



HAL
open science

Démarche d'Élicitation, Calibration, Agrégation et Dé-biaisage “ EiCAD ” du jugement d'expert pour l'évaluation probabiliste de la fiabilité structurelle des digues fluviales

Michel Hathout, Claudio Carvajal, Marc Vuillet, Laurent Peyras, Youssef Diab

► **To cite this version:**

Michel Hathout, Claudio Carvajal, Marc Vuillet, Laurent Peyras, Youssef Diab. Démarche d'Élicitation, Calibration, Agrégation et Dé-biaisage “ EiCAD ” du jugement d'expert pour l'évaluation probabiliste de la fiabilité structurelle des digues fluviales. 10èmes journées Fiabilité des Matériaux et des Structures, Mar 2018, Bordeaux, France. hal-01829568

HAL Id: hal-01829568

<https://hal.science/hal-01829568>

Submitted on 4 Jul 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Démarche d'Elicitation, Calibration, Agrégation et Dé-biaisage « EiCAD » du jugement d'expert pour l'évaluation probabiliste de la fiabilité structurelle des digues fluviales

HATHOUT Michel^{1,2}, CARVAJAL Claudio², VUILLET Marc¹, PEYRAS Laurent², DIAB Youssef^{2,3}

¹ Lab'Urba, Université Paris Est, Ecole des ingénieurs de la ville de Paris, 80 rue Rébeval, 75019 Paris

² Irstea, unité de recherche RECOVER, 3275 route Cézanne – CS 40061 – 13182 Aix-en-Provence

³ Lab'Urba, Université Paris Est (Marne-la-Vallée), bureau J29, Bât Lavoisier, 5 Bvd Descartes, 77420 Champs-sur-Marne Ecole des ingénieurs de la ville de Paris, 80 rue Rébeval, 75019 Paris

michel.hathout@irstea.fr; claudio.carvajal@irstea.fr; marc.vuillet@eivp-paris.fr;
laurent.peyras@irstea.fr; youssef.diab@u-pem.fr

RÉSUMÉ.

En génie civil, l'utilisation et l'exploitation du jugement d'expert est un sujet important pour l'évaluation de la fiabilité structurelle des ouvrages dont le comportement de défaillance est mal connu et difficilement quantifiable. Dans cette communication l'application des démarches d'élicitation, de calibration, d'agrégation et de dé-biaisage des avis experts pour évaluer la fiabilité des digues fluviales est étudiée. Ces démarches sont peu exploitées dans le domaine du génie civil alors que leur mise en œuvre dans d'autres domaines comme la statistique, l'économie et la psychologie a été témoin d'un intérêt croissant. La communication propose une démarche quantitative d'exploitation et de traitement du jugement d'expert pour l'évaluation des probabilités de défaillance des digues fluviales. Une application au cas d'une digue fluviale existante est présentée.

ABSTRACT.

In civil engineering, utilization and exploitation of the expert judgment is an important issue in evaluating the structural reliability of civil works whose failure behavior is poorly known and difficult to quantify. In this communication, the application of the elicitation, calibration, aggregation and debiasing steps of expert judgment to evaluate the reliability of river Levees is studied. These steps are been little exploited in the field of civil engineering while their implementation in other domains such as statistics, economics and psychology has witnessed a growing interest. The communication proposes a quantitative methodology of exploitation and treatment of the expert judgment for evaluating of probabilities of failure of the river levees. An application of this methodology in the case of an existing river levee is presented.

MOTS-CLÉS : *digues, incertitudes, jugement expert, biais heuristiques, probabilités subjectives*

KEYWORDS: *Levees, uncertainties, expert judgment, heuristic biases, subjective probabilities*

Introduction

Les digues fluviales de protection contre les inondations sont des ouvrages à grand linéaire et d'une importance stratégique incontournable pour la sécurité des personnes et des biens. Une meilleure connaissance de leur fiabilité structurelle constitue un enjeu majeur pour les ingénieurs et pour les gestionnaires de ces ouvrages. L'évaluation de leur fiabilité structurelle est particulièrement délicate en raison de faibles quantités de données disponibles, de la complexité des mécanismes de défaillances et de l'hétérogénéité des ouvrages. Les digues

sont des ouvrages anciens, mal connus, mal entretenus et mal documentés (il est rare d'avoir des plans de construction et les rapports d'éventuels travaux de réparation ou de confortement) (Mériaux *et al.*, 2001). Leur comportement et leur risque de défaillance demeurent dans une large mesure mal connus et difficilement quantifiables vis-à-vis des différents mécanismes de rupture essentiels tels que le glissement du talus, l'érosion interne, l'affouillement et la surverse. De ce fait, l'évaluation par des approches quantitatives, de la fiabilité structurelle des digues fluviales, demeure -à ce jour- complexe et lourde. Ceci conduit à une prédominance du recours au jugement d'expert dans les études de fiabilité et plus précisément dans les études d'analyse de risques. Les ingénieurs experts doivent procéder à l'interprétation des données disponibles et à la prise en compte de leurs incertitudes, pour pouvoir ensuite évaluer la fiabilité structurelle des digues en termes de probabilité de défaillance.

L'ILH (2013) (International Levees Handbook) souligne l'avantage et l'importance d'utiliser le jugement expert pour réaliser une analyse de performance des digues lorsque l'évaluation par des approches quantitatives est difficile à réaliser (Sharp *et al.*, 2013). Toutefois il ne préconise aucune démarche spécifique permettant une bonne utilisation et exploitation du jugement d'expert dans un contexte incertain. Serre *et al.*, (2008) propose un modèle déterministe d'évaluation de la performance des digues basé sur une méthode d'aide à la décision multicritère : l'affectation à base de règles. Vuillet *et al.* (2013) propose un modèle d'évaluation de la performance des digues permettant de prendre en compte les incertitudes et le manque de données affectant l'évaluation de la sûreté des digues sur la base d'indicateurs de performance. Le modèle demeure semi-quantitatif et il ne permet pas de raisonner en probabilité de rupture des ouvrages. Par ailleurs, ce modèle développé repose sur l'utilisation d'avis d'un seul expert ou d'un consensus d'experts. Il ne propose pas de solution pour la combinaison de plusieurs avis experts or l'évaluation de la fiabilité des digues repose sur plusieurs domaines de compétences (géotechnique, hydraulique fluviale, génie civil, etc.). En outre, malgré leurs compétences, leurs connaissances approfondies et leurs expertises, les experts ne sont pas à l'abri de biais propres aux opinions et aux jugements probabilistes (O'Hagan *et al.*, 2006). En effet, il y a toujours une transformation qui peut engendrer une différence entre la connaissance détenue et la façon dont elle est restituée et exprimée (Bateson, 1979). Ces différences s'expriment principalement par la présence de biais pouvant intervenir à différents moments du processus d'élicitation et pour lesquels un traitement spécifique devrait être établie (Kahneman *et al.*, 1982).

La littérature du génie civil est moins riche en méthodes de traitement du jugement expert que la littérature statistique, économique et psychologique. Ces dernières distinguent trois démarches d'exploitation du jugement d'expert : les démarches d'élicitation, les démarches de calibration et d'agrégation et les démarches de dé-biaisage des avis experts. Cependant l'utilisation et l'adaptation des différentes méthodes d'exploitation du jugement d'expert pourraient être pertinentes dans le domaine des digues fluviales pour l'évaluation de leur fiabilité structurelle. L'objectif de cette communication est de présenter une démarche d'élicitation, de calibration, d'agrégation et de dé-biaisage des avis experts pour l'évaluation des probabilités de défaillance des digues fluviales vis-à-vis des mécanismes de rupture. Dans un premier temps, la communication présente le développement d'un protocole d'élicitation des avis experts fournis en termes de probabilité de défaillance vis-à-vis le mécanisme de rupture par glissement de talus. Dans un deuxième temps, elle présente l'adaptation du modèle de calibration de Cooke (1991) aux domaines des digues afin de pondérer les avis experts probabilistes. Ensuite, elle présente le résultat d'application du protocole d'agrégation des avis experts de Lichtendahl *et al.* (2013) permettant de déterminer, à partir du jugement d'expert une distribution probabiliste calibrée et agrégée des probabilités de défaillance pour

plusieurs profils d'une digue existante. Enfin, la communication présente le développement d'un protocole de dé-biaisage des avis experts basé sur les notions du modèle du dé-biaisage des avis experts de Clemen et Lichtendahl (2002).

Démarches méthodologiques d'exploitation du jugement expert pour les digues

Les études d'analyse de risques des digues doivent inclure l'évaluation probabiliste des différents scénarios de défaillance, et plus particulièrement de ceux conduisant à une rupture de digue (ILH, 2013). L'évaluation de la fiabilité structurelle des ouvrages de génie civil s'exprime habituellement en termes de probabilité de défaillance (EN1990, 2002). Ainsi, l'évaluation de la fiabilité structurelle des digues constitue un élément majeur dans les études d'analyse de risques pour évaluer des probabilités de défaillance des digues vis-à-vis des mécanismes de rupture. Dès-lors, la démarche développée se focalise sur l'élicitation par jugement expert des probabilités de défaillance vis-à-vis des mécanismes de rupture de digues. Dans le but de prendre en compte les incertitudes associées aux digues, la démarche privilège les élicitations expertes sous forme d'intervalle d'incertitude et d'une valeur plus vraisemblable (Vuillet et al., 2013 ; Soll et Klayman, 2003). Les études d'analyse de risques des digues traitent une diversité d'avis experts associés à des différentes disciplines (géotechnique, hydraulique fluviale, génie civile,...). Plusieurs ingénieurs ou experts de plusieurs disciplines avec plusieurs niveaux d'expériences (junior, confirmé, senior) sont donc susceptibles d'intervenir dans ces études d'analyse de risques des digues. Ainsi, les élicitations probabilistes sont fournies par un panel d'experts contenant plusieurs experts dont le choix est basé sur deux critères : l'interdisciplinarité (avoir des experts de champs de compétence différents et complémentaires) et la diversité d'opinions (avoir des opinions représentant tous les aspects du problème) (Morgan et Henrion, 1990).

Dans le cadre de l'évaluation de la fiabilité des digues fluviales à partir du jugement expert, la démarche s'appuie ensuite sur des méthodes de calibration et d'agrégation pour évaluer les avis de différents experts en vue de les agréger. A la fin, la démarche vise à traiter quantitativement le biais de sur-confiance/sous-confiance susceptibles d'entacher les évaluations expertes calibrées et agrégées.

Démarche d'élicitation des avis experts pour les digues fluviales

La mise en œuvre de la démarche d'élicitation des avis experts débute par l'identification d'un panel comportant plusieurs experts ayant une activité liée directement au domaine des digues fluviales (Koehler et harvey, 2004). Chaque expert du panel est questionné individuellement sur des probabilités de défaillance relatives à plusieurs profils de digue à évaluer vis-à-vis de différents mécanismes de rupture. Un formulaire de questionnement et de recueil des jugements experts sous format probabiliste est développé pour ce but. Pour chaque variable à éliciter, le formulaire comporte trois éléments principaux : l'élément de question, l'élément d'informations et l'élément de réponse. Dans l'exemple de la figure (Fig. 1), concernant le mécanisme de glissement, ces éléments se déclinent de la manière suivante :

- L'élément de question : correspond à la question sur laquelle les experts sont interrogés. Il contient une précision sur le mécanisme à évaluer, le talus du profil à évaluer (aval ou amont) et la crue pour laquelle la probabilité de défaillance est demandée. Cette question, pour le mécanisme de rupture par glissement, prend la forme suivante : « *Compte tenu de toutes les informations à disposition, quelle est la probabilité P_{fc} que la digue se rompe par glissement du talus aval en cas d'occurrence d'une crue ayant un niveau d'eau atteignant la crête de digue? $P_{f-totale} = P_{fc} * P_{crue}$ »*

- L'élément d'informations : fournit toutes les données nécessaires aux experts pour formuler leurs réponses et donner leurs jugements. Ces informations correspondent aux « *informations de repérage objectif* » (Soll, 1996). Les éléments d'information constituant le formulaire sont : les données géométriques, les données hydrauliques, les données géotechniques, la coupe type du profil à évaluer, les lois des probabilités des caractéristiques géotechniques du profil à évaluer et les données issues d'une analyse déterministe représentée par le facteur de sécurité FS de la digue vis-à-vis du mécanisme de rupture par glissement du talus.
- L'élément de réponse : correspond à la façon dont l'expert communique son jugement sur la probabilité de défaillance (P_{fc}) du profil à évaluer. L'expert est sollicité tout d'abord pour éliciter un intervalle d'incertitude [quantile 5%, quantile 95%] puis, pour préciser la valeur la plus vraisemblable issue de l'intervalle élicité comme proposé par Soll et Klayman (2003).

Fiche 01		Profil – Digue Théorique 1	
Compte tenu de toutes les informations à disposition, quelle est la probabilité P_{fc} que la digue se rompe par glissement du talus aval en cas d'occurrence d'une crue ayant un niveau d'eau atteignant la crête de digue ?			
$P_{f-Total} = P_{fc} * P_{crue}$			
Données Géométriques <i>Cf. Figure ci-dessous</i>	- Hauteur = 5 m - Côte en crête = 205 m NGF - Largeur en crête = 6 m - Fruit amont = 3/2 (H/V) - Fruit aval = 3/2(H/V)		
Données Hydrauliques	- Niveau d'eau de la crue atteignant la crête de la digue - Ligne piézométrique calculée en régime permanent		
Données Géotechniques	Corps et Fondation		
	- Matériaux sable-limoneux - Perméabilité $K = 10^{-6} m/s$ - Poids volumique : $\gamma_h = 18 kN/m^3$		
Lois des probabilités des données d'entrée			
	L'angle de frottement (corps - fondation) Loi probabiliste : loi normale Moyenne : 20° - Ecart-type : 5°		La cohésion (corps - fondation) Loi probabiliste : loi normale tronquée à 0 Moyenne : 10 kPa - Ecart-type : 5 kPa, Min = 0 kPa
Analyse déterministe	FS calculé (avec les valeurs moyennes) Talus aval 1,3		
Analyse probabiliste	Probabilité de défaillance (P_{fc}) au glissement	P_{fc} : quantile 5%	
		P_{fc} : quantile 95%	
		P_{fc} : Valeur la plus vraisemblable (q50%)	

Fig. 1 : Exemple d'une fiche d'un profil d'une digue issue du formulaire de questionnaire développé pour évaluer la probabilité de défaillance au glissement

Démarche de calibration des avis experts pour les digues fluviales

La mise en œuvre de la démarche de calibration des avis experts est basée sur le modèle de Cooke (1991). Ce modèle consiste à évaluer et pondérer les avis experts par rapport à des variables dont les vraies valeurs sont connues. Il repose sur une distinction entre des variables d'intérêt et des variables de calibration. Les variables d'intérêt correspondent à des variables dont l'analyste cherche à connaître les valeurs. Le jugement expert est utilisé donc pour

construire une distribution probabiliste subjective concernant la réalisation de ces variables. Les variables de calibration correspondent à des variables susceptibles d'éclairer l'analyste sur la capacité des experts à fournir une estimation juste et précise. Elles sont autant que possible similaire à la variable d'intérêt, à cette différence près qu'il existe pour ces variables des résultats connus (expérimentaux, observés, analytiques,...). Dans le domaine des digues, il est nécessaire donc d'identifier des potentielles variables de calibration pertinentes par rapport aux variables d'intérêt. Plusieurs candidats sont envisageables, tels que : les facteurs de sécurité résultant d'un calcul géotechnique, les propriétés de matériaux mesurées par des essais, etc. Ces potentielles variables de calibration restent pour la quasi-totalité d'entre elles dans une forme déterministe (et non probabiliste).

Dans cette étude, il a été adopté comme variable de calibration la probabilité de défaillance de profils théoriques de digues vis-à-vis du mécanisme de rupture par glissement. Cette probabilité peut être évaluée d'une manière quantitative par une analyse de fiabilité. Cette analyse repose d'une part, sur la définition du mécanisme de glissement par un modèle géotechnique de stabilité de talus, puis d'une autre part sur la définition des incertitudes sur les données d'entrée (cohésion et angle de frottement) comme des variables aléatoires. Ensuite, une propagation des incertitudes par couplage mécano-fiabiliste (classiquement par simulation de Monte Carlo) est effectué afin d'évaluer la distribution de la variable de sortie (le facteur de sécurité - Fs) et de déterminer la probabilité de défaillance (considérée comme la probabilité que le facteur de sécurité soit inférieur à 1).

Le modèle de Cooke (1991) permet par le calcul du poids de calibration individuel w_e de chaque expert, lequel peut s'exprimer également comme un poids de calibration relatif w'_e par rapport à l'ensemble des experts du panel. Le poids de calibration se détermine à partir du score de calibration C_e et du score d'entropie K_e selon les formules suivantes (Cooke, 1991) :

$$w_e = C_e \cdot K_e \rightarrow w'_e = \frac{C_e \cdot K_e}{\sum_{e=1}^z C_e \cdot K_e} \quad [1]$$

Où : w_e poids de calibration de l'expert e ; w'_e poids de calibration relatif de l'expert e au sein d'un panel d'experts ; C_e score de calibration de l'expert e ; K_e score d'entropie de l'expert e et z nombre d'experts

Le score de calibration C_e permet d'évaluer la qualité (la justesse) de l'information donnée par l'expert (par comparaison entre une distribution obtenue de l'expert et une distribution cible). Il se détermine par le complément d'une probabilité d'erreur évaluée à partir du test statistique khi2 (test d'interdépendance). Ce score est évalué par la formule suivante (Cooke, 1991) :

$$C_e = 1 - P(2 * n * I_e(c, p)) = 1 - \chi^2(2 * n * I_e(c, p)) \quad [2]$$

Où : $P(\)$ probabilité d'une variable aléatoire suivant une loi khi2 ; $\chi^2(\)$ fonction de répartition d'une variable aléatoire selon la loi khi2 ; n nombre de variables de calibration ; c vecteur de calibration représentant la portion de vraies valeurs comprises dans chaque intervalle interquantile $c = \{c_1, c_2, \dots, c_j\}$; p vecteur de probabilité théorique représentant les probabilités théoriques de chaque intervalle interquantile $p = \{p_1, p_2, \dots, p_j\}$; $I_e(c, p) = \sum_{i=1}^j c_i \ln\left(\frac{c_i}{p_i}\right)$ information relative entre le vecteur de probabilités théoriques p et le vecteur de calibration c .

Le score d'entropie K_e permet de mesurer la quantité d'informations contenues dans les distributions probabilistes données par les experts. Elle s'appuie sur une mesure de distances

entre le vecteur de probabilités subjectives s et le vecteur de probabilités théoriques p . Ce score est évalué par la formule suivante :

$$K_e = \frac{1}{n} \sum_{i=1}^n I_{e,i}(p, s) \quad [3]$$

Où : s vecteur de probabilité subjective représentant les probabilités subjectives de chaque intervalle interquantile $s = \{s_1, s_2, \dots, s_j\}$. Il est utilisé pour mesurer la distance entre la distribution de l'expert et une distribution uniforme non informative ; $I_e(p, s) = \sum_{i=1}^j p_i \ln \left(\frac{p_i}{s_i} \right)$ l'information relative entre le vecteur de probabilités subjectives s et le vecteur de probabilités théoriques p .

Démarche d'agrégation des avis experts pour les digues fluviales

La mise en œuvre de la démarche d'agrégation des avis experts s'effectue dans cette étude par la moyenne des quantiles de Lichtendahl *et al.* (2013). L'objectif est de conserver la forme initiale des élicitations expertes données sous format probabiliste avec une valeur considérée comme la plus vraisemblable et un intervalle d'incertitude. Cette démarche correspond à une somme pondérée des quantiles calibrés des avis experts. Les quantiles agrégés ($q_{5\%}$, $q_{50\%}$, $q_{95\%}$) sont évalués par les formules suivantes :

$$q_{5\%} = \sum w'_e \cdot q_{5\%(e)} \quad ; \quad q_{50\%} = \sum w'_e \cdot q_{50\%(e)} \quad ; \quad q_{95\%} = \sum w'_e \cdot q_{95\%(e)} \quad [4]$$

Où : $q_{5\%}$ quantile agrégé correspondant à une probabilité de 5% ; $q_{50\%}$ quantile agrégé correspondant à une probabilité de 50% ; $q_{95\%}$ quantile agrégé correspondant à une probabilité de 95%.

Démarche de dé-biaisage des avis experts pour les digues fluviales

La mise en œuvre de la démarche de dé-biaisage des avis experts s'effectue dans cette étude par le principe du modèle dé-biaisage des avis experts de Clemen et Lichtendahl (2002). L'objectif est d'identifier quantitativement la tendance de l'avis expert à la sous-confiance et/ou à la sur-confiance susceptible d'entacher l'estimation de la valeur la plus vraisemblable et son intervalle d'incertitude. Cette démarche correspond à déterminer par calcul itératif trois coefficients de correction. Le premier coefficient β vise à corriger la valeur élicitée comme « la plus vraisemblable » (correspondant au quantile 50% dans notre cas) par la formule suivante (Clemen et Lichtendahl, 2002) :

$$R_i^* = \beta R_i \quad [5]$$

Où : R_i^* valeur la plus vraisemblable dé-biaisée ; β coefficient de correction concernant la valeur la plus vraisemblable ; R_i élicitation experte calibrée et agrégée de la valeur la plus vraisemblable ($q_{50\%}$).

Le deuxième coefficient α_L vise à corriger la valeur élicitée de la borne inférieure de l'intervalle d'incertitude (correspondant au quantile 5% dans notre cas) par la formule suivante (Clemen et Lichtendahl, 2002) :

$$L_i^* = (\beta - \alpha_L)R_i + \alpha_L L_i \quad [6]$$

Où : L_i^* borne inférieure dé-biaisée de l'intervalle d'incertitude ; α_L coefficient de correction concernant la borne inférieure de l'intervalle d'incertitude ; L_i élicitation experte calibrée et agrégée de la borne inférieure de l'intervalle d'incertitude ($q_{5\%}$).

Le troisième coefficient α_U vise à corriger la valeur élicitée de la borne supérieure de l'intervalle d'incertitude (correspondant au quantile 95% dans notre cas) par la formule suivante (Clemen et Lichtendahl, 2002) :

$$U_i^* = (\beta - \alpha_U)R_i + \alpha_U U_i \quad [7]$$

Où : U_i^* borne supérieure dé-biaisée de l'intervalle d'incertitude ; α_U coefficient de correction concernant la borne supérieure de l'intervalle d'incertitude ; U_i élicitation experte calibrée et agrégée de la borne supérieure de l'intervalle d'incertitude ($q_{95\%}$).

Le schéma explicatif ci-dessous (Fig.2) illustre les données d'entrée et de sortie des différentes phases de la démarche EiCAD développée :

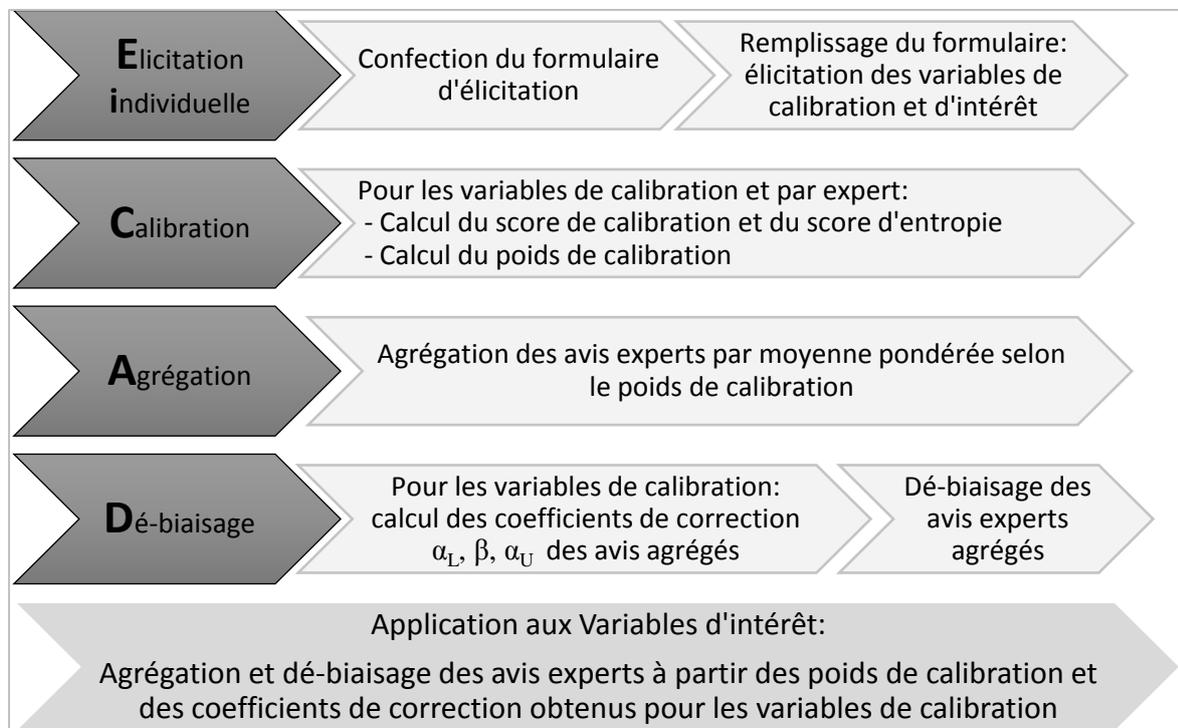


Fig.2 : Schéma explicatif de la démarche développée EiCAD

Résultats

La digue étudiée est une digue en remblai destinée à la protection d'une agglomération urbaine contre des phénomènes épisodiques de crue susceptibles de provoquer des inondations. Sur la zone d'étude, la digue présente une hauteur variant globalement entre 1 m et 6 m. La mise en œuvre de la démarche développée est réalisée par un panel d'experts de trois ingénieurs ayant différents profils et expériences (géotechnique, hydraulique fluviale, génie civil). Ils fournissent leurs jugements sur les variables de calibration et les variables d'intérêt correspondant dans les cas aux probabilités de défaillance des ouvrages vis-à-vis des mécanismes de rupture de digue. Les variables de calibration correspondent à des probabilités de défaillance au glissement pour 24 profils des digues théoriques. L'étude

s'applique ensuite sur 12 variables d'intérêt correspondant à 6 profils de la digue étudiée vis-à-vis de mécanisme de rupture par glissement et 6 profils de la digue étudiée vis-à-vis de mécanisme de rupture par érosion interne.

A l'issue de la phase d'élicitation des avis experts, la répartition des vraies valeurs sur l'ensemble des intervalles d'incertitude élicités par chaque expert du panel est présentée dans le tableau (Tab. 1) ci-dessous :

Tab. 1 : Répartition des vraies valeurs de variables de calibration (n°1-n°24) dans les intervalles interquartiles élicités par les experts

Intervalle interquartile		[0%, 5%]	[5%, 50%]	[50%, 95%]	[95%, 100%]
Expert n°1					
(P_{fc}) (n°1 à n°24)	n	9	10	5	0
	%	37.5%	41.7%	20.8%	0%
Expert n°2					
(P_{fc}) (n°1 à n°24)	n	10	5	4	5
	%	41.7%	20.8%	16.7%	20.8%
Expert n°3					
(P_{fc}) (n°1 à n°24)	n	10	6	8	0
	%	41.7%	25%	33.3%	0%

Le tableau (Tab. 2) suivant présente pour chaque expert du panel le résultat de la calibration des avis experts : score de calibration (C_e), score d'entropie (K_e), poids de calibration individuel (w_e) et poids de calibration relatif (w'_e).

Pour ce qui concerne le calcul du score d'entropie, il est à noter que les valeurs de probabilité de défaillance élicitées peuvent avoir des valeurs très faibles et exprimées généralement en (10^{-x}). Dès-lors, un changement de variable est envisagé de manière à garder une valeur d'écart constante entre (10^{-x}) et ($10^{-(x-1)}$). Le changement de variable adopté correspond simplement à appliquer le logarithme décimal aux probabilités élicitées : $\log_{10}(P_{fc})$

Tab. 2 : Résultats de la phase de calibration des avis experts par rapport aux variables de calibration

	Score de calibration	Score d'entropie	Poids de calibration individuel	Poids de calibration relatif
n° expert	C_e	K_e	w_e	w'_e
1	0.25	0.81	0.21	0.57
2	0.01	0.91	0.01	0.02
3	0.13	1.10	0.15	0.41
			Somme	1

Le score de calibration élevé obtenu par l'expert n°1 reflète que ces intervalles d'incertitude élicités contiennent plus de vraies valeurs de calibration par rapport aux intervalles

d'incertitudes des autres experts du panel (15 variables sur 30 \in [5%, 95%] (Tab.1)). Au contraire, le score de calibration faible obtenu par l'expert n°2 reflète que ces intervalles d'incertitude élicités contiennent moins de vraies valeurs de calibration par rapport aux intervalles d'incertitude des autres experts du panel (9 variables sur 30 \in [5%, 95%] (Tab.1)).

Le score d'entropie élevé obtenu par l'expert n°3 reflète que ces intervalles d'incertitude élicités sont les plus précis (plus serrés) par rapport à ceux donnés par les autres experts du panel. Au contraire, le score d'entropie faible obtenu par l'expert n°1 reflète que ces intervalles d'incertitude élicités sont les moins précis (plus larges) par rapport à ceux donnés par les autres experts du panel.

Il est donc à noter que le score de calibration présente une influence plus significative par rapport au score d'entropie sur le résultat final de la phase de calibration des avis experts. Ainsi, l'avis de l'expert n°1 ayant eu le score de calibration le plus élevé a eu le meilleur poids de calibration relatif ($w'_1 = 57\%$) alors que l'avis de l'expert n°2 ayant eu le score de calibration le moins élevé a eu le poids de calibration relatif le plus faible ($w'_2 = 2\%$).

Le résultat de l'agrégation des avis experts se présente par une seule distribution probabiliste relative à chaque profil de digue évalué par les experts. Le tableau suivant (Tab. 3) présente la répartition des vraies valeurs de variable de calibration dans les intervalles interquantiles à l'issue de l'application de la phase d'agrégation des avis experts :

Tab. 3 : Répartition des vraies valeurs de variable de calibration (n°1-n°24) dans les intervalles interquantiles agrégés

Intervalle interquantile		[0%, 5%]	[5%, 50%]	[50%, 95%]	[95%, 100%]
(P _{fc}) (n°1 à n°24)	n	10	11	3	0
	%	41.7%	45.8%	12.5%	0%

La phase d'agrégation des avis experts calibrés permet d'avoir un intervalle d'incertitude [5%,95%] contenant plus de vraies valeurs par rapport à des intervalles d'incertitude des avis experts bruts. Toutefois à l'intérieur de cet intervalle d'incertitude la répartition des vraies valeurs n'est pas équilibrée car 45.8% de vraies valeurs se trouvent dans l'intervalle interquantile [5%, 50%] et 12.5% de vraies valeurs se trouvent dans l'intervalle interquantile [50%, 95%]. Avec une répartition totale de 87.5% des vraies valeurs dans l'intervalle [0%, 50%], cela note une tendance des avis experts du panel à la sur-confiance.

Le résultat du dé-biaisage des avis experts se présente aussi par une seule distribution probabiliste relative à chaque profil de digue évalué par les experts. Le graphique suivant illustre le résultat du dé-biaisage des avis experts calibrés pour les 24 variables de calibration relatives aux profils des digues théoriques (lignes continues) ainsi que le résultat du dé-biaisage pour les 12 variables d'intérêt relatives aux profils de la digue d'étude (lignes pointillés) où les vraies valeurs se présentent par un cercle :

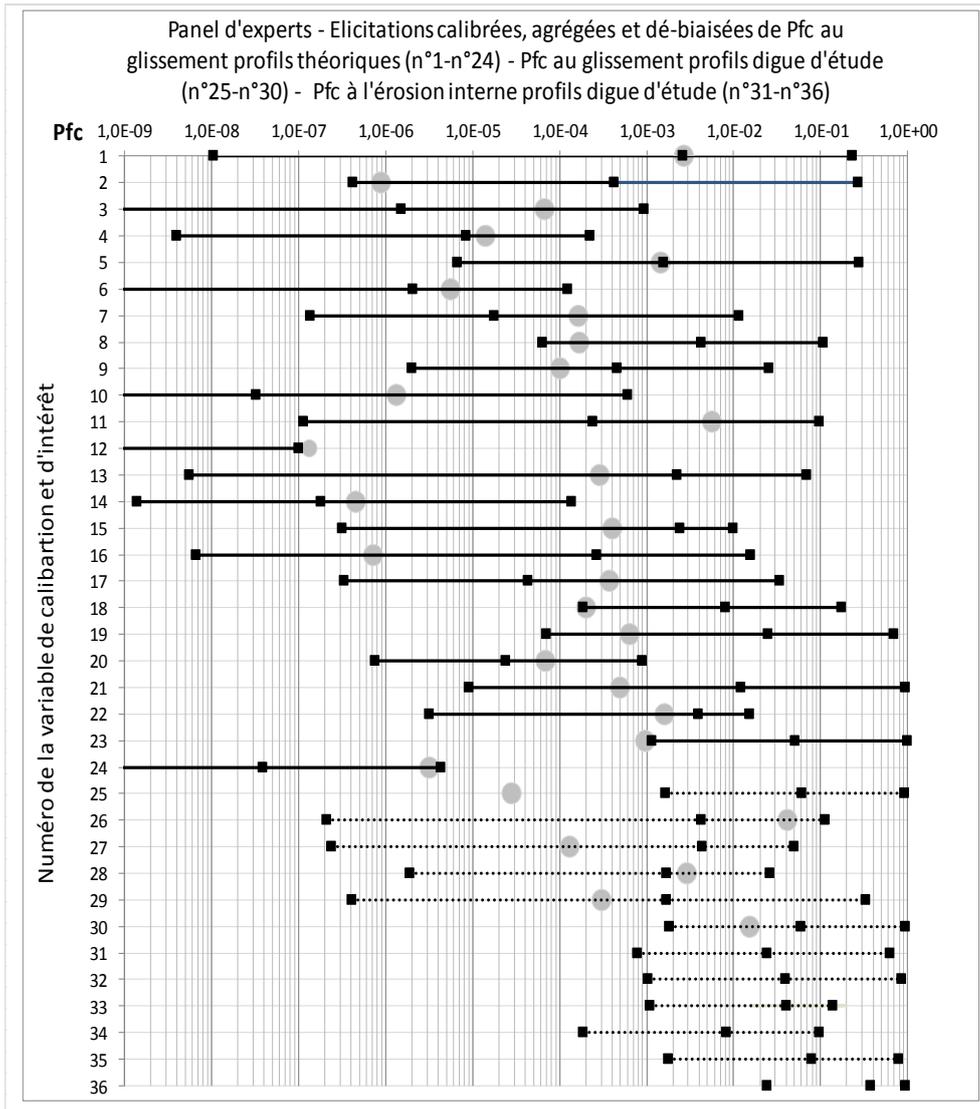


Fig. 3 : Distributions calibrées, agrégées et dé-biaisées des avis experts

Les coefficients de correction obtenus issus de la phase de dé-biaisage correspondent à $\beta = 1.38$, $\alpha_L = 2.11$ et $\alpha_U = 2$ qui conduit à une répartition des vraies valeurs des variables de calibration dans les intervalles interquantiles présentées dans le tableau suivant :

Tab. 4 : Répartition des vraies valeurs de variables de calibration (n°1-n°24) dans les intervalles interquantiles dé-biaisés

Intervalle interquantile		[0%, 5%]	[5%, 50%]	[50%, 95%]	[95%, 100%]
(P _{fc}) (n°1-n°24)	n	1	11	11	1
	%	4%	46%	46%	4%

A l'issue de cette phase, les nouvelles distributions probabilistes contiennent plus de vraies valeurs dans l'intervalle d'incertitude (92%) par rapport aux résultats obtenus des phases

précédentes, avec une répartition des vraies valeurs très proche de la répartition ciblée {5%, 45%, 45%, 5%}. Dans cette application, la calibration, agrégation et dé-biaisage des jugements experts conduit à des résultats assez satisfaisants pour la valeur la plus vraisemblable ($q_{50\%}$), avec cependant des intervalles d'incertitude très larges.

Parmi les différentes variables d'intérêt considérées dans cette étude, il est possible d'évaluer de manière quantitative (par une analyse de fiabilité) les probabilités de défaillance vis-à-vis du mécanisme de rupture par glissement. Une analyse de fiabilité a été réalisée pour les variables n°25 à n°30, lesquelles correspondent aux probabilités de défaillance au glissement des six profils de la digue d'étude (ce sont des variables d'intérêt et elles n'ont donc pas été utilisées dans la phase de calibration). Cette analyse permet ainsi de comparer ici, juste pour quelques variables (n°25 à n°30), les résultats issus de la démarche de calibration, agrégation et dé-biaisage des avis d'experts avec les résultats d'une analyse quantitative (plus lourde et pas forcément applicable pour tous les mécanismes de rupture considérés). Dans la figure (Fig. 2), il est remarquable que la majorité des distributions finales des avis experts calibrées, agrégées et dé-biaisées des variables n°25 à n°30 contiennent la vraie valeur dans l'intervalle d'incertitude [5%, 95%]. Ce résultat apparaît assez satisfaisant compte tenu de la faible taille de l'échantillon (6 valeurs), de la complexité d'une analyse quantitative de fiabilité et des faibles valeurs de probabilité à éliciter par jugement expert. Les intervalles d'incertitude restent toutefois très larges.

La démarche a été testée également pour des variables de calibration déterministes (correspondant à des facteurs de sécurité au glissement). Les poids de calibration obtenus pour ces variables déterministes ont présenté des différences par rapport à ceux obtenus des variables de calibration probabilistes (correspondant à des probabilités de défaillance). Cela montre simplement que certains experts maîtrisent mieux les facteurs de sécurité et que d'autres experts maîtrisent mieux les probabilités de défaillance. Il convient ainsi que les variables de calibration soient du même type que les variables d'intérêt.

Conclusion

La réalisation des études de fiabilité de digues par des approches quantitatives est une procédure difficile à effectuer compte tenu de la complexité des mécanismes concernés et du manque de données fiables. Dans ce cadre, la communication présente une démarche d'évaluation des probabilités de défaillance de digues basée sur l'Elicitation individuelle, la calibration, l'agrégation et le dé-biaisage du jugement expert (EiCAD). L'issue de cette démarche correspond à un poids de calibration et des coefficients de correction qui, une fois déterminés pour chaque expert (par des analyses quantitatives de fiabilité), est possible de les réutiliser ultérieurement pour d'autres études d'analyse de risques de digues. Plus généralement, la démarche développée pour le cadre des digues fluviales peut être adaptée et appliquée dans d'autres domaines où le recours au jugement expert est la seule source pour l'obtention d'informations exploitables pour effectuer une analyse de fiabilité.

La démarche développée présente un point de départ important pour combler la lacune relative à l'exploitation du jugement expert dans le domaine des digues fluviales. Elle permet d'évaluer la fiabilité structurelle des ouvrages en termes de probabilité de défaillance à partir du jugement d'expert. La démarche EiCAD proposée a permis, par le poids de calibration, à identifier les meilleures élicitations expertes et de les exploiter. Elle accorde un poids faibles de calibration aux élicitations expertes qui sont peu justes et moins informatives. Ainsi, les élicitations les plus pertinentes ont plus de poids. De même, la démarche proposée a permis, par les coefficients de correction, à corriger quantitativement la tendance des avis experts à la

sur-confiance/sous-confiance susceptible d'entacher les élicitations de la valeur la plus vraisemblable et son intervalle d'incertitude.

Références

- Bateson, G. (1979) «Mind and Nature: a necessary unity». Bantam Books, New-York.
- Clemen, R.T., Lichtendahl, K.C. (2002). “Debiasing expert overconfidence: a Bayesian calibration model”. PSAM6, San Juan, Puerto Rico
- Cooke, R. (1991). “Experts in uncertainty: opinion and subjective probability in science”. Oxford University Press, 336 p.
- EUROPEAN COMMITTEE FOR STANDARDISATION. EN 1990: 2002 (2002). Eurocode-Basis of Structural Design. European ommittee for Standardisation, Brussels.
- Kahneman. D., Slovic, P., Tversky, A. (1982). «Judgement under Uncertainty, Heuristics and Biases». Cambridge University Press, Cambridge.
- Koehler, J., Harvey, N. (2004). “Blackwell Handbook of judgment & Decision Making, Blackwell Publishing”, UK, 675p.
- Lichtendahl, K.C Jr., Grushka-Cockayne, Y., Winkler, RL. (2013). “Is it better to average probabilities or quantiles?”, *Management Sci.* 59:1594–1611.
- Mériaux, P., Royet, P., Folton, C. (2001). “Surveillance, entretien et diagnostic des digues de protection contre les inondations”, Cemagref Editions, 191 p.
- Morgan, M.G., Henrion, M. (1990). “Uncertainty - A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis”, Cambridge, Royaume-Uni, Cambridge University Press.
- O'Hagan, A., Buck, C. E., Daneshkah, A., Eiser, J. R., Garthwaite, P.H., Jenkinson, D.J., Oakley, J.E., Rakow, T. (2006). *Uncetains Judgments : Eliciting Experts' Probabilities*, West Sussex, UK, 338 p.
- Serre, D., Peyras, L., Tourment, R., Diab, Y., (2008). “Levee Performance Assessment Methods Integrated in a GIS to Support Planning Maintenance Actions”, *ASCE - Journal of Infrastructure Systems*, 10.1061/(ASCE)1076-0342(2008)14:3(201), 201-213.
- Sharp, M., Wallis, M., Deniaud, F., Hersch-Burdick, R., Tourment, R., Matheu, E., Seda-Sanabria, Y., Wersching, S., Veylon, G., Durand, E., Smith, P. (2013). “The international Levee Handbook”, C731. CIRIA, London.
- Soll, J. B. (1996). “Determinants of overconfidence and miscalibration: The roles of random error and ecological structure”, *Organizational Behavior and Human Decision Processes*, 65, 117-137.
- Soll, J.B., Klayman, J. (2003). “Overconfidence in intervale estimates”. *Journal of Experimental Psychology: Learning, memory and cognition*, 30, 299-314.
- Vuillet, M., Peyras, L., Carvajal, C., Serre, D., Diab, Y. (2013). “Levees performance evaluation based on subjective probability”, *European Journal of Environment and Civil Engineering*, Volume 17, issue 5/2013, pp 329-349.