Réflexions sur l'analyse d'incertitudes dans un contexte industriel : informations disponibles et enjeux décisionnels

Merlin M. Keller, Alberto A. Pasanisi, Éric É. Parent

Journal de la Société Française de Statistique
Vol. 152 No. 4 (2011)

Réflexions sur l'analyse d'incertitudes dans un contexte industriel : information disponible et enjeux décisionnels

Title: On uncertainty analysis in an industrial context:
Or, how to combine available information with decisional stakes

Merlin Keller[1], Alberto Pasanisi[1] et Eric Parent[2]

Résumé : L'analyse d'incertitudes a pour objet de quantifier le degré de connaissance affectant la valeur d'une quantité d'intérêt, caractéristique du fonctionnement d'un système physique, et liée à des enjeux décisionnels. La plupart des approches rencontrées en ingénierie font appel à l'inférence statistique, et se divisent en trois grandes classes.

Les techniques dites "plug-in" fournissent une estimation ponctuelle de la quantité d'intérêt, valable uniquement en présence d'un grand nombre de données. Lorsque le nombre de données est réduit, il est préférable de faire appel aux procédures de Bayes, qui déduisent une valeur optimale de la grandeur d'intérêt d'une loi *a priori* décrivant l'incertitude sur les paramètres du modèle et d'une fonction de coût formalisant les enjeux décisionnels. Enfin, les approches purement descriptives visent à décrire l'incertitude sur la quantité d'intérêt, plutôt qu'à en fournir une estimation ponctuelle.

De nombreuses heuristiques ont été proposées pour contourner le problème du choix d'une fonction de coût pour l'estimation de la quantité d'intérêt dans un cadre bayésien. Nous considérons en particulier celle qui consiste à remplacer dans la définition de la quantité d'intérêt la distribution réelle de la variable de sortie du système, qui est en général inconnue, par sa distribution prédictive. Nous montrons que cette approche amène implicitement à utiliser un estimateur bayésien, relatif à une fonction de coût qui dépend entièrement de l'expression de la quantité d'intérêt. Ce résultat démontre qu'une estimation ponctuelle sous incertitude repose nécessairement sur le choix, conscient ou non, d'une fonction de coût.

Nous illustrons notre propos sur un jeu de données réelles de hauteurs et débits d'un cours d'eau, et discutons plus généralement la pertinence de chaque approche en fonction des enjeux de l'étude, et de la connaissance plus ou moins explicite dont dispose l'analyste.

Abstract: Uncertainty analysis aims to quantify the degree of knowledge affecting the value of a quantity of interest, characteristic of the behavior of a physical system, and related to decisional stakes. The approaches most encountered in engineering practice involve statistical inference, and fall into three broad classes.

So-called "plug-in" techniques provide a point estimate of the quantity of interest, which is only valid in presence of a large number of observations. When the data are scarce, one may favor Bayes procedures, which deduce an optimal value of the quantity of interest from an *a priori* distribution, characterizing the uncertainty on the model parameters, and a cost function, formalizing the stakes motivating the analysis. Finally, purely descriptive approaches aim at describing the uncertainty affecting the quantity of interest, rather than providing a point estimate.

Many heuristics have been proposed to avoid the choice of a cost function for point estimation in a Bayesian context. In particular, we consider the approach that consists in replacing, inside the definition of the quantity of interest, the true distribution of the output variable of the physical system, which is generally unknown, by its predictive distribution. We show that this approach leads implicitly to adopt a Bayesian estimator, relative to a cost function that depends

[1] EDF R&D Dépt. Management des Risques Industriels, 6 quai Watier, 78401 Chatou
E-mail : merlin.keller@edf.fr and E-mail : alberto.pasanisi@edf.fr
[2] AgroParisTech, UMR 518 Mathématiques et Informatique Appliquées, 16 rue Claude Bernard, 75005 Paris
E-mail : eric.parent@agroparistech.fr

 ISSN: 2102-6238

entirely on the expression of the quantity of interest at hand. Thus, this result demonstrates that point estimation under uncertainty necessarily requires the choice, conscious or not, of a cost function.

Our argument is illustrated on a real dataset of height and discharge measures from a river section. More generally, we discuss the relevance of the above-cited approaches, in relation to the stakes motivating the study, and the amount of information available to the analyst.

Mots-clés : analyse d'incertitudes, théorie de la décision, incertitude épistémique, Estimation bayésienne, estimation prédictive
Keywords: uncertainty analysis, decision theory, epistemic uncertainty, Bayesian estimation, predictive estimation
Classification AMS 2000 : 62P30, 62C10

1. Modèle déterministe

Le cadre général de l'analyse d'incertitudes, tel qu'il a été formalisé par [14] entre autres, est celui d'un modèle déterministe, décrit par une équation de la forme :

$$\mathrm{Y} = G(\mathrm{X}), \tag{1}$$

où X est le vecteur des entrées, $G(\cdot)$ est un code déterministe (une fonction) et Y la sortie du système, que nous supposons scalaire par souci de simplicité. Le fonctionnement de nombre de systèmes physiques se met sous la forme (1), où le degré de complexité de la fonction G (équation différentielle ordinaire, aux dérivées partielles, etc.) varie suivant les applications.

Exemple. Pour illustrer notre propos, nous considérons ici un modèle hydraulique simplifié de relation débit/hauteur pour un tronçon de rivière, résultant de la résolution des équations de Saint-Venant en 1D sous hypothèse d'écoulement stationnaire et de section rectangulaire très large. La fonction G se présente alors sous une forme analytique paramétrée :

$$Z_c = Z_v + \left\{ Q \,/\, \left(BK_s \sqrt{(Z_m - Z_v)/L} \right) \right\}^{0.6}, \tag{2}$$

où Z_c représente la cote de la surface de la rivière en aval (en m), Z_m et Z_v les cotes du fond de la rivière en amont et en aval respectivement (en m), Q le débit (en m^3/s), B la largeur du cours d'eau (en m), K_s le coefficient de Strickler caractérisant la rugosité du lit de la rivière, et L la longueur (en m) du tronçon considéré (voir Figure 1). La connaissance des valeurs des constantes B, K_s, Z_m, Z_v et L détermine ici complètement la transformation G.

Bien que les données que nous utiliserons dans la suite soient réelles, nous insistons sur le fait que cet exercice très simplifié a exclusivement une valeur d'exemple ; ses conclusions ne sont pas représentatives des méthodes de gestion du risque hydraulique chez EDF R&D.

2. Probabilisation des entrées

Supposons à présent que l'on ait muni le vecteur d'entrée X d'une structure de variable aléatoire, de fonction de répartition $F(x)$. Il s'ensuit que Y est également une variable aléatoire, dont la fonction de répartition $H(t) = \mathbb{P}(\mathrm{Y} \leq t) = \mathbb{P}(G(\mathrm{X}) \leq t)$ est entièrement déterminée par la donnée du couple $\Theta = (F, G)$. Et pour souligner que tous les résultats de nos calculs dépendent de ces inconnues, nous noterons dans la suite $H(t) = H(t|\Theta)$.

 ISSN: 2102-6238

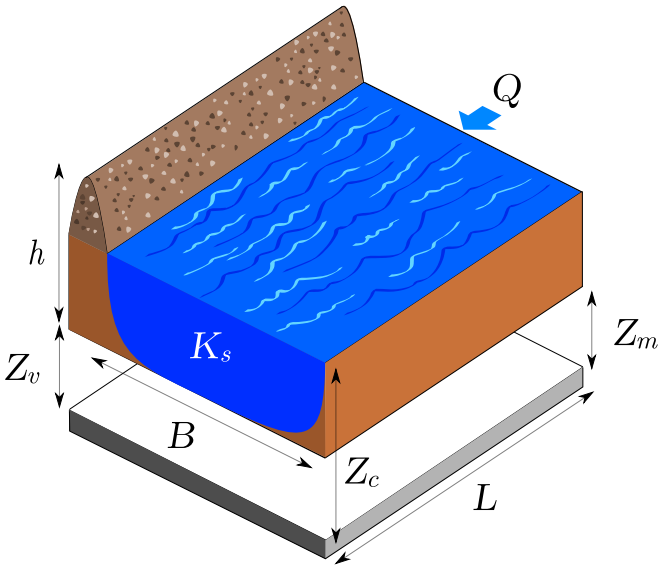

FIGURE 1. *Modèle simplifié d'un tronçon de rivière*

La première composante de l'inconnue Θ est la loi F. Celle-ci traduit souvent la variabilité naturelle des grandeurs physiques modélisées par les entrées X. Le fait que Y soit alors également une variable aléatoire est souvent interprété comme une "propagation d'incertitude" par le système physique de ses entrées à ses sorties (voir à ce sujet [8] ainsi que [5]).

D'un point de vue théorique, le calcul de la distribution $H(t|\Theta)$ fait intervenir le jacobien de la transformation G, qui est la deuxième composante de l'inconnue Θ. Il s'agit d'un problème classique de transformation de variables en calcul des probabilités. Certes, en pratique, obtenir la loi de Y peut s'avérer difficile, en particulier si le code $G(\cdot)$ est numériquement coûteux. Dans ce cas, des techniques d'intégration de Monte-Carlo avec réduction de la variance (cf. par exemple [21] ainsi que [20]), ou d'émulation de code (cf. [22] et [14]) peuvent s'avérer nécessaires. C'est pourquoi ce calcul est souvent considéré comme l'un des enjeux majeurs de l'analyse d'incertitudes.

Enfin, notons que $\Theta = (F, G)$ récapitule quels modèles statistiques et déterministes sont à considérer. Il s'agit donc d'un *paramètre* dans le sens statistique du terme, qui plus est d'un paramètre fonctionnel, puisque F et G sont toutes deux des fonctions. En pratique cependant, il arrive le plus souvent que l'on se limite à choisir les fonctions F et G dans des familles paramétriques, de telle sorte que l'espace fonctionnel contenant Θ devienne de dimension p finie. Dans ce cas, on définit directement Θ comme un vecteur de $\mathbb{R}^p$.

Le paramètre Θ régit la distribution des entrées comme des sorties du système physique considéré, il caractérise donc entièrement le fonctionnement de ce système ; Θ est encore appelé *état de la nature*. En accord avec cette interprétation phénoménologique, Θ est une constante : il ne peut prendre plusieurs valeurs distinctes, la nature ne pouvant se trouver dans plusieurs états à la fois. Pour le moment jusqu'à la section 4 , la valeur de Θ est supposée parfaitement connue de l'analyste, une situation qualifiée *d'information parfaite*.

 ISSN: 2102-6238

Exemple. Dans le modèle hydraulique (2), nous modélisons Q, représentant le débit maximal annuel, par une variable aléatoire de loi de Gumbel $Gu(\mu,\rho)$, de fonction de répartition

$$F(x) = \exp\{-\exp[\rho(\mu - x)]\}. \tag{3}$$

Justifiée par la théorie des valeurs extrêmes [7], la loi de Gumbel, appartenant à la famille des distributions GEV (Generalized Extreme Value), est couramment utilisée par les ingénieurs pour modéliser la distribution de probabilité de maxima annuels de débits, hauteurs de pluie, vitesses de vent, etc. La loi de la sortie Z_c, qui représente la cote maximale annuelle, est alors explicite :

$$H(t|\Theta) = \exp\left(-\exp\left\{\rho\left(\mu - BK_s\sqrt{\frac{Z_m - Z_v}{L}}(t - Z_v)^{5/3}\right)\right\}\right). \tag{4}$$

Ici l'état de la nature Θ est défini comme le vecteur : $(B, K_s, Z_m, Z_v, L, \mu, \rho) \in \mathbb{R}^7$, qui suffit à déterminer entièrement les fonctions F et G.

3. Quantité d'intérêt

Une fois déterminée la distribution de la sortie Y, il devient possible d'en évaluer n'importe-quelle grandeur caractéristique Φ (moyenne, variance, quantile, etc...). Celle-ci dépend de la distribution H, qui dépend elle-même uniquement de Θ, donc nous noterons $\Phi(\Theta)$ pour mettre en évidence que connaître l'état de la nature est indispensable pour mener ce calcul.

$\Phi(\Theta)$ résume la distribution de la variable aléatoire Y, qui représente une grandeur physique caractéristique de l'état du système physique étudié. Elle est présentée la plupart du temps comme directement liée aux enjeux décisionnels ayant motivé l'analyse, et porte pour cette raison le nom de "quantité d'intérêt" dans ce schéma, que nous pouvons qualifier de "normatif". On considère alors que son évaluation constitue l'enjeu central de l'analyse d'incertitudes en situation d'information parfaite, puisqu'elle est susceptible de guider le choix d'une certaine décision, comme le soulignent entre autres [8] ou [1].

Dans les études fiabilistes par exemple, la quantité d'intérêt prend généralement la forme d'une probabilité de dépassement de seuil $\mathbb{P}(Y \geq t)$, ou d'un quantile, bien que des expressions plus complexes, combinant les probabilités jointes de plusieurs évènements, puissent être utilisées (voir [9] pour une définition plus générale de la notion de fiabilité). Des enjeux économiques sont souvent associés à la quantité d'intérêt. Dès lors, il est conceptuellement satisfaisant de supposer qu'elle minimise les coûts socio-économiques des décisions envisageables au terme de l'étude :

$$\Phi(\Theta) = \arg\min_d \; c(d;\Theta), \tag{5}$$

où $c(d;\Theta)$ représente le coût d'une certaine "décision" d lorsque l'état du système physique est donné par Θ.

Exemples.

1. Pour fixer les idées, supposons que la décision d à prendre représente la hauteur d'une digue en construction. On peut alors envisager d'utiliser à cette fin le quantile $\Phi(\Theta) = q_\beta$ d'ordre β de Y, défini par : $q_\beta = H^{-1}(\beta|\Theta)$.

 ISSN: 2102-6238

En effet, on peut montrer que celui-ci peut s'exprimer comme la quantité d minimisant le coût linéaire par morceaux :

$$c(d;\Theta) = \mathbb{E}[|\mathrm{Y}-d|(\beta 1_{\{\mathrm{Y}>d\}} + (1-\beta)1_{\{\mathrm{Y}<d\}})|\Theta]$$

(voir [17] pour plus de détails), que l'on peut également écrire (grâce à une intégration par parties) :

$$c(d;\Theta) = \int_{-\infty}^{d} H(t|\Theta)\,\mathrm{d}t - d\times\beta. \tag{6}$$

q_β peut donc s'interpréter comme la hauteur de digue optimale si les coûts de construction et de débordement de la digue sont tous deux supposés linéaires.

2. La décision d peut aussi simplement consister en le choix d'une estimation. Ainsi, on montre facilement que l'espérance $\Phi(\Theta) = \mathbb{E}[\psi(\mathrm{Y})|\Theta]$ de toute fonction ψ de la sortie Y peut être obtenue en minimisant le coût quadratique :

$$c(d;\Theta) = \mathbb{E}[(d-\psi(Y))^2|\Theta]. \tag{7}$$

Par exemple, la probabilité de dépassement $\mathbb{P}(\mathrm{Y}>t|\Theta) = 1-H(t|\Theta)$ peut être définie comme la valeur de d minimisant le coût quadratique $c(d;\Theta) = \mathbb{E}[(d-1_{\{\mathrm{Y}>t\}})^2|\Theta]$. Elle s'interprète alors comme une estimation de l'indicateur binaire de dépassement du seuil t par Y, pour un coût d'erreur quadratique.

3. Poursuivant l'exemple bâti sur les équations (2) et (3), nous considérons les quantités d'intérêt suivantes :
   - La valeur h_{opt} minimisant la fonction de coût suivante (proposée et étudiée en détail par [4]) :

$$c(d;\Theta) \quad = \quad I_0\times(Z_v+d) + C_0\times\mathbb{E}[1_{\{Z_c>d\}}(Z_c-Z_v-d)^2|\Theta], \tag{8}$$

   Dans l'expression (8) du coût de construction d'une digue de protection fluviale de hauteur d (en m), I_0 représente un coût d'investissement marginal et C_0 un coût de dommage marginal. La fonction $\Phi(\Theta)$ peut être évaluée numériquement, en utilisant par exemple une approximation par quadrature ou par simulation Monte-Carlo.
   - La probabilité de dépassement $\Phi(\Theta) = P(Z_c > Z_v+h|\Theta) = 1-H(Z_v+h|\Theta)$, qui peut se calculer explicitement d'après (4). Celle-ci peut s'interpréter comme la probabilité de débordement d'une digue de hauteur h (en m) (voir Figure 1) ;
   - Le quantile $\Phi(\Theta) = q_\beta = H^{-1}(\beta|\Theta)$, donné par la formule

$$q_\beta = Z_v + \left\{\left(\mu - \frac{1}{\rho}\log\log\frac{1}{\beta}\right)\Big/\left(K_s B\sqrt{\frac{Z_m-Z_v}{L}}\right)\right\}^{0.6}. \tag{9}$$

   Si T est un entier naturel, la crue dite de période de retour T, $q_{1-1/T}$, s'interprète comme la hauteur de la crue T-annuelle, c'est-à-dire qu'il faut attendre en moyenne T années (sous des hypothèses de stationnarité et d'indépendance) avant d'observer une crue $Z_c > q_{1-1/T}$.

Notons que l'évaluation de Φ ne pose toujours pas de difficulté conceptuelle particulière pour peu que Θ soit connu : elle prend alors une valeur bien déterminée. Dans un contexte d'information parfaite, l'analyse d'incertitudes n'est donc qu'un problème purement probabiliste, plus ou moins complexe, de transformation de variables aléatoires.

 ISSN: 2102-6238

4. Incertitude épistémique

En pratique cependant, le paramètre Θ, c'est-à-dire le couple (F,G) ou le vecteur le définissant (voir Section 2), n'est jamais parfaitement connu. Autrement dit, l'ingénierie ordinaire s'effectue toujours en situation d'*information imparfaite*. L'incertitude qui affecte l'état de la nature est dite *épistémique*, c'est-à-dire par manque de connaissance du système physique. Il devient alors nécessaire d'*estimer* les inconnues à partir de l'information disponible.

Ainsi, F peut être estimée à l'aide d'un échantillon $\mathrm{x} = (x_1, \ldots, x_n)$ de mesures d'entrées du système physique étudié, assimilées à des réalisations indépendantes de la variable d'entrée X. De même, G peut être estimée à partir de couples $(\tilde{\mathrm{x}}, \tilde{\mathrm{y}}) = ((\tilde{x}_1, \tilde{y}_1), \ldots, (\tilde{x}_m, \tilde{y}_m))$ de mesures d'entrée/sortie du système, en supposant que chaque couple vérifie la relation (1), à d'éventuelles erreurs de mesure près. Dans la suite, nous notons $D = (\mathrm{x}, (\tilde{\mathrm{x}}, \tilde{\mathrm{y}}))$ l'ensemble des données disponibles.

L'incertitude sur Θ est qualifiée de *réductible* par la collecte de nouvelles informations. De façon plus formelle, la théorie des statistiques asymptotiques, telle qu'on la trouve par exemple dans [25], étudie les conditions sous lesquelles on peut faire diminuer arbitrairement l'erreur d'estimation sur Θ en augmentant la taille n de l'échantillon.

Par opposition, l'incertitude affectant les grandeurs modélisées par X est dite *naturelle*, ou encore *par essence*, car, sous les hypothèses de modélisation adoptées, elle résulte de la nature imprévisible du phénomène observable. Elle est par là même *irréductible par construction*, car elle ne peut être diminuée par l'ajout d'une quelconque information.

Exemple. Dans le modèle hydraulique (2), afin de se focaliser sur un seul paramètre essentiel pour G, nous supposons connues les quantités Z_m, Z_v, B et L, qui résument la géométrie du tronçon de rivière considéré (en pratique, ces grandeurs sont en fait mesurées avec une certaine précision). La transformation G est donc complètement déterminée par le coefficient de Strickler K_s, qui est *a priori* inconnu. Ce dernier est donc affecté d'une incertitude *épistémique*, dûe à une connaissance imparfaite de sa valeur pour le système physique considéré.

Par opposition, le débit représenté par la variable Q, varie de manière imprévisible d'une année sur l'autre. Il s'agit donc d'une grandeur incertaine par nature dont on ne pourra jamais prédire exactement la valeur (sauf à changer de modèle et introduire d'autres relations, par exemple de transfert pluie-débit). Enfin, la distribution F de la variable Q est elle-même, comme G, affectée d'une incertitude épistémique (les coefficients μ et ρ ne sont pas connus parfaitement). On notera dans la suite de cet exemple $\Theta = (K_s, \mu, \rho)$ car ce triplet définit ici complètement les inconnues F et G du modèle.

5. Estimation "plug-in"

Sur le plan pratique, il existe de nombreuses méthodes permettant d'obtenir des estimations ponctuelles $\widehat{\Theta} = \widehat{\Theta}(D)$ des paramètres du modèle, à partir des données D disponibles. On peut de manière grossière les regrouper en deux classes : les approches paramétriques (par maximum de vraisemblance, moindres carrés, ...) et non paramétriques (estimateurs à noyau, empiriques, ...) (cf. par exemple [28] et [29]). On peut alors estimer toute quantité d'intérêt $\Phi(\Theta)$ par "plug-in"

 ISSN: 2102-6238

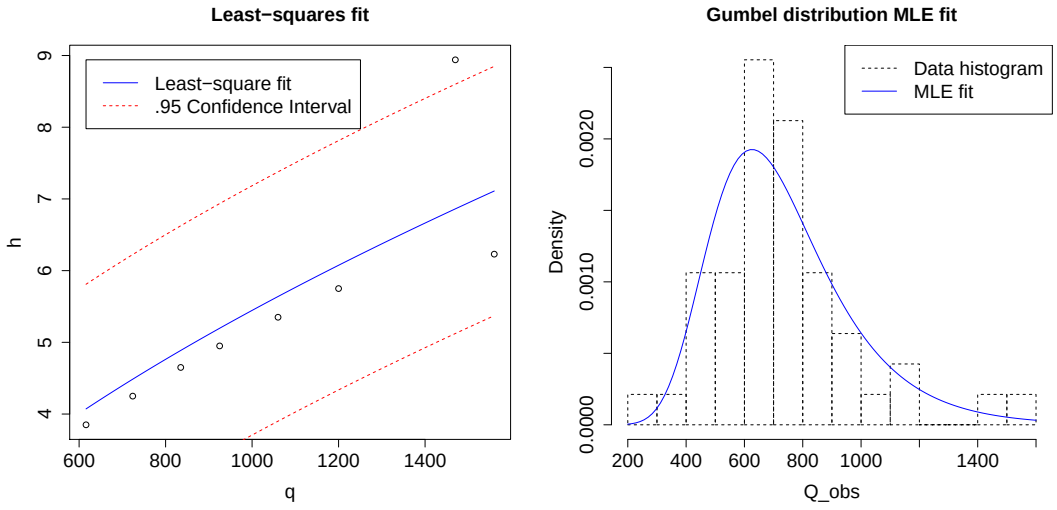

FIGURE 2. *Gauche : Ajustement par moindre carrés du modèle (2) à $m = 8$ couples de mesures débit/hauteur. Droite : estimation par maximum de vraisemblance du modèle de Gumbel à partir de $n = 47$ mesures de débits maximaux annuels.*

(substitution) :

$$\widehat{\Phi} = \Phi(\widehat{\Theta}). \tag{10}$$

Il est donc possible de contourner le problème de l'incertitude sur l'inconnnue Θ en lui attribuant une valeur estimée, que l'on espère proche de la vraie valeur, et en effectuant tous les calculs comme si l'on était en situation d'information parfaite.

Exemple. Nous avons mis en œuvre l'approche "plug-in" sur un jeu de données composé de :

- $n = 47$ mesures $(q_i)_{1\leq i\leq n}$ de débits maximaux annuels effectués sur un cours d'eau ;
- $m = 8$ couples $(\tilde{q}_j, \tilde{y}_j)_{1\leq j\leq m}$ de mesures débit/hauteur effectués sur le même site.

Nous avons supposé que les valeurs de débit q_i et $\tilde{q}_j$ étaient observées sans erreurs, mais que les hauteurs d'eau $\tilde{y}_j$ étaient mesurées avec une erreur additive, modélisée par un bruit blanc gaussien, d'écart-type σ inconnu. Nous avons alors ajusté le modèle (2) par moindre carrés ordinaires (MCO) aux couples $(\tilde{q}_j, \tilde{y}_j)$, et estimé le modèle de Gumbel (3) par maximum de vraisemblance (MV) sur la base des q_i. On obtient ainsi les estimations suivantes (voir Figure 2) :

- Coefficient de Strickler :

$$\widehat{K_s}^{\text{MCO}} = 59.33;$$

- Paramètre de localisation de la loi des débits :

$$\widehat{\mu}^{\text{MV}} = 626.14;$$

- Paramètre d'échelle inverse de la loi des débits :

$$\widehat{\rho}^{\text{MV}} = 5.24 \times 10^{-3}.$$

 ISSN: 2102-6238

Par "plug-in" de ces valeurs dans les équations (4), (9) et (8), on obtient alors des estimations ponctuelles des quantités d'intérêt suivantes :

- Probabilité de dépassement d'une digue de hauteur $h = 7.5\,m$:

$$\widehat{\mathbb{P}}(Z_c \geq Z_v + h) = 3.52 \times 10^{-3};$$

- Crue centennale :

$$\widehat{q_{99\%}} = Z_v + 6.96\,m;$$

- Hauteur optimale d'une digue, calculée relativement au coût (8), pour $I_0/C_0 = 1/1\,000$:

$$\widehat{h_{opt}} = 8.18\,m.$$

Limites de l'approche "plug-in". L'approche décrite ci-dessus permet de calculer relativement simplement des estimateurs de n'importe quelle quantité d'intérêt voulue, et donne des résultats facilement interprétables. Elle est de plus rigoureuse si l'on dispose d'un nombre suffisant de données, pour peu que l'on utilise des estimateurs qui convergent presque sûrement (voir à ce sujet [15]). Cela revient à dire que, lorsque le nombre d'observations tend vers l'infini, on se rapproche de la situation d'information parfaite.

En dehors de cette situation idéale cependant, l'erreur commise en estimant Θ, et donc Φ, ne peut être négligée. Se posent alors plusieurs questions :

- Comment quantifier l'erreur d'estimation, *i.e.* l'écart entre la vraie valeur Θ et son estimateur, que nous noterons conventionnellement $\widehat{\Theta} - \Theta$, ainsi que l'erreur résultante $\widehat{\Phi} - \Phi$ sur la quantité d'intérêt ?
- Quelles sont les conséquences de ces erreurs sur la décision éventuelle que doit guider l'étude ?

La réponse à la première question est connue mais rarement menée jusqu'à son terme par les ingénieurs : les outils de la statistique classique tels que bootstrap [10], delta-method [24], etc., permettent de construire des intervalles de confiance, asymptotiques ou non.

Les réponses à la seconde question, on s'en doute, sont multiples. Elles se distinguent avant tout par le degré d'implication de l'analyste vis-à-vis des enjeux décisionnels qui ont motivé l'étude. Dans la suite, nous tentons de classer les principales approches utilisées suivant ce degré d'accès de l'analyste statisticien à la formulation des enjeux décisionnels.

6. Procédures de Bayes

Dès lors que la définition de Φ intègre une formalisation des enjeux décisionnels sous-jacents à l'étude sous la forme d'une fonction de coût (5), la théorie de l'estimation bayésienne, développée par exemple dans [17], offre un cadre cohérent permettant de répondre aux deux questions posées ci-dessus. Étant donnée une loi $\pi(\Theta)$ sur les inconnues du modèle, la décision optimale au sens de Bayes correspond au minimum du coût intégré :

$$\widehat{\Phi}^{Bayes} = \arg\min_d \int c(d;\Theta)\pi(\Theta|D)\mathrm{d}\Theta, \tag{11}$$

 ISSN: 2102-6238

où $\pi(\Theta|D) \propto \mathscr{L}(D|\Theta)\pi(\Theta)$ donne formellement la loi de l'inconnue Θ conditionnellement aux données, $\mathscr{L}(D|\Theta)$ étant la vraisemblance de celles-ci. Pour traduire cette équation en mots, (11) réalise une analyse de sensibilité où chaque coût potentiel $c(d;\Theta)$ (associé à une décision sous-optimale par méconnaissance de l'état de la nature) est pondéré par la probabilité de l'état de la nature $\pi(\Theta|D)$, évaluée conditionnellement aux données dont dispose l'analyste.

D'un point de vue strictement fréquentiste, de tels estimateurs bénificient d'excellentes propriétés, notamment celle de constituer, avec leurs limites, l'ensemble des règles de décision dites *admissibles*, ou *non dominées* (voir par exemple [18]).

Deux courants se distinguent alors :

- Les statisticiens "classiques" à la suite de [27] reconnaissent que cette façon de procéder par pondération $\pi(\Theta)$ fournit simplement une construction mathématique d'estimateurs intéressants sans vouloir donner un sens à la distribution $\pi(.)$ sinon celui de la commodité calculatoire ;
- Les statisticiens bayésiens font un pas de plus en matière d'interprétation. Ils munissent aussi l'espace des modèles d'une structure de probabilité, mais contrairement aux précédents, ils donnent à $\pi(\Theta)$ le sens d'un pari probabiliste quant aux valeurs plausibles des inconnues avant considération des résultats expérimentaux (loi *a priori*).

Exemple. Poursuivant l'exemple hydraulique dans un cadre bayésien, nous avons utilisé les lois *a priori* "faiblement informatives" suivantes sur les paramètres $\Theta = (K_s, \mu, \rho)$ du modèle, ainsi que sur l'écart-type σ des erreurs de mesure :

- Coefficient de Strickler : $K_s \sim \mathscr{U}([10;100])$;
- Précision des erreurs de mesures : $1/\sigma^2 \sim \mathscr{E}(1)$;
- Paramètre de localisation de la loi des débits : $\mu \sim \mathscr{G}(1;500)$;
- Paramètre d'échelle de la loi des débits : $1/\rho \sim \mathscr{G}(1;200)$,

où $\mathscr{U}([a;b])$ est la loi uniforme sur l'intervalle $[a;b]$, $\mathscr{E}(\lambda)$ est la loi exponentielle de taux λ et $\mathscr{G}(\alpha;\beta)$ est la distribution Gamma de paramètre de forme α et d'échelle inverse β. La loi jointe *a posteriori* (conditionnellement aux données) de ces paramètres n'est pas explicite, en revanche il est possible d'en simuler un échantillon par diverses méthodes [19]. Nous avons choisi pour ce faire la méthode d'acceptation-rejet (décrite dans l'ouvrage ci-dessus notamment). Notons que, à la différence d'une estimation ponctuelle, l'approche bayésienne fournit pour le vecteur des paramètres une distribution complète de valeurs possibles, comme l'illustre la Figure 3.

A partir d'un échantillon de la loi *a posteriori* de Θ, il devient alors facile de résoudre numériquement le problème d'optimisation (11) en se donnant une grille de valeurs possibles pour d et en choisissant celle pour laquelle le coût intégré, évalué par Monte-Carlo, est minimal. Nous avons appliqué cette procédure au problème de calcul d'une hauteur optimale de digue, calculée relativement au coût (8), pour un rapport investissement / coût de dommage pris égal à $I_0/C_0 = 1/1\,000$. On trouve alors :

$$\widehat{h_{opt}}^{Bayes} = 10.76\,m.$$

Ce résultat est à comparer à celui obtenu par "plug-in" dans la Section précédente : $\widehat{h_{opt}} = 8.18\,m$. De manière informelle, la prise en compte des incertitudes épistémiques sur les paramètre du modèles, se traduisant en incertitudes sur les coûts de dommage potentiels, entraîne dans le cas

 ISSN: 2102-6238

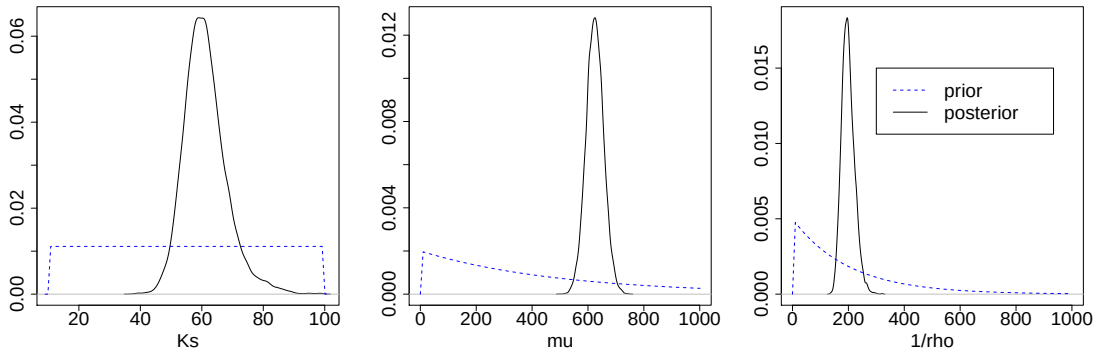

FIGURE 3. *Distributions marginales* a priori *(pointillés) et* a posteriori *(trait) des paramètres du modèle hydraulique, approchées par un estimateur à noyau à partir d'un échantillon de* 10000 *valeurs. De gauche à droite :* K_s, μ, $1/\rho$.

présent un relèvement de la hauteur optimale de digue d'environ 2.5 m, et mène donc à une décision plus "prudente". Une telle observation se retrouve dans [4] pour une fonction de coût et un modèle similaires. Notons toutefois que nous avons considéré ici une fonction de coût très dissymétrique, et pénalisant énormément les erreurs de sous-estimations, d'où l'écart important (voire même exagéré) avec l'estimateur plug-in.

7. Estimation ponctuelle basée sur la densité prédictive

Peut-on proposer une estimation ponctuelle "raisonnable" de Φ dans un cadre bayésien, c'est-à-dire lorsqu'on décrit l'incertitude sur Θ par une distribution de probabilité *a priori* $\pi(\Theta)$, mais en l'absence d'une fonction de coût ?

Tout dépend bien sûr de ce que l'on entend par "raisonnable". Diverses heuristiques ont été proposées pour répondre à cette question, notamment à travers l'emploi de caractéristiques de localisation de la loi *a posteriori* de Φ, telles que le mode [2], la moyenne ou encore la médiane [3]. Notons que ces deux dernières propositions reviennent à utiliser l'estimateur de Bayes (11) pour la fonction de coût quadratique $c(d;\Theta) = (d - \Phi(\Theta))^2$, ou valeur absolue $c(d;\Theta) = |d - \Phi(\Theta)|$, respectivement.

7.1. Densité prédictive

Une autre heuristique, proposée dans [12], est basée sur la *distribution prédictive* de la sortie Y, qui est sa distribution conditionnellement aux données D. La fonction de répartition de cette distribution, que nous noterons $\widehat{H}^{pred}(t)$, n'est autre que la moyenne *a posteriori* de $H(t|\Theta)$:

$$\begin{aligned} \widehat{H}^{pred}(t) &= \mathbb{E}[H(t|\Theta)\,|D] \\ &= \int H(t|\Theta)\pi(\Theta|D)\,\mathrm{d}\Theta. \end{aligned} \tag{12}$$

$\widehat{H}^{pred}(t)$ s'interprète comme le pari qu'engagerait l'analyste concernant les valeurs possibles de Y, mis à jour au vu des données D. Mathématiquement, cette probabilité épistémique peut également

 ISSN: 2102-6238

s'écrire comme l'estimateur de Bayes de $H(t|\Theta)$ pour la fonction de coût quadratique. Notons que dans aucun cas elle ne décrit la distribution de la sortie d'un système physique observable. Pour s'en convaincre, il suffit de considérer p copies indépendantes $(\mathrm{Y}_1, \ldots, \mathrm{Y}_p)$ de la sortie Y. Leur fonction de répartition prédictive jointe s'écrit :

$$\begin{aligned}\widehat{H}^{pred}(t_1, \ldots, t_p) &= \mathbb{P}[\mathrm{Y}_1 < t_1, \ldots, \mathrm{Y}_p < t_p | D] \\ &= \int \left\{ \Pi_{i=1}^{p} H(t_i|\Theta) \right\} \pi(\Theta|D)\, \mathrm{d}\Theta.\end{aligned}$$

Cette expression n'étant pas factorisable, les distributions prédictives des Y_i ne sont pas indépendantes, ce qui traduit le fait que l'on ne peut pas porter un jugement probabiliste prédictif sur le couple $(\mathrm{Y}_1, \mathrm{Y}_2)$ en considérant Y_1 puis Y_2 indépendamment. Pourtant, leurs distributions d'échantillonnage sont, elles, indépendantes (par définition), ce qui s'interprète en disant que, lorsqu'on connaît l'état de la nature Θ, les réplicats Y_1 et Y_2 sont indépendants.

Il est donc important de bien garder à l'esprit que la distribution prédictive est un outil mathématique permettant de coder l'incertitude sur les valeurs de la sortie Y du système, du point de vue épistémique d'un analyste, et non une description de la variabilité effective du système physique étudié. Ainsi, comme l'incertitude sur les Y_i contient une part qui leur est commune, à savoir la méconnaissance de Θ, il est parfaitement normal que les Y_i ne soient pas indépendants d'un point de vue prédictif. Ils le sont cependant, conditionnellement à l'état de la nature, quel qu'il soit.

7.2. Estimateur prédictif

Dans [12], il est proposé d'estimer $\Phi(\Theta)$, vue comme une fonction $\Phi(H(\cdot|\Theta))$ de la distribution $H(\cdot|\Theta)$, en remplaçant cette dernière, inconnue, par la prédictive $\widehat{H}^{pred}(\cdot)$:

$$\widehat{\Phi}^{pred} = \Phi(\widehat{H}^{pred}(\cdot)). \tag{13}$$

Il s'agit donc d'une variante de l'estimation "plug-in" (10), où la substitution s'effectue sur la distribution $H(\cdot|\Theta)$ de la sortie plutôt que sur l'état de nature Θ. Cette utilisation de la densité prédictive à des fins d'estimation est également suggérée dans [8], où le calcul de la densité prédictive (12) est interprété comme une "propagation d'incertitude" de Θ à Y, les incertitudes épistémiques et par nature étant de fait traitées de la même manière.

Notons cependant que, à la différence de l'estimateur "plug-in", $\widehat{\Phi}^{pred}$ ne bénéficie d'aucune interprétation physique, puisqu'il ne caractérise la sortie d'aucun système physique observable. De ce point de vue, la définition (13) peut apparaître comme incohérente en termes d'interprétation, puisqu'elle applique la fonction Φ, censée donner une caractéristique à fort enjeu de la distribution de sortie d'un système physique, à une distribution qui a perdu toute interprétation phénoménologique.

Quelles sont les conséquences du choix d'une telle heuristique au plan décisionnel ? Dans [6], un premier élément de justification théorique à l'estimateur défini par (13) est apporté, en montrant que, pour toute quantité d'intérêt pouvant s'écrire : $\Phi(\Theta) = \mathbb{E}[g(\mathrm{Y})|\Theta]$, $\widehat{\Phi}^{pred}$ coïncide avec la moyenne *a posteriori* de $\Phi(\Theta)$, soit l'estimateur de Bayes pour le coût quadratique $c(d;\Theta) = (d - \Phi(\Theta))^2$. On peut généraliser ce résultat en montrant que, pour une large classe de fonctions d'intérêt, $\widehat{\Phi}^{pred}$ coïncide également avec un estimateur de Bayes :

 ISSN: 2102-6238

Théorème 1. ***(Estimation basée sur la prédictive)*** *Soit* $\Phi(H(\cdot|\Theta))$ *une fonction d'intérêt de la distribution de* Y. *On suppose que :*

1. *En tant que fonction opérant sur les fonctions de répartitions* $H(\cdot)$, Φ *peut-être définie comme minimisant une certaine fonction de coût :*

$$\Phi(H(\cdot)) = \arg\min_d \; c(d; H(\cdot)).$$

2. *La fonction de coût* $c(d; H(\cdot))$ *est* linéaire *en* $H(\cdot)$, *c'est-à-dire que*

$$\sum_j p_j \times c(d; H_j(\cdot)) = c\Big(d; \Big\{\sum_j p_j \times H_j(\cdot)\Big\}\Big),$$

pour toutes fonctions de répartitions $\{H_j(\cdot)\}$ *et tous poids* $\{p_j\}$ *tels que :* $p_j \geq 0$ *et* $\sum_j p_j = 1$.

Alors, l'estimateur $\widehat{\Phi}^{pred}$ *défini par (13) coïncide avec l'estimateur de Bayes relatif au coût* $c(d; H(\cdot|\Theta))$.

Preuve. Par définition (13) de $\widehat{\Phi}^{pred}$, par la première hypothèse du théorème, puis par la définition (12) de $\widehat{H}^{pred}(\cdot)$ on a :

$$\begin{aligned}\widehat{\Phi}^{pred} &= \Phi(\widehat{H}^{pred}(\cdot)) \\ &= \arg\min_d \; c(d; \widehat{H}^{pred}(\cdot)) \\ &= \arg\min_d \; c(d; \mathbb{E}[H(\cdot|\Theta)|D]).\end{aligned}$$

La deuxième hypothèse du théorème (linéarité du coût) entraîne alors que :

$$\begin{aligned}\widehat{\Phi}^{pred} &= \arg\min_d \; \mathbb{E}[c(d; H(\cdot|\Theta))|D] \\ &= \arg\min_d \; \int c(d; H(\cdot|\Theta))\pi(\Theta|D)d\Theta.\end{aligned}$$

Cette dernière expression coïncide bien avec la définition (11) de l'estimateur de Bayes, relatif au coût $c(d; H(\cdot|\Theta))$ □

Les conditions d'application de ce théorème sont très générales ; en particulier, nous montrons dans les exemples ci-dessous qu'il s'applique à toutes les quantités d'intérêt définies en Section 3, et qui sont les plus utilisées en pratique dans les études industrielles.

L'interprétation de ce résultat est claire : utiliser l'estimateur $\widehat{\Phi}^{pred}$ revient implicitement à employer l'estimateur de Bayes relatif à une fonction de coût entièrement déterminée par l'expression de la quantité d'intérêt. Ces conclusions nous paraissent en contradiction flagrante avec le but premier de la méthode, qui est de proposer une estimation ponctuelle de Φ en l'absence d'une fonction de coût. Ainsi, en pratique, la méconnaissance volontaire des propriétés théoriques de l'estimateur $\widehat{\Phi}^{pred}$ signifie que son utilisateur délègue le choix d'une fonction de coût à l'heuristique qu'il a adoptée.

 ISSN: 2102-6238

Exemples. Continuant l'analyse bayésienne commencée en Section 6, nous avons appliqué le résultat ci-dessus à l'estimation basée sur la densité prédictive des quantités d'intérêt suivantes :

- La probabilité de débordement $\mathbb{P}(Z_c \geq Z_v + h|\Theta)$ d'une digue de hauteur $h = 4.5\,m$: cette probabilité peut s'écrire en fonction de la distribution de Z_c sous la forme

$$\mathbb{P}(Z_c \geq Z_v + h|\Theta) = 1 - H(Z_v + h|\Theta).$$

L'estimateur prédictif de cette quantité d'intérêt est donc la "probabilité prédictive de débordement", égale à :

$$\begin{aligned}\widehat{\mathbb{P}}^{pred}(Z_c \geq Z_v + h) &= 1 - \widehat{H}^{pred}(Z_v + h) \\ &= 5.9 \times 10^{-3}.\end{aligned}$$

On remarque que cette probabilité prédictive n'est en fait pas autre chose que l'espérance *a posteriori* de la probabilité sachant Θ, soit encore l'estimateur bayésien relatif à la fonction de coût quadratique : $c(d; H(\cdot|\Theta)) = d^2 - 2d(1 - H(Z_v + h|\Theta))$ (à laquelle on a retiré le terme $H(Z_v + h|\Theta)^2$, qui n'intervient pas dans l'optimisation). On retrouve donc les conclusions du théorème dans ce cas, puisque $\mathbb{P}(Z_c \geq Z_v + h|\Theta)$ minimise la fonction de coût $c(d; H(\cdot|\Theta))$ ci-dessus, et que celle-ci est bien linéaire en $H(\cdot|\Theta)$.
Notons que l'estimateur prédictif est recommandé dans [9] pour le calcul de fiabilité, justifiant ce choix sous l'hypothèse d'un coût de dommage constant. Cependant nous voyons qu'en tant qu'estimateur bayésien, il est fondé sur le coût quadratique, qui pénalise également la sur- et la sous-estimation du risque de défaillance. Dans des situations où la défaillance peut entraîner des coûts très importants, on pourra donc lui privilégier des fonctions de coûts dissymétrique par souci de conservatisme.

- La crue centennale $q_{99\%}$: celle-ci s'écrit en fonction de $H(\cdot|\Theta)$ sous la forme

$$q_{99\%} = H^{-1}(99\%|\Theta),$$

donc l'estimateur prédictif associé est la "crue centennale prédictive" :

$$\begin{aligned}\widehat{q_{99\%}}^{pred} &= (\widehat{H}^{pred})^{-1}(99\%) \\ &= 8.56\,m.\end{aligned}$$

Mais, comme rappelé en section 3, $q_{99\%}$ peut être défini comme minimisant le coût linéaire par morceaux : $c(d; H(\cdot|\Theta)) = \int_{-\infty}^{d} H(t|\Theta)dt - 99\%d$. Ce coût est en outre linéaire en $H(\cdot|\Theta)$, de sorte que, par application du théorème 1, $\widehat{q_{99\%}}^{pred}$ n'est autre que l'estimateur bayésien relatif à $c(d; H(\cdot|\Theta))$.
Or, nous avons vu que ce coût peut s'interpréter comme étant celui résultant de la construction d'une digue de protection fluviale, pour un rapport coût d'investissement/coût de dommage égal à : $I_0/C_0 = 1/99$, et pour un coût de dommage linéaire en la hauteur de crue, plutôt que quadratique (voir [17] pour plus de détails). Ainsi, on peut interpréter ce résultat en disant que la conjonction d'une norme de sûreté (la digue doit pouvoir résister à la crue centennale) et d'une heuristique d'estimation (l'estimateur prédictif) mène à adopter implicitement une certaine évaluation des coûts de dommages (linéaires), qui peut-être très éloignée des coûts réels (si ceux-ci sont quadratiques par exemple).

 ISSN: 2102-6238

- La hauteur optimale de digue h_{opt} minimisant le coût (8) : ce coût est bien linéaire en $H(\cdot|\Theta)$ (par linéarité de l'espérance d'une variable aléatoire relativement à sa loi). Donc par application du théorème 1, $\widehat{h_{opt}}^{pred}$ coïncide avec l'estimateur bayésien relatif à cette même fonction de coût, soit l'estimateur calculé dans la section précédente :

$$\widehat{h_{opt}}^{pred} = \widehat{h_{opt}}^{Bayes} = 10.76m.$$

8. Approches descriptives

Lorsque l'analyste ne dispose pas de fonction de coût explicite, la responsabilité lui est souvent laissée d'en proposer une, reflétant sa connaissance des enjeux décisionnels.

À défaut de pouvoir quantifier cette fonction de coût précisément, et sous l'hypothèse qu'elle est deux-fois dérivable en d autour de $d = \Phi(\Theta)$, il pourra se contenter de la fonction de coût quadratique (7), que l'on peut voir comme résultant d'un développement limité au deuxième ordre de la "vraie" fonction de coût $c(d,\Theta)$, dont la forme exacte est inconnue, en d autour de $d = \Phi(\Theta)$. Ce choix équivaut donc à opter pour une estimation par moindres carrés.

Cependant, il se peut que l'analyste ne dispose pas d'éléments suffisants pour justifier ce choix, ou bien estime qu'une telle approche n'est pas appropriée. C'est le cas notamment lorsque l'étude n'a pas pour objectif premier de guider une décision, mais plutôt de vérifier le respect d'une norme réglementaire, que l'on peut formaliser sans perte de généralité par le respect d'une borne supérieure sur Φ :

$$\Phi < M. \tag{14}$$

Les études de fiabilité structurelle notamment visent le plus souvent à vérifier qu'une probabilité de défaillance, (ou, de manière équivalente, un quantile d'ordre donné) est en deçà d'une certaine limite réglementaire. Les exemples de ce type d'études abondent, avec des domaines d'application aussi variés que la sûreté d'un site de stockage de déchets nucléaires [13], le risque de rupture des réacteurs d'une navette spatiale [16] ou l'impact sur un habitat naturel d'un polluant [26], pour n'en citer que quelques-uns.

Etant donné l'incertitude épistémique sur la valeur de Φ, il est impossible de déterminer de manière certaine si la condition (14) est vérifiée. Le fait de fournir une réponse basée sur une estimation ponctuelle de Φ reviendrait dans ce cas pour l'analyste à effectuer un pari sur le respect de la norme, basé le cas échéant sur une fonction de coût traduisant les enjeux d'un tel pari. S'il refuse de s'engager dans cette voie-là, il doit alors au moins quantifier l'incertitude sur Φ, sous la forme :

- soit d'un intervalle de confiance contenant la vraie valeur de Φ dans approximativement 95% des cas (par exemple), s'il adopte un point de vue fréquentiste ;
- soit de la distribution *a posteriori* complète de Φ, assortie d'indicateurs tels que la probabilité *a posteriori* épistémique de respect de la condition (14), s'il adopte un point de vue bayésien.

Enfin, notons que certaines normes incluent explicitement de telles descriptions de l'incertitude sur les grandeurs considérées. C'est le cas en écotoxicologie par exemple, où la dose maximale autorisée de polluant dans un habitat, notée $HC_{5,5}$, est définie comme la borne inférieure d'un intervalle de confiance symétrique à 90 % sur la dose de polluant minimale affectant 5 % au moins des espèces vivant dans le milieu considéré (voir [26] pour une définition plus précise).

 ISSN: 2102-6238

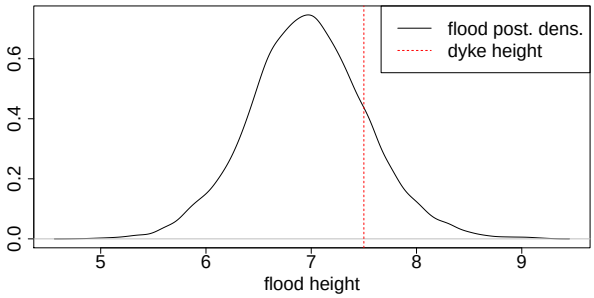

FIGURE 4. *Densité* a posteriori *du quantile* $q_{99\%}$ *(moins la cote* Z_v *de fond de la rivière) approchée par un estimateur à noyau à partir d'un échantillon de* 10000 *valeurs. La droite verticale correspond à la hauteur de la digue.*

Exemple. On se donne pour but de vérifier que la crue centennale $q_{99\%}$ est inférieure à l'altitude d'une digue de hauteur $h = 7.5\,m$, ce qui s'interprète en disant qu'elle ne sera débordée en moyenne qu'une fois par siècle. Rappelons que l'estimation "plug-in" par maximum de vraisemblance donne : $\widehat{q_{99\%}} = Z_v + 6.96\,m$, ce qui suggère que le risque de débordement est bien en deça de la limite souhaitée. Cependant, étant donnée l'incertitude sur la valeur réelle de $q_{99\%}$, on peut légitimement se demander dans quelle mesure cette première réponse est fiable.

Continuant l'analyse bayésienne commencée dans les Sections précédentes, nous avons calculé la distribution *a posteriori* du quantile $q_{99\%}$ (voir Figure 4). Ce quantile dépasse la hauteur de la digue avec une probabilité *a posteriori* de :

$$\mathbb{P}(q_{99\%} > Z_v + 7.5 | D) = 0.17,$$

qui est loin d'être négligeable, suggérant que l'estimation par "plug-in" amène à sous-estimer le risque de débordement considéré. Au vu de ce résultat, le décideur pourra donc juger prudent d'envisager des travaux d'agrandissement de la digue afin de réduire ce risque.

Ce cas précis illustre le danger qu'il y a à ne considérer que des estimateurs ponctuels sans tenir compte de l'incertitude affectant les grandeurs d'intérêt considérées, et la nécessité pour l'analyste de fournir le maximum d'informations au commanditaire de l'étude afin de lui permettre une prise de décision la plus renseignée possible.

9. Discussion

En conclusion de cette réflexion sur la pratique de l'ingénieur sous incertitudes, il nous semble important de mettre en avant les points suivants, qui devraient éclairer l'analyste dans le choix d'une méthode.

Dans un premier temps, le critère le plus important est la quantité d'information disponible. En effet, en situation d'information parfaite ou presque, c'est-à-dire lorsqu'on dispose d'assez de données pour estimer les inconnues du modèle avec grande précision, on peut se contenter d'une estimation "plug-in" de la quantité d'intérêt, basée sur n'importe quelle méthode d'estimation ponctuelle garantissant une convergence presque sûre.

 ISSN: 2102-6238

Si l'on ne dispose que de peu de données, il est en revanche impératif de prendre en compte les incertitudes affectant les quantités d'intérêt à évaluer. Un deuxième critère est alors l'accès à une fonction de coût formalisant les enjeux décisionnels de l'étude, auquel cas le choix d'une approche bayésienne est à privilégier. Si le problème ne peut être formalisé à l'aide d'une fonction de coût, il demeure important de fournir au destinataire de l'étude une description détaillée de l'incertitude sur les quantités d'intérêt étudiées, ainsi que nous l'avons illustré dans le cas de l'estimation d'une hauteur de crue centennale.

Par ailleurs, par souci de simplicité, nous n'avons pas étudié complètement la seconde composante du paramètre Θ : l'incertitude porte aussi sur le code souvent complexe G. Quel poids accorder à son estimation au vu des données disponibles ? Dans notre exemple nous avons traduit cette incertitude par une erreur de mesure indépendante, mais les techniques statistiques avancées, par exemple par [22] ou [14], considérent G comme la réalisation d'un objet probabiliste complexe (processus stochastique, le plus souvent gaussien), ce qui permet d'introduire une structure plus réaliste des erreurs de modèle.

Enfin, une autre conclusion qui ressort de cette réflexion est la nécessité de séparer les modèles mathématiques de variable aléatoire utilisés lors d'une analyse d'incertitudes d'une interprétation événementielle purement physique, afin de conserver une interprétation cohérente tout au long de l'étude. Ces recommandations semblent aller de soi, cependant nous avons vu que le risque de mélanger les incertitudes épistémique et par nature est relativement courant. Une telle confusion peut mener à l'utilisation d'heuristiques aux conséquences mal maîtrisées, ainsi que nous l'avons montré dans le cas de l'estimation ponctuelle d'une quantité d'intérêt basée sur la densité prédictive.

Le risque de confusion entre les différents types d'incertitude apparaît le plus clairement dans l'emploi de la densité prédictive, mais il s'étend en fait à toute l'analyse bayésienne, ainsi que le note [11]. Ce dernier préconise l'utilisation de l'analyse par intervalles pour propager l'incertitude épistémique, combinée avec une modélisation probabiliste des incertitudes par nature. Une telle approche hybride présente l'avantage de séparer explicitement les deux sources d'indétermination, et évite la question du choix d'une distribution *a priori*.

Plus généralement, il existe d'autres représentations de l'incertitude épistémique que celle, probabiliste, que propose la théorie bayésienne, comme par exemple la logique floue, développée dans [30], ou la théorie de Dempster-Shafer, introduite par [23]. Cependant, de telles approches soulèvent d'autres problèmes difficiles. Notamment, la simplicité conceptuelle avec laquelle l'approche bayésienne permet de prendre en compte un contexte décisionnel via des fonctions de coût, ainsi que les propriétés d'optimalité des estimateurs de Bayes, n'ont pas, à notre connaissance, d'équivalents dans un cadre extra-probabiliste.

Il nous semble donc que, si le choix d'un "bon" outil pour représenter l'incertitude épistémique soulève encore nombre de débats dans la communauté des ingénieurs, la piste de la théorie statistique de la décision, malgré son grand âge, offre des garanties de cohérence toujours d'actualité :

- sur le plan théorique, grâce aux fondements axiomatiques de l'approche bayésienne,
- sur le plan pratique grâce à la théorie du calcul des probabilités.

 ISSN: 2102-6238

Remerciements.

Nous remercions Pietro Bernardara (EDF R&D) d'avoir mis à notre disposition le jeu de données qui nous a permis d'illustrer notre propos, ainsi que Nicolas Bousquet (EDF R&D) pour sa relecture détaillée et pour ses nombreuses remarques et suggestions. Ce travail a été partiellement financé par le Ministère de l'Economie, des Finances et de l'Industrie (DGCIS) dans le cadre du projet CSDL (Complex Systems Design Lab) du Pôle de Compétitivité System@tic Paris-Région. Cette réflexion a également été aiguisée par de nombreuses notes internes EDF et communications personnelles de Jacques Bernier, ingénieur de recherche EDF à la retraite.

Références

[1] T. AVEN : *Foundations of Risk Analysis*. Wiley, 2003.

[2] V. BARNETT : *Comparative Statistical Inference*. New York : John Wiley, 2 édition, 1982.

[3] J. O. BERGER : *Statistical Decision Theory and Bayesian Analysis*. Springer, 2nd édition, 1985.

[4] J. BERNIER : Décisions et comportement des décideurs face au risque hydrologique / Decisions and attitude of decision makers facing hydrological risk. *Hydrological Sciences Journal*, 43(3):301–316, 2003.

[5] S. E. CHICK : Subjective probability and bayesian methodology. *In* S. G. HENDERSON et B. L. NELSON, éditeurs : *Simulation*, volume 13 de *Handbooks in Operations Research and Management Science*, chapitre 9, pages 225–257. Elsevier, 2006.

[6] R. CHRISTENSEN et M. D. HUFFMAN : Bayesian Point Estimation Using the Predictive Distribution. *The American Statistician*, 39(4):319–321, 1985.

[7] S. COLES : *An Introduction to Statistical Modelling of Extreme Values*. Springer-Verlag, 2001.

[8] E. de ROCQUIGNY : La maîtrise des incertitude dans un contexte industriel, Seconde partie : revue des méthodes de modélisation statistique, physique et numérique. *Journal de la SFdS*, 147(3):73–106, 2006.

[9] A. DER KIUREGHIAN : Analysis of structural reliability under parameter uncertainties. *Probabilistic Engineering Mechanics*, 23(4):351–358, 2008.

[10] B. EFRON et R. J. TIBSHIRANI : *An Introduction to the Bootstrap*. Chapman & Hall, New York, 1993.

[11] S. FERSON et L. GINZBURG : Different methods are needed to propagate ignorance and variability. *Reliability Engineering and System Safety*, 54:133–144, 1996.

[12] S. GEISSER : *Predictive inference : an introduction*. Chapman and Hall, 1993.

[13] J. C. HELTON : Probability, conditional probability and complementary cumulative distribution functions in performance assessment for radioactive waste disposal. *Reliability Engineering and System Safety*, 54(3):145–163, 1996.

[14] M. C. KENNEDY et A. O'HAGAN : Bayesian calibration of computer models. *Journal of the Royal Statistical Society, Series B, Methodological*, 63(3):425–464, 2001.

[15] E. L. LEHMANN et G. CASELLA : *Theory of Point Estimation*. Springer, 1998.

[16] Krzysztofowicz R. MARANZANO C. J. : Bayesian reanalysis of the challenger o-ring data. *Risk Analysis*, 28(4):1053–1067, 2008.

[17] E. PARENT et J. BERNIER : *Le raisonnement Bayésien* (in French). Springer, 2007.

[18] C. P. ROBERT : *The Bayesian choice : A Decision theoretic Motivation*. Springer, 2007.

[19] C. P. ROBERT et G. CASELLA : *Monte Carlo Statistical Methods*. Springer-Verlag, New York, 1998.

[20] G. RUBINO et B. TUFFIN : *Rare Event Simulation using Monte Carlo Methods*. John Wiley & Sons, Chichester, 2008.

[21] R. Y. RUBINSTEIN et D. P. KROESE : *Simulation and the Monte Carlo Method*. Wiley, New York, 1981.

[22] J. SACKS, W. J. WELCH, T. J. MITCHELL et H. P. WYNN : Design and Analysis of Computer Experiments. *Statistical Science*, 4(4):409–423, 1989.

[23] G. SHAFER : *A Mathematical Theory of Evidence*. Princeton University Press, 1976.

[24] J. ULMO et J. BERNIER : *Éléments de décision statistique* (in French). PUF, 1973.

 ISSN: 2102-6238

[25] A. W. van der VAART : *Asymptotic Statistics (Cambridge Series in Statistical and Probabilistic Mathematics)*. Cambridge University Press, June 2000.

[26] F.A.M. VERDONCK : Determining environmental standards using bootstrapping, bayesian and maximum likelihood techniques : a comparative study. *Analytica Chimica Acta*, 446(1–2):429–438, 2001.

[27] A. WALD : Contributions to the Theory of Statistical Estimation and Testing Hypotheses. *The Annals of Mathematical Statistics*, 10(4):299–326, 1939.

[28] L. WASSERMAN : *All of Statistics : A Concise Course in Statistical Inference*. Springer-Verlag, 2004.

[29] L. WASSERMAN : *All of Nonparametric Statistics*. Springer-Verlag, 2006.

[30] L. A. ZADEH : Toward a generalized theory of uncertainty (GTU)–an outline. *Inf. Sci*, 172(1–2):1–40, 2005.

 ISSN: 2102-6238