



**HAL**  
open science

## Estimation du vieillissement et de l'efficacité de la maintenance d'actifs industriels à partir de données censurées par intervalle

Théo Cousino, Florent Brissaud, Leïla Marle, Laurent Doyen, Gaudoin Olivier

### ► To cite this version:

Théo Cousino, Florent Brissaud, Leïla Marle, Laurent Doyen, Gaudoin Olivier. Estimation du vieillissement et de l'efficacité de la maintenance d'actifs industriels à partir de données censurées par intervalle. Congrès Lambda Mu 23 “ Innovations et maîtrise des risques pour un avenir durable ” - 23e Congrès de Maîtrise des Risques et de Sécurité de Fonctionnement, Institut pour la Maîtrise des Risques, Oct 2022, Paris Saclay, France. hal-03878136

**HAL Id: hal-03878136**

**<https://hal.science/hal-03878136>**

Submitted on 29 Nov 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Estimation du vieillissement et de l'efficacité de la maintenance d'actifs industriels à partir de données censurées par intervalle

## Estimation of ageing and maintenance efficiency of industrial tools, considering interval censored data

COUSINO Théo

Research & Innovation Center  
for Energy (RICE), GRTgaz ; Univ.  
Grenoble Alpes, CNRS, Grenoble INP  
France  
[theo.cousino@univ-grenoble-alpes.fr](mailto:theo.cousino@univ-grenoble-alpes.fr)

BRISAUD Florent

Research & Innovation Center  
for Energy (RICE), GRTgaz  
France  
[florent.brissaud@grtgaz.com](mailto:florent.brissaud@grtgaz.com)

MARLE Leïla

Research & Innovation Center  
for Energy (RICE), GRTgaz  
France  
[leila.marle@grtgaz.com](mailto:leila.marle@grtgaz.com)

DOYEN Laurent

Univ. Grenoble Alpes, CNRS, Grenoble  
INP  
France  
[laurent.doyen@univ-grenoble-alpes.fr](mailto:laurent.doyen@univ-grenoble-alpes.fr)

GAUDOIN Olivier

Univ. Grenoble Alpes, CNRS, Grenoble  
INP  
France  
[olivier.gaudoin@univ-grenoble-alpes.fr](mailto:olivier.gaudoin@univ-grenoble-alpes.fr)

**Résumé** - La fiabilité d'un système dépend des effets conjoints du vieillissement et de l'efficacité des maintenances. Son évaluation nécessite l'application de modèles probabilistes et l'estimation statistique de leurs paramètres au vu des données. Les modèles de maintenance imparfaite considérés sont les modèles d'âge virtuel de type ARA (*Arithmetic Reduction of Age*). On s'intéresse ici au cas où les données sont censurées par intervalle : on sait qu'une défaillance s'est produite entre deux maintenances préventives, mais on ne sait pas exactement quand. Pour estimer les paramètres des modèles d'âge virtuel dans cette situation, cet article considère deux hypothèses différentes sur l'effet des maintenances correctives consécutives aux défaillances dont les dates sont censurées. Dans les deux cas, des pseudo dates de défaillance sont générées en remplacement des dates censurées. L'étude compare sur des données simulées la perte de qualité des estimations quand on utilise les dates fictives à la place des vraies dates.

**Mots clés** - *maintenance imparfaite, retour d'expérience, fiabilité, vieillissement, censure par intervalle, estimation statistique*

**Abstract** - The reliability of a system depends on the joint effects of ageing and maintenance efficiency. Its assessment requires the application of probabilistic models and the statistical estimation of their parameters from the observation of field data. The considered imperfect maintenance models are the virtual age models of type

ARA (*Arithmetic Reduction of Age*). We are interested here in the case where the data are interval censored: we know that a failure occurred between two preventive maintenance dates, but we do not know exactly when. To estimate the parameters of virtual age models in this situation, this article considers two different hypotheses on the effect of the corrective maintenance following failures whose dates are censored. In both cases, pseudo failure dates are generated to replace the censored dates. The study compares on simulated data the loss of quality of estimates when using the fictitious dates instead of the real dates.

**Keywords** - *imperfect maintenance, field feedback data, reliability, ageing, interval censoring, statistical estimation*

### I. INTRODUCTION

GRTgaz possède et exploite en France le plus long réseau de transport de gaz naturel à haute pression d'Europe. Ses actifs industriels incluent plus de 32.500 km de canalisation, 26 stations de compression, environ 5.000 postes de détente, près de 4.800 postes de sectionnement/coupeure et des postes d'injection de biométhane en fort développement. La mission de GRTgaz est d'acheminer, pour le compte des expéditeurs et dans des conditions de coût et de sécurité optimales, le gaz

naturel aux distributeurs et clients industriels raccordés au réseau de transport. RICE (*Research and Innovation Center for Energy*) est la direction de GRTgaz en charge de la R&D. Les activités de RICE s'organisent autour des enjeux d'excellence opérationnelle, de sécurité industrielle et de transition énergétique et écologique (tels que l'hydrogène et les biométhanés). Depuis plusieurs années, RICE développe des outils de modélisation des actifs industriels afin d'optimiser leur gestion, notamment en termes de politiques de maintenance.

Les actifs industriels de GRTgaz sont soumis à des phénomènes de vieillissement et subissent tout au long de leur vie des défaillances et des actions de maintenance corrective et préventive. La maintenance corrective (MC) survient après la détection d'une défaillance et a pour but de remettre le matériel en état de fonctionner. La maintenance préventive (MP) a pour objectif de détecter d'éventuelles défaillances, de ralentir le vieillissement du matériel et ainsi de retarder l'occurrence de prochaines défaillances. La fiabilité d'un système dépend des effets conjoints du vieillissement et de l'efficacité des maintenances [1].

Le vieillissement intrinsèque est modélisé à l'aide de lois de probabilité pour la durée de fonctionnement d'un système non maintenu. La plus usuelle d'entre elles est la loi de Weibull. Une maintenance « parfaite » est telle que le système après maintenance est « comme neuf » (*AGAN : As Good As New*). À l'opposé, une maintenance est dite « minimale » si le système après maintenance est dans le même état d'« usure » qu'avant maintenance (*ABAO : As Bad As Old*). En pratique, l'effet d'une maintenance est souvent compris entre ces deux extrêmes, on parle alors de maintenance imparfaite.

De nombreux modèles de maintenance imparfaite existent dans la littérature. Parmi eux, GRTgaz a retenu comme modèles pertinents pour ses applications industrielles les modèles d'âge virtuel *ARA* (*Arithmetic Reduction of Age*) développés par le Laboratoire Jean Kuntzmann (LJK) [2, 3]. Ces modèles supposent que l'effet des maintenances est de « rajeunir » les matériels, c'est-à-dire de les remettre dans un état de santé meilleur que leur état présent. GRTgaz exploite notamment ces modèles dans des modélisations en réseaux de Petri pour optimiser ses politiques de maintenance [4, 5].

L'évaluation de la fiabilité des matériels et l'optimisation de leur maintenance nécessite l'application de ces modèles et l'estimation statistique de leurs paramètres, au vu des données de défaillances et de maintenances correctives et préventives. L'outil logiciel VAM [6] a été développé dans ce but par le LJK et a été utilisé par GRTgaz pour certains actifs industriels de transport du gaz, telles que les soupapes des postes de détente. Cependant, certaines caractéristiques des actifs industriels de GRTgaz ne sont pas prises en compte par ces modèles d'estimation et les méthodes d'inférence classiques, notamment la censure des données. Le traitement des données censurées est classique quand les données sont des réalisations de variables aléatoires indépendantes et de même loi [7]. Ça l'est beaucoup moins quand on s'intéresse à des processus aléatoires ponctuels, et en particulier aux modèles d'âge virtuel.

Il existe plusieurs types de censure dans les différents cas de données de GRTgaz, cet article traite de la censure par

intervalle. C'est le cas quand, par exemple, on sait qu'une défaillance s'est produite entre deux maintenances préventives, mais on ne sait pas exactement quand. La maintenance corrective est alors réalisée conjointement à la maintenance préventive qui a permis de détecter la défaillance. Par exemple, le cas des soupapes pour lesquelles le mode de défaillance « non-ouverture » ne peut être détecté que par des maintenances préventives périodiques, généralement effectuées chaque année. Ce problème a été abordé dans le cas de maintenances minimales [8], mais n'a pas encore été traité pour des maintenances imparfaites. Il est donc important de proposer des méthodes permettant d'estimer les paramètres des modèles d'âge virtuel quand les données sont censurées par intervalle.

Dans cet article, nous faisons deux hypothèses différentes sur l'effet des maintenances correctives consécutives aux défaillances dont les dates sont censurées. Dans le premier cas, on suppose qu'une MC de type *ARA* est faite immédiatement après la défaillance. Dans le second cas, plus proche de la réalité, on suppose que la MC correspondant à la défaillance censurée est effectuée à la date de la prochaine MP. Dans les deux cas, on génère des pseudo dates de défaillance en remplacement des dates censurées, selon deux règles différentes (déterministe et aléatoire). Nous comparons sur des données simulées la perte de qualité des estimations quand on utilise les dates fictives à la place des vraies dates.

La section II présente les modèles d'âge virtuel *ARA* dans un cadre sans censure, ainsi que la manière de les simuler et d'estimer leurs paramètres. La section III décrit les deux méthodes de génération des pseudo-dates de défaillance et les deux hypothèses de modélisation effectuées pour le traitement de la censure. La section IV présente et analyse les résultats de simulation obtenus à l'aide de VAM.

## II. CADRE MATHÉMATIQUE

### A. Les modèles d'âge virtuel *ARA*

Un actif industriel est mis en fonctionnement à la date 0, il subit des défaillances et des actions de MC à des dates aléatoires juste après les défaillances, et des actions de MP à des dates planifiées. On suppose que les durées de maintenances sont négligeables. Les notations utilisées dans cet article sont, pour tout  $i \geq 1$  et  $t \geq 0$  :

- $K_t$ , le nombre de maintenances (correctives et préventives) entre 0 et  $t$  compris.
- $C_i$ , le  $i$ -ème instant de maintenance (corrective ou préventive) avec  $C_0 = 0$ .
- $U_i$ , le type de la  $i$ -ème maintenance (corrective ou préventive), 1 si MP et 0 si MC.
- $W_i = C_i - C_{i-1}$ , la  $i$ -ème durée inter-maintenances (correctives et préventives).
- $T_i$ , le  $i$ -ème instant de défaillance.
- $M_t$ , le nombre de MP entre 0 et  $t$ .
- $N_t$ , le nombre de MC entre 0 et  $t$ .
- $\rho_p$ , l'efficacité des MP.
- $\rho_c$ , l'efficacité des MC.

- $\alpha$ , paramètre d'échelle de l'intensité initiale.
- $\beta$ , paramètre de forme de l'intensité initiale.
- par convention, on notera avec des minuscules les réalisations des variables aléatoires ci-dessus.

Illustrons un exemple représentant l'évolution des défaillances.

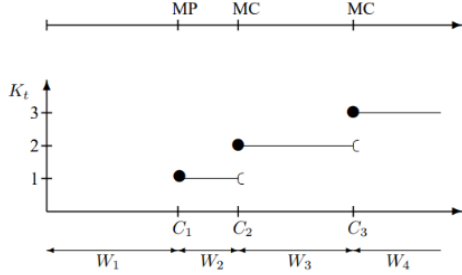


Fig 1 – Modélisation du processus des défaillances

Les modèles d'âge virtuel [9] considèrent que chaque maintenance réduit l'âge virtuel d'une quantité dépendante du passé du processus de défaillances et d'un paramètre d'efficacité. L'âge virtuel repose sur le fait qu'après chaque maintenance le système « rajeunit », ce qui influence sa durée de vie passée (d'où la notion d'âge « virtuel »). A la date  $t$ , le système maintenu est équivalent à un système neuf non maintenu d'âge  $V(t)$  où  $V(t)$  est l'âge virtuel du système à la date  $t$ .

Posons,  $\forall 0 \leq t < C_1, V_0(t) = t$ , l'âge virtuel à la date  $t$  avant la première maintenance. Il est, par hypothèse, identique au temps calendaire écoulé depuis la mise en service. Posons également  $V_i(t)$ , l'âge virtuel à la date  $t$  entre la  $i$ -ème et la  $(i+1)$ -ème maintenance. Parmi les modèles d'âge virtuel, les modèles ARA (*Arithmetic Reduction of Age*) [2, 3] ont un âge virtuel qui évolue comme le temps calendaire entre deux maintenances. Ils vérifient donc  $V_i(t) = V_i(C_i) + t - C_i$ . Plusieurs sous modèles existent, ici nous nous concentrons sur les modèles  $ARA_1$  et  $ARA_\infty$ .

Dans le modèle MC  $ARA_1$ -MP  $ARA_1$ , l'effet des maintenances préventives et correctives est de réduire l'âge virtuel d'une quantité proportionnelle au temps écoulé depuis la dernière maintenance. Ce coefficient de réduction diffère selon que la maintenance est corrective ( $\rho_c$ ) ou préventive ( $\rho_p$ ). Nous avons :  $\forall i \geq 1$ ,

$$V_i(C_i) = V_{i-1}(C_{i-1}) + W_i - \rho_p^{U_i} \rho_c^{1-U_i} W_i \quad (1)$$

Dans le modèle MC  $ARA_\infty$ -MP  $ARA_\infty$ , l'effet des maintenances préventives et correctives est de réduire globalement l'âge virtuel proportionnellement à sa valeur juste avant la maintenance. Nous avons :  $\forall i \geq 1$ ,

$$V_i(C_i) = (V_{i-1}(C_{i-1}) + W_i) (1 - \rho_p)^{U_i} (1 - \rho_c)^{1-U_i} \quad (2)$$

Dans les deux exemples ci-dessus, les MC et MP sont supposées de même type ( $ARA_1$  ou  $ARA_\infty$ ). Il est possible de supposer qu'elles sont de types différents, par exemple MC  $ARA_1$ -MP  $ARA_\infty$ .

Notons que les modèles MC *AGAN*-MP *AGAN* et MC *ABAO*-MP *ABAO* correspondent respectivement à  $\rho_c = \rho_p = 1$  et  $\rho_c = \rho_p = 0$ .

L'intensité conditionnelle de défaillance  $\lambda_t$ , qui caractérise complètement la distribution du processus de défaillance [1], est :

$$\lambda_t = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} P(N_{t+\Delta t} - N_t = 1 | \mathcal{H}_t^-) \quad (3)$$

où  $\mathcal{H}_t^-$  est le passé du processus, c'est-à-dire la filtration naturelle engendrée par les événements arrivés juste avant  $t$  non compris. La fonction de répartition conditionnelle des instants successifs de défaillance peut s'exprimer en fonction de l'intensité conditionnelle de défaillance, comme l'indique la propriété suivante :  $\forall i \geq 1$ ,

$$P(T_{N_{C_i}+1} \leq t | \mathcal{H}_{C_i}^-) = 1 - \exp\left(-\int_{C_i}^t \lambda_s ds\right) \quad (4)$$

où  $T_{N_{C_i}+1}$  est le prochain instant de défaillance après la  $i$ -ème maintenance [1]. À partir de cette quantité, nous allons pouvoir caractériser les modèles de maintenances. Dans les modèles ARA, l'intensité conditionnelle de défaillance est de la forme :  $\forall C_i \leq t \leq C_{i+1}$

$$\lambda_t = h(V_i(C_i) + t - C_i) \quad (5)$$

Où  $h$  est le taux de défaillance du système neuf, que l'on nomme l'intensité initiale. Nous ferons l'hypothèse que la date de la première défaillance suit une loi de Weibull où  $h(t) = \alpha \beta t^{\beta-1}$ . Notons que cette paramétrisation est différente de la paramétrisation usuelle où  $h(t) = \frac{\beta}{\eta} \left(\frac{t}{\eta}\right)^{\beta-1}$ .

Les deux modèles précédemment introduits vérifient donc  $\forall t \geq 0$  [3] :

- dans le modèle MC  $ARA_1$ -MP  $ARA_1$ ,

$$\lambda_t = h\left(t - \sum_{k=1}^{K_t^-} \rho_p^{U_k} \rho_c^{1-U_k} W_k\right) \quad (6)$$

- dans le modèle MC  $ARA_\infty$ -MP  $ARA_\infty$ ,

$$\lambda_t = h\left(t - C_{K_t^-} + \sum_{i=1}^{K_t^-} (1 - \rho_p)^{(M_{t^-}) - (M_{C_{i-1}})} (1 - \rho_c)^{(N_{t^-}) - (N_{C_{i-1}})} W_i\right) \quad (7)$$

Dans la suite, nous allons proposer plusieurs méthodes pour estimer les paramètres  $(\alpha, \beta, \rho_c, \rho_p)$  lorsqu'on a des données censurées par intervalle.

## B. Simulation

Nous pouvons simuler des jeux de données ayant des caractéristiques proches des données rencontrées en pratique en tenant compte des différents modèles de maintenances.

Nous utilisons le principe de simulation par inversion de la fonction de répartition. Après la  $i$ -ème maintenance ayant eu lieu à la date  $C_i$ , si l'on veut simuler le prochain instant de défaillance  $T_{N_{C_i}+1}$  pour tout  $i \geq 0$ , on doit prendre en compte tout le passé du processus et ainsi calculer l'inverse de  $P(T_{N_{C_i}+1} \leq t | \mathcal{H}_{C_i}^-)$  avec  $t > C_i$ . Il suffit d'appliquer la fonction ainsi calculée à une variable aléatoire de loi uniforme sur  $[0, 1]$  pour obtenir une méthode de simulation du prochain instant potentiel de défaillance. Soit  $U \sim U[0, 1]$ , alors nous avons l'égalité en loi :

$$\widetilde{T}_{N_{C_i}+1} = V_i^{-1}\left(H^{-1}(H(V_i(C_i)) - \log(U))\right) \quad (8)$$

où  $\tilde{T}$  est la date potentielle de défaillance et  $H(t) = \int_0^t h(s) ds$  pour tout  $t \geq 0$ . Puisque l'âge virtuel entre la  $i$ -ème et la  $(i+1)$ -ème maintenance est croissant, son inverse généralisé existe.  $V_i^{-1}(\cdot)$  est l'inverse généralisé de  $V_i(\cdot)$  et  $H^{-1}(\cdot)$  est l'inverse généralisé de  $H(\cdot)$ . Cette potentielle date de défaillance sera effectivement observée uniquement si aucune maintenance préventive n'est planifiée avant, on aura alors  $C_{i+1} = \widetilde{T}_{N_{C_i+1}}$ . Cependant, dans le cas contraire, le prochain instant de maintenance sera une maintenance préventive en  $C_{i+1}$  et il faudra alors re-simuler une nouvelle potentielle date de défaillance en prenant en compte en plus dans l'histoire du processus cette nouvelle date de maintenance préventive en  $C_{i+1}$ .

### C. Maximum de vraisemblance

Pour estimer les paramètres  $(\alpha, \beta, \rho_c, \rho_p)$ , nous utilisons le maximum de vraisemblance. Il convient d'être prudent ici, car il ne s'agit pas de défaillances indépendantes, nous ne pouvons donc pas écrire la vraisemblance comme produit des densités marginales, mais comme produit des densités conditionnelles.

En l'absence de censure par intervalle, la vraisemblance associée à l'observation de  $n$  instants de maintenance entre 0 et  $t$  est :

$$L_t(\alpha, \beta, \rho_p, \rho_c; n, \mathbf{c}_n, \mathbf{u}_n) = \prod_{i=1}^n \left( \lambda_{c_i}^{\mathbb{1}_{\{u_i=0\}}} \exp\left(-\int_{c_{i-1}}^{c_i} \lambda_s ds\right) \right) \exp\left(-\int_{c_n}^t \lambda_s ds\right) \quad (9)$$

$$\mathbf{c}_n = (c_1, \dots, c_n) \text{ et } \mathbf{u}_n = (u_1, \dots, u_n) \quad [3]$$

Les paramètres d'un modèle d'âge virtuel peuvent être estimés en maximisant cette dernière quantité *via* la méthode du gradient.

### III. PRISE EN COMPTE DE LA CENSURE PAR INTERVALLE

Dans notre cas d'étude, les dates de défaillance ne sont pas observées. La seule information disponible est le nombre de défaillances survenues entre deux dates de MP successives, ce qui correspond à une censure par intervalle. En pratique, pour les données GRTgaz, il n'est pas pertinent de considérer plusieurs défaillances dans un intervalle inter-MP.

Pour prendre en compte la censure, on génère de pseudo dates de défaillance en remplacement des dates censurées et on estime les paramètres par maximum de vraisemblance, en faisant comme si ces pseudo dates étaient de vraies dates de défaillance. Par ailleurs, nous faisons deux hypothèses différentes sur l'effet des maintenances correctives consécutives aux défaillances dont les dates sont censurées. L'impact de la censure par intervalle sera évalué en quantifiant, à l'aide de simulations de Monte Carlo, la perte de qualité des estimations quand on utilise les dates fictives à la place des vraies dates.

#### A. Generation des pseudo dates de défaillance

On génère des pseudo dates de défaillance en remplacement des dates censurées selon deux règles différentes.

- Première règle (déterministe) : On place de façon équirépartie les dates manquantes de défaillance dans l'intervalle inter-MP correspondant.

- Seconde règle (aléatoire) : Quand il n'y a qu'une date manquante dans un intervalle inter-MP, on tire aléatoirement une pseudo-date selon une loi uniforme dans cet intervalle. S'il y a plusieurs dates manquantes, la première est tirée de cette façon. Pour refléter la croissance de l'intensité initiale, la seconde est tirée selon une loi uniforme entre cette première date et la date de la MP suivante. Et ainsi de suite.

La figure 2 illustre les deux règles précédentes.

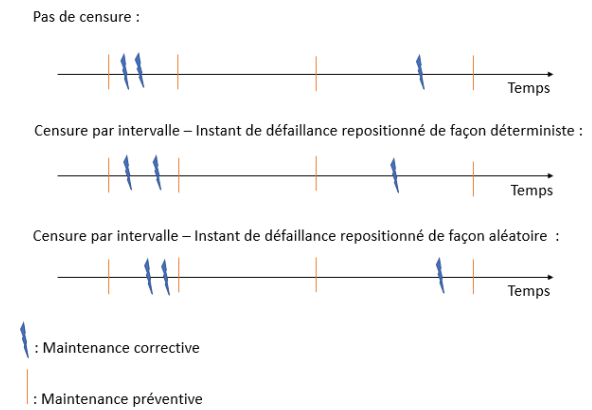


Fig 2 - Illustration des deux règles de génération des pseudo-dates de défaillance

#### B. Première hypothèse de modélisation

Dans un premier temps, on suppose qu'une MC de type ARA est faite immédiatement après chaque défaillance. Comme les dates de défaillance ne sont pas observées, on les remplace par les pseudo-dates telles qu'expliquées ci-dessus. Cela revient donc à déplacer l'effet des MC des vraies dates inconnues aux pseudo-dates générées.

#### C. Seconde hypothèse de modélisation

Dans la réalité industrielle de GRTgaz, si par exemple une soupape d'une unité industrielle tombe en panne à une date inconnue, elle n'est réparée qu'au moment où on détecte la panne, c'est-à-dire à l'instant de la prochaine MP. En attendant d'être réparée, elle continue de se détériorer. Cela nous amène à considérer une seconde hypothèse de modélisation dans laquelle on distingue deux types de MP. Par conséquent, on sort du cadre de la section II.A. Cette deuxième hypothèse est définie de la manière suivante :

- S'il n'y a pas eu de défaillance dans un intervalle inter-MP, la MP qui suit est de type 1 ( $U_i = 1$ ).

- S'il y a eu une défaillance, il n'y a pas de réparation immédiate et le système continue de se détériorer, ce que l'on modélise par une MC ABAO ( $U_i = 0$ ) à la (pseudo)date de défaillance. La MC associée à cette défaillance est effectuée à la date de la prochaine MP. L'effet de cette MC se cumule avec celui de la MP, ce qui est modélisé par une MP d'un autre type, le type 2 ( $U_i = 2$ ).

Cette hypothèse de modélisation amène à construire un nouveau modèle d'âge virtuel qui généralise le modèle présenté en section II.A au sens où il y a maintenant 3 types de maintenances  $U_i$ , un type de MC et deux types de MP. Comme



la MC est supposée ABAO, ce modèle a néanmoins autant de paramètres que le précédent. Les efficacités des deux types de MP sont notées  $\rho_1$  et  $\rho_2$ .

Par exemple, l'intensité de défaillance du modèle MC ABAO-MP1 ARA<sub>1</sub>-MP2 ARA<sub>1</sub> est  $\forall t \geq 0$ ,

$$\lambda_t = h \left( t - \sum_{k=1}^{K_t^-} \rho_1 \mathbb{1}_{\{U_k=1\}} \rho_2 \mathbb{1}_{\{U_k=2\}} W_k \right) \quad (10)$$

La figure 3 illustre cette nouvelle hypothèse avec les deux types de MP représentés par des couleurs différentes.

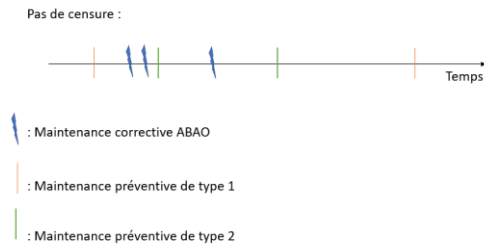


Fig 3 - Illustration de la deuxième hypothèse de modélisation

#### IV. RESULTATS OBTENUS

Nous évaluons la performance des méthodes proposées en comparant par simulations de Monte-Carlo la perte de qualité des estimations quand on utilise les dates fictives à la place des vraies dates. Un jeu de données simulé correspond à 100 systèmes identiques et indépendants (dans la réalité, un système correspond à une unité industrielle) observés sur une durée de 20 ans. Chaque système a les mêmes dates de MP, dont la période est de un an. En fonction du modèle et des valeurs des paramètres choisis, le pourcentage de MC et MP par jeu de données varie. On estime les paramètres sur ce jeu de données en effectuant le produit des 100 vraisemblances associées à chacun des systèmes. On répète la procédure 5000 fois afin d'obtenir 5000 valeurs des estimations de chaque paramètre. Dans les algorithmes de maximisation de la vraisemblance, les valeurs des paramètres sont contraintes : entre 0,1 et 5 pour  $\beta$  et entre -10 pour 1 pour les efficacités de maintenance afin d'éviter des problèmes de convergence. On s'intéresse aux distributions empiriques des estimations.

Nous avons testé toutes les combinaisons correspondantes à :

- Effets de MC et MP de type ARA<sub>1</sub> et ARA<sub>∞</sub>.
- Valeurs des paramètres d'efficacité de maintenance pour les MC et MP : 0.25 et 0.75.
- Valeurs de  $\beta$  : 1.5, 2, 2.5, 3.
- La valeur de  $\alpha$  est choisie de façon à ce que la date moyenne de la première défaillance est proche de 5 ans.

Les sous-sections suivantes présentent deux exemples de résultats pour chaque hypothèse de modélisation.

##### A. Première hypothèse de modélisation

Le cas étudié pour la première hypothèse de modélisation est celui du modèle MC ARA<sub>1</sub>-MP ARA<sub>1</sub>. Pour les valeurs des paramètres choisies, un jeu de données contient en moyenne 953 MC, ce qui représente environ un tiers du nombre total de maintenances.

Le tableau 1 présente les moyennes empiriques des 5000 estimations. La figure 4 présente les caractéristiques des distributions empiriques des estimations sous forme de boxplots. Ils nous donnent des informations sur les quartiles afin de visualiser comment sont réparties les différentes valeurs possibles des estimations des paramètres du modèle. Les traits horizontaux représentent la vraie valeur des paramètres. Sur chaque graphique, nous visualisons 3 boîtes à moustache. La 1ère correspond au cas sans censure, la 2ème au cas avec censure et repositionnement déterministe, la 3ème au cas avec censure et repositionnement aléatoire.

Modèle de maintenance	MC ARA <sub>1</sub> -MP ARA <sub>1</sub> .			
	$\alpha$	$\beta$	$\rho_c$	$\rho_p$
Vraies valeurs des paramètres	$5 \times 10^{-7}$	2	0.25	0.75
Moyenne empirique des 5000 estimations				
Sans censure	$5 \times 10^{-7}$	2.1	0.16	0.72
Censure par intervalle et Repositionnement déterministe	$1 \times 10^{-7}$	2.3	0.77	0.89
Censure par intervalle et Repositionnement aléatoire	$9 \times 10^{-7}$	1.9	-0.54	0.57

Table 1 – Moyennes empiriques des estimations des paramètres pour un modèle MC ARA<sub>1</sub>-MP ARA<sub>1</sub>

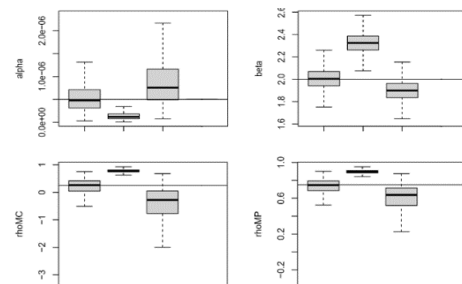


Fig 4 – Distribution empirique des estimations

Sans surprise, les meilleures estimations sont obtenues dans le cas sans censure et elles sont de très bonne qualité. On voit clairement que les repositionnements aléatoire et déterministe donnent de mauvais résultats. Dans le cas déterministe, les estimateurs sont fortement biaisés. Dans le cas aléatoire, ils sont moins biaisés mais leur variance est beaucoup plus forte.

##### B. Seconde hypothèse de modélisation

Le cas étudié pour la deuxième hypothèse de modélisation est celui du modèle MC ABAO - MP1 ARA<sub>∞</sub>- MP2 ARA<sub>1</sub>. Pour les valeurs des paramètres choisis, un jeu de données contient en moyenne 1010 MC, représentant un tiers du nombre total de maintenances.

Le tableau 2 présente les moyennes empiriques des 5000 estimations. La figure 5 présente les caractéristiques des distributions empiriques des estimations par le biais de boxplots.

Modèle de maintenance	MC ABAO -MP1 ARA <sub>∞</sub> -MP2 ARA <sub>1</sub> .			
	$\alpha$	$\beta$	$\rho_1$	$\rho_2$
Vraies valeurs des paramètres	$5 \times 10^{-7}$	2	0.25	0.75
Moyenne empirique des 5000 estimations				
Sans censure	$6 \times 10^{-7}$	1.9	0.24	0.71
Censure par intervalle et Repositionnement déterministe	$1 \times 10^{-6}$	1.9	0.23	0.22
Censure par intervalle et Repositionnement aléatoire	$7 \times 10^{-7}$	2.0	0.27	0.64

Table 2 – Moyennes empiriques des estimations des paramètres pour un modèle MC ABAO -MP1 ARA<sub>∞</sub>-MP2 ARA<sub>1</sub>

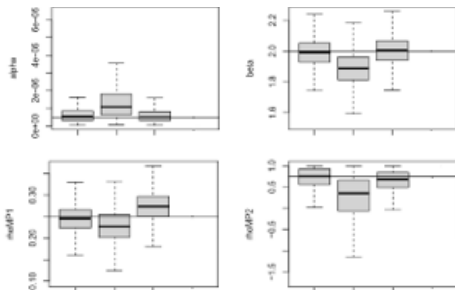


Fig 5 – Distribution empirique des estimations

Comme pour la précédente hypothèse de modélisation, les meilleures estimations sont obtenues dans le cas sans censure. Mais cette fois, les estimations pour le repositionnement aléatoire sont presque aussi bonnes, sauf pour  $\rho_1$ . Le repositionnement déterministe donne de moins bons résultats.

## V. CONCLUSION

L'ensemble des expérimentations effectuées tend à montrer que les estimations sont de bien meilleure qualité pour la deuxième hypothèse de modélisation où les effets des MC sont reportés que pour la première, et meilleures pour le repositionnement aléatoire que pour le repositionnement déterministe. La qualité des estimations est liée au grand nombre de systèmes et au grand nombre de défaillances dans chaque jeu de données simulé (mais qui sont représentatifs de la réalité industrielle dans le cas étudié). Dans le cas où on aurait beaucoup moins de données, on pourrait avoir recours à d'autres méthodes d'estimation, telles que l'estimation bayésienne [10].

Cette première étude a consisté à remplacer les dates manquantes par de pseudo-dates de défaillance. Pour éviter de générer des pseudos-dates, on peut prendre en compte les données manquantes à l'aide d'une méthode de type *Expectation-Maximization* (EM) [8]. On peut également calculer une fonction de vraisemblance qui prenne directement en compte la censure.

Pour un cas d'étude réel, un autre problème sera de choisir à la vue des données le modèle d'effet de maintenance, la distribution du premier instant de défaillance et les possibles covariables les mieux adaptés. Des travaux sont en cours afin d'effectuer ce choix via des critères comme l'AIC (Akaike Information Criterion).

## REMERCIEMENTS

Ce travail a été en partie soutenu par le LabEx PERSYVAL-Lab (ANR-11-LABX-0025-01) financé par le programme français Investissement d'avenir.

## VI. BIBLIOGRAPHIE

- [1] Lindqvist B.H. (2006). On the statistical modeling and analysis of repairable systems. *Statistical Science*, 21 (4), pp 532-551.
- [2] Doyen L., Gaudoin O. (2004). Classes of imperfect repair models based on reduction of failure intensity or virtual age.
- [3] Doyen L., Gaudoin O. (2011). Modelling and assessment of ageing and efficiency of corrective and planned preventive maintenance. *IEEE Transactions on Reliability*, 60 (4), pp 759-769.
- [4] Brissaud F., Folleau C. (2020). Modeling Imperfect Maintenance with Stochastic Petri Nets. *Proceedings of the 30th European Safety and Reliability Conference and the 15th Probabilistic Safety Assessment and Management Conference*.
- [5] Brissaud F., Folleau C., De Cournaud B. (2020). Modélisation de maintenances imparfaites avec réseaux de Petri stochastiques à prédicats. *Actes du 22ème Congrès de Maîtrise des Risques et de Sécurité de Fonctionnement*.
- [6] Doyen L., Drouilhet R. VAM, an R package for Virtual Age Models. <https://rpackages.dyndoc.fr/VAM>
- [7] Klein J.P., Moeschberger M.L. (2003). *Survival analysis: techniques for censored and truncated data*. Springer.
- [8] Peng Y., Wang Y., Zi Y.Y., Tsui K.L., Zhang C. (2017). Dynamic reliability assessment and prediction for repairable systems with interval-censored data. *Reliability Engineering & System Safety*, 159, pp 301-309.
- [9] Kijima, M. (1989). Some results for repairable systems with general repair. *Journal of Applied Probability*, 26, pp 89-102.
- [10] Corset F., Doyen L., Gaudoin O., (2012). Bayesian Analysis of ARA Imperfect Repair Models. *Communication in Statistics - Theory and Methods*.