

# Modélisation des incertitudes et variabilités en EQRS

Période : avril 2012 à août 2012

Alain-Claude ROUDOT

Université de Bretagne Occidentale – Laboratoire d'Évaluation du Risque Chimique pour le Consommateur – Brest – France

Mots clés : incertitude, variabilité, modélisation, méthode probabiliste, méthode possibiliste, nombre flou

En évaluation quantitative du risque sanitaire (EQRS), on est souvent amené à modéliser divers mécanismes: transfert de contaminants du milieu vers l'homme, transposition des données animales chez l'homme etc. Dans tous les cas, il existe de nombreuses sources d'incertitudes épistémiques<sup>(1)</sup> sur les paramètres concernés ou de variabilités<sup>(2)</sup> liées à la diversité biologique. Pour assurer la protection sanitaire de la population, le gestionnaire du risque est donc amené à utiliser des facteurs d'incertitude de valeurs potentiellement élevées pour garantir cette protection. Pour autant, rien ne permet de chiffrer d'éventuels dépassements ou de déterminer la fraction de population éventuellement à risque. Cependant, de nombreuses méthodes plus fines existent aujourd'hui, permettant une meilleure prise en compte des variabilités et incertitudes (1, 2). La méthode la plus usitée en EQRS est la méthode dite probabiliste qui transforme une donnée « peu sûre » en distribution de probabilités (courbe centrée sur la valeur de la donnée précédente et d'écart-type non nul): la valeur ne sera plus déterminée une fois pour toutes mais tirée aléatoirement de cette distribution, permettant ainsi la prise en compte d'une variabilité de cette donnée représentant la variabilité (ou l'étendue d'incertitude) des données utilisées. On obtient alors un résultat, dit probabiliste, permettant de mesurer les dépassements des seuils sanitaires et donc d'avoir une meilleure vision de la réalité. Le calcul nécessitera alors de faire un grand nombre de simulations en tirant, à chaque fois, une valeur aléatoirement de cette distribution en utilisant notamment la méthode dite de Monte Carlo<sup>(3)</sup>. Le résultat sera également obtenu sous la forme d'une distribution de probabilités. Comme l'incertitude épistémique et la variabilité sont de nature différentes, il est aujourd'hui préconisé de les traiter successivement en utilisant la méthode de Monte Carlo du second ordre<sup>(4)</sup> (3). Une autre possibilité consiste à se baser sur la théorie de la logique floue<sup>(5)</sup>, en transformant chaque valeur fixe incertaine, en un nombre dit flou<sup>(6)</sup>. Enfin, la méthode hybride a fait son apparition appliquant la méthode probabiliste pour les variabilités, et la méthode floue pour les incertitudes.

## Risques sanitaires liés à l'exposition au plomb et à l'arsenic *via* la consommation d'aliments auto-produits, ainsi que *via* l'ingestion de poussières et particules du sol: un outil probabiliste

Bacigalupo C, Hale B. Human health risks of Pb and As exposure *via* consumption of home garden vegetables and incidental soil and dust ingestion: A probabilistic screening tool. *Sci Tot Environ* 2012;423:27-38.

### Résumé

L'autoconsommation<sup>(7)</sup> peut être une source majeure de contamination alimentaire au plomb et à l'arsenic. Un modèle basé sur la méthode de Monte Carlo du second ordre est proposé afin de prévoir les centiles d'exposition élevée liés à l'autoconsommation et de calculer le quotient de danger<sup>(8)</sup> pour le plomb et l'excès de risque de cancer<sup>(9)</sup> pour l'arsenic. Ce travail a deux objectifs: montrer l'importance de l'autoconsommation dans le risque et examiner l'influence des incertitudes et variabilités dans le résultat final. Les données expérimentales proviennent des agences sanitaires ainsi que de la littérature scientifique. Les équations modélisées proviennent de l'USEPA et de Santé Canada. Les résultats montrent que pour les forts consommateurs de produits du jardin l'autoconsommation de légumes intervient au même niveau que l'ingestion de terre et

poussières pour l'exposition à l'arsenic, et de quatre à dix fois moins pour le plomb. Il est estimé que la différence provient essentiellement de la méthode d'analyse du risque différente pour les effets cancer (As) (effet à seuil de dose) et non-cancer (Pb) (effet sans seuil de dose). Au niveau méthodologique, l'utilisation de la méthode de Monte Carlo du second ordre présente des limitations telles que notamment le temps de calcul prohibitif pour obtenir des résultats stables et la difficulté d'obtention de distributions de probabilités robustes. Ceci se rajoute aux problèmes d'évaluation de consommation et contamination sur le long terme ainsi que de détermination de la biodisponibilité des métaux dans les sols.

### Commentaire

Cet article est discuté ici pour la présentation de l'évaluation du risque *via* un modèle basé sur la méthode de Monte Carlo du second ordre. Bien que théorisée depuis une dizaine d'année dans le domaine alimentaire et environnemental par l'USEPA (3), cette méthode a été très relativement peu utilisée jusqu'à aujourd'hui. En effet, elle nécessite de séparer les facteurs d'incertitudes et de variabilités ce qui crée un modèle à deux niveaux imbriqués nécessitant des puissances de calcul élevées. L'évolution du matériel informatique permet aujourd'hui d'appliquer cette méthode. Cette application montre que si la

boucle de variabilité peut être correctement évaluée, le temps de calcul ne permet toujours pas d'évaluer correctement la boucle des incertitudes. Le résultat présente donc certaines instabilités qui auraient pu être limitées en rééquilibrant en partie ces deux boucles. Toutefois, à ce point près, ce travail illustre l'application stricte de la méthodologie développée à l'USEPA pour la prise en compte des variabilités et incertitudes. En particulier, par rapport à la méthode probabiliste classique (méthode simple de Monte Carlo), il permet de relativiser la notion de 95<sup>e</sup> percentile, trop souvent considérée comme un seuil fixe, en lui adjoignant un intervalle de confiance qui peut être relativement étendu. Malheureusement, cet article ne discute pas de l'interprétation possible des résultats méthodologiques obtenus: en pratique d'EQRS, doit-on considérer la médiane du 95<sup>e</sup> percentile du quotient de danger, ou son propre 95<sup>e</sup> percentile, par exemple? Autrement dit, à quel niveau d'interprétation doit-on faire propager les variabilités et incertitudes?

### Modélisation de l'incertitude en évaluation du risque par la transformation probabilité-possibilité

Tazid A, Palash D. Modeling of uncertainty in dose assessment using probability-possibility transformation. *Int J Comput Applic* 2012; **48** (12):1-7.

#### Résumé

Cette publication présente plusieurs méthodes mathématiques candidates à la modélisation des incertitudes en évaluation du risque et notamment la théorie des possibilités<sup>(10)</sup> issue de la logique floue (4), et de la méthode probabiliste (5). Les auteurs étudient également les conséquences du passage d'une méthode à l'autre. Cette partie théorique est ensuite mise en application sur un exemple d'exposition d'un enfant au plomb suite à une contamination des sols. Le modèle utilisé est issu de l'USEPA, et trois scénarios sont proposés: probabiliste, possibiliste et hybride. La première méthode, où les paramètres sont soit fixes soit sous forme de probabilités, donne le résultat sous la forme d'une courbe de probabilité d'atteindre une exposition donnée, la méthode possibiliste où les paramètres sont exprimés sous une forme fixe ou de nombres flous, conduit à un résultat sous forme d'un nombre flou représentant l'exposition (son support représentant la variabilité maximale, et sa hauteur la probabilité d'apparition des expositions), et la méthode hybride, avec des paramètres fixes, probabilistes et flous, mène à deux courbes représentant, pour chaque exposition, l'intervalle des probabilités maximum et minimum. Les trois méthodes proposées, bien que ne donnant pas les mêmes résultats sont compatibles entre elles. Cependant les méthodes probabilistes seront choisies par les gestionnaires préférant évaluer la probabilité de survenue d'un effet, à sa seule possibilité.

#### Commentaire

Cette publication présente une première partie très mathématique qui peut rebuter le lecteur non théoricien.

Cependant, elle n'est pas nécessaire pour comprendre la seconde partie, applicative, qui est tout à fait accessible et permet de voir l'importance du choix de la méthode d'évaluation des incertitudes sur le résultat. On peut toutefois regretter les données réglementaires obsolètes pour le plomb et l'absence de commentaires et discussion des résultats, notamment sur les raisons entraînant le choix de l'approche probabiliste ou possibiliste dans la définition des paramètres, mais on peut se reporter à deux autres publications des mêmes auteurs (6, 7). Alors que la méthode probabiliste, sous l'influence de la diffusion de logiciels *ad hoc*, se développe en EQRS, on peut noter l'apport important de la théorie hybride, qui permet de différencier variabilités et incertitudes épistémiques pour une plus grande précision du modèle. On peut noter que cette méthode hybride est directement utilisable *via* le recours aux logiciels Hyrisk ou Ramas par exemple.

#### CONCLUSION GÉNÉRALE

La méthode probabiliste est souvent considérée comme un paradigme indépassable pour la modélisation prenant en compte des variabilités et des incertitudes. S'il est vrai qu'elle permet de prendre en compte ces éléments et de les propager jusqu'au résultat final, ce qui est un progrès par rapport à l'utilisation de facteurs d'incertitude très dépendants de jugements d'experts, il est également vrai que les méthodes de prise en compte pratique de ces éléments évoluent rapidement. Aujourd'hui, tant en évaluation de l'exposition, qu'en calcul de valeurs de référence ou qu'en évaluation du risque, la modélisation des variabilités et incertitudes devient de plus en plus prégnante. L'utilisation de la méthode de Monte Carlo du second ordre, de la méthode possibiliste ou d'une méthode hybride, relève du strict choix de l'utilisateur, cependant le problème majeur va reposer sur l'interprétation des résultats obtenus. Ces méthodes donnent des résultats compatibles mais différents, et doivent donc être interprétés différemment. Si les résultats obtenus par un de ces modèles complexes sont réduits à une valeur fixe (la moyenne par exemple) au moment de l'évaluation des risques, on peut clairement se poser la question de l'utilité de ces modèles. D'un autre côté, on peut également se poser la question de la signification pratique du 95<sup>e</sup> percentile de la variabilité, du 95<sup>e</sup> percentile d'une exposition ou d'une valeur de référence. Au vu des possibilités de modélisation aujourd'hui possibles et robustes, il s'avère donc nécessaire de réfléchir à la signification pratique de ces propagations d'incertitudes et variabilités, et de proposer, le cas échéant, des procédures d'interprétation de ces données modélisées. Cette réflexion a été entamée par l'USEPA dans son programme de recherche en toxicologie numérique, et récemment remise en avant par le NRC. (8)

## Lexique

- (1) Incertitude épistémique: incertitude liée à un manque de connaissance
- (2) Variabilité: variation naturelle d'un paramètre, qu'elle soit connue ou non
- (3) Méthode de Monte Carlo: méthode consistant à remplacer les paramètres d'un calcul par des distributions de probabilité. Le calcul est effectué par simulation consistant à tirer aléatoirement une valeur de chaque distribution, à effectuer le calcul sur ces valeurs et à répéter le processus un grand nombre de fois (généralement 10 000 fois). Le résultat aura la forme d'une distribution.
- (4) Méthode de Monte Carlo du second ordre: méthode de Monte Carlo dans deux boucles imbriquées: pour chaque valeur aléatoire de la première boucle il est effectué un ensemble de tirages aléatoires de la deuxième boucle (10 000 tirages par exemple). Le nombre d'itérations de la première boucle est généralement moindre que dans une procédure de Monte Carlo simple (généralement 200 fois) pour une question de durée de calcul. Dans le cas indiqué, le nombre total de tirages est de 2 000 000.
- (5) Logique floue: la logique classique (booléenne) propose deux choix: vrai ou faux. La logique multivariée propose plusieurs choix: vrai, presque vrai, presque faux, faux, par exemple. La logique floue repose sur la généralisation de cette dernière avec une variation continue des possibilités entre vrai et faux. Elle est basée sur des courbes d'appartenance: par exemple presque vrai pourrait être représenté par 80 % de chance que ce soit vrai et 20 % de chance que ce soit faux.
- (6) Nombre flou: ensemble basé sur la logique floue. Un nombre flou est défini par son support qui définit sa variabilité (par exemple les valeurs réelles comprises entre 1 et 3 appartiennent au nombre flou A) et son noyau, valeur pour laquelle l'appartenance est 1 (par exemple le noyau de A est 2, c'est-à-dire que 2 a 100 % de chance d'appartenir à A).
- (7) Autoconsommation: consommation alimentaire issue de denrées autoproduites par le consommateur
- (8) Quotient de danger: rapport entre l'exposition et la valeur toxicologique de référence
- (9) Excès de risque de cancer: également appelé excès de risque individuel: produit de la dose par la pente de la droite d'extrapolation de la courbe effet-dose vers les faibles valeurs.
- (10) Théorie des possibilités: basée sur la logique floue, cette théorie permet de représenter une information incomplète ou incertaine

## Publications de référence

- (1) Cox LA. Confronting deep uncertainties in risk analysis. *Risk Anal* 2012 ;doi: 10.1111/j.1539-6924.2012.01792.x
- (2) Guyonnet D, Côme B, Perrochet P, *et al.* Comparing two methods for addressing uncertainty in risk assessment. *J Environ Eng* 1999;125:660-6
- (3) US EPA. Risk assessment guidance for superfund. Volume 3

Part A- Process for conducting probabilistic risk assessment. Office and emergency and remedial response. 2001 EPA 540-R-02-002

- (4) Zadeh LA. Fuzzy sets. *Inf Cont* 1965;8 (3):338-53
- (5) Dubois D, Prade H. Possibility theory. Plenum Press 1988
- (6) Palash D, Tazid A. Uncertainty modeling in risk analysis: a fuzzy set approach. *Int J Comput Applic* 2012;43 (17):35-9.
- (7) Palash D, Tazid A. A hybrid method to deal with aleatory and epistemic uncertainty in risk assessment. *Int J Comput Applic* 2012;42 (11):37-44.
- (8) NRC. Science & Decisions. Advancing risk assessment. NAP, Washington 2012, 423 p

## Revue de la littérature

Mesa-Frias M, Chalabi Z, Vanni T, *et al.* Uncertainty in environmental health impact assessment: quantitative methods and perspectives. *Int J Environ Health Res* 2012;DOI:10.1080/09603123.2012.678002

## Autres publications identifiées

Qin XS. Assessing environmental risks through fuzzy parameterized probabilistic analysis. *Stoch Environ Res Risk Assess* 2012;26 (1):43-56

*Cette publication montre une application de la méthode possibiliste en contamination environnementale. Son principal intérêt réside dans le fait que l'étude est menée jusqu'à l'évaluation du risque environnemental, dans un domaine où les méthodes probabilistes sont généralement choisies.*

Ben-Haim Y. Why risk analysis is difficult and some thoughts on how to proceed. *Risk Anal* 2012; doi: 10.1111/j.1539-6924.2012.01859.x.

*Il s'agit d'une réflexion sur les incertitudes en analyse de risque et sur leur prise en compte. La question de la robustesse des nouvelles méthodes d'évaluation est notamment envisagée.*

Tatari F, Akbarzadeh MR, Sabahi A. Fuzzy-probabilistic multi agent system for breast cancer risk assessment and insurance premium assignment. *J Biomed Inf* 2012 ; doi : 10.1016/j.jbi.2012.05.004

*Bien que le domaine d'application soit un peu éloigné de l'EQRS, ce travail montre une autre méthode de propagation des incertitudes: la théorie des agents basée sur la création d'agents informatiques intelligents et autonomes.*

## Conflits d'intérêts

Les auteurs déclarent:

- n'avoir aucun conflit d'intérêt;  
 avoir un ou plusieurs conflits d'intérêt.