

Rapport de projet OS13

Analyse de politique de maintenance

TRAN QUOC NHAT HAN & ADRIEN WARTELLE

13 janvier 2019

Sommaire

1	Maintenance basée sur l'âge	1
1.1	Rappel	1
1.2	Modéliser la durée de vie du système	2
1.3	La politique de maintenance basée sur l'âge	6
2	Maintenance basée sur dégradation	8
2.1	Rappel	8
2.2	Modéliser la dégradation du système	8
2.3	La politique de maintenance basée sur dégradation	10
3	Annexe	16
3.1	L'importation de données de pannes	16
3.2	Le premier histogramme de distribution de pannes	16
3.3	Estimer le mixage de la loi Exponentielle et Gamma	16
3.4	Optimiser le coût moyen sur une durée de temps	18
3.5	Importer les valeurs de dégradation	18
3.6	Premiers traces de dégradation	19
3.7	Estimation de paramètres de dégradations	19
3.8	Calcul analytique de maintenance préventive	20
3.9	Optimisation de maintenance basée sur dégradations	22

Résumé

A partir de données liées au fonctionnement d'un système, nous allons déterminer un modèle approprié, lié à l'état de celui-ci : durée de vie ou niveau de défaillance. Nous allons ensuite, grâce au modèle, optimiser une politique de maintenance : basée sur l'âge ou conditionnelle.

1 Maintenance basée sur l'âge

1.1 Rappel

Considérons un système non maintenu. En l'observant, nous obtenons une liste des dates de panne, grâce auxquelles nous pouvons construire une politique de remplacement systématique basée sur l'âge : *Nous remettons à neuf le système lorsqu'il tombe en panne ou après une durée t_0 si il a survécu jusque là.*

Le but est de minimiser le coût moyen cumulé.

$$\mathbb{E}(C) = \frac{\mathbb{E}(C(S))}{\mathbb{E}(S)} \quad (1)$$

Où S est la variable aléatoire représentant la date de remplacement et $C(S)$ est le coût de maintenance cumulé à l'instant S (sachant que $C(S)$ vaut $c_c(= 1200)$ lors d'une maintenance corrective et $c_p(= 800)$ lors d'une maintenance préventive).

1.2 Modéliser la durée de vie du système

L'importation de données de `FailureTimes_5.csv` (l'annexe 3.1) montre des dates de pannes d'ordres grandement variées (300 à 27000) (l'annexe 3.2).

Mettre à l'exponentiel des valeurs extrêmes résultera en des valeurs nulles (si négatives) ou Inf (si positives) pour l'ordinateur, ce qui est indésirable. Alors nous devons forcément les réduire en les divisant par un scalaire `scale`, prenons par exemple 1000. (Figure 1)

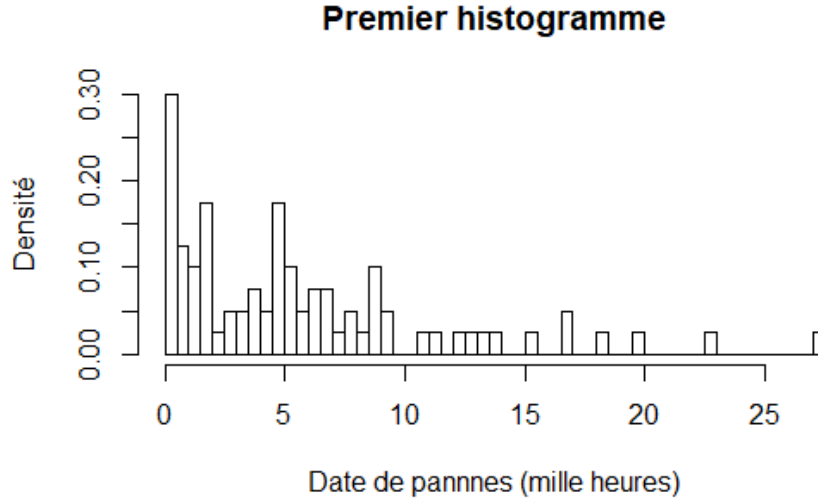


FIGURE 1 – Le premier histogramme de distribution de pannes

Nous pouvons remarquer que les valeurs sont positives (étant données que ce sont des temps) et que la distribution semble posséder deux parties importantes. En effet, les pannes se concentrent autour de 2 sommets, l'un à $[0; 0, 5]$ et l'autre à $[4, 5; 5]$. Ceci nous fait penser naturellement à un mixage de deux lois. Le premier sommet est suivi d'une pente forte, et la distribution locale de l'autre sommet a une forme de pic. Nous avons donc pensé estimer un mixage de loi *Exponentielle* et *Gamma* afin de modéliser les données. En effet la première partie correspondrait à une loi exponentielle tandis que la seconde à une loi gamma.

La fonction de densité du mixage avec le paramètre $\theta = (p_1, p_2, \lambda, \alpha, \beta)$ est donnée par :

$$\begin{aligned} f_{\theta}(x) &= p_1 f_1(x) + p_2 f_2(x) \\ &= p_1 \lambda e^{-\lambda x} + p_2 \frac{\beta^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \end{aligned} \quad (2)$$

Où f_1, f_2 désignent respectivement $\exp(\lambda)$ et $\Gamma(\alpha, \beta)$; $p_1, p_2 > 0$: $p_1 + p_2 = 1$.

Nous allons utiliser l'algorithme EM, la méthode la plus efficace pour estimer l'estimateur du maximum de vraisemblance (MLE) de mixage fini.

Soit X la variable aléatoire de durée de vie du système. Soient (x_1, \dots, x_N) les observations.

Soit la matrice de probabilité d'appartenance (ζ_{ki}) : ζ_{ki} vaut la probabilité que x_i suive la loi f_k .

$$\zeta_{ki} = \frac{p_k f_k(x_i)}{p_1 f_1(x_i) + p_2 f_2(x_i)} \forall k = \overline{1, 2} \forall i = \overline{1, N} \quad (3)$$

La fonction de vraisemblance :

$$\ln \Lambda = \sum_{i=1}^N \ln f_{\theta}(x_i) = \sum_{i=1}^N \ln (p_1 f_1(x_i) + p_2 f_2(x_i)) \quad (4)$$

Nous cherchons à maximiser $\ln \Lambda$ en la dérivant selon λ, α, β .

Pour λ :

$$\begin{aligned} \frac{\partial}{\partial \lambda} \ln \Lambda &= \sum_{i=1}^N \frac{p_1 e^{-\lambda x_i} - p_1 \lambda x_i e^{-\lambda x_i}}{p_1 f_1(x_i) + p_2 f_2(x_i)} \\ &= \sum_{i=1}^N \frac{p_1 f_1(x_i)}{p_1 f_1(x_i) + p_2 f_2(x_i)} \left(\frac{1}{\lambda} - x_i \right) \\ &= \frac{1}{\lambda} \sum_{i=1}^N \zeta_{1i} - \sum_{i=1}^N \zeta_{1i} x_i = 0 \\ \Leftrightarrow \lambda &= \frac{\sum_{i=1}^N \zeta_{1i}}{\sum_{i=1}^N \zeta_{1i} x_i} \end{aligned} \quad (5)$$

Pour β :

$$\begin{aligned}
\frac{\partial}{\partial \beta} \ln \Lambda &= \sum_{i=1}^N \frac{p_2 x_i^{\alpha-1}}{\Gamma(\alpha)} \frac{\alpha \beta^{\alpha-1} e^{-\beta x_i} - \beta^\alpha x_i e^{-\beta x_i}}{p_1 f_1(x_i) + p_2 f_2(x_i)} \\
&= \sum_{i=1}^N \frac{p_2 f_2(x_i)}{p_1 f_1(x_i) + p_2 f_2(x_i)} \left(\frac{\alpha}{\beta} - x_i \right) \\
&= \frac{\alpha}{\beta} \sum_{i=1}^N \zeta_{2i} - \sum_{i=1}^N \zeta_{2i} x_i = 0 \\
&\Leftrightarrow \beta = \alpha \frac{\sum_{i=1}^N \zeta_{2i}}{\sum_{i=1}^N \zeta_{2i} x_i}
\end{aligned} \tag{6}$$

Pour α :

$$\begin{aligned}
\frac{\partial}{\partial \alpha} \ln \Lambda &= \sum_{i=1}^N \frac{p_2 e^{-\beta x_i}}{p_1 f_1(x_i) + p_2 f_2(x_i)} \left(\frac{\beta (\ln \beta + \ln x_i) (\beta x_i)^{\alpha-1}}{\Gamma(\alpha)} - \beta^\alpha x_i^{\alpha-1} \frac{\Psi(\alpha)}{\Gamma(\alpha)} \right) \\
&= \sum_{i=1}^N \frac{p_2 f_2(x_i)}{p_1 f_1(x_i) + p_2 f_2(x_i)} (\ln \beta + \ln x_i - \Psi(\alpha)) \\
&= \left(\sum_{i=1}^N \zeta_{2i} \right) \ln \beta + \sum_{i=1}^N \zeta_{2i} \ln x_i - \Psi(\alpha) \left(\sum_{i=1}^N \zeta_{2i} \right) = 0 \\
&\Leftrightarrow 0 = \ln \alpha + \ln \frac{\sum_{i=1}^N \zeta_{2i}}{\sum_{i=1}^N \zeta_{2i} x_i} + \frac{\sum_{i=1}^N \zeta_{2i} \ln x_i}{\sum_{i=1}^N \zeta_{2i}} - \Psi(\alpha) \quad (\beta \text{ substitué par (6)}) \\
&\Leftrightarrow 0 = \ln \alpha - \Psi(\alpha) - c
\end{aligned}$$

Où $c = \ln \left(\frac{\sum_{i=1}^N \zeta_{2i} x_i}{\sum_{i=1}^N \zeta_{2i}} \right) - \frac{\sum_{i=1}^N \zeta_{2i} \ln(x_i)}{\sum_{i=1}^N \zeta_{2i}}$; Ψ est la fonction digamma.

Selon la méthode de Newton-Rashphon, nous pouvons résoudre α numériquement avec ce formul itératif :

$$\alpha_{r+1} = \alpha_r - \frac{\ln \alpha_r + \Psi(\alpha_r) - c}{\frac{1}{\alpha_r} - \Psi'(\alpha_r)}$$

[1] propose un autre formule convergeant plus vite :

$$\frac{1}{\alpha_{r+1}} = \frac{1}{\alpha_r} + \frac{\ln(\alpha_r) - \Psi(\alpha_r) - c}{\alpha_r^2 \left(\frac{1}{\alpha_r} - \Psi'(\alpha_r) \right)} \tag{7}$$

Avec Ψ' la fonction trigamma. L'itération part avec $\alpha_0 = \frac{0.5}{c}$.

Au final, pour p_k :

$$p_k = \frac{\sum_{i=1}^N \zeta_{ki}}{N} \quad \forall k = 1, 2 \tag{8}$$

Etant donné (3), (5), (6), (7) et (8), nous utilisons l'algorithme EM :

1. **Initialisation** : Choisir θ_0 .
2. **Etape E** : Evaluer (ζ_{ki}) sachant θ_c en utilisant (3).
3. **Etape M** : Calculer θ_{c+1} à l'aide des équations (5), (6), (7) et (8).
Note : Pour α , l'itération se termine quand $|\alpha_{r+1} - \alpha_r| < \varepsilon_\alpha$ où ε_α est un réel positif fixé à l'initialisation.
4. **Evaluation** : Si $\|\theta_{c+1} - \theta_c\| < \varepsilon_\theta$ (ε_θ est un réel positif fixé à l'initialisation), l'algorithme s'arrête et $\theta = \theta_{c+1}$.
 Sinon, reviens à l'étape E avec $c = c + 1$.

Avant de lancer l'algorithme, nous avons essayé d'obtenir un ensemble de paramètres initiaux θ_0 qui soient cohérents avec la distribution des données. Nous avons choisi :

$$(p_{1_0}, p_{2_0}, \lambda_0, \alpha_0, \beta_0) = (0.5; 0.5; 1; 10; 2)$$

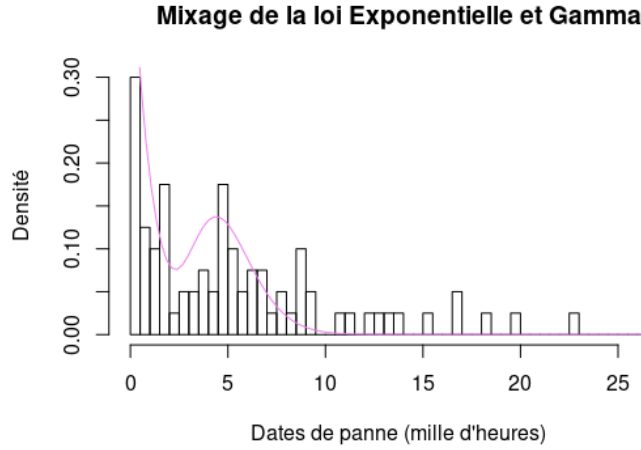


FIGURE 2 – Mixage de la loi Exponentielle et Gamma : **paramètres initiaux**

Après l'utilisation de l'algorithme EM, nous avons obtenu le résultat :

$$(p_1, p_2, \lambda, \alpha, \beta) = (0.2194518; 0.7805482; 1.56738; 1.665659; 0.2332427)$$

Nous avons tracé la fonction de densité f_θ trouvé (figure 3) et réalisé un test de Kolmogorov-Smirnov qui donne $p - value = 0,9663111$ signifiant 96,63% de nous tromper si nous rejetons ce modèle. Nous l'acceptons alors, quoiqu'il ne génère pas 2 sommets comme la remarque initiale. Le code est trouvable à l'annexe 3.3.

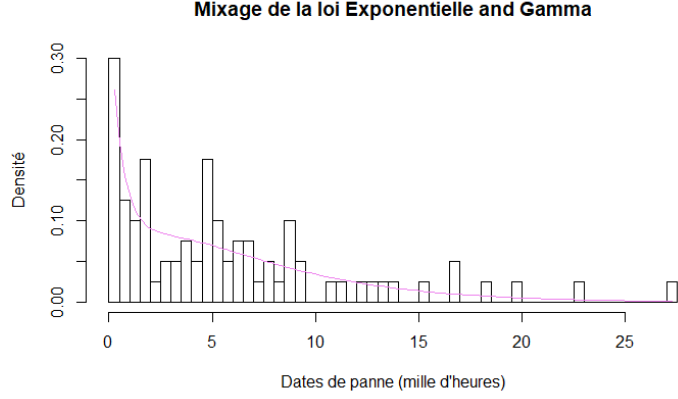


FIGURE 3 – Mixage de la loi Exponentielle et Gamma :paramètres finaux

1.3 La politique de maintenance basée sur l'âge

Avec la fonction f_θ trouvée, nous allons pouvoir une politique optimale (selon le modèle) basée sur l'âge.

Nous avons par définition : $S = \min(X, t_0)$.

Autrement dit, $S = X\mathbb{I}_{\{X < t_0\}} + t_0\mathbb{I}_{\{X \geq t_0\}}$.

Cela se traduit avec le coût : $C(S) = C_c\mathbb{I}_{\{X < t_0\}} + C_p\mathbb{I}_{\{X \geq t_0\}}$, avec C_c, C_p les coûts de maintenances correctives et préventives respectivement.

Le coût moyen :

$$\begin{aligned}\mathbb{E}(C(S)) &= c_c\mathbb{E}(\mathbb{I}_{\{X < t_0\}}) + c_p\mathbb{E}(\mathbb{I}_{\{X \geq t_0\}}) \\ &= c_cP(X < t_0) + c_pP(X \geq t_0) \\ &= c_cF_\theta(t_0) + c_p(1 - F_\theta(t_0)) \\ &= (c_c - c_p)F_\theta(t_0) + c_p\end{aligned}\tag{9}$$

La durée moyenne :

$$\begin{aligned}\mathbb{E}(S) &= \mathbb{E}(X\mathbb{I}_{\{X < t_0\}}) + t_0\mathbb{E}(\mathbb{I}_{\{X \geq t_0\}}) \\ &= \int_0^{t_0} x f_\theta(x) dx + t_0P(X \geq t_0) \\ &= xF_\theta(x) \Big|_0^{t_0} - \int_0^{t_0} F_\theta(x) dx + t_0(1 - F_\theta(t_0)) \\ &= t_0 - \int_0^{t_0} F_\theta(x) dx\end{aligned}\tag{10}$$

De (9) et (10), nous détaillons le coût moyen sur une durée de temps (1) :

$$\mathbb{E}(C) = \frac{\mathbb{E}(C(S))}{\mathbb{E}(S)} = \frac{(c_c - c_p)F_\theta(t_0) + c_p}{t_0 - \int_0^{t_0} F_\theta(x) dx}\tag{11}$$

L'annexe (3.4) montre comment chercher l'optimum numériquement. La valeur minimum est $t_0 = 27,29639$ (mille heures), correspondant à un coût moyen

de 210,6402. Nous constatons que t_0^{min} est très proche du maximum de durée de vie, indiquant que l'optimisation de t_0 est inutile car le système ne vieillit que très peu¹.

1. la loi de durée de vie trouvée est très proche d'une loi exponentielle car c'est celle-ci qui pré-pondère et cette loi implique un système toujours neuf (d'âge 0), et donc sans besoin de maintenance basée sur l'âge. On peut facilement montrer que l'optimum de t_0^* dans le cas exponentiel est infini.

2 Maintenance basée sur dégradation

2.1 Rappel

En observant de multiples systèmes identiques, nous effectuons des mesures de dégradation sur des intervalles de temps réguliers tout au long de leurs durée de vie.

La valeur limite de dégradation est $L = 20$. C'est-à-dire que lorsque le niveau de dégradation dépasse L , le système tombe en panne et nous ne pourrions plus le mesurer.

On souhaite mettre en place une politique de maintenance conditionnelle, basée sur un seuil M inférieur à L et l'intervalle de temps ΔT entre les inspections répétées. Appelons X_t le niveau de dégradation à l'instant t (instant d'une inspection) :

- Si $X_t < M$, nous laissons le système tel quel.
- Si $M \leq X_t < L$, un remplacement préventif est réalisé au coût c_p . Et puis X_t est remis à 0.
- Si $X_t \geq L$, un remplacement correctif est fait au coût c_c . Et puis X_t est remis à 0.

Le but est minimiser le coût moyen sur une durée de temps (1) en choisissant bien le seuil M et l'intervalle d'inspection ΔT .

2.2 Modéliser la dégradation du système

Avec les données de `DegradLevel_2.csv`, nous traçons leurs processus de dégradation (figure (2.2)). (Les annexes (3.5) et (3.6))

Comme les temps d'inspection sont de l'ordre du millier, il faudrait les diviser par un scalaire (par exemple, $scale = 1000$) afin d'assurer la précision des calculs numériques.

Nous voyons que les accroissements sont positifs, suggérant un modèle de processus Gamma.

Soit $X(t)$ la variable aléatoire de dégradation du système. Supposons que $X(t) - X(s) \sim \Gamma(a(t-s), b) \forall t > s > 0$. Nous allons estimer les paramètres a, b en modélisant la distribution des incréments entre deux moments successifs.

Fixons $t - s = \delta = 0.8$, car les mesures données sont effectués au bout de chaque intervalle de 0.8.

L'histogramme des incréments :

A l'aide du librairie MASS : $(a; b) = (\frac{\alpha}{\delta}; \beta) = (2.843101; 1.140354)$ où (α, β) sont les paramètres estimés par MASS. Le code est mis à l'annexe (3.7).

Un test rapide de Kolmogorov-Smirnov nous donne $p - value = 0.8651934$, indiquant le modèle est largement acceptable.

D'autant plus, si nous refaisons les calculs ci-dessus avec $\delta = 1.6, 2.4, 3.2, etc.$, nous voyons les valeurs de a et b ne varient pas trop. Alors, nous choisissons le couple (a, b) avec p le plus grand. (δ trop grand réduira nombreux de données, résultant une perte importante de précision)

Au final, nous choisissons $X(t) - X(s) \sim \Gamma(a(t-s), b) \forall t > s > 0, \delta = 1.6$ et $(a, b) = (3.0207719; 1.2091646)$.

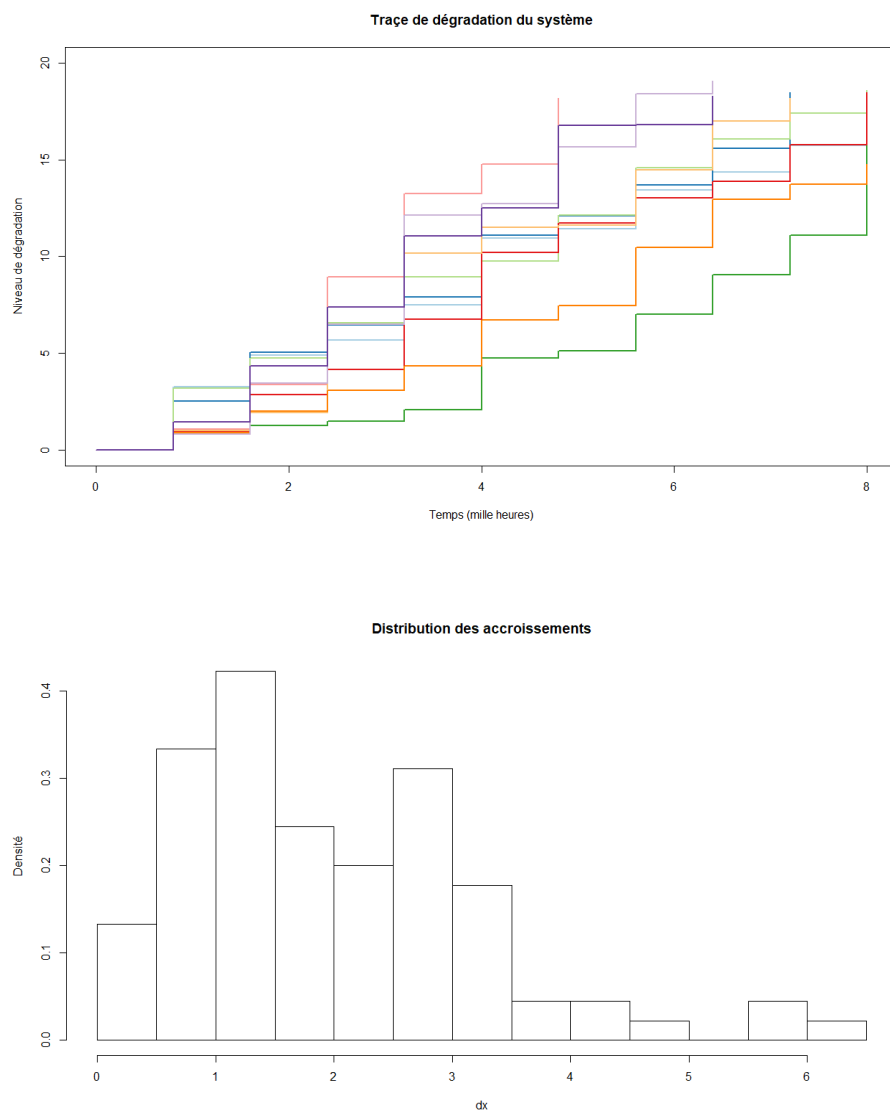


FIGURE 4 – Histogramme des incréments de l'intervalle $\delta = 0.8$

δ	a	b	p
0.8	2.8431009	1.1403540	0.8651934
1.6	3.0207719	1.2091646	0.9919939
2.4	3.0187497	1.1907719	0.1279252
3.2	2.997763	1.185547	0.300610
4.0	3.5036580	1.4061063	0.6307951

TABLE 1 – Calcul (a, b) avec différents δ

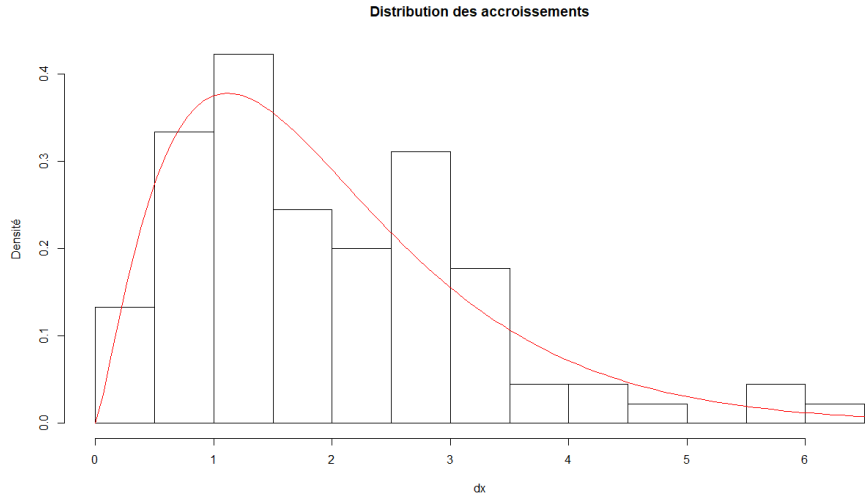


FIGURE 5 – Histogramme et la courbe de densité estimé des incréments de l'intervalle $\delta = 0.8$

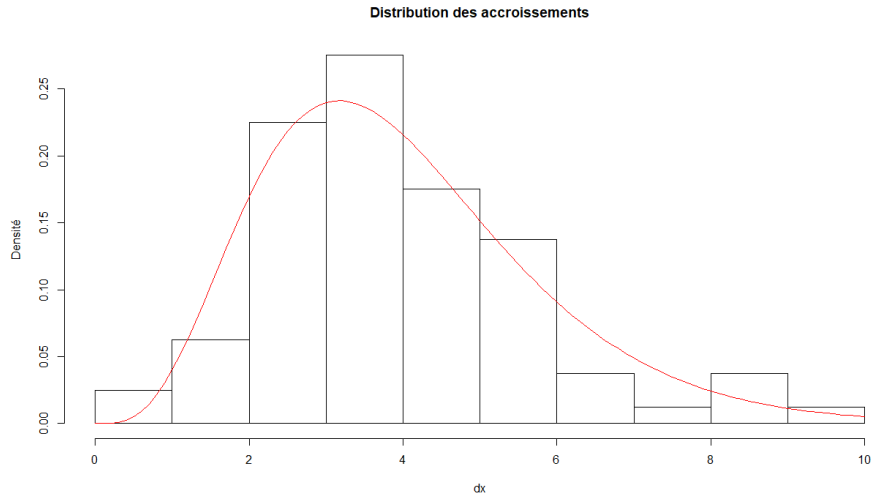


FIGURE 6 – Histogramme et la courbe de densité estimé des incréments de l'intervalle $\delta = 1.6$

2.3 La politique de maintenance basée sur dégradation

Cette politique, outre que $c_c = 1200$ et $c_p = 800$, introduit ainsi le coût d'inspection répétitive $c_i = 10$.

Soit S la variable aléatoire représentant la date de remplacement ; $C(S)$ est le coût de maintenance cumulé à l'instant S , et $N(S)$ le nombre d'inspections depuis la dernière remplacement jusqu'à S .

$$C(S) = c_i N(S) + c_p \mathbb{I}_{\{L > X_{N(S)} \geq M\}} + c_c \mathbb{I}_{\{X_{N(S)} \geq L\}} \quad (12)$$

Posons $F_t = \Gamma(at, b)$ et les instants d'inspections $t_1, t_2, \dots, t_{N(S)}$ avec $t_{j+1} - t_j = \Delta T$. Puisque $t_0 = 0$, nous aurons $t_k = k\Delta T$ et $X_{t_k} \sim F_{k\Delta T}$.

Comme le calcul analytique de $C(S)$ et $N(S)$ devient grossier (l'annexe (3.8)), nous passons à la méthode Monté Carlo, qui nous permet d'estimer la vraie valeur en répétant un nombre suffisant de simulations.

- L'algorithme (1) calcule la durée de vie d'une simulation.
- L'algorithme (2) calcule le coût d'une simulation.
- L'algorithme (2) simule de multiples processus et retourne une valeur approximative de $\mathbb{E}(C)$.
- L'algorithme (4) trouve l'optimum sur une zone de recherche.

Algorithme 1 : Mesurer la durée de vie d'un processus

Entrées : p le processus

ΔT l'intervalle d'inspection

M le seuil de maintenance préventive

Output : S la durée de vie

```

1  $len \leftarrow$  longueur de  $p$ ;
2 si  $p_{len} < M$  alors
   | // La simulation n'est pas suffisante longue
3   | retourner  $\emptyset$ 
4 sinon
5   |  $n_s \leftarrow$  le dernier index de  $p$  tel que  $p_{n_s} < M$ ;
6   | retourner  $n_s + 1$ 
7 fin
```

Algorithme 2 : Mesurer le coût de maintenance d'un processus

Entrées : p le processus

M le seuil

Output : $C(S)$ le coût de maintenance

```
1  $cout \leftarrow 0$ ;  
2  $len \leftarrow$  longueur de  $p$ ;  
3 si  $p_{len} \geq M$  alors  
4    $n_s \leftarrow$  le dernier index de  $p$  tel que  $p_{n_s} < M$ ;  
   //  $n_s$  correspond le nombre d'inspection  
5    $cout \leftarrow cout + c_i * n_s$ ;  
   // A l'inspection suivante  
6   si  $p_{n_s+1} \geq L$  alors  
   // Le système est tombé en panne  
7    $cout \leftarrow cout + c_c$ ;  
8   sinon  
   // Le système fonctionne encore  
9    $cout \leftarrow cout + c_p$ ;  
10  fin  
11 fin  
   // Le coût est zéro si le niveau de dégradation ne dépasse  
   pas encore  $M$   
12 retourner  $cout$ 
```

Algorithme 3 : Simuler le processus et calculer $\widehat{\mathbb{E}(C)}$

Entrées : ΔT l'intervalle d'inspection

M le seuil

$nSim$ le nombre de simulation

$nStep$ le nombre d'inspection maximale

Output : $\widehat{\mathbb{E}(C)}$ la valeur approximative

```
1 Générer la matrice de l'incrément  $simStep$  de taille  $nSim \times nStep$  aux  
  valeurs aléatoires selon la loi  $\Gamma(a\Delta T, b)$ ;  
2 Calculer la matrice de processus  $simProc$  de taille  $nSim \times nStep$  en  
  faisant la somme cumulative de  $simStep$  horizontalement;  
3  $simTime \leftarrow$  les durées de vie de processus en appliquant l'algorithme  
  (1);  
4  $simCost \leftarrow$  les coûts de processus en appliquant l'algorithme (2);  
5  $meanSimCost \leftarrow$  la moyenne de tous les coûts non nuls de  $simCost$ ;  
6  $meanSimTime \leftarrow$  la moyenne de durée de vie non nuls de  $simTime$ ;  
7  $meanCostTime \leftarrow \frac{meanSimCost}{meanSimTime}$ ;  
8 si  $meanSimCost$  est NaN ou  $meanSimTime$  est NaN alors  
9    $meanCostTime \leftarrow MAX\_DOUBLE$ ;  
10 fin  
11 retourner  $meanCostTime$ 
```

Algorithme 4 : Optimiser la politique de maintenance basée sur dégradation

Entrées : $[L_{\Delta T}, U_{\Delta T}]$ l'intervalle de recherche de ΔT
 $[L_M, U_M]$ l'intervalle de recherche de M
 τ la tolérance
 $d_{\Delta T}$ le nombre de points à évaluer pour ΔT
 d_M le nombre de points à évaluer pour M
 $e_{\Delta T}$ le ratio de réduction sur l'intervalle de ΔT
 e_M le ratio de réduction sur l'intervalle de M

Output : ΔT^* l'intervalle d'inspection optimal
 M^* le seuil optimal
 o^* la valeur objective optimale

// Suivre le principe de diachotomie

```

1  $\Delta T^* \leftarrow 0$ ;
2  $M^* \leftarrow 0$ ;
3  $o^* \leftarrow 0$ ;
4 tant que  $\left(\frac{L_{\Delta T} - U_{\Delta T}}{d_{\Delta T}}\right)^2 + \left(\frac{L_M - U_M}{d_M}\right)^2 \geq \tau^2$  faire
    // Evaluer
5  $I_{\Delta T} \leftarrow$  l'ensemble de  $d_{\Delta T}$  points égaux-distances de  $[L_{\Delta T}, U_{\Delta T}]$ ;
6  $I_M \leftarrow$  l'ensemble de  $d_M$  points égaux-distances de  $[L_M, U_M]$ ;
7  $Z \leftarrow$  résultats de l'application de l'algorithme (3) pour chaque point
   sur la grille  $I_{\Delta T} \times I_M$ ;
   // Chercher le minimum
8  $(i^*, j^*) \leftarrow$  l'indice du minimum de  $Z$ ;
9  $\Delta T^* \leftarrow$  l'élément  $i^*$ -ième de  $I_{\Delta T}$ ;
10  $M^* \leftarrow$  l'élément  $j^*$ -ième de  $I_M$ ;
11  $o^* \leftarrow Z_{i^*, j^*}$ ;
   // Ajuster la nouvelle zone de recherche autour du minimum
   actuel
12  $l_{\Delta T} \leftarrow L_{\Delta T} - U_{\Delta T}$ ;
13  $l_M \leftarrow L_M - U_M$ ;
14  $L_{\Delta T} \leftarrow \min\left(\Delta T^* + \frac{l_{\Delta T}}{e_{\Delta T}}, L_{\Delta T}\right)$ ;
15  $U_{\Delta T} \leftarrow \max\left(\Delta T^* - \frac{l_{\Delta T}}{e_{\Delta T}}, U_{\Delta T}\right)$ ;
16  $L_M \leftarrow \min\left(M^* + \frac{l_M}{e_M}, L_M\right)$ ;
17  $U_M \leftarrow \max\left(M^* - \frac{l_M}{e_M}, U_M\right)$ ;
18 fin
19 retourner  $(\Delta T^*, M^*, o^*)$ 

```

Puisque l'algorithme (3) utilise du calcul matriciel pour accélérer le traitement, $nStep$ doit être fixé judicieusement pour ne pas avoir "trop" de simulations où $X_{t_{nStep}} < M$ (ce qui va forcer les valeurs de $C(S)$ et S à zéro selon les algorithmes (1) et (2)). Soit α la probabilité de cet événement.

$$\begin{aligned} P(X_{t_{nStep}} < M) &= \alpha \\ \Leftrightarrow F_{nStep * \Delta T}(M) &= \alpha \\ \Leftrightarrow \Gamma(a * nStep * \Delta T, b)(M) &= \alpha \end{aligned}$$

Nous testons quelques valeurs de F . Comme F est croissant, nous l'évaluons directement à $M = 20$.

$nStep * \Delta T$	$\Gamma(a * nStep * \Delta T, b)(20)$
8	0.5284405
10	0.1319799
12	0.01300913
13	0.002925565
14	0.000533971

TABLE 2 – Estimation de $nStep * \Delta T$

La fonction distribution décroît selon $nStep * \Delta T$. Et pourtant, $\alpha \leq 1\%$ nous suffit. Et plus grand $nStep$, plus lente la traitement. C'est pourquoi, basé sur la table (2), nous emploirons

$$nStep = \left\lfloor \frac{13}{\Delta T} \right\rfloor$$

Le code implémenté se situe à l'annexe 3.9.

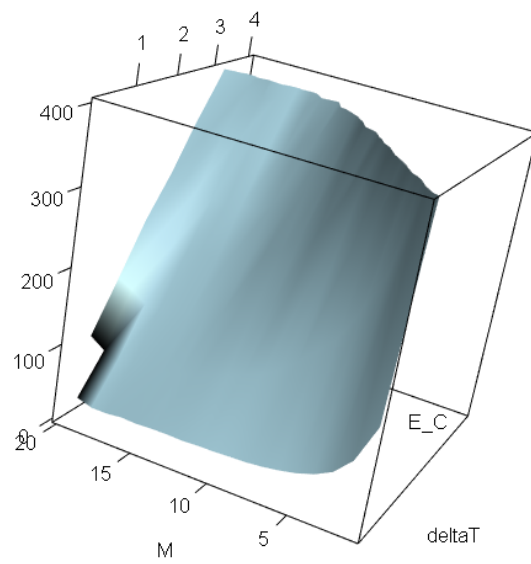
Bien que le résultat n'est pas stable, nous pouvons arriver à une valeur approximative après plusieurs itérations de l'algorithme 4 sur l'intervalle $[0.001, 4] \times [0, 20]$:

$$\begin{aligned} (\Delta T^*, M^*, o^*) &= (0.001007627 \text{ (mille heures)}, \\ &\quad 18.58027, \\ &\quad 10.10541 \text{ (euros/mille heures)}) \end{aligned}$$

Ce résultat nous semble logique car :

- ΔT^* faible (grâce à coût c_i qui est faible) nous permet d'observer mieux et juste à temps le système.
- M^* proche de L va "remplacer" la plupart des coûts c_c par c_p pour économiser.

FIGURE 7 – Représentation graphique de $E(C)$



3 Annexe

3.1 L'importation de données de pannes

```
1 pannes = read.csv(  
2   file = "FailureTimes_5.csv",  
3   header = TRUE,  
4   sep = ",",  
5   dec = ".",  
6   colClasses = c("NULL", NA)  
7 )  
8 scale = 1000  
9 data = pannes$Heures / scale  
10 N = length(data)
```

3.2 Le premier histogramme de distribution de pannes

```
1 hist(  
2   data,  
3   breaks = 40,  
4   probability = TRUE,  
5   xlab = "Date de pannes (mille heures)",  
6   ylab = "Densité",  
7   main = "Premier histogramme"  
8 )
```

3.3 Estimer le mixage de la loi Exponentielle et Gamma

```
1 # Fitting mixture of Exp and Gamma  
2 # Algorithm EM  
3 # Initialisation  
4 k = 2 # number of components  
5 p = c(0.5, 0.5)  
6 lambda = 1  
7 alpha = 10  
8 beta = 2  
9 f = list(  
10   '1' = function(x) {  
11     dexp(x, rate = lambda)  
12   },  
13   '2' = function(x) {  
14     dgamma(x, shape = alpha, rate = beta)  
15   }  
16 )  
17 # Illustration initial  
18 f_theta = function(x) {  
19   p[[1]] * f[[1]](x) + p[[2]] * f[[2]](x)  
20 }  
21 h_theta = hist(  
22   data,  
23   breaks = 40,  
24   probability = TRUE,  
25   main = "Mixage de la loi Exponentielle et Gamma",  
26   xlab = "Dates de panne (mille d'heures)",  
27   ylab = "Densité"  
28 )  
29 curve(  
30   f_theta(x),  
31   add = TRUE,  
32   col = "violet",  
33   from = min(h_theta$mids),  
34   to = max(h_theta$mids)  
35 )  
36 epsilon = list(  
37   alpha = 1e-4,  
38   theta = 1e-4
```



```

39 )
40 zeta = matrix(
41   0,
42   nrow = k,
43   ncol = N
44 )
45 # Norm
46 normVec = function(x) sqrt(sum(x^2))
47 # New value
48 p_new = p
49 alpha_new = alpha
50 beta_new = beta
51 lambda_new = lambda
52 repeat {
53   ## E Step
54   # Calculate each proba
55   for (l in 1:k) {
56     zeta[l,] = p[[l]] * f[[l]](data)
57   }
58   # Normalize proba
59   zeta = t(t(zeta) / rowSums(t(zeta)))
60   ## M step
61   # Lambda
62   lambda_new = sum(zeta[1,]) / sum(zeta[1,] * data)
63   # Alpha
64   c = log(sum(zeta[2,] * data) / sum(zeta[2,])) - sum(zeta[2,] * log(data)) / sum(zeta[2,])
65   alpha_new = 0.5 / c
66   alpha_temp = 0
67   repeat {
68     alpha_temp = 1 / (1 / alpha_new + (log(alpha_new) - digamma(alpha_new) - c) / (alpha_new^2 * (1 / alpha_new - trigamma(alpha_new))))
69     if (abs(alpha_temp - alpha_new) < epsilon$alpha) {
70       break
71     } else {
72       alpha_new = alpha_temp
73     }
74   }
75   alpha_new = alpha_temp
76   # Beta
77   beta_new = alpha_new * sum(zeta[2,]) / sum(zeta[2,] * data)
78   # p
79   for (l in 1:k) {
80     p_new[[l]] = mean(zeta[l,])
81   }
82   ## Evaluation
83   if (normVec(c(alpha, beta, lambda, p[[1]], p[[2]]) - c(alpha_new, beta_new, lambda_new, p_new[[1]], p_new[[2]])) < epsilon$theta) {
84     break
85   } else {
86     alpha = alpha_new
87     beta = beta_new
88     lambda = lambda_new
89     p = p_new
90   }
91 }
92 # Final value update
93 alpha = alpha_new
94 beta = beta_new
95 lambda = lambda_new
96 p = p_new
97 # Illustration final
98 f_theta = function(x) {
99   p[[1]] * f[[1]](x) + p[[2]] * f[[2]](x)
100 }
101 h_theta = hist(
102   data,
103   breaks = 40,
104   probability = TRUE,
105   main = "Mixage de la loi Exponentielle et Gamma",
106   xlab = "Dates de panne (mille d'heures)",
107   ylab = "Densité"
108 )
109 curve(

```

```

110 f_theta(x),
111 add = TRUE,
112 col = "violet",
113 from = min(h_theta$mids),
114 to = max(h_theta$mids)
115 )
116 # Kolmogorov-Smirnov test
117 F_theta = function(x) {
118   p[[1]] * pexp(x, rate = lambda) + p[[2]] * pgamma(x, shape = alpha, rate
119   = beta)
120 }
121 test = ks.test(data, F_theta, exact = TRUE)

```

3.4 Optimiser le coût moyen sur une durée de temps

```

1 # Finding optimal t_0
2 # Given F_theta
3 c_c = 1200
4 c_p = 800
5 E_C_S = function(x) {
6   (c_c - c_p) * F_theta(x) + c_p
7 }
8 E_S = function(x) {
9   x - integrate(F_theta, 0, x)$value
10 }
11 E_C = function(x) {
12   E_C_S(x) / E_S(x)
13 }
14 o = optimize(
15   E_C,
16   c(min(data), max(data)),
17   tol = 1e-5
18 )
19 d = seq(
20   min(data),
21   max(data),
22   0.01
23 )
24 plot(
25   d,
26   lapply(
27     d,
28     E_C
29   ),
30   main = "Coût moyenne sur une durée de temps",
31   xlab = "t_0",
32   ylab = "",
33   type = "l"
34 )

```

3.5 Importer les valeurs de dégradation

```

1 # Import degradation
2 table = read.csv(
3   file = "DegradLevel_2.csv",
4   header = TRUE,
5   sep = ",",
6   dec = "."
7 )
8 scale = 1000
9 time = c(0, table$Temps / scale)
10 nbProcess = length(table) - 2
11 process = matrix(
12   ,
13   nrow = nbProcess,
14   ncol = 1 + length(table[[3]])
15 )
16 for (i in 1:nbProcess) {

```

```

17 |     process[i,] = c(0, table[[i + 2]]) # degrad = 0 at t = 0
18 | }

```

3.6 Premiers traçes de dégradation

```

1 | # Plot process curves
2 | L = 20
3 | # Colormap
4 | library(RColorBrewer)
5 | color = brewer.pal(nbProcess, "Paired")
6 | # First process
7 | plot(
8 |     NULL,
9 |     type = "n",
10 |    main = "Traçes de dégradation du système",
11 |    xlim = c(min(time), max(time)),
12 |    xlab = "Temps (mille heures)",
13 |    ylim = c(0, L),
14 |    ylab = "Niveau de dégradation"
15 | )
16 | for (i in 1:nbProcess) {
17 |     lines(
18 |         x = time,
19 |         y = process[i,],
20 |         type = "s",
21 |         col = color[[i]],
22 |         lwd = 2
23 |     )
24 | }

```

3.7 Estimation de paramètres de dégradations

```

1 | ## Estimate parameters for process
2 | lag = 2
3 | delta = 0.8 * lag
4 | d = t(diff(t(process), lag))
5 | # Concatenate into one vector
6 | increments = vector(
7 |     mode = "numeric",
8 |     length = length(d)
9 | )
10 | n = dim(d)[[1]]
11 | l = dim(d)[[2]]
12 | for (i in 1:n) {
13 |     increments[((i - 1) * l + 1):(i * l)] = d[i,]
14 | }
15 | # Filter out NA values
16 | increments = increments[!is.na(increments)]
17 | # Histogram
18 | hiso_degrad = hist(
19 |     increments,
20 |     breaks = 10,
21 |     probability = TRUE,
22 |     main = "Distribution des accroissements",
23 |     xlab = "dx",
24 |     ylab = "Densité"
25 | )
26 | # Estimate gamma distribution
27 | library(MASS)
28 | estim = fitdistr(
29 |     increments,
30 |     dgamma,
31 |     list(
32 |         shape = 1,
33 |         rate = 1
34 |     )
35 | )
36 | # Draw estimated density

```

```

37 a = estim$estimate[[1]] / delta
38 b = estim$estimate[[2]]
39 curve(
40   dgamma(x, a * delta ,b),
41   add = TRUE,
42   col = "red"
43 )
44 # Kolmogorov-Smirnov test
45 test = ks.test(increments, "pgamma", a * delta, b)
46 c(delta, a, b, test$p.value) # Print values
47 # PDF
48 f = function(x, t) {
49   dgamma(x, shape = a * t, rate = b)
50 }
51 # CDF
52 F = function(x, t) {
53   pgamma(x, shape = a * t, rate = b)
54 }
55 # Generator
56 rG = function(n, t) {
57   rgamma(n, shape = a * t, rate = b)
58 }

```

3.8 Calcul analytique de maintenance préventive

Rappelons que $X_{t_k} \sim F_{k\Delta t}$ et (12) :

$$C(S) = c_i N(S) + c_p \mathbb{I}_{\{L > X_{N(S)} \geq M\}} + c_c \mathbb{I}_{\{X_{N(S)} \geq L\}}$$

D'où

$$\begin{aligned} \mathbb{E}(C(S)) &= c_i \mathbb{E}(N(S)) + c_p \mathbb{E}\left(I_{\{L > X_{t_{N(S)}} \geq M\}}\right) \\ &\quad + c_c \mathbb{E}\left(I_{\{X_{t_{N(S)}} \geq L\}}\right) \end{aligned}$$

Détaillons le nombre moyen d'inspection $\mathbb{E}(N(S))$:

$$\mathbb{E}(N(S)) = \sum_{k=1}^{\infty} k P(N(S) = k) = P(N(S) = 1) + \sum_{k=2}^{\infty} k P(N(S) = k)$$

Or

$$\begin{aligned} P(N(S) = 1) &= P(X_{t_1} \geq M) \\ &= 1 - F_{\Delta T}(M) \end{aligned}$$

$$\begin{aligned} P(N(S) = k) &= P(X_{t_k} \geq M \wedge X_{t_{k-1}} < M) \\ &= P(\Delta X \geq M - X_{t_{k-1}} \wedge X_{t_{k-1}} < M) \\ &= \int_0^M P(\Delta X \geq M - \xi) P(X_{t_{k-1}} = \xi) d\xi \\ &= \int_0^M (1 - F_{\Delta T}(M - \xi)) f_{(k-1)\Delta T}(\xi) d\xi \\ &= \int_0^M f_{(k-1)\Delta T}(\xi) d\xi - \int_0^M F_{\Delta T}(M - \xi) f_{(k-1)\Delta T}(\xi) d\xi \\ &= F_{(k-1)\Delta T}(M) - \int_0^M F_{\Delta T}(M - \xi) f_{(k-1)\Delta T}(\xi) d\xi \end{aligned}$$

Par conséquence,

$$\mathbb{E}(N(S)) = 1 - F_{\Delta T}(M) + \sum_{k=2}^{\infty} k \left(F_{(k-1)\Delta T}(M) - \int_0^M F_{\Delta T}(M - \xi) f_{(k-1)\Delta T}(\xi) d\xi \right) \quad (13)$$

En suite, nous exprimons la probabilité de maintenance préventive :

$$\begin{aligned} \mathbb{E} \left(\mathbb{I}_{\{L > X_{t_{N(S)}} \geq M\}} \right) &= \sum_{k=1}^{\infty} P(L > X_{t_k} \geq M | X_{t_{k-1}} < M) \\ k=1 : P(L > X_{t_1} \geq M) &= F_{\Delta T}(L) - F_{\Delta T}(M) \\ \forall k \geq 2 : \\ P(L > X_{t_k} \geq M | X_{t_{k-1}} < M) &= \frac{P(L > X_{t_k} \geq M \wedge X_{t_{k-1}} < M)}{P(X_{t_{k-1}} < M)} \\ &= \frac{P(X_{t_{k-1}} < M \wedge \Delta X \geq M - X_{t_{k-1}} \wedge \Delta X < L - X_{t_{k-1}})}{P(X_{t_{k-1}} < M)} \\ &= \frac{\int_0^M P(\Delta X \geq M - \xi \wedge \Delta X < L - \xi) P(X_{t_{k-1}} = \xi) d\xi}{F_{(k-1)\Delta T}(M)} \\ &= \frac{\int_0^M (F_{\Delta T}(L - \xi) - F_{\Delta T}(M - \xi)) f_{(k-1)\Delta T}(\xi) d\xi}{F_{(k-1)\Delta T}(M)} \end{aligned}$$

Ceci induit que :

$$\begin{aligned} \mathbb{E} \left(\mathbb{I}_{\{L > X_{t_{N(S)}} \geq M\}} \right) &= F_{\Delta T}(L) - F_{\Delta T}(M) \\ &+ \sum_{k=2}^{\infty} \frac{\int_0^M (F_{\Delta T}(L - \xi) - F_{\Delta T}(M - \xi)) f_{(k-1)\Delta T}(\xi) d\xi}{F_{(k-1)\Delta T}(M)} \end{aligned} \quad (14)$$

Puis, de même manière, la probabilité de maintenance corrective sera :

$$\begin{aligned} \mathbb{E} \left(\mathbb{I}_{\{X_{t_{N(S)}} \geq L\}} \right) &= \sum_{k=1}^{\infty} P(X_{t_k} \geq L | X_{t_{k-1}} < M) \\ k=1 : P(X_{t_1} \geq L) &= F_{\Delta T}(L) \\ \forall k \geq 2 : \\ P(X_{t_k} \geq L | X_{t_{k-1}} < M) &= \frac{P(X_{t_k} \geq L \wedge X_{t_{k-1}} < M)}{P(X_{t_{k-1}} < M)} \\ &= \frac{P(\Delta X \geq L - X_{t_{k-1}} \wedge X_{t_{k-1}} < M)}{P(X_{t_{k-1}} < M)} \\ &= \frac{\int_0^M P(\Delta X \geq L - \xi) P(X_{t_{k-1}} = \xi) d\xi}{P(X_{t_{k-1}} < M)} \\ &= \frac{\int_0^M (1 - F_{\Delta T}(L - \xi)) f_{(k-1)\Delta T}(\xi) d\xi}{F_{(k-1)\Delta T}(M)} \end{aligned}$$

D'où :

$$\begin{aligned} \mathbb{E} \left(\mathbb{I}_{\{X_{t_{N(S)}} \geq L\}} \right) &= F_{\Delta T}(L) \\ &+ \sum_{k=2}^{\infty} \frac{\int_0^M (1 - F_{\Delta T}(L - \xi)) f_{(k-1)\Delta T}(\xi) d\xi}{F_{(k-1)\Delta T}(M)} \end{aligned} \quad (15)$$

Attention, parce que nous considérons S l'instant de la dernière inspection :

$$S = t_{N(S)} = N(S) \Delta T \Rightarrow \mathbb{E}(S) = \mathbb{E}(N(S)) \Delta T$$

Alors :

$$\begin{aligned} \mathbb{E}(C) &= \frac{\mathbb{E}(C(S))}{\mathbb{E}(S)} \\ &= \frac{c_i \mathbb{E}(N(S)) + c_p \mathbb{E} \left(\mathbb{I}_{\{L > X_{t_{N(S)}} \geq M\}} \right) + c_c \mathbb{E} \left(\mathbb{I}_{\{X_{t_{N(S)}} \geq L\}} \right)}{\mathbb{E}(N(S)) \Delta T} \\ &= \frac{c_i}{\Delta T} + \frac{1}{\Delta T} \frac{c_p \mathbb{E} \left(\mathbb{I}_{\{L > X_{t_{N(S)}} \geq M\}} \right) + c_c \mathbb{E} \left(\mathbb{I}_{\{X_{t_{N(S)}} \geq L\}} \right)}{\mathbb{E}(N(S))} \end{aligned}$$

Substituer par (13), (14) et (15) nous donnera le formule exacte de coût moyen par une durée de temps.

3.9 Optimisation de maintenance basée sur dégradations

```

1 # Monte carlo to optimize degradation
2 set.seed(2019)
3 L = 20
4 cost = c(10, 800, 1200) # Inspection, Preventive, Corrective
5 C_S = function(
6   p,
7   M
8 ) {
9   c = 0
10  p = p[!is.na(p)]
11  len = length(p)
12  if (p[[len]] >= M) {
13    n_s = length(p[p < M]) # Last index before >= M
14    c = c + cost[[1]] * n_s # Inspection
15    if (p[[n_s + 1]] >= L) {
16      c = c + cost[[3]] # Corrective
17    } else {
18      c = c + cost[[2]] # Preventive
19    }
20  }
21  return(c)
22 }
23 S = function(
24   p,
25   deltaT,
26   M
27 ) {
28   p = p[!is.na(p)]
29   len = length(p)
30   if (p[[len]] < M) return(0)
31   else return(length(p[p < M]) + 1)
32 }
33 E_C = function(

```

```

34 |     deltaT,
35 |     M,
36 |     nSim = 200,
37 |     nStep = floor(13 / deltaT),
38 |     silence = TRUE
39 | ) {
40 |     simStep = matrix(
41 |         rG(nSim * nStep, deltaT),
42 |         nrow = nSim,
43 |         ncol = nStep
44 |     )
45 |     simProc = t(apply(
46 |         t(simStep),
47 |         2,
48 |         cumsum
49 |     ))
50 |     # Count cost
51 |     simCost = apply(
52 |         t(simProc),
53 |         2,
54 |         C_S,
55 |         M
56 |     )
57 |     # Count time
58 |     simTime = apply(
59 |         t(simProc),
60 |         2,
61 |         S,
62 |         deltaT,
63 |         M
64 |     )
65 |     meanSimCost = mean(simCost[simCost > 0])
66 |     meanSimTime = mean(simTime[simTime > 0])
67 |     meanCostTime = 0
68 |     if (is.nan(meanSimCost) || is.nan(meanSimTime)) meanCostTime = .Machine$
69 |         double.xmax
70 |     else meanCostTime = meanSimCost / meanSimTime
71 |     # For debug purpose
72 |     if (!silence)
73 |         cat(
74 |             "dT =", deltaT,
75 |             "M =", M,
76 |             "mCost =", meanSimCost,
77 |             "mTime =", meanSimTime,
78 |             "mCostTime =", meanCostTime,
79 |             "\n"
80 |         )
81 |     return(meanCostTime)
82 | }
83 | # Optimize
84 | optim_degrad = function(
85 |     lower_deltaT,
86 |     upper_deltaT,
87 |     lower_M,
88 |     upper_M,
89 |     tol = 1e-2,
90 |     div_deltaT = 5,
91 |     div_M = 5,
92 |     cut_deltaT = 2,
93 |     cut_M = 2
94 | ){
95 |     deltaT = 0
96 |     M = 0
97 |     obj = 0
98 |     while (((upper_deltaT - lower_deltaT) / div_deltaT)^2 + ((upper_M -
99 |         lower_M) / div_M)^2) >= tol^2 {
100 |         interval_deltaT = seq(lower_deltaT, upper_deltaT, length.out = div_
101 |             deltaT)
102 |         interval_M = seq(lower_M, upper_M, length.out = div_M)
103 |         cat("deltaT =", interval_deltaT,
104 |             "\nM =", interval_M,

```

```

105         interval_M,
106         Vectorize(E_C)
107     )
108     # First minimum position
109     posMin = arrayInd(which.min(Z), dim(Z))
110     deltaT = interval_deltaT[[posMin[[1]]]]
111     M = interval_M[[posMin[[2]]]]
112     obj = Z[posMin]
113     cat("deltaT = ", deltaT,
114         "\nM = ", M,
115         "\nobj = ", obj,
116         "\n")
117     # Adjust zone of optimize
118     l_deltaT = upper_deltaT - lower_deltaT
119     l_M = upper_M - lower_M
120     upper_deltaT = min(deltaT + l_deltaT / cut_deltaT, upper_deltaT)
121     lower_deltaT = max(deltaT - l_deltaT / cut_deltaT, lower_deltaT)
122     upper_M = min(M + l_M / cut_M, upper_M)
123     lower_M = max(M - l_M / cut_M, lower_M)
124 }
125 return(list(
126     deltaT = deltaT,
127     M = M,
128     obj = obj
129 ))
130 }
131 val = optim_degrad(0.001,4,0,20)
132 # Illustration
133 interval_deltaT = seq(0.1, 4.1, 0.5)
134 interval_M = seq(1, 20, 1)
135 Z = outer(
136     interval_deltaT,
137     interval_M,
138     Vectorize(E_C)
139 )
140 library(rgl)
141 persp3d(
142     interval_deltaT,
143     interval_M,
144     Z,
145     xlab = "deltaT",
146     ylab = "M",
147     zlab = "E_C",
148     zlim = c(0, 400),
149     col = "lightblue"
150 )
151 rgl.snapshot(filename = "img/E_C_degrad.png")

```

Références

- [1] Minka, Thomas P. (2002). "Estimating a Gamma distribution"
<https://tminka.github.io/papers/minka-gamma.pdf>