (2018) proposes a new model, the doubly geometric process (GDP), which overcomes two limitations of the GP: the GDP can model stochastically non-monotonic inter-arrival times of recurrent event processes and recurrent event processes where the inter-arrival time distributions have varying shape parameters.

Most of the existing models depict the effectiveness of repair either with a fixed parameter (for example, the ratio of the GP (Lam, 1988) or the repair efficiency parameter ρ in the ARI models and the ARA models (Doyen & Gaudoin, 2004)), which is referred to Model type A, or with the number of many parameters (for example, the degree of the n th repair, A_n , in the virtual age models (Kijima, 1989) or the parameters in each renewal process in the super-imposed renewal process (SRP) (p264, Hoyland and Rausand (2004))), which is referred to Model type B. Model type A may be useful for the one-component system scenarios where the effectiveness of repair on failures may be the same and Model type B is useful for the scenarios with many failure data, which can be used to estimate all the parameters in the model. However, neither of them may be suitable for modelling the failure process of a multicomponent system for two reasons: (1) the effectiveness of repair on failures of different components is normally different and it may therefore be inadequate to use one parameter to depict those repair effectiveness, (2) in the real world, failure data are too sparse to obtain stable estimates for models with many parameters. This paper aims to develop new models to address these two drawbacks.

1.1 Modelling the failure process of a series system

Let's consider the failure process of a series system composed of multiple components. Whenever a component fails, it is replaced with a new identical component. Suppose the replacement time is negligible. There are the following two scenarios in terms of data availability.

Component level. Once the system fails, both the time to the k th failure (denoted by T_k for $k = 1, 2, \dots, n$) and the component (denoted by C_k) that causes the system to fail are recorded.

System level. Once the system fails, only the time of the failure is recorded. But the component that causes the system to fail is not known.

On modelling the failure process of a system when failure data are only available at the system level, the widely used failure process for such a scenario is a super-imposed renewal process (SRP), which

is formed by the union of all the failures of the series system. It has been shown, for example, by Drenick (1960), unless the individual renewal processes are HPPs (homogeneous Poisson processes), an SRP is not a renewal process. When the number of components is large and time is large enough, an SRP behaves as an HPP (p264, Hoyland and Rausand (2004)). Zhang, Tian, Escobar, and Meeker (2017) proposes a procedure for estimating the component lifetime distribution based on system level data when the components are identical and assumed to be structured in series. The SRP may be useful as it is close to the real practice: a failed component is normally replaced with a new identical one and most real-world systems can be regarded as series systems. However, the application of the SRP may be hindered by the fact that it needs to know the renewal process of each individual component, that is, both T_k and C_k should be available for modelling. It is known that *“one major problem associated with reliability data is, ironically, the lack of sufficient data to properly run statistical analyses, as many authors mentioned repeatedly”* (Louit, Pascual, & Jardine, 2009), data, C_k , about which component causes the system to fail may not be recorded and may therefore be sparse. As such, it is not possible to estimate the lifetime distribution of each individual component and then obtain the failure intensity function of the system based on the system structure information. Hence, there is a need to develop failure process models based on the limited number of failure data of a system even if the causes of system failures is not identified. That is, a failure process model is needed for the scenario when C_k is not available, T_k is available, and n is small.

Assuming that C_k is not available and T_k is available, Wu and Scarf (2017) propose two models, Model I and Model II, to model the failure process of a real series system. With artificially generated data, they found that Model II outperforms Model I and four other models (i.e., the renewal process, the geometric process, the non-homogeneous Poisson process with the power law intensity, and the generalised renewal process). Model II regards the failure process of a real series system equivalent to that of a virtual system composed of two subsystems: subsystems 1 and 2. Subsystem 1 contains one virtual component (VC) and subsystem 2 is composed of a pre-specified number m of VCs. Broadly speaking, whenever the real system fails, minimal repair is assumed to be conducted on subsystem 1 and the oldest VC in subsystem 2 is assumed to be replaced with a new identical VC. The number of VCs in subsystem 2 is the integer number that has the maximum value of the maximum likelihood values for possible m 's (Wu & Scarf, 2017).

A drawback of Model II is that it includes too many parameters. For example, if both the intensity functions of the failure processes of the two subsystems are two-parameter power law functions, then Model II has four unknown parameters that must be estimated from failure data. This number may be too large if Model II has to be built on a small number of failure data, which may hinder it from a wide application.

On the basis of the above discussion, we conclude that models with good performance and a small number of parameters are needed for modelling the failure process of a series system for the scenario only if T_k ($k = 1, 2, \dots, n$) are available.

1.2 The proposed method

To model the failure process of a series system, this paper proposes a new model, referred to as the ESI (Exponential Smoothing of Intensity) model, and investigates one of its special case models, which is called the MAI (Moving Average of Intensity) model. The ESI model simply regards the failure process of a real-world system equivalent to that of a virtual system composed of a number of different VCs and the MAI assumes that those VCs are identical. In either model, we assume: if the real system fails, the oldest VC is replaced. We then develop the models for the virtual system instead of the failure process of each individual VC.

As can be seen, both the ESI and the MAI models simulate the failure process of a multi-component system but do not need to know the renewal process of each component. As such, they may find a wider application than the SRP.

It should be noted that the MAI model does not have any additional parameter over a given intensity function such as the power law intensity function, whereas the ESI model has one more parameter than that of the MAI model.

1.3 Summary

The rest of the paper is structured as following. Section 2 introduces the models, investigates their properties, gives a parameter estimation method. Section 3 compares the proposed models with nine existing models on both artificially generated data and real-world data. Section 4 discusses methods of extending the proposed models and potential applications. Section 5 concludes the paper and plans future work.

2 Model development

Let $\{T_k : k = 1, 2, \dots\}$ be the successive failure times of a series system composed of multiple components, starting from $T_0 = 0$, and N_t be the number of failures up to time t . Whenever a component fails, it is replaced with a new identical component. Suppose the replacement time is negligible. Let \mathcal{H}_{t-} denote the history of the failure process up to t (exclusive of t). The failure process of the system can be defined equivalently by the random processes $\{T_k\}_{k \geq 1}$ or $\{N_t\}_{t \geq 0}$ and

is characterised by the intensity function,

$$\lambda(t|\mathcal{H}_{t-}) = \lim_{\Delta t \downarrow 0} \frac{P\{N_{t+\Delta t} - N_t = 1 | \mathcal{H}_{t-}\}}{\Delta t}, \quad (1)$$

where $P\{N_{t+\Delta t} - N_t = 1 | \mathcal{H}_{t-}\}$ is the probability that the system fails within the interval $(t, t + \Delta t)$, given the history of failures up to time t , \mathcal{H}_{t-} (Cox & Lewis, 1966).

2.1 The exponential smoothing of intensity model

Model II proposed in (Wu & Scarf, 2017) outperforms the four existing models (i.e., the renewal process, the geometric process, the non-homogeneous Poisson process and the generalised renewal process), which encourages us to propose a new model, referred to as the exponential smoothing of intensity (ESI) model.

Similar to Wu and Scarf (2017), in this paper, we distinguish three systems: i) the real system, which may be a car or an air-conditioner; ii) the system, that is, the mathematical model of the system, e.g. a series system with a number of non-repairable, non-identical components; and iii) the virtual system, which consists of VCs.

Strictly, to introduce the ESI model, we make the following assumptions with descriptions where needed, which are similar to the assumptions of the virtual subsystem of model II in Wu and Scarf (2017).

- (a) Suppose the failure process of a real series system of multiple components. Once the system fails, the failed component can be immediately identified and replaced with a new identical one. The occurrences of the failures of the components in the real system are assumed to statistically independent.
- (b) Denote $t(> 0)$ as a time variable starting from the time when the real system is started at time $t = 0$.
- (c) The real system is assumed to be equivalent to a virtual system that is composed of m different VCs in series. The virtual system itself does not fail and is assumed to fail as the same time when the real system fails. The occurrences of the failures of the VCs are assumed to statistically independent.
- (d) The failure process of the real system is modelled with the intensity process of the virtual system as described below.

- (d.i) Before the first failure of the real system, the VCs have intensity functions $\frac{1}{m}\lambda_0(t)$, $\frac{1}{m}\rho\lambda_0(t)$, \dots , $\frac{1}{m}\rho^{m-1}\lambda_0(t)$, respectively, where $\rho \in (0, 1]$, $k \in \{1 \dots m\}$, and $\lambda_0(t)$ is a continuous function of time with $\lambda_0(t) > 0$. Hence, $\frac{1}{m}\rho^{m-1}\lambda_0(t) < \frac{1}{m}\rho^{m-2}\lambda_0(t) < \dots < \frac{1}{m}\lambda_0(t)$. $\frac{1}{m}\lambda_0(t)$ is the largest intensity function and its associated VC may be more likely to fail than the other VCs with smaller intensity functions.
- (d.ii) After the first failure of the real system occurs at T_1 , the VC with intensity $\frac{1}{m}\lambda_0(t)$ is replaced with a new VC with intensity $\frac{1}{m}\rho^{m-1}\lambda_0(t - T_1)$ and the intensity function of a VC from the other $m - 1$ VCs becomes the product of its previous intensity and ρ^{-1} . To be more specifically, the intensity function of a component from the $m - 1$ VCs changes from $\frac{1}{m}\rho^{m-k}\lambda_0(t)$ to $\frac{1}{m}\rho^{m-k-1}\lambda_0(t)$ for $k = 1, \dots, m - 1$, respectively. Similarly, if the real system fails at T_{N_t} and $2 \leq N_t < m$, the VC with intensity function $\frac{1}{m}\lambda_0(t)$ is replaced with a new VC with intensity $\frac{1}{m}\rho^{m-1}\lambda_0(t - T_N)$ and the intensity function of a VC from the other $m - 1$ VCs becomes the product of its previous intensity and ρ^{-1} . Table 1 shows the changes of the intensity functions of the VCs after the first and the second failures of the real system, respectively.

Table 1: Changes of the intensity functions of the VCs after the first 3 failures of the real system

$t \in (0, T_1]$	$t \in (T_1, T_2]$	$t \in (T_2, T_3]$	$t \in (T_3, T_4]$
$\frac{1}{m}\lambda_0(t)$	$\frac{1}{m}\rho^{m-1}\lambda_0(t - T_1)$	$\frac{1}{m}\rho^{m-2}\lambda_0(t - T_1)$	$\frac{1}{m}\rho^{m-3}\lambda_0(t - T_1)$
$\frac{1}{m}\rho\lambda_0(t)$	$\frac{1}{m}\lambda_0(t)$	$\frac{1}{m}\rho^{m-1}\lambda_0(t - T_2)$	$\frac{1}{m}\rho^{m-2}\lambda_0(t - T_2)$
$\frac{1}{m}\rho^2\lambda_0(t)$	$\frac{1}{m}\rho\lambda_0(t)$	$\frac{1}{m}\lambda_0(t)$	$\frac{1}{m}\rho^{m-1}\lambda_0(t - T_3)$
\dots	\dots	\dots	\dots
$\frac{1}{m}\rho^{m-1}\lambda_0(t)$	$\frac{1}{m}\rho^{m-2}\lambda_0(t)$	$\frac{1}{m}\rho^{m-3}\lambda_0(t)$	$\frac{1}{m}\rho^{m-4}\lambda_0(t)$

- (d.iii) If the real system fails at T_{N_t} and $N_t > m$, the VC with intensity $\frac{1}{m}\lambda_0(t - T_{N_t - m})$ (which is the oldest VC in the virtual system) is replaced with a new VC with intensity $\frac{1}{m}\rho^{m-1}\lambda_0(t - T_N)$ and the intensity function of a VC from the other $m - 1$ VCs becomes the product of its previous intensity and ρ^{-1} . Table 2 shows the changes of the intensity functions of the VCs after the T_{N_t} th and the T_{N_t+1} th failures of the real system, respectively.

The assumption that the VC with the intensity function $\frac{1}{m}\lambda_0(t)$ fails first (for $1 \leq N_t \leq m$) or the oldest VC fails first (for $N_t > m$) is made on the virtual system. For a real system, due to the randomness of the failure occurrences, it is not possible to know exactly which component might have caused the system to fail. As such, a model with some assumptions has to be developed to approximate the real situation.

Table 2: Changes of the intensity functions of the VCs after the T_{N_t} th and the T_{N_t+1} th failures of the real system

$t \in (T_{N_t-1}, T_{N_t}]$	$t \in (T_{N_t}, T_{N_t+1}]$	$t \in (T_{N_t+1}, T_{N_t+2}]$
$\frac{1}{m}\lambda_0(t - T_{N_t-m})$	$\frac{1}{m}\rho^{m-1}\lambda_0(t - T_{N_t})$	$\frac{1}{m}\rho^{m-2}\lambda_0(t - T_{N_t})$
$\frac{1}{m}\rho\lambda_0(t - T_{N_t-m+1})$	$\frac{1}{m}\lambda_0(t - T_{N_t-m+1})$	$\frac{1}{m}\rho^{m-1}\lambda_0(t - T_{N_t+1})$
$\frac{1}{m}\rho^2\lambda_0(t - T_{N_t-m+2})$	$\frac{1}{m}\rho\lambda_0(t - T_{N_t-m+2})$	$\frac{1}{m}\lambda_0(t - T_{N_t-m+2})$
\dots	\dots	\dots
$\frac{1}{m}\rho^{m-1}\lambda_0(t - T_{N_t-1})$	$\frac{1}{m}\rho^{m-2}\lambda_0(t - T_{N_t-1})$	$\frac{1}{m}\rho^{m-3}\lambda_0(t - T_{N_t-1})$

Assumptions (d.i), (d.ii) and (d.iii) are inspired by the principle of the simple exponential smoothing model that has broad applications in time series forecasting (see Chapter 4 in the book by Makridakis, Wheelwright, and Hyndman (2008), for example). For a given time series Y_t , let \hat{Y}_{t+1} denote the forecast of Y_t . With the simple exponential smoothing method, \hat{Y}_{t+1} can be estimated by

$$\hat{Y}_{t+1} = \alpha_0 \sum_{k=0}^{t-1} (1 - \alpha_0)^k Y_{t-k} + (1 - \alpha_0)^t Y_0. \quad (2)$$

where α_0 is the smoothing factor, and $0 < \alpha_0 < 1$ (Makridakis et al., 2008). The exponential smoothing assigns exponentially decreasing weights as the observation get older, which agrees with the fact that older components may have larger intensity functions than younger ones.

Assumptions (d.ii) and (d.iii) give recently installed components relatively more weight in forecasting than the older components. They imply that the replacement method relates to not only the actual elapsed time of a VC but also the times of failures of the real system.

It is noted that the geometric process and its extensions make a similar assumption that the replacement relates to the times of failures of the real system (see Lam (1988); Bordes and Mercier (2013); Wu (2018), for example).

Then, from the Assumptions (a–d), the failure intensity of the virtual system after the N_t -th failure is given by

$$\lambda(t|\mathcal{H}_{t-}) = \begin{cases} \frac{1}{m} \sum_{k=0}^{m-1} \rho^{m-k-1} \lambda_0(t), & \text{for } N_t = 0, \\ \frac{1}{m} \left(\sum_{k=0}^{N_t-1} \rho^{m-k-1} \lambda_0(t - T_{N_t-k}) + \sum_{k=N_t}^{m-1} \rho^{m-k-1} \lambda_0(t) \right), & \text{for } 1 \leq N_t < m, \\ \frac{1}{m} \sum_{k=0}^{m-1} \rho^{m-k-1} \lambda_0(t - T_{N_t-k}), & \text{for } N_t \geq m. \end{cases} \quad (3)$$

where $\rho > 0$. t is the time since the system starts, $t \in (0, +\infty)$. We refer to the count process with

intensity $\lambda(t|\mathcal{H}_{t-})$ as an ESI model, or more precisely, an ESI(m) model.

$\lambda(t|\mathcal{H}_{t-})$ in (3) is interpreted below.

- Before the first system failure, i.e., for the case when $N_t = 0$, according to Assumptions (c) and (d.i), the intensity function of the virtual system is $\frac{1}{m} \sum_{k=0}^{m-1} \rho^{m-k-1} \lambda_0(t)$, which is the sum of the intensity functions of the m VCs in series structure.
- For the case when $1 \leq N_t < m$, based on Assumption (d.ii), the ages of some VCs are t and those of the other VCs are younger than t due to replacement. Those VCs with age t have failure rate functions $\frac{1}{m} \rho^{m-k-1} \lambda_0(t)$ for $N_t \leq k \leq m-1$, respectively. Hence, the failure intensity function of the part composed of those VCs aged t is $\frac{1}{m} \sum_{k=N_t}^{m-1} \rho^{m-k-1} \lambda_0(t)$. Those VCs with age younger than t have failure intensity functions $\frac{1}{m} \rho^{m-k-1} \lambda_0(t - T_{N_t-k})$ for $0 \leq k \leq N_t - 1$, respectively. Hence, the failure intensity function of the part composed of those VCs with age younger than t is $\frac{1}{m} \sum_{k=0}^{N_t-1} \rho^{m-k-1} \lambda_0(t - T_{N_t-k})$.

As such, the failure intensity function of the entire virtual system for $1 \leq N_t < m$ is given by

$$\frac{1}{m} \left(\sum_{k=0}^{N_t-1} \rho^{m-k-1} \lambda_0(t - T_{N_t-k}) + \sum_{k=N_t}^{m-1} \rho^{m-k-1} \lambda_0(t) \right).$$

- For the case when for $N_t \geq m$, according to Assumption (d.iii), the failure intensity functions of the VCs are $\rho^{m-k-1} \lambda_0(t - T_{N_t-k})$, respectively. Hence, the failure intensity function of the virtual system is $\frac{1}{m} \sum_{k=0}^{m-1} \rho^{m-k-1} \lambda_0(t - T_{N_t-k})$.

Remarks

- If $m = 1$, then $\lambda(t|\mathcal{H}_{t-}) = \lambda_0(t - T_{N_t})$ and the process is the renewal process.
- It should be noted that $\lambda(t|\mathcal{H}_{t-})$ does not require the monotonicity of $\lambda_0(t)$ in t . Similar requirements are needed by some existing models, for example, the ARI_m and ARA_m models assume that the initial intensity function is strictly increasing (Doyen & Gaudoin, 2004).

If the intensity $\lambda_0(t)$, which is also called the initial intensity function in the cases of ARA and ARI, is an increasing function, then the intensity of the failure process model in some existing models such as ARI_m and ARA_m (Doyen & Gaudoin, 2004), Kijima I and Kijima II models (Kijima, 1989) converges to infinite for $t \rightarrow \infty$. Unlike those models, $\lambda(t|\mathcal{H}_{t-})$ in the ESI model does not converge to infinity for $t \rightarrow \infty$ even if $\lambda_0(t)$ is an increasing function in t , which is proved in Lemma 1.

Denote $G_k(t)$ as the probability distribution of time to the k -th failure, T_k , of the component with intensity function $\lambda_0(t)$. Then we can obtain the following lemma.

Lemma 1 *For a given m , if $\int_0^\infty \lambda_0(u)dG_k(u) < \infty$, then $E[\lambda(t|\mathcal{H}_{t-})] < \infty$ for $t \rightarrow \infty$.*

It is easy to establish Lemma 1: if $t \rightarrow \infty$, $\lambda(t|\mathcal{H}_{t-}) = \frac{1}{m} \sum_{k=0}^{m-1} \rho^{m-k-1} \lambda_0(t - T_{N_t-k})$. Hence, the

$$\text{expected value } E \left[\frac{1}{m} \sum_{k=0}^{m-1} \rho^{m-k-1} \lambda_0(t - T_{N_t-k}) \right] = \frac{1}{m} \sum_{k=0}^{m-1} \rho^{m-k-1} \int_0^\infty \lambda_0(u)dG_k(u) < \infty.$$

An implication of Lemma 1 is its ability to model the failure process of multi-component systems. This is because: The intensity function of a one-component system may become infinite whereas the intensity function of a multi-component system may become equilibrium (Drenick, 1960), as long as any failed component in the system is replaced with a new identical component. It should be reminded that most of the existing models including Kijima's virtual age models, i.e, Kijima I and Kijima II (Kijima, 1989), are not able to model such a failure process as their intensity functions converge to infinity for $t \rightarrow \infty$.

One can also easily establish the following lemma.

Lemma 2 • *If $\lambda_0(t)$ is an increasing function in t , then $\lambda(t|\mathcal{H}_{t-})$ is an increasing function for $t \in (T_{N_t}^+, T_{N_{t+1}}^-)$;*

• *If $\lambda_0(t)$ is a decreasing function in t , then $\lambda(t|\mathcal{H}_{t-})$ is a decreasing function for $t \in (T_{N_t}^+, T_{N_{t+1}}^-)$.*

Lemma 2 implies that: if $\lambda_0(t)$ is an increasing function in t , then the intensity of the process is increasing in t between adjacent failures. This result is exactly the case for many systems as they are ageing or deteriorating between adjacent failures. However, it should be noted that $\lambda(t|\mathcal{H}_{t-})$ is a non-monotone function over time $t(\in (0, +\infty))$ even if $\lambda_0(t)$ is a monotone function.

Unlike Model II (Wu & Scarf, 2017) in which all the VCs in subsystem 2 are identical, the VCs in the ESI model are different in the sense that they have different intensity functions.

The ESI model is a simpler model than Model II as it contains fewer parameters. For example, for the ESI model, if the power law intensity function is assumed to each VC, then the ESI model only contains three parameters.

2.2 The moving average of intensity model (MAI)

The ESI model has one more parameter than $\lambda_0(t)$. We may remove this additional parameter by setting it to be one, that is, let $\rho = 1$, which means Assumptions (d.i), (d.ii) and (d.iii) become

(d.i') before the first failure of the system, the VCs are identical with intensity function $\frac{1}{m}\lambda_0(t)$, and
(d.ii') assume whenever the real system fails, the oldest VC is replaced with a component with intensity $\frac{1}{m}\lambda_0(t)$.

It should be noted that although the oldest VC has the same form of the intensity function as $\frac{1}{m}\lambda_0(t)$, the argument t in $\frac{1}{m}\lambda_0(t)$ of the oldest VC is larger than the one in the new component. That is, the ages of the m components are different.

If $\rho = 1$ in the ESI model (3), then the failure intensity of the system after the N_t -th failure ($N_t \geq 1$) reduces to

$$\lambda(t|\mathcal{H}_{t-}) = \begin{cases} \lambda_0(t), & \text{for } N_t = 0, \\ \frac{1}{m} \left(\sum_{k=0}^{N_t-1} \lambda_0(t - T_{N_t-k}) + (m - N_t)\lambda_0(t) \right), & \text{for } 1 \leq N_t < m, \\ \frac{1}{m} \sum_{k=0}^{m-1} \lambda_0(t - T_{N_t-k}), & \text{for } N_t \geq m. \end{cases} \quad (4)$$

where $t \in (0, +\infty)$.

We refer to the process with intensity $\lambda(t|\mathcal{H}_{t-})$ in (4) as a moving average of intensity (MAI) model, more precisely, an MAI(m) model. The MAI model assumes that a real system is equivalent to a virtual system composed of m identical VCs structured in series. If the real system fails, the virtual system is assumed to fail and a new VC with intensity function $\frac{1}{m}\lambda_0(t)$ is used to replace the oldest VC.

Remarks

- Similar to the ESI model, if $m = 1$, then $\lambda(t|\mathcal{H}_{t-}) = \lambda_0(t - T_{N_t})$ and the process is a renewal process.
- If $m \rightarrow \infty$, then for a given t , $\frac{1}{m} \left(\sum_{k=0}^{N_t-1} \lambda_0(t - T_{N_t-k}) + (m - N_t)\lambda_0(t) \right) = \frac{1}{m} \sum_{k=0}^{N_t-1} \lambda_0(t - T_{N_t-k}) + (1 - \frac{N_t}{m})\lambda_0(t) \rightarrow \lambda_0(t)$. That is, $\lambda(t|\mathcal{H}_{t-}) \rightarrow \lambda_0(t)$, which implies that the MAI model process can be approximated by a NHPP.
- From the above two bullet points, one can see that the MAI model can model perfect, imperfect and minimal maintenance, where m is the indicator of maintenance effectiveness. Whenever the real system fails, one out of the m VCs is replaced. A small m implies a large proportion of the components in a real system is repaired whereas a large m implies a small proportion of the components is repaired.

- As can be seen, the number of the parameters in the MAI model is the same as that in $\lambda_0(t)$, which is vitally important as models with a small number of parameters are favourable for real-world situations, as discussed above.
- Recalling the failure process model Model II (Wu & Scarf, 2017), which is discussed in Section 1.2, we may regard the MAI model as Model II when its subsystem 1 is taken away and only its subsystem 2 is remained.

Example 1. Let $\lambda_0(t) = 0.05t^{1.5}$, then Figure 1 shows $\lambda_0(t)$ and an example of $\lambda(t|\mathcal{H}_{t-})$ with $m = 15$, where the X -axis represents t and the Y -axis represents the failure intensity.

Example 2. Similar to Example 1, let $\lambda_0(t) = 0.2t^{-0.5}$, then Figure 2 shows $\lambda_0(t)$ and an example of $\lambda(t|\mathcal{H}_{t-})$ with $m = 15$.

The $\lambda(t|\mathcal{H}_{t-}) \geq \lambda_0(t)$, represented by the zigzag lines in Figure 1 and Figure 2, are plotted with Monte Carlo simulation. Figure 1 and Figure 2 show that $\lambda(t|\mathcal{H}_{t-}) \leq \lambda_0(t)$ and $\lambda(t|\mathcal{H}_{t-}) \geq \lambda_0(t)$, respectively, and both quickly become equilibrium when t becomes bigger. The equilibrium for the failure intensity as t becomes bigger in Figures 1 and 2 may be interpreted below.

Take Figure 1 as an example on equilibrium. Suppose a series system composed of multiple components, each of which has an increasing failure rate. At the beginning, all of the components are new. As times goes by, the failure rate of each component becomes larger and therefore the failure intensity function of the system increases. A time period later, the system is composed of a mix of components with different ages and failure rate functions, some of which are large and some are small. The failure intensity function of the system is then fluctuated around a fixed value.

If $\lambda_0(t)$ is an increasing function in t , then $\lambda_0(t - T_{N_t-k}) \leq \lambda_0(t)$. Similarly, if $\lambda_0(t)$ is a decreasing function in t , then $\lambda_0(t - T_{N_t-k}) \geq \lambda_0(t)$. Hence, we can obtain Lemma 3.

Lemma 3 • If $\lambda_0(t)$ is an increasing function in t , then $\lambda(t|\mathcal{H}_{t-}) \leq \lambda_0(t)$;

- If $\lambda_0(t)$ is an decreasing function in t , then $\lambda(t|\mathcal{H}_{t-}) \geq \lambda_0(t)$;
- If $\lambda_0(t) = \lambda_0$, where λ_0 is a constant, then $\lambda(t|\mathcal{H}_{t-}) = \lambda_0$.

2.3 Expected number of failures in time interval $(0, t]$

In this section, we derive the expected number of failures in time interval $(0, t]$ and give a Monte Carlo simulation example of estimating times to failures, which is useful on how the proposed models may be used in lifetime forecasting.

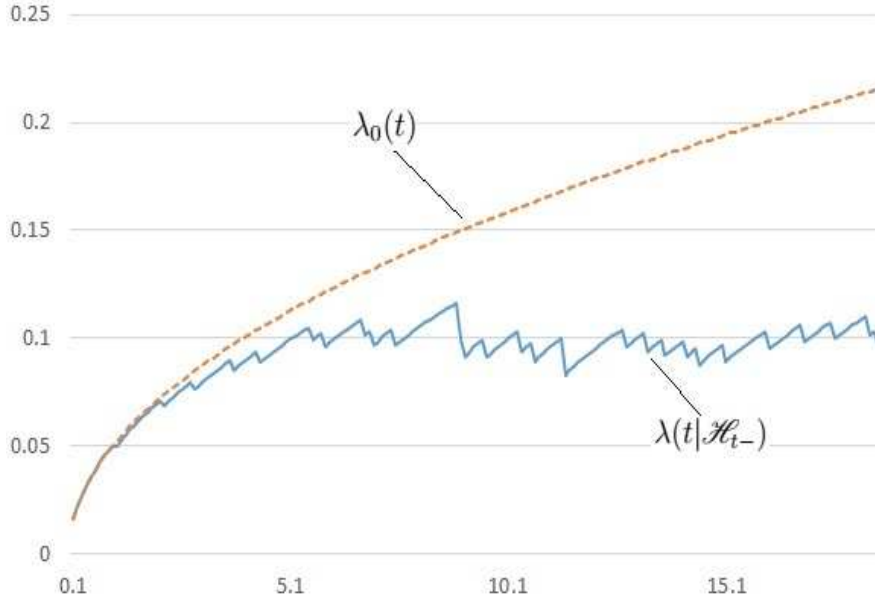


Figure 1: $\lambda(t|\mathcal{H}_{t-})$ with $m = 15$ and $\lambda_0(t)$, where $\lambda_0(t) = 0.05t^{1.5}$.

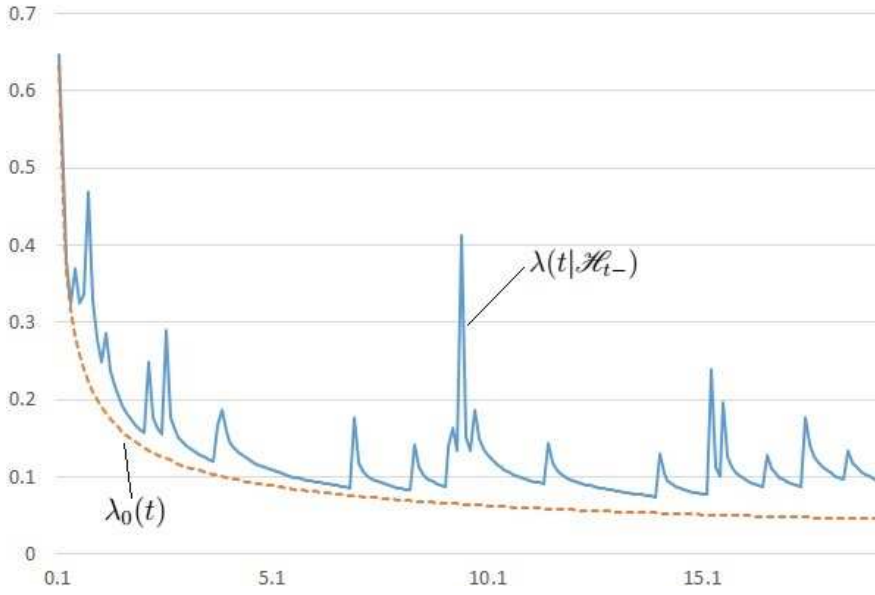


Figure 2: $\lambda(t|\mathcal{H}_{t-})$ with $m = 15$ and $\lambda_0(t)$, where $\lambda_0(t) = 0.2t^{-0.5}$.

2.3.1 Exploring the expected value of failures

Denote $F_n(t) = \Pr\{T_n < t\}$ and $\Lambda_0(t) = \int_0^t \lambda_0(u)du$. From Eq. (3), we obtain

$$\Pr\{T_1 < t\} = 1 - \exp\left(-\frac{1}{m} \sum_{k=0}^{m-1} \rho^{m-k-1} \Lambda_0(t)\right), \quad (5)$$

For $2 \leq n < m + 1$ and $t_1 < t_2 < \cdots < t_{n-1} < t$, we have

$$\Pr\{T_n < t | T_{n-1} = t_{n-1}, \dots, T_1 = t_1\} = 1 - \exp\left(-\frac{1}{m} \sum_{k=0}^{n-2} \rho^{m-k-1} \Lambda_0(t - t_{n-k-1}) - \frac{1}{m} \sum_{k=n-1}^{m-1} \rho^{m-k-1} \Lambda_0(t)\right), \quad (6)$$

For $n \geq m + 1$, we obtain

$$\Pr\{T_n < t | T_{n-1} = t_{n-1}, \dots, T_{n-m} = t_{n-m}\} = 1 - \exp\left(-\frac{1}{m} \sum_{k=0}^{m-1} \rho^{m-k-1} \Lambda_0(t - t_{n-k-1})\right), \quad (7)$$

where $t_{n-m} < t_{n-m+1} < \cdots < t_{n-1} < t$.

According to Ibragimov (2009), T_n forms a Markov process of order m . Further, we have

$$f(t_1, t_2, \dots, t_n) = f(t_n | t_{n-1}, \dots, t_{n-m}) \cdots f(t_{m+1} | t_m, \dots, t_1) f(t_1, \dots, t_m) \quad (8)$$

where $f(t_k | t_{k-1}, \dots, t_{k-m}) = \frac{\partial P(T_k < t_k | T_{k-1} = t_{k-1}, \dots, T_{k-m} = t_{k-m})}{\partial t_k}$, $k = m + 1, \dots, n$, $f(t_1, \dots, t_m) = \frac{\partial^m F_{1, \dots, m}(t_1, \dots, t_m)}{\partial t_1 \cdots \partial t_m}$, and $F_{1, \dots, m}(t_1, \dots, t_m) = P(T_1 < t_1, \dots, T_m < t_m)$. Since

$$f(t_1, t_2, \dots, t_m) = f(t_m | t_{m-1}, \dots, t_1) \cdots f(t_2 | t_1) f(t_1), \quad (9)$$

we have

$$f(t_1, t_2, \dots, t_n) = f(t_n | t_{n-1}, \dots, t_{n-m}) \cdots f(t_{m+1} | t_m, \dots, t_1) f(t_m | t_{m-1}, \dots, t_1) \cdots f(t_2 | t_1) f(t_1). \quad (10)$$

Since $t_1 < \cdots < t_{n-1} < t$, we obtain,

$$\begin{aligned} F_n(t) &= \int_0^t \cdots \int_0^t f(t_n | t_{n-1}, \dots, t_1) f(t_1, \dots, t_{n-1}) dt_1 \cdots dt_n \\ &= \int_0^t \cdots \int_0^t f(t_n | t_{n-1}, \dots, t_{n-m}) f(t_1, \dots, t_{n-1}) dt_1 \cdots dt_n \\ &= \int_0^t \cdots \int_0^t f(t_1, \dots, t_{n-1}, t_n) dt_1 \cdots dt_n. \end{aligned} \quad (11)$$

Apparently, $N_t \leq n \iff T_n \leq t$, hence, we can obtain $\Pr\{N_t = n\} = F_n(t) - F_{n+1}(t)$. As such, $\mathbb{E}[N_t] = \sum_{n=1}^{\infty} F_n(t)$.

Due to the complicating expression of $F_n(t)$, it is difficult to obtain a more closed form of expression of $\mathbb{E}[N_t]$. A Monte Carlo simulation method may therefore be sought, which is illustrated in the example in the following subsection.

2.3.2 An example of forecasting the time to the first fifty failures

Let u be a uniformly distributed random variable within interval $(0, 1)$. Denote $a = \sum_{k=0}^{m-1} \rho^{m-k-1}$,
 $b_n = -2 \sum_{k=0}^{n-2} \rho^{m-k-1} t_{n-k-1}$, $b_m = -2 \sum_{k=0}^{m-1} \rho^{m-k-1} t_{n-k-1}$, $c_0 = \frac{m \ln(1-u)}{\alpha}$, $c_n = \sum_{k=0}^{n-2} \rho^{m-k-1} t_{n-k-1}^2$,
 $c_m = \sum_{k=0}^{m-1} \rho^{m-k-1} t_{n-k-1}^2$. Assume $\Lambda(t) = \alpha t^2$. From Eq. (5), one can obtain

$$t_1 = \sqrt{\frac{-m \ln(1-u)}{\frac{1}{m} \sum_{k=0}^{m-1} \rho^{m-k-1}}} \quad (12)$$

From Eq. (6), we have

$$\begin{aligned} -\ln(1-u) &= \frac{1}{m} \sum_{k=0}^{n-2} \rho^{m-k-1} \Lambda_0(t - t_{n-k-1}) + \frac{1}{m} \sum_{k=n-1}^{m-1} \rho^{m-k-1} \Lambda_0(t) \\ &= \frac{\alpha}{m} \sum_{k=0}^{n-2} \rho^{m-k-1} (t - t_{n-k-1})^2 + \frac{\alpha}{m} \sum_{k=n-1}^{m-1} \rho^{m-k-1} t^2 \\ &= \left(\frac{\alpha}{m} \sum_{k=0}^{m-1} \rho^{m-k-1} \right) t^2 - \left(\frac{2\alpha}{m} \sum_{k=0}^{n-2} \rho^{m-k-1} t_{n-k-1} \right) t + \frac{\alpha}{m} \sum_{k=0}^{n-2} \rho^{m-k-1} t_{n-k-1}^2 \end{aligned} \quad (13)$$

Then, for $2 \leq n < m+1$,

$$t_n = \frac{-b_n + \sqrt{b_n^2 - 4a(c_0 + c_n)}}{2a} \quad (14)$$

Similarly, from Eq. (7), we have

$$\begin{aligned} -\ln(1-u) &= \frac{1}{m} \sum_{k=0}^{m-1} \rho^{m-k-1} \Lambda_0(t - t_{n-k-1}) \\ &= \frac{\alpha}{m} \sum_{k=0}^{m-1} \rho^{m-k-1} (t - t_{n-k-1})^2 \\ &= \left(\frac{\alpha}{m} \sum_{k=0}^{m-1} \rho^{m-k-1} \right) t^2 - \left(\frac{2\alpha}{m} \sum_{k=0}^{m-1} \rho^{m-k-1} t_{n-k-1} \right) t + \frac{\alpha}{m} \sum_{k=0}^{m-1} \rho^{m-k-1} t_{n-k-1}^2 \end{aligned} \quad (15)$$

Then, for $n \geq m + 1$,

$$t_n = \frac{-b_m + \sqrt{b_m^2 - 4a(c_0 + c_m)}}{2a}. \quad (16)$$

One can generate pseudo random numbers for the uniformly distribution variable u and then use Eqs. (12), (14), and (16) to generate t_1, t_2, \dots . For example, Table 3 shows the first 50 failure times (i.e., t_1, t_2, \dots, t_{50}) when $\Lambda(t) = 0.00458t^2$, $\rho = 1$, and $m = 4$. One may repeat the process for many times and use the distribution of the values $\{t_1, t_2, \dots\}$ for forecasting. For example, with the same parameter settings of $\Lambda(t)$, ρ and m , we repeat the data generation process for 5,000 times and find that the mean and standard deviation of the total time of the first 50 failures (i.e., t_1, t_2, \dots, t_{50}) are 270.71 and 116.30, respectively.

Table 3: Monte-Carlo simulation of the first 50 failure times, $\Lambda(t) = 0.00458t^2$, $\rho = 1$, and $m = 4$.

18.33	18.48	22.30	28.22	37.32	40.92	46.41	52.79	59.26	64.30
70.21	76.30	82.20	87.79	93.72	99.64	105.46	111.24	117.13	122.99
128.81	134.65	140.51	146.35	152.19	158.03	163.88	169.72	175.57	181.41
187.26	193.10	198.94	204.79	210.63	216.48	222.32	228.16	234.01	239.85
245.70	251.54	257.38	263.23	269.07	274.92	280.76	286.60	292.45	298.29

The above example assumes $\Lambda(t) = \alpha t^2$. For more complicated forms of $\Lambda(t)$, more complicated numerical methods may be used to obtain t_n ($n = 1, 2, \dots$) based on Eqs. (5), (6) and (7).

2.4 Likelihood and optimisation

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 is merely a sum of the likelihood of individual ones.

It 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.

The probability density function of the time to the first failure is given by

$$f(t) = \frac{1}{m} \sum_{j=0}^{m-1} \rho^{m-j-1} \lambda_0(t) \exp \left(-\frac{1}{m} \int_0^t \sum_{j=0}^{m-1} \rho^{m-j-1} \lambda_0(u) du \right). \quad (17)$$

Between the first and second system failures, one of the VCs is replaced with a new one and the others are not replaced. Hence, the newly installed one has age $t - t_1$ and the other $m - 1$ components have age t . As such, the intensity function of the new component is $\frac{1}{m} \rho^{m-1} \lambda_0(t)$ and the intensity functions of the other components are $\frac{1}{m} \sum_{j=1}^{m-1} \rho^{m-j-1} \lambda_0(t)$. Similarly, between the $(k - 1)$ th and

k th system failures (where $k < m$), $k - 1$ VCs are replaced and the others are not replaced. Hence, the newly installed ones have age $t - t_{m-j-1}$ for $j = 0, \dots, k - 2$ and the others have age t . As such, the intensity functions of the installed components are $\sum_{j=0}^{k-2} \rho^{m-j-1} \lambda_0(t - t_{k-j-1})$ and the intensity functions of the other components are $\frac{1}{m} \sum_{j=k-1}^{m-1} \rho^{m-j-1} \lambda_0(t)$. That is, the conditional probability density function of t , given the times to the first k failures, t_1, \dots, t_k (where $2 \leq k < m$) is given by

$$f(t|t_1, \dots, t_k) = \frac{1}{m} \left(\sum_{j=0}^{k-2} \rho^{m-j-1} \lambda_0(t - t_{k-j-1}) + \sum_{j=k-1}^{m-1} \rho^{m-j-1} \lambda_0(t) \right) \exp \left(-\frac{1}{m} \sum_{j=0}^{k-2} \rho^{m-j-1} \int_{t_{k-1}-t_{k-j-1}}^{t-t_{k-j-1}} \lambda_0(u) du - \frac{1}{m} \sum_{j=k-1}^{m-1} \rho^{m-j-1} \int_0^t \lambda_0(u) du \right) \quad (18)$$

and the conditional probability density function of t , given t_1, \dots, t_k (for $k \geq m$) is given by

$$f(\tau|t_1, \dots, t_k) = \frac{1}{m} \left(\sum_{j=0}^{m-1} \rho^{m-j-1} \lambda_0(t - t_{k-j-1}) \right) \exp \left(-\frac{1}{m} \sum_{j=0}^{m-1} \int_{t_{k-1}-t_{k-j-1}}^{t-t_{k-j-1}} \rho^{m-j-1} \lambda_0(u) du \right). \quad (19)$$

Hence, given a dataset of n successive failure times t_1, \dots, t_n , the likelihood function is

$$L = \frac{1}{m} \left(\sum_{j=0}^{m-1} \rho^{m-j-1} \lambda_0(t_1) \right) \exp \left(-\frac{1}{m} \sum_{j=0}^{m-1} \rho^{m-j-1} \int_0^{t_1} \lambda_0(u) du \right) \times \prod_{k=2}^m \left[\frac{1}{m} \left(\sum_{j=0}^{k-2} \rho^{m-j-1} \lambda_0(t_k - t_{k-j-1}) + \sum_{j=k-1}^{m-1} \rho^{m-j-1} \lambda_0(t_k) \right) \right] \times \prod_{k=2}^m \exp \left(-\frac{1}{m} \sum_{j=0}^{k-2} \int_{t_{k-1}-t_{k-j-1}}^{t_k-t_{k-j-1}} \rho^{m-j-1} \lambda_0(u) du - \frac{1}{m} \sum_{j=k-1}^{m-1} \int_{t_{k-1}}^{t_k} \rho^{m-j-1} \lambda_0(u) du \right) \times \prod_{k=m+1}^n \left[\frac{1}{m} \left(\sum_{j=0}^{m-1} \rho^{m-j-1} \lambda_0(t_k - t_{k-j-1}) \right) \exp \left(-\frac{1}{m} \sum_{j=0}^{m-1} \int_{t_{k-1}-t_{k-j-1}}^{t_k-t_{k-j-1}} \rho^{m-j-1} \lambda_0(u) du \right) \right]. \quad (20)$$

By maximising $\log(L)$, one can find optimal parameters $\hat{\alpha}$, $\hat{\beta}$ and $\hat{\rho}$, respectively.

As a special case, let $\lambda_0(t) = \alpha\beta t^{\beta-1}$ and let's derive the likelihood of the MAI model. Substituting $\lambda_0(t) = \alpha\beta t^{\beta-1}$ into Eq. (20), and let $\frac{\partial \log(L)}{\partial \alpha} = 0$ and $\frac{\partial \log(L)}{\partial \beta} = 0$, respectively, we can obtain the

estimators of parameters α and β by solving the following equations.

$$\hat{\alpha} = n \left[t_1^{\hat{\beta}} + \frac{1}{m} \sum_{k=2}^m \sum_{j=0}^{k-2} \left((t_k - t_{k-j-1})^{\hat{\beta}} - (t_{k-1} - t_{k-j-1})^{\hat{\beta}} \right) + \frac{1}{m} \sum_{k=2}^m (m - k + 1) (t_k^{\hat{\beta}} - t_{k-1}^{\hat{\beta}}) + \frac{1}{m} \sum_{k=m+1}^n \sum_{j=0}^{m-1} \left((t_k - t_{k-j-1})^{\hat{\beta}} - (t_{k-1} - t_{k-j-1})^{\hat{\beta}} \right) \right]^{-1}, \quad (21)$$

and

$$\begin{aligned} & n/\hat{\beta} + \log(t_1) - \hat{\alpha} t_1^{\hat{\beta}} \log(t_1) + \sum_{k=2}^m \frac{\sum_{j=0}^{k-2} (t_k - t_{k-j-1})^{\hat{\beta}-1} \log(t_k - t_{k-j-1}) + \sum_{j=k-1}^{m-1} t_k^{\hat{\beta}-1} \log(t_k)}{\sum_{j=0}^{k-2} (t_k - t_{k-j-1})^{\hat{\beta}-1} + \sum_{j=k-1}^{m-1} t_k^{\hat{\beta}-1}} \\ & + \sum_{k=m+1}^n \frac{\sum_{j=0}^{m-1} (t_k - t_{k-j-1})^{\hat{\beta}-1} \log(t_k - t_{k-j-1})}{\sum_{j=0}^{m-1} (t_k - t_{k-j-1})^{\hat{\beta}-1}} \\ & = \frac{\hat{\alpha}}{m} \sum_{k=2}^m \sum_{j=0}^{k-2} \left[(t_k - t_{k-j-1})^{\hat{\beta}} \log(t_k - t_{k-j-1}) - (t_{k-1} - t_{k-j-1})^{\hat{\beta}} \log(t_{k-1} - t_{k-j-1}) \right] \\ & + \frac{\hat{\alpha}}{m} \sum_{k=2}^m (m - k + 1) \left(t_k^{\hat{\beta}} \log(t_k) - t_{k-1}^{\hat{\beta}} \log(t_{k-1}) \right) \\ & + \frac{\hat{\alpha}}{m} \sum_{k=m+1}^n \sum_{j=0}^{m-1} \left((t_k - t_{k-j-1})^{\hat{\beta}} \log(t_k - t_{k-j-1}) - (t_{k-1} - t_{k-j-1})^{\hat{\beta}} \log(t_{k-1} - t_{k-j-1}) \right), \quad (22) \end{aligned}$$

respectively.

Substituting $\hat{\alpha}$ in Eq. (21) into Eq. (22), we can obtain an equation that only contains $\hat{\beta}$. Due to the complexity of the equation, one may pursue a heuristic algorithm such as the genetic algorithm to find the solution. Once $\hat{\beta}$ is obtained, substituting it back to Eq. (21), we can obtain $\hat{\alpha}$.

One can easily calculate the 2×2 observed information matrix $I(\hat{\alpha}, \hat{\beta})$. By taking the inverse of $I(\hat{\alpha}, \hat{\beta})$, he/she can obtain the variance and covariance matrix for $\hat{\alpha}$ and $\hat{\beta}$ and then obtain the 2 sided confidence intervals for them.

3 Simulation and case studies

In this section, we use both artificially generated data and real-world data to compare the performance of the ESI and MAI models with nine existing models. The nine models are: RP (renewal process), GP (geometric process), NHPP-PL (non-homogeneous Poisson process with the power law intensity), Kijima I, Kijima II, ARI_m, ARA_m, BBIP (Bounded Bathtub Intensity Process) and Model II. The

reader is referred to Kijima (1989) for the details of Kijima I and Kijima II models, to Doyen and Gaudoin (2004) for the introduction of the ARI_m and ARA_m models, to Guida and Pulcini (2009) for the introduction of the BBIP model and to Wu and Scarf (2017) for the introduction of Model II. The method to estimate the parameters in a model is the maximum likelihood estimation method.

For the purpose of measuring and comparing the performance of the models, we calculate the AIC, AIC_c , and BIC values for each model:

$$AIC = -2\log(L) + 2q, \quad (23)$$

$$AIC_c = -2\log(L) + 2q + \frac{2(q+2)(q+3)}{n-q-2}, \quad (24)$$

and

$$BIC = -2\log(L) + q\log(n), \quad (25)$$

respectively, 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. The term $2q$, $\frac{2(q+2)(q+3)}{n-q-2}$, and $q\log(n)$ in the AIC, AIC_c and BIC penalise a model with a large number of parameters, respectively. The reader is referred to Burnham and Anderson (2002) for details on model performance measures.

Let $\lambda_0(t)$ be the power law, that is, $\lambda_0(t) = \alpha\beta t^{\beta-1}$, where $\alpha > 0$ and $\beta > 1$. The values of q of the eleven models are shown in Table 4: models RP, GP and MAI include two parameters, respectively, whereas Model II and BBIP, in both of which $q = 4$, incur the highest penalty on their AIC, AIC_c , and BIC values, respectively.

Table 4: The number of unknown parameters in each model.

	RP	NHPP-PL	GP	Kijima I	Kijima II	ARI_m	ARA_m	BBIP	Model II	ESI	MAI
q	2	2	3	3	3	3	3	4	4	3	2

3.1 Simulation study

In this subsection, we fit the 11 models to artificially generated datasets, and then compare the AIC, AIC_c and BIC values of those models.

The following steps, which are similar to the numerical example steps in Wu and Scarf (2017), compare the performance of the models.

- (1) Assume a series system composed of m components. If a component fails, a new identical component is used to replace the failed one. Replacement time can be neglected.

- (2) We consider three cases $n = 15, 30, 45$ for the number of failures. For each case, we assume that $m = 5, 15, 25$, respectively, for the number of components in a system.
- (3) The time to failure of each component follows a Weibull distribution $1 - e^{-\left(\frac{t}{\gamma}\right)^\eta}$, where the parameter η and the parameter γ are randomly selected from the uniform distribution, on the intervals $\eta \in (0.5, 4)$ and $\gamma \in (12, 60)$, respectively.
- (4) Using the artificially generated data implied by the above, we fit each of the 11 models: RP, NHPP, GP, Kijima I, Kijima II, ARI_m , ARA_m , BBIP, Model II, ESI and MAI, by maximising their respective log-likelihoods. The reader is referred to Lam (2007) for the likelihoods of the RP, the GP and the NHPP-PL, respectively, to Yanez, Joglar, and Modarres (2002) for the likelihood of the Kijima I (the likelihood of Kijima II can be similarly obtained), to Doyen and Gaudoin (2004) for the likelihood of ARI_m (the likelihood of ARA_m can be similarly obtained), to Guida and Pulcini (2009) for the likelihood of the BBIP, and to Wu and Scarf (2017) for the likelihood of Model II. In all the models, we assume $\lambda_0(t) = \alpha\beta t^{\beta-1}$.

We repeat the above steps (1) to (4) for 30 repetitions, in each of which the values of γ and η are different. Calculations are done with a statistical package R (which can be downloaded from www.r-project.org/). Table 5 shows the mean values and the variances (that are shown in the brackets under means) of the $-\log(L)$ values of the models over the 30 repetition for each combination.

Table 5: The means and standard deviations of $(-\log(L))$ from 30 repetitions.

		Estimated values of $(-\log(L))$										
		Nine existing models									New models	
		$q = 2$			$q = 3$			$q = 4$			$q = 3$	$q = 2$
n	m	RP	NHPP-PL	GP	Kijima I	Kijima II	ARI_m	ARA_m	BBIP	Model II	ESI	MAI
15	5	45.33	42.40	43.08	41.65	41.17	40.82	41.20	42.18	40.60	40.82	41.00
		(5.07)	(4.47)	(4.21)	(4.25)	(4.35)	(4.21)	(4.29)	(4.21)	(4.23)	(4.25)	(4.28)
	15	28.73	24.79	26.72	24.18	23.96	24.15	24.25	24.32	23.14	23.84	24.02
		(4.21)	(3.38)	(3.23)	(3.57)	(3.68)	(3.72)	(3.72)	(3.23)	(3.65)	(3.77)	(3.67)
	25	23.16	19.35	21.62	18.32	18.33	18.36	18.31	18.96	17.53	18.01	18.24
		(3.95)	(3.58)	(3.04)	(3.84)	(3.74)	(3.83)	(3.80)	(3.47)	(3.96)	(3.89)	(3.87)
30	5	91.80	86.67	88.15	86.02	85.46	85.30	85.51	86.34	85.09	85.22	85.33
		(9.10)	(8.13)	(8.10)	(8.08)	(7.99)	(8.33)	(8.16)	(8.10)	(8.10)	(8.06)	(8.17)
	15	57.69	50.53	53.60	49.96	49.34	50.02	49.92	49.98	48.42	49.06	49.23
		(7.70)	(6.03)	(5.99)	(6.19)	(6.14)	(6.15)	(6.12)	(5.98)	(6.23)	(6.13)	(6.18)
	25	45.03	36.95	41.76	36.46	35.97	36.26	36.26	36.52	35.22	35.67	35.84
		(7.32)	(5.57)	(5.45)	(5.66)	(5.81)	(5.92)	(5.79)	(5.44)	(5.83)	(5.93)	(5.95)
45	5	136.12	127.52	130.17	126.61	125.76	125.60	125.71	127.20	125.15	125.41	125.56
		(13.97)	(11.82)	(12.19)	(11.69)	(11.63)	(11.75)	(11.66)	(11.84)	(11.31)	(11.68)	(11.62)
	15	84.94	74.17	79.01	73.65	73.12	73.99	73.52	73.49	72.15	72.75	72.96
		(11.66)	(7.87)	(8.77)	(7.95)	(8.75)	(8.24)	(7.93)	(7.86)	(7.73)	(8.06)	(8.02)
	25	64.92	52.24	59.85	51.72	51.31	51.69	51.62	51.68	50.57	51.09	51.30
		(12.03)	(6.77)	(9.13)	(6.94)	(7.07)	(7.20)	(7.11)	(6.71)	(7.00)	(7.02)	(7.09)

From Table (5), we have the following observations.

- The value $-\log(L)$ of Model II is the smallest among the 11 models in all the 9 cases, the value $-\log(L)$ of the ESI is the second smallest among the 11 models in all the 9 cases, and the value $-\log(L)$ of the MAI is the third smallest among the 11 models in 8 out of the 9 cases.
- One can also calculate the values of AIC, AIC_c and BIC of the 11 models, respectively, and compare them. It can be found that the MAI model has the smallest AIC, AIC_c and BIC among the 11 models in all the 9 cases. The ESI model has the second smallest AIC, AIC_c and BIC among the 11 models in all the 9 cases.
- Among all three models with $q = 2$, the MAI has the smallest value $-\log(L)$ in all the 9 cases.
- Among all six models with $q = 3$, the ESI has the smallest value $-\log(L)$ in all the 9 cases.

From the above observations, one can conclude that (1) Model II outperforms the other 10 models in term of the likelihood, (2) the MAI and ESI models outperform the other 9 models in terms of the AIC, AIC_c and BIC, (3) the MAI has the best performance among the models with $q = 2$, and (4) the ESI has the best performance among the models with $q = 3$.

3.2 Case study

In this section, 15 datasets collected from the real world are used to compare the performance of the 11 models. All the datasets are failure data of certain systems. Their names and sample sizes are shown in Table 6, where n is the sample size. The last column shows the model used on the dataset in the corresponding publication shown in the *Data source* column, i.e., column 4. Details of the datasets are explained below.

- The hydraulic system failure data. Datasets 1 and 6 in are times between failures of hydraulic systems, analysed by Kumar and Klefsjö (1992), who use the NHPP-PL to fit those datasets.
- The aircraft engine failure data. Datasets 7 and 10 in are times between failures of the aircraft engines, published in Proschan (1963).
- The compressor failure data. The failure data related to the U.S.S. Halfbeak No. 3 main propulsion motor, which are used by Ascher and Feingold (1984) to perform a time terminated analysis.
- Datasets 11 and 12. Those two datasets contain times-between-failures of a compressor and a main propulsion motor. Yanez et al. (2002) use Kijima I to fit the datasets.

- Datasets 13 and 14. Those two datasets are miles-between-failures of two bus engines, analysed by Guida and Pulcini (2009), who propose an NHPP model, called the BBIP, with intensity function $a_1e^{-t/a_2} + a_3(1 - e^{-t/a_4})$, where $a_1, a_2, a_3, a_4 > 0$.
- Dataset 15. Dataset 15 is times-between-failures of a diesel engine, analysed by Lee (1980), who proposes an NHPP model with the Weibull and log linear rate (NHPP-WLL): $a_1a_2t^{a_2-1}e^{a_3t}$, where $a_1, a_2 > 0$ and $-\infty < a_3 < \infty$.

Table 6: The real-world datasets.

No.	Dataset	n	Data source	Model
1	Hydraulic system (LHD 1)	23	Kumar and Klefsjö (1992)	NHPP-PL
2	Hydraulic system (LHD 3)	25	Kumar and Klefsjö (1992)	NHPP-PL
3	Hydraulic system (LHD 9)	27	Kumar and Klefsjö (1992)	NHPP-PL
4	Hydraulic system (LHD 11)	28	Kumar and Klefsjö (1992)	NHPP-PL
5	Hydraulic system (LHD 17)	26	Kumar and Klefsjö (1992)	NHPP-PL
6	Hydraulic system (LHD 20)	23	Kumar and Klefsjö (1992)	NHPP-PL
7	Air conditioner (TBF 7909)	24	Proschan (1963)	HPP
8	Air conditioner (TBF 7912)	30	Proschan (1963)	HPP
9	Air conditioner (TBF 7913)	27	Proschan (1963)	HPP
10	Air conditioner (TBF 7914)	23	Proschan (1963)	HPP
11	Compressor	24	Yanez et al. (2002)	Kijima I
12	Main propulsion motor	24	Yanez et al. (2002)	Kijima I
13	Powertrain System 510	55	Guida and Pulcini (2009)	BBIP
14	Powertrain System 514	35	Guida and Pulcini (2009)	BBIP
15*	Diesel engine	56	Lee (1980)	NHPP-WLL

* In dataset 15, there is a value 0, which is replaced with 0.5 in this paper.

We fit the 11 models on the 15 datasets, respectively. with the sole aim to compare their likelihood and the other performance measures: AIC, AIC_c and BIC. The results are shown in Table 7, in which the underlined number is the smallest one in each row.

3.2.1 The results

The last four rows in Table 7 show the average values resulted from the modelling outcomes on the 15 datasets. For example, in column 2 where RP is used, $-\log(L)=205.09$ (located in the 4th row from the bottom), which means that the average of the values $-\log(L)$ on the 15 datasets is 205.09. Similarly, the RP's average values of AIC, AIC_c and BIC on the 15 datasets are 414.17, 415.20, and 416.89, respectively.

From Table 7, we have the observations that are summarised in Table 8 and that are also concluded in the following.

Table 7: $-\log(L)$ of each model on the real-world datasets.

No.	Estimated value of $(-\log(L))$										
	Nine existing models									New models	
	$q = 2$		$q = 3$					$q = 4$		$q = 3$	$q = 2$
	RP	NHPP	GP	Kijima I	Kijima II	ARI _m	ARA _m	BBIP	Model II	ESI	MAI
1	129.99	128.50	129.50	128.46	128.50	128.44	128.46	129.09	<u>128.02</u>	128.32	129.38
2	148.72	146.96	148.72	146.96	145.47	145.31	145.32	146.22	144.94	<u>144.20</u>	144.94
3	166.55	163.64	165.39	163.52	163.65	163.48	163.49	164.36	163.96	<u>163.45</u>	163.97
4	158.05	157.09	157.99	157.09	155.89	155.54	155.86	156.23	155.21	<u>154.91</u>	155.44
5	151.20	149.33	150.96	149.32	149.12	148.38	<u>148.36</u>	149.81	148.98	148.87	149.14
6	137.27	136.86	137.12	136.73	135.65	135.55	135.65	136.61	<u>134.77</u>	135.29	135.80
7	125.37	126.30	<u>124.48</u>	125.37	125.37	125.05	125.37	125.96	125.37	125.37	125.37
8	151.94	150.43	151.14	150.41	150.42	150.37	150.42	150.64	<u>150.20</u>	150.24	151.33
9	143.96	144.22	143.10	143.96	143.10	<u>141.19</u>	142.75	143.81	141.80	142.07	142.15
10	119.60	119.66	119.21	119.52	119.56	118.68	119.57	119.47	<u>117.98</u>	118.58	119.55
11	191.06	189.30	190.95	189.32	188.90	188.70	<u>187.82</u>	189.22	188.78	188.12	188.85
12	183.88	182.44	182.70	181.63	182.45	<u>181.55</u>	181.85	183.29	182.37	182.45	182.91
13	543.26	543.57	543.19	542.28	<u>541.88</u>	543.32	543.58	542.52	542.35	542.35	542.87
14	356.18	357.05	356.06	355.88	355.07	354.55	357.06	<u>353.77</u>	354.74	353.98	355.04
15	369.29	368.31	369.14	368.31	368.00	<u>367.06</u>	367.68	367.70	367.89	367.75	368.07
$-\log(L)^*$	205.09	204.24	204.64	203.92	203.53	203.14	203.55	203.91	203.16	<u>203.06</u>	203.65
AIC*	414.17	412.48	415.29	413.83	413.07	412.29	413.10	415.83	414.32	412.13	<u>411.31</u>
AIC _c *	415.20	413.50	417.07	415.62	414.85	414.08	414.88	418.64	417.13	413.91	<u>412.33</u>
BIC*	416.89	415.19	419.36	417.90	417.14	416.36	417.16	421.25	419.74	416.20	<u>414.02</u>

* The value with * on its right upper corner represents the mean of the value.

In Table 8, the last row shows the best model with its frequency in each column. For example, 4×Model II in column 2 means that Model II is the best performed model as it has the smallest $-\log(L)$ in 4 cases, whereas the other models have smaller frequencies of performing as the best.

- Compared the values of the likelihood, Model II outperforms the others as it has the smallest $-\log(L)$, or the largest $\log(L)$. Both the ESI and the ARI_m models have the second largest value of the likelihood, as shown in the second column.
- If one simply compares the values of the likelihood among those model with 3 parameters, i.e., $q = 3$, then he/she can find that the ESI has the largest likelihood in 8 out of the 15 cases. That is, the ESI outperforms the other 3-parameter models in terms of the likelihood, as shown in the third column.
- If one compares the AIC values of the 11 models, then he/she can find that the ESI has the smallest AIC value in 10 out of the 15 cases. That is, the MAI outperforms the other 10 models in term of AIC, as shown in the fourth column.

- If one compares the values of the AIC_c and BIC of the 11 models, then he/she can find that the MAI has the smallest values of the AIC_c and BIC in 11 out of the 15 cases. That is, the MAI outperforms the other models in terms of both AIC_c and BIC, as shown in the fifth column.

It can also be found, surprisingly, the Kijima models including both Kijima I and Kijima II, which are widely studied and applied in the literature, do not perform so well as either ARA_m or ARI_m .

Table 8: Results of the performance comparison from Table 7.

No.	ESI		MAI	
	$-\log(L)$	$-\log(L)$ of the "q = 3" models	AIC	AIC_c & BIC
1	Model II	ESI	NHPP-PL	NHPP-PL
2	ESI	ESI	MAI	MAI
3	ESI	ESI	NHPP-PL	NHPP-PL
4	ESI	ESI	MAI	MAI
5	ARA_m	ARA_m	MAI	MAI
6	Model II	ESI	MAI	MAI
7	GP	GP	MAI	MAI
8	Model II	ESI	NHPP-PL	NHPP-PL
9	ARI_m	ARI_m	MAI	MAI
10	Model II	ESI	MAI	MAI
11	ARA_m	ARA_m	MAI	MAI
12	ARI_m	ARI_m	NHPP-PL	NHPP-PL
13	Kijima II	Kijima II	MAI	MAI
14	BBIP	ESI	ESI	MAI
15	ARI_m	ARI_m	MAI	MAI
Frequency	4×Model II	8×ESI	10×MAI	11×MAI

3.2.2 Model checking

According to Cook and Lawless (2007) (p93), the martingale residual process is defined as

$$\hat{\epsilon}(t) = \bar{N}(t) - \int_0^t \lambda(u|\mathcal{H}_{u-}) du, \quad (26)$$

where $\bar{N}(t)$ is the observable counting process. Plugging $\lambda(t|\mathcal{H}_{t-})$ in (3) into (26), one can obtain

$$\hat{\epsilon}(t) = \begin{cases} \bar{N}(t) - \frac{1}{m} \sum_{k=0}^{m-1} \rho^{m-k-1} \Lambda_0(t), & \text{if } N_t = 0, \\ \bar{N}(t) - \frac{1}{m} \left(\sum_{k=0}^{N_t-1} \rho^{m-k-1} \Lambda_0(t - T_{N_t-k}) + \sum_{k=N_t}^{m-1} \rho^{m-k-1} \Lambda_0(t) \right), & \text{if } 1 \leq N_t < m, \\ \bar{N}(t) - \frac{1}{m} \sum_{k=0}^{m-1} \rho^{m-k-1} \Lambda_0(t - T_{N_t-k}), & \text{if } N_t \geq m, \end{cases} \quad (27)$$

where $\Lambda_0(t) = \int_0^t \lambda_0(u) du$.

The process $\{\epsilon(t), t \geq 0\}$ should follow the normal distribution and have uncorrelated increments if the failure process model with intensity $\lambda(u|\mathcal{H}_{u-})$ is correctly specified. We take the MAI model on dataset 2 as an example in model checking.

With the Cramer-von Mises test, we test the normality of the residuals $\epsilon(t)$ of the MAI model on dataset 2 and have the result that the p-value of the test is greater than a significance level of 0.05, therefore we fail to reject the null hypothesis that the residuals is normally distributed. We also test the serial correlation of the residuals with the Breusch-Godfrey test and fail to reject the null hypothesis that there is no serial correlation at a significance level of 0.05.

To further check the MAI model, we compare the performance of models RP, NHPP and MAI on dataset 2. The parameters of the MAI on this dataset are shown in row 2 in Table 9. Row 3 in this table shows the estimated root-mean-square error (RMSE) of the difference (i.e., the residuals $\hat{\epsilon}(t)$ in Eq. (26)) between the mean cumulative function (MCF) of the real-world data and the expected cumulative number of failures of the models RP, NHPP and MAI, respectively, from which one can see that the MAI has the smallest RMSE among the three models. We also plot the mean cumulative function of the real-world data (i.e, empirical MCF) and the expected cumulative number of failures of the models RP, NHPP and MAI, as shown in Figure 3. It can be seen that the MAI fits better than the other two models.

The reason that we compare the ESI and MAI models with other nine models is merely to show the performance of the ESI and MAI models. It should be noted, however, using the NHPP model to approximate the failure process of a series system in which a failed component is replaced may violate the assumption of NHPP. This is because NHPP can be applied to model the failure process of a system on which only minimal repair is conducted. In addition, both the renewal process and the geometric process assume that the gap times $T_k - T_{k-1}$ are statistically independent, which may not hold in the setting in this paper because only a failed component is replaced and the other components in the system are remained.

Table 9: The estimated parameters of the RP, NHPP-PL and MAI on dataset 2.

Models	RP	NHPP-PL	MAI
Parameters	$\alpha = 6.64 \times 10^{-3}, \beta = 1.012$	$\alpha = 1.27 \times 10^{-4}, \beta = 1.49$	$\alpha = 4.58^{-6}, \beta = 1.98, m = 18$
Estimated RMSE	1.62	1.46	1.07

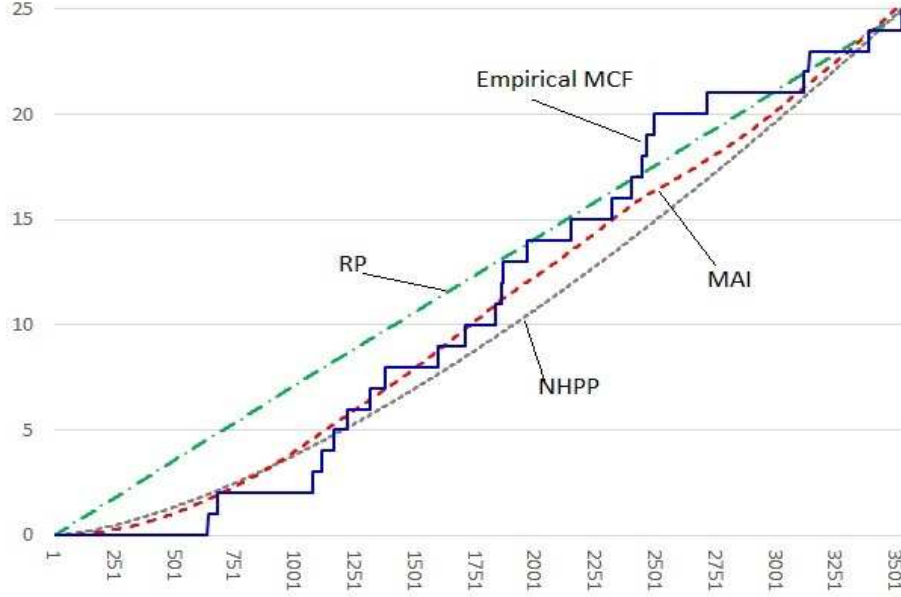


Figure 3: Cumulative air-conditioning failures (empirical MCF on dataset 2) and estimated mean cumulative functions.

4 Discussion

In this section, we discuss possible extensions and applications of the proposed models.

If the intensity function $\lambda_0(t)$ of the GP, ARI_m , ARA_m , Kijima I and Kijima II are the same, then they should have the same number of parameters. As such, to compare the performance of those models, one may simply compare their log-likelihood values and does not need to compare their AIC, AIC_c and BIC values. From the second and third columns of Table 8, one can find that the model ARI_m is the second best model. As such, the ARI_m may be used as the basis intensity function in the MAI model and the failure intensity of the system after the N_t -th failure is given by

$$\lambda(t|\mathcal{H}_{t-}) = \begin{cases} \lambda_0(t), & \text{if } N_t = 0, \\ \frac{1}{m} \left(\sum_{k=0}^{N_t-1} \lambda_{\text{ARI}_m}(t - T_{N_t-k}) + (m - N_t) \lambda_{\text{ARI}_m}(t) \right), & \text{if } 1 \leq N_t < m, \\ \frac{1}{m} \sum_{k=0}^{m-1} \lambda_{\text{ARI}_m}(t - T_{N_t-k}), & \text{if } N_t \geq m. \end{cases} \quad (28)$$

where $\lambda_{\text{ARI}_m}(t) = \lambda_0(t) - \rho \sum_{j=0}^{\text{Min}(m-1, N_t-1)} (1 - \rho)^j \lambda_0(T_{N_t-j})$, see (Doyen & Gaudoin, 2004) for details about the ARI_m model. $\lambda_0(t)$ in Eq. (28) is an increasing function.

The above model has the same number of parameters as that of the ARI_m , which may be interesting

and will be investigated in our future work.

The ESI and MAI models in the preceding section depict the failure process of a multi-component system. A possible application in maintenance is on development of maintenance policy. Since the models are not able to exactly pinpoint which component's failure causes the system to fail, an inspection may be executed on possible failures of the system to locate the failure cause. As such, the ESI and MAI models may be useful in development of inspection policies.

Other applications of the ESI and MAI models may be in warranty claim analysis, in which stochastic process models are importance tools to make forecasting. One may estimate a stochastic process model such as an ESI model or an MAI model based on historical warranty claim data and then use the model to project the claim amount in future. In such a scenario, models ESI and MAI offers modelling methods that make a more precise forecasting. Other examples may include lifecycle costing that also relies on stochastic process modelling methods with better performance to make more precise forecasting.

5 Conclusions and future work

This paper proposes a new failure process model, ESI (Exponential Smoothing of Intensity model), and investigates its special case model, MAI (Moving Average of Intensity model). The models are developed for modelling the failure process of a repairable multi-component system.

According to the comparison between the two models (ESI and MAI) and nine existing models on artificially generated data, with respect to the model performance metrics AIC, AIC_c and BIC, the MAI outperforms the other ten models and the ESI is the second best model. Based on the comparison of the eleven models on the fifteen real-world datasets, the MAI outperforms the other ten models in eleven out of the fifteen cases. If one uses the log-likelihood as a model performance measure, then the ESI model is the second best performed model among the eleven models on both the artificially generated data and the fifteen real-world datasets.

More importantly, the MAI model has the same number of parameters as the NHPP if both follow the same life-time distribution before the first failure. Additionally, it is worth highlighting that the MAI outperforms the NHPP on the artificially generated data and most cases of the real world datasets.

It is also worth highlighting that the MAI model can model perfect, imperfect and minimal maintenance and has the ordinary renewal process and the NHPP as special cases.

Our future work aims at investigating the probabilistic and statistical properties of the ESI and MAI models. Another area of our future work is to provide an in-depth analysis of the performance

of the models proposed in this paper, in a similar vein as those papers by Azaiez and Bier (1995, 1996).

Acknowledgment

The authors are indebted to the three reviewers for their suggestions for improving the clarity of the presentation.

References

- Ascher, H., & Feingold, H. (1984). *Repairable systems reliability modeling, inference, misconceptions and their causes*. New York: Marcel-Dekker.
- Azaiez, M. N., & Bier, V. M. (1995). Perfect aggregation for a class of general reliability models with bayesian updating. *Applied Mathematics and Computation*, 73(2-3), 281–302.
- Azaiez, M. N., & Bier, V. M. (1996). Aggregation error in bayesian analysis of reliability systems. *Management science*, 42(4), 516–528.
- Bordes, L., & Mercier, S. (2013). Extended geometric processes: Semiparametric estimation and application to reliability. *Journal of the Iranian Statistical Society*, 12(1), 1-34.
- Burnham, K. P., & Anderson, D. (2002). *Model selection and multi-model inference: A practical informatio-theoric approach*. Springer Science and Business Media.
- Cha, J. H., Finkelstein, M., & Levitin, G. (2018). Bivariate preventive maintenance of systems with lifetimes dependent on a random shock process. *European Journal of Operational Research*, 266(1), 122–134.
- Cook, R. J., & Lawless, J. (2007). *The statistical analysis of recurrent events*. Springer Science & Business Media.
- Cox, D., & Lewis, P. (1966). *The statistical analysis of series of events*. London: John Wiley and Sons.
- Doyen, L., & Gaudoin, O. (2004). Classes of imperfect repair models based on reduction of failure intensity or virtual age. *Reliability Engineering & System Safety*, 84(1), 45–56.
- Doyen, L., Gaudoin, O., & Syamsundar, A. (2017). On geometric reduction of age or intensity models for imperfect maintenance. *Reliability Engineering and System Safety*, 168, 40-52.
- Drenick, R. F. (1960). The failure law of complex equipment. *Journal of the Society for Industrial & Applied Mathematics*, 8(4), 680–690.

- Guida, M., & Pulcini, G. (2009). Reliability analysis of mechanical systems with bounded and bathtub shaped intensity function. *IEEE Transactions on Reliability*, 58(3), 432–443.
- Hoyland, A., & Rausand, M. (2004). *System reliability theory: models, statistical methods, and applications*. NJ: Wiley-Interscience.
- Huang, Y.-S., Huang, C.-D., & Ho, J.-W. (2017). A customized two-dimensional extended warranty with preventive maintenance. *European Journal of Operational Research*, 257(3), 971–978.
- Ibragimov, R. (2009). Copula-based characterizations for higher order markov processes. *Econometric Theory*, 25(3), 819–846.
- Kijima, M. (1989). Some results for repairable systems with general repair. *Journal of Applied Probability*, 26(1), 89–102.
- Kumar, U., & Klefsjö, B. (1992). Reliability analysis of hydraulic systems of lhd machines using the power law process model. *Reliability Engineering & System Safety*, 35(3), 217–224.
- Lam, Y. (1988). Geometric processes and replacement problem. *Acta Mathematicae Applicatae Sinica*, 4, 366–377.
- Lam, Y. (2007). *The geometric process and its applications*. Singapore: World Scientific.
- Lee, L. (1980). Testing adequacy of the weibull and log linear rate models for a poisson process. *Technometrics*, 22(2), 195–199.
- Louit, D. M., Pascual, R., & Jardine, A. K. (2009). 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), 1618–1628.
- Makridakis, S., Wheelwright, S. C., & Hyndman, R. J. (2008). *Forecasting methods and applications*. John Wiley & Sons.
- Olde Keizer, M., Teunter, R., & Veldman, J. (2017). Joint condition-based maintenance and inventory optimization for systems with multiple components. *European Journal of Operational Research*, 257(1), 209–222.
- Proschan, F. (1963). Theoretical explanation of observed decreasing failure rate. *Technometrics*, 5(3), 375–383.
- Wu, S. (2018). Doubly geometric processes and applications. *Journal of the Operational Research Society*, 69(1), 66–77.
- Wu, S., & Scarf, P. (2017). Two new stochastic models of the failure process of a series system. *European Journal of Operational Research*, 257(3), 763–772.
- Yanez, M., Joglar, F., & Modarres, M. (2002). Generalized renewal process for analysis of repairable systems with limited failure experience. *Reliability Engineering & System Safety*, 77(2), 167–180.

- Yang, L., Ye, Z.-S., Lee, C.-G., Yang, S.-F., & Peng, R. (2018). A two-phase preventive maintenance policy considering imperfect repair and postponed replacement. *European Journal of Operational Research*. doi: <https://doi.org/10.1016/j.ejor.2018.10.049>
- Zhang, W., Tian, Y., Escobar, L. A., & Meeker, W. Q. (2017). Estimating a parametric component lifetime distribution from a collection of superimposed renewal processes. *Technometrics*, *59*(2), 202–214.