

# Représentation et propagation de la connaissance imprécise : ce que les théories de l'incertain peuvent apporter aux sciences environnementales

Catherine Mercat-Rommens\*, Éric Chojnacki\*, Cédric Baudrit\*\*

Dans les sciences environnementales, l'analyse d'incertitude est devenue une étape incontournable de la restitution des résultats, tant en termes de rigueur scientifique qu'en termes de crédibilité sociale. Cependant, sa réalisation soulève généralement de nombreuses difficultés en raison d'une connaissance souvent incomplète ou imprécise ce qui nécessite alors fréquemment le recours à une connaissance plus subjective, le jugement d'expert. Dans ce contexte, les techniques statistiques classiques à disposition de l'environnementaliste ne sont pas toujours suffisantes pour rendre compte de la connaissance disponible et l'exemple rencontré dans le cas de l'étude dite « du Nord-Cotentin » montre certaines limites des outils classiques basés sur la théorie des probabilités. Pour dépasser ces limites, les possibilités d'évolution des outils pour une meilleure adaptation à la connaissance disponible dans le domaine des sciences environnementales et notamment dans le domaine de la radioécologie ont été recherchées en s'appuyant sur les théories de l'incertain.

## **Le cas du Nord-Cotentin : les limites de l'approche classique de traitement de l'incertitude**

### *Contexte de l'évaluation des risques*

Entre 1997 et 1999, dans un contexte de polémique scientifique, le Groupe Radioécologie Nord-Cotentin (GRNC), comité d'experts scientifiques, reconstitue le risque de leucémie radio-induite par les rejets des installations nucléaires du Nord-Cotentin. Il s'agissait de répondre à l'inquiétude des populations locales suite à la publication des travaux de

---

\* Institut de Radioprotection et de Sûreté Nucléaire, CE Cadarache, 13115 Saint-Paul-lez-Durance cedex, France

\*\* Institut de Recherche en Informatique de Toulouse, 118 route de Narbonne, F31062 Toulouse cedex 4, France

l'équipe du Professeur Viel et au débat suscité, d'une part, par l'existence d'une tendance à un excès de cas de leucémies constatées chez les moins de 25 ans, dans le canton de Beaumont-Hague (Viel *et al.*, 1995 ; Guizard *et al.*, 1997 ; Guizard *et al.*, 2001) et, d'autre part, par le lien, suggéré en 1997, entre l'apparition de la leucémie et la consommation de produits de la mer locaux et la fréquentation des plages (Pobel et Viel, 1997).

La méthode d'évaluation du risque de leucémie radioinduite décrite dans (Laurier *et al.*, 2000 ; Rommens *et al.*, 2000) est basée sur la succession de différentes étapes qui permettent la quantification des transferts de radioactivité depuis les rejets des installations nucléaires du Nord-Cotentin jusqu'à l'organe-cible pour le risque de leucémie, la moelle osseuse. Ces étapes sont les suivantes :

- l'inventaire des rejets radioactifs dans l'environnement des quatre installations nucléaires de 1966, date de mise en fonctionnement de la première installation nucléaire dans l'environnement, à 1996, fin de la période couverte par les études épidémiologiques ;
- l'inventaire des mesures de radioactivité dans l'environnement réalisées par les principaux laboratoires ayant réalisé ce type de mesure dans le Nord-Cotentin,
- la confrontation des résultats de mesures environnementaux aux estimations des modèles radioécologiques de transfert dans les milieux et les produits de la chaîne alimentaire pour établir la meilleure estimation possible des niveaux d'activité dans l'environnement,
- la reconstitution des modes de vie de la population des jeunes de 0 à 24 ans du canton de Beaumont-Hague ayant donné lieu à des expositions aux rayonnements ionisants d'origine naturelle et artificielle,
- le calcul des doses individuelles à la moelle osseuse et le calcul de la dose collective après reconstitution de la population locale,
- le calcul des risques individuels de leucémie radioinduite et le calcul du risque collectif qui peut alors être mis en perspective de l'excès de cas observé par les études épidémiologiques.

Cet enchaînement des étapes nécessaires pour aboutir au résultat en termes de risque fait intervenir de nombreuses compétences pluri-disciplinaires dont il faut harmoniser les contributions. Dans le contexte polémique de la mission confiée au GRNC par les ministères de l'environnement et de la santé, la pression sociale imposait des délais d'obtention de résultats incompatibles avec un traitement des incertitudes tel qu'on l'entend classiquement sous le vocable d'analyse des incertitudes. L'estimation des risques a donc été réalisée par une approche déterministe en recherchant à chaque étape les valeurs les plus représentatives des données d'entrée (rejets radioactifs) et des paramètres de transfert et

d'exposition et en recherchant l'accord de l'ensemble de parties prenantes du comité d'experts du GRNC autour de ces choix de valeurs. Par cette approche, le nombre de cas de leucémie estimé du fait des installations nucléaires locales (environ 0,002 cas) est faible comparé à l'incidence de leucémie observée dans les études épidémiologiques (environ 2 cas en excès). Mais les incertitudes autour de cette estimation, a priori nombreuses, n'ont pas été quantifiées et une partie des experts du GRNC a considéré ne pas pouvoir conclure en l'absence de prise en compte des incertitudes (GRNC, 1999).

#### *Analyse d'incertitude du GRNC: l'approche probabiliste*

Dans une deuxième étape, entre 2000 et 2002, le GRNC a alors entrepris la réalisation d'une analyse d'incertitude autour du calcul de risque de leucémie pour les jeunes du Nord-Cotentin (GRNC, 2002). Initialement l'analyse d'incertitudes a été réalisée avec les moyens classiquement utilisés dans les domaines de la radioécologie et de la radioprotection :

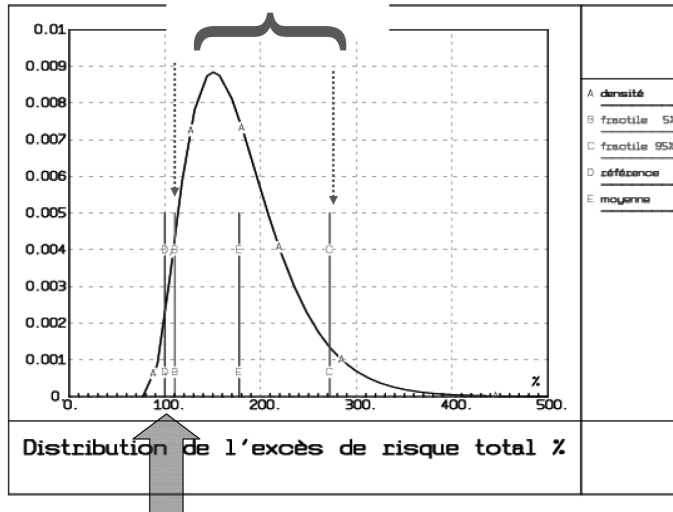
- les incertitudes associées aux paramètres des modèles sont représentées par des distributions de probabilités (Rommens *et al.*, 2001)
- la technique de propagation des incertitudes utilisée est la plus usuelle, la méthode de Monte-Carlo (nombreux tirages aléatoires dans les distributions de chaque paramètre pour déduire une distribution du risque).

Les premiers résultats de calcul par cette approche probabiliste ont montré ses limites dans l'étude Nord-Cotentin pouvant se résumer autour des deux principales constatations suivantes :

- La distribution des valeurs de résultat est très étroite (figure 1). En effet, de façon paradoxale, les incertitudes sur le résultat apparaissent d'autant plus faibles (la distribution des valeurs du risque est d'autant plus étroite) que le nombre de paramètres incertains est élevé.
- La valeur de risque estimée par le premier calcul déterministe du GRNC correspond au percentile 2 de la distribution obtenue par l'approche probabiliste et est notamment en dehors de la gamme de valeurs habituellement considérée comme réaliste, à savoir le domaine entre le percentile 5 et le percentile 95.

Figure 1 : la distribution des résultats obtenue par l'approche probabiliste

La distribution du risque est très étroite  
(facteur 2 à 3 entre le fractile 95 % et le fractile 5 %)



Le calcul déterministe correspond au fractile 2 %

Ces résultats s'expliquent, d'une part, par l'hypothèse d'indépendance des paramètres faite par défaut de connaissance permettant de quantifier les dépendances existantes et, d'autre part, par les choix faits lors du traitement des incertitudes pour représenter la connaissance. L'hypothèse d'indépendance entraîne notamment une diminution de l'occurrence simultanée des valeurs extrêmes (correspondant aux queues des distributions de probabilité) puisque celles-ci ont de moins en moins de chance de survenir au fur et à mesure que le nombre de paramètres incertains augmente. Ce phénomène entraîne alors un resserrement de la distribution. La suite de cet article s'intéresse plus particulièrement à la deuxième cause des résultats obtenus, à savoir les problèmes associés à la représentation de la connaissance. En effet, pour prendre en compte les sources d'incertitudes concernant les paramètres d'un modèle une première difficulté réside dans la façon de quantifier ces incertitudes. En fonction du problème, différents modes de représentation (modélisation aléatoire ou floue) et de quantification (à partir des fractiles ou des densités) des incertitudes sont possibles, mais le choix du mode de représentation et de quantification va influencer sur le résultat final. Le problème de la détermination (mode de représentation et de quantification) des sources d'incertitudes est donc une étape cruciale pour une bonne prise en compte des incertitudes.

### Réflexion sur la représentation de la connaissance et introduction des possibilités dans l'analyse d'incertitude du GRNC

Pour illustrer les problèmes de représentation de la connaissance, considérons le cas d'un paramètre appelé la « profondeur racinaire » ( $P_r$ ) intervenant dans l'équation du transfert des polluants radioactifs du sol vers la plante. Ce paramètre est utilisé par le modèle du GRNC dans l'équation suivante :

$$A_{rac} = A_{sol} \times \frac{F_{rac}}{R_0 P_r} \quad (1)$$

Avec  $A_{rac}$  ( $Bq \cdot kg^{-1}$  frais) : activité massique des végétaux due au transfert racinaire,

$A_{sol}$  ( $Bq \cdot m^{-2}$ ) : activité surfacique du sol,

$F_{rac}$  ( $Bq \cdot kg^{-1}$  frais de végétal par  $Bq \cdot kg^{-1}$  sec de sol en place) : facteur de transfert sol/plante,

$R_0$  ( $kg \cdot sec \cdot m^{-3}$ ) : masse volumique du sol en place,

$P_r$  (m) : profondeur de l'horizon racinaire.

Pour des plantes cultivées, la profondeur de l'horizon racinaire est déterminée par la profondeur de labour et cette information est considérée comme triviale dans le domaine agricole. Pour l'étude Nord-Cotentin, l'ensemble des informations disponibles converge vers un labour « d'environ 20 cm » (IRSN 2001). Toute la difficulté a été ensuite de formuler le « environ 20 cm » en une distribution mathématique et donc de formaliser le « environ ». La solution simplificatrice choisie par les radioécologistes a été de fixer un minimum et un maximum réalistes (respectivement 10 cm et 30 cm) et de proposer une loi triangulaire.

Cependant, si l'on se reporte à l'équation (1), le paramètre  $P_r$  est utilisé au dénominateur de l'équation et c'est donc l'inverse de la fonction triangulaire qui est finalement utilisée dans les calculs d'incertitude. La distribution inverse obtenue est très différente d'une distribution triangulaire :

$$\begin{aligned} f(u) &= (30-1/u)/(100u^2) \text{ si } 1/30 < u < 1/20 \\ &= (1/u-10)/(100u^2) \text{ si } 1/20 < u < 1/10 \\ &\text{avec une valeur maximum en } f(1/20) = 40. \end{aligned}$$

Le problème posé peut alors être illustré de la façon suivante ; si l'on avait demandé aux radioécologistes de fournir la distribution de l'inverse du paramètre  $P_r$ , auraient-ils...

- ... proposé la distribution de probabilités correspondant à l'inverse de celle choisie par le GRNC pour le paramètre  $P_r$  (cas 1) ?
- ... proposé une distribution de probabilités triangulaire de minimum 1/30, mode 1/20 et maximum 1/10 (cas 2) ?

... proposé une distribution de probabilités triangulaire de mode 1/20 (0,05) et des minimum/maximum de +/- 50 % (cas 3)?

L'influence de ces différents choix sur la concentration obtenue dans les légumes racines par le transfert racinaire ( $A_{rac}$ ) a été quantifiée.  $A_{rac}$  nominal étant l'activité obtenue par un calcul déterministe avec  $Pr = 20$  cm, pour le cas 1 :  $A_{rac}$  est comprise dans l'intervalle  $[0,745 \times A_{rac} \text{ nominal}; 1,520 \times A_{rac} \text{ nominal}]$  ; pour le cas 2 :  $A_{rac}$  est comprise dans l'intervalle  $[0,815 \times A_{rac} \text{ nominal}; 1,741 \times A_{rac} \text{ nominal}]$  ; pour le cas 3 :  $A_{rac}$  est comprise dans l'intervalle  $[0,658 \times A_{rac} \text{ nominal}; 1,342 \times A_{rac} \text{ nominal}]$ .

L'objectif de cet exemple n'est pas de trancher quant au choix à faire pour la distribution de valeurs du paramètre  $Pr$  à utiliser dans les calculs car le choix dépend du contexte et de l'objectif de l'étude. L'intérêt de cet exemple est de montrer quelles peuvent être les conséquences sous-jacentes à ce choix et qu'au-delà de cet exemple illustratif, les questions de fond qui sont posées sont de savoir comment s'affranchir de l'influence du choix d'une métrique dans le résultat de calcul d'incertitude? Et de façon plus générale, comment prendre en compte le savoir de type subjectif (« environ » étant un exemple de formulation) dans les calculs d'incertitudes.

La réponse à cette dernière question a été le recours à la théorie des possibilités (Dubois et Prade, 1985). En effet, les incertitudes rencontrées dans les travaux du GRNC, relèvent principalement de l'incomplétude et de l'imprécision de la connaissance et le jugement d'expert est finalement apparu comme indispensable afin de permettre de compenser les méconnaissances. La théorie des possibilités ayant été développée afin de parvenir à un raisonnement approximatif qui ait une certaine ressemblance avec le raisonnement humain (Zadeh, 1975) (Bouchon-Meunier, 1995), cette théorie est apparue comme le moyen de traiter simultanément les connaissances numériques et symboliques comme le jugement d'expert (Dubois *et al.*, 2000).

Ces considérations ont conduit le GRNC à mettre en œuvre une méthode dite possibiliste qui combine probabilité et possibilité pour évaluer les incertitudes du calcul du risque de leucémie radioinduite. Cette démarche est cohérente avec le fait que le choix des distributions de paramètres traduit souvent le défaut de connaissance et l'intervention du jugement d'expert. Cette méthode a permis d'atténuer certains problèmes posés par l'hypothèse d'indépendance des paramètres. En particulier les incertitudes sur le résultat ne décroissent plus lorsque le nombre de paramètres incertains augmente et la distribution des valeurs de risque est plus large (GRNC, 2002). Deux distributions de valeurs du risque, l'une issue de la méthode probabiliste, l'autre issue de la méthode possibiliste, ont ainsi été obtenues pour le risque de développer une

leucémie pour les jeunes de 0 à 24 ans du Nord-Cotentin du fait des rejets radioactifs des installations nucléaires locales. Cette double approche a permis de conclure que les résultats des calculs d'incertitudes ne remettent pas en cause l'ordre de grandeur du risque de référence calculé initialement par le GRNC. La gamme la plus large des valeurs possibles du risque montre qu'il est peu probable que les rejets de routine des installations nucléaires du Nord-Cotentin puissent expliquer l'incidence élevée du nombre de leucémies observées, pour la période 1978-96, chez les jeunes de 0 à 24 ans du canton de La Hague (4 cas observés contre 2 attendus) (GRNC, 2002 ; Rommens *et al.*, 2002).

### **Développement d'un nouveau cadre de traitement des incertitudes**

Pour généraliser le travail initié dans le cadre des travaux du GRNC et proposer un cadre rigoureux de traitement de la connaissance imprécise appliqué au domaine de l'évaluation des risques, une collaboration entre le Bureau de Recherches Géologiques et Minières (BRGM), l'Institut National de l'Environnement Industriel et des Risques (INERIS), l'Université Paul Sabatier de Toulouse (Institut de Recherche en Informatique de Toulouse, IRIT) et l'Institut de radioprotection et de Sécurité Nucléaire (IRSN) a été mise en place. Les objectifs du travail pour l'IRSN étaient les suivants : prolonger la réflexion initiée par le GRNC sur l'utilisation d'outils mathématiques autres que les distributions de probabilités pour représenter et propager les incertitudes et réfléchir à une prise en compte plus rigoureuse des dépendances entre les données d'entrée de l'analyse d'incertitudes. Il s'agissait notamment d'évaluer l'intérêt du recours aux théories plus récentes de Dempster-Shafer (Dempster, 1967 ; Shafer, 1976) et d'étudier les possibilités et les limites de leur application au domaine de la radioécologie. Ce travail a été conduit en deux temps : d'abord le développement d'un cadre mathématique rigoureux pour combiner les différents outils mathématiques et ensuite l'application à un exemple du domaine de la radioécologie (Baudrit *et al.*, 2005a). D'autres exemples d'applications environnementales sont proposés dans (Baudrit, 2005) mais ils ne sont pas développés dans ce papier.

Les différentes théories mathématiques et les outils disponibles pour traiter les incertitudes sont examinés successivement en explicitant, d'une part, leur intérêt pour la radioécologie et en recherchant, d'autre part, les limites de leur utilisation.

### **Présentation des différents outils**

L'incertitude sur les paramètres des modèles utilisés dans l'évaluation des risques peut avoir deux origines. L'information dont on dispose réellement peut être de nature aléatoire due à une variabilité naturelle résultant de phénomènes stochastiques (pluviométrie, débit d'un cours d'eau, nombre de jours d'ensoleillement). Elle peut également être de nature imprécise due à un manque d'information (profondeur racinaire, période d'élimination par décroissance biologique dans un végétal, date de récolte).

### **Calcul d'intervalles**

Le calcul d'intervalles est fréquemment utilisé pour réaliser des calculs d'erreur. Les incertitudes sont alors représentées sous la forme d'intervalles de valeurs et le calcul d'erreur revient à faire un calcul de pire cas et de meilleur cas. L'intérêt de cette méthode réside dans sa simplicité. L'inconvénient majeur de ce type de calcul est son caractère peu informatif en termes de quantification de l'incertitude; notamment il ne permet pas de quantifier dans quelle mesure, un seuil risque d'être dépassé.

### **Théorie des probabilités**

La théorie des probabilités est celle traditionnellement utilisée par les statisticiens. Les incertitudes sont alors représentées par des distributions de probabilités et la technique de propagation la plus usuelle est la méthode de Monte-Carlo. C'est une théorie bien adaptée pour traiter la variabilité mais qui ne rend pas bien compte de l'imprécision et de l'incomplétude. En effet, utiliser une forme de distribution de probabilité précise (normale, lognormale) lorsque la connaissance est incomplète repose sur un choix subjectif et en partie arbitraire.

### **P-box**

Une distribution de probabilité peut aussi être définie par sa fonction de répartition, appelée aussi probabilité cumulée. Un modèle naturel pour donner une approximation d'une distribution de probabilité mal connue est alors de considérer une paire  $F_*$  et  $F^*$  de fonctions de répartition haute et basse qui encadre la distribution de probabilités mal connue et qui définit alors une famille de probabilités (figure 2). L'écart entre les deux fonctions de répartition reflète le caractère incomplet de l'information. Cet intervalle, nommé p-box (Ferson, 2003) correspond à une extension du calcul d'intervalle évoqué précédemment.



Figure 2: représentation d'une p-box

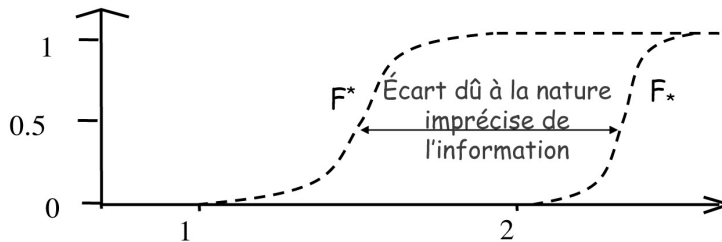
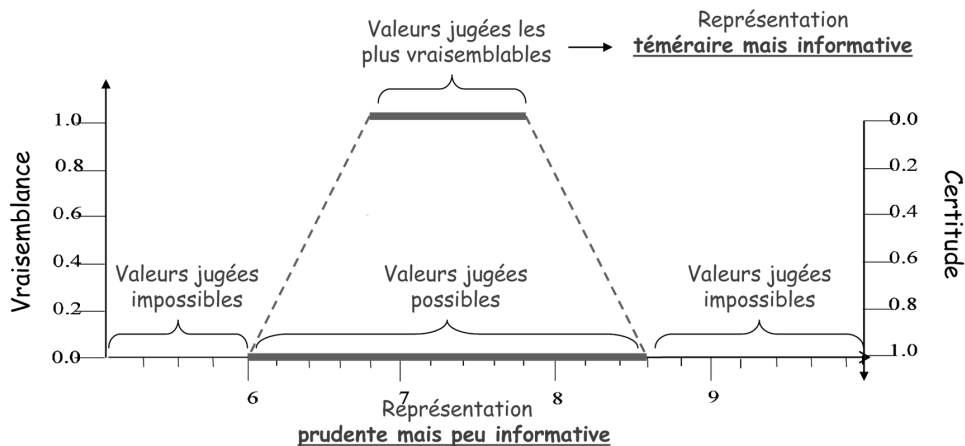


Figure 3: représentation d'une distribution de possibilité



### ***Théorie des possibilités***

Une distribution de possibilité est un outil mathématique qui reflète naturellement le format des informations données par un expert, lorsque celui-ci propose un intervalle de valeurs dans lequel il est sûr que la valeur recherchée se trouve et propose à l'intérieur de cet intervalle, des préférences d'intervalles qui lui semblent plus vraisemblables. Ces préférences permettent alors de définir des intervalles emboîtés de plus en plus restreints qui définissent une information plus précise mais au demeurant moins certaine (figure 3).

### ***Théorie des fonctions de croyance***

La théorie des fonctions de croyance de Shafer (Shafer, 1976) fournit des outils mathématiques permettant à la fois de traiter de l'information

de nature aléatoire et imprécise. Les informations sont alors représentées comme des masses (poids de probabilité) affectées à des intervalles de valeurs qui peuvent être emboîtés ou non et l'ensemble de la connaissance est synthétisé par une distribution de masse appelée aussi fonction de croyance. Toute distribution de probabilité et toute distribution de possibilité peuvent être représentées à l'aide d'une telle fonction ce qui a pour avantage de travailler dans un cadre commun pour traiter l'information qu'elle soit de nature aléatoire (variabilité naturelle) ou imprécise (jugement d'expert).

### ***Cadre général de propagation des incertitudes***

Le cadre général qui rassemble l'ensemble de ces outils est le domaine des probabilités imprécises (Walley, 1991). L'existence de ce cadre unifié est particulièrement importante d'un point de vue technique car il permet de combiner les différents types de connaissance et donc de propager les différents types d'incertitudes associées (variabilité, imprécision, incomplétude de la connaissance). Ce cadre est aussi utile en termes de communication car il permet de traduire les résultats de ces outils mathématiques complexes en distribution de probabilités, outils d'usage plus fréquent pour les scientifiques (Baudrit et al., 2005b). Dans ce cadre, Guyonnet et al. (2003) ont développé une méthode hybride combinant le calcul possibiliste avec la technique dite de Monte-Carlo, méthode qui a été utilisée dans notre application à la radioécologie.

### ***Application à la radioécologie***

L'un des principaux résultats opérationnels de ce travail a été de proposer une typologie des outils mathématiques en fonction de la connaissance disponible et d'étudier l'application de cette typologie à la connaissance disponible en radioécologie. Cette typologie générale et la typologie appliquée qui en résulte sont présentées sur la figure 4.

La figure 4 présente quels sont les outils mathématiques les plus adaptés en fonction de la connaissance disponible; par exemple pour le paramètre de profondeur racinaire  $Pr$  évoqué précédemment, la source d'information principale étant la littérature, la représentation proposée est une distribution de possibilité dont la forme est triangulaire compte tenu des informations disponibles pour établir des valeurs caractéristiques de cette distribution. Pour d'autres paramètres pour lesquels la connaissance repose sur de nombreux résultats de mesures qui montrent que l'incertitude principale est la variabilité naturelle (pour ce travail, plus de 200 mesures de la masse volumique du sol étaient disponibles), la représentation proposée est une distribution de probabilité. La forme et les valeurs caractéristiques de la distribution sont alors obtenues à partir du traitement statistique des résultats de mesures.

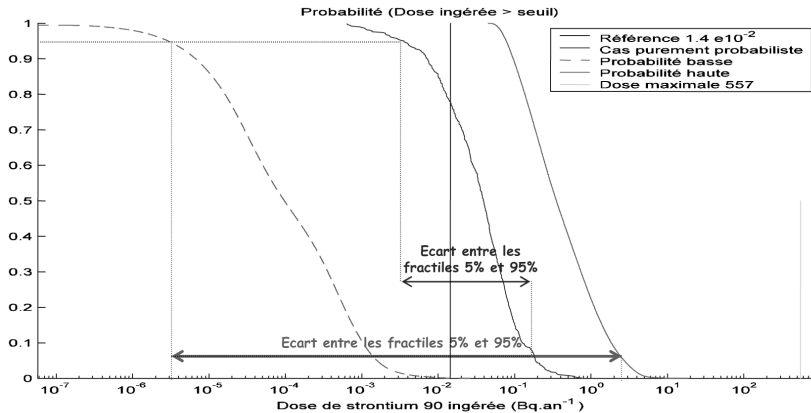
Figure 4: Typologie Connaissance/Outils mathématiques et application dans le domaine de la radioécologie

Modèle mathématique	Connaissance	
Probabilités	Echantillon conséquent Connaissance théorique	
P-boxes	Modèle paramétrique imprécis Echantillon de petite taille	
Possibilités	Avis d'expert [min, max] [min, mode, max]	
Fonction de croyance	Toute information de nature aléatoire et imprécise	

↓

Exemple de paramètre	Source d'information	Représentation proposée
Profondeur racinaire	littérature	Possibilité triangulaire
Masse volumique du sol	nombreuses mesures	Probabilité normale
Rapport de captation	littérature	Possibilité uniforme (= intervalle)
Décroissance biologique	Avis d'experts, quelques mesures, littérature	Possibilité trapézoïdale

Figure 5: Influence de la représentation de l'incertitude autour des paramètres sur le résultat en termes de dose ingérée



Cette typologie a été utilisée pour réaliser un calcul de transfert d'un polluant radioactif (le strontium 90, <sup>90</sup>Sr) depuis un dépôt atmosphérique jusqu'à l'homme, au travers de la consommation d'un aliment (le lait de vache). Les équations du modèle de transfert et les valeurs des paramètres sont présentées dans (Mercat-Rommens *et al.* 2004). On considère dans ce calcul que le dépôt de polluant radioactif (en Becquerel (Bq)) est réparti de façon homogène sur l'année et est intégralement fait sous

forme de dépôt sec. Les végétaux contaminés par le dépôt sont les fourrages constitutifs de la ration alimentaire de la vache laitière : l'herbe et le maïs ensilage. Le lait de vache est ensuite contaminé par le métabolisme alimentaire de la vache. Finalement, l'homme est contaminé au travers de l'incorporation alimentaire de lait de vache et le résultat final du calcul proposé est la dose annuelle par ingestion de lait de vache contaminé en  $^{90}\text{Sr}$  exprimé en  $\text{Bq. an}^{-1}$ . Ce calcul a été réalisé successivement par différentes approches (figure 5) :

- Approche déterministe en utilisant une valeur fixe pour chaque paramètre qui conduit au résultat dit « de référence »,
- Approche par intervalle qui conduit à un résultat sous la forme d'un intervalle entre une dose dite « minimale » et une dose dite « maximale »,
- Approche probabiliste (« cas purement probabiliste ») avec chaque paramètre représenté par une distribution de probabilité qui conduit à un résultat représenté par une distribution de probabilité unique,
- Approche dite « hybride » utilisant la typologie appliquée à la radioécologie évoquée précédemment et conduisant à un résultat représenté par une probabilité basse et une probabilité haute (Guyonnet *et al.*, 2003 ; Baudrit *et al.*, 2005b).

Les résultats de la figure 5 permettent de comparer les incertitudes obtenues sur le résultat en termes de dose ingérée de  $^{90}\text{Sr}$  selon les différentes approches. L'approche probabiliste conduit à sous-estimer l'incertitude totale d'environ un facteur 10 par rapport à l'approche utilisant la typologie de représentation qui tient compte du type de connaissance disponible sur les paramètres du modèle. Cet écart pourrait générer une prise de décision différente dans le cas où la limite admissible pour la dose ingérée se situerait autour de  $0,4 \text{ Bq. an}^{-1}$  (valeur fictive proposée ici à titre didactique). L'approche probabiliste conclurait alors que la probabilité de dépasser ce seuil est quasi-impossible car inférieure à la valeur d'usage de 5 %, alors que cette même probabilité peut atteindre 45 % lorsque l'on prend en compte l'imprécision de la connaissance autour des paramètres.

## Conclusion et perspectives

Ce travail a mis en évidence que l'utilisation systématique de distributions de probabilités pour représenter la connaissance environnementale est souvent subjective et arbitraire et peut conduire à sous-estimer le risque encouru par les populations. Pour dépasser ce problème, les possibilités de recours à différentes théories de l'incertain ont été étudiées (probabilités, possibilités, fonctions de croyance, p-

boxes). Ce recours a nécessité des développements mathématiques pour proposer des outils utilisables pour la représentation et la propagation des connaissances disponibles en sciences environnementales. Ce travail a permis de promouvoir la cohérence entre la manière dont l'information est mathématiquement représentée et l'information réellement disponible et de proposer des méthodes alternatives aux probabilités pour représenter et propager les incertitudes. Il faut cependant poursuivre le travail sur ces théories pour développer les moyens d'une utilisation plus aisée et le recours à d'autres fonctionnalités comme l'analyse de sensibilité floue, les calculs inverses, De plus un effort important est attendu pour faciliter l'utilisation de ces outils dans le cadre de la prise de décision. Les tentatives de mise à disposition de la communauté scientifique des ces outils au travers du logiciel SUNSET© de l'IRSN (Chojnacki and Ounsy, 1996) ou du didacticiel Hyrisk du BRGM (Guyonnet et al., 2005) doivent être poursuivies.

## Bibliographie

- BAUDRIT C., 2005. *Représentation et propagation de connaissances imprécises et incertaines : Application à l'évaluation des risques liés aux sites et sols pollués*. Thèse de l'université Paul Sabatier, Toulouse III, soutenue le 19 octobre 2005, 198 p.
- BAUDRIT C., CHOJNACKI E., MERCAT-ROMMENS C., 2005a. *Influence de la Représentation Mathématique de la connaissance sur les résultats de calcul de risque*. Rapport IRSN, DEI/SESURE n° 2005-14. Fontenay-aux-Roses, 48 p.
- BAUDRIT C., GUYONNET D., DUBOIS D., 2005b. Post-processing the hybrid method for addressing uncertainty in risk assessments. In *Journal of Environmental Engineering*, Vol. 131, Issue 12, pp 1750-1754.
- BOUCHON-MEUNIER B., 1995. *La logique floue et ses applications*. Paris, Addison-Wesley, 257 p.
- CHOJNACKI E., OUNSY A., 1996. Description of the IPSN method for the uncertainty and sensitivity analysis and the associated software: SUNSET. In *ASME/JSME ICONE 4 proceedings*, Louisiana, USA, Vol.3, pp 545-550.
- DEMPSTER A.P., 1967. Upper and Lower Probabilities Induced by a Multivalued Mapping. In *Annals of Mathematical Statistics*, Vol. 38, pp 325-339.
- DUBOIS D., NGUYEN H.T., PRADE H., 2000. Possibility theory, probability and fuzzy sets : misunderstandings, bridges and gaps. In *Fundamentals of Fuzzy Sets*, Dubois, D. Prade, H., Eds : Kluwer, Boston, Mass, pp 343-438.
- FERSON S., GINZBURG L., KREINOVICH V., MYERS D.M., SENTZ K., 2003. *Construction Probability Boxes and Dempster-Shafer structures*. Sandia National Laboratories, Technical report SANDD2002-4015, USA.
- GUYONNET D., BOUGINE B., DUBOIS D., FARGIER H., COME B., J.-P. CHILES J.-P., 2003. Hybrid approach for addressing uncertainty in risk assessments. In *Journal of Environmental Engineering*, Vol. 126, pp 68-78.
- GUYONNET D., MÉNARD Y., BAUDRIT C., DUBOIS D., 2005. *HyRisk - Traitement Hybride des incertitudes en Évaluation des Risques*. Rapport BRGM/RP 53714, [www.brgm.fr/hyrisk/HyRisk\F4\F.pdf](http://www.brgm.fr/hyrisk/HyRisk\F4\F.pdf).

MERCAT-ROMMENS C., RENAUD P., VRAY F., CHOJNACKI E., 2004. *Exemple simplifié proposé comme cas d'étude pour la représentation de la connaissance*. Rapport IRSN DEI/SESURE n° 2004-02, Fontenay-aux-Roses, 39 p.

ROMMENS C., CHOJNACKI E., MERLE-SZEREMETA A., BRÉNOT J. SUGIER A., 2002. La nécessaire prise en compte des incertitudes dans les évaluations de risque: l'exemple du GRNC. In *Environnement, Risque et Santé*, Vol. 1, n° 5-6, pp 276-282.

SHAFER G., 1976. *A Mathematical Theory of Evidence*. Princeton University Press.

WALLEY P., 1991. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall.

ZADEH L.A., 1975. Fuzzy logic and approximate reasoning. In *Synthese* Vol. 30, pp 407-428.