

Inférence statistique pour un modèle markovien de dégradation avec covariables dépendantes du temps *

Title: Statistical inference of a discrete-time Markovian degradation model with time-dependent covariates

Christian Paroissin¹, Emmanuel Remy² et Véronique Verrier²

Résumé : On propose un modèle markovien de dégradation à temps discret. L'espace d'états décrit l'évolution de la dégradation d'un composant industriel. Des covariables sont intégrées dans les probabilités de transition comme dans un modèle de régression logistique. Si les covariables dépendent du temps, alors le modèle devient non-homogène en temps. On s'intéresse ici à l'inférence statistique pour un tel modèle sur la base de l'observation de plusieurs composants à un ou plusieurs instants et de l'observation des covariables à chaque unité de temps. On s'intéresse également au problème de la sélection des covariables influentes pour chacune des probabilités de transition.

Abstract: We propose a discrete-time Markov chain over a finite state space to describe the degradation evolution of an industrial component. Covariates are integrated in the transition probabilities as in a logistic regression. If covariates evolve with time, the model turns to be time-inhomogeneous. Statistical inference for such a model is considered here for a fleet of components observed at several times when covariates are measured at each unit of discrete time. The problem of determining which covariates impact a given transition probability (called hereafter covariates selection) is also discussed.

Mots-clés : inspections visuelles, modèle multi-états, sélection de covariables

Keywords: covariates selection, multi-states model, visual inspection data

Classification AMS 2000 : 62M05, 62N05

1. Description du problème

La méthodologie proposée a pour objectif de mieux comprendre et prévoir le vieillissement de composants installés sur des centrales de production d'électricité d'EDF. Ces composants assurent une fonction d'étanchéité. S'il y a perte de la fonction d'étanchéité (on parlera alors de "panne"), des jets de vapeur peuvent survenir et endommager les composants situés autour. De fait, des inspections visuelles sont régulièrement effectuées sur ces composants pour contrôler leur état : on dispose ainsi d'une information qualitative sur le niveau de dégradation des composants, qui sont immédiatement remplacés lorsqu'ils tombent en panne ou lorsqu'ils sont constatés trop dégradés.

* Ce travail a reçu le soutien de l'Agence Nationale de la Recherche (ANR) dans le cadre du projet AMMSI, sous le numéro ANR-2011-BS01-021.

¹ Université de Pau et des Pays de l'Adour, Pau, France.

E-mail : cparoiss@univ-pau.fr

² EDF R&D, Chatou, France.

E-mail : emmanuel.remy@edf.fr and E-mail : veronique.verrier@edf.fr

Dans cet article, le modèle proposé est décrit dans le cadre le plus général possible, le cas d'étude étant plus spécifique. Durant les inspections, le niveau de dégradation des composants est classé selon une échelle ordinale E (i.e. E est un ensemble fini et muni d'une relation d'ordre). Sans perte de généralité, on pose $E = \{0, \dots, q\}$ où 0 correspond à l'état de parfait fonctionnement, q à l'état de panne et les autres états à des niveaux intermédiaires de dégradation (dans le cas d'étude, on a $q = 3$). On suppose que la date de fin de l'étude ne coïncide pas avec une date d'inspection : on a donc une censure à droite pour tous les composants encore en marche à la date de fin de l'étude. Par ailleurs, un niveau de dégradation pour un composant sera associé soit à une date d'inspection (y compris la date de fin de l'étude), soit à un instant de panne. Enfin, une politique de remplacement \mathcal{R} est appliquée : à chaque niveau de dégradation est associée une stratégie de remplacement. Pour le cas d'étude, la politique de remplacement est résumée dans le tableau 1. Ce tableau montre bien la différence entre les états 1 et 2. En effet, un composant dans

TABLEAU 1. Niveaux de dégradation et stratégie de remplacement associée

État	Niveau de dégradation	Remplacement ?
0	Parfait état	Non
1	Dégradation mineure	Non
2	Dégradation majeure	Oui
3	Panne	Oui

l'état 1 signifie que, durant l'inspection, une dégradation a été observée mais qu'elle n'a pas été jugée suffisamment avancée pour que ce composant soit remplacé de façon préventive (l'équipe de maintenance a considéré que le composant ne tombera probablement pas en panne avant la prochaine inspection). Inversement, un composant dans l'état 2 signifie que, durant l'inspection, une dégradation critique a été observée entraînant son remplacement immédiat par un composant neuf.

Les modèles multi-états (markoviens ou semi-markoviens) sont utilisés pour des études en fiabilité lorsqu'on dispose de mesures qualitatives de la dégradation d'un composant (voir, par exemple, Kallen (2009) et Kallen and van Noordwijk (2006)). Ces modèles sont utilisés principalement en biostatistique et en épidémiologie (Hougaard, 1999; Commenges, 1999; Andersen and Keiding, 2002), mais également dans d'autres domaines comme, par exemple, l'ingénierie (Guo et al., 2010; Natvig and Mørch, 2003). Bien que des analogies puissent être établies, il existe des différences de contexte (par exemple, du fait de la considération de politiques de remplacement ou de maintenance en fiabilité). Cela implique qu'une méthode statistique pour un modèle multi-états adaptée pour l'un des domaines ne puisse pas nécessairement être applicable (du moins, directement) dans un autre domaine.

On propose de modéliser l'évolution de la dégradation de ce type de composant par une chaîne de Markov à temps discret et à valeurs dans E . On considérera d'abord le cas homogène, i.e. le cas où les probabilités de transition ne dépendent pas du temps. Un tel modèle fait donc intervenir q paramètres (on rappelle qu'ici $q = 3$) correspondant aux probabilités de transition en un pas de temps d'un état de dégradation vers l'état plus dégradé suivant (on notera que l'état de "panne" q est absorbant). Cependant, les probabilités de transition peuvent dépendre de certaines covariables qui varient dans le temps (ces covariables peuvent correspondre, par exemple, aux modes

d'exploitation de la centrale). Dans notre étude, ces covariables sont observées mensuellement, et c'est la raison pour laquelle nous avons considéré un modèle de dégradation à temps discret. Les covariables sont intégrées dans les probabilités de transition comme dans une régression logistique, et ce second modèle devient non-homogène en temps. Comme mentionné précédemment, des modèles similaires ont déjà été utilisés dans la littérature (Kallen, 2009; Kallen and van Noortwijk, 2006), mais il s'agit la plupart du temps de modèle à temps continu.

Pour le modèle proposé (section 2) et pour le type de données manipulées (voir ci-dessus), on écrit la fonction de vraisemblance (section 3) qui permet ensuite de calculer numériquement un estimateur des paramètres du modèle. Les temps de séjour dans les états n'étant pas observés, cela induit que les durées sont censurées (à gauche, à droite ou par intervalle), à l'exception des durées jusqu'à panne (i.e. temps d'atteinte de l'état q). Les différents modèles (définis selon les covariables incluses dans les probabilités de transition) seront comparés à l'aide du critère d'information d'Akaike (AIC). Enfin, on présente une application du modèle sur des données simulées (uniquement pour le modèle sans covariable), puis au jeu de données EDF (section 4) pour lequel une discussion sur la sélection du meilleur sous-modèle (au regard de l'AIC) est proposée.

2. Modèle markovien à temps discret

Nous considérons donc un modèle markovien non-homogène (en temps) à valeurs dans E , l'ensemble des niveaux possibles de dégradation de chaque composant. On note $X = (X_t)_{t \geq t_0}$ cette chaîne de Markov, où t_0 est la date de mise en service du composant (date connue). On suppose qu'initialement, le composant est en parfait état de fonctionnement : $X_{t_0} = 0$. Puisqu'aucune action de maintenance n'est opérée, il s'agit d'une chaîne de naissance pure (il n'y a pas de transition possible entre un état de dégradation donné vers un état moins dégradé) et les probabilités de transition possibles entre les états sont décrites par la figure 1. Une autre description est donnée par la matrice de transition P_t suivante à l'instant t :

$$P_t = \begin{pmatrix} 1 - p_t(0) & p_t(0) & 0 & 0 \\ 0 & 1 - p_t(1) & p_t(1) & 0 \\ 0 & 0 & 1 - p_t(2) & p_t(2) \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

On rappelle que, pour tout $t \geq t_0$ et pour tout $(i, j) \in E^2$, $P_t(i, j) = \mathbb{P}[X_{t+1} = j | X_t = i]$.

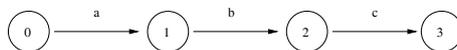


FIGURE 1. Probabilités de transition du modèle de dégradation à 4 états ($q = 3$)

On suppose que s covariables sont observées tous les mois (dans notre cas d'étude, $s = 4$). En fait, on peut considérer que c'est soit l'effet ponctuel d'une covariable qui influence une probabilité de transition, soit son effet intégré dans le temps. On peut donc dissocier stress ponctuel (dans le premier cas) et stress cumulatif (dans le second cas), les deux cas n'étant pas exclusifs.

On a donc $r = 2s$ covariables observées tous les mois (s stress ponctuels et s stress cumulés). Cependant, selon les experts du composant EDF, seuls les effets cumulatifs d'une covariable sont *a priori* pertinents pour le cas d'étude présenté ici. On note $(z(t))_{t \geq t_0}$ le vecteur des covariables disponibles : pour tout $t \geq t_0$, $z(t) = (z_1(t), \dots, z_r(t)) \in \mathbb{R}^r$. De plus, pour tout $j \in \{1, \dots, r\}$, $z_j(t_0) = 0$ et $t \mapsto z_j(t)$ est croissante si cela correspond à un effet cumulatif d'une covariable.

Pour tout $i \in \{0, \dots, q-1\}$, on note $\Omega_i \subset \{1, \dots, r\} = \Omega = \Omega^p \cup \Omega^c$ l'ensemble des covariables influençant la probabilité de transition de l'état i vers l'état $i+1$, avec $\Omega^p = \{1, \dots, s\}$ (resp. $\Omega^c = \{s+1, \dots, r\}$) l'ensemble des covariables correspondant à des stress ponctuels (resp. cumulatifs). On considère que les covariables sont reliées aux probabilités de transition par une fonction de lien h , comme dans une régression logistique :

$$\forall i \in \{0, \dots, q-1\}, \forall t \geq t_0, \quad p_t(i) = \frac{\exp(\alpha_i + \sum_{j \in \Omega_i} \beta_{ij} z_j(t))}{1 + \exp(\alpha_i + \sum_{j \in \Omega_i} \beta_{ij} z_j(t))} := h\left(\alpha_i + \sum_{j \in \Omega_i} \beta_{ij} z_j(t)\right), \quad (1)$$

où $\alpha_i \in \mathbb{R}$ (intercept) et $\underline{\beta}_i = (\beta_{ij})_{j \in \Omega_i} \in \mathbb{R}^{|\Omega_i|}$. Pour tout $i \in \{0, \dots, q-1\}$, on pose $\underline{\theta}_i = (\alpha_i, \underline{\beta}_i) \in \Theta_i = \mathbb{R} \times \mathbb{R}^{|\Omega_i|}$ et $\Theta = \Theta_0 \times \dots \times \Theta_{q-1}$ (espace des paramètres). Si on considère l'ensemble l'ensemble des covariables, alors le modèle complet correspond au cas où $\Omega_0 = \dots = \Omega_{q-1} = \Omega$ et comporte donc $q(r+1)$ paramètres (dans notre cas d'étude, cela donnerait 27 paramètres). Si on considère qu'un seul type de covariables (i.e. uniquement les stress ponctuels ou uniquement les stress cumulatifs), alors le modèle complet correspond au cas où $\Omega_0 = \dots = \Omega_{q-1} = \Omega^p$ ou Ω^c et comporte donc $q(s+1)$ paramètres (dans notre cas d'étude, cela donnerait 15 paramètres). Quant au modèle sans covariable, il fait intervenir q paramètres et on a $P_t = P$ pour tout $t \geq t_0$. L'intercept peut prendre n'importe quelle valeur réelle, mais il doit être *a priori* négatif si les probabilités de transition sont petites. Pour les autres paramètres liés aux covariables, si ces covariables correspondent à des stress, on pourrait s'attendre à ce que plus le stress est important, plus la probabilité de transition est grande : ceci implique que le paramètre devant cette covariable soit positif. Enfin, une covariable sans influence sur une probabilité de transition devrait être associée à un paramètre égal à zéro.

Il sera important pour la suite de considérer les probabilités de transition entre deux instants $t > t_0$ et $t' > t$ non nécessairement consécutifs. On a :

$$\forall (i, j) \in E^2, \quad P_t(i, j) = \left[\prod_{u=t}^{t'-1} P_u \right] (i, j) \stackrel{(hom.)}{=} P^{t'-t}(i, j),$$

où $\stackrel{(hom.)}{=}$ signifie qu'on se place dans le cadre homogène, i.e. celui du modèle sans covariable. On peut vérifier qu'en prenant $t' = t+1$, on retrouve les expressions précédentes. On remarquera que, si $j < i$, ces probabilités sont nulles. De plus, dans le cas particulier où $j = i$, l'expression ci-dessus se simplifie de la manière suivante :

$$\forall i \in E, \quad P_t(i, i) = \prod_{u=t}^{t'-1} P_u(i, i) \stackrel{(hom.)}{=} P(i, i)^{t'-t}$$

(autrement dit, dans ces cas-là, on peut s'abstenir de faire des calculs matriciels au profit de calculs entre scalaires, moins coûteux en temps).

Remarque 1.

1. Afin d'établir une relation entre probabilités de transition et covariables, nous avons considéré comme fonction de lien h l'inverse de la fonction logit :

$$\forall x \in \mathbb{R}, \quad \text{logit}^{-1}(x) = \frac{e^x}{1 + e^x} \in [0, 1].$$

D'autres fonctions de lien auraient pu être également considérées. Par exemple, on peut utiliser l'inverse de la fonction probit, i.e. la fonction de répartition de loi normale centrée-réduite.

2. Quel que soit le choix de la fonction de lien, pour tout $i \in E \setminus \{q\}$ et pour tout $t \geq 0$, $p_t(i) \in]0, 1[$: $p_t(i)$ ne peut donc pas être égale à 0 ou 1.

Cas particulier : temps de défaillance pour le modèle sans covariable Pour tout $i \in E \setminus \{q\}$, soit S_i le temps de séjour dans l'état i : c'est une variable aléatoire de loi géométrique de paramètre $p(i)$. Ainsi, le temps moyen de séjour dans l'état i est égal à $1/p(i)$. Notons, pour tout $i \in E \setminus \{0\}$, T_i le temps d'atteinte de l'état i : $T_i = S_0 + \dots + S_{i-1}$. Clairement, le temps de panne T_q est presque sûrement fini et son espérance (MTTF) est égale à :

$$MTTF = \mathbb{E}[T_q] = \frac{1}{p(0)} + \dots + \frac{1}{p(q-1)}.$$

Un cas simple apparaît quand toutes les probabilités de transition sont égales : $p(0) = \dots = p(q-1) = p$; dans ce cas, la loi marginale de T_i est la loi binomiale négative de paramètres p et i . Lorsque tous les paramètres sont différents, la loi de T_2 et celle de T_3 sont données par le lemme suivant.

Lemme 1. La loi de T_2 est donnée par : $\forall k \in \{2, 3, \dots\}$,

$$\mathbb{P}[T_2 = k] = \frac{p(0)p(1)}{p(1) - p(0)} \left[[1 - p(0)]^{k-1} - [1 - p(1)]^{k-1} \right]$$

et celle de T_3 par : $\forall k \in \{3, 4, \dots\}$,

$$\mathbb{P}[T_3 = k] = \frac{p(0)p(1)p(2)}{p(1) - p(0)} \left[\frac{[1 - p(2)]^{k-1} - [1 - p(0)]^{k-1}}{p(0) - p(2)} - \frac{[1 - p(2)]^{k-1} - [1 - p(1)]^{k-1}}{p(1) - p(2)} \right].$$

Pour $q \geq 4$, on peut également calculer la loi de T_q par récurrence. En effet, on peut facilement montrer que, pour tout $q \in \mathbb{N} \setminus \{0, 1\}$ et pour tout $k \in \{q, q+1, \dots\}$,

$$\mathbb{P}[T_q = k] = p(q-1) \sum_{k'=1}^{k-q+1} (1 - p(q-1))^{k'} \mathbb{P}[T_{q-1} = k - k'].$$

3. Inférence statistique

Après avoir décrit tous les types d'observations possibles pour notre cas d'étude, nous allons calculer la fonction de vraisemblance dans le cadre général. Ce travail peut être adapté à d'autres situations (nombre d'états et/ou politique de remplacement différents), la méthodologie restant la même.

3.1. Types d'observation

Puisque nous considérons une chaîne de naissance pure (que des covariables soient incluses ou pas), il est facile d'établir une liste exhaustive de toutes les différentes observations possibles. Pour notre cas d'étude, il y a 18 cas différents qui sont reportés dans le tableau 2. L'état *cens* signifie qu'à la fin de la période de l'étude, le composant n'était pas en panne (censure à droite). Il y a donc trois "états finaux" : *cens* (non-défaillant à la fin de l'étude), 2 (dégradation majeure constatée et donc remplacement), 3 (panne). Il est possible de dénombrer le nombre de situations différentes

TABLEAU 2. Liste exhaustive des différents cas possibles pour $q = 3$

Cas	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
t_1	<i>cens</i>	2	3	0	1	0	1	0	1	0	0	0	1	1	1	0	0	0
t_2	-	-	-	<i>cens</i>	<i>cens</i>	2	2	3	3	1	1	1	1	1	1	1	1	1
t_3	-	-	-	-	-	-	-	-	-	<i>cens</i>	2	3	<i>cens</i>	2	3	1	1	1
t_4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	<i>cens</i>	2	3
Effectif sur le cas d'étude EDF	28	13	3	21	2	26	5	0	0	2	2	2	2	0	0	0	0	0

dans un cadre plus général. Supposons qu'un composant est remplacé si, lors d'une inspection, il est dans l'un des n_r états les plus dégradés, i.e. dans un des états $\{q - n_r + 1, \dots, q\}$ (pas de remplacement dans les autres cas). Alors, par un simple calcul combinatoire, on montre qu'il y a exactement $2 \times 3^{q-n_r} \times (n_r + 1)$ cas distincts (on retrouve bien la valeur obtenue directement pour le cas d'étude : $q = 3$, $n_r = 2$ et donc il y a $2 \times 3^{3-2} \times (2 + 1) = 18$ cas distincts).

3.2. Fonction de vraisemblance

On note $H(E, \mathcal{R})$ le nombre de cas distincts qui dépend donc à la fois de l'espace d'états E et de la politique de remplacement \mathcal{R} appliquée. Dans notre cas d'étude, on a $H(E, \mathcal{R}) = 18$ (voir plus haut). Pour tout $h \in H(E, \mathcal{R})$, on note m_h le nombre de temps d'inspection ou de défaillance (ici, $m_h \in \{1, \dots, 4\}$) et n_h le nombre de composants dans l'échantillon correspondant à ce cas. La fonction de vraisemblance s'écrit, en toute généralité, sous la forme suivante :

$$L(\theta | \text{données}) = \prod_{h=1}^{H(E, \mathcal{R})} \prod_{i=1}^{n_h} \mathbb{P} \left[X_{t_{0,i}}^{(h,i)} = 0, X_{t_{1,i}}^{(h,i)} \in A_1, \dots, X_{t_{m_h,i}}^{(h,i)} \in A_{m_h} \right],$$

où $X^{(h,i)}$ sont des copies i.i.d. de la chaîne de Markov X décrite précédemment et où A_1, \dots, A_{m_h} sont des sous-ensembles convenablement choisis pour décrire le h -ième cas. Par exemple, pour le cas 1 du tableau 2, on a $m_1 = 1$ et $A_1 = \{0, 1, 2\}$ et pour le cas 2, on a $m_2 = 1$ et $A_1 = \{2\}$, etc. Pour tout $h \in H(E, \mathcal{R})$ et pour tout $i \in \{1, \dots, n_h\}$, on a :

$$\mathbb{P} \left[X_{t_{0,i}}^{(h,i)} = 0, X_{t_{1,i}}^{(h,i)} \in A_1, \dots, X_{t_{m_h,i}}^{(h,i)} \in A_{m_h} \right] = \prod_{j=1}^{m_h} \mathbb{P} \left[X_{t_j}^{(h,i)} \in A_j | X_{t_{j-1}}^{(h,i)} \in A_{j-1} \right]$$

avec $A_0 = \{0\}$, parce qu'on a toujours $X_{t_0} = 0$. Chaque contribution à la fonction de vraisemblance correspond à la probabilité de la trajectoire observée pour chacun des composants.

Nous ne donnons à titre d'exemple que deux cas, le cas 9 et le cas 11 du tableau 2 (tous les autres cas figurant en annexe à la fin de l'article). Pour un composant dans le cas 9, on a $\{X_{t_0} = 0\} \cup \{X_{t_1} = 1\} \cup \{X_{t_2} = 3\}$. On peut alors calculer la probabilité d'observer cet événement :

$$\mathbb{P}[X_{t_0} = 0; X_{t_1} = 1; X_{t_2} = 3] = \left[\prod_{u=t_0}^{t_1-1} P_u \right] (0, 1) \left[\prod_{u=t_1}^{t_2-1} P_u \right] (1, 3).$$

Si on considère le modèle sans covariable, on a la simplification suivante :

$$\mathbb{P}[X_{t_0} = 0; X_{t_1} = 1; X_{t_2} = 3] \stackrel{(hom.)}{=} P^{t_1-t_0}(0, 1)P^{t_2-t_1}(1, 3).$$

De manière alternative, on peut aussi utiliser l'information suivante : si, à l'instant t_2 , le composant est en panne, alors on en déduit qu'à l'instant $t_2 - 1$, le composant est dans l'état 2 (puisqu'une panne est détectée immédiatement). On a donc également :

$$\begin{aligned} \mathbb{P}[X_{t_0} = 0; X_{t_1} = 1; X_{t_2} = 3] &= \left[\prod_{u=t_0}^{t_1-1} P_u \right] (0, 1) \left[\prod_{u=t_1}^{t_2-1-1} P_u \right] (1, 2)P_{t_2-1}(2, 3) \\ &\stackrel{(hom.)}{=} P^{t_1-t_0}(0, 1)P^{t_2-t_1-1}(1, 2)P(2, 3). \end{aligned}$$

Considérons maintenant un composant dans le cas 11. On a $\{X_{t_0} = 0\} \cup \{X_{t_1} = 0\} \cup \{X_{t_2} = 1\} \cup \{X_{t_3} = 2\}$. La probabilité d'avoir une telle observation est égale à :

$$\begin{aligned} \mathbb{P}[X_{t_0} = 0; X_{t_1} = 0; X_{t_2} = 1; X_{t_3} = 2] &= \prod_{u=t_0}^{t_1-1} P_u(0, 0) \left[\prod_{u=t_1}^{t_2-1} P_u \right] (0, 1) \left[\prod_{u=t_2}^{t_3-1} P_u \right] (1, 2) \\ &\stackrel{(hom.)}{=} P(0, 0)^{t_1-t_0} P^{t_2-t_1}(0, 1)P^{t_3-t_2}(1, 2). \end{aligned}$$

3.3. Estimateur du maximum de vraisemblance

On obtient alors l'estimateur du maximum de vraisemblance en optimisant la fonction L

$$\hat{\theta}_n = \operatorname{argmax}_{\theta \in \Theta} L(\theta | \text{données}),$$

avec $n = \sum_{h \in H(E, \mathcal{R})} n_h$. La théorie du maximum de vraisemblance paramétrique permet de supposer raisonnablement l'approximation de la loi de $\hat{\theta}_n$ convenablement renormalisée par la loi normale centrée et de matrice de variance-covariance $I^{-1}(\theta)$:

$$\sqrt{n} (\hat{\theta}_n - \theta) \stackrel{approx.}{\sim} \mathcal{N}(0, I^{-1}(\theta))$$

la matrice d'information de Fisher $I(\theta)$ pouvant être estimée par la hessienne du logarithme de la vraisemblance évaluée en $\hat{\theta}_n$. Ainsi, si on dispose d'un échantillon de grande taille, on peut construire des intervalles de confiance asymptotiques pour chacun des paramètres (ou des ellipsoïdes de confiance pour plusieurs paramètres). En présence d'échantillon de taille petite ou moyenne, on pourra préférer l'obtention d'intervalles de confiance par bootstrap.

3.4. Comparaison de modèles

Il est bien connu que, plus on considère de covariables et donc de paramètres, plus la fonction de vraisemblance à l'optimum est grande (voir Saporta (2011), par exemple). Ceci conduirait mécaniquement à considérer le modèle complet, si on ne se basait que sur la fonction de vraisemblance comme critère de sélection de modèle. Il est donc nécessaire d'introduire un critère pénalisé. Le critère d'information d'Akaike (AIC) peut être utilisé pour comparer deux modèles (emboîtés ou pas). On rappelle que ce critère pénalise la vraisemblance par un terme qui décroît avec le nombre de paramètres (donc ici avec le nombre de covariables). Si, pour un modèle donné, l'ensemble des covariables est :

$$\Theta = \prod_{i=0}^{q-1} \left[\mathbb{R} \times \mathbb{R}_+^{|\Omega_i|} \right], \quad \text{alors} \quad \text{AIC}(\Theta) = -2 \log L(\hat{\theta}_n | \text{données}) + 2 \sum_{i=0}^{q-1} (1 + |\Omega_i|). \quad (2)$$

Entre deux modèles, on retient celui qui donne le plus faible critère AIC. En pratique, quand on compare deux modèles, la différence entre les deux valeurs de l'AIC peut être parfois très petite. Or, le critère dépendant de l'estimateur, c'est donc aussi une variable aléatoire et il est donc légitime de se poser la question de la significativité de la différence entre les deux valeurs observées de l'AIC. Une analyse comparative plus fine des critères de sélection de modèles a été proposée par Commenges et al. (2007, 2008) et pourrait être mise à profit pour des développements complémentaires ultérieurs.

3.5. Estimation de la loi du temps de panne

Une fois un modèle ajusté, il est intéressant d'estimer la loi du temps de panne. Dans le cas du modèle sans covariable, nous avons donné précédemment l'expression de la loi de T_q (section 2). Pour le modèle avec covariables, la situation est plus compliquée car les covariables dépendent du temps. Pour cela, une solution consiste à regarder un comportement moyen des covariables (qui interviennent dans le modèle ajusté) par un lissage sur les trajectoires des covariables observées, par exemple à l'aide d'une régression linéaire (simple ou polynomiale) ou d'une interpolation par splines (avec la contrainte de croissance de la trajectoire moyennisée pour des covariables associées à des stress cumulatifs). On peut alors effectuer des simulations de Monte Carlo pour obtenir un échantillon de temps de panne et en déduire une estimation non-paramétrique de la fiabilité, par exemple.

4. Applications

Afin de valider la méthodologie proposée, on a simulé des données "proches" de celles recueillies par EDF, uniquement pour le modèle sans covariable. Puis, on a appliqué cette méthodologie au jeu de données réelles EDF qui a motivé cette étude. N'ayant pas d'expression explicite pour l'estimateur du maximum de vraisemblance, il est nécessaire d'utiliser une méthode numérique d'optimisation ; dans les applications, on a utilisé l'algorithme BFGS avec contraintes de borne (Gilbert., 2008).

4.1. *Données simulées*

On a simulé des données en prenant $p(0) = 0,0081$, $p(1) = 0,0275$ et $p(2) = 0,0020$, sur la base des mêmes temps d'observation que dans le jeu de données EDF (et donc un échantillon de même taille, i.e. un échantillon constitué de 106 composants). Le tableau 3 contient les estimations des paramètres ainsi que l'intervalle de confiance ponctuel à 95% pour chacun des paramètres (en supposant la normalité asymptotique de l'estimateur du maximum de vraisemblance). On obtient sensiblement les mêmes résultats avec les deux fonctions de lien, logit et probit. On observe

TABLEAU 3. *Données simulées sans covariable : estimation des paramètres*

Paramètre	Valeur	Estimation	Intervalle de confiance asymptotique
$p(0)$	0,0081	0,0070	[0,0052;0,0095]
$p(1)$	0,0275	0,0271	[0,0174;0,0420]
$p(2)$	0,0020	0,0020	[0,0007;0,0054]

que les estimations ponctuelles sont globalement proches des vraies valeurs. Les vraies valeurs sont toutes dans l'intervalle de confiance asymptotique et, même si on peut les considérer plutôt proches de la borne inférieure, on considère que la méthodologie développée dans cet article est tout à fait applicable aux données EDF.

L'utilisation d'intervalles de confiance basés sur un résultat asymptotique est souvent délicate. En effet, la validité des conditions d'application est en pratique discutable : l'échantillon est-il de taille suffisante ? les durées ne sont-elles pas trop censurées ? Afin d'appréhender ces points, on a réalisé une seconde étude basée sur des données simulées. On a simulé 500 échantillons de 106 composants puis de 212 composants, avec dans les deux cas au moins 5 durées exactes observées (i.e. au moins 5 pannes). Pour chaque échantillon, on a calculé les estimateurs des paramètres et on a testé la normalité ces estimateurs, après transformation logit, à l'aide du test de Shapiro-Wilk et du test de Lilliefors (Capéraà and van Cutsem, 1988). On donne les p -valeurs de ces tests dans le tableau 4 : on constate que la normalité des estimateurs peut être acceptée, même pour un échantillon de taille modérée (c'est moins le cas pour $\hat{\alpha}_2$, mais il s'agit nécessairement du paramètres qui est le moins "bien" estimé). Sur la figure 2, on a représenté les histogrammes et les densités des lois normales (avec moyenne et variance empiriques) pour chaque estimateur et chaque situation.

TABLEAU 4. *Données simulées : normalité des estimateurs ?*

Test	p -valeur pour $\hat{\alpha}_0$	p -valeur pour $\hat{\alpha}_1$	p -valeur pour $\hat{\alpha}_2$
Échantillons de taille 106			
Shapiro-Wilk	0,095	0,976	<1e-7
Lilliefors	0,113	0,895	0,012
Échantillons de taille 212			
Shapiro-Wilk	0,366	0,520	0,008
Lilliefors	0,899	0,382	0,101

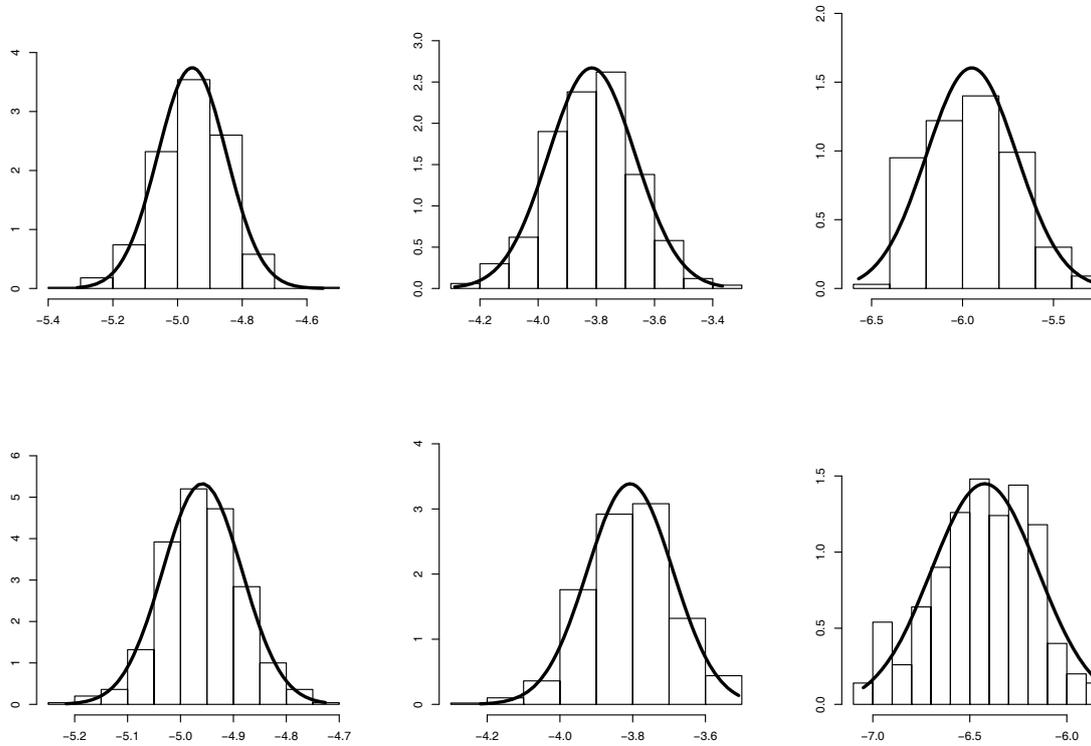


FIGURE 2. Données simulées : normalité des estimateurs ? En ligne : échantillon de taille 106 ou 212. En colonne : lois empiriques de $\hat{\alpha}_0$, $\hat{\alpha}_1$ et $\hat{\alpha}_2$.

4.2. Données réelles

On commence par ajuster le modèle sans covariable. Ce modèle servira de référence par rapport à un modèle avec covariables. De plus, les estimations obtenues pour le modèle sans covariable pourront servir de point de départ pour la recherche numérique du maximum de vraisemblance dans un modèle avec covariables.

4.2.1. Ajustement au modèle sans covariable

Le tableau 5 donne l'estimation ponctuelle ainsi que l'intervalle de confiance asymptotique à 95% pour les trois paramètres du modèle sans covariable (que l'on considérera avec précaution, vu les résultats obtenus ci-dessus par simulation). On peut en déduire une estimation du temps moyen avant perte de la fonction d'étanchéité du composant. On obtient ici une durée moyenne égale à 649 mois (soit environ 54 ans). Le temps pour atteindre un niveau de dégradation majeure est estimé à environ 159 mois, soit environ 13 ans. En utilisant le lemme 1, on peut estimer la loi du

TABLEAU 5. Données réelles : estimation des paramètres du modèle sans covariable

Paramètre	Estimation	Intervalle de confiance asymptotique
$p(0)$	0,0081	[0,0062;0,0105]
$p(1)$	0,0275	[0,0195;0,0384]
$p(2)$	0,0020	[0,0008;0,0050]

temps de panne ainsi que sa fonction de survie, voir la figure 3. Enfin, on peut en déduire une estimation du temps médian : on obtient un temps médian estimé à 512 mois, soit environ 42 ans.

4.2.2. Ajustement au modèle avec covariable(s)

Il n'est pas évident que toutes les covariables aient une influence sur toutes les probabilités de transition. On doit chercher le meilleur sous-modèle selon un critère donné, par exemple le critère d'Akaike (AIC). D'après les experts du matériel EDF étudié, seul le stress cumulatif des covariables est à considérer. On a donc au total 15 paramètres possibles dans le modèle complet. Cela donne $2^{15} = 32768$ sous-modèles possibles (si on considère à la fois les effets ponctuels et cumulés des covariables, on a deux fois plus de paramètres et donc il y a $2^{27} = 134217728$ sous-modèles à tester). Une recherche exhaustive est donc impossible à mener. Une recherche à l'aide d'une méthode ascendante, descendante ou pas à pas n'est pas non plus envisageable. En effet, les temps de calculs peuvent être très longs : de quelques minutes pour le modèle sans covariable à plus d'une douzaine d'heures dans certains cas, les calculs étant lancés sur le cluster (machine virtuelle 4 cœurs, 24 Go de RAM).

Soit $\Omega_0^{(v)}$, $\Omega_1^{(v)}$ et $\Omega_2^{(v)}$ l'ensemble des covariables pour chacune des probabilités de transition. Chercher le meilleur sous-modèle revient à déterminer ces trois sous-ensembles de $\{1, 2, 3, 4\}$. On propose donc l'approche suivante.

- Étape 0 : on ajuste le modèle homogène : $\Omega_0^{(v)} = \Omega_1^{(v)} = \Omega_2^{(v)} = \emptyset$ et on calcule le critère d'Akaike $AIC^{(0)}$.
- Étape v : on cherche à ajouter, à toutes les probabilités de transition, une covariable non encore incluse et on compare les critères AIC. Considérons $j \notin \Omega_0^{(v-1)} \cup \Omega_1^{(v-1)} \cup \Omega_2^{(v-1)}$. On pose $\Omega_i^{(v,j)} = \Omega_i^{(v-1)} \cup \{j\}$, pour tout $i \in \{0, 1, 2\}$. On estime les paramètres du modèle avec pour covariables $\Omega_0^{(v,j)}$, $\Omega_1^{(v,j)}$ et $\Omega_2^{(v,j)}$, ainsi que le critère d'Akaike $AIC^{(v,j)}$ (la pénalité dans l'AIC ne concerne que les paramètres non nuls, i.e. on ne prend pas en compte les paramètres qui atteindraient 0). Deux cas peuvent alors se produire :
 1. soit, pour tout $j \notin \Omega_0^{(v-1)} \cup \Omega_1^{(v-1)} \cup \Omega_2^{(v-1)}$, $AIC^{(v,j)} > AIC^{(v-1)}$: dans ce cas, on retient le sous-modèle obtenu à l'étape $v-1$;
 2. soit il existe $j^* \notin \Omega_0^{(v-1)} \cup \Omega_1^{(v-1)} \cup \Omega_2^{(v-1)}$ tel que $AIC^{(v,j^*)} \leq AIC^{(v-1)}$ et $AIC^{(v,j^*)} < AIC^{(v,j)}$ pour $j \neq j^*$: dans ce cas, le sous-modèle obtenu à l'issue de l'étape v est

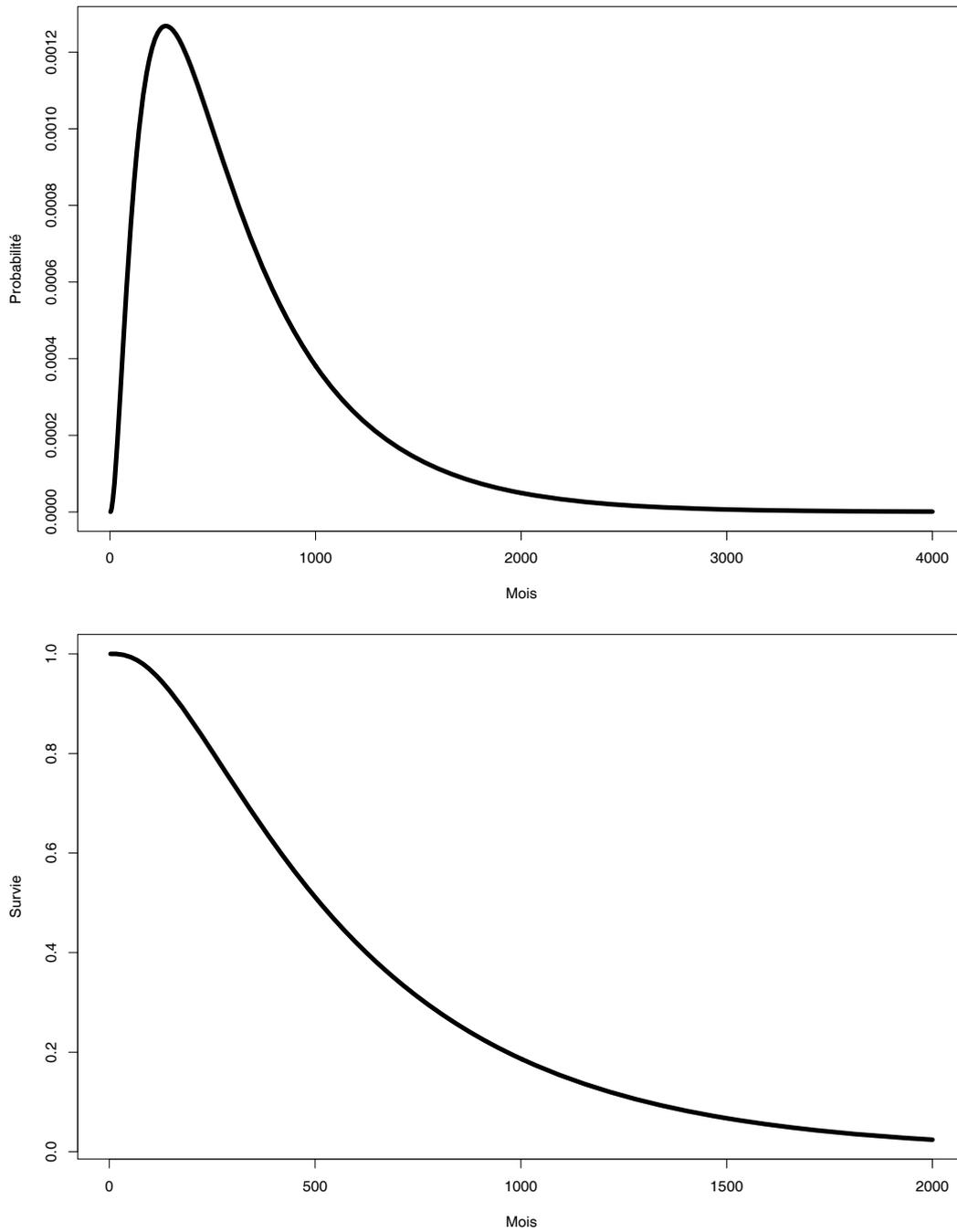


FIGURE 3. *Données réelles : estimation de la loi et de la fonction de survie du temps de panne (modèle sans covariable)*

déterminé, pour tout $i \in \{0, 1, 2\}$, par :

$$\Omega_i^{(v)} = \begin{cases} \Omega_i^{(v-1)} \cup \{j^*\} & \text{si } \widehat{\beta}_{ij^*} \neq 0 \\ \Omega_i^{(v-1)} & \text{si } \widehat{\beta}_{ij^*} = 0, \end{cases}$$

où $\widehat{\beta}_{ij^*}$ est l'estimation du coefficient de β_{ij^*} .

Évidemment, cette exploration de l'ensemble des sous-modèles ne donne aucune garantie de trouver le meilleur sous-modèle. Toutefois, il est raisonnable de supposer que le modèle obtenu n'en est pas trop loin.

Le tableau 6 indique les différentes étapes obtenues en appliquant la méthodologie décrite ci-dessus. Le meilleur sous-modèle avec une covariable est celui obtenu avec la covariable 2. Mais celle-ci n'intervient que dans la probabilité de transition de l'état 0 vers l'état 1 car $\widehat{\beta}_{12}^{(1)} = \widehat{\beta}_{22}^{(1)} = 0$ tandis que $\widehat{\beta}_{02}^{(1)} \neq 0$. Le critère AIC étant inférieur à celui du modèle sans covariable, on retient donc ce sous-modèle à l'issue de la première étape. On cherche alors à intégrer une des autres covariables restantes. Parmi les sous-modèles restants, celui qui donne le critère AIC le plus petit est celui où les covariables 2 et 3 interviennent dans $p_t(0)$ tandis que la covariable 3 intervient dans $p_t(1)$ et $p_t(2)$. Cependant, tous les nouveaux paramètres de ce sous-modèle saturant la contrainte de positivité : $\widehat{\beta}_{03}^{(2)} = \widehat{\beta}_{13}^{(2)} = \widehat{\beta}_{23}^{(2)} = 0$. Donc, on n'accepte pas ce nouveau sous-modèle, l'algorithme de sélection s'arrête et on retient le sous-modèle obtenu à la première étape.

TABLEAU 6. Données réelles : étapes de la recherche du meilleur sous-modèle

Étape	Modèle	Log-vraisemblance	Nb param. non nuls	AIC
0	sans aucune covariable	-168,6184	3	343,2368
1	covariable 1 pour $p_t(0)$, $p_t(1)$ et $p_t(2)$	-167,4615	4	343,2368
	covariable 2 pour $p_t(0)$, $p_t(1)$ et $p_t(2)$	-167,0329	4	342,0658
	covariable 3 pour $p_t(0)$, $p_t(1)$ et $p_t(2)$	-167,4214	4	342,8428
	covariable 4 pour $p_t(0)$, $p_t(1)$ et $p_t(2)$	-167,6825	5	345,3650
2	covariables 1 et 2 pour $p_t(0)$ covariable 2 pour $p_t(1)$ et $p_t(2)$	-167,0117	5	344,0234
	covariables 2 et 3 pour $p_t(0)$ covariable 3 pour $p_t(1)$ et $p_t(2)$	-167,0329	4	342,0658
	covariables 2 et 4 pour $p_t(0)$ covariable 4 pour $p_t(1)$ et $p_t(2)$	-166,9935	5	343,9870

Dans le tableau 7, on donne l'estimation des paramètres du sous-modèle retenu à l'issue de cette procédure de sélection de modèles. On notera que 0 est dans l'intervalle de confiance asymptotique pour β_{02} , ce qui pourrait conduire à accepter le test de Student pour la nullité de ce coefficient. Cependant, comme expliqué précédemment, il n'est pas évident de pouvoir raisonnablement utiliser un résultat asymptotique avec 106 composants de manière systématique. Donc, si on applique un test de significativité, il faudra considérer avec précaution le résultat

obtenu. Il est alors intéressant de calculer, de manière complémentaire, des intervalles de confiance par bootstrap (mais cela peut s'avérer très coûteux en temps). La technique du Jackknife (Tassi, 2004) semble être une alternative à la fois simple à mettre en œuvre et pertinente. En effet, cette méthode permet de garantir que les ré-échantillons auront un "taux de censure" comparable à celui du jeu de données d'origine ; en particulier, il serait "désastreux" de considérer un échantillon bootstrapé ne contenant aucune durée exacte et la technique du Jackknife assure qu'il y aura pas plus d'une durée exacte en moins dans chaque échantillon bootstrapé. On constate que les intervalles de confiance obtenus par bootstrapé sont beaucoup resserrés que ceux asymptotiques : cette fois-ci, 0 n'est pas dans l'intervalle de confiance obtenue par bootstrap pour β_{02} . Il est donc plus prudent de considérer que la covariable 2 est influente sur l'apparition d'une dégradation d'un composant, puis, une fois dans un état de dégradation mineure, aucune covariable ne contribue à une détérioration accrue des composants. Ce point devra être validé par les experts matériels EDF.

TABLEAU 7. *Données réelles : estimation des paramètres du meilleur sous-modèle avec covariable*

Paramètre	Estimation	Intervalle de confiance asymptotique	Intervalle de confiance par bootstrap
α_0	-5,1711	[-5,6689; -4,6732]	[-5,2214; -5,1336]
β_{02}	1,3662	[-0,1007; 2,8331]	[1,2608; 1,4879]
α_1	-3,5428	[-3,8930; -3,1927]	[-3,5630; -3,5982]
α_2	-6,1788	[-7,0770; -5,2807]	[-6,3932; -6,1458]

5. Conclusion

Dans cet article, nous avons proposé un modèle markovien à temps discret pour la dégradation d'un composant soumis à une politique de remplacement donnée. Les probabilités de transition peuvent être fonction de covariables dépendantes du temps, rendant ainsi le modèle non homogène. L'estimateur du maximum de vraisemblance est considéré. Bien que générale, la méthodologie proposée n'est pas opérationnelle immédiatement pour n'importe quelle situation (i.e. pour n'importe quel nombre d'états et pour n'importe quelle politique de remplacement). En effet, pour le moment, des développements informatiques propres sont à mener pour chacune des situations, à moins d'écrire des procédures automatiques permettant de gérer tous les cas possibles. La méthodologie proposée est évaluée à travers des données simulées. Enfin, une application à des données réelles d'EDF est également étudiée. Cette méthodologie répond à la problématique industrielle et est bien adaptée aux données. Des améliorations pourront être apportées, tant sur des aspects théoriques (par exemple, sur une utilisation plus fine du critère AIC –Commenges et al. (2007, 2008)– pour comparer des modèles ou bien sur les tests de significativité des coefficients – les intervalles de confiance obtenus sont très différents selon la méthode utilisée pour le jeu de données étudié) que pratiques (mise en œuvre automatique dans un cadre plus général que celui du cas d'étude).

Références

Andersen, P. K. and Keiding, N. (2002). Multi-state models for event history analysis. *Statistical methods in medical research*, 11 :91–115.

- Capéraà, P. and van Cutsem, B. (1988). *Méthodes et modèles en statistique non paramétrique : exposé fondamental*. Dunod.
- Commenges, D. (1999). Multi-state models in epidemiology. *Lifetime Data Analysis*, 5 :315–327.
- Commenges, D., Joly, P., Gégout-Petit, A., and Liqueur, B. (2007). Choice between semi-parametric estimators of Markov and non-Markov multi-state models from generally coarsened observations. *Scandinavian Journal of Statistics*, 34 :33–52.
- Commenges, D., Sayyareh, A., Letenneur, L., Guedj, J., and Bar-Hen, A. (2008). Estimating a difference of Kullback-Leibler risks using a normalized difference of AIC. *Annals of Applied Statistics*, 2(3) :1123–1142.
- Gilbert, J.-C. (2008). Optimisation différentielle. *Techniques de l'Ingénieur*, AF 1252.
- Guo, F., Rakha, H., and Park, S. (2010). Multistate model for travel time reliability. *Transportation Research Record : Journal of the Transportation Research Board*, 2128 :46–54.
- Hougaard, P. (1999). Multi-state models: A review. *Lifetime Data Analysis*, 5 :239–264.
- Kallen, M. (2009). A comparison of statistical models for visual inspection data. In *H. Furuta, D.M. Frangopol and M. Shinozuka (Eds). Safety, Reliability and Risk of Structures, Infrastructures and Engineering Systems, Proceedings of the Tenth International Conference on Structural Safety and Reliability (ICOSSAR'2009)*, pages 3235–3242. Taylor & Francis, London.
- Kallen, M. and van Noordwijk, J. (2006). Statistical inference for Markov deterioration models of bridge conditions in the Netherlands. In *D.M. Frangopol P.J.S. Cruz and L.C. Neves (Eds)*, Proceedings of the Third International Conference on Bridge Maintenance, Safety and Management (IABMAS), pages 16–19. Taylor & Francis Group.
- Natvig, B. and Mørch, H. (2003). An application of multistate reliability theory to an offshore gas pipeline network. *International Journal of Reliability, Quality and Safety Engineering*, 10 :361–381.
- Saporta, G. (2011). *Probabilités, analyse des données et statistique*. Technip.
- Tassi, P. (2004). *Méthodes statistiques*. Technip.

Annexe

En complément des cas 9 et 11 déjà présentés en section 3.2, nous donnons ci-dessous les expressions explicites de ces probabilités correspondant aux dix-huit différents cas.

Cas 1 $X_{t_0} = 0$ et $X_{t_1} \in A_1 = \{0, 1, 2\}$. On a :

$$\mathbb{P}[X_{t_0} = 0; X_{t_1} \in \{0, 1, 2\}] = \sum_{i=0}^2 \mathbb{P}[X_{t_1} = i | X_{t_0} = 0] = \sum_{i=0}^2 \left[\prod_{u=t_0}^{t_1-1} P_u \right] (0, i) \stackrel{(hom.)}{=} \sum_{i=0}^2 P^{t_1-t_0}(0, i).$$

Cas 2 $X_{t_0} = 0$ et $X_{t_1} = 2$. On a :

$$\mathbb{P}[X_{t_0} = 0; X_{t_1} = 2] = \mathbb{P}[X_{t_1} = 2 | X_{t_0} = 0] = \left[\prod_{u=t_0}^{t_1-1} P_u \right] (0, 2) \stackrel{(hom.)}{=} P^{t_1-t_0}(0, 2).$$

Cas 3 $X_{t_0} = 0$ et $X_{t_1} = 3$. On a :

$$\mathbb{P}[X_{t_0} = 0; X_{t_1} = 3] = \mathbb{P}[X_{t_1} = 3 | X_{t_0} = 0] = \left[\prod_{u=t_0}^{t_1-1} P_u \right] (0, 3) \stackrel{(hom.)}{=} P^{t_1-t_0}(0, 3).$$

De manière alternative, on peut aussi utiliser l'information suivante : si, à l'instant t_1 , le composant est en panne, alors on en déduit qu'à l'instant $t_1 - 1$, le composant est dans l'état 2 (puisque une panne est détectée immédiatement). On a donc également :

$$\mathbb{P}[X_{t_0} = 0; X_{t_1} = 3] = \left[\prod_{u=t_0}^{t_1-2} P_u \right] (0, 2) P_{t_1-1}(2, 3) \stackrel{(hom.)}{=} P^{t_1-t_0-1}(0, 2) P(2, 3).$$

Cas 4 $X_{t_0} = 0, X_{t_1} = 0$ et $X_{t_2} \in A_1 = \{0, 1, 2\}$. On a :

$$\begin{aligned} \mathbb{P}[X_{t_0} = 0; X_{t_1} = 0; X_{t_2} \in \{0, 1, 2\}] &= \prod_{u=t_0}^{t_1-1} P_u(0, 0) \left(\sum_{i=0}^2 \left[\prod_{u=t_1}^{t_2-1} P_u \right] (0, i) \right) \\ &\stackrel{(hom.)}{=} P(0, 0)^{t_1-t_0} \left(\sum_{i=0}^2 P^{t_2-t_1}(0, i) \right). \end{aligned}$$

Cas 5 $X_{t_0} = 0, X_{t_1} = 1$ et $X_{t_2} \in A_1 = \{1, 2\}$. On a :

$$\begin{aligned} \mathbb{P}[X_{t_0} = 0; X_{t_1} = 1; X_{t_2} \in \{1, 2\}] &= \left[\prod_{u=t_0}^{t_1-1} P_u \right] (0, 1) \left(\sum_{i=0}^2 \left[\prod_{u=t_1}^{t_2-1} P_u \right] (1, i) \right) \\ &\stackrel{(hom.)}{=} P^{t_1-t_0}(0, 1) \left(\sum_{i=1}^2 P^{t_2-t_1}(1, i) \right). \end{aligned}$$

Cas 6 $X_{t_0} = 0, X_{t_1} = 0$ et $X_{t_2} = 2$. On a :

$$\mathbb{P}[X_{t_0} = 0; X_{t_1} = 0; X_{t_2} = 2] = \prod_{u=t_0}^{t_1-1} P_u(0, 0) \left[\prod_{u=t_1}^{t_2-1} P_u \right] (0, 2) \stackrel{(hom.)}{=} P(0, 0)^{t_1-t_0} P^{t_2-t_1}(0, 2).$$

Cas 7 $X_{t_0} = 0, X_{t_1} = 1$ et $X_{t_2} = 2$. On a :

$$\mathbb{P}[X_{t_0} = 0; X_{t_1} = 1; X_{t_2} = 2] = \left[\prod_{u=t_0}^{t_1-1} P_u \right] (0, 1) \left[\prod_{u=t_1}^{t_2-1} P_u \right] (1, 2) \stackrel{(hom.)}{=} P^{t_1-t_0}(0, 1) P^{t_2-t_1}(1, 2).$$

Cas 8 $X_{t_0} = 0, X_{t_1} = 0$ et $X_{t_2} = 3$. On a :

$$\mathbb{P}[X_{t_0} = 0; X_{t_1} = 0; X_{t_2} = 3] = \prod_{u=t_0}^{t_1-1} P_u(0, 0) \left[\prod_{u=t_1}^{t_2-1} P_u \right] (0, 3) \stackrel{(hom.)}{=} P(0, 0)^{t_1-t_0} P^{t_2-t_1}(0, 3).$$

En reprenant la discussion du cas 3, on a aussi :

$$\begin{aligned} \mathbb{P}[X_{t_0} = 0; X_{t_1} = 0; X_{t_2} = 3] &= \prod_{u=t_0}^{t_1-1} P_u(0, 0) \left[\prod_{u=t_1}^{t_2-2} P_u \right] (0, 2) P_{t_2-1}(2, 3) \\ &\stackrel{(hom.)}{=} P(0, 0)^{t_1-t_0} P^{t_2-t_1-1}(0, 2) P(2, 3). \end{aligned}$$

Cas 9 $X_{t_0} = 0, X_{t_1} = 1$ et $X_{t_2} = 3$. On a :

$$\mathbb{P}[X_{t_0} = 0; X_{t_1} = 1; X_{t_2} = 3] = \left[\prod_{u=t_0}^{t_1-1} P_u \right] (0, 1) \left[\prod_{u=t_1}^{t_2-1} P_u \right] (1, 3) \stackrel{(hom.)}{=} P^{t_1-t_0}(0, 1) P^{t_2-t_1}(1, 3).$$

En reprenant la discussion du cas 3, on a aussi :

$$\begin{aligned} \mathbb{P}[X_{t_0} = 0; X_{t_1} = 1; X_{t_2} = 3] &= \left[\prod_{u=t_0}^{t_1-1} P_u \right] (0, 1) \left[\prod_{u=t_1}^{t_2-1-1} P_u \right] (1, 2) P_{t_2-1}(2, 3) \\ &\stackrel{(hom.)}{=} P^{t_1-t_0}(0, 1) P^{t_2-t_1-1}(1, 2) P(2, 3). \end{aligned}$$

Cas 10 $X_{t_0} = 0, X_{t_1} = 0, X_{t_2} = 1$ et $X_{t_3} = \{1, 2\}$. On a :

$$\begin{aligned} \mathbb{P}[X_{t_0} = 0; X_{t_1} = 0; X_{t_2} = 1; X_{t_3} \in \{1, 2\}] &= \prod_{u=t_0}^{t_1-1} P_u(0, 0) \left[\prod_{u=t_1}^{t_2-1} P_u \right] (0, 1) \left(\sum_{i=1}^2 \left[\prod_{u=t_2}^{t_3-1} P_u \right] (1, i) \right) \\ &\stackrel{(hom.)}{=} P(0, 0)^{t_1-t_0} P^{t_2-t_1}(0, 1) \left(\sum_{i=1}^2 P^{t_3-t_2}(1, i) \right). \end{aligned}$$

Cas 11 $X_{t_0} = 0, X_{t_1} = 0, X_{t_2} = 1$ et $X_{t_3} = 2$. On a :

$$\begin{aligned} \mathbb{P}[X_{t_0} = 0; X_{t_1} = 0; X_{t_2} = 1; X_{t_3} = 2] &= \prod_{u=t_0}^{t_1-1} P_u(0, 0) \left[\prod_{u=t_1}^{t_2-1} P_u \right] (0, 1) \left[\prod_{u=t_2}^{t_3-1} P_u \right] (1, 2) \\ &\stackrel{(hom.)}{=} P(0, 0)^{t_1-t_0} P^{t_2-t_1}(0, 1) P^{t_3-t_2}(1, 2). \end{aligned}$$

Cas 12 $X_{t_0} = 0, X_{t_1} = 0, X_{t_2} = 1$ et $X_{t_3} = 3$. On a :

$$\begin{aligned} \mathbb{P}[X_{t_0} = 0; X_{t_1} = 0; X_{t_2} = 1; X_{t_3} = 3] &= \prod_{u=t_0}^{t_1-1} P_u(0, 0) \left[\prod_{u=t_1}^{t_2-1} P_u \right] (0, 1) \left[\prod_{u=t_2}^{t_3-1} P_u \right] (1, 3) \\ &\stackrel{(hom.)}{=} P(0, 0)^{t_1-t_0} P^{t_2-t_1}(0, 1) P^{t_3-t_2}(1, 3). \end{aligned}$$

En reprenant la discussion du cas 3, on a aussi :

$$\begin{aligned} \mathbb{P}[X_{t_0} = 0; X_{t_1} = 0; X_{t_2} = 1; X_{t_3} = 3] &= \prod_{u=t_0}^{t_1-1} P_u(0, 0) \left[\prod_{u=t_1}^{t_2-1} P_u \right] (0, 1) \left[\prod_{u=t_2}^{t_3-2} P_u \right] (1, 2) P_{t_3-1}(2, 3) \\ &\stackrel{(hom.)}{=} P(0, 0)^{t_1-t_0} P^{t_2-t_1}(0, 1) P^{t_3-t_2-1}(1, 2) P(2, 3). \end{aligned}$$

Cas 13 $X_{t_0} = 0, X_{t_1} = 1, X_{t_2} = 1$ et $X_{t_3} = \{1, 2\}$. On a :

$$\begin{aligned} \mathbb{P}[X_{t_0} = 0; X_{t_1} = 1; X_{t_2} = 1; X_{t_3} \in \{1, 2\}] &= \left[\prod_{u=t_0}^{t_1-1} P_u \right] (0, 1) \prod_{u=t_1}^{t_2-1} P_u(1, 1) \left(\sum_{i=1}^2 \left[\prod_{u=t_2}^{t_3-1} P_u \right] (1, i) \right) \\ &\stackrel{(hom.)}{=} P^{t_1-t_0}(0, 1) P(1, 1)^{t_2-t_1} \left(\sum_{i=1}^2 P^{t_3-t_2}(1, i) \right). \end{aligned}$$

Cas 14 $X_{t_0} = 0, X_{t_1} = 1, X_{t_2} = 1$ et $X_{t_3} = 2$. On a :

$$\begin{aligned} \mathbb{P}[X_{t_0} = 0; X_{t_1} = 1; X_{t_2} = 1; X_{t_3} = 2] &= \left[\prod_{u=t_0}^{t_1-1} P_u \right] (0, 1) \prod_{u=t_1}^{t_2-1} P_u(1, 1) \left[\prod_{u=t_2}^{t_3-1} P_u \right] (1, 2) \\ &\stackrel{(hom.)}{=} P^{t_1-t_0}(0, 1) P(1, 1)^{t_2-t_1} P^{t_3-t_2}(1, 2). \end{aligned}$$

Cas 15 $X_{t_0} = 0, X_{t_1} = 1, X_{t_2} = 1$ et $X_{t_3} = 3$. On a :

$$\begin{aligned} \mathbb{P}[X_{t_0} = 0; X_{t_1} = 1; X_{t_2} = 1; X_{t_3} = 3] &= \left[\prod_{u=t_0}^{t_1-1} P_u \right] (0, 1) \prod_{u=t_1}^{t_2-1} P_u(1, 1) \left[\prod_{u=t_2}^{t_3-1} P_u \right] (1, 3) \\ &\stackrel{(hom.)}{=} P^{t_1-t_0}(0, 1) P(1, 1)^{t_2-t_1} P^{t_3-t_2}(1, 3). \end{aligned}$$

En reprenant la discussion du cas 3, on a aussi :

$$\begin{aligned} \mathbb{P}[X_{t_0} = 0; X_{t_1} = 1; X_{t_2} = 1; X_{t_3} = 3] &= \left[\prod_{u=t_0}^{t_1-1} P_u \right] (0, 1) \prod_{u=t_1}^{t_2-1} P_u(1, 1) \left[\prod_{u=t_2}^{t_3-2} P_u \right] (1, 2) P_{t_3-1}(2, 3) \\ &\stackrel{(hom.)}{=} P^{t_1-t_0}(0, 1) P(1, 1)^{t_2-t_1} P^{t_3-t_2}(1, 2) P(2, 3). \end{aligned}$$

Cas 16 $X_{t_0} = 0, X_{t_1} = 0, X_{t_2} = 1, X_{t_3} = 1$ et $X_{t_4} = \{1, 2\}$. On a :

$$\begin{aligned} \mathbb{P}[X_{t_0} = 0; X_{t_1} = 0; X_{t_2} = 1; X_{t_3} = 1; X_{t_4} \in \{1, 2\}] &= \prod_{u=t_0}^{t_1-1} P_u(0, 0) \left[\prod_{u=t_1}^{t_2-1} P_u \right] (0, 1) \prod_{u=t_2}^{t_3-1} P_u(1, 1) \left(\sum_{i=1}^2 \left[\prod_{u=t_3}^{t_4-1} P_u \right] (1, i) \right) \\ &\stackrel{(hom.)}{=} P(0, 0)^{t_1-t_0} P^{t_2-t_1}(0, 1) P(1, 1)^{t_3-t_2} \left(\sum_{i=1}^2 P^{t_4-t_3}(1, i) \right). \end{aligned}$$

Cas 17 $X_{t_0} = 0, X_{t_1} = 0, X_{t_2} = 1, X_{t_3} = 1$ et $X_{t_4} = 2$. On a :

$$\begin{aligned} \mathbb{P}[X_{t_0} = 0; X_{t_1} = 0; X_{t_2} = 1; X_{t_3} = 1; X_{t_4} = 2] &= \prod_{u=t_0}^{t_1-1} P_u(0, 0) \left[\prod_{u=t_1}^{t_2-1} P_u \right] (0, 1) \prod_{u=t_2}^{t_3-1} P_u(1, 1) \left[\prod_{u=t_3}^{t_4-1} P_u \right] (1, 2) \\ &\stackrel{(hom.)}{=} P(0, 0)^{t_1-t_0} P^{t_2-t_1}(0, 1) P(1, 1)^{t_3-t_2} P^{t_4-t_3}(1, 2). \end{aligned}$$

Cas 18 $X_{t_0} = 0, X_{t_1} = 0, X_{t_2} = 1, X_{t_3} = 1$ et $X_{t_4} = 3$. On a :

$$\begin{aligned} \mathbb{P}[X_{t_0} = 0; X_{t_1} = 0; X_{t_2} = 1; X_{t_3} = 1; X_{t_4} = 3] &= \prod_{u=t_0}^{t_1-1} P_u(0, 0) \left[\prod_{u=t_1}^{t_2-1} P_u \right] (0, 1) \prod_{u=t_2}^{t_3-1} P_u(1, 1) \left[\prod_{u=t_3}^{t_4-1} P_u \right] (1, 3) \\ &\stackrel{(hom.)}{=} P(0, 0)^{t_1-t_0} P^{t_2-t_1}(0, 1) P(1, 1)^{t_3-t_2} P^{t_4-t_3}(1, 3). \end{aligned}$$

En reprenant la discussion du cas 3, on a aussi :

$$\begin{aligned} \mathbb{P}[X_{t_0} = 0; X_{t_1} = 0; X_{t_2} = 1; X_{t_3} = 1; X_{t_4} = 3] &= \prod_{u=t_0}^{t_1-1} P_u(0, 0) \left[\prod_{u=t_1}^{t_2-1} P_u \right] (0, 1) \prod_{u=t_2}^{t_3-1} P_u(1, 1) \left[\prod_{u=t_3}^{t_4-2} P_u \right] (1, 2) P_{t_4-1}(2, 3) \\ &\stackrel{(hom.)}{=} P(0, 0)^{t_1-t_0} P^{t_2-t_1}(0, 1) P(1, 1)^{t_3-t_2} P^{t_4-t_3-1}(1, 2) P(2, 3). \end{aligned}$$