

Dispersion de la résistance en fatigue : analyse comparée de différentes approches

C. MATTRAND^a, G. CAUSSE^b, T. YALAMAS^b

a. Université Clermont Auvergne, CNRS, SIGMA Clermont, Institut Pascal, F-63000 Clermont-Ferrand, France

cecile.mattrand@sigma-clermont.fr

b. PHIMECA ENGINEERING, 34 rue de Sarliève, F-63800 Cournon d'Auvergne, France

causse@phimeca.com, yalamas@phimeca.com

Résumé :

Le travail présenté dans ce papier s'intéresse à la détermination des courbes de conception en fatigue pour des matériaux qui peuvent être modélisés par une tendance linéaire (en échelle logarithmique) dans le domaine de l'endurance limitée. Il s'agit plus particulièrement de positionner une approche probabiliste bayésienne par rapport à l'approche statistique / fréquentiste utilisée dans les normes et les recommandations. L'approche probabiliste bayésienne montre qu'il est possible de réduire le degré de conservatisme des approches traditionnelles dans le cas de petits échantillons et/ou si une information a priori sur la dispersion des données peut être apportée. Les courbes issues des approches classiques peuvent devenir très pénalisantes (voire impossibles à appliquer) dans ces cas-là.

Mots clefs : fatigue, courbes de conception, approche statistique, approche bayésienne

1 Introduction

Les mécanismes de fatigue à l'origine de la rupture des matériaux métalliques sont variés et multi-échelles. Une fissure de fatigue peut s'amorcer par micro-plasticité cyclique lorsque le matériau présente peu de défauts internes par exemple ou au contraire à partir d'un défaut pré-existant. C'est le cas des matériaux métalliques avec des discontinuités métallurgiques (carbures, inclusions ou porosités par exemple) ou avec des défauts liés au processus de fabrication (rayures, entailles, etc.). Les différents mécanismes d'amorçage possibles sont souvent en compétition et peuvent coexister. Par ailleurs le processus de fatigue s'initie à l'échelle de la microstructure (échelle du grain), par nature hétérogène et aléatoire, et s'étend jusqu'à l'échelle de la structure où les effets de la géométrie et du chargement appliqué sont à considérer. Ceci rend difficile la prévision du comportement en fatigue. Même sous des conditions expérimentales maîtrisées et supposées identiques, les résultats d'essais en fatigue menés sur des éprouvettes de forme simple et constituées d'un même matériau montrent des dispersions en termes de durée de vie ou de tenue résiduelle. Il est ainsi aujourd'hui communément admis que le phénomène de fatigue est par essence incertain.

En milieu industriel, le dimensionnement en fatigue des composants ou des structures métalliques s'appuie en général sur l'utilisation de courbes de fatigue, dites aussi courbes $S - N$ ou courbes de Wöhler. Pour prendre en compte les variabilités inhérentes au phénomène de fatigue mais également celles liées au nombre limité d'essais de fatigue, des courbes de conception peuvent être établies pour assurer un certain niveau de fiabilité lors du dimensionnement. Ces courbes de conception intègrent ainsi le niveau de probabilité de survie visé, noté P , qui capture les dispersions « physiques » du phénomène de fatigue mais également un niveau de confiance, noté γ , qui prend en compte les incertitudes « statistiques » lors de l'estimation des paramètres du modèle de fatigue (Basquin, Wöhler, etc.). Elles sont notées $S - N - P_\gamma$. Des approches statistiques / fréquentistes peuvent être utilisées pour estimer analytiquement le fractile des courbes $S - N$ (c'est-à-dire à probabilité de survie P donnée) ainsi que la borne inférieure de son intervalle de tolérance (unilatéral ou bilatéral) $S - N - P_\gamma$ qui intègre les incertitudes sur la calibration du modèle. Ces approches sont souvent employées dans les normes et recommandations et sont donc celles qui sont en pratique utilisées dans le milieu industriel même si elles peuvent être très conservatrices (le degré de conservatisme et son impact lors du dimensionnement ne sont pas évalués). Une autre façon d'obtenir les courbes $S - N - P_\gamma$ est de recourir aux approches probabilistes qui s'appuient souvent sur la simulation. Le lecteur pourra par exemple consulter le tableau 4.2 de la référence [1] qui présente une brève revue des méthodes permettant d'obtenir des intervalles de tolérance d'une loi log-normale, loi généralement utilisée pour représenter la dispersion soit sur le niveau de contrainte à nombre de cycles fixé, soit sur la durée de vie à contrainte imposée.

Dans ce papier, nous proposons d'analyser les courbes $S - N - P$ et $S - N - P_\gamma$ issues de normes et recommandations [2-6] et celles pouvant être obtenues par approche probabiliste bayésienne. Nous restreignons l'étude aux matériaux métalliques qui peuvent raisonnablement être approchés par un modèle de Basquin ou de Wöhler (sur une plage de contraintes donnée) et ne considérons pas la zone à endurance illimitée de la courbe $S - N$, c'est-à-dire les essais censurés (données non-rompues) ne sont pas pris en compte. L'influence du nombre d'essais disponibles, de la stratégie d'échantillonnage (nombre de niveaux de contraintes / nombre d'essais) ainsi que l'influence de la connaissance *a priori* sur la variance seront analysées à partir de données de fatigue générées artificiellement par un processus dont les propriétés sont connues. L'idée est alors de positionner les deux approches afin d'illustrer le potentiel intérêt d'une approche bayésienne pour construire les courbes de conception en fatigue. Ce travail s'inscrit dans le cadre du groupe de travail « Prise en compte des incertitudes dans le dimensionnement en fatigue » en lien avec la commission Fatigue de la SF2M et le GST Mécanique et Incertain de l'AFM.

2 Approche statistique / fréquentiste

L'approche statistique / fréquentiste est utilisée dans les normes [2-4], cadre de ce travail. L'incertitude est supposée portée par la durée de vie, c'est-à-dire le nombre de cycles N , sur la plage des contraintes où les mesures sont réalisées. Elle est modélisée par une loi log-normale : $\log N \sim \mathcal{N}(\mu_N(S), \sigma_N(S))$ avec $\mu_N(S)$ la moyenne (modèle de tendance) qui dépend du niveau de contrainte S , et, $\sigma_N(S)$ l'écart-type de la variable aléatoire $\log N$ qui, en toute rigueur, peut varier avec S . Dans le contexte des normes et recommandations [2-6] la moyenne et l'écart-type sont cependant définis tels que :

$$\mu_N(S) = A + B \cdot \log S \text{ (modèle de Basquin) ou } \mu_N(S) = A + B \cdot S \text{ (modèle de Wöhler)} \quad (1)$$

Et

$$\sigma_N(S) = \sigma \quad (2)$$

L'écart-type est supposé constant, c'est-à-dire qu'une éprouvette, si elle pouvait être testée à différents niveaux de contraintes S , aurait un comportement en fatigue globalement bon ou au mauvais [7], et que la dispersion est la même quel que soit le niveau de contrainte. Cette propriété est appelée homoscedasticité. Le problème peut alors être réécrit comme un problème de régression linéaire :

$$\log N = M(A, B, \log S) + \varepsilon \quad (3)$$

Où $\varepsilon \sim \mathcal{N}(0, \sigma)$ est l'écart résiduel (encore appelé résidu) entre la valeur d'un essai et le modèle de tendance $M(A, B, \log S)$. Les résidus sont par ailleurs supposés indépendants entre eux. $M(A, B, \log S)$ constitue une partie déterministe et correspond dans la suite au modèle de Basquin.

Dans le contexte statistique, les paramètres A , B et l'écart-type σ sont estimés par la méthode des moindres carrés ordinaires (\equiv cas particulier du maximum de vraisemblance, les résidus étant ici gaussiens et indépendants). On notera \hat{A} , \hat{B} et $\hat{\sigma}$ les estimations obtenues à partir de n essais. D'après [2], il semble préférable de privilégier le nombre d'essais par niveau de contraintes plutôt que le nombre de niveaux à iso nombre d'essais n , afin de mieux estimer la dispersion (c'est-à-dire σ).

Les courbes $S - N - P$, c'est-à-dire sans prise en compte des incertitudes statistiques liées aux estimations, sont alors obtenues comme :

$$\widehat{\log N}_p = \hat{A} + \hat{B} \cdot \log S - u_p \cdot \hat{\sigma} \quad (4)$$

Où u_p est le fractile de la loi normale centrée réduite pour la probabilité P .

Les courbes de conception $S - N - P_\gamma$ intègrent les incertitudes liées aux estimations \hat{A} , \hat{B} et $\hat{\sigma}$ dues au nombre limité de données n . Elles s'écrivent :

$$\widehat{\log N}_{p,\gamma} = \hat{A} + \hat{B} \cdot \log S - \mathbf{K} \cdot \hat{\sigma} \quad (5)$$

Où \mathbf{K} est un coefficient issu des intervalles de tolérance sur le fractile à P probabilité de survie. Il peut ainsi être vu comme un « coefficient d'abattement » du fractile u_p permettant d'assurer un certain niveau de confiance lors du dimensionnement et ainsi un certain conservatisme (donc une certaine fiabilité même si non évaluée). Les écarts résiduels étant normalement distribués, le coefficient \mathbf{K} peut être obtenu à partir de la loi de Student non-centrale (intervalle unilatéral) [3-4] et [8] ou à partir de la loi du chi-deux (intervalle bilatéral) [8]. Les recommandations de l'IIW [5-6] proposent une formulation basée sur la borne minorante de l'intervalle de confiance (uni ou bilatéral) de la moyenne μ_N et sur la borne majorante de l'intervalle de confiance (uni ou bilatéral) de l'écart-type. Il semblerait que le coefficient \mathbf{K} soit légèrement plus pénalisant dans ce cas-là.

Selon la philosophie de l'IIW et en considérant ici un intervalle de tolérance bilatéral, le coefficient K s'écrit :

$$K = \left[\frac{t_{\frac{1+\gamma}{2}, k+1}}{\sqrt{n}} + u_P \cdot \sqrt{\frac{k+1}{\chi_{\frac{1-\gamma}{2}, k+1}^2}} \right] \cdot \sqrt{\frac{k}{k+1}} \quad (6)$$

Avec $k = n - 2$ le nombre de degrés de liberté associés à l'estimation des deux paramètres (A et B), si la pente B n'est pas fixée *a priori*. t désigne le fractile de la loi de student, χ^2 celui de la loi du chi-deux et γ correspond au niveau de confiance.

3 Approche probabiliste bayésienne

3.1 Bref état de l'art

Les méthodes probabilistes semblent opportunes pour modéliser les dispersions observées au niveau des courbes $S - N$ et font l'objet de nombreuses recherches. A titre d'exemple, Fouchereau *et al.* [9] proposent un modèle de courbes $S - N$ probabilistes valable sur l'intégralité de la durée de vie du matériau, c'est-à-dire intégrant à la fois la phase d'amorçage et celle de propagation de fissures longues, et cela sans analyse spécifique de la microstructure. Leur modèle est basé sur une loi des mélanges. Très récemment, D'Angelo et Nussbaumer [10] ont proposé des méthodes probabilistes pour obtenir les courbes $S - N$ probabilisées pour des assemblages soudés soumis à des chargements d'amplitude constante ou variable ; les données censurées sont prises en compte. La méthodologie s'appuie sur les simulations de Monte Carlo et met en jeu les distributions asymptotiques des paramètres (obtenues *via* la théorie du maximum de vraisemblance), ce qui est en principe une des limites de leur méthode dans le cas de petits échantillons. Ronold et Lostberg [11] s'intéressent à la prise en compte des incertitudes statistiques dans l'estimation des courbes de conception, surtout quand le nombre de données est limité. Ils remettent en cause les valeurs P, γ fournies par les recommandations. Une approche basée sur la méthode de fiabilité FORM (*First Order Reliability Method*) est alors proposée pour déterminer quelle valeur γ permet d'assurer un niveau de fiabilité donné (au niveau la structure étudiée). Dans certains cas, l'hypothèse de log-normalité de la dispersion du nombre de cycles est remise en cause, comme par exemple dans [12] où le phénomène de fatigue est dû au contact de roulement (*rolling contact fatigue*). De même, l'hypothèse de constance de la variance quel que soit le niveau de contraintes, peu réaliste pour les faibles niveaux, peut être levée, comme cela a été réalisé dans la thèse de Guédé [7] où des analyses de fiabilité sur des structures sollicitées en fatigue thermique sont ensuite menées.

Des approches probabilistes peuvent également être mises en œuvre dans un cadre bayésien pour déterminer les distributions *a posteriori* des paramètres du modèle de tendance et de la variance, comme cela a été fait dans [13]. L'approche proposée par Edwards et Pacheco permet de prendre en compte les essais non-rompus mais ne permet pas d'estimer les intervalles de confiance des courbes courbes $S - N - P$. L'actualisation bayésienne est aussi employée dans [14] pour comparer différents modèles de tendance en fatigue. Enfin, deux récents articles proposés par Dekker *et al.* [1] et Guida et Penta [15] montrent l'intérêt des approches bayésiennes pour déterminer les courbes $S - N - P_\gamma$ par rapport aux approches classiques. Des méthodologies sont notamment présentées pour construire les distributions *a priori* des paramètres et notamment celle de l'écart-type σ avec pour objectif d'agréger

différentes sources d'information possibles (jugement d'expert, essais antérieurs sur matériau similaire, etc.). Dans le cas d'inférence de courbes de conception en fatigue à partir de peu de données, fournir une information *a priori* sur l'écart-type peut en effet être particulièrement utile. Ceci permet notamment d'espérer réduire le degré de conservatisme des approches actuelles.

Le travail présenté dans ce papier s'inscrit dans ce contexte. En complément des références [1], [13-15], il s'agit de donner une illustration du comportement de la méthode bayésienne par rapport aux approches normatives en choisissant des lois *a priori* sur σ uniformes, progressivement de plus en plus informatives. Egalement, une méthode est proposée pour établir la borne de l'intervalle de confiance du fractile sous la même forme que celle présentée dans l'Eq. 5, c'est-à-dire *via* la détermination d'un coefficient K . Enfin, l'impact de la stratégie d'échantillonnage (nombre de niveaux de contraintes / nombre d'essais) est illustré.

3.2 Démarche proposée

L'actualisation bayésienne est utilisée pour modéliser la dispersion sur l'estimation du vecteur de paramètres $\theta = (A, B, \sigma)$ afin d'établir les courbes $S - N - P_\gamma$. Une fois supposée une densité *a priori* sur les paramètres, notée dans la suite $\pi_\theta(\theta)$, celle-ci est actualisée par l'ajout d'observations, ici la donnée des couples $(\log S_i, \log N_i)$ où $i \in \{1, \dots, n\}$ correspond au numéro de l'essai réalisé. On notera $(\mathbf{log S}, \mathbf{log N})$ la matrice de taille $n \times 2$ regroupant l'ensemble des observations. La distribution dite *a posteriori* du vecteur θ est alors donnée par :

$$p_{\theta|(\mathbf{log S}, \mathbf{log N})}(\theta|(\mathbf{log S}, \mathbf{log N})) = \frac{L((\mathbf{log S}, \mathbf{log N})|\theta) \cdot \pi_\theta(\theta)}{\int_{\theta} L((\mathbf{log S}, \mathbf{log N})|\theta) \cdot \pi_\theta(\theta) d\theta} \quad (7)$$

Où $L(\cdot)$ est la fonction de vraisemblance de l'échantillon à la loi log-normale, ici supposée pour modéliser la durée de vie N . Elle s'écrit donc :

$$L((\mathbf{log S}, \mathbf{log N})|\theta) = \prod_{i=1}^n \frac{1}{\sigma \cdot \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{\log N_i - M(A, B, \log S_i)}{\sigma} \right)^2} \quad (8)$$

En augmentant la taille de l'échantillon de données, la distribution *a posteriori* des paramètres converge vers celle du maximum de vraisemblance. Dans le cas de petits échantillons, c'est l'information *a priori* contenue dans $\pi_\theta(\theta)$ qui, au contraire, impacte le plus la distribution *a posteriori*. Le lecteur pourra se référer aux documents [1] et [15-16] pour un aperçu sur le choix de lois non informatives ou informatives possibles.

Dans un souci de facilité de mise en œuvre, nous choisissons dans ce travail un *a priori* uniforme. Plus précisément, le domaine de définition des lois uniformes des paramètres du modèle de tendance A et B est choisi assez large afin de rendre ces lois *a priori* les moins informatives possibles. Concernant, la loi uniforme *a priori* de l'écart-type, la borne supérieure b_σ est choisie comme une variable d'ajustement. Une valeur très grande conduira alors à une loi *a priori* très peu informative. Une borne plus faible traduira au contraire l'intégration d'une certaine information et rendra la loi *a priori* plus informative. De plus, nous n'apportons pas d'information concernant la dépendance entre les

paramètres A , B et σ et supposons donc leur indépendance *a priori*. Ainsi, nous pouvons écrire d'une manière générique :

$$\pi_{\theta}(\theta) = U_A[\cdot, \cdot] \times U_B[\cdot, \cdot] \times U_{\sigma}[0, b_{\sigma}] \quad (9)$$

Où les bornes seront précisées lors de l'application réalisée en section 4.

Les courbes $S - N - P$ et $S - N - P_{\gamma}$ s'obtiennent comme dans l'approche basée sur les normes et recommandations à partir des Eq. 4 et 5, mais les estimations de \hat{A} , \hat{B} et $\hat{\sigma}$ et le calcul du coefficient \mathbf{K} sont issus de la simulation.

En effet, N_{sim} réalisations $\theta^* = (A^*, B^*, \sigma^*)$ du vecteur $\theta = (A, B, \sigma)$ sont générées à partir de la distribution *a posteriori* de ces paramètres (Eq. 7 – 9). Les Chaînes de Markov Monte Carlo (MCMC) peuvent classiquement être mises en œuvre pour s'affranchir du calcul de l'intégrale au dénominateur de l'Eq. 7. Il est également possible de recourir à la récente méthode BUS (*Bayesian Updating with Structural Reliability Methods*) proposée par Straub et Papaioannou, voir [17] pour le principe de mise en œuvre de la méthode. Les avantages de cette technique sont qu'il n'y a pas de phase transitoire comme dans les chaînes MCMC et que les réalisations sont indépendantes entre elles. Le principal inconvénient de cette technique est que les temps de génération d'échantillons par la simulation de Monte Carlo (BUS-MC) sont longs lorsque la taille n de l'échantillon de données ($\log \mathbf{S}, \log \mathbf{N}$) est grande. Les auteurs proposent alors de coupler cette technique avec des méthodes de fiabilité plus performantes que les simulations MC pour accélérer les temps de calcul, comme la méthode des *Subset Simulation*.

A partir des N_{sim} réalisations $\theta^* = (A^*, B^*, \sigma^*)$, l'estimation de l'écart-type médian $\hat{\sigma}$ est directe. La courbe médiane (c'est-à-dire la courbe $S - N$ à probabilité de survie $P = 50\%$, soit les coefficients \hat{A} , \hat{B}), est estimée comme suit :

1. Génération de N_{sim} triplets $\theta^* = (A^*, B^*, \sigma^*)$ (Eq. 7 – 9)
2. A chaque niveau discrétisé de contrainte $S_l \in [S_{\min}, S_{\max}]$, calcul des N_{sim} log-durées de vie : $\log N_l^* = A^* + B^* \cdot \log S_l$ avec $l \in \{1, \dots, n_S\}$ et n_S le nombre de niveaux de contraintes considérés (potentiellement différent de celui des essais)
3. Estimation de \hat{A} et \hat{B} à partir de la régression linéaire réalisée sur les $N_{\text{sim}} \times n_S$ couples $(\log S_l, \log N_l^*)$.

Le coefficient \mathbf{K} est ici aussi calculé à partir de la simulation. On considère le cas d'un intervalle de tolérance bilatéral mais la procédure suivante s'applique également au cas d'un intervalle de tolérance unilatéral. On cherche une borne $T_{\text{inf}} = \hat{A} + \hat{B} \cdot \log S - \mathbf{K} \cdot \hat{\sigma}$ du supposé « vrai » fractile $\log N_P$ telle que :

$$P(T_{\text{inf}} \leq \log N_P \leq T_{\text{sup}}) = \gamma \text{ ou encore } P(\log N_P \leq T_{\text{inf}}) = \frac{1-\gamma}{2} \quad (10)$$

L'estimation de \mathbf{K} (et donc de T_{inf}) est obtenue par bootstrap à partir de la loi *a posteriori* de θ et du problème approché suivant :

$$P(\widehat{\log N}_P \leq T_{\text{inf}}^*) = \frac{1 - \gamma}{2} \quad (11)$$

Avec $\widehat{\log N}_P = \hat{A} + \hat{B} \cdot \log S - u_p \cdot \hat{\sigma}$ et $T_{\text{inf}}^* = A^* + B^* \cdot \log S - \mathbf{K} \cdot \sigma^*$. L'Eq. 11 peut alors être réécrite comme :

$$P(Z(S) \leq \mathbf{K}) = \frac{1 + \gamma}{2} \quad (12)$$

Avec $Z(S) = \frac{A^* + B^* \cdot \log S - \hat{A} - \hat{B} \cdot \log S + u_p \cdot \hat{\sigma}}{\sigma^*}$. Le coefficient \mathbf{K} est donc le fractile de $Z(S)$ pour la probabilité $\frac{1 + \gamma}{2}$. Il varie avec le niveau de contrainte S . Par souci de simplicité, on choisit ici \mathbf{K} comme la valeur maximale obtenue pour les n_S niveaux $S_l \in [S_{\text{min}}, S_{\text{max}}]$ considérés. Ceci permet d'obtenir une droite pour $T_{\text{inf}} = \widehat{\log N}_{P, \gamma}$.

4 Analyses réalisées et résultats

4.1 Analyses

Afin de juger la pertinence de la démarche bayésienne proposée et la positionner par rapport à l'approche basée sur les normes et recommandations, des échantillons de fatigue sont ici générés artificiellement. Un modèle de Basquin de paramètres $A = 11,47$ et $B = -3$ est utilisé. Le nombre de cycles N est la variable dépendante, distribuée suivant une loi log-normale d'écart-type constant $\sigma = 0,3$ (quel que soit le niveau de contrainte S). L'étude est réalisée sur la plage de contraintes $[S_{\text{min}} = 50 \text{ MPa}, S_{\text{max}} = 200 \text{ MPa}]$. Trois stratégies d'échantillonnage sont analysées pour un même nombre total d'essais $n = 6^1, 9, 12, 18$ ou 36 . Ainsi, on considère 3, 6 ou 9 niveaux de contraintes. Une fois un essai réalisé (généralisé aléatoirement ici), celui-ci est conservé quel que soit la stratégie d'échantillonnage et lorsque n est progressivement augmenté, comme dans une approche expérimentale.

Dans la suite, on s'intéresse aux courbes $S - N$ pour une probabilité de survie $P = 97,7\%$ (ce qui correspond à 2 écarts-types par rapport à la courbe médiane) ainsi qu'aux courbes de conception pour une probabilité de survie $P = 97,7\%$ et un niveau de confiance $\gamma = 90\%$ (intervalle bilatéral). Les couples $(\hat{A} - u_p \cdot \hat{\sigma}, \hat{B})$ ou $(\hat{A} - \mathbf{K} \cdot \hat{\sigma}, \hat{B})$ obtenus soit par l'approche classique (Eq. 4 – 6) soit par l'approche bayésienne (section 3.2) sont ensuite recalculés à pente imposée $B = -3$ pour faciliter la comparaison des résultats par rapport à la solution de référence $(11,47 - u_{0,977} \cdot 0,3 = 10,8714, -3)$. Ainsi seuls les coefficients ci-après nommés $A_{97,7\%}$ (sans prise en compte des incertitudes d'estimation) ou $A_{97,7\% \gamma = 90\%}$ (avec prise en compte des incertitudes d'estimation) sont analysés et comparés à la valeur de référence 10,8714.

¹ Seules deux stratégies sont testées dans ce cas-là : 3 niveaux / 2 essais et 6 niveaux / 1 essai.

Concernant la simulation, $N_{\text{sim}} = 15000$ réalisations sont générées pour assurer un coefficient de variation de l'ordre de 5%. Les méthodes MCMC (les 20000 premiers points sont supprimés pour assurer la stationnarité de la chaîne et 1/40 points sont conservés pour réduire la corrélation) et BUS-MC ont été mises en œuvre. Elles donnent des résultats similaires. Seuls les résultats obtenus avec la méthode BUS-MC sont ici présentés. La loi *a priori* du vecteur de paramètres $\theta = (A, B, \sigma)$ est choisie telle que :

$$\pi_{\theta}(\theta) = U_A[1,30] \times U_B[-10, -1] \times U_{\sigma}[0, b_{\sigma}]$$

Avec $b_{\sigma} \in \{0,2; 0,3; \dots; 0,9\}$ pour intégrer progressivement une information *a priori* sur la variance.

4.2 Résultats

Les distributions marginales *a priori* et celles *a posteriori* des paramètres (A, B, σ) obtenues à partir de la méthode BUS-MC sont illustrées aux Fig. 1 et 2, pour $b_{\sigma} = 0,8$ et les deux échantillons $M_6_3_2$ ($n = 6$, 3 niveaux de contraintes / 2 essais) et $M_36_3_12$ ($n = 36$, 3 niveaux de contraintes / 12 essais). L'effet de la taille de l'échantillon sur les distributions *a posteriori* est évident, surtout pour celle de l'écart-type qui s'éloigne de la distribution uniforme *a priori* lorsque n augmente.

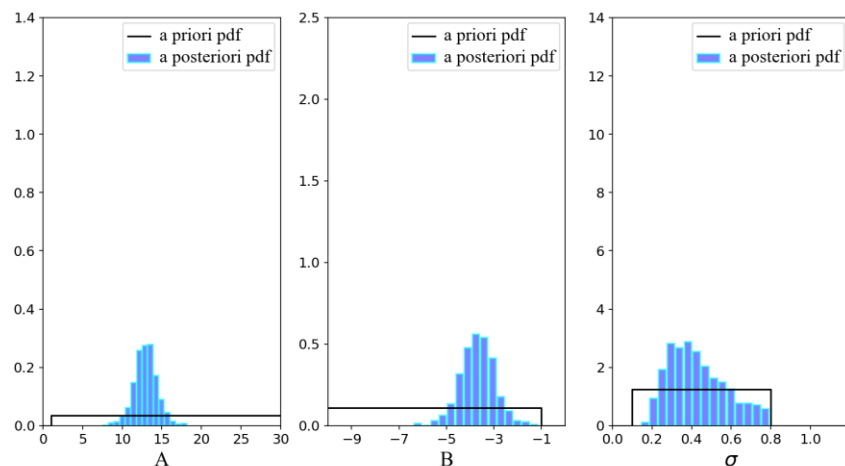


Figure 1 : Distributions *a priori* et *a posteriori* des paramètres. $b_{\sigma} = 0,8$; $M_6_3_2$.

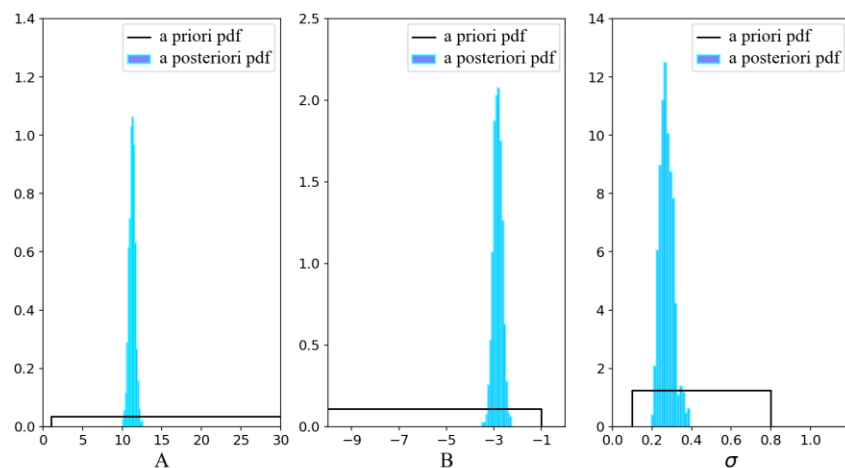


Figure 2 : Distributions *a priori* et *a posteriori* des paramètres. $b_{\sigma} = 0,8$; $M_36_3_12$.

Les Fig. 3 à 5 positionnent l'approche statistique / fréquentiste (« Normes /Recom. ») et l'approche bayésienne (« Bayes M_x_x_x ») (même code couleur) par rapport à la valeur de référence (« Valeur Ref. ») (en noire) pour chaque stratégie d'échantillonnage. Les coefficients A sans ou avec prise en compte des incertitudes sur l'estimation des paramètres sont reportés ainsi que l'estimation de l'écart-type.

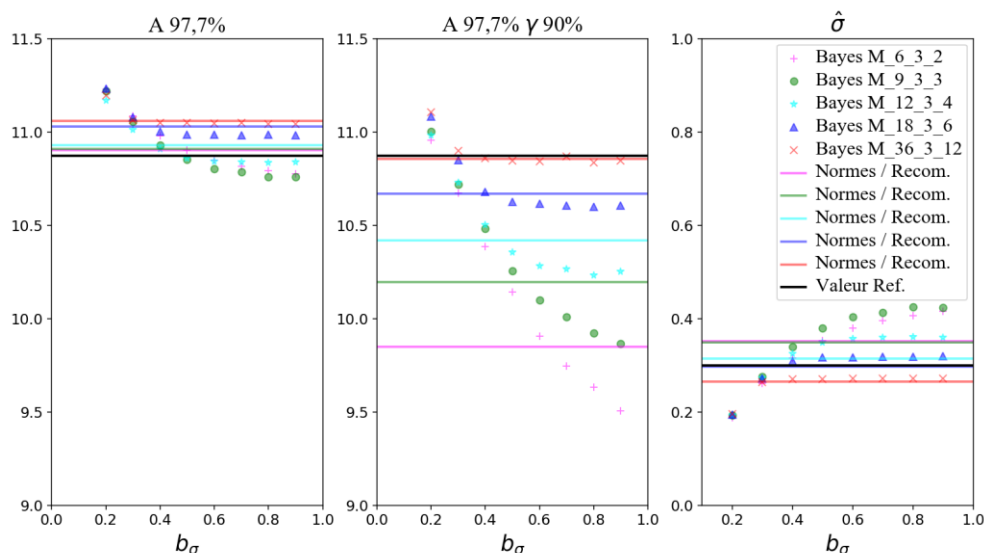


Figure 3 : Comparaison des approches statistique et bayésienne en fonction de b_σ , 3 niveaux de contraintes, $n = 6, 9, 12, 18, 36$.

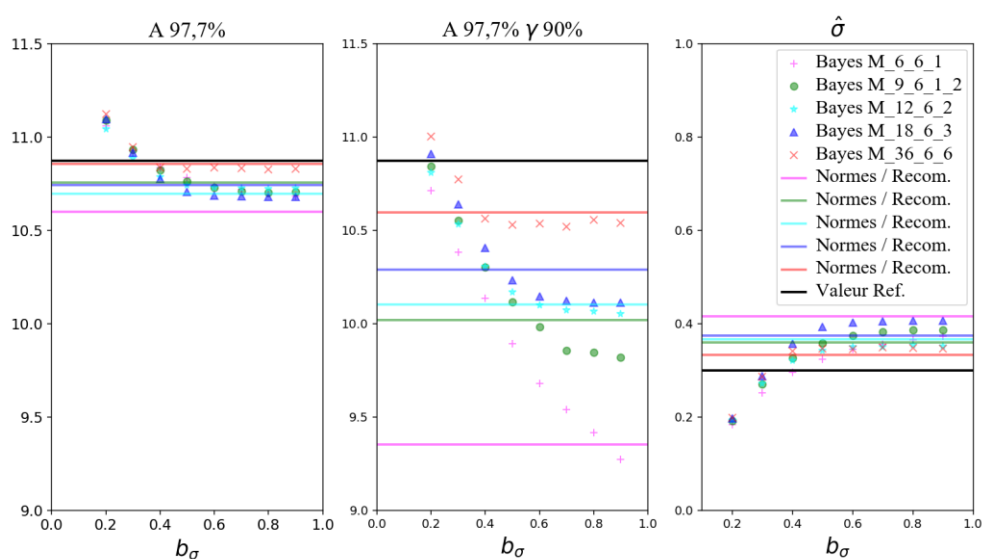


Figure 4 : Comparaison des approches statistique et bayésienne en fonction de b_σ , 6 niveaux de contraintes, $n = 6, 9, 12, 18, 36$.

De cette brève étude, trois observations peuvent être relevées. Tout d'abord, les échantillons de données ici générés conduisent potentiellement à une estimation non conservative de la courbe $S - N$ pour une probabilité de survie $P = 97,7\%$ par rapport à la référence. C'est le cas pour l'estimation *via* la norme pour la stratégie 3 niveaux de contraintes ou encore l'estimation bayésienne avec une borne b_σ trop faible par exemple. En répétant l'analyse à partir de nouveaux échantillons de données (résultats non présentés ici), il apparaît que toutes les courbes $S - N - 97,7\%$ obtenues par

l'approche normative ou bayésienne (si $b_\sigma > 0.3$) peuvent être conservatives, c'est-à-dire quel que soit le nombre de données ou la stratégie d'échantillonnage. Ceci reflète donc l'impact de l'échantillonnage et l'importance d'intégrer un niveau de confiance pour estimer les courbes $S - N - P$ dans une optique de conception.

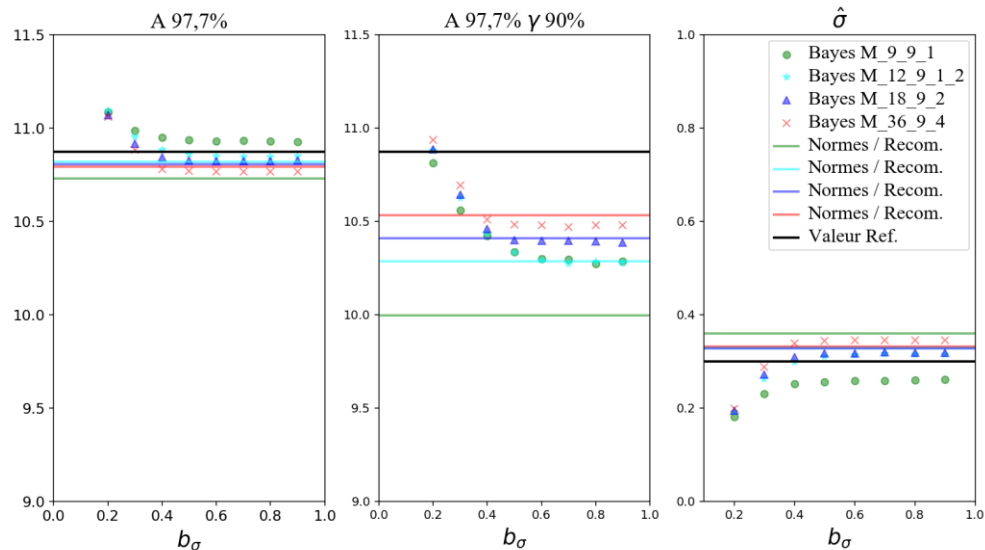


Figure 5 : Comparaison des approches statistique et bayésienne en fonction de b_σ , 9 niveaux de contraintes, $n = 9, 12, 18, 36$.

L'analyse des courbes de conception $S - N - P_\gamma$ avec $P = 97,7\%$ et $\gamma = 90\%$ montre que l'approche normative est systématiquement conservatrice mais que le degré de conservatisme diminue lorsque la taille de l'échantillon augmente et cela quel que soit la stratégie d'échantillonnage. Ceci est à la fois dû à une meilleure estimation de la tendance mais également de la dispersion, ce qui est plutôt rassurant. Concernant l'approche bayésienne, si aucune ou très peu d'information est fournie concernant l'écart-type (via la borne b_σ) alors les résultats obtenus sont plus conservatifs, c'est-à-dire plus éloignés de la valeur de référence, que ceux fournis par la norme (sauf pour l'échantillon $M_9_9_1$ ici où l'écart-type $\hat{\sigma}$ est fortement sous-estimé, résultat qui diffère si on change d'échantillon). Cependant, dès lors qu'une information est prise en compte sur l'écart-type, les courbes de conception obtenues par l'approche bayésienne sont moins conservatrices que l'approche normative. Ceci est d'autant plus marqué pour les petits échantillons de données $n = 6$ ou 9 , ce qui souligne l'intérêt potentiel de l'approche bayésienne par rapport à l'approche normative qui conduit à une forte sous-estimation des fractiles dans ces cas-là (et donc à des courbes potentiellement peu utilisables en pratique). Par ailleurs, l'approche bayésienne reste conservatrice si b_σ n'est pas inférieure à la vraie valeur (mais inconnue) $0,3$. Une perspective de ce travail serait donc de proposer une méthode pour ajuster le paramètre b_σ à partir de la physique du problème et des données pouvant être au préalable recueillies.

A partir des échantillons testés ici, il semblerait enfin que la stratégie avec le moins de niveaux à iso nombre total d'essais n (ici 3 niveaux) permette d'obtenir des courbes de conception $S - N - P_\gamma$ globalement plus proches du processus sous-jacent, que l'on ait recours à la stratégie bayésienne ou à l'approche basée sur les normes et les recommandations. Ceci appuierait donc l'analyse en [2]. Ce résultat mériterait néanmoins d'être confirmé à partir d'autres essais.

5 Conclusion

Cet article illustre l'intérêt potentiel que peuvent avoir les approches probabilistes bayésiennes en fatigue des matériaux en complément des approches classiques. Une démarche a été proposée pour associer un niveau de confiance aux courbes de fatigue $S - N - P$ issues de la simulation bayésienne, d'une manière analogue à la pratique actuelle. Des analyses basées sur des échantillons aux propriétés connues ont ensuite été réalisées pour comparer les approches basées sur les normes et les recommandations et l'approche bayésienne. Il ressort qu'en l'absence (ou peu) d'information sur la dispersion des données, l'approche bayésienne est plus conservatrice que l'approche traditionnelle, quel que soit la taille de l'échantillon et la stratégie d'échantillonnage. Ceci pourrait donc rassurer les concepteurs quant à l'utilisation de la méthode. A l'inverse, si la distribution *a priori* de l'écart résiduel est plus informative, alors un gain est obtenu dans l'estimation du fractile de la courbe $S - N - P$ pour un même niveau de confiance donné, notamment dans le cas des petits échantillons. Les perspectives de ce travail exploratoire sont nombreuses. Tout d'abord, les résultats mériteraient d'être confirmés à partir d'échantillons issus de processus aux propriétés différentes et dans le cas d'échantillons indépendants entre eux (pas de conservation des données d'une stratégie d'échantillonnage à une autre et lorsque la taille augmente). Il serait également intéressant de proposer une méthode permettant de donner une valeur (ou un intervalle) pour la borne maximale de la loi *a priori* de la variance basée sur la physique du phénomène et/ou des données (afin de restreindre son amplitude de variation) et/ou pour conférer un niveau de fiabilité donné. Enfin, une réflexion sur les choix des couples (probabilité de survie, niveau de confiance) pourrait être menée pour mieux comprendre et évaluer le degré de conservatisme associé aux courbes de conception, comme dans [11]. Ceci semble nécessaire pour assurer un niveau de fiabilité cible lors du dimensionnement d'une structure.

Remerciements

Les auteurs remercient l'ensemble des participants au groupe de travail « Prise en compte des incertitudes dans le dimensionnement en fatigue » en lien avec la commission Fatigue de la SF2M et le GST Mécanique et Incertain de l'AFM pour leurs réflexions sur le sujet. Ils remercient particulièrement Mr Géraud Blatman et Mr Chu Mai, ingénieurs – chercheurs du groupe EDF France, pour avoir proposé l'estimation du coefficient K introduit au paragraphe 3.2 (Eq. 10 à 12).

Références

- [1] S. Dekker, G. Wurzel, R. Alderliesten, A Bayesian tolerance interval estimation for fatigue strength substantiation of rotorcraft dynamic components, International Journal of Fatigue 92 (2016) 333-344
- [2] Standard Practice for Statistical Analysis of Linear or Linearized Stress-Life (S-N) and Strain-Life (ϵ -N) Fatigue Data, ASTM E739-10
- [3] Metallic materials – Fatigue testing – Statistical planning and analysis of data, BS ISO 12107: (2003)
- [4] Produits métalliques – Essais de Fatigue – Traitement statistique des données, NF – A 03-405
- [5] A. Hobbacher, Recommendations For Fatigue Design of Welded Joints and Components, IIW 1823-07 (2008)

- [6] M. Huther, Fatigue Testing and Evaluation of Data for Design, IIS/IIW – Com XIII, Lappeenranta, (2003)
- [7] Z. Guédé, Approche probabiliste de la durée de vie des structures sollicitées en fatigue thermique, Thèse, Université Blaise Pascal Clermont II, 2005
- [8] NIST/SEMATECH e-Handbook of Statistical Methods, <http://www.itl.nist.gov/div898/handbook/>, 2012 (section 7.2.3.6 Tolerance intervals for a normal distribution)
- [9] R. Fouchereau, G. Celeux, P. Pamphile, Probabilistic modeling of S-N curves, International Journal of Fatigue 68 (2014) 217–223
- [10] L. D’Angelo, A. Nussbaumer, Estimation of fatigue S-N curves of welded joints using advanced probabilistic approach, International Journal of Fatigue 97 (2017) 98–113
- [11] K. O. Ronold, I. Lotsberg, On the estimation of characteristic S-N curves with confidence, Marine Structures 27 (2012) 29–44
- [12] Y. X. Zhao, H. B. Liu, Weibull modeling of the probabilistic S-N curves for rolling contact fatigue, International Journal of Fatigue 66 (2014) 47–54
- [13] G. Edwards, L.A. Pacheco, A Bayesian method for establishing fatigue design curves, Structural Safety 2 (1984) 27-38
- [14] I. Babuska, Z. Sawlan, M. Scavino, B. Szabo, R. Tempone, Bayesian inference and model comparison for metallic fatigue data, Comput. Methods Appl. Mech. Engrg 304 (2016) 171-196
- [15] M. Guida, F. Penta, A Bayesian analysis of fatigue data, Structural Safety 32 (2010) 64–76
- [16] C. Elster, K. Klauenberg, M. Walzel et al., A guide to Bayesian inference for regression problems, technical report of project NEW04 “Novel mathematical and statistical approaches to uncertainty evaluation”, 2015
- [17] D. Straub, I. Papaioannou, Bayesian Updating with Structural reliability methods, Journal of Engineering Mechanics 141(3) (2015) 04014134