

TENIR COMPTE DE L'HISTOIRE DANS LES JUSTIFICATIONS DE STABILITE : L'APPROCHE BAYESIENNE

Taking history into account in the justifications for stability: the bayesian approach

Luc DEROO, Michel LINO, Sadri MEVEL

ISL

Jean-Jacques FRY

EDF-CIH

MOTS CLEFS

Diagnostic. Barrages. Approche bayésienne. Jugement de l'expert. Erosion interne.

KEY WORDS

Safety assessment. Dams. Bayesian approach. Expert judgment. Internal erosion.

RÉSUMÉ

En France, la majorité des barrages et digues en service sont des ouvrages anciens. L'ancienneté des barrages est souvent une difficulté : la mémoire précise de leur constitution a pu se perdre, les archives ont pu s'altérer, les normes de construction et de justification sont obsolètes. La méconnaissance de la constitution ou l'inadéquation de la constitution aux standards actuels peuvent faire douter de la sécurité. Mais l'ancienneté est aussi un gage de confiance : les barrages anciens ont subi de nombreuses épreuves de chargement, qui sont autant de tests de robustesse ; leur comportement à l'occasion de ces épreuves est un enseignement précieux quant à leur capacité à résister à de futures sollicitations.

Dans tout diagnostic de sécurité d'un barrage en service, l'analyse du comportement historique est une composante essentielle. Cependant, cette analyse est avant tout qualitative. Peut-on trouver le moyen de la quantifier? Cette question est explorée dans cette communication, en utilisant les outils de l'approche bayésienne. L'idée générale est la suivante :

- cherchant à poser un diagnostic concernant une pathologie donnée (exemple : érosion interne non stabilisée) ;

- l'ingénieur établit un jugement a priori, bâti à partir de sa connaissance de la constitution de l'ouvrage ;

- puis il considère les remplissages précédents comme autant de tests pour lesquels le comportement du barrage est caractérisé (résultats de la surveillance et de l'auscultation) ;

- et, tirant profit des statistiques du parc mondial et en utilisant l'approche Bayésienne, il exploite ces tests pour obtenir un jugement a posteriori.

La démarche est comparable à celle adoptée par le médecin qui utilise des tests médicaux pour améliorer son évaluation de la probabilité qu'un patient soit atteint d'une maladie donnée. On retrouve en effet la notion de probabilités a priori et a posteriori, et l'idée de tests calibrés sur des cohortes de patients ; la qualité des tests (sensibilité, spécificité, ratio de vraisemblance) est alors le paramètre clé qui permet de faire évoluer de manière quantifiée le diagnostic.

ABSTRACT

In France, the majority of dams and levees in service are old structures. The age of dams is often a difficulty: the precise memory of their constitution may have been lost, archives may have been altered, construction and justification standards are obsolete. Lack of knowledge of the constitution or the inadequacy of the constitution to current standards may raise doubts about safety. But seniority is also a guarantee of confidence: old dams have undergone numerous loading tests, which are all robustness tests; their behaviour during these tests is a valuable lesson as to their ability to withstand future loadings.

In any safety diagnosis of a dam in service, the analysis of historical behaviour is an essential component. However, this analysis is above all qualitative. Can we find a way to quantify it? This question is explored in this paper using the tools of the Bayesian approach. The general idea is as follows:

- seeking to diagnose a given pathology (example: unstabilized internal erosion) ;

- the engineer makes an a priori judgment, based on his knowledge of the structure's constitution ;

- then, he or she considers the previous fills as tests, for which the behaviour of the dam is characterized (results of monitoring and auscultation) ;

- and, taking advantage of world portfolio statistics and using the Bayesian approach, he or she uses these tests to obtain an a posteriori judgment.

The approach is similar to that adopted by a physician who uses medical tests to improve his or her assessment of the likelihood that a patient will have a particular disease. We find the notion of a priori and a posteriori probabilities, and the idea of calibrated tests on patient cohorts; the quality of the tests (sensitivity, specificity, likelihood ratio) is then the key parameter that allows the diagnosis to evolve in a quantified way.

1. INTRODUCTION

Dans le domaine des barrages, l'évaluation de la sûreté utilise régulièrement le « jugement de l'ingénieur », car les méthodes de calcul ne sont pas toujours disponibles et car, lorsqu'elles le sont, elles ne représentent parfois qu'imparfaitement la réalité du comportement de la structure.

Ce « jugement de l'ingénieur » s'appuie sur un double retour d'expérience :

- le retour d'expérience du parc : le comportement du parc mondial des barrages, ou celui, plus limité, des barrages qu'il a eu l'occasion de connaître ;
- confronté au retour d'expérience du comportement du barrage qu'il examine.

Ce jugement est généralement peu formalisé, et il peut être biaisé.

Cet article examine dans quelle mesure les méthodes bayésiennes pourraient limiter le biais du jugement expert. Cette démarche paraît légitime : car ces méthodes sont pratiquées dans un domaine qui utilise des démarches similaires : le diagnostic médical.

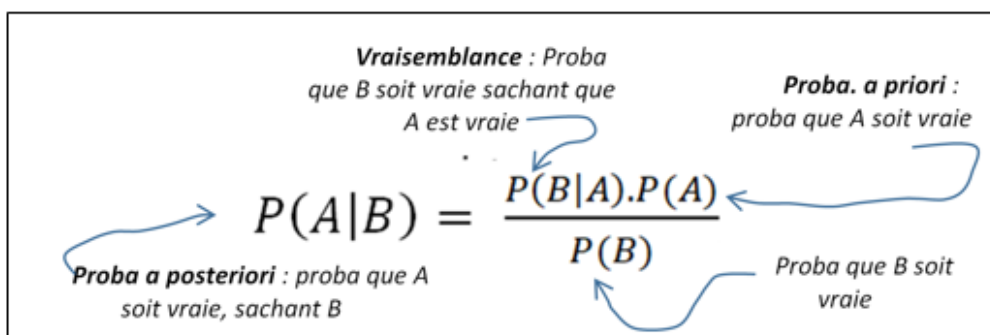
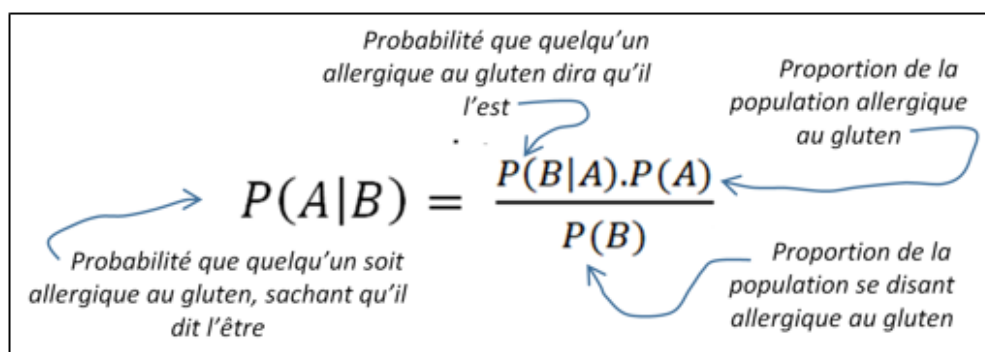
L'article présente successivement :

- §2, les concepts principaux de l'approche bayésienne et son utilisation dans le cadre des biostatistiques médicales ;
- §3, la transposition au domaine des barrages à retenue permanente pour le diagnostic à l'érosion interne ;
- §4, une note sur les limites d'applications, et les prolongements qu'il serait utile d'envisager.

2. CONCEPTS

2.1. Approche bayésienne, principe général

Cet article n'a pas pour ambition de donner un aperçu général de l'approche bayésienne. Nous nous contentons d'en donner une illustration. Le théorème de Bayes s'écrit : $P(A|B) = \frac{P(B|A).P(A)}{P(B)}$, où $P(A|B)$ est la probabilité conditionnelle de A, sachant B. On peut l'illustrer ainsi :



2.2. Utilisation en diagnostic médical, définitions

Dans le domaine médical, les statistiques bayésiennes permettent d'évaluer les performances d'un Test utilisé pour poser un diagnostic [1].

Dans le cas le plus simple, un test donne un résultat soit positif, soit négatif. Le test n'est cependant pas parfait : le test peut échouer à déceler la maladie (« faux négatif ») ou être positif chez un patient qui n'est pas atteint de la maladie (« faux positif »).

On note M l'événement "le patient est atteint de la maladie" et NM son complémentaire "le patient n'est pas atteint de la maladie", on note $T+$ l'événement "le test est positif". La grandeur recherchée est $p(M|T+)$, la probabilité que le patient soit malade sachant que le test est positif. Le Tableau 1 donne la liste des termes les plus utilisés.

Tableau 1 : Liste des termes les plus utilisés

Terme	Définition / Notation
Malade / Non malade	M : Malade ; NM : Non malade
Test	Test diagnostic (T), qui peut être positif ($T+$) ou négatif ($T-$)
VP : vrai positif	Malade (M) et Testé positif ($T+$)
VN : vrai négatif	Non malade (NM) et Testé négatif ($T-$)
FP : faux positif	Non malade (NM) et Testé positif ($T+$)
FN : faux négatif	Malade (M) et Testé négatif ($T-$)
Pv : prévalence	Par définition, la prévalence de la maladie est $p(M)$. On l'appelle également « taux de base » en statistique bayésienne.
o : Cote (ou Odds)	$o = P(M) / (P(NM))$: probabilité (Malade) / probabilité (Non Malade) $P(M) = o / (1+o)$
Sensibilité d'un test	Capacité du test à identifier un malade. $Se = P(T+ M) = P(M \text{ et } T+)/P(M)$. Mesure sur une population : $VP / (VP+FN)$
Spécificité	Capacité du test à ne pas alerter faussement les non malades $Sp = P(T- NM) = P(NM \text{ et } T-)/P(NM)$. Mesure sur une population : $VN / (VN+FP)$
Probabilité pré-test	Eventualité de la maladie évaluée par le médecin avant de connaître le résultat du test diagnostic : $P(M)$
Probabilité post-test	Eventualité de la maladie évaluée par le médecin après le résultat du test diagnostic : $P(M T+)$ si le test est positif
RV+ ou RV- : Ratios de vraisemblance	$RV+ = Se / (1-Sp)$; $RV- = (1-Se) / Sp$ Capacité des tests à faire évoluer la probabilité pré-test, de maladie ($RV+$) ou non-maladie ($RV-$). $RV+=1$: le test n'apporte aucune information ; $RV+>>1$: très bon test. $RV-=1$: le test n'apporte aucune information ; $RV-<<1$: très bon test.
Cote post-test	Cote post-test = odds pré-test * $RV+$ (si $T+$) ou cote post-test * $RV-$ (si $T-$) Si deux tests successifs positifs, on peut combiner : $o_{\text{post}} = o * RV1 * RV2$
Evidence	Evidence = $\log(P) - \log(1-P)$, avec logarithmes en base décimale (pour une probabilité de très basse probabilité 10^{-4} , l'évidence est -4 ; pour une probabilité de très haute probabilité, 0,9999, l'évidence est 4). Le décalage de 1 point d'évidence correspond au décalage d'une classe de probabilité dans la notation utilisée ci-dessous.

Le théorème de Bayes permet d'écrire $p(M|T+) = p(T+|M).p(M) / [p(T+|M).p(M) + p(T+|NM).(1-p(M))]$

Il est également possible d'écrire cela par : $p(M|T+) = Pv * Se / [Pv * Se + (1-Pv) * (1 - Sp)]$

2.3. Classes de probabilité

Tableau 2 : Classes de probabilités

Appréciation de la fiabilité de la barrière ou de l'occurrence d'un événement	Classe de fiabilité	Probabilité annuelle (Grille 1)	Probabilité d'un événement pendant la vie de l'ouvrage (*)	Evidence = Valeur numérique de la classe
Jamais jusqu'à présent mis en défaut, sur un très grand nombre d'installations	F	$P < 10^{-6}$	$P < 10^{-4}$	- 4
Mise en défaut ou événement possible mais extrêmement peu probable	E	$P < 10^{-5}$	$P < 1/1000$	- 3
Mise en défaut ou événement très improbable	D	$P < 10^{-4}$	$P < 1/100$	- 2
Mise en défaut ou événement improbable	C	$P < 1/1000$	$P < 1/10$	- 1
Mise en défaut ou événement probable	B	$P < 1/100$	$P < 0,6$	- 0,2
Mise en défaut ou événement courant	A	$1/100 \leq P$	$0,6 \leq P$	0

(*) la durée de vie étant conventionnellement prise égale à 100 ans.

Les classes de probabilités annuelles peuvent notamment être approchées à partir de :

- données statistiques, lorsque l'échantillon est disponible (exemple : Defra [2]) ;
- retour d'expérience mondial tel qu'apprécié par le jugement d'expert (Tableau 3).

Elles peuvent parfois être reliées à d'autres considérations ; par exemple, les coefficients partiels utilisés par les Eurocode sont rattachés à des probabilités de défaillance (cf. §5).

Tableau 3 : Utilisation du REX mondial pour établir les classes de probabilité

Événement possible mais extrêmement peu probable	Événement très improbable	Événement improbable	Événement probable	Événement courant
N'est pas impossible au vu des connaissances actuelles, mais non rencontré au niveau mondial sur un très grand nombre d'installations	S'est déjà produit dans ce secteur d'activités, mais a fait l'objet de mesures correctives réduisant significativement sa probabilité	S'est déjà produit dans ce secteur d'activités, sans que les éventuelles corrections apportées depuis apportent une garantie de réduction significative de probabilité	S'est produit et/ou peut se produire pendant la durée de vie de l'installation	S'est produit sur le site considéré et/ou peut se produire à plusieurs reprises pendant la durée de vie de l'installation

3. DANS LE DOMAINE DU DIAGNOSTIC DES BARRAGES

3.1. Une analogie naturelle

Il est tentant d'utiliser les approches bayésiennes dans le domaine des barrages, car les analogies sont frappantes !

Analyse médicale (source : [1])	Le domaine des barrages
<p>Test diagnostic : tout moyen permettant d'obtenir une information utile au médecin pour l'assister dans sa démarche diagnostic face à un patient</p> <p>Le test diagnostic peut être :</p> <ol style="list-style-type: none"> 1. Un examen paraclinique (ex : un dosage biologique) 2. Un examen d'imagerie médicale 3. Une épreuve fonctionnelle 4. Un signe physique révélé par l'examen clinique 5. Un signe fonctionnel rapporté par le patient 6. Une information obtenue par l'interrogatoire 	<p>Médecin = l'ingénieur Patient = l'ouvrage</p> <p>Le test diagnostic peut-être :</p> <ol style="list-style-type: none"> 1. Des mesures d'auscultation : piézométrie, débits, ... ; une mesure au laboratoire sur un prélèvement in-situ 2. Des reconnaissances géotechniques, géophysiques pour connaître l'état de la constitution interne 3. Une épreuve de mise en eau de l'ouvrage 4. Les observations des inspections en VTA ou ETC 5. Observations lors des tournées de l'exploitant 6. Le comportement lors d'un épisode passé

Avec les barrages, comme en médecine, il est impossible de tout connaître du fonctionnement interne du patient. Avec les barrages, comme en médecine, certaines maladies restent mal connues, et les remèdes incomplets. Avec les barrages, comme en médecine, le jugement humain intervient dans l'appréciation du risque.

3.2. Un exemple : la stabilité du talus aval du barrage de *Douteux-MaisSansSouci*.

3.2.1. Le barrage

Le barrage (fictif !) de *Douteux-MaisSansSouci* a été construit dans les années 1950. Il s'agit d'un barrage en remblai, de hauteur 19 m. Les archives de la construction ont été perdues. Des plans de projet laissent penser que le barrage est un ouvrage en argiles sableuses, avec un filtre-drain cheminé, et un tapis aval. La cote de retenue normale est à 17 m dans le référentiel du barrage. Le bassin versant est relativement modeste, et l'évacuateur, construit assez loin en rive droite, est de forte capacité : les Plus Hautes Eaux sont à 17,50 m. L'ensemble du barrage est fondé sur la frange d'altération de terrains marneux.

Le barrage est exploité *sans souci* depuis sa mise en eau en 1952. De 1957 à 1996, le propriétaire avait même pris la liberté de rehausser de 50 cm la cote d'exploitation du barrage en installant un batardeau rustique par-dessus l'évacuateur. L'auscultation est pratiquée par trois piézomètres dans le talus aval et un seuil de mesure des débits de percolation en fond de vallée ; les mesures faites tous les mois depuis 1982 montrent une nappe partiellement rabattue dans le talus aval, et des débits de fuite de l'ordre de 30 l/min, sans évolution. Le pied aval du barrage est humide (et l'a toujours été, d'après l'exploitant), ce qui tend à confirmer que la nappe est au-dessus du tapis aval.

La retenue du barrage dépasse 7 hm³ et il y a des enjeux à l'aval en cas de rupture. Ce contexte a motivé la réalisation d'une étude de stabilité récente, utilisant le référentiel du CFBR. Une campagne de reconnaissance a été conduite. Les reconnaissances ont confirmé la fondation sur les marnes altérées, mais n'ont pas pu mettre clairement en évidence le tapis, et ont été l'occasion de prélever des échantillons du remblai du barrage.

Les calculs de stabilité du barrage ont été conduits conformément aux Recommandations du CFBR, en considérant des valeurs raisonnablement conservatives pour les paramètres. La pente du talus aval du barrage est raide (2,2H/1V), et les résultats, sans surprise, ne sont pas bons. Avec un coefficient partiel de 1,25 sur les paramètres géotechniques, le coefficient de modèle est 0,93 sous RN ; avec 1,1, le coefficient de modèle est 1,03 aux PHE. La note de calcul conclut en rappelant la pente raide, le caractère homogène du barrage et la possible absence de tapis : l'état du barrage est qualifié de *douteux*.

3.2.2. La « maladie »

La « maladie » que l'on cherche à dépister est l'*instabilité du talus aval pour une cote de retenue inférieure ou égale aux PHE (hors séisme)*.

Cette maladie peut être plus précisément définie comme suit. Un barrage est « malade » si sa stabilité est mise en défaut pour une cote de retenue inférieure aux PHE : le barrage va se rompre à l'occasion d'une prochaine montée de la cote de retenue, pour une cote < PHE (sans qu'il y ait « vieillissement » significatif). Notons qu'un barrage « malade » peut en théorie avoir déjà résisté à une cote de retenue égale aux PHE sans se rompre, car les facteurs autres que la cote de retenue peuvent jouer un rôle (notamment les autres facteurs qui influencent le régime des pressions interstitielles dans la recharge aval ou en fondation).

3.2.3. L'approche bayésienne

Un barrage « malade » est un barrage qui, au cours de son existence, subit une rupture par instabilité mécanique en grand de son talus aval, en dehors de toute sollicitation sismique, et pour une cote de retenue < PHE. L'approche par le calcul fournit une information : le barrage a une probabilité assez élevée d'être malade : c'est la probabilité a priori, $p(M)$. L'information supplémentaire est que le barrage a subi une épreuve de mise en eau longue et réussi, c'est l'épreuve T, de mise en eau, avec un résultat -, négatif.

L'approche bayésienne s'écrit : $p(M|T-) = p(T-|M).p(M) / [p(T-|M).p(M) + p(T-|NM).(1-p(M))]$.

Dans ce calcul, les deux termes qui pèsent sont $p(M)$ et $P(T-|M)$. Une manière de les approcher est décrite ci-dessous. Il faut considérer les valeurs quantitatives avec la plus grande prudence ; la méthode n'est pas encore calibrée.

$P(M)$ est la probabilité qu'un barrage soit « malade », indépendamment de toute considération quant à son comportement lors du test de mise en eau. C'est la probabilité a priori, évaluée sur la base de la constitution du barrage, et pas de son comportement. Ici, une manière de l'approcher est d'utiliser le résultat des calculs de stabilité.

Le résultat du calcul de stabilité est mauvais. Certes, le coefficient de stabilité « global » est supérieur à 1 (1,25*0,93 sous RN, ce qui fait 1,16 ; 1,1*1,03 sous PHE, ce qui fait 1,13). Mais cette marge de sécurité est faible : par une approche du type de celle développée dans les Eurocodes ([3][4]), la probabilité de rupture est estimée à 5.10^{-3} par an (classe B), alors que si le coefficient de stabilité « global » avait été de 1,5, cette probabilité serait de 5.10^{-6} (classe E).

$P(T-|M)$ est la probabilité qu'un barrage « malade » (i.e. qui va connaître une rupture pour une cote < PHE) ait auparavant passé avec succès l'épreuve d'une mise en eau de longue durée à RN et PHE, sans désordre, avec piézométrie et débit de drainage stabilisés.

Première approche : l'avis d'expert, sur la base du Tableau 3. Dans le monde, les cas documentés de rupture ou d'accident sérieux sur les talus aval de grands barrages en remblai se sont produits au premier remplissage, et ne correspondent donc pas à ce cas (détails de cette analyse non présentés ici). Cela n'exclut pas des exceptions, non rapportées. On note dans la classe « événement très improbable », voire « événement improbable » (P pendant la vie de l'ouvrage entre 1/10 et 1/1000). Pour l'exercice présenté ici, on choisit de poser $P(T-|M) \sim 1\%$.

Deuxième approche : base de données constituées pour DEFRA sur le parc des barrages anglais, [2]. Ces données sont utilisées dans l'encadré ci-contre. Elles sont évidemment un peu fragiles, car peu nombreuses. Elles ne contredisent pas l'ordre de grandeur mentionné précédemment.

$P(T-|NM)$ est la probabilité qu'un barrage « non malade » subisse un chargement aux PHE dans les mêmes conditions que le barrage de *Douteux-MaisSansSouci* : avec une piézométrie et des débits stables, *sans souci* apparent. Ce que l'on cherche ici, c'est la proportion des barrages du parc qui sont *sans souci* apparent aux PHE, sachant qu'ils ne sont pas malades. Une borne inférieure de $P(T-|NM)$ est $P(T-)$, proportion des barrages du parc qui sont sans souci apparent aux PHE. Cette statistique n'est pas disponible, mais on peut provisoirement considérer que c'est le cas de la grande majorité des barrages. Le calcul est effectué avec une probabilité de 80% ; ce chiffre joue peu.

Calcul sur la base des données DEFRA [2]

Données :

- Nombre de « emergency drawdown » par « slope instability », entre 1975 et 2000 : 12 pour des barrages de moins de 5 ans (sur une population d'environ 300 barrages) ; 1 pour des barrages de plus de 5 ans (population 2107 barrages).

Hypothèses de travail :

- Sans « emergency drawdown », il y aurait eu rupture
- Faute de disposer d'une meilleure appréciation, on considère que : tous les barrages de plus de 5 ans ont subi un test de mise en eau complet, T.
- On postule que le comportement des 2107 barrages et celui des 300 nouveaux barrages (post 1975), vis-à-vis du test T, est comparable

Calculs :

- Nombre de barrages malades, rapporté à une population de 300 barrages : $NbM = 1/2107*300+12 = 12$
- Nombre de barrages malades ayant subi avec succès le test de mise en eau : 1 pour 2107, soit, pour une population de 300 barrages : $1/2107*300$
- $P(T-|M) = Nb(T-|M) / NbM = 1/2107*300 / 12 = 1,2\%$.

Le test est performant : $Se = 99\%$; $Sp = 80\%$; $RV = (1-Se)/Sp = 0,0125$.

Et la probabilité a posteriori est : $p(M|T-) = 0,01 * p(M) / [0,01 * p(M) + 0,8 * (1-p(M))] \sim 0,01 * p(M)$

Il y a décalage de deux points d'évidence. La confiance dans le barrage a été multipliée par 100. Le barrage de *Douteux-MaisSansSouci* est ainsi aussi fiable qu'un barrage en fin de construction (i.e. qui n'a pas encore passé l'épreuve de première mise en eau), de même nature mais avec des pentes ou des matériaux qui auraient donné un coefficient « global » de 1,4 par le calcul conforme aux Recommandations CFBR.

Rappelons ici que ce calcul est fragile : il dépend beaucoup de la probabilité $P(T-|M)$, et il y aurait lieu d'approfondir beaucoup l'évaluation de ce chiffre. Mais il suffit à démontrer ceci : à supposer que l'on parvienne à bien calibrer la valeur de $P(T-|M)$, alors la méthode apporte une contribution substantielle au jugement sur la sécurité d'un ouvrage. Notons que $P(T-|M)$ est une probabilité qui ne dépend pas du barrage de *Douteux-MaisSansSouci* : elle est établie à partir du comportement de l'ensemble du parc des barrages, de même type que *Douteux-MaisSansSouci*.

La conclusion (§3) apporte des commentaires sur l'interprétation de ces résultats.

4. LE CAS DE L'ÉROSION INTERNE

4.1. Un cas d'intérêt majeur

Le cas précédent présente un intérêt mais ne répond pas à une nécessité : en matière de stabilité du talus aval, on dispose de moyens permettant de procéder aux calculs de stabilité, et il n'y a pas vraiment *besoin* de nouvelles approches.

En matière d'érosion interne, au contraire, il y a un besoin : lorsqu'il n'y a pas de filtres et drains conformes, il y a toujours suspicion d'érosion, et la preuve de la résistance par le calcul est difficile. La détection de l'érosion interne par l'auscultation est aléatoire et par la surveillance demande que l'on ait de la chance en lien avec la dynamique du développement de l'érosion.

L'interprétation du comportement passé est souvent utilisé lors du jugement d'expert, comme facteur de confiance quand la première mise en eau s'est bien passée. Les mises en eau sont en effet des épreuves fonctionnelles très efficaces, à condition que ce premier remplissage soit assez long et assez haut pour bien éprouver l'ensemble des chemins d'érosion, car les accidents d'érosion se produisent d'abord au premier remplissage et plus rarement dans les 5 premières années. Mais aussi, a contrario, comme facteur d'inquiétude lorsque des incidents se sont produits pour des cotes inférieures à la cote des PHE.

Une appréciation quantitative de la valeur des tests du passé permet de progresser, en tentant de bien prendre en compte la valeur explicative de ces épreuves. En effet, toutes les épreuves du passé ne se valent pas, et toutes les « maladies » d'érosion interne ne sont pas aussi bien diagnostiquées par une épreuve donnée.

4.2. Une définition de la « maladie »

Le bulletin 164 de la CIGB propose une terminologie relative à l'érosion interne.

L'érosion interne peut se décliner sous ses quatre formes de base : érosion de conduit, érosion régressive, érosion par suffusion et érosion de contact – ainsi que par des combinaisons entre ces formes.

Dans tous les cas, le processus d'érosion interne couvre plusieurs phases : initiation, continuation (non-filtration), progression. Le mécanisme d'érosion interne, qui se met en place avec l'initiation, peut être auto-interrompu à n'importe laquelle de ces phases, et alors il n'y aura pas rupture. Nous nous intéressons ici aux cas d'érosion interne qui ne sont pas auto-interrompus et qui, sauf intervention humaine, iront à la rupture. Par simplification, nous les appelons cas d'Erosion Interne Non Stabilisée (EINS).

Les différents mécanismes d'érosion interne ont un autre point commun : celui de se produire le long d'un chemin de percolation. L'altitude de ce chemin a une importance, car elle permet de comprendre dans quelle mesure une épreuve de mise en eau passée a sollicité ce chemin.

La « maladie » que l'on examine ici est EINS-z : l'érosion interne non stabilisée provoquée par une cote de retenue inférieure ou égale à la cote z.

4.3. Approche bayésienne

4.3.1. Notations

On note ici :

- $P(\text{EINS-z})$, probabilité que se développe un mécanisme d'érosion interne non stabilisé ;
- TestZT : la cote « z » a été atteinte (ou pratiquement atteinte) pendant longtemps.

« Longtemps » signifie : un temps suffisamment long pour que l'érosion se soit développée, soit quelques heures à quelques jours pour l'érosion de conduit et pour l'érosion le long des ouvrages traversants ; quelques jours ou semaines pour l'érosion régressive ; quelques années (!) pour la suffusion et l'érosion de contact.

« pratiquement atteinte » signifie que l'on peut admettre une part d'extrapolation, notamment pour les mécanismes d'érosion régressive, suffusion et érosion de contact.

D'autres tests, moins performants mais plus fréquemment pratiqués peuvent être considérés :

- TestZ : la cote « z » a été atteinte (ou pratiquement atteinte) pendant suffisamment de temps pour que l'érosion ait normalement le temps de se développer, mais pas suffisamment longtemps pour que l'on soit placé dans les circonstances du TestZT ;
- TestT : la cote « z » n'a pas été atteinte. Mais une cote de retenue inférieure, déjà significative pour les chemins d'érosion les plus sensibles, a été atteinte pendant longtemps. Test valable uniquement le long d'un chemin déjà actif.

Par ailleurs, on évalue le comportement lors du Test par une Classe, exprimée ci-dessous :

Tableau 4 : Echelle des symptômes d'érosion

Echelle	Intitulé	Classe
0	Pas d'écoulement	F _c
1	Écoulement diffus	
2	Écoulement local	E _c
3	Écoulement évolutif	D _c
4	Preuve d'érosion	C _c
5	Forte érosion	B _c
6	Erosion majeure	
7	Rupture évitée par action volontaire	A _c
8	Rupture	

Et on note $P(E|ZTi) = P(EINS-z | TestZT = Classe_i)$ la Probabilité que se développe un mécanisme d'érosion interne non stabilisé, sachant que le TestZT a atteint la classe i (la classe i précisément, ni plus, ni moins).

L'approche bayésienne s'écrit de la manière suivante :

$$P(E|ZTi) = P(ZTi|E) * P(E) / (P(ZTi|E) * P(E) + P(ZTi|nE) * (1-P(E))) \sim P(ZTi|E)/P(ZTi|nE) * P(E).$$

$P(E)$ est le *prior* ou probabilité pré-test : probabilité que l'ouvrage soit affecté d'une érosion interne non stabilisée ; évaluation faite avant toute observation du comportement. $P(E)$ est la probabilité évaluée par l'approche analytique.

Note pour le cas de l'érosion de conduit : l'érosion de conduit laisse rarement des traces : quand elle est initiée, elle s'arrête rarement en route. Donc on peut avoir 0 masse d'érosion observée, puis une rupture soudaine. Ainsi, un mécanisme d'érosion ne prendra pas nécessairement le temps de parcourir toute la palette des symptômes.

4.3.2. Mise au point des tables de vraisemblance

Cette étape est le point central de la méthode. Il s'agit d'établir, à partir de la population mondiale ou d'échantillons représentatifs, les deux termes suivants : $P(ZTi|E)$ et $P(ZTi|nE)$.

Pour rendre la méthode plus précise, il sera utile de distinguer des sous-populations dans la population mondiale. Typiquement :

- Pour ne pas considérer l'érosion interne comme une seule maladie, mais comme une famille de maladies différentes (pour reprendre l'exemple du §2.1, distinguer différentes formes d'allergie au gluten permet de rendre le test plus spécifique) ;
- Pour tenir compte des facteurs de risques particuliers de certains sous-ensembles de population (pour reprendre l'exemple du §2.1, distinguer les populations à risque en raison de leur patrimoine génétique permet d'affiner le prior).

Mais, même sans ces précisions, l'analyse apporte déjà des indications.

La mise au point des tables de vraisemblance résulte pour l'essentiel d'un jugement général sur la répartition des symptômes d'érosion dans la population mondiale des barrages. A ce stade, les chiffres donnés dans le Tableau 5 restent très généraux et soumis à discussion. Ils résultent d'avis d'experts, basés sur le Tableau 3, et du REX Erinoh.

Ces chiffres donnés ne sont donc pas certains, et devront faire l'objet d'un étalonnage futur. En ordre de grandeur, ils sont jugés vraisemblables, pour les barrages à retenue permanente.

Tableau 5 : table de vraisemblance pour le **TestZT** (A une cote donnée, en situation stabilisée : cote atteinte régulièrement, pendant longtemps), pour les processus d'érosion « rapides » : érosion de conduit et érosion régressive

Echelle	0-1	2	3	4	5-6	7
Classe comportementale	Fc	Ec	Dc	Cc	Bc	Ac
	Écoulements diffus stables	Écoulement concentré stable	Écoulement évolutif	Preuve d'érosion	Forte érosion	Erosion conduisant à la rupture
P(Classe EINS-z) = VP / (VP+FN)	Extrêmement peu probable (a)		Très improbable (c)		Situation la plus courante	
	Classe E ⇨ 0.10% (b)	Classe E ⇨ 0.10% (b)	Classe D ⇨ 1%	Classe D ⇨ 1%	Classe A 18% (d)	Classe A 80% (d)
P(Classe nonEINS-z) = 1- VN / (VN+FP)	Situation la plus courante (f)			Improbable (e)	Improbable (e)	Extrêmement peu probable
	Classe A ⇨ 32%	Classe A ⇨ 32%	Classe A ⇨ 30%	Classe C ⇨ 5%	Classe D ⇨ 1%	-
Déplacement évidence	-2.5	-2.5 (g)	-1.5	-0.7	1	-

(a) : la démarche de qualification utilise la méthode suivante. Cette case correspond à : « un ouvrage subissant une érosion interne non stabilisée, pour une cote de retenue Z donnée, manifeste comme seuls symptômes – à cette cote Z tenue suffisamment longtemps - des écoulements diffus stables ou des écoulements concentrés stables ». Cela est quasiment impossible (au vu des connaissances actuelles sur la cinétique de ces phénomènes) et cela est non rencontré au niveau mondial.

(b) : On adopte alors la cotation E, extrêmement peu probable, et la probabilité associée.

(c) : Ces cas correspondent à un barrage pour lequel un mécanisme d'érosion interne s'est développé, mais a été stabilisé (par différents mécanismes possibles) : filtration, auto-cicatrisation, effondrement d'un toit, ... ; puis le barrage a connu une longue phase de stabilité à la cote considérée. Les auteurs n'ont pas connaissance de barrages EINS dans cette catégorie => classement en « très improbable ». Par contre, de nombreux barrages non EINS ont connu des écoulements évolutifs.

(d) : considérant les notes (a) et (c), il reste 98% à distribuer entre ces deux chiffres. On utilise le REX Erinoh. Notons que le nombre de cas de ruptures est faible, et rend les chiffres incertains. Pour 12 cas de rupture (ERINOH), nous avons : 8 Ac, 2 Bc causés par une fondation évolutive (remontée de karsts, subsidence pour causes externes), 2 cas non déterminés. Le Test ZT aboutit presque toujours à la note Ac, sauf lorsque des phénomènes extérieurs éventuellement lents viennent créer de nouveaux chemins (Briare, Baldwin Hills). Cela correspond historiquement à environ 80% des cas. Les notes Bc constituent des alertes sérieuses. Cela correspond historiquement à environ 20% des cas.

(e) : dans la population des barrages qui ne subissent pas d'érosion interne généralisée, la probabilité que la classe « Preuve d'érosion » ou « Forte érosion » soit atteinte est faible, parce que les ouvrages subissant une Forte érosion ou une Preuve d'érosion sont peu nombreux dans la population des barrages => classement « improbable » par lecture du Tableau 3. Autre approche, la base DEFRA [2] fournit un moyen de comparaison : 5 « serious incident » + 38 « concerns leading to works » sur un parc de 2107 barrages en remblai ; les pourcentages retenus ici sont supérieurs.

(f) : la plupart des barrages se trouvent dans ces trois classes. Compte-tenu de la note e), il reste 94% à distribuer, qui ont été répartis également entre ces trois classes.

(g) : le déplacement d'évidence est calculé comme $\log(0,1\% / 32\%) = -2,5$. La prise en compte de l'information historique permet de diminuer la note comportementale de 2 à 3 classes.

4.3.3. Résultats

La synthèse est exprimée par le tableau ci-dessous :

Selon la classe d'érosion constatée à l'occasion de l'épreuve ZT, on calcule un « déplacement d'évidence » :

Classe d'érosion	Fc	Ec	Dc	Cc	Bc	Ac
Information provenant de l'épreuve ZT	+2.5	+2.5	+1.5	+1	-1	-3

Compte-tenu du caractère préliminaire de la table de vraisemblance ci-dessus, les valeurs numériques de ce tableau sont données uniquement à titre d'illustration ; elles demandent calibration future. Intuitivement, les résultats paraissent solides :

- Plus-value de confiance substantielle (+1,5) qui peut être accordée aux ouvrages ayant subi une épreuve complète de mise en eau (Test ZT) démontrant l'absence de phénomènes évolutifs (classes Ec ou Fc). Modulations :
 - très grande plus-value de confiance si on a à la fois l'épreuve complète et l'auscultation avec absence de dérive (+2,5) ;
 - moindre plus-value si, à retenue permanente, des phénomènes d'érosion sont observés (+1), résultat qui peut paraître paradoxal, mais qui correspond à une érosion interne stabilisée.
- L'historique devient un facteur d'inquiétude (-1) si une forte érosion est constatée (Bc) alors même qu'une épreuve de mise en eau complète a eu lieu (Test ZT).

L'approche comportementale proposée ici génère des déplacements d'évidence de +2,5 à -3. Ces déplacements sont significatifs, et peuvent dans de nombreux cas suffire à changer la conclusion opérationnelle d'un diagnostic.

5. LIMITES, COMMENTAIRES

5.1. L'évaluation de la probabilité $P(T+|M)$: les tables de vraisemblance

L'approche bayésienne repose sur des statistiques qui ne sont pas encore disponibles, et en particulier, $P(T+|M)$: probabilité qu'un barrage « malade » soit diagnostiqué comme tel par un test.

Cependant, le nombre de « maladies » et le nombre de tests ne sont pas si nombreux : il devrait être possible de bâtir des tables génériques, qui pourraient être progressivement affinées. Il sera nécessaire de bâtir des tables différentes. Par exemple pour les différents types d'érosion interne, car les symptômes observés n'ont pas la même signification selon les mécanismes mis en oeuvre.

Les outils pour bâtir ces tables sont :

- Les données statistiques, et on voit que la base DEFRA anglaise, basée sur l'accidentologie constatée sur une trentaine d'années, apporte des éléments de quantification intéressants ;
- Le jugement d'expert, dans une approche assez similaire à celle utilisée dans les analyses de risque.

Le grand intérêt de la méthode, c'est que les tables de vraisemblance sont indépendantes du cas d'étude spécifique : elles sont mises au point en dehors de tout cas particulier, et peuvent ainsi bénéficier de l'accumulation du savoir et de la confrontation entre de nombreux avis d'experts. Lors de l'étude d'un cas particulier, il reste alors à choisir la bonne table de vraisemblance parmi celles qui sont disponibles. Cela limite grandement le « biais » de jugement de l'ingénieur sur un cas particulier.

La grande difficulté de la méthode ... c'est qu'il faut un effort collectif pour aboutir à ces tables.

5.2. Le vieillissement

L'approche proposée ici valorise les épreuves du passé. Il faut pour cela que ces épreuves soient représentatives. Si, entre temps, le barrage a *significativement* vieilli, ou s'il a été *significativement* conforté, alors le résultat de l'épreuve est caduc et le test doit être renouvelé fréquemment. Ce jugement vaut notamment pour les remblais dans leurs premières années de service (effets de la consolidation qui peut ouvrir des fissures par tassement), les digues et petits barrages, qui « vieillissent » par agressions extérieures (fouisseurs, végétations), mais aussi pour les parties d'ouvrages anciens sensibles au vieillissement lent : érosion interne de suffusion ou de contact, corrosion de l'acier, altération du mortier ou du bois des ouvrages inscrits dans les remblais. Il vaut également pour les dispositifs de drainage sensibles au colmatage.

En dehors de ces cas, ou pour les barrages entretenus (i.e. maintenus en conditions), le vieillissement pourrait n'avoir qu'un impact secondaire. Ce qui expliquerait cette observation générique : sauf exceptions et hors séisme, les barrages ont cédé à leur niveau maximum d'exploitation.

Le travail à mener consiste alors à préciser (pour chaque catégorie de table de vraisemblance) ce que signifie un vieillissement *significatif*.

5.3.L'importance de la probabilité a priori (ou « taux de base »), $p(M)$

Un test, même performant, ne permet pas de conduire seul au diagnostic. En effet, la probabilité résultante dépend beaucoup du « taux de base ».

Supposons par exemple un test, soit positif, soit négatif. Les études ont montré que, lorsque le patient est porteur de la maladie, le test est positif dans 90 % des cas : $P(T+|M) = 0,9$. Pour un patient non atteint de la maladie, le test est positif dans seulement 1% des cas (faux positif) : $P(T+|NM)=1\%$. Alors, on obtient : sensibilité 90%, spécificité 99%, ratio de vraisemblance = 90. C'est un bon test.

L'application donne : $P(M|T+) = 0,9 p(M) / [0,89 p(M) + 0,01]$.

Si la maladie est rare et qu'elle ne concerne que 1 personne sur 100 000, alors $p(M)=10^{-5}$. Dans ce cas, $p(M|T+) = 9 \cdot 10^{-4}$. Ce qui signifie que, sur 1 000 000 personnes qui seraient testées, on détecterait 9 VP, 1 FN, 10 000 FP et près de 1 millions de VN. Etre détecté positif laisse une probabilité d'être malade aussi basse que 1 pour mille !

Dans ce cas, le Test apporte certes une information (être détecté positif multiplie par 100 la probabilité d'être malade), mais n'apporte pas de conclusion (être détecté positif laisse une probabilité de 999/1000 de ne pas être malade). Dans l'évaluation de la probabilité d'être malade, l'information principale est portée par $p(M)$, la probabilité a priori.

Appliqué au cas du barrage de Douteux-MaisSansSouci, cela signifie la chose suivante : certes le test apporte des indications précieuses, mais la probabilité de rupture est largement pilotée par la probabilité a priori $p(M)$, qui dépend de la constitution du barrage – et est donc approchée par les calculs de stabilité. Il met bien en évidence que, dans ce cas, le point-clé est le calcul de stabilité – si on réussit à le rendre représentatif.

5.4.Une approche basée sur des données incomplètes

L'approche de Bayes est une approche probabiliste. Elle est pertinente à condition de pouvoir bâtir des tables de probabilités solides. Le moyen le plus fiable d'y parvenir est d'utiliser des statistiques basées sur une population importante. La population des barrages mondiaux est importante : 60 000 grands barrages dans la base de la CIGB ; plus de 2000 (petits) barrages construits au Royaume-Uni depuis 1975, plus de 2000 barrages de classes A à C en France, et des dizaines de milliers de barrages plus petits.

La constitution progressive de bases d'incidents pourrait donc fournir une excellente base de travail. Mais les données demeureront incomplètes. Elles devront être complétées par le jugement d'expert pour établir les tables de vraisemblance, comme ce qui a été fait dans le cas de l'érosion interne. Ce n'est pas parfait, et cela demeure contestable – mais c'est en tous cas une des sources principales du jugement de l'expert, qui peut être rationalisée et actualisée régulièrement.

Par ailleurs, on note que le besoin d'établir des conclusions sur la base de petites populations existe également en médecine, domaine dans lequel l'approche bayésienne est considéré comme une approche utile dans le cas où les populations sont insuffisantes pour procéder à des analyses statistiques classiques. Il reste à pleinement tirer profit de la production scientifique abondante sur le sujet des approches bayésiennes dans le cas de maladies rares.

5.5.Note concernant la relation entre coefficient partiels et probabilité de défaillance

Les Eurocodes utilisent la méthode des coefficients partiels, par laquelle « on donne aux variables de base (actions, résistances et propriétés géométriques) des valeurs de calcul en utilisant des coefficients partiels [...] ». Une question importante a été la calibration des coefficients partiels à considérer.

L'annexe C de l'Eurocode 0 précise que deux méthodes ont été utilisées pour calibrer ces coefficients :

1. « essentiellement » en se basant sur « la longue expérience acquise dans la construction traditionnelle » ;
2. par des méthodes fiabilistes, qui « ont été utilisées par la suite pour le développement des Eurocodes ».

Les méthodes fiabilistes consistent à se fixer une probabilité maximale de défaillance à ne pas dépasser, probabilité de défaillance qui peut être également exprimée sous la forme d'un indice de fiabilité : $P_f = \Phi(-\beta)$, avec Φ la distribution cumulée de la fonction normale standard.

La signification de β est la suivante :

On définit la fonction de performance de la structure $g = R(\text{résistance}) - E(\text{effet des actions})$; la probabilité de défaillance s'écrit alors $P_f = \text{Prob}(g < 0)$, et, si g est normalement distribuée, $P_f = \text{Prob}(g < \mu_g - \beta \sigma_g)$, avec μ_g la moyenne de g et σ_g l'écart type de g . « Pour d'autres distributions de g , β n'est qu'une mesure conventionnelle de la fiabilité ».

Sous ces réserves, les Eurocodes donnent les résultats suivants :

1. Pour des éléments de type RC2, l'indice cible de fiabilité pour les ELU, aux Eurocodes, est 4,7. Soit une probabilité de défaillance de l'ordre de 10^{-6} par an.
2. Pour des éléments de type RC1 (classe de conséquence CC1 : conséquences faibles en cas de rupture) et RC3 (classe de conséquence élevées en cas de rupture), les indices cibles sont 4,2 et 5,2. Ces indices sont atteints en multipliant les actions par un coefficient partiel KFI de 0,9 et 1,1.
3. Ainsi, de manière conventionnelle (exacte seulement si g est une loi normale), augmenter de 10% le coefficient de sécurité divise par 10 la probabilité de rupture ; diminuer de 10% le coefficient de sécurité multiplie par 10 la probabilité de rupture.

L'EC1 attire l'attention sur le point suivant : les probabilités de défaillances ne correspondent pas à des taux réels de défaillance. Ce sont des valeurs opérationnelles « aux fins de calibration des codes et de comparaison des niveaux de fiabilité des structures ».

RÉFÉRENCES ET CITATIONS

- [1] *Biostatistique*, sous la direction de Regis Beuscart, Omniscience, 2009
- [2] DEFRA - Research Contract "Reservoir Safety – Floods And Reservoir Safety Integration", Ref. XU0168 Rev A05, August 2002, p27
- [3] Eurocode 0 – Base de calcul des structures
- [4] J. Calgaro – Introduction aux Eurocodes – Sécurité des constructions et base de la théorie de la fiabilité – Eyrolles, 1996