

RÉACTUALISATION BAYÉSIENNE D'UN MODÈLE DE DÉGRADATION EN FONCTION DU RETOUR D'EXPÉRIENCE

BAYESIAN INFERENCE IN A DEGRADATION MODEL DEPENDING ON THE EXPERIENCE FEEDBACK

G. Bouchard, G. Celeux
INRIA
Dept. de mathématiques, bât. 425,
Université Paris-Sud
91405 Orsay Cedex
Guillaume.Bouchard@inria.fr
Gilles.Celeux@math.u-psud.fr

F Billy, F. Josse
EDF- R&D
6 quai Watier
78401 Chatou cedex
francois.billy@edf.fr
florent.josse@edf.fr

Résumé

EDF dispose de pièces métalliques sujettes à la fissuration par corrosion sous contrainte. Les études physiques ont montré que leur temps de fissuration dépasse largement leur durée d'utilisation. En cas de fissuration avant cette date, nous souhaitons savoir de quelle manière le modèle physique doit être remis en question. Un modèle bayésien incluant les connaissances sur les variables physiques influentes et les données du retour d'expérience est proposé. L'estimation des paramètres *a posteriori* se fait par échantillonnage de Gibbs. L'importance de la réactualisation en fonction du scénario de fissuration est quantifiée et le rôle de chacun des paramètres est identifié.

Summary

Some studies carried out by the french Électricité de France company on the failure process of metallic components showed that the expected date of failure largely exceeds their period of use. In the case of failure before this date, we wish to know how the physical model is contradicted. A bayesian model mixing knowledge on the physical variables and the experience feedback is proposed. The *a posteriori* parameters are estimated by Gibbs sampling. The importance of the reactualization according to the failure scenario is quantified and the role of each parameter is identified.

Dans un contexte de fiabilité, la dégradation d'éléments métalliques intervenant dans le fonctionnement de centrales nucléaires est modélisée. A partir du modèle physique de dégradation existant, un modèle probabiliste de fissuration a été défini dans un contexte bayésien, justifié par la présence d'information *a priori*. L'utilisation du formalisme des modèles graphiques a été d'un intérêt considérable dans la mise au point d'un modèle en accord avec les experts du domaine. L'échantillonnage de Gibbs a permis d'estimer la loi *a posteriori* du modèle. L'exploitation du modèle sur les données réelles répond aux objectifs d'anticipation des dégradations et de calcul du risque de défaillance.

Bien que spécifique du modèle étudié, la démarche utilisée peut être appliquée à de nombreuses applications industrielles. Nous détaillons les principaux aspects novateurs de ce travail.

Sources d'information hétérogènes Au sein d'un même modèle, les données issues d'expériences en laboratoire, de mesures sur site, de connaissances d'expert et d'observations de défaillances sont prises en compte de manière cohérente dans un même modèle. Grâce aux indépendances conditionnelles contenues dans le modèle graphique, il a été possible d'incorporer au cas par cas un modèle pour chaque type de donnée : température, contrainte, sensibilité du matériau, cinétique de propagation et retour d'expérience. De plus, le choix de distributions de proba-

bilité appropriées (par exemple les lois *a priori* conjuguées) permet de simplifier considérablement l'inférence.

Modélisation d'événements rares Une particularité du problème est que le nombre d'observations de défaillance est nul. Cela nécessite de définir des scénarios virtuels pour lesquels les observations sont en forte contradiction avec les connaissances *a priori*. Ce type d'approche permet d'anticiper des événements rares mais aux lourdes conséquences.

Quantification de l'effet des contrôles négatifs Être capable d'observer comment un risque évolue en fonction des contrôles effectués est d'un intérêt majeur dans les applications industrielles. Il permet de pouvoir apprécier de manière quantitative l'effet des contrôles négatifs (c'est à dire sans observation de défaillance). Dans tous les domaines où toute défaillance est à proscrire, les statistiques classiques demeurent impuissantes puisqu'on dispose de très peu d'observations. L'utilisation des statistiques bayésiennes nous a permis de montrer que le risque de défaillance est significativement plus grand si on ne prend pas en compte ces contrôles négatifs.

Echantillonnage de Gibbs et calcul de risque structurel Un aspect particulièrement novateur est l'utilisation d'un calcul de risque structurel (calcul FORM) au cours de la simulation. En effet, ce type de modèle était généralement utilisé pour calculer un risque de défaillance en fonction de variables aléatoires, mais l'utilisation du retour d'expérience pour réactualiser la valeur de ce risque n'avait jamais été étudié.

1 Présentation du problème

1.1 Contexte

EDF se doit d'anticiper d'éventuelles dégradations de ses composants, surtout lorsque des événements passés les rendent plausibles. L'alliage 600 (alliage Fe-Ni-Cr), utilisé pour la fabrication de pièces particulièrement sensibles des tranches nucléaires, s'est avéré sujet à la fissuration par corrosion sous contrainte. Tous les composants en alliage 600, y compris ceux sur lesquels les modèles ne prévoient pas de fissuration avant longtemps, doivent faire l'objet d'inspections destinées à vérifier l'absence de fissures et de travaux d'anticipation d'EDF, afin de pouvoir réagir en cas de dégradation constatée.

Le temps au bout duquel l'amorçage de fissures dans l'alliage 600 est susceptible de se produire est estimé au moyen d'une équation dans laquelle interviennent 3 indices : l'indice matériau, caractérisant la sensibilité du matériau à la corrosion, l'indice contrainte, dépendant de l'ensemble des contraintes subies par le composant (fabrication, assemblage, géométrie, fonctionnement), l'indice température, dépendant de la température de fonctionnement de la tranche.

L'estimation des trois indices est entachée d'une certaine imprécision, d'autant plus que leur connaissance est issue d'essais de laboratoire, ce qui fait du temps d'amorçage une variable aléatoire. Les résultats du modèle sont donc entachés de nombreuses incertitudes. Son application à un composant particulier (un tuyau cylindrique épais dans lequel circule un fluide), présent à environ 2800 exemplaires sur l'ensemble des 58 tranches françaises, conduit à des temps d'amorçage très élevés, compris entre 100 et 500 ans.

L'approche développée consiste à tenir compte des résultats des contrôles (présence ou absence de fissures) sur les composants en condition réelle d'exploitation afin de réactualiser les paramètres du modèle physique qui sont au nombre de trois : la température, la contrainte et l'indice matériau.

L'étude de l'influence des résultats de contrôle, désignés par la suite par Retour d'Expérience (REX), sur les paramètres du modèle permet non seulement de déterminer quels sont les composants critiques et avec quelle probabilité, mais aussi d'évaluer l'influence d'une stratégie d'inspection sur le risque global de détection d'une fissure.

Un modèle probabiliste a été utilisé afin de concilier : des événements rares et très improbables : par exemple la détection d'une fissure sur un composant âgé de 30 ans, et l'ensemble des informations « normales » : tous les composants contrôlés sains.

1.2 Méthode

Le modèle décrit permet d'estimer la loi de probabilité du temps d'amorçage de fissures. Il met en œuvre une approche de type bayésien afin de prendre en compte les connaissances a priori. On note que : le modèle est partiellement hiérarchique dans la mesure où certains paramètres définissant les lois a priori sont eux même aléatoires : leur valeur est attachée à celle d'autres paramètres. Ce type de construction hiérarchique permet d'atténuer l'influence de paramètres a priori dont les valeurs ne sont pas connues avec une grande précision. les distributions de probabilité de la plupart des variables du modèle sont définies conditionnellement à d'autres variables.

Les spécificités de l'approche portent sur les quatre points suivants :

1. les données ne sont pas homogènes : elles comprennent différents types de matériau, d'environnement et de contrainte,
2. le temps à fissuration est donné par un modèle physique : ce modèle n'est pas remis en cause. Il repose sur des critères physiques, et on a considéré que les sources d'incertitude provenaient uniquement des paramètres du modèle,
3. la présence d'une forte connaissance a priori sur les paramètres du modèle, en fonction d'essais de laboratoire, donnant des temps d'amorçage extrêmement longs,
4. les informations issues du REX : ce sont essentiellement des données de survie, donc censurées ; il y a très peu d'observations de fissure puisqu'on veut modéliser l'influence de leur éventuelle détection.

Afin de rendre la modélisation cohérente, il a fallu prendre en compte des liens entre paramètres de même nature. Par exemple, nous savons que la température de fonctionnement de certaines installations est plus élevée que d'autres. On ne peut admettre une réactualisation de la température qui fasse que cette propriété ne serait plus respectée : ces règles de cohérence ont été modélisées au moyen de variables de décalage communes à l'ensemble des paramètres.

L'estimation de la loi jointe de l'ensemble des variables aléatoires du modèle est réalisée par l'échantillonnage de Gibbs[2], qui consiste à simuler successivement les variables du modèle conditionnellement à toutes les autres. La distribution des variables ainsi obtenue suit, au bout d'un grand nombre d'itérations, la loi a posteriori recherchée.

L'utilisation de lois conjuguées (c'est-à-dire de lois qui conservent leur forme analytique après intégration des données aux connaissances a priori) permet de simplifier les simulations. Cependant la prise en compte de scénarios peu précis (par exemple 10 fissures sur n'importe lesquels des 2800 composants) requiert un calcul de risque structurel pour chaque composant. La méthode FORM (First Order Reliability Method) a été utilisée pour conduire ces calculs[3]. L'imbrication de ces deux techniques (calcul FORM et échantillonnage de Gibbs) constitue un intérêt technique majeur de ce travail.

2 Modèle de fissuration

2.1 Modèle physique

La fissuration des composants métalliques sous contrainte est un phénomène qui a fait l'objet de nombreuses études à EDF. Des études antérieures ont défini un modèle d'évaluation du temps de fissuration, c'est-à-dire le temps qui s'est écoulé entre la date de mise en service du composant et l'amorçage d'une fissure. Une fissuration est divisée en deux phases distinctes : l'*amorçage* et sa *propagation* jusqu'à une profondeur jugée importante relativement au risque de fuite.

Modèle des indices L'amorçage représente le temps d'apparition d'une fissure telle que sa profondeur soit à la limite des performances techniques des contrôles (1mm de profondeur par exemple). Il est calculé par une formule appelée *modèles des indices* utilisé dans plusieurs études antérieures [1]. Il peut se résumer à la formule donnant le temps de fissuration théorique t_{fiss} :

$$t_{fiss} = \frac{t_{ref}}{I_{mat} I_{\theta} I_{\sigma}} \quad (1)$$

où t_{ref} est le temps de fissuration de référence, I_{mat} un indice déterminant la sensibilité du matériau, I_{θ} un indice déterminé par température de fonctionnement et I_{σ} un indice dépendant de la contrainte que subit le composant. Le modèle des indices cumule de manière multiplicative l'influence de trois facteurs favorisant la fissuration.

t_{ref} est une constante de normalisation et on a choisi de ne pas la remettre en cause par le retour d'expérience. Dans cette étude, il est fixé à 10000 heures.

I_{mat} n'a pas de forme explicite. Suite à plusieurs études on admet que sa loi de probabilité est une loi lognormale dont les paramètres dépendent directement du matériau (trois matériaux possibles) et de la présence ou non de détensionnement (traitement thermique de la pièce diminuant les contraintes de soudure, de la lors de la construction de la tranche [1]). Il y a donc en tout six types de lois différentes en fonction du type de matériau. Nous travaillerons dans la suite avec la variable $\ell = \log(I_{mat})$ qui suit une loi normale.

I_{θ} est une fonction de la température de fonctionnement : $I_{\theta} = e^{-\left(\frac{1}{\theta} - \frac{1}{\theta_0}\right)}$ avec $\theta_0 = 598^{\circ}K$ la température de référence. Dans les conditions d'exploitation, la température peut varier sensiblement d'un composant à l'autre, et elle n'est pas directement mesurable. Ainsi, θ est considéré aléatoire et est modélisé par une loi normale.

I_{σ} dépend de la contrainte exercée sur le composant : $I_{\sigma} = \left(\frac{\sigma}{\sigma_{ref}}\right)^4$ avec $\sigma_{ref} = 450 MPa$ la contrainte de référence. Nous différencierons dans la suite σ et σ_{calc} la contrainte réelle de la contrainte estimée.

Propagation d'une fissure La propagation représente le temps que met la fissure pour atteindre une profondeur donnée. Dans cette étude, on a retenu la valeur de 6 mm. Remarquons que dans le cadre de cette étude, notre approche a été très conservatrice car il n'y a pas de perte fonctionnelle à cette profondeur. On justifie cependant une opération de maintenance à partir de cette profondeur, c'est-à-dire à la mi-épaisseur du composant.

On utilise la loi de cinétique suivante[1] :

$$\frac{da}{dt} = C. [q(a) - q_0]^m, \quad (2)$$

où

- a est la profondeur de fissure,
- q est le facteur d'intensité de contrainte en fond de fissure,
- q_0 est le seuil de propagation,
- C est un coefficient aléatoire,
- m caractérise la vitesse de la cinétique.

La valeur m est considérée comme une constante pour tous les composants étudiés.

En intégrant l'équation 2, on peut exprimer le temps écoulé entre l'apparition de la fissure (date d'amorçage t_{fiss}) et l'instant t_{def} où la profondeur de la fissure a_{def} est suffisamment grande pour justifier une opération de maintenance.

$$t_{def} - t_{fiss} = \frac{1}{C} \int_{a_{fiss}}^{a_{def}} \frac{da}{[K(a) - K_{ISCC}]^m} = \frac{\lambda}{C} \quad (3)$$

où a_{fiss} est la profondeur de fissure à son amorçage supposé constant. On considère que m est une constante. Le facteur d'intensité $K(a)$ dépend des caractéristiques de chaque composant. On dispose de valeurs de $K(a_i)$ pour des profondeurs a_i allant de 1 mm à 11 mm. L'intégrale λ est calculée par intégration numérique pour chaque composant.

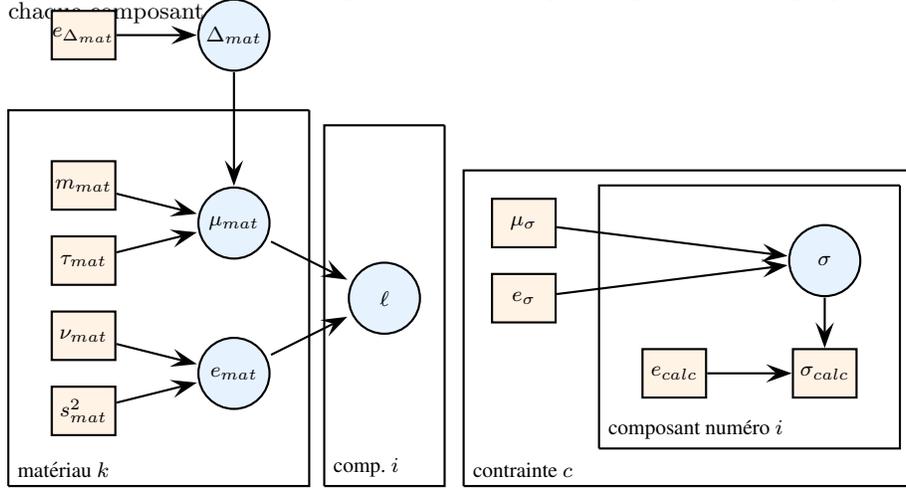


FIG. 1 – Exemples de modèles graphiques. Le modèle de droite représente la modélisation de la contrainte, celui de gauche la modélisation de la sensibilité du matériau. Des cadres délimitent les variables qui doivent être répétées plusieurs fois. Il faut par exemple remarquer que ℓ ne peut pas être à l'intérieur du cadre matériau car il y a N (et non $N \times K$) variables ℓ définies.

Le temps de fissuration est obtenu en sommant le temps d'amorçage et le temps de propagation. Nous donnons la formule finale utilisée dans notre modèle :

$$t_{def} = \frac{t_{ref}}{e^{\ell} e^{-(\frac{1}{\theta} - \frac{1}{\theta_0})} \left(\frac{\sigma}{\sigma_{ref}}\right)^4} + \frac{\lambda}{e^x}, \quad (4)$$

où les quantités t_{ref} , θ_0 , σ_{ref} et λ sont des constantes. Les paramètres aléatoires sont le logarithme de l'indice matériau ℓ , la température θ , la contrainte σ et le

facteur de cinétique χ . Ils sont spécifiques à chaque composant et on suppose qu'ils suivent une loi normale. Le choix d'une loi normale est justifiée par des études antérieures.

– sensibilité du matériau ℓ

$$\ell \sim \mathcal{N}(\mu_{mat}(k), e_{mat}^2(k)), \quad (5)$$

– température θ

$$\theta \sim \mathcal{N}(\mu_{\theta}(j), e_{\theta}^2(j)), \quad (6)$$

– contrainte σ

$$\sigma \sim \mathcal{N}(\mu_{\sigma}(c), e_{\sigma}^2(c)), \quad (7)$$

– coefficient de cinétique C :

$$\chi \sim \mathcal{N}(\mu_{\chi}(c), e_{\chi}(c)). \quad (8)$$

Les paramètres sont indicés par le type de matériau k , le palier j ou le type de contrainte c du composant considéré. En effet, les matériaux, même au sein d'une tranche nucléaire donnée, peuvent être différents, et obéissent à une loi de probabilité différente. De même, à chaque type de palier j correspond une température différente. Il n'y a que deux types de contrainte c : la paroi *interne* du tube et la paroi *externe*. À chaque paroi est associée une contrainte σ et une cinétique de fissuration χ . Ainsi, tout composant a deux dates de fissuration : une sur la paroi interne et une sur la paroi externe. Ces dates dépendent de cinq paramètres ℓ , θ , σ^{int} , σ^{ext} , χ^{int} , χ^{ext} . par la formule (4). Nous noterons N le nombre de composants susceptibles de se fissurer, K le nombre de matériaux différents (quatre dans notre cas), et J le nombre de paliers différents (quatre dans cette étude). Nous avons constaté que la paroi interne donne de plus fortes probabilités de fissuration. Les résultats donnés à la fin ne portent que sur celle-ci bien que les calculs soient toujours effectués sur les deux parois.

Estimation de σ Nous disposons d'une estimation σ_{calc} de la contrainte pour chaque composant. Cette valeur, censée approximer σ est entachée d'une erreur et σ_{calc} peut donc différer notablement de la vraie valeur, inconnue, σ . On considère donc que la contrainte calculée σ_{calc} suit une loi normale autour de la valeur réelle σ avec un écart-type e_{calc} donné et mis à 30MPa).

$$\sigma_{calc} \sim \mathcal{N}(\sigma, e_{calc}^2). \quad (9)$$

La modélisation de σ est représentée sur la gauche de la figure 1. Les valeurs de μ_{σ} et de e_{σ} doivent être fixées de manière à représenter la distribution des contraintes si on n'avait pas de méthode de calcul pour les estimer. La connaissance de ces contraintes est très faible, ce qui est traduit dans la modélisation par un écart-type e_{σ} assez élevé.

Modélisation bayésienne Les experts ont une connaissance relativement incertaine des paramètres μ et e de chacune des lois que nous avons introduites. Cette connaissance est modélisée par de nouvelles lois de probabilités : en utilisant les lois conjuguées des paramètres de la loi normale, les moyennes sont modélisées par des lois normales, et les variances par des lois inverse gamma :

$$\mu_{mat}(k) \sim \mathcal{N}(m_{mat}(k) + \Delta_{mat}, \tau_{mat}^2(k)), \quad (10)$$

$$e_{mat}^2(k) \sim \mathcal{IG}\left(\frac{\nu_{mat}(k)}{2}, \frac{s_{mat}^2(k)}{2}\right), \quad (11)$$

$$\mu_{\theta}(j) \sim \mathcal{N}(m_{\theta}(j) + \Delta_{\theta}, \tau_{\theta}^2(j)), \quad (12)$$

$$e_{\theta}^2(j) \sim \mathcal{IG}\left(\frac{\nu_{\theta}(j)}{2}, \frac{s_{\theta}^2(j)}{2}\right), \quad (13)$$

$$\sigma_{calc}(i) \sim \mathcal{N}(\sigma(i) - \Delta_{\sigma}, e_{calc}^2), \quad (14)$$

$$e_{calc}^2 \sim \mathcal{IG}\left(\frac{\nu_{calc}}{2}, \frac{s_{calc}^2}{2}\right), \quad (15)$$

$$\mu_{\chi} \sim \mathcal{N}(m_{\chi}, \tau_{\chi}^2), \quad (16)$$

$$e_{\chi}^2 \sim \mathcal{IG}\left(\frac{\nu_{\chi}}{2}, \frac{s_{\chi}^2}{2}\right). \quad (17)$$

Les hyperparamètres m , τ , ν et s sont fixés par l'utilisateur du modèle de manière à refléter ses connaissances sur modèle physique. Les variables Δ ont été introduites pour refléter les corrélations qui existent entre les paramètres de moyennes de différentes catégories. Nous expliquons ce point dans le paragraphe suivant.

Cohérence des paramètres Supposons dans un premier temps que les valeurs Δ soient nulles. Dans ce cas, si on observe des fissures uniquement sur un seul type de matériau, une seule catégorie de tranche ou un composant particulier, alors les paramètres réactualisés ne concernent que la catégorie du composant, et non l'ensemble des composants du parc. Ceci vient du fait que la modélisation suppose que chaque catégorie a des paramètres indépendants des autres catégories. On souhaiterait par exemple qu'une réactualisation de $\mu_{mat}(k)$ à une valeur supérieure engendre du même coup une augmentation de tous les $\mu_{mat}(k')$ avec $k' = 1, \dots, k - 1, k + 1, \dots, K$. Afin de prendre en compte le fait qu'entre deux matériaux différents les valeurs de μ_{mat} sont liées, nous introduisons une nouvelle variable commune à tous les types de matériau qui influent sur les différentes lois de μ_{mat} . Cette variable notée Δ_{mat} représente un *décalage* additif systématique de tous les paramètres $\mu_{mat}(k)$ pour $k = 1, \dots, K$:

$$\mu_{mat}(k) = m_{mat}(k) + \Delta_{mat} \quad (18)$$

Ce décalage hypothétique vaut évidemment 0 *a priori* mais peut prendre une toute autre valeur suivant les données du REX. il est clair que tous les indices matériau seront affectés par un tel décalage. Précisons que cette variable n'a pas de sens physique, et que son introduction dans le modèle n'a pour justification que la cohérence des paramètres *a posteriori* et la volonté de l'exploitant de vouloir propager ou non un événement de REX à l'ensemble des tranches. Exactement de la même manière, nous introduisons des décalages systématiques Δ_θ et Δ_σ dans la mesure de la température et dans la mesure de la contrainte. Pour obtenir des décalages positifs, on soustrait le décalage par rapport à σ .

Les trois nouvelles variables aléatoires Δ_{mat} , Δ_θ et Δ_σ , requièrent une loi de probabilité. En moyenne, le décalage vaut 0, c'est-à-dire qu'il n'y a aucun décalage au vu des connaissances *a priori* que l'on a. Le choix naturel est de prendre la loi normale de moyenne nulle et d'écart-type e_Δ qui reste à définir, mais qui est fixe (i.e. c'est un hyperparamètre) :

$$\Delta_{mat} \sim \mathcal{N}(0, e_{\Delta_{mat}}^2), \quad (19)$$

$$\Delta_\theta \sim \mathcal{N}(0, e_{\Delta_\theta}^2), \quad (20)$$

$$\Delta_\sigma \sim \mathcal{N}(0, e_{\Delta_\sigma}^2). \quad (21)$$

Les valeurs $e_{\Delta_{mat}}$, e_{Δ_θ} et e_{Δ_σ} sont fixées par un expert suivant sa connaissance sur la force des relations qui lient les indices matériau entre eux, les températures de tranches différentes ou les calculs de contraintes. Par, exemple, si l'expert pense que les indices matériau n'ont aucun rapport entre eux, il donnera une petite valeur à $e_{\Delta_{mat}}$. Inversement, s'il pense qu'une réactualisation de μ_{mat} pour un certain type de sensibilité doit entraîner systématiquement un décalage comparable pour les autres indices matériau, alors il prendra une valeur de $e_{\Delta_{mat}}$ relativement grande par rapport à s_{mat} . En revanche, aucune dépendance dans la cinétique de propagation χ n'a été modélisée pour des raisons physiques : la propagation interne et la propagation externe sont totalement différentes.

Pas d'incertitude sur le modèle physique Le modèle physique donne l'instant de fissuration d'un composant en fonction des valeurs exactes des trois paramètres ℓ , θ et σ . De la même manière, la valeur exacte de χ donne le temps exact écoulé entre l'amorçage et l'atteinte d'un seuil de profondeur critique. On considère donc que ces deux modèles (amorçage et propagation) sont systématiquement justes et que la seule source d'erreur provient de la mauvaise évaluation des paramètres. Cette hypothèse est évidemment invérifiable puisque les paramètres ℓ , θ , σ et χ ne sont jamais observés avec certitude. Dans une approche différente, on aurait pu modéliser le temps de fissuration comme une variable aléatoire autour de la valeur donnée par le modèle. Cela conduirait à un niveau d'incertitude supplémentaire dans le modèle. Ainsi, une fissure avait une probabilité non nulle de s'amorcer en un temps

<p>Répéter un grand nombre de fois :</p> <ul style="list-style-type: none"> - $\Delta_{mat} . \sim \mathcal{N}(m'_{\Delta_{mat}}, e'^2_{\Delta_{mat}})$ - $\Delta_{\theta} . \sim \mathcal{N}(m'_{\Delta_{\theta}}, e'^2_{\Delta_{\theta}})$ - Pour chaque type de contrainte c - $\Delta_{\sigma} . \sim \mathcal{N}(m'_{\Delta_{\sigma}}, e'^2_{\Delta_{\sigma}})$ - $e^2_{calc} . \sim \mathcal{IG}(\frac{\nu'_{calc}}{2}, \frac{s'^2_{mes}}{2})$ - $\mu_{\chi} . \sim \mathcal{N}(m'_{\chi}, e'^2_{\chi})$ - $e^2_{\chi} . \sim \mathcal{IG}(\frac{\nu'_{\chi}}{2}, \frac{s'^2_{\chi}}{2})$ - Pour chaque sensibilité k, - $\mu_{mat} . \sim \mathcal{N}(m'_{mat}, e'^2_{mat})$ - $e^2_{mat} . \sim \mathcal{IG}(\frac{\nu'_{mat}}{2}, \frac{s'^2_{mat}}{2})$ - Pour chaque type de tranche j, - $\mu_{\theta} . \sim \mathcal{N}(m'_{\theta}, e'^2_{\theta})$ - $e^2_{\theta} . \sim \mathcal{IG}(\frac{\nu'_{\theta}}{2}, \frac{s'^2_{\theta}}{2})$ - Pour chaque composant numéro i, - Répéter plusieurs fois et dans le désordre - $\ell . \sim \mathcal{N}_{t2}(\mu_{mat}, e^2_{mat}, (K^c_{\ell}, d(c))_{c=\{int,ext\}})$ - $\theta . \sim \mathcal{N}_{t2}(\mu_{\theta}, e^2_{\theta}, (K^c_{\theta}, d(c))_{c=\{int,ext\}})$ - Pour chaque type de contrainte c - $\sigma . \sim \mathcal{N}_t(\mu_{\sigma}, e'^2_{\sigma}, K_{\sigma}, d)$ - $\chi . \sim \mathcal{N}_t(\mu_{\chi}, e'^2_{\chi}, K_{\chi}, d(c))$ - Pour chaque type de contrainte c - $d . \sim \mathcal{M}_d(S, (p_i)_{i=1\dots N})$
--

TAB. 1 – Plan d'échantillonnage de Gibbs

très court. Cependant, l'observation d'une seule fissure pourrait alors provenir de l'alea du modèle physique, et la réactualisation des paramètres serait donc naturellement plus faible. C'est donc dans un souci de *conservatisme* que nous n'introduisons pas d'incertitude à ce niveau. En outre, il n'y a pas de raison objective, au sens physique, pour que le temps de fissuration soit aléatoire : ce sont les conditions de fabrication et d'exploitation du composant qui ne sont pas constantes, et qui sont responsables de l'alea sur le temps de fissuration.

2.2 Modèle graphique

Un grand nombre de variables entrent en jeu dans la modélisation de la fissuration. Pour représenter les dépendances entre ces variables, un modèle graphique est très utile. (figure 4). Les variables entourées par des carrés sont fixes, c'est-à-dire qu'elles sont soit observées, soit fixées par l'utilisateur. Les variables entourées d'un cercle sont aléatoires et nous cherchons leur loi conditionnellement aux variables fixes. Les cadres sont définis de manière à représenter la multiplication des variables. Leur en-tête permet en effet de définir le nombre de fois qu'il faut les dupliquer pour représenter toutes les variables.

Nous considérerons que $d(i)$ prend la valeur 1 pour les composants fissurés et 0 pour les composants non fissurés, de façon à avoir $\sum_i d_i = S$. À ce stade de l'étude, il s'agit d'estimer, dans une optique d'inférence bayésienne, les lois *a posteriori* des variables du modèle sachant des scénarios pour les lois *a priori* définies dans le modèle graphique de la figure 4. Ces lois seront obtenues par intégration de la loi jointe *a posteriori* du modèle. L'ensemble des variables simulées est :

$$X = (\Delta_{mat}, \Delta_{\theta}, \Delta_{\sigma}, \mu_{mat}, \mu_{\theta}, \mu_{\chi}, e_{mat}, e_{\theta}, e_{calc}, e_{\chi}, \ell, \theta, \sigma, \chi, d). \quad (22)$$

Il nous faut donc estimer la loi de X . Cela se fait, comme décrit ci-dessous, par l'échantillonnage de Gibbs [2]. Partant de là, les estimateurs bayésiens des variables sont obtenus comme sous-produit de l'échantillonnage de Gibbs par intégration de Monte-Carlo sur la loi *a posteriori* approximée.

Notons que σ_{calc} est fixe, car c'est une donnée observée. La loi de X est calculée *conditionnellement* à cette variable. Nous avons introduit sa modélisation précédemment uniquement pour pouvoir obtenir la loi conditionnelle de X . Intuitivement, on peut comprendre qu'à partir des valeurs de σ et σ_{calc} , on aura une idée de l'erreur de calcul e_{calc} , et ainsi de toutes les variables dont elle dépend.

Lois conditionnelles Soit v une variable aléatoire du modèle, alors

$$P(v|\cdot) \propto P(X) \quad (23)$$

où $P(v|\cdot)$ représente la probabilité conditionnelle à l'ensemble des variables de X sauf v . On exploite la structure du modèle graphique dirigé grâce à la formule suivante[2] :

$$P(v|V_{-v}) = P(v|par[v]) * \prod_{w \in enf[v]} P(w|par[w]). \quad (24)$$

où $par(x)$ désigne l'ensemble des parents d'une variable x , et $enf(x)$ l'ensemble de ses enfants. Cette formule découle de l'application en chaîne de la loi de Bayes $P(A|B) = \frac{P(A \cap B)}{P(B)}$. La convergence de l'échantillonnage de Gibbs est théoriquement assurée par l'ergodicité (irréductibilité et apériodicité) de la chaîne de Markov. Nous devons donc simuler un très grand nombre de fois des réalisations de la loi jointe de X , portant sur les variables inobservées. L'utilisation de lois conjuguées permet de calculer les lois conditionnelles analytiquement. Dans notre cas, le choix de lois *a priori* normales et gamma pour les moyennes et les variances permet d'obtenir *a posteriori* une loi normale dont il suffit de caractériser les paramètres. Par exemple, L'application de la formule (24) donne :

$$\mu_\theta | \cdot \sim \mathcal{N}(m'_\theta, e'^2_\theta) \quad (25)$$

avec

$$\frac{1}{e'^2_\theta} = \frac{1}{\tau_\theta^2} + \frac{|T_j|}{e_\theta^2} \quad (26)$$

$$m'_\theta = e'^2_\theta \left[\frac{m_\theta + \Delta_\theta}{\tau_\theta^2} + \sum_{i \in T_j} \frac{\theta(i)}{e_\theta^2} \right]. \quad (27)$$

où $|T_j|$ est le nombre de composants dans le palier de type j . Les lois conditionnelles des autres variables μ , e et Δ sont identifiées à des lois normales de la même manière, et ne sont pas reportées ici pour plus de clarté.

Les paramètres ℓ , θ , σ et χ sachant d suivent des lois normales tronquées. Nous étudions ici le cas de la loi conditionnelle de ℓ . Cette loi peut être calculée en considérant les différentes valeurs possibles de $d(int)$ et $d(ext)$. Supposons que $d(int) = 1$ et $d(ext) = 0$, c'est-à-dire que le composant est fissuré sur sa paroi interne mais pas sur sa paroi externe.

$$\begin{aligned} & P(\ell|d(int) = 1, d(ext) = 0, \dots) \\ & \propto P(\ell|\mu_{mat}, e_{mat})P(d(int) = 1|.)P(d(ext) = 0|.) \\ & \propto \exp\left(-\frac{1}{2} \frac{(\ell - \mu_{mat})^2}{e_{mat}^2}\right) \times \\ & I_{t_{théo}(\ell, \theta, \chi, \sigma(int)) > t_{max}} \times I_{t_{théo}(\ell, \theta, \chi, \sigma(ext)) < t_{max}} \\ & \propto \exp\left(-\frac{1}{2} \frac{(\ell - \mu_{mat})^2}{e_{mat}^2}\right) I_{K_l^{int} < \ell < K_l^{ext}} \end{aligned} \quad (28)$$

avec pour $c = \{int, ext\}$,

$$K_\ell^c = \log\left(\frac{\frac{t_{ref}}{e^{\frac{\lambda}{\sigma_{ref}}}}}{e^{-E_a(\frac{1}{\theta} - \frac{1}{\theta_0})} (\frac{\sigma(c)}{\sigma_{ref}})^4}}\right). \quad (29)$$

La densité de ℓ correspond exactement à la densité d'une loi normale tronquée en K_l^- à gauche et en K_l^+ à droite. Pour les autres cas de défaillance, le raisonnement est le même. La simulation de ces lois tronquées se fait un algorithme d'acceptation-rejet adapté[4].

Probabilités de fissuration et calcul FORM La variable d a cependant un statut particulier. En effet la valeur que prend d sachant ℓ , θ , σ et χ est déterministe, et on ne peut donc pas simuler suivant cette loi conditionnelle. Il faut considérer un niveau hiérarchique supplémentaire, et étudier la loi (notée \mathcal{M}_d) de d sachant S , μ_{mat} , μ_θ , μ'_σ , μ_χ , e_{mat} , e_θ , e'_σ et e_χ . Pour éviter d'alourdir les notations, nous désignons ces variables par “.” (cf. par exemple le tableau 1). μ'_σ et e'_σ sont l'espérance et l'écart-type de σ sachant e_{calc} , μ_σ et e_σ . Nous appellerons cette loi. La distribution de \mathcal{M}_d s'écrit de la manière suivante :

$$P(d|S, \cdot) \propto P(S|d)P(d|.) \quad (30)$$

par double application de la règle de Bayes, et par indépendance de S et des autres variables conditionnellement à d . Nous avons $P(S|d) = I_{(\sum_i d_i=S)}$ et $P(d|.) = \prod_{i=1}^N P(d_i)$. Les probabilités $p_i = P(d_i = 1|.)$ valent :

$$p_i = \int \int \int \int_{t_{théo}(\ell, \theta, \sigma, \chi) > t_{max}} dP_\ell dP_\theta dP_\sigma dP_\chi. \quad (31)$$

Les probabilités P_ℓ , P_θ , P_σ et P_χ sont les lois des variables ℓ , θ , σ et χ conditionnellement à leurs variables parentes μ et e . Le calcul analytique de cette intégrale est extrêmement difficile, voire impossible, mais c'est un problème d'intégration que l'on rencontre couramment en fiabilité. Puisque ce sont des lois normales, on peut avoir recours à des méthodes d'approximation appelées First and Second Order Reliability Methods (FORM et SORM)[3]. Ces méthodes approchent la frontière d'intégration $t_{théo}(\ell, \theta, \sigma, \chi) = t_{max}$ par une droite ou une fonction quadratique. Nous avons utilisé la méthode FORM effectuant l'approximation linéaire.

Plan d'échantillonnage Nous avons donc obtenu les lois de chacune des variables conditionnellement aux autres. Le plan d'échantillonnage complet est donné dans le tableau 1. \mathcal{N}_t et \mathcal{N}_{t2} désignent les lois normales avec une et deux troncatures. Les simulations des variables ℓ , θ , σ , χ dépendent fortement de l'ordre dans lequel elles sont simulées. Ainsi, ces variables sont simulées à chaque fois dans une ordre différent, et ceci plusieurs fois par itération (quatre fois dans nos expériences) afin d'augmenter la vitesse de convergence de la chaîne de Markov vers sa loi stationnaire.

3 Résultats

Un programme en C++ avec une interface Matlab a été réalisé. Il permet d'analyser la sensibilité des paramètres *a priori* et de détecter les éventuelles situations critiques susceptibles de favoriser l'apparition d'une fissure.

3.1 Comportement des simulations

Comme dans toutes les techniques basées sur les simulations par chaîne de Markov, les valeurs successives des variables sont corrélées, et le nombre d'échantillons

	1 fissure	1 fiss.+cn	10 fissures	10 fiss.+cn
matériau 1	3.15	2.83	4.35	3.91
matériau 2	3.19	2.82	4.42	3.38
matériau 3	3.12	2.74	4.48	4.40
matériau 4	3.10	2.82	4.32	3.81
température 1	0.160	0.110	0.332	0.272
température 2	0.551	0.112	1.31	-0.251
température 3	0.571	0.431	0.596	0.543
température 4	0.173	0.0928	0.3	0.291
contrainte	0.631	0.601	1.59	1.18
cinétique	1.10	0.962	2.45	2.13

TAB. 2 – Quantification de la réactualisation des paramètres en fonction du scénario. Chaque valeur représente le décalage normalisé par l'écart-type *a priori* entre les moyennes *a priori* et *a posteriori*.

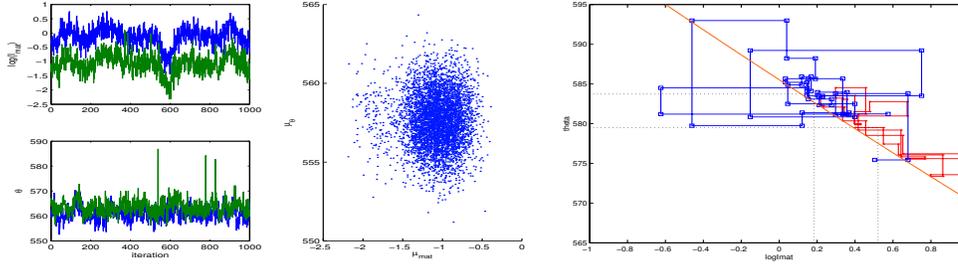


FIG. 2 – Illustrations de l'échantillonnage de Gibbs. A droite sont représentées 1000 valeurs de ℓ et θ simulées pour le composant le plus critique (en bleu, courbe supérieure) et un composant moins critique (vert). θ en fonction de ℓ est donné sur la figure du milieu, illustrant le fait qu'*a posteriori*, ces deux quantités sont proches de l'indépendance. La figure de droite montre la frontière de fissuration et quelques étapes de simulation pour le modèle simplifié aux paramètres ℓ et θ , les autres paramètres étant fixes. Les simulations correspondent à un scénario de 10 fissurations parmi 14 composants (bleu) et 14 fissurations parmi 14 (rouge). Ce dernier scénario n'ayant que des fissures, les valeurs sont simulées au dessus de la limite de fissuration.

issus de l'échantillonnage de Gibbs doit être suffisamment important. Ceci est illustré par les graphes de gauche et de droite de la figure 2. Nous avons empiriquement utilisé 50000 itérations, où plus exactement 5000 échantillons espacés de 10 itérations chacun. Au préalable, 500 itérations sont effectuées sans échantillonnage pour éliminer la dépendance à l'initialisation. Pour chaque variable simulée, la moyenne des valeurs échantillonnées est choisie comme estimateur *a posteriori*. Pour toutes les applications considérées, la corrélation entre les variables μ_{mat} , μ_{θ} , $\Delta\sigma$ et $\Delta\chi$ est calculée et est inférieure à 0.02. Nous considérons donc qu'elles sont *a posteriori* proche de l'indépendance. La figure 2, au milieu, illustre cette absence de corrélation entre μ_{mat} et μ_{θ} . L'utilisation des densités marginales pour prendre des décisions d'inspection est donc possible, sans risque de non conservatisme du risque calculé.

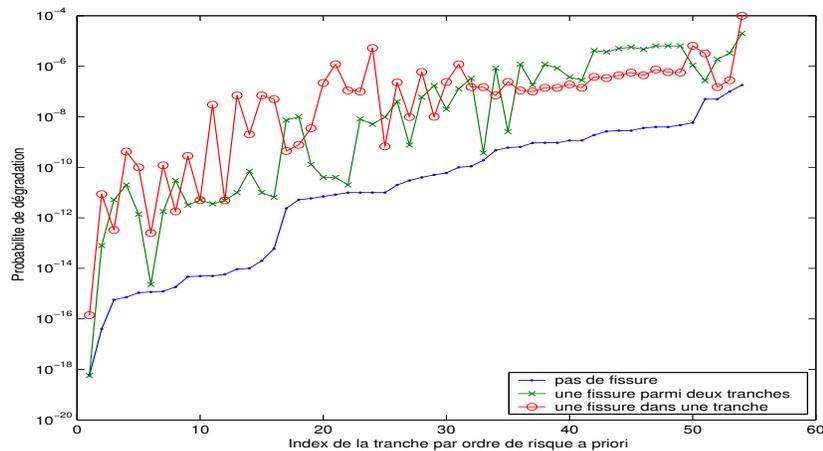


FIG. 3 – Risque de dégradation pour 54 tranches comportant chacune une cinquantaine de composants. Elles sont triées sur la première courbe (.) par risque *a priori*. La courbe verte(x) représente le scénario d'une fissure sur 2 tranches contrôlées et la courbe rouge(o) le scénario d'une fissure sur 1 tranche contrôlée.

3.2 Exploitation des résultats

Actuellement, aucune fissure n'a jamais été détectée lors des contrôles effectués. Les résultats présentés ici ne sont donc basés que sur des scénarios hypothétiques de

fissuration, que nous pouvons modifier afin de bien appréhender le comportement du modèle.

Les hyperparamètres du modèle correspondent dans la plupart des cas à des moyennes et variances de grandeurs physiques connues (comme la température), ou de quantité étudiées (comme l'indice matériau). Ils ont donc été fixés par des experts à des valeurs plausibles. Certains paramètres comme l'écart-type du décalage commun Δ_{mat} , Δ_θ ou Δ_σ sont plus difficiles à interpréter. La valeur utilisée a donc été fixée après plusieurs essais du modèle, en fonction de la cohérence des valeurs *a posteriori*. Plusieurs essais ont montré qu'un décalage commun cohérent est obtenu en fixant $e_{\Delta_{mat}} = \frac{1}{2}\tau_{mat}$, $e_{\Delta_\theta} = \frac{1}{2}\tau_\theta$ et $e_{\Delta_\sigma} = \frac{1}{2}s_{calc}$. Afin d'illustrer l'intérêt de ce paramètre de cohérence, dans la première expérience, nous relâchons la contrainte sur les températures en fixant $e_{\Delta_\theta} = \frac{1}{20}\tau_\theta$.

Mesure de réactualisation des paramètres Si les connaissances *a priori* sont exactes, les simulations ne donnent aucune fissuration, confirmant le fait que les connaissances actuelles des experts donnent une très faible probabilité à l'amorçage des fissures. Elle est de l'ordre de 10^{-7} pour le composant le plus critique au bout de 30 ans de fonctionnement.

Les paramètres étant inchangés, différents scénarios de fissuration sont introduits. Nous avons choisis 108 composants contrôlés, sur lesquels nous observons soit une unique fissure, soit 10 fissures. Pour chaque cas, un contrôle éventuel donnant 0 fissure sur 689 autres composants est ajouté. Ces 689 composants sont tous de la catégorie de température 2. Pour les quatre scénarios ainsi obtenus, les lois *a posteriori* ont été estimées. La date des contrôles est fixée arbitrairement au 15 juin 2004.

Pour plus de clarté dans la présentation des résultats, une mesure de *décalage* entre l'*a priori* et *a posteriori* a été introduite : $dec = \frac{\mu_1 - \mu_0}{\sigma_0}$ où μ_0 et σ_0 sont les moyennes et les écarts-types *a priori* et μ_1 l'estimation de la moyenne *a posteriori*. Ces décalages reportés dans le tableau 2 quantifient les effets de la réactualisation pour différents scénarios avec les mêmes paramètres.

Le tableau 2 est assez riche en informations :

1. On peut constater qu'une seule fissure entraîne une très forte réactualisation des paramètres (de l'ordre de trois écarts-types). Ceci est dû au fait que l'observation d'une fissure est contradictoire avec les connaissances *a priori*.
2. Cette forte réactualisation se trouve diminuée par l'ajout des contrôles de non-fissuration. En effet, ces contrôles sont cohérents avec la loi *a priori* du modèle. Pouvoir quantifier de cette manière l'effet des contrôles négatifs est très important d'un point de vue applicatif.
3. Il y a une très forte différence de réactualisation entre les variables associées au matériau et les variables de température. En effet, la variable Δ_{mat} , en prenant de fortes valeurs permet d'obtenir *a posteriori* les plus grandes probabilités de fissuration. Ainsi, cette différence est essentiellement due au fait que e_{Δ_θ} est notablement différente de $e_{\Delta_{mat}}$. En intervenant sur ce paramètre *a priori*, on peut contrecarrer cette différence. Précisons qu'en plus des effets dus à $e_{\Delta_{mat}}$ et e_{Δ_θ} , les autres hyperparamètres (μ , τ , ν , et s), l'équation du modèle physique et les observations sont trois origines possibles des différences de réactualisation.
4. La valeur négative du décalage de la température numéro 2 n'est pas un aléa d'estimation, mais un comportement prévisible du modèle : tous les contrôles négatifs ayant été effectués pour ce type de température, celle-ci lui donner une petite valeur est cohérent avec les observations, et dans le cas d'un scénario catastrophique (ici 10 fissures), la réactualisation des autres paramètres est si importante que l'absence de fissure ne peut être compensée que par une température plus faible qu'*a priori*.

Les résultats que nous obtenons sont purement virtuels puisqu'aucune fissuration n'a jamais été constatée, mais de tels scénarios permettent à un expert de fixer dynamiquement les paramètres *a priori*. Cette tâche, cruciale en analyse bayésienne [5] est en général délicate. Elle est ici réalisée en analysant la sensibilité du modèle face aux changements des différents paramètres. L'expert peut ainsi mesurer leur influence sur la prise de décision, et déterminer par tâtonnements les valeurs qui correspondent à ses conjectures afin de définir un scénario de référence.

Probabilités de fissuration des tranches On a décidé ici d'étudier une réactualisation dans son ensemble, et donc de regrouper la totalité des composants d'une tranche. En effet, lorsqu'on effectue une inspection, l'ensemble des composants d'une tranche sont inspectés simultanément. C'est donc au niveau des tranches que se définit la stratégie d'inspection. Une tranche est dégradée si un de ses composants est fissuré. La probabilité de dégradation de chaque tranche a été calculée en utilisant les paramètres *a priori*, et les paramètres *a posteriori* pour le scénario précédent (détection d'une fissure parmi 108 composants, soit deux tranches contrôlées) et pour un scénario similaire portant sur la tranche pour laquelle le risque est plus faible : détection d'une fissure parmi 58 composants.

Les risques de dégradation *a priori* et *a posteriori* pour l'année 2010 ont été calculés pour 54 tranches, et représentés par ordre de risque de dégradation *a priori* sur la figure 3. On peut visualiser de quelle manière ce classement des tranches par risque de fissuration se trouve modifié par les deux scénarios.

On observe que pour toutes les tranches, le risque réactualisé est au moins égal au risque *a priori* : la détection des fissures ne peut que dégrader les estimations actuelles du modèle physique. On remarque aussi des fluctuations importantes pour certaines tranches, en raison de leur population spécifique (palier, sensibilité du matériau) qui les expose ou non à la réactualisation due à l'événement modélisé. Les valeurs de risque obtenues montrent que la réactualisation est importante : pour chaque composant, le risque de fissuration en 2010 augmente de plusieurs décades, en général 3, c'est-à-dire qu'il est multiplié par 10^3 . Le modèle remplit donc son rôle qui est d'indiquer un risque nettement plus élevé si un événement de REX improbable survient.

À partir de scénarios de fissuration, il est possible avec notre modèle de décider des installations à contrôler en priorité : ce sont celles qui fournissent les plus fortes probabilités de fissuration *a posteriori*.

4 Discussion

- Ce travail nous a permis d'incorporer à la fois
- des connaissances *a priori* sur un modèle physique,
 - des données issues du REX.

Lorsque les connaissances *a priori* sont trop faibles pour garantir un niveau de risque donné, les données conformes au modèle (c'est à dire les contrôles négatifs) issues du REX peuvent s'avérer bénéfiques pour minimiser les risques.

L'étude effectuée modélise un très grand nombre de variables. La visualisation des dépendances par un modèle graphique a permis d'isoler chaque variable d'intérêt de construire le modèle de densité jointe par étapes successives. L'analyse bayésienne est bien adaptée à la résolution du problème de réactualisation des paramètres en fonction du REX.

Le modèle peut cependant être trop simpliste, car il suppose des indépendances entre variables qui sont probablement loin de la réalité. De plus, on considère que le modèle de fissuration est parfaitement juste, ce qui n'est évidemment pas vérifié dans la réalité, mais modéliser un temps de fissuration aléatoire est par construction moins conservatif que notre approche.

La connaissance *a priori* est réactualisée par des mesures sur site, et l'étude de la distribution *a posteriori* permet de réévaluer les temps d'amorçage de chaque composant. Les résultats présentés montrent qu'en conjecturant la détection d'une seule fissure sur les 2800 composants, la réactualisation des paramètres du modèle des indices est significative : la rareté d'une information pénalisante est bien prise en compte et traduite quantitativement par le modèle statistique.

Références

- [1] EDF. Zones inconel du circuit primaire principal., 1999. rapport interne.
- [2] W. R. Gilks, S. Richardson, and D. J. Spiegelhalter. *Markov Chain Monte Carlo in Practice*. Chapman and Hall, London, 1996 (ISBN : 0-412-05551-1). This book thoroughly summarizes the uses of MCMC in Bayesian analysis. It is a core book for Bayesian studies.

- [3] Robert E. MELCHERS. *Integration and Simulation Methods*, chapter Second-Moment and Transformation Methods, pages 64–93. Wiley, 2001.
- [4] Christian P. ROBERT. Simulation of truncated normal variables. *Statistics and computing*, 5 :121–125, 1995.
- [5] Christian P. ROBERT. *The Bayesian Choice : from Decision-Theoretic Motivations to Computational Implementation (2001)*. Springer-Verlag, NY, 2001.

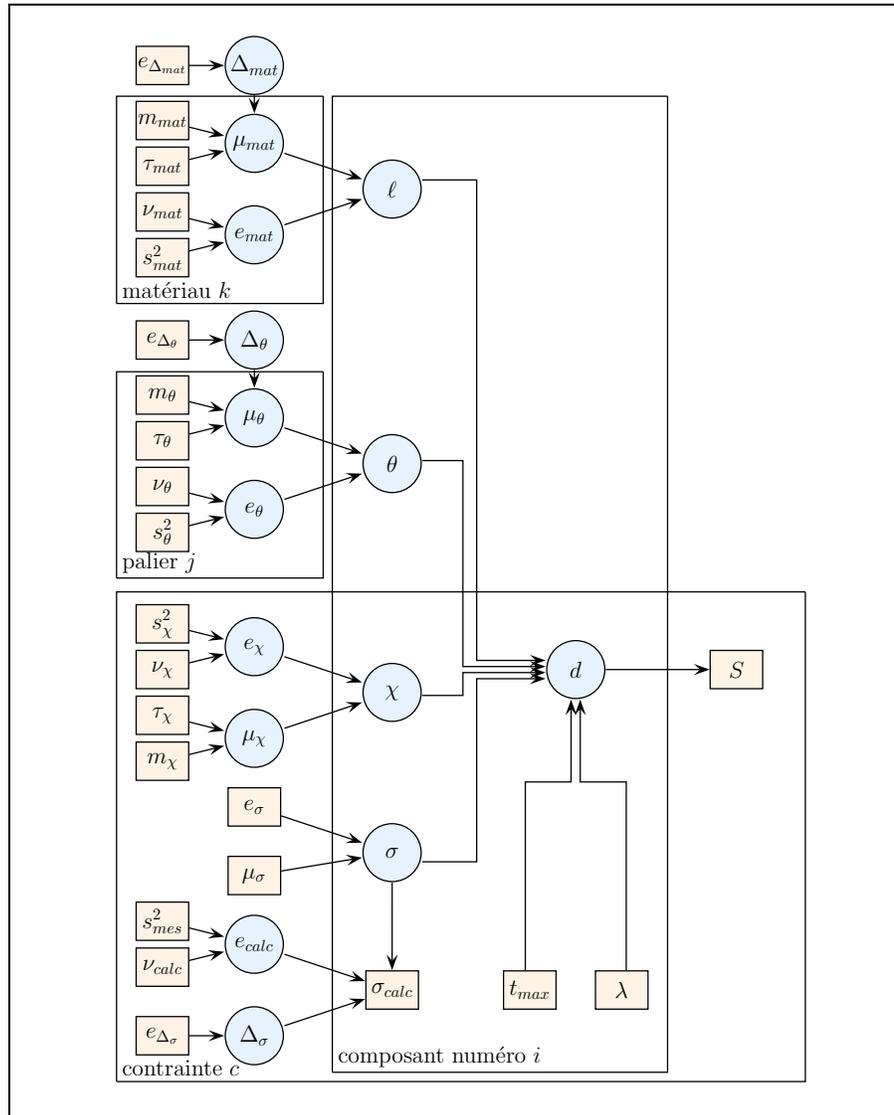


FIG. 4 – Modèle graphique complet. Le ronds représentent toutes les variables simulées par échantillonnage de Gibbs.