Properties of the Weibull-ARA\(_\infty\) virtual age model and application in maintenance policy optimization

Yann Dijoux, Mitra Fouladirad, Tuan Nguyen

AMMSI Troyes 2015
1. Context

2. Modelling the maintenance process

3. The Weibull-ARA$_\infty$ model

4. Inference on the WARA$_\infty$ model on an observation window

5. Optimal preventive maintenance strategy

6. Conclusion
Context

1. Context
2. Modelling the maintenance process
3. The Weibull-ARA\(_\infty\) model
4. Inference on the WARA\(_\infty\) model on an observation window
5. Optimal preventive maintenance strategy
6. Conclusion
Context

• Complex industrial systems subjected to corrective maintenances (CM, repair), carried out after a failure.

• Maintenance are imperfect. Virtual age models are employed to characterize the general wear-out of the systems.
Motivations

- Adjusting an optimal preventive maintenance policy (periodic or dynamic) is rarely addressed in the literature considering imperfect maintenance only.

- The systems are not necessarily new at the beginning of the observations.
  - Second-hand unit.
  - Previous maintenances times are not recorded, observations are missing.
  - Consistency of the observations (new maintenance policy, systems after a burn-in period)

→ To develop theoretical properties of a classical virtual age model.

→ To develop inference procedures when the initial age of the system is unknown.

→ To present an optimal preventive maintenance strategy.
Modelling the maintenance process

1. Context
2. Modelling the maintenance process
3. The Weibull-$\text{ARA}_\infty$ model
4. Inference on the $\text{WARA}_\infty$ model on an observation window
5. Optimal preventive maintenance strategy
6. Conclusion
Modelling the maintenance process

- Failure times: $\{T_i\}_{i \geq 1}$
- Inter-failure times: $X_i = T_i - T_{i-1}$, $i \geq 1$
- Counting failure process: $\{N_t\}_{t \geq 0}$, $N_t =$ number of failures occurred at time $t$
- Durations of repair are not taken into account.
- Two failures cannot occur at the same time.
Stochastic modelling

→ Considering $H_{t-}$ as the history of the failure process up to time $t$, the failure intensity $\lambda_t$ is defined as:

$$\lambda_t = \lim_{dt \rightarrow 0} \frac{1}{dt} P(N_{t+dt} - N_{t-} = 1 | H_{t-})$$

→ For a self-excited point process: $H_{t-} = \sigma \left( \{N_s\}_{0 \leq s < t} \right)$ and $\lambda_t$ completely defines the failure process.

→ Before the first failure, the failure intensity is a deterministic and continuous function of time $\lambda(t)$, called initial intensity, the failure rate of $T_1$.

Considering industrial or software systems, a Weibull distribution is frequently used.

$$\lambda(t) = \alpha \beta t^{\beta-1} , \; \alpha > 0 , \; \beta > 0$$
Classical models

**Minimal Repair or As Bad As Old model (ABAO)**
- Each maintenance leaves the system in the same state as it was before failure.
- The failure process is a Non Homogeneous Poisson Process (NHPP).

\[ \lambda_t = \lambda(t) \]

**Perfect repair or As Good As New model (AGAN)**
- Each maintenance perfectly repairs the system and leaves it as if it were new.
- The failure process is a Renewal Process (RP).

\[ \lambda_t = \lambda(t - T_{N_t^-}) \]

Reality is between the case ABAO and AGAN.
Virtual age models

After the \( i^{th} \) repair, the system performs as a new one having survived until \( A_i \).

\[
\forall i \geq 0, \quad \forall t \geq 0 P(X_{i+1} > t | X_1, \ldots, X_i, A_i) = P(Y > A_i + t | Y > A_i) = \frac{S(A_i + t)}{S(A_i)}
\]

where \( Y \) has the same distribution as \( X_1 \).

\[
\lambda_t = \lambda(t - T_{N_t^-} + A_{N_t^-})
\]

The \( A_i \) are called the **effective ages**. \( A_0 = 0 \).

- **ABAO** : \( A_i = T_i \)
- **AGAN** : \( A_i = 0 \)

The Brown-Proshchan model: repairs are either perfect (AGAN) with a probability \( \rho \), or minimal (ABAO) with a probability \( 1-\rho \).

- The Arithmetic Reduction of Age model with memory 1 (\( ARA_1 \)):
  \[
  A_i = (1 - \rho) T_i
  \]

\( \rightarrow \) A virtual age model is characterized by the initial intensity and by the evolution of the effective ages.
Failure intensities of virtual age models

Figure: BP model and a Weibull initial intensity 
\( (\alpha = 0.001, \beta = 3, \rho = 0.5) \)

Figure: ARA_1 model and a Weibull initial intensity 
\( (\alpha = 0.001, \beta = 3, \rho = 0.5) \)

\( \rightarrow \rho \in [0, 1] \) describes the maintenance efficiency.

\( \rightarrow \) Exponential distributions are not adapted as initial intensities (no ageing).
The $ARA_\infty$ assumption

$\rightarrow$ Arithmetic Reduction of Age model with infinite memory.


Assumption: The age of a system after maintenance is proportional to its age just before maintenance.

$$A_i = (1 - \rho)(A_{i-1} + X_i)$$

$$A_i = (1 - \rho)^i A_0 + \sum_{j=1}^{i} (1 - \rho)^{i+1-j} X_j$$

$\rightarrow$ Existence of a potential stationary regime of the process (Last and Szekli (1998)).
The Weibull-$ARA_\infty$ model

1. Context
2. Modelling the maintenance process
3. The Weibull-$ARA_\infty$ model
4. Inference on the $WARA_\infty$ model on an observation window
5. Optimal preventive maintenance strategy
6. Conclusion
Definition and simulation

We consider a Weibull initial intensity $\lambda(t) = \alpha \beta t^{\beta-1}$ and the $ARA_\infty$ assumption.

The corresponding model is denoted Weibull-$ARA_\infty$ model or $WARA_\infty$ model.

Simulation of effective ages under $ARA_\infty$ assumption

$$A_{i+1} = (1 - \rho) \Lambda^{-1} (\Lambda(A_i) + \xi_{i+1})$$

where $\Lambda$ is the cumulative initial intensity and $\Lambda^{-1}$ its inverse, and where $\xi_{i+1}$ is an exponential r.v. $\mathcal{E}(1)$ independent of $\{\xi_j\}_{j=1..i}$

$\Lambda(t) = \alpha t^\beta$ considering a Weibull initial intensity.
Effective ages for the $\text{WARA}_\infty$ model

Proposition 1: General expression of $A_n$

$$A_n = (1 - \rho) \alpha^{-\frac{1}{\beta}} \left( \sum_{i=1}^{n} (1 - \rho)^{\beta(n-i)} \xi_i \right)^{\frac{1}{\beta}}$$

with $\{\xi_j\}_{j=1}^n$ sample of exponential distribution of parameter 1.

$\rightarrow$ Proof by induction.

$\rightarrow$ A series of exponential distributions with different parameters follow an hypoexponential distribution.

**Notation** $q = (1 - \rho)^\beta$

**Notation** $q$-Pochhammer series:

$$(x, x)_k = \prod_{i=1}^{k} (1 - x^i), x \in \mathbb{R}, k \in \mathbb{N}$$
Effective age distributions for the $\text{WARA}_\infty$ model

**Proposition 2: Survival function of $A_n$**

$$R_{A_n}(t) = \sum_{k=1}^{n} \frac{1}{(q, q)_{n-k}(\frac{1}{q}, \frac{1}{q})_{k-1}} e^{-\frac{\alpha k}{q k}}$$

$\rightarrow$ “Hypo-Weibull” distribution.

$\rightarrow$ Given the effective $A_n$, the distribution of the next inter-failure time can be determined.

**Proposition 3: (Marginal) Survival function of $X_{n+1}$**

$$R_{X_{n+1}}(t) = \sum_{k=1}^{n} \frac{\alpha \beta}{q^k (q, q)_{n-k}(\frac{1}{q}, \frac{1}{q})_{k-1}} \int_0^{\infty} x^{\beta-1} e^{-\alpha(x+t)} dx + \alpha(1-q^{-k})x^\beta$$
Effective age distributions for the $\text{WARA}_\infty$ model

Proposition 4: Limiting survival distribution of $A_n$

$$R_{\text{A}_\infty}(t) = \sum_{k=1}^{\infty} \frac{1}{(q, q)_\infty (\frac{1}{q}, \frac{1}{q})_{k-1}} e^{-\frac{\alpha t}{q^k}}$$

→ Explicit expression of $R_{X_\infty}$.

Proposition 5: Expected value of the age

$$E[A_n] = \alpha^{-1/\beta} \Gamma(\frac{1}{\beta} + 1) \sum_{k=1}^{n} \frac{q^k}{(q, q)_{n-k}(\frac{1}{q}, \frac{1}{q})_{k-1}}$$

→ Derive $E[X_{n+1}], E[A_\infty], E[X_\infty]$

→ Derive the distribution of the age of the system just before the failure and its expected value in the transient and steady regime: $A_n^-, E[A_n^-], A_\infty^-, E[A_\infty^-]$. 

17 / 37
Inference on the WARA∞ model on an observation window

1. Context
2. Modelling the maintenance process
3. The Weibull-ARA∞ model
4. Inference on the WARA∞ model on an observation window
5. Optimal preventive maintenance strategy
6. Conclusion
The actual observations

The process is recorded on an observation window $[s, s + t]$ and no information on the failure process is available prior to $s$.

No model associated to imperfect maintenance on an observation window has been developed.

Under the $\text{ARA}_\infty$ assumption, the only necessary information to derive the likelihood function is the initial virtual age $a_0$.

$$
\mathcal{L}_{s,t,a_0}(x_1, \ldots, x_n) = \prod_{i=1}^{n} \lambda(a_{i-1} + x_i) \times \exp(-\sum_{i=1}^{n+1} [\Lambda(a_{i-1} + x_i) - \Lambda(a_{i-1})])
$$

with $a_i = (1 - \rho)^i a_0 + \sum_{j=1}^{i} (1 - \rho)^{i+1-j} x_j$
Choice of initial age

→ In the literature, the initial age $a_0$ is assumed to be 0 (new system).

→ This assumption is relatively valid for a large dataset (renewal aspects of the $\text{WARA}_\infty$ model).

→ For a small dataset, the actual ageing of the system should be taken into account.
Steady regime assumption

Figure: Survival functions of $A_n$ with $(\alpha = 1, \beta = 2, \rho = 0.2)$

→ If it is likely that few maintenances have occurred, it is realistic to assume that the system is already in its stationary state.

→ The first effective age is assumed to follow the distribution of $A_\infty$. 
Proposition 6: Likelihood function under steady regime

\[ L^\infty(t_1, t_2, \ldots, t_n, t) = \]

\[ = - \int_{(1-\rho)t_1}^{\infty} \prod_{i=1}^{n} \lambda(a_{i-1} + x_i) \exp \left( - \sum_{i=1}^{n+1} \Lambda(a_{i-1} + x_i) - \Lambda(a_{i-1}) \right) dR_{A_{\infty}}(h) \]

with \( a_i = (1 - \rho)^i h + \sum_{j=1}^{i} (1 - \rho)^{i+1-j} x_j \).

\[ \longrightarrow \text{No explicit expression of the ML estimators}. \]
Simulations

Configuration (Example)

- $\alpha = 1$, $\beta = 4.5$, $\rho = 0.2$, $n \in \{10, 20, 30, 50\}$.
- The failure times $(t_1, \ldots, t_n) \in [s, s + t]$ are generated so that there are an average of 100 failures before $s$.

Objective: Assess the surplus value of the new model

- First model (Simulated): Assume that the system is in steady regime.
- Second model: Assume that the system is As Good As New at the beginning of the observations.

$\rightarrow$ Comparison criteria: Bias and MSE.
Results

Figure: Bias and MSE with $A_0 \sim A_\infty$ (plain) and $A_0 = 0$ (dashed) of $\alpha$, $\beta$ and $\rho$
Analysis

• As the number of observations increases, the empirical bias and MSE decrease to 0.

• The estimation from the first model is always more efficient than the second model ($A_0 = 0$).

• For small $\rho$, the added value of the new modelling is significant.

• As $\rho$ tends to 1, results in terms of MSE and Bias become similar for both models.
Optimal preventive maintenance strategy

1. Context
2. Modelling the maintenance process
3. The Weibull-$ARA_{\infty}$ model
4. Inference on the $WARA_{\infty}$ model on an observation window
5. Optimal preventive maintenance strategy
6. Conclusion
Context

A repairable system is observed with the following assumptions on the history of the process:

- The ageing and the maintenance efficiency are consistent with the $\text{WARA}_\infty$ model.

- Sufficient maintenances have occurred in the past of the process so that the system is assumed to be in its stationary regime.

- Parameters of the model are assumed to be known.

→ A planned preventive maintenance (PM) policy is established on the system for maintenance cost reduction.
Assumptions

- After a maintenance (corrective or preventive) has restored the system, a new PM is scheduled after a duration $a$.

- If a failure is observed before a duration $a$, a corrective maintenance is carried out with cost $C_c$.

- Otherwise a PM is carried out with cost $C_p$ ($C_p < C_c$).

- The PM efficiency and the CM efficiency are identical: $\text{ARA}_\infty$ with same parameter $\rho$.

$\rightarrow$ Multiple strategies on the choice of the "age-based" duration $a$ have been investigated.
First policy: Age-dependent policy

→ The duration \( a \) is constant during the whole process.

→ The renewal aspects of the WARA\(_\infty\) model ensure that this strategy makes sense and that the long-run average cost per unit of time for an infinite horizon is finite.

\[
C(a) = \lim_{t \to \infty} \frac{C_a(t)}{t} < \infty
\]

→ Two choices for \( a \) have been studied:

- The optimal age \( a^* \) minimizing the cost function \( C(a) \) is obtained by Monte Carlo simulations.

- As the process has similarities with a Renewal Process with generic distribution \( X_\infty \), it is possible to approach the optimal solution by optimizing the classical age-based maintenance strategy and to obtain an age \( \hat{a} \).

\[
\hat{a} = \arg \min_{a} \frac{C_p + (C_c - C_p)(1 - R_{X_\infty}(a))}{\int_0^a R_{X_\infty}(u)du}
\]
Evolution of the cost for a static policy

Figure: $\alpha = 1, \beta = 4.5, C_c = 10C_p$
Second policy: Dynamic policy

The duration is adaptative and depends on the past of maintenance process.
Dynamic policy (II)

→ At the beginning, the initial age follows the limiting distribution $A_{\infty}$.

→ Given the initial age $u$, the age after the $i$th maintenance is

$$A_i(u) = (1 - \rho)^i u + \sum_{j=1}^{i} (1 - \rho)^{i-j+1} x_j$$

→ The distribution of the next inter-failure time can be computed.

$$R_{Z_{i+1}}(z) = -\int_{0}^{\infty} \frac{e^{-\alpha(A_i(u)+z)^{\beta}}}{e^{-\alpha A_i(u)^{\beta}}} dR_{A_{\infty}}(u)$$

→ The optimal PM should be carried out after a duration $a_{i+1}^*$.

$$a_{i+1}^* = \arg \min_{a} \frac{C_p + (C_c - C_p)(1 - R_{Z_{i+1}}(a))}{\int_{0}^{a} R_{Z_{i+1}}(u) du}$$
Third policy: Failure limit policy

→ After a maintenance (corrective or preventive) has restored the system, a new PM is scheduled when the virtual age of the system exceeds a threshold $A$.

→ "Virtual age limit" policy.
Comparing the costs

Table: Optimal maintenance strategies (\(\alpha = 1, C_c = 10C_p\))

<table>
<thead>
<tr>
<th>(\beta)</th>
<th>(\rho)</th>
<th>Static age</th>
<th>Failure limit</th>
<th>Variant I age</th>
<th>Dynamic</th>
<th>no PM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.5</td>
<td>0.2</td>
<td>17.72 0.18</td>
<td>17.69</td>
<td>20.81 0.84</td>
<td>20.76</td>
<td>22.01</td>
</tr>
<tr>
<td>1.5</td>
<td>0.5</td>
<td>12.30 0.26</td>
<td>12.28</td>
<td>13.51 0.62</td>
<td>13.11</td>
<td>15.68</td>
</tr>
<tr>
<td>1.5</td>
<td>0.8</td>
<td>9.71 0.33</td>
<td>9.71</td>
<td>9.97 0.48</td>
<td>9.72</td>
<td>12.74</td>
</tr>
<tr>
<td>3</td>
<td>0.2</td>
<td>15.50 0.09</td>
<td>15.48</td>
<td>25.00 0.23</td>
<td>23.64</td>
<td>40.70</td>
</tr>
<tr>
<td>3</td>
<td>0.5</td>
<td>7.54 0.20</td>
<td>7.54</td>
<td>8.02 0.26</td>
<td>7.62</td>
<td>20.72</td>
</tr>
<tr>
<td>3</td>
<td>0.8</td>
<td>4.92 0.30</td>
<td>4.92</td>
<td>4.92 0.30</td>
<td>4.92</td>
<td>13.90</td>
</tr>
<tr>
<td>4.5</td>
<td>0.2</td>
<td>12.51 0.10</td>
<td>12.51</td>
<td>16.51 0.15</td>
<td>16.10</td>
<td>46.87</td>
</tr>
<tr>
<td>4.5</td>
<td>0.5</td>
<td>5.49 0.23</td>
<td>5.48</td>
<td>5.51 0.22</td>
<td>5.49</td>
<td>21.39</td>
</tr>
<tr>
<td>4.5</td>
<td>0.8</td>
<td>3.46 0.37</td>
<td>3.46</td>
<td>3.49 0.34</td>
<td>3.46</td>
<td>13.68</td>
</tr>
</tbody>
</table>
• The failure limit (virtual age limit) policy (III) is the most efficient.
• The age-dependent policy (I) is almost as powerful.
• In practice, the age-dependent policy seems simpler to implement than the failure limit policy.
• The dynamic policy (II) is locally optimal, but is outperformed by the previous static policies.
• The approximation of the age-dependent policy (I’) offers decent results.
• As this approximation does not take into account the dependency between consecutive inter-failure times, its validity is poor with small $\rho$. 
Conclusion

1. Context
2. Modelling the maintenance process
3. The Weibull-ARA∞ model
4. Inference on the WARA∞ model on an observation window
5. Optimal preventive maintenance strategy
6. Conclusion
Future work

- Add estimation procedures within the optimal PM strategy.
- Dissociate the maintenance efficiencies.
- Take into account the downtime costs.
- Application to real dataset
- Develop goodness of fit procedures for the $\text{WARA}_\infty$ model.