

Analyse statistique de modèles de fiabilité en environnements dynamiques

Lise Guérineau

EDF R&D



Mesure de l'Information
des Réseaux Electriques
Fiabilité et Disponibilité
des Réseaux

Université de Bretagne Sud



Laboratoire de
Mathématiques
de Bretagne
Atlantique

Sous la direction de Evans Gouno
Encadrants industriels : Philippe Carer & Romain Lattes

- ▶ Planification et exploitation du réseau électrique
- ▶ Objectifs : **fiabilité & économie**
- ▶ Différentes approches pour satisfaire au mieux ces objectifs :
 - ▶ Une approche *déterministe*, qui effectue un compromis entre fiabilité et économie
 - ▶ Une approche *probabiliste* visant à satisfaire un critère de fiabilité
- ▶ Exigences « nouvelles » en matière de qualité d'alimentation
 - ▶ approche *probabiliste* centrée sur des objectifs de fiabilité

Variabilité de la disponibilité du réseau électrique

Contexte
industriel

Contraintes
climatiques
variables

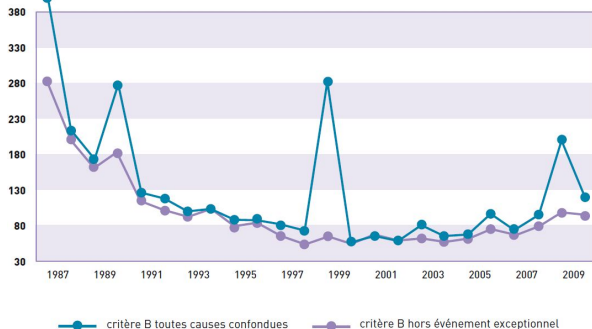
Modèle général
Contraintes de
température et
d'humidité
Contraintes de
surtensions

Système évolutif

Situation
Modèle
Observation
Inférence
Résultats

Conclusion

Evolution de la durée moyenne de coupure par client en minutes





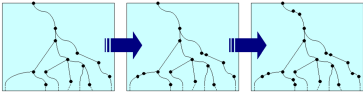
Facteurs d'influence sur la fiabilité du réseau électrique

- ▶ **Stratégie actuelle** : évaluation de la fiabilité hors évènements exceptionnels
 - ▶ Stratégie cohérente avec les exigences actuelles en matière de qualité d'alimentation
 - ▶ La définition des évènements exceptionnels fait débat
 - ▶ Jusqu'où peut-on ignorer les défaillances sans pour autant masquer une forme de vulnérabilité du réseau face aux agressions liées à l'environnement ?
- ▶ **Stratégie visée** : intégration des conditions environnementales dans l'évaluation de la fiabilité
 - ▶ Permet d'affirmer que le niveau de fiabilité observé est acceptable ou non au regard des contraintes climatiques subies par le réseau
 - ▶ Requiert la connaissance des lois de défaillance des matériels électriques en fonction de l'environnement

Objectif

- ▶ Lois de défaillance pour les matériels électriques sous l'effet de contraintes de différentes natures
- ▶ Méthodes adaptées à l'observation du réseau (parfois incomplète)

Cas d'étude industriels

Type de Contrainte	Composant/système
Température/humidité 	Composants du réseau moyenne tension
Foudre 	Transformateurs moyenne tension/basse tension
Contrainte structurelle 	Réseau générique (nœuds interconnectés par des liens)

Contexte industriel

Contraintes
climatiques
variables

Contraintes climatiques variables

Modèle général
Contraintes de
température et
d'humidité
Contraintes de
surtensions

Système évolutif

Système évolutif

Situation
Modèle
Observation
Inférence
Résultats

Conclusion

Conclusion

Contexte industriel

Contraintes
climatiques
variables

Contraintes climatiques variables

Modèle général
Contraintes de
température et
d'humidité
Contraintes de
surtensions

Systeme évolutif

Systeme évolutif

Situation
Modèle
Observation
Inférence
Résultats

Conclusion

Conclusion

Plan

Contexte industriel

Contraintes climatiques variables

Système évolutif

Conclusion

Contexte industriel

Contraintes
climatiques
variables

Contraintes climatiques variables

Modèle général
Contraintes de
température et
d'humidité
Contraintes de
surtensions

Système évolutif

Système évolutif

Situation
Modèle
Observation
Inférence
Résultats

Conclusion

Conclusion

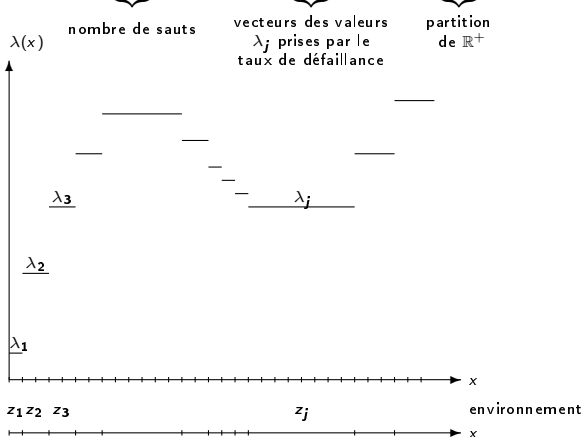
Fiabilité d'un composant
sous l'influence de contraintes
climatiques variables

Modèle général

Hypothèses

1. X durée de bon fonctionnement telle que :

$$X \sim PE(\underbrace{m}_{\text{nombre de sauts}}, \underbrace{\lambda}_{\text{vecteurs des valeurs } \lambda_j \text{ prises par le taux de défaillance}}, \underbrace{\tau}_{\text{partition de } \mathbb{R}^+})$$

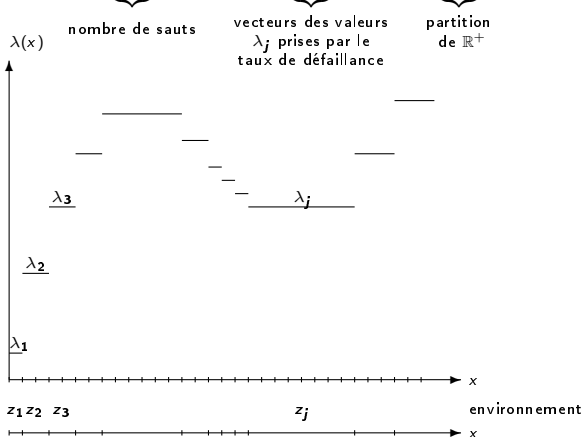


2. Relation paramétrique sur les valeurs du taux de défaillance : $\lambda_j = \lambda_0 e^{\beta' \varphi_j}$
 φ_j dépend de z_j (éventuellement de z_1, \dots, z_{j-1})

Hypothèses

1. X durée de bon fonctionnement telle que :

$$X \sim PE\left(\underbrace{m}_{\text{nombre de sauts}}, \underbrace{\lambda}_{\text{vecteurs des valeurs } \lambda_j \text{ prises par le taux de défaillance}}, \underbrace{\tau}_{\text{partition de } \mathbb{R}^+} \right)$$

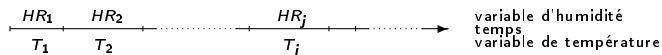


2. Relation paramétrique sur les valeurs du taux de défaillance : $\lambda_j = \lambda_0 e^{\beta' \varphi_j}$
 φ_j dépend de z_j (éventuellement de z_1, \dots, z_{j-1})

Exemples

Composants sous contraintes de température et d'humidité

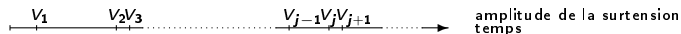
1. Taux de défaillance constant par morceaux



2. Modèle de Peck sur chaque valeur prise par le taux de défaillance : $\lambda_j = \lambda_0 \exp \left\{ -\frac{E_a}{K} \left(\frac{1}{T_j} - \frac{1}{T_0} \right) \right\} \left(\frac{HR_j}{HR_0} \right)^\eta$

Composants sous contraintes d'orage

1. Taux de défaillance constant entre chaque impact de foudre



2. Modèle de type « puissance inverse » paramétrant l'effet cumulé des surtensions : $\lambda_j = \lambda_0 \left(\frac{V_1 \cdots V_j}{V_{ref}^j} \right)^\eta$

$X_i \sim PE(m_i, \lambda_i, \tau_i)$ donc :

$$\lambda_i(x) = \sum_{j=1}^{m+1} \lambda_{i,j} \mathbb{1}_{[\tau_{j-1}; \tau_j]}(x)$$

$$\begin{aligned} R_i(x_i) &= \exp \left\{ - \int_0^{x_i} \lambda(u) du \right\} \\ &= \exp \left\{ - \sum_{j=1}^{m_i+1} \lambda_{i,j} t_{i,j} \right\} \end{aligned}$$

$$\text{avec } t_{i,j} = \begin{cases} 0 & , \text{ si } x_i < \tau_{j-1} \\ x_i - \tau_{j-1} & , \text{ si } \tau_{j-1} \leq x_i < \tau_j \\ \tau_j - \tau_{j-1} & , \text{ si } x_i \geq \tau_j \end{cases}$$

$$f_i(x_i) = \left(\sum_{j=1}^{m_i+1} \lambda_{i,j} \mathbb{1}_{[\tau_{j-1}, \tau_j]}(x_i) \right) \exp \left\{ - \sum_{j=1}^{m_i+1} \lambda_{i,j} t_{i,j} \right\}$$

On observe pour le composant i :

- ▶ la durée de bon fonctionnement X_i
- ▶ l'indicateur de censure δ_i (1 si X_i est observé, 0 sinon)

La contribution de (X_i, δ_i) à la vraisemblance vaut :

$$\begin{aligned}L_i(\boldsymbol{\lambda}^i) &= f_i(x_i)^{\delta_i} R_i(x_i)^{1-\delta_i} \\ &= \left(\sum_{j=1}^{m_i+1} \lambda_{i,j} \mathbb{1}_{[\tau_{j-1}, \tau_j]}(x_i) \right)^{\delta_i} \exp \left\{ - \sum_{j=1}^{m_i+1} \lambda_{i,j} t_{i,j} \right\}\end{aligned}$$

On pose $\delta_{ij} = \begin{cases} 1, & \text{si le composant } i \text{ est défaillant dans } [\tau_{j-1}, \tau_j] \\ 0, & \text{sinon} \end{cases}$

On obtient $L(\boldsymbol{\lambda}) = \prod_{i=1}^n \prod_{j=1}^{m_i+1} (\lambda_{i,j})^{\delta_{ij}} \exp \{-\lambda_{i,j} t_{i,j}\}$

On a la relation $\lambda_{i,j} = \lambda_0 e^{\beta' \varphi_j^i}$, d'où :

$$\log L(\boldsymbol{\lambda}) = \sum_{i=1}^n \sum_{j=1}^{m_i+1} \left(\delta_{ij} \log \lambda_0 + \delta_{ij} \beta' \varphi_j^i - \lambda_0 e^{\beta' \varphi_j^i} t_{i,j} \right)$$

Les équations de vraisemblance sont :

$$\left\{ \begin{array}{l} \frac{\partial}{\partial \lambda_0} \log L(\lambda_0, \beta) = \frac{k}{\lambda_0} - \sum_{i=1}^n \sum_{j=1}^{m_i+1} e^{\beta' \varphi_j^i} t_{i,j} = 0 \\ \frac{\partial}{\partial \beta} \log L(\lambda_0, \beta) = \sum_{i=1}^n \sum_{j=1}^{m_i+1} (\delta_{ij} \varphi_j^i - \lambda_0 \varphi_j^i e^{\beta' \varphi_j^i} t_{i,j}) = 0 \end{array} \right. \quad (1) \quad (2)$$

avec k , le nombre de défaillances observées : $k = \sum_{i=1}^n \delta_i$.

- ▶ Estimation des paramètres β et λ_0 par une méthode numérique
- ▶ Estimation des intervalles de confiance asymptotiques :

$$[\hat{\lambda} + 1, 96H(\hat{\lambda}_0, \hat{\beta})^{-1}; \hat{\lambda} - 1, 96H(\hat{\lambda}_0, \hat{\beta})^{-1}]$$

- ▶ Certains composants du réseau électrique sont sensibles
 - ▶ au phénomène de diffusion d'humidité
 - ▶ aux fortes chaleurs

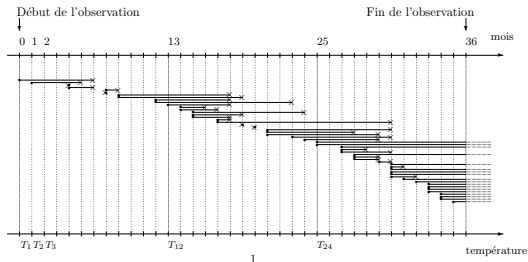


composant



- ▶ Ces facteurs environnementaux agissent sur la rigidité diélectrique et peuvent favoriser l'occurrence de pannes

- ▶ Dates d'installation et dates de défaillance des composants sur le réseau
- ▶ Valeurs de la température au cours de la période d'observation
- ▶ Valeurs de l'humidité en fonction de l'âge du composant



Contexte
industriel

Contraintes
climatiques
variables

Modèle général

Contraintes de
température et
d'humidité

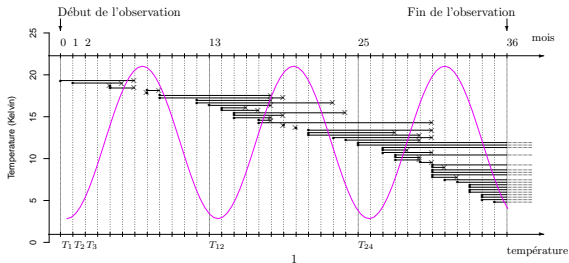
Contraintes de
surtensions

Système évolutif

Situation
Modèle
Observation
Inférence
Résultats

Conclusion

- ▶ Dates d'installation et dates de défaillance des composants sur le réseau
- ▶ Valeurs de la température au cours de la période d'observation
- ▶ Valeurs de l'humidité en fonction de l'âge du composant



Contexte
industriel

Contraintes
climatiques
variables

Modèle général

Contraintes de
température et
d'humidité

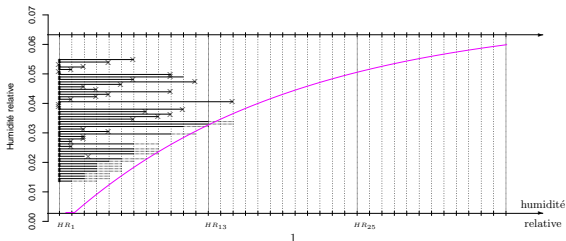
Contraintes de
surtensions

Système évolutif

Situation
Modèle
Observation
Inférence
Résultats

Conclusion

- ▶ Dates d'installation et dates de défaillance des composants sur le réseau
- ▶ Valeurs de la température au cours de la période d'observation
- ▶ Valeurs de l'humidité en fonction de l'âge du composant



Contexte
industriel

Contraintes
climatiques
variables

Modèle général

Contraintes de
température et
d'humidité

Contraintes de
surtensions

Système évolutif

Situation
Modèle
Observation
Inférence
Résultats

Conclusion

- ▶ Tous les composants évoluent suivant le même profil climatique
 - ▶ Covariables de différentes natures
 - ▶ température : covariable externe
 - ▶ humidité : covariable interne
- ⇒ Deux composants ayant la même date d'installation ont la même fiabilité

Le modèle

1. Taux de défaillance constant par morceaux

Pour un composant posé à la date d :

$$\lambda_d(x) = \sum_{j=1}^{m+1} \lambda_{d+j} \mathbb{1}_{[\tau_{j-1}, \tau_j]}(x)$$

2. Valeurs mensuelles du taux de défaillance données par un modèle de Peck

$$\begin{aligned} \lambda_{d+j} &= \lambda_0 \exp \left\{ -\frac{E_a}{K} \left(\frac{1}{T_{d+j}} - \frac{1}{T_{\text{ref}}} \right) \right\} \left(\frac{HR_j}{HR_{\text{ref}}} \right)^\eta \\ &= \lambda_0 e^{\beta' \varphi_{d+j}} \end{aligned}$$

avec

$$\beta = \begin{pmatrix} E_a \\ \eta \end{pmatrix}, \varphi_{d+j} = \begin{pmatrix} -\frac{1}{K} \left(\frac{1}{T_{d+j}} - \frac{1}{T_{\text{ref}}} \right) \\ \log \left(\frac{HR_j}{HR_{\text{ref}}} \right) \end{pmatrix}$$

λ_0 , le taux de défaillance sous la température T_{ref} et sous l'humidité HR_{ref}

1. Taux de défaillance constant par morceaux

Pour un composant posé à la date d :

$$\lambda_d(x) = \sum_{j=1}^{m+1} \lambda_{d+j} \mathbb{1}_{[\tau_{j-1}, \tau_j]}(x)$$

2. Valeurs mensuelles du taux de défaillance données par un modèle de Peck

$$\begin{aligned} \lambda_{d+j} &= \lambda_0 \exp \left\{ -\frac{E_a}{K} \left(\frac{1}{T_{d+j}} - \frac{1}{T_{\text{ref}}} \right) \right\} \left(\frac{HR_j}{HR_{\text{ref}}} \right)^\eta \\ &= \lambda_0 e^{\beta' \varphi_{d+j}} \end{aligned}$$

avec

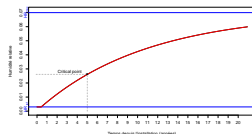
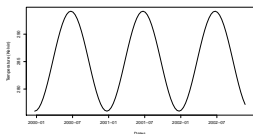
$$\beta = \begin{pmatrix} E_a \\ \eta \end{pmatrix}, \varphi_{d+j} = \begin{pmatrix} -\frac{1}{K} \left(\frac{1}{T_{d+j}} - \frac{1}{T_{\text{ref}}} \right) \\ \log \left(\frac{HR_j}{HR_{\text{ref}}} \right) \end{pmatrix}$$

λ_0 , le taux de défaillance sous la température T_{ref} et sous l'humidité HR_{ref}

Hypothèses

- ▶ 50 composants observés pendant 3 années civiles
- ▶ $\tau_j - \tau_{j-1} = 1$ mois
- ▶ 10 composants installés la première année, 15 la deuxième, 25 la dernière; tirage aléatoire des mois de pose (uniforme)
- ▶ valeurs des paramètres λ_0, E_a, η fixées

Facteurs environnementaux



Contexte
industriel

Contraintes
climatiques
variables

Modèle général

Contraintes de
température et
d'humidité

Contraintes de
surtensions

Système évolutif

Situation
Modèle
Observation
Inférence
Résultats

Conclusion

Simulation avec $\lambda_0 = 0.08$, $E_a = 1.5$, $\eta = 1$

Groupe de travail
FIMA
23 mai 2013

Lise Guérineau

Contexte
industriel

Contraintes
climatiques
variables

Modèle général

Contraintes de
température et
d'humidité

Contraintes de
surtensions

Système évolutif

Situation

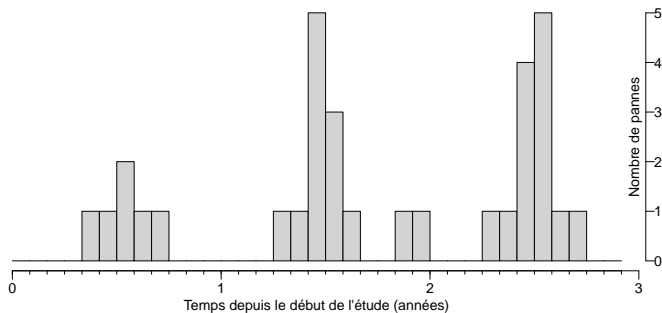
Modèle

Observation

Inférence

Résultats

Conclusion



Résultats de la méthode sur des données simulées

- ▶ Pour chaque combinaison : simulation de 500 échantillons de 50 durées de bon fonctionnement
- ▶ Algorithme de Newton-Raphson initialisé avec :
 $\lambda_0 = 0, 1$
 $E_a = 1$
 $\eta = 3$

	Exact	Moyenne	Biais	EQM
λ_0	0,08	0,08378	0,00378	0,00046
E_a	0,88	0,90365	0,02365	0,05653
η	1,23	1,02521	0,20479	0,30846
λ_0	0,08	0,08737	0,00737	0,00039
E_a	0,1	0,22461	0,12461	0,04501
η	2,56	2,34348	0,21652	0,39207
λ_0	0,08	0,09250	0,01250	0,00061
E_a	1,17	1,08988	0,08012	0,05193
η	3,35	2,88426	0,46574	0,92119
λ_0	0,08	0,08652	0,00652	0,00032
E_a	0,05	0,19523	0,14523	0,04766
η	1,79	1,56142	0,22858	0,33988
λ_0	0,08	0,08866	0,00866	0,00075
E_a	2,1	1,68384	0,41616	0,25014
η	2,95	2,57029	0,37971	1,81504
λ_0	0,88	0,09260	0,01260	0,00077
E_a	1,83	1,52107	0,30893	0,16340
η	3,47	2,91897	0,55103	1,61020

- ▶ 13 006 composants observés pendant $m = 31$ mois
- ▶ 263 composants défectueux, 12 743 censurés (toujours en service à la fin de la période d'observation)
- ▶ Température : moyennes mensuelles enregistrées sur le lieu de l'étude
- ▶ Modélisation déterministe du taux d'humidité interne au composant à l'aide de lois physiques (Fick)

Données réelles : répartition des pannes pendant la période d'observation

Groupe de travail
FIMA
23 mai 2013

Lise Guérineau

Contexte
industriel

Contraintes
climatiques
variables

Modèle général

Contraintes de
température et
d'humidité

Contraintes de
surtensions

Système évolutif

Situation

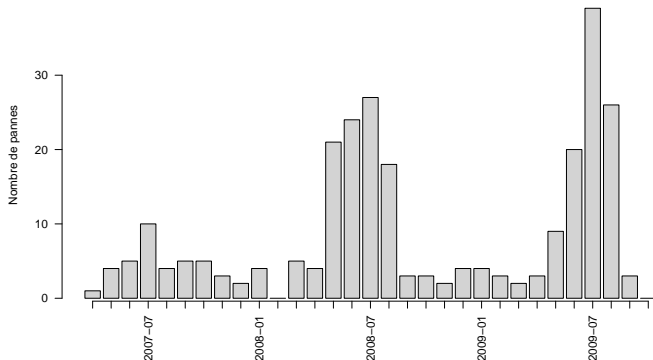
Modèle

Observation

Inférence

Résultats

Conclusion



Estimation

	EMV	$IC_{0,95}$
λ_0	$3,1838 \cdot 10^{-6}$	$[1,1923 \cdot 10^{-6}; 5,1752 \cdot 10^{-6}]$
E_a	0,5076	$[0,3440; 0,6711]$
η	1,8450	$[1,5339; 2,1561]$

Contexte
industriel

Contraintes
climatiques
variables

Modèle général

Contraintes de
température et
d'humidité

Contraintes de
surtensions

Système évolutif

Situation

Modèle

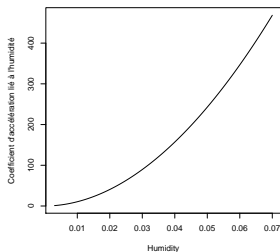
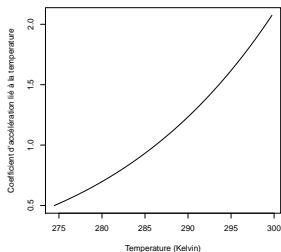
Observation

Inférence

Résultats

Conclusion

Facteurs d'accélération



Résultats de la méthode sur des données réelles

Groupe de travail
FIMA
23 mai 2013

Lise Guérineau

Contexte
industriel

Contraintes
climatiques
variables

Modèle général

Contraintes de
température et
d'humidité

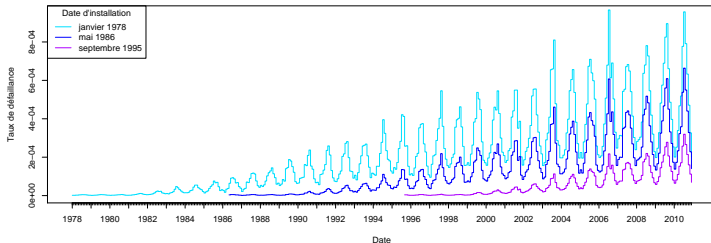
Contraintes de
surtensions

Système évolutif

Situation
Modèle
Observation
Inférence
Résultats

Conclusion

$\lambda(x)$ pour des dates d'installation successives



Conclusion

- ▶ Modèle de fiabilité pour des composants sensibles à la combinaison température \times humidité
- ▶ Incorporation de variables interne/externe
- ▶ Application sur des données réelles :
 - ▶ exhiber des grandeurs physiques caractérisant le vieillissement du composant
 - ▶ simuler l'évolution d'un parc sous l'hypothèse d'une composition initiale (dates d'installation), de l'évolution de la température et de l'humidité,
 - ▶ déterminer les composants les plus à risque sur le réseau en calculant les valeurs courantes de taux de défaillance
 - ▶ éclairer les décisions concernant le renouvellement des matériels
- ▶ Perspectives :
 - ▶ Effet de mémoire pour la température
 - ▶ Incorporation des connaissances des experts sur les mécanismes physiques en jeu dans la défaillance (modèle bayésien)

Certains équipements électriques sont très sensibles à la foudre ; leurs défaillances peuvent conduire à des coupures d'électricité

- ▶ Les transformateurs moyenne tension/basse tension situés en haut des poteaux sont vulnérables face aux impacts de foudre
- ▶ Les parafoudres protègent ces transformateurs en écrêtant les surtensions engendrées par la foudre
- ▶ Ces dispositifs de protection ne sont pas destinés à protéger les transformateurs de tout impact de foudre : un compromis est établi entre le niveau d'isolement, le niveau de protection des parafoudres et le risque de défaillance admissible



Contexte
industriel

Contraintes
climatiques
variables

Modèle général
Contraintes de
température et
d'humidité

Contraintes de
surtensions

Système évolutif

Situation
Modèle
Observation
Inférence
Résultats

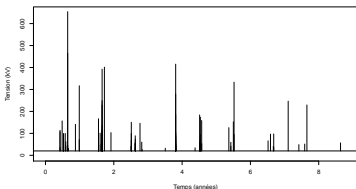
Conclusion

Durées de bon fonctionnement des matériels

- ▶ date de pose
- ▶ date de panne / date de censure

Surtensions atmosphériques

Dates τ_1, \dots, τ_m et valeurs V_1, \dots, V_m des surtensions atmosphériques subies par chacun des transformateurs



- ▶ Evènements majeurs suivis de répliques
- ▶ On modélise ce phénomène par des processus auto-excités

Processus auto-excité

$\{N(t), t \geq 0\}$ processus de comptage d'intensité :

$$\begin{aligned}\lambda(t) &= \mu + \int_0^t g(t-s) dN(s) \\ &= \underbrace{\mu}_{\substack{\text{intensité « de base »} \\ \uparrow \\ \text{évènements majeurs}}} + \underbrace{\sum_{t_i < t} g(t-t_i)}_{\substack{\text{partie auto-excitée} \\ \uparrow \\ \text{influence des évènements passés}}}\end{aligned}$$

où $\mu > 0$, $g : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ fonction décroissante, par exemple :

- ▶ $g(t) = \frac{K}{(t+c)^p}$
- ▶ $g(t) = \alpha e^{-\beta t}$

Processus auto-excité

$\{N(t), t \geq 0\}$ processus de comptage d'intensité :

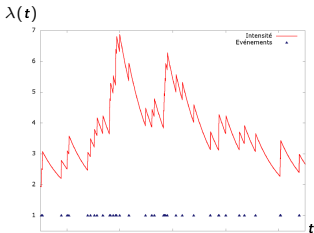
$$\begin{aligned}\lambda(t) &= \mu + \int_0^t g(t-s) dN(s) \\ &= \underbrace{\mu}_{\substack{\text{intensité « de base »} \\ \uparrow \\ \text{évènements majeurs}}} + \underbrace{\sum_{t_i < t} g(t-t_i)}_{\substack{\text{partie auto-excitée} \\ \uparrow \\ \text{influence des évènements passés}}}\end{aligned}$$

où $\mu > 0$, $g : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ fonction décroissante, par exemple :

▶ $g(t) = \frac{K}{(t+c)^p}$

▶ $g(t) = \alpha e^{-\beta t}$

processus de
Hawkes



- ▶ **Estimation** : Ozaki (1979) a proposé une procédure d'estimation¹ des paramètres μ, α, β reposant sur l'observation d'une trajectoire

- ▶ **Adéquation** : Ogata (1988) a proposé un test d'adéquation au processus de Hawkes

Si τ_1, \dots, τ_m sont les dates de sauts d'un processus de Hawkes, alors pour $j = 1, \dots, m$:

⇒ $\Lambda(\tau_j)$ sont les dates de sauts d'un PPH(1)

⇒ $\Lambda(\tau_j) - \Lambda(\tau_{j-1}) \sim \mathcal{E}(1)$

⇒ $1 - \exp\{-\Lambda(\tau_j) - \Lambda(\tau_{j-1})\} \sim \mathcal{U}[0, 1]$

En pratique, on va chercher à identifier la loi des $\hat{\Lambda}(\tau_j) - \hat{\Lambda}(\tau_{j-1})$, où :

$$\hat{\Lambda}(\tau_j) = \hat{\mu}\tau_j + \frac{\hat{\alpha}}{\hat{\beta}} \sum_{k=2}^j \sum_{i=1}^{k-1} \left(e^{-\hat{\beta}(\tau_k - \tau_i)} - e^{-\hat{\beta}(\tau_{k-1} - \tau_i)} \right)$$

1. algorithme de Newton-Raphson estimant le maximum de vraisemblance

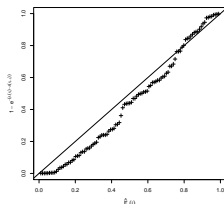
Exemple sur une chronique

Estimateurs du maximum de vraisemblance :

$$\hat{\mu} = 0,014$$

$$\hat{\alpha} = 94,3$$

$$\hat{\beta} = 163$$



$n = 103$

Test de Kolmororov-Smirnov

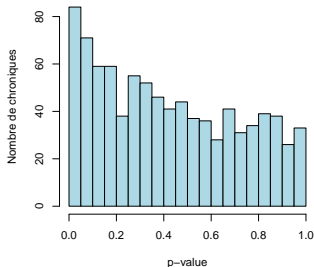
\mathcal{H}_0 : les dates des surtensions sont les dates de saut d'un processus de Hawkes

Statistique de test : $D = 0,1353$

$P(D > 0,150) = 0,02$	Au risque 2%, on ne peut pas rejeter \mathcal{H}_0
$P(D > 0,134) = 0,05$	Au risque 5%, on peut rejeter \mathcal{H}_0

$p - value = 0,046$

- ▶ On s'intéresse à un ensemble de $n = 892$ transformateurs d'une zone géographique donnée
- ▶ Pour la chronique $\tau_1^i, \dots, \tau_{m_i}^i$ des dates de surtension ayant affecté chaque transformateur i :
 1. Estimation par maximum de vraisemblance des paramètres μ, α, β
 2. Test d'adéquation au processus de Hawkes

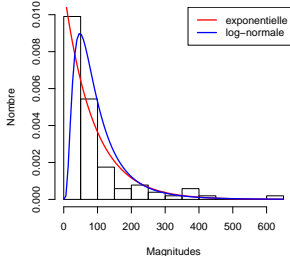


⇒ Le plus petit risque à partir duquel on peut rejeter l'hypothèse \mathcal{H}_0 (p-value) est supérieur à 0,05 pour 90% des chroniques

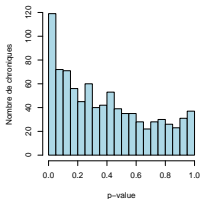
Exemple sur une chronique

On considère les amplitudes V_1, \dots, V_m des surtensions apparues sur un transformateur :

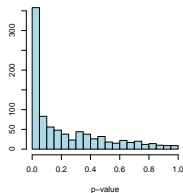
- ▶ Ajustement à une loi exponentielle :
p-value = 0,19
- ▶ Ajustement à une loi log-normale
p-value = 0,25



- ▶ On s'intéresse à un ensemble de $n = 892$ transformateurs d'une zone géographique donnée
- ▶ Pour la chronique $V_1^i, \dots, V_{m_i}^i$ des amplitudes des surtensions ayant affecté chaque transformateur i :
 1. Estimation par maximum de vraisemblance des paramètres d'une loi log-normale, d'une loi exponentielle
 2. Test d'adéquation à la loi exponentielle/log-normale



Loi log-normale



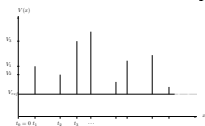
Loi exponentielle

⇒ On considérera :

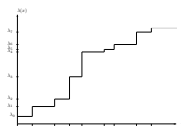
$$V_j^i \sim \text{LogN}(\mu_i, \sigma_i)$$

1. Taux de défaillance constant entre les chocs de foudre successifs

$$\lambda(x) = \sum_{j=0}^{+\infty} \lambda_j \mathbb{1}_{[\tau_j; \tau_{j+1}[}(x)$$



(a)



(b)

2. Valeurs successives du taux de défaillance liées par :

$$\lambda_j = \lambda_{j-1} \left(\frac{V_j}{V_{ref}} \right)^\eta, \quad j \in \mathbb{N}^*$$

Autrement dit,

$$\lambda_j = \lambda_0 \left(\frac{V_1 \cdots V_j}{V_{ref}^j} \right)^\eta, \quad j \in \mathbb{N}^*$$

Contexte
industriel

Contraintes
climatiques
variables

Modèle général
Contraintes de
température et
d'humidité

Contraintes de
surtensions

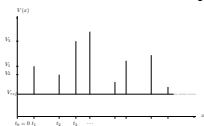
Système évolutif

Situation
Modèle
Observation
Inférence
Résultats

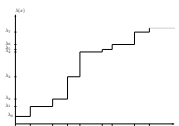
Conclusion

1. Taux de défaillance constant entre les chocs de foudre successifs

$$\lambda(x) = \sum_{j=0}^{+\infty} \lambda_j \mathbb{1}_{[\tau_j; \tau_{j+1}[}(x)$$



(a)



(b)

2. Valeurs successives du taux de défaillance liées par :

$$\lambda_j = \lambda_{j-1} \left(\frac{V_j}{V_{ref}} \right)^\eta, \quad j \in \mathbb{N}^*$$

Autrement dit,

$$\lambda_j = \lambda_0 \left(\frac{V_1 \cdots V_j}{V_{ref}^j} \right)^\eta, \quad j \in \mathbb{N}^*$$

On fixe les paramètres du modèle :

- ▶ $\lambda_0 = 6, 8 \cdot 10^{-5} \Leftrightarrow$ sans contrainte de foudre, $E[X] = 40$ ans
- ▶ $\eta = 2 \cdot 10^{-2}, \eta = 1, 3 \cdot 10^{-3}$

Pour $i = 1, \dots, n$:

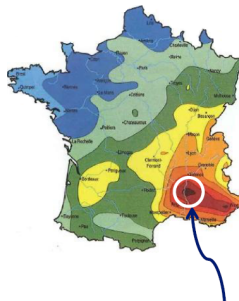
1. On simule les impacts de foudre subis par le transformateur i :
 - ▶ dates d'impact sur une période de 40 ans (Hawkes)
 - ▶ amplitudes associées (log-normale)
2. On calcule l'évolution du taux de défaillance du transformateur i jusqu'à la dernière date d'impact
3. On simule une durée de bon fonctionnement suivant la loi ainsi calculée

En moyenne, 1% des données sont censurées

	Exact	Moyenne	Biais	EQM
λ_0	$6,8 \cdot 10^{-5}$	$6,63 \cdot 10^{-5}$	$1,66 \cdot 10^{-6}$	$1,88 \cdot 10^{-7}$
η	$2 \cdot 10^{-2}$	$2,07 \cdot 10^{-2}$	$7,4 \cdot 10^{-4}$	$9,35 \cdot 10^{-3}$

En moyenne, 1/3 des données sont censurées

	Exact	Moyenne	Biais	EQM
λ_0	$6,8 \cdot 10^{-5}$	$7,08 \cdot 10^{-5}$	$2,85 \cdot 10^{-6}$	$2,30 \cdot 10^{-7}$
η	$1,30 \cdot 10^{-3}$	$1,06 \cdot 10^{-3}$	$2,35 \cdot 10^{-4}$	$9,10 \cdot 10^{-3}$



Carte de
niveau
céramique

Echantillon : $n = 969$ transformateurs de Drome-Ardèche

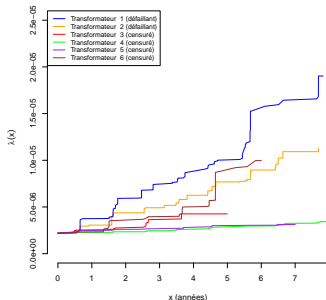
	Effectif
Données de panne	8
Données de patrimoine	961
TOTAL	969

Estimateurs du maximum de vraisemblance (IC à 95 %) :

$$\hat{\lambda}_0 = 2,2 \cdot 10^{-6} \quad (\pm 2,6 \cdot 10^{-6})$$

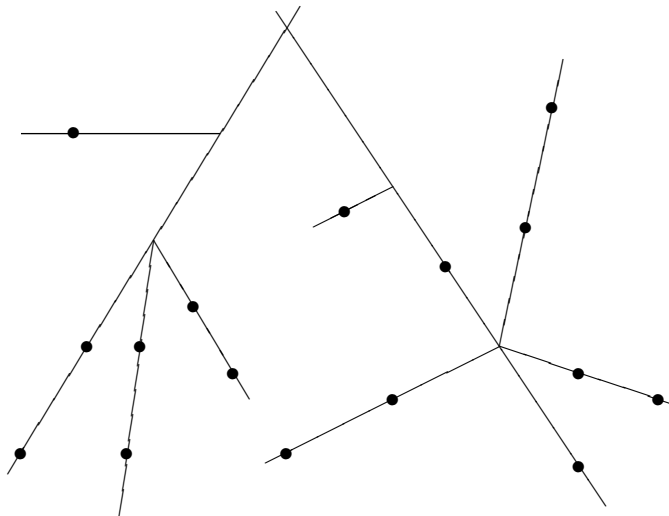
$$\hat{\eta} = 2,074 \cdot 10^{-2} \quad (\pm 2,1 \cdot 10^{-2})$$

Taux de défaillance pour 6 transformateurs :



- ▶ Modélisation de la fiabilité des transformateurs « haut de poteau »
- ▶ Estimation de paramètres qualifiant la fiabilité du matériel
- ▶ Validation du modèle par confrontation avec des résultats expérimentaux
 - ▶ EDF R&D étudie des marqueurs chimiques permettant de quantifier le vieillissement des transformateurs
 - ▶ Confrontation des résultats des marqueurs avec l'estimation de la fiabilité fournie par notre modèle
- ▶ Contribution à l'aide à la décision sur les choix de renouvellement des transformateurs
 - ▶ Classement des transformateurs actifs suivant leur état de vieillissement
 - ▶ Alerte sur les transformateurs dans un état critique

Système composé de nœuds (●) et de liens (—)



Contexte
industriel

Contraintes
climatiques
variables

Modèle général
Contraintes de
température et
d'humidité
Contraintes de
surtensions

Système évolutif

Situation

Modèle

Observation

Inférence

Résultats

Conclusion

Evolution du système suite aux maintenances correctives

Groupe de travail
FIMA
23 mai 2013

Lise Guérineau

Contexte
industriel

Contraintes
climatiques
variables

Modèle général

Contraintes de
température et
d'humidité

Contraintes de
surtensions

Système évolutif

Situation

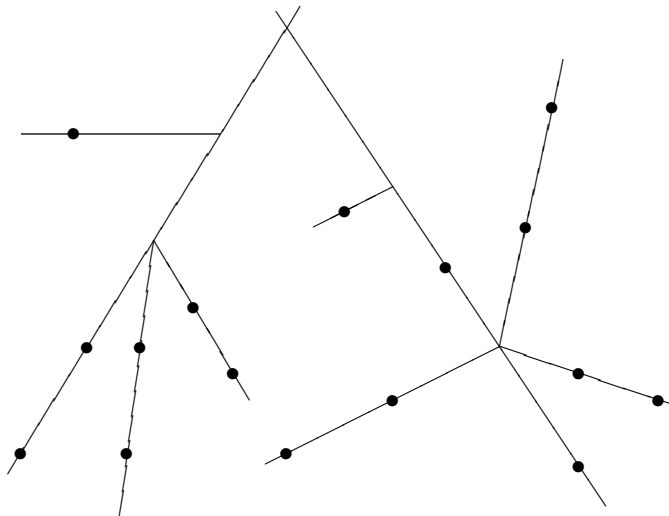
Modèle

Observation

Inférence

Résultats

Conclusion

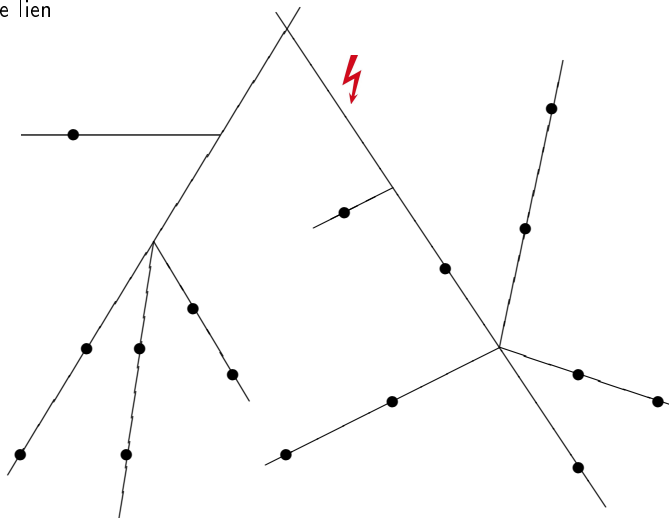


Evolution du système suite aux maintenances correctives

Groupe de travail
FIMA
23 mai 2013

Lise Guérineau

panne lien



Contexte
industriel

Contraintes
climatiques
variables

Modèle général
Contraintes de
température et
d'humidité
Contraintes de
surtensions

Système évolutif

Situation

Modèle

Observation

Inférence

Résultats

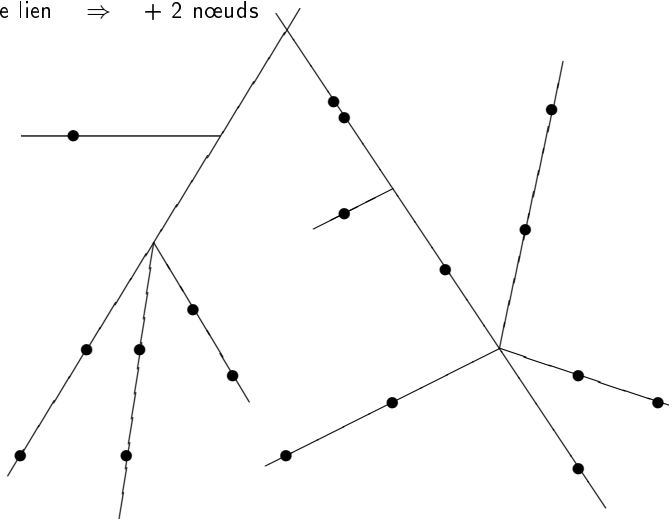
Conclusion

Evolution du système suite aux maintenances correctives

Groupe de travail
FIMA
23 mai 2013

Lise Guérineau

panne lien \Rightarrow + 2 nœuds



Contexte
industriel

Contraintes
climatiques
variables

Modèle général
Contraintes de
température et
d'humidité
Contraintes de
surtensions

Système évolutif

Situation

Modèle

Observation

Inférence

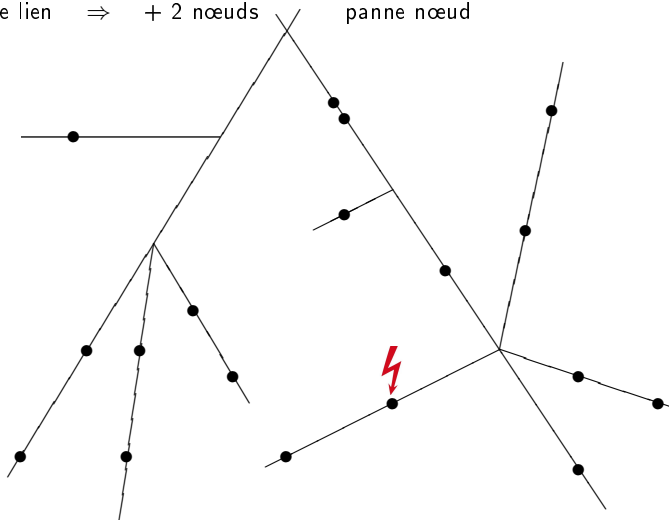
Résultats

Conclusion

Evolution du système suite aux maintenances correctives

Groupe de travail
FIMA
23 mai 2013
Lise Guérineau

panne lien \Rightarrow + 2 nœuds panne nœud



Contexte
industriel

Contraintes
climatiques
variables

Modèle général
Contraintes de
température et
d'humidité
Contraintes de
surtensions

Système évolutif

Situation

Modèle

Observation

Inférence

Résultats

Conclusion

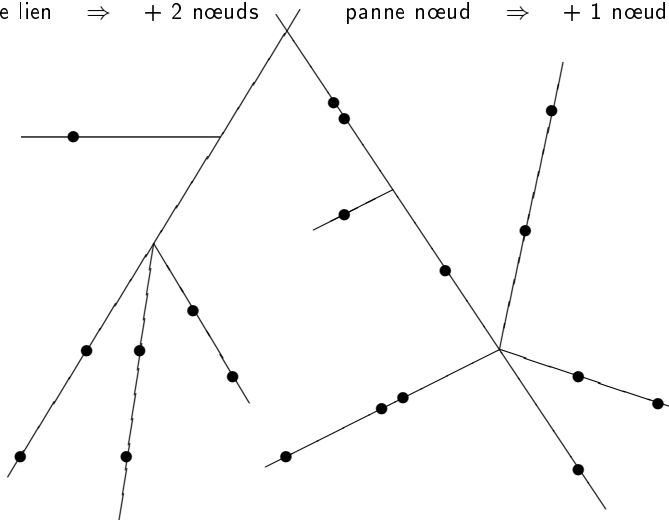
Evolution du système suite aux maintenances correctives

Groupe de travail
FIMA
23 mai 2013

Lise Guérineau

panne lien \Rightarrow + 2 nœuds

panne nœud \Rightarrow + 1 nœud



Contexte
industriel

Contraintes
climatiques
variables

Modèle général
Contraintes de
température et
d'humidité
Contraintes de
surtensions

Système évolutif

Situation

Modèle

Observation

Inférence

Résultats

Conclusion

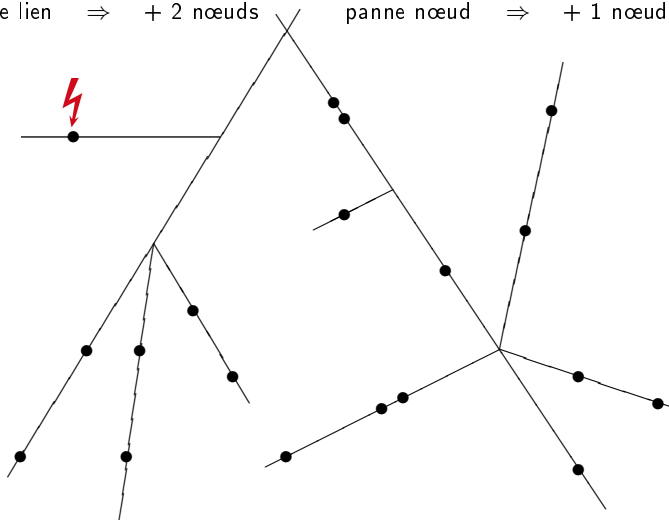
Evolution du système suite aux maintenances correctives

Groupe de travail
FIMA
23 mai 2013

Lise Guérineau

panne lien \Rightarrow + 2 nœuds

panne nœud \Rightarrow + 1 nœud



Contexte
industriel

Contraintes
climatiques
variables

Modèle général
Contraintes de
température et
d'humidité
Contraintes de
surtensions

Système évolutif

Situation

Modèle

Observation

Inférence

Résultats

Conclusion

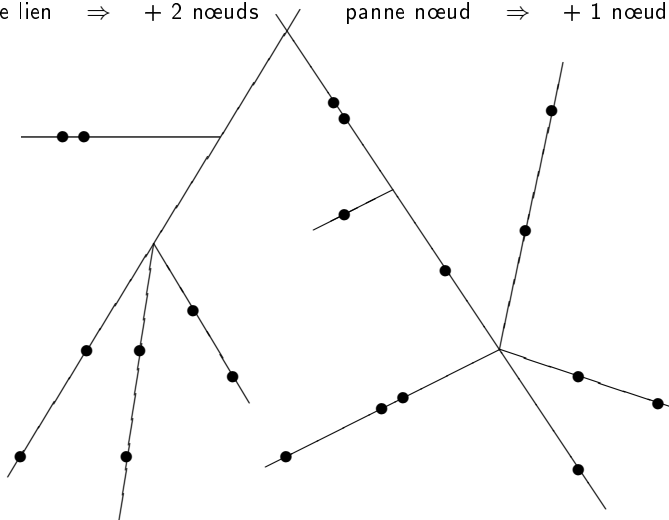
Evolution du système suite aux maintenances correctives

Groupe de travail
FIMA
23 mai 2013

Lise Guérineau

panne lien \Rightarrow + 2 nœuds

panne nœud \Rightarrow + 1 nœud



Contexte
industriel

Contraintes
climatiques
variables

Modèle général
Contraintes de
température et
d'humidité
Contraintes de
surtensions

Système évolutif

Situation

Modèle

Observation

Inférence

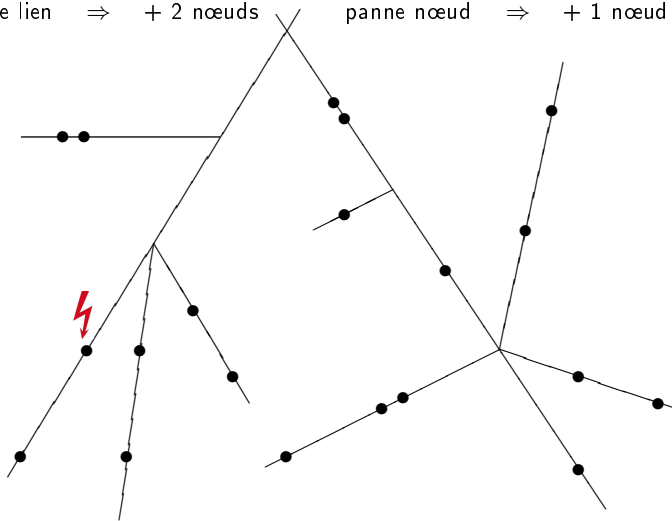
Résultats

Conclusion

Evolution du système suite aux maintenances correctives

panne lien \Rightarrow + 2 nœuds

panne nœud \Rightarrow + 1 nœud



Contexte
industriel

Contraintes
climatiques
variables

Modèle général
Contraintes de
température et
d'humidité
Contraintes de
surtensions

Système évolutif

Situation

Modèle

Observation

Inférence

Résultats

Conclusion

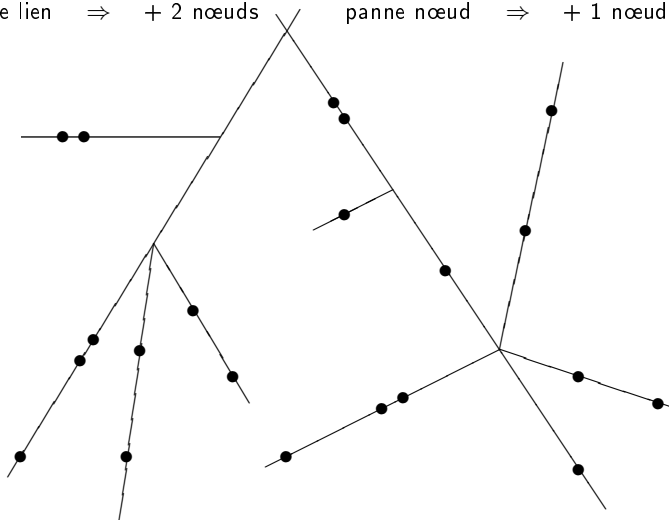
Evolution du système suite aux maintenances correctives

Groupe de travail
FIMA
23 mai 2013

Lise Guérineau

panne lien \Rightarrow + 2 nœuds

panne nœud \Rightarrow + 1 nœud



Contexte
industriel

Contraintes
climatiques
variables

Modèle général
Contraintes de
température et
d'humidité
Contraintes de
surtensions

Système évolutif

Situation

Modèle

Observation

Inférence

Résultats

Conclusion

Contexte

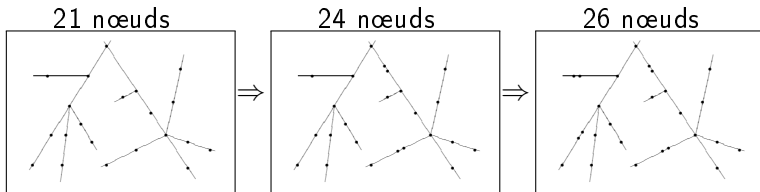
L'introduction de nouveaux nœuds, nécessaire pour assurer la disponibilité du système, dégrade sa fiabilité

Objectif

Estimer la fiabilité des nœuds et du lien dans un contexte d'observation incomplète

Données disponibles

« photos » du système en différents instants :



Contexte
industriel

Contraintes
climatiques
variables

Modèle général
Contraintes de
température et
d'humidité
Contraintes de
surtensions

Système évolutif

Situation

Modèle

Observation

Inférence

Résultats

Conclusion

Connaissances experts

- ▶ Le flux de pannes sur les liens est constant

Contexte
industriel

Contraintes
climatiques
variables

Modèle général
Contraintes de
température et
d'humidité
Contraintes de
surtensions

Système évolutif

Situation

Modèle

Observation

Inférence

Résultats

Conclusion

Connaissances experts

- ▶ Le flux de pannes sur les liens est constant

Modèle

- ▶ Le nombre de pannes lien $(X_t)_{t \geq 0}$ est un PPH

Contexte
industriel

Contraintes
climatiques
variables

Modèle général
Contraintes de
température et
d'humidité
Contraintes de
surtensions

Système évolutif

Situation

Modèle

Observation
Inférence
Résultats

Conclusion

Connaissances experts

- ▶ Le flux de pannes sur les liens est constant
- ▶ La probabilité de défaillance du système est proportionnelle au nombre de nœuds

Modèle

- ▶ Le nombre de pannes lien $(X_t)_{t \geq 0}$ est un PPH

Contexte
industriel

Contraintes
climatiques
variables

Modèle général
Contraintes de
température et
d'humidité
Contraintes de
surtensions

Système évolutif

Situation

Modèle

Observation
Inférence
Résultats

Conclusion

Connaissances experts

- ▶ Le flux de pannes sur les liens est constant
- ▶ La probabilité de défaillance du système est proportionnelle au nombre de nœuds

Modèle

- ▶ Le nombre de pannes lien $(X_t)_{t \geq 0}$ est un PPH
- ▶ Le nombre de pannes nœud $(Y_t)_{t \geq 0}$ est un **processus de naissance** influencé par $(X_t)_{t \geq 0}$

Contexte
industriel

Contraintes
climatiques
variables

Modèle général
Contraintes de
température et
d'humidité
Contraintes de
surtensions

Système évolutif

Situation

Modèle

Observation

Inférence

Résultats

Conclusion

Connaissances experts

- ▶ Le flux de pannes sur les liens est constant
- ▶ La probabilité de défaillance du système est proportionnelle au nombre de nœuds
- ▶ Mécanisme d'évolution du système fonction des maintenances correctives

Modèle

- ▶ Le nombre de pannes lien $(X_t)_{t \geq 0}$ est un PPH
- ▶ Le nombre de pannes nœud $(Y_t)_{t \geq 0}$ est un **processus de naissance** influencé par $(X_t)_{t \geq 0}$

Contexte
industriel

Contraintes
climatiques
variables

Modèle général
Contraintes de
température et
d'humidité
Contraintes de
surtensions

Système évolutif

Situation

Modèle

Observation
Inférence
Résultats

Conclusion

Connaissances experts

- ▶ Le flux de pannes sur les liens est constant
- ▶ La probabilité de défaillance du système est proportionnelle au nombre de nœuds
- ▶ Mécanisme d'évolution du système fonction des maintenances correctives

Modèle

- ▶ Le nombre de pannes lien $(X_t)_{t \geq 0}$ est un PPH
- ▶ Le nombre de pannes nœud $(Y_t)_{t \geq 0}$ est un **processus de naissance** influencé par $(X_t)_{t \geq 0}$
- ▶ $(Z_t)_{t \geq 0}$ processus qui compte le nombre de nœuds sur le système
$$Z_t = z_0 + 2X_t + Y_t$$

Contexte
industriel

Contraintes
climatiques
variables

Modèle général
Contraintes de
température et
d'humidité
Contraintes de
surtensions

Système évolutif

Situation

Modèle

Observation
Inférence
Résultats

Conclusion

$(Z_t)_{t \geq 0}$ est un processus markovien de sauts de noyau de transition :

$$P(Z_{t+h} - Z_t = j | Z_t = i) = \begin{cases} \lambda i h + o(h) & , \text{ si } j = i + 1 \\ \nu h + o(h) & , \text{ si } j = i + 2 \\ 1 - (\lambda i + \theta) h + o(h) & , \text{ si } j = i \\ o(h) & , \text{ sinon} \end{cases}$$

$(Z_t)_{t \geq 0}$ est un processus de naissance avec des immigrations par paires

Différents schémas d'observation

	$(X_t, Y_t)_{t \geq 0}$	$(Z_t)_{t \geq 0}$
Schéma d'observation continu observation sur $[0, C]$		×
Schéma d'observation discret : observations aux dates t_1, \dots, t_n		

Contexte
industriel

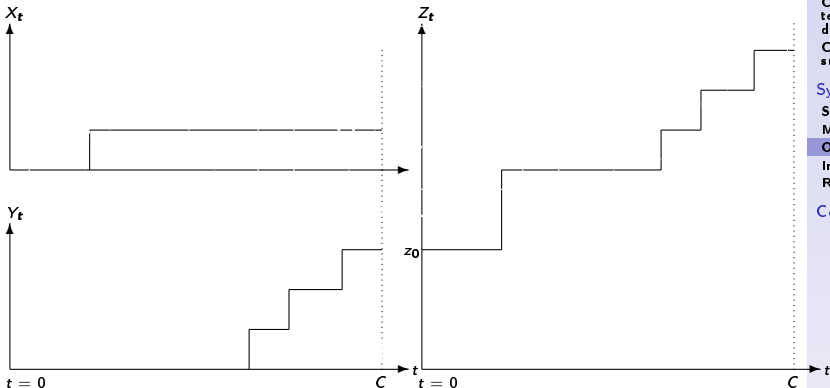
Contraintes
climatiques
variables

Modèle général
Contraintes de
température et
d'humidité
Contraintes de
surtensions

Système évolutif

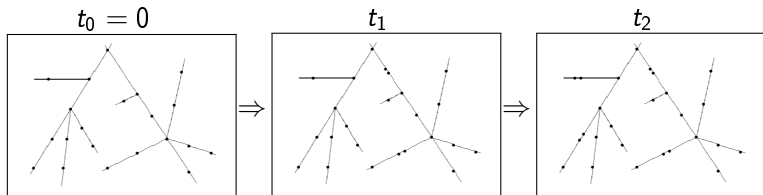
Situation
Modèle
Observation
Inférence
Résultats

Conclusion



Différents schémas d'observation

	$(X_t, Y_t)_{t \geq 0}$	$(Z_t)_{t \geq 0}$
Schéma d'observation continu observation sur $[0, C]$		×
Schéma d'observation discret : observations aux dates t_1, \dots, t_n	×	



$[0, t_1]$ | $[t_1, t_2]$
 1 panne lien, 1 panne nœud | 0 panne lien, 2 pannes nœud

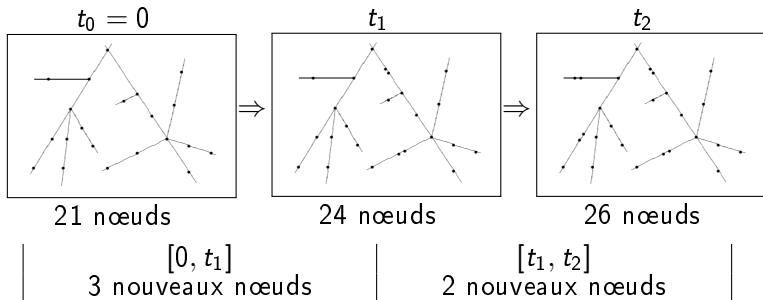
Observation

Nombre de nœuds initial z_0 . Pour $i = 1, \dots, n$:

$[t_{i-1}, t_i]$
 Δx_i pannes lien, Δy_i pannes nœud

Différents schémas d'observation

	$(X_t, Y_t)_{t \geq 0}$	$(Z_t)_{t \geq 0}$
Schéma d'observation continu observation sur $[0, C]$		×
Schéma d'observation discret : observations aux dates t_1, \dots, t_n	×	×



Observation

Nombre de nœuds initial z_0 . Pour $i = 1, \dots, n$:

$$\Delta z_i \text{ nouveaux nœuds}$$

- ▶ Résultats existants sur l'estimation des processus markoviens de saut (Billingsley, 1961) ou sur l'estimation des processus de comptage (Andersen, 1995)

$$\Rightarrow L(\nu, \lambda) = \nu^{X_C} \lambda^{Y_C} e^{-(\lambda R_C + \nu C)},$$

avec $R_t = \int_0^t Z_u du$

- ▶ Estimateurs du maximum de vraisemblance :

$$\hat{\nu} = X_C / C \quad \hat{\lambda} = Y_C / R_C$$

- ▶ Hessienne associée à la log-vraisemblance

$$\begin{pmatrix} -X_C / \nu^2 & 0 \\ 0 & -Y_C / \lambda^2 \end{pmatrix}$$

Schéma d'observation discret du processus

$$(X_t, Y_t)_{t \geq 0}$$

On note pour $i = 1, \dots, n$:

$L_i = t_i - t_{i-1}$, les longueurs des intervalles d'observation,

$\underline{\Delta X}_i$, le vecteur $(\Delta X_1, \dots, \Delta X_i)$

$\underline{\Delta Y}_i$, le vecteur $(\Delta Y_1, \dots, \Delta Y_i)$

$\underline{\Delta Z}_i$, le vecteur $(\Delta Z_1, \dots, \Delta Z_i)$

- ▶ $(X_t)_{t \geq 0}$ processus de Poisson de paramètre ν

$$\Rightarrow P(\Delta X_i = \Delta x_i) = \frac{(\nu L_i)^{\Delta x_i}}{(\Delta x_i)!} e^{-\nu L_i}$$

- ▶ $(Y_t)_{t \geq 0}$ processus dont l'évolution après t ne dépend du passé que par la valeur Z_t

On fait l'hypothèse que les nœuds installés suite aux maintenances correctives dans l'intervalle $[t_{i-1}, t_i]$ ne peuvent eux-même être défaillants avant t_i

$$\Rightarrow P(\Delta Y_i = \Delta y_i \mid \underline{\Delta Z}_{i-1} = \underline{\Delta z}_{i-1}) \approx \binom{\bar{z}_{i-1} + \Delta y_i - 1}{\Delta y_i} e^{-\lambda L_i \bar{z}_{i-1}} (1 - e^{-\lambda L_i})^{\Delta y_i}$$

- ▶ Loi jointe de $(\Delta X_i, \Delta Y_i), i = 1, \dots, n$:

$$\begin{aligned} & P(\Delta X_i = \Delta x_i, \Delta Y_i = \Delta y_i \mid \underline{\Delta Z_{i-1}} \underline{\Delta z_{i-1}}) \\ &= P(\Delta X_i = \Delta x_i) P(\Delta Y_i = \Delta y_i \mid \underline{\Delta Z_{i-1}} = \underline{\Delta z_{i-1}}) \end{aligned}$$

- ▶ Log-vraisemblance

$$\begin{aligned} \log f(\underline{\Delta x}, \underline{\Delta y}; \nu, \lambda) &\propto \left(\sum_{i=1}^n \Delta x_i \right) \log \nu - \nu t_n \\ &\quad - \lambda \sum_{i=1}^n \bar{z}_{i-1} L_i + \sum_{i=1}^n \Delta y_i \log(1 - e^{-\lambda L_i}) \end{aligned}$$

- ▶ Solution explicite en fréquentiste seulement pour $L_i = L, i = 1, \dots, n$:

$$\hat{\nu} = \sum_{i=1}^n \Delta x_i / t_n, \quad \hat{\lambda} = -\log \left(\frac{\sum_{i=1}^n \bar{z}_{i-1}}{\sum_{i=1}^n (\bar{z}_{i-1} + \Delta y_i)} \right) / L$$

- ▶ Solutions explicites en bayésien sous l'approximation

$$1 - e^{-\lambda L_i} \approx \lambda L_i :$$

- ▶ A priori $\nu \sim \mathcal{G}(\alpha, \beta)$, $\lambda \sim \mathcal{G}(\varphi, \Gamma)$

$$\hat{\nu} = \left(\sum_{i=1}^n \Delta x_i + \alpha \right) / (\beta + t_n) \quad , \quad \hat{\lambda} = \left(\sum_{i=1}^n \Delta y_i + \varphi \right) / \left(\Gamma + \sum_{i=1}^n L_i \bar{z}_{i-1} \right)$$

- ▶ A priori non informatif $\pi(\nu, \lambda) \propto \nu^{-1/2} \lambda^{-1/2}$

$$\hat{\nu} = \left(\sum_{i=1}^n \Delta x_i + 1/2 \right) / t_n \quad , \quad \hat{\lambda} = \left(\sum_{i=1}^n \Delta y_i + 1/2 \right) / \sum_{i=1}^n L_i \bar{z}_{i-1}$$

- ▶ Solutions explicites en bayésien pour $L_j = L$:

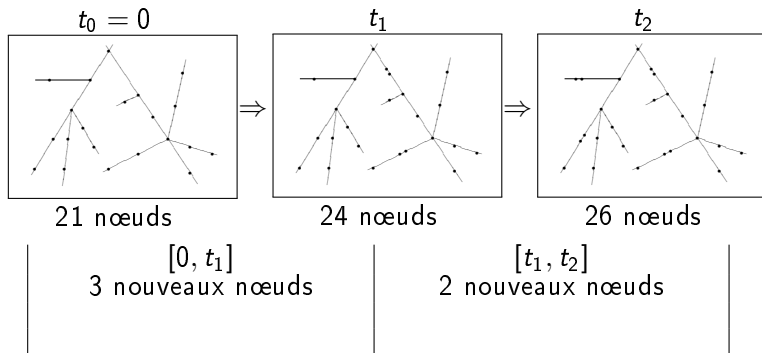
- ▶ A priori $\nu \sim \mathcal{G}(\alpha, \beta)$, $\xi \sim \mathcal{B}(\varphi, \Gamma)$, avec $\xi = e^{-\lambda L}$

$$\hat{\nu} = \left(\sum_{i=1}^n \Delta x_i + \alpha \right) / (\beta + nL) \quad , \quad \hat{\lambda} = -\log \left(\frac{\sum_{i=1}^n \bar{z}_{i-1} + \varphi}{\sum_{i=1}^n (\bar{z}_{i-1} + \Delta y_i) + \varphi + \Gamma} \right) / L$$

- ▶ A priori non informatif $\pi(\nu, \lambda) \propto \nu^{-1/2} (1 - e^{-\lambda L})^{-1/2}$

$$\hat{\nu} = \left(\sum_{i=1}^n \Delta x_i + 1/2 \right) / nL \quad , \quad \hat{\lambda} = -\log \left(\frac{\sum_{i=1}^n \bar{z}_{i-1} + 1}{\sum_{i=1}^n (\bar{z}_{i-1} + \Delta y_i) + 3/2} \right) / L$$

Schéma d'observation discret du processus $(Z_t)_{t \geq 0}$



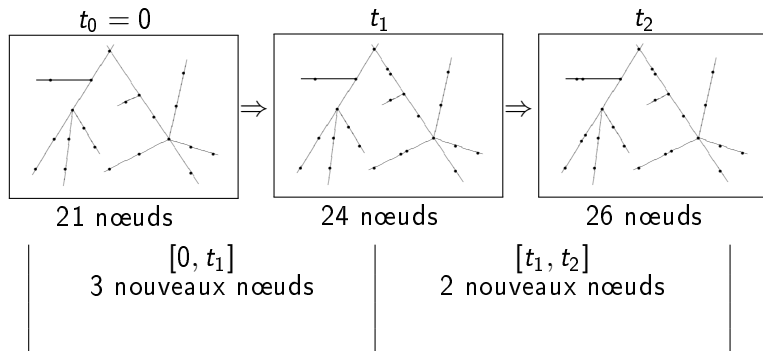
Observation

Nombre de nœuds initial z_0 . Pour $i = 1, \dots, n$:

$$\begin{array}{c} [t_{i-1}, t_i] \\ \Delta z_i \text{ nouveaux nœuds} \end{array}$$

Schéma d'observation discret du processus $(Z_t)_{t \geq 0}$

Un problème de données manquantes



Observation

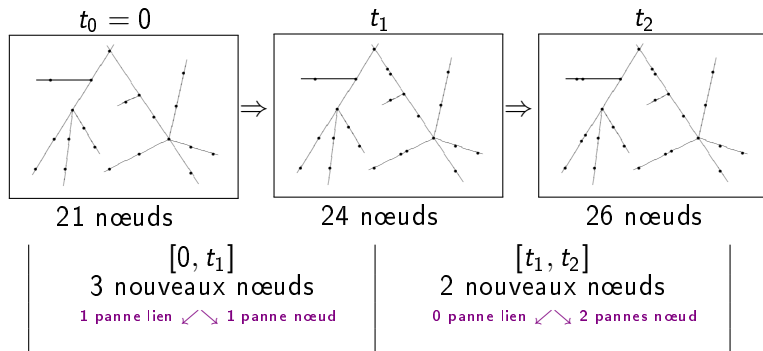
Nombre de nœuds initial z_0 . Pour $i = 1, \dots, n$:

$$\Delta z_i \text{ nouveaux nœuds}$$

Schéma d'observation discret du processus $(Z_t)_{t \geq 0}$

Un problème de données manquantes

$(\Delta X, \Delta Y)$ manquant



Observation

Nombre de nœuds initial z_0 . Pour $i = 1, \dots, n$:

$$\begin{aligned} & [t_{i-1}, t_i] \\ & \Delta z_i \text{ nouveaux nœuds} \\ & \Delta x_i \text{ pannes lien } \swarrow \searrow \Delta y_i \text{ pannes nœuds} \end{aligned}$$

Contexte
industriel

Contraintes
climatiques
variables

Modèle général
Contraintes de
température et
d'humidité
Contraintes de
surtensions

Système évolutif

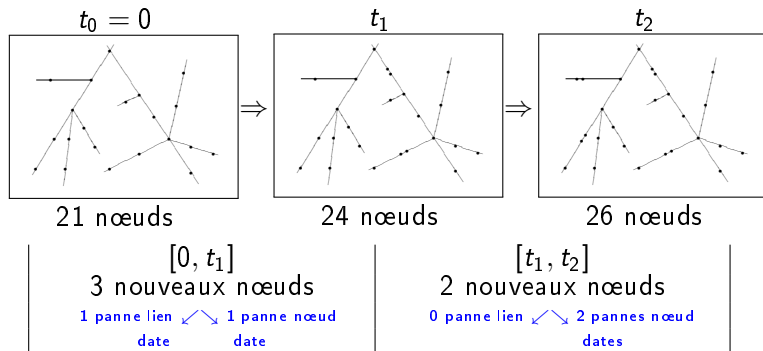
Situation
Modèle
Observation
Inférence
Résultats

Conclusion

Schéma d'observation discret du processus $(Z_t)_{t \geq 0}$

Un problème de données manquantes

$(X_t, Y_t)_{0 \leq t \leq t_n}$ manquant



Contexte
industriel

Contraintes
climatiques
variables

Modèle général
Contraintes de
température et
d'humidité
Contraintes de
surtensions

Système évolutif

Situation
Modèle
Observation
Inference

Résultats

Conclusion

Observation

Nombre de nœuds initial z_0 . Pour $i = 1, \dots, n$:

$$\Delta z_i \text{ nouveaux nœuds}$$
$$(X_t, Y_t)_{t_{i-1} \leq t \leq t_i}$$

- ▶ L'algorithme EM est adapté à notre situation de données manquantes
- ▶ Objectif : estimer ν et λ à partir de l'observation de z_0 et de $\underline{\Delta z} = (\Delta z_1, \dots, \Delta z_n)$

Principe

L'algorithme EM consiste à remplacer la **maximisation de la log-vraisemblance des données observées incomplètes** par la **maximisation de l'espérance conditionnelle de la log-vraisemblance des données complètes non observées sachant les données incomplètes observées**

Les données manquantes sont discrètes

Vraisemblance des données complètes

$$f(\underline{\Delta x}, \underline{\Delta y}; \nu, \lambda) \propto e^{-(\nu t_n + \lambda \sum_{i=1}^n \overline{\Delta z}_{i-1} L_i)} \nu^{\sum_{i=1}^n \Delta x_i} \prod_{i=1}^n (1 - e^{-\lambda L_i})^{\Delta y_i}$$

Densité des données complètes (non observées) sachant les données incomplètes (observées)

$$P(\underline{\Delta x}, \underline{\Delta y} \mid \underline{\Delta z}; \nu, \lambda) = \frac{f(\underline{\Delta x}, \underline{\Delta y}; \nu, \lambda)}{\sum_{(\underline{u}, \underline{v}) \in \mathcal{A}} f(\underline{\Delta u}, \underline{\Delta v}; \nu, \lambda)}$$

$$\mathcal{A} = \{(\underline{x}, \underline{y}) \in \mathbb{N}^n \times \mathbb{N}^n \mid 2\underline{x} + \underline{y} = \underline{z}\}$$

Initialisation : commencer avec $\nu^{(0)}$ et $\lambda^{(0)}$

A la $(p + 1)^{\text{ème}}$ itération, $\theta^{(p)}$ et $\lambda^{(p)}$ étant disponibles :

- **Etape E** : calculer $\mathbb{E}(\log f(\underline{\Delta X}, \underline{\Delta Y}; \nu, \lambda) \mid \underline{\Delta Z}; \nu^{(p)}, \lambda^{(p)})$
Nécessite le calcul de :

$$\mathbb{E}(\Delta X_i \mid \underline{\Delta Z}; \nu^{(p)}, \lambda^{(p)}) = \sum_{(\underline{\Delta u}, \underline{\Delta v}) \in A} \Delta u_i P(\underline{\Delta u}, \underline{\Delta v} \mid \underline{\Delta Z}; \nu^{(p)}, \lambda^{(p)})$$

$$\mathbb{E}(\Delta Y_i \mid \underline{\Delta Z}; \nu^{(p)}, \lambda^{(p)}) = \sum_{(\underline{\Delta u}, \underline{\Delta v}) \in A} \Delta v_i P(\underline{\Delta u}, \underline{\Delta v} \mid \underline{\Delta Z}; \nu^{(p)}, \lambda^{(p)})$$

- **Etape M** :

remplacer Δx_i , par $\mathbb{E}(\Delta X_i \mid \underline{\Delta Z}; \nu^{(p)}, \lambda^{(p)})$
 Δy_i , par $\mathbb{E}(\Delta Y_i \mid \underline{\Delta Z}; \nu^{(p)}, \lambda^{(p)})$ dans :

- ▶ l'estimateur du maximum de vraisemblance
- ▶ l'estimateur de Bayes

obtenus en considérant l'observation $(\underline{\Delta X}, \underline{\Delta Y})$

Soit $\underline{\Delta z}$ une réalisation avec $z_0 = 3$ et $n = 3$

t_j	25	50	75
Z_j	4	1	2

Ensemble \mathcal{A} : configurations possibles des données complètes sachant les données incomplètes ci-dessus

i		1	2	3	
Date t_j		25	50	75	
Observation Δz_j		4	1	2	
Configuration	1	Δx_j	0	0	0
		Δy_j	4	1	2
	2	Δx_j	1	0	0
		Δy_j	2	1	2
	3	Δx_j	2	0	0
		Δy_j	0	1	2
	4	Δx_j	0	0	1
		Δy_j	4	1	0
	5	Δx_j	1	0	1
		Δy_j	2	1	0
	6	Δx_j	2	0	1
		Δy_j	0	1	0

Contexte
industriel

Contraintes
climatiques
variables

Modèle général
Contraintes de
température et
d'humidité
Contraintes de
surtensions

Système évolutif

Situation
Modèle
Observation
Inférence

Résultats

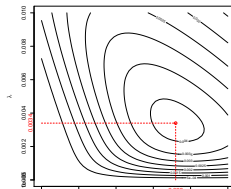
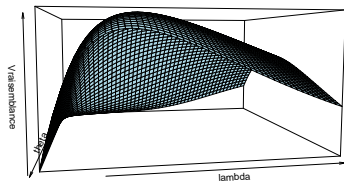
Conclusion

Résultats d'une maximisation classique de la vraisemblance

La fonction `nlm` de  donne en 8 iterations :

$$\hat{\theta} = 0.0361952 \quad \text{et} \quad \hat{\lambda} = 0.00334622$$

Fonction de vraisemblance pour les données observées incomplètes



Contexte
industriel

Contraintes
climatiques
variables

Modèle général
Contraintes de
température et
d'humidité
Contraintes de
surtensions

Système évolutif

Situation
Modèle
Observation
Inference

Résultats

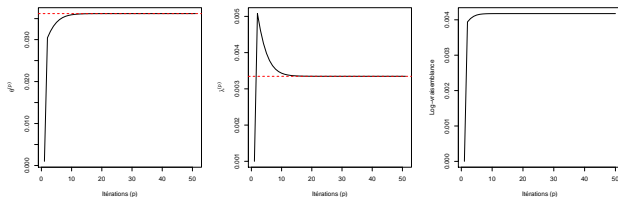
Conclusion

L'algorithme EM donne en 20 itérations :

$$\hat{\nu} = 0.0361945 \quad \text{and} \quad \hat{\lambda} = 0.00334674$$

Itérations de l'algorithme EM

— : estimation obtenue



Contexte
industriel

Contraintes
climatiques
variables

Modèle général
Contraintes de
température et
d'humidité
Contraintes de
surtensions

Système évolutif

Situation
Modèle
Observation
Inférence

Résultats

Conclusion

- ▶ Une explosion combinatoire peut apparaître sur l'ensemble \mathcal{A} , rendant le calcul de la densité conditionnelle des données complètes sachant les données incomplètes impossible
- ▶ Des versions stochastiques de l'algorithme EM telles que SEM (Celeux & Diebolt, 1986) ou MCEM (Wei & Tanner, 1990) ne permettraient pas de résoudre ce problème car elles requièrent des simulations à partir de la densité :

$$P(\underline{\Delta x}, \underline{\Delta y} \mid \underline{\Delta z}; \nu, \lambda) = \frac{f(\underline{\Delta x}, \underline{\Delta y}; \nu, \lambda)}{\sum_{(\underline{u}, \underline{v}) \in \mathcal{A}} f(\underline{\Delta u}, \underline{\Delta v}; \nu, \lambda)} \quad (1)$$

- ▶ Une procédure permettant de simuler suivant la loi de densité (1) sans avoir à calculer le dénominateur permettrait de résoudre le problème

- ▶ Il permet de simuler suivant une loi de densité f en ne connaissant qu'une fonction proportionnelle à f
- ▶ L'algorithme de Metropolis-Hasting construit une chaîne de Markov réversible dont la loi stationnaire est la loi cible
- ▶ On définit une loi de proposition, de noyau de transition Q
- ▶ Initialisation : choisir $u^{(0)}$
- ▶ A chaque itération, on calcule la valeur $u^{(k+1)}$ sachant $u^{(k)}$:
 1. Générer une proposition u^* à l'aide du noyau $Q(\cdot, u_k)$
 2. Calculer le ratio d'acceptation $\rho = \min \left\{ 1, \frac{P(u^*)Q(u_k, u^*)}{P(u_k)Q(u^*, u_k)} \right\}$
 3. Accepter ou rejeter la proposition suivant cette probabilité ρ . La valeur choisie est notée $u^{(k+1)}$

Initialisation : commencer avec $\nu^{(0)}$ et $\lambda^{(0)}$

A la $(p + 1)^{\text{ème}}$ itération, connaissant $\nu^{(p)}$ and $\lambda^{(p)}$:

1. Etape SE-Metropolis :

Simuler $(\underline{\Delta X}^*, \underline{\Delta Y}^*) \sim$

$P(\cdot, \cdot \mid \underline{\Delta Z}; \nu^{(p)}, \lambda^{(p)})$

(Metropolis-Hastings)

1. Etape MCE-Metropolis :

Simuler un M-échantillon de

loi $P(\cdot, \cdot \mid \underline{\Delta Z}; \nu^{(p)}, \lambda^{(p)})$:

$(\underline{\Delta X}^{(1)}, \underline{\Delta Y}^{(1)}, \dots, \underline{\Delta X}^{(M)}, \underline{\Delta Y}^{(M)})$

(Metropolis-Hastings)

et calculer la moyenne des

M échantillons obtenus

$$\underline{\Delta X}^* = \frac{1}{M} \sum_{\ell=1}^M \underline{\Delta X}^{(\ell)}$$

$$\underline{\Delta Y}^* = \frac{1}{M} \sum_{\ell=1}^M \underline{\Delta Y}^{(\ell)}$$

En déduire $\underline{\Delta Z}^* = 2\underline{\Delta X}^* + \underline{\Delta Y}^*$

2. Etape M

remplacer Δx_i , par $\mathbb{E}(\Delta X_i \mid \underline{\Delta Z}; \nu^{(p)}, \lambda^{(p)})$
 Δy_i , par $\mathbb{E}(\Delta Y_i \mid \underline{\Delta Z}; \nu^{(p)}, \lambda^{(p)})$ dans :

- ▶ l'estimateur du maximum de vraisemblance
- ▶ l'estimateur de Bayes

obtenus en considérant l'observation $(\underline{\Delta X}_n, \underline{\Delta Y}_n)$

Résultats pour l'algorithme SEM-Metropolis

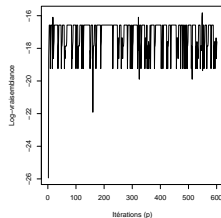
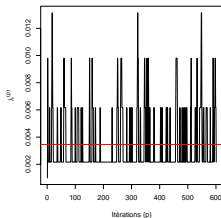
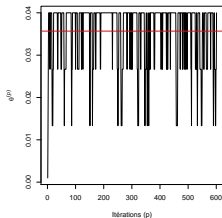
100 itérations SE+M pour la période de chauffe

La moyenne des 500 itérés des paramètres suivants donne :

$$\hat{\theta} = 0,03568 \quad \text{and} \quad \hat{\lambda} = 0,003459$$

Itérations de l'algorithme SEM-Metropolis

— estimation obtenue



Contexte
industriel

Contraintes
climatiques
variables

Modèle général
Contraintes de
température et
d'humidité
Contraintes de
surtensions

Système évolutif

Situation
Modèle
Observation
Inference
Résultats

Conclusion

Résultats pour l'algorithme MCEM-Metropolis

Groupe de travail
FIMA
23 mai 2013

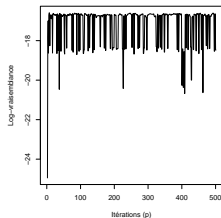
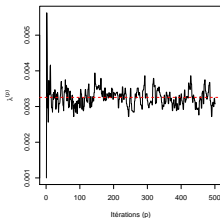
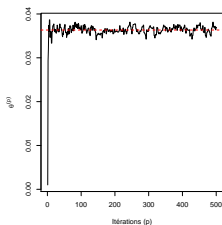
Lise Guérineau

100 itérations MCE+M pour la période de chauffe
La moyenne des 400 itérés des paramètres suivants donne :

$$\hat{\theta} = 0,03634 \text{ et } \hat{\lambda} = 0,003257$$

Itérations de l'algorithme MCEM-Metropolis

— estimation obtenue



Contexte
industriel

Contraintes
climatiques
variables

Modèle général
Contraintes de
température et
d'humidité
Contraintes de
surtensions

Système évolutif

Situation
Modèle
Observation
Inference

Résultats

Conclusion

Synthèse des résultats sur le cas d'école

Groupe de travail
FIMA
23 mai 2013

Lise Guérineau

Contexte
industriel

Contraintes
climatiques
variables

Modèle général
Contraintes de
température et
d'humidité
Contraintes de
surtensions

Système évolutif

Situation
Modèle
Observation

Inférence
Résultats

Conclusion

Algorithme	$\hat{\nu}$	$\hat{\lambda}$
NR ²	0,03620	0,003346
EM	0,03619	0,003347
SEM-Metropolis	0,03568	0,003459
MCEM-Metropolis	0,03634	0,003257

- ▶ Simulation d'observations discrètes sous l'hypothèse de nœuds inhibés :

Fixer \bar{z}_0 .

Pour $i = 1, \dots, n$:

- ▶ Tirer $\Delta x_i \sim \mathcal{P}(\nu L)$, $\Delta y_i \sim \mathcal{Bneg}(\bar{z}_{i-1}, e^{-\lambda L})$
- ▶ Calculer $\bar{z}_i \leftarrow \bar{z}_{i-1} + \eta \Delta x_i + \Delta y_i$.

- ▶ Dates d'observation équidistantes : $\Delta = 25$
- ▶ Nombre d'observations variable : $n = 6, 12, 24, 60$
- ▶ Paramètres ν, λ choisis de manière à couvrir différentes situations : flux de Poisson et flux de naissance plus ou moins importants

Résultats sur des données simulées

Estimation de la moyenne (EQM) des estimateurs du maximum de vraisemblance obtenus par l'algorithme MCEM-Metropolis à partir de 500 simulations ($\Delta = 25$, $z_0 = 3$, $\lambda^{(0)} = 0,01$, $\nu^{(0)} = 0,02$)

Paramètres du modèle	n	$\hat{\nu}$	$\hat{\lambda}$
$\nu = 0,08$ $\lambda = 0,0001$	6	0,0694 ($6.6 \cdot 10^{-4}$)	0,002408 ($1.3 \cdot 10^{-5}$)
	12	0,0723 ($3.0 \cdot 10^{-4}$)	0,000796 ($9.4 \cdot 10^{-7}$)
	24	0,0739 ($1.9 \cdot 10^{-4}$)	0,000310 ($1.3 \cdot 10^{-7}$)
	60	0,0752 ($1.1 \cdot 10^{-4}$)	0,000086 ($1.1 \cdot 10^{-8}$)
$\nu = 0,024$ $\lambda = 0,0038124$	6	0,02267 ($1.6 \cdot 10^{-4}$)	0,0030658 ($6.7 \cdot 10^{-6}$)
	12	0,02309 ($3.4 \cdot 10^{-3}$)	0,0019955 ($5.7 \cdot 10^{-6}$)
	24	0,02472 ($8.3 \cdot 10^{-5}$)	0,0016979 ($5.3 \cdot 10^{-6}$)
$\nu = 0,036$ $\lambda = 0,011508$	60	0,01563 ($4.3 \cdot 10^{-3}$)	0,0030895 ($9.8 \cdot 10^{-5}$)
	6	0,04472 ($7.2 \cdot 10^{-4}$)	0,006233 ($4.2 \cdot 10^{-5}$)
	12	0,03653 ($3.2 \cdot 10^{-4}$)	0,008024 ($1.7 \cdot 10^{-5}$)

Résultats sur des données simulées

Estimation de la moyenne (EQM) des estimateurs du maximum de vraisemblance obtenus par l'algorithme SEM-Metropolis à partir de 500 simulations ($\Delta = 25$, $z_0 = 3$, $\lambda^{(0)} = 0,01$, $\nu^{(0)} = 0,02$)

Paramètres du modèle	n	$\hat{\nu}$	$\hat{\lambda}$
$\nu = 0,08$ $\lambda = 0,0001$	6	0,0790 ($5.6.10^{-4}$)	0,000825 ($1.7.10^{-6}$)
	12	0,0789 ($2.5.10^{-4}$)	0,000162 ($3.0.10^{-8}$)
	24	0,0801 ($1.2.10^{-4}$)	0,000037 ($4.1.10^{-9}$)
	60	0,0802 ($4.5.10^{-5}$)	0,000007 ($8.7.10^{-9}$)
$\nu = 0,024$ $\lambda = 0,0038124$	6	0,02214 ($3.5.10^{-3}$)	0,0029671 ($1.3.10^{-5}$)
	12	0,02303 ($3.4.10^{-3}$)	0,0019657 ($5.5.10^{-6}$)
	24	0,02409 ($3.2.10^{-3}$)	0,0016653 ($3.3.10^{-6}$)
$\nu = 0,036$ $\lambda = 0,011508$	60	0,01626 ($4.1.10^{-3}$)	0,0030483 ($8.8.10^{-6}$)
	6	0,04642 ($1.9.10^{-3}$)	0,005915 ($4.8.10^{-5}$)
	12	0,03410 ($2.5.10^{-3}$)	0,008007 ($6.8.10^{-5}$)

- ▶ Développement de méthodes d'estimation de la fiabilité sous l'influence de contraintes dynamiques :
 - ▶ composants sous contraintes d'environnement
 - ▶ système sous l'influence de contraintes structurelles
- ▶ Démarche :
 - ▶ Validation des méthodes sur des données simulées
 - ▶ Application sur des cas d'intérêt industriel
- ▶ Pour EDF, les résultats obtenus peuvent :
 - ▶ au niveau **local**, contribuer à la priorisation des actions de maintenance pour des systèmes dont la fiabilité ne dépend pas que de l'âge
 - ▶ venir alimenter un modèle plus **global** dont l'objectif serait de paramétrer la fiabilité du réseau électrique dans son ensemble par différents états de l'environnement (Billinton & Allan, 1984)



P. Andersen, O. Borgan, R. Gill, and Keiding N.,
Statistical Models Based on Counting Processes,
Springer Series in Statistic, 1993.



P. Billingsley,
Statistical inference for Markov processes,
University of Chicago Press, 1961.



Billinton,
Reliability Evaluation of Power Systems,
Plenum Press, 1984.



G. Celeux and J. Diebolt,
"L'algorithmem : un algorithme d'apprentissage probabiliste pour la reconnaissance
de mélanges de densités",
Revue de statistique appliquée, vol. 34, pp. 35–52, 1986.



A. Dempster, N. Laird, and D. Rubin,
"Maximum likelihood from incomplete data with the em algorithm",
Journal of the Royal Statistical Society. Series B (Methodological), vol. 39, no. 1, pp.
1–38, 1977.



Y. Ogata,
"Statistical models for earthquake occurrences and residual analysis for point processes",
Journal of the American Statistical Association, vol. 83, no. 401, pp. 9–27, 1988.



T. Ozaki,
"Maximum likelihood estimation of hawkes self-exciting point processes",
Annals of the Institute of Statistical Mathematics, vol. 31, pp. 145–155, 1979.



G.C.G. Wei and M.A. Tanner,
"A monte carlo implementation of the em algorithm and the poor man's data
augmentation algorithms",
Journal of the American Statistical Association, vol. 85, no. 411, pp. 699–704, 1990.