



Nov
2019

PRISE EN COMPTE DES NOTIONS D'INCERTITUDE DANS LA GESTION DES SITES ET SOLS POLLUES

État des lieux et perspectives

RAPPORT FINAL



En partenariat avec :



REMERCIEMENTS

Cet ouvrage a bénéficié de la relecture et des conseils de :
Yves Duclos (ADEME), Franck Marot (ADEME), Hélène Roussel (ADEME), Guillaume Gay (MTES),
Nathalie Guiserix (Renault),

CITATION DE CE RAPPORT

Guyonnet, D., Dubois, D., de Fouquet, C., Ghestem, J.-P., Leprond, H., Atteia, O., Côme, J.-M., 2019. Prise en compte des notions d'incertitude dans la gestion des sites et sols pollués. Etat des lieux et perspectives. Rapport ADEME. Collection Expertises. Octobre 2019. 61 pp.

Cet ouvrage est disponible en ligne www.ademe.fr/mediatheque

Toute représentation ou reproduction intégrale ou partielle faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause est illicite selon le Code de la propriété intellectuelle (art. L 122-4) et constitue une contrefaçon réprimée par le Code pénal. Seules sont autorisées (art. 122-5) les copies ou reproductions strictement réservées à l'usage privé de copiste et non destinées à une utilisation collective, ainsi que les analyses et courtes citations justifiées par le caractère critique, pédagogique ou d'information de l'œuvre à laquelle elles sont incorporées, sous réserve, toutefois, du respect des dispositions des articles L 122-10 à L 122-12 du même Code, relatives à la reproduction par reprographie.

Ce document est diffusé par l'ADEME

20, avenue du Grésillé

BP 90406 | 49004 Angers Cedex 01

Numéro de contrat : 19COM004814

Étude réalisée pour le compte de l'ADEME par : BRGM

Coordination technique - ADEME : ROUSSEL Hélène

Direction/Service : Villes et Territoires Durables / Friches Urbaines
et Sites Pollués



SOMMAIRE

RÉSUMÉ	5
ABSTRACT	6
1. Introduction	7
1.1. Contexte et objectifs de ce travail	7
1.2. Le cycle de la donnée en gestion des sites et sols pollués	7
1.3. La notion d'incertitude dans l'histoire des sciences	8
1.4. Évolutions récentes des notions d'incertitude en gestion des SSP.....	10
1.4.1. Pratiques courantes en gestion des SSP	10
1.4.2. Approches récentes en matière de représentation et de propagation des incertitudes	12
2. Incertitudes le long du cycle de la donnée en gestion des SSP	14
2.1. Introduction	14
2.2. Études documentaires et échantillonnage	14
2.2.1. Etudes historiques et documentaires	14
2.2.2. Échantillonnage des sols	15
a) Contexte et sources d'incertitude	15
b) Gestion des incertitudes en phase d'échantillonnage des sols	16
2.2.3. Échantillonnage des eaux souterraines.....	20
2.2.4. Échantillonnage de gaz du sol et d'air intérieur.....	20
2.3. Préparation et analyses en laboratoire	22
2.3.1. Introduction	22
2.3.2. Principales sources d'incertitude	23
2.3.3. Gestion des incertitudes	24
a) Obligations pour les laboratoires.....	24
b) Approche de l'incertitude.....	24
c) Essais inter-laboratoires.....	26
2.4. Représentation et interprétation des données	27
2.4.1. Paramètres précis, aléatoires, imprécis,	27
2.4.2. Incertitudes et données spatialisées	31
2.4.3. Critères d'aide au choix du mode de représentation de l'information	32
2.4.4. Propagation des incertitudes dans des modèles prédictifs	33
2.4.5. Incertitudes liées au modèles prédictifs.....	35
a) Le schéma conceptuel	35
b) Modèles analytiques.....	36
c) Modèles numériques	37
2.4.6. Comparer des résultats de calcul de modèles prédictifs avec un seuil précis	38
2.4.7. Comparaison avec un seuil non-précis	38
2.4.8. Utilisation de données spatialisées.....	39
a) Tracer une carte de concentrations d'un site et l'utiliser	40
b) Le fond géochimique urbain : quelle définition ?.....	43
2.5. Décision et communication	44
2.5.1. Une approche multi-partenaire	44
2.5.2. Communiquer sur l'incertitude d'origine épistémique.....	45
2.6. Utilisation pour des mesures de gestion ou de surveillance	47

2.7. Bancarisation des données / métadonnées	47
3. Perspectives de recherche.....	48
3.1. Introduction	48
3.2. Échantillonnage	48
3.3. L'incertitude liée aux analyses de laboratoire	48
3.4. Représentation et utilisation d'information incertaine.....	49
3.4.1. Représentation d'information incertaine	49
3.4.2. Seuils non-précis	49
3.4.3. Respect d'un seuil sous contraintes	49
3.4.4. Méthodes géostatistiques	49
a) Estimation de teneurs fortement contrastées dans les sols : poursuivre l'amélioration des modèles	49
b) Estimation du fond pédo-péochimique anthropisé	49
3.5. Communication avec les parties prenantes.....	50
3.6. Utilisation pour la gestion et la surveillance.....	51
3.7. Bancarisation des données/metadonnées	51
4. Conclusions.....	52
Références bibliographiques	53
Index des tableaux et figures	59
Sigles et acronymes	60

RÉSUMÉ

A l'occasion des 4^{èmes} rencontres nationales de la recherche sur les sites et sols pollués, les 26 et 27 novembre 2019 à Paris, l'ADEME a souhaité une synthèse relative à la notion d'incertitude dans le domaine de la gestion des sites et sols pollués (SSP) et des propositions de pistes de recherche future. Le présent rapport, qui n'a aucune prétention d'exhaustivité compte tenu de l'étendue de la problématique, s'attache à identifier des sources d'incertitude aux différentes étapes du cycle de la donnée en contexte de SSP, à décrire des approches permettant d'appréhender ces incertitudes et à proposer des pistes de recherche.

Le cycle de la donnée considéré va de l'échantillonnage des milieux en contexte de SSP, à l'utilisation des informations collectées pour définir des choix de gestion, en passant par l'analyse, la représentation, l'interprétation, etc. Si on peut définir de manière générique l'incertitude comme étant le fait « *de ne pas être certain* », les travaux de ces trente dernières années illustrent l'importance de la notion de « *source d'incertitude* ». En particulier, le fait de ne pas être certain quant à tels ou tels phénomènes influençant les impacts sur les milieux ou les risques pour la santé, dépend-il d'une variabilité liée au caractère aléatoire de ces phénomènes et qui peut être décrite (incertitude d'origine dite « *stochastique* »), ou alors au caractère incomplet et/ou imprécis de la connaissance que nous avons de ces phénomènes (incertitude d'origine dite « *épistémique* ») ? Pour certains auteurs, la confusion entre ces deux sources d'incertitude constitue l'une des erreurs les plus courantes en matière de gestion des incertitudes.

Il ressort notamment de cette synthèse que s'il n'existe pas de méthode « *passe-partout* » pour gérer les incertitudes de manière générique en contexte de sites et sols pollués, on peut néanmoins promouvoir une certaine cohérence entre le choix des méthodes et la nature des informations dont on dispose dans la pratique. Cette recherche de cohérence devrait favoriser des choix de gestion plus fiables et robustes. Parmi les étapes du cycle de la donnée en contexte de SSP, il convient de considérer également la communication avec les parties prenantes. En effet, la gestion des incertitudes ne peut s'appuyer uniquement sur des approches purement calculatoires, mais devrait également faire appel aux sciences humaines et sociales pour que les incertitudes soient intégrées dans une compréhension partagée des enjeux liés aux sites pollués.

ABSTRACT

On the occasion of the 4th national research meeting on contaminated sites and soils, on November 26th and 27th 2019 in Paris, ADEME called for a synthesis on uncertainties in the area of contaminated site and soil management, including proposals for future research. This report, which has no claim to completeness given the scope of the problem, seeks to identify sources of uncertainty at different stages of the data cycle in a context of contaminated site management, to describe approaches for addressing these uncertainties and also to identify some avenues of research.

The data cycle ranges from the sampling stage (soil, water, air, ...), to using the collected information for defining site management options, through analysis, representation, interpretation, etc. While uncertainty can be generically defined as "*not being certain*", research in the past thirty years or so illustrates the importance of "*uncertainty sources*". In particular, does uncertainty relative to such and such a phenomena affecting impacts or health risks, originate from variability related to the random character of these phenomena and which can be described (uncertainty of so-called "*stochastic*" origin), or rather from the incomplete and/or imprecise nature of our knowledge regarding these phenomena (uncertainty of "*epistemic*" origin)? For some authors, the confusion between these two sources of uncertainty is one of the most common shortcomings in the field of uncertainty management.

It is clear from this synthesis that while there exists no "*one-size-fits-all*" method for addressing uncertainties in a context of contaminated site management, it is possible however to promote consistency between the choice of uncertainty methods and the nature of the information that is actually available in practice. Such an effort should favour more reliable and robust contaminated site management options. Among the different stages of the data cycle in this context, it is also important to consider communication with stakeholders. Indeed, managing uncertainties cannot rely solely on purely computational approaches, but should also call on the humanities and social sciences to integrate uncertainties into a shared understanding of issues associated with these sites.

1. Introduction

1.1. Contexte et objectifs de ce travail

Il est des questions scientifiques qui demeurent d'actualité malgré le passage des années. La gestion des incertitudes est de celles-là. Abordée déjà lors de la première édition des rencontres nationales de la recherche sur les sites et sols pollués, organisée par l'ADEME les 12 et 13 décembre 2002, cette question demeure un axe de recherche important que l'ADEME souhaite continuer à approfondir à l'occasion des 4^{èmes} rencontres, les 26 et 27 novembre 2019.

Ces rencontres ayant pour objectif de faire le point sur les récents acquis en matière de gestion des sites et sols pollués, mais aussi d'identifier des besoins de recherche et de transfert de résultats de la recherche vers les acteurs opérationnels, l'ADEME a souhaité proposer des présentations de synthèse qui éclairent sur l'état actuel des connaissances et qui ouvrent des perspectives. A noter que la présente synthèse traite la thématique des incertitudes, sans aucune prétention d'exhaustivité compte tenu de l'étendue de cette question. Par ailleurs, le contenu de cette synthèse est fortement influencé par les travaux antérieurs de ses auteurs.

Parmi les notions qui sous-tendent ce document, il y a celle apparue depuis une trentaine d'années, selon laquelle la manière de représenter les incertitudes devrait être adaptée si possible au type d'information disponible et à l'objet d'étude. En effet, si l'incertitude peut être définie de manière très pragmatique comme « *le fait de ne pas être certain* », **ce qui prime en matière de gestion des incertitudes, ce sont les « sources d'incertitude »**, qui peuvent être de natures très variées, **ainsi que leurs conséquences pour la gestion des sites et sols pollués**. Un accent tout particulier est donc mis sur les sources d'incertitudes présentes aux différentes étapes du cycle de l'information (la donnée environnementale) dans un contexte de gestion des sites et sols pollués.

1.2. Le cycle de la donnée en gestion des sites et sols pollués

Lord Kelvin (1824 - 1907) aurait dit : « *si on ne peut pas le mesurer, on ne peut pas l'améliorer* ». Cette maxime vaut également en matière de gestion des sites et sols pollués, où la mesure tient une place toute particulière compte tenu notamment de la variété des paramètres pouvant influencer la mobilité et le devenir des polluants ainsi que les expositions et nuisances associées, en fonction de la nature de ces polluants mais aussi des différentes matrices impactées (sols, eaux, gaz, végétaux, etc.).

Un site pollué étant une entité éminemment changeante au cours du temps, l'acquisition de données à des fins de gestion n'est pas nécessairement un processus linéaire, avec un début et une fin, mais plutôt un processus itératif, avec plusieurs étapes. C'est ce qui est illustré de manière schématique à la **Figure 1**, qui se réfère de manière générale à la méthodologie nationale de gestion des sites et sols pollués (DGPR, 2017). En partant d'un certain état des milieux et suite aux études historiques et documentaires, l'investigateur va définir un plan d'échantillonnage pour ensuite effectuer l'échantillonnage proprement dit. L'échantillon est préparé en vue de son analyse, puis les résultats sont représentés et interprétés. Ensuite viennent des étapes de concertation et de communication avec diverses parties prenantes, l'utilisation des résultats pour justifier des mesures de gestion (évacuation de terres polluées, dépollution in situ, ...) et leur utilisation dans un contexte de surveillance des sites ou de collecte de nouvelles données. Tout ce cycle est sous-tendu par un effort de bancarisation de l'information afin notamment de conserver la mémoire de l'état des sites.

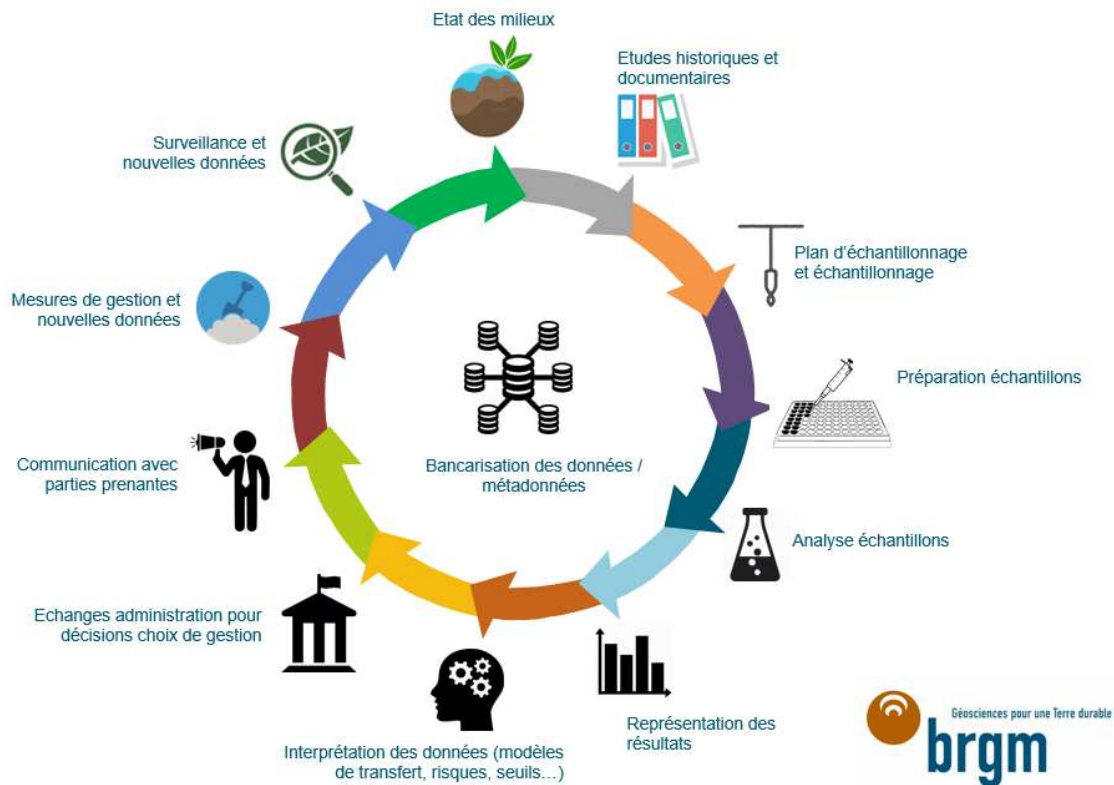


Figure 1 - Cycle de la donnée environnementale en contexte de gestion de sites et sols pollués.

Chaque étape de ce cycle, soit génère de l'incertitude, ou alors est influencée par des sources d'incertitude. Dans ce qui suit, ces étapes sont abordées en mettant en avant des résultats récents et des questionnements dans le but de faire émerger des pistes d'actions futures spécifiques.

1.3. La notion d'incertitude dans l'histoire des sciences

Le souci de formaliser la notion d'incertitude remonte principalement au XVII^{ème} siècle avec les contributions de Pascal, Fermat, et Huygens relatives aux jeux de hasard et à la façon de partager les gains espérés entre les joueurs (Hacking, 1975). C'est la naissance de la théorie des probabilités. Néanmoins, dès le début, cette notion est duale. En effet, soit on modélise les fréquences observées dans un jeu après répétition des coups (calcul des chances sur les jeux de hasard), soit on évalue le degré de croyance d'un individu en l'occurrence d'un événement ou la vérité d'une proposition (comme en témoigne le pari de Pascal). Dans le premier cas, on modélise de façon objective un phénomène aléatoire. Le terme « probabilité », comme attribut de l'opinion, est utilisé dans le second cas. Par la suite, Bernoulli s'intéressera aux deux aspects, même s'il est surtout connu pour avoir prouvé la loi des grands nombres, laquelle concerne la version dite *fréquentiste* des probabilités. Cette loi établit que la moyenne empirique, calculée sur les valeurs d'un échantillon, converge vers l'espérance mathématique lorsque la taille de l'échantillon tend vers l'infini.

Appliquée essentiellement aux jeux de hasard (phénomènes aléatoires), les indicateurs de probabilité développés sont des indicateurs uniques, par opposition aux probabilités imprécises dans le cas d'une origine dite épistémique de l'incertitude (liée au caractère incomplet ou imprécis de l'information disponible) et dont il sera question plus loin (voir section 0) avec notamment la théorie des fonctions de croyance (Shafer, 1976). Mais comme l'a montré l'article de Shafer (1978), Bernoulli a posé également quelques jalons de cette théorie, en considérant le problème de la fusion de témoignages incertains indépendants, pour lequel il ne suppose pas que la somme des probabilités (et non des chances) de propositions contraires soit égale à 1. L'additivité des degrés de croyance n'est pas aussi naturelle que celle des probabilités objectives basées sur les fréquences. La vision additive des probabilités épistémiques résulte

de la conception déterministe du monde (courante dans cette époque-là) qui conduit à penser qu'on peut énumérer des situations équi-possibles et les supposer équiprobables (selon le principe dit de raison insuffisante). Par exemple, Laplace envisageait ainsi des probabilités non-fréquentielles.

L'intérêt pour un calcul de probabilités non-additif et la fusion des témoignages va s'estomper au XVIII^{ème} siècle (si l'on excepte des travaux tels que ceux de Lambert publiés en 1764 ; Shafer, 1978) avec le développement des sciences dures, puis des statistiques au XIX^{ème} siècle. Au XX^{ème} siècle, on assiste à la fois à la formalisation mathématique du calcul des probabilités, issue des statistiques (donc fréquentielles) avec les travaux de Kolmogorov et à la (re)naissance des probabilités subjectives avec ceux de Ramsey (1926) et De Finetti (1937), pour qui « la probabilité n'existe pas ». Dans ce dernier cas, on donne une définition opérationnelle des probabilités subjectives comme degrés de croyance additifs, en considérant que la probabilité subjective d'un événement est mesurée par le prix qu'un individu accepte de payer pour acheter ou vendre un billet de loterie qui rapporte un euro si l'événement se produit. Dans ce cadre, tout degré de croyance non-additif est irrationnel car menant à une perte sûre d'argent. A noter que l'utilisation d'un gain d'un euro dans ce raisonnement présente l'avantage que le prix que le joueur est prêt à payer est directement assimilable à une valeur de probabilité subjective.

Cette représentation probabiliste des degrés de croyance basée sur les paris échangeables a été formalisée en économie dans le cadre de la théorie de la décision par Leonard Savage, au début des années 1950 (Savage, 1954). Il a proposé des postulats de rationalité gouvernant l'ordre de préférence entre des décisions dont le résultat dépend de l'état mal connu du monde. Si ces postulats sont respectés, l'incertitude du décideur peut être représentée par une mesure de probabilité subjective et les décisions possibles classées selon l'espérance mathématique de l'utilité de leurs conséquences. Ce critère de l'utilité espérée est formellement identique à celui proposé par Von Neumann et Morgenstern (1953) quand les fréquences d'occurrence des états du monde sont connues.

Le résultat de Savage a contribué à établir la probabilité subjective comme outil idéal pour représenter la croyance, même quand l'information statistique sur les fréquences est manquante ou la décision non répétable. Mais elle a été mise à mal par les tests empiriques d'Ellsberg, qui dans les années 1960 a montré que, dans leur ensemble, les individus soumis à des choix en présence d'ignorance partielle (tirages dans une urne dont on ne connaît pas la composition) ne suivent pas le critère de l'utilité espérée au sens d'une probabilité subjective unique, mais adoptent une attitude prudente face à l'information incomplète (Ellsberg, 1961). Cela a donné lieu à toute une littérature où l'information incomplète des décideurs est modélisée par des fonctions plus générales que les probabilités et les critères de décision sont de type « intégrales de Choquet » (Chateauneuf et Cohen, 2009).

Les probabilités subjectives sont notamment à la base des méthodes dites *bayésiennes* (Gelman *et al.*, 2003) où l'on fait le choix de modéliser les attentes en début de processus à l'aide de distributions de probabilité uniques, quitte à réviser ces premiers jugements en donnant des poids de plus en plus faibles aux *a priori*, au fur et à mesure que de nouvelles observations sont disponibles. Cette révision est réalisée à l'aide du théorème de Bayes (1763), dit théorème des probabilités conditionnelles.

En intelligence artificielle, la même tendance a été observée depuis les années 1970 pour la modélisation de l'incertitude dans des domaines comme la formalisation du raisonnement dans l'incertain, la fusion d'informations incertaines, par exemple.

Trois théories sont apparues pour traiter l'incertitude due à l'information incomplète, à savoir, dans l'ordre de généralité croissant :

- la théorie des possibilités (Dubois et Prade, 1988), originellement conçue par Zadeh, pour représenter l'information linguistique relative à des valeurs numériques à l'aide d'ensembles flous (Zadeh, 1978). Plus généralement la représentation d'informations d'experts sous la forme d'intervalles de confiance emboîtés prend la forme de distributions de possibilité ;
- la théorie des fonctions de croyance de Dempster-Shafer (Shafer, 1976), qui affecte des probabilités non pas aux états du monde, mais aux états de connaissance sur le monde (on ne donne plus la probabilité qu'un fait soit vrai ou faux, mais on attribue aussi un poids à la possibilité d'ignorance). Dans ce cas, la somme des degrés de croyance en un fait et son contraire n'est plus égale à 1. Cela permet de représenter soit des témoignages incertains, soit des statistiques avec des observations incomplètes ;

- la théorie des probabilités imprécises de Peter Walley, qui reprend la notion de probabilité subjective mesurée par des prix d'achat de jeux incertains (gambles en anglais), mais en supprimant la condition que les prix de vente doivent être égaux aux prix d'achats (Walley, 1991). On justifie alors l'usage de familles de mesures de probabilité dont on extrait des probabilités supérieures et inférieures. Ce cadre capture également le cas de modèles probabilistes incomplètement spécifiés (voir Augustin *et al.*, 2014).

Ces approches permettent de bien distinguer entre incertitude d'origine épistémique versus stochastique, distinction qui prévalait donc déjà au début de l'histoire des probabilités. Ce cadre élargi pour l'incertitude pose de nouveaux problèmes pour l'élicitation et la propagation de l'incertitude. L'existence de ces cadres formels de plus en plus acceptés pour la représentation de l'incertain et qui diffèrent surtout par leur niveau de généralité mathématique, incitent à revoir les méthodes d'analyse de risque car, au lieu de tout représenter par des distributions de probabilité uniques, on doit choisir la représentation en fonction de la nature et la qualité de l'information obtenue par les experts ou par les observations disponibles. L'idée est de coller au mieux à l'information sans en ajouter arbitrairement. De plus, lorsqu'un expert exprime une information incertaine, est-ce une fréquence mal connue, ou un pur degré de croyance? Comment gérer la différence entre fréquence (événements répétables) et croyance (événements uniques ; Flage *et al.* 2016) ?

1.4. Évolutions récentes des notions d'incertitude en gestion des SSP

1.4.1. Pratiques courantes en gestion des SSP

La méthodologie nationale en matière de gestion des sites et sols pollués (DGPR, 2017), évoque plusieurs types d'incertitude : d'échantillonnage, analytiques, ..., mais ne préconise pas de méthodologies permettant d'appréhender ces incertitudes et leurs conséquences en termes de mesures de gestion. Il y est toutefois question d'estimation quantitative d'incertitudes dans le cas d'analyses spatiales de la répartition de paramètres (par ex. de concentrations en polluants), à l'aide de la méthode du krigeage (Chilès et Delfiner, 1999).

En matière d'évaluation des risques environnementaux et/ou sanitaires associés à des sites et sols pollués, une pratique courante dans un cadre opérationnel (par ex. de bureau d'étude) consiste à estimer, souvent par jugement d'expert, des gammes de valeurs plausibles pour les paramètres influençant le risque, puis de rechercher des bornes [min, max] de risque à partir des bornes des paramètres. Cette approche présente au moins un inconvénient significatif. Lorsqu'il s'agit de comparer le risque estimé à un seuil, supposé ici « précis » (voir la section 2.4.6) et en supposant que des bornes de risque [min, max] aient été identifiées de manière rigoureuse (en balayant sur l'ensemble des valeurs possibles des paramètres), on peut distinguer trois situations illustrées dans la **Figure 2**. Si la valeur max du risque estimé se situe en-dessous du seuil de risque acceptable (**Figure 2a**), alors on peut conclure que le risque estimé est acceptable. Si au contraire le seuil se situe en-dessous de la valeur min du risque estimé (**Figure 2b**), alors le risque est inacceptable. L'ambiguïté apparaît dans la situation qui est la plus courante dans la pratique, où le seuil de risque se situe quelque part entre les bornes min et max du risque estimé (**Figure 2c**). Dans ce cas on ne peut pas conclure, en l'absence de discrimination d'évidence à l'intérieur de l'intervalle de risque [min, max], contrairement au cas d'une distribution de probabilité par exemple.

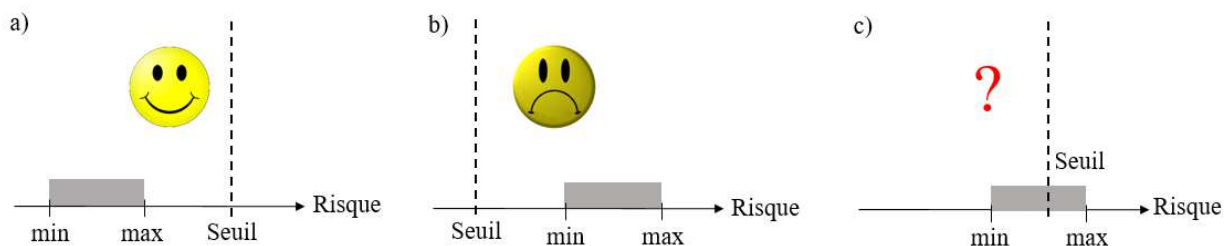


Figure 2 - Trois situations de comparaison d'intervalles [min, max] de risque avec un seuil de risque précis

Pour pallier ce problème, certains recommandent de définir des distributions de probabilité en appliquant un principe dit « d'entropie maximale » (Cover et Thomas, 2006) en référence à un cadre « Bayésien » et aux distributions de probabilité dites subjectives : puisqu'on ne dispose pas d'informations relatives aux distributions de probabilité de certains des paramètres influençant le risque, supposons dans ce cas la distribution la plus simple ; à savoir la distribution uniforme. Mais cette approche présente l'inconvénient d'être non-sécuritaire car elle a tendance à artificiellement minimiser le risque estimé, en favorisant les résultats proches des valeurs moyennes.

Le problème est illustré de manière schématique à la **Figure 3**. Supposons qu'on ait un « modèle de risque », noté C, qui résulte de la somme de deux facteurs de risques notés A et B et qu'on s'intéresse à la valeur moyenne de C. Tout ce qu'on sait sur A est qu'il est situé quelque part entre 0 et 1, tandis que B est quelque part entre 1 et 2. En appliquant un simple calcul d'intervalle (**Figure 3a**) on déduit que C est situé quelque part entre 1 et 3, sans qu'on puisse exprimer de préférence au sein de cet intervalle. Dans ce cas on ne peut pas identifier de valeur moyenne. Si maintenant on suppose que A et B sont représentés par des distributions de probabilité uniformes sur les supports définis par les précédents intervalles (**Figure 3b**), alors la somme de A et B est une distribution triangulaire de mode 2 et de support [1-3]. Dans ce cas la valeur moyenne est égale au mode, c'est-à-dire 2. On voit donc que l'hypothèse de distributions uniformes fournit une valeur précise de la moyenne, mais cette précision peut être illusoire si l'hypothèse de distributions uniformes n'est pas justifiée par l'information disponible.

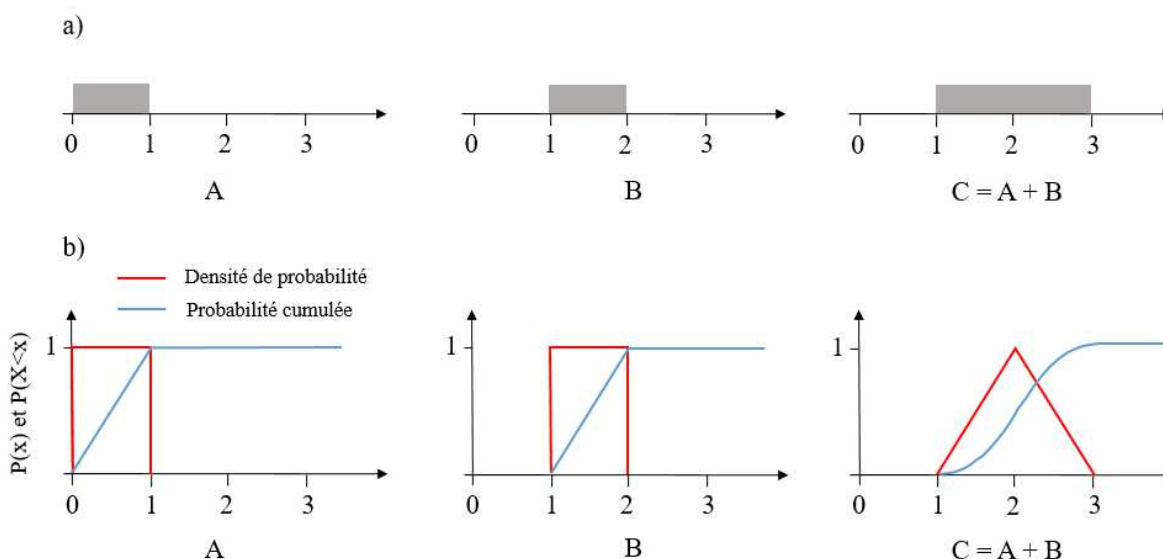


Figure 3 - Illustration de la différence entre un calcul d'intervalle (a) et la somme de deux distributions de probabilité uniformes (b)

Note : la densité de probabilité représente la probabilité $P(x)$ que la variable prenne une certaine valeur, tandis que la probabilité cumulée représente la probabilité $P(X < x)$ que la variable soit inférieure à une certaine valeur x .

Un calcul probabiliste plus « objectif » et cohérent par rapport à l'information disponible sur A et B, tiendrait compte de toutes les distributions possibles sur les supports définis par les intervalles, tel qu'illustré dans la Figure 4. Ce calcul arrive au même résultat que le calcul d'intervalle de la **Figure 3a**.

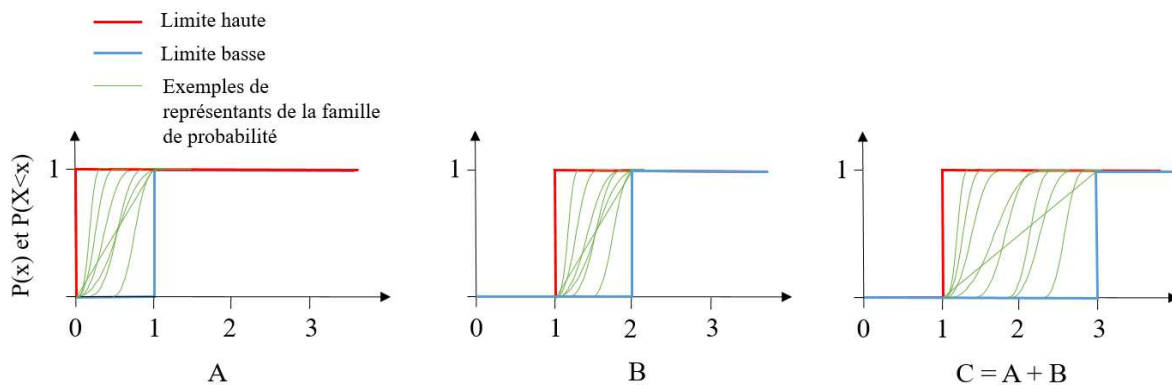


Figure 4 - Calcul effectué à partir des familles de distributions de probabilité définies par l'information disponible

Le caractère non-sécuritaire de l'approche subjectiviste apparaît également lorsqu'on propage, par exemple en utilisant la méthode dite « Monte Carlo » (voir la section 2.4.4), plusieurs distributions de probabilité subjectives sous hypothèse d'indépendance entre les paramètres. En raison de la multiplicité des probabilités : $P(A \text{ et } B) = P(A) \times P(B)$, des scénarios combinant des valeurs très peu probables des paramètres, ont très peu de chances d'être tirées au sort. Or ce sont ces scénarios « limite » qui sont généralement les plus porteurs de risque.

Ces illustrations simples font apparaître une question fondamentale en matière d'incertitude relative à une grandeur X qui est, comme il a été rappelé précédemment, « *le fait de ne pas être certain* » :

- est-ce que je ne suis pas certain de la valeur de X parce que cette valeur est influencée par le hasard (variable aléatoire), mais je connais sa distribution de probabilité ?
- ou est-ce parce qu'il me manque des informations sur X et/ou que mes informations sont imprécises ?

Cette question est au cœur de certaines évolutions de ces 30 dernières années en matière de représentation et de propagation des incertitudes en évaluation des risques, avec notamment des applications dans le domaine des sites et sols pollués.

1.4.2. Approches récentes en matière de représentation et de propagation des incertitudes

Dans les années 90, Scott Ferson (Ferson, 1996 ; Ferson et Ginzburg, 1996), qui travaillait dans le domaine du risque radioactif, a mis en exergue le besoin de distinguer entre variabilité aléatoire et ignorance en matière d'évaluation des risques. On parle également de sources d'incertitude stochastiques (liées à la variabilité aléatoire) versus épistémiques (liées à la méconnaissance). Tandis que les analystes étaient bien armés pour traiter la variabilité aléatoire, en appliquant la théorie des probabilités classique, diverses théories de l'information ont émergé depuis les années 60 pour appréhender l'information incomplète et/ou imprécise : théorie des possibilités (Zadeh, 1978 ; Dudois et Prade, 1988), théorie des fonctions de croyance (Shafer, 1976), etc., qui ont des liens forts et sont généralement des cas particuliers de la théorie, plus globale, dite des « probabilités imprécises » (Walley, 1991). Certaines de ces approches seront illustrées dans le chapitre suivant.

Dans le domaine de la gestion des risques liés aux sites et sols pollués, les chercheurs se sont demandés au début des années 2000 comment combiner des informations à la fois stochastiques et épistémiques en vue de les propager dans des modèles d'évaluation des risques. Les méthodes dites « hybride » (Guyonnet *et al.*, 2003 ; Baudrit *et al.*, 2005) et IRS (Independent Random Sets ; Baudrit *et al.*, 2006) ont été développées, avec notamment un soutien de l'ADEME dans le cadre de la thèse de Cédric Baudrit (Baudrit, 2005) et d'applications aux domaines du risque d'exposition sanitaire à des polluants du sol ou de la contamination des eaux souterraines (Baudrit *et al.*, 2007). Ces méthodes ont été programmées sous la

forme d'un outil « R », en accès libre et dont l'élaboration a bénéficié du soutien de l'ADEME (Guyonnet *et al.*, 2015 ; Rohmer *et al.*, 2017).

Suite à ces développements, les méthodes hybrides s.l. ont été appliquées à des domaines très divers. En ne tenant compte que des publications les plus récentes, on peut citer l'application aux risques sanitaires (Chutia et Datta, 2017), l'analyse de cycle de vie (Igos *et al.*, 2018), l'évaluation des services écosystémiques (Baustert *et al.*, 2018), la dispersion atmosphérique (Abdo *et al.*, 2017) ou encore les risques naturels (Tacnet *et al.*, 2017, Dupouy *et al.*, 2017, Tacnet *et al.*, 2014). Mais également à des secteurs assez éloignés des questions environnementales, comme par exemple le fonctionnement des satellites (Toscani *et al.*, 2018), l'efficacité du transport ferroviaire (Wanke *et al.*, 2018), la construction (Islam *et al.*, 2017), la fiabilité des systèmes électriques (Jiashen, 2018) ou le risque financier (Rebiasz *et al.*, 2017).

Dans des domaines plus proches de la gestion des sites et sols pollués, où l'analyse géostatistique tient une place importante, certains travaux ont cherché à tenir compte du caractère incomplet de la connaissance relative au variogramme (Loquin et Dubois, 2012). D'autres ont combiné les approches possibilistes et Bayésiennes (Tang *et al.*, 2017) pour estimer le risque écologique associé à la présence de certains métaux (Cu, Cd et Hg) dans des sols.

On comprend de ce qui précède, qu'il n'existe pas de méthode unique pour traiter la notion d'incertitude. Tout dépend de la question posée et de l'information dont on dispose pour répondre à cette question. Dans ce qui suit, on aborde la notion d'incertitude en partant de celle d'information collectée dans un contexte de gestion des sites et sols pollués.

- L'étude de l'histoire des sciences montre que dès le début du développement de la théorie des probabilités, cette notion est duale : soit on modélise des fréquences observées, soit on évalue le degré de croyance d'un individu en l'occurrence d'un événement.
- Si l'incertitude peut être définie comme *le fait de ne pas être certain*, ce qui prime en matière de gestion des incertitudes, ce sont les *sources d'incertitude* (les informations disponibles). Ce rapport considère les incertitudes le long du cycle de la donnée en gestion des sites et sols pollués.
- Différentes sources d'incertitude se prêtent à différents modes de représentation de l'incertitude :
 - Si on dispose de valeurs suffisamment précises et en nombre suffisant pour décrire un phénomène aléatoire (incertitude d'origine stochastique), alors les méthodes de probabilité ou statistiques classiques (voire géostatistiques dans le cas d'une variabilité spatiale) s'appliquent.
 - Si on a affaire au contraire à des données éparses ou imprécises, ou encore à du jugement d'expert (incertitude d'origine épistémique, car liée au caractère incomplet/imprécis de l'information disponible), alors d'autres outils peuvent être plus adaptés pour représenter l'incertitude (théorie des possibilités, fonctions de croyance, probabilités imprécises, ...).
- Dans le domaine de la gestion des sites et sols pollués, des approches cohérentes des incertitudes sont nécessaires pour une restitution fiable des informations et des résultats des modélisations prédictives et des évaluations de risques, susceptible de générer des prises de décision robustes et en adéquation avec les informations disponibles.

2. Incertitudes le long du cycle de la donnée en gestion des SSP

2.1. Introduction

Pour faciliter la lecture du document, le cycle de la donnée en gestion des sites et sols pollués de la **Figure 1** est repris ci-dessous pour y inclure les numéros de section correspondant aux différentes étapes du cycle.

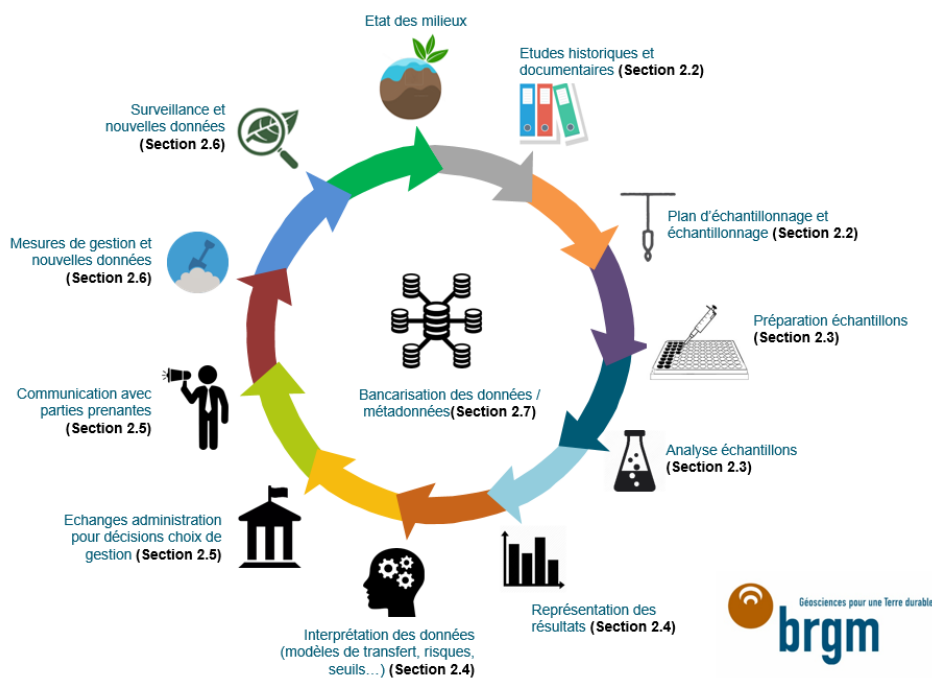


Figure 5 - Cycle de la donnée environnementale en contexte de gestion de sites et sols pollués

Tandis que ce chapitre présente une synthèse sur les incertitudes en jeu et sur les travaux réalisés pour les traiter, le chapitre 3 regroupe des perspectives de recherche identifiées pour les différentes étapes.

2.2. Études documentaires et échantillonnage

2.2.1. Etudes historiques et documentaires

Dans la pratique de la gestion des sites et sols pollués, les études historiques et documentaires servent notamment à orienter les reconnaissances ultérieures. Elles consistent essentiellement à recenser les activités et pratiques exercées, la localisation des installations sur le site, l'évolution de l'emprise du site au cours du temps, les incidents ou accidents passés, l'utilisation de remblais et les polluants susceptibles de se retrouver aussi bien sur le site que dans les milieux avoisinants, etc. Si ces études ne sont pas réalisées de manière satisfaisante, en négligeant par exemple certaines activités historiques sur le site ou les types de produits stockés, etc., les erreurs qui en découleront pourront avoir des conséquences très dommageables sur les choix de gestion ultérieurs.

Les études historiques et documentaires ont donc une influence majeure sur les incertitudes qui affectent un projet. Elles servent, à un stade très amont d'un projet, à déterminer si le site est potentiellement pollué ou si la pollution est avérée. En effet, la première prestation décrite dans la norme NF X 31-620-2 (AFNOR, 2018) est LEVE pour « *Levée de doute* », ce qui illustre bien l'importance accordée aux incertitudes dans la démarche à adopter, dès la phase initiale d'une étude.

2.2.2. Echantillonnage des sols

a) Contexte et sources d'incertitude

Si les laboratoires d'analyse chimique environnementale sont depuis longtemps habitués à estimer leurs incertitudes de mesure (voir section 2.3), la situation est différente pour d'autres acteurs du domaine des sites et sols pollués (bureaux d'études, entreprises de travaux, ...). Ce constat n'est pas lié à un manque de volonté de la part des acteurs, mais plutôt à différents facteurs liés à la complexité d'un site pollué (hétérogénéité du sol, étendue spatiale de la pollution, contraintes d'accès à la zone polluée, etc.) et à un manque d'outils et de référentiels de mesure permettant l'évaluation précise voire la quantification des incertitudes en phase notamment d'échantillonnage.

La profession s'accorde, néanmoins, autour de règles tacites visant à évaluer de manière qualitative les influences de certaines pratiques.

Sans rechercher l'exhaustivité, les principales sources d'incertitude susceptibles d'influer sur l'échantillonnage et donc, plus globalement, sur la gestion d'un site pollué, peuvent être identifiées et listées de la manière suivante :

- Lors de l'étude historique et documentaire, où l'identification d'activités, de substances (voir <http://ssp-infoterre.brgm.fr/matrice-activites-polluants>) et de pratiques susceptibles de générer une pollution dépend des informations collectées, de leur qualité et des possibilités de croisement entre elles ;
- Lors de la description du contexte environnemental où l'hétérogénéité intrinsèque des milieux est une source importante d'incertitudes, ainsi que la connaissance souvent très partielle d'un milieu et à une échelle trop générale (par exemple, le sens d'écoulement d'une nappe au cœur d'une agglomération ou encore la délimitation d'un panache de pollution hors site) ;
- Pour la phase de diagnostic, c'est-à-dire lors de la phase d'acquisition de données au droit du site, les incertitudes sont liées à plusieurs aspects tels que la capacité d'accès à une zone précise du site, la réalisation d'investigations ponctuelles alors que les milieux sont dynamiques (eaux souterraines, gaz du sol, air intérieur), la représentativité des investigations réalisées suivant la stratégie d'échantillonnage et souvent dans un calendrier contraint (profondeur de prélèvement, nombre d'échantillons constitués, quantité de matrice prélevée, ...) ou encore les moyens mis en œuvre et les opérations techniques de prélèvement (protocoles d'échantillonnage, conditionnement, transport) ;
- Les délais d'enchaînement entre les études et les travaux de réhabilitation d'un site sont également générateurs d'incertitudes puisque cette déconnexion, qui peut prendre plusieurs mois ou plusieurs années, ne tient pas compte de la migration d'une pollution. Celle-ci va ainsi être caractérisée à un moment donné et des scénarios de gestion vont être établis mais leur mise en œuvre effective peut s'effectuer dans de toutes autres conditions et avec une extension bien plus importante de la pollution que celle étudiée ;
- Lors des travaux de réhabilitation, il est courant de définir une marge de sécurité à partir des hypothèses de dimensionnement d'une unité de traitement ou d'un chantier afin d'avoir un peu de latitude au cas où certains facteurs appréhendés à partir d'essais en laboratoire et/ou de terrain (rayon d'action, débit d'extraction, cinétique de biodégradation, ...) ne seraient pas pleinement atteints ;
- Enfin, lors de la phase de réception, les incertitudes liées à l'échantillonnage sont également très importantes puisqu'elles dépendent directement du nombre de points de contrôle et des modalités et protocoles employés, voire également de la variabilité des concentrations dans certains contextes (phénomène d'effets rebonds à prendre en compte).

b) Gestion des incertitudes en phase d'échantillonnage des sols

Obligations liées à l'échantillonnage des sols ?

Il faut souligner en premier lieu qu'il n'existe pas à l'heure actuelle, pour la phase d'échantillonnage et plus globalement pour l'ensemble du processus de gestion d'un site pollué, de guide méthodologique qui traite spécifiquement de la gestion des incertitudes tout au long d'un projet. Néanmoins, l'identification des incertitudes et leur influence sur des résultats sont des notions présentes dans la méthodologie nationale de gestion des sites et sols pollués d'avril 2017, mais également dans la plupart des prestations définies dans la norme NF X 31-620 (prestations globales et élémentaires des parties 2 et 3 de la norme : A200 et suivants, A270, B120 et B130, ...).

Il est ainsi exigé qu'un prestataire en sites et sols pollués, lors de prélèvements, mesures, observations et/ou analyses sur un milieu, réalise une discussion des incertitudes sur les résultats obtenus compte de la méthode ou de l'équipement utilisé. De la même manière, lors de la phase d'interprétation des résultats (prestation A270), une présentation des limites et incertitudes découlant des investigations réalisées et une discussion de leurs influences sur les résultats sont demandées.

Au niveau international, la notion d'incertitude en phase d'échantillonnage est abordée dans au moins deux normes : NF ISO 18400-102 (§ 5.3 ; ISO, 2017a) relative au choix et à l'application des techniques d'échantillonnage et NF ISO 18400-104 (§ 5.8 ; ISO, 2017b) qui traite, en particulier, des stratégies susceptibles d'être déployées pendant l'échantillonnage. Ce dernier document présente d'ailleurs une annexe (informatrice) dédiée à l'évaluation et à la modification de l'incertitude de l'échantillonnage et propose différentes méthodes pour mieux la caractériser (méthode de l'échantillonnage en double, méthode OCLI, ...).

Enfin, il peut être mentionné l'existence de normes spécifiques sur certaines pratiques de l'échantillonnage, comme la norme ISO 22155 (ISO, 2016) : « Qualité du sol – Dosage des hydrocarbures aromatiques et halogénés volatils et de certains éthers par chromatographie en phase gazeuse - Méthode par espace de tête statique », qui vise à définir des exigences pour les méthodes d'échantillonnage des sols en vue de la recherche de composés volatils, pour limiter les pertes (par exemple par volatilisation) et donc les incertitudes lors de cette phase.

Essais d'inter-comparaison des techniques d'échantillonnage des sols

Au niveau français, deux études se sont plus spécifiquement concentrées sur la comparaison de protocoles de prélèvement et d'échantillonnage des sols pour l'analyse de composés organiques (notamment les composés volatils). Il s'agit de :

- Etude réalisée par Burgeap pour le compte de l'ADEME (ADEME-BURGEAP, 2015) ;
- Etude réalisée par le BRGM pour le compte du MTES (Favereaux et al., 2019).

Une photographie d'une des zones investiguées lors de l'essai d'inter-comparaison du BRGM en 2018 est présentée en **Figure 6**. Ces études se sont efforcées à déployer sur le terrain, dans différents contextes de sol, des techniques d'investigations distinctes (tarière pleine, carottier sous gaine, carottier à gouges) couplées à plusieurs méthodes d'échantillonnages (flacons, cylindres d'échantillonnage et flacons pré-remplis de méthanol). Des illustrations des techniques de forage et méthodes d'échantillonnage mises en œuvres lors de l'essai d'inter-comparaison du BRGM en 2018 sont présentées en **Figure 7** et **Figure 8**.

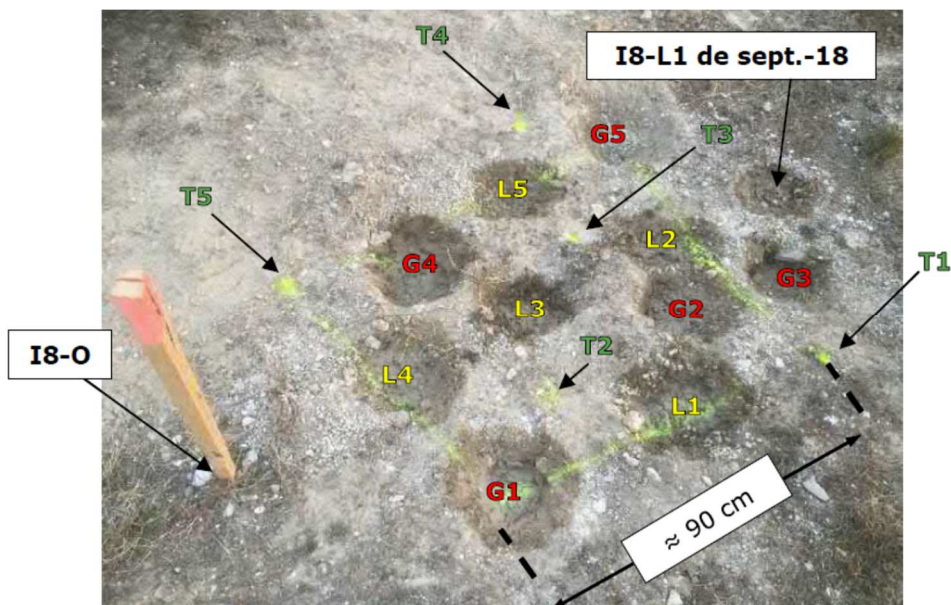


Figure 6 – Vue d'une zone investiguée lors de l'essai d'inter-comparaison du BRGM (2018)

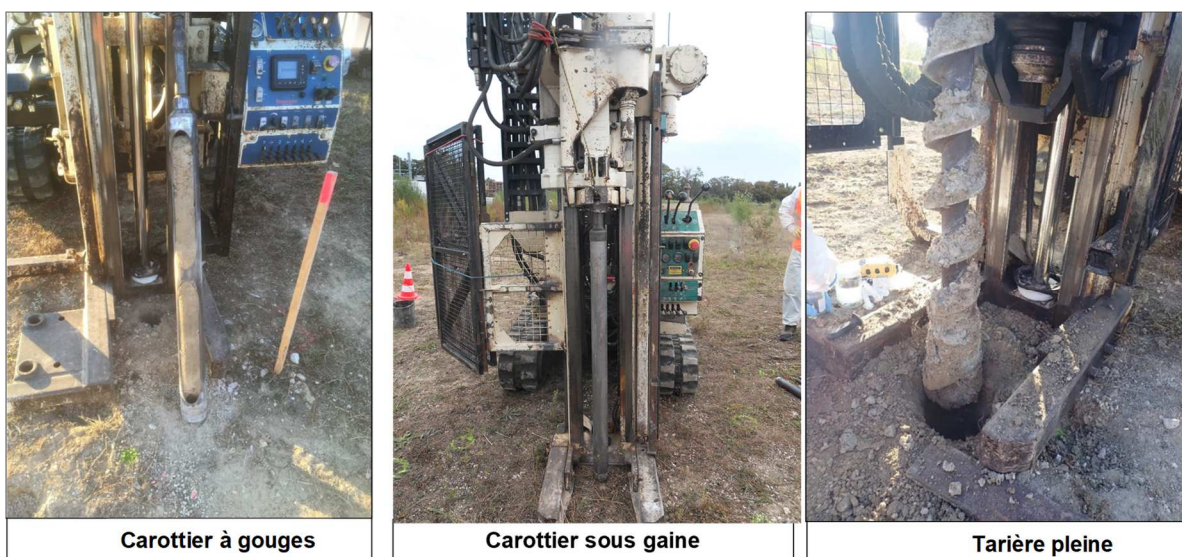


Figure 7 – Techniques de forage mises en œuvre lors de l'essai d'intercomparaison du BRGM (2018)

Les résultats de ces études ne permettent malheureusement pas de quantifier précisément les incertitudes liées aux méthodes de prélèvements et aux pratiques d'échantillonnage : il n'existe en effet à l'heure actuelle aucun abaque ou méthode permettant de corréler une technique de prélèvement ou méthode d'échantillonnage donnée à une incertitude quantifiée de manière précise. Néanmoins, malgré des résultats ponctuellement différents selon le contexte (en fonction de la typologie des matrices notamment : sables, limons, marnes), les études réalisées conduisent à recommander l'utilisation d'un atelier de sondages carottés sous gaine (Liner) pour caractériser et délimiter une source de composés volatils, mais également pour déterminer un profil de concentrations sur la zone investiguée. Elle permet de mieux conserver l'intégrité des échantillons de sols et éviter leur remaniement et la perte de composés volatils.



Echantillonnage en Flacon (F)



Echantillonnage en kit Méthanol (M)



Echantillonnage en Petit Cylindre (CP)



Echantillonnage en Grand Cylindre (GC)

Figure 8 – Méthodes d'échantillonnage mises en œuvre lors de l'essai d'inter-comparaison du BRGM (2018)

Concernant l'échantillonnage des sols, si les deux études sont plus contrastées, elles mettent néanmoins en évidence l'intérêt de la méthode du flacon pré-rempli de méthanol pour permettre une caractérisation plus fine des composés volatils (en terme d'accès à la variabilité des concentrations), ainsi que celui du cylindre inox de 16 ml pour certains types de sols. L'utilisation du flacon, qui reste une méthode très utilisée, semble davantage réservée à l'identification de la présence de polluants plutôt qu'à la caractérisation précise des concentrations. Cette méthode nécessite de suivre un protocole rigoureux destiné à limiter autant que possible les pertes par déstructuration du sol et exposition à l'air. Le flacon peut être utilisé pour la recherche de composés moins volatils (HCT C10-C40).

Les études réalisées montrent également que la position du prélèvement sur une carotte de sol est primordiale et nécessite une bonne expérience du préleveur et la constitution de plusieurs échantillons de sols pour analyses. En effet, des variations ont été observées au sein même d'une carotte de 1 m de longueur supposée homogène en terme de concentration et de lithologie. Ces variations sont liées à l'hétérogénéité du sol et aux différences de comportement/mobilité des polluants et doivent être prises en compte dans la stratégie d'échantillonnage.

Ainsi, une exploitation croisée de l'ensemble des résultats de ces deux études (ADEME et BRGM), et d'autres essais éventuellement disponibles dans le cadre de travaux de recherche, est encore à réaliser afin d'élaborer un guide de bonnes pratiques avec prise en compte des contextes de terrain (polluants recherchés, lithologies, gammes de concentrations, ...). L'aboutissement de ce travail qui va démarrer est envisagé pour fin 2020/début 2021.

Quelques recommandations pour limiter les incertitudes liées à l'échantillonnage des sols

Si les incertitudes globales combinent des éléments d'incertitude découlant à la fois des incertitudes d'analyse associées aux modes opératoires des laboratoires (y compris la phase de préparation) et des incertitudes liées à la phase d'échantillonnage, il est reconnu (cf. Norme NF EN ISO 18400-104) que ces dernières sont généralement les plus importantes et les plus difficiles à estimer. Afin de réduire autant que possible ces incertitudes, plusieurs recommandations peuvent être proposées ¹:

1. Densifier le nombre de points de prélèvement, notamment dans un secteur ou une zone donnée. Les différents documents consultés s'accordent pour recommander la constitution de clusters d'échantillons, afin de ne pas laisser reposer une stratégie de gestion sur un seul ou quelques échantillons épars ;
2. Assurer une continuité entre les opérateurs de terrain. En effet, la multiplication des opérateurs sur le terrain peut conduire à une hétérogénéité des protocoles et modalités d'échantillonnage et donc à une plus grande variabilité des résultats (exemple : échantillonnage ciblé exclusivement sur la pollution ou sur un faciès plus global de sols, échantillonnage ponctuel versus constitution d'un échantillon composite, ...)
3. En lien avec le deuxième point, définir avant l'intervention sur site un protocole d'échantillonnage adapté au contexte et aux polluants recherchés, puis conserver le même protocole de prélèvement et, si possible, le même laboratoire d'analyse (même méthode de préparation de l'échantillon et même méthode analytique) ;
4. Raisonner en tenant compte de la granulométrie des particules de sol sur lesquelles va porter la préparation de l'échantillonnage au laboratoire et l'analyse. S'il est reconnu classiquement que la pollution est concentrée dans les fractions fines (inférieures à 2 mm), la pratique du tamisage des sols avant prétraitement n'est pas toujours adaptée et peut avoir de fortes répercussions lors de la réalisation d'un bilan massique en interpolant les concentrations mesurées sur la fraction fine à la totalité de la masse d'un échantillon et donc à la cubature de toute une zone polluée (projet CAFRAGES piloté par l'ADEME en cours). A l'inverse, cette même fraction inférieure à 2 mm peut aussi être considérée comme trop grossière pour la prise en compte de l'exposition par ingestion de particules de sol, où d'après la littérature on va s'intéresser plutôt à la fraction inférieure à 250 µm, voire plus fine encore ;
5. Avoir une vision intégrée de la situation environnementale au travers d'un schéma conceptuel adapté : pour cela il convient d'associer et de corrélérer les résultats obtenus sur les sols et les eaux aux données disponibles sur les autres milieux (par exemple : gaz du sol en présence de volatils, voire d'air intérieur) ce qui permet de mieux comprendre les phénomènes de migration d'une pollution et d'être plus précis lors de calculs de risques sanitaires ;
6. Conjuguer mesures de terrain et prélèvement d'échantillons pour l'analyse en laboratoire. Ce type d'approche permet de multiplier les mesures de terrain et de les fiabiliser par des analyses en laboratoire en optimisant le programme d'investigation. Il y a lieu de bien vérifier certains aspects comme la corrélation entre les méthodes et de bien veiller à l'adéquation et au réglage des appareils de terrain au regard du contexte (spectromètre à fluorescence X par exemple) ;
7. Réaliser des interventions enchaînées les unes par rapport aux autres avec un laps de temps entre elles le plus réduit possible, afin de limiter les évolutions du ou des milieux. Cet aspect est particulièrement sensible dans le cas de milieux dynamiques comme les eaux souterraines, les eaux superficielles, les gaz du sol et l'air ;
8. Raisonner sur des chroniques de résultats pour dégager une tendance et interpréter la situation au regard de celle-ci plutôt que de conclure sur la base d'un seul jeu de valeurs à l'issue d'une seule campagne. Si cette affirmation paraît évidente, elle ne tient pas compte d'autres contraintes d'un projet, par exemple, une durée souvent très courte d'intervention sur site. Néanmoins, elle mérite d'être soulignée tant elle est importante dans différents contextes (par exemple, calcul de risques sanitaires, réception d'une installation de traitement in situ, ...). Pour autant, selon les incertitudes qui persistent à la réception d'une étude ou de travaux, la poursuite de l'action par un programme de surveillance de la qualité de certains milieux peut aider à consolider les résultats et sécuriser la gestion d'une pollution.

¹ La hiérarchisation proposée a vocation à assurer une meilleure compréhension des points développés. Elle ne préjuge pas de l'importance des incertitudes associées (pas de quantification croissante ou décroissante des incertitudes).

2.2.3. Échantillonnage des eaux souterraines

Dans le domaine des eaux souterraines, l'échantillonnage est spécifique pour deux raisons majeures : la distribution des concentrations est orientée par les écoulements et la nappe peut avoir une épaisseur importante. En ce qui concerne le premier point, l'incertitude est principalement générée par la variabilité multi-échelle de la conductivité hydraulique (Bayer et al., 2011) qui se traduit souvent par la notion de dispersion à large échelle (Gelhar et al., 1992). De très nombreux travaux théoriques existent qui sont rarement transposables d'un site à l'autre (Dagan, 1990). Ainsi, en ce qui concerne l'échantillonnage, la « continuité spatiale » théorique d'un panache ne permet de placer les forages que très grossièrement, et il peut exister des différences d'un ou plusieurs ordres de grandeurs sur quelques mètres de distance. Dans certains cas, il existe aussi une variabilité temporelle des concentrations liée à la variabilité de la direction de l'écoulement (Goode and Konikow, 1990).

De façon encore plus marquée, la plupart des études démontrent une variabilité verticale qui peut atteindre plusieurs ordres de grandeur, notamment sur les sites pollués proches des zones sources. Cette variabilité provient à la fois de l'hétérogénéité du milieu souterrain, mais aussi de la mise en place de la pollution sous forme de phase organique (Côme et al., 2006). En présence d'une variabilité verticale la technique d'échantillonnage va bien sûr avoir une conséquence sur les concentrations mesurées. C'est pour cela que sur tous les sites de recherche et de nombreux sites nord-américains des piézomètres multi-niveaux sont utilisés (Puls and Paul, 1997).

En ce qui concerne les ordres de grandeur de la variabilité, quelques rapports et publications indiquent les grandes lignes. Proche des zones sources, les concentrations peuvent varier d'un facteur 100 sur 1 à 2 m d'épaisseur (Rivett et al., 2001). L'interface entre un panache et la zone non contaminée est souvent de quelques cm sur la verticale (Klenk and Grathwohl, 2002). En latéral, le traçage sur un site hétérogène a démontré des variations de plus d'un ordre de grandeur sur quelques décimètres (Harvey and Gorelick, 2000).

Différents comparatifs ont été effectués sur les méthodes d'échantillonnages et leur rôle dans l'incertitude sur les concentrations des eaux prélevées (Barnier et al., 2013; Dubearnes et al., 2005; INERIS, 2018; Kulkarni et al., 2015; Quiot and Guerin, 2018; Rollin et al., 2008). Mais il ne semble pas qu'il y ait de synthèse de ces diverses études permettant de générer des données consolidées sur les incertitudes.

2.2.4. Échantillonnage de gaz du sol et d'air intérieur

L'échantillonnage de gaz des sols ou d'air intérieur dans un bâtiment pose des problématiques spécifiques en termes d'incertitudes, induites notamment par la forte variabilité à la fois spatiale et temporelle des concentrations en polluants. La variabilité spatiale est principalement générée par la diffusion de la pollution, la variabilité multi-échelle des propriétés du milieu souterrain mais également de la répartition de la pollution sous forme de phase organique. Concernant les ordres de grandeur de cette variabilité, la littérature indique que les concentrations dans les gaz du sol peuvent varier d'un facteur 100 sur quelques mètres d'épaisseur (cf. par exemple **Figure 9**).

La variabilité temporelle des concentrations gazeuses est induite par les mécanismes en jeu dans le système étudié, qui comprend a minima le sol, la nappe et l'air extérieur, mais aussi dans le cas d'un bâtiment, une dalle béton et l'air intérieur.

Pour les concentrations dans les gaz du sol ou dans l'air intérieur, 4 échelles temporelles sont à considérer :

- horaire (effet immédiat associé à une variation de pression / ventilation) ;
- journalière (cycle diurne associé aux variations jour/nuit des conditions météorologiques) ;
- pluri-journalière (liée au rééquilibrage des pressions faisant suite à la pluie, etc.) ;
- saisonnière (liée aux variations du transport diffusif ou convectif).

Sur un plan conceptuel, le transport des gaz du sol est induit par les phénomènes hydro-météorologiques suivants : infiltration de la pluie, variation du niveau de nappe, variation de la pression atmosphérique et vitesse du vent. Et les facteurs à considérer en priorité sont : la température, la teneur en eau du sol, le niveau de la nappe d'eau souterraine, la pression atmosphérique et la vitesse du vent.

Concernant l'air intérieur, les facteurs à considérer en priorité sont ceux décrits pour les gaz du sol, ceux liés au bâtiment, à sa ventilation, son chauffage (variabilité du tirage thermique), etc., mais également les facteurs associés aux comportements des occupants (tabagisme, diffuseur d'ambiance, stockage ou utilisation de produits d'entretien, etc.). Les Figures ci-dessous illustrent des variations temporelles de concentrations mesurées dans les gaz du sol et l'air intérieur qui excèdent un facteur 10 sur des échelles de temps de quelques mois pour les gaz des sols et de seulement quelques heures pour l'air intérieur.

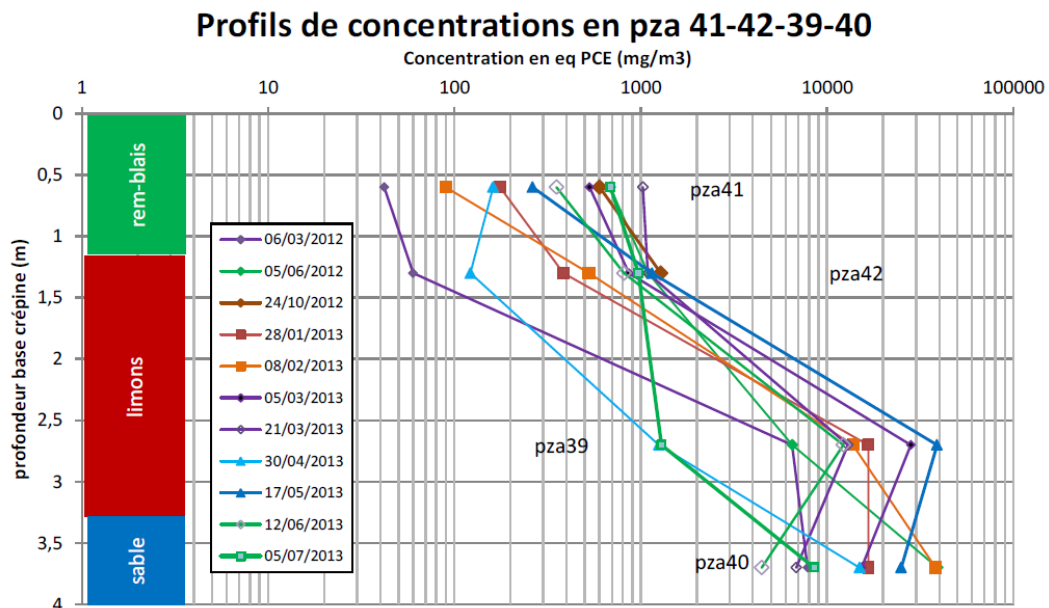


Figure 9 – Variabilité spatio-temporelle des gaz du sol, mesures sur un quadruplet de piézairs (site atelier FLUXOBAT, Traverse et al., 2013)

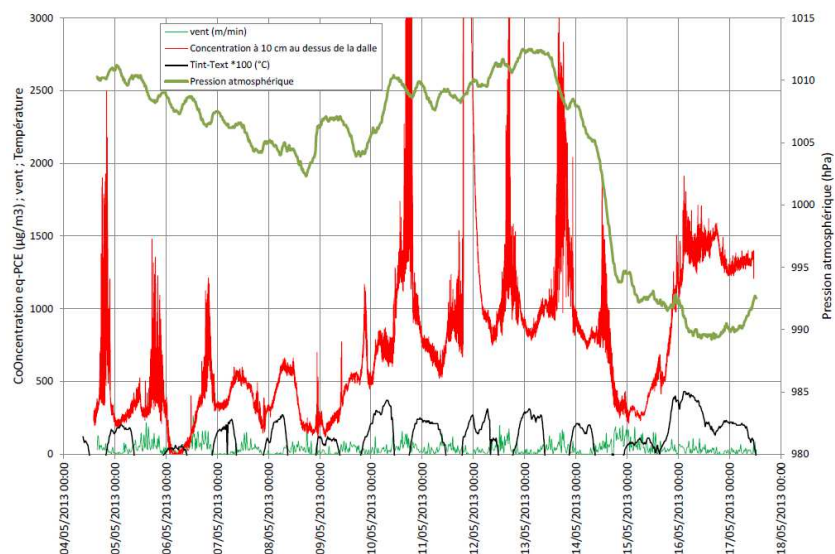


Figure 10 – Influence du vent et de la pression sur les concentrations dans l'air intérieur (site atelier FLUXOBAT, Traverse et al., 2013)

- En l'absence de guide méthodologique faisant foi, mais aussi de la multiplicité des facteurs d'influence, les incertitudes associées à la phase d'échantillonnage des sols pollués sont rarement abordées de manière très quantitative dans la pratique courante de la gestion des sites et sols pollués.
- Les sources d'incertitudes susceptibles d'influer sur l'échantillonnage vont de la qualité des études historiques et documentaires, à la prise en compte de l'hétérogénéité et de la dynamique des milieux (pollution dans le sol, nappes d'eaux souterraines, gaz du sol, ...), aux techniques de prélèvement et de conditionnement, etc.
- Concernant le prélèvement de sols pollués par des composés volatils (COHV, BTEX), des essais d'inter-comparaison de techniques permettent de mettre en évidence les atouts de certaines techniques telles que l'utilisation de flacons pré-remplis de méthanol.
- D'autres bonnes pratiques permettent de réduire les incertitudes associées à la phase d'échantillonnage des sols pollués, telles que l'association de mesures de terrain (qui peuvent être plus nombreuses) et de mesures en laboratoire pour la validation, la prise en compte de la granulométrie des fractions analysées, pour éviter d'appliquer certains résultats d'analyse à des volumes de sols inadaptés, ou encore le bon enchaînement des interventions pour éviter des délais durant lesquels la situation sur le terrain peut évoluer de manière significative.
- Dans les eaux souterraines on observe souvent de fortes variations spatiales des concentrations, tout particulièrement dans les zones proches des zones sources, où les concentrations peuvent varier d'un facteur 100 sur seulement 1 à 2 m d'épaisseur
- Dans le cas des concentrations dans le gaz du sol ou l'air intérieur, les variations spatiales peuvent s'accompagner de fortes variations temporelles, en lien notamment avec des phénomènes hydro-météorologiques.

2.3. Préparation et analyses en laboratoire

2.3.1. Introduction

L'incertitude déterminée par le laboratoire ne concerne en général que le protocole analytique appliqué sur l'échantillon dit « pour analyse » (ou le sous-échantillon le cas échéant), donc elle ne prend pas en compte les phases de préparation physique. Elle prend en compte à l'inverse la totalité du protocole analytique, à savoir l'extraction (le cas échéant) et l'analyse.

Préparation physique

Pour rappel, les étapes de préparation physique qui sont à l'heure actuelle pas ou peu prises en compte sont les suivantes :

- **Le sous-échantillonnage** : une partie de l'échantillon du flacon initial peut être prélevée suivant les quantités nécessaires à l'analyse et les pratiques des laboratoires. Par exemple pour les eaux, le volume traité est en général plus faible que le volume demandé, ce qui permet de réaliser une analyse en double en cas de difficultés. Pour les sols, le laboratoire ne va en général préparer ou analyser qu'une partie de la masse reçue, afin de garder une souche brute, ou d'optimiser le temps de préparation. Les normes d'analyse et les normes de préparation des solides précisent ces conditions afin de réaliser un sous-échantillonnage représentatif. L'échantillon initial peut se retrouver ainsi divisé en plusieurs sous-échantillons (par exemple un sous-échantillon pour l'analyse des composés volatils, un autre pour l'analyse des HAP, un autre pour le COT, etc.).
- **La préparation physique** : elle est fonction des paramètres à analyser et de la matrice. Elle peut être réalisée sur l'ensemble de l'échantillon reçu, ou sur un sous-échantillon. Les sols peuvent par exemple passer par des étapes de séchage, quartage, homogénéisation, broyage et les échantillons d'eau par une étape de filtration. Ainsi, l'échantillon préparé pour l'analyse est physiquement différent de l'échantillon initial (granulométrie, teneur en eau, élimination des MES...).

Analyse

L'incertitude de mesure déterminée par le laboratoire ne concerne donc majoritairement que le protocole analytique appliqué sur l'échantillon préparé, à savoir l'extraction (le cas échéant) et l'analyse. Les laboratoires d'analyse chimique environnementale sont depuis longtemps habitués à estimer leurs incertitudes de mesure. Cependant les exigences liées aux systèmes d'assurance qualité (accréditation) ou encore aux textes réglementaires se sont très fortement renforcées depuis les années 2000. Les incertitudes sont des informations indispensables à associer à un résultat d'analyse pour estimer le niveau de confiance sur le résultat et l'exploiter correctement (au même titre que la méthode utilisée, la limite de quantification, ...). Les incertitudes :

- permettent de comparer des résultats à des valeurs seuils (réglementaires ou pas) notamment en permettant d'évaluer le risque de dépassement de cette valeur. Dans ce cas, des documents normatifs et/ou réglementaires doivent décrire le mode d'utilisation de l'incertitude pour effectuer la comparaison ;
- sont des informations indispensables pour dimensionner une surveillance en tendance : elles impactent en effet la capacité à détecter des futures tendances à la hausse ou à la baisse, en même temps que la fréquence d'acquisition des données, la durée de surveillance, l'intensité des variations à observer (plus l'incertitude sera forte, plus les chances d'observer des tendances seront faibles et plus il sera nécessaire d'acquérir des données ou de prolonger l'observation) ;
- elles permettent de comparer des résultats entre eux afin de savoir si les différences observées sont significatives (par exemple, contrôles effectués sur des résultats provenant de deux laboratoires ayant travaillé sur le même échantillon ou analyses à deux périodes différentes). Sans information sur les incertitudes, de telles comparaisons sont impossibles.

2.3.2. Principales sources d'incertitude

Dans un contexte de laboratoire d'analyse, de très nombreuses sources peuvent contribuer à l'incertitude. On citera les exemples suivants :

Général

- Imprécision dans la définition de « l'objet à analyser » (mesurande).

Etape de préparation :

- mauvaise représentativité de la prise d'essai pour analyse, par rapport à l'échantillon reçu (quartage, broyage insuffisants) ;
- pertes entre l'échantillonnage et l'analyse (dégradation, volatilisation, précipitation, ...) ;
- risques de contamination (par l'opérateur, l'atmosphère, les réactifs, ...). Ces risques sont plus importants pour certaines familles de paramètres (composés plastifiants dont les phtalates, HAP, COV, quelques métaux, ...) et dès lors que les concentrations recherchées sont faibles. Il peut également se produire des contaminations croisées entre échantillons de teneurs très différentes ; on peut citer aussi l'importance capitale de la préparation des légumes (lavage, épluchage), tant le peu de terre résiduelle peut impacter fortement la qualité du résultat.

Etape d'analyse :

- biais liés aux étapes notamment d'extraction (ex : mauvaise estimation des rendements d'extraction, estimation complexe en l'absence de matériau de référence) ;
- les erreurs d'étalonnage (pureté des étalons, ...) ;
- la métrologie des instruments utilisés (température, volume, masse, ...) ;
- les erreurs dans la méthode (dilutions,...).

En général, les sources d'incertitude peuvent être regroupées en 5 classes (méthode « 5M) : Milieu, Méthode, Matière, Main d'œuvre et Moyen.

2.3.3. Gestion des incertitudes

Comme indiqué dans le paragraphe 2.3.1 les évaluations d'incertitude dans les laboratoires concernent principalement le protocole « analytique » et peu les phases de préparation physique.

a) Obligations pour les laboratoires

La quasi-totalité des laboratoires d'analyses environnementales travaille dans un cadre d'accréditation (COFRAC au niveau français). Les laboratoires doivent montrer leur conformité aux exigences de la norme NF EN IEC 17025, qui impose aux laboratoires d'évaluer leurs incertitudes et d'identifier les contributions à ces incertitudes. Elle demande que ces incertitudes soient transmises via le rapport d'essai si :

- elle est importante pour la validité ou l'application des résultats d'essai ;
- les instructions du client l'exigent ;
- ou si l'incertitude de mesure affecte la conformité aux limites d'une spécification.

Outre ce cadre général d'assurance qualité, différents textes réglementaires en contexte environnemental ou sanitaire ont renforcé les exigences pour les laboratoires concernant les incertitudes de mesure. La Directive Cadre Européenne sur l'Eau (2000/60/CE) a pris en compte l'ampleur de l'importance de la qualité des mesures, de leur fiabilité et de leur comparabilité (voir Ghestem, 2011). Elle a ainsi donné lieu à l'élaboration d'une directive spécifique sur l'assurance et le contrôle qualité (2009/90/CE). Les laboratoires opérant dans le cadre de ces directives doivent garantir des incertitudes en-deçà de certaines valeurs par rapport aux valeurs seuils NQE (normes de qualité environnementale pour les eaux de surface) ou NQ (normes de qualité pour les eaux souterraines). Au niveau national, l'arrêté d'agrément des laboratoires du ministère en charge de l'environnement a repris cette exigence et en a fait l'une des conditions d'agrément des laboratoires. Cependant ceci ne concerne que les substances et les matrices (eau, sédiment, biote) couvertes par cet agrément.

b) Approche de l'incertitude

Tout processus de mesure est soumis à des facteurs d'influence que l'on appelle « sources ». Ces facteurs d'influence sont plus ou moins bien maîtrisés et maîtrisables par l'opérateur. Ainsi, le résultat de mesure n'est pas une valeur unique. Des variations entre des mesures répétées se produisent parce que les grandeurs d'influence qui peuvent affecter le résultat de mesure ne sont pas maintenues parfaitement constantes. Il en découle un doute sur le résultat final. L'estimation de l'incertitude a pour objectif de caractériser ce doute.

De façon schématique, on peut décomposer un résultat de mesure en trois composantes (**Figure 11**) :

- la valeur « vraie » du mesurande, qui est le plus souvent inconnue ou très difficilement accessible ;
- une composante liée à des erreurs systématiques qui se reproduisent à l'identique à chaque nouvelle mesure ;
- une composante liée à des erreurs aléatoires qui fluctuent de manière imprévisible à chaque nouvelle mesure.

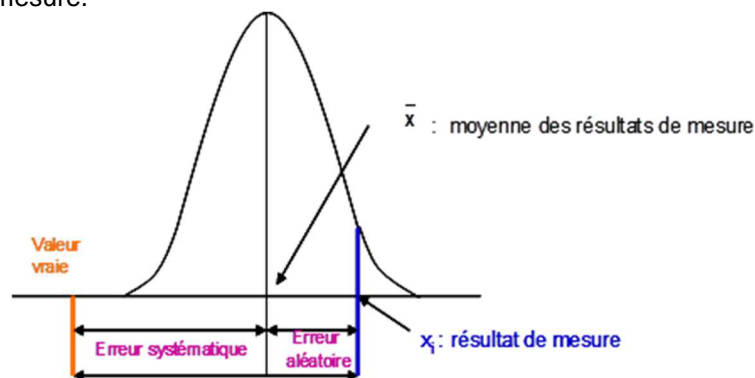


Figure 11 - Représentation schématique de l'incertitude de laboratoire (Ghestem et Lachenal, 2008).

Les résultats se répartissent de façon plus ou moins dispersée (fidélité) autour d'une valeur moyenne qui est elle-même plus ou moins éloignée de la valeur « vraie » (justesse ou biais). Les schémas ci-dessous, prenant comme modèle une cible centrée sur la valeur « vraie », illustrent ces concepts.

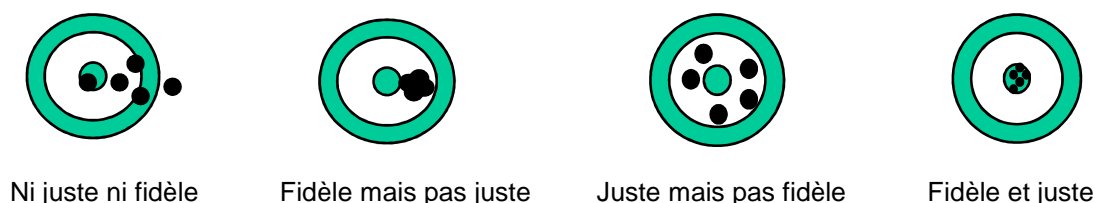


Figure 12 – Différents types d'incertitude associés aux mesures de laboratoire (Ghestem et Lachenal, 2008).

Une estimation correcte de l'incertitude doit prendre en compte les sources d'erreurs aléatoires et systématiques. L'incertitude dépend :

- de la concentration : elle est beaucoup plus élevée aux concentrations proches de la limite de quantification qu'aux concentrations supérieures ;
- de la méthode utilisée : l'incertitude est par exemple souvent plus élevée pour les méthodes manuelles que pour les méthodes automatisées mais elle peut également être très différente entre deux méthodes automatisées mais basées sur des principes différents ;
- du laboratoire : l'incertitude dépend de l'effort que met le laboratoire à améliorer ses méthodes d'analyses (contrôles plus nombreux, réalisation de doubles, utilisation de matériaux de référence certifiés, qualification du personnel,...) ;
- de la matrice : les incertitudes ne seront pas les mêmes, toutes conditions étant identiques par ailleurs (paramètre, concentration, méthode, ...), selon qu'on a affaire à de l'eau potable, résiduaire ou saline dans le cas d'une matrice liquide, ou alors à du sable ou à de l'argile dans le cas d'une matrice solide. En pratique, les laboratoires évaluent des incertitudes pour un groupe d'échantillons de caractéristiques voisines ;
- du mode d'estimation de l'incertitude utilisée. En théorie cela ne devrait pas être le cas, mais ce facteur peut jouer non seulement sur le plan de la maîtrise qu'a le laboratoire du calcul d'incertitude mais également de la méthode d'estimation utilisée (pour cette raison, l'incertitude ne devrait pas être, sauf exception, un critère de sélection des laboratoires ou alors ce critère devrait être utilisé avec une grande prudence).

Toutes ces informations devraient être fournies avec l'incertitude associée afin de pouvoir en faire une interprétation correcte. Par exemple, l'absence d'information sur le niveau de concentration auquel est estimée l'incertitude rend la donnée inutilisable.

Un point très important concerne le mode d'expression de l'incertitude et la notion de facteur d'élargissement. Dans son estimation d'incertitude, le laboratoire aboutit à une « incertitude type », notée u , souvent issue d'une évaluation d'écart-type. L'incertitude « finale » est généralement exprimée sous la forme d'une incertitude élargie U , qui est un multiple k de l'incertitude-type u (en général $k = 2$). Avec $k = 2$, la probabilité que la valeur « vraie » du résultat se trouve dans l'intervalle « résultat $\pm U$ » est de 95% (hypothèse d'une distribution normale). Elle ne serait que de 68 % si l'incertitude était sous la forme « résultat $\pm u$ ». Sans information sur ce facteur d'élargissement, l'incertitude est inexploitable.

Estimer une incertitude de mesure est un travail complexe qui nécessite souvent l'acquisition de nombreuses données. Des documents normatifs donnent les principes et les méthodes permettant d'estimer ces incertitudes. Les plus notables dans le domaine de l'eau sont les normes NF EN ISO 11352 (AFNOR, 2013) et EURACHEM / CITAM (LNE, 2012). Les méthodes sont basées soit sur une prise en compte et une quantification des incertitudes liées à chaque source d'incertitude puis une combinaison des incertitudes élémentaires, soit de façon plus pragmatique sur une estimation globale des incertitudes à partir des estimations de « fidélité » et « biais » du processus. C'est cette dernière approche qui est largement répandue actuellement dans les laboratoires à travers l'application de la norme NF EN 11352.

c) Essais inter-laboratoires

Chaque année en France et en Europe, de nombreux essais inter-laboratoires (EIL) sont organisés par des organismes de comparaison inter-laboratoires (OCIL), en particulier en France par les OCILs AGLAE et BIPEA. Ces essais permettent aux laboratoires de vérifier la fiabilité de leurs résultats de façon régulière par un contrôle externe. Les résultats obtenus par les laboratoires à ces essais sont utilisés comme critères d'obtention d'agrément par les ministères de l'environnement et de la santé.

Outre leur objectif d'évaluation de l'aptitude du laboratoire à fournir des résultats fiables, ces essais inter-laboratoires sont très riches en termes d'informations sur la dispersion des résultats obtenus par les laboratoires. Cette dispersion est parfois qualifiée « d'incertitude profession » et donne une évaluation de la « maîtrise » de l'analyse de ce paramètre pour les gestionnaires.

Dans le domaine des analyses d'eau, AQUAREF en collaboration avec l'ANSES et AGLAE, a fait une synthèse des dispersions observées pour plus de 150 paramètres (Moreau *et al.*, 2019). Pour chaque paramètre, les évaluations des dispersions observées aux essais inter-laboratoires AGLAE sont fournies au niveau de la valeur seuil du milieu concerné. Les milieux sont : eaux superficielles, eaux souterraines et eaux destinées à la consommation humaine. Les dispersions sont fournies avec un facteur multiplicatif de 2 (correspondant de façon très simplifiée à un niveau de confiance à 95 %). Ainsi, pour prendre l'exemple du plomb dans les eaux souterraines (norme = 10 µg/L), 95 % des laboratoires ont obtenu un résultat d'analyse situé dans l'intervalle $10 \pm 1,8$ µg/L.

Ces informations pourraient être distribuées aux gestionnaires afin de les renseigner sur le niveau de fiabilité des résultats. Des synthèses complémentaires pourraient être réalisées sur d'autres paramètres d'intérêt pour la gestion des sites pollués, ou pour d'autres matrices (sols, sédiments, air, plantes potagères) dans la mesure de la disponibilité de données d'essais inter-laboratoires.

- Dans le domaine de l'analyse, les sources d'incertitude sont multiples et liées aux étapes de préparation, avec le cas échéant une mauvaise représentativité de la prise d'essai pour analyse, des pertes entre l'échantillonnage et l'analyse (dégradation, volatilisation, ...) et à l'étape d'analyse elle-même (contamination par l'opérateur ou l'atmosphère, erreurs d'étalonnage, etc.).
- Les incertitudes actuellement estimées par les laboratoires prennent uniquement (ou majoritairement) en compte les étapes « analytiques » appliquées sur un échantillon préparé. Elles ne prennent pas ou peu en compte les étapes de préparation physique des échantillons.
- Les incertitudes de mesure déterminées par le laboratoire sont des informations indispensables à associer à un résultat d'analyse afin de rendre compte de la dispersion possible du résultat.
- Il est nécessaire d'être précis sur le mode d'expression de l'incertitude et notamment sur le facteur d'élargissement utilisé, qui selon les normes en vigueur est pris égal à 2.
- Les incertitudes varient en fonction de la concentration du paramètre, de la méthode d'analyse ou encore de la matrice étudiée.
- Les essais inter-laboratoires sont des outils indispensables pour l'amélioration de la qualité des données car ils permettent d'accéder à des informations assimilables à des « incertitudes de la profession ». Il est essentiel que ces essais soient maintenus et développés quand ils n'existent pas.

2.4. Représentation et interprétation des données

2.4.1. Paramètres précis, aléatoires, imprécis, ...

Paramètre précis :

Un paramètre peut être de nature précise : par exemple la vitesse de la lumière est égale à 299792458 m/s. Quel que soit le lieu ou le référentiel de mesure et sous réserve d'erreurs expérimentales, la mesure de la vitesse de la lumière donnera cette valeur ou une valeur approchée. Le degré de précision nécessaire (notamment le nombre de chiffres significatifs) dépend du domaine d'utilisation de la valeur (par ex. calcul de la distance d'une étoile, applications GPS, ...).

Paramètre aléatoire :

Peu de paramètres dans la nature ont des valeurs aussi précisément connues. Par exemple si on s'intéresse par exemple à la hauteur de pluie à Orléans sur la période 1992-2018, les valeurs mesurées sur 10 jours (valeurs décadaires) sont présentées dans l'histogramme de fréquences de la **Figure 13a**, qui suggère que la variable aléatoire « hauteur de pluie décadaire » suit une distribution de type log-normale. La **Figure 13b** présente les mêmes données sous la forme d'une distribution des fréquences relatives cumulées. En assimilant cette distribution des fréquences cumulées à une distribution de probabilité, on peut dire par exemple que si on choisit au hasard une décade sur la période 1992-2018, on a 50 % de chances que la hauteur de pluie sur cette décade soit inférieure à 13,2 mm. Ou encore 95 % de chances qu'elle soit inférieure à 50,6 mm.

Pour utiliser ce type d'information dans un modèle de risque en contexte de gestion des sites et sols pollués on peut, soit utiliser directement les valeurs mesurées, ou alors utiliser une distribution de probabilité calée sur les données mesurées. La **Figure 14a** montre le résultat d'un calage automatique (à l'aide d'un algorithme d'optimisation) d'une distribution de probabilité log-normale sur les données mesurées. On note que le calage n'est pas très satisfaisant. Un meilleur calage est obtenu à l'aide d'une distribution de probabilité dite Gamma (**Figure 14b**), qui est définie par deux paramètres : un paramètre de forme (dans ce cas : 0,86) et un paramètre d'intensité (dans ce cas : 22,54).

Ce calage d'une loi de probabilité unique à des données mesurées est possible, d'une part, parce que des données sont disponibles en nombre suffisant et, d'autre part, parce que la hauteur de pluie est une variable qui est bien définie et qui se mesure aisément. Dans d'autres situations on peut avoir affaire, soit à des mesures éparses et lacunaires, ou alors à des mesures en nombre suffisant mais peu précises. La **Figure 15a** montre le calage d'une distribution de probabilité normale à des données de concentration en arsenic mesurées sur un terril minier. Le meilleur calage, obtenu à l'aide d'un algorithme de minimisation de fonction objectif (une somme de carrés des différences), est obtenu pour une valeur moyenne = 83,5 mg/Kg et un écart type = 35,7 mg/Kg. Mais les valeurs mesurées montrant une certaine dispersion autour de cette distribution, on pourrait choisir de caler ces données à l'aide d'une famille de probabilité dite « paramétrique » (Baudrit *et al.*, 2008) tel qu'illustré en **Figure 15b**, où les valeurs mesurées sont délimitées par des distributions de probabilité haute et basse obtenues en assignant à la moyenne un intervalle de valeurs possibles [68-100] mg/Kg et à l'écart-type un intervalle [34-36] mg/Kg. Ces limites délimitent une « famille » de distributions de probabilité qui est analogue aux « boîtes de probabilité » (p-box) de Ferson et Hajagos (2004).

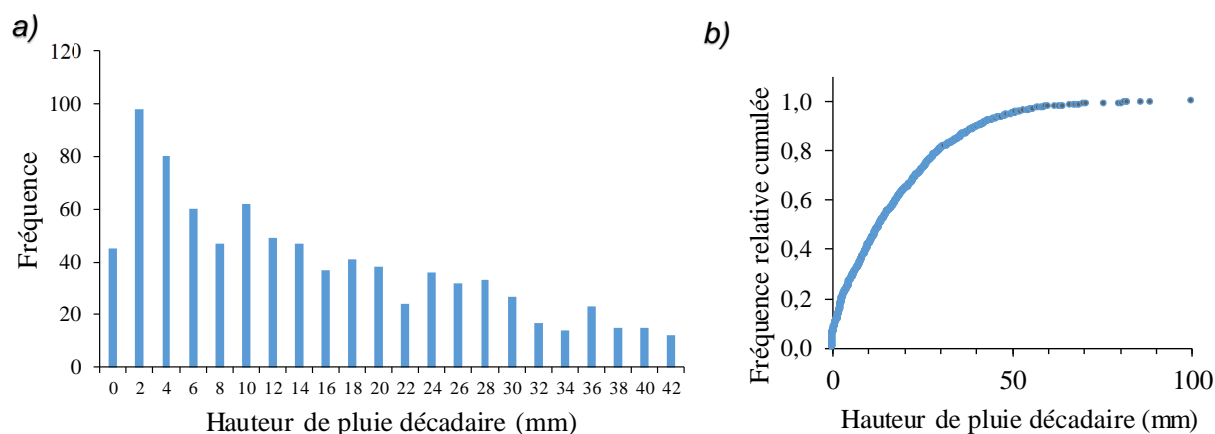


Figure 13 - Histogramme des fréquences obtenu sur des mesures décadaires d'hauteurs de pluie, pour la période 1992-2018 (a) et distribution des fréquences relatives cumulées (b).

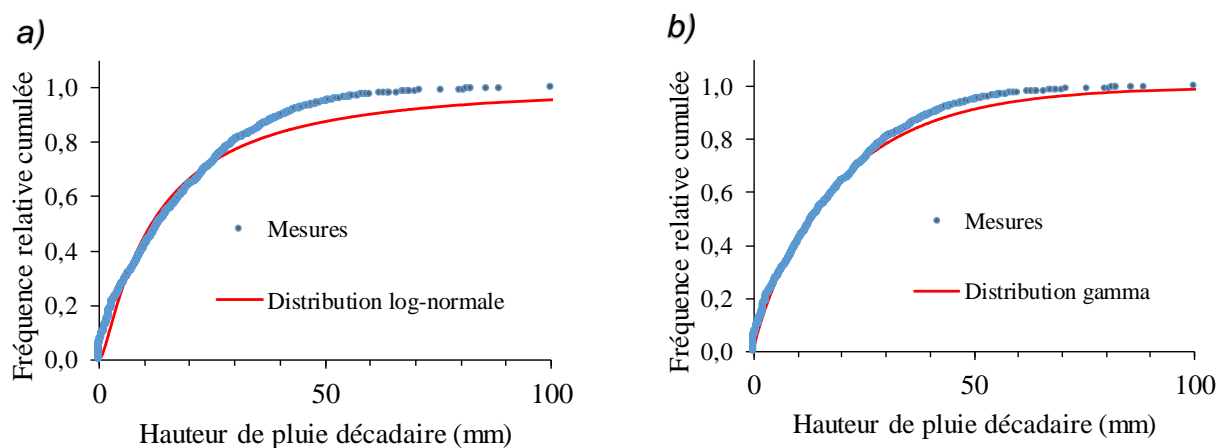


Figure 14 - Calage automatique de la distribution des fréquences relatives à l'aide d'une distribution de probabilité log-normale (a) et calage automatique à l'aide d'une distribution de probabilité Gamma.

Comme il est montré dans le chapitre suivant, ces types de représentations peuvent être exploitées pour divers types d'analyses, telles que l'évaluation des risques environnementaux ou sanitaires. Elles peuvent également servir à déterminer des seuils de coupure à partir de bilans massiques, en appliquant le principe dit « *de Pareto* », qui propose que 80% des effets soient la conséquence de 20% des causes. Appliqué à la gestion des sites pollués, ce principe implique que 80% de la charge polluante soit souvent associée à seulement 20% du volume de sols pollués (Lion et al., 2016).

Paramètres imprécis:

Dans ce qui précède on a abordé successivement les paramètres précis (exemple de la vitesse de la lumière), puis les variables aléatoires représentées par des distributions de probabilité classiques (uniques) et enfin les familles de distributions de probabilité. Mais dans certains cas on n'a pas affaire à des données en nombre significatif, qu'elles soient précises ou non, mais plutôt à de l'information très partielle et/ou imprécise, comme c'est le cas lorsqu'on fait appel à du jugement d'expert. Dans ce cas, on peut se poser la question de savoir si la représentation de l'information par une distribution de probabilité unique est justifiée d'un point de vue scientifique et méthodologique.

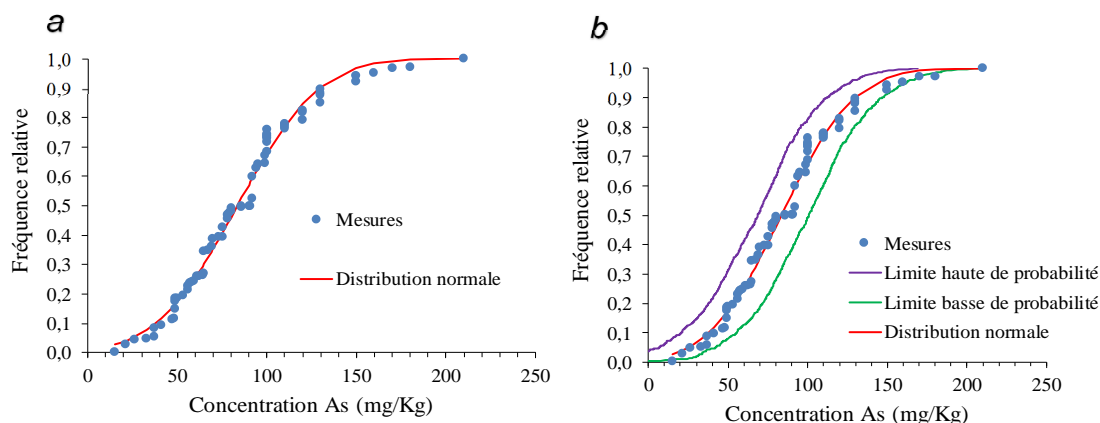


Figure 15 - Calage de mesures de la concentration en arsenic dans un terril minier à l'aide d'une distribution de probabilité normale (a) et délimitation de ces mesures à l'aide d'une famille paramétrique de distributions normales (b).

Dans le chapitre précédent il a été rappelé comment de l'information mal connue pouvait être représentée par un simple intervalle [min-max], avec des bornes définies le cas échéant par jugement d'expert (Figure 2). Or bien souvent un expert peut fournir les bornes d'un intervalle, mais également exprimer des préférences au sein de cet intervalle. En précisant par exemple que les valeurs proches de la borne inférieure (ou supérieure) lui paraissent plus plausibles que les autres, etc. La théorie des possibilités (Zadeh, 1978 ; Dudois et Prade, 1988) permet de formaliser ce type d'information dans un cadre mathématique qui rejoint la notion de probabilités imprécises illustrée dans la Figure 15b.

Prenons l'exemple d'un paramètre assez mal défini : le Facteur d'Atténuation (FA) dans un contexte de risque d'intrusion de vapeur dans les bâtiments. Ce facteur FA est le rapport entre la concentration en un polluant dans l'air intérieur d'un bâtiment et la concentration en ce même polluant dans le gaz du sol, au voisinage immédiat du bâtiment. Pour une certaine concentration en polluant dans le gaz du sol, plus la valeur de FA est élevée, plus la concentration en polluant dans l'air intérieur est élevée. Bien qu'il s'agisse d'un paramètre généralement mal connu, l'USEPA a publié des résultats de retours d'expérience sur de nombreux sites aux Etats-Unis (qui sont repris dans le projet FLUXOBAT ; Traverse et al., 2013). En France, les travaux notamment de Derycke et al. (2018) sur les données acquises dans le cadre de la démarche « établissements sensibles », montrent que les valeurs de FA dépendent, outre la nature du polluant, également de l'âge des bâtiments : le FA est plus élevé dans le cas de bâtiments anciens (les dalles sont plus poreuses et donc les vapeurs les pénètrent plus aisément). Pour une situation donnée et compte tenu des incertitudes « épistémiques » relatives à ce paramètre, un expert pourrait conseiller :

- un intervalle de valeurs qu'il estime contenir la valeur de FA avec certitude ;
- un intervalle de valeurs (voire une valeur unique) qu'il estime plus plausible(s) que les autres.

Cette information peut être représentée par une distribution dite de possibilités (**Figure 16**), qui associe des intervalles (emboîtés) de valeurs du paramètre FA à une mesure de possibilité (notée généralement μ). Pour un intervalle donné, appelé « alpha-coupe », $1 - \mu$ est une limite inférieure de la probabilité que la valeur soit située dans l'intervalle. Ainsi, pour le support de la distribution (l'intervalle $[2 \times 10^{-4} ; 6 \times 10^{-4}]$), on a : $1 - 0 = 1$; c'est-à-dire qu'on est « sûrs » que la valeur est comprise dans le support (la probabilité est comprise entre 1 et 1). Pour le noyau par contre : $1 - 1 = 0$, on n'est pas du tout « sûrs » que la valeur soit égale au noyau (la probabilité est comprise entre 0 et 1). La **Figure 17a** et **b** illustre comment l'information contenue dans la **Figure 16** se traduit par une famille de distributions de probabilités cumulées : toutes celles comprise entre les limites haute (appelée Possibilité ; Π , ou Plausibilité, PI dans la théorie des fonctions de croyance de Shafer, 1976) et basse (appelée Nécessité ; N ou Crédibilité ; Cr ; Shafer).

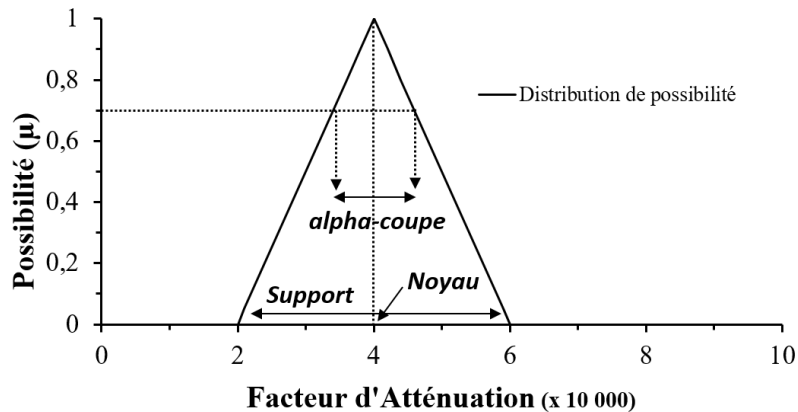


Figure 16 - Distribution de possibilité triangulaire correspondant à un jugement d'expert sur la valeur du facteur d'atténuation FA.

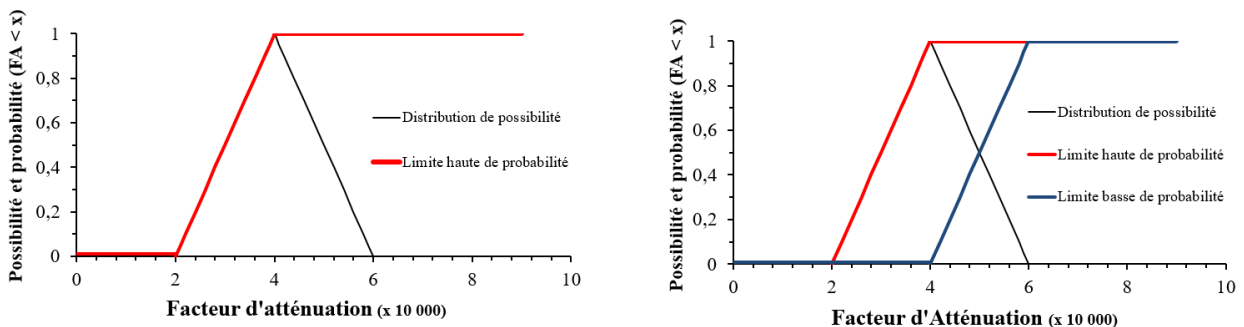


Figure 17 - Correspondance entre distribution de possibilité et famille de distributions de probabilités cumulées.

Ce type d'intervalle emboîté (on parle aussi d'intervalle nuancé) se prête particulièrement à la représentation de jugement d'expert, car un expert est généralement cohérent avec lui-même : l'intervalle des valeurs qu'il considère comme étant les plus vraisemblables sera forcément inclus dans l'intervalle des valeurs qu'il considère comme étant possibles.

Un « inconvénient » de cette approche dans un contexte d'aide à la décision réside dans le fait que le caractère incomplet de l'information est conservé dans l'analyse. Si certains chercheurs recommandent d'utiliser dans ce cas des distributions de probabilité uniques, dites « subjectives » ou « a priori », en invoquant un cadre Bayésien (Lindley, 1971), de nombreux chercheurs (voir par ex. Lloyd et Ries, 2007) contestent cette approche lorsqu'elle est réalisée dans un contexte où de nouvelles données ne permettent pas de « corriger » les distributions a priori, en appliquant le théorème de Bayes des probabilités conditionnelles. Dans ce cas la méthode Bayésienne introduit une confusion entre variabilité aléatoire et méconnaissance.

2.4.2. Incertitudes et données spatialisées

Les informations spatialisées entachées d'incertitude sont de différents types :

- les résultats de mesures de concentrations ou de teneurs (erreurs de mesures au sens large, mesures indirectes comme les sondes radioactives pour les métaux ou par ionisation pour les composés organiques, mesures d'une fraction de carotte seulement, etc.) ;
- les cartes tracées à partir de ces mesures, fournissant une estimation (ou plus simplement une extrapolation) à partir de ces données ;
- les informations sur le milieu, comme l'épaisseur d'un niveau de remblais (inconnue ou supérieure à la valeur atteinte par la reconnaissance), la profondeur atteinte par la pollution, la cote topographique de tout ou partie d'un site, les formations présentes dans les sols et leur localisation ou leurs limites, etc.

On examine ici le premier cas, les suivants étant traités au paragraphe 2.4.7. La représentation des données « de teneur » peut être statistique ou cartographique. Même lorsque les données de teneurs peuvent être supposées « exactes », donc en l'absence d'erreurs de mesures, leur résumé statistique peut ne pas être représentatif des valeurs sur l'ensemble de la zone d'étude. En effet :

- la zone d'étude peut ne pas être homogène pour la grandeur étudiée (exemple : la pollution en une substance est localisée dans une partie du site seulement, ou dans un seul type de remblais ou de formation pédologique) ;
- la reconnaissance peut être préférentielle (guidée par l'histoire de site), privilégiant les valeurs fortes ou faibles ;
- à cause de la présence de corrélation spatiale, un sur-échantillonnage local modifie la distribution (l'histogramme) des données (puisque des valeurs proches « se ressemblent » plus) et en particulier leur moyenne et variance. Ceci peut être corrigé par une pondération des données en fonction de la densité locale de la reconnaissance ;
- dans le cas de données à support variable (longueur de carotte ou longueur analysée variables par exemple), la variabilité dépendant du support, la population statistique des teneurs n'est plus homogène. Les caractéristiques liées à la dispersion (variance et coefficient de variation, quantiles) ne reflètent alors pas celle des teneurs en place, pour un support donné.

En présence d'une limite de quantification, il est utile que la limite (ou les limites) de quantification soient précisées. Le taux de données non quantifiées est une information importante. L'influence de ce taux et des valeurs limites sur le résumé statistique est facilement mise en évidence en ramenant les valeurs non quantifiées successivement à 0 puis à la limite (ou aux limites) de quantification. Les valeurs non quantifiées peuvent aussi être rendues visibles par un symbole ou une couleur spécifiques sur les cartes de données ou en « coloriant » l'ensemble des données associées sur l'histogramme de la variable.

Les doublons (répliquats) sont utiles pour visualiser l'incertitude de mesure (au sens large) et la quantifier (avec des hypothèses qui ne sont pas toutes vérifiables ; de Fouquet, 2016). La dispersion des mesures apparaît sur le report cartographique des données, par simple superposition des symboles (cercles, carrés ou triangles, ...). Lorsque les mesures peuvent être réparties par lots (1^{er} et 2^{ème} sous-échantillon, ou deux appareils de mesure par exemple), les statistiques simples et bivariées de ces lots (avec les nuages de corrélation, la superposition des histogrammes, les statistiques des différences) permettent de détecter la présence d'un écart-systématique éventuel (biais) et fournissent l'amplitude de la variance d'erreur de mesure. Les variogrammes² expérimentaux simples et croisés des données par lots permettent de contrôler certaines hypothèses du calcul statistique et d'en préciser les résultats (Chilès et Delfiner, 1999).

Pour les traitements ultérieurs (étude géostatistique pour le tracé de cartes), il est préférable de conserver les différentes valeurs mesurées (répliquats) plutôt que d'en prendre la moyenne, afin de conserver l'information sur l'incertitude de mesure (au sens large).

² Le variogramme représente, au facteur $\frac{1}{2}$ près, la moyenne de l'écart-carré entre deux valeurs en fonction de leur distance, en tenant compte de l'orientation. Le variogramme croisé est, toujours au facteur $\frac{1}{2}$ près, la moyenne du produit des différences des deux variables (la covariance de leurs accroissements) en fonction de la distance, en tenant compte de l'orientation.

Dans le cas de mesures directes des teneurs, plus fiables mais plus coûteuses et de mesures indirectes plus nombreuses car moins coûteuses mais aussi moins fiables, l'étude exploratoire statistique (nuages de corrélation, régression empirique) et variographique (variogrammes expérimentaux simples et croisés) permet de préciser la liaison entre ces variables, sous réserve que les points de mesure communs soient « suffisamment nombreux » (pour les calcul des variogrammes notamment) et échantillonnent l'ensemble des deux populations statistiques. Pour les cartographies ultérieures, la « transformation » des covariables en teneurs (à l'aide de régressions linéaires, par exemple) est à éviter, car les deux variables n'ont généralement pas la même variabilité.

Les représentations statistiques ou cartographiques des données, complétées par l'étude exploratoire et variographique, permettent de représenter les incertitudes sur les valeurs, voire de les quantifier si les répliquats ou les différentes mesures aux mêmes points expérimentaux sont en nombre suffisant. Cette redondance, qui peut sembler coûteuse, permet seule de quantifier les incertitudes sur les données (valeurs mesurées).

2.4.3. Critères d'aide au choix du mode de représentation de l'information

Comme illustré plus haut, différents modes de représentation de l'information et de l'incertitude associée se prêtent à différents types d'information (variables aléatoires ou paramètres imprécis, données spatialisées ou non, etc.) et il n'existe pas de solution « fourre-tout ». L'organigramme de la **Figure 18** propose de mettre ces deux dimensions (type d'information incertaine et mode de représentation) en relation, en essayant de rechercher dans la mesure du possible une cohérence entre les deux. On cherche notamment à éviter de « créer » de l'information qui en fait n'existe pas, comme c'est le cas lorsqu'on suppose a priori une distribution de probabilité unique pour représenter de l'information incomplète/imprécise. Cet organigramme n'est qu'une ébauche et mériterait d'être retravaillé pour en améliorer le caractère opérationnel.

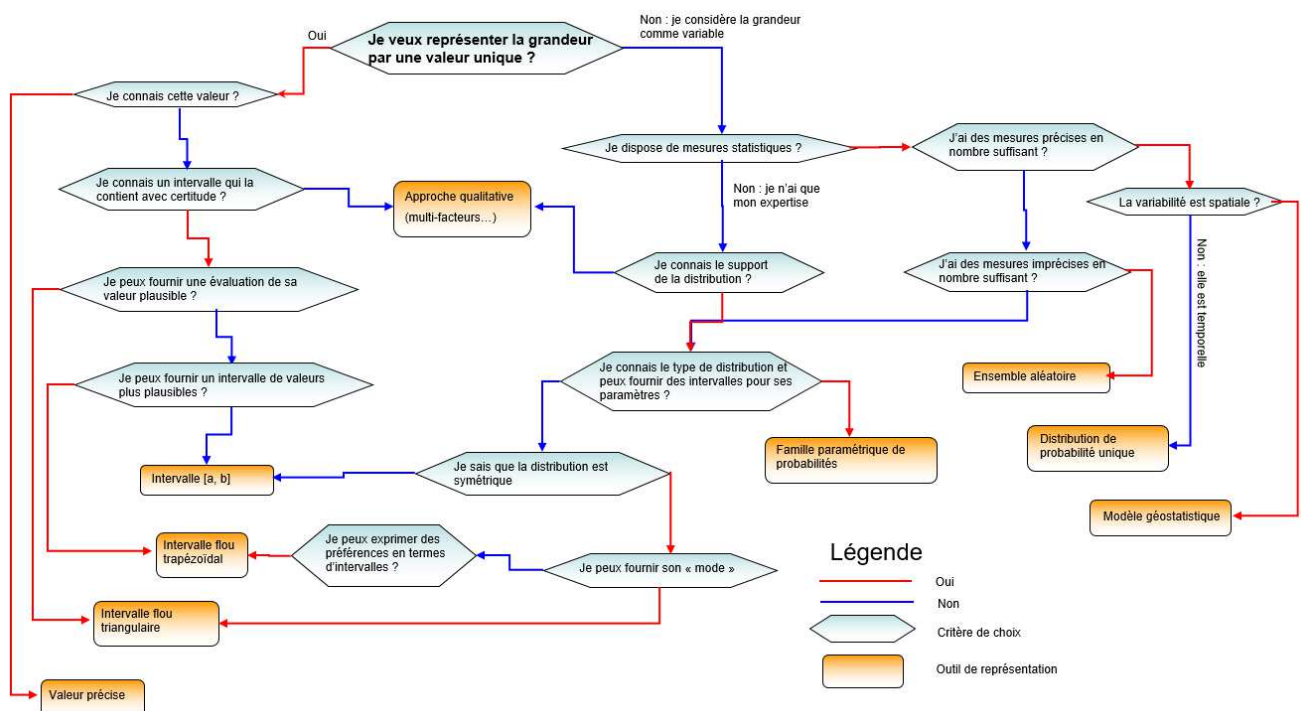


Figure 18 - Organigramme mettant en relation différents types d'information et des modes de représentation de cette information (modifié de Dubois et Guyonnet, 2011).

- Différents types d'information (précises, aléatoires, imprécises, spatialisée ou non, ...) se prêtent à différents modes de représentation.
- La théorie classique des probabilités traite des variables ou des fonctions aléatoires, tandis que l'approche bayésienne permet de tenir compte de l'incertitude de modélisation. D'autres théories de l'information permettent de représenter des informations imprécises (théorie des possibilités, des fonctions de croyance, des probabilités imprécises, ...).
- Les intervalles dits « nuancés » de la théorie des possibilités, se prêtent particulièrement à la représentation de dire d'expert lorsque l'expert exprime un intervalle de valeurs (support) au sein duquel il est certain de trouver la valeur d'un paramètre, ainsi qu'une valeur (ou intervalle plus restreint que le support) jugée la plus vraisemblable. Ces intervalles se traduisent par des « familles » de distributions de probabilité, l'information disponible n'étant pas suffisante pour permettre d'en sélectionner une seule.
- L'information spatialisée est appréhendée à l'aide des outils de la géostatistique. Même lorsque les données (teneurs du sol par exemple) sont supposées « exactes », des incertitudes subsistent en raison de la reconnaissance qui n'est pas exhaustive ou alors préférentielle, de l'hétérogénéité de la zone d'étude, de sur-échantillonnage en présence de corrélation spatiale qui modifie la distribution des données, etc.

2.4.4. Propagation des incertitudes dans des modèles prédictifs

Les modèles prédictifs sont utilisés en gestion des sites et sols pollués pour obtenir notamment des estimations de l'évolution future de l'état des milieux (sols, air, eaux, ...) ou des impacts sanitaires, etc. Différents types de modèles peuvent être sélectionnés en fonction des objectifs et du contexte (voir section suivante). Pour illustrer la notion de propagation des incertitudes liées aux paramètres d'entrée des modèles, on se place ici dans le contexte générique où on dispose d'un modèle de calcul de risque (environnemental ou sanitaire) qui est fonction d'un certain nombre de paramètres, notés X , Y , ..., Z . Si l'incertitude relative à ces paramètres peut être représentée par des distributions de probabilité uniques (car on dispose de données permettant de justifier ces distributions), alors cette incertitude peut être propagée aux résultats fournis par le modèle en utilisant par exemple la méthode couramment utilisée dite « Monte Carlo » (voir par ex. Gobet, 2013).

Cette méthode est illustrée de manière schématique à la **Figure 19**. Elle consiste à générer un nombre aléatoire (χ) pour chaque paramètre (cas où ces paramètres sont supposés statistiquement indépendants) et à « échantillonner » les distributions à ce niveau de probabilité : c'est-à-dire trouver x_i tel que $P(X < x_i) = \chi$. Le calcul de risque est effectué avec les valeurs x_i , y_i , ... ainsi échantillonnées. En répétant cette procédure un grand nombre de fois, on obtient un échantillonnage représentatif des distributions des valeurs de risque calculées et on peut alors dresser un diagramme des fréquences relatives cumulées. Dans la pratique, on classe toutes les valeurs de risque calculé par ordre croissant et on assigne à chaque valeur une fréquence de $1/n$, où n est le nombre d'itérations. Le diagramme de fréquence cumulée est assimilé à une distribution de probabilité cumulée et les différents quantiles de fréquences cumulées peuvent ensuite être comparés au seuil de risque, pour déterminer si le risque calculé est acceptable ou non.

Il faut souligner qu'en raison de l'effort de calcul nécessaire, la méthode dite Monte Carlo s'applique surtout aux modèles dits « analytiques » (solutions mathématiques des équations décrivant les phénomènes influençant le risque) plutôt qu'aux modèles dits « numériques » (méthodes aux différences finies ou aux éléments finis, etc. ; voir section suivante).

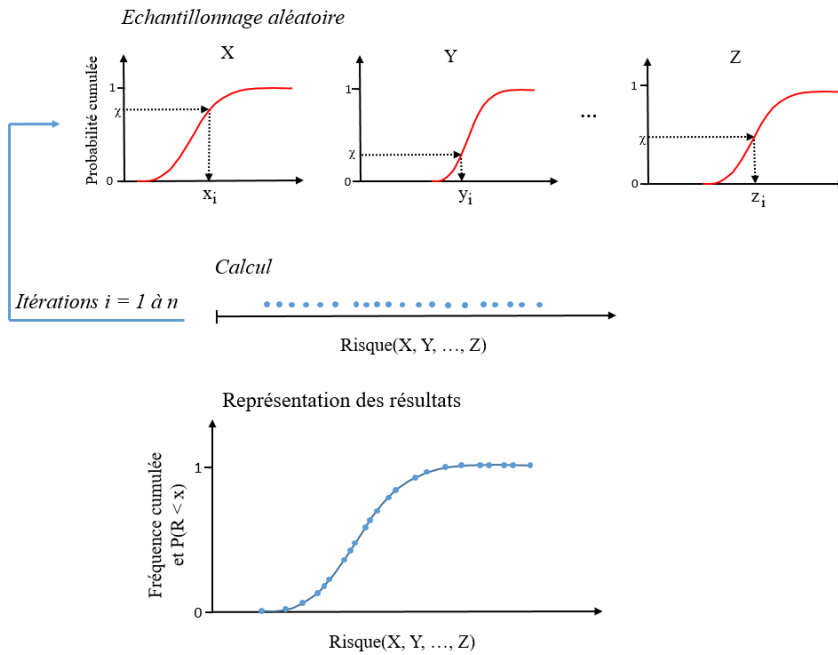


Figure 19 - Illustration schématique de la propagation d'incertitude dans un modèle de risque à l'aide de la méthode dite Monte Carlo.

Si un ou plusieurs paramètres ne sont pas représentés par des distributions de probabilité uniques (par manque d'information), mais plutôt par un des autres outils de représentation illustrés dans la section précédente (familles de probabilités, intervalles nuancés ou non, ...), alors on peut procéder de manière similaire mais au lieu d'obtenir une distribution unique pour le résultat du calcul de risque, on obtient une famille de distributions. C'est ce qui est illustré à la **Figure 20** pour le cas simple d'un modèle de risque dont tous les paramètres sauf un (Z), sont représentés par des distributions de probabilité uniques, tandis que pour Z on sait juste que sa valeur se situe quelque part entre une valeur min et une valeur max (intervalle classique). Dans ce cas, la procédure consiste (Baudrit *et al.*, 2006), comme pour le cas Monte Carlo classique, à échantillonner de manière aléatoire les distributions de probabilité, mais cette fois il faut tenir compte de toutes les valeurs possibles de Z sur l'intervalle $[Z_{\min}; Z_{\max}]$. D'un point de vue opérationnel on peut distinguer deux cas de figure : (i) soit il s'agit d'un modèle simple (monotone) où les valeurs min et max du calcul de risque sont obtenues à partir des valeurs min et max sur Z, ou alors (ii) le modèle est plus complexe et il faut alors faire appel à un algorithme d'optimisation pour trouver les valeurs min et max du calcul de risque en balayant sur toutes les valeurs de l'intervalle sur Z.

Dans les deux cas, le résultat du calcul est ce qu'on appelle un « intervalle aléatoire ». On obtient la famille de distributions de probabilité correspondante en classant toutes les valeurs min par ordre croissant, idem pour toutes les valeurs max et en assignant une fréquence relative égale à $1/n$ (n est le nombre d'itérations) à chaque paire $[\min, \max]$. On définit ainsi les bornes de probabilité inférieure (la crédibilité ; Cr, Shafer, 1976) et supérieure (la plausibilité ; PI, Shafer) de la proposition « le risque calculé est inférieur à une certaine valeur » ; $P(R < x)$ de la **Figure 20**.

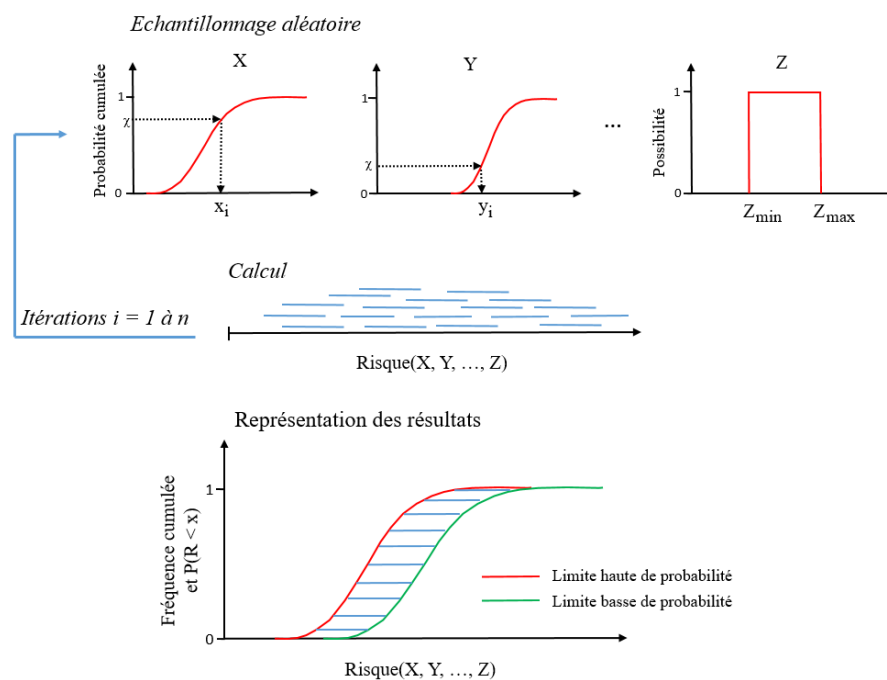


Figure 20 - Illustration schématique de la propagation conjointe d'information aléatoire et incomplète/imprécise.

2.4.5. Incertitudes liées au modèles prédictifs

Dans la section précédente il a été question d'incertitude relative aux paramètres d'un modèle prédictif. Mais se pose également la question de l'incertitude relative au modèle lui-même : modèle conceptuel, manière de représenter les phénomènes, schémas de résolution, etc. Dans ce cas, une approche peut consister à utiliser plusieurs modèles jugés « éligibles » et à synthétiser les résultats comme illustré précédemment, en attribuant le cas échéant des poids différents aux résultats des différents modèles, selon le degré de confiance que l'on souhaite leur accorder.

a) Le schéma conceptuel

Comme le souligne la méthodologie nationale de gestion des sites et sols pollués (DGPR, 2017), les deux principaux outils que sont l'inventaire de l'état des milieux et le plan de gestion, doivent s'appuyer sur une phase préliminaire d'étude et de réflexion qui servent à réaliser un bilan factuel du site étudié sous la forme d'un état des lieux qui est formalisé par un schéma conceptuel. Ce schéma vise à présenter, de manière très synthétique, l'état de pollution des milieux et les modes de contamination potentiels au regard des activités et des usages qui existent sur le site étudié et dans son environnement. Un exemple de schéma conceptuel, pour le cas d'une pollution par un composé organochloré (plus lourd que l'eau ; DNAPL) est présenté à la **Figure 21**.

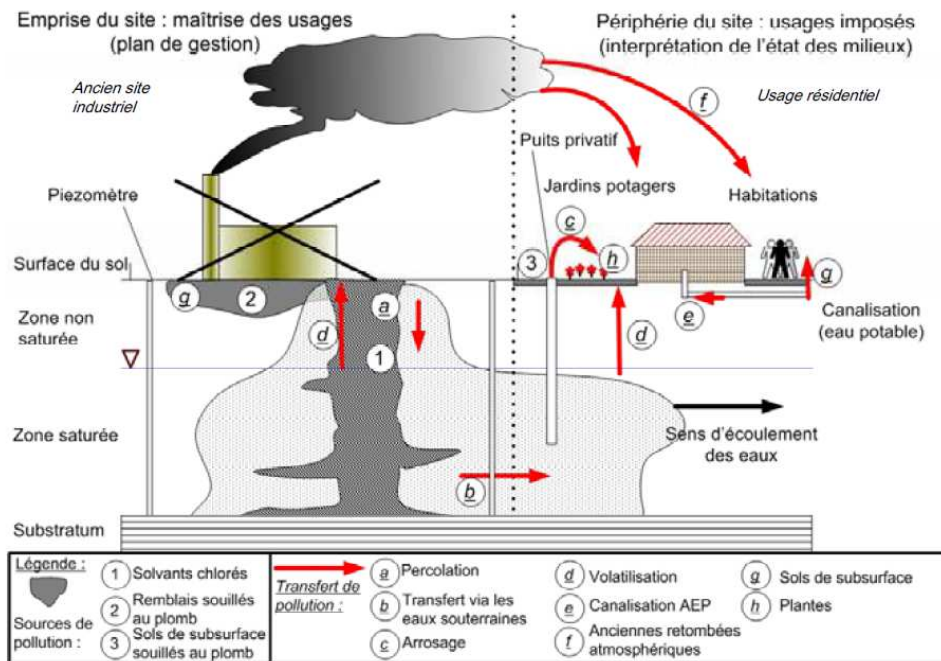


Figure 21 - Exemple de schéma conceptuel illustrant les sources, voies de transfert, etc.
 Source : <http://ssp-infoterre.brqm.fr/schema-conceptuel/>

Naturellement, des incertitudes existent également au stade du schéma conceptuel. Par exemple et en se référant à la **Figure 21**, la direction d'écoulement pourrait être mal connue (en raison par exemple d'une caractérisation insuffisante des hauteurs piézométriques). Ou les volumes de DNAPL déversés pourraient être sous-estimés et par conséquent la durée du « terme source » ou les volumes à dépolluer seraient alors supérieurs aux prévisions. Il existe également des incertitudes relatives aux mécanismes physico-chimiques qui influencent le devenir des polluants sous différentes formes (phase pure, dissoute, gazeuse, adsorbée, ...).

Mais l'élaboration du schéma conceptuel est l'occasion de faire le point sur ce que l'on sait d'un site, ce qui peut mettre le cas échéant en exergue des « zones d'ombre » et susciter des investigations complémentaires, ce qui aura pour effet de réduire les incertitudes et d'augmenter la fiabilité des choix de gestion proposés.

b) Modèles analytiques

Les modèles analytiques sont des solutions mathématiques aux équations décrivant les phénomènes qui influencent les impacts et les risques. Par exemple, dans le cas du transfert de polluants dans le sous-sol, il s'agit de solutions des équations aux dérivées partielles qui décrivent le devenir des polluants sous l'effet de mécanismes tels que le transport par l'eau mobile (advection), la diffusion-dispersion, l'adsorption sur la phase solide, la volatilisation, etc. Un des avantages significatifs des modèles analytiques est d'éviter la dispersion numérique associée aux modèles maillés (section suivante), qui constitue également une source d'incertitude. Un autre avantage est la rapidité de calcul comparé aux modèles maillés, notamment pour les problèmes transitoires (pas besoin d'itérer sur un pas de temps). En raison de cette rapidité de calcul, les modèles analytiques sont particulièrement bien adaptés à des schémas de traitement des incertitudes comme la méthode Monte Carlo (voir section 2.4.4). Mais ces avantages s'accompagnent d'inconvénients et notamment la contrainte de géométries et de conditions aux limites du système modélisé très simplifiées. Malgré leurs limites, les modèles analytiques sont très utiles en première approche et notamment pour « vérifier » les résultats des modèles numériques.

Les exemples d'applications de modèles analytiques au domaine des sites et sols pollués sont très nombreuses. Dans le domaine des transferts, on peut citer par exemple les solutions analytiques

appliquées à la biodégradation des polluants au cœur et en limite des panaches (Hunkeler et al., 2010, Gutierrez et al., 2009), des solutions pour le transport de solutés dans les eaux souterraines en 3D et pour différentes géométries (par ex., Guyonnet, 2008 ; **Figure 22**, Guyonnet et Neville, 2004) ou encore des approches semi-analytiques développées par exemple pour le cas de polluants influencés par les phénomènes d'oxydo-réduction et notamment les BTEX (voir Atteia et Höhener, 2012).

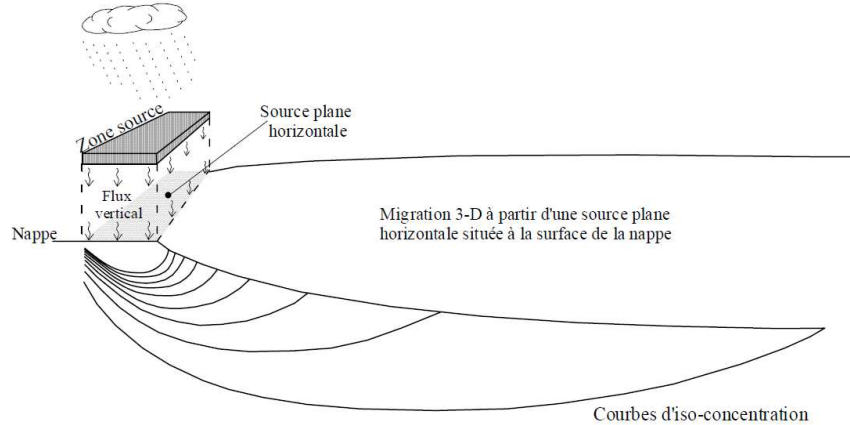


Figure 22 - Schéma conceptuel du modèle analytique MISP (Guyonnet, 2008)

c) Modèles numériques

Les modèles numériques utilisés le plus couramment, au niveau international, dans le domaine des sites et sols pollués pour simuler la migration de polluants dans les eaux souterraines sont probablement les modèles Visual MODFLOW (de Waterloo Hydrogeologic Inc. ; **Figure 23**) et FEFLOW (de DHI Group). Mais un nombre important de modèles numériques existent pour traiter tel ou tel problème particulier du domaine. Le modèle MARTHE développé par le BRGM (Thiéry, 2014), s'il est moins développé que les précédents en termes de fonctionnalités graphiques, possède néanmoins des atouts en termes de contrôle de la dispersion numérique par exemple.

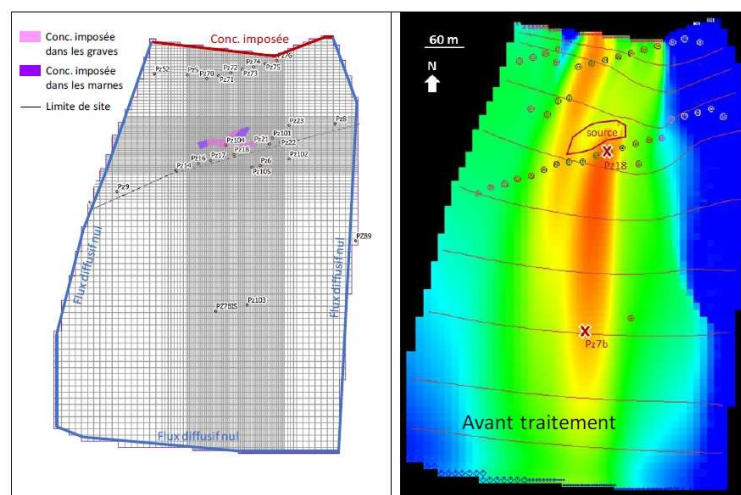


Figure 23 - Exemple de modélisation numérique du transfert d'un polluant (le dichloroéthylène) dans les eaux souterraines à l'aide du modèle numérique Visual MODFLOW (Chastanet et al., 2019)

Par rapport à la prise en compte des incertitudes, l'inconvénient des modèles maillés est le temps de calcul associé à une simulation. En effet, s'il faut multiplier le nombre de simulations pour balayer sur des fourchettes de valeurs possibles des paramètres influençant le modèle, ce temps de calcul peut devenir rédhibitoire. Dans le cadre d'une étude soutenue par l'ADEME, Chastanet et al. (2019) ont proposé de pallier cette difficulté en faisant appel à un « métamodèle » qui reproduit le comportement du modèle numérique complet mais pour un temps de calcul moindre et de traiter les incertitudes à l'aide du métamodèle ainsi calibré.

2.4.6. Comparer des résultats de calcul de modèles prédictifs avec un seuil précis

La comparaison est illustrée dans la **Figure 24** pour les deux cas considérés dans la section 2.4.4 : représentation probabiliste classique (**Figure 24a**) et hybride (**Figure 24b**). Dans le cas a), on compare typiquement le seuil à un quantile 90 % ou 95 % et le risque est jugé acceptable si le seuil est supérieur aux valeurs de risque correspondantes : la probabilité que le risque soit inférieur au seuil est d'au moins 90 %. Dans le cas b), la situation est moins simple car la comparaison à un seuil fournit deux valeurs : une limite haute et une limite basse. Dans le cas de la **Figure 24b**, la probabilité que le risque calculé soit inférieur au seuil est comprise entre 0,95 (Plausibilité ; PI) et 0,6 (Crédibilité ; Cr). Dans la section relative à la communication avec les parties prenantes (section 2.5.), on examinera comment aborder ce type de résultat.

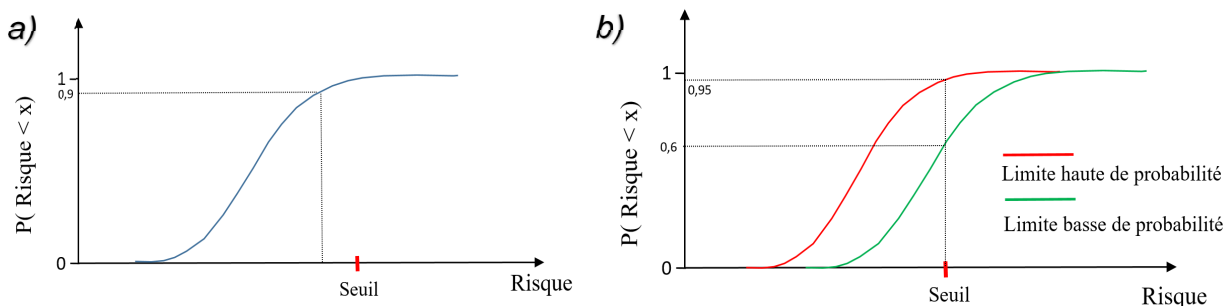


Figure 24 - Comparaison d'un résultat de calcul de risque à un seuil de risque précis.

2.4.7. Comparaison avec un seuil non-précis

Compte tenu des incertitudes associées à la détermination de seuils de risque, il peut paraître surprenant que ces seuils soient représentés dans l'extrême majorité des cas par des valeurs précises. Or comme l'ont montré plusieurs auteurs, cela n'est pas obligatoire. Par exemple, dans une analyse des risques sanitaires induits par la consommation de nitrates, Lee *et al.* (1994) ont utilisé un seuil « flou » pour définir un Quotient de Danger Acceptable (QDA), tel qu'illustré dans la **Figure 25**. Pour une valeur de QDA < 1, une apparition de méthémoglobinémie chez des personnes exposées à la consommation de nitrates n'est pas attendue, mais elle l'est si QDA > 1,5. Entre ces deux valeurs, cette apparition est jugée possible, avec un degré de possibilité qui varie entre ces deux pôles.

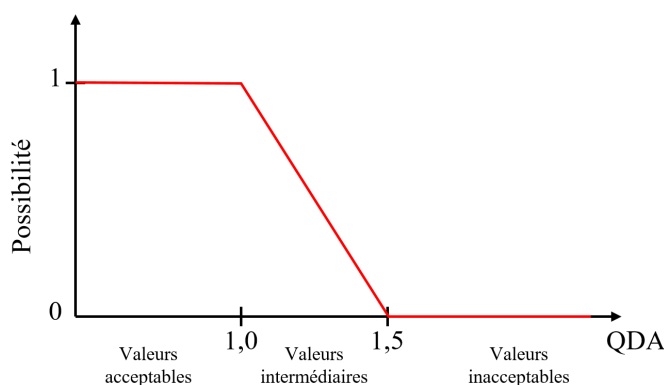


Figure 25 - Seuil d'acceptabilité flou (QDA = quotient de danger acceptable).

Ce type de représentation (intégrant une transition entre ce qui paraît totalement acceptable et ce qui ne l'est pas du tout) pourrait être appliqué à des seuils dans nombre d'autres situations, où l'utilisation de « seuils couperets » engendre parfois des difficultés de gestion. Dans le cas de la **Figure 25**, il appartient au décideur de spécifier quel degré de possibilité exiger pour juger de l'acceptabilité du Quotient de Danger calculé (voir Côme *et al.*, 1997). Une réflexion analogue pourrait s'appliquer à l'élaboration de valeurs toxicologiques utilisées en analyse des risques. Par exemple pour les valeurs de PNEC (concentrations prédites sans effet) ou de VTR (valeurs toxicologiques de référence) qui, malgré l'utilisation dans le cadre de leur élaboration de facteurs dits « d'incertitude », sont exprimés généralement par des valeurs précises.

On peut également envisager ce mode de représentation pour les fonds pédo-géochimiques urbains (FPGU), dont l'identification peut avoir des conséquences importantes en matière de gestion des sites pollués (par exemple pour la destination des terres excavées) en permettant d'identifier des teneurs en éléments dans les sols qui relèvent d'anomalies de celles qui relèvent d'un fond anthropisé ou non. Compte tenu des incertitudes relatives à la définition de ces fonds (Sauvagat, 2019), on peut s'interroger sur la pertinence de valeurs « précises » pour ces fonds et une représentation comme celle de la **Figure 26** pourrait être plus adaptée. Sachant qu'en zone urbaine, tous les sols sont plus ou moins influencés par l'activité anthropique (nous sommes entrés dans l'ère de l'Anthropocène...), on peut poser la question de la définition scientifique de la notion même de fonds pédo-géochimique urbain (voir section 2.4.8). En effet, il est peut-être plus adapté de parler d'une « base de définition d'anomalies de teneurs du sol », à partir de laquelle des mesures de gestion pourront être définies.

Rappelons enfin que, dans un domaine géographique donné, la proportion des concentrations dépassant un seuil fixé dépend du support, c'est-à-dire de la géométrie (donc du volume) des unités de sélection considérées. À tout seuil de concentration devrait donc être associée l'indication du support idoine (litre, m³ ou 10m³ par exemple).

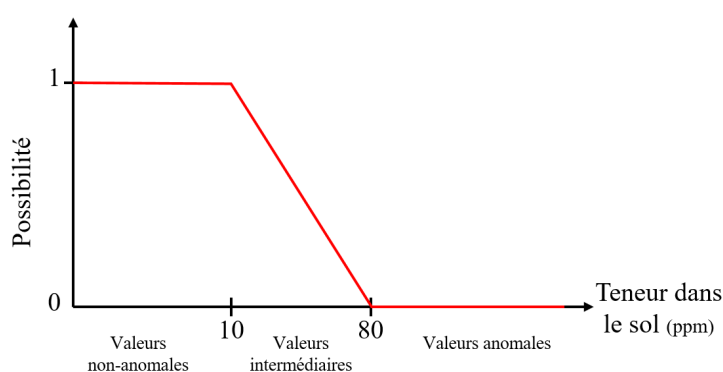


Figure 26 - Fond pédo-géochimique urbain flou.

- De l'information imprécise peut être propagée au travers d'un modèle de calcul de risque, au même titre que de l'information aléatoire.
- Une méthode pratique combine la technique bien connue dite « Monte Carlo » avec une recherche de limites [min - max].
- Ces résultats peuvent être comparés à des seuils qui peuvent être précis (le cas général) mais également imprécis.

2.4.8. Utilisation de données spatialisées

L'utilisation de cartes de concentrations calculées à partir d'une information très lacunaire est d'abord examinée en distinguant le cas d'un site de celui de l'évaluation du fond pédo-géochimique anthropisé, dont la définition même n'est pas encore définitive (voir discussion précédente). Le cas de mesures imprécises est ensuite examiné, en particulier celui d'une reconnaissance par des mesures indirectes, nombreuses mais peu précises, pour la conduite de la dépollution par excavation des terres. La notion de « support » pour l'estimation est enfin discutée.

a) Tracer une carte de concentrations d'un site et l'utiliser

Quel que soit le procédé de dépollution, sa mise en œuvre sur de grands sites nécessite de disposer d'une ou de plusieurs cartes de concentrations (composé ou familles de composés organiques, éléments métalliques). Les données sont fournies par des mesures de concentrations, éventuellement complétées par la description des formations pédologiques. Les concentrations n'étant pas connues de manière exhaustive (ni en surface, ni en profondeur), les zones à dépolluer sont donc délimitées à partir d'un résultat d'interpolation ou d'estimation, et non d'après une carte « exacte ».

• Rappels sur les méthodes d'interpolation ou d'estimation

Une estimation est dite linéaire si la valeur calculée est une combinaison linéaire (une moyenne pondérée) des données. Le paramètre z désignant la concentration et x_i les coordonnées des données (le centre d'une carotte par exemple), la concentration estimée au point x s'écrit : $z^*(x) = \sum_i \ell_i z(x_i)$, l'étoile indiquant qu'il s'agit d'une estimation, et non de la valeur exacte. Les méthodes d'interpolation (estimation par le plus proche voisin, pondération par une puissance inverse de la distance par exemple) reposent sur l'intuition selon laquelle les concentrations dans deux échantillons proches sont semblables. Les pondérateurs attribués aux données les plus proches du point ou du bloc à estimer sont donc plus élevés.

Parmi les méthodes d'interpolation, le krigeage est l'estimateur linéaire optimal, au sens où, dans le modèle probabiliste explicite, l'estimation est sans biais et la variance de l'erreur d'estimation est minimale. Au lieu d'être déterminés par la seule géométrie des données et des points à estimer, les pondérateurs de krigeage dépendent (i) du support de la variable à estimer (« point », moyenne sur une maille de géométrie donnée), ainsi que (ii) de la variabilité spatiale de la variable étudiée. Pour une même géométrie des données et de la « cible » à estimer, les poids différeront selon la plus ou moins grande régularité spatiale de la variable.

La variabilité spatiale est quantifiée par le variogramme, qui reflète la variabilité des teneurs en deux points en fonction de leur distance (**Figure 27**). Des outils variographiques complémentaires (nuée et carte variographiques) sont commodes pour détecter des anomalies et faire apparaître d'éventuelles directions préférentielles (anisotropie). Le variogramme expérimental est calculé à partir des données, puis ajusté par un modèle résumant les principales caractéristiques de la variabilité spatiale.

Lorsque les données sont peu nombreuses, le variogramme expérimental, disponible à des distances et suivant des directions peu nombreuses, peut différer fortement du variogramme « exhaustif », inaccessible, qui serait calculé avec l'ensemble des concentrations de la zone d'étude. La technique de validation croisée peut aider au choix du modèle de variogramme qui sera utilisé pour le krigeage. Une étude de sensibilité des résultats d'estimation au modèle ajusté est une façon simplifiée de tenir compte de l'incertitude de modélisation.

Différentes techniques permettent de préciser la structure spatiale, en présence d'un variogramme « erratique » ... sous réserve de la disponibilité de données en nombre suffisant. Utiles pour préciser la variabilité spatiale, les transformations de données nécessitent une modélisation spécifique (anamorphose et estimation par krigeage disjonctif ou espérance conditionnelle), afin d'éviter des biais d'estimation.

Il existe de nombreuses variantes du krigeage, selon que l'estimation est monovariante ou multi-variante (plusieurs concentrations simultanément, ou prise en compte de co-variables pour améliorer la précision). Des modèles spécifiques, comme le modèle d'écraçage (Rivoirard *et al.* ; Donati *et al.*, 2018) ont été développés pour résoudre les problèmes posés par l'influence de quelques très fortes valeurs.

Le krigeage fournit également la carte de l'écart-type (la racine carrée de la variance) de l'erreur d'estimation, qui peut être corrigée pour tenir compte de l'effet proportionnel, généralement présent : la variabilité spatiale dépend de la moyenne locale de la concentration (cette correction a été effectuée sur la **Figure 28**).

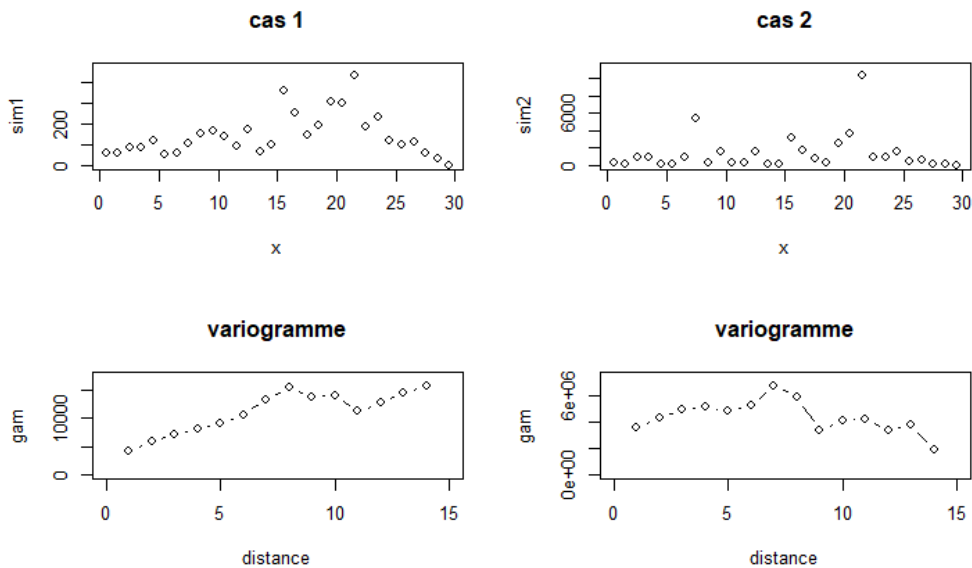


Figure 27 - Deux transects de concentrations (en haut) et leur variogramme respectif (en bas).

Note : la plus grande irrégularité du cas 2 par rapport au cas 1 se reflète par l'amplitude du variogramme (10^6 au lieu de 10^4), sa plus forte discontinuité relative à l'origine, et sa portée plus courte.

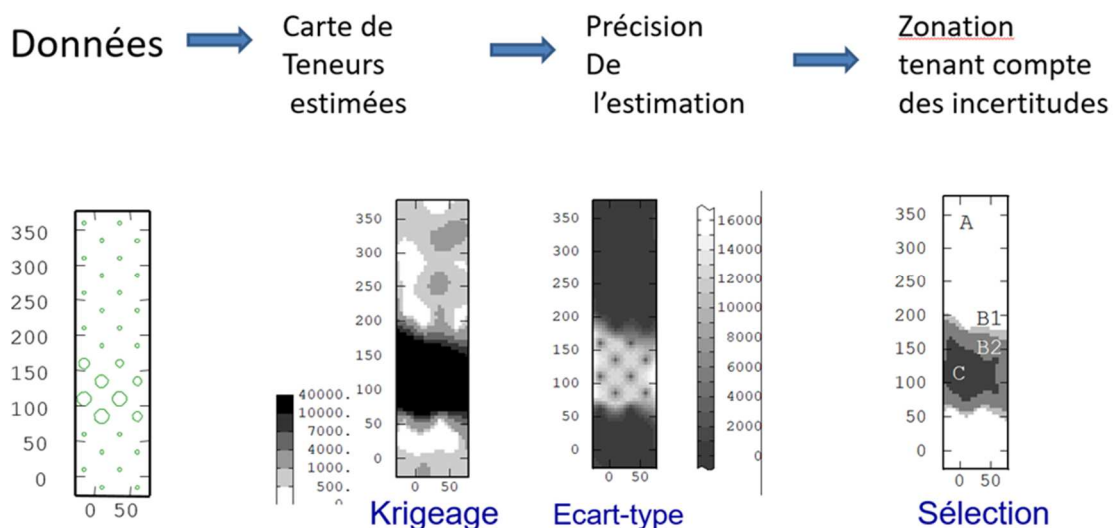


Figure 28 - Des données à la carte de la variabilité spatiale (projet LOQUAS, Benoit et al., 2010).

Le passage de l'écart-type de l'erreur d'estimation à un intervalle de probabilité nécessite de poser une hypothèse sur l'histogramme de l'erreur. Les estimateurs non-linéaires (krigeage disjonctif, espérance conditionnelle) ont été construits pour répondre de façon rigoureuse à cette question. Ces méthodes rigoureuses restent peu utilisées dans le contexte des sites et sols pollués, pour plusieurs raisons : (i) les bureaux d'étude sont très contraints en temps et pour certains, manquent d'expertise en géostatistique ; (ii) les hypothèses sur lesquelles ces méthodes « avancées » sont fondées ne sont pas toujours vérifiées expérimentalement, ce qui nécessite de les adapter aux cas traités. En pratique, une solution approchée pourrait être déduite empiriquement de l'étude de l'erreur standardisée (c'est-à-dire l'erreur rapportée à son écart-type de krigeage), obtenue par validation croisée ; mais la validité de telles approximations est encore à examiner attentivement.

- **Sélection d'après la carte des teneurs estimées ?**

La délimitation des zones dans lesquelles la concentration dépasse un seuil s fixé doit tenir compte de l'erreur d'estimation. En effet, sélectionner d'après la teneur estimée revient à sélectionner les blocs v_i dont la teneur estimée dépasse le seuil s , ce qui s'écrit $z^*(v_i) \geq s$. La teneur étant la somme de son estimation et de l'erreur d'estimation associée $r(v_i)$, $z(v_i) = z^*(v_i) + r(v_i)$, les blocs dont la teneur réelle dépasse le seuil s , sont ceux tels que $z(v_i) \geq s$, c'est-à-dire $z^*(v_i) \geq s - r(v_i)$, où l'erreur d'estimation $r(v_i)$ est inconnue, mais dont le modèle géostatistique précise que la moyenne (probabiliste) est nulle, et que la variance est la variance de krigeage. A l'aide d'une hypothèse sur l'histogramme de l'erreur standardisée, on peut choisir, à un risque statistique près, les quantiles q_ζ (négatif, et écrit $-|q_\zeta|$ dans la suite) et q_η (positif) de l'erreur standardisée définissant un intervalle, non nécessairement symétrique autour de la valeur estimée $z^*(v_i)$, intervalle auquel la valeur réelle inconnue a, dans le modèle, une forte probabilité d'appartenir : notant $\sigma_K(v_i)$ l'écart-type de krigeage, à un risque statistique que l'on se fixe près, $z^*(v_i) - |q_\zeta| \sigma_K(v_i) \leq z(v_i) \leq z^*(v_i) + q_\eta \sigma_K(v_i)$. La comparaison des bornes de cet intervalle au seuil s , fait apparaître trois cas (Cori, 2005) :

- $s \leq z^*(v_i) - |q_\zeta| \sigma_K(v_i)$; le seuil est inférieur à la borne inférieure de l'intervalle : le bloc se situe dans la zone polluée ;
- $z^*(v_i) + q_\eta \sigma_K(v_i) \leq s$; le seuil est supérieur à la borne supérieure de l'intervalle : le bloc n'est pas pollué ;
- $z^*(v_i) - |q_\zeta| \sigma_K(v_i) \leq s \leq z^*(v_i) + q_\eta \sigma_K(v_i)$; le seuil appartient à l'intervalle de probabilité : à cause de l'incertitude d'estimation, on ne peut savoir si le bloc est pollué ou non. Le bloc se situe dans la zone d'incertitude (**Figure 28**).

Selon les dimensions de la zone d'incertitude, il peut être utile de préciser deux parties (notées B1 et B2 sur la **Figure 28**) suivant que la concentration est inférieure au seuil ou le dépasse. Si elle est techniquement possible (délais, coûts), une reconnaissance complémentaire devrait se focaliser à l'intérieur et en périphérie (externe) de cette zone.

Les estimateurs non-linéaires (espérance conditionnelle et conditionnement uniforme en modèle gaussien anamorphosé, krigeage disjonctif) fournissant une estimation rigoureuse de la distribution locale (de l'histogramme local), permettent ces calculs. Cependant, les hypothèses du modèle le plus classique (le modèle gaussien anamorphosé) ne sont pas toujours vérifiées, notamment dans le cas de substances dont les concentrations sont fortement dispersées, comme certains polluants organiques. Les recherches sont donc à poursuivre.

En pratique, il est utile d'améliorer la précision des estimations, par exemple en tenant compte de mesures indirectes moins précises mais plus nombreuses. Mais les « points expérimentaux » communs aux deux types de mesures doivent être suffisamment nombreux pour rendre possible l'étude de la corrélation entre les deux types de mesures (comparaison des moyennes et des moyennes par classe de teneurs), ainsi que le calage d'un modèle variographique bivariable.

- **À propos du support de sélection**

Comme mentionné précédemment, la délimitation de la zone polluée dépend du support de sélection, rarement précisé. En effet, la proportion des carottes de sondages dont la concentration est supérieure au seuil s diffère de celle des « petits » blocs (par exemple de 2.5m x 2.5m x 1.5m) dépassant ce seuil, elle-même différente de celle des « grands » blocs (par exemple de 10m x 10m x 3m, soit 32 fois plus gros) dépassant le seuil s . Le support influe sur l'ordre de grandeur même des résultats.

Le support de sélection devrait donc être précisé en même temps que le seuil s . Il devrait en être tenu compte en cas de contrôle après dépollution, si la teneur de l'échantillon « de contrôle » dépasse le seuil. Les estimateurs géostatistiques non linéaires permettent en effet de calculer « la probabilité que la concentration d'un bloc dépasse le seuil, connaissant la teneur d'un échantillon dans le bloc ». En pratique, l'influence du support apparaît contournée par le choix du seuil, plus ou moins élevé selon le contexte.

b) Le fond géochimique urbain : quelle définition ?

Le fond pédo-géochimique en milieu anthropisé est généralement calculé à l'aide de critères statistiques élaborés pour la détection d'anomalies : quantiles d'ordre élevé (90 ou 95 %), critère construit comme la somme d'une valeur centrale et d'un indicateur de dispersion (vibrisse supérieure, écart à la médiane). Par défaut, les concentrations inférieures au critère retenu sont supposées correspondre au « fond ».

Ce calcul présente plusieurs inconvénients : le calcul statistique usuel ne tient pas compte de la localisation des données, ni de la corrélation spatiale des variables. Par ailleurs, comment le paramétrage est-il argumenté ? Par exemple, les critères sont-ils calculés avec l'ensemble des données du domaine d'étude, ou par « voisinages glissants », et dans ce cas, comment ce voisinage est-il déterminé ? Comment les informations disponibles sur le milieu anthropisé (occupation des sols actuelle et passée, formation géologique du substratum notamment) sont-elles prises en compte ?

Sous certaines hypothèses, les critères statistiques suivent un ordre partiel (Sauvaget, 2019), mais ces hypothèses ne sont pas nécessairement vérifiées (distribution symétrique ou faiblement dispersées par exemple). Le choix d'un unique critère statistique risque donc d'être empirique.

Une autre démarche consiste à définir le fond pédogéochimique anthropisé d'après les différentes échelles de variabilité de la substance dans le domaine d'étude, par analyse krigeante, également connue sous la dénomination de « krigeage factoriel » (Sauvaget, 2019). La décomposition de la variable en composantes d'échelle repose sur une analyse variographique, et peut tenir compte des co-variables décrivant le milieu anthropisé. Le résultat dépend du domaine d'étude, ce qui est conforme à l'intuition : une anomalie sur un grand domaine peut être partiellement rapportée au « fond anthropisé » sur un domaine plus petit, qui serait entièrement sous l'influence d'une source de pollution donnée. On retrouve ici l'incertitude liée à la définition de la grandeur d'étude.

Le calcul, par analyse krigeante, d'une carte du fond et des anomalies nécessite de préciser à quelle composante rattacher « la moyenne » de la variable étudiée, ce qui introduit une indétermination sur le résultat. Selon le choix de modélisation, les « anomalies » peuvent ainsi être de moyenne nulle (dans le modèle). Différentes définitions du fond et des anomalies, qui diffèrent localement par une constante (dont la valeur change avec le voisinage) sont donc possibles. La réflexion sur cette indétermination mérite d'être poursuivie, afin de mieux argumenter les cartes obtenues.

Enfin, l'utilisation des cartes ainsi obtenues est un sujet à approfondir :

- choix de la maille, et du support : estimation d'une grandeur « ponctuelle » (au nœud de la grille), ou en moyenne spatiale par maille (valeur régularisée). Il est envisageable de retenir un support plus grand pour l'estimation « du fond » que pour l'estimation des « anomalies » ;
- pour le fond, faut-il considérer un maximum local (et dans ce cas, dans quel voisinage ?) des valeurs estimées ?
- comment tenir compte des variances d'erreur d'estimation du fond et des anomalies fournies par l'analyse krigeante, variances qui dépendent de la localisation des données, de la maille estimée ainsi que du ou des variogrammes (variogrammes simples et croisés de la variable et des covariables).

- L'étude d'un site nécessite des données de teneurs, que l'analyse exploratoire géostatistique valorise, suite à une phase d'analyse de la qualité et de tri des données, par une description synthétique de leur répartition spatiale. Les estimateurs géostatistiques (krigeage et ses variantes) améliorent la précision des cartes, mais ne pallient pas le manque de données et/ou leur qualité insuffisante. Une reconnaissance insuffisante d'un site a comme conséquence des incertitudes qui ne sont pas toujours toutes identifiées.
- Les mesures rapides (indirectes) sur site peuvent permettre d'améliorer à moindre coût la précision des estimations, sous réserve que les points communs aux deux types de mesures soient suffisamment nombreux et que leur liaison soit suffisamment élevée. La prise en compte de co-variables comme les formations pédologiques améliore les estimations, mais nécessite une étude exploratoire détaillée.
- La délimitation des zones à dépolluer doit tenir compte des incertitudes d'estimation et ne peut donc être effectuée directement d'après la carte des teneurs estimées. Différentes méthodes de comparaison à un seuil de teneur sont disponibles, mais restent peu utilisées, du fait des contraintes de temps auxquelles les bureaux d'études sont soumis et parce que ces méthodes nécessitent une expertise en géostatistique, notamment dans les cas complexes.
- La détermination d'un seuil de dépollution devrait préciser "le support" (le volume) associé. En pratique, la référence au support de sélection apparaît contournée par le choix du seuil, plus ou moins élevé selon le contexte du site.

2.5. Décision et communication

2.5.1. Une approche multi-partenaire

Trop souvent, les approches adoptées pour appréhender et communiquer les incertitudes en gestion des sites et sols pollués sont avant tout techniques et calculatoires. Mais des approches plus « *partenariales* » peuvent également être envisagées dans une recherche de meilleure valorisation d'une « *intelligence collective* ». En effet, dans un état de l'art relatif à l'implication des populations dans la gestion des sites et sols pollués (INERIS-IRSN, 2008), il est montré comment cette implication peut aider de manière significative le processus de gestion et cela malgré les incertitudes inhérentes, en raison d'une meilleure compréhension partagée des enjeux par les différentes parties prenantes

Dans une étude réalisée pour l'ADEME, Hazebrouck *et al.* (2008) proposent des modalités et outils de communication conçus pour un contexte de site et sols pollués, mais qui sont également adaptables à d'autres types de risques environnementaux (voir la plateforme COMRISK ; <https://comrisk.fr>). Des approches multipartenaires de gestion de friches urbaines ont aussi été proposées notamment par le Baltic Urban Lab (2018). Ces exemples soulignent que la manière de représenter le risque et les incertitudes associées doit nécessairement tenir compte de la nature des interlocuteurs qui reçoivent l'information. En effet, dans une étude sociologique réalisée sur la base de questionnaires et d'entretiens auprès de populations souffrant des nuisances associées à la présence d'une source de pollution atmosphérique, Harpet *et al.* (2005) concluent que « *communiquer sur les incertitudes auprès de personnes qui sont « certaines » d'être des victimes n'est pas une option...* ». Dans ce contexte, une étape préalable à toute communication avec des parties-prenantes non-techniques, est la mise en commun d'éléments de langage compris par tous : vocabulaire, formulations, images, concepts, ...

Au-delà de la restitution de résultats d'étude (courriers, réunion publique, ...), la tenue de réunions de travail avec les parties prenantes (incluant notamment les riverains) apparaît aujourd'hui de plus en plus nécessaire, y-compris en amont de l'étude afin de présenter en détail la démarche, les hypothèses à retenir qui pourront alors être discutées et par la suite prendre le temps de discuter des résultats et des limites de connaissance.

2.5.2. Communiquer sur l'incertitude d'origine épistémique

La communication entre et avec les parties prenantes d'un dossier de SSP, que ce soit les gestionnaires du site (propriétaire, exploitant, promoteur, ...) de l'Administration (Préfet, Bureau de l'environnement, DREAL, ARS, ...), de la société civile (représentants de collectivités, associations, riverains, ...), est un aspect incontournable de la gestion des sites et sols pollués. Par rapport à la communication sur les incertitudes, l'expérience montre qu'il est difficile de communiquer sur des critères de dépassement de seuils de risque exprimés dans un cadre probabiliste classique (à partir de distributions uniques). On peut donc s'attendre à encore plus de difficultés dès lors qu'il s'agit de communiquer sur des degrés de probabilité *imprécis* de dépassement de seuils... (cas des incertitudes d'origine épistémique).

Pour pallier cette difficulté, on peut s'inspirer de ce qui est pratiqué dans le domaine de la météorologie, dans lequel des scientifiques s'appuient sur un grand nombre de mesures pour effectuer des prévisions et les communiquer à un public très large. Différence significative avec le domaine des sites et sols pollués : les météorologues ont très rapidement un retour sur la fiabilité de leurs prévisions, alors que dans le domaine des sites et sols pollués, où les phénomènes sont généralement très lents, un retour sur les prévisions est généralement obtenu seulement après de nombreuses années. Il est néanmoins très instructif d'examiner comment cette communauté communique sur les incertitudes. En effet, comparé à la communauté de la gestion des sites et sols pollués, la communauté des météorologues est particulièrement expérimentée en matière de communication.

Les météorologues s'appuient sur des prévisions dites « d'ensemble » (Palmer et al., 2005), qui consistent à effectuer des calculs de flux atmosphériques à l'échelle de la planète à partir de plusieurs états initiaux jugés « plausibles ». Ces différents calculs sont accompagnés d'un estimateur de fiabilité qui n'est pas appelé « probabilité » (terme qui évoque une certaine exhaustivité de l'exploration des alternatives), mais plutôt « indice de confiance ». À noter que ce terme, beaucoup plus souple que celui de probabilité, semble bien perçu par le public avec lequel les météorologues communiquent (journaux télévisés, radio, presse). L'indice de confiance, typiquement une valeur entre 1 et 5, fournit au public une estimation qualitative de la fiabilité de la prévision (5 très fiable, 1 pas fiable, ...). Le public qui reçoit une prévision assortie d'un « indice de confiance », n'a pas besoin de connaître les détails mathématiques sous-jacents à son élaboration. Il lui suffit de savoir que l'expert a confiance ou non dans sa prévision.

S'inspirant de cet exemple, Dubois et Guyonnet (2011) ont utilisé la notion d'indice de confiance pour proposer un estimateur unique du degré de dépassement d'un seuil (précis ou non). S'appuyant sur les travaux de Hurwicz (1951), ils proposent de calculer un indice de confiance (IC) à partir d'une moyenne pondérée des limites de probabilité haute et basse de la **Figure 24b** précédente :

$$IC = \alpha PI + (1 - \alpha) Cr$$

Où α est un facteur de pondération.

Une valeur $\alpha = 0$ revient à supposer $IC = Cr$: on prend l'indicateur le plus défavorable, ce qui peut sembler exagérément pessimiste et majorant puisqu'on ignore toutes les informations disponibles suggérant un risque moins sévère. Une valeur $\alpha = 1$ revient à supposer $IC = PI$, ce qui est exagérément optimiste et minorant et difficile à justifier dans un contexte de sites et sols pollués où il y a une attitude d'aversion au risque. Il a été proposé conventionnellement d'assigner à α une valeur 1/3, ce qui revient à donner un poids de 1/3 à la limite optimiste du degré de dépassement et de 2/3 à la limite pessimiste. Il est suggéré que cette approche permet de définir un indicateur unique « raisonnablement majorant » du degré de dépassement du seuil. Cet indice est illustré dans la **Figure 29**, qui reprend la **Figure 24b**.

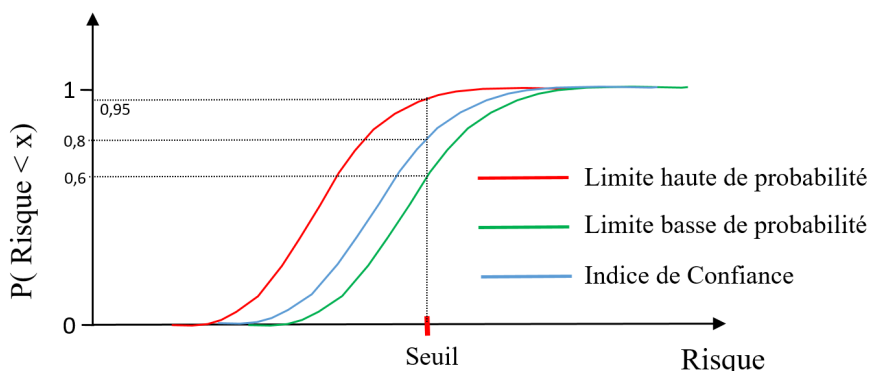


Figure 29 - Famille de distributions de probabilité et indice de confiance.

On pourrait objecter à cette approche d'introduire une part de subjectivité, alors que la critique formulée précédemment par rapport à l'approche dite « Bayésienne » avait justement trait à son caractère subjectif. Mais ici la subjectivité est introduite à une étape d'aide à la décision, pas en amont de l'analyse de risque en confondant variabilité aléatoire et méconnaissance, ce qui est le cas lorsqu'on « décrète » des distributions de probabilité uniques sans justification par des données.

Lorsqu'on a affaire à des sources d'informations incomplètes ou imprécises, la restitution d'une estimation de risque sous la forme d'une famille de distributions de probabilité comme celle représentée schématiquement dans la **Figure 29**, présente l'avantage d'illustrer le champ des valeurs de risque jugées possibles. La distance qui sépare les limites de probabilité haute et basse dans la **Figure 29** est directement le reflet de notre méconnaissance vis-à-vis de certains paramètres influençant le risque. Si cette distance est jugée trop importante, il peut être décidé de la réduire en améliorant la connaissance (par de la mesure) relative à tel ou tel paramètre d'influence (une analyse de sensibilité peut aider à identifier les paramètres les plus influents). Ce type de décision ne peut pas découler d'une analyse où il aura été choisi de représenter les paramètres mal connus par des distributions de probabilité uniques (par exemple des distributions uniformes) car il n'y a dans ce cas aucun moyen de distinguer, dans la variabilité du résultat calculé, ce qui découle de la variabilité aléatoire de ce qui découle du caractère incomplet de l'information (Dubois et Guyonnet, 2011).

À noter que le seuil dans la **Figure 29** est « précis », mais que le même exercice de comparaison pourrait être réalisé dans le cas d'un seuil « flou » (cf. **Figure 25**). Ce cas de figure mériterait d'être testé dans la pratique, notamment pour établir s'il représente une plus-value en termes de souplesse par rapport au cas du seuil précis. En effet, il est fréquent dans la pratique de trouver des situations où un léger dépassement de seuil sur un ou deux paramètres peut remettre en question une solution de gestion. Mais si on considère toutes les incertitudes associées à l'élaboration de certains seuils (par exemples d'exposition sanitaire), l'utilisation de zones de transition entre ce qui paraît totalement acceptable et ce qui ne l'est pas du tout, peut sembler une solution pragmatique.

- Tenir compte du caractère incomplet/imprécis de l'information disponible permet de communiquer sur l'étendue des résultats jugés plausibles.
- Lorsqu'il s'agit de comparer ces résultats à un seuil d'acceptabilité (précis ou non-précis), un indicateur unique peut être estimé (appelé ici « indice de confiance »).
- Compte tenu des nombreuses incertitudes affectant l'élaboration des seuils d'acceptabilité, l'utilisation de seuils de risque non-précis mériterait d'être explorée.
- La communication sur les risques ne peut pas se limiter à des approches techniques et/ou calculatoires : des approches plus sociologiques et partenariales, faisant appel à l'intelligence collective, sont nécessaires.

2.6. Utilisation pour des mesures de gestion ou de surveillance

Les mesures de gestion dans le domaine des sites et sols pollués sont nombreuses et incluent notamment l'excavation de terres polluées, les opérations de dépollution, de mise en sécurité, et/ou de surveillance des milieux, au cours desquelles de nouvelles données seront collectées. Une des principales parties prenantes impactées directement (notamment financièrement) par les incertitudes sont les entreprises responsables de la mise en œuvre des mesures de gestion. Dans une étude pour l'association RECORD sur les méthodes d'estimation et d'optimisation des quantités de terres polluées à traiter (Jeannée et Demougeot-Renard, 2016), un questionnaire a été envoyé à des professionnels de la dépollution des sites pollués pour recueillir un retour d'expérience relatif aux écarts observés entre les estimations résultant des diagnostics et la dépollution proprement dite. Les résultats de ce questionnaire illustrent combien les incertitudes associées aux phases de diagnostic, d'analyse, ou d'interprétation, peuvent influencer les objectifs des travaux de dépollution et entraîner des risques financiers (voire judiciaires) pour les entreprises.

Parmi les réponses concrètes on note en particulier que les packs analytiques sont souvent insuffisants pour valider de manière fiable les filières de traitement, que les diagnostics sont souvent trop peu précis en termes d'épaisseurs de terrains investigués, l'absence d'informations sur, par exemple, la présence d'eau dans les sondages, celle d'éléments volatils, etc. Ici également, l'utilisation d'objectifs de dépollution alternatifs à des « seuils couperet » (voir la section 2.4.7) pourrait permettre de mieux gérer les incertitudes.

2.7. Bancarisation des données / métadonnées

Le stockage des données est un aspect essentiel de la gestion des sites pollués, pour l'information des parties prenantes par rapport aux risques associés à ces sites et pour conserver la mémoire en lien avec les éventuelles servitudes. Le BRGM réalise à la demande de l'État le recensement et la diffusion d'informations relatives aux anciennes activités industrielles. L'administration est quant à elle chargée du suivi et de la communication sur les sites et sols pollués ou potentiellement pollués appelant une action des pouvoirs publics, à titre préventif ou curatif. Il s'agit respectivement des bases de données BASIAS et BASOL, accessibles au public. Sur les 318 000 sites BASIAS qui ont connu par le passé une activité industrielle, environ 7 000 font l'objet d'une surveillance particulière et sont référencés dans BASOL, en raison de leur état.

Plus récemment, le BRGM développe, avec le soutien de l'ADEME, la base de données BDSolU, qui contient des analyses chimiques de sols prélevés en milieu urbain, ainsi que des informations qui servent à décrire leur prélèvement. Ces données améliorent la connaissance de la qualité géochimique des sols en milieu urbain et des valeurs des fonds pédo-géochimiques anthropisés. Le site internet dédié permet la bancarisation en ligne des analyses de sols produites par les fournisseurs déclarés et la consultation des données mises à disposition du public. La question de l'incertitude est abordée dans la base BDSolU, avec notamment un champ prévu pour que les laboratoires renseignent sur les incertitudes d'analyse (section 2.3.). D'autres sources d'incertitude, par exemple sur les coordonnées x, y des points de prélèvement, pourraient également être abordées.

Parmi les questions clé en matière de bancarisation des données et de diffusion en direction des parties prenantes, il y a celle des passerelles entre différentes bases, ce qui touche à l'interopérabilité des données (leur utilisation pour différentes applications ou plateformes qui peuvent communiquer entre elles). Un domaine d'interopérabilité qui est particulièrement actuel dans le domaine de la gestion des sites et sols pollués est celui du BIM (Building Information Modeling ; voir Molleron, 2019), qui permet notamment une visualisation 3(4)D de la donnée tout le long du cycle de vie d'un bâtiment, d'une infrastructure, etc. Parmi les enjeux de recherche et développement figure le développement de passerelles entre les informations relatives à la qualité des sols pollués et le BIM.

3. Perspectives de recherche

3.1. Introduction

Ce bref et très partiel survol de la notion d'incertitude dans le cycle de la donnée associée à la gestion des sites et sols pollués illustre combien les questions sont nombreuses et importantes pour une aide à la décision efficace et cohérente avec les informations disponibles. Les incertitudes sont inhérentes à toute étude environnementale et aucune méthodologie ne permettra de les éliminer en totalité. Certaines incertitudes sont d'ailleurs non quantifiables, car on ne peut quantifier ce que l'on ne connaît pas du tout : il peut s'agir par exemple de problèmes mal maîtrisés d'un point de vue scientifique, comme les effets synergétiques de mélanges de polluants en termes de toxicité, ou encore le comportement et les effets de certains polluants émergents.

Mais on peut tout au moins réduire les incertitudes par la mise en œuvre de techniques adaptées (d'échantillonnage, d'analyse, etc.) et essayer de les prendre en compte de manière cohérente pour alimenter le processus d'aide à la décision. Les sections suivantes présentent quelques pistes de recherche, en repartant du schéma de la Figure 1.

3.2. Échantillonnage

Dans le domaine de l'échantillonnage des milieux (sols, eaux, gaz, ...), il faut poursuivre les travaux destinés à améliorer la représentativité des échantillons et des analyses, d'une part en optimisant les plans d'échantillonnage par rapport à une connaissance du comportement des polluants et d'une pré-connaissance de leur répartition spatiale compte tenu des études historiques et documentaires et d'autre part en améliorant les méthodes de terrain qui permettent d'éviter l'évolution physico-chimique des échantillons durant leur acheminement vers le laboratoire d'analyse.

Le développement de méthodes indirectes (géophysique, ...) est à poursuivre pour optimiser l'implantation des sondages de prélèvement. Les approches développées ces dernières années pour caractériser in situ des flux polluants, plutôt que des concentrations, méritent d'être développées pour permettre d'optimiser le traitement des zones les plus contributrices à la pollution et représentant le plus de risques pour des ressources en eaux souterraines par exemple.

3.3. L'incertitude liée aux analyses de laboratoire

Pour ce volet, parmi les points offrant des perspectives de travaux futurs on notera :

- la synthèse de données « d'incertitude profession », à destination des gestionnaires pour les paramètres et matrices d'intérêt en contexte sites et sols pollués ;
- l'organisation d'essais inter-laboratoire pour des paramètres « orphelins » (pour lesquels il n'existe pas d'essais à l'heure actuelle sur le marché) et préparation de matériaux de référence pour des paramètres à enjeu. L'objectif serait de vérifier la justesse et la comparabilité des données pour ces paramètres ;
- estimer de façon spécifique les parts d'incertitude liées à la préparation physique de l'échantillon au laboratoire (quartage, broyage, ...) versus la part liée à l'analyse.

3.4. Représentation et utilisation d'information incertaine

3.4.1. Représentation d'information incertaine

Comme suggéré précédemment, l'organigramme de la Figure 18 mériterait d'être précisé pour mieux identifier les conditions d'utilisation des différents outils mathématiques servant à représenter et propager de l'information incertaine. Mais au-delà de cet outil, une réflexion pourrait être engagée autour de la représentation de donnée d'expert, qui est une source essentielle d'information dans la pratique de la gestion des sites et sols pollués. Il semblerait particulièrement opportun de mener cette réflexion en lien avec la gestion des différentes bases de données développées au cours des années dans ce contexte (sur les sols ; BASIAS, BASOL, ASPITET, DONESOL, BDSolU, sur les plantes ; BAPPET, BAPPOP, sur la qualité de l'air, etc.), car le jugement d'expert se nourrit de telles sources d'informations.

3.4.2. Seuils non-précis

La section 2.4.6. a illustré la notion de seuil non-précis (ou seuil flou). Cette notion pourrait être « testée » dans des contextes concrets de gestion des risques liés aux sites et sols pollués. Dans un premier temps, en termes de leur construction (collecte des informations, interprétation et représentation de seuils de risques non-précis). Et dans un second temps en termes de communication avec les parties prenantes (section 2.5) pour différentes situations concrètes.

3.4.3. Respect d'un seuil sous contraintes

Souvent en gestion des sites et sols pollués, on cherche à identifier un objectif de qualité des sols pour lequel le niveau de risque, estimé à l'aide d'un « modèle », est en-deçà d'un seuil de risque jugé acceptable (par exemple le seuil de 10^{-5} pour les substances cancérigènes). Le problème inverse consiste à déterminer cette qualité de manière automatisée, en tenant compte des contraintes liées aux facteurs qui influencent le risque. Si certains de ces facteurs sont connus de manière incomplète/imprécise et représentés à l'aide de distributions de possibilités, il s'agit d'un problème d'optimisation sous contraintes floues. S'il existe plusieurs exemples dans la littérature concernant l'application de ce type d'approche, ce n'est pas encore le cas à notre connaissance pour le contexte des sites et sols pollués.

3.4.4. Méthodes géostatistiques

a) Estimation de teneurs fortement contrastées dans les sols : poursuivre l'amélioration des modèles

Des travaux antérieurs ont déjà mis en évidence l'utilité du « krigeage disjonctif discret » pour l'estimation de teneurs fortement contrastées. Développé dans le cas mono-variable, cet estimateur doit être adapté au cas multi-variable : plusieurs substances, cas de co-variables comme la composition des sols ou les mesures rapides sur sites.

D'autres tests sont à effectuer, pour différentes pollutions organiques ou métalliques. Un guide est à rédiger pour rendre cette méthode applicable par les bureaux d'études, en tenant compte des particularités des pollutions de sols, comme la non-stationnarité verticale des teneurs souvent observée.

Enfin, les comparaisons de modèles (écrêtage, krigeage disjonctif avec différents modèles de loi spatiale) sont à effectuer pour différents types de pollutions (métallique, organique). Les résultats seront utiles comme aide à leur mise en œuvre par les bureaux d'études.

b) Estimation du fond pédo-péochimique anthropisé

L'estimation du fond pédo-géochimique anthropisé à partir de la décomposition des teneurs selon différentes échelles de variabilité donne un sens précis au fond pédo-géochimique. Plusieurs questions restent en suspens :

- Comment passer de la carte à un « seuil » pour la gestion des terres ?

- Quelle valeur donner au seuil : une moyenne spatiale, le maximum local ou un quantile sur une maille ? Quel doit être le support ?
- Comparer la décomposition actuelle par analyse krigeante à une décomposition fondée sur la discrétisation des teneurs par classe (KDD), en examinant le « surplus de teneur » estimé pour caractériser les « anomalies ».

En effet, dans l'estimation du « fond » par analyse krigeante, la moyenne des anomalies est nulle (Sauvaget, 2019). Cette propriété est intéressante : la teneur peut être inférieure au fond géochimique estimé, ce qui est utile pour la gestion des terres excavées. Mais d'autres décompositions de la teneur en composantes associées à différentes « échelles » de variabilité spatiale existent, comme celle mise en œuvre dans le KDD. Dans ce modèle, la teneur peut être vue comme un ajout de valeurs successives (intervalles de 0 à 10ppm, de 10 à 50ppm, de 50 à 150ppm) pondérées par les indicatrices (connues et égales à 0 ou 1 sur les données), chaque indicatrice de classe ayant son propre variogramme. On obtient ainsi une description de la variabilité spatiale de la teneur par classe d'intervalles de valeurs, plus « détaillée » que la décomposition par analyse krigeante, et qui pourrait « résoudre » la question de la moyenne des composantes (nulle pour les « anomalies » de l'analyse krigeante). La comparaison de ces deux décompositions de la teneur par échelle de variabilité (analyse krigeante et KDD) devrait permettre d'affiner le sens « physique » donné au fond pédo-géochimique, en précisant quels « intervalles de valeurs » successifs contribuent à quelle échelle de variabilité spatiale. Schématiquement, là où l'analyse krigeante tient compte des seules échelles de variabilité spatiale, l'approche proposée par krigeage disjonctif considère simultanément les classes de valeurs (basses, intermédiaires ou fortes) avec leur variabilité spatiale propre. De plus, la modélisation multivariable est à considérer, comme pour l'analyse krigeante.

Dans le cas de données présentant une importante hétérogénéité de répartition (nombreuses à certains endroits mais éparses à d'autres, avec des effets de clusterisation), la question d'une combinaison de méthodes par krigeage et par estimateurs flous mériterait également d'être étudiée. Un tel axe de recherche serait l'occasion de revisiter les travaux de Loquin et Dubois (2012) sur le krigeage à partir de variogrammes et de données imprécis. On peut également s'interroger sur les apports potentiels de l'IA (intelligence artificielle) dans ce contexte : en exploitant des quantités de données considérables, l'IA pourrait-elle aider à identifier des corrélations permettant de mieux faire ressortir des anomalies et par conséquent ce qui peut être assimilé au fond pédo-géochimique (au sens « non anomal ») ?

3.5. Communication avec les parties prenantes

Si l'utilisation de seuils non-précis (avec une zone de transition) peut présenter certains avantages, leur degré de « *recevabilité* » dans un contexte de communication avec les parties prenantes pose également question. On pense, d'une part, aux personnes potentiellement exposées qui subissent le risque, mais également aux personnels de l'administration qui doivent décider des mesures de gestion. Dans cet axe de recherche, il s'agirait donc d'introduire une dimension SHS dans la démarche. En particulier, il faut développer des outils d'aide à la concertation, à la compréhension partagée et à l'appropriation des enjeux liés à la gestion des sites et sols pollués et aux incertitudes associées. Cela concerne par exemple les servitudes associées à des sites pollués (interdiction de consommation de légumes autoproduits, non-consommation de l'eau, etc.). Si de telles contraintes ne sont pas « comprises », elles ne peuvent pas être acceptées ou mises en application par les populations concernées.

Des approches de type « jeu sérieux » ou de simulation participative commencent à être de plus en plus utilisés pour la concertation avec les parties prenantes et pour mieux intégrer l'intelligence collective dans la gestion des sites dégradés. C'est le cas par exemple pour les conflits d'usage de l'eau (Barnaud *et al.*, 2007). La simulation participative croise la modélisation et les approches participatives, pour produire des apprentissages sociaux sur une problématique environnementale (Becu *et al.*, 2017). Les jeux de rôles multi-acteurs (mise en situation collective et apprentissage social) permettent d'aborder des questions comme : quel type(s) de décision(s) individuel(s) ou collectif(s) de quel(s) acteur(s) ont quelles influences connues aujourd'hui sur la gestion des sites et sols pollués (ou de son équivalent dans le jeu) ? L'outil peut aider à identifier les différents jugements et perceptions qu'ont les acteurs du territoire d'étude par rapport à leur problématique actuelle liée à un site pollué.

3.6. Utilisation pour la gestion et la surveillance

Parmi les avantages d'une restitution de la méconnaissance dans la propagation des incertitudes, il y a les éventuelles conséquences en termes de gestion et de surveillance. Si l'écart entre les limites inférieures et supérieures de probabilité d'être conforme aux objectifs de qualité ou de risque, l'étude peut basculer vers une nouvelle phase d'acquisition afin de réduire les incertitudes. Cette démarche s'accompagnera d'une phase d'analyse de sensibilité afin d'identifier quels sont les paramètres dont une meilleure connaissance permet de réduire le plus les incertitudes globales. Des approches de type coûts-bénéfices pourraient être développées pour accompagner ce type de démarche.

Une étape essentielle en matière de gestion est l'arrêt des opérations de dépollution (quand, comment, selon quels critères ?...). La prise en compte des incertitudes dans la définition des critères d'arrêt de ces opérations est un objectif pertinent.

Mais à l'issue des études et de la mise en œuvre des mesures de gestion, il est possible que des incertitudes incompressibles demeurent quant à l'acceptabilité de la qualité des milieux et des risques. Dans ce cas des mesures compensatoires pourront être considérées telles que les mesures constructives (par exemple pour la gestion du risque d'intrusion de vapeur).

3.7. Bancarisation des données/metadonnées

Les informations relatives aux incertitudes associées aux différentes étapes de l'acquisition de la donnée environnementale font souvent défaut dans les bases de données. Si une des raisons évidentes est la relative complexité du problème, il est vraisemblable que l'obligation supposée a priori d'appliquer un cadre probabiliste classique freine également l'évolution de cette question. En effet, l'application de ce cadre suppose que l'incertitude relative à chaque étape de l'acquisition puisse être caractérisée (dans le cas Gaussien) par un écart-type autour d'une valeur moyenne et que les interactions entre les étapes soient suffisamment connues (et quantifiables) pour arriver à une incertitude probabiliste globale. Face à la difficulté (voire l'impossibilité) de cette démarche, la tendance est, soit de passer sous silence la question des incertitudes (cas le plus fréquent), ou alors d'en fournir une indications très partielle (en se limitant par exemple à l'incertitude analytique puisqu'elle fait l'objet de normes ; cf. section 2.3).

Mais d'autres approches peuvent être explorées. Notamment dans le cadre d'une « *réhabilitation* » de la notion d'intervalle en métrologie environnementale, l'investigateur de terrain (qui est la personne la plus à même de qualifier l'incertitude relative à sa mesure), pourrait fournir un intervalle sur la base d'un jugement d'expert, dont l'étendue du support dépendrait des conditions de prélèvement (bonnes versus mauvaises). Ainsi, dans la base de données, une mesure ne serait pas représentée par une valeur unique, mais par un intervalle avec le cas échéant une ou des préférence(s) au sein de cet intervalle. Alternativement, la qualité du prélèvement et de la mesure pourraient être qualifiée par un indicateur de type « indice de confiance » (par exemple un chiffre compris entre 1 ; pas fiable à 5 ; très fiable).

La question de la représentation des incertitudes dans les données bancarisées pose celle de l'interopérabilité de ces bases de données avec d'autres outils. On pense par exemple aux passerelles entre le BIM (Building Information Modelling), qui se développe de plus en plus pour gérer et représenter les données relatives aux constructions et infrastructures tout au long de leur cycle de vie, avec les informations sur la qualité des sols dans un contexte de friches industrielles par exemple. Si des réflexions concernant l'interopérabilité de ces différentes bases sont actuellement en cours (par ex. le projet MINDD ; <http://www.minnd.fr/>), la question de la représentation des incertitudes relatives aux données mérite sans doute d'être creusée.

4. Conclusions

Comme il a été dit en introduction, ce document n'a aucune prétention d'exhaustivité et ne fait qu'effleurer la surface du problème de la gestion des incertitudes dans le domaine des sites et sols pollués. Il tente néanmoins de poser certaines questions fondamentales, telles que : *quelle est la différence entre savoir et ne pas savoir dans ce contexte ?* De la réponse à cette question dépend en effet le choix des méthodes et outils pour la représentation et la propagation des incertitudes, dont plusieurs exemples ont été présentés dans ce rapport.

C'est d'ailleurs là une question qui paraît pertinente à un moment où le mythe de « *l'expert omniscient* » a clairement vécu. En effet, la réalité sociétale de la gestion des sites et sols pollués aujourd'hui et de manière plus générale celle des sites « *dégradés* » (par ex. les anciens sites miniers) est bien souvent celle d'une remise en cause de « *l'expertise officielle* » et la mise en place d'une « *expertise profane* », qui peut d'ailleurs dans bien des cas se révéler très pertinente (d'où l'intérêt de développer l'intelligence collective...). Or face à la complexité des phénomènes en jeu dans ce domaine, l'expert a certainement le droit de ne pas (tout) savoir et il est dans ce cas recommandé d'en faire état. En effet, l'affichage de certitude en dépit de réelles incertitudes dans un contexte d'aversion au risque (ce qui est le cas du contexte des sites et sols pollués) peut se révéler particulièrement contre-productif dès lors que les faits contredisent la confiance affichée (or les faits sont souvent tenaces...).

Dans ce type de situation, la « *défiante* » évoquée plus haut vis-à-vis de l'expertise officielle s'accompagnera souvent par la recherche de responsabilités. Or une meilleure prise en compte des incertitudes, notamment d'origine épistémique (liées aux défauts de connaissance), associée à une meilleure compréhension partagée des enjeux liés à la gestion des sites et sols pollués, pourrait permettre d'éviter une évolution vers plus de défiance vis-à-vis des acteurs opérationnels du domaine.

REFERENCES BIBLIOGRAPHIQUES

- Abdo, H., Flaus, J.M., Masse, F., 2017. Uncertainty quantification in risk assessment - Representation, propagation and treatment approaches: Application to atmospheric dispersion modelling. *Journal of Loss Prevention in the Process Industries*, 49, 551-571.
- ADEME – BURGEAP, 2015. Etude de comparaison de protocoles de prélèvement, d'échantillonnage et de préparation des échantillons pour l'analyse des COV dans les sols, recommandations pour une homogénéisation des (bonnes) pratiques.
- AFNOR (2013). Norme NF ISO 11352. Qualité de l'eau - Estimation de l'incertitude de mesure basée sur des données de validation et de contrôle qualité. AFNOR Norm'Info.
- AFNOR (2018). Qualité du sol - Prestations de services relatives aux sites et sols pollués. Norme AFNOR NF X31-620. AFNOR, décembre 2018.
- Atteia, O., Höhener, P., 2012. Fast semi-analytical approach to approximate plumes of dissolved redox-reactive pollutants in heterogeneous aquifers. 1. *BTEX Advances in Water Resources* 46, 63–73.
- Augustin, T., Coolen, F., de Cooman, G., Troffaes, M. (Eds.), 2014. *Introduction to Imprecise Probabilities*, New York: Wiley.
- Baltic Urban Lab, 2018. Towards integrated and partnership-based planning of brownfield areas. www.balticurbanlab.eu
- Barnaud, C., Promburom, T., Trébuil, G., Bousquet, F., 2007. An evolving simulation/gaming process to facilitate adaptive watershed management in northern mountainous Thailand. *Simulation & Gaming*, 38(3), p. 398-420.
- Barnier, C., Palmier, C., Atteia, O., 2013. Validation of a multilevel sampling device to determine the vertical variability of chlorinated solvent in a contaminated aquifer. *Environ. Technol. (United Kingdom)* 34. doi:10.1080/09593330.2012.758665
- Baudrit, C., Dubois, D. and Perrot, N., 2008. Representing parametric probabilistic models tainted with imprecision. *Fuzzy sets and systems*, 159, 1913–1928.
- Baudrit, C., Guyonnet, D., Dubois, D., 2007. Joint propagation of variability and partial ignorance in a groundwater risk assessment. *Journal of Contaminant Hydrology*, 93: 72-84.
- Baudrit, C., Dubois, D., Guyonnet, D., 2006. Joint propagation and exploitation of probabilistic and possibilistic information in risk assessment models. *IEEE Transactions on Fuzzy Systems*, vol.14, No.5, pp.593-608.
- Baudrit, C., Guyonnet, D., Dubois, D., 2005. Post-processing the hybrid approach for addressing uncertainty in risk assessments. *Journal of Environmental Engineering*, Volume 131, Issue 12, pp. 1750-1754.
- Baudrit, C., 2005. Représentation et propagation de connaissances imprécises et incertaines : application à l'évaluation des risques liés aux sites et aux sols pollués. Thèse de doctorat de l'Université de Toulouse III – Paul Sabatier.
- Baustert, P., Othoniel, B., Rugani, B., Leopold, U., 2018. Uncertainty analysis in integrated environmental models for ecosystem service assessments: Frameworks, challenges and gaps. *Ecosystem Services*, 33, 110-123.
- Bayer, P., Huggenberger, P., Renard, P., Comunian, A., 2011. Three-dimensional high resolution fluvio-glacial aquifer analog: Part 1: Field study. *J. Hydrol.* 405, 1–9. doi:10.1016/j.jhydrol.2011.03.038
- Bayes, T., 1763. An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London*, 53, p.370-418.

- Becu, N., Amalric, M., Anselme, B., Beck, E., Bertin, X., Delay, E., Long, N., Marilleau, N. et al., 2017. Participatory simulation to foster social learning on coastal flooding. *Environmental Modelling & Software*, 98, 1-11.
- Benoit Y., de Fouquet C., Fricaudet B., Carpentier C., Gourry J.-C. LOQUAS D23. Guide méthodologique & évaluation technico-économique. Rapport R2011-002-CFOU, octobre 2010, 61p. Ecole des Mines de Paris, Fontainebleau.
- Chastanet, J., Quintard, M., Côme, J.-M., Ababou, R., Esrael, D., Marcoux, M., Tribouillard, N., 2019. Evaluation de la sensibilité des modèles de prédiction d'épuisement et de remédiation des sources de pollution par des composés organiques : approches déterministes et probabilistes pour sécuriser les choix de gestion. Synthèse opérationnelle. Rapport ADEME - Projet ESPER. Collection expertise. Mai 2019. 29 pp.
- Chateauneuf, A., Cohen, M., 2009. Cardinal extensions of the EU model based on Choquet integral. Dans : D. Bouyssou, D. Dubois, M. Pirlot, and H. Prade (Eds). *Decision-making process: concepts and methods*. Chapter 3. London: ISTE & Wiley, 401–433.
- Chilès, J.-P., Delfiner, P., 1999. *Geostatistics: Modeling Spatial Uncertainty*. Wiley, New York.
- Chutia, R., Datta, D., 2017. Probability-credibility health risk assessment under uncertain environment. *Stochastic Environmental Research and Risk Assessment*, 31, 2, 449-460.
- Côme, B., Guyonnet, D., Magnouni, S., 1997. Apports potentiels de la théorie des possibilités à l'évaluation des risques environnementaux. *Déchets Sciences et Techniques*, n°7, 40-46.
- Côme, J.-M., Haessler, F., Kaskassian, S., Quintard, M., Razakarisoa, O., Ropars, Schaeffer, G., 2006. MACAOH : Caractérisation dans les aquifères d'une source constituée d'organo-chlorés aliphatiques. ADEME.
- Cori, A., 2005. Utilisation de données environnementales pour la cartographie du NO₂ sur la région Rouennaise. Rapport de stage d'option, 1ère partie. Ecole des Mines de Paris & AirNormand/Remappa.
- Cover, T. M., Thomas, J. A., 2006. *Elements of information theory*. 2nd Edition, Wiley.
- Dagan, G., 1990. Transport in heterogeneous porous formations: Spatial moments, ergodicity, and effective dispersion. *Water Resour. Res.*, 26(6), 1281-1290.
- Derycke, V., Coftier, A., Zornig, C., Leprond, H., Scamps, M., Gilbert, D., 2018. Environmental assessments on schools located on or near former industrial facilities: Feedback on attenuation factors for the prediction of indoor air quality. *Science of the Total Environment*, 626, 754-761.
- DGPR, 2017. *Méthodologie nationale de gestion des sites et sols pollués*. Direction générale de la Prévention des Risques. Bureau du Sol et du Sous-Sol.
- Donati, M., de Fouquet, C., 2018. How to describe organic contamination in soils: A model comparison for chlorinated solvent concentrations at industrial site scale. *Science of The Total Environment*. 633, p. 1480–1495.
- Dubearnes, B., Schmitt, J., Begassat, P., 2005. *La Representativite Des Echantillons D'Eau Prelevés En Forages De Controle : Synthèse Et Conclusions*, Rapport ADEME.
- Dubois, D., Prade, H., 1988. *Possibility theory*. New York Plenum Press.
- Dubois, D., Guyonnet, D., 2011. Risk-informed decision-making in the presence of epistemic uncertainty. *International Journal of General Systems*, 40(2), 145-167.
- Dupouy, G., Tacnet, J.M., Bourrier, F., Berger, F., Crimier, N., Mekhnacha, K., Memier, M., Moulet-Vargas, E., 2017. Propagating spatial and thematic uncertainties in mountain natural hazard assessment process. *Risk, Reliability and Safety: Innovating Theory and Practice*. Walls, L., Revie, M., Bedford, T. (Editors), 261-268.
- Ellsberg, D., 1961. Risk, ambiguity and the Savage axioms. *Q. J. Econ.* Vol. 75, p. 643–669.

- Favéreaux, S., Belbeze, S., Balon, P., 2019. Essai collaboratif sur l'échantillonnage de sols pollués – Exploitation des résultats et recommandations. Rapport final. BRGM/RP- 69237-FR, 113 p.
- Ferson, S., Hajagos, J.G., 2004. Arithmetic with uncertain numbers: rigorous and (often) best possible answers. *Reliability Engineering & System Safety*, 85(1-3), 135-152.
- Ferson, S., Ginzburg, L.R., 1996. Different methods are needed to propagate ignorance and variability. *Reliability Engineering and Systems Safety*, 54, 133-144.
- Ferson, S., 1996. What Monte Carlo methods cannot do. *Human and Environmental Risk Assessment*, 2, 990-1007.
- Flage, R., Dubois, D., Aven, T., 2016. Combined analysis of unique and repetitive events in quantitative risk assessment. *International Journal of Approximate Reasoning*, Vol. 70, p. 68-78.
- de Finetti, B. 1937. La prévision : ses lois logiques, ses sources subjectives. *Ann. Inst. Poincaré*, 7, p. 1–68.
- de Fouquet, C., 2016. Can measurement errors be characterized from replicates? Dans : *GEOSTATISTICS Valencia 2016*. Book Series: Quantitative Geology and Geostatistics 19: 45-57.
- Gelhar, L.W., Welty, C., Rehfeldt, K.R., 1992. A critical review of data on field - scale dispersion in aquifers. *Water Resour. Res.*, 29(6), 1867-1869.
- Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 2003. *Bayesian data analysis*, Chapman CRC.
- Ghestem, J.-P., Lachenal, J., 2008. Incertitude sur l'échantillonnage et le prélèvement d'eaux : synthèse bibliographique, AQUAREF. Rapport BRGM/RP-56885-FR, 56 pp.
- Ghestem, J.-P., 2011. Estimation des incertitudes de mesure dans les programmes de surveillance DCE : situation actuelle et impact des exigences européennes. Rapport AQUAREF – BRGM/RP-60611-FR, 49 pp.
- Gobet, E., 2013. *Méthodes de Monte-Carlo et processus stochastiques - du linéaire au non linéaire*. Éditions de l'École polytechnique.
- Goode, D.J., Konikow, L.F., 1990. Apparent dispersion in transient groundwater flow. *Water Resour. Res.*, 26(10), 2339-2351.
- Gutierrez-Neri, M., Ham, P.A.S., Schotting, R.J., Lerner, D.N., 2009. Analytical modelling of fringe and core biodegradation in groundwater plumes. *J. Contam. Hydrol.* 107, 1–9.
- Guyonnet, D., Leroux, R., Dubois, D., Fargier, H., 2015. Risk-R : un outil « open source » de propagation des incertitudes en évaluation des risques. Rapport BRGM/RP-64871-FR.
- Guyonnet, D., 2008. MISP_v1. Un modèle analytique pour estimer l'impact de sources de pollution sur les eaux souterraines. Guide d'utilisation. Rapport BRGM/RP-56152-FR.
- Guyonnet, D., Neville, C., 2004. Dimensionless analysis of two analytical solutions for 3-D solute transport in groundwater. *J. Contam. Hydrol.* 75, 141–153.
- Guyonnet, D., Bourguine, B., Dubois, D., Fargier, H., Côme, B., Chilès, J.-P., 2003. Hybrid approach for addressing uncertainty in risk assessments. *Journal of Environmental Engineering* 129, 68-78.
- Hacking, I., 1975. *The Emergence of Probability*. Cambridge University Press, Cambridge, UK.
- Harpet, C., Guyonnet, D., Dor, F., 2005. Risk perception and communication on risks: a field survey. Conférence ConSoil 2005, 3 - 7 Octobre 2005, Bordeaux, France.
- Harvey, C., Gorelick, S.M., 2000. Rate-limited mass transfer or macrodispersion: Which dominates plume evolution at the Macrodispersion Experiment (MADE) site? *Water Resour. Res.* 36(3), 637-650.
- Hazebrouck, B., Baumont, G., Legout, C., 2008. Supports de communication pour l'implication des populations dans l'évaluation et la gestion d'un site ou sol pollués. Rapport INERIS-IRSN pour l'ADEME, Mars 2008.

- Hunkeler, D., Höhener, P., Atteia, O., 2010. Comments on “Analytical modelling of fringe and core biodegradation in groundwater plumes.” by Gutierrez-Neri et al. in *J. Contam. Hydrol.* 107: 1–9. *Journal of Contaminant Hydrology*, 117, 1–6.
- Hurwicz, L., 1951. Optimality criteria for decision making under ignorance. Cowles Commission discussion paper, Statistics No. 370.
- Igos, E., Benetto, E., Meyer, R., Baustert, P., Othoniel, B., 2018. How to treat uncertainties in life cycle assessment studies? *International Journal of Life Cycle Assessment*, 24, 794-807.
- INERIS, 2018. Contamination Des Eaux Souterraines Dans Le Contexte Des Icpe Et Des Sites Pollues: Comparaison D'Outils Et De Protocoles D'Echantillonnage. Ineris-Drc-17-164249-11465a 180.
- INERIS-IRSN, 2008. Implication des populations dans l'évaluation et la gestion des sites et sols pollués. Etat de l'art. En partenariat avec la Cire Ile de France. B. Hazebrouck, G. Baumont, C. Legout. INERIS DRC-07-61078-13306B. Mars 2008.
- Islam, M.S., Nepal, M.P., Skitmore, M., Attarzadeh, M., 2017. Current research trends and application areas of fuzzy and hybrid methods to the risk assessment of construction projects. *Advanced Engineering Informatics*, 33, 112-131.
- ISO, 2017a. Norme ISO/WD 18400-102 Qualité du sol - Échantillonnage : choix et application des techniques d'échantillonnage
- ISO, 2017b. Norme ISO/WD 18400-104 Qualité du sol - Échantillonnage (2019) : stratégie.
- ISO, 2016. « Qualité du sol – Dosage des hydrocarbures aromatiques et halogénés volatils et de certains éthers par chromatographie en phase gazeuse - Méthode par espace de tête statique.
- Jeannée, N., Demougeot-Renard, H., 2016. Méthodes d'estimation et d'optimisation des quantités de terres polluées à traiter. Etude Record n° 14-0515/1A – Rapport final.
- Jiashen, T., 2018. Uncertainty analysis of transmission line end-of-life failure model for bulk electric system reliability studies. *IEEE Transactions on Reliability*, 67, 3.
- Klenk, I.D., Grathwohl, P., 2002. transverse vertical dispersion in groundwater and the capillary fringe. *J. Contam. Hydrol.*
- Kulkarni, P., Newell, C., Krebs, C., McHugh, T., Sanford, B., 2015. Methods for Minimization and Management of Variability in Long-Term Groundwater Monitoring Results.
- Lee, Y.W., Dahab, M.F., Bogardi, I., 1994. Fuzzy decision-making in groundwater nitrate risk management. *Water Resources Bulletin*, 30(1), 135-148.
- Lindley, D.V., 1971. Making decisions. Wiley-Interscience, London.
- Lion, F., Colombano, S., Aubert, N., Boissard, G., 2016. Définir une stratégie de dépollution : approche basée sur la masse de polluant et la capacité de relargage d'une pollution. Rapport final BRGM/RP-64350-FR.
- Lloyd, S.M., Ries, R., 2007. Characterising, propagating and analysing uncertainty in life-cycle assessment, a survey of quantitative approaches. *Journal of Industrial Ecology*, 11, 161–1790.
- LNE, 2012. Guide EURACHEM / CITAC : Quantifier l'incertitude des mesures analytiques - 3e édition.
- Loquin, K., Dubois, D., 2012. A fuzzy interval analysis approach to kriging with ill-known variogram and data. *Soft Computing*, 16, 767-784.
- Molleron, 2019. BIM: the Dunkerque refinery cleaned up using an innovative digital tool. AQUACONSOIL - Sustainable Use and Management of Soil, Sediment and Water Resources, 15th International Conference, 20–24 May 2019, Antwerp, Belgium.
- Moreau, P., Ghestem, J.P., Lalère, B., Lepot, B., Charpentier, R., Querio, L., Guarini, P., Rosin, C., 2019. Surveillance réglementaire des eaux environnementales et sanitaires : évaluation de la dispersion

analytique à partir des données d'essais interlaboratoires – Rapport AQUAREF– BRGM/RP-68705-FR - 60p.

- von Neumann, J., Morgenstern, O., 1953. *Theory of Games and Economic Behavior*. Princeton, NJ. Princeton University Press.
- Palmer, T.N., Shutts, G.J., Hagedorn, R., Doblas-Reyes, F.J., Jung T., Leutbecher, M. 2005. Representing Model Uncertainty in Weather and Climate Prediction. *Annual Review of Earth and Planetary Sciences*, 33,163–193.
- Puls, R.W., Paul, C.J., 1997. Multi-layer sampling in conventional monitoring wells for improved estimation of vertical contaminant distributions and mass. *J. Contam. Hydrol.* 25, 85–111.
- Quiot, F., Guerin, V., 2018. Surveillance de la qualité des eaux souterraines appliquée aux ICPE et sites pollués.
- Ramsey, F.P., 1980. Truth and probability. Dans: J. Kyburg and H.E. Smokler (Eds). *Studies in Subjective Probability*, p. 23–52. Krieger Pub. Co, Huntington, N.Y. Originellement publié en 1926.
- Rebiasz, B., Gawel, B., Skalna, I., 2017. Joint treatment of imprecision and randomness in the appraisal of the effectiveness and risk of investment projects. *Information Systems Architecture and Technology - ISAT 2016, PT IV*. Wilimowska, Z., Borzemski, L., Grzech, A., Swiatek, J. (Editors). Book Series: *Advances in Intelligent Systems and Computing*, 524, 21-31.
- Rivett, M.O., Feenstra, S., Cherry, J.A., 2001. A controlled field experiment on groundwater contamination by a multicomponent DNAPL: Creation of the emplaced-source and overview of dissolved plume development. *J. Contam. Hydrol.* 49(1-2), 111-149.
- Rohmer, J., Manceau, J.-C., Guyonnet, D., Boulahya, F., 2017. HYRISK: Hybrid methods for addressing uncertainty in risk assessments. *Projet CRAN.R*; cran.r-project.org.
- Rollin, C., Quiot, F., Masselot, G., 2008. Surveillance des eaux souterraines dans le contexte des sites pollués.
- Rivoirard, J., Demange, C., Freulon, X., Lécureuil, A., Bellot, N., 2013. A top-cut model for deposits with heavy-tailed grade distribution, *Math Geosc.* 45(8), p. 967 - 982.
- Sauvaget, B., 2019. Constitution de référentiels géochimiques locaux pour les sols et proches sous-sols urbains : de la base de données à l'interprétation géostatistique. Thèse de doctorat de l'Université PSL.
- Savage, L.J., 1954. *The Foundations of Statistics*. Wiley, 2nd edition, Dover Publications Inc., New York, 1972.
- Shafer, G., 1978. Non-additive probabilities in the work of Bernoulli and Lambert. *Archive for History of Exact Sciences*, 19, p. 309–370.
- Shafer, G., 1976. *A Mathematical Theory of Evidence*. Princeton University Press.
- Tacnet, J.M., Carlados, S., Dezert, J., Han, D.Q., Batton-Hubert, M., 2017. New MCDM methods under uncertainty applied to integrated natural risks management. *IEEE International Conference on Computational Intelligence and virtual environments for measurement systems and applications (CIVEMSA)*, 193-198.
- Tacnet, J.M., Dezert, J., Curt, C., Batton-Hubert, M., Chojnacki, E., 2014. How to manage natural risks in mountain areas in a context of imperfect information? New frameworks and paradigms for expert assessments and decision-making. *Environmental Systems Decisions*, 34, 288–311.
- Tang, P.Z., Liu, J.Z., Lu, H.W., Wang, Z., He, L., 2017. Information-based network environ analysis for ecological risk assessment of heavy metals in soils. *Ecological Modelling*, 344, 17-28.
- Thiéry, D., 2014. Modélisation 3D du transport de masse avec le code de calcul MARTHE version 7.3. BRGM/RP 63869.

- Toscani, N., Grassi, F., Spadacini, G., Pignari, S.A., 2018. A Possibilistic approach for the prediction of the risk of interference between power and signal lines onboard satellites. *Mathematical Problems in Engineering*. Article N° 7921048, doi: 10.1155/2018/7921048.
- Traverse, S., Schäfer, G., Chastanet, J., Hulot, C., Perronnet, K., Collignan, B., Cotel, S., Marcoux, M., Côme, J.-M., Correa, J, Gay, G., Quintard, M, Pepin, L., 2013. Projet FLUXOBAT. Evaluation des transferts de COV du sol vers l'air intérieur et extérieur. Guide méthodologique. 257 pp.
- Walley, P., 1991. *Statistical reasoning with imprecise probabilities*. Chapman and Hall.
- Wanke, P., Azad, M.A.K., 2018. Efficiency in Asian railways: a comparison between data envelopment analysis approaches. *Transportation Planning and Technology*, 41, 6, 573-599.
- Zadeh, L., 1978. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1, 3-28.

INDEX DES TABLEAUX ET FIGURES

FIGURES

Figure 1 - Cycle de la donnée environnementale en contexte de gestion de sites et sols pollués....	8
Figure 2 - Trois situations de comparaison d'intervalles [min, max] de risque avec un seuil de risque précis	10
Figure 3 - Illustration de la différence entre un calcul d'intervalle (a) et la somme de deux distributions de probabilité uniformes (b)	11
Figure 4 - Calcul effectué à partir des familles de distributions de probabilité définies par l'information disponible.....	12
Figure 5 - Cycle de la donnée environnementale en contexte de gestion de sites et sols pollués.	14
Figure 6 – Vue d'une zone investiguée lors de l'essai d'inter-comparaison du BRGM (2018)	17
Figure 7 – Techniques de forage mises en œuvre lors de l'essai d'intercomparaison du BRGM (2018)	17
Figure 8 – Méthodes d'échantillonnage mises en œuvre lors de l'essai d'inter-comparaison du BRGM (2018)	18
Figure 9 – Variabilité spatio-temporelle des gaz du sol, mesures sur un quadruplet de piézaires (site atelier FLUXOBAT, Traverse et al., 2013)	21
Figure 10 – Influence du vent et de la pression sur les concentrations dans l'air intérieur (site atelier FLUXOBAT, Traverse et al., 2013).....	21
Figure 11 - Représentation schématique de l'incertitude de laboratoire (Ghestem et Lachenal, 2008).....	24
Figure 12 – Différents types d'incertitude associés aux mesures de laboratoire (Ghestem et Lachenal, 2008).....	25
Figure 13 - Histogramme des fréquences obtenu sur des mesures décennales d'hauteurs de pluie, pour la période 1992-2018 (a) et distribution des fréquences relatives cumulées (b).	28
Figure 14 - Calage automatique de la distribution des fréquences relatives à l'aide d'une distribution de probabilité log-normale (a) et calage automatique à l'aide d'une distribution de probabilité Gamma.	28
Figure 15 - Calage de mesures de la concentration en arsenic dans un terril minier à l'aide d'une distribution de probabilité normale (a) et délimitation de ces mesures à l'aide d'une famille paramétrique de distributions normales (b).	29
Figure 16 - Distribution de possibilité triangulaire correspondant à un jugement d'expert sur la valeur du facteur d'atténuation FA.	30
Figure 17 - Correspondance entre distribution de possibilité et famille de distributions de probabilités cumulées.....	30
Figure 18 - Organigramme mettant en relation différents types d'information et des modes de représentation de cette information (modifié de Dubois et Guyonnet, 2011).	32
Figure 19 - Illustration schématique de la propagation d'incertitude dans un modèle de risque à l'aide de la méthode dite Monte Carlo.....	34
Figure 20 - Illustration schématique de la propagation conjointe d'information aléatoire et incomplète/imprécise.....	35
Figure 21 - Exemple de schéma conceptuel illustrant les sources, voies de transfert, etc.	36
Figure 22 - Schéma conceptuel du modèle analytique MISP (Guyonnet, 2008)	37
Figure 23 - Exemple de modélisation numérique du transfert d'un polluant (le dichloroéthylène) dans les eaux souterraines à l'aide du modèle numérique Visual MODFLOW (Chastanet et al., 2019).....	37
Figure 24 - Comparaison d'un résultat de calcul de risque à un seuil de risque précis.	38
Figure 25 - Seuil d'acceptabilité flou (QDA = quotient de danger acceptable).	38
Figure 26 - Fond pédo-géochimique urbain flou.....	39
Figure 27 - Deux transects de concentrations (en haut) et leur variogramme respectif (en bas). ..	41
Figure 28 - Des données à la carte de la variabilité spatiale (projet LOQUAS, Benoit et al., 2010). ..	41
Figure 29 - Famille de distributions de probabilité et indice de confiance.	46

SIGLES ET ACRONYMES

ADEME	Agence de l'Environnement et de la Maîtrise de l'Énergie
BRGM	Bureau de Recherches Géologiques et Minières
IRIT	Institut de Recherche en Informatique de Toulouse
Mines ParisTech	Ecole des Mines de Paris

L'ADEME EN BREF

L'Agence de l'Environnement et de la Maîtrise de l'Energie (ADEME) participe à la mise en œuvre des politiques publiques dans les domaines de l'environnement, de l'énergie et du développement durable. Elle met ses capacités d'expertise et de conseil à disposition des entreprises, des collectivités locales, des pouvoirs publics et du grand public, afin de leur permettre de progresser dans leur démarche environnementale. L'Agence aide en outre au financement de projets, de la recherche à la mise en œuvre et ce, dans les domaines suivants : la gestion des déchets, la préservation des sols, l'efficacité énergétique et les énergies renouvelables, les économies de matières premières, la qualité de l'air, la lutte contre le bruit, la transition vers l'économie circulaire et la lutte contre le gaspillage alimentaire.

L'ADEME est un établissement public sous la tutelle conjointe du ministère de la Transition Ecologique et Solidaire et du ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation.

<https://www.ademe.fr/>

LES COLLECTIONS DE L'ADEME



ILS L'ONT FAIT

L'ADEME catalyseur : Les acteurs témoignent de leurs expériences et partagent leur savoir-faire.



EXPERTISES

L'ADEME expert : Elle rend compte des résultats de recherches, études et réalisations collectives menées sous un regard.



FAITS ET CHIFFRES

L'ADEME référent : Elle fournit des analyses objectives à partir d'indicateurs chiffrés régulièrement mis à jour.



CLÉS POUR AGIR

L'ADEME facilitateur : Elle élabore des guides pratiques pour aider les acteurs à mettre en œuvre leurs projets de façon méthodique et/ou en conformité avec la réglementation



HORIZONS

L'ADEME tournée vers l'avenir : Elle propose une vision prospective et réaliste des enjeux de la transition énergétique et écologique, pour un futur désirable à construire ensemble.



PRISE EN COMPTE DES INCERTITUDES DANS LA GESTION DES SITES ET SOLS POLLUÉS

Résumé Ce document est une synthèse relative à la gestion des incertitudes dans le domaine des sites et sols pollués (SSP). Le document, qui n'a aucune prétention d'exhaustivité compte tenu de l'étendue de la problématique, s'attache à identifier des sources d'incertitude aux différentes étapes du cycle de la donnée en contexte de SSP, à décrire certaines approches qui sont proposées pour traiter ces incertitudes et à proposer des pistes de recherche future.

Le cycle de la donnée considéré va de l'échantillonnage des milieux en contexte de SSP, à l'utilisation des informations collectées pour définir des choix de gestion, en passant par l'analyse, la représentation, l'interprétation, etc.

Il ressort notamment de cette synthèse que s'il n'existe pas de méthode « passe-partout » pour gérer les incertitudes de manière générique en contexte de sites et sols pollués, il est possible néanmoins de promouvoir une cohérence entre le choix des méthodes de gestion des incertitudes et la nature des informations dont on dispose dans la pratique. Cette recherche de cohérence devrait favoriser des choix de gestion plus fiables et robustes.

