

THESE – UNIVERSITE AIX-MARSEILLE

Ecole doctorale
ED 251 Sciences de l'environnement

Pour obtenir le grade de
Docteur de l'Université d'Aix-Marseille

Spécialité
Géosciences de l'environnement

Analyse d'incertitude en situation accidentelle.
Transfert de radionucléides dans l'environnement et
évaluation de l'exposition humaine par voie alimentaire.

Par
Mouhamadou Moustapha Sy

Thèse dirigée par Marie Simon-Cornu
Soutenance publique le 21 Mars 2016

Devant le jury d'examen composé de :

Pr. Patrice Bertail	Professeur, Université Paris X	Président du Jury
Pr. Bernard Bonin	Directeur de Recherche, CEA, Saclay	Rapporteur
Pr. Eric Parent	Professeur, AgroParistech, Paris	Rapporteur
Dr. Philippe Ciffroy	Ingénieur – Chercheur, EDF R&D, Chatou	Examineur
Dr. Franck Torre	Maître de conférences, Université Aix-Marseille	Examineur
Dr. Marie Simon-Cornu	Ingénieur-Chercheur, IRSN, Cadarache	Directrice de thèse

Remerciements

Je tiens tout d'abord à exprimer mes chaleureux remerciements à la région Provence-Alpes-Côte d'Azur (PACA) et à l'Institut de Radioprotection et de Sûreté Nucléaire (IRSN) pour avoir financé mes travaux de recherche. En particulier, je souhaite remercier le SERIS dirigé par Jacqueline Garnier-Laplace puis par François Paquet pour m'avoir accueilli durant ces trois années.

Je souhaite témoigner toute ma reconnaissance à Marie Simon-Cornu pour la qualité de son encadrement, de ses conseils, pour l'orientation donnée à ce travail. Un projet de thèse est non seulement l'occasion de développer ses connaissances scientifiques et ses compétences techniques mais surtout de grandir humainement. Durant ces trois années, j'ai énormément appris à ton contact. Ta confiance et ton soutien constants m'ont beaucoup apporté et m'ont permis de m'accrocher dans les moments les plus difficiles. Merci pour tout.

Je remercie chaleureusement Bernard Bonin et Éric Parent pour avoir accepté d'évaluer mon manuscrit de thèse en tant que rapporteur. Un grand merci pour la justesse de vos commentaires et la pertinence de vos suggestions. J'adresse mes sincères remerciements à Patrice Bertail, Philippe Ciffroy et Franck Torre, pour avoir accepté de faire partie de mon jury de thèse.

Mes remerciements les plus chaleureux vont à Sophie Ancelet, Claire Della Vedova, Catherine Keller, Jessica Tressou, Éric Chojnacki, Philippe Ciffroy et Marc-André Gonze pour avoir accepté de faire partie de mon comité de pilotage. Un grand merci pour votre disponibilité lors de chaque rencontre à Cadarache et pour la pertinence de vos commentaires. Vos conseils m'ont été particulièrement précieux dans mes différents moments de doute.

J'adresse mes sincères remerciements aux membres du projet COMET. En particulier, je souhaite exprimer ma reconnaissance à Astrid Liland, Almudena Real, Brenda Howard, Claire Wells, Ali Hosseini, Nick Beresford et Håvard Thørring pour les discussions particulièrement intéressantes sur les paramètres radioécologiques et la contamination de la chaîne alimentaire.

J'adresse toute ma reconnaissance aux ingénieurs de l'IRSN et du SERIS que j'ai côtoyés durant ces trois années. Mes sincères remerciements à Sidi Souvi pour toutes nos discussions sur la Mauritanie et le Sénégal. Je souhaite exprimer mes remerciements les plus chaleureux à Rodolphe Gilbin, Béatrice Gagnaire, Pascale Henner, Pierre Hurtevent, Laurent Garcia-Sanchez, Frédéric Alonzo, Arnaud Martin-Garin, Christelle Adam et Laureline Février. Pascale et Pierre, un grand merci pour cette « grande quête » qu'aura été l'interception sèche ; « Indiana Jones dans les « préfas » de l'IRSN » ça pète bien comme titre, non ? Mes sincères remerciements à Rodolphe et Laurent pour la relecture de mon manuscrit et pour vos différents conseils.

Un grand merci à Christelle Antonelli, Philippe Calmon, Philippe Renaud, Olivier Masson, Laetitia, Julien, Hugo Lepage, Sabrina, Franck, David, Frédérique Eyrolle-Boyer, Jackie Tav, Maud, Loïc et à tous les membres de l'ancien LERCM. Je souhaite témoigner toute ma reconnaissance à Christelle et à Philippe Calmon pour la relecture de mon manuscrit. Vos commentaires m'ont été très précieux.

J'adresse mes remerciements les plus chaleureux et toute ma gratitude aux membres de l'ex-LM2E. Ces trois années passées à vos côtés ont été très agréables. Merci pour tout votre soutien, vos conseils et surtout votre gentillesse durant ces années. J'ai particulièrement apprécié le temps passé avec vous tous. Véro, ah Véro, ta sollicitude, ton aide et ta présence ont été très précieuses durant ces années. Karine, merci pour ta gentillesse à mon égard et ton soutien. Tes multiples victoires aux (semi-) marathons suscitent encore toute mon admiration. Un grand merci pour ta relecture exhaustive de mon manuscrit. Merci beaucoup Valérie pour tes conseils en statistiques et sur les différentes questions liées à l'inférence bayésienne et aux analyses de sensibilité. Etre à proximité d'une mathématicienne de ton niveau a été d'une aide très précieuse. Séverine, merci pour ta gentillesse, ton aide sur la biomasse des plantes et surtout pour ta compassion. Claire et Fabien Vanni, merci du fond du cœur pour votre gentillesse. Claire, nos différents échanges sur les statistiques et l'analyse de données m'ont énormément servi. Mon « voisin » Christophe, merci pour ta sympathie et les bons moments passés les midis. Un grand merci pour ton aide avec SYMBIOSE. Patrick, merci beaucoup pour ta gentillesse et pour toutes nos discussions sur l'OM et sur le sport en général. Jean-Michel, merci pour tes belles cartes et ton « art géographique », merci pour ton humour et ta blague sur les « petits escargots ». Marc-André, un grand merci pour tes commentaires sur les processus radioécologiques et la pertinence de tes interrogations sur la modélisation bayésienne. J'ai énormément appris durant notre collaboration sur l'interception humide. Un grand merci pour ces moments de travail enrichissants que tu as rendus particulièrement agréables. Un grand merci aux athlètes émérites du LM2E, Ahn, Hugo, et Thomas. J'ai passé de très bons moments avec vous. Je vous souhaite bon courage pour la suite et merci pour les petits déjeuners improvisés. Merci Ahn pour la relecture de mon manuscrit et la pertinence de tes commentaires. Un grand merci à Névénick pour sa gentillesse et surtout pour la tasse Winnie l'ourson, cet immense héritage laissé entre mes mains.

J'adresse mes chaleureux remerciements à tous les stagiaires, doctorants et post-doctorants rencontrés durant ces trois années. Un grand merci à Ahn, Hugo, Thomas, JC, Benji, Névénick, Donovan, Mario, Sylvain, Maud, Fred Chaux, Léa Beaumelle, Loïc, Jackie, Florian, Kewin, Antoine, Yvan, Adrien, Adeline, Benoît, Guillaume, Amine et Rym, Laure, Cécile, Ingrid et Ahmedou pour tous les bons moments passés ensemble. Merci Florian, Kewin et Antoine pour les délires ensemble lors des journées des thèses 2013, 2014 et 2015.

Un grand merci à tous les « footeux » rencontrés à Cadarache durant ces trois années. En particulier, j'adresse mes remerciements les plus chaleureux aux Globetrotteurs (vainqueurs du tournoi

de Cadarache en 2014 et troisième l'édition suivante) : Ayoub, Clément, Eden, Fred Chau, Florent, Julien, Petit Lu, Ludo, Louis Dumas, Louis, Rémy, Samuel, Simon, Yohann. Merci également à tous les joueurs de l'équipe FLA de l'AS CEA championne en durant la saison 2014/2015.

Je souhaite exprimer toute ma gratitude à mes anciens collègues de l'INRA-Metarisk et de l'ANSES pour leur aide pour l'audition pour cette thèse et pour m'avoir offert l'opportunité de me jeter dans le bain de l'évaluation des risques alimentaires. Un grand merci à Sandrine, Amélie, Max, Philippe, Stéphan, Laura-Flore, Camille, Antoine, Jessica, Isabelle, Éric, Alain, Stéphane, Lydie, Juliette, Laurie, Jean-Marc et Clémence.

J'adresse mes sincères remerciements à mes amis pour leur soutien, les discussions sur tout et surtout n'importe quoi. Merci à Moctar et Mounissa, Badou et Lika, aux autres membres de l'Equipe06, Rassoul, Loss, Nafissa, Penda et à tous ceux que je n'ai pas cités. Merci d'être là. Mes remerciements vont également à Fabien Garic qui, au fil des années, est devenu un véritable ami.

Un grand merci à ma belle-famille, qui m'a témoigné beaucoup de gentillesse, son soutien et une aide constante toutes ces années. Merci à Mouhamadou Lamine Gaye pour ses précieux conseils sur la recherche scientifique et à Aminata Diop pour sa gentillesse et toutes ses prières. Merci à Zeyna, Moustapha et Cheikh pour tous les instants de rigolade qu'ils génèrent.

Merci à mes frères et sœurs. Un grand merci à Madiye Dimé et à mon frère Pape Demba pour m'avoir montré, par l'exemple, le chemin à suivre. Merci à Bineta Gaye, Dialé, Amy Sy, Amy Kandji, Kinta Correa, à sa mère Marie, et à tous mes petits neveux pour votre amour. Je souhaite témoigner toute ma reconnaissance à mes parents sans qui bien des choses dans ma vie n'auraient pas de sens. Vous êtes une source intarissable de motivation pour moi. Merci de me pousser tous les jours et de me témoigner votre amour.

Mes derniers mots vont à ma merveilleuse femme, Amina. Il est difficile de poser des mots sur tout ce que tu m'apportes. Merci de partager ma vie. Merci pour ton aide. Merci pour ton soutien. Merci pour ton amour. Merci pour tout.

A mes Parents,

A ma femme,

A ma famille,

A mes amis,

A tous les êtres chers qui nous ont quittés.

Résumé

Les accidents nucléaires majeurs de Tchernobyl (Avril 1986) et de Fukushima (Mars 2011) ont généré une importante contamination environnementale avec des quantités conséquentes de substances radioactives libérées dans l'atmosphère. L'évaluation des risques, en situation d'urgence nucléaire, est entachée d'incertitudes sur le transfert de substances radioactives dans les écosystèmes terrestres et vers l'homme à travers la chaîne alimentaire, pouvant altérer la fiabilité des décisions. L'ampleur des répercussions des accidents de Tchernobyl et de Fukushima a mis en évidence la difficulté de gérer les conséquences post-accidentelles de tels événements et surtout d'appréhender les différentes sources d'incertitudes dans les prises de décision. L'objectif de cette thèse est de développer une méthodologie pour la prise en compte des incertitudes dans les modèles d'évaluation de risque environnemental et alimentaire en vue d'améliorer les outils d'aide à la décision en situation accidentelle. Dans cette optique, différents modèles bayésiens hiérarchiques visant à saisir, dans un cadre unique de modélisation, l'*incertitude* et la *variabilité* sur des paramètres radioécologiques d'intérêt en situation post-accidentelle (facteurs d'interception sèche et d'interception humide et constante de décroissance biomécanique) ont été développés. Les paramètres de ces modèles ont été estimés par inférence bayésienne sur des données collectées à partir d'une revue étendue de la littérature. L'influence sur les modèles d'évaluation de risque, de l'*incertitude* et de la *variabilité* autour de ces paramètres radioécologiques a ensuite été évaluée par simulations probabilistes et analyses de sensibilité, appliquées à deux études de cas : un accident hypothétique simulant un dépôt standardisé de radionucléides et l'accident survenu à la centrale nucléaire de Fukushima. Les travaux réalisés dans cette thèse contribuent à améliorer la connaissance autour des processus clés de transfert de radionucléides dans l'environnement et la paramétrisation des modèles radioécologiques d'évaluation de risque en ligne avec les axes de recherche prioritaires définis par la communauté scientifique en radioécologie.

Mots-clés : *Incertitude/Variabilité, Modélisation bayésienne, Accidents nucléaires majeurs, Radioécologie, Evaluation du risque alimentaire.*

Abstract

Major nuclear accidents of Chernobyl (April, 1986) and Fukushima (March, 2011) have led to a huge environmental contamination with important amounts of radionuclides released in the atmosphere. Risk assessment, in case of nuclear emergency, is confronted to uncertainties on the transfer of radioactive substances in terrestrial ecosystems and to human population through the food chain, which could affect the reliability of decisions. The extent of the repercussions of Chernobyl and Fukushima accidents highlighted the difficulty of managing the consequences of such disasters and specifically to accommodate the different sources of uncertainty within decision-making processes. The objective of this research project is to develop a methodology to take into account uncertainties within environmental and food risk assessment models in order to improve decision support tools used for accidental situations. In this regard, different hierarchical Bayesian models aiming at capturing, within a unique modelling framework, *uncertainty* and *variability* about radioecological parameters of great importance for accidental situation (dry and wet interception fractions and weathering loss parameter) were developed. Model parameters were estimated by Bayesian inference applied on databases obtained by an extended literature review. The impact on the risk assessment models of *uncertainty* and *variability* about these radioecological parameters was then assessed by stochastic simulations and sensitivity analyses applied on two case-studies: a hypothetical accident simulating a standardized deposition of radionuclides and the accident of Fukushima nuclear power plant. The works developed in this project contribute to enhance knowledge on key processes governing environmental transfer of radionuclides and to improve the parameterization of the radioecological risk assessment models with respect to the research lines outlined by the scientific community in radioecology.

Keywords: *Uncertainty/Variability, Bayesian modelling, Major nuclear accidents, Radioecology, Food risk assessment.*

Liste des acronymes et abréviations

Ag = Argent

AIEA (IAEA en anglais) = Agence Internationale de l'Energie Atomique

AMORAD = Amélioration des MODèles de prévision de la dispersion et d'évaluation de l'impact des RADionucléides au sein de l'environnement (projet ANR-11-RSNR-0002)

ANR = Agence Nationale de la Recherche

ASN = Autorité de Sûreté Nucléaire

ASTRAL = ASsistance Technique en Radioprotection post-Accidentelle

Ba = Baryum

BIOMOVS = BIOsperic MOdel Validation Study

BPA = Basic Probability Assignment

CEA = Commissariat à l'Energie Atomique

CIPR (ICRP en anglais) = Commission Internationale de Protection Radiologique

CI = Confidence/Credible Interval

CODIRPA = Comité Directeur pour la gestion de la phase Post-Accidentelle

COMET = COordination and iMplementation of a pan-Europe instrument for radioecology

Cs = Cesium

DAG = Directed Acyclic Graph

DIC = Deviance Information Criterion

EDOFU = External DOsimetry in FUKushima region

EFAST (voir FAST) = Extended FAST

EFSA = European Food Security Authority

ERA = European Radioecology Alliance

FAO = Food and Agriculture Organization

FAST = Fourier Amplitude Sensitivity Testing

FORM = First Order Reliability Method

ICRP = voir CIPR

INES = International Nuclear Events Scale

IPCS (PISC en français) = International Programme on Chemical Safety

IRSN = Institut de Radioprotection et de Sûreté Nucléaire

IUR = International Union of Radioecology

LAI = Leaf Area Index

MC1D / 2D = Monte Carlo une dimension et deux dimensions

MCMC = Monte Carlo par Chaînes de Markov

NMA = Niveaux Maximaux Admissibles

NRC = National Research Council

NRPA = Norwegian Radiation Protection Authority

OCDE = Organisation de Coopération et de Développement Economiques

OECD = voir OCDE

OMS = voir WHO

PICS = voir IPCS

PE = Périmètre d'Eloignement

SCCS = Scientific Committee on Consumer Safety

SCENIHR = Scientific Committee on Emerging and Newly Health Risks

SCHER = Scientific Committee on Health and Environmental Risks

SGDSN = Secrétariat Général de la Défense et de la Sécurité Nationale

SFEN = Société Française de l'Energie Nucléaire

SORM = Second Order Reliability Method

SRA = Strategic Research Agenda

Sr = Strontium

STAR = STrategy for Allied Radioecology (network of excellence NoE)

TECDOC = Technical Document

Te = Technetium

TRS = Technical Report Series

UNSCEAR = United Nations Scientific Committee on the Effects of Atomic Radiation

WHO (OMS en français) = World Health Organization

ZPP = Zone de Protection des Populations

ZST = Zone de Surveillance renforcée du Territoire

Glossaire

Becquerel = Unité de mesure de la radioactivité (Bq)

Dose absorbée = Quantité d'énergie absorbée par la matière rapportée à la masse de matière.

Dose efficace = Somme des doses reçues par les différents organes du corps, ramenée au corps entier intégrant les différences entre type de rayonnements et prenant en compte la « sensibilité » de chaque organe.

Dose équivalente = Dose moyenne reçue par un organe intégrant les différences entre types de rayonnement

Gray = Unité de mesure de la dose absorbée (Gy).

Perte aux champs (Field-Loss ou quelques fois « weathering » en anglais) = Décroissance biomécanique.

Radionucléide = Isotope radioactif, appelé aussi parfois radio-isotope, d'un élément.

Resuspension = Phénomène de remise en suspension des particules déposées.

Sievert = Unité de mesures des doses efficace et équivalente (Sv).

Terme Source = Définit la quantité et la chronique des substances radioactives rejetées en fonctionnement normal, lors d'un incident ou d'un accident.

Nomenclature

EE_i	= L'ensemble des effets élémentaires du paramètre X_i (méthode de Morris).
$\eta(.,.)$	= Fonction intervenant dans la probabilité d'acceptation.
\mathcal{F}	= σ -algèbre.
F	= Facteurs de transfert généralement utilisés en radioécologie pour quantifier le passage de substances radioactives d'un compartiment à l'autre de l'environnement et donc de la chaîne alimentaire.
F_X	= Fonction de répartition de la variabilité aléatoire X .
$[E, \overline{F}]$	= p-box encadrant une probabilité imprécise.
$Bel(.)$	= Mesure de crédibilité définie dans le cadre de la théorie de Dempster-Shafer.
$g(.)$	= Modèle de type boîte noire (formalisme adopté dans ce manuscrit pour la notion de modèle).
$h(. .)$	= Loi instrumentale conditionnelle (Algorithme de Metropolis Hastings par acceptation-rejet).
$K(. .,.)$	= Matrice de transition des chaînes de Markov.
Kd	= Coefficient de distribution dans le sol.
λ	= Constante de décroissance (biomécanique λ^W ou λ^{WL} ou, radioactive λ^r , métabolique λ_{PR}^E et λ_{PL}^E , ou de migration dans le sol λ^{SMig} et λ^{Mig} etc.).
$\mu_X, \mathbb{E}(X)$	= Espérance mathématique de la variable aléatoire X .
$N(.)$	= Mesure de nécessité définie dans le cadre de la théorie des possibilités.
Ω	= Univers des possibles contenant l'ensemble des événements aléatoires.

$P(.)$	= Mesure de probabilité.
$[P, \overline{P}]$	= Probabilités basse et haute encadrant la probabilité imprécise.
$\pi(.)$	= Distribution de possibilité.
$\Pi(.)$	= Mesure de possibilité définie dans le cadre de la théorie des possibilités.
$Pl(.)$	= Mesure de plausibilité définie dans le cadre de la théorie de Dempster-Shafer.
$\rho(.,.)$	= Coefficient de corrélation linéaire de Pearson.
ρ	= Densité du sol.
R_X	= Distribution des rangs des réalisations de la variable aléatoire X .
SI_i	= Indice de sensibilité de premier ordre du paramètre X_i .
θ	= L'ensemble des paramètres dans un modèle paramétrique.
Θ	= Espace des paramètres.
TSI_i	= Indice de sensibilité total du paramètre $X_i : i = 1, \dots, p$.
$Var(X)$	= Variance de la variable aléatoire X .
X_1, \dots, X_p	= L'ensemble des p paramètres incertains et/ou variables des modèles $g(.)$.
Y	= Sortie des modèles de type boîte noire $g(.)$.
Z	= Couche des variables latentes dans les modèles hiérarchiques.

Valorisation Scientifique

Articles Acceptés

- Simon-Cornu M., Beaugelin-Seiller K., Boyer P., Calmon P., Garcia-Sanchez L., Mourlon C., Nicoulaud V., **Sy M. M.**, Gonze M-A., 2015. Evaluating variability and uncertainty in radiological impact assessment with the probabilistic database of SYMBIOSE. *Journal of Environmental Radioactivity*. Volume 139, pp 91-102.
- **Sy M.M.**, Ancelet S., Henner P., Hurtevent P., Simon-Cornu M., 2015. Foliar interception of radionuclides in dry conditions: A meta-analysis using a Bayesian modeling approach. *Journal of Environmental Radioactivity*. Volume 147, pp 63-75. Cf. [Chapitre 6](#).
- **Sy M.M.**, Gonze M.A., Nicoulaud-Gouin V., Métivier J.M., Simon-Cornu M., 2016. Uncertainty analysis in post-accidental risk assessment models: an application to the Fukushima accident. *Annals of Nuclear Energy*. *Sous presse*. Cf. [Chapitre 9](#).

Articles Soumis

- **Sy M.M.**, Hosseini, A., Liland, A., Simon-Cornu, M., Thorrington, H., (Soumis le 28 Janvier 2016). Uncertainty and variability in food chain exposure by radioactive fallouts. What can Hierarchical Bayesian modelling bring to radioecology community *Journal of Environmental Radioactivity*. Cf. [Chapitre 8](#).
- Gonze M-A., **Sy M.M.**, (Soumis le 22 Janvier 2016). Interception of Wet Deposited Atmospheric Pollutants by Herbaceous Vegetation: Data Review and Modelling. *Science of the Total Environment*. Cf. [Chapitre 7](#).

Communications Scientifiques

Communications Orales :

- **Sy M. M.**, Simon-Cornu M. (2014). Risk assessment of radioactive cesium *via* consumption of leafy vegetables after an environmental contamination (*e.g.* Fukushima). *Society for Risk Analysis Annual Meeting (SRA - 2014)*. Denver (USA). 7 – 11 December 2014.
- **Sy M. M.**, Gonze M-A., Simon-Cornu M. (2015). Uncertainty analysis in post-accidental risk assessment models: An application to the Fukushima accident. *The 7th European Review Meeting on Severe Accident Research (ERMSAR - 2015)*. Marseille (France). 24 - 26 March 2015

- **Sy M. M.**, Simon-Cornu M. (2015). Incertitude et variabilité en situation accidentelle : Exposition humaine par voie alimentaire après l'accident de Fukushima. *22ème Congrès annuel de l'Ecole Doctorale Sciences de L'Environnement – ED 251*. Cassis (France). 21 - 22 April 2015.

Communications Affichées :

- **Sy M.M.**, Simon-Cornu M. (2013). Quantification et propagation des incertitudes par inférence statistique et simulation de Monte Carlo du second ordre. *Journée des thèses de l'IRSN*. Presqu'île de Giens (France). 14 – 17 Octobre 2013.
- **Sy M.M.**, Henner P., Hurtevent P., Simon-Cornu M. (2014). Interception de radionucléides par les légumes-feuilles en temps sec : Méta-analyse par une approche de modélisation bayésienne. *21ème Congrès annuel de l'Ecole Doctorale Sciences de L'Environnement – ED 251*. Aix-en-Provence (France). 15 - 16 April, 2014.
- Métivier J.M., Gonze M.A., Murlon Ch., **Sy M.M.**, Nicoulaud V., Simon-Cornu M. (2015). How is landuse from remote sensing helpful for terrestrial radioecological expertise? Application to the Fukushima nuclear accident (Japan). *International Geoscience and Remote Sensing Symposium (IGARSS 2015)*. Milan, Italy. 26 – 31 July, 2015.

Encadrement de stage

Co-encadrement de **Rym Chérif-Allouche**, Stage de Master 2 (Janvier – Juillet, 2015) pour l'obtention du Diplôme de Master en Sécurité Sanitaire des Aliments d'Aix-Marseille Université.

Sujet : Etude du transfert des radionucléides vers les produits animaux suite à un accident nucléaire.

Table des matières

Remerciements	iii
Résumé	viii
Abstract.....	ix
Liste des acronymes et abréviations	x
Glossaire	xiii
Nomenclature.....	xiv
Valorisation Scientifique	xvi
Table des matières	xviii
Liste des Figures	xxiii
Liste des Tableaux	xxv
1. Introduction	1
1.1 Contexte	1
1.2 Problématique	3
1.3 Objectifs et démarche méthodologique.....	4
1.4 Organisation du document	5
Partie I : Etude bibliographique.....	7
2. Radioactivité et risque nucléaire	9
2.1 La radioactivité : un peu d’histoire	9
2.2 Les sources de radioactivité	10
2.2.1 Sources naturelles	10
2.2.2 Sources artificielles.....	11
2.3 Quelques grandeurs d’intérêt	12
2.4 Gestion de crise suite à un accident nucléaire	13
2.4.1 Identification des phases de l’accident	13

2.4.2	Voies d'exposition.....	14
2.4.3	Gestion des denrées alimentaires contaminées.....	15
3.	Contamination environnementale et exposition humaine <i>via</i> la chaîne alimentaire..	20
3.1	Transfert de radionucléides dans l'environnement en situation accidentelle	20
3.1.1	Dispersion atmosphérique et dépôt des rejets.....	21
3.1.2	Transfert foliaire	23
3.1.3	Radionucléides dans les sols et transfert racinaire	29
3.1.4	Transfert aux produits animaux	31
3.2	Modélisation du transfert et de l'exposition par ingestion.....	34
3.2.1	Transfert à l'herbe et aux légumes feuilles.....	35
3.2.2	Transfert aux produits animaux	38
3.2.3	Exposition des populations par ingestion	40
4.	Incertitude et variabilité en radioécologie.....	42
4.1	Incertitude et variabilité : De quoi parle-t-on ?.....	42
4.2	Incertitude et variabilité dans les modèles radioécologiques d'évaluation de risque	44
4.2.1	Séparation Incertitude/ Variabilité.....	44
4.2.2	Sources d'incertitude et de variabilité dans les modèles radioécologiques....	45
5.	Analyse des sources d'incertitude et analyse de sensibilité.....	50
5.1	Introduction et démarche générale.....	50
5.2	Quantification des incertitudes.....	53
5.2.1	Approche probabiliste.....	54
5.2.2	Approches alternatives	67
5.3	Propagation d'incertitude.....	73
5.3.1	Méthodes de Monte Carlo	73
5.3.2	Bootstrap.....	80

5.3.3	Méthode alternative : méthode « hybride »	81
5.4	Méthodes d'analyse de sensibilité	87
5.4.1	Méthodes locales d'analyse de sensibilité	88
5.4.2	Méthodes de criblage	89
5.4.3	Méthodes globales d'analyse de sensibilité	91
5.4.4	Conclusion sur les méthodes d'analyse de sensibilité	95
	Partie II : Quantification des incertitudes sur les paramètres radioécologiques	99
6.	Interception en conditions sèches	103
6.1	Objectifs et méthodes	103
6.2	Article: Foliar interception of radionuclides in dry conditions: a meta-analysis using a Bayesian modeling approach	105
6.3	Résultats, limites et perspectives	119
7.	Interception humide	122
7.1	Objectifs et méthodes	122
7.2	Article: Interception of Wet Deposited Atmospheric Pollutants by Herbaceous Vegetation: Data Review and Modelling.	125
7.3	Résultats, limites et perspectives	158
8.	Décroissance biomécanique	162
8.1	Objectifs	162
8.1.1	Inférence bayésienne sur la constante de décroissance biomécanique	162
8.1.2	Propagation des incertitudes	163
8.2	Article: Uncertainty and variability in food chain exposure by radioactive fallouts. What can Hierarchical Bayesian modelling bring to radioecology community?	165
8.3	Résultats, limites et perspectives	202
	Partie III : Propagation des incertitudes	205
9.	Application à l'accident de Fukushima	207

9.1	Article: Uncertainty analysis in post-accidental risk assessment models: An application to the Fukushima accident	208
9.2	Résultats, limites et perspectives	222
	Partie IV: Conclusion Générale	225
10.	Conclusion Générale	227
10.1	Pluridisciplinarité et choix méthodologiques	227
10.2	Bilan des travaux	229
10.3	Limites et perspectives	232
10.3.1	Régionalisation des paramètres radioécologiques	232
10.3.2	Modélisation du transfert aux produits animaux	233
10.3.3	Incertitude et Variabilité sur les paramètres contextuels	235
10.3.4	Concilier modèle et mesures environnementales	238
	Références bibliographiques.....	242
	Annexes	263
A	Accidents de Tchernobyl et de Fukushima	265
A.1	Classement des accidents nucléaires	265
A.2	Dépôts de césium	266
A.3	Plans d'évacuation Fukushima	266
B	SYMBIOSE et le modèle ASTRAL	268
B.1	ASTRAL.....	268
B.2	La plateforme de modélisation SYMBIOSE	269
C	Lois de probabilités usuelles	271
C.1	Lois discrètes	271
C.2	Lois continues	273
D	Méthodes de Monte Carlo et méthode Hybride	276
D.1	Définition du cas d'étude	276

D.2 Résultats.....	282
D.3 Conclusion	283

Liste des Figures

Fig. 1.1 : Cadre général de l'analyse de risques. Adapté de (IPCS, 2009).	4
Fig. 2.1 : Pouvoir de pénétration des rayonnements α , β et γ (source 2015: Médiathèque site internet de l'IRSN)	10
Fig. 2.2 : Représentation schématique du zonage post-accidentel (source : Médiathèque site internet IRSN http://www.irsn.fr/FR/connaissances/Installations_nucleaires/La_surete_Nucleaire/risque-nucleaire/zonage-post-accidentel/Pages/sommaire.aspx#.VqiMe0roiUk).....	17
Fig. 3.1 : Schéma des voies d'exposition et du transfert de radionucléides dans les écosystèmes terrestres suite aux rejets atmosphériques accidentels. Adapté de (Thiessen <i>et al.</i> 1999).....	21
Fig. 3.2 : Schéma illustrant les dépôts sec et humide sur une surface agricole. Source Médiathèque site IRSN	22
Fig. 5.1 : Démarche générale d'analyse d'incertitude et d'analyse de sensibilité. Adapté de (Sudret, 2007).....	51
Fig. 5.2 : Illustration conceptuel d'un modèle.....	51
Fig. 5.3 : Illustration des visions fréquentiste et bayésienne (source : http://www.christian-faure.net/2013/08/31/deux-visages-du-calcul-des-probabilites-bayesiens-et-frequentistes/)..	61
Fig. 5.4 : Illustrations schématiques de la modélisation et de l'inférence hiérarchiques. (a) illustre un DAG représentant un modèle hiérarchique simple et (b) illustre les démarches de modélisation et d'inférence bayésiennes. (Adapté de (Ancelet, 2008)).....	66
Fig. 5.5 : Description de l'algorithme de la méthode de Monte Carlo 1D.	75
Fig. 5.6 : Schéma illustrant la méthode de Monte Carlo 2D. Source : (Pouillot et Delignette-Muller, 2010).....	76
Fig. 5.7 : Description de l'algorithme de Metropolis-Hastings.....	78
Fig. 5.8 : Description de l'échantillonneur de Gibbs.	79
Fig. 5.9: Description de la méthode itérative de propagation "hybride".	82

Fig. 5.10: Illustration schématique de la méthode hybride et présentation du résultat obtenu. Source (Guyonnet <i>et al.</i> 2005).....	83
Fig. 5.11 : Synthèse des méthodes d'analyse de sensibilité issue de (Iooss et Lemaitre, 2014).....	96
Fig. 7.1 : (a) DAG des modèles mécanistes utilisant un jeu de données dupliqué pour l'estimation commune des paramètres <i>SLA</i> et <i>L</i> . (b) DAG du modèle de mélange basé sur les deux modèles mécanistes. Le nouveau paramètre α est le paramètre de mélange. Les carrés à fond blanc correspondent aux données fixées (observations et paramètres fixés), les ellipses à fond gris sont les paramètres et les ellipses à fond blanc les variables latentes. Les flèches pleines désignent des relations stochastiques et celles en pointillées des relations déterministes.....	160
Fig. 10.1 : Cartes du rendement de l'herbe de pâturage en Europe. Le rendement est exprimé en tonnes (de matière sèche) par hectare ce qui équivaut à des kg (matière sèche) par m ² . (a) illustre des classes qui se chevauchent et (b) des classes disjointes. Illustration (Source : Peeters (1993) adapté de (Lee, 1983)).....	233
Fig. A. 1 : Echelle INES des accidents nucléaires.....	265
Fig. A. 2 : Cartes des dépôts de césium lors des accidents de Fukushima (à gauche) et de Tchernobyl (Source Médiathèque IRSN: http://www.irsn.fr/FR/connaissances/Environnement/expertises-incident-accidents/comparaison-tchernobyl-fukushima/Pages/sommaire.aspx#.VqiNY0roiUk).....	266
Fig. A. 3 : Plan d'évacuation au mois de juillet 2012 (Source : http://www.meti.go.jp/english/)	267
Fig. D. 1 : Histogrammes des distributions de probabilités des paramètres incertains et variables utilisées dans les scénarios A, B ou C. Les courbes rouges représentent les densités des distributions de probabilité.	278
Fig. D. 2: Distributions de possibilité des paramètres incertains ((a) <i>PSs</i> , (b) <i>PSh</i> et (c) <i>Ks</i>) utilisées dans le scénario C. Les segments désignent les α -coupes avec α égal à 0.6, 0.4 et 0.2 dans les figures (a), (b) et (c) respectivement.....	279
Fig. D. 3 : Fonctions de répartition empiriques des logarithmes en base 10 des concentrations de Cs-137 dans les légumes feuilles. (a) illustre les résultats issus des scénarios A et B, (b) les résultats des scénarios A et C et (c) les résultats des scénarios A, B et C.....	281

Liste des Tableaux

Tab. 2.1 : Demi-vies des isotopes du césium, de l'iode et du strontium généralement majoritaires dans les rejets accidentels (Nucleonica, 2014).....	12
Tab. 2.2 : Niveaux Maximaux Admissibles fixés par les règlements EURATOM N° 3954/87 et 770/90 (Adapté de (CODIRPA, 2012)).....	16
Tab. 4.1 : Sources d' <i>incertitude</i> et de <i>variabilité</i> potentielles identifiées dans les modèles d'évaluation du transfert de radionucléides vers l'homme par ingestion de denrées contaminées.....	46
Tab. D. 1 : Définition des distributions de probabilité et de possibilité associées aux paramètres incertains et variables. N = Loi normale ; U = Loi uniforme ; Tr = Loi de probabilité triangulaire/ distribution de possibilité triangulaire.....	277
Tab. D. 2 : Résultats des simulations des méthodes MC1D, MC2D et Hybride. Les concentrations sont exprimées en Bq.kg ⁻¹ de poids frais.	280
Tab. D. 3 : Indices d'Özkaynak <i>et al.</i> (2009) calculés pour les méthodes MC1D, MC2D et Hybride.....	280

« Nous souhaitons la vérité, et ne trouvons en nous qu'incertitude. »

Blaise Pascal, Pensées (1670)

1. Introduction

1.1 Contexte

Depuis la mise en service des premiers réacteurs dans les années 1950 et 1960 en France, l'industrie nucléaire a connu un essor tel qu'elle est devenue la principale source de production d'électricité en France avec près de 70% de la production nationale provenant des centrales nucléaires. Les différentes applications de l'énergie nucléaire, en plus de la production d'électricité, notamment médicales et militaires, lui confèrent un rôle important dans les activités socio-économiques en Europe et particulièrement en France. Avec les enjeux soulevés par le besoin accru d'énergie et par le réchauffement climatique ainsi que ses conséquences actuelles et futures sur l'environnement, la question de l'exploitation de l'énergie nucléaire pour la production d'électricité est au centre des débats publics. L'industrie nucléaire apparaît pour ses partisans comme une solution à la nécessité de réduire les émissions de CO₂ et au besoin d'indépendance énergétique pour bon nombre de pays (source SFEN : <http://www.sfen.org/fr>). Cependant, le risque d'accident, qui est une réalité inhérente à toute activité industrielle, a suscité au cours des dernières décennies des avis controversés ayant, suite aux accidents de Three Mile Island (1979) et de Tchernobyl (1986), considérablement freiné le développement de l'énergie nucléaire partout dans le monde (OCDE, 2010).

Les accidents majeurs survenus au cours des quatre dernières décennies (Tchernobyl, 1986 et Fukushima, 2011) ont engendré une importante contamination environnementale avec d'importantes quantités de radionucléides rejetés dans l'atmosphère pouvant présenter un risque sanitaire non négligeable pour les populations humaines. Le transfert de radionucléides dans l'environnement (et vers l'homme) s'inscrit dans un cadre disciplinaire né avec le développement de l'énergie nucléaire et communément appelé « radioécologie » (radioecology dans la littérature anglo-saxonne). La radioécologie est une science pluridisciplinaire introduite à la 1^{ère} Conférence Internationale sur l'utilisation pacifique de l'énergie atomique de Genève en 1955 (Odum, 1955; Vandecasteele, 2004) qui vise à étudier le comportement et le devenir des radionucléides dans l'environnement et les différents

processus intervenant dans leur transport à travers les milieux naturels et agricoles vers les végétaux, les animaux et les populations humaines (Vandecasteele, 2004).

Un des volets importants de la radioécologie concerne l'analyse des risques suite à un accident nucléaire majeur. La fiabilité de l'analyse des risques en situation accidentelle réside non seulement dans la capacité des acteurs à disposer d'informations pertinentes sur les différentes phases (phase d'urgence, phase post-accidentelle) et sur les caractéristiques des populations directement impactées mais aussi dans l'identification des différentes voies de contamination (CODIRPA, 2012). La consommation de produits issus des activités agricoles (voie alimentaire) constitue le mode majeur d'exposition radiologique interne durant la phase post-accidentelle, l'exposition au panache et aux rejets aériens n'existant plus (CODIRPA, 2012).

L'ampleur des répercussions de la catastrophe de Fukushima a entraîné une prise de conscience sur la difficulté de gérer les conséquences accidentelles et post-accidentelles d'un tel événement (CODIRPA, 2012). Apporter des réponses adaptées dans le cadre de l'évaluation des risques dans des contextes où l'information est le plus souvent incomplète et les délais très courts est particulièrement complexe et peut affecter considérablement la qualité des évaluations de risque (FAO/WHO, 2011). En effet, les acteurs de la gestion de risque sont confrontés à un ensemble d'incertitudes dues par exemple au manque de données et/ou à une compréhension imprécise des phénomènes et processus en jeu. Les évaluations de risques menées après l'accident de Fukushima (Bocquet, 2013; UNSCEAR, 2013; WHO, 2012) et une récente étude (Bedwell *et al.* 2015) ont ainsi mis en évidence l'influence des incertitudes liées aux paramètres contextuels (terme source, conditions climatiques) sur l'évaluation des doses reçues par les populations japonaises par ingestion de denrées contaminées. La pertinence et l'efficacité des décisions prises durant la phase post-accidentelle sont donc assujetties à la bonne compréhension que les acteurs de la gestion de crise ont (hors période de crise) de l'existence, de la nature des principales sources d'incertitude. Dans le cas spécifique de l'évaluation de l'exposition des populations humaines par voie alimentaire, les incertitudes portent sur les processus de transfert des radionucléides dans l'environnement et sur l'ensemble des processus anthropiques qui gouvernent leur transport vers l'homme *via* la chaîne alimentaire.

1.2 Problématique

Les travaux engagés dans ce projet de recherche entrent dans le cadre général de l'analyse des risques (illustré par [Fig. 1.1](#)) et portent spécifiquement sur l'évaluation de l'exposition aux substances radioactives des populations humaines par l'ingestion de denrées contaminées suite au rejet accidentel de ces substances dans l'environnement. Dans ce contexte, ces travaux de recherche participent d'une manière générale à la réflexion sur les enjeux méthodologiques relatifs à l'analyse des incertitudes associées au transfert des radionucléides dans les écosystèmes terrestres en situation accidentelle et à la mise en évidence de leur impact sur les résultats de calcul de dose ingérée. En France, cette réflexion s'inscrit dans l'axe continental du projet ANR, nommé AMORAD, impliquant entre autres l'IRSN ([IRSN, 2015a](#)) et dans l'action élaborée par le groupe de travail « Hypothèses » du CODIRPA en réponse à la problématique suivante ([CODIRPA, 2012](#)) : quels sont les paramètres et les hypothèses à retenir, dans le cadre des évaluations prédictives des conséquences radiologiques et dosimétriques d'une exposition accidentelle des populations, pour permettre aux autorités de décider des premières actions de protection au début de la phase de transition post-accidentelle ? L'action proposée par ce groupe de travail se concentre notamment sur les évaluations prédictives concernant les doses reçues au cours d'une période relativement courte (ex : 1 mois) suivant la fin des rejets radioactifs, le choix des populations retenues pour ces calculs de dose et la contamination des denrées agricoles (incluant les hypothèses relatives aux différents processus de transfert de radionucléides dans les écosystèmes terrestres). Au niveau européen, le développement de méthodes robustes pour la quantification des incertitudes dans les modèles d'évaluation des risques est affiché comme un sujet de recherche prioritaire dans le cadre du projet COMET ([Liland *et al.* 2013](#)) et apparaît comme une des thématiques d'intérêt dans la liste des défis scientifiques en radioécologie proposés dans l'agenda stratégique de recherche (SRA) de l'Alliance Européenne en Radioécologie (ERA) ([Hinton *et al.* 2013](#)).

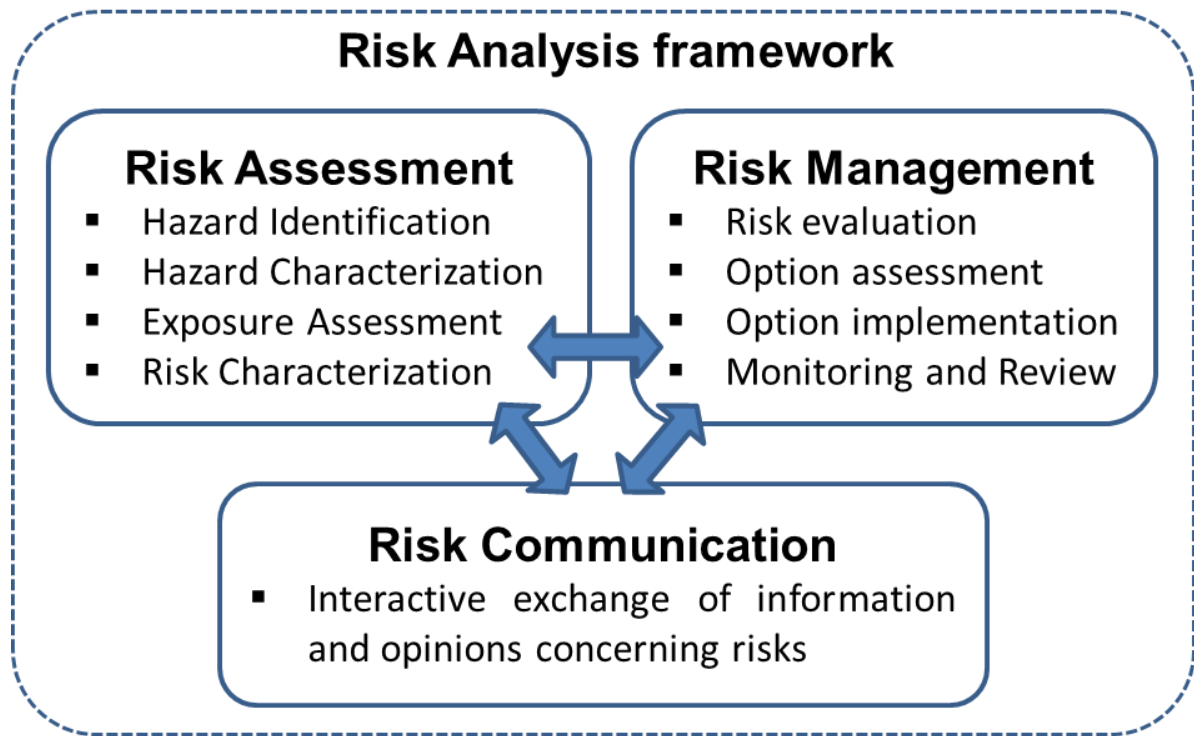


Fig. 1.1 : Cadre général de l'analyse de risques. Adapté de (IPCS, 2009).

1.3 Objectifs et démarche méthodologique

Sur la base des points soulevés précédemment, l'objectif principal de ce projet de recherche a été de développer une méthodologie et des méthodes pour la prise en compte des incertitudes dans les modèles d'évaluation des risques alimentaires suite à un accident majeur en vue d'améliorer les outils d'aide à la décision en situation accidentelle (ex : la plateforme SYMBIOSE développée par l'IRSN (cf. [Annexe B](#))). La démarche méthodologique adoptée a principalement consisté à identifier les différentes sources d'incertitudes, à les caractériser quantitativement, à les hiérarchiser et à évaluer leur impact sur les résultats des calculs d'exposition par voie alimentaire des populations humaines en phase post-accidentelle.

De manière plus spécifique, une première étape du travail a consisté à répertorier les sources d'incertitude et de variabilité susceptibles d'affecter la fiabilité des modèles prédictifs. Le deuxième volet concerne une part essentielle du travail qui a consisté à collecter des informations ou données pertinentes pour caractériser les processus et phénomènes d'intérêt. Ce travail de recherche de données a été conduit d'une part en effectuant une revue étendue de la littérature et d'autre part au travers des échanges de connaissances *via* les groupes de travail du projet COMET et du réseau Européen d'excellence STAR et enfin *via*

l'exhumation de la « littérature grise » de l'IRSN. Pour les différents processus étudiés, ce travail a permis de compiler des bases de données assez riches. Enfin, le troisième volet a consisté à identifier les méthodes adéquates permettant de quantifier ces sources d'incertitude et de variabilité *i.e.* de proposer une représentation quantitative des paramètres des modèles sur la base de l'information collectée.

Une seconde étape du travail a été entreprise dans l'optique d'évaluer l'impact de ces sources d'incertitude et de variabilité sur les sorties des modèles radioécologiques d'évaluation de risque (et d'impact dosimétrique). Les méthodes définies ont été mises en application sur deux études de cas, l'un portant sur un accident hypothétique et l'autre sur l'accident de Fukushima.

1.4 Organisation du document

Le présent manuscrit s'articule en de quatre parties.

La première partie du document, qui regroupe les [Chapitres 2 à 5](#), est dédiée à l'étude bibliographique qui a permis de faire un état des connaissances sur les différents concepts et disciplines qui sous-tendent le sujet de thèse. Les [Chapitres 2 et 3](#) sont relatifs à la radioactivité, aux activités nucléaires, au risque d'accident qui en découle et aux processus et phénomènes (physiques, chimiques, biologiques et anthropiques) décrivant le transfert des radionucléides dans les écosystèmes terrestres. Le [Chapitre 4](#) introduit les notions d'*incertitude* et de *variabilité*, et a pour but de définir le formalisme adopté dans le cadre de ce projet de recherche, ainsi que les contours de leur utilisation dans les différents développements présentés dans la suite. Le [Chapitre 5](#), dernier volet de cette étude bibliographique, est consacré aux différentes approches et méthodes statistiques associées permettant la quantification, la propagation et la hiérarchisation des sources d'*incertitude* et de *variabilité*.

Dans la deuxième partie du document, il est question des travaux menés sur la quantification des sources d'*incertitude* et de *variabilité* relatives aux phénomènes radioécologiques gouvernant le transfert de radionucléides, émis dans l'atmosphère, à travers la chaîne alimentaire *via* les écosystèmes terrestres en général et agricoles en particulier. En particulier, nous nous sommes intéressés à la modélisation des processus expliquant l'interception et la rétention des radionucléides par le couvert végétal en proposant une

caractérisation de l'incertitude et de la variabilité par la construction de modèles hiérarchiques reposant sur le paradigme bayésien. Ainsi les [Chapitres 6](#) et [7](#) traitent respectivement l'interception en conditions sèche et humide et le [Chapitre 8](#) s'intéresse à la décroissance biomécanique. Nous proposons également dans le [Chapitre 8](#) un cas d'étude sur la propagation des incertitudes (paramètres radioécologiques) appliquée à un accident hypothétique dans lequel un dépôt standardisé de quatre isotopes d'intérêt est défini. Ce cas d'étude a été mis en œuvre dans le cadre du groupe de travail « Human Food Chain » du projet européen COMET.

La troisième partie du document est consacrée à l'évaluation prédictive de l'impact de l'incertitude et de la variabilité sur les niveaux de concentrations attendus dans différents produits alimentaires. Une analyse probabiliste de risque appliquée à l'accident de Fukushima est proposée.

La quatrième et dernière partie du manuscrit est dédiée au bilan des travaux réalisés au cours de ces trois années et aux perspectives qu'ils ouvrent.

Les [Chapitres 6](#), [7](#), [8](#) et [9](#) sont présentés sous la forme d'articles. L'article présenté dans le [Chapitre 6](#) a été accepté le 8 Mai 2015 dans la revue *Journal of Environmental Radioactivity*. L'article présenté dans le [Chapitre 7](#) a été soumis le 22 Janvier 2016 à la revue *Science of the Total Environment*. L'article présenté dans le [Chapitre 8](#) a été soumis le 28 Janvier 2016 dans la revue *Journal of Environmental Radioactivity*. L'article relatif au [Chapitre 9](#) a été accepté le 17 Décembre 2015 dans la revue *Annals of Nuclear Energy*.

Partie I : Etude bibliographique

2. Radioactivité et risque nucléaire

2.1 La radioactivité : un peu d'histoire

En 1895, les travaux du physicien allemand Wilhelm Röntgen sur les rayons cathodiques ont permis de découvrir un rayonnement traversant la matière qu'il nomma à l'époque « rayons X ». La découverte de la radioactivité tient son origine de l'intérêt du physicien français Henri Becquerel pour les travaux de Röntgen et des études menées sur la phosphorescence faisant intervenir des sels d'uranium. Les travaux d'Henri Becquerel ont conduit en 1896 à la découverte d'un rayonnement, plus complexe que les « rayons X », impactant des plaques photographiques à travers une couche de papier noir. En 1898, les études menées par Pierre et Marie Curie, reprenant les travaux de Becquerel, ont permis d'isoler des éléments jusque-là inconnus (le polonium et le radium) émettant spontanément des rayonnements. Ce type de rayonnement fut rebaptisé « radioactivité » par Marie Curie.

La radioactivité est un phénomène physique naturel qui existe depuis l'origine de l'Univers et la formation des atomes constituant tout corps qu'il soit solide, liquide ou gazeux. Un atome est constitué d'un noyau, composé de protons et de neutrons et concentrant l'essentiel de sa masse, autour duquel se distribuent des électrons formant un nuage. Les atomes sont présents dans la nature sous forme stable ou instable (isotope radioactif) du fait d'un déséquilibre entre protons et neutrons. La radioactivité désigne alors l'ensemble des phénomènes physiques par lesquels les atomes instables libèrent, par désintégration de leur noyau, une quantité d'énergie sous forme de rayonnement et/ou de particules pour corriger naturellement cette instabilité. On distingue alors plusieurs types de rayonnements qui diffèrent par leur nature et leur pouvoir de pénétration dans la matière (Fig. 2.1) ; le rayonnement particulaire alpha (α) correspond à une perte de protons et de neutrons, le rayonnement particulaire bêta (β) revient à transformer un neutron en proton ou vice-versa et le rayonnement électromagnétique gamma (γ) correspond à une émission de photons (particules composant la lumière).

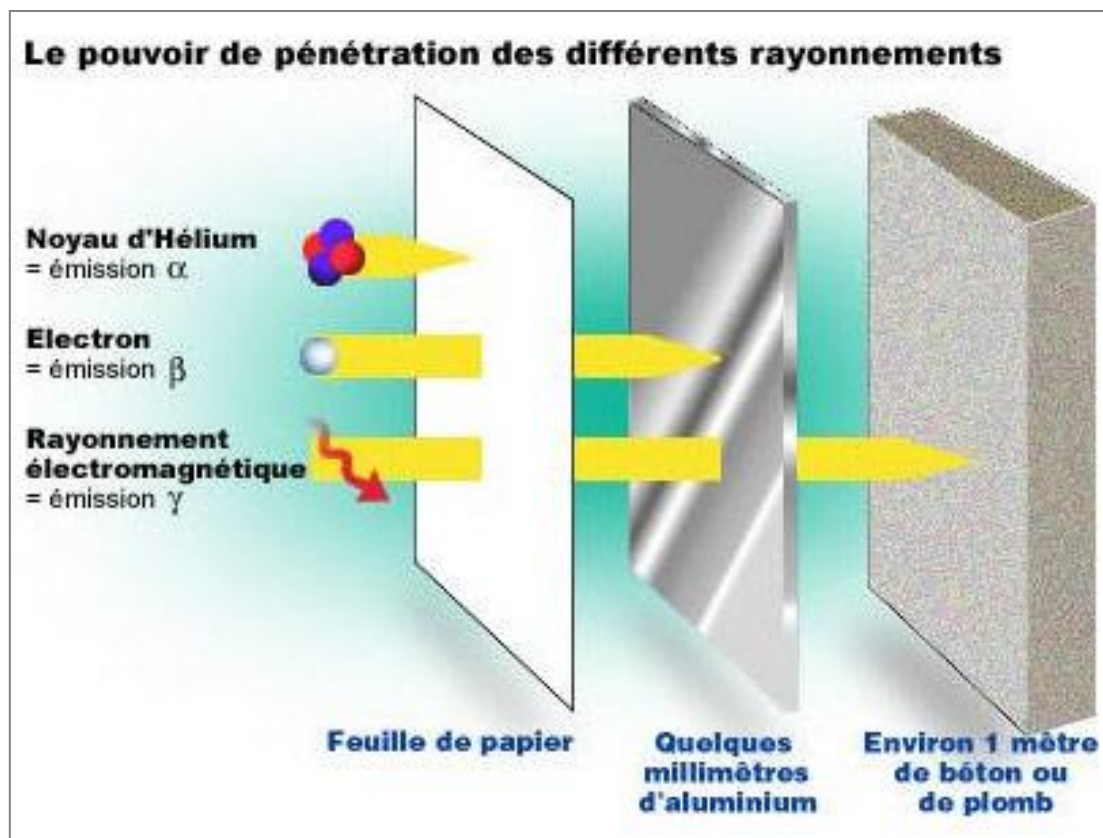


Fig. 2.1 : Pouvoir de pénétration des rayonnements α , β et γ (source 2015: Médiathèque site internet de l'IRSN)

2.2 Les sources de radioactivité

Les niveaux de radioactivité observés actuellement dans l'environnement sont le résultat des radionucléides naturellement présents sur Terre depuis son origine mais également de la décroissance des radionucléides produits artificiellement et introduits depuis le milieu du XXème siècle par les différentes activités anthropiques.

2.2.1 Sources naturelles

La radioactivité d'origine naturelle peut provenir de différentes sources. Les principales sont :

- Le rayonnement cosmique, qui provient soit des ions issus des galaxies soit du soleil par le biais du « vent solaire » constitué principalement de protons. Le champ magnétique terrestre forme avec l'atmosphère un écran protecteur de sorte qu'une

partie négligeable de ce rayonnement arrive au niveau de la mer. À l'inverse, l'importance du rayonnement cosmique augmente avec l'altitude.

- La radioactivité atmosphérique, qui est principalement due à la présence de radon ²²² sous forme gazeuse (issu de l'uranium 238) dans l'atmosphère.
- Le rayonnement tellurique, qui correspond à la radioactivité du sol qui est principalement émise par les éléments radioactifs présents dans l'écorce terrestre tels que l'uranium et le thorium.
- La radioactivité des eaux, dépendant des propriétés chimiques des terrains lessivés en surface. Certaines eaux souterraines sont riches en gaz radon dissous.
- La radioactivité induite par les produits d'activation présents dans le corps humain tels que le carbone 14 (C-14).

2.2.2 Sources artificielles

Contrairement aux radionucléides naturels, les radionucléides artificiels désignent l'ensemble des substances radioactives introduites dans l'environnement du fait des différentes activités nucléaires initiées depuis le début du XXe siècle. Cette radioactivité artificielle provient principalement :

- Des applications médicales. Certains radionucléides peuvent être utilisés comme source de rayonnements pour des radiographies ou comme source d'irradiation pour des radiothérapies.
- Des retombées des essais atmosphériques d'armes nucléaires (des années 1940 aux années 1980).
- Des rejets autorisés en fonctionnement normal des installations nucléaires. Les radionucléides rejetés par les installations nucléaires sont multiples et différent en fonction du type d'installation et de sa phase de fonctionnement (production, arrêt pour maintenance, démantèlement *etc.*).
- Des rejets accidentels. Les catastrophes majeures de Tchernobyl et de Fukushima ont engendré une importante contamination environnementale avec des quantités importantes de radionucléides rejetés dans l'environnement ([Steinhauser et al. 2014](#)).

2.3 Quelques grandeurs d'intérêt

La finalité de la radioécologie est d'étudier l'impact sur l'environnement et sur l'homme des substances radioactives présentes dans les différents compartiments de l'environnement. Différentes unités sont disponibles et permettent de quantifier cet impact.

L'activité d'une substance radioactive est mesurée par le becquerel (Bq) qui correspond à 1 désintégration par seconde, l'ancienne unité étant le Curie ; $1 \text{ Ci} = 3.7 \cdot 10^{10} \text{ Bq} = 3.7 \cdot 10^{10}$ transformations nucléaires par seconde. Selon le compartiment environnemental d'intérêt, cette activité peut être exprimée par unité de masse (Bq/kg sol, faune ou flore), de volume (Bq/m³ ou Bq/L d'air ou d'eau) ou de surface (Bq/m² de sol). Les noyaux des éléments radioactifs se désintègrent en fonction du temps en donnant naissance à d'autres éléments. Ainsi, l'activité des substances radioactives évolue au cours du temps selon une période radioactive appelée demi-vie correspondant au temps nécessaire pour que l'activité diminue de moitié. Il est important de noter que cette demi-vie ne dépend pas de l'environnement (température, pression) mais est une propriété intrinsèque au radionucléide considéré (Tab. 2.1).

Tab. 2.1 : Demi-vies des isotopes du césium, de l'iode et du strontium généralement majoritaires dans les rejets accidentels (Nucleonica, 2014).

Isotope	Demi-vie
Cs-134	2 ans
Cs-137	30 ans
I-131	8 jours
Sr-90	28 ans

La dose absorbée par la matière correspondant à la quantité d'énergie absorbée par unité de masse est exprimée en grays (Gy). Un (1) Gy correspond à 1 Joule d'énergie par kg de matière. L'accroissement de la dose absorbée par intervalle de temps désigne le débit de dose absorbée (exprimée en Gy/s ou Gy/h) représentant la quantité d'énergie absorbée par unité de masse par unité de temps.

L'impact du rayonnement sur l'homme est quantifié par la dose efficace (somme des doses reçues par les différents organes du corps, ramenée au corps entier intégrant les différences

entre type de rayonnements et prenant en compte la « fragilité » de chaque organe) et la dose équivalente (dose moyenne reçue par un organe intégrant les différences entre types de rayonnement) exprimées en Sievert (Sv).

2.4 Gestion de crise suite à un accident nucléaire

Suite à un accident nucléaire majeur, les acteurs de la gestion de crise sont amenés à prendre des décisions et à mener des actions pour assurer la protection des populations humaines de manière adéquate. L'efficacité de ces actions réside dans la capacité des décideurs à identifier précisément les différentes phases de l'accident, les principales voies d'exposition, les caractéristiques des populations directement impactées et dans la mise en œuvre de réponses adaptées. En France, la directive interministérielle du 7 avril 2005 a chargé l'Autorité de Sûreté Nucléaire (ASN) d'établir le cadre pour la mise en œuvre des dispositions nécessaires afin de répondre aux situations post-accidentelles consécutives à un accident nucléaire (CODIRPA, 2012). A la suite de l'accident de Fukushima, le gouvernement français a choisi de renforcer le dispositif existant en tirant les leçons de cette catastrophe pour une meilleure prévention et une réaction plus efficace face à un événement de même nature (source : <http://www.risques.gouv.fr/actu-risques-crisis/actualites/nouveau-plan-national-accident-nucleaire-ou-radiologique-majeur>). Un plan national de réponse à un « Accident nucléaire ou radiologique majeur » a ainsi été élaboré, puis publié en février 2014. Dans les sections suivantes, les éléments énoncés font exclusivement référence à la doctrine établie et adoptée à l'échelle nationale.

2.4.1 Identification des phases de l'accident

En situation accidentelle et suite aux rejets de radionucléides dans l'environnement, on distingue la phase d'urgence et la phase post-accidentelle (CODIRPA, 2012).

2.4.1.1 La phase d'urgence

La phase d'urgence est caractérisée par la nécessité d'agir dans des délais très courts pour faire face aux potentiels rejets susceptibles d'entraîner une exposition non négligeable des

populations et pour engager les premières actions de protection de la population (CODIRPA, 2012, SGDSN, 2014a, b). Elle se compose de :

- Une période de menace, résultant de défaillances dans l'installation et durant laquelle l'exploitant met en œuvre des actions visant à restaurer un niveau de sûreté satisfaisant.
- Une période de rejets radioactifs dans l'environnement, qui correspond à une phase critique durant laquelle la menace n'a pas pu être confinée au sein de l'installation.
- Une période de sortie de la phase d'urgence, annonçant la fin des rejets radioactifs et l'absence de nouvelle menace de rejet.

2.4.1.2 La phase post-accidentelle

La phase post-accidentelle succède à la phase d'urgence et donc correspond à la fin des rejets et au retour de l'installation à un état sûr (CODIRPA, 2012). Elle se compose de :

- Une phase de transition, pouvant durer jusqu'à quelques mois après l'accident. Les risques d'exposition des populations peuvent être très importants.
- Une période de long terme, pouvant durer jusqu'à quelques décennies après l'accident. Le risque d'exposition des populations est réel du fait d'une contamination durable des territoires mais ce risque est néanmoins faible.

Cette définition de la phase post-accidentelle est discutable. En effet, dans le cas d'un accident nucléaire ayant généré des rejets atmosphériques de très longue durée (*i.e.* de plusieurs semaines, comme par exemple lors de l'accident de Fukushima) les préoccupations relatives à la phase post-accidentelle doivent être prises en compte et les différentes actions et décisions s'y rapportant menées dès la phase d'urgence (*i.e.* avant la fin des rejets). Ce décalage temporel représente une limite principalement liée au fait que la doctrine CODIRPA a été établie pour des accidents de niveau 6 sur l'échelle INES.

2.4.2 Voies d'exposition

L'exposition de l'Homme, de la faune et de la flore aux substances radioactives peut être externe ou interne.

L'exposition externe fait référence soit au contact cutané avec des substances radioactives, soit à l'irradiation du fait des rayonnements émis par les radionucléides dispersés dans l'air ou

déposés au sol. Cette exposition est communément appelée irradiation externe. Les effets qui découlent de cette exposition peuvent être très importants dans la mesure où les rayonnements des émetteurs gamma dispersés dans l'air (sous forme de gaz ou de fines particules en suspension appelées « aérosols ») peuvent impacter des organes en profondeur du fait de leur pouvoir pénétrant.

L'exposition interne résulte de l'entrée de substances radioactives dans les organismes par l'inhalation de particules radioactives présentes dans l'air et par l'ingestion de denrées contaminées suite aux retombées radioactives. Les radionucléides incorporés irradient les tissus et les organes pendant toute la durée de leur présence dans l'organisme ; certains radionucléides peuvent se concentrer préférentiellement dans certains organes, comme par exemple les iodes radioactifs dans la glande thyroïde. La persistance dans le corps dépend de la période radioactive du radionucléide et de son élimination plus ou moins rapide par les voies naturelles.

Selon la phase de l'accident considérée, l'importance des différentes voies d'exposition peut varier. Durant la phase de rejets, l'inhalation de l'air contaminé et l'irradiation externe lors du passage du panache radioactif constituent les principales voies d'exposition. A la fin des rejets atmosphériques, l'irradiation externe aux dépôts constitue la principale voie d'exposition externe et la consommation de denrées alimentaires contaminées constitue la principale voie d'exposition interne. Dans le contexte de la gestion post-accidentelle les décisions doivent donc être prises en prenant en considération cette voie d'exposition d'intérêt.

2.4.3 Gestion des denrées alimentaires contaminées

2.4.3.1 Contexte international et déclinaison française

Des niveaux de références européens et internationaux relatifs aux denrées alimentaires, appelés niveaux maximaux admissibles (NMA), constituant des limites réglementaires pour la commercialisation des denrées alimentaires à l'intérieur de l'Europe, ont été établis pour des produits destinés au bétail et à la consommation humaine (Tab. 2.2).

Tab. 2.2 : Niveaux Maximaux Admissibles fixés par les règlements EURATOM N° 3954/87 et 770/90 (Adapté de (CODIRPA, 2012)).

Denrées alimentaires (NMA en Bq.kg ⁻¹)				
	Aliments pour nourrissons	Produits laitiers	Autres denrées alimentaires	Liquides destinés à la consommation
Isotope Sr (Sr-90)	75	125	750	125
Isotope Iode (I-131)	150	500	2000	500
Isotope Pu (Pu-239)	1	20	80	20
Radionucléides (période > 10 j) Cs-134/-137	400	1000	1250	125
Catégories d'animaux		Cs-134/-137 (NMA en Bq.kg⁻¹)		
Porcs		1250		
Volailles, agneaux et veaux		2500		

Pour la France, le CODIRPA (2012) a proposé pour la protection des populations vis-à-vis de la consommation de denrées contaminées de définir, en amont ou durant la phase post-accidentelle, des zones visant à éloigner les populations des zones contaminées, à bannir les produits dépassant les seuils d'activité autorisée et à limiter la consommation de certains produits présentant potentiellement un risque (CODIRPA, 2012).

Ce zonage post-accidentel, une des composantes essentielles de la gestion post-accidentelle, serait mis en œuvre en cas d'accident ayant des conséquences en France, par le préfet sur la base de ces NMA (SGDSN, 2014a). Le système Français distingue principalement deux zones illustrées par Fig. 2.2.

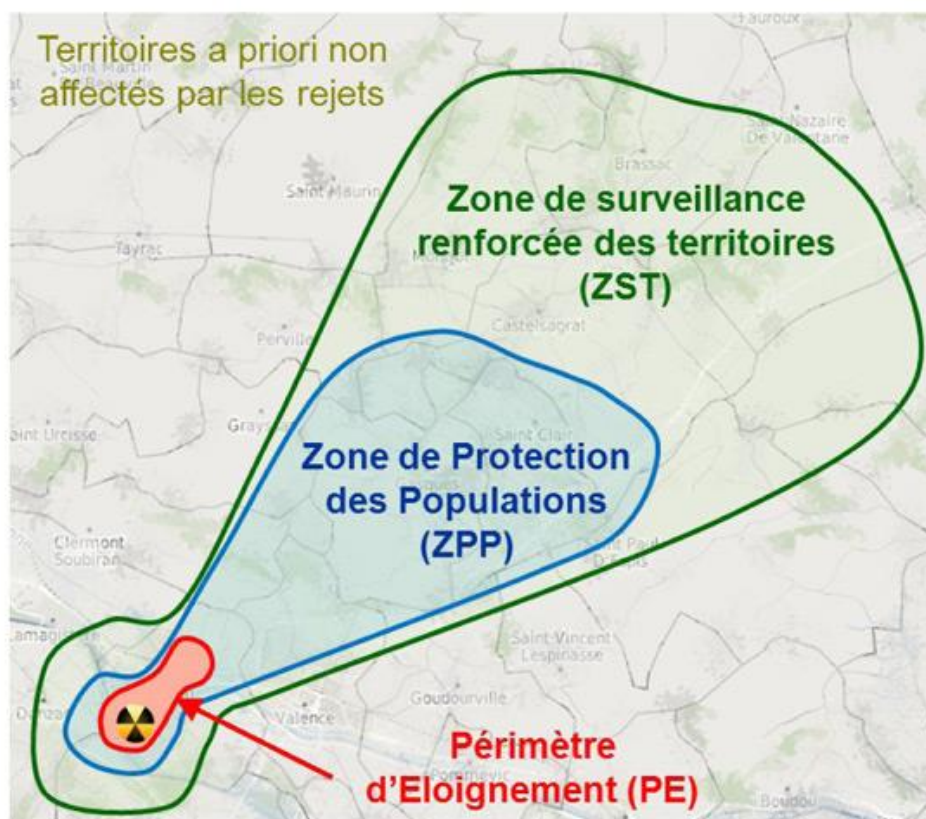


Fig. 2.2 : Représentation schématique du zonage post-accidentel (source : Médiathèque site internet IRSN

http://www.irsn.fr/FR/connaissances/Installations_nucleaires/La_surete_Nucleaire/risque-nucleaire/zonage-post-accidentel/Pages/sommaire.aspx#.VqiMe0roiUk).

2.4.3.2 Zone de protection des populations

En cas d'accident nucléaire, la zone de protection des populations (ZPP) correspondrait au périmètre au sein duquel sont mises en œuvre des actions visant à réduire l'exposition à la radioactivité ambiante des populations locales (SGDSN, 2014a). Cette zone est définie dans un objectif de radioprotection de la population vivant dans les territoires les plus contaminés, sur la base des évaluations prévisionnelles des doses susceptibles d'être reçues au cours du premier mois suivant la fin des rejets (CODIRPA, 2012). Les denrées issues de la ZPP (d'origine agricole comme provenant de la pêche, de la chasse ou de la cueillette) sont interdites de consommation et de mise sur le marché, quel que soit leur niveau de contamination (CODIRPA, 2012). La ZPP inclut un périmètre d'éloignement (PE) défini comme la zone dans laquelle les prévisions de dose efficace sur le premier mois après la fin des rejets dépasseraient 10 mSv. Il faut noter que ces évaluations prévisionnelles ne prennent pas en compte l'ingestion de denrées d'origine locale.

2.4.3.3 Zone de surveillance renforcée des territoires

La zone de surveillance renforcée des territoires (ZST) définirait en situation accidentelle une zone allant au-delà de la ZPP englobant l'ensemble des périmètres au sein desquels, pour une catégorie de production agricole donnée susceptible d'être produite et récoltée au cours des mois suivant la fin des rejets, les NMA pourraient être dépassés (CODIRPA, 2012). La commercialisation de denrées alimentaires serait limitée et particulièrement contrôlée dans la ZST. Contrairement à la ZPP, la ZST est définie dans un objectif de préservation des activités économiques ; ce qui garantit la commercialisation des produits conformes uniquement.

La pertinence du zonage mis en œuvre serait réévaluée régulièrement compte tenu d'une connaissance de plus en plus précise, grâce aux différentes campagnes de mesures sur le terrain, de l'état réel de la contamination de l'environnement (CODIRPA, 2012) ; le premier zonage étant établi sur la base de calculs prédictifs reposant sur la modélisation des transferts de radionucléides dans l'environnement. Cela souligne l'importance, dans le contexte de la gestion post-accidentelle, d'une bonne compréhension des processus en jeu et de leur caractérisation à travers les modèles mathématiques utilisés. La section suivante présente la modélisation du transfert de radionucléides émis lors d'un accident et du calcul d'impact dosimétrique associé.

3. Contamination environnementale et exposition humaine *via* la chaîne alimentaire

Dans le cadre de ces travaux de recherche, nous nous sommes intéressés à la contamination des populations humaines, suite aux rejets accidentels de substances radioactives dans l'environnement, en nous focalisant principalement sur l'exposition humaine induite par la consommation de denrées alimentaires contaminées. Dans cette section, les différents processus intervenant dans le transfert de radionucléides de l'atmosphère sont introduits de manière succincte. Les différents scénarios de transfert considérés dans ce travail de recherche et les modèles d'évaluation de risque correspondant sont également évoqués.

3.1 Transfert de radionucléides dans l'environnement en situation accidentelle

Le transfert de radionucléides dans les écosystèmes terrestres est gouverné par un ensemble de processus qui interviennent en interaction dans les différents compartiments de la biosphère comme illustré par [Fig. 3.1](#).

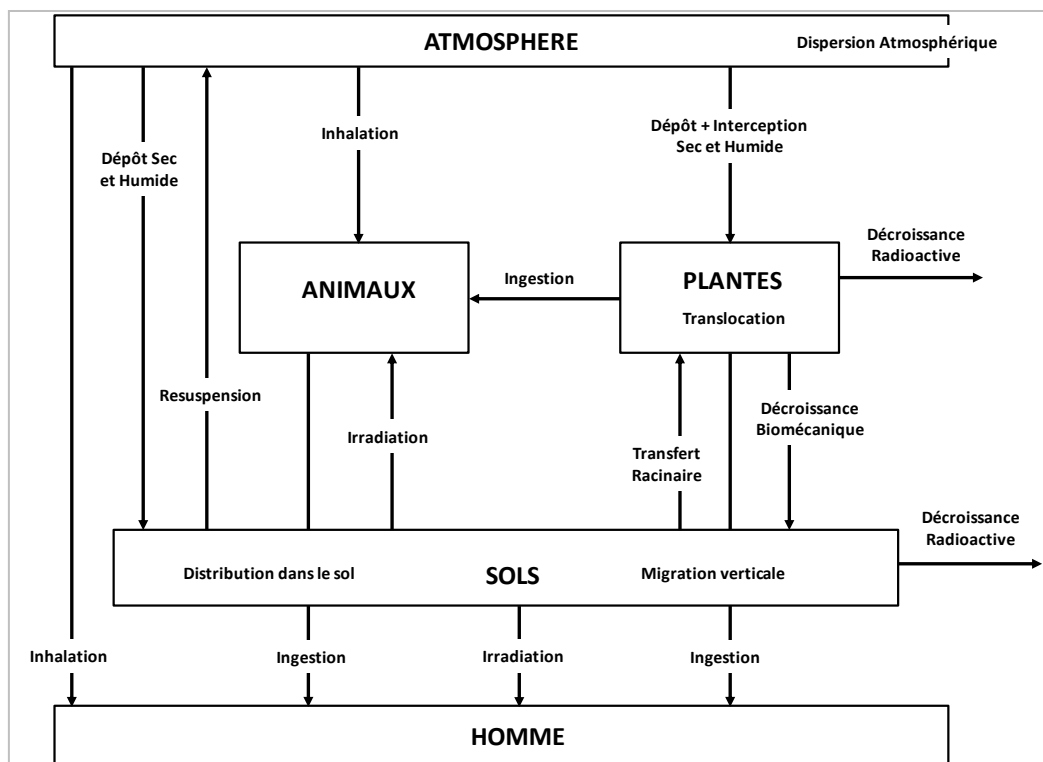


Fig. 3.1 : Schéma des voies d'exposition et du transfert de radionucléides dans les écosystèmes terrestres suite aux rejets atmosphériques accidentels. Adapté de (Thiessen *et al.* 1999).

3.1.1 Dispersion atmosphérique et dépôt des rejets

Suite à un accident dans une installation nucléaire, des substances radioactives peuvent être rejetées en masse formant un panache radioactif qui sera dispersé dans l'atmosphère sur des distances plus ou moins grandes en fonction de la taille et des caractéristiques des substances rejetées (Bocquet, 2013), des conditions de l'accident (explosion, incendie), des conditions atmosphériques (météo). Par exemple, les particules les plus lourdes sont transportées à des distances relativement proches de l'installation concernée. La dispersion atmosphérique des polluants atmosphériques, y compris les radionucléides rejetés dans l'atmosphère sous forme gazeuse et particulaire, est soumise à l'action de différents mécanismes physiques (ex : l'advection par le vent, la diffusion turbulente, *etc.*), aux conditions météorologiques (vent et précipitations) et aux propriétés physico-chimiques de l'atmosphère (Roustan, 2005). Les particules dispersées dans l'atmosphère se déposent ensuite à la surface du sol et sur le couvert végétal. Le dépôt est considéré différemment suivant les caractéristiques physiques des radioéléments rejetés (sous forme gazeuse ou particulaire) du fait de la prise en compte de

mécanismes différents (Sehmel, 1980). Le dépôt de radionucléides intervient de deux manières : par dépôt sec et par dépôt humide (Fig. 3.2).

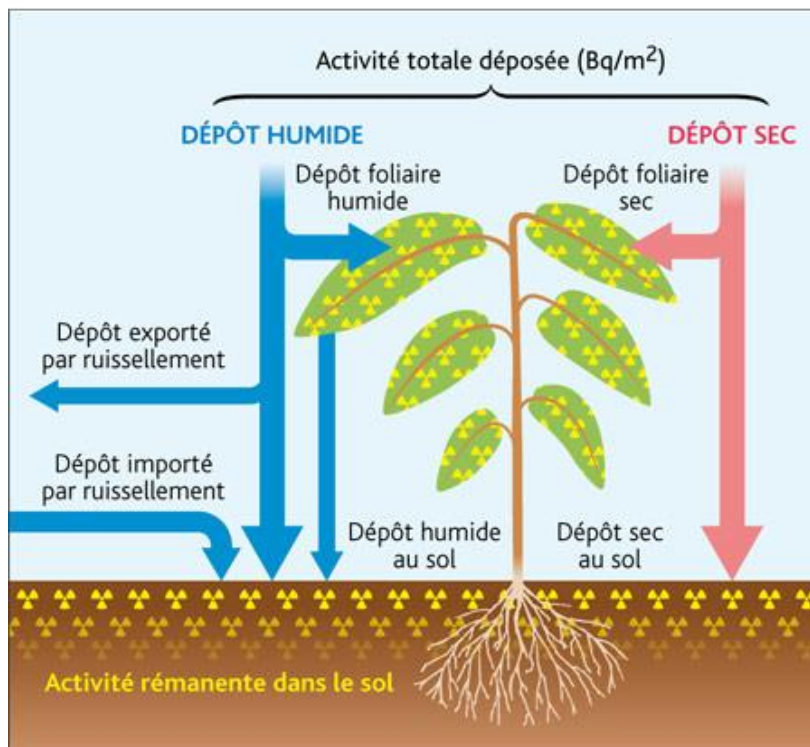


Fig. 3.2 : Schéma illustrant les dépôts sec et humide sur une surface agricole. Source Médiathèque site IRSN

3.1.1.1 Dépôt sec

Le dépôt sec représente l'ensemble des mécanismes physiques (comme la diffusion brownienne, la sédimentation ou l'impaction) entraînant, en l'absence de précipitations, le dépôt de particules amenées par l'écoulement atmosphérique à proximité de la surface terrestre (Petroff, 2005). Le dépôt sec est en pratique entièrement caractérisé par un flux de dépôt sec déterminé à partir de la concentration de radionucléides dans l'air et de la vitesse de dépôt sec de ces radionucléides (Calec, 2013). Il faut noter que la vitesse de dépôt sec dépend des caractéristiques de la particule et de la surface réceptrice considérées.

3.1.1.2 Dépôt humide

Le dépôt humide désigne, contrairement au dépôt sec, l'ensemble des processus traduisant le retrait d'un polluant de l'atmosphère par le biais des précipitations (Roustan, 2005). Le dépôt humide est déterminé par le phénomène de captation des aérosols par les gouttes d'eau

lors de la formation des nuages appelé lavage (« in-cloud scavenging » ou « rainout » dans la littérature anglophone) et par le phénomène de lessivage (« below-cloud scavenging » ou « washout » dans la littérature anglophone) représentant le rabatement des aérosols par les gouttes lors des précipitations (Calec, 2013 ; Depuydt, 2013).

L'étude des processus de dépôt et leur caractérisation en fonction des surfaces réceptrices (surtout pour le dépôt sec, le dépôt humide dépendant peu des surfaces réceptrices (Calec, 2013)) restent des sujets de recherche d'une grande complexité et d'actualité suite à l'accident de Fukushima. Ces sujets ont notamment été abordés dans plusieurs travaux de recherche au sein de l'IRSN (e.g. Calec, 2013 ; Depuydt, 2013 ; Petroff, 2005 ; Roupsard, 2013).

3.1.2 Transfert foliaire

Les radionucléides déposés sur le couvert végétal sont par la suite incorporés dans les plantes à travers leur système foliaire. Le terme transfert foliaire est alors employé. Le transfert des radionucléides par la voie foliaire revêt une importance particulière en analyse des risques suite à un accident nucléaire car les feuilles des végétaux représentent la principale interface entre les radionucléides et les produits consommés par les populations humaines (produits cultivés ou issus de l'élevage). Les substances radioactives sont incorporées dans les plantes par le biais de différents phénomènes physiques, chimiques et biologiques (Coughtrey and Thorne, 1983).

3.1.2.1 Interception foliaire

L'interception foliaire représente la fraction de radionucléides initialement déposée à la surface du couvert végétal qui est absorbée par les feuilles des plantes (IAEA, 2009). Durant l'interception foliaire un certain nombre de facteurs interagissent (Pröhl, 2009) tels que l'état de développement de la plante, sa capacité à retenir l'eau, les propriétés de la particule ou encore la quantité de pluie. En pratique, l'interception par le système foliaire est quantifiée différemment selon la paramétrisation adoptée et les informations disponibles. En effet, on distingue selon (Pröhl, 2009) :

- Le ratio entre l'activité initialement retenue par les feuilles suite au dépôt (A_i exprimée en $\text{Bq}\cdot\text{m}^{-2}$ de surface de sol) et l'activité totale déposée sur le sol (A_t exprimée en $\text{Bq}\cdot\text{m}^{-2}$ de surface de sol) communément noté f .

$$f = \frac{A_i}{A_t} \quad (3.1)$$

- La normalisation du paramètre f soit par la biomasse aérienne des plantes B , exprimée en kg de masse sèche ou de masse fraîche par m^2 de sol, soit par l'indice de surface foliaire (ou leaf area index (LAI) en anglais) qui est le rapport en $m^2 \cdot m^{-2}$ entre la surface occupée par les feuilles sur le sol et la surface du sol. Ainsi, on obtient les nouveaux paramètres f_B et f_{LAI} qui permettent d'intégrer le stade de développement de la plante dans la définition de l'interception foliaire.

$$f_B = \frac{f}{B} \text{ et } f_{LAI} = \frac{f}{LAI} \quad (3.2)$$

L'interception foliaire étant un processus clé dans le transfert des radionucléides a fait l'objet de plusieurs études visant à élaborer des modèles pour la prédiction du facteur d'interception en situation de dépôts sec et humide. Différents modèles empiriques, *i.e.* ne reposant pas sur une caractérisation explicite des mécanismes physico-chimiques, ont ainsi été élaborés depuis les années 1970. [Chadwick et Chamberlain \(1970\)](#) ont établi une dépendance non-linéaire entre le facteur d'interception et la biomasse à l'issue d'un dépôt sec à travers l'équation ci-dessous :

$$f = 1 - e^{-\mu \times B} \quad (3.3)$$

B est la biomasse aérienne (en $kg \cdot m^{-2}$) et μ est un coefficient d'absorption (exprimé en m^2 de sol par kg sec ou frais) traduisant le ratio entre l'activité massique initiale sur les feuilles ($Bq \cdot kg^{-1}$) et l'activité surfacique au sol ($Bq \cdot m^{-2}$). En considérant la forme de l'équation (3.3) et en accord avec les propriétés de l'exponentielle, on remarque que pour des valeurs suffisamment petites de la biomasse aérienne ($B \approx 0.3 \text{ kg} \cdot m^{-2}$ selon [Chadwick et Chamberlain \(1970\)](#)), l'interception est sensiblement égale au produit $\mu \times B$ et donc l'interception massique f_B et le coefficient d'absorption μ sont équivalents. [Vandecasteele et al. \(2001\)](#) ont proposé une formulation analogue préconisant d'utiliser l'indice de surface foliaire.

$$f = 1 - e^{-\mu \times LAI} \quad (3.4)$$

ce qui implique :

$$f_{LAI} = \frac{1 - e^{-\mu \times LAI}}{LAI} \quad (3.5)$$

Le LAI avait déjà été intégré dans un modèle proposé par Muller et Pröhl (1993) pour l'estimation de la fraction interceptée suite au dépôt humide de radionucléides. Le modèle ECOSYS de Muller et Pröhl (1993) prend en compte, en plus de l'index foliaire, la capacité S (exprimée en mm) de stockage d'eau de la plante, la quantité R (en mm) de pluie tombée et un facteur k (sans unité) qui quantifie l'aptitude de la particule à s'accrocher aux feuilles. Ce modèle, décrit ci-dessous, semble apporter une bonne estimation de l'interception du dépôt humide pour le blé, les fèves et l'herbe (Kinnersley *et al.* 1997 ; Pröhl, 2009).

$$f = \min\left(1; \frac{LAI \cdot k \cdot S}{R} \left[1 - e^{-\frac{\ln 2}{3 \cdot k \cdot S} R}\right]\right) \quad (3.6)$$

Pour une meilleure prise en compte du stade de développement des espèces végétales d'intérêt il est suggéré par plusieurs auteurs (*e.g.* Muller and Pröhl, 1993 ; Pröhl, 2009 ; Vandecasteele *et al.* 2001) de privilégier l'indice de surface foliaire car celui-ci rend compte plus fidèlement de l'évolution des stades de développement des plantes (floraison, feuillaison, sénescence *etc.*). Cependant, obtenir une mesure précise de l'indice de surface foliaire a un coût matériel qui limite son intérêt dans une optique opérationnelle. L'interception foliaire est un processus complexe à modéliser du fait de sa dépendance à des facteurs tels que le développement des plantes, l'ouverture ou non de leurs stomates, les caractéristiques chimiques de la surface des feuilles, les propriétés physico-chimiques des particules considérées et les données liées aux précipitations si on considère le dépôt humide. Nous nous attacherons plus en détail à la caractérisation de ce processus dans la suite du manuscrit.

3.1.2.2 Décroissance biomécanique

Une fois les particules radioactives déposées à la surface des feuilles, interceptées, l'activité dans les feuilles peut être réduite par l'action de plusieurs mécanismes environnementaux, biologiques et anthropiques (Miller and Hoffman, 1983). On peut noter l'action de phénomènes physiques tels que l'érosion éolienne, le lessivage par l'eau de pluie, la volatilisation et l'évaporation des radionucléides, l'action de mécanismes biologiques tels que l'abrasion des cuticules, le vieillissement des tissus, la chute des feuilles ou des actions telles que le pâturage (IAEA, 2009, 2010). Ces phénomènes engendrent une perte de l'activité qualifiée de « perte au champ » (« field loss » en anglais) car les particules sont entraînées sur le sol. Cependant, il est important de noter que l'activité dans les plantes peut également être réduite du fait de la croissance des plantes. Ce phénomène appelé communément dilution

biologique est souvent inclus dans la « perte aux champs » même s'il ne génère pas vraiment de perte (IAEA, 2010). Tous ces mécanismes peuvent être regroupés derrière le terme « décroissance biomécanique ». Dans la suite, nous utiliserons également le terme perte aux champs. La décroissance biomécanique est généralement représentée, dans les modèles radioécologiques, par une constante de lessivage λ^W (exprimée en jour⁻¹) décrivant la vitesse à laquelle la concentration des plantes est réduite par les mécanismes environnementaux évoqués précédemment. La constante de lessivage est décrite comme suit :

$$\lambda^W = \frac{\log 2}{T^W} \quad (3.7)$$

T^W est la demi-vie environnementale correspondant au temps nécessaire pour que la moitié de l'activité soit réduite du fait des différents processus environnementaux. En pratique, la constante de lessivage est estimée par régression (non-linéaire ou linéaire sur la transformée logarithmique des données) appliquée :

- soit aux concentrations de radionucléides dans les plantes considérées mesurées en fonction du temps et exprimées en Bq.kg⁻¹,
- soit aux concentrations de radionucléides dans les plantes mesurées en fonction du temps et exprimées en Bq.m⁻².

Dans le premier cas, la constante de lessivage estimée intègre implicitement la dilution biologique alors que dans le second cas, la réduction du fait de la croissance des plantes n'est pas prise en compte. La prise en compte de la dilution biologique n'est pas toujours effective dans les modèles radioécologiques. La revue de la littérature réalisée par l'AIEA (IAEA, 2009) a permis de collecter une base de données de valeurs de décroissance biomécanique qui a révélé une certaine variabilité entre les éléments, entre les plantes et entre les stades de développement des différentes plantes considérées. Dans la suite du manuscrit nous nous sommes intéressés plus en détail à la caractérisation de cette variabilité.

3.1.2.3 Translocation

Suite à l'interception (sèche ou humide) initiale et à la captation des radionucléides déposés sur le couvert végétal, la complexité du transfert foliaire résulte de la grande variété de processus biologiques affectant l'efficacité du transfert foliaire tels que la fixation et la pénétration des radionucléides à travers les cuticules des plantes, l'internalisation par les cellules des feuilles et les processus de transport et de distribution intervenant dans la plante

(Hurtevent *et al.* 2013). L'ensemble de ces mécanismes désigne le processus de translocation consécutif à l'interception. La translocation représente ainsi la redistribution du radionucléide déposé à la surface des feuilles vers des parties de la plante non impactées lors du dépôt et (IAEA, 2010) et vers l'organe d'intérêt étant la partie portée par la plante et consommée par l'homme. Par exemple, il est important de noter que pour les légumes-feuilles (salades, épinards *etc.*) l'organe désigne les feuilles ; ce qui implique que le processus de translocation n'est pas considéré pour ces espèces végétales. Différentes paramétrisations proposées dans la littérature pour quantifier la translocation sont répertoriées dans le rapport technique de l'AIEA (IAEA, 2009). Ainsi, le facteur de translocation (F_{trans}) exprime soit :

- Le ratio entre la concentration dans l'organe consommable de la plante à la récolte (Bq/kg) et la concentration dans la plante entière ou dans le système foliaire à la récolte (Bq/kg),
- Le ratio entre la concentration dans les parties consommables de la plante à la récolte (Bq/kg) et la concentration dans le système foliaire au moment du dépôt (Bq/kg) ou l'activité sur les feuilles par mètre carré au dépôt (en Bq/m²), ou
- Le ratio entre l'activité surfacique dans les parties consommables de la plante mesurée à la récolte (en Bq/m²) et l'activité surfacique dans les feuilles mesurée au dépôt (en Bq/m²).

Les différentes paramétrisations utilisées dans la littérature et la variabilité des valeurs répertoriées par l'AIEA ont révélé les difficultés liées à la caractérisation des différents mécanismes décrivant la translocation des radionucléides dans les plantes. La revue réalisée par l'AIEA a surtout souligné le manque de connaissances vis-à-vis de l'influence du stade de développement des plantes sur le facteur de translocation.

3.1.2.4 Remise en suspension

Les particules déposées à la surface du sol peuvent de nouveau intégrer l'air ambiant par l'action d'un écoulement d'air turbulent (Roustan, 2005) pouvant être généré par le vent ou par les agriculteurs ou les animaux en mouvement à l'intérieur des parcelles agricoles ou par l'action de la pluie heurtant le sol (appelée « rain splash »). Ces particules peuvent alors de nouveau se déposer sur les feuilles des plantes et constituent alors une source secondaire de contamination. Ce phénomène correspond à la remise en suspension des particules. Le terme « resuspension » est également rencontré dans la littérature. La remise en suspension des particules déposées sur les surfaces agricoles, et sur tout type de surface en général, est un

processus particulièrement complexe intégrant un grand nombre de mécanismes tels que l'adhérence et la géométrie des particules considérées, la rugosité, l'humidité de la surface de dépôt et le temps écoulé depuis le dépôt (IAEA, 2009 ; Roustan, 2005).

En pratique deux coefficients de transfert sont utilisés pour caractériser la resuspension des polluants : le facteur de resuspension F_{MES} et le taux de mise en suspension T_{MES} . Concernant les radionucléides, le facteur de resuspension, exprimé en m^{-1} , est défini comme le ratio entre la concentration volumique de radionucléides remis en suspension ($Bq.m^{-3}$) et la concentration surfacique de radionucléides déposée ($Bq.m^{-2}$). Le taux de mise en suspension, exprimé par unité de temps (ex : par seconde) désigne le ratio entre le flux de polluant ($Bq.m^{-2}.s^{-1}$) et la concentration surfacique de polluants ($Bq.m^{-2}$). Ce dernier facteur est recommandé par l'AIEA (IAEA, 2009) et par Roustan (2005) dans la modélisation mécaniste du transport de polluants. Garger *et al.* (1999) ont réalisé une revue de la littérature dans laquelle huit modèles empiriques (exponentiels, hyperboliques et/ou combinant les deux approches) modélisant le facteur de resuspension F_{MES} ont été testés sur des données issues de l'accident de Tchernobyl. Roustan (2005) suggère de la prudence dans le choix de l'un ou l'autre modèle du fait de l'impact potentiel de ce choix sur le comportement du dépôt sur le long terme. Pour les milieux ruraux, le modèle hyperbolique proposé par l'IAEA (1992) apparaît comme l'un des plus pertinents selon la comparaison de modèles effectuée dans le cadre du projet BIOMOVIS II (IAEA, 2009).

Du fait de la diversité des mécanismes impliqués, le phénomène de remise en suspension des particules déposées à la surface des sols agricoles et des autres surfaces est très complexe à modéliser. Des modèles dits « théoriques » ont été proposés et ont permis de diviser le phénomène en une étape d'arrachement de la particule à la surface et une étape de transport de la particule dans l'écoulement de proche paroi (Roustan, 2005). Cependant, ces modèles ont semblé rencontrer des difficultés dans la prédiction de la remise en suspension lors de la première partie du phénomène (Roustan, 2005). La modélisation empirique présente également des limites, car ces modèles sont souvent peu pertinents pour des situations autres que celles dans lesquelles ils ont été obtenus (Garger *et al.* 1999 ; Roustan, 2005). On peut de plus noter que la gamme de variation des valeurs de F_{MES} reportées par l'AIEA (IAEA, 2009) est particulièrement étendue couvrant plusieurs ordres de grandeur (de $7 \cdot 10^{-4}$ à $1.2 \cdot 10^{-10} m^{-1}$).

3.1.3 Radionucléides dans les sols et transfert racinaire

Les radionucléides déposés à la surface du sol peuvent migrer au cours du temps dans les couches plus profondes des sols agricoles, contaminer durablement ces sols et constituer une source secondaire de contamination pour les différentes cultures à travers le réseau racinaire (Coughtrey and Thorne, 1983). Le transfert sol-plante est régi par un ensemble de processus physiques, chimiques et biologiques expliquant la distribution et la mobilité des radionucléides dans les sols (par ex : les mécanismes de sorption, d'adhésion (Duc, 2002), et de transport vertical (IAEA, 2009)) et leurs interactions avec le système racinaire des plantes décrivant le transfert vers les plantes (IAEA, 2009). Ces interactions dépendent entre autres des caractéristiques chimiques des radionucléides et de la composition granulométrique et chimique des sols (Coughtrey and Thorne, 1983). Face à la complexité des mécanismes en jeu, des approches empiriques ont généralement été adoptées.

3.1.3.1 Distribution dans le sol

La distribution des radionucléides dans les sols est expliquée par les mécanismes de sorption. La sorption définit l'accumulation de matière sur un solide (*i.e.* à l'échelle atomique, le transfert d'ions ou de molécules de la phase liquide vers la phase solide (Duc, 2002)). Elle est souvent quantifiée à l'aide du coefficient de distribution solide-liquide, K_d (en $L.kg^{-1}$), désignant le rapport à l'équilibre de la concentration du radionucléide « sorbé » sur le solide (en $Bq.kg^{-1}$) sur la concentration du radionucléide restant en solution (en $Bq.L^{-1}$) (IAEA, 2009).

$$K_d = \frac{\text{Concentration en phase solide (Bq.kg}^{-1}\text{)}}{\text{Concentration en phase liquide (Bq.L}^{-1}\text{)}} \quad (3.8)$$

Cette paramétrisation synthétise tous les mécanismes de sorption /désorption mis en jeu sous l'hypothèse d'un échange réversible à l'équilibre entre la phase liquide et la phase solide (IAEA, 2009). Des études mécanistes et phénoménologiques visant à améliorer la caractérisation des mécanismes de sorption sont en cours au sein de l'IRSN avec un travail de thèse lancé en 2013, dans le cadre du projet AMORAD (IRSN, 2015a), sur la modélisation dynamique de la (bio)disponibilité des éléments chimiques dans les sols et le développement

d'une approche comparative modèles-expériences appliquée aux transferts de strontium et de césium dans la rhizosphère.

3.1.3.2 Mobilité dans le sol

La mobilité des radionucléides dans le sol est régie par des processus tels que le transport convectif de l'eau ruisselante, la dispersion due aux variations spatiales des vitesses de convection, le mouvement par diffusion des fluides, les interactions physico-chimiques avec la matrice du sol (IAEA, 2010). Deux approches largement répandues ont été identifiées par l'AIEA (IAEA, 2010) pour modéliser la mobilité des radionucléides dans le sol. Les hypothèses et les équations correspondant aux deux approches ne seront pas détaillées dans la section.

- La première utilise des modèles à compartiments en série. Elle est basée sur la représentation du sol en plusieurs couches superposées (compartiments en série) à l'aide d'équations différentielles (IAEA, 2009). Cette approche consiste à déterminer une constante de migration verticale entre les différents compartiments exprimée en centimètres par ans ($\text{cm}\cdot\text{a}^{-1}$).
- La seconde utilise des modèles de convection-dispersion et repose sur l'estimation des deux paramètres clés de ces modèles, la vitesse de convection ($\text{cm}\cdot\text{a}^{-1}$) et le coefficient de dispersion ($\text{cm}^2\cdot\text{a}^{-1}$). Ces modèles permettent pour un instant t donné de produire un profil vertical de la distribution des radionucléides dans le sol.

L'approche reposant sur les équations de convection-dispersion est de plus en plus répandue selon IAEA (2009) qui recommande son utilisation pour une meilleure caractérisation des processus impliqués. Cependant le nombre limité de données de vitesses de convection et de coefficients de dispersion compilé par l'AIEA pour des écosystèmes précis n'encourage pas vraiment l'utilisation de ces méthodes dans les évaluations de risques impliquant par exemple des écosystèmes tropicaux (IAEA, 2009).

3.1.3.3 Transfert racinaire

Le transfert de radionucléides du sol vers les plantes, correspondant au passage de radionucléides à travers le système racinaire, dépend principalement de la forme chimique des dépôts, des propriétés des radionucléides, du temps écoulé après les dépôts, des propriétés du sol, du type de plantes, des pratiques agricoles (irrigation, utilisation de fertilisants, *etc.*) et des

conditions climatiques (IAEA, 2010). En pratique, le transfert racinaire est quantifié par un facteur de transfert défini comme le ratio entre la concentration dans les plantes (en Bq.kg⁻¹ poids sec) et la concentration (en Bq.kg⁻¹ de masse sèche) dans la couche de sol en contact avec les racines (ou horizon racinaire). Par souci d'homogénéité dans le calcul des facteurs de transfert, une couche de sol standardisée a été définie par l'union internationale de radioécologie (IAEA, 2009 ; IUR, 1992) ; 10 cm pour l'herbe et 20 cm pour les autres espèces végétales. L'AIEA a compilé, à travers ses réseaux d'experts et sur la base d'études réalisées depuis les années 1960, une base de données conséquente de facteurs de transfert racinaire constituant une source de données très intéressante pour les évaluations d'impact environnemental (IAEA, 2009, 2010).

En situation accidentelle, le transfert racinaire intervient principalement sur la contamination sur le long terme, plusieurs mois voire quelques années après la fin des rejets. Dans les premiers jours et les premières semaines suivant la fin des rejets, l'influence du transfert racinaire sur les niveaux de concentrations présents dans les denrées alimentaires est négligeable, l'essentiel de la contamination étant issue du transfert foliaire. L'impact du transfert racinaire sur le long terme est pertinent lorsque l'on s'intéresse aux radionucléides tels que le Cs-134, le Cs-137 et le Sr-90 qui ont des périodes radioactives longues. Cependant, les différentes contre-mesures telles que le labour des surfaces agricoles ou le décapage des sols contaminés peuvent potentiellement limiter cette influence au cours du temps.

3.1.4 Transfert aux produits animaux

Les radionucléides déposés à la surface du sol, des plantes et présents dans l'atmosphère peuvent constituer une source de contamination pour les animaux d'élevage (ex : bovins *etc.*) et les animaux sauvages (ex : sangliers). En effet, ces animaux peuvent être contaminés par deux voies d'exposition principalement : par inhalation et par ingestion (eau, aliments et sols) (Coughtrey and Thorne, 1983 ; IAEA, 2010) ; l'ingestion étant identifiée comme la plus importante voie d'exposition en situation post-accidentelle (IAEA, 2010). Lorsque l'on évalue le transfert de radionucléides aux produits animaux, il est important d'identifier les facteurs influençant les niveaux de concentration dans les rations pour les animaux, et leur transfert vers les tissus et les organes après ingestion.

3.1.4.1 Ingestion des radionucléides

L'incorporation de radionucléides est principalement régie par la présence de contamination dans les régimes alimentaires des animaux d'élevage. L'alimentation des dépend principalement de processus anthropiques à savoir les pratiques zootechniques adoptées par les éleveurs (Calmon, 2009). Cette alimentation s'effectue à partir de différents produits dans des proportions variables suivant le type d'élevage. Les aliments peuvent être produits localement (dans la région même) et être disponibles en consommation immédiate (herbe de pâture) ou différée (fourrages, ensilages), peuvent provenir d'une autre région également contaminée mais à un niveau différent, ou peuvent être importés et ne pas contenir de contamination (Calmon, 2009). L'incorporation de radionucléides dans l'alimentation des animaux domestiques dépend fortement des disponibilités liées essentiellement à la région, au climat, à la saison, et donc aux différents calendriers agricoles.

3.1.4.2 Transfert dans les tissus et dans les organes

Les radionucléides ingérés sont absorbés à travers le tractus gastro-intestinal avant d'entrer dans la circulation sanguine. L'absorption gastro-intestinale est un processus clé définissant la distribution des éléments et les niveaux de contamination dans les différents tissus et organes des animaux d'élevage (IAEA, 2009). Le processus d'absorption gastro-intestinale peut être variable en fonction de l'animal considéré ; ce qui témoigne des différences physiologiques entre animaux, les ruminants ayant un estomac avec quatre compartiments et les non-ruminants étant monogastriques. Une manière simple de quantifier l'absorption gastro-intestinale est de calculer le coefficient d'absorption apparente, A_a , exprimé en pourcentage et désignant le ratio entre la contamination présente dans l'alimentation et ingérée (en Bq.kg^{-1}) par les animaux et la contamination présente dans les matières fécales (en Bq.kg^{-1}) (Mayes *et al.* 1996). Ce facteur ne prend pas en compte la sécrétion endogène de la radioactivité absorbée du corps vers les matières fécales qui peut être importante pour certains radionucléides (IAEA, 2010 ; Mayes *et al.* 1996). Une approche alternative a été proposée par Mayes et Lamb (1989) pour l'étude de l'absorption gastro-intestinale du césium par les ruminants. Cette alternative est le coefficient d'absorption vraie, noté A_t dans la littérature, qui peut être défini comme la fraction de radionucléides ingérés qui est transférée à travers la paroi de l'intestin (Mayes *et al.* 1996). Le coefficient d'absorption vraie est jugé plus pertinent que le coefficient d'absorption apparente par Mayes *et al.* (1996), car il fournit une

meilleure information sur les mécanismes d'absorption, il explique mieux la biodisponibilité du césium. [Mayes et al. \(1996\)](#) ont proposé trois approches (par bilan fécal, par renouvellement du plasma sanguin, par différenciation isotopique entre Cs-134 et Cs-137) permettant d'estimer ce coefficient pour l'absorption du césium en prenant en compte différentes composantes biologiques. Nous ne discuterons pas ces méthodes ici.

Les radionucléides absorbés sont ensuite distribués dans le reste de l'organisme *via* la circulation sanguine. Le transfert de radionucléides aux produits animaux dépend d'un ensemble de facteurs incluant les aspects physiologiques des organismes, et les caractéristiques physico-chimiques des radionucléides. Le transfert de radionucléides aux produits animaux est principalement quantifié suivant deux alternatives utilisant deux types de coefficients ([IAEA, 2009](#)) :

- Le facteur de transfert (*ETF*) défini pour le lait (en jour/L ou en jour/kg) et la viande (en jour/kg) comme le ratio à l'équilibre entre la concentration dans le lait ou la viande et la concentration journalière ingérée par l'animal d'intérêt ([IAEA, 2009](#)). Le facteur de transfert est l'approche majoritairement utilisée depuis son introduction par [Ward et al. \(1965\)](#). Une base de données de facteurs de transfert intégrant entre autres des études consécutives à l'accident de Tchernobyl et des données obtenues en conditions expérimentales a été compilée par l'AIEA ([IAEA, 2009](#)).
- Le facteur de concentrations (*CR* pour « concentration ratio » en anglais) correspond au ratio à l'équilibre entre la concentration dans le produit exprimée en Bq.kg⁻¹ (poids frais) et la concentration dans la ration alimentaire de l'animal considéré exprimée en Bq.kg⁻¹ (poids sec). Il faut noter que le facteur de transfert peut être dérivé à partir du facteur de concentrations en le divisant par la ration journalière (en kg par jour) et inversement on peut dériver le ratio de concentration en le multipliant par la ration journalière ([IAEA, 2009](#)). C'est une approche de plus en plus utilisée dans la communauté en radioécologie ([Beresford et al. 2015](#) ; [Brown et al. 2013](#)). Le facteur de concentration est un paramètre radioécologique « générique » qui peut pour un élément donné être considéré comme commun à différentes espèces animales ([IAEA, 2009](#)) : *e.g.* des valeurs définies pour les taureaux peuvent être utilisées pour les bisons. Ce paramètre peut notamment être utilisé pour des animaux sauvages *via* une extrapolation à partir des facteurs de concentration obtenus pour des animaux domestiques plus largement étudiés ([Beresford et al. 2015](#)).

3.2 Modélisation du transfert et de l'exposition par ingestion

Dans cette section, nous étudierons à travers la modélisation les différents processus décrivant le transfert de radionucléides sur les surfaces agricoles et vers l'homme en situation accidentelle. La modélisation est une démarche scientifique qui vise à représenter toutes sortes de situations d'objets et de structures du monde réel (Delahaye, 2015). C'est une opération par laquelle on établit le modèle d'un système complexe, afin de mesurer les effets sur ce système des variations de tel ou tel de ses éléments composants (Trésor de la Langue Française Informatisé cité par Roustan (2005)). Dans le cadre de la modélisation d'un phénomène physique, un modèle est une représentation abstraite idéalisée de ce phénomène (Sudret, 2007) capable d'en expliquer ou d'en reproduire dynamiquement le fonctionnement (Roustan, 2005) en combinant un ensemble d'hypothèses et de paramètres (Zio and Pedroni, 2013). Dans ce manuscrit, un modèle sera considéré comme étant toute équation ou tout ensemble d'équations mathématiques décrivant un phénomène donné, combinant hypothèses et paramètres et permettant de calculer des variables de sortie à partir de variables d'entrée ; les paramètres pouvant être eux-mêmes modélisés. Les modèles étudiés dans cette thèse sont à visée opérationnelle, i.e. définis relativement à un objectif précis et adaptés de manière à répondre à un besoin (« *fit for purpose* »). En ce sens, nous pouvons voir un modèle comme une caricature fonctionnelle de la réalité. Nous utiliserons indifféremment le terme « modèle » pour les modèles statistiques, mécanistes, déterministes, probabilistes, par équations différentielles.

Dans ce travail la méthodologie s'est construite en se focalisant sur l'herbe et sur les légumes feuilles (choux, épinards et salades *etc.*) ayant directement intercepté le panache radioactif. L'intérêt porté à la contamination de l'herbe vient du fait qu'elle peut constituer durant la phase post-accidentelle une source de contamination *via* la consommation de lait et de viande d'animaux mis au pâturage. Le choix des légumes feuilles s'explique non seulement par la richesse relative des données disponibles (études réalisées à l'IRSN, littérature internationale, mesures réalisées par les autorités japonaises au printemps 2011), mais également parce que la consommation de tels aliments est connue pour contribuer de manière significative à la dose reçue dans les mois qui suivent l'accident (Simpkins, 2005; Tschiersch *et al.* 2009). Les travaux ont porté sur les isotopes radioactifs du césium (Cs-134/-

137), de l'iode (I-131) et du strontium (Sr-90) qui sont généralement majoritaires dans les rejets accidentels de centrales nucléaires (Gonze *et al.* 2014 ; Lelieveld *et al.* 2012). Les modèles servant à prédire les concentrations de radionucléides dans l'herbe, les légumes feuilles et dans les produits animaux tels que le lait et la viande et les modèles, visant à calculer les doses reçues suite à l'ingestion de ces produits, sont exposés dans les sections suivantes. Tous ces modèles sont implémentés dans la plateforme SYMBIOSE dédiée à l'évaluation de l'impact du transfert de radionucléides dans l'environnement et à (cf. Annexe B).

3.2.1 Transfert à l'herbe et aux légumes feuilles

Soit $[C]_{H,F-F}^r(t)$ la concentration dans le système foliaire de l'herbe induite par l'interception du radionucléide r par le système foliaire à la date t (Bq.kg⁻¹ poids frais), $[C]_{H,F-R}^r(t)$ la concentration du radionucléide r dans le système foliaire de l'herbe induite par absorption racinaire de la contamination présente dans le sol à la date t (Bq.kg⁻¹ poids frais) et $[C]_{H,R-F}^r(t)$ la concentration dans le système racinaire de l'herbe induite par la migration de la contamination depuis le système foliaire à la date t . Soient $[C]_{LF,F}^r(t_{Rec})$ et $[C]_{LF,R}^r(t_{Rec})$ les concentrations (Bq.kg⁻¹ poids frais) du radionucléide r dans les légumes feuilles considérés induites par les voies foliaire et racinaire respectivement à la date de récolte t_{Rec} . La contamination de l'herbe à la date t (Bq.kg⁻¹ poids frais), $([C]_H^r(t))$, et la contamination des légumes-feuilles à la date de récolte (Bq.kg⁻¹ poids frais), $([C]_{LF}^r(t_{Rec}))$, sont décrites par les équations suivantes.

$$[C]_H^r(t) = [C]_{H,F-F}^r(t) + [C]_{H,F-R}^r(t) \quad (3.9)$$

$$[C]_{LF}^r(t_{Rec}) = [C]_{LF,F}^r(t_{Rec}) + [C]_{LF,R}^r(t_{Rec}) \quad (3.10)$$

Pour l'herbe et pour les légumes feuilles, les concentrations sont modélisées en tenant compte de la contamination directe par voie foliaire et de la contamination indirecte par voie racinaire.

3.2.1.1 Voie foliaire

Les concentrations du radionucléide r dans l'herbe et dans les légumes feuilles induites par la voie foliaire sont déterminées à partir des équations de bilan de masse prenant en compte le flux du radionucléide r déposé et initialement retenu par les systèmes foliaires de l'herbe et des légumes feuilles, la période radioactive de r et la décroissance biomécanique due aux différents processus environnementaux et à la croissance des plantes considérées. Ces équations de bilan de masse sont données ci-dessous (équation (3.11) pour l'herbe et équation (3.12) pour les légumes feuilles) :

$$\begin{cases} \frac{d}{dt} [C]_{H,F-F}^r = \frac{D^r}{B_H} (K^r \times IF^s + (1 - K^r) \times IF^h(H^p)) - (\lambda^{FL} + \lambda^r) \times [C]_{H,F-F}^r + \frac{\lambda^L}{\beta} \times [C]_{H,R-F}^r \\ \frac{d}{dt} [C]_{H,R-F}^r = \left(\frac{(1-a) \times (\lambda^{FL})^2}{\lambda^L} [C]_{H,F-F}^r \right) - \lambda^r \times \beta \times [C]_{H,F-F}^r - \lambda^L \times [C]_{H,R-F}^r \end{cases} \quad (3.11)$$

$$\frac{d}{dt} [Cs]_{LF}^r = D^r \times (K^r \times MIF^s + (1 - K^r) \times MIF^h(H^p)) - (\lambda^{FL} + \lambda^r) \times [Cs]_{LF}^r \quad (3.12)$$

$$MIF^h(H^p) = (MIF^h(5) - MIF^s) \times \frac{H^p}{5} + MIF^s ; \quad \forall 0 < H^p < 5 \quad (3.13)$$

Les variables D^r et K^r représentent le flux de dépôt exprimé en (Bq.m⁻² par unité de temps) et la proportion de dépôt sec respectivement. La variable B_H désigne la biomasse aérienne de l'herbe (kg.m⁻²) et β représente le ratio entre les biomasses fraîches des systèmes foliaire et racinaire, supposé constant dans le temps et dans l'espace. Le paramètre λ^{FL} correspond à la constante de décroissance biomécanique (« field loss ») exprimée en jour⁻¹ et λ^r est la constante de décroissance radioactive (en jour⁻¹) spécifique à chaque isotope. La quantité λ^L est la constante de décroissance « long terme » (en jour⁻¹) intégrant la décroissance biomécanique et la décroissance induite par les échanges entre les systèmes foliaire et racinaire. Le paramètre a est un pondérateur de croissance (sans unité). Les quantités IF^s et IF^h (sans unité) désignent les facteurs d'interception sèche et humide respectivement et les quantités MIF^s et MIF^h correspondent aux facteurs d'interception massique exprimés en m².kg⁻¹. Les paramètres IF^h et MIF^h dépendent du paramètre H^p exprimant la quantité totale de pluie (en mm) tombée durant les épisodes de dépôt. Dans la plateforme SYMBIOSE, les facteurs d'interception humide $IF^h(H^p)$ et $MIF^h(H^p)$ sont déterminés grâce à une interpolation linéaire entre les facteurs d'interception correspondant aux hauteurs de pluie références fixées à 0, 5, 10 et 15 mm. Par exemple pour une hauteur de pluie H^p comprise entre 0 et 5 mm le facteur d'interception humide massique correspondant est donné

par l'équation (3.13). Pour rappel, dans les deux cas (herbe et légumes feuilles) le modèle n'intègre pas la translocation (cf. [Section 3.1.2.3](#)) car « l'organe » consommé est la feuille.

3.2.1.2 Voie racinaire

Les systèmes racinaires de l'herbe et des légumes feuilles sont modélisés de manière similaire. Dans la suite, nous considérerons donc les mêmes équations pour les deux espèces végétales. Evaluer les concentrations des radionucléides dans l'herbe et les légumes feuilles induites par la voie racinaire nécessite d'étudier les caractéristiques des sols sur lesquels ces cultures sont plantées et les mécanismes décrivant les échanges sols-plantes. Le sol est considéré ici comme un modèle bicouche, *i.e.* une superposition de deux couches accueillant les mécanismes de migration.

- La « couche de surface » (CS) désigne la couche superficielle représentant l'interface entre l'atmosphère et les couches sous-jacentes du sol, et
- La « couche racinaire » (CR) dans laquelle les racines des plantes puisent les nutriments nécessaires à la croissance de celles-ci. CR est composée d'un « horizon disponible » (HD) représentant la partie de l'horizon racinaire contenant la contamination (bio) disponible pour la plante et disponible vis-à-vis du processus de migration verticale. Elle comprend la solution du sol et les particules du sol fixant réversiblement le polluant.

Les mécanismes décrivant la migration des radionucléides à travers les différentes couches du sol définies plus haut sont décrits par les équations (différentielles du premier ordre) suivantes.

$$[C]_{CS}^r = \frac{S_{CS}^r}{\rho \times h_{CS}} \quad (3.14)$$

$$[C]_{HD}^r = \frac{S_{HD}^r}{\rho \times h_{CR}} \quad (3.15)$$

$$\lambda^{Mig} = \frac{1}{h_{CR}} \times \frac{R^p}{\theta + \rho \times Kd} \quad (3.16)$$

$$\lambda^{SMig} = \alpha \times \lambda^{Mig} \quad (3.17)$$

$$\frac{d}{dt} S_{CS}^r = D^r - \lambda^{SMig} \times S_{CS}^r - \lambda^r \times S_{CS}^r \quad (3.18)$$

$$\frac{d}{dt} S_{HD}^r = \lambda^{SMig} \times S_{CS}^r - \lambda^{Mig} \times S_{HD}^r - \lambda^r \times S_{HD}^r \quad (3.19)$$

Les variables $[C]_{CS}^r$ et $[C]_{HD}^r$ désignent les concentrations dans la couche de surface et dans l'horizon disponible respectivement et les quantités S_{CS}^r et S_{HD}^r définissent les stocks dans la couche de surface et dans l'horizon disponible de la couche racinaire (en mol.m^{-2}) respectivement, ρ représente la densité du sol (en kg.m^{-3}), h_{CS} et h_{CR} correspondent aux épaisseurs de la couche de surface et de la couche racinaire, θ (L.m^{-3}) exprime la teneur en eau de couche racinaire, R^p est une donnée météorologique (exprimée en mètres par unité de temps) décrivant la quantité effective d'eau précipitée atteignant le sol, Kd est le coefficient de distribution dans le sol (en L.kg^{-1}), λ^{SMig} représente la constante de migration entre la couche de surface et la couche racinaire (en jour^{-1}) et λ^{Mig} désigne la constante de migration (en jour^{-1}) entre la couche racinaire et les couches sous-jacentes. Ces deux paramètres sont supposés proportionnels entre eux.

Ainsi, la concentration du radionucléide r dans le compartiment racinaire de l'herbe ou des légumes feuilles à la date t (à la récolte pour les légumes feuilles), $[C]_{,R}^r(t)$ (le point représente l'herbe ou les légumes feuilles), est dérivée de la concentration issue de l'horizon disponible comme décrit par l'équation suivante.

$$[C]_{,R}^r(t) = f^{sf} \times TF^{Up} \times [C]_{HD}^r(t) \quad (3.20)$$

Le paramètre TF^{Up} est défini comme le ratio entre la concentration dans les plantes et la concentration dans la couche du sol d'intérêt (en kg poids sec de sol par kg poids sec de biomasse aérienne) et représente le facteur de transfert sol-plante à l'équilibre (IAEA 2010), et f^{sf} désigne le rapport poids sec poids frais (exprimé en kg de poids sec par kg de poids frais de biomasse aérienne). Cette formulation suppose qu'un équilibre est atteint entre la concentration dans le système foliaire et la concentration de l'horizon disponible de la couche racinaire.

3.2.2 Transfert aux produits animaux

La modélisation du transfert aux produits animaux est mise en œuvre dans l'objectif d'évaluer la contamination des produits animaux en proposant un modèle de calcul de l'évolution temporelle de la concentration de radionucléides dans des produits animaux d'origine agricole destinés à l'alimentation humaine. Dans la plateforme SYMBIOSE, la métabolisation du radionucléide ingéré se fait *via* un pool dont les temps caractéristiques de métabolisation sont rapides (pool rapide) et un pool dont les caractéristiques de

métabolisation sont lentes (pool lent). Soit $[C]_{prod,PR}^r(t)$ et $[C]_{prod,PL}^r(t)$ les concentrations dans le produit d'intérêt induites par les pools rapide et lent respectivement. La concentration dans le produit d'intérêt (noté *prod*) à la date t ($[C]_{prod}^r(t)$) est donc déterminée en tenant compte des deux pools métaboliques suivant :

$$[C]_{prod}^r(t) = [C]_{prod,PR}^r(t) + [C]_{prod,PL}^r(t) \quad (3.21)$$

Le transfert de radionucléides dans ces deux pools est décrit par les deux équations différentielles suivantes :

$$\begin{cases} \frac{d}{dt} [C]_{prod,PR}^r = (TC^{Ing} \times \lambda_{PR}^E \times \beta^{PRL} \times ETF_{prod}) - (\lambda_{PR}^E + \lambda^r) \times [C]_{prod,PR}^r \\ \frac{d}{dt} [C]_{prod,PL}^r = (TC^{Ing} \times \lambda_{PL}^E \times (1 - \beta^{PRL}) \times ETF_{prod}) - (\lambda_{PL}^E + \lambda^r) \times [C]_{prod,PL}^r \end{cases} \quad (3.22)$$

La variable TC^{Ing} correspond au flux de radionucléides ingérés (en kg par unité de temps). Le paramètre λ^r désigne la constante de décroissance radioactive (en jour⁻¹) et les paramètres λ_{PR}^E et λ_{PL}^E représentent les taux d'élimination (en jour⁻¹) dans les pools rapide et lent respectivement. Le paramètre β^{PRL} désigne le coefficient de partage métabolique entre les voies lente et rapide. Le paramètre ETF_{prod} correspond au facteur de transfert à l'équilibre entre l'alimentation de l'animal et le produit animal dépendant du radionucléide et du produit animal d'intérêt (en jour.kg⁻¹ ou en jour.L⁻¹).

Le modèle de transfert aux produits animaux est un modèle dynamique ne reposant pas sur une hypothèse d'équilibre instantané entre la concentration dans l'organe et le flux ingéré. Il permet donc d'accéder à des échelles de temps inférieures à celles d'un modèle à l'équilibre. Il présente néanmoins quelques limites. Le modèle ne tient compte ni de l'âge, ni de la durée de vie des animaux. Il ne suit donc pas de cohortes avec des âges différents. La date de naissance est considérée comme antérieure ou égale à la date de démarrage du calcul ; ce qui peut être maximisant du fait de la comptabilisation pour la cohorte d'intérêt des flux ingérés antérieurs à sa naissance. Enfin, la prévision à l'échelle journalière est plus incertaine qu'aux échelles hebdomadaires et mensuelles, du fait notamment de la non prise en compte des classes d'âge.

3.2.3 Exposition des populations par ingestion

Un des aspects de l'évaluation des risques en situation post-accidentelle consiste à prédire la dose potentiellement reçue l'ingestion de ces denrées contaminées en estimant la quantité de radionucléides ainsi ingérée par jour. La dose efficace due au radionucléide r ($DE_i^r(t)$) cumulée à la date t pour un individu d'une classe d'âge i donnée est calculée suivant l'expression ci-dessous :

$$DE_i^r(t) = DE_i^r(t-1) + \sum_{j=1} [C]_j^r(t) \times IR_j^i \times DC_i^r \quad (3.23)$$

La quantité $[C]_j^r(t)$ est la concentration (en Bq.kg⁻¹ poids frais) du radionucléide r prédite dans le produit alimentaire j composant la ration alimentaire. La quantité IR_j^i correspond à la consommation journalière du produit j (en kg par jour). Dans la plateforme SYMBIOSE des valeurs par défaut correspondant aux standards français sont appliqués (Favier *et al.* 1995; Bertrand, 1993). Le paramètre DC_i^r est le coefficient de dose efficace (exprimé en Sv.Bq⁻¹) permettant de convertir la contamination ingérée en dose (exprimée en Sv). Les valeurs du coefficient de dose utilisées dans SYMBIOSE sont basées sur celles proposées au niveau international par la Commission Internationale pour la Protection Radiologique (ICRP, 1996).

Box 3.1 : Notes aux lecteurs

- Dans ces travaux de thèse nous nous sommes concentrés sur :
 - ❖ le transfert de radionucléides à l'herbe de pâturage, aux légumes feuilles et au lait et à la viande.
 - ❖ l'exposition humaine par consommation de denrées contaminées durant la phase post-accidentelle allant de quelques jours à quelques mois.
- Les modèles développés dans ce travail concernent principalement :
 - ❖ les facteurs d'interception sèche et humide aux **Chapitres 6 et 7** respectivement.
 - ❖ la constante de décroissance biomécanique au **Chapitre 8**.
- Les autres processus présentés dans ce chapitre n'ont pas fait l'objet de développements spécifiques. Les modèles définis dans la plateforme SYMBIOSE et les valeurs par défaut de SYMBIOSE sont considérées. Ce choix est discuté au **Chapitre 10**.
 - ❖ Nous avons défini la quantité de radionucléides déposée comme notre donnée d'entrée. Les modèles de calculs de **dispersion atmosphérique** ne sont donc pas considérés.
 - ❖ La **translocation** n'est pas considérée car ce sont les feuilles des plantes interceptant directement le panache qui sont consommées.
 - ❖ La **resuspension** et le **transfert sol-plantes** sont pertinents sur le long terme (environ à partir d'un an et plus). Dans ce travail nous nous intéressons à la phase post-accidentelle (i.e. quelques semaines, ou mois après l'accident). Dans les applications présentées aux **Chapitre 8 et 9**, le sol est modélisé par le modèle bicouche défini dans la **Section 3.2.1.2**.
 - ❖ Pour le **transfert aux produits animaux** les facteurs de transfert à l'équilibre sont définis à partir de [Simon-Cornu et al. \(2015\)](#).

4. Incertitude et variabilité en radioécologie

Dans l'application des premières mesures de protection des populations, durant une crise, les décideurs souffrent le plus souvent de contraintes de temps, du manque de données et d'une compréhension imprécise des processus en action ; ce qui peut affecter considérablement la qualité des évaluations de risque (FAO/WHO, 2011). Il est alors important que les acteurs de la gestion de crise aient pris conscience au préalable (*i.e.* hors période de crise) de l'existence, de la nature des principales sources des incertitudes impactant la fiabilité des modèles de prévision (IAEA, 1989). Deux concepts sont principalement dans la terminologie généralement adoptée dans la littérature (Vose, 2008): l'*incertitude* reflétant une connaissance imprécise ou généralement un manque de connaissance autour du phénomène étudié et la *variabilité* caractérisant l'hétérogénéité naturelle d'origine physique, chimique, biologique *etc.* inhérente au phénomène d'intérêt.

4.1 Incertitude et variabilité : De quoi parle-t-on ?

Les définitions généralement associées aux notions d'*incertitude* et de *variabilité* restent relativement vagues et varient subtilement d'une terminologie à une autre (Simon-Cornu *et al.* 2015) ou d'un domaine d'application à un autre.

Dans le contexte général de l'évaluation des risques, un certain nombre de définitions ont été considérées pour caractériser le mot « incertitude » et les implications liées à ces définitions (EFSA, 2015). Le comité scientifique de l'EFSA s'est récemment intéressé, à travers un document de référence sur l'incertitude dans les évaluations scientifiques (EFSA, 2015), à un ensemble de formalismes adoptés pour définir l'*incertitude* dans le contexte de l'évaluation des risques. L'*incertitude* est définie par l'Organisation Mondiale de la Santé (OMS) et le Programme International sur la Sécurité des Substances Chimiques (PISC), dans le cadre des risques liés aux substances chimiques, comme la connaissance imparfaite sur l'état présent ou futur d'un organisme, d'un système ou d'une (sous-) population d'intérêt (IPCS, 2004 ; WHO/IPCS, 2008). De manière similaire, l'EFSA (2011) définit l'*incertitude* comme « l'incapacité de déterminer le véritable état des composantes d'un système ». Des définitions faisant état de l'absence d'information ou d'une information incomplète (NRC,

2009), de l'expression d'une connaissance « inadéquate » (SCHER, SCENIHR and SCCS, 2013) ou du manque de connaissance (EFSA, 2012 ; IAEA, 1989) résultant de limitations dans les connaissances scientifiques (EFSA, 2006) ont également été données dans le contexte des évaluations des risques (alimentaires et non-alimentaires).

L'*incertitude* peut provenir de différentes sources et peut ainsi être de plusieurs types. Cullen et Frey (1999) définissent, dans le cadre de l'évaluation de l'exposition aux polluants chimiques, deux types d'incertitude : l'*incertitude de modèle* et l'*incertitude* sur les entrées du modèle appelée *incertitude paramétrique*. L'*incertitude de modèle* représente le manque de connaissance sur l'adéquation d'un modèle donné à un scénario donné. L'*incertitude de modèle* peut faire référence à l'existence d'hypothèses alternatives plausibles pour expliquer le phénomène étudié (Zio and Pedroni, 2013). L'*incertitude paramétrique* correspond au manque de connaissance ou à une connaissance imprécise sur les valeurs des paramètres et variables du modèle. L'OMS et le PISC ont identifié un troisième type d'*incertitude* appelé *incertitude de scénario* (WHO/IPCS, 2008) qui caractérise l'incertitude liée à la définition du scénario d'exposition (Kettler et al. 2015). L'EFSA (2008) a défini une classification à 5 catégories d'*incertitude* en identifiant deux types d'incertitude supplémentaires : l'*incertitude sur l'objectif de l'évaluation* et l'*incertitude liée aux performances de l'évaluation*. Ces catégories additionnelles d'*incertitude* sont principalement liées à la définition de l'objectif visé par l'évaluation de l'exposition et aux résultats de l'évaluation de l'exposition au regard de l'objectif visé.

Plusieurs auteurs (IAEA, 1989 ; Cullen and Frey, 1999 ; Kirchner and Steiner, 2008) s'accordent sur une conception générale de la *variabilité* et la présentent, comme l'hétérogénéité inhérente à un phénomène physique, chimique, biologique *etc.* du fait de fluctuations temporelles, spatiales ou encore interindividuelles (ex : le poids et/ou la taille des français ayant entre 17 et 25 ans). Il est cependant commun qu'il existe des divergences sur la manière effective de définir la notion de *variabilité* et de l'aborder dans les analyses. En effet, Cullen et Frey (1999) voient la *variabilité* comme la « vraie » dispersion propre au cas d'étude et ne la considèrent pas comme contribuant à l'incertitude alors que dans plusieurs études, (IAEA, 1989 ; Kirchner and Steiner, 2008), elle est considérée comme une catégorie additionnelle d'*incertitude*. Toutes ces définitions convergent vers un formalisme caractérisant l'ensemble des variations autour des réalisations d'un phénomène, irréductibles même avec l'apport de nouvelles informations. Il appartient au modélisateur, dans de

nombreux cas, de spécifier si une dispersion reflète de l'*incertitude* ou de la *variabilité* tant la différence entre ces deux concepts peut quelque fois être ténue (Simon-Cornu *et al.* 2015).

Dans les développements présentés dans la suite la *variabilité* correspondra à la « vraie » dispersion propre à nos paramètres au sein du système (ou scénario) étudié (pour éviter la confusion la source de *variabilité* sera précisée si elle est connue précisément ; celle-ci pouvant être spatiale, temporelle, interindividuelle, inter-espèce ou intra-espèce, *etc.*). Le formalisme adopté dans cette étude pour la notion d'*incertitude* fait référence à l'*incertitude épistémique* sur les scénarios d'exposition, sur les modèles radioécologiques, sur leurs hypothèses et sur les différents paramètres des modèles. Bien que de nombreuses recommandations internationales préconisent de couvrir exhaustivement toutes les sources d'incertitude (FAO/WHO, 2011 ; WHO/IPCS, 2008), l'*incertitude paramétrique* est la seule à faire l'objet d'un intérêt approfondi avec le développement de méthodes plus ou moins sophistiquées (Zio and Pedroni, 2013).

4.2 Incertitude et variabilité dans les modèles radioécologiques d'évaluation de risque

Nous nous sommes intéressés dans le cadre de ces travaux de thèse à la modélisation de l'évaluation de la dose reçue par les populations humaines en situation accidentelle suite à l'ingestion de denrées agricoles contaminées. La modélisation de l'évaluation de la dose ingérée peut faire penser à la construction d'un pont avec les éléments constitutifs qui s'imbriquent au fur et à mesure que la construction avance. En effet, cette modélisation fait intervenir des modèles permettant d'étudier le transfert de radionucléides sur les parcelles agricoles, puis vers les animaux domestiques et enfin vers l'homme. Chaque maillon (modèle) de cette chaîne (alimentaire) est sujet à un ensemble de sources d'*incertitude* de différents types et de *variabilité*.

4.2.1 Séparation Incertitude/ Variabilité

Un certain nombre de recommandations sont émises dans la littérature en vue de définir une démarche méthodologique adaptée dans le cadre d'une analyse d'incertitude en évaluation des risques (Frey and Rhodes, 1998). Ce projet de recherche soulève un ensemble d'enjeux

méthodologiques liés notamment à la quantification des incertitudes à travers différentes méthodes d'analyse et de traitement de données. L'un de ces enjeux concerne les questions théoriques et pratiques liées à la distinction entre *incertitude* et *variabilité*. Il est par exemple suggéré par plusieurs auteurs (Frey and Rhodes, 1998 ; Hoffman and Hammonds, 1994 ; Nauta, 2000 ; EFSA, 2015) de séparer l'*incertitude* et la *variabilité* car, comme l'explique Vose (2008), cela est plus correct d'un point de vue mathématique, produit une estimation raisonnable de l'incertitude totale et nous renseigne sur la part de l'incertitude totale due à la variabilité ou à l'incertitude (Kirchner and Steiner, 2008). De plus, la séparation entre incertitude et variabilité revêt une certaine importance quand on s'intéresse à une (sous-) population spécifique (Frey, 1993 ; Frey and Rhodes, 1998) : le risque lié au méthyle-mercure plus important pour les forts consommateurs de poisson par exemple. On retrouve dans la littérature, depuis quelques années, quelques applications séparant *incertitude* et *variabilité* avec l'utilisation de méthodes statistiques avancées (ex : réseaux bayésiens hiérarchiques, méthode de Monte Carlo à deux dimensions, etc.) (e.g. Pouillot et al. 2003; Crépet, 2007; Pouillot et al. 2007; Pouillot and Delignette-Muller, 2010). De telles applications sont relativement récentes dans le domaine de l'évaluation de risque radiologique mais se développent peu à peu (Jang et al. 2009; Simon-Cornu et al. 2015). La prise en compte des incertitudes de scénario et de modélisation représente également un enjeu intéressant dans cette thèse car elles pourraient contribuer bien plus que l'*incertitude* paramétrique à l'*incertitude* totale (Draper et al. 1999; Linkov and Burmistrov, 2003).

4.2.2 Sources d'incertitude et de variabilité dans les modèles radioécologiques

L'objectif de cette section est d'identifier les différentes sources d'*incertitude* et de *variabilité* intervenant dans les modèles d'évaluation de risque considérés et introduits dans la Section 3.2. Il s'agira par la suite d'identifier les catégories d'*incertitude* dans lesquelles chacune des sources peut trouver sa place. Cette section est le lieu permettant d'établir le formalisme adopté dans ce projet de recherche et me permet également de définir le cadre d'application dans lequel je me suis inscrit durant ces travaux de thèse. Il est important de noter que certaines sources d'*incertitude/variabilité* peuvent appartenir à plusieurs catégories (Kettler et al. 2015 ; WHO/IPCS, 2008). Le tableau suivant résume les sources identifiées.

Tab. 4.1 : Sources d'*incertitude* et de *variabilité* potentielles identifiées dans les modèles d'évaluation du transfert de radionucléides vers l'homme par ingestion de denrées contaminées.

Types d' <i>incertitude</i> et <i>Variabilité</i>	Sources d' <i>incertitude</i> et <i>variabilité</i>
<ul style="list-style-type: none"> ▪ <i>Incertitude de modèle</i> 	<ul style="list-style-type: none"> - Hypothèses simplificatrices et conservatives et prise en compte ou non de certains processus (ex : resuspension) - Utilisation des facteurs de transfert
<ul style="list-style-type: none"> ▪ <i>Incertitude de scénario</i> 	<ul style="list-style-type: none"> - Terme source - Nature des dépôts (sec ou humide) - Conditions climatiques et météorologiques (ex : précipitations, vent <i>etc.</i>) - Pratiques agricoles et zootechniques (ex : culture sous serres, ensilage, <i>etc.</i>) - Types de culture (ex : cultures maraichères ou grandes cultures, <i>etc.</i>)
<ul style="list-style-type: none"> ▪ <i>Incertitude paramétrique</i> 	<ul style="list-style-type: none"> - Toutes informations imprécises, incomplètes sur les données d'entrée et les paramètres du modèle <i>etc.</i>
<ul style="list-style-type: none"> ▪ <i>Variabilité</i> 	<ul style="list-style-type: none"> - Variabilité inter et intra espèces végétales - Différences physico-chimiques entre radionucléides - Caractéristiques physiologiques et métaboliques entre individus - Différences interindividuelles de comportement alimentaire - Variabilité spatio-temporelle

4.2.2.1 Incertitudes de modèle

L'*incertitude de modèle* fait référence comme défini plus haut à la bonne adéquation entre le phénomène étudié et la représentation abstraite qui en est faite par le modèle. Notons qu'une distinction est faite entre *incertitude de modèle* et les écarts observés entre le modèle et les réalisations du phénomène étudié (*erreurs de modèle*). L'*incertitude de modèle* intervient en amont de la démarche de modélisation *i.e.* lors de la définition des concepts et des hypothèses. Dans notre cas précis, elle peut provenir d'une représentation trop simplifiée des mécanismes et des processus entrant en jeu du fait d'hypothèses de modélisation simplificatrices. Dans la pratique, les processus physiques, chimiques et biologiques décrivant les transferts de substances radioactives dans les parcelles agricoles, vers les produits animaux et vers l'homme sont représentés quantitativement à l'aide de facteurs de transfert. Ces facteurs de transfert sont une source d'*incertitude* (de modèle) car ils représentent de façon empirique un ensemble de mécanismes complexes et ne permettent pas forcément de saisir les différentes fluctuations issues de ces mécanismes ou de leurs interactions (Hinton *et al.* 2013). Parmi les hypothèses simplificatrices, on peut noter l'hypothèse d'équilibre sur laquelle reposent les facteurs de transfert et qui très souvent n'est pas vérifiée dans la réalité. En guise d'exemple on peut également noter la représentation des sols agricoles en une ou deux couches plus ou moins épaisses et superposées. Le choix de l'une ou de l'autre configuration (une ou deux couches) peut générer une incertitude supplémentaire.

4.2.2.2 Incertitudes de scénario

Les modèles considérés sont composés d'un ensemble de paramètres dépendant plus ou moins fortement du scénario accidentel. On peut néanmoins distinguer des paramètres dits « contextuels » qui font exclusivement référence aux conditions se produisant durant l'accident et durant la phase post-accidentelle et des paramètres dits « faiblement scénario-dépendants » exportables dans le cadre d'un autre scénario accidentel (ex : de Tchernobyl à Fukushima). Le terme source (chronique et quantités rejetées) est l'une des premières sources d'*incertitude* contextuelles pouvant significativement impacter les évaluations de risque car il définit ce qui peut atteindre la surface du sol et le couvert végétal (Bedwell *et al.* 2015 ; Bocquet, 2013 ; Mallet *et al.* 2007). La nature des dépôts définie par les conditions climatiques et météorologiques au cours des épisodes de dépôt font également partie des sources contextuelles d'*incertitude*. La caractérisation des sources d'*incertitude* contextuelles

est complexe car il est difficile de connaître exactement les conditions météorologiques durant l'accident ou de savoir par exemple si un nuage a lessivé le panache radioactif. Les pratiques agricoles et zootechniques locales peuvent être également considérées comme une source d'*incertitude de scénario* car elles peuvent varier d'une localité à une autre du fait par exemple des conditions climatiques. L'agriculture sous serres peut par exemple être privilégiée dans les territoires où il fait froid ce qui peut limiter la quantité de radionucléides déposée sur les cultures.

4.2.2.3 Incertitudes paramétriques

L'incertitude paramétrique est omniprésente dans les modèles étudiés car elle accompagne tout modèle faisant intervenir des paramètres dont les valeurs sont connues plus ou moins précisément. De plus l'*incertitude* de scénario et l'*incertitude* de modèle peuvent être caractérisées *via* l'incertitude paramétrique car celles-ci sont souvent matérialisées dans les modèles par le biais de paramètres additionnels. La proportion de dépôt sec, qui est un paramètre contextuel caractérisant la nature du dépôt, est représentée par des valeurs comprises entre 0 et 1, la valeur 1 étant associée à un dépôt intégralement sec. Le choix de modéliser le transfert aux produits animaux par un facteur de transfert à l'équilibre peut être vu comme un choix de modélisation simplifiée générant une source d'*incertitude* de modèle.

4.2.2.4 Variabilité

La variabilité, dans les modèles considérés, reflète le caractère fluctuant des phénomènes étudiés au cours du temps ou dans l'espace. Elle permet de caractériser les différences entre éléments d'un même système. Par exemple, les différences entre espèces végétales ou les différences entre plantes d'une même espèce (du fait de la pilosité des feuilles, de leur structure, de l'orientation des plantes *etc.*) peuvent influencer sur l'interception et le devenir des substances radioactives. Ces substances radioactives peuvent elles-mêmes être la source d'une variabilité expliquée principalement par leurs caractéristiques physico-chimiques (ex : formes particulaire ou gazeuse, cations ou anions *etc.*). Les radionucléides peuvent agir de manière variable en fonction des caractéristiques physiologiques variables entre individus d'une (sous-)population, ou d'une cohorte animale ; ce qui traduit la variabilité interindividuelle. Les variations des coefficients de dose entre classes d'âge illustrent bien cette variabilité interindividuelle.

Box 4.1 : Notes aux lecteurs

- Dans ces travaux nous nous sommes concentrés sur la caractérisation de l'*incertitude* et de la *variabilité* paramétriques sur :
 - ❖ le facteur d'**interception** sèche et humide dans les **Chapitres 6 et 7**.
 - ❖ la constante de **décroissance biomécanique** dans le **Chapitre 8**.
- L'*incertitude* et la *variabilité* paramétriques sur les paramètres ci-dessous sont caractérisées principalement par [Simon-Cornu et al. \(2015\)](#) sur la base des données collectées par l'AIEA (IAEA, 2009, 2010).
 - ❖ Coefficient de **distribution** dans le sol
 - ❖ Facteur de **transfert sol-plantes**
 - ❖ Facteur de **transfert aux produits animaux** à l'équilibre
- L'*incertitude* et la *variabilité* sur les processus suivants ne sont pas caractérisées. Nous avons cependant conscience que ces sources d'incertitudes sont généralement influentes dans les évaluations de risque.
 - ❖ **Resuspension**
 - ❖ Paramètres de **dispersion atmosphérique**
- L'*incertitude* de *scénario* sur les paramètres contextuels et l'incertitude de modèle ne sont pas spécifiquement caractérisées. Des pistes de développements futurs sont proposées dans le **Chapitre 10**.

5. Analyse des sources d'incertitude et analyse de sensibilité

5.1 Introduction et démarche générale

L'*incertitude* et la *variabilité*, telles que définies dans la section précédente, sont des composantes inévitables dans la modélisation du comportement d'un système (Zio and Pedroni, 2013). La caractérisation, la représentation, la propagation et l'interprétation appropriées de l'*incertitude* et de la *variabilité* constituent des éléments fondamentaux de l'analyse des risques pour tout système (Zio and Pedroni, 2013). Les décideurs doivent être informés non seulement des connaissances scientifiques disponibles, mais également des incertitudes pertinentes dans la base de connaissances (van Leeuwen and Vermeire, 2007). Dans le cadre de l'analyse des risques environnementaux, sanitaires et alimentaires, plusieurs organismes internationaux tels que l'OMS, la FAO, la Commission du Codex Alimentarius et l'AIEA recommandent l'analyse quantitative (ou à défaut qualitative) des sources d'*incertitude* en vue d'obtenir des évaluations de risque plus réalistes (Codex Alimentarius, 2007 ; FAO/WHO, 2011 ; IAEA, 1989). Les méthodes d'analyse d'incertitude et d'analyse de sensibilité peuvent être implémentées suivant la démarche générique établie par De Rocquigny *et al.* (2006a, b) et illustrée Fig. 5.1.

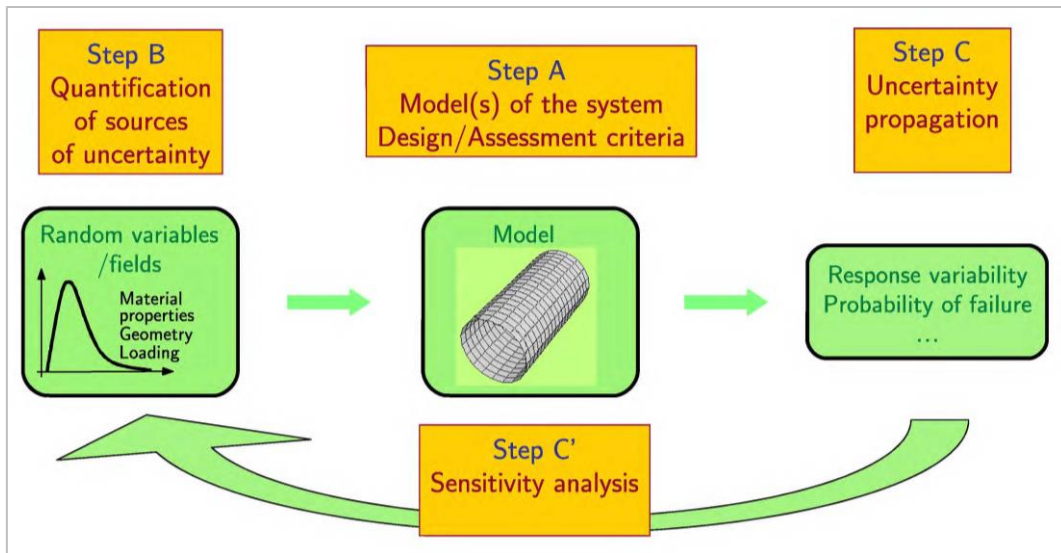


Fig. 5.1 : Démarche générale d'analyse d'incertitude et d'analyse de sensibilité. Adapté de (Sudret, 2007).

Avant de décrire ces notions plus en détails, plaçons-nous dans un cadre de modélisation plus formel. Nous avons admis dans la [Section 3.2](#) que le terme « modèle » devra faire référence dans la suite du manuscrit à toute équation ou tout ensemble, voire système, d'équations mathématiques combinant hypothèses et paramètres dans le but de représenter conceptuellement un phénomène (physique, chimique, biologique *etc.*) et permettant de calculer des variables de sortie à partir de variables d'entrée. De plus, rappelons que les paramètres d'un modèle peuvent également être déterminés à l'aide de modèles. Dans ce contexte considérons qu'un modèle est défini par le triplet $(X, g(\cdot), Y)$ où $g(\cdot)$ est une fonction représentant la structure mathématique du modèle, X correspond aux n variables d'entrée et Y aux m variables de sortie (Fig. 5.2). Par souci de simplicité nous considérerons une seule sortie, *i.e.* $m = 1$.

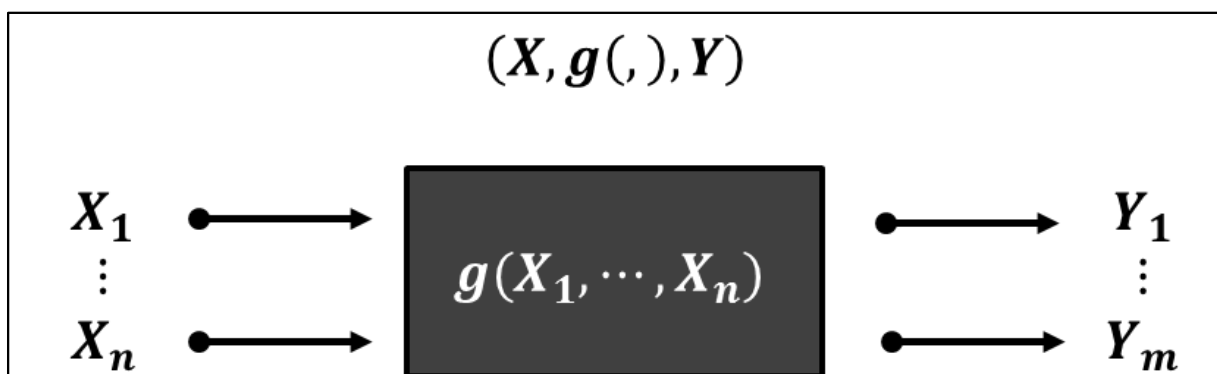


Fig. 5.2 : Illustration conceptuel d'un modèle

Une analyse d'*incertitude* permet principalement de répondre à la question (Monod *et al.* 2006): quel est le niveau d'incertitude de la sortie Y du modèle $g(.)$ induit par l'incertitude sur les entrées X ? Dans le cadre des modèles d'évaluation de risque, les principales étapes d'une analyse d'*incertitude* sont d'identifier et de caractériser les principales sources d'*incertitude* sur les entrées susceptibles d'affecter l'analyse et d'évaluer quantitativement leur impact sur la sortie des modèles en proposant une représentation de l'*incertitude* sur la sortie. Les analyses d'*incertitude* sont, de manière classique, mises en œuvre dans un cadre probabiliste particulièrement adapté pour apporter des réponses pertinentes aux questions posées (Da Veiga, 2007) ; ce qui passe par la définition de distributions de probabilité pour caractériser l'incertitude sur les entrées et la sortie des modèles. Dans ce contexte, lorsque l'on souhaite conduire une analyse d'incertitude, nous pouvons identifier trois étapes principales au lieu des quatre définies par Monod *et al.* (2006) ; les 2^{ème} et 3^{ème} étapes identifiées par Monod *et al.* (2006) pouvant être regroupées.

- La 1^{ère} étape, la quantification de l'*incertitude*, correspond à l'étape B du graphe de Sudret (2007), qui suit la phase de conception du modèle (étape A cf. Fig. 5.1). Elle consiste, dans le cadre probabiliste, à définir les distributions de probabilités caractérisant l'incertitude sur les entrées du modèle. C'est une étape cruciale dans l'analyse (Favre *et al.* 2013). Elle peut être complexe car très souvent très peu d'informations sont disponibles (Sudret, 2007). On peut alors recourir à des données issues d'études antérieures ou à un avis d'expert (Oakley and O'Hagan, 2007).
- La 2^{ème} étape, la propagation de l'*incertitude*, utilise, dans le cadre probabiliste, des méthodes de simulation stochastique (Monte Carlo, Hypercubes Latins, *etc.*) (IAEA, 1989). La principale difficulté à ce niveau repose sur les temps de calcul qui peuvent être particulièrement longs dans certains cas.
- La 3^{ème} étape désigne l'étape de caractérisation de l'incertitude sur la sortie du modèle. A ce niveau on cherche à représenter mathématiquement ou graphiquement l'incertitude sur la sortie.

Le terme analyse de sensibilité est utilisée dans beaucoup de domaines pour répondre à des questions associées à des problèmes, des méthodes et des disciplines quelque peu différents (Saltelli *et al.* 2004). Selon Monod *et al.* (2006), l'analyse de sensibilité permet, à la différence de l'analyse d'incertitude, de répondre à la question : quelles sont les principales sources d'incertitude parmi les différentes entrées incertaines ? Elle permet donc de quantifier l'impact de la variation de chaque entrée incertaine sur la variation de la sortie (Da Veiga,

2007). L'intérêt de l'analyse de sensibilité est multiple car elle constitue à la fois une aide à la validation d'un code de calcul, à la justification de la représentation proposée d'un système (en évaluant à la conception du modèle son « identifiabilité » (Dobre, 2010) et la robustesse des sorties) ainsi qu'à l'orientation des efforts de recherche (Faivre *et al.* 2013). Saltelli *et al.* (2004) ont identifié quatre objectifs fondamentaux d'une analyse de sensibilité.

- Hiérarchisation des entrées ("*Factors Prioritisation*"). L'objectif est d'identifier les entrées influentes et de les classer.
- Identification des entrées non influentes ("*Factors Fixing*"). Le but est de simplifier le modèle sans altérer les prédictions en fixant les entrées non influentes.
- Réduction de Variance ("*Variance Cutting*"). Cet objectif consiste à déterminer les entrées qui peuvent être fixées sans réduire la quantité d'information disponible.
- Cartographie des entrées (Iooss, 2009) ("*Factors Mapping*"). Cet objectif correspond à l'identification des entrées influentes suivant les réalisations de la sortie Y dans des régions d'intérêt.

L'analyse d'incertitude et l'analyse de sensibilité constituent des démarches scientifiques précieuses dans les processus de prise de décision en analyse de risque (industriel, environnement, alimentaire, *etc.*) car elles permettent entre autres de renseigner le degré de confiance placée dans les évaluations de risque. Dans la suite de cette section, nous introduisons les approches adoptées dans le cadre de la quantification de l'incertitude, les méthodes utilisées pour la propagation des incertitudes et les méthodes de calcul d'indices de sensibilité.

5.2 Quantification des incertitudes

La quantification (ou caractérisation) des incertitudes est une étape clé dans la démarche générale d'analyse quantitative des incertitudes qui consiste à proposer une représentation mathématique de l'incertitude sur les entrées du modèle. La quantification des incertitudes connaît un intérêt croissant avec des développements de plus en plus nombreux dans des domaines aussi divers que variés tels que l'industrie, les technologies de l'information et de la communication, la microbiologie alimentaire, l'écologie, l'économie et la finance *etc.* (Bannör and Scherer, 2014 ; EFSA, 2015 ; Kann and Weyant, 2000 ; Oberkampf *et al.* 2004 ; Zio and Pedroni, 2013). Des approches utilisant des paradigmes différents ont été proposées dans la littérature (Sudret, 2007). Cependant, cette étape repose, quel que soit le paradigme

adopté, sur la définition d'une distribution ou d'une mesure de l'incertitude décrivant l'ensemble des valeurs possibles ou probables des entrées du modèle et sur le besoin d'exploiter le maximum d'information disponible (De Rocquigny, 2009). Ainsi, la principale difficulté réside dans la quantité d'information disponible. Cette information peut provenir de différentes sources incluant des observations directes du paramètre d'intérêt, l'avis d'un expert, des arguments physiques (ex : si c'est une proportion on sait que ces valeurs sont définies entre 0 et 1) ou par observations indirectes *i.e.* par modélisation (De Rocquigny, 2009). Dans la suite, les approches généralement utilisées sont présentées.

5.2.1 Approche probabiliste

La théorie des probabilités constitue l'outil mathématique traditionnellement utilisé pour exprimer, caractériser et analyser les différentes sources d'incertitude dans les évaluations de risque (Zio and Pedroni, 2013). Selon Winkler et Hays (1975) (cité par IAEA (1989)), la théorie des probabilités peut être considérée comme le langage mathématique de l'*incertitude*. L'approche probabiliste offre le cadre théorique et les outils techniques appropriés pour la prise en compte d'informations plus ou moins précises provenant de sources différentes. Une distribution de probabilité décrit le comportement d'un paramètre incertain, renseigne son domaine de définition et la probabilité que ce paramètre a de prendre telle ou telle autre valeur (Caniou, 2012). Cette dernière caractéristique fait référence à la réalisation aléatoire d'un événement ; ici que le paramètre soit égal à une valeur donnée ou soit inclus dans une gamme de valeurs. Deux interprétations de la théorie des probabilités sont utilisées en pratique : l'interprétation basée sur les fréquences relatives appelée approche fréquentiste et l'interprétation subjective appelée approche bayésienne (Zio and Pedroni, 2013). Avant d'introduire ces deux notions quelques concepts de la théorie des probabilités sont rappelés.

5.2.1.1 Rappels théorie des probabilités

Une introduction aux principaux concepts de base de la théorie des probabilités est fournie ici, en appui aux développements décrits plus loin sur la quantification des incertitudes et l'analyse de sensibilité. Pour une introduction complète plus détaillée, on pourra se référer à Saporta (2006).

Soit (Ω, \mathcal{F}, P) un espace probabilisé où Ω désigne l'univers des possibles, \mathcal{F} est une σ -algèbre sur Ω et P correspond à une mesure de probabilité attribuant à tout événement $A \subset \Omega$ la probabilité $P(A) \in [0,1]$ de se réaliser. Une mesure de probabilité a les propriétés suivantes :

- $P(\Omega) = 1$ et $P(\emptyset) = 0$
- $P(\bar{A}) = 1 - P(A)$; \bar{A} désignant l'événement contraire
- $A \subset B \Rightarrow P(A) \leq P(B)$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- Soit B_i un système complet d'événements alors $\forall A, P(A) = \sum_i P(A \cap B_i)$

5.2.1.1.1 Conditionnement et indépendance

Considérons B un événement de probabilité non nulle. La probabilité conditionnelle de A sachant B (notée $P(A|B)$) désigne la probabilité de réalisation de l'événement A conditionnellement à la réalisation de l'événement B . Elle est définie comme suit :

$$P(A|B) = \frac{P(A \cap B)}{P(B)}; P(B) \neq 0 \quad (5.1)$$

On peut alors introduire la notion d'indépendance entre événements qui désigne la réalisation d'un événement indépendamment de la réalisation d'un autre événement. Ainsi l'événement A est dit indépendant de l'événement B si la connaissance de B n'affecte pas la réalisation de A *i.e.* $P(A|B) = P(A)$. L'indépendance peut aussi se traduire par le fait que la probabilité de réalisation mutuelle de deux événements résulte du produit des probabilités de réalisations des deux événements, *i.e.* $P(A \cap B) = P(A) \times P(B)$.

5.2.1.1.2 Variables (et vecteurs) aléatoires

Le concept de variable aléatoire formalise la notion de grandeur variant selon le résultat d'une expérience à caractère aléatoire (Saporta, 2006). Elle représente donc une grandeur mathématique prenant différentes valeurs (ou différents états) de manière imprévisible avec une certaine probabilité d'occurrence. Une variable aléatoire, que nous notons X est une fonction associant à chaque événement élémentaire de l'univers des possibles $\omega \in \Omega$ une valeur réelle comme suit :

$$X: \omega \rightarrow X(\omega) \in D_X \subset \mathbb{R} \quad (5.2)$$

D_X désigne le domaine de définition de X *i.e.* représente l'ensemble des valeurs (ou états) que X peut prendre après réalisation de l'expérience.

On peut distinguer différents types de variables aléatoires.

- Binaires : ne peuvent prendre que deux états ;
- Catégorielles qualitatives : peuvent prendre un nombre fini d'états possibles ;
- Catégorielles ordonnées : peuvent prendre un nombre fini d'états possibles naturellement ordonnés ;
- Discrètes : prennent un nombre dénombrable (fini ou infini) de valeurs (ex : nombres de client arrivant à un guichet dans une journée) ;
- Continues : prennent un nombre indénombrable de valeurs réelles possibles (ex : Tailles des individus).

Notons qu'un vecteur aléatoire de taille n sera donc tout n -uplet (X_1, \dots, X_n) tel que les $X_i, i = 1, \dots, n$ sont des variables aléatoires et qu'un échantillon de taille N est l'ensemble des N réalisations indépendantes et identiquement distribuées de la variable aléatoire X .

5.2.1.1.3 Distribution (ou Loi) de probabilité des variables aléatoires réelles

Une loi de probabilité est une fonction, P_X , qui décrit l'ensemble des valeurs (ou états) qu'une variable aléatoire X peut prendre et qui renseigne la probabilité de prendre telle ou telle valeur. Une variable aléatoire est entièrement caractérisée par sa loi de probabilité. Les variables aléatoires continues sont entièrement caractérisées par leur fonction de répartition (cumulative distribution function (CDF) en anglais) ou leur densité de probabilité (PDF pour probability density function en anglais).

a) Fonctions de répartition

La fonction de répartition F_X d'une variable aléatoire X , est une fonction monotone croissante continue à gauche définie par :

$$F_X(x) = P(X < x) \quad (5.3)$$

b) Densités de probabilité

Une loi de probabilité P_X admet une densité de probabilité f_X si pour tout intervalle I de l'ensemble des réels \mathbb{R} on a :

$$P_X(I) = \int_{-\infty}^{+\infty} \mathbb{I}_I(x) f_X(x) dx \quad (5.4)$$

Ce qui implique que F_X est dérivable de dérivée f_X :

$$P_X(a < X < b) = \int_a^b f_X(x) dx = F_X(b) - F_X(a) \quad (5.5)$$

Une densité de probabilité est une fonction positive d'intégrale égale à 1.

5.2.1.1.4 Moments d'une variable aléatoire

Une loi de probabilité peut également être caractérisée par certaines valeurs typiques associées aux notions de tendance centrale, de dispersion et de forme de la distribution appelées *moments* (Saporta, 2006). Pour un nombre entier n positif, le moment d'ordre n de la variable aléatoire X , noté $\mathbb{E}(X^n)$, est définie sous des hypothèses de régularité par :

$$\mathbb{E}(X^n) = \int_{\Omega} X^n(\omega) P(d\omega) = \int_{\mathbb{R}} x^n P_X(dx) \quad (5.6)$$

a) Espérance mathématique

Le moment d'ordre 1, ou *espérance mathématique*, d'une variable aléatoire réelle X est notée $\mathbb{E}(X)$. C'est la moyenne des valeurs que l'on attend, si l'on répète un grand nombre de fois la même expérience aléatoire. Elle est définie pour une variable aléatoire X de l'espace probabilisé (Ω, \mathcal{F}, P) dans \mathbb{R} muni de sa tribu $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ par :

$$\mathbb{E}(X) = \int_{\Omega} X(\omega) P(d\omega) = \int_{\mathbb{R}} x P_X(dx) \quad (5.7)$$

b) Variance et écart-type

La variance de X , notée $Var(X)$, est une grandeur mathématique reflétant la dispersion de l'ensemble des valeurs possibles de la variable aléatoire X . Une variance élevée indique que la dispersion des valeurs possibles de X est large alors qu'une variance faible indique que les réalisations de la variable aléatoire X sont proches de l'espérance mathématique. La variance est définie par :

$$Var(X) = \int_{\mathbb{R}} (x - \mathbb{E}(X))^2 f_X(x) dx \quad (5.8)$$

Une autre mesure de la dispersion est l'écart-type, noté souvent σ , et défini comme la racine carrée de la variance ($\sigma = \sqrt{Var(X)}$).

c) Moments décrivant la forme de la distribution

Parmi les moments remarquables on peut également distinguer le skewness (moment centré réduit d'ordre 3 noté γ_1) et le kurtosis (moment centré réduit d'ordre 4 noté β_2) décrivant respectivement l'asymétrie et l'aplatissement de la distribution de la variable aléatoire X . Ils sont respectivement définis par :

$$\gamma_1 = \mathbb{E} \left[\left(\frac{X - \mathbb{E}(X)}{\sigma} \right)^3 \right] \quad (5.9)$$

$$\beta_2 = \mathbb{E} \left[\left(\frac{X - \mathbb{E}(X)}{\sigma} \right)^4 \right] \quad (5.10)$$

5.2.1.1.5 Dépendance entre variables et corrélation

L'analyse d'incertitude et l'analyse de sensibilité peuvent faire appel à la notion de corrélation, *i.e.* la notion de dépendance entre variables aléatoires, et par conséquent à la dépendance entre distributions de probabilité. La dépendance décrit ici tout type de relation entre deux ou plusieurs ensembles de valeurs (*i.e.* dans ce contexte deux ou plusieurs ensembles des réalisations de variables aléatoires) : *e.g.* la corrélation entre la taille et le poids d'un individu. La corrélation peut être linéaire ou non linéaire. Il est important de noter que la corrélation entre deux variables aléatoires n'implique pas de relation de cause à effet entre les réalisations de ces deux variables. L'intensité de la corrélation entre deux variables est caractérisée par des coefficients de corrélation définis dans l'intervalle $[-1,1]$. Lorsque la valeur du coefficient de corrélation est proche de 1 (ou de -1) il traduit une corrélation forte positive (ou négative) et lorsque sa valeur est proche de 0 il traduit une corrélation presque nulle (*i.e.* absence de relation entre les variables considérées). Les principaux coefficients de corrélation sont les coefficients dits de (Bravais-) Pearson, de Spearman et de Kendall.

a) Coefficient de corrélation de Pearson

Le coefficient de corrélation de Bravais-Pearson, communément appelé coefficient de corrélation de Pearson renseigne la corrélation linéaire entre deux variables. Si l'on considère

X_1 et X_2 deux variables aléatoires, le coefficient de corrélation de Pearson noté généralement $\rho(X_1, X_2)$ est défini par :

$$\rho(X_1, X_2) = \frac{\mathbb{E}[(X_1 - \mathbb{E}(X_1))(X_2 - \mathbb{E}(X_2))]}{\sigma_{X_1} \sigma_{X_2}} \quad (5.11)$$

L'une des limites du coefficient de corrélation de Pearson repose sur le fait qu'il renseigne exclusivement une relation considérée comme linéaire entre les deux variables d'intérêt ; ce qui n'est généralement pas le cas dans la réalité lorsque l'on s'intéresse à des phénomènes naturels par exemple.

b) Coefficient de corrélation de Spearman

Le coefficient de corrélation de Spearman entre deux variables aléatoires X_1 et X_2 est un coefficient non-paramétrique calculé à partir des distributions des rangs des réalisations des deux variables aléatoires. Il est égal au coefficient de Pearson calculé sur les rangs des variables X_1 et X_2 et défini par :

$$\rho_S(X_1, X_2) = \rho(R_{X_1}, R_{X_2}) \quad (5.12)$$

Les quantités R_{X_1} et R_{X_2} désignent les distributions des rangs des valeurs des variables X_1 et X_2 . Pour chaque valeur, le rang représente sa position dans l'échantillon (après classement croissant). Le coefficient de Spearman est particulièrement pratique car il ne fait pas l'hypothèse d'une relation linéaire ; ce qui en fait une mesure adaptée lorsque l'on étudie des phénomènes n'ayant aucune raison d'être liés linéairement comme c'est généralement le cas dans la réalité.

c) Coefficient de corrélation de rang de Kendall

Le coefficient de corrélation de rang de Kendall (appelé *tau* de Kendall et introduit par [Kendall \(1955\)](#)) évalue le degré de similitude entre deux ensembles de rangs associés à deux variables aléatoires. Ce coefficient est basé sur les relations deux-à-deux entre les éléments constitutifs de deux échantillons ([Caniou, 2012](#)). Soit $(x_1^{(i)}, x_2^{(i)})$ et $(x_1^{(j)}, x_2^{(j)})$ deux paires de réalisations issues de la loi jointe entre X_1 et X_2 avec i et j les positions des observations dans l'échantillon. Les deux paires sont *concordantes* si et seulement si les ordres des deux variables sont les mêmes (i.e. $x_1^{(i)} > x_1^{(j)}$ et $x_2^{(i)} > x_2^{(j)}$, ou $x_1^{(i)} < x_1^{(j)}$ et $x_2^{(i)} < x_2^{(j)}$) et *discordantes* sinon ([Caniou, 2012](#)).

L'estimateur du τ de Kendall est défini par :

$$\hat{\tau} = \frac{N_C - N_D}{\frac{1}{2}N(N-1)} \quad (5.13)$$

Où N_C et N_D désignent respectivement le nombre de paires concordantes et discordantes et le dénominateur correspond au nombre total de paires comparables.

5.2.1.2 Interprétations fréquentiste et bayésienne

Dans cette partie nous nous intéressons à la probabilité de réalisation de l'événement A , notée $P(A)$. Suivant l'interprétation fréquentiste, la probabilité $P(A)$ définit la fréquence de réalisation de l'événement A si l'expérience venait à se reproduire infiniment (Zio and Pedroni, 2013). L'interprétation fréquentiste est une vision qui repose sur la loi des grands nombres et qui, ainsi, considère la probabilité d'un événement comme la limite de sa fréquence d'occurrence lorsque le nombre de répétitions de l'expérience tend vers l'infini (Saporta, 2006). Un exemple typique permettant d'illustrer cette vision est le lancer de pièces de monnaie où l'on s'intéresse à la probabilité de tomber sur *pile* (ou sur *face*) lorsque l'on répète un nombre fini de fois.

A contrario, l'interprétation subjective des probabilités sur laquelle repose l'approche bayésienne considère la probabilité $P(A)$ comme le *degré de croyance* accordée à la réalisation de l'événement A (Villon, 2004 ; Zio and Pedroni, 2013). Cette vision est souvent associée à une notion de *pari rationnel* dans laquelle $P(A)$ est définie par les conditions auxquelles un individu rationnel est prêt à parier sur la réalisation d'un tel événement (Pasanisi, 2004) ou à une notion de choix sur la base d'un événement standard : choisir une boule favorable dans une urne contenant $P(A) \times 100\%$ de boules favorables (Zio and Pedroni, 2013). La figure ci-après (Fig. 5.3) illustre la distinction entre les visions.

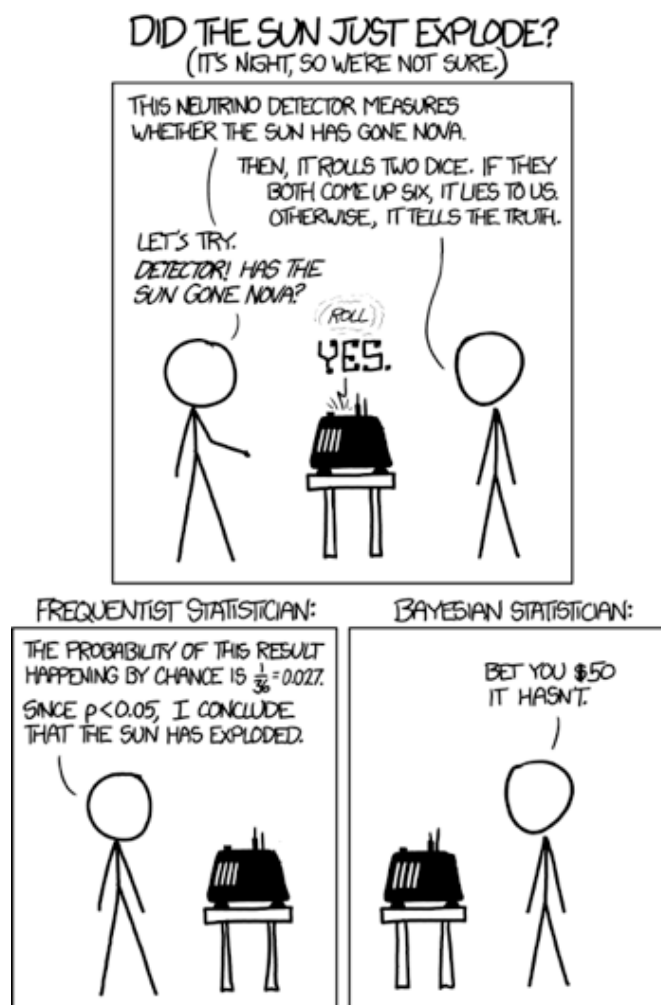


Fig. 5.3 : Illustration des visions fréquentiste et bayésienne (source : <http://www.christian-faure.net/2013/08/31/deux-visages-du-calcul-des-probabilites-bayesiens-et-frequentistes/>).

5.2.1.3 Inférence statistique

Il est fréquent que les distributions de probabilité devant quantifier les incertitudes sur les différents paramètres d'intérêt soient inconnues. Une solution consiste à ajuster une loi de distribution usuelle sur les données disponibles en inférant (estimant) les paramètres de ces distributions usuelles. Cette méthode appelée inférence statistique est décrite ci-dessous.

L'inférence statistique repose sur le cadre théorique introduit dans la section précédente et offre les outils techniques permettant d'identifier et de prendre en compte les sources d'incertitudes autour de phénomènes étudiés et en particulier de quantifier l'incertitude et la variabilité accompagnant les réalisations de ces phénomènes. Deux types de méthodes « s'opposent » en matière d'inférence statistique sur la base des interprétations fréquentiste et subjective de la notion de probabilité (Commeau, 2012). Nous avons d'une part une approche

d'estimation ponctuelle dite fréquentiste qui considère l'information comme venant exclusivement des données à disposition et que les paramètres incertains sont des quantités fixes. D'autre part, nous avons une approche dite bayésienne qui prend en compte, en plus de l'information apportée par les données, les connaissances *a priori* sur le phénomène étudié et qui considère, tous les paramètres incertains comme des quantités aléatoires (Parent and Bernier, 2007 ; Boreux *et al.* 2010).

5.2.1.3.1 Inférence fréquentiste

L'inférence au sens fréquentiste consiste à effectuer un ajustement d'une distribution de probabilité usuelle telle que la loi normale ou la loi gamma sur les données constituant les réalisations du phénomène observé. Cet ajustement se fait dans l'optique d'obtenir une distribution paramétrique usuelle proche de la vraie distribution des données (qui elle est inconnue) pouvant être utilisée ultérieurement à des fins prédictives.

Plaçons-nous dans un cadre de modélisation où un phénomène est la réalisation d'une certaine loi de probabilité, notée $P(X|\theta)$, engendrée par un modèle paramétrique dépendant du paramètre θ de dimension k , à partir de laquelle sont générées N observations $X_i: i = 1, \dots, N$. Il existe différentes méthodes d'inférence fréquentiste, dont la maximisation de la vraisemblance, illustrée ici. Cette méthode consiste à trouver l'estimateur $\hat{\theta}$ du paramètre fixé θ qui maximise la vraisemblance de l'échantillon (*i.e.* la probabilité des observations conditionnellement à l'unique valeur vraie du paramètre θ du modèle statistique qu'on utilise pour leur description (Pasanisi, 2004)). Ce problème d'optimisation peut être présenté comme suit :

$$\hat{\theta} = \operatorname{argmax}_{\theta \in \Theta} P(X, \theta) = \operatorname{argmax}_{\theta \in \Theta} \prod_{i=1}^N P(X_i, \theta) \quad (5.14)$$

D'un point de vue pratique, on dispose d'outils techniques tels que la librairie *fitdistrplus* du logiciel R (Delignette-Muller and Dutang, 2012 ; R CoreTeam, 2015) offrant un panel de fonctions permettant d'appliquer entre autres la méthode du maximum de vraisemblance (Saporta, 2006) pour estimer les paramètres des distributions. Ces outils permettent également de produire des critères de bon ajustement (en anglais *goodness-of-fit*) en réalisant des tests statistiques comme ceux d'Anderson-Darling, Cramer-Von Mises et Kolmogorov-Smirnov (Anderson and Darling, 1952; 1954) visant à évaluer l'écart entre les fonctions de répartition théorique et empirique.

5.2.1.3.2 Inférence bayésienne

a) Démarche classique

L'inférence bayésienne repose fondamentalement sur la théorie des probabilités conditionnelles et le théorème de Bayes (Bayes and Price, 1763), décrit par l'équation ci-dessous :

$$P(B|A) = \frac{P(A|B) \times P(B)}{P(A)} ; P(A) \text{ et } P(B) \neq 0 \quad (5.15)$$

où $P(A)$ et $P(B)$ correspondent aux probabilités que les événements A et B se réalisent respectivement, $P(A|B)$ et $P(B|A)$ sont respectivement les probabilités que l'événement A se réalise en sachant B et que B se réalise sachant A .

Plaçons-nous dans le même cadre de modélisation que l'approche fréquentiste où un phénomène est la réalisation d'une distribution de probabilité notée $P(Y|\theta)$. Contrairement à l'approche fréquentiste, dans le formalisme bayésien, le paramètre θ est considéré comme étant aléatoire et donc comme la réalisation d'une loi de probabilité $P(\theta)$ dite loi *a priori* (*prior* en anglais). L'inférence bayésienne consiste à mettre à jour l'état de la connaissance sur le paramètre θ en combinant d'une part l'information fournie par la loi *a priori* et d'autre part celle apportée par les données observées issues de la loi $P(Y|\theta)$ (vraisemblance du modèle paramétrique). Ce nouvel état de la connaissance va correspondre à la réalisation d'une certaine loi de probabilité dite loi *a posteriori* (*posterior* en anglais) et notée $P(\theta|Y)$. Le formalisme bayésien peut alors se présenter, par application du théorème de Bayes, sous la forme ci-dessous :

$$P(\theta|Y) = \frac{P(Y|\theta) \times P(\theta)}{\int P(Y|\theta) \times P(\theta) d\theta} \quad (5.16)$$

Le dénominateur du membre de droite correspond à la loi marginale des observations $P(Y)$, qui est indépendante du paramètre θ . Il est également usuel de représenter ce formalisme sous sa forme simplifiée comme suit :

$$P(\theta|Y) \propto P(Y|\theta) \times P(\theta) \quad (5.17)$$

Le symbole \propto représentant la proportionnalité entre le membre de gauche et celui de droite dans l'équation (5.9).

Il est possible d'obtenir une loi *a posteriori* ayant une forme analytique correspondant à des lois usuelles telles que la loi normale de Laplace-Gauss. Cela est possible quand la vraisemblance et la loi *a priori* appartiennent à la même famille de lois conjuguées. Cette possibilité porte le nom de *conjugaison*. Dans ce cas précis la loi de probabilité *a posteriori* est de la même forme que la loi *a priori*. Il est cependant rare qu'une distribution de probabilité devant « expliquer » une réalité physique prenne la forme d'une loi de probabilité usuelle appartenant à une famille de lois conjuguées. Le calcul analytique d'inférence bayésienne devient alors particulièrement fastidieux voire impossible à réaliser. L'évolution technologique et le développement de machines de plus en plus performantes au cours des soixante dernières années ont permis à l'approche bayésienne de connaître un essor considérable avec le développement d'outils numériques de simulation se basant principalement sur la théorie des Chaines de Markov et des simulations de type Monte Carlo (appelés MCMC pour Markov Chain Monte Carlo en anglais) (Parent and Bernier, 2007). Parmi les plus connus, citons l'*algorithme de Metropolis-Hastings* (Metropolis et al. 1953; Hastings, 1970) et l'*échantillonneur de Gibbs* (Geman and Geman, 1984; Casella and George, 1992). Ces algorithmes permettent la simulation de chaînes de Markov dont la distribution stationnaire correspond à la loi *a posteriori* recherchée. Les deux méthodes sont présentées plus loin.

b) Démarche hiérarchique

Les praticiens peuvent être généralement confrontés, dans l'étude de phénomènes naturels ou de mécanismes physiques, biologiques ou chimiques, à des problèmes de modélisation complexes et dont la complexité est probablement gouvernée par l' (inter)action des processus sous-jacents très souvent difficiles à identifier et à spécifier. Ainsi, dans un cadre probabiliste de modélisation, des structures probabilistes plus riches sont nécessaires pour être pertinentes face à cette complexité des phénomènes étudiés (Ancelet, 2008). Une démarche de modélisation et d'inférence hiérarchiques apparaît comme une approche intéressante pour s'accommoder de cette complexité.

L'essence de la modélisation hiérarchique repose sur le développement de relations stochastiques multi-niveaux permettant de saisir les composantes importantes des phénomènes étudiés et de s'accommoder de leur complexité en disséquant le problème en une multitude de relations simples à 1 niveau (Carlin et al. 2006 ; Clark, 2005). Comme évoqué par Ancelet (2008) cette démarche n'est pas si nouvelle car elle reprend les idées de la

méthode cartésienne : « *Diviser chacune des difficultés que j'examinerais, en autant de parcelles qu'il se pourrait et qu'il serait requis pour les mieux résoudre* ». René Descartes. *Discours de la méthode* (1637).

Berliner (1996) a proposé d'étendre le formalisme bayésien classique du problème, pouvant être vu comme une structure de modélisation à deux entités composées des données observées (Y), des paramètres (θ) (équation (5.18)), en définissant trois entités (décrites dans l'équation (5.19)) composées des données observées, des paramètres et des variables latentes décrivant les processus sous-jacents (Z).

$$P(\theta|Y) \propto P(Y|\theta) \times P(\theta) \quad (5.18)$$

$$P(\theta|Z, Y) \propto P(Y|Z, \theta) \times P(Z|\theta) \times P(\theta) \quad (5.19)$$

Dans notre cas précis, la notion de hiérarchie vient du fait que l'on peut être en présence de données constituées de différents types de regroupements. Par exemple étudier les cultures maraichères contaminées suite à un accident nucléaire suivi de dépôts atmosphériques sur le couvert végétal peut nous amener à traiter des données relatives à des espèces végétales différentes (*e.g.* légumes feuilles, légumineuses) qui elles-mêmes présentent des variétés différentes (épinard, choux, *etc.*); constituant ainsi différents niveaux d'organisation. D'autres hiérarchies dans les données sont également possibles, comme les méta-analyses (les données sont regroupées par étude) ou les mesures répétées sur des individus (Commeau, 2012).

La modélisation hiérarchique en général, et bayésienne hiérarchique spécialement, est particulièrement appropriée lorsque l'on souhaite caractériser l'*incertitude* relative à des phénomènes naturels et la *variabilité* induite par les différents niveaux d'organisation. De plus, cette approche jouit d'une certaine flexibilité dans la modélisation de problèmes complexes car elle se prête bien à une représentation graphique des dépendances conditionnelles (toujours orientées dans le même sens (Ancelet, 2008)) entre les différentes entités du problème à l'aide graphes acycliques orientés (ou DAG pour Direct Acyclic Graph selon Spiegelhalter *et al.* (1993)). Le concept de DAG et les approches de modélisation et d'inférence bayésiennes hiérarchiques sont illustrées par la Fig. 5.4. Généralement dans un DAG :

- les entités constituent des nœuds
- les enfants d'un nœud sont constitués par tous les nœuds dépendants de ce nœud
- les parents d'un nœud sont tous les nœuds dont dépend le nœud

- les variables latentes constituent l'ensemble des nœuds avec parents et enfants
- les paramètres sont des nœuds sans parents
- les observables sont généralement des nœuds sans enfants (appelés feuilles)
- les paramètres et les variables latentes sont représentés par des ellipses
- les observables (les données) sont représentés par des formes rectangulaires
- différents types de relation (telles que la chaîne, l'héritage, la naissance) peuvent être représentées (voir (Parent and Bernier, 2007)).

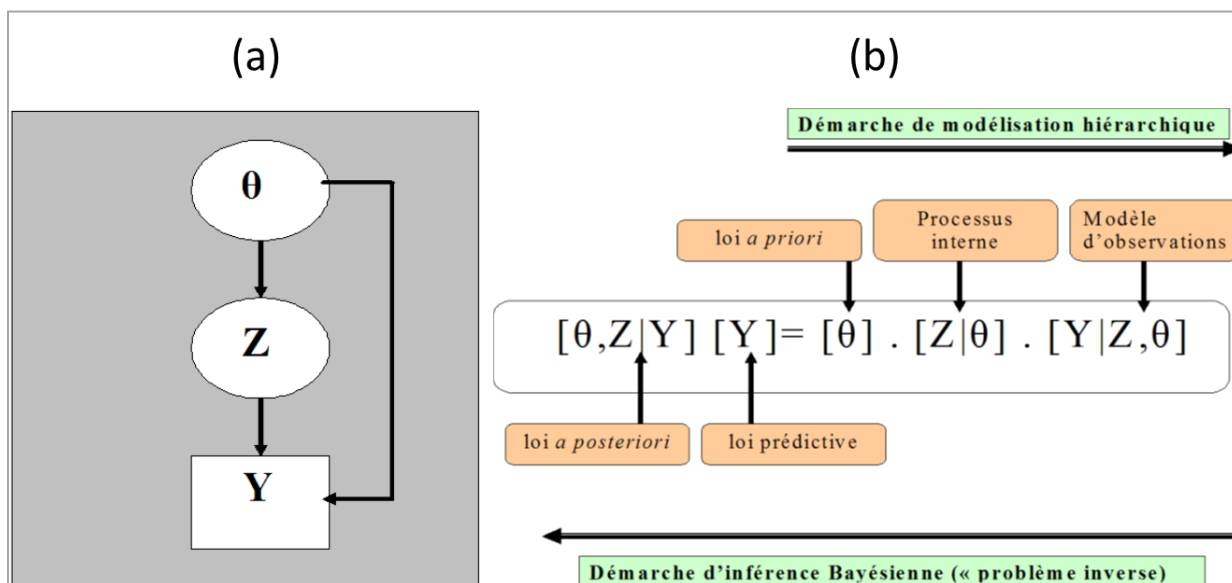


Fig. 5.4 : Illustrations schématiques de la modélisation et de l'inférence hiérarchiques. (a) illustre un DAG représentant un modèle hiérarchique simple et (b) illustre les démarches de modélisation et d'inférence bayésiennes. (Adapté de (Ancelet, 2008)).

L'approche bayésienne hiérarchique, appliquée de manière générale dans des disciplines telles que l'écologie ou la radioécologie, représente une stratégie intermédiaire entre les modèles traditionnels ne prenant pas en compte toutes les sources de variabilité et les modèles hautement paramétrés qui tentent de toutes les caractériser (Jiao *et al.* 2009). On peut donc voir l'approche bayésienne hiérarchique comme une hiérarchie imbriquée de relations stochastiques pour laquelle l'un des défis importants est d'estimer les paramètres présents dans les couches latentes décrivant le processus interne (cf. Fig. 5.4 (b)).

Je vous renvoie, dans le but de fouiller plus en détails la construction des modèles hiérarchiques et la philosophie derrière la modélisation hiérarchique et en particulier la modélisation hiérarchique bayésienne, aux ouvrages (anglo-saxons et francophones) de référence suivants : (Boreux *et al.* 2010 ; Gelman *et al.* 2004 ; Jackman, 2009 ; Parent et Bernier, 2007).

5.2.1.4 Conclusion sur l'approche probabiliste

La démarche de quantification des incertitudes dans le cadre de l'analyse des risques industriels, financiers, environnementaux, sanitaires et alimentaires entre autres, a été principalement menée depuis plusieurs décennies dans un cadre purement probabiliste. L'approche probabiliste (objective ou subjective) repose sur un cadre théorique bien défini se prêtant de manière adéquate aux préceptes décrivant le cadre général de l'analyse de risque (cf. Fig. 1.1). De plus, cette approche offre des outils techniques garantissant une certaine flexibilité pour la prise en compte des sources d'*incertitude* et de *variabilité* dans les différentes évaluations de risque. Les analyses de risque et d'incertitude reposant sur les probabilités peuvent cependant être contestées car la disponibilité des informations pertinentes et la nature des sources d'*incertitude* à prendre en compte peuvent rendre particulièrement difficile la caractérisation des incertitudes (Baudrit, 2005 ; Zio and Pedroni, 2013). Différentes approches alternatives garantissant une distinction entre *incertitude* due au manque de connaissances ou à l'imprécision et l'*incertitude* due à l'hétérogénéité naturelle des phénomènes, introduites depuis les années 1960, prennent de plus en plus de l'ampleur en analyse quantitative de risque (Zio and Pedroni, 2013). Dans la section suivante, quelques-unes de ces approches alternatives dédiées à la caractérisation des incertitudes dans un cadre probabiliste « inhabituel » ou dans un cadre « extra-probabiliste » (*i.e.* utilisant une axiomatique non-probabiliste) sont brièvement présentées.

5.2.2 Approches alternatives

Avant d'introduire les différentes approches il est intéressant de définir les concepts de probabilités haute et basse qui seront utiles pour la compréhension des approches alternatives dans la suite. Soit l'espace probabilisé (Ω, \mathcal{F}, P) et \mathcal{P} une famille de mesures de probabilité sur Ω . Pour tout ensemble mesurable $A \subseteq \Omega$, les probabilités haute et basse définissent les bornes *sup* et *inf* de la probabilité inconnue de réalisation de l'événement A définies par :

$$\text{Probabilité haute} \quad \bar{P}(A) = \sup_{P \in \mathcal{P}} P(A) \quad (5.20)$$

$$\text{Probabilité basse} \quad \underline{P}(A) = \inf_{P \in \mathcal{P}} P(A) \quad (5.21)$$

Ces probabilités haute et basse fournissent, dans l'optique de caractériser l'incertitude entourant une probabilité imprécise ou inconnue, un encadrement de cette probabilité.

5.2.2.1 Représentation par paires de probabilités cumulées

L'utilisation de paires de probabilités cumulées hautes et basses, appelées p-boxes, pour représenter l'*incertitude* due à l'imprécision ou au manque de connaissances a été introduite entre la fin des années 1990 et le début des années 2000 par [Ferson et al. \(2003\)](#).

La représentation de l'incertitude par les p-boxes repose sur le principe visant à approcher une probabilité en l'encadrant par une paire $(\underline{F}, \overline{F})$ de fonctions de répartition haute et basse ([Baudrit, 2005](#)). La p-box, correspondant à l'intervalle $[\underline{F}, \overline{F}]$, représente la classe des mesures de probabilité dont la fonction de répartition est comprise entre les bornes \underline{F} et \overline{F} comme suit :

$$\underline{F}(x) \leq F(x) \leq \overline{F}(x) \quad , \quad \forall x \in \mathbb{R} \quad (5.22)$$

Ainsi, la p-box fournit un encadrement d'une probabilité inconnue (ou imprécise) en rendant compte du caractère incomplet de l'information ([Baudrit, 2005](#)). [Baudrit \(2005\)](#) a proposé une synthèse des représentations mathématiques cohérentes avec la nature de l'information. Dans cette synthèse, il est recommandé de représenter à l'aide de p-boxes les informations disponibles : soit le minimum, la moyenne et le maximum des valeurs possibles, soit la forme de la loi, soit la forme de la loi et les bornes des valeurs des paramètres. Un ensemble de méthodes permettant de définir des p-boxes à partir de ce type d'information sont décrites par [Ferson et al. \(2003\)](#).

5.2.2.2 Représentation par les fonctions de croyance

La notion de *fonctions de croyance* ("belief functions" en anglais) et leur utilisation pour la caractérisation de l'*incertitude* tiennent leur origine de la théorie de Dempster-Shafer ("Dempster-Shafer theory of evidence" en anglais) introduite par les travaux de [Dempster \(1967\)](#) puis développée par ceux de [Shafer \(1976\)](#). Cette théorie est particulièrement intéressante dans l'incorporation et la représentation d'une information incomplète car elle permet la prise en compte et le traitement d'informations intermédiaires entre un intervalle (ex : $[0,1]$) et une distribution de probabilité dont les caractéristiques sont bien définies ([Zio and Pedroni, 2013](#)).

La Théorie de Dempster-Shafer implique la définition du triplet (Ω, I, m) où Ω correspond à l'univers des possibles (par analogie à la théorie des probabilités), I désigne l'ensemble des éléments focaux ("set of focal elements" en anglais) représentant une collection dénombrable d'ensembles de Ω et m (BPA pour Basic Probability Assignment en anglais) désigne la fonction (ou distribution) de masse sur Ω ($m : \mathcal{P}(\Omega) \rightarrow [0,1]$) définie pour tout ensemble $A \subseteq \Omega$ par :

$$\begin{cases} m(A) > 0 & \text{si } A \in I \\ m(A) = 0 & \text{si } A \notin I \\ \sum_{A \in I} m(A) = 1 \end{cases} \quad (5.23)$$

Contrairement aux distributions de probabilités qui dans le cas discret attribuent à chaque singleton un poids, la fonction de masse attribue un poids à l'ensemble des éléments focaux pouvant être des ensembles, des intervalles ou des singletons (Baudrit, 2005). La valeur de $m(A)$ exprime le degré d'appartenance à l'élément focal A . Il faut noter que ce degré d'appartenance ne renseigne pas comment cette appartenance est distribuée dans l'élément focal A (Zio and Pedroni, 2013). L'ignorance totale (probabilité que l'on ne sache rien) est illustrée par $m(\Omega) = 1$.

Fondamentalement la Théorie de Dempster-Shafer repose sur deux mesures de vraisemblance, le *degré de crédibilité* (belief) noté $Bel(\cdot)$ et le *degré de plausibilité* noté $Pl(\cdot)$ permettant de caractériser le degré de croyance en la réalisation d'un événement. Pour $A \subseteq \Omega$, le *degré de crédibilité* et le *degré de plausibilité* sont définis par :

$$\begin{aligned} Bel(A) &= \sum_{E, E \subseteq A} m(E) \\ Pl(A) &= \sum_{E, E \cap A \neq \emptyset} m(E) \end{aligned} \quad (5.24)$$

Le degré de crédibilité $Bel(A)$ évalue à quel point il est certain que l'information disponible implique une appartenance à l'élément focal A et le degré de plausibilité $Pl(A)$ évalue à quel point l'information disponible ne contredit pas l'appartenance à l'élément focal A . Ces deux mesures sont des fonctions duales respectant, pour $A \subseteq \Omega$ et A^c l'ensemble des éléments focaux n'appartenant pas à A ,

$$Pl(A) = 1 - Bel(A^c) \text{ et } Bel(A) = 1 - Pl(A^c) \quad (5.25)$$

La fonction de masse permet de définir une famille de probabilités $\mathcal{P}_{Bel,Pl}$ telle que l'écart entre les mesures Bel et Pl encadrant la probabilité P (i.e. $\forall P \in \mathcal{P}_{Bel,Pl}, Bel \leq P \leq Pl$) représente le caractère imprécis de l'information.

5.2.2.3 Représentation par les distributions de possibilité

Les distributions de possibilité sont des outils mathématiques permettant de proposer une représentation de l'incertitude. Elles reposent sur la Théorie des possibilités initialement introduite par Zadeh (1978) puis développée à partir du milieu des années 1980 par Dubois et Prade (1985). La théorie des possibilités est liée à travers les distributions de possibilité (sous certaines conditions) à une modélisation basée sur l'utilisation d'objets mathématiques tels que les nombres, intervalles et ensembles flous. La citation suivante évoque ce lien : *"Possibility distributions are linked to fuzzy intervals [...]. Fuzzy sets logic is essentially a mathematical tool imagined for coping with uncertainty. Formalized in the 60's, first applied to control problems, then as a modelling technique, fuzzy rules prove to be valuable tools to model complex systems in presence of uncertainties, as well as to build approximate models of deterministic "well defined" systems"*. (Pasanisi, 2014). Cependant, nous n'introduisons pas la théorie sur les ensembles flous. Quelques références sont présentées ci-après pour les lecteurs intéressés (Goguen, 1967 ; Kaufman, 1973 ; Zadeh, 1965).

Soit $\pi(\cdot)$, la fonction qui assigne à chaque élément a d'un ensemble $A \subseteq \Omega$ une valeur $\pi(a) \in [0,1]$ caractérisant le *degré de possibilité* que l'élément a appartienne à A . La distribution de possibilité, $\pi_b(\cdot)$, est une représentation de ce que l'on sait sur la possibilité d'appartenance de b à l'ensemble A au sens que $\pi_b(a) = 0$ signifie que $a = b$ est impossible et que $\pi_b(a) = 1$ signifie que $a = b$ est tout simplement normal (Dubois, 2006). La distribution de possibilité $\pi_b(\cdot)$ est normalisée telle que $\exists a \in A : \pi_b(a) = 1$. La forme la plus simple de distribution de possibilité sur l'ensemble A est la fonction indicatrice sur un sous-ensemble E de A telle que $\mathbb{I}_b(a) = 1$ si $b \in E$ et 0 sinon (Dubois 2006).

La représentation des sources d'incertitude sur la base des distributions de possibilité est rendue possible par les mesures duales de *possibilité* et de *nécessité* (Π et N respectivement) définies pour tout ensemble $A \subseteq \Omega$ à partir d'une distribution de possibilité comme décrit ci-dessous.

$$\begin{aligned} \Pi(A) &= \sup_{a \in A} \pi(a) \\ N(A) &= 1 - \Pi(A^c) = 1 - \sup_{a \notin A} \pi(a) \end{aligned} \quad (5.26)$$

Le degré de *nécessité* $N(A)$ indique dans quelle mesure l'événement A est certain et le degré de *possibilité* $\Pi(A)$ indique dans quelle mesure l'événement contraire A^c n'est pas possible (Baudrit, 2005). Les propriétés de ces mesures sont décrites par Dubois (2006).

A l'image des mesures de *crédibilité* et de *plausibilité* (Bel et Pl), les mesures de *nécessité* et de *possibilité* permettent de définir, sur la base de la distribution de possibilité $\pi(\cdot)$, une famille de probabilité $\mathcal{P}_{N,\Pi}$ contenant l'ensemble des mesures de probabilité bornées par les mesures N et Π et telles que la probabilité imprécise que l'événement A se réalise $P(A)$ est encadrée comme suit $N(A) \leq P(A) \leq \Pi(A)$. Ainsi, l'écart entre ces deux mesures renseigne le caractère imprécis de l'information.

5.2.2.4 Conclusion sur les approches alternatives

Dans cette section nous avons fait une brève description de trois approches alternatives appropriées dans le cadre de la quantification des incertitudes reflétant l'imprécision ou le manque de connaissances. Les trois approches présentées proposent un encadrement de la probabilité mal connue de réalisation d'un événement par des mesures duales par analogie aux probabilités haute et basse (\overline{P} et \underline{P}) introduites au début de la section. On peut donc déceler un lien entre ces approches et l'approche probabiliste classique. Pour une description détaillée de ces approches, de leur lien avec l'approche probabiliste classique et des défis numériques soulevés par leur implémentation je vous renvoie aux travaux de Baudrit (2005). Les travaux de Baudrit (2005) ont apporté une contribution riche sur des approches jusque-là peu explorées dans la littérature francophone en permettant d'avoir un aperçu clair sur les liens théoriques mais également pratiques entre ces différentes approches d'une part et avec l'approche probabiliste d'autre part.

Ces approches connaissent un intérêt croissant car elles apportent une perspective nouvelle dans la prise en compte simultanée et distincte de l'*incertitude* et de la *variabilité* (Arunraj *et al.* 2013 ; Baraldi *et al.* 2011 ; Kentel and Aral, 2005). Dans le cadre de l'évaluation des risques liés aux activités nucléaires on peut citer les travaux de Baudrit (2005), de Baccou *et al.* (2008) et de Destercke (2008). Un exemple illustratif est proposé en Annexe D.

Box 5.1 : Notes aux lecteurs

- Dans les trois applications présentées aux **Chapitres 6, 7 et 8** l'*incertitude* et la *variabilité* sur les paramètres radioécologiques d'intérêt (facteur d'**interception sèche** et d'**interception humide** et la **constante de décroissance biomécanique**, respectivement) sont quantifiées à travers le développement de modèles hiérarchiques bayésiens.
- L'*incertitude* sur les processus de **transfert racinaire** et de **transfert aux produits animaux** a été quantifiée sur la base des données du TRS472 (IAEA, 2010) disponibles pour les paramètres correspondant (**coefficient de distribution dans le sol**, **facteur de transfert sol-plantes** et **facteur de transfert à l'équilibre aux produits animaux**). Cette caractérisation de l'*incertitude* s'est faite via la démarche décrite par [Simon-Cornu et al. \(2015\)](#).

5.3 Propagation d'incertitude

Différentes méthodes de propagation des incertitudes dépendant des types d'information disponibles sur les paramètres d'entrée des modèles existent dans la littérature. Ces méthodes dépendent également fortement du cadre théorique adopté lors de l'étape de caractérisation des différentes sources d'*incertitude* et de *variabilité*. La notion de modèle ici fait toujours référence au formalisme défini et adopté dans les sections précédentes.

5.3.1 Méthodes de Monte Carlo

Les méthodes de Monte Carlo fournissent des solutions approximatives à une très grande variété de problèmes mathématiques en effectuant un échantillonnage statistique sur les données du problème. Elles peuvent être définies comme des méthodes de simulation qui utilisent des séquences de nombres aléatoires. Elles reposent sur la loi des grands nombres ([Saporta, 2006](#)).

5.3.1.1 Un peu d'histoire

Les Méthodes de Monte Carlo ont connu plusieurs développements historiques au cours des derniers siècles dans les sciences mathématiques. Les jeux de hasard ont motivé les mathématiciens des 17^{ème} et 18^{ème} siècles à regarder de près les résultats des essais successifs formant une séquence aléatoire, comme le montrent les travaux du [Comte de Buffon \(1777\)](#), qui a décrit une méthode de calcul de π (pi) basée sur la réalisation d'expériences répétées. Près d'un siècle et demi plus tard, en observant que la moyenne d'une variable aléatoire continue prend la forme d'une intégrale, les statisticiens des 19^{ème} et 20^{ème} siècles ont pu déduire une solution approximative d'une intégrale dans un problème qui contient des difficultés intrinsèques de calcul à partir de nombre tirés aléatoirement.

Les techniques de simulation de Monte Carlo existent formellement depuis le début des années 1940 avec des recherches sur la fusion nucléaire ([Metropolis and Ulam, 1949](#)). Elles tiennent leur nom du physicien Nicholas Metropolis du fait du nombre important de casinos (dans lesquels la roulette peut être considérée comme un bon exemple de générateur de nombres aléatoires) dans la Principauté de Monaco. Cependant c'est seulement avec l'évolution de la technologie et de la puissance des ordinateurs que ces techniques ont trouvé

une plus large utilisation, la notion de performance numérique étant particulièrement importante dans la modélisation de systèmes complexes.

5.3.1.2 Simulation de Monte Carlo 1D et 2D

Comme évoqué précédemment, l'utilisation de distributions de probabilité pour quantifier l'*incertitude* et la *variabilité* associées aux entrées des modèles dans le cadre d'une analyse de risque est l'une des approches les plus largement utilisées (notamment en évaluation des risques alimentaires (Nauta, 2000; Rigaux, 2013)). Cependant, le nombre important de paramètres à prendre compte dans les modèles rend complexe voire impossible l'utilisation des distributions de probabilités dans les calculs analytiques (Crépet, 2007; Robert and Casella, 2011). La simulation de Monte Carlo est un outil statistique particulièrement puissant pour approcher les solutions de problèmes mathématiques complexes (Tuffin, 2010). Cette technique permet de produire à chaque itération un scénario « plausible » en appliquant le modèle considéré à un échantillon aléatoire des entrées, tiré dans leurs distributions de probabilité. Ainsi la distribution de probabilité obtenue reflète les possibles valeurs des sorties du modèle (Vose, 2008).

En fonction du problème rencontré, le modélisateur peut être amené à considérer que les distributions de probabilité liées aux entrées du modèle correspondent soit à de l'*incertitude* due au manque de connaissances, soit à de la *variabilité*, soit à la combinaison des deux. Dans les deux premiers cas la simulation de Monte Carlo d'ordre 1 (MC1D pour Monte Carlo 1 dimension dans la suite) permet de propager l'*incertitude* et / ou la *variabilité* induites par les entrées à travers le modèle et d'obtenir la distribution d'*incertitude* et/ou de *variabilité* sur la sortie du modèle. La méthode est décrite par Fig. 5.5.

Considérons le modèle $Y = f(X_1, \dots, X_n)$

Monte Carlo 1D

1. Définir $i = 1$
2. Pour chaque paramètre $x_j : j = 1, \dots, n$ (n désigne le nombre de paramètres) tirer la $i^{\text{ème}}$ valeur x_j^i à partir de la distribution correspondante ($P_X^j(x_j)$)
3. Calculer la valeur $y^i = f(x_1^i, \dots, x_n^i)$
4. Si $i < N$ (N la taille de l'échantillon) $\rightarrow i = i + 1$ et retourner à l'étape 2
Sinon \rightarrow aller à 5
5. Post-traitement des résultats (Calcul des statistiques descriptives, calculer la fonction de répartition empirique $F_Y(y) \approx \frac{1}{N} \sum_{i=1}^N \mathbb{I}\{y^i \leq y\}$)

Fig. 5.5 : Description de l'algorithme de la méthode de Monte Carlo 1D.

Cependant, une telle approche présente quelques inconvénients car elle peut induire des biais (en agréant l'*incertitude* et la *variabilité*) dans les résultats et de plus ne permet pas de mettre en relief la part de l'*incertitude* ou de la *variabilité* traduisant notre incapacité à prédire précisément la sortie du modèle (Cullen and Frey, 1999; Nauta, 2000; Vose, 2008) ; cela peut conduire à des estimations erronées et rend les résultats difficiles à interpréter (Cullen and Frey, 1999). L'utilisation des résultats dans la gestion du risque est contrariée car impliquant des mesures différentes dans les deux cas (Rigaux, 2013). Par conséquent, de plus en plus il est préconisé de privilégier une approche à deux dimensions permettant d'évaluer l'importance de l'incertitude et celle de la variabilité sur les sorties des modèles considérés en les évaluant « séparément » (Hoffman and Hammonds, 1994; Hoffman and Thiessen, 1996; Cullen and Frey, 1999; Nauta, 2000). Ainsi, la simulation Monte Carlo d'ordre ou de dimension 2 (abrégée dans la suite en MC2D) (Nauta, 2000; Pouillot *et al.* 2007; Pouillot and Delignette-Muller, 2010) apporte un début de solution aux problèmes soulevés ci-dessus et

permet de propager séparément l'incertitude et la variabilité des paramètres dans le modèle (Crépet, 2007; Vose, 2008). Le MC2D consiste à réaliser deux simulations de Monte Carlo imbriquées (Crépet, 2007). En effet à chaque itération, correspondant à une simulation suivant la dimension d'incertitude (la dimension d'incertitude sera N_U), un nombre N_V de valeurs sont simulées suivant la dimension de variabilité ; au final on se retrouve avec $N_U \times N_V$ valeurs de la sortie du modèle. La structure et le fonctionnement de la méthode de simulation MC2D sont illustrés par la Fig. 5.6.

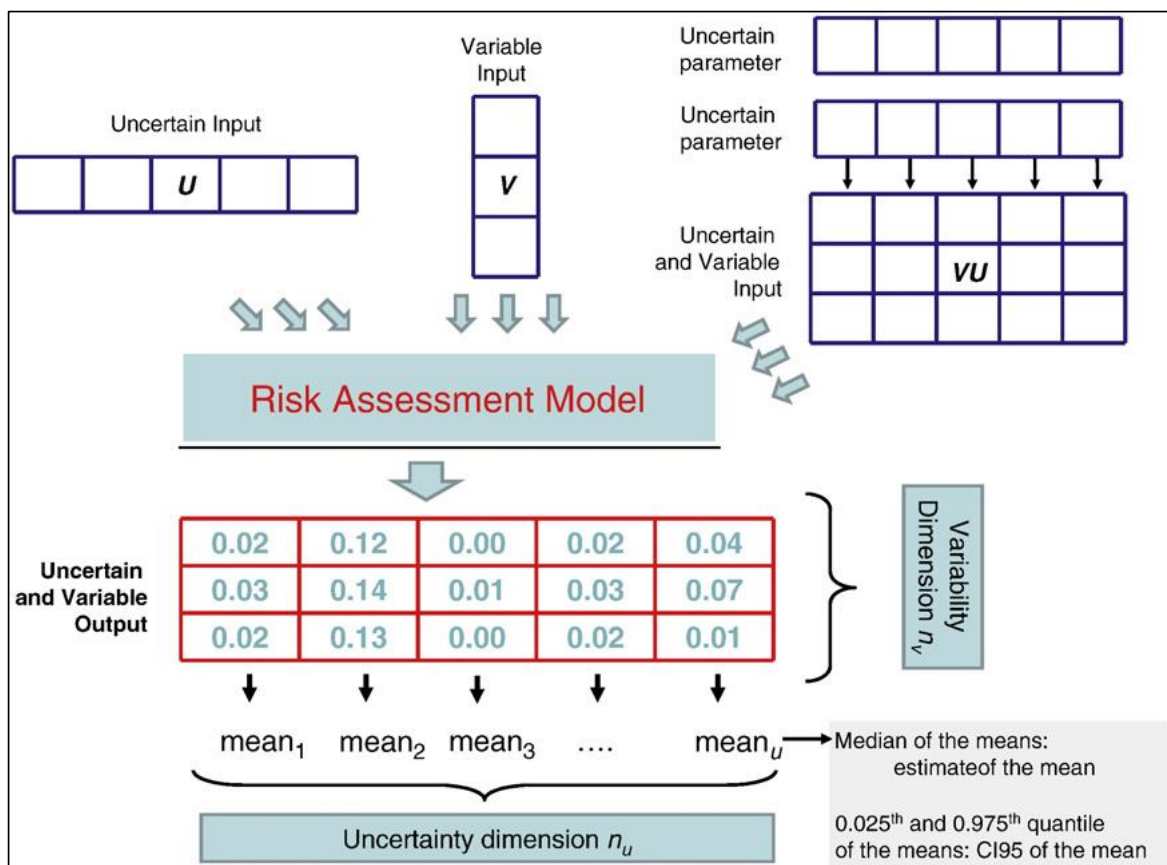


Fig. 5.6 : Schéma illustrant la méthode de Monte Carlo 2D. Source : (Pouillot et Delignette-Muller, 2010).

5.3.1.3 Simulation de Monte Carlo par Chaînes de Markov (MCMC)

Les algorithmes MCMC sont des méthodes mathématiques faisant appel à la fois à la théorie des chaînes de Markov et aux simulations de type Monte Carlo qui permettent de générer aléatoirement une suite de variables corrélées (Vose, 2008 ; Robert and Casella, 2011) dans le but d'atteindre une loi de probabilité cible, inconnue, ayant vraisemblablement

engendré ces variables. Soit $f(\theta) = P(\theta|x) \propto P(x|\theta) \times P(\theta)$ la distribution *a posteriori* du paramètre θ . L'idée derrière les méthodes MCMC est d'obtenir des tirages aléatoires dans une loi de probabilité dont l'expression est connue à une constante près (*i.e.* $f(\theta) = c \times h(\theta)$) (Ancelet, 2008 ; Pasanisi, 2004). De manière plus précise, étant donné la loi cible f , on construit un noyau de Markov K^1 de loi stationnaire² f puis on génère une chaîne de Markov $(\theta^{(t)})$ en utilisant ce noyau de telle sorte que la loi limite de la chaîne $(\theta^{(t)})$ soit la distribution f (Robert and Casella, 2011). La construction du noyau de transition représente le point essentiel dans l'application des méthodes MCMC et donne lieu à l'utilisation de différentes méthodes dont les plus connues sont l'algorithme de Metropolis-Hastings et l'échantillonnage de Gibbs.

5.3.1.3.1 Algorithme de Metropolis-Hastings

La méthode de Metropolis-Hastings (Metropolis *et al.* 1953; Hastings, 1970) est historiquement la première des méthodes MCMC (Parent and Bernier, 2007). Elle a initialement été utilisée pour résoudre des problèmes numériques par les physiciens vers les années 50 avant de toucher avec l'évolution des outils informatiques une communauté plus large dans laquelle on retrouve les statisticiens. Elle se fonde sur le choix d'une distribution de transition instrumentale conditionnelle que nous notons $h(.|.)$ qui en pratique est facile à simuler (Robert and Casella, 2011). L'algorithme de Metropolis-Hastings se présente comme dans la Fig. 5.7.

A chaque itération, un candidat est simulé suivant la loi de proposition conditionnelle $h(.|.)$ et celui-ci est accepté suivant la probabilité d'acceptation $P(u \leq \eta(\theta^{(t)}, \theta^*)) = \eta(\theta^{(t)}, \theta^*)$ et l'état de la chaîne au moment $k+1$ devient $\theta^{(k+1)} = \theta^*$ ou ce candidat est rejeté et la chaîne de Markov reste à l'état courant $\theta^{(k)}$ avec une probabilité $P(u > \eta(\theta^{(t)}, \theta^*)) = 1 - \eta(\theta^{(t)}, \theta^*)$.

¹ Le noyau de Markov ou noyau de transition est la loi de probabilité caractérisant les passages de la chaîne $(\theta^{(t)})$ entre deux états consécutifs *i.e.* la famille $[P_{t,t+1} = K(\theta^{(t)}, \theta^{(t+1)}) = P(\theta^{(t+1)} | \theta^{(t)}, \dots, \theta^{(0)})$, t un entier naturel] est une distribution de probabilité.

² La loi de probabilité f est dite stationnaire lorsque la chaîne de Markov satisfait l'équation $f.K = f$. *i.e.* les passages aux différents états de la chaîne laissent inchangée la loi.

Algorithme de Metropolis-Hastings

1. Initialisation de la chaîne $k = 0 : \theta^{(0)}$

2. Etape k :

- Générer un candidat θ^* pour $\theta^{(k+1)} : \theta^* \sim h(\theta^* | \theta^{(k)})$
- Tester le candidat θ^* :

Simuler $u \sim U(0,1)$

$$\theta^{(k+1)} = \begin{cases} \theta^* & \text{si } u \leq \eta(\theta^{(k)}, \theta^*) \\ \theta^{(k)} & \text{sinon} \end{cases}; \text{ avec } \eta(a, b) = \min \left\{ \frac{f(b) \times h(a|b)}{f(a) \times h(b|a)}, 1 \right\}$$

$a, b \in \mathbb{R}$

3. Revenir en 2 jusqu'à la convergence

Fig. 5.7 : Description de l'algorithme de Metropolis-Hastings.

L'algorithme de Metropolis-Hastings est un outil particulièrement puissant car quelle que soit la loi instrumentale $h(\cdot | \cdot)$ utilisée on peut construire un noyau de Metropolis-Hastings tel que f est sa loi stationnaire (Robert and Casella, 2011). En revanche cet algorithme peut mettre un temps très long à converger dès lors que la loi instrumentale n'explore pas le support (domaine de définition) de la distribution f (en un nombre raisonnable d'itérations). De plus la méthode Metropolis-Hastings s'avère inefficace dans le cas de problèmes impliquant plusieurs paramètres (Vose, 2008). Il existe, cependant, d'autres algorithmes MCMC plus adaptés à la simulation de distributions multidimensionnelles. Dans la section suivante, je présente l'un de ces algorithmes.

5.3.1.3.2 Echantillonneur de Gibbs

L'échantillonnage de Gibbs a été introduit par Geman and Geman (1984) et utilisé par Gelfand and Smith (1990) pour le calcul de densités marginales dans des modèles multidimensionnels à plusieurs paramètres. Geman and Geman (1984) ont élaboré un algorithme dans le cadre de modèles stochastiques complexes faisant intervenir un grand nombre de paramètres tels que des modèles de reconstitution d'image ou des réseaux de neurones ; cet algorithme est connu sous le nom d'échantillonneur de Gibbs (Gelfand and Smith, 1990).

Considérons le vecteur de paramètres $\theta = (\theta_1, \dots, \theta_d)$ de l'espace des paramètres Θ de dimension d . L'échantillonnage de Gibbs consiste à actualiser à chaque itération la

distribution jointe $P(\theta_1, \dots, \theta_d)$ des composantes du vecteur θ en tirant tour à tour dans chacune des *distributions conditionnelles complètes*³ des paramètres (Boreux *et al.* 2010). La méthode est décrite par Fig. 5.8.

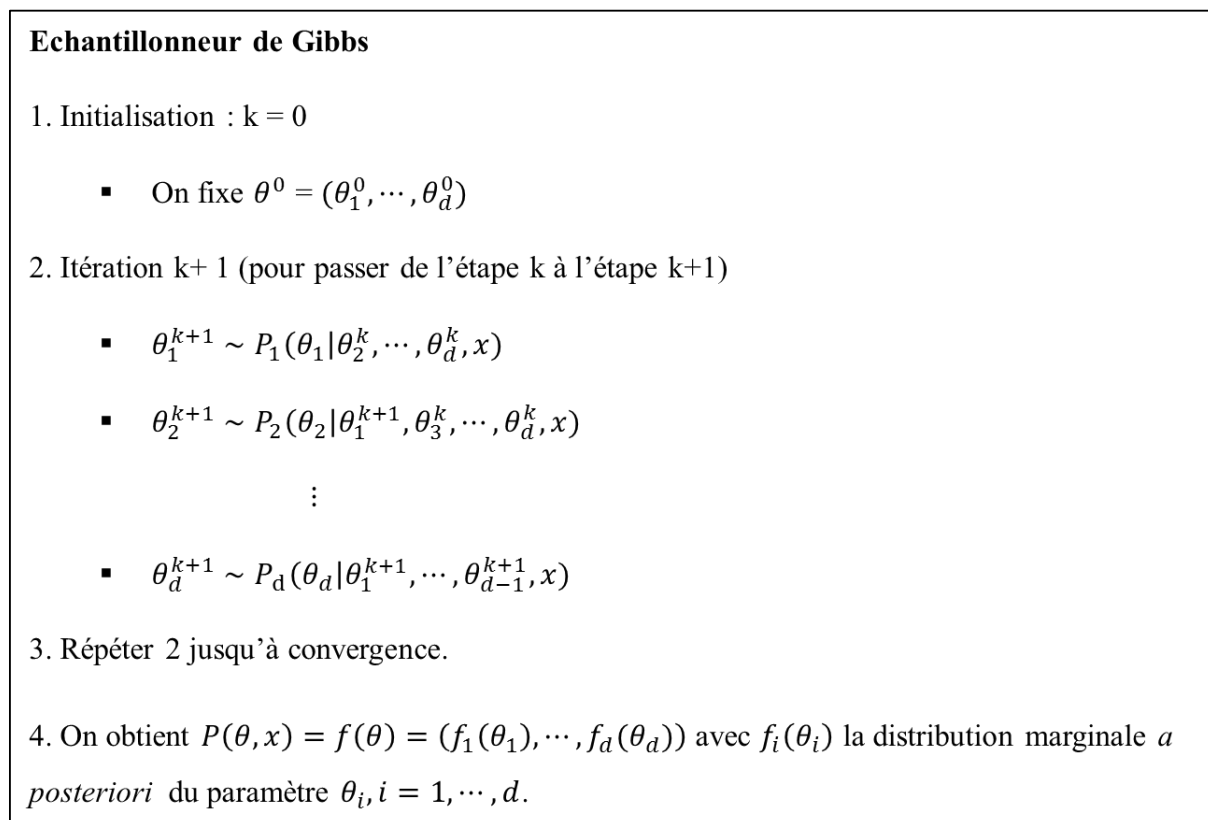


Fig. 5.8 : Description de l'échantillonneur de Gibbs.

A l'origine de chaque étape, les valeurs simulées à l'étape précédente sont utilisées pour spécifier la première conditionnelle complète (Parent and Bernier, 2007), puis les valeurs de conditionnement sont progressivement (paramètre par paramètre) actualisées pour générer les nouvelles valeurs du paramètre considéré.

L'algorithme de Gibbs est très efficace et s'applique bien à des modèles à structure hiérarchique (Robert and Casella, 2011) car il permet de simplifier l'inférence en transformant un problème de simulation d'une distribution jointe dans un espace de dimension d ($d > 1$) en une simulation de d distributions en dimension 1 (Parent and Bernier, 2007).

³ On appelle conditionnelle complète la distribution d'une composante du vecteur θ sachant toutes les autres composantes, sachant les observations et les hypothèses (même si celles-ci ne sont en général pas précisées) (Boreux *et al.* 2010). Par exemple la distribution conditionnelle complète de θ_i est $P_i(\theta_i | \theta_1, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_d, x)$. $x = (x_1, \dots, x_N)$.

Malgré la puissance et l'efficacité des méthodes MCMC évoquées dans cette section, il arrive du fait de la complexité des modèles étudiés (plusieurs niveaux hiérarchiques, lois *a priori* inappropriées ou dimensions variables des espaces d'états *etc.*) que ces outils ne produisent pas les performances escomptées (Parent and Bernier, 2007). Il est fréquent de retrouver dans la littérature des approches combinant les deux méthodes, ou les associant à d'autres techniques de simulation telles que l'échantillonnage pondéré ou « importance sampling » en anglais (Parent and Bernier, 2007; Robert and Casella, 2011).

5.3.2 Bootstrap

La propagation d'incertitude dans un cadre probabiliste utilisant les méthodes usuelles de tirage aléatoire paraît dans de nombreux cas difficilement réalisable. Des techniques de ré-échantillonnage telles que la méthode du bootstrap peuvent être des alternatives crédibles aux méthodes de Monte Carlo classiques. Le bootstrap est une technique de ré-échantillonnage initialement introduite par Efron (1979) puis par Efron et Tibshirani (1993) dans le but d'approcher par simulation (Monte Carlo) la distribution d'un estimateur « plug-in » lorsque l'on ne connaît pas la loi de l'échantillon. On peut également ajouter le livre de Davison et Hinkley (1997) parmi les références sur le bootstrap. Le bootstrap est un outil très efficace, s'apparentant aux méthodes Monte Carlo, ne reposant pas sur des hypothèses lourdes (car c'est une méthode numérique), alternatif aux approximations asymptotiques, pour l'évaluation des estimateurs dans le cadre du faible nombre de données. Parmi les principales applications on peut noter le calcul d'intervalles de confiance, les tests d'hypothèses, la réduction de biais et de la variance (Frey and Rhodes, 1998).

Deux types de bootstrap sont listés ci-après :

- Le bootstrap non-paramétrique s'apparente à du ré-échantillonnage. Différentes techniques de bootstrap non-paramétrique dépendant de l'échantillonnage utilisé peuvent être considérées (bootstrap ordinaire (par tirage aléatoire ou par permutation), bootstrap balancé, bootstrap antithétique).
- Le bootstrap paramétrique pour lequel les échantillons bootstrap sont construits sur la base d'une distribution théorique dont les paramètres de position et de dispersion sont estimés à partir de l'échantillon de données disponibles.

Le principe derrière le bootstrap non-paramétrique repose sur la loi des grands nombres et par application sur la convergence de la fonction de répartition empirique $F_n(\cdot)$ vers la vraie fonction de répartition $F(\cdot)$ lorsque n tend vers l'infini. Considérons un modèle d'évaluation de risque, composé de p paramètres incertains (X_1, \dots, X_p) et d'une sortie Y , définie par $Y = g(X_1, \dots, X_p)$. Supposons que chaque paramètre $X_i : i = 1, \dots, p$ est associé à un échantillon de taille n ($x_i^j : i = 1, \dots, p$ et $j = 1, \dots, n$). L'idée est de réaliser, pour chaque paramètre X_i , un tirage par remise d'un nombre n_b (qui peut être supérieur à n) de données dans l'échantillon ($x_i^j : i = 1, \dots, p$ et $j = 1, \dots, n$) de calculer les sorties $y^{(b)} = \frac{1}{n_b} \sum_{k=1}^{n_b} y^k$ avec $y^k = g(x_1^k, \dots, x_p^k)$ et ($k = 1, \dots, n_b$) et de répéter cette opération un nombre B de fois. Les B valeurs obtenues vont donc constituer l'échantillon bootstrap qui caractérisera l'incertitude sur la sortie du modèle.

Du fait de sa facilité de mise en œuvre et de sa flexibilité, la méthode du bootstrap est appliquée dans des disciplines nombreuses et variées (en traitement du signal, analyse de risque, *etc.*) et notamment dans le contexte de l'analyse des risques environnementaux et sanitaires (Frey and Rhodes, 1998 ; Haas *et al.* 1993 ; Pouillot *et al.* 2007). En particulier, Frey et Rhodes (1998) proposent une application dans laquelle le bootstrap paramétrique est utilisé dans une simulation 2D permettant de séparer l'*incertitude* et la *variabilité*.

5.3.3 Méthode alternative : méthode « hybride »

La méthode de propagation d'incertitude décrite dans cette section est motivée par la caractérisation simultanée de sources d'informations « hybrides » caractérisées d'une part par l'approche probabiliste classique et d'autre part par une approche « extra-probabiliste » (distributions de possibilité en l'occurrence). La méthode présentée ici repose sur la combinaison de la simulation stochastique de type Monte Carlo et du calcul de nombre ou d'intervalles flous (ou *Fuzzy arithmetic* en anglais). Les distributions de possibilité (cf. 5.2.2.3) sont liées aux nombres flous. Il est, en effet, possible de définir une unique distribution de possibilité à partir d'un nombre flou ; ce nombre flou pouvant être considéré comme un emboîtement d'intervalles (de confiance), généralement appelés α -coupes ($\alpha \in [0,1]$ définissant le degré de possibilité d'appartenir à l'intervalle), tel que chacun est associé à un poids caractérisant l'appartenance à l'intervalle en question (Baudrit, 2005 ; Dubois and Guyonnet, 2011). Ainsi, cette méthode de propagation « hybride » permet de prendre en

compte de manière distincte l'*incertitude* due à un manque de connaissance et la *variabilité* (Baudrit, 2005).

Considérons le modèle de type boîte noire $Y = g(X_1, \dots, X_k, X_{k+1}, \dots, X_n)$ composé de n paramètres dont les k premiers sont considérés comme variables et les $n - k$ restants sont considérés comme incertains du fait du manque de connaissance ou de l'imprécision sur l'information les caractérisant. Considérons la loi jointe $(P_1(X_1), \dots, P_k(X_k))$ définie sur les k paramètres variables et les distributions de possibilité $(\pi_{X_{k+1}}, \dots, \pi_{X_n})$ définies sur les paramètres incertains restants. La simulation hybride proposée par (Guyonnet *et al.* 2003 ; Baudrit *et al.* 2005) est basée sur deux étapes : une étape de simulation de type Monte Carlo permettant de propager la variabilité des k paramètres et une étape d'arithmétique floue pour propager l'incertitude sur les $n - k$ paramètres incertains restants. La méthode hybride est un processus itératif décrit par la Fig. 5.9. Une illustration schématique de la méthode et du type de résultat attendu est proposée par la Fig. 5.10.

Méthode « Hybride »

1. Définir un ensemble de niveaux de possibilité $\alpha_i \in [0,1]$ de taille n_α
 - Par exemple $\alpha = \begin{cases} \alpha_1 = 0.2 \\ \alpha_2 = 0.6 \\ \alpha_3 = 0.9 \end{cases}$ ou assigner des valeurs de 0 à 1 avec un pas de 0.1
2. Générer aléatoirement N réalisations de la loi jointe $(P_1(X_1), \dots, P_k(X_k))$
 - On obtient une matrice $M(x_1, \dots, x_k)$ de nombres réels de dimension $N \times k$
3. Choisir un niveau de possibilité α_i dans l'ensemble défini en (1.)
Choisir les α -coupes correspondantes pour les $n - k$ paramètres incertains
 - On obtient $n - k$ α -coupes
4. Déterminer les valeurs *inf* et *sup* des sorties du modèle en combinant les $n - k$ α -coupes (2^{n-k} combinaisons possibles) et les lignes de la matrice de réalisations aléatoires $(M(x_1^j, \dots, x_k^j): j = 1, \dots, N)$.
 - Valeur inf : $L(y_{\alpha_i}^j) = \min(g(x_1^j, \dots, x_k^j, x_{k+1}^{\alpha_i}, \dots, x_n^{\alpha_i}))$
 - Valeur sup : $U(y_{\alpha_i}^j) = \max(g(x_1^j, \dots, x_k^j, x_{k+1}^{\alpha_i}, \dots, x_n^{\alpha_i}))$
 - On obtient deux vecteurs de taille N avec les valeurs *inf* et *sup*
5. Répéter les étapes (3.) et (4.) pour tous les niveaux de possibilité α_i
 - On obtient ainsi une famille de N distributions de possibilité (générant au total $N \times n_\alpha$ intervalles flous)
6. Attribuer à chaque intervalle flou un poids égal à $(1/N \times (n_\alpha - 1))$.

Fig. 5.9: Description de la méthode itérative de propagation "hybride".

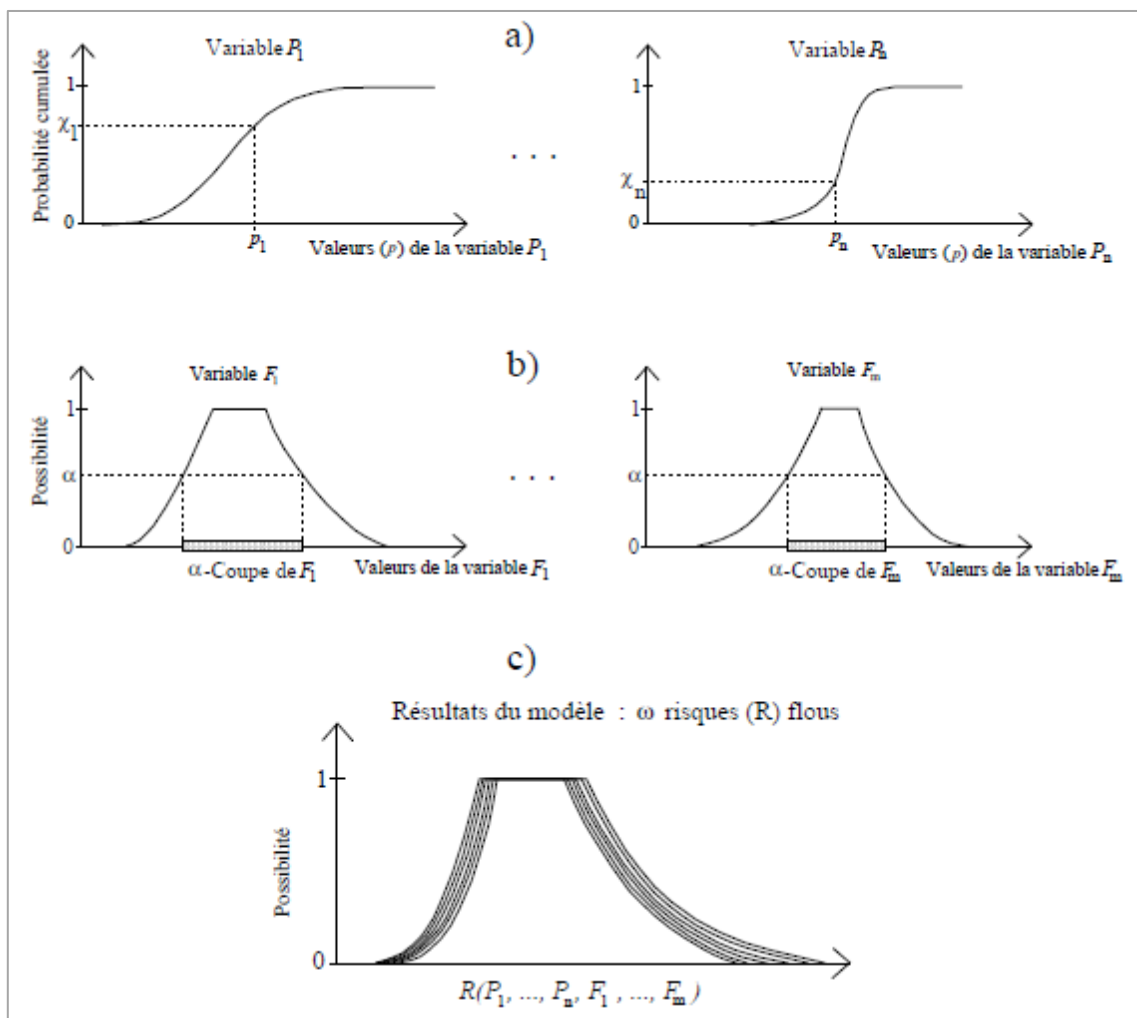


Fig. 5.10: Illustration schématique de la méthode hybride et présentation du résultat obtenu.
 Source (Guyonnet *et al.* 2005).

Le résultat obtenu se présente sous la forme d'une famille de distributions de possibilité que l'on peut qualifier « d'originale » (*i.e.* pas classique) (cf. Fig. 5.10 - (c)). Le résultat du calcul hybride s'appelle un intervalle flou aléatoire (Guyonnet *et al.* 2005). Une étape importante repose alors sur le développement d'une méthode qui permette de restituer les résultats de manière compréhensible pour un public étranger à ce genre d'approches. Baudrit (2005) et Guyonnet *et al.* (2005) proposent un post-traitement utilisant les fonctions de croyance de la théorie de Dempster-Shafer. En effet, Baudrit *et al.* (2004, 2005) ont montré comment les fonctions de croyance de Dempster-Shafer s'appliquaient naturellement au problème de la comparaison du résultat hybride à un critère de tolérance (Guyonnet *et al.* 2005). Cette méthode de post-traitement utilisée par (Pasanisi, 2014) est décrite ci-après.

En pratique, le résultat produit par la méthode « hybride » consiste en q distributions de possibilité $(\pi^{(i)}: i = 1, \dots, q)^4$; chacune étant définie à l'aide des n α -coupes emboîtées. Plaçons-nous dans le cadre d'une évaluation de risque où l'on s'intéresse à la probabilité ou non de dépasser un seuil y^* . Les intervalles de type $] - \infty; y^*]$ (ou inversement $[y^*; +\infty[$) sont les ensembles d'intérêt. Il est possible de définir q mesures de nécessité et de possibilité telles que :

$$\begin{cases} \Pi^{(i)}(] - \infty; y^*]) = \sup_{y \in] - \infty; y^*]} \pi^{(i)}(y) \\ N^{(i)}(] - \infty; y^*]) = \inf_{y \notin] - \infty; y^*]} \pi^{(i)}(y) \end{cases} \quad (5.27)$$

Les paires de mesures de possibilité-nécessité obtenues permettent ensuite de définir, sur la base des mesures duales de crédibilité et de plausibilité (*Bel* et *Pl* respectivement) de la théorie de Dempster-Shafer, un encadrement de la probabilité (inconnue ou imprécise) d'être inférieur au seuil y^* , $P(Y \leq y^*)$ décrit par :

$$Bel(] - \infty; y^*]) \leq P(Y \leq y^*) \leq Pl(] - \infty; y^*]) \quad (5.28)$$

avec

$$\begin{cases} Pl(] - \infty; y^*]) = \frac{1}{q} \sum_{i=1}^q \Pi^{(i)}(] - \infty; y^*]) \\ Bel(] - \infty; y^*]) = \frac{1}{q} \sum_{i=1}^q N^{(i)}(] - \infty; y^*]) \end{cases} \quad (5.29)$$

En résumé, la méthode « hybride » introduite dans cette section apporte une réponse simple au problème de la combinaison d'informations de natures aléatoire et imprécise dans l'estimation du risque (Guyonnet *et al.* 2005). D'autres méthodes sont également envisageables lorsque l'on souhaite combiner des informations de natures différentes. Baudrit (2005) propose dans ses travaux trois méthodes en plus de la méthode hybride.

- La méthode de *Propagation de p-box sans hypothèse de dépendances* (“*Dependency Bounds Convolution*” en anglais) utilisant les paires de probabilités haute et basse ;
- La méthode des *Ensembles Aléatoires Indépendants* (“*Independent Random Sets*” en anglais) englobant possibilité et probabilités dans le cadre des fonctions de croyance ;
- La méthode de *Propagation d'ensembles aléatoires sans hypothèse de dépendances* (“*Conservative Random Sets*” en anglais) dans le cadre des fonctions de croyance.

⁴ Nous choisissons ici q pour éviter la confusion avec la mesure de *nécessité* N .

L'ensemble de ces méthodes offrent aux modélisateurs et aux décideurs impliqués dans les analyses de risque (toutes applications confondues) un panel intéressant d'alternatives aux méthodes de simulation de Monte Carlo 1D et 2D auxquelles il est principalement reproché une prise en compte inadéquate de l'*incertitude*. L'*incertitude* et la *variabilité* sont confondues dans le cas du MC1D. De plus, le MC2D présenterait selon Baudrit (2005) des limites car dans le cas d'un modèle paramétrique gaussien avec μ et σ les paramètres de la distribution représentant respectivement l'*incertitude* et la *variabilité*, la simulation MC2D suppose l'indépendance entre les deux paramètres ; ce qui ne refléterait pas l'*incertitude* sur la sortie du modèle. La citation ci-dessous illustre ce point.

« De plus, comment prendre en compte la structure des dépendances entre les variables aléatoires associées aux paramètres θ . Par exemple, si un des modèles paramétriques est une loi normale $\mathcal{F}(\mu, \sigma)$, comment justifier l'(in)dépendance entre μ et σ ? Il paraît illusoire de penser que le 2D Monte-Carlo permet de prendre en compte l'information de nature imprécise. Il fournit un ensemble de m distributions de probabilité autant que de valeurs distinctes de θ choisies. Certains auteurs ont voulu synthétiser cet ensemble en une probabilité unique ce qui revient à confondre imprécision et variabilité, chose que nous ne voulons pas faire. Ainsi, il est plus cohérent d'utiliser les p -boxes pour représenter ce type d'information. » (Baudrit, 2005).

Box 5.2 : Notes aux lecteurs

Dans ces travaux nous nous sommes principalement placés dans un cadre probabiliste de propagation des incertitudes. Ce choix s'explique, globalement, par le fait que les différentes sources d'incertitude et de variabilité sont représentées à travers les variations possibles des valeurs des paramètres (radioécologiques) considérés.

- Dans l'article présenté au **Chapitre 8** les simulations sont réalisées suivant la méthode **MC1D**.
- Dans l'article présenté au **Chapitre 9** les simulations sont réalisées suivant la méthode **MC1D** en chaque point du paysage, la variabilité spatiale est donc considérée comme une deuxième dimension ; ce qui s'assimile à du **MC2D**.
- En **Annexe D**, nous présentons un exemple illustrant la propagation des incertitudes suivant les méthodes **MC1D**, **MC2D** et **Hybride**.

5.4 Méthodes d'analyse de sensibilité

L'analyse de sensibilité vise fondamentalement à identifier dans la modélisation d'un phénomène simple ou complexe les facteurs pour lesquels les fluctuations et/ou l'imprécision influencent le comportement du phénomène et donc son incertitude et/ou sa variabilité. Nous nous intéressons dans le cadre de l'analyse de sensibilité aux « paramètres dont l'incertitude et/ou la variabilité influencent l'incertitude et/ou la variabilité sur la sortie du modèle ». Cette formulation étant assez lourde il nous arrivera de juste évoquer des « paramètres influents ».

Les objectifs formels d'une analyse de sensibilité sont d'identifier les facteurs les plus (et également les moins) influents, de les hiérarchiser en fonction de leur influence sur le phénomène pour finalement orienter l'effort d'acquisition de nouvelles connaissances en conséquence. Il existe une grande diversité de méthodes d'analyse de sensibilité appliquées dans des domaines et des disciplines particulièrement variés tels que l'économie, la finance, la physique, l'hydraulique, l'écologie, la sûreté et la fiabilité des installations nucléaires ou l'analyse des risques environnementaux *etc.* (Ellouze, 2009) : on peut citer par exemple les méthodes graphiques, d'analyse par seuil, utilisant les arbres de régression, par approche bayésienne, fiabilistes de type FORM et SORM ou celles basées sur la décomposition de la variance ou sur la régression linéaire *etc.* Dans le contexte de l'évaluation des risques alimentaires, ces méthodes sont précieuses et de plus en plus utilisées du fait du nombre de paramètres à prendre en compte, et des différents scénarios d'exposition possibles entre autres.

Nous distinguons principalement dans ce projet trois approches d'analyse de sensibilité pouvant regrouper un certain nombre de méthodes : les approches locale, de criblage et globale. Dans l'optique d'atteindre les objectifs mentionnés ci-dessus, ces trois approches reposent sur le calcul de mesures quantitatives (ou d'indices) de la sensibilité permettant de caractériser l'influence des variations d'un paramètre sur les variations de la sortie. Dans la suite, nous introduirons ces trois différentes approches et les principales méthodes associées en explorant de manière un peu plus détaillée la méthode de Morris et les méthodes globales utilisées dans le cadre de ces travaux de thèse.

Pour mémoire, nous considérerons à travers le mot « modèle » une boîte noire de type $Y = g(X_1, \dots, X_p)$ où Y représente la sortie du modèle, $g(\cdot)$ la boîte noire et $(X_i : i = 1, \dots, p)$ l'ensemble des paramètres incertains et/ou variables.

5.4.1 Méthodes locales d'analyse de sensibilité

Les méthodes locales d'analyse de sensibilité reposent sur l'évaluation des variations de la sortie du modèle induites par les variations locales des paramètres d'entrée. De manière plus précise, une analyse de sensibilité locale consiste à faire varier chaque paramètre d'entrée au voisinage d'une valeur nominale et ensuite évaluer l'impact de cette variation sur la sortie en calculant les dérivées partielles de la sortie en fonction de chaque paramètre.

De manière formelle, les méthodes d'analyse de sensibilité locales proposent d'estimer autour du point de référence $x^0 = (x_1^0, \dots, x_p^0)$ les dérivées partielles définies ci-dessous et de les classer ; les paramètres les plus influents ont alors les dérivées partielles les plus élevées.

$$S_i = \frac{\partial g}{\partial X_i|_{X=x^0}} \text{ avec } i = 1, \dots, p \quad (5.30)$$

Il est important de noter que ces dérivées partielles sont dépendantes des unités dans lesquelles sont mesurés les paramètres, unités qui peuvent être différentes. Il faut alors calculer un indice de sensibilité locale normalisé (adimensionné) défini par :

$$\widehat{S}_i = \frac{x_i^0}{g(x^0)} S_i = \frac{x_i^0}{g(x^0)} \frac{\partial g}{\partial X_i|_{X=x^0}} \quad (5.31)$$

Les méthodes d'analyse de sensibilité ont notamment été utilisées en cinétique chimique où il est classique d'étudier l'influence des variations de paramètres cinétiques autour d'une valeur donnée (Da Veiga, 2007). L'avantage des méthodes locales d'analyse de sensibilité réside dans le fait que l'on peut réaliser une exploration préliminaire du modèle avec un temps de calcul réduit ; ce qui peut être utile lorsque les analyses globales nécessitent des temps de calcul coûteux (Ciric, 2012). Cependant, puisque ces méthodes sont construites de sorte à étudier pour chaque paramètre l'impact des variations autour du voisinage de valeurs nominales, deux limites peuvent alors être identifiées. En effet, la sensibilité des méthodes locales aux valeurs nominales (des valeurs nominales mal connues ou peu pertinentes) et une

exploration limitée du support des paramètres d'entrée peuvent avoir un impact sur les résultats et donc les conclusions des analyses locales de sensibilité.

5.4.2 Méthodes de criblage

Les méthodes de criblage ("*screening methods*" en anglais) permettent d'explorer rapidement le comportement des sorties d'un code de calcul coûteux en faisant varier un grand nombre de ses paramètres d'entrée (Iooss and Mahévas, 2013). Ces variations sont basées sur la discrétisation des paramètres d'entrée en plusieurs valeurs nommées niveaux (Iooss and Mahévas, 2013). On parle alors de méthodes déterministes car elles ne nécessitent pas de spécifier les distributions de probabilité des paramètres (Iooss, 2011). Les méthodes de criblage présentent l'intérêt d'avoir un faible coût de calcul ; ce qui en fait des outils particulièrement utiles pour des modèles avec un très grand nombre de paramètres et lorsque l'on souhaite avoir un aperçu préliminaire des paramètres influents mais également ceux qui influent le moins. On peut distinguer plusieurs méthodes regroupées en trois classes par Iooss and Mahévas, (2013) :

- Les méthodes de criblage à très grande dimension regroupant des méthodes telles que les plans *supersaturés*, la méthode de criblage par groupe et la méthode des bifurcations séquentielles (Iooss, 2011). Les méthodes de criblage par groupe et les bifurcations séquentielles consistent à identifier itérativement les paramètres influents en créant un certain nombre de groupes (2 pour les bifurcations séquentielles).
- Les plans factoriels usuels comprenant les plans factoriels fractionnaires, les plans de résolution trois et quatre (RIII et RIV respectivement), les plans one-at-a-time (OAT).
- La méthode de Morris qui est présentée dans la suite.

5.4.2.1 La méthode de Morris

La méthode de Morris a pour objectif d'examiner les variations dans les valeurs d'une sortie qui sont attribuées de manière non ambiguë aux variations des paramètres d'entrée (Morris, 1991). La méthode de Morris repose sur le calcul, pour chaque paramètre d'entrée X_i , du nombre de ratios entre l'effet d'un incrément sur la sortie et l'incrément choisi. Ces ratios sont appelés *effet élémentaire* par Morris (1991). Pour chaque paramètre, ces effets élémentaires sont ensuite moyennés en vue d'évaluer l'importance de chaque paramètre

d'entrée (Campolongo *et al.* 2007). La méthode de Morris consiste d'abord à définir un plan d'expérience de type OAT (cf. Section 5.4.2) dans lequel chaque paramètre varie une fois dans son domaine de variation discrétisé. L'étape suivante consiste à calculer les effets élémentaires pour chaque paramètre en répétant un nombre défini de fois (r) les plans d'expérience OAT en définissant aléatoirement un point de départ à chaque répétition. Cette approche aléatoire a été suggérée par Morris (1991) en vue de tenir compte des interactions lorsque deux ou plusieurs paramètres varient simultanément (Girard *et al.* 2014).

Considérons le modèle défini précédemment. L'effet élémentaire pour le paramètre X_i à la répétition k considérant un incrément δ est défini par :

$$EE_i^k = \frac{g(X_1^k, \dots, X_i^k + \delta, \dots, X_p^k) - g(X_1^k, \dots, X_i^k, \dots, X_p^k)}{\delta} \quad (5.32)$$

Une fois les effets élémentaires calculés pour chaque paramètre, l'analyse de sensibilité repose sur le calcul de statistiques descriptives à partir de la distribution des valeurs des effets élémentaires. Morris (1991) propose, comme mesure de la sensibilité de la sortie du modèle à chaque paramètre, d'estimer la moyenne arithmétique μ_i et l'écart-type σ_i définis par :

$$\begin{cases} \mu_i = \frac{1}{r} \sum_{k=1}^r EE_i^k \\ \sigma_i = \sqrt{\frac{1}{r} \sum_{k=1}^r (EE_i^k - \mu_i)^2} \end{cases} \quad (5.33)$$

La moyenne μ_i évalue l'influence globale du paramètre X_i sur la sortie et l'écart-type σ_i permet d'estimer la totalité des effets d'ordre supérieur (*i.e.* les interactions entre paramètres d'entrée ou la non-linéarité) (Nicoulaud-Gouin *et al.* 2014).

Sur la base des valeurs μ_i et σ_i estimées pour chaque paramètre $X_i : i = 1, \dots, p$ la méthode de Morris permet d'identifier trois groupes de paramètres d'entrée :

- Les paramètres avec un effet négligeable caractérisés par une faible valeur μ_i
- Les paramètres avec un effet linéaire mais sans effet d'interaction caractérisés par une valeur faible de σ_i et une valeur élevée de μ_i et
- Les paramètres ayant un effet non-linéaire et/ou un effet d'interaction correspondant à des valeurs élevées de μ_i et σ_i .

Pour un paramètre X_i , lorsque des valeurs tantôt négatives tantôt positives des effets élémentaires sont estimées (comme par exemple pour des modèles non-monotones) la valeur de l'indice de sensibilité μ_i peut alors être proche de 0 ; ce qui suppose probablement à tort que le paramètre considéré n'influence pas la sortie. [Campolongo et al. \(2007\)](#) ont introduit, comme une alternative dans le cas de modèles non-monotones ([Girard et al. 2014](#)), l'indice μ_i^* définissant, pour chaque paramètre, la moyenne arithmétique des valeurs absolues des effets élémentaires comme décrit ci-dessous.

$$\mu_i^* = \frac{1}{r} \sum_{k=1}^r |EE_i^k| \quad (5.34)$$

Une analyse simultanée des indices μ_i et μ_i^* est généralement recommandée pour évaluer l'influence des entrées sur la sortie car le signe (négatif) n'intervient plus lorsque μ_i^* est estimé ([Campolongo et al. 2007](#)). On peut également considérer dans l'analyse des effets, les intervalles de confiance sur les indices de Morris produits par calcul bootstrap.

L'un des avantages de la méthode de Morris se situe dans un coût de calcul faible nécessitant une évaluation du modèle par répétition pour chaque effet élémentaire ([Brevault et al. 2013](#)). L'intérêt de la méthode de Morris, en plus d'un coût de calcul faible, est qu'elle s'applique à n'importe quel type de modèle (*i.e.* modèles non-monotones); ce qui en fait une méthode particulièrement intéressante pour l'analyse de modèles complexes. La méthode de Morris comme les autres méthodes de criblage est particulièrement intéressante lorsque le nombre de paramètres est plutôt élevé ([Girard et al. 2014](#) ; [King and Perera, 2013](#) ; [Ruano et al. 2012](#)). Elle permet notamment de faire une analyse préliminaire permettant d'identifier les paramètres non influents et ainsi de les fixer pour une analyse plus fine.

5.4.3 Méthodes globales d'analyse de sensibilité

Contrairement aux méthodes locales d'analyse de sensibilité, les méthodes globales ont pour objectif d'évaluer l'influence des paramètres d'entrée sur la sortie du modèle en explorant pour chaque paramètre tout le support. De plus, les méthodes globales calculant des indices de sensibilité ou mesures d'importance permettent de proposer une hiérarchisation précise des paramètres influents contrairement aux méthodes de criblage qui hiérarchisent surtout les entrées non influentes ([Iooss and Mahévas, 2013](#)). Les approches globales permettent d'identifier les interactions entre les paramètres dans les modèles non-linéaires

et/ou non-additifs car elles évaluent l'effet d'un paramètre sur la sortie lorsque tous les autres paramètres varient (Cariboni *et al.* 2007). Dans la suite nous nous intéressons aux méthodes basées sur la régression linéaire et celles basées sur la décomposition de la variance.

5.4.3.1 Méthodes basées sur la régression linéaire

Considérons le même modèle de type boîte noire $Y = g(X_1, \dots, X_p)$ défini par p paramètres d'entrée. Soit q le nombre de simulations de Monte Carlo et (x_1^j, \dots, x_p^j) la $j^{\text{ème}}$ réalisation aléatoire issue de la distribution jointe des paramètres X_i ($i = 1, \dots, p$). Plaçons-nous dans le cadre d'une propagation d'incertitude de type MC1D (cf. Section 5.3.1.2) et calculons pour chaque réalisation de la loi jointe des paramètres d'entrée $(x_1^j, \dots, x_p^j) : j = 1, \dots, q$ la sortie $y^j : j = 1, \dots, q$ associée. On obtient l'échantillon (y^1, \dots, y^q) caractérisant les q réalisations de la sortie Y .

Les méthodes d'analyse de sensibilité basées sur la régression linéaire visent à estimer la sortie par une régression linéaire sur les paramètres d'entrée *i.e.* d'approcher la variable de sortie par une combinaison linéaire des paramètres d'entrée (Circic, 2012 ; Frey and Patil, 2002). Pour chaque sortie la combinaison linéaire des paramètres d'entrée est définie par :

$$y^j = \beta_0 + \sum_{i=1}^p \beta_i \times x_i^j + \varepsilon^j \quad (5.35)$$

Les quantités $(\beta_0, \dots, \beta_p)$ désignent les coefficients de la régression linéaire et $\varepsilon^j : j = 1, \dots, q$ les erreurs résiduelles. L'effet des entrées sur la sortie peut être analysé, à partir de la relation linéaire définie par l'équation (5.35), sur la base des coefficients de régression, de l'erreur standard sur les coefficients de régression et sur le niveau de significativité des coefficients de régression estimés (Frey and Patil, 2002).

Il est d'usage de vérifier au préalable si la relation entre la sortie et les paramètres d'entrée est linéaire (Frey and Patil, 2002) en évaluant la valeur du coefficient de détermination R^2 (Saporta, 2006). Dans le cas d'une relation linéaire trois indices de sensibilité (ou mesures d'importance) peuvent être retenus :

- Le *coefficient de corrélation linéaire de Pearson* (introduit dans la Section 5.2.1.1.5-(a)) calculé entre la sortie et chaque paramètre d'entrée et défini par $CC = \rho(Y, X_i) : i =$

$1, \dots, p$. Nous ne décrivons pas le calcul car la formule est décrite dans la [Section 5.2.1.1.5-\(a\)](#).

- Le *coefficient de régression standard* (SRC) calculé entre la sortie et chaque paramètre d'entrée et défini par $SRC(Y, X_i) = \beta_i \sqrt{\frac{Var(X_i)}{Var(Y)}}$ avec $i = 1, \dots, p$ et β_i le coefficient de régression issu de l'ajustement linéaire décrit par l'équation (5.35).
- Le *coefficient de corrélation partielle* (PCC) calculé entre la sortie et chaque paramètre d'entrée et défini par $PCC(Y, X_i) = \rho(Y - \widehat{Y^{(-i)}}, X_i - \widehat{X}_i)$, où $\widehat{Y^{(-i)}}$ désigne la prévision du modèle linéaire ne tenant pas compte du paramètre X_i et \widehat{X}_i est la prévision du modèle linéaire qui exprime le paramètre X_i en fonction des autres paramètres d'entrée.

Lorsque la relation entre la sortie et les paramètres d'entrée n'est pas linéaire mais monotone, LES coefficients présentés ci-dessus doivent être préférentiellement calculés pour un ajustement linéaire sur les rangs et non plus sur les valeurs des paramètres. Cela revient, par exemple pour le coefficient de corrélation linéaire de Pearson, à calculer le coefficient de corrélation de Spearman (cf. [5.2.1.1.5-\(b\)](#)) entre la sortie et chaque paramètre d'entrée.

L'avantage des méthodes d'analyse de sensibilité basées sur la régression linéaire se situe dans la facilité de leur mise en application, et dans un coût de calcul faible si l'on ne tient pas compte de la phase de simulation de type MC1D. Ces méthodes reposent sur une hypothèse de linéarité ou *a minima* de monotonie entre la sortie et les paramètres d'entrée. Ce n'est, cependant, pas souvent le cas l'on modélise des systèmes complexes dans lesquels la relation entre paramètres peut être non-linéaire et où les interactions potentielles entre les paramètres d'entrée peuvent ajouter un niveau de complexité supplémentaire. Dans ce contexte il est intéressant de privilégier des méthodes indépendantes de la structure du modèle telles que les méthodes basées sur la décomposition de la variance que nous introduisons dans la section suivante.

5.4.3.2 Méthodes basées sur la décomposition de la variance

Les méthodes d'analyse de sensibilité basées sur la décomposition de la variance sont des méthodes indépendantes des modèles (*i.e.* de leur structure) qui permettent d'évaluer l'importance des effets des paramètres d'entrée sur la sortie en proposant une décomposition

de la variance de la sortie du modèle. Cette décomposition de la variance repose sur la méthode de décomposition fonctionnelle introduite par [Hoeffding \(1948\)](#) et par la suite étendue à l'analyse de sensibilité par [Sobol \(1993\)](#) qui démontra son unicité ([Iooss and Mahévas, 2013](#)). Si les paramètres d'entrée sont aléatoires et mutuellement indépendantes on obtient la décomposition de variance suivante :

$$Var(Y) = \sum_{i=1}^p V_i(Y) + \sum_{i<j} V_{ij}(Y) + \sum_{i<j<k} V_{ijk}(Y) + \dots + V_{12\dots p}(Y) \quad (5.36)$$

où

$$V_i(Y) = Var[\mathbb{E}(Y|X_i)] \text{ et } V_{ij}(Y) = Var[\mathbb{E}(Y|X_i X_j)] - V_i(Y) - V_j(Y) \quad (5.37)$$

On peut alors définir des indices de sensibilité, ou mesures d'importance basées sur la variance appelés indices de Sobol, à partir des équations (5.36) et (5.37). On distingue alors des indices de sensibilité du premier ordre, du deuxième ordre (pour les interactions entre deux paramètres), du troisième ordre (interactions entre trois paramètres) *etc.* Ces indices de sensibilité sont définis par :

$$\left\{ \begin{array}{l} SI_i = \frac{Var[\mathbb{E}(Y|X_i)]}{Var(Y)} = \frac{V_i(Y)}{Var(Y)} \\ SI_{ij} = \frac{V_{ij}(Y)}{Var(Y)} \\ SI_{ijk} = \frac{V_{ijk}(Y)}{Var(Y)} \\ \vdots \\ SI_{12\dots p} = \frac{V_{12\dots p}(Y)}{Var(Y)} \end{array} \right. \quad (5.38)$$

Les indices de sensibilité décrits ci-dessus sont compris entre 0 et 1 et leur somme vaut 1. Le nombre d'indices augmentant avec le nombre de paramètres (le nombre d'indices est égal à $2^K - 1$), [Homma et Saltelli \(1996\)](#) ont introduit les indices de sensibilité totaux représentant pour un paramètre d'entrée X_i la somme des indices de sensibilité de tout ordre impliquant le paramètre X_i . Les indices totaux sont définis par :

$$TSI_i = \sum_{k \neq i} SI_k = SI_i + \sum_{j \neq i} SI_{ij} + \sum_{j \neq i, k \neq i, j < k} SI_{ijk} + \dots \quad (5.39)$$

Le calcul des indices de Sobol peut nécessiter un nombre élevé d'explorations du modèle ; ce qui rend inappropriées les méthodes de Monte Carlo initialement développées par [Sobol \(1993\)](#) et par [Saltelli \(2002\)](#) pour leur estimation ([Iooss, 2011](#)). Différentes méthodes alternatives ont été développées en vue de réduire le coût de l'estimation des indices de Sobol. Les méthodes quasi Monte Carlo permettent de réduire d'un facteur 10 ce coût ([Iooss, 2011](#)).

Les méthodes FAST (Cukier *et al.* 1978) et EFAST (Saltelli *et al.* 1999) basées sur l'utilisation de la transformée de Fourier constituent des méthodes encore moins coûteuses que les approches de type Monte Carlo (Iooss, 2011). L'utilisation de plans d'expérience optimisés (Ellouze, 2009) ou d'approches non-paramétriques (Mara *et al.* 2015) peut également être une alternative intéressante.

L'existence de dépendance entre paramètres d'entrée constitue, en plus du coût élevé de l'estimation des indices de Sobol, l'une des principales difficultés dans les analyses de sensibilité (globales). La question des paramètres non indépendants est un problème d'actualité pour lequel de plus en plus d'applications et méthodes sont développées (Da Veiga *et al.* 2009; Mara *et al.* 2015 ; Saltelli and Tarantola, 2002).

5.4.4 Conclusion sur les méthodes d'analyse de sensibilité

Nous avons introduit dans cette section un ensemble d'approches et de méthodes permettant d'étudier l'effet des variations des paramètres d'entrée sur la sortie du modèle considéré. Les méthodes de criblage et les méthodes locales présentent l'intérêt de ne pas être excessivement gourmandes en temps de calcul, offrant ainsi la possibilité de faire un tour d'horizon préliminaire pour identifier les paramètres non influents et les fixer à leurs valeurs nominales. Cependant, les méthodes locales dépendent fortement des valeurs nominales et ne permettent pas d'explorer les domaines des paramètres d'entrée dans leur intégralité. Les méthodes globales sont les méthodes le plus souvent privilégiées car elles explorent en entier, contrairement aux méthodes locales, les supports des paramètres d'entrée. Les méthodes basées sur la régression linéaire offrent de très bons résultats lorsque la relation entre la sortie et les paramètres d'entrée est soit linéaire soit monotone. Ces méthodes souffrent néanmoins du fait qu'il est rare d'observer ce type de relations dans les modèles d'évaluation de risque environnemental, sanitaire et alimentaire. Les méthodes basées sur la décomposition de la variance offrent un champ plus large d'applications car elles sont indépendantes de la structure des modèles. En revanche, elles sont particulièrement coûteuses en temps de calcul ce qui peut limiter leur attrait.

La diversité des approches et des méthodes d'analyse de sensibilité peut perdre le modélisateur dans sa recherche d'une méthode fiable et surtout pertinente au regard du

modèle considéré. Dans le cas des méthodes globales, le choix d'une approche requiert dans un premier temps d'identifier les objectifs de l'analyse du modèle et la précision recherchée dans le diagnostic (Iooss and Mahévas, 2013). Plusieurs facteurs incluant le nombre et la nature (discrète ou continue) des entrées et le temps de calcul nécessaire à une exploration du modèle doivent être pris en compte lorsque l'on souhaite choisir une méthode d'analyse de sensibilité. A titre d'exemple, Ciric (2012) propose une comparaison entre les méthodes de Morris et EFAST dans ses travaux de recherche. Plusieurs auteurs ont proposé des règles d'aide à la décision offrant aux modélisateurs la possibilité de choisir en fonction des particularités citées ci-dessus la méthode idoine (Cariboni *et al.* 2007 ; De Rocquigny *et al.* (2008) ; Iooss et Lemaitre, 2014). Une synthèse des méthodes d'analyse de sensibilité orientée par le coût de calcul, indexé par le nombre de paramètres et la complexité du modèle, est proposée ci-dessous par Iooss et Lemaitre (2014).

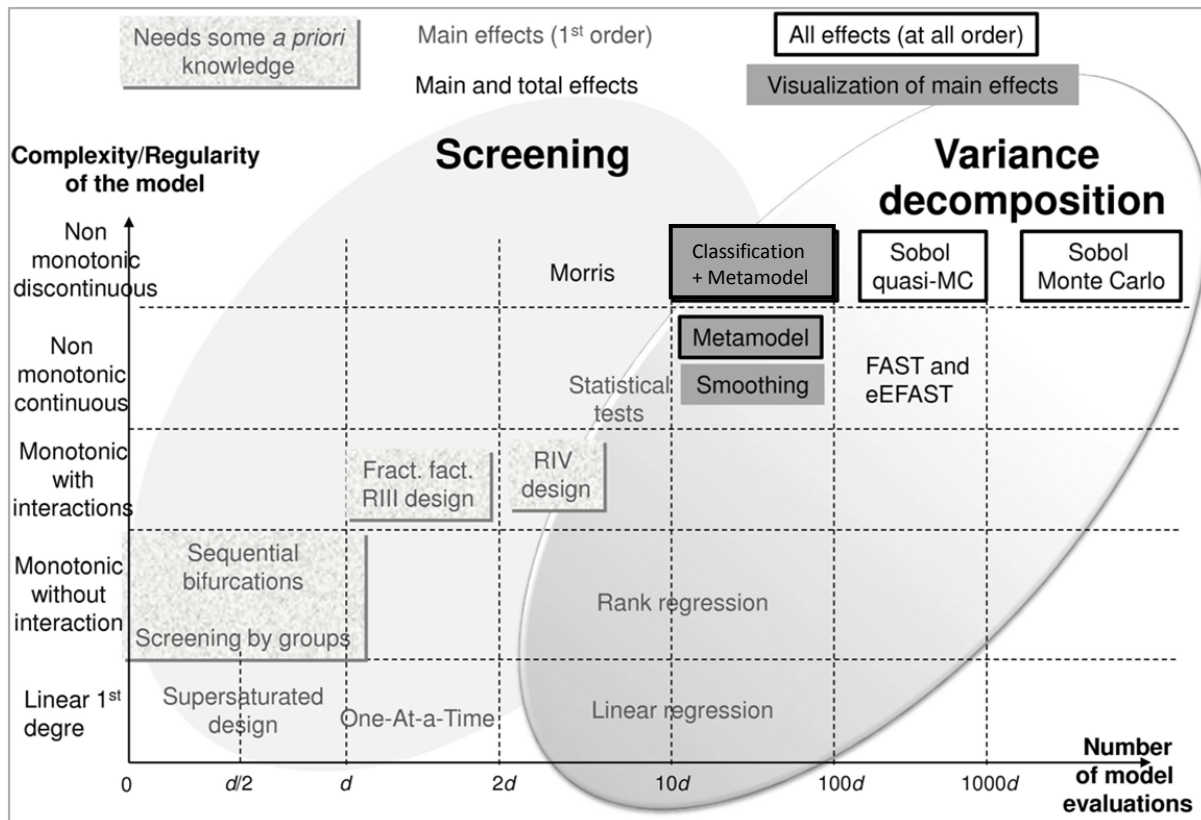


Fig. 5.11 : Synthèse des méthodes d'analyse de sensibilité issue de (Iooss et Lemaitre, 2014).

Box 5.3 : Notes aux lecteurs

- Dans le cadre des cas d'applications réalisés dans ce projet de recherche (**Chapitres 8 et 9**) nous nous sommes principalement focalisés sur l'utilisation de la **méthode de Morris** et sur le calcul des mesures d'importance issues de la régression linéaire (principalement le **coefficient de corrélation de Spearman**).
- Notre positionnement a globalement été d'utiliser les méthodes d'analyse de sensibilité se prêtant bien à nos applications en termes de temps de calcul, de complexité des modèles et du nombre de paramètres incertains/variables.
- Les temps de calcul induits par les méthodes de décomposition de la variance et l'estimation des indices de Sobol combinés aux temps de calcul propres à la plateforme SYMBIOSE rendent difficile voire impraticable l'utilisation de ces méthodes dans notre cas.

Partie II : Quantification des incertitudes sur les paramètres radioécologiques

Dans cette partie, nous nous intéressons à la modélisation de l'incertitude et de la variabilité sur des paramètres radioécologiques gouvernant l'interception et la rétention des substances radioactives émises en situation accidentelle dans l'atmosphère. Ces paramètres radioécologiques sont le facteur d'interception sèche, le facteur d'interception humide et la constante de décroissance biomécanique introduits dans les [Sections 3.1.2.1](#) et [3.1.2.2](#) du [Chapitre 3](#).

Cette modélisation a nécessité de collecter le plus grand nombre de données à notre portée, que ce soit dans les études publiées ou la littérature grise de l'IRSN et du CEA. Dans cette modélisation nous adoptons principalement une approche de type méta-analyse permettant de combiner des données issues de plusieurs sources. Le recours à une approche de type méta-analyse, de plus en plus fréquent en évaluation des risques, permet d'intégrer et d'estimer efficacement différentes sources de variabilité issues de la combinaison de plusieurs études ([Rigaux et al. 2013](#)) et convient bien à une modélisation bayésienne hiérarchique. Dans le contexte de la radioécologie, [Hosseini et al. \(2013\)](#) proposent une approche permettant d'utiliser différentes sources de données dans un cadre bayésien pour estimer les paramètres d'un modèle quelconque. Cette approche vise à dresser une méthodologie suivant quatre situations typiques dans lesquels les modélisateurs peuvent se retrouver ([Hosseini et al. 2013](#)).

- Le modélisateur dispose de suffisamment de données représentatives issues de l'étude sur le phénomène d'intérêt. Dans ce cas, il est possible de déduire leur distribution et d'estimer les paramètres de celle-ci par la méthode du maximum de vraisemblance.
- Le modélisateur ne dispose ni de données représentatives, ni de données de la littérature. Une approche d'extrapolation peut dans ce cas être utilisée pour déduire la loi de distribution en rapport avec le cas d'intérêt. Pareille approche est proposée par [Brown et al. \(2013\)](#), en radioécologie, pour déduire des facteurs de concentration pour les animaux sauvages à partir de données sur des animaux domestiques de la même espèce (*e.g.* sangliers et cochons).
- Un nombre limité de données représentatives est à disposition du modélisateur pour le cas d'étude considéré et aucune donnée subsidiaire en relation avec le cas d'intérêt n'est disponible. [Hosseini et al. \(2013\)](#) préconisent de déduire les distributions des paramètres à partir des données à disposition en leur associant une large incertitude. Les auteurs précisent, néanmoins qu'il est rare de ne trouver aucune donnée ou information pouvant être exploitées en sus des données du cas d'intérêt.

- Enfin il existe des situations où le modélisateur dispose de peu (ou pas) de données pour le cas d'étude mais a la possibilité de trouver, dans la littérature ou à partir d'extrapolation, des données en lien avec le cas d'étude. Dans ce cas précis, [Hosseini et al. \(2013\)](#) proposent différents modèles bayésiens considérant l'information additionnelle comme partie intégrante du jeu de données ou non.

Les trois études présentées dans les trois chapitres suivants s'inscrivent globalement dans la dernière situation. Les données collectées dans la littérature vont constituer la base de connaissances sur laquelle des modèles bayésiens hiérarchiques ont été développés et ajustés.

6. Interception en conditions sèches

Dans ce chapitre, nous étudions l'interception de radionucléides par les feuilles de plusieurs espèces végétales (*e.g.* herbe de pâturage, légumes feuilles, légumes racine, légumes fruit et tubercules) suite à un dépôt survenu en conditions sèches (*i.e.* en l'absence de précipitation). Les travaux présentés dans la suite ont fait l'objet d'un article rédigé en collaboration avec Sophie Ancelet, Pascale Henner, Pierre Hurtevent et Marie Simon-Cornu. L'article, accepté le 8 Mai 2015 dans la revue *Journal of Environmental Radioactivity*, est reproduit dans la [Section 6.2](#).

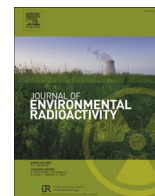
6.1 Objectifs et méthodes

L'interception sèche est en général quantifiée en calculant le ratio entre la quantité de radionucléides déposés à la surface des feuilles des plantes et la quantité totale de radionucléides déposés sur le sol et sur les plantes (cf. [Section 3.1.2.1](#)). Plusieurs revues de la littérature ont permis de répertorier des données relatives à l'interception sèche de radionucléides (*e.g.* [Anspaugh, 1987](#) ; [Chamberlain and Garland, 1991](#) ; [IAEA, 2009, 2010](#) ; [Pröhl and Hoffman, 1996](#) ; [Vandecasteele *et al.* 2001](#)). L'ensemble de ces revues, (notamment celles réalisées par l'AIEA ([IAEA, 2009, 2010](#))) a mis en évidence la quantité limitée d'informations relatives à l'interception sèche de radionucléides par des plantes agricoles telles que les légumes feuilles, les légumes-racines ; la majorité des données étant disponibles pour l'herbe. Ce manque de données concerne également certaines formes des radionucléides (avec des caractéristiques physiques différentes) car assez peu de données relatives à l'interception de radionucléides sous forme gazeuse sont disponibles dans la littérature.

Ainsi, dans cette étude, l'objectif était de synthétiser l'information disponible sur le facteur d'interception sèche pour différents radionucléides et différentes espèces végétales (herbe de pâturage et plantes cultivées), de proposer une meilleure caractérisation de l'incertitude sur le facteur de transfert et de la variabilité induite par les différentes espèces végétales d'intérêt, les différents radionucléides et les différences entre radionucléides en termes de caractéristiques physiques.

Un total de 145 couples (facteur d'interception, biomasse aérienne) a été recensé. Par biomasse aérienne, nous entendons la masse des feuilles ramenée à la surface occupée sur le sol. Sur les 145 données, 114 ont été extraites de la « littérature grise » de l'IRSN ayant fait l'objet de publications. Cette base de données concerne des radionucléides sous forme particulaire (^{110m}Ag , ^{133}Ba , Cs-134, Cs-137, Sr-85 et Te-123m) et d'iode gazeux (I-131). Des modèles non-linéaires hiérarchiques ont été ajustés par inférence bayésienne sur ces données. La structure hiérarchique des différents modèles testés a notamment permis d'évaluer l'influence de la variabilité entre espèces végétales et entre formes physiques des radionucléides (particulaire ou gazeuse) sur le facteur d'interception. Les qualités d'ajustement et prédictives des modèles ont été testées sur la base de différents critères quantitatifs.

6.2 Article: Foliar interception of radionuclides in dry conditions: a meta-analysis using a Bayesian modeling approach



Foliar interception of radionuclides in dry conditions: a meta-analysis using a Bayesian modeling approach



Mouhamadou Moustapha Sy ^{a,*}, Sophie Ancelet ^b, Pascale Henner ^c, Pierre Hurtevent ^c, Marie Simon-Cornu ^a

^a Institut de Radioprotection et de Sûreté Nucléaire (IRSN), Laboratoire de Modélisation pour l'Expertise Environnementale (LM2E), Cadarache, Bâtiment 159, St Paul-lez-Durance, 13115, France

^b Institut de Radioprotection et de Sûreté Nucléaire (IRSN), PRP-HOM, SRBE, Laboratoire d'Epidémiologie (LEPID) Fontenay-aux-Roses, 92262, France

^c Institut de Radioprotection et de Sûreté Nucléaire (IRSN), Laboratoire de Biogéochimie, Biodisponibilité et Transferts des radionucléides (L2BT), Cadarache, Bâtiment 183, St Paul-lez-Durance, 13115, France

ARTICLE INFO

Article history:

Received 4 February 2015

Received in revised form

30 April 2015

Accepted 8 May 2015

Available online

Keywords:

Foliar interception

Dry deposition

Meta-analysis

Hierarchical modeling

Bayesian statistics

ABSTRACT

Uncertainty on the parameters that describe the transfer of radioactive materials into the (terrestrial) environment may be characterized thanks to datasets such as those compiled within International Atomic Energy Agency (IAEA) documents. Nevertheless, the information included in these documents is too poor to derive a relevant and informative uncertainty distribution regarding dry interception of radionuclides by the pasture grass and the leaves of vegetables. In this paper, 145 sets of dry interception measurements by the aboveground biomass of specific plants were collected from published scientific papers. A Bayesian meta-analysis was performed to derive the posterior probability distributions of the parameters that reflect their uncertainty given the collected data. Four competing models were compared in terms of both fitting performances and predictive abilities to reproduce plausible dry interception data. The asymptotic interception factor, applicable whatever the species and radionuclide to the highest aboveground biomass values (e.g. mature leafy vegetables), was estimated with the best model, to be 0.87 with a 95% credible interval (0.85, 0.89).

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Direct deposition of radionuclides onto the plant canopy may cause relatively high activity concentrations in feed and foods (IAEA, 2009). The exchanges at the atmosphere-plant interface and the fate of contaminants subsequent to the deposition of radionuclides onto the plant canopy (Thiessen et al., 1999) are governed by complex physical and chemical processes leading to large sources of uncertainty when assessing nuclear accident consequences. Hence, the interception of the radioactive plume by vegetation is highlighted as a further research priority in the field of radioecology (Thiessen et al., 1999; Tschiersch et al., 2009).

Interception is defined as the proportion of deposited radionuclides retained by vegetation (Pröhl, 2009). The interception factor can thus be defined as the ratio of the radioactivity intercepted by plant leaves to the total radioactivity deposited on both ground and

plant leaves (both expressed in Becquerel per square meter of ground, $\text{Bq}\cdot\text{m}^{-2}$). One can also consider interception factors normalized either by the biomass (expressed in kilogram of fresh or dry weight per square meter of ground) or by the leaf area index (LAI). LAI is generally considered as the one-sided green leaf area per unit ground surface area (leaf area/ground area, $\text{m}^2\cdot\text{m}^{-2}$) (Colle and Madoz-Escande, 2007). Such normalized interception factors are one way to account for plant phenology, even if they are insufficient to take into account all conditions and factors involved in the interception process, such as aerodynamic conditions, amount and intensity of the precipitation, type of deposits (dry or wet), the ability of plants' leaves to retain water or the radionuclides' properties (Pröhl, 2009; Vandecasteele et al., 2001). Radionuclides are different in terms of chemical and physical characteristics. Thus, interception of a gas may be significantly different from interception of a submicronic particle, itself different from interception of micronic particles, resulting from differences between the dry deposition velocities of gases, submicronic and micronic particles (Sehmel, 1980; Sehmel and Hodgson, 1978; Tschiersch et al., 2009) onto the plant canopy.

* Corresponding author.

E-mail address: moustapha.sy@irsn.fr (M.M. Sy).

The International Atomic Energy Agency (IAEA) TRS472 technical report (IAEA, 2010) provides only limited data about interception of radionuclides by crops leaves (i.e., pasture grass and vegetables) after a dry deposition event. Thus, further data must be collected to provide information about interception of radionuclides by crop leaves after a dry deposition of radionuclides. This is a wide topic and the main difficulty is that it requires simultaneous analysis of heterogeneous, and thus potentially not directly comparable, data.

Different sources of such heterogeneity can be seen in the literature.

- Heterogeneity depending on the radionuclides characteristics: it is unusual to see in the literature works that combine simultaneously particulate and gaseous radionuclides to study their (dry or wet) interception by plant canopy.
- Parameterization heterogeneity: as previously mentioned different ways of studying the interception factors are present in the literature f , f_B and f_{LAI} . Using these data in the same modeling framework should require a preliminary data treatment.
- Measurement heterogeneity: some covariates such as the aboveground biomass or the LAI are sometime measured on different and miscellaneous bases. For instance, foliar biomass values are found in dry weight or in fresh weight in the literature.
- Heterogeneity due to the botanic characteristics of plants: most authors focused in their experimental studies, on interception, on a single type of plants rather than combining in a unique experimental framework plants that are different in terms of the botanic specificities of the plants. The level of dry interception may be strongly dependent on the peculiarities of the plant canopy (Kinnersley and Scott, 2001; IAEA, 2010; Thiessen et al., 1999; Tschiersch et al., 2009).
- Heterogeneity due to the plants use in daily life by humans: plants could be considered to be different, and thus not combined in the same models because they are used for different purposes in terms of daily life. Indeed pasture grass are consumed by animals, whereas vegetable leaves, roots, or fruits are mainly consumed by humans.

Bayesian meta-analysis is a widely used approach in medicine, epidemiology and food microbiology (Cohen et al., 2013; Dohoo et al., 2007; Qian et al., 2004; Rigaux et al., 2013; Tweedie et al., 1996). Independently from the meta-analysis method, the Bayesian paradigm is an approach based on the theory of conditional probabilities and the Bayes' theorem (Kruschke, 2010; Parent and Bernier, 2007). When used to fit a given parametric model, this statistical approach consists in updating the state of knowledge about the unknown parameter(s) of the model of interest by combining prior information (e.g. drawn from expert advice or historical data) and observed information (likelihood of the observed data given the model). This new state of knowledge corresponds to a certain probability distribution called posterior. The Bayesian meta-analysis method consists here of combining potentially heterogeneous data, from different studies into a unique and flexible parametric model, usually employing Bayesian hierarchical models (Boreux et al., 2010; Bradlow et al., 1999; Brophy et al., 2003; Kruschke, 2010; Helser and Lai, 2004; Parent and Bernier, 2007). Performing a meta-analysis under the Bayesian framework offers the possibility to derive PDFs, called posterior probability distributions, reflecting the uncertainty about unknown parameters of interest given the different combined datasets. Such an approach is relevant for accounting for different sources of variability and uncertainty (Helser and Lai, 2004; Hosseini et al., 2013; Li et al., 2012).

Deterministic models have been previously proposed to describe the interception process of radionuclides onto the plant canopy by accounting for some potentially important factors, such as plant phenology, and/or between-species variability. Chadwick and Chamberlain (1970) established a non-linear relationship between the interception factor and standing plant biomass (of dry matter) after a dry deposition over grassland. The non-linear Chapman–Richards growth model (Li et al., 2012; Richards, 1959) was proposed as an alternative structure to express the relationship between the interception factor and the aboveground plant biomass after a dry deposition scenario. This model is a generalized form of the Chadwick and Chamberlain model, and includes an unknown parameter, the asymptotic interception factor (not necessary equal to 1 as in the model proposed by Chadwick and Chamberlain) corresponding to the unknown value of the plateau reached by the interception factor for high plant biomass. Such a parameter allows taking consideration of a saturation phenomenon when plants achieve their maximal biomass. A similar model was proposed by Vandecasteele et al. (2001) to describe the relationship between the interception factor and dry plant biomass for cereal cultures.

The main purpose of this paper is to better characterize the uncertainties relative to dry interception of radionuclides by pasture grass and vegetables leaves. A dataset including dry interception measurements and aboveground biomass of plants was collected from published scientific papers. Even though these data show some heterogeneity in regard to the previously mentioned reasons they all provide, in this study, a crucial amount of information on dry interception of radionuclides by plant canopy when handled in the same modeling framework. A Bayesian meta-analysis was then performed, with different hierarchical models based on non-linear structures and accounting for the between-species variability proposed and fitted under the Bayesian paradigm. The competing models were compared in terms of fitting and predictive performances. Cross-validation was performed regarding the predictive performances of the models and quantitative indices were produced. The best model was then employed to predict new dry interception measurements of radionuclides by pasture grass or vegetables leaves.

2. Materials and methods

2.1. Data

Data collection focused on interception by the leaves of crops, more specifically, pasture grass and vegetables, measured at harvest after a dry deposition scenario of radionuclides (whatever their physical form) and with a short deposition-harvest delay (2 days maximum).

Data, were collected from different sources: 145 interception factors measured after a dry deposition scenario by (i) radionuclide aerosols (^{110m}Ag , ^{133}Ba , ^{134}Cs , ^{137}Cs , ^{85}Sr and ^{123m}Te), and (ii) gaseous iodine (^{131}I). Thirty-one data were extracted from published scientific papers (Chamberlain and Garland, 1991; Watterson and Nicholson, 1996) and 114 were extracted from IRSN reports also associated with published papers (Madoz-Escande et al., 2001, 2002, 2003; Madoz-Escande and Santucci, 2004; Rauret et al., 1995; Réal and Miara, 1990; Réal et al., 1994).

As mentioned above, two types of data were finally collated: 135 data for which the interception measurements were performed at the deposition and 10 data for which the measurements were performed at harvest even though the delay between the deposition and the harvest was rather short: 1 day for the lettuces (7 data) and 2 days for the data on the leaves of potatoes (3 data). For consistency purposes, data corresponding to measurements at

harvest were corrected to the deposition date by using the radioactive decays of ^{134}Cs , $^{110\text{m}}\text{Ag}$ and ^{85}Sr since the interception is the ratio between the concentration that is initially intercepted by the plant canopy and the total concentration deposited on plant canopy and ground. This correction resulted in a minor increase of the interception fraction values (at deposition) estimated between 0.3 and 0.6%, 0.1–0.2% and 1.1–2.1% for $^{110\text{m}}\text{Ag}$, ^{134}Cs and for ^{85}Sr particles respectively. In addition, we also corrected the corresponding fresh biomass values by using a foliar growth rate of 0.03 for lettuces and of 0.024 for the leaves of potatoes. The foliar growth rate of lettuces was estimated from the data (deposition-harvest time + biomass values at deposition and at harvest) reported by Madoz-Escande and Santucci (2004) and the foliar growth rate used for the potatoes leaves was estimated from radish data also reported by Madoz-Escande and Santucci (2004). The correction of the biomass values resulted in a decrease of about 3.2% for lettuces and about 4.7% for the leaves of potatoes.

Thus, the collected interception factors corresponded to the leaves of 5 different plants (green beans, grass, lettuces, radishes and potatoes). Indeed, the paucity of data for the interception factors of vegetables, discussed in Chamberlain and Garland (1991); IAEA, 2010; Pröhl and Hoffman, 1996; Vandecasteele et al., 2001 motivated our focus on “such different” plants. The relevancy of these plants also lay partly in their significant contribution to the total dose ingested by human population in a post-accidental situation (De Lurdes Dinis and Fiúza, 2007; Simpkins, 2005; Tschiersch et al., 2009).

The aboveground biomasses relative to grass were measured in dry weight (Chamberlain and Garland, 1991; Madoz-Escande et al., 2003). Consequently, a strategy to convert those dry weight biomasses into fresh weight biomasses and to take into account the underlying uncertainty related to this conversion was to use dry matter content values as those reported in the TRS472 for pasture grass (IAEA, 2010). This strategy is presented in Section 2.2.

Fig. 1 displays a scatterplot of the observed interception factors and the corresponding standing biomass (in fresh weight and dry weight) for the five plant species (graphs on the top) and the seven pollutants (graphs on the bottom). Even if a common global feature appears, our motivation was to investigate differences, between species and between radionuclides physical forms, in the relationship between the interception factor and the standing fresh biomass by using a hierarchical modeling framework.

2.2. The models

2.2.1. The data model

Let F_{ijk} be the interception factor of plant species $i = 1, \dots, 5$ for radionuclide physical form $j = 1, 2$ in replicate $k = 1, \dots, n_{ij}$. Let B_{ijk} be the corresponding aboveground biomass expressed in either kg of fresh weight per square-meter or in kg of dry weight per square-meter (considering only leaves and stems, excluding fruits if any). As F_{ijk} denotes the ratio of the radioactivity intercepted by plant leaves to the total radioactivity deposited on ground and plant leaves, it takes values over the (0,1) open interval. The *logit* mathematical function has been considered to transform this random variable into the following random variable F_{ijk}^* taking real values (i.e., between $-\infty$ and $+\infty$),

$$F_{ijk}^* = \text{logit}(F_{ijk}) = \log\left(\frac{F_{ijk}}{1 - F_{ijk}}\right) \quad (1)$$

As it is shown in equation (1), given any number over the open interval (0,1), the *logit* function returns a real value (i.e. between negative infinity and positive infinity). This function is extensively used in economics and statistics, especially in logistic regression analyses (Cramer, 2003).

We considered the following Gaussian model for the logit-transformed interception factors as described by the equations (2)–(7). This model is based on the logit-transformation of a simplified version of the Chapman–Richards growth model (Li et al., 2012; Richards, 1959) that has been previously used in radioecology for modeling the interception factor (Vandecasteele et al., 2001). An asymptotic interception factor and the rate of convergence to this asymptotic interception factor that are biased estimators (due to the logit-transformation) of the maximum interception and the absorption coefficient of Vandecasteele et al. (2001) are defined respectively.

$$F_{ijk}^* = \mu_{ijk} + \varepsilon_{ijk} \quad (2)$$

$$\mu_{ijk} = \text{logit}\left(A_i \times \left(1 - \exp^{-G_j \times \widetilde{B}_{ijk}}\right)\right) \quad (3)$$

$$\widetilde{B}_{ijk} = \frac{B_{ijk}}{DMC} \quad (4)$$

$$\varepsilon_{ijk} \sim \text{i.i.d} N(0, S^2) \quad (5)$$

$$A_i \sim \text{i.i.d} \text{Beta}(\mu_A * \tau_A, (1 - \mu_A) * \tau_A), \text{ and} \quad (6)$$

$$G_j \sim \text{i.i.d} \text{Gamma}(\mu_G * \tau_G, \tau_G) \quad (7)$$

Where μ_{ijk} is the expected value of the logit-transformed interception factor F_{ijk}^* ; A_i is the asymptotic interception factor, i.e., the value of the expected logit-transformed interception factor when the fresh biomass is very high. On the basis of expert advice (regarding considerations of vegetal physiology) and preliminary ANOVAs, it was assumed that A_i depends only on the plant, and then not on the physical form of the radionuclide. Thus, each plant species has its own A_i , which can be seen as a plant species effect. However, the whole asymptotic interception factors, that take values over the (0,1) open interval, are assumed to be similar between species because the maximal efficiency of the vegetation is considered to be similar for all the plant species of interest when the canopy entirely covers the ground. Thus, the whole asymptotic interception factors are assumed to stem from a common beta distribution with parameters that are function of $\mu_A > 0$ and $\tau_A > 0$ describing the between-species variability (see equation (6)). For each plant species, the expected asymptotic interception factor, A_i , is given by μ_A . Dissimilarity (or degree of resemblance) between plant species in terms of asymptotic interception factor is taken into account and is determined by the inter-species variance called V_A . V_A is defined by the formula $V_A = \frac{(1 - \mu_A) \times \mu_A}{1 + \tau_A}$. The smaller the variance, the more similar are the plant species in terms of asymptotic interception. The beta distribution is chosen due to its flexibility to describe random variables that take their values over the interval (0,1). Indeed, the beta distribution is a probability distribution that can take different shapes depending on the values of its two parameters.

G_j , the rate of convergence of μ_{ijk} to its asymptotic interception factor A_i when the fresh biomass is very high, is a strictly positive parameter expressed in $\text{m}^2 \text{kg}^{-1}$ of fresh weight (fw). In other words, the higher G_j the faster μ_{ijk} converges to A_i . Based on our prior knowledge of the interception process, this rate of convergence is assumed to depend only on the physical form of the radionuclides. Thus, it is assumed to be constant whatever the plant species. The rate of convergence is analogous to the absorption coefficient defined by Chadwick and Chamberlain (1970). Each

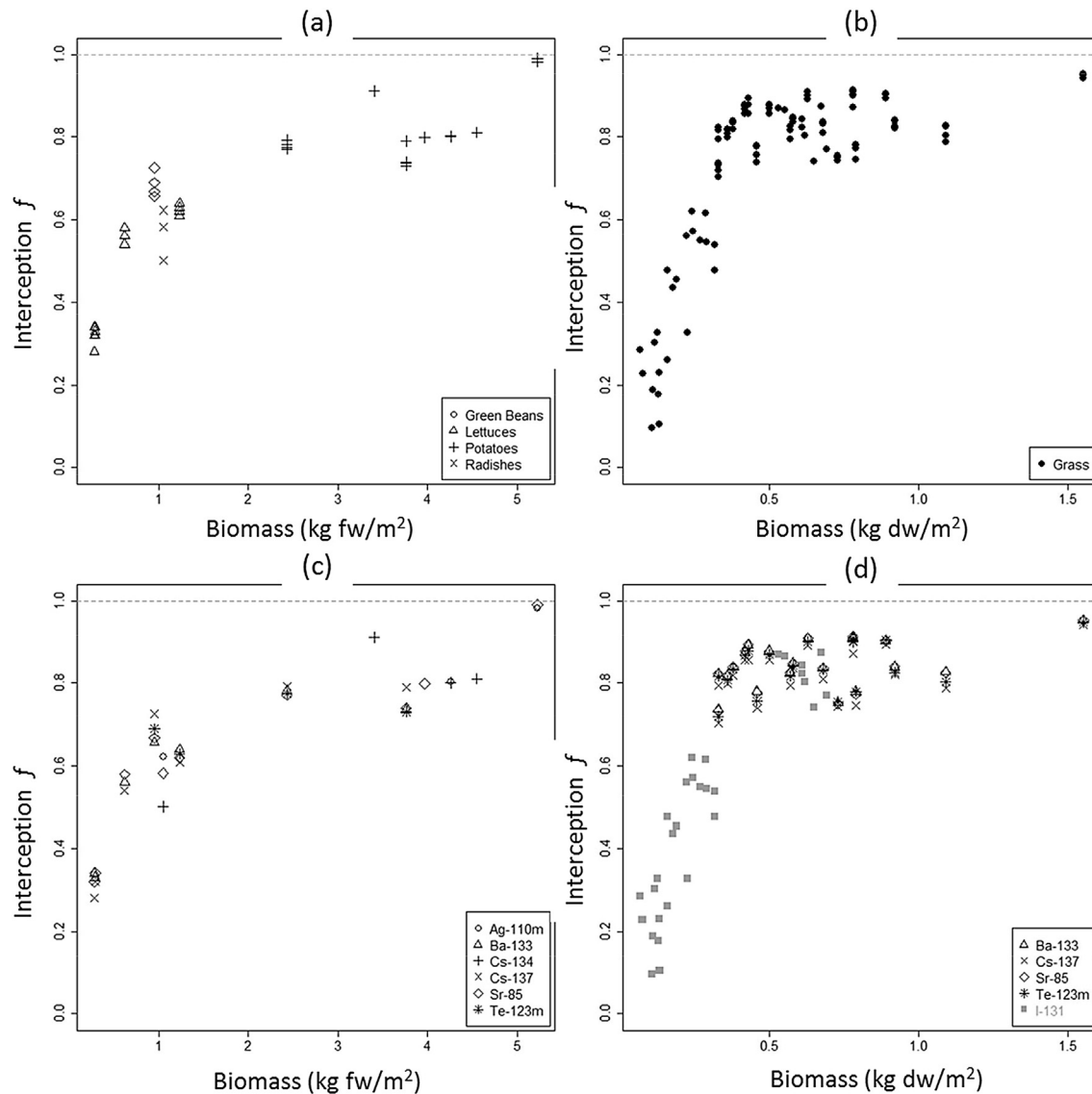


Fig. 1. Scatterplots of the observed interception factors f (Y axis) and the associated standing biomass (X axis) for all the plants (figures (a) and (b) on the top) and the 7 pollutants composed by radionuclides isotopes in particulate and gaseous forms (figure (c) and (d) on the bottom). On the left (figures (a) and (c)) and on the right (figures (b) and (d)) the fresh biomasses and the dry biomasses are represented respectively.

radionuclide physical form (i.e., gaseous and particular radionuclides) gives rise to a specific rate of convergence G_j . However, the whole rates of convergence, that take positive values, are assumed to be similar between radionuclides and thus to stem from a common gamma distribution with parameters function of $\mu_G > 0$ and $\tau_G > 0$ describing the between-radionuclides variability (see equation (7)). For each radionuclide physical form, the expected rate of convergence, G_j , is given by μ_G . Dissimilarity (or degree of resemblance) between radionuclides physical forms in terms of rates of convergence is determined by the inter-radionuclides variance V_G , which is defined by the formula $V_G = \frac{\mu_G}{\tau_G}$.

DMC refers to the dry matter content that is used to convert dry weight biomasses in fresh weight biomasses. For fresh weight biomasses a value of DMC fixed at 1 is considered whereas a DMC value between 0 and 1 is considered for the dry weight biomass values corresponding only to the pasture grass in our dataset. Thus, \bar{B}_{ijk} that is given by equation (4) corresponds to the aboveground biomass expressed in a fresh weight basis.

ε_{ijk} are random variables that denote the discrepancy between the logit-transformed interception factor, F_{ijk}^* , and the expected

values $\text{logit}(\mu_{ijk})$. Such a discrepancy may be due both to the natural variability (also called aleatory uncertainty) of the observed interception factors but also to modeling error (as a part of what could be called epistemic uncertainty). The non-linear regression models, as the one considered here, classically involve such random variables that are supposed to be independent and identically distributed (i.i.d), and centered on zero (Smyth, 2006). Moreover, symmetric probability distributions are classically defined for these terms and a natural common choice of symmetric distribution is the Gaussian distribution. S^2 denotes the unknown variance of such a discrepancy also called residual variability, as due to remaining unexplained variability. Note that, according to Li et al. (2012) the whole ε_{ijk} do not have to be necessarily Gaussian. However, the discussion on the choice of the Gaussian distribution is widely beyond the scope of this article.

2.2.2. The competing models

Equations (2)–(7) define the full hierarchical model of interest (M_4) that allows describing the variability of the interception process that is due to the plant species and the physical form of

radionuclides. Nevertheless, the following hypotheses have also been considered and compared to the previous ones:

a) No between-species variability. In this case, the asymptotic interception factors are considered equal for all plant species i.e. $A_i = A$ for $i = 1, \dots, 5$. and/or.

b) No variability due to the physical forms of radionuclides. In this case, the rates of convergence are considered equal for all radionuclides physical forms i.e. $G_j = G$ for $j = 1, 2$.

We also defined three simpler versions of model M_4 (M_1 , M_2 , and M_3) to consider all the assumptions cited above. M_1 is the most parsimonious model, including only 3 unknown parameters A , G , and S . It is a non-hierarchical model assuming no between-species variability and no variability due to the physical forms of radionuclides. M_2 is a hierarchical model only accounting for a potential between-species variability of the asymptotic interception factors. M_3 is a hierarchical model only accounting for a potential variability of the rates of convergence due to the physical forms of radionuclides. The four models are explicitly defined in Table 1.

All the models are represented by a Directed Acyclic Graph (DAG) in Fig. 2. A DAG gives a simple representation of the different variables (also called nodes in this context) involved in a model and the conditional dependencies between them. For more details on DAGs see Parent and Bernier (2007) and Rigaux et al. (2013). The conditional dependencies may be deterministic or stochastic and are represented by arrows on the DAG. To avoid confusion, nodes without parents and represented by an ellipse in the DAG are called parameters, nodes without children are called observable random variables (or outputs) and intermediate nodes (between parameters and outputs) are called intermediate (or latent) random variables in the following.

2.3. Bayesian inference

The inference of models (M_m ; $m = 1, \dots, 4$) was performed under the Bayesian framework that is particularly appropriate for fitting nonlinear models (De La Cruz-Mesía and Marshall, 2006; Li et al., 2012; Wakefield et al., 1994) and hierarchical models (Banerjee et al., 2004; Helsner and Lai, 2004; Parent and Bernier, 2007) such as those presented above.

The Bayesian model requires assigning prior distributions on the set of unknown parameters. Such prior distributions allow for

incorporation of external information (expert knowledge, literature data ...) and reflect the modeler's uncertainty about these unknown parameters before collecting new data. Our prior choices were motivated by mathematical reasons and by some prior knowledge about the values or range of possible values of some parameters. A flat uniform prior distribution between 0 and 100 was assigned to the strictly positive G parameter for models M_1 and M_2 due to a lack of knowledge about this parameter. The same prior distribution was assigned to μ_G in models M_3 and M_4 whereas a uniform distribution between 0 and 1000 was defined for the hyper-parameter τ_G . The asymptotic interception factor, A , of models M_1 and M_3 takes values over the (0,1) interval but no further knowledge is available. Consequently, a flat uniform distribution between 0 and 1 was assigned to A (M_1 and M_3) and to μ_A (M_2 and M_4). A flat uniform distribution between 0 and 1000 was chosen for τ_A (M_2 and M_4), and following (Gelman, 2006; Lambert et al., 2005), a uniform prior distribution between 0 and 100 was assigned to the residual standard deviation parameter, S . A uniform probability distribution between 0.1 and 0.33 was defined for the dry matter content parameter, DMC , related to pasture grass. The lower and upper bounds of this uniform prior distribution were fixed from the water content values reported for pasture grass in the TRS472 by IAEA (2010). The prior distributions are summarized in Table 1.

The target of Bayesian inference is the posterior distribution of the set of parameters, that reflects our uncertainty about these parameters after combining the prior distributions with the likelihood (that reflects the information brought by the observed data) through the well-known Bayes' theorem (Bayes and Price, 1763). Markov Chain Monte-Carlo (MCMC) sampling algorithms (Robert and Casella, 2004) were applied as implemented in the software JAGS 3.3.0 (Plummer, 2003) to obtain a sample from the joint posterior distribution of all unknown parameters of the models.

For each model, three independent Markov chains (starting from different random values) were run with a burn-in-period of 10,000 iterations to allow the Markov chains to reach their stationary equilibrium distribution, and 50,000 additional iterations after the burn-in-period. To reduce autocorrelation in the posterior samples, we kept only those values generated every 10 iterations. Therefore, our results are based on thinned samples of size $N = 15,000 (= 5000 \times 3)$ for each parameter of interest. The

Table 1
Prior probability distributions assigned to all unknown parameters for all models and models' specifications.

Models	Model formulation	Parameters	Unit ^a	Prior distribution
M_1	$F_{ijk}^* \sim iid N(\mu_{ijk}, S^2)$ $\mu_{ijk} = \text{logit} \left(A \left(1 - \exp^{-G^* B_{ijk}} \right) \right)$	A	–	Uniform(0,1)
		DMC	–	Uniform(0.1,0.33)
		G	$m^2.kg^{-1}$ fw	Uniform(0,100)
		S	–	Uniform(0,100)
M_2	$F_{ijk}^* \sim iid N(\mu_{ijk}, S^2)$ $\mu_{ijk} = \text{logit} \left(A_i \left(1 - \exp^{-G^* B_{ijk}} \right) \right)$	μ_A	–	Uniform(0,1)
		τ_A	–	Uniform(0,1000)
		DMC	–	Uniform(0.1,0.33)
		G	$m^2.kg^{-1}$ fw	Uniform(0,100)
M_3	$F_{ijk}^* \sim iid N(\mu_{ijk}, S^2)$ $\mu_{ijk} = \text{logit} \left(A \left(1 - \exp^{-G_j^* B_{ijk}} \right) \right)$	A	–	Uniform(0,1)
		DMC	–	Uniform(0.1,0.33)
		μ_G	$m^2.kg^{-1}$ fw	Uniform(0,100)
		τ_G	$m^2.kg^{-1}$ fw	Uniform(0,1000)
M_4	$F_{ijk}^* \sim iid N(\mu_{ijk}, S^2)$ $\mu_{ijk} = \text{logit} \left(A_i \left(1 - \exp^{-G_j^* B_{ijk}} \right) \right)$	μ_A	–	Uniform(0,1)
		τ_A	–	Uniform(0,1000)
		DMC	–	Uniform(0.1,0.33)
		μ_G	$m^2.kg^{-1}$ fw	Uniform(0,100)
		τ_G	$m^2.kg^{-1}$ fw	Uniform(0,1000)
		S	–	Uniform(0,100)

^a fw = fresh weight.

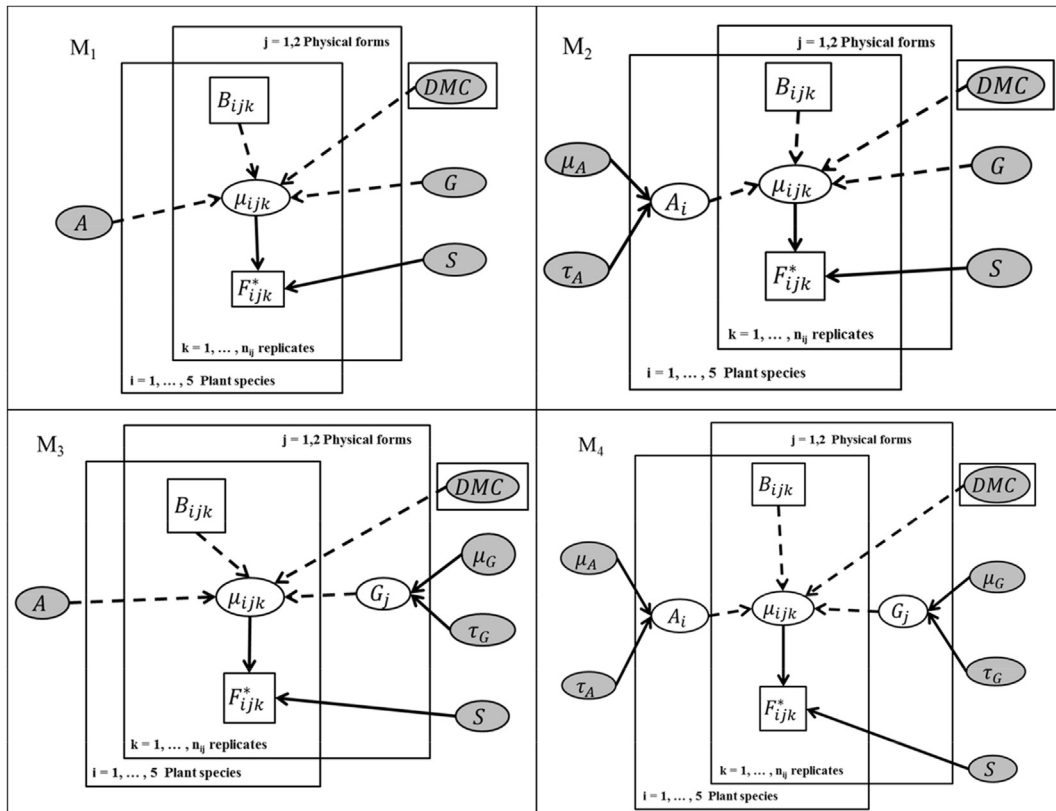


Fig. 2. Directed Acyclic Graphs (DAG) of models M_1 , M_2 , M_3 and M_4 . Ellipses in light gray indicate parameters; ellipses in white indicate intermediate variables and rectangles indicate. Since the parameter DMC is fixed at 1 for fresh biomasses and is a $U(0,1,0.33)$ for dry biomasses, the node corresponding to DMC is represented simultaneously by a light gray ellipse and a white rectangle. Dashed edges are deterministic links and solid edges are stochastic links.

convergence of the MCMC runs was assessed by graphical inspection of the chains that they mixed well for all parameters, and by computing the Gelman-Rubin convergence diagnostic for all parameters. This convergence diagnostic is based, for each parameter, on the analysis of the within-chain and the between-chains variances. For further details see Brooks and Gelman (1998).

Various empirical quantities were computed from the posterior samples obtained, such as, posterior means, posterior medians, and posterior standard deviations estimates or the 95% credible intervals (95%CI) that represent the range of values between the 2.5th and the 97.5th percentiles of the posterior distribution. These empirical quantities are displayed in Table 2.

2.4. Model selection

The Deviance Information Criterion (DIC) proposed by Spiegelhalter et al. (2002) was first computed from the posterior samples to compare the fitting performances of the competing models (see Section 2.1). DIC is a Bayesian alternative to penalized likelihood criteria, such as Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). It provides a compromise between having a model that fits the data well (goodness-of-fit characterized by the posterior mean of the deviance) and one that is as parsimonious as possible, i.e., including few parameters, characterized by an estimate of the effective number of parameters. The smaller the DIC, the better the associated model in terms of fitting performances (see Table 3).

To compare the ability of our models to reproduce plausible logit-transformed interception factor values, posterior predictive model checks were performed. An iterative procedure based on the

leave-one-out cross-validation (LOOCV) approach (Kohavi, 1995; Hastie et al., 2009) was implemented.

Recall that the available data consists of 145 couples of an observed logit-transformed interception factor value associated with its corresponding standing aboveground biomass. Let us set $D_n = (F_n^*, B_n) : n = 1, \dots, N = 145$. For a given model $M_m : m = (1, \dots, 4)$, the LOOCV procedure used is described in details in Appendix A.

Several quantitative indices were then produced from the samples drawn from the LOOCV procedure. The average width and coverage probabilities related to the 95% credible intervals were derived from the obtained samples. The posterior predictive loss criterion (PPLC) proposed by Gelfand and Ghosh (1998) was calculated using the samples drawn from the LOOCV procedure to compare the predictive performances of the models. The PPLC is defined as:

$$PPLC(M_m, k) = \sum_{s=1}^{145} \text{Var}(\tilde{F}_{(s)}^*) + \frac{k}{k+1} \sum_{s=1}^{145} (E(\tilde{F}_{(s)}^*) - F_s^*)^2$$

Where $E(\tilde{F}_{(s)}^*)$ and $\text{Var}(\tilde{F}_{(s)}^*)$ are the estimated mean and variance of the predictive distribution of a new interception factor under model $M_m : m = (1, \dots, 4)$ and given a learning sample $F_{(s)}^*$ that includes observed logit-transformed dry interception values (see the LOOCV procedure in Appendix A). Such predictive mean and variance are estimated from posterior predictive samples $\tilde{F}_{(s)}^*$ obtained from the learning set; $k (= 1, 10, 100, \infty)$ is a weighting coefficient related to the discrepancy between the observed logit-transformed dry interception values and the mean value of the predicted

Table 2

Summary statistics (i.e., posterior means, medians, standard deviations and 95%CI) of the approximate posterior distributions of the parameters and latent variables for all models.

Models	Nodes type	Nodes	Corresponding groups	Unit ^b	Post. mean	Post. Median	Post. Sd ^a	95%CI		
M ₁	Parameters	A	–	–	0.89	0.89	0.01	[0.87;0.91]		
		DMC	–	–	0.26	0.25	0.03	[0.20;0.31]		
		G	–	–	m ² .kg ⁻¹	1.22	1.22	0.14	[0.96;1.52]	
		S	–	–	–	0.62	0.62	0.04	[0.55;0.70]	
M ₂	Parameters	μ _A	–	–	0.88	0.89	0.02	[0.85;0.92]		
		τ _A	–	–	538.55	546.51	273.76	[64.13;975.06]		
		V _A	–	–	3.6e-4	1.8e-4	7.4e-4	[9.4e-5;1.6e-3]		
		DMC	–	–	0.26	0.26	0.03	[0.2;0.32]		
		G	–	–	m ² .kg ⁻¹	1.25	1.24	0.15	[0.97;1.55]	
		S	–	–	–	0.62	0.62	0.04	[0.55;0.70]	
	Latent Variables	A ₁	Green beans	–	–	0.89	0.89	0.03	[0.83;0.93]	
		A ₂	Grass	–	–	0.89	0.89	0.01	[0.87;0.92]	
		A ₃	Lettuces	–	–	0.88	0.88	0.02	[0.84;0.91]	
		A ₄	Potatoes	–	–	0.89	0.89	0.02	[0.84;0.93]	
		A ₅	Radishes	–	–	0.88	0.88	0.03	[0.82;0.93]	
M ₃	Parameters	A	–	–	0.87	0.87	0.01	[0.85;0.89]		
		DMC	–	–	0.19	0.19	0.03	[0.14;0.25]		
		μ _G	–	–	m ² .kg ⁻¹	1.08	1.01	0.47	[0.58;2.04]	
		τ _G	–	–	m ² .kg ⁻¹	18.51	12.44	25.58	[0.91;72.31]	
		V _G	–	–	m ² .kg ⁻¹	0.38	0.08	2.50	[0.01;1.79]	
		S	–	–	–	0.56	0.55	0.03	[0.50;0.63]	
	Latent Variables	G ₁	Particles	–	–	m ² .kg ⁻¹	1.30	1.29	0.14	[1.05;1.60]
		G ₂	Gaseous Iodine	–	–	m ² .kg ⁻¹	0.69	0.68	0.12	[0.48;0.95]
M ₄	Parameters	μ _A	–	–	0.87	0.87	0.02	[0.83;0.90]		
		τ _A	–	–	560.27	574.95	269.03	[68.42;979.60]		
		V _A	–	–	3.7e-4	1.9e-4	9.2e-4	[1.1e-4;1.7e-3]		
		DMC	–	–	0.20	0.19	0.03	[0.14;0.26]		
		μ _G	–	–	m ² .kg ⁻¹	1.10	1.03	0.49	[0.58;2.08]	
		τ _G	–	–	m ² .kg ⁻¹	18.08	12.30	20.40	[0.84;70.61]	
		V _G	–	–	m ² .kg ⁻¹	0.55	0.08	5.60	[0.01;2.01]	
		S	–	–	–	0.56	0.55	0.03	[0.50;0.63]	
		Latent Variables	A ₁	Green beans	–	–	0.87	0.87	0.03	[0.82;0.92]
	A ₂		Grass	–	–	0.87	0.87	0.01	[0.85;0.89]	
	A ₃		Lettuces	–	–	0.87	0.87	0.02	[0.84;0.90]	
	A ₄		Potatoes	–	–	0.87	0.87	0.02	[0.82;0.92]	
	A ₅		Radishes	–	–	0.87	0.87	0.03	[0.81;0.91]	
	Latent Variables	G ₁	Particles	–	–	m ² .kg ⁻¹	1.32	1.30	0.16	[1.03;1.68]
G ₂		Gaseous Iodine	–	–	m ² .kg ⁻¹	0.70	0.69	0.13	[0.48;0.97]	

^a Post.sd is the posterior standard deviation.

^b m².kg⁻¹ = m².kg⁻¹ fw.

posterior samples $F_{(s)}^{\sim}$ ($s = 1, \dots, 145$). The best model will show the smallest PPLC.

The best model was selected by comparing DIC (see Table 3), PPLC (see Table 4) and predictive performances of the 4 competing models.

2.5. Further posterior predictive illustrations

Once the best model was selected, as described above, posterior predictive distributions were produced, as an illustration of the predictive results from this best model, for a set of standing biomass values ranging from 0 to 6 kg fw m⁻² (with step 0.01 kg fw m⁻²).

The interception factor distributions were then compared with the literature (Réal and Miara, 1990; Watterson and Nicholson, 1996). Interception factor values measured by Réal and Miara

(1990) and Watterson and Nicholson (1996) correspond to ¹³⁴Cs and ^{110m}Ag particles, and silica particles intercepted by mature lettuces, respectively. These data have not been considered in previous analyses because the aboveground standing biomass of the mature lettuces is unavailable in both studies, which adds an additional layer of difficulty. Considering the standing biomass unavailability, we defined a model that generates plausible standing biomass values of mature lettuces, connected to the measures from Réal and Miara (1990) and Watterson and Nicholson (1996). The Bayesian model was

$$B_{let} = (N_{let} * W_{let}) / 1000, \tag{8}$$

Table 4

Quantitative criteria for comparing the predictive performances of models M₁, M₂, M₃ and M₄. The posterior predictive loss criterion (PPLC), coverage ratios and 95%CI average widths are provided.

Models	PPLC ^a (-)	Coverage ratio (in %)	Average width 95%CI (-)
M ₁	85.17	97.2	2.47
M ₂	85.74	97.2	2.46
M ₃	68.86	96.5	2.22
M ₄	69.93	97.2	2.23

^a PPLC for w = 10,100 and +∞ were also calculated and led to the same conclusions.

Table 3

DIC values produced during the inference of models M₁ to M₄. Mean deviance (Dbar) and penalization (pD) are also provided.

Models	Dbar	pD	DIC
M ₁	271.4	4.1	275.5
M ₂	270.9	4.7	275.6
M ₃	239.4	5.4	244.8
M ₄	239.8	6.2	246.0

$$N_{let} = q * N_{let}^{outdoor} + (1 - q) * N_{let}^{indoor}, \quad (9)$$

$$N_{let}^{outdoor} \sim \text{Poisson}(5); N_{let}^{indoor} \sim \text{Poisson}(12), \quad (10)$$

$$q \sim \text{Beta}(2, 2), \quad (11)$$

$$W_{let} \sim \text{Normal}(400, 50), \quad (12)$$

where B_{let} is the product of the number N_{let} of lettuces per square meter of ground and the weight of one lettuce, W_{let} (in grams fw). Division by 1000 was applied to express B_{let} in kg fw.m⁻². Both N_{let} and W_{let} were assumed to be random variables because they are uncertain quantities due to potential observational errors, heterogeneity of agricultural practices, varieties of lettuce crops, etc. N_{let} was modeled with a mixture of two Poisson distributions that allow distinguishing indoor and outdoor agricultural practices for lettuce culture. The number of lettuces per square meter of ground for indoor, N_{let}^{indoor} and outdoor, $N_{let}^{outdoor}$, cultures was defined by a Poisson distribution with mean values elicited from an expert in vegetable crops at 12 and 5 lettuces per m² of ground, respectively. The mixture ratio parameter, q , was assumed to follow a beta distribution with shape parameters = 2. This distribution formalizes our avoiding preferential treatment to any specific agricultural practice. A Gaussian distribution was used to express the weight of any individual lettuce. The mean value of W_{let} was assumed to be 400 g with standard deviation = 50 g, again from expert opinion. This standing biomass model is used jointly but independently of the models M_1 to M_4 , and could be applied in further works (or fields) where the simulation of standing biomass values of mature lettuces (or other crops) is of interest.

3. Results

3.1. Posterior results

Some summary statistics (i.e., posterior mean, median, standard deviation, and 95%CI) of the posterior distributions computed from the obtained posterior samples (Section 2.3) are shown in Table 2 for the model parameters. The posterior mean of the asymptotic interception factor, A , which is common to all species for M_1 and M_3 was estimated at 0.89 and 0.87, respectively. The same values were obtained for the posterior median for M_1 and M_3 , respectively. These values estimate the expected logit-transform interception factor for the highest biomass values. The posterior means and posterior medians estimated for the five plant-specific asymptotic interception factors, $A_i, i = 1, \dots, 5$ in M_2 are similar to those obtained for M_1 and M_3 . The posterior means and medians of the asymptotic interception factors estimated for M_4 are equivalent to the estimation for M_3 , approximately 0.87. When all species are assumed to have the same asymptotic interception (M_1 and M_3), grass may have a slightly larger influence than the other plant species because the data quantity available for this species is somewhat higher. The narrow range of the associated 95% CIs, [0.87; 0.91] for M_1 and [0.85; 0.89] for M_3 , shows that the number of available interception factors is sufficient to estimate the common asymptotic interception of the interception factors with small uncertainty.

Accounting for the inter-species variability of the asymptotic interception factor does not induce strong changes in asymptotic interception factors. This is confirmed by low values of the posterior means of the parameter V_A , approximately 3.6×10^{-4} ($9.4 \times 10^{-5}, 1.6 \times 10^{-3}$) and 3.7×10^{-4} ($1.1 \times 10^{-4}, 1.7 \times 10^{-3}$) for models M_2 and M_4 , respectively, where the 95%CI is shown in brackets. This suggests that distinguishing the inter-species variability in our models would not be relevant.

The posterior mean of the rate of convergence, G , is rather close for M_1 and M_2 , approximately 1.22 (0.96, 1.52) m².kg⁻¹ fw and 1.25 (0.97, 1.55) m².kg⁻¹ fw, respectively. Nevertheless, the posterior estimates are different when variability arising from the physical form of the radionuclides is considered (M_3 and M_4). Posterior means (and medians) approximately 1.3 m².kg⁻¹ fw and 0.7 m².kg⁻¹ fw were obtained for particulate radioisotopes and gaseous iodine, respectively. This indicates that whatever the considered plant species, the interception process could be approximately twice as efficient for radionuclides in particular form as for gaseous radionuclides. Furthermore, the necessity of modeling the variability between the physical form of radionuclides in our models becomes more clear when we consider the posterior distribution of the parameter V_G , approximately at 0.38 (0.01, 1.79) m².kg⁻¹ fw and 0.55 (0.01, 2.01) m².kg⁻¹ fw for M_3 and M_4 , respectively.

The posterior mean (and median) of the residual standard error, S , was approximately 0.62 (0.55, 0.70) for M_1 and M_2 , 0.56 (0.50, 0.63) for M_3 and M_4 . Thus, the residual standard error decreases and is more precisely estimated when the variability of the rates of convergence due to the physical form of the radionuclides is accounted for. On the other hand, taking into account the between-species variability of the asymptotic interception factors does not seem to provide any changes in the posterior distribution of this parameter.

3.2. Model selection

The results provided in Section 3.1 highlight the influence of the variability of the rates of convergence due to the physical form of the radionuclides in the interception process. This is confirmed by the DIC values shown in Table 3, with the hierarchical models, M_3 and M_4 , offering the best fitting performances to the dataset described in Section 2.1 (with a DIC difference around 30).

Model selection was completed by comparing the predictive performances of the 4 models performing a LOOCV procedure and calculating quantitative criteria such as PPLC and coverage indices, for all models as described in Section 2.4. Fig. 3 shows the 95% CIs of the posterior predictive distributions of the logit-transformed interception factors versus the observed logit-transformed interception factors (not included during the inference process) for all the models. Four posterior predictive distributions do not match the corresponding observed logit-transformed interception factors for any model. The two maximal data correspond to the two highest interception factor values measured by Réal et al. (1993) for ^{110m}Ag and ⁸⁵Sr particles for mature lettuces and the two minimal data correspond to the two smallest interception factor values related to gaseous ¹³¹I interception by pasture grass and drawn from Chamberlain and Garland (1991). These discrepancies are explained mainly by the difficulty of the logit transformation to deal with interception factor values that are in the close neighborhood of the extremes 0 and 1 (which are not handled by the logit transformation). These deficiencies are more noticeable for the highest predicted values that behave to us, like “outliers”: indeed interception factor values having their corresponding standing biomass values in the same range do not exceed 0.9. In addition, one must keep in mind that at each step of the LOOCV procedure, the models are fitted on a learning set and the predictions computed from this learning set are compared to the logit-transformed interception factor value that is excluded from the fit. Thus, our models show some difficulty to provide good predictions for these values. However, the posterior predictions obtained for M_1 to M_4 are satisfactory as shown by the coverage ratios, more than approximately 96%, and the average widths of the 95%CIs, summarized in Table 4.

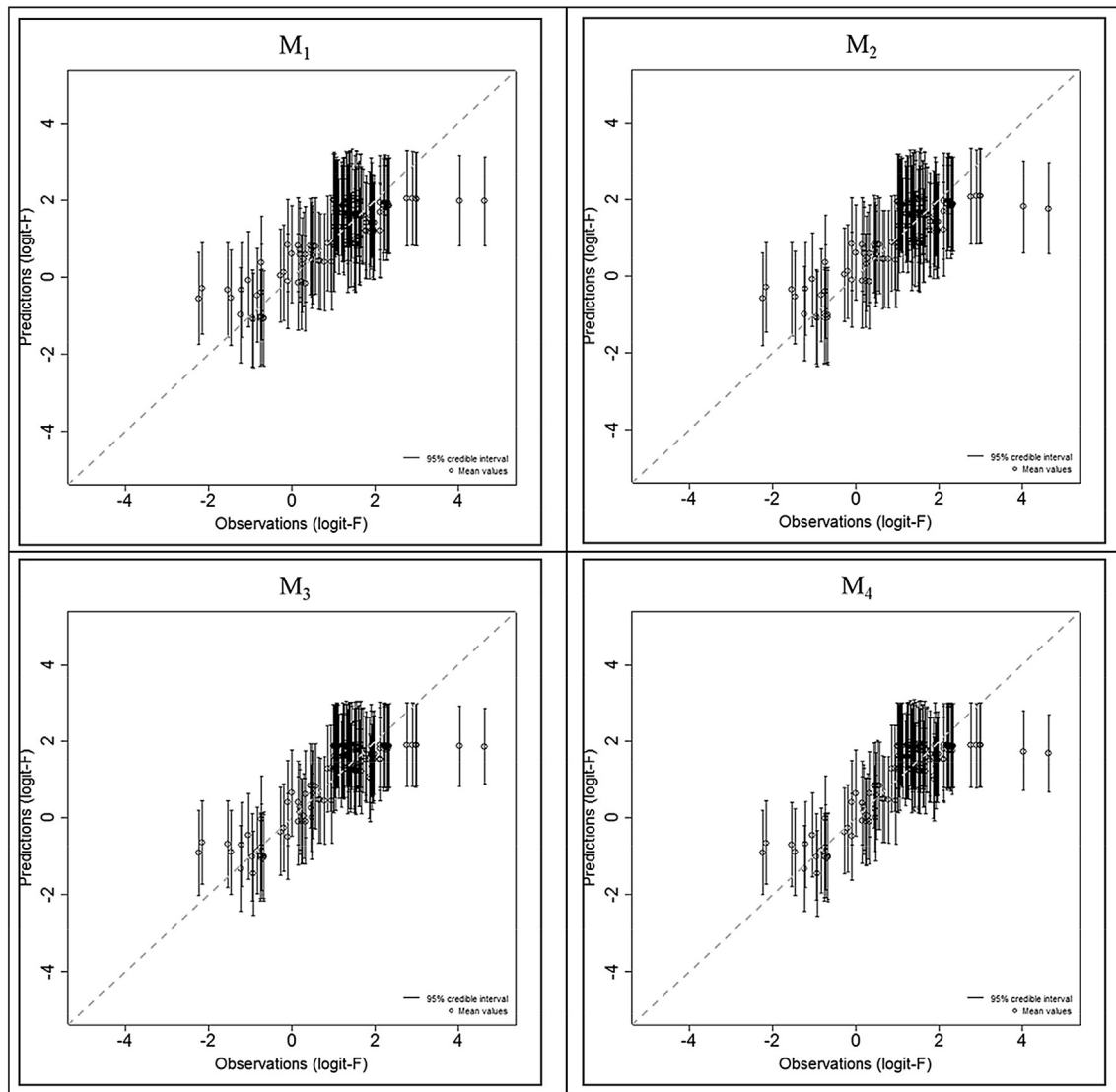


Fig. 3. Posterior predictions versus observed logit-transformed interception factors for models M_1 to M_4 . The dashed line is the first bisector. Mean values of posterior predictions are represented by circular points and the 95% CIs by vertical segments.

The PPLC values (see Table 4) (calculated for $k = 1$) are larger for the models that do not consider variability induced by the physical form of radionuclides. This is also true for $k = 10, 100$ and ∞ . Consequently, from the graphical representation derived from the LOOCV procedure, and the calculated quantitative criteria, hierarchical models, M_3 and M_4 appear to offer the best predictive performances. The predicted logit-transformed interception factors and the subsequent predicted interception factors are not only more accurate but also more precise when drawn from M_3 and M_4 than from M_1 and M_2 . M_3 and M_4 also produce less uncertain logit-transformed interception factors, as demonstrated by the average width of the 95% CIs.

Even though the fitting and predictive performances of M_3 and M_4 seem similar, there is no clear evidence of improvement by modeling the between-species asymptotic interception variability to describe the observed interception factors values. This is also suggested by the posterior distributions estimated for V_A (M_2 and M_4) as discussed in Section 3.1. Consequently, we selected M_3 as the best model in terms of fitting and predictive performances to describe and produce reliable and relevant values of interception factors.

3.3. Posterior simulations from M_3

Fig. 4 shows the prior and posterior probability distributions of the parameters and latent variables (i.e., rates of convergence) of M_3 . The correlations coefficients between the posterior distributions of the hyper-parameters (A , DMC , μ_G , τ_G and S) were computed and the results (not shown) are satisfactory. See Figure S-3 of the supplementary materials.

The available data contribute to strongly update the prior beliefs of all the parameters A , μ_G , τ_G and S for which flat uniform priors have been considered. Indeed, the range of the posterior distribution shrinks, meaning a reduction in uncertainty. For DMC , we observe an updating of the prior belief characterized by the reduction of the uncertainty with a slight shift of the location of the most probable value (and also the extreme values) proposed in the TRS472 (IAEA, 2010) from 0.215 to approximately 0.19 with a 95%CI (0.14, 0.25). In Bayesian analysis, the definition of the prior belief is an important issue that has to be discussed in details. In our analysis even if very weakly informative prior distributions were considered, complementary sensitivity analyses were performed to

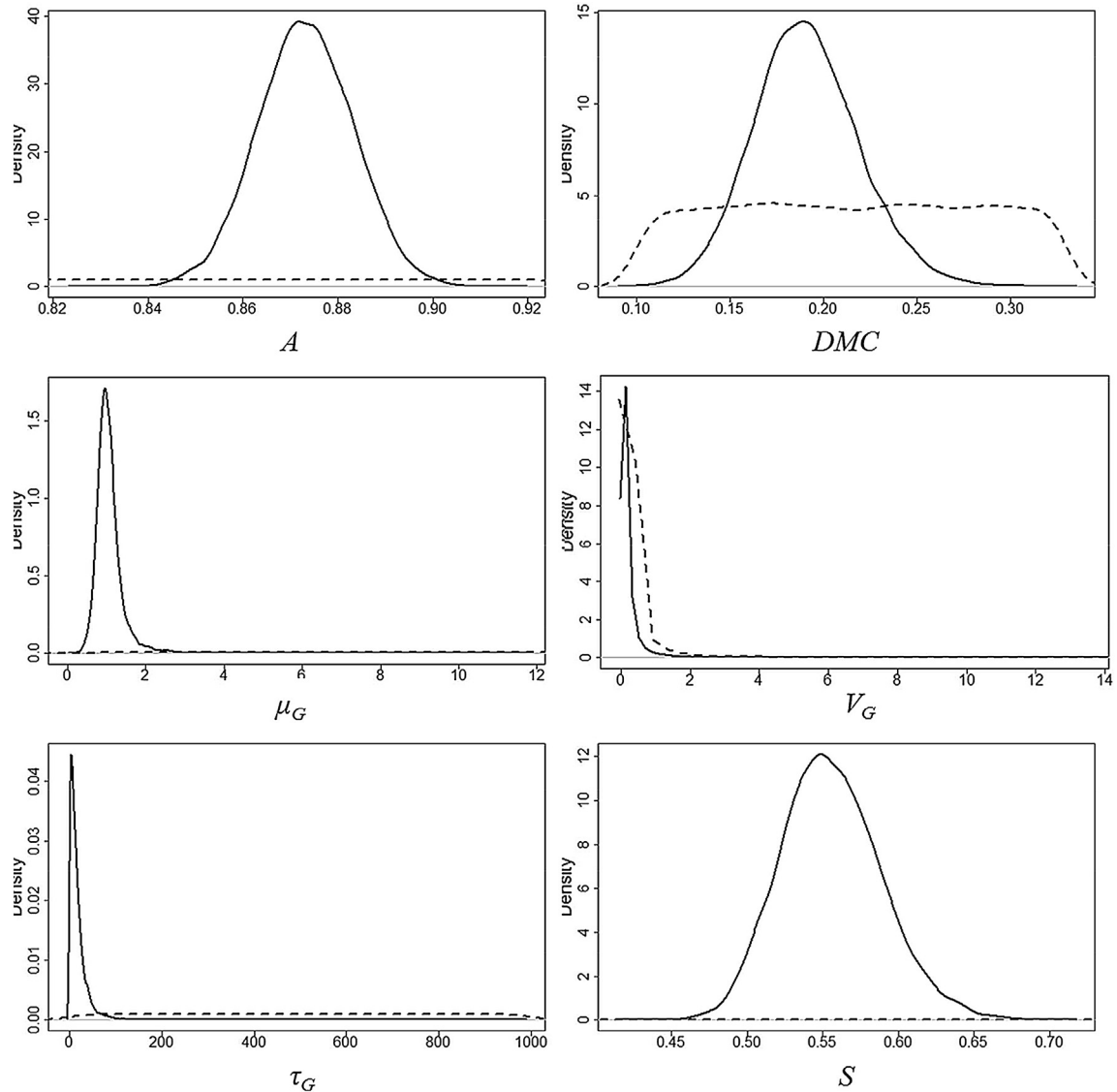


Fig. 4. Prior and posterior probability distributions of the parameters (A, DMC, μ_G , τ_G and S) of the selected model M_3 . Dashed lines correspond to the priors and the solid lines to the posteriors. For illustrative purpose V_G is also represented.

investigate the role on this analysis of the defined priors. (See [supplementary materials](#)).

It is interesting to give an illustrative view of the predictive results from the M_3 and compare these to “real-life”. Logit-transformed interception factors values were predicted for a set of standing biomass values ranging from 0 to 6 kg fw.m⁻² using model M_3 , considering the posterior distributions of the asymptotic interception factor and the rate of convergence corresponding to radioactive particles, and the residual standard deviation. The inverse of the logit function was applied to obtain predicted interception factors. For each standing biomass value (x axis), the predicted distribution of the interception factor was represented by the 2.5, 25, 75 and 97.5 percentiles, as shown in [Fig. 5](#). These predicted interception factor distributions were then compared to the interception factor values measured by [Réal and Miara \(1990\)](#) and by [Watterson and Nicholson \(1996\)](#). Recall that the aboveground biomass values of the mature lettuces were missing and thus simulated according to the probabilistic model defined in [Section 2.5](#), and are also shown in [Fig. 5](#) as black horizontal segments whose y coordinate corresponds to each associated measured

interception factor. The predictive results are in good agreement with the measured data. Only two observed interception factors, corresponding to 13% of the lettuce data (2 on 15 observations) and approximately 2% of the whole data (2 on the 145 observations) used in the meta-analysis for the models inference, are larger than our predictions, and outside the “95% uncertainty region” (the area between the 2.5 and the 97.5 percentile uncertainty curves) while all the 6 observed data of [Watterson and Nicholson \(1996\)](#) are inside this 95% uncertainty region.

4. Discussion

It has been recommended by international organizations, such as World Health Organization (WHO), Food and Agriculture Organization of the United Nations (FAO) and the Codex Alimentarius Commission, that sources of uncertainty should be quantified (or, failing that, qualified) to obtain more realistic food risk assessments for many hazards ([Codex Alimentarius, 2007](#); [FAO/WHO, 2011](#)). In the context of the assessment of dosimetric impact following releases of radionuclides, [IAEA \(1989\)](#) has provided similar

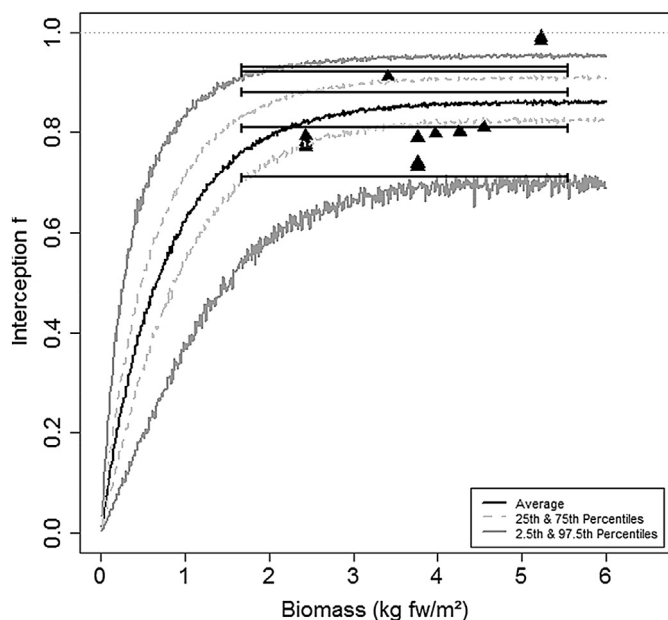


Fig. 5. Posterior predictions obtained from model M_3 . Black curve corresponds to the mean predictions values, gray curves correspond to the 50% (light gray) and to the 95% (dark gray) uncertainty bandwidths. The black segments represent the six (6) data for which aboveground biomass values are unavailable. Among these values, 2 are equal; resulting in 5 segments. Black triangles refer to the observations (interception fraction values, standing biomass values) corresponding to the lettuces leaves in our dataset.

recommendations. This paper focused on the uncertainties around the atmosphere-plant interface and, specifically, on the current lack of knowledge regarding the dry interception factor.

Data were collected from published scientific papers and, a Bayesian meta-analysis performed to provide relevant characterization of the uncertainties relative to dry interception of radionuclides by pasture grass and the vegetables leaves in the context of heterogeneous data from different sources.

Several assumptions were made to represent the variability of interception factors that could be induced by physiological specificities of plants, and by the physical forms of radionuclides (i.e., particles or gaseous radionuclides). Four models based on the non-linear Chapman-Richards growth equation, were proposed to model dry interception as a function of biomass. These are Bayesian hierarchical¹ models and cover all the possible assumptions between *i*) dry interception is totally independent of vegetation type and radionuclide physical form, and *ii*) dry interception is plant and element specific. The hierarchical modeling proposed here introduces good flexibility for combining plant and element specific parameters.

Although Bayesian statistics are well-known and widely used in various fields their use in radioecology is relatively uncommon (Hinton et al., 2013; Hosseini et al., 2013). The benefit of Bayesian inference lies in the flexibility and efficiency to address features of hierarchical models (Bernardo, 2011; Draper, 1995; Li et al., 2012; Wakefield et al., 1994). This statistical approach is well-adapted for characterizing (parametric) uncertainties when data paucity is encountered (Allenby et al., 2005; Parent and Bernier, 2007), and/or when the dataset is unbalanced, i.e. not obtained through a (balanced) design (Wolfinger and Kass, 2000).

¹ Only M_1 is not hierarchical since neither the plant-species variability nor the variability due to the physical form of the radionuclide is taken into account (see Fig. 2).

The interception factor has often been assumed to be Gaussian (Chamberlain and Garland, 1991; Vandecasteele et al., 2001). However, this choice is not theoretically appropriate, as the interception factor is a dimensionless parameter defined as a ratio over the interval (0, 1). We proposed an alternative consisting of applying a logit transformation on both the interception factor to be explained and the expected interception factor given by the Chapman-Richards equation to make the Gaussian choice relevant (Kieschnick and McCullough, 2003; Paolino, 2001).

Due to this logit transformation, one constraint is that the model is only applicable to values in the open interval (0,1). This excludes null interception efficiency, which means that the model is not applicable when vegetation is negligible over the ground surface. In the context of bare soil, the only contamination pathway of plants subsequent to the deposition process is by root uptake (De Lurdes Dinis and Fiúza, 2007; IAEA, 2009). A deterministic value equal to 0 can then be considered for the interception factor in such a case. The use of the logit transformation also excludes an interception factor of 1, i.e., when vegetation completely covers the ground and subsequently intercepts 100% of the deposited radionuclides. A deterministic value equal to 1 can be considered (as for bare soil context). An alternative Beta-Regression model (Bayes et al., 2012; Ferrari and Cribari-Neto, 2004) could also be considered to describe the relationship between the interception factors and plant biomass in future work, if this specific case of full interception by the vegetation had to be dealt with.

The models M_1 to M_4 implicitly assume, by their (hierarchical) construction, the exchangeability of data and the exchangeability of the latent variables (Jackman, 2009). When doing (Bayesian) meta-analysis, this concept of (data and parameters) exchangeability is an important assumption since it mathematically characterizes the judgment of the similarity between data from different related sources (Bernardo, 1996). Specifically, in our analysis, we have implemented Bayesian meta-analysis using a hierarchical modeling framework. The implications of (data and parameters) exchangeability can be seen at different levels regarding the considered models.

- For M_1 , the exchangeability refers to the whole data (whatever the sources) since we investigated by construction the fact that there is no evidence of distinguishing plants species and the physical forms of the radionuclides.
- For M_2 , M_3 and M_4 the exchangeability refers to both the latent variables and the replicates related to each underlying group (groups specific to the physical form of the radionuclides and/or plant-specific groups). In these models we investigated for instance the fact that for a given group there is no evidence to distinguish at an individual level the element of this group. By group we mean for instance lettuces, grass, particles or gaseous radioiodine etc.

For more details on this concept of exchangeability and on its implications on the (Bayesian) hierarchical statistical models see Jackman (2009).

The best model in terms of fitting and predictive performance is M_3 in which all plant species have the same asymptotic interception factors but the rates of convergence are assumed to be specific to the radionuclides physical form (even if they share some similarities). For M_3 , the posterior mean of the rates of convergence, G_1 and G_2 , were approximately 1.58 (1.34, 1.86) $\text{m}^2 \cdot \text{kg}^{-1} \text{fw}$ (95%CI) for particles and 0.83 (0.70, 0.98) $\text{m}^2 \cdot \text{kg}^{-1} \text{fw}$ for gaseous radioiodine.

A factor of approximately two was estimated between the rates of convergence G_1 and G_2 . This reflects that for the different plant species the interception process could be twice as efficient, when their standing biomasses are low (less than about 1.25 $\text{kg fw} \cdot \text{m}^{-2}$),

for radioactive particles as for gaseous radioiodine. The aerodynamic processes, which impact differently for gas and particles (Tschiersch et al., 2009), and the structure of the plant leaves combined with planting density during the radionuclide deposition onto plant surfaces may explain this difference.

There have been multiple reviews (IAEA, 2010; Pröhl and Hoffman, 1996) about absorption coefficient values (generally called μ or α and expressed in $\text{m}^2 \cdot \text{kg}^{-1} \text{ fw}$ or in $\text{m}^2 \cdot \text{kg}^{-1} \text{ dw}$)² for radionuclides. This parameter is close, but is not directly comparable to the rate of convergence, G , defined in this study. The asymptotic interception factor A , is assumed to be unknown and thus estimated in our work, but elsewhere is assumed to be fixed and equal to 1 (IAEA, 2010; Pröhl and Hoffman, 1996). Since A is estimated to be less than 1 in our work, this may contribute to larger estimated values for G than those obtained with a fixed asymptotic interception factor at 1. For the main dataset with gaseous ¹³¹I, the absorption coefficient is estimated by Chadwick and Chamberlain (1970) at $2.8 \text{ m}^2 \cdot \text{kg}^{-1} \text{ dw}$, which is slightly outside our estimated 95% CI (2.92, 4.09) in $\text{m}^2 \cdot \text{kg}^{-1} \text{ dw}$ for G_2 . The absorption coefficients reported by IAEA (2010) and Pröhl and Hoffman (1996) for dry interception of (non-radioactive) particles (such as *lycopodium*, *quartz* and *silica sand* particles) by grass range between 0.2 and $3.0 \text{ m}^2 \cdot \text{kg}^{-1} \text{ dw}$, which are clearly outside the 95% CI (5.58, 7.75) estimated here for G_1 .

This study provides an interesting improvement in the field of interception since lack of information has been exhibited for particles with average aerodynamic diameter below $40 \mu\text{m}$ in most reviews (Chamberlain and Garland, 1991; Pröhl and Hoffman, 1996). The dry interception factor values measured for radioactive particles, and used in the meta-analysis, correspond to particles having average aerodynamic diameters ranging between $0.75 \mu\text{m}$ and $3.5 \mu\text{m}$. In section 3.3, the 6 dry interception factors measured by Watterson and Nicholson (1996) and by Réal and Miara (1990) for particles with average aerodynamic diameter ranging from $1 \mu\text{m}$ to $22 \mu\text{m}$ were shown to lie within the 95% CI region predicted from M_3 . However, some cautions should be exercised in applying these results. Since our data were not obtained through a balanced experimental design, generalization needs to be carefully implemented, even though the Bayesian (meta-analysis) framework employed offers robust and relevant inference performance.

Another improvement brought by this study lies in the fact that biomass values measured in a dry weight basis and biomass values measured in a fresh weight basis can be handled simultaneously by considering in the analysis the plant-specific dry matter contents. In addition, this relatively innovative feature could be extended to other fields in radioecology where accounting for the uncertainty induced by dry and fresh weights is a source of limitation.

This work aims at enhancing the field of (dry) foliar interception where continuous scientific research effort, particularly for uncertainty characterization, is required (Hinton et al., 2013), and where there is paucity of knowledge for different plant species, except pasture grass, and different types of radionuclides. The meta-analysis applied here, using Bayesian hierarchical models, provides relevant posterior probability distributions describing the uncertainty for some characteristics of the interception process as well as useful tools to reproduce plausible interception factors given a plant biomass. One of the further developments expected from this work is to upgrade the probabilistic database of SYMBI-OSE (Gonze et al., 2011; Simon-Cornu et al., 2015), which is a modeling platform dedicated to the simulation of the transfer of

radionuclides in ecosystems developed by the French Institute of Radioprotection and Nuclear Safety (IRSN). This would improve the way uncertainties are quantified and considered in risk assessment models.

Acknowledgments

This work is a fulfillment of author's PhD thesis, which is supported by the French region PACA. Authors acknowledge Laurent Camoin (from the Chamber of Agriculture of the Bouches-du-Rhône, France) for the support in the elicitation of agricultural parameters. Authors are also grateful to the referees' suggestions and comments that strongly contribute to enhance the quality of this manuscript.

Appendix A. Leave-one-out cross-validation (LOOCV) procedure

The procedure applied for the Leave-one-out cross validation is defined as follows.

At each iteration s : $s = (1, \dots, N = 145)$

- Step 1: Divide the dataset into two sample sets: a validation set composed of a unique couple of data $D_s = (F_s^*, B_s)$, and a learning set that includes the $N-1$ remaining couples of data, $D_{(s)} = (F_{(s)}^*, B_{(s)})$
- Step 2: Fit the model M_m to the learning set $D_{(s)}$ with an MCMC algorithm.
- Step 3: Generate 10,000 values of logit-transformed interception factors ($F_{(s)}^*$) from B_s (the biomass value corresponding to the validation sample) according to model M_m using the posterior samples obtained at step 2 from the learning set.

At the end of the procedure, we obtain 145 samples of 10,000 logit-transformed interception factors. Each sample $\{F_{(s)}^* : s = 1, \dots, 145\}$ has then been compared to the observed logit-transformed values F_s^* , which have not been used to fit the model (see Fig. 3).

Appendix B. Supplementary data

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.jenvrad.2015.05.007>.

References

- Allenby, G.M., Rossi, P.E., McCulloch, R.E., 2005. Hierarchical Bayes Models: a Practitioners Guide. Available at: SSRN. <http://dx.doi.org/10.2139/ssrn.655541>.
- Banerjee, S., Gelfand, A.E., Carlin, B.P., 2004. Hierarchical Modeling and Analysis for Spatial Data. CRC Press.
- Bayes, C.L., Bazán, J.L., García, C., 2012. A new robust regression model for proportions. *Bayesian Anal.* 7 (4), 841–866.
- Bayes, M., Price, M., 1763. An essay towards solving a problem in the doctrine of chances. By the Late Rev. Mr. Bayes, F. R. S. Communicated by Mr. Price, in a Letter to John Canton. A. M. F. R. S. *Philos. Trans.* 53, 370–418. <http://dx.doi.org/10.1098/rstl.1763.0053>.
- Bernardo, J.M., 1996. The concept of exchangeability and its applications. *Far East J. Math. Sci.* 4, 111–121.
- Bernardo, J.M., 2011. *Bayesian Statistics. International Encyclopedia of Statistical Science.* Springer, pp. 107–133.
- Boreux, J.-J., Parent, E., Bernier, J., 2010. *Pratique du calcul bayésien.* Springer, Paris.
- Bradlow, E.T., Wainer, H., Wang, X., 1999. A Bayesian random effects model for testlets. *Psychometrika* 64 (2), 153–168.
- Brooks, S.P., Gelman, A., 1998. General methods for monitoring convergence of iterative simulations. *J. Comput. Graph. Stat.* 7 (4), 434–455.
- Brophy, J.M., Belisle, P., Joseph, L., 2003. Evidence for use of coronary stents: a hierarchical Bayesian meta-analysis. *Ann. Intern. Med.* 138 (10), 777–786.
- Chadwick, R.C., Chamberlain, A.C., 1970. Field loss of radionuclides from grass. *Atmos. Environ.* (1967) 4 (1), 51–56.

² When absorption coefficients were expressed in $\text{m}^2 \cdot \text{kg}^{-1} \text{ dw}$ (dry weight) in the literature, the dry matter content conversion factor cited in Section 2.1 was applied.

- Chamberlain, A.C., Garland, J.A., 1991. Interception of Radioactive Fallout by Vegetation. Atomic Energy Research Establishment, Harwell. AERE-R-13826.
- Codex Alimentarius, 2007. Working Principles for Risk Analysis for Food Safety for Application by Governments. Codex Alimentarius.
- Cohen, D., Raffin, M., Canitano, R., Bodeau, N., Bonnot, O., Périsset, D., Consoli, A., Laurent, C., 2013. Risperidone or aripiprazole in children and adolescents with autism and/or intellectual disability: a Bayesian meta-analysis of efficacy and secondary effects. *Res. Autism Spectr. Disord.* 7 (1), 167–175. <http://dx.doi.org/10.1016/j.rasd.2012.08.001>.
- Colle, C., Madoz-Escande, C., 2007. Revue bibliographique et expertise des valeurs de paramètres de transfert foliaire dans la biosphère. 2^{ème} partie : l'indice de surface foliaire (LAI). IRSN report DEI/SECRE n°:07-015.
- Cramer, J.S., 2003. *Logit Models from Economics and Other Fields*. Cambridge. Book DOI: <http://dx.doi.org/10.107/CB09780511615412>.
- De La Cruz-Mesía, R., Marshall, G., 2006. Non-linear random effects models with continuous time autoregressive errors: a Bayesian approach. *Stat. Med.* 25 (9), 1471–1484.
- De Lurdes Dinis, M., Fiúza, A., 2007. Exposure Assessment to Radionuclides Transfer in Food Chain. Multiple Stressors: a Challenge for the Future. Springer, pp. 309–323.
- Dohoo, I., Stryhn, H., Sanchez, J., 2007. Evaluation of underlying risk as a source of heterogeneity in meta-analyses: a simulation study of Bayesian and frequentist implementations of three models. *Prev. Vet. Med.* 81 (1–3), 38–55.
- Draper, D., 1995. Inference and hierarchical modeling in the social sciences. *J. Educ. Behav. Stat.* 20 (2), 115–147.
- FAO/WHO, 2011. *FAO/WHO Guide for Application of Risk Analysis Principles and Procedures during Food Safety Emergencies*. Food and Agriculture Organization of the United Nations and World Health Organization, Rome.
- Ferrari, S., Cribari-Neto, F., 2004. Beta regression for modelling rates and proportions. *J. Appl. Stat.* 31 (7), 799–815.
- Gelfand, A.E., Ghosh, S.K., 1998. Model choice: a minimum posterior predictive loss approach. *Biometrika* 85 (1), 1–11.
- Gelman, A., 2006. Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian anal.* 1 (3), 515–534.
- Gonze, M.-A., Mourlon, C., Garcia-Sanchez, L., Le Dizès, S., Nicoulaud, V., Gerber, P.P., Vermorel, F., 2011. SYMBIOSE: a Modeling Platform for Environmental Chemical Risk Assessment. International Conference on Radioecology & Environmental Radioactivity (ICRER-2011). McMaster University, Canada.
- Hastie, T., Tibshirani, R., Friedman, J., 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, second ed. Springer, New York.
- Helser, T.E., Lai, H.-L., 2004. A Bayesian hierarchical meta-analysis of fish growth: with an example for North American largemouth bass, *Micropterus salmoides*. *Ecol. Model.* 178 (3–4), 399–416.
- Hinton, T.G., Garnier-Laplace, J., Vandenhove, H., Dowdall, M., Adam-Guillermin, C., Alonzo, F., Barnett, C., Beaugelin-Seiller, K., Beresford, N.A., Bradshaw, C., Brown, J., Eyrolle, F., Fevrier, L., Gariel, J.C., Gilbin, R., Hertel-Aas, T., Horemans, N., Howard, B.J., Ikäheimonen, T., Mora, J.C., Oughton, D., Real, A., Salbu, B., Simon-Cornu, M., Steiner, M., Sweeck, L., Vives I Batlle, J., 2013. An invitation to contribute to a strategic research agenda in radioecology. *J. Environ. Radioact.* 115, 73–82.
- Hosseini, A., Stenbergh, K., Avila, R., Beresford, N.A., Brown, J.E., 2013. Application of the Bayesian approach for derivation of PDFs for concentration ratio values. *J. Environ. Radioact.* 126, 376–387.
- IAEA, 1989. *Evaluating the Reliability of Predictions Made Using Environmental Transfer Models*. Vienna. IAEA Safety Series no. 100: 106.
- IAEA, 2009. *Quantification of Radionuclide Transfers in Terrestrial and Freshwater Environments for Radiological Assessments*. IAEA-TECDOC-1616. International Atomic Energy Agency, Vienna.
- IAEA, 2010. *Handbook of Parameter Values for the Prediction of Radionuclide Transfer in Terrestrial and Freshwater Environments*. IAEA-TRS-472. International Atomic Energy Agency, Vienna.
- Jackman, S., 2009. *Bayesian Analysis for the Social Sciences*. John Wiley & Sons, Ltd, Chichester.
- Kieschnick, R., McCullough, B.D., 2003. Regression analysis of variates observed on (0, 1): percentages, proportions and fractions. *Stat. Model.* 3 (3), 193–213.
- Kinnersley, R., Scott, L., 2001. Aerial contamination of fruit through wet deposition and particulate dry deposition. *J. Environ. Radioact.* 52 (2), 191–213.
- Kohavi, R., 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: *IJCAI'95 Proceedings of the 14th International Joint Conference on Artificial Intelligence – Volume 2*, pp. 1137–1143.
- Kruschke, J.K., 2010. *Doing Bayesian Data Analysis: a Tutorial with R and BUGS*, first ed. Academic Press/Elsevier, Burlington, MA.
- Lambert, P.C., Sutton, A.J., Burton, P.R., Abrams, K.R., Jones, D.R., 2005. How vague is vague? A simulation study of the impact of the use of vague prior distributions in MCMC using WinBUGS. *Stat. Med.* 24 (15), 2401–2428.
- Li, R., Stewart, B., Weiskittel, A., 2012. A Bayesian approach for modelling non-linear longitudinal/hierarchical data with random effects in forestry. *Forestry* 85 (1), 17–25.
- Madoz-Escande, C., Henner, P., Gonze, M.A., Bonhomme, T., Grasset, G., 2001. Contamination foliaire par aérosols radioactifs : Analyse et modélisation de l'influence d'un épisode pluvieux et du stade végétatif de la plante au moment du dépôt. cas du césium, strontium, baryum et tellure. IRSN Report DPRE/SERLAB/2001-43.
- Madoz-Escande, C., Henner, P., Bonhomme, T., Poncet-Bonnard, D., 2002. Contamination foliaire par aérosols radioactifs de légumes-feuilles et légumes-racines : Analyse et modélisation de l'influence d'un épisode pluvieux, cas du césium, strontium, baryum et tellure. IRSN Report DPRE/SERLAB/2002-52.
- Madoz-Escande, C., Bonhomme, T., Poncet-Bonnard, D., 2003. Contamination foliaire de prairie par aérosols radioactifs cas du césium, du strontium, du baryum et du tellure.- Influence des épisodes pluvieux -. IRSN Report DEI/SECRE/2003-017.
- Madoz-Escande, C., Santucci, P., 2004. RADEMIC 2000-2003-Synthèse des contaminations foliaires par aérosols radioactifs. IRSN Report DEI/SECRE/2004 – 03.
- Parent, E., Bernier, J., 2007. *Le raisonnement bayésien: modélisation et inférence*. Springer, Paris.
- Paolino, Ph., 2001. Maximum likelihood estimation of models with beta-distributed dependent variables. *Polit. Anal.* 9 (4), 325–346.
- Plummer, M., 2003. JAGS: a program for analysis of Bayesian graphical models using Gibbs sampling. In: *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003)*. Vienna. March, pp. 20–22.
- Pröhl, G., Hoffman, F.O., 1996. Radionuclide Interception and Loss Processes in Vegetation. Modelling of Radionuclide Interception and Loss Processes in Vegetation and of Transfer in Semi-natural Ecosystems: Second Report of the VAMP Terrestrial Working Group. IAEA, Vienna, pp. 9–47.
- Pröhl, G., 2009. Interception of dry and wet deposited radionuclides by vegetation. *J. Environ. Radioact.* 100 (9), 675–682.
- Qian, S.S., Donnelly, M., Schmelling, D.C., Messner, M., Linden, K.G., Cotton, C., 2004. Ultraviolet light inactivation of protozoa in drinking water: a Bayesian meta-analysis. *J. Water Res.* 38 (2), 317–326.
- Rauret, G., Vallejo, V.R., Cancio, D., Réal, J., 1995. Transfer of radionuclides in soil-plant systems following aerosol simulation of accidental release: design and first results. *J. Environ. Radioact.* 29 (2), 163–184.
- Réal, J., Miara, P., 1990. Effet d'un rejet radioactif sévère sur les plantes cultivées. Génération et caractérisation d'aérosols produits par un four à absorption atomique modifié. Programme associé EDF/CEA, Rapport Définitif. Institut de Protection et de Sureté Nucléaire.
- Réal, J., Rauret, G., Cancio, D., 1993. Rapport du contrat CCE TARRAS.
- Réal, J., Hamoniaux, M., Rouxel, R., 1994. Transfert à la pomme de terre et à la tomate d'aérosols de césium, strontium et argent représentatifs d'un accident nucléaire (D. SERE, IPSN).
- Richards, F.J., 1959. A flexible growth function for empirical use. *J. Exp. Bot.* 10 (2), 290–301.
- Rigaux, C., Denis, J.B., Albert, I., Carlin, F., 2013. A meta-analysis accounting for sources of variability to estimate heat resistance reference parameters of bacteria using hierarchical Bayesian modeling: estimation of D at 121.1 °C and pH 7. z_T and z_{pH} of *Geobacillus stearothermophilus*. *Int. J. Food Microbiol.* 161 (2), 112–120.
- Robert, C.P., Casella, G., 2004. *Monte Carlo Statistical Methods*. Springer, New York.
- Sehmel, G.A., Hodgson, W.H., 1978. *Model for Predicting Dry Deposition of Particles and Gases to Environmental Surfaces*. Battelle Pacific Northwest Labs, Richland, WA (USA).
- Sehmel, G.A., 1980. Particle and gas dry deposition: a review. *Atmos. Env.* 1967 14 (9), 983–1011.
- Simon-Cornu, M., Beaugelin-Seiller, K., Boyer, P., Calmon, P., Garcia-Sanchez, L., Mourlon, C., Nicoulaud, V., Sy, M.M., Gonze, M.A., 2015. Evaluating variability and uncertainty in radiological impact assessment with the probabilistic database of SYMBIOSE. *J. Environ. Radioact.* 139, 91–102.
- Simpkins, A.A., 2005. Method for estimating ingestion doses to the public near the Savannah River site following an accidental atmospheric release. *Health Phys.* 88 (2), 133–138.
- Smyth, G.K., 2006. Nonlinear regression. *Encycl. Environmetrics* 4, 1405–1411.
- Spiegelhalter, D.J., Best, N.G., Carlin, B.P., Linde, A.V.D., 2002. Bayesian measures of model complexity and fit. *J. Roy. Stat. Soc. Ser. B Stat. Methodol.* 64 (4), 583–639.
- Thiessen, K.M., Thorne, M.C., Maul, P.R., Pröhl, G., Wheeler, H.S., 1999. Modelling radionuclide distribution and transport in the environment. *Environ. Pollut.* 100 (1–3), 151–177.
- Tschiersch, J., Shinonaga, T., Heuberger, H., 2009. Dry deposition of gaseous radioiodine and particulate radiocaesium onto leafy vegetables. *Sci. Tot. Environ.* 407 (21), 5685–5693.
- Tweedie, R.L., Scott, D.J., Biggerstaff, B.J., Mengersen, K.L., 1996. Bayesian meta-analysis, with application to studies of ETS and lung cancer. *Lung Cancer* 14 (Suppl. 1), S171–S194.
- Vandecasteele, C.M., Baker, S., Förstel, H., Muzinsky, M., Millan, R., Madoz-Escande, C., Tormos, J., Sauras, T., Schulte, E., Colle, C., 2001. Interception, retention and translocation under greenhouse conditions of radiocaesium and radiostrontium from a simulated accidental source. *Sci. Tot. Environ.* 278 (1–3), 199–214.
- Wakefield, J., Smith, A., Racine-Poon, A., Gelfand, A.E., 1994. Bayesian analysis of linear and non-linear population models by using the Gibbs sampler. *Appl. Stat.* 201–221.
- Watterson, J.D., Nicholson, K.W., 1996. Dry deposition and interception of 4–22 µm diameter particles to a lettuce crop. *J. Aerosol Sci.* 27 (5), 759–767.
- Wolfinger, R.D., Kass, R.E., 2000. Nonconjugate Bayesian analysis of variance component models. *Biometrics* 56 (3), 768–774.

6.3 Résultats, limites et perspectives

Les différents modèles présentés dans cet article ont permis d'étudier l'incertitude sur le facteur d'interception en condition sèche en proposant une caractérisation de la variabilité spécifique aux espèces végétales et aux caractéristiques physiques des radionucléides considérés.

Sur la base des critères quantitatifs calculés, le modèle prenant en compte la variabilité induite par la forme physique du radionucléide est celui qui s'ajuste le mieux aux données et qui prédit le mieux le comportement du facteur d'interception sèche en fonction de la biomasse aérienne. Dans la mesure où la variabilité inter-espèces végétales est évaluée uniquement sur la base du paramètre d'interception maximale, on peut penser que lorsque le couvert végétal est assez dense pour recouvrir presque entièrement la surface du sol, la différence entre espèces végétales joue peu sur les potentielles variations des valeurs du facteur d'interception. L'influence observée de la variabilité due à la forme physique des radionucléides reflète une différence attendue de comportement entre, d'un côté, les particules et, de l'autre, l'iode gazeux. Cette différence est traduite par l'écart entre les valeurs des variables latentes G_1 et G_2 représentant les taux de convergence vers l'interception maximale pour les particules et l'iode gazeux, respectivement. Il est pour l'instant difficile pour nous de généraliser cette différence de comportement à tous les radionucléides sous forme gazeuse et surtout d'apporter une explication claire des facteurs décrivant cet écart. Pareil écart pourrait également être expliqué par une différence entre les formes chimiques des radionucléides considérés dans l'étude. Une étude plus approfondie basée sur une modélisation mécaniste des processus impliqués dans l'interception sèche, prenant par exemple en compte les vitesses de dépôt sec sur les surfaces végétales, la taille des particules, l'ouverture des stomates des plantes, la pilosité à la surface des feuilles, permettrait d'apporter quelques éléments d'explication.

L'une des innovations apportées par les modèles d'interception sèche développés dans cet article concerne la prise en compte de l'incertitude sur le ratio poids sec/poids frais. Généralement, les mesures de biomasse sont présentées soit en poids sec, soit en poids frais, ce qui génère une certaine hétérogénéité sur des données renseignant la même information. Ainsi, les modèles développés offrent la possibilité de prédire des valeurs d'interception sèche que l'on soit en poids sec ou en poids frais en prenant en compte l'incertitude sur le ratio

poids sec/poids frais. Notre choix de modéliser l'interception sèche en fonction de la biomasse fraîche est principalement motivé par le fait qu'il est nécessaire de connaître le stade de développement des plantes au moment du dépôt ; ce stade étant renseigné par la biomasse aérienne fraîche. Dans notre cas précis, toutes les valeurs de biomasse exprimées en poids sec sont relatives à l'herbe. Il est néanmoins important de préciser que les valeurs de biomasse fraîche présentent une variabilité saisonnière et spatiale du fait des variations de température et d'humidité, ce qui pousse la plupart des auteurs à privilégier des biomasses exprimées en poids sec. Dans le cadre de ces travaux de thèse, et en particulier dans le cadre des activités engagées dans le groupe de travail "Human Food Chain" du projet COMET, la caractérisation de la variation spatiale et saisonnière des rendements en herbe présente un grand intérêt car elle peut influencer l'interception de radionucléides et par conséquent influencer sur les niveaux d'activité observables dans des produits comme le lait ou la viande. Ce point sera discuté dans la [Section 10.3.1](#).

7. Interception humide

Ce chapitre est dédié à l'étude de l'interception de radionucléides déposés sur le couvert végétal (*e.g.* herbe, mauvaises herbes et trèfle) suite à un dépôt survenu en conditions humides (*i.e.* pluies naturelles ou simulées). Les autres types de précipitation (neige, brouillard, *etc.*) n'ont pas été considérés. Les travaux présentés dans la suite ont fait l'objet d'un article rédigé en collaboration avec Marc-André Gonze. L'article qui a été soumis le 22 Janvier 2016 dans la revue *Science of the Total Environment* est reproduit dans la [Section 7.2](#).

7.1 Objectifs et méthodes

L'interception humide est un processus impliquant un ensemble de mécanismes pouvant varier en fonction de la nature des précipitations, leur quantité, leur durée dans le temps, *etc.* Le facteur d'interception humide quantifie l'interception de radionucléides par différentes espèces végétales suite au dépôt par précipitations. La caractérisation de l'incertitude et de la variabilité autour de ce processus nécessite une base de connaissances renseignant, en plus des caractéristiques des radionucléides déposés et des espèces végétales réceptrices, les caractéristiques des précipitations entraînant les radionucléides émis dans l'atmosphère à la surface du sol et des plantes. Le constat fait sur la quantité limitée de données disponibles relatives à l'interception sèche (cf. chapitre précédent) peut également être fait pour le facteur d'interception humide.

L'un des principaux enjeux dans cette étude est d'améliorer la connaissance sur le processus d'interception en synthétisant l'information disponible et en exploitant des données jusque-là inutilisées pour la modélisation de l'interception humide (*e.g.* celles issues de l'étude publiée par [Bengtsson *et al.* \(2014\)](#)).

Dans cette étude, nous proposons tout d'abord la base de données la plus riche à notre connaissance réunissant des informations publiées dans la littérature. Ces données sont issues d'expérimentations en conditions contrôlées menées en intérieur ou en extérieur, renseignant au moins les formes physico-chimiques des substances d'intérêt et la hauteur de pluie. Suite à une première étude exploratoire, un modèle empirique ayant une structure hiérarchique a été ajusté par inférence bayésienne appliquée aux données collectées (facteur d'interception

massique). Ce modèle empirique, initialement introduit par [Hoffman et al. \(1989, 1992\)](#), exprime le facteur d'interception massique humide en fonction de la biomasse aérienne des plantes B , exprimée en poids sec, de la hauteur H (en mm) et de l'intensité I (en mm par unité de temps) de la pluie sur la base d'une régression linéaire sur les transformées logarithmiques des données. Ici, la notion de hiérarchie fait référence à la prise en compte de la variabilité spécifique des caractéristiques physico-chimiques des substances considérées dans l'étude (valence des substances). De plus, nous avons travaillé sur l'emboîtement des modèles empiriques suivant que l'on prend en compte ou non l'intensité ou la biomasse, (*i.e.* $HBI \supset HB \supset H$ ou $HBI \supset HI \supset H$).

L'inférence bayésienne appliquée à ces modèles empiriques hiérarchiques permet non seulement de caractériser l'incertitude relative au processus d'interception mais également de quantifier les variations induites par les différences physico-chimiques entre substances.

Les modèles empiriques ne permettent pas toujours d'expliquer toutes les variations relatives aux actions et/ou à l'interaction entre les mécanismes impliqués dans un processus physique. A titre d'exemple, l'écart observé entre particules et iode gazeux dans le cadre de l'interception sèche n'est pas vraiment expliqué de manière précise (cf. chapitre précédent). Ainsi, l'une des innovations dans cette étude a consisté à caractériser explicitement les mécanismes impliqués dans l'interception humide de substances (interception des gouttes de pluie à la surface des feuilles, drainage après saturation de la surface en eau, évaporation et absorption/fixation des substances sur la surface des feuilles) en proposant des modèles hiérarchiques mécanistes (cf. (ii) seconde signification de la notion de hiérarchie). Ce type de modélisation, de plus en plus répandu dans des domaines comme l'écologie (*e.g.* [Adams et al. 2013](#) ; [Cuddington et al. 2013](#) ; [Miehle et al. 2009](#)), est notamment recommandé par la communauté scientifique en radioécologie pour mieux prédire le comportement des radionucléides dans la biosphère ([Hinton et al. 2013](#)). Deux approches ont été proposées en lien avec la modélisation du processus d'absorption :

- une approche considérant un équilibre durant la période d'exposition de la plante entre la concentration fixée à sa surface et la concentration non fixée dans la pellicule d'eau ; ce qui fait de l'absorption un mécanisme réversible et instantané ayant un temps caractéristique inférieur à la durée d'exposition.
- une approche cinétique stipulant que le processus d'absorption est un processus cinétique lent avec un temps caractéristique pouvant dépasser la durée de l'exposition.

Dans tous les cas, les paramètres des modèles ont été estimés par inférence bayésienne ; ce qui permet comme précédemment de proposer une caractérisation de l'incertitude et de la variabilité substance-spécifique. Les différents modèles hiérarchiques développés dans cette étude ont été comparés sur la base de critères quantitatifs liés à leurs qualités prédictives.

7.2 Article: Interception of Wet Deposited Atmospheric Pollutants by Herbaceous Vegetation: Data Review and Modelling.

M.-A. Gonze*¹ and M.M. Sy¹

¹ *Institute of Radiation Protection and Nuclear Safety, CE Cadarache-Bat 153, BP3 – 13115 St-Paul-lez-Durance cedex, France.*

*corresponding author: marc-andre.gonze@irsn.fr

ABSTRACT

Better understanding and predicting interception of wet deposited pollutants by vegetation remains a key issue in risk assessment studies of atmospheric pollution. We develop different alternative models, following either empirical or mechanistic descriptions, on the basis of an exhaustive dataset consisted of 440 observations obtained in controlled experiments, from 1970 to 2014, for a wide variety of herbaceous plants, radioactive substances and rainfall conditions. The predictive performances of the models and the uncertainty/variability on the parameters are evaluated under Hierarchical Bayesian modelling framework. It is demonstrated that the variability of the interception fraction is satisfactorily explained and quite accurately modelled by a process-based alternative in which absorption of ionic substances onto the foliage surfaces is determined by their electrical valence. Under this assumption, the 95% credible interval of the predicted interception fraction encompasses 81% of the observations, including situations where either plant biomass or rainfall intensity are unknown. This novel approach is a serious candidate to challenge existing empirical relationships in radiological or chemical risk assessment tools.

Keywords: Wet deposition, interception by plant, process-based modelling, Bayesian inference

1. Introduction

Rainfall plays a crucial role in the deposition of airborne pollutants onto the ground surfaces and vegetation. Where rain occurs in the presence of airborne contamination, the wash-out (entrainment of particles or dissolution of gaseous species by falling droplets) and rain-out (formation of rain drops around particles) of material from the atmosphere can be responsible for deposit levels higher than those arising from dry deposition. A recent example was given by the Fukushima nuclear accident which induced considerable radioactive deposits in March 2011, far to several tens of kilometers to the Northwest of the nuclear site. This heavy ground contamination was mainly attributed to the passage of heavy rainfall at time where high radioactive releases were dispersed in the atmospheric boundary layer (Katata et al., 2012a,b; Terada et al., 2012; Mathieu et al., 2012; Korsakissok et al., 2013; Draxler et al., 2014). Knowledge of the proportion of the deposit which is initially intercepted and retained by plants, also called wet interception fraction, is an important element of risk assessment studies, because it determines the level of contamination that can be expected in crops and other vegetation, for weeks, months or even years after deposition. In emergencies or post-accidental situations in which the consequences have to be evaluated quickly but as realistically as possible, it is of paramount importance to take into consideration the type of deposited substances, the peculiarities of the precipitation event(s) that cause deposits and the characteristics of vegetation covers in the contaminated territories. Generic approaches which do not consider the specificities of the situation might lead to significant underestimations or overestimations, by more than one order of magnitude.

Rather comprehensive reviews of the mechanisms involved in the interception of wet-deposited radioactive contaminants by crops and herbaceous species were provided by Hoffman et al. (1992), Kinnersley et al. (1997) and Pröhl (2009). The interception fraction, f (unitless), is defined as the ratio of the pollutant initially retained by the above-ground vegetation immediately

after the deposition event and the total amount deposited onto the soil-plant system. This key parameter depends first on the capacity of a plant foliage (*i.e.* leaves, stems, etc.) to intercept and store precipitation. It is usually quantified by the water storage capacity, denoted S (mm of water), expressed in liters of water per unit surface area. As discussed for example by Keim et al. (2006) and Muzylo et al. (2009), storage capacity increases with above-ground plant biomass and, to some extent, on the structure of the plant foliage as determined by the geometry and orientation with respect to the horizontal direction of the leaves (e.g. typical values for grass are slightly smaller than other crop plants due to their preferential vertical orientation). Once this storage limit is reached, subsequent precipitation run off the plant surface to the ground, and so with increasing rainfall height, the intercepted water and contamination are expected to become a decreasing fraction of the total deposited amount. The effect of mechanical movement of the foliage, induced by wind or heavy rainfall intensities, may reduce water storage capacity, and so should also be considered as controlling factors. Secondly, the interception fraction depends on the ability of the plant foliage to fix or absorb the chemical substances contained in water droplets during the time for which the plant is exposed. The physico-chemical affinity of the substances for the foliage surfaces is mainly governed by their electrical valence because foliage cuticle is negatively charge. As plant cuticle acts as a cation-exchanger at pH-values above 3 (Keppel, 1966; Franke, 1967; Schönherr, 1977; Schaefer and Reiners, 1990), it results in a strong absorption of cations and an inhibited absorption of anions. This has been confirmed by numerous experimental studies reported in the literature for a variety of plants and substances (e.g. as reviewed by Pröhl and Hoffman, 1996 and Pröhl, 2009). The experimental work by Kinnersley et al. (1997) who investigated the kinetics of Cesium attachment to crop leaf surfaces from a contaminated solution demonstrated that active chemical accumulation of Cesium occurred at plant surfaces and increased with the time of exposure, long after maximum water storage capacity has been reached. These results are compatible with Keppel's

(1966) assertion that cuticle acts as a cation exchanger. These results actually suggest that, for any precipitation amount greater than the water storage capacity, the interception fraction should increase with rainfall duration or decrease with rainfall intensities. This is contradictory with the assertions by Hoffman et al. (1995) that rainfall intensity appeared to be of minor importance in determining interception.

A few operational models have been proposed in the literature that account for the dependence of the interception fraction or upon the above-mentioned

$$f = f_0 \times \left(\frac{H}{1 \text{ mm}} \right)^\alpha \times \left(\frac{B}{1 \text{ kg d.w. m}^{-2}} \right)^\beta \times \left(\frac{I}{1 \text{ mm h}^{-1}} \right)^\gamma \quad (1)$$

Where f_0 is the interception fraction estimated for a unit biomass, unit rainfall amount and intensity. These authors evaluated f_0 for each kind of substance and each kind of plant (e.g. clover, fescue and a mixture of them) while α , β and γ parameters were kept constant among plants. These coefficients were estimated by performing a multi-linear regression (MLR) analysis of the log-transformed data. This relationship very accurately reproduced observations for anions ($R^2 > 0.9$) but proved less adequate for cations and inert particles ($R^2 < 0.7$). Although its overall predictability was satisfactory, this empirical approach did not involve any mechanistic concept and could not be extrapolated straightforwardly to another plants or substances than those considered in the field experiments. Another interesting formulation

$$f = \min \left\{ 1, \frac{LAI \times L \times k'}{H} \times \left[1 - \exp \left(- \frac{\ln(2)}{3} \times \frac{H}{L \times k'} \right) \right] \right\} \quad (2)$$

Where LAI ($\text{m}^2 \text{ m}^{-2}$) is the single-sided leaf area index, L (in mm of water) is the equivalent thickness of water film on the leaves and k' (unitless) is the chemical affinity coefficient, whose value depends on the valence of the deposited substance. Different values of k' were attributed to anions, monovalent and divalent cations, although too few data were available and published in the early 1990's to ascertain the distinction between monovalent and divalent cations. In this equation, the chemical and mechanical contributions to pollutant interception are combined into the single term $L \times k'$, while precipitation is only described in terms of its total amount. Unlike Hoffman's relationship, it thus implicitly assumes that accumulation of wet-deposited radionuclides on the foliage surfaces is independent of rainfall intensity and exposure time, $T = H/I$ (in hours).

Although these two approaches provided some insights into the modelling of the interception fraction, none of them relies on a detailed and comprehensive parameterisation of the underlying mechanisms, that is: (i) interception of water droplets by plant foliage, (ii)

factors. Most approaches rely on the use of a constant value, in some cases, tabulated against the type of substance considered. The first approach was originally proposed by Hoffman et al. (1989, 1992) based on an empirical regression analysis of a detailed data acquired for various herbaceous plant species and deposited substances (i.e. dissolved ions and micronic particles). They proposed to express the interception fraction, f , as a function of the standing biomass density, B (in kg dry weight m^{-2}), rain amount, H (in mm), and rain intensity, I (in mm h^{-1}), as follows:

was that proposed after Chernobyl accident by Müller and Pröhl (1988, 1993) and Pröhl (1990) for agricultural plants, including grass, vegetables and other crops, and implemented in the European Decision Support System for Off-site Nuclear Emergency (Ehrhardt and Weis, 2000). In contrast to Hoffman's approach, this parameterisation attempted to account for some of the driving biological and physico-chemical factors. The model parameters were mainly derived from data acquired for herbaceous species by Angeletti et al. (1977a) and Hoffman et al. (1989, 1992). The formulation relies on the use of a "water-film-leaf-area" and a chemical-dependent factor that accounts for the affinity of the chemical element for the foliage surfaces. The interception fraction was expressed as follows:

drainage of water and substances after saturation has been reached, (iii) in-storm evaporation of water (and volatile substances) and (iv) absorption or fixation of the substances to the foliage surfaces. Their respective role in the overall retention of wet-deposited substances has still to be elucidated. An important missing piece in our understanding especially concerns the way rainfall duration or intensity can influence wet interception. As interception is determined by competition between drainage and absorption, low interception is expected when drainage or rainfall occurs faster than the characteristic time required for absorption/fixation. On the contrary, if absorption is faster than saturation, or the cation exchange capacity of the leaf surfaces is low, interception should not much depend on rainfall duration. These questions motivate the development and testing of a process-based model.

In this paper, we propose to improve the mathematical parameterisation of the wet interception process and compare different alternative approaches, for a variety of radioactive substances, herbaceous plants and precipitation events. This research relies on a set of more

than four hundred observations obtained in controlled experiments for a variety of plants and substances. This dataset which constitutes the most exhaustive one ever worked on for studying wet interception of radioactive substances by plants gathers data used in the 1990's for the development of the two above-mentioned interception models as well as data never exploited for model development, such as those published by Angeletti et al. (1977b), Kinnersley et al. (1997) and Bengtsson et al. (2014). The selected experiments and the proposed models are described in section 2. In section 3, data are first explored and briefly discussed. The model parameters are then inferred for each alternative model, using the Bayesian modelling approach that enables to capture within a cohesive framework the uncertainty on parameter values. Finally, their respective performances and capability to predict the wet interception fraction for the various substances considered were assessed through quantitative criteria, compared and discussed.

2. Material and methods

2.1. Selected data

In the selection of experimental data, the focus was put on controlled experiments conducted either in greenhouses or outdoor plots, in which at least the physico-chemical form of the pollutants and the rainfall height were known and reported in the publications. In general, field measurements of wet interception of Chernobyl fallouts were not included in this study because either environmental conditions were poorly characterized or the estimation of the wet interception fraction was too much uncertain. The selected experiments are summarized in Tab. 1. They provide 440 observations obtained for different types of substances (e.g., anions or cations in soluble forms, water and small inert polystyrene particles) and herbaceous plants (e.g., clover, grass and weed species, or a mixture of them). These data were published by: Chadwick and Chamberlain (1970: 9 data), Angeletti and Levi (1977a,b: 80 data), Hoffman et al. (1989, 1992: 309 data), Jacob et al. (1993: 3 data), Kinnersley et al. (1997: 15 data) and Bengtsson et al. (2014: 24 data). As indicated in Tab. 1, the interception fraction, denoted f , was not reported and could not be estimated in some of these experiments (i.e. 65 data). The leaf area index, LAI (unitless), and the leaf area interception fraction, defined as $f_{LAI} = f/LAI$ ($m^2 kg^{-1}$ dry weight), were rarely measured and reported. Both quantities were available for experiments conducted on grass vegetation only, by Angeletti and Levi (1977a), Kinnersley et al. (1997) and Bengtsson et al. (2014). These observations (Iodine: 5, Cesium: 27, Strontium: 17 and H_2O : 5) are not sufficient to properly assess the sensitivity of the interception fraction to the area index and precipitation characteristics, for a given substance. The above ground dry biomass was preferred to the leaf area index, as this parameter was reported for 375 observations. Fresh biomasses were only reported by Angeletti and Levi (1977a) who indicated fresh to dry

biomass ratios ranging from 5.4 to 6.9 for grass. The mass interception fraction, defined as $f_B = f/B$ ($m^2 kg^{-1}$ dry weight), was systematically measured or could be estimated, like in Angeletti's experiments, from the measured ratio of the concentration in the plant to that in rainfall water.

Thus, the selected experiments include 363 observations of the mass interception fraction for which dry biomass, rainfall height and rainfall intensity (or duration) were all known and 77 remaining observations for which either dry biomass or rainfall intensity were not available (never both of them at the same time). The first group of observations was used for calibrating and testing models (i.e. "calibration" data subset) and the second group was used for evaluating the predictive performance of the models when information regarding either plant biomass or rainfall characteristics was missing (i.e. "validation" data subset).

A detailed description of each experimental conditions and method can be found in the Supplementary Information (section S.1).

The 440 measurements of the mass interception fraction, f_B , are synthetized in Figure 1 and Figure 2. The interception fraction is plotted against the rainfall height, H (mm), the rainfall intensity, I ($mm h^{-1}$), and the dry above-ground plant biomass, B ($kg dw m^{-2}$), for three arbitrary families of substances: anions (I), cations (Ba, Be, Cd, Cr, Cs, Pb and Sr) and inert micronic particles. In Figure 2, observations have been further classified into 3 different groups depending on the rainfall amount: low range (0.4-3 mm), medium range (3-12 mm) and high range (12-75 mm). The different kinds of vegetation are not discriminated here because no significant difference was found between grass, clover and weeds, except in the experiments conducted by Angeletti et al. (1977b) for which a systematic increase of the mass interception fraction for clover, by about 50 to 100%, was reported for both Iodine and Strontium. Hoffman et al. (1992) could also notice in their experiments some differences between vegetation types which included marked differences in leaf structure and texture (i.e. clover, fescue or a mixture of them), but as this was not systematic, they concluded that the influence of vegetation type was much less important than the influence of biomass and rainfall characteristics. For data obtained with an intermittent precipitation (Angeletti and Levi, 1977b; Hoffman et al., 1989), an "equivalent" rainfall amount and duration were estimated by cumulating event-based quantities, and an "equivalent" intensity was estimated as the ratio of the cumulated amount and duration. As can be seen from Figure 1 and Figure 2, the set of observations covers, for each kind of deposited substance, a wide range of variation for both rainfall amount and intensity: from 1 to 75 mm (with a median value of 8 mm) and from 1 to 122 $mm h^{-1}$ (with a median value of 29 $mm h^{-1}$). Although not shown, these observations also cover a wide range of rainfall duration, ranging from 30 seconds in the experiments conducted

by Bengtsson et al. (2014) (with intensities equal to 120 mm h⁻¹), to more than 4 hours in some experiments conducted by Angeletti and Levi (1977b) and Kinnersley et al. (1997). However, the range of durations is confined to rather short exposition durations, as 95% of the data correspond to durations shorter than 1 hour and 50% to durations shorter than 10 minutes. Biomasses range from 0.02 to 0.93 kg dw with 95% of the values being comprised between 0.05 and 0.35 kg dw, and 50% between 0.09 and 0.21 kg dw (with a median value of 0.125 kg dw). These data are thus confined within a

$$\log(f_B) = \log(f_{B0}) + \alpha \times \log\left(\frac{H}{1 \text{ mm}}\right) + \beta \times \log\left(\frac{B}{1 \text{ kg dw m}^{-2}}\right) + \gamma \times \log\left(\frac{I}{1 \text{ mm h}^{-1}}\right) + \varepsilon \quad (3)$$

Where the residual variability, ε , is assumed to stem from a normal distribution - (i.e. $\mathcal{N}(0, \Sigma)$) and where f_B is further constrained not to exceed $1/B$ (i.e. $f \leq 1$). The objective is to measure the predictive performance of this empirical model while decreasing its level of complexity. Its complexity can be decreased by ignoring the dependence of the interception fraction upon H , B or I , or no longer differentiating the substances in the estimation of the regression coefficients and the residual variability. As described in Table 2, we evaluated the predictive performances of eight alternative approaches that all derive from the model originally proposed by Hoffman et al. (1989) and. The first five models, prefixed with “HBI”, explicitly account for the dependence of f_B upon H , B and I , while the others, prefixed with “HB”, “HI” and “H”, do not account for I and/or B terms.

The different HBI-type alternative approaches differ from each other in the way they account for the dependence of the regression coefficients upon substances: for some of them, Cesium and/or polystyrene particles were estimated separately from divalent cations (i.e. Be, Sr), while for others they were grouped together. In the so-called “HBI-Biomass” approach, β coefficient was constant and independent of the substance considered. Comparing the different alternatives enables to quantify to what extent interception is sensitive to H , B and I and to what extent the substances behave differently.

2.3. Process-based model

In order to get further insights into our understanding of wet interception, it is proposed to develop and test a newly-developed approach, based on a detailed description of the underlying mechanisms that participate to the interception of wet-deposited substances by grass vegetation. Such a process-based model offers the opportunity to explain a part of the variability observed on the interception fraction and relate it to that of environmental factors.

The model relies on the resolution of mass balance equations written for both water and pollutants, with

relatively narrow range of biomasses (less than one order of magnitude).

2.2. Multi-linear regression model

To address the sensitivity of the interception fraction to the type of deposited substances, plant properties and rainfall characteristics, it is first proposed to analyze data through an empirical regression model, derived from Eq. (1) given in the introduction and originally proposed by Hoffman et al. (1989, 1992):

explicit parameterisations of the underlying hydrological, biological and chemical mechanisms. The interception fraction is computed as the fraction of the deposited mass (per unit surface of soil) which remains on the plant at the end of a contaminating precipitation event. Two alternative approaches are derived from this initial model, depending on the mathematical parameterisation of the interaction of the deposited pollutants with the plant surfaces. In the first one, we assume that absorption/fixation is instantaneous and reversible, while in the second one, it is treated as a kinetic and irreversible process.

Plant model. The model makes use of simple macroscopic parameters for describing the structure of herbaceous vegetation: the (single-sided) foliage area index, LAI (m² m⁻²), the foliage capacity, S (mm), which is defined as the amount of water left on a saturated foliage under zero evaporation conditions and after foliage drainage have ceased (Gash and Morton, 1978), and the free throughfall coefficient, denoted p (unitless), which is the fraction of rainfall amount that reaches the ground without hitting the plant. For dense (non sparse) herbaceous layer, the value of p is close to one. Following van Dijk (2001), the foliage capacity can be linearly related to the foliage area index through the so-called specific foliage storage, L (mm), as follows:

$$S = LAI \times L \quad (4)$$

This parameter which quantifies the maximum thickness of the water film that can be retained by leaves typically ranges between 0.2 and 0.4 mm for grass vegetation (Clark, 1940; Merriam, 1961). Because aerial organs such as stems and reproductive organs can also intercept contaminated water droplets, it is more appropriate to use in Eq. (4) the (single-sided) plant area index which characterizes the whole aerial surface. In the followings, we will implicitly assume that LAI at least includes the contribution of grass stems (the contribution of other organs can be neglected for grass vegetation). Some relationships have been proposed in the literature to link the foliage area index to the foliar dry biomass that are often linear or quasi-linear (Müller and Pröhl, 1993;

Jouven, 2006a; Jouven 2006b), where the coefficient of proportionality is the so-called specific leaf area. By analogy, we introduce in the model a specific foliage area, denoted SLA ($m^2 kg^{-1} dw$), such as:

$$LAI = SLA \times B \quad (5)$$

This parameter which implicitly includes the contribution of both grass foliage and stems can be estimated as the weighted average of both specific areas:

$$SLA = \%Leaf \times SLA_{leaf} + (1 - \%Leaf) \times SLA_{stem} \quad (6)$$

Where $\%Leaf$ represents the contribution of leaves in the total aboveground biomass. For grass species, $\%Leaf$ amounts to about 2/3 of the total biomass typically (Jouven et al., 2006a,b). The vegetation sub-model thus relies on the use of 4 agronomical parameters – B , SLA , L and $p - S$ and LAI being calculated with Eq. (4) and Eq. (5), respectively. SLA , L and p can be reasonably assumed constant throughout the vegetative stages, from plant emergence until the point at which leaf area begins to senesce (Storkey, 2004, 2005, 2006; Storkey et al., 2007). When plant switches to reproductive stages, the specific area index begins to decline due to both the dying of the foliage and the increase of biomass due to the growth of storage organ.

Hydrological sub-model. The hydrological sub-model relies on a simple mass balance between rain, evaporation and water drainage, which writes as follows:

$$\frac{d}{dt}W = (1 - p) \times I - E - D \quad (7)$$

Where W (mm) is the plant water storage, $E(t)$ ($mm h^{-1}$) is the plant loss by evaporation and $D(t)$ ($mm h^{-1}$) is the drainage or dripping rate. This balance is maintained over time t throughout the rainfall duration. In this equation, W is constrained not to exceed the storage capacity S (mm) because we assume that water in excess drains off the plant instantaneously. As rain starts, the foliage water storage usually undergoes three stages: wetting, saturation and drying (after rain has ceased). Saturation may not be reached before rainfall stops if rainfall height is too low or duration too short. Saturation time will be denoted T_S (hr). Following Rutter et al. (1971), the evaporation of the intercepted water can write:

$$E = Min\left(\frac{W}{S}, 1\right) \times E_0 \quad (8)$$

Where E_0 ($mm h^{-1}$) is the potential rate of evaporation (i.e. wet foliage evaporation). Typical values of E_0 for herbaceous vegetation and temperate climates can be calculated using Penman-Monteith equation with the foliage resistance set to zero: they range from 0 to 0.45 $mm h^{-1}$. We notice that, if $E_0 \ll (1-p) \times I$, saturation is reached at the following time:

$$T_S \approx \frac{S}{(1-p) \times I} \quad (9)$$

To simplify the hydrological model, we assume that there is no dripping during the wetting phase and that water in excess evacuates instantaneously once saturation is reached. Under these assumptions, the drainage term writes:

$$D = \begin{cases} 0 & \text{if } T \leq T_S \\ D_S = (1-p) \times I - E_0 & \text{if } T > T_S \end{cases} \quad (10)$$

The numerical solving of Eq. (7), (8) and (10) enables to calculate the temporal evolution of W , E and D variables.

Pollutant sub-model. The innovative aspect of the present model mainly resides in the pollutant sub-model. It relies on the assumption that, once intercepted, the pollutant either remains available for drainage (and evaporation if it is volatile) or get fixed to the foliage surfaces (i.e. cuticles) or even incorporated in the living biomass. The pollutants which are initially available can be progressively transferred to the “non-available” pool through an either reversible or irreversible process. This “absorption” process aggregates all physical, chemical and biological underlying mechanisms that are likely to decrease the availability of intercepted pollutants with respect to drainage. We thus propose to adopt the following mass conservation equations:

$$\begin{cases} \frac{d}{dt}W \times C_w = (1-p) \times I \times C_{rain} - D \times C_w - T^{abs} \\ \frac{d}{dt}LAI \times C_{lai} = T^{abs} \end{cases} \quad (11a-b)$$

Where C_{rain} ($mol l^{-1}$) is the concentration of dissolved or particulate pollutants in rain water (constant), C_w ($mol l^{-1}$ of water solution) denotes the concentration of available pollutants in the aqueous film, C_{lai} ($mol m^{-2}$ of foliage) quantifies the surface concentration in the unavailable pool and T^{abs} ($Bq m^{-2}$ of ground surface h^{-1}) is the net absorption flux. The drainage term evacuates materials at a rate that is proportional to their concentration. Although “absorption” of inert particles is likely to involve other elementary mechanisms than those for dissolved ions, we propose to adopt the same conceptual and mathematical approach, but with mass being expressed in kilogrammes rather than moles here. Although volatilization could be easily added to the mass balance, it was not accounted for because the substances used in the reviewed experiments, including Iodine, were reported as non-volatile forms.

Two different mathematical parameterisations were adopted for the absorption term.

1- *Equilibrium reversible approach* (referred hereafter as to EM). In this first approach, we assume that absorption is a reversible and instantaneous process

whose characteristic time is much smaller than rainfall duration. We thus consider that concentrations in the two pools are in equilibrium during the time for which the plant is exposed:

$$C_{lai} = CR \times C_w \quad (12)$$

$$\frac{d}{dt}(W + LAI \times CR) \times C_w = (1 - p) \times I \times C_{rain} - D \times C_w \quad (13)$$

2- *Kinetic irreversible approach* (referred hereafter as to KM). In this second approach, we hypothesize that absorption is a slow kinetic process, with a characteristic time that may exceed rainfall duration. This process is also assumed irreversible, where we assume that, once attached to the foliage surfaces (or incorporated to the biomass), the pollutant cannot be remobilized and drained off the plant anymore, at least until the end of the contamination event. The absorption rate is assumed to follow a first order kinetics:

$$T^{abs} = \frac{J}{K} \times C_w \times LAI \quad (14)$$

Where the characteristic velocity, J/K (mm h^{-1}), depends on the physico-chemical form of the deposited substance. We notice that this formulation is equivalent to a Michaelis-Menten model of enzymatic kinetics, in the case where the substrate concentration is much smaller than the Michaelis constant, i.e. $C_w \ll K$. This holds for trace elements. This approach implies that the proportion of the intercepted fraction increases with increasing precipitation duration as absorption increases with the time elapsed after rainfall start. Such behaviour is consistent with the observations by Kinnersley et al. (1997) who demonstrated that accumulation of Cesium at plant surfaces increased with the time of exposure, long after water storage capacity has been reached.

Integrating Eq (11a-b), or Eq. (12) and (12) if equilibrium is assumed, enables to predict the time evolution of $C_w(t)$ and $C_{lai}(t)$ from $t=0$ to $t=T$. Time integration until the sampling time, after rainfall has ceased, is not required because drainage stops and pollutant inventory in plant remains unchanged.

Interception fraction. Based on the estimations of the hydrological quantities and pollutant concentrations at time $t=T$, the interception fractions are then calculated for any substance as the ratio of the amount in plant (including evaporated water during precipitation for water interception) to the total amount deposited by rain, that is:

$$f \approx \begin{cases} 1 - p & \text{if } T \leq T_S \\ \frac{(1-p)}{T \times \lambda} \times [1 - (1 - \lambda \times T_S) \times \exp(-\lambda \times (T - T_S))] & \text{if } T > T_S \end{cases} \quad (19)$$

Where:

Where CR (l m^{-2} or mm) is the concentration ratio at equilibrium. This parameter depends on the physico-chemical form of the deposited substance. Combining Eq. (11a-b) and (12) leads to the following ordinary differential equation:

$$f_w = \frac{W(T) + \int_0^T E(t)dt}{H} \quad (15)$$

$$f = \frac{(W \times C_w + LAI \times C_{lai})(T)}{H \times C_{rain}} \quad (16)$$

The mass interception fractions are deduced from the interception fractions by dividing them by the biomass.

It is possible to derive analytical solutions to the previous differential equations, under the assumption that $E_0 \ll (1-p) \times I$. As this condition covers almost all the experiments reviewed in this work, apart from potentially very low precipitation intensities ($\leq 2 \text{ mm h}^{-1}$) in very strong evaporative conditions ($\sim 0.5 \text{ mm h}^{-1}$), these analytical solutions will be preferred to the numerical solutions in the following study.

For water, we can demonstrate that the interception fraction obeys to the following expression:

$$f_w \approx \begin{cases} 1 - p & \text{if } T \leq T_S \\ \frac{LAI \times L}{H} + \frac{E_0}{I} & \text{if } T > T_S \end{cases} \quad (17)$$

This equation indicates that f_w increases as rainfall intensity decreases or rainfall duration increases. This expression is very similar to the modified version of the Horton's model (Merriam, 1961):

$$f_w = \frac{S}{H} + \frac{k \times e}{I} \quad (18)$$

Where e (mm h^{-1}) is the evaporation rate per unit surface of foliage and k (unitless) is a surface factor defined as the ratio of the evaporation surface to the projected area, i.e. $k \approx LAI$.

The interception fraction as predicted by EM for pollutants expresses as:

$$\lambda = \frac{D_S}{LAI \times (L + CR)} \quad (20)$$

And T_S and D_S are given by Eq. (9) and (10), respectively. The dependence upon rainfall intensity or rainfall duration is relatively complex and will be further examined in the next section. In non-evaporating conditions (i.e. $E_0=0$), we can check that f does no longer depend upon I or T .

$$f \approx \begin{cases} 1-p & \text{if } T \leq T_S \\ \frac{(1-p)}{T \times \lambda} \times \left[\frac{D_S}{LAI \times L} \times T_S + \frac{\alpha}{LAI \times L} \times T + \left(\frac{D_S}{\alpha + D_S} - \frac{D_S}{\alpha + (1-p) \times I} \right) \times (1 - \exp(-\lambda \times (T - T_S))) \right] & \text{if } T > T_S \end{cases} \quad (21)$$

Where:

$$\begin{aligned} \alpha &= \frac{J}{K} \times LAI \\ \lambda &= \frac{\alpha + D_S}{LAI \times L} \end{aligned} \quad (22)$$

$$f \approx \begin{cases} 1-p & \text{if } T \leq T_S \\ \frac{(1-p)}{T \times \lambda} \times \left[1 - \frac{E_0}{(1-p) \times I} \times \exp(-\lambda \times (T - T_S)) \right] & \text{if } T > T_S \end{cases}$$

Where:

$$\lambda = \frac{(1-p) \times I - E_0}{LAI \times L} \quad (23)$$

In non-evaporating conditions, we further notice that:

$$f \approx LAI \times L / H.$$

Model parameters. To run EM and KM models when biomass (B) and rainfall characteristics (H , I and T) are known, it is necessary to provide estimates of the following parameters: the potential evaporation rate (E_0), the free throughfall coefficient (p), the specific foliage area (SLA), the equivalent film thickness (L) and the substance-specific parameter, i.e. the concentration ratio at equilibrium (CR) or the absorption velocity (J/K). Unfortunately, E_0 could not be estimated for any of the experiments concerned in this study because the meteorological conditions prevailing at time of deposition were not reported. In-storm evaporation rates in out-door conditions and temperate climates typically range from 0 to 0.45 mm h⁻¹ for grass vegetation. As climatic conditions may have been very specific for

No simple analytical solution can be derived with KM, unless we neglect evaporation with respect to rainfall intensity during the wetting phase (i.e. $t < T_S$). If evaporation is neglected then $W(t) = (1-p) \times I \times t$ and Eq. (11a-b) can be integrated analytically. The following expression provides in this case an accurate estimation of the interception fraction (i.e. with a relative error smaller than 5%) except for very low precipitation intensity (≤ 2 mm h⁻¹):

The functional dependence of f upon rainfall intensity and duration is even more complex than previously.

For non-reactive pollutants (i.e. $CR=0$ mm h⁻¹ or $J/K=0$ mm h⁻¹), the two approaches are identical as they both reduce to the following expression:

especially in-door experiments, we cannot exclude that relatively high evaporation rates occurred. We thus decided to consider in this study two extreme situations, by setting its value to either 0 or 0.6 mm h⁻¹. The free throughfall coefficient was set to one because the selected experiments were not dealing with sparse vegetation.

The three other parameters were estimated through Bayesian inference (see below). Bayesian inference of EM or KM parameters was carried out using the analytical solutions: Eq. (17) for water, Eq. (19) to (22) for other substances. To enable a rigorous comparison between EM and KM models, the plant parameters, L and SLA , were inferred while assuming a common value for both models. For each model, four alternative approaches were tested which are listed in Table 4. They differ from each other in the way they account for the dependence of CR or J/K upon substances: for some of them, C_s and/or particles were estimated separately from divalent cations, while for others they were aggregated together.

2.4. Model calibration and testing

Inference of models parameters. The estimation of MLR, EM and KM models parameters was implemented by adopting the Bayesian approach in order to capture the uncertainty on their values and to

characterize the variability observed among the multiple substances of concern in this study. Under the Bayesian paradigm, which is fundamentally based on the Bayes rule (Parent and Bernier, 2007) and the theory of conditional probabilities, the characterization of both uncertainty and variability about the parameters estimates is achieved by combining information drawn from both prior knowledge (e.g. historical related data and/or expert advice) and observed information (that are the stochastic realizations of the modeled process) on the parameters of interest (Carlin et al. 2006). This approach provides according to Carlin et al. (2006) a cohesive framework under which a new state-of-knowledge characterized by a probability distribution called posterior is reached (Parent and Bernier, 2007).

One important aspect within the Bayesian framework relies upon the characterization of the prior knowledge, generally by using probability distribution functions. Prior distributions, which reflect the modeler's uncertainty about the unknown parameters and by extrapolation the uncertainty about the considered processes, enable the incorporation of external additional information such as expert judgment or relevant literature data (Sy *et al.* 2015). In this study, the prior distributions allocated to the unknown parameters were defined in regard to mathematical reasons and knowledge about the values and the range of plausible values of the parameters of concern. The strategy was to assign prior distributions with low intrinsic information to parameters for which accurate information is lacking.

The prior distributions assigned to the MLR and EM/KM model parameters are briefly discussed in the Supplementary Information (section S.2).

The posterior distributions of the model parameters were computed using Markov Chain Monte Carlo (MCMC) sampling algorithm as implemented in JAGS 3.3.0 (Plummer, 2003). For each parameter of interest, three independent chains starting from randomly defined positions were computed considering 10^5 runs after a burn-in period fixed at $5 \cdot 10^4$. To reduce within-chain autocorrelation, a selection step was defined at 20 *i.e.* one run is kept every 20 iterations. The convergence of the chains was assessed by graphical visualization and by computing the measure of Gelman-Rubin which is based on the comparison of the between-chains variance to the

within-chain variance (Brooks and Gelman, 1998). For each applied model, empirical descriptive statistics (posterior mean, median, 95% credible interval and standard deviation) were calculated on the basis of the obtained posterior distribution samples of size $N = 1.5 \cdot 10^4$ ($5 \cdot 10^3 \times 3$ chains).

Quantification of models performances. The competing models (MLR, EM and KM) were compared by quantifying their respective predictive performances using the following procedure. Suppose N is the total number of mass interception fraction data for which the corresponding values of the aboveground biomass, rainfall height and rainfall intensity are known (*i.e.* calibration data subset, $N = 363$). The predictive performances of each model were evaluated by implementing an iterative procedure, called leave-one-out cross validation (Hastie et al. 2009). At each s^{th} iteration, this approach consists in:

- Splitting the dataset in a learning subset $\left(\left(f_{B(s)}^*, B_{(s)}^*, H_{(s)}^*, I_{(s)}^*\right)\right)$ consisting of $N-1$ data and a validation subset $\left(\left(f_{B(s)}^*, B_{(s)}^*, H_{(s)}^*, I_{(s)}^*\right)\right)$ consisting of the remaining data,
- Fitting the model on the learning subset $\left(\left(f_{B(s)}^*, B_{(s)}^*, H_{(s)}^*, I_{(s)}^*\right)\right)$ and
- Predicting posterior distributions of the mass interception fraction $\widehat{f_{B(s)}^*}$ for the calibrated model, with respect to the aboveground biomass, the rainfall height and intensity values corresponding to the validation subset. A number of $2 \cdot 10^4$ runs was defined with a thin equal to 4. Thus, it resulted in a sample of $5 \cdot 10^3$ posterior values of interception fraction.

On the basis of the posterior distributions drawn from the LOOCV (leave-one-out cross validation) procedure, the posterior predictive loss criterion (PPLC) value associated to each model was computed. PPLC is a quantitative criterion initially proposed by Gelfand and Ghosh (1998) which measures the gap between the observations and the corresponding predictions. PPLC is defined as:

$$PPLC(M, k) = \sum_{s=1}^N Var\left(\widehat{f_{B(s)}^*}\right) + \frac{k}{k+1} \sum_{s=1}^N \left(\mathbb{E}\left(\widehat{f_{B(s)}^*}\right) - f_{B(s)}^*\right)^2 \quad (24)$$

Where $\mathbb{E}\left(\widehat{f_{B(s)}^*}\right)$ and $Var\left(\widehat{f_{B(s)}^*}\right)$ are the estimated mean and variance of the predictive distribution of a new mass interception fraction $\left(\widehat{f_{B(s)}^*}\right)$ under the model M and given a learning subset $f_{B(s)}^*$ that includes the observed values. Such predictive mean and variance are estimated

from posterior predictive samples $\left(\widehat{f_{B(s)}^*}\right)$ drawn using the learning subset. The parameter k ($=1, 10, 100, \infty$) is a weighting coefficient which calibrate the relative loss due to the gap between the expected wet mass interception fraction predicted value and the observed wet mass interception fraction value and the gap due to the variance of the predicted posterior samples $\left(\widehat{f_{B(s)}^*}\right)$.

The best model will show the smallest PPLC value. The applied models were compared with respect to the PPLC values.

2.5. Model validation

To further check and compare the performance of the competing models, we evaluated their predictive performances in case where either biomass or rainfall intensity was unknown. This kind of exercise is of practical interest as the precise environmental conditions prevailing at time of deposition are rarely known, especially in case of unexpected situations. The models predictions were tested against the validation data subset which was not used in the calibration step. The calculations were accomplished by performing Monte Carlo simulations with parameter uncertainties estimated from the Bayesian inference process.

To estimate the mass interception fraction with the MLR approach, we used either HI or HB models, depending on the missing parameter (i.e. B and I , respectively). For the EM and KM models where both B and I are mandatory parameters, we assumed that the missing parameter was randomly distributed in accordance with the empirical density distributions established from the calibration data subset. The justification for introducing the empirical density functions is that HI and HB model parameters, themselves, rely implicitly on the same distributions. Comparison of the respective model performances is thus unbiased. In the sampling process, we further assumed that I was uncorrelated with H , because correlation between calibration data was very low (i.e. correlation coefficient equal to 0.11).

3. Results and discussion

3.1. Exploratory analysis of selected data

The decrease of the mass interception fraction, f_B , with increasing rainfall height, H , is marked for all kinds of substances and all kinds of experiments (see Figure 1), apart from those by Hoffman et al. (1989) where an intermittent rain was applied. In the latter experiments, the observed values were nearly constant regardless of how much rain was applied: the cumulative amount was increased from 15 to 75 mm at constant rainfall intensity with 20-30 min between applications. Such effect was attributed by the authors to the process of cumulative evaporation of water on the plant surface. As pointed out by Pröhl and Hoffman (1996), this also suggests that once initial retention occurs, subsequent rain is rather ineffective in removing previously deposited materials. For single deposition experiments, the fastest decrease is observed for Iodine where f_B falls down to values as small as $0.1 \text{ m}^2 \text{ kg}^{-1} \text{ dw}$ as a quasi-inverse function of rain amount. The experiments by Angeletti and Levi's (1977a) clearly demonstrate that the decrease for water is slower than for Iodine. This might be due, at least for low-to-moderate rainfall heights (i.e. 1 to 4 mm), to the

effect of evaporation from saturated leaf surfaces during the 1-hour daytime deposit: as suggested by Eq. (17), evaporation increases the total amount of water intercepted by plant. Interception of cations and small particles is even less sensitive to the rainfall height. As already discussed in the literature, the enhancement of cations retention by plant is due to their active absorption during the contact of the contaminated water with the cuticle which acts as a cation-exchanger (Keppel, 1966; Franke, 1967; Schönherr, 1977; Schaefer and Reiners, 1990). In separate experiments with gummed paper (not plant) and in the absence of electrical forces, Hoffman et al. (1989, 1992b) demonstrated that retention of Beryllium (divalent cation) was comparable to that of Iodine. They also observed that ions were much less efficiently retained by paper than insoluble particles. They suggested that interception and retention of particles were controlled by their rapid settling out from the rain droplets and their attachment to the gummed paper. Similar mechanisms are likely to play a role in the interception by plant surfaces, but this has still to be elucidated.

The dependence of the mass interception fraction upon rainfall intensity, I , and the way intensity can influence interception was much less discussed in the literature. Figure 1 seems to indicate, however, that f_B might be also affected by rainfall intensity, at least for cations, as it decreases by almost one order of magnitude when intensity is increased from 1 to 100 mm h^{-1} . To ascertain the dependence of f_B upon I , it is necessary to analyze observations in the same range of rainfall height as this factor strongly influences interception. In Figures 2a-c, observations have been classified into 3 different groups depending on the rainfall amount: low range (0.4-3 mm), medium range (3-12 mm) and high range (12-75 mm). Regression slopes (i.e. $\log(f_B) \sim \log(I)$) have been computed for each rainfall range and each kind of substance (except water). These calculations confirm that sensitivity to I is more pronounced for cations than anions. When intensity is increased from 1 to 10 (resp. 100 mm h^{-1}), f_B decreases by a factor of 2 (resp. 5) for cations and by about 30% (resp. 50%) for Iodine. The regression slopes also suggest that the dependence might be stronger for low rainfall heights. This might explain why values measured by Bengtsson et al. (2014) at a very high rainfall intensity (120 mm h^{-1}) are much smaller than those obtained by Angeletti et al. (1977a,b) and Kinnersley et al. (1997) for similar rainfall heights (about 1 mm). The situation is less conclusive for polystyrene particles because data are missing in the low and moderate intensity ranges.

The dependence of the mass interception fraction upon dry biomass, B , is displayed in Figure 2d-f for the 9 groups identified above. Although experiments cover a relatively narrow range of biomasses, the negative regression slopes seem to indicate that the mass interception fraction might slightly decrease with B . This implies that the normalization of the interception

fraction, f , by the dry biomass does reduce the dependence upon biomass but does not suppress it. Pröhl et al. (2009) suggested that the leaf area index, or even the fresh biomass, might be more relevant for quantifying interception and explaining the observed variability. Unfortunately, the data set does not enable to introduce these factors into the analysis.

This preliminary exploratory analysis raises two fundamental questions: (i) to what extent does the dependence of mass interception fraction upon rainfall intensity and dry biomass need to be considered to accurately predict interception, and (ii) to what extent Cesium and micronic particles behave differently from divalent cations? These questions are answered and discussed in the following section.

3.2. Bayesian inference of the MLR models parameters

Inter-comparison of alternative approaches.

The posterior predictive loss criterion (PPLC) values associated to each alternative MLR model are reported in Table 2. The calculations clearly demonstrate that the model which most accurately reproduces the observations is the most elaborate one, HBI, where all dependences are accounted for. Four other alternative models still guarantee satisfactory predictive performances: HBI-Cs, HBI-Poly, HBI-Biomass and HI. Combining Cesium and particles together, like in HBI-PolyCs, significantly decreases the quality of the model. This tends to demonstrate that monovalent cations and particles must be definitely treated as two different kinds of substance. Another interesting result is that neglecting the dependence of the mass interception fraction upon dry biomass does only slightly degrade the model performance; this implies that normalizing f_B by M remains a reasonable assumption, at least in the range of biomasses investigated. Unfortunately, the dataset does not enable to explore the behavior of f_B in the extreme biomass ranges, because data from Bengtsson et al. (2014) which cover a very large range of biomasses were obtained in extreme rainfall conditions (i.e. very low H and very high I). On the contrary, neglecting the dependence of the mass interception fraction upon rainfall intensity, through the use of either HB or H approaches, drastically affects the quality of the model. This clearly demonstrates the role of rainfall intensity in the interception process, at least for substances that interact with the leaf surfaces like cations. Although less accurate, HB and H have some practical importance for the assessment of the interception fraction in case where rainfall intensity is unknown. An application which aims at demonstrating the interest of adopting such models when rainfall intensity and/or biomass are missing is provided in section 3.4.

Model calibration and testing. The inferred distributions (i.e. median and 95% credible interval) of both the regression coefficients (i.e. f_{B0} , α , β and γ) and the residual error (i.e. Σ) are reported in Table 3 for HBI,

HI, HB and H models. These statistics are provided for each group of substances. To help interpretation of the tabulated values, median predictions of f_B by the HBI model are displayed in Figure 3 for varying rainfall heights (i.e. 1, 3, 10, 30 and 100 mm), varying rainfall intensities (i.e. from 1, 3, 10, 30 and 100 mm h⁻¹) and a representative biomass of 0.125 kg dw m⁻² (i.e. median of the experimental values). These ranges of rainfall amount and intensity correspond to rainfall durations, $T=H/I$, varying from 36 seconds to 100 hours.

The calculated values of the regression coefficients indicate that the contribution of rainfall height to the variability of the mass interception fraction (as given by α value) predominates over the two other factors for Iodine, while for divalent cations, the most influential factor is rainfall intensity (as given by γ value). The situation is less conclusive for Cesium and inert particles, as α and γ have quite similar values. Dependence of f_B upon H is, in a decreasing order: I > Cs > particles > Be/Sr, while dependence upon I follows a nearly inverse order: Cs > Be/Sr > particles > I. Predictions of the HBI model as depicted in Figure 3 show that, for light rainfalls (i.e. 3 mm in 1 hour), f_B decreases as follows: Be/Sr (5.21 m² kg⁻¹) > particles (4.05 m² kg⁻¹) > Cs (3.47 m² kg⁻¹) > I (1.19 m² kg⁻¹). For higher precipitation amounts and intensities (i.e. 30 mm in 1 hour), the values are even more contrasted: particles (1.45 m² kg⁻¹) > Be/Sr (1.33 m² kg⁻¹) > Cs (0.2 m² kg⁻¹) > I (0.11 m² kg⁻¹). Dependence of the mass interception fraction upon the standing biomass (as given by β value for HBI or HB model) is contrasted among the substances. For Iodine and particles, the calculations indicate that f_B strongly decreases with increasing B , while dependence is unclear for cations. Calculations with the HBI model and coefficients given in Table 3 show that, when B is increased from 0.05 to 0.35 kg m⁻² (i.e. 5th and 95 percentiles of the observations distribution), f_B is decreased by a factor of about 2 for Iodine and particles, while it does not change by more than 30% for cations, whatever rainfall intensity and rainfall height are (3 or 30 mm h⁻¹ during one hour). The observations reported in Figure 2d-f further suggest that sensitivity of the mass interception fraction upon the biomass might be dependent on the rainfall height, because the regression slopes reported in this figure slightly vary with rainfall height. This dependence results from interactions between plant and rainfall characteristics which are not accounted for in this empirical model, as no interaction term $-\log(B)\times\log(H)$ – was included in Eq. (3). However, this point must be considered with caution because the number of data is probably not sufficient to estimate interaction terms with confidence.

Regarding the standard deviation of the residual variability, we notice that, for water and Iodine, Σ remains relatively small compared to other substances and relatively constant among the four models. This implies that interception of water and Iodine by herbaceous plants is quite accurately reproduced by these

empirical models, even when using the least sophisticated H model. The situation is different for cations because the residual variability here increases when rainfall intensity is no more considered (like in the HB or H approaches). This means that for such substances rainfall intensity is an important controlling factor which explains some part of the variability.

The predicted mass interception fraction (i.e. median and 95% credible interval) is plotted against observed values in Figure 4, for each of the four alternative models. These results confirm that observations are satisfactorily reproduced by the HBI and HI approaches for all substances, as 89% and 85% of the predicted median values, respectively, are within a factor of 2 compared to observed values. This proportion falls down to, respectively, 77% and 73% for the HB and H models, due to poorer performances for cations, as only 57% and 53% of the related observations (i.e. 152) are predicted within a factor of 2.

3.3. Bayesian inference of the process-based models parameters

In this section, it is proposed to apply the same methodology as previously, to the process-based modelling approaches, EM and KM.

Inter-comparison of alternative approaches.

The posterior predictive loss criterion (PPLC) values associated to each of the 8 alternatives are reported in Table 2, for both the evaporating and non-evaporating situations ($E_0=0 \text{ mm h}^{-1}$ and $E_0=0.6 \text{ mm h}^{-1}$). These calculations demonstrate that the overall performance of the kinetic approach is by far better than that of the equilibrium approach, whatever the alternative is. This means that, for some of the substances investigated, fixation to the foliage surfaces increases with the duration of the exposition. Cross-validation also shows that the highest PPLC value is, here again, obtained when the substance-dependent parameters are estimated separately, like in the EM and KM models, while the lowest one is obtained when aggregating Cesium, divalent cations and particles, like in the EM-PolyCs and KM-PolyCs models. In agreement with MLR study, present analysis confirms that particles must be treated separately from ions and that the electrical valence is likely to play a role for dissolved cations. The calculations also show that accounting for in-storm evaporation slightly improves the quality of the predictions, at least for water (as E_0 directly influences the estimation of the interception fraction). This point will be further analyzed in the next section.

Although not drastically different, the cross-validation performance of the EM model is less satisfactory than that of the regression model (i.e. 37 against 21 for MLR). This may be explained by the fact that the estimation of the interception fraction with the EM/KM approach makes use of fewer degrees of freedom (DF) than in the detailed MLR approach: 7 DF for EM/KM models (2 plant parameters + 1 parameter

per substance, evaporation rate being the parameter specific for water) against 20 DF in HBI model (4 parameters per substance) or even 10 DF in the least sophisticated H model (2 parameters per substance).

Model calibration and testing. The posterior distributions (i.e. median and 95% credible interval) of the EM/KM models parameters and residual errors (i.e. Σ) are reported in Table 5 and Table 6 for the non-evaporating and evaporating conditions, respectively. To help interpretation of the tabulated values, predicted values of the median mass interception fraction are displayed in Figure 5 (water with the Horton's like model), Figure 6 (Iodine), Figure 7 (cations and particles with EM) and Figure 8 (cations and particles with KM). As mentioned in section 2.3, predictions of the EM and KM models for non-interacting substances, like Iodine, are rigorously identical. In these figures, f_B is plotted against rainfall height and rainfall intensity for a representative biomass of $0.125 \text{ kg dw m}^{-2}$.

As given in Table 5 and Table 6, the posterior distributions of the two plant parameters, SLA and L , are much narrower than the prior distribution (as detailed in the Supplementary Information, section S.2). This clearly indicates that the available calibration data strongly contribute in updating the prior beliefs on these two parameters, and that plant structural parameters strongly influence the interception process. The two distributions are very asymmetric, with a positive skew for the foliage area and a negative one for the leaf storage. We also notice that both posterior distributions are shifted towards the lower part of the prior distribution. This is particularly acute for the specific foliage area as the 95% credible interval, which ranges from 12 to about $21 \text{ m}^2 \text{ kg}^{-1} \text{ dw}$, is below the values usually reported for grasses and other herbaceous plants in the literature (see Supplementary Information, section S.2). The reason could be that most of the literature values actually quantify the area of the (laminar) leaves but do not include the contribution of other organs, like stems, although they participate to the interception of contaminated droplets by herbaceous plants. Including stem organs into the reasoning (like in our analysis) is likely to decrease both: (i) the total "plant" area, as $SLA_{stem} < SLA_{leaf}$, and (ii) the plant water storage capacity, as water droplets intercepted by stems may be more easily evacuated than from leaves due to their preferential vertical orientation and smooth surfaces. This might explain why output ranges are shifted to lower values than reported in the literature.

The mass interception fraction of water predicted by the Horton's like model in the absence of in-storm evaporation decreases as $\sim 1/H$ but remains independent of rainfall intensity (see Figure 5). When evaporation is considered, the predicted value is increased by E_0/I , according to Eq (5). This implies that, in such a case, f_B decreases with rainfall intensity but increases with rainfall duration. The contribution of evaporation

becomes significant and even predominant over direct interception by plant canopy, for high precipitation amounts and low intensities (or long duration). In-storm evaporation is a sensitive parameter for predicting water interception in long-lasting deposition events or strongly evaporative conditions. Strong evaporation might have occurred in the experiments by Angeletti et al. (1977a) because measurements were made in the daytime and greenhouse conditions. Model predictions with evaporation better reproduce observations, as demonstrated by the residual standard deviation which is decreased by more than a factor of 3 when considering in-storm evaporation (i.e. 0.4 against 1.37).

The results obtained for iodine do not significantly differ from those provided by the MLR approach in Figure 3b. The process-based models predict that f_B decreases roughly as $1/H$ and is slightly increased in evaporative conditions and low rainfall intensities (i.e. smaller than 3 mm h^{-1} , typically). For such non-interacting substances, model predictions are accurate and the residual standard deviation is very low which means that variability is satisfactorily explained by such process-based models.

Bayesian calculations for the other substances confirms that particles interact more with the leaf surfaces than divalent cations which themselves are more reactive than monovalent cations. The affinity of the substances, as quantified by either CR or J/K parameters, is in an increasing order: $\text{Cs} < \text{Be/Sr} < \text{particles}$. These findings are consistent with the observations related in the literature as reviewed in the introduction. The calculations further indicate that evaporation is not a sensitive factor for such substances, as posterior distributions and residual standard deviation are very similar between the two evaporation scenarios. As is clearly illustrated in Figure 7 and Figure 8, EM and KM provide drastically different predictions. The main differences reside in the fact that: EM predicts a strong decrease of the mass interception fraction with rainfall height and a weak dependence upon rainfall intensity while KM predictions are much more sensitive to rainfall intensity. With the EM model, f_B is roughly proportional to $1/H$ for moderate to high precipitation height (i.e. greater than about $CR \times LAI$, according to Eq. (19)), while with KM model, the decrease is much less pronounced and even vanishes for high H and moderate I values. In the latter approach, sensitivity to rainfall intensity is greater for less reactive substances: f_B for Cs decreases by nearly 2 orders of magnitude when I is increased from 1 to 200 mm h^{-1} to be compared with the value obtained for particles which decreases by only one order of magnitude. It is easy to demonstrate from Eq. (21) that f_B starts to decrease when I becomes greater than the absorption rate, $\alpha = J/K \times LAI$, i.e. when rainfall duration becomes smaller than the characteristic time required for absorption. When exposure time is too short, the substance cannot be efficiently absorbed by the foliage surfaces and is thus evacuated by water drainage. The

competition between drainage and absorption mechanisms is accounted for in both models, but in the KM approach, this balance depends on their respective characteristic rates (i.e. Ds versus α) while in the EM approach, it depends rather on their total capacity to mobilize (or evacuate) deposited substances (i.e. L versus CR). Thus, in the EM approach, interception by plant does not or only weakly depend on rainfall duration and intensity. As can be noticed in Figure 7 and Figure 8, both modelling approaches fail to reproduce observations by Bengtsson et al. (2014) for Cs and Sr, which are over-predicted by a factor of four. We might challenge the validity of the process-based models in case of very short durations of the contaminating precipitation, like in the above-mentioned experiments (1 mm in 36 seconds). We notice however that the values obtained by Bengtsson and co-authors for Strontium ($0.7 \pm 0.25 \text{ m}^2 \text{ kg}^{-1} \text{ dw}$) are much smaller than those reported by Hoffman et al. (1989) for Beryllium ($2 \pm 0.64 \text{ m}^2 \text{ kg}^{-1} \text{ dw}$), similar rainfall conditions (2.5 mm in 75 seconds) and a same biomass range (from 0.06 to 0.22 kg m^{-2}). The reason for this discrepancy thus remains unclear. Figure 7 and Figure 8 finally suggest that the KM predictions for reactive substances might be more accurate than EM's. This is more or less confirmed by the smaller residual variability of KM predictions, especially for divalent cations.

Dependence of the mass interception fraction upon biomass is displayed in Figure 9 for reactive substances only, because the process-based models predict a nearly constant value for water and Iodine. Both models predict that f_B decreases with B over the whole range, but more specifically for biomasses greater than about 0.1 kg dw m^{-2} and light precipitation (i.e. either low heights or intensities depending on the model considered). Sensitivity to biomass increases with increasing affinity of the substances for the foliage surfaces: $\text{Cesium} < \text{divalent cations} < \text{particles}$. This means that, for high biomasses and weak precipitation, the interception fraction for reactive substances, f , increases less rapidly than the plant biomass, which in turn implies that the normalization with B is less relevant in such situations. For the EM model, the dependence of f_B upon B originates in the negative "exponential mathematical term" in Eq. (19), that is: $-(1 - \lambda \times T_S) \times \exp(-\lambda \times (T - T_S))$. This term vanishes for non-reactive substances but decreases with increasing biomass and/or increasing concentration ratio, CR . For the KM model, it can be shown that the dependence upon B resides mainly in the first multiplicative factor of Eq. (21), $1/(T \times \lambda)$, which is not strictly proportional to B (like in the EM model) but rather proportional to $B / (J/K \times B/I + cte)$. This explains why f_B decreases with B for high values of J/K , high biomasses and low intensities. From these figures, we cannot identify which model better fits data. However, it is worth noting that KM model predicts a nearly constant value of f_B for $H=1 \text{ mm}$ and $I=100 \text{ mm h}^{-1}$ (black curve), this result being consistent with data produced by Bengtsson and co-

authors (triangles) for Cesium and Strontium (although values are overestimated by the models).

The predicted distributions of the mass interception fraction (median and 95% credible interval) are plotted against observed values in Figure 10, Figure 11 and Figure 12, for respectively water, Iodine and other substances. Results are given for the evaporating scenario. All substances taken together, the 95% credible interval of the predicted f_B 's encompasses 94% of the observations for both models. Interception of water and Iodine is very satisfactorily reproduced by the process-based models. The variability of f_B for reactive substances is quite accurately predicted by the process-based approaches, except data from Bengtsson et al. (2014). However, the performance of the KM model for this kind of species is slightly better than that of EM, as respectively 82% and 73% of the predicted median values are within a factor of 2 compared to the observed values. For cations, this proportion falls down to respectively 74% and 57% which definitely indicates that KM is by far more appropriate than EM for cationic species, although less accurate than the HBI approach for which 83% of data are within a factor of 2. We may notice that the credible interval of the KM predicted values for divalent cations are drastically reduced with respect to that of the EM's. The two approaches remain satisfactory for particles, as 85% and 78% of the predicted values falls within a factor of 2 for KM and EM, respectively. We thus conclude that the behaviour of cations, and to some extent micronic particles, is better represented by a kinetic absorption formulation than an instantaneous and reversible formulation. These findings are consistent with the results obtained by Kinnersley et al. (1997) who investigated the kinetics of Cesium attachment to (bean) leaf surfaces from a contaminated solution. The aim of their leaf dipping experiments, in which freshly sampled bean plant leaves were immersed in a demineralised water then rinsed (to remove available fraction of Cesium), was to test the dependence of the absorption process on exposure time and the concentration of Cesium in solution. They demonstrated that the level of Cesium not removed by rinsing, therefore presumably fixed or internalised by the leaf, increased with time of exposure and Cesium concentration without any saturation could be reached, even at the highest concentrations (i.e. 10 mM or $4 \cdot 10^9$ Bq/l) and exposure durations (i.e. 40 min). These experiments served to demonstrate that active chemical accumulation of Cesium occurred at plant surfaces and would account for the increase in surface contamination long after maximum water storage capacity has been reached. Although their leaf-dipping experiments were not representative of realistic deposition conditions, it is possible to estimate effective kinetic rates, J/K , from their experiments in which the total amount of Cesium absorbed onto the leaf surface was proportional to the initial solution concentration (i.e. experiments with 1 and 10mM of Cesium). The estimated values ranged from

about 0.13 mm/h to 1.48 mm/h, depending on the exposure time and solution concentration. These values are consistent with the 95% credible interval of J/K parameter inferred for Cs in our Bayesian calculations, that is: 0.06 to 2.73 mm/h (cf. Table 6).

3.4. Models predictability

The predicted posterior distribution (median and 95% confidence interval) of the mass interception fraction by the HB/HI, EM and KM models (with $E_0=0.6$ mm h⁻¹) are compared to the 77 observations for which either biomass or rainfall intensity were unknown in Figure 13. The observations include divalent and trivalent cations that were not present in the calibration dataset, that is: Barium (as Ba²⁺), Lead (as Pb²⁺), Cadmium (as Cd²⁺) and Chromium (as Cr³⁺). The latter was treated as a divalent ion as models parameters could not be inferred for valence greater than 2. The 12 observations for which rainfall intensity was not available were predicted with the HB model (Chadwick and Chamberlain, 1970; Jacob et al., 1993) while the remaining 65 ones with missing values of biomasses were predicted with the HI model (Angeletti and Levi, 1977b; Hoffman et al., 1989; Hoffman et al., 1995). Among these 65 data 36 observations correspond to intermittent rainfall conditions (Angeletti and Levi, 1977b; Hoffman et al., 1989), for which an "equivalent" rainfall amount, duration or intensity was approximated as explained in section 2.1

As shown in Figure 13a and Figure 13c, the HB/HI and KM models agree satisfactorily with observations for all substances and experiments, apart from data by Hoffman et al. (1989) in intermittent rainfall conditions. The two models predict a median value which is within a factor of 2 for 69 and 68% of the observations, respectively. And the 95% credible interval of the predicted f_B value encompasses 64% of the observations for both of them. The corresponding indicators for the EM model falls down to 52% and 49%, respectively, due to poorer performances for divalent cations, especially Beryllium and Strontium (see data from Hoffman et al., 1989; Angeletti et al, 1977b and Chadwick and Chamberlain, 1970).

Under intermittent rainfall conditions, some of the observed values by Hoffman et al. (1989) for Iodine and Beryllium are under-predicted by a factor of 2 to 6. For Iodine, it can be shown that the discrepancy between predicted and observed values increases when the number of consecutive rainfall applications (of same intensity in Hoffman's experiments) increases from 2 to 5 (i.e. when cumulated H increases from 25 to 75 mm). This suggests that interception of Iodine in case of multiple consecutive events j may be significantly greater than that observed or predicted for a unique equivalent event, of equivalent amount $H=\sum_j H_j$ and duration $T=\sum_j T_j$. Accounting for inter-storm evaporation in the KM model does not improve the quality of the predictions. We thus hypothesize that some unknown mechanism must

promote fixation or even incorporation of Iodine into the biological tissues, during inter-storm periods, this fraction being no longer available for drainage in the following applications. Such mechanism would explain why, in Figure 1a, values of f_B for cumulated $H > 25$ mm remain of the same level than those measured just after the first 15 mm application. For Beryllium, one cluster of data is unpredicted by a factor of 3; they correspond to rainfall applications at very high rainfall intensity, 103.8 mm h⁻¹; while the other observations obtained for a 19.8 mm h⁻¹ intensity are predicted within a factor of 2. The predictions of the KM models could be greatly improved if we hypothesize that absorption of the dissolved cations by the wet foliage tissues still continues during inter-storm periods, while drainage stops. Provided that this mechanism remains active, which seems a reasonable assumption, the contribution of the inter-storm periods in the equivalent rainfall duration must be accounted for, which in turn increases the estimated value of the interception fraction as longer durations promotes fixation with respect to drainage. An even more realistic approach would consist in explicitly simulating the succession of wetting and drying periods based on the numerical solving of the differential equations described in section 2.3. This study might enable to test the adequacy of the process-based approach to such complex situations.

These calculations thus demonstrate that the predictions of the MLR and KM approaches are equally accurate in case of a single deposition event, and in a context where either rainfall intensity or herbaceous plant characteristics are unknown (i.e. plant species, dry or fresh biomass, foliage area index, etc.). These results are consistent with the conclusions driven from the cross-validation analysis based on calibration data in previous sections. However, we might argue that the process-based model which relies on only one substance-dependent parameter (against 3 parameters for HB/HI models) remains a more satisfactory approach.

4. Conclusion

In this study, we demonstrated that interception of wet-deposited pollutants by herbaceous plants could be quite accurately described and predicted by a mechanistic model, for a variety of chemical species, plant characteristics and rainfall conditions. This innovative model relies on a detailed parameterisation of the underlying hydrological, biological and chemical mechanisms that participate to interception. The main assumption is that absorption or fixation of dissolved cations to the foliage surfaces might be controlled by a kinetic rate that depends on their electrical valence. In contrast to the existing empirical formulations, the present approach provides key insights into our understanding of the interception process and the variability of the interception fraction, notably regarding its sensitivity to rainfall intensity (for cations). We derived an analytical formulation of the mass interception

fraction (see Eq. (21) and (22)) and estimated ranges of parameter values through Bayesian inference. This new formulation is a serious candidate to challenge existing empirical relationships in radiological and chemical risk assessment tools. The set of experimental data which served as the basis for evaluating the model, however, features some residual variability that remained unexplained. Part of this variability may originate in the simplification of some mathematical parameterisations, regarding in particular drainage and fixation mechanisms. It may also result from some interspecific variability because the different families of plants investigated in the experiments – grasses, clovers, weeds – were not discriminated in our analysis. Another important reason may be that the mass interception fraction was actually predicted from the above-ground dry biomass, although interception is controlled by rather the fresh biomass or the plant area index. Unfortunately, these factors could not be examined because they were rarely reported in the reviewed literature. These questions have still to be elucidated. Another perspective would be to extrapolate the present approach to other vegetation covers than herbaceous plants, including crops or even forest canopies.

Supporting Information

Supporting Information includes two sections dealing with the description of the experiments considered in this study (section S.1) and the determination of the prior distributions for the models parameters (section S.2).

References

- Angeletti L., Levi E., 1977a. Etude comparative des facteurs de transfert de l'eau, de l'iode et du strontium sur le ray-grass et le trèfle (A comparative study of transfer factors of water, iodine and strontium on ryegrass and clover). Rapport CEA-R-4860.
- Angeletti L., Levi E., 1977b. Etude du dépôt humide et de la rétention foliaire de l'iode et du strontium sur le ray-grass et le trèfle (A study of the wet deposit and foliar uptake of iodine and strontium on ryegrass and clover). Rapport CEA-R-4897.
- Angeletti L., 1980. La contamination des pâturages par l'iode-13. Rapport CEA-R-5056.
- Bengtsson S.B., Gårdenäs A.I., Eriksson J., Vinichuk M., Rosén K., 2014. Interception and retention of wet-deposited radioCesium and radiostrontium on a ley mixture of grass and clover. *Science of the Total Environment*, 497-498: 412-419.
- Brooks, S.P., Gelman, A., 1998. General methods for monitoring convergence of iterative simulations. *J. Comput. Graph. Stat.* 7 (4), 434-455.
- Calmon P., Gonze M.A. and Mourlon C., 2015. Modelling the early-phase redistribution of radioCesium fallouts in a evergreen coniferous forest after Chernobyl and Fukushima accidents, (just accepted). DOI: 10.1016/j.scitotenv.2015.04.084

- Carlin, B.P., Clark, J.S., Gelfand, A.E., 2006. Elements of hierarchical Bayesian inference. In: Clark, J.S., Gelfand, A.E., 2006. Hierarchical modelling for the environmental sciences, Oxford University Press Inc., New York.
- Chadwick R.C. and Chamberlain A.C., 1970. Field loss of radionuclides from grass. Atmospheric Environment Pergamon Press 1970. Vol. 4, pp. 51-56.
- Clark, O. R., 1940. Interception of rainfall by prairie grasses, weeds, and crop plants, *Ecol. Monographs*, 10(2): 243-277.
- Cruz P., Duru M., Therond O., Theau J.P., Ducourtieux C., Jouany C., Al Haj Khaled R. and Ansquer P., 2002. Une nouvelle approche pour caractériser les prairies naturelles et leur valeur d'usage. [A new approach to the characterization of natural grasslands and their use value]. *Fourrages*, 172: 335-354.
- Draxler R., Arnold D., Chino M., Galmarini S., Hort M., Jones A., Leadbetter S., Malo A., Maurer C., Rolph G., Saito K., Servranckx R., Shimbori T., Solazo E. and Wotawa G., 2014. World Meteorological Organization's model simulations of the radionuclide dispersion and deposition from the Fukushima Daiichi nuclear power plant accident, *Journal of Environmental Radioactivity*, DOI: 10.1016/j.jenvrad.2013.09.014.
- Ehrhardt J. and Weis A., 2000. RODOS: Decision Support System for Off-site Nuclear Emergency Management in Europe. European Commission, Brussels, Report EUR 19144.
- Franke W., 1967. Mechanisms of foliar penetration of solutions. *Plant Physiology* 18, 281-300.
- Gash J.H.C. and Morton A.J. (1978). An application of the Rutter model to the estimation of the interception loss from the Thetford forest. *Journal of Hydrology*, 38, 49-58.
- Gelfand, A.E., Ghosh, S.K., 1998. Model choice: a minimum posterior predictive loss approach. *Biometrika* 85 (1), 1-11.
- Gonze M.-A., Mourlon C., Calmon P., Manach E., Debayle C., Baccou J., 2015. Modelling the dynamics of ambient dose rates induced by radioCesium in the Fukushima terrestrial environment. *Journal of Environmental Radioactivity*. <http://dx.doi.org/10.1016/j.jenvrad.2015.06.003>.
- Hastie, T., Tibshirani, R., Friedman, J., 2009. The elements of statistical learning: Data Mining, Inference, and Prediction, second ed. Springer, New York.
- Hoffman F. O., Frank M. L., Blaylock B. G., von Bernuth R. D., Deming E. J., Graham R. V., Mohrbacher D. A., Waters A. E., 1989. Pasture grass interception and retention of ¹³¹I, ⁷Be and insoluble microspheres deposited in rain. Oak Ridge National Laboratory report ORNL-6542. Environmental Sciences Division Publication No. 3247.
- Hoffman F.O., Thiessen K.M., Frank M.L., Blaylock B.G., 1992a. Quantification of the interception and initial retention of radioactive contaminants deposited on pasture grass by simulated rain. *Atmospheric Environment*, 26A (18): 3313-3321.
- Hoffman F.O., Thiessen K.M., Frank M.L., Blaylock B.G., 1992b. Determining the collection efficiency of gummed paper for the deposition of radioactive contaminants in simulate drain. *Health Physics*, 62: 439-442.
- Hoffman F.O., Thiessen K.M., Rael R.M., 1995. Comparison of interception and initial retention of wet-deposited contaminants on leaves of different vegetation types. *Atmospheric Environment*, 29: 1771-1775.
- Horton, R.E., 1919. Rainfall interception. *Monthly Weather Rev.* 47, 603-623.
- International Atomic Energy Agency, 2009. Quantification of radionuclide transfer in terrestrial and freshwater environments for radiological assessments IAEA-TECDOC-1616.
- International Atomic Energy Agency, 2010. Handbook of parameter values for the prediction of radionuclide transfer in terrestrial and freshwater environments. IAEA, technical report series n° 472. Available on http://www-pub.iaea.org/MTCD/publications/PDF/trs472_web.pdf (last accessed on April 22 2015).
- Keim R.F., Skaugset A.E., Weiler M., 2006. Storage of water on vegetation under simulated rainfall of varying intensity. *Advances in Water Resources* 29, 974-986
- Keppel H., 1966. Cation exchange phenomena in plant leaves. *Proc. Isotopes in Plant Nutrition and Physiology*, IAEA, 5-9 sept 1966, SM - 77/28.
- Jacob P., Müller H., Pröhl G., Voigt G., Berg D., Paretzke H.G., Regulla D., 1993. Environmental behaviour of radionuclides deposited after the reactor accident of Chernobyl and related exposures. *Radiat. Environm. Biophys.* 32, 193-207.
- Jouven M., Carrère P. and Baumont R., 2006a. Model predicting dynamics of biomass, structure and digestibility of herbage in managed permanent pastures. 1. Model description. *Grass and Forage Science*, 61: 112-124
- Jouven M., Carrère P. and Baumont R., 2006b. Model predicting dynamics of biomass, structure and digestibility of herbage in managed permanent pastures. 1. Model evaluation. *Grass and Forage Science*, 61: 125-133
- Katata G., Ota M., Terada H., Chino M. and Nagai H., 2012a. Atmospheric discharge and dispersion of radionuclides during the Fukushima Daiichi Nuclear Power Plant accident. Part I: Source term estimation and local-scale atmospheric dispersion in early phase of the accident. *Journal of Environmental Radioactivity*, 109: 103-113.
- Katata G., Terada H., Chino M., Nagai H. and Chino M., 2012b. Numerical reconstruction of high dose rate zones due to the Fukushima Dai-ichi Nuclear power Plant Accident. *Journal of Environmental Radioactivity*, 111: 2-12.

- Keim R.F., Skaugset A.E., M. Weiler M., 2006. Storage of water on vegetation under simulated rainfall of varying intensity. *Water Resources* 29 (2006) 974–986
- Kinnersley R.P., Goddard A.J.H., Minski M.J., Shaw G., 1997. Interception of Cesium-contaminated rain by vegetation. *Atmospheric Environment* 31(8), 1137-1145.
- Korsakissok, I.; Mathieu, A. and Didier, D., 2013. Atmospheric dispersion and ground deposition induced by the Fukushima Nuclear power plant accident: a local-scale simulation and sensitivity study. *Atmospheric Environment*, 70: 267–279.
- Massman W.J., 1983. The derivation and validation of a new model for the interception of rainfall by forests. *Agric. Meteorol.* 28, 261–286.
- Mathieu, A.; Korsakissok, I.; Quélo, D.; Groel, J.; Tombette, M.; Didier, D.; Quentric, E.; Saunier, O.; Benoit, J.-P. and Isnard, O., 2012. Atmospheric dispersion and deposition of radionuclides from the Fukushima Daiichi nuclear power plant accident. *Elements*, 8(3): 195–200.
- Merriam R.A., 1961. Surface water storage on annual ryegrass. *Journal of Geophysical Research* 66(6) :1833-1838.
- Müller, H. and G. Pröhl, 1988. Cesium transport in food chains – comparison of model predictions and observations. In *Reliability of Radioactive transfer Models* (edited by Desmet G.), 104-113. Elsevier Applied Sciences, New-York.
- Müller H., Pröhl G., 1993. ECOSYS-87: A dynamic model for assessing radiological consequences of nuclear accidents. *Health Physics*, 64(3): 232-252.
- Muzylo A., Llorens P., Valente F., Keizer J.J., Domingo F., Gash J.H.C., A review of rainfall interception modelling. *Journal of Hydrology* 370 (2009) 191–206
- National Cancer Institute, 1997. Exposure of the American People to Iodine-131 from Nevada Nuclear-Bomb Tests. <http://www.cancer.gov/about-cancer/causes-prevention/risk/radiation/i131-report-and-appendix>. (last accessed on October 9, 2015).
- Parent, E., Bernier, J., 2007. *Le raisonnement bayésien: modélisation et inférence*, Springer, Paris.
- Plummer, M., 2003. JAGS : a program for analysis of Bayesian graphical models using Gibbs sampling. In: *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003)*. Vienna. March, pp 20-22.
- Pröhl G., 1990. Modellierung der Radionuklidausbreitung in Nahrungsketten nach Deposition von Strontium-90, Cäsium-137 und Jod-131 auf landwirtschaftlichgenutzte Flächen. München–Neuherberg, GSF-Report 29/90.
- Pröhl G., Hoffman F.O., 1996. The Interception, Initial and Post-deposition Retention by Vegetation of Dry and Wet Deposited Radionuclides. IAEA-TECDOC-857.
- Pröhl G., 2009. Interception of dry and wet deposited radionuclides by vegetation. *Journal of Environmental Radioactivity*, 100: 675-682.
- Rutter A.J., Kershaw K.A., Robins P.C., Morton A.J., 1971. A predictive model of rainfall interception in forests. I. Derivation of the model from observations in a plantation of corsican pine. *Agric. Meteorol.* 9, 267–384.
- Schönherr J., 1977. Plant cuticles are polyelectrolytes with isoelectric points around them. *Plant Physiology* 59, 145-150
- Schaefer D.A., Reiners W.A., 1990. Throughfall chemistry and canopy processing mechanisms. In *Acidic Precipitation: Vol. 3 Sources, Deposition and Canopy Interactions* (ed., S.E. Lindberg, A.L. Page and S.A. Norton), Springer, New York, 242-284.
- Spiegelhalter, D.J., Best, N.G., Carlin, B.P., Linde, A.V.D., 2002. Bayesian measures of model complexity and fit. *J. Roy. Stat. Soc. Ser. B Stat. Methodol.* 64 (4), 583-639.
- Storkey J., 2004. Modelling seedling growth rates of 18 temperate arable weed species as a function of the environment and plant traits. *Annals of Botany*, 93:681-689.
- Storkey J., 2005. Modelling assimilation rates of 14 temperate arable weed species as a function of the environment and leaf traits. *Weed Research*, 45: 361–370.
- Storkey J., 2006. A functional group approach to the management of UK arable weeds to support biological diversity. *Weed Research*: 46, 513–522.
- Storkey J. and Cussans J.W. (2007) Reconciling the conservation of in-field biodiversity with crop production using a simulation model of weed growth and competition. *Agricultural, Ecosystems & Environment*, 122:173-182.
- Sy, M.M., Ancelet, S., Henner, P., Hurtevent, P., Simon-Cornu, M., 2015. Foliar interception of radionuclides in dry conditions: a meta-analysis using a Bayesian modelling approach. *J. Environ. Radioact.* 147, 63-75.
- Terada H., Katata G., Chino M. and Nagai, H., 2012. Atmospheric discharge and dispersion of radionuclides during the Fukushima Daiichi Nuclear Power Plant accident. Part II: verification of the source term and analysis of regional-scale atmospheric dispersion. *Journal of Environmental Radioactivity*, 112: 141–154.

Table 1. Summary of wet interception experiments considered in this study, for different families of plants: grass (G), clover (C), weeds (W) or a mixture of them, different chemical elements, water (H2O) and spherical balls (Poly) of micronic diameter range. This data set consists of 440 data.

Plant (number of data)	Substance (number of data)	Availability of parameters (A: available, a: deduced from available parameters, NA : Not available)								Reference
		<i>B</i>	<i>LAI</i>	<i>H</i>	<i>I</i>	<i>T</i>	<i>f</i>	<i>f_B</i>	<i>f_{LAI}</i>	
CGW (9)	Cr (2) Pb (2) Sr (4) Poly 1µm (1)	A	NA	A	NA	NA	A	A	NA	Chadwick and Chamberlain (1970)
G (15)	I (5) Sr (5) H ₂ O (5)	A	A	A	A	A	A	A	A	Angeletti and Levi (1977a)
CG (25)	I (10) Sr (5) H ₂ O (10)	A	A	A	A	A	A	A	A	Angeletti and Levi (1977a)
G (14)	I (7) Sr (7)	A	NA	A	A	A	A	a ^a	NA	Angeletti and Levi (1977b)
G (6)	I (3) Sr (3)	A	NA	a ^b	a ^b	a ^b	A	a ^a	NA	Angeletti and Levi (1977b)
C (14)	I (7) Sr (7)	NA	NA	A	A	A	NA	a ^a	NA	Angeletti and Levi (1977b)
C (6)	I (3) Sr (3)	NA	NA	a ^b	a ^b	a ^b	NA	a ^a	NA	Angeletti and Levi (1977b)
C (96) G (96) GW (72)	I (88) Be (88) Poly 3µm (88)	A	NA	A	A	A	A	A	NA	Hoffman et al. (1989)
G (30)	I (10) Be (10) Poly 3µm (10)	NA	NA	a ^b	a ^b	a ^b	NA	A	NA	Hoffman et al. (1989)
C (5) G (5) W (5)	I (3) Be (3) Sr (3) Cr (3) Cd (3)	NA	NA	A	A	A	NA	A	A	Hoffman et al. (1995)
G (3)	I (1) Cs (1) Ba (1)	A	NA	A	NA	NA	A	A	NA	Jacob et al. (1993)
G (15)	Cs (15)	a ^c	a ^d	A	A	A	A	A	A	Kinnersley et al. (1997)
G (24)	Cs (12) Sr (12)	A	A	A	A	A	A	A	A	Bengtsson et al. (2014)

^a Calculated as f_C/H , where f_C is the ratio of the concentration in the above-ground vegetation immediately after the deposition event (in Bq kg⁻¹ dry weight) and the concentration in rainfall water (in Bq kg⁻¹ water).

^b Rainfall height, intensity and duration correspond to a series of rainfall events.

^c Calculated as f/f_B

^d Calculated as f/f_{LAI}

Table 2. Multi-linear regression models and their computed value of the posterior predictive loss criterion (PPLC).

Id	Model	Hypothesis (dependence upon substance)	PPLC
HBI	$f_B(H, B, I)$	Fully dependent	20.60
HBI-Biomass	$f_B(H, B, I)$	^a No dependence for β	24.00
HBI-Poly	$f_B(H, B, I)$	^b Poly = Be, Sr	23.74
HBI-Cs	$f_B(H, B, I)$	^c Cs = Be, Sr	23.80
HBI-PolyCs	$f_B(H, B, I)$	^{b,c} Poly = Cs = Be, Sr	29.59
HI	$f_B(H, I)$	Fully dependent	24.58
HB	$f_B(H, B)$	Fully dependent	46.24
H	$f_B(H)$	Fully dependent	49.31

^a β is estimated for all kinds of substances together (no differentiation).

^b No differentiation between polystyrene particles (Poly) and divalent cations (Be, Sr).

^c No differentiation between Cesium (Cs) and divalent cations (Be, Sr).

Table 3. Descriptive statistics (Median and 95% credible intervals) of the posterior distributions estimated for the multi-linear regression models of the log-transformed value of the mass interception fraction.

Model	Parameter	Type of substance				
		H ₂ O	I	Cs	Be, Sr	Poly 3 μ m
HBI	$\log_{10}(f_{B0})$	0.675 ^a [0.087, 1.309]	0.212 [0.056, 0.372]	1.22 [0.878, 1.504]	0.869 [0.705, 1.039]	0.392 [0.187, 0.593]
	α	-0.527 [-1.08, -0.026]	-0.922 [-0.999, -0.847]	-0.572 [-0.787, -0.340]	-0.132 [-0.218, -0.049]	-0.253 [-0.340, -0.165]
	β	-0.230 [-0.823, 0.396]	-0.389 [-0.518, -0.257]	0.124 [-0.081, 0.295]	-0.143 [-0.26, -0.021]	-0.465 [-0.586, -0.349]
	γ	-0.330 [-0.842, 0.232]	-0.104 [-0.166, -0.043]	-0.646 [-0.763, -0.517]	-0.459 [-0.529, -0.389]	-0.188 [-0.291, -0.086]
	Σ	0.140 [0.099, 0.221]	0.158 [0.139, 0.183]	0.219 [0.166, 0.313]	0.217 [0.192, 0.247]	0.163 [0.141, 0.192]
	HI	$\log_{10}(f_{B0})$	0.902 [0.782, 1.024]	0.597 [0.479, 0.717]	1.047 [0.840, 1.240]	0.998 [0.875, 1.123]
α		-0.538 [-1.077, -0.048]	-0.966 [-1.054, -0.879]	-0.559 [-0.787, -0.325]	-0.147 [-0.232, -0.064]	-0.319 [-0.43, -0.208]
γ		-0.316 [-0.812, 0.230]	-0.116 [-0.19, -0.045]	-0.620 [-0.738, -0.494]	-0.456 [-0.525, -0.385]	-0.209 [-0.328, -0.085]
Σ		0.139 [0.099, 0.218]	0.184 [0.162, 0.213]	0.235 [0.18, 0.325]	0.22 [0.195, 0.251]	0.216 [0.187, 0.253]
HB		$\log_{10}(f_{B0})$	0.161 [-0.039, 0.62]	0.085 [-0.041, 0.202]	0.1 [-0.173, 0.324]	0.134 [-0.013, 0.308]
	α	-0.818 [-0.968, -0.638]	-0.957 [-1.03, -0.884]	-0.264 [-0.733, 0.221]	-0.154 [-0.276, -0.034]	-0.299 [-0.383, -0.299]
	β	-0.705 [-0.948, -0.27]	-0.39 [-0.51, -0.273]	-0.102 [-0.372, 0.155]	-0.204 [-0.354, -0.043]	-0.477 [-0.589, -0.362]
	Σ	0.139 [0.098, 0.222]	0.165 [0.145, 0.19]	0.567 [0.438, 0.777]	0.336 [0.298, 0.383]	0.172 [0.15, 0.202]
	H	$\log_{10}(f_{B0})$	0.872 [0.717, 0.992]	0.443 [0.366, 0.518]	0.236 [0.01, 0.474]	0.348 [0.239, 0.456]
α		-0.811 [-0.967, -0.623]	-0.993 [-1.079, -0.908]	-0.295 [-0.756, 0.195]	-0.198 [-0.327, -0.07]	-0.368 [-0.479, -0.255]
Σ		0.141 [0.099, 0.223]	0.189 [0.166, 0.218]	0.562 [0.435, 0.766]	0.338 [0.299, 0.386]	0.224 [0.194, 0.262]

^a 95% credible interval indicated in brackets

Table 4. Process-based models and their computed value of the posterior predictive loss criterion (PPLC) for either $E_0=0$ or 0.6 mm h^{-1} .

Name	Model	Hypothesis (dependence upon substance)	$E_0=0 \text{ mm h}^{-1}$	$E_0=0.6 \text{ mm h}^{-1}$
			PPLC	PPLC
EM	Equilibrium	^a Fully dependent	74.78	62.04
EM-Poly	Equilibrium	^b Poly = Be, Sr	74.84	62.31
EM-Cs	Equilibrium	^c Cs = Be, Sr	79.63	69.37
EM-PolyCs	Equilibrium	^{b,c} Poly = Cs = Be, Sr	80.90	71.28
KM	Kinetic	^a Fully dependent	46.14	37.01
KM-Poly	Kinetic	^b Poly = Be, Sr	47.49	38.34
KM-Cs	Kinetic	^c Cs = Be, Sr	48.59	40.21
KM-PolyCs	Kinetic	^{b,c} Poly = Cs = Be, Sr	51.02	42.59

^a Substance-dependent parameters (CR and J/K) are estimated for each kind of substance, separately

^b No differentiation between polystyrene particles (Poly) and divalent cations (Be, Sr).

^c No differentiation between Cesium (Cs) and divalent cations (Be, Sr).

Table 5. Descriptive statistics (Median and 95% credible intervals) of the posterior distributions estimated for the equilibrium (EM) and kinetic (KM) models of the mass interception fraction, with $E_0=0 \text{ mm h}^{-1}$.

Model	Parameter	Type of substance				
		H ₂ O	I	Cs	Be, Sr	Poly 3 μm
	$SLA \text{ (m}^2 \text{ kg}^{-1}\text{)}$			13.48 ^a [12.05, 22.18]		
	$L \text{ (mm)}$			0.20 [0.12, 0.23]		
EM	CR (mm)	n.a. ^b	0 ^c	0.08 [0.003, 0.44]	0.67 [0.39, 0.95]	0.97 [0.58, 1.27]
		Σ 1.37 [0.98, 2.08]	0.43 [0.38, 0.50]	1.51 [1.17, 2.05]	0.97 [0.86, 1.11]	0.57 [0.49, 0.66]
KM	$J/K \text{ (mm h}^{-1}\text{)}$	n.a. ^b	0 ^c	0.63 [0.13, 3.29]	3.63 [2.20, 4.85]	6.95 [4.17, 8.98]
		Σ 1.37 [0.98, 2.08]	0.43 [0.38, 0.50]	1.26 [0.98, 1.72]	0.62 [0.55, 0.71]	0.46 [0.40, 0.54]

^a 95% credible interval indicated in brackets.

^b Water obeys to the Horton's like model (cf. Eq. (17)) which does not involve CR and J/K .

^c For Iodine, zero was imposed.

Table 6. Descriptive statistics (Median and 95% credible intervals) of the posterior distributions estimated for the equilibrium (EM) and kinetic (KM) models of the mass interception fraction, with $E_0=0.6 \text{ mm h}^{-1}$.

Model	Parameter	Type of substance				
		H ₂ O	I	Cs	Be, Sr	Poly 3 μm
	$SLA \text{ (m}^2 \text{ kg}^{-1}\text{)}$			13.41 ^a [12.05, 20.78]		
	$L \text{ (mm)}$			0.19 [0.12, 0.22]		
EM	CR (mm)	n.a. ^b	0	0.04 [0.002, 0.26]	0.65 [0.40, 0.91]	0.96 [0.62, 1.26]
		Σ 0.40 [0.29, 0.61]	0.41 [0.36, 0.47]	1.36 [1.06, 1.86]	0.96 [0.85, 1.09]	0.57 [0.49, 0.67]
KM	$J/K \text{ (mm h}^{-1}\text{)}$	n.a. ^b	0 ^c	0.43 [0.06, 2.73]	3.65 [2.31, 4.83]	7.04 [4.51, 9.11]
		Σ 0.40 [0.29, 0.61]	0.41 [0.36, 0.47]	1.22 [0.94, 1.67]	0.612 [0.54, 0.70]	0.46 [0.40, 0.54]

^a 95% credible interval indicated in brackets.

^b Water obeys to the Horton's type model (cf. Eq. (17)) which does not involve CR and J/K .

^c For Iodine, zero was imposed.

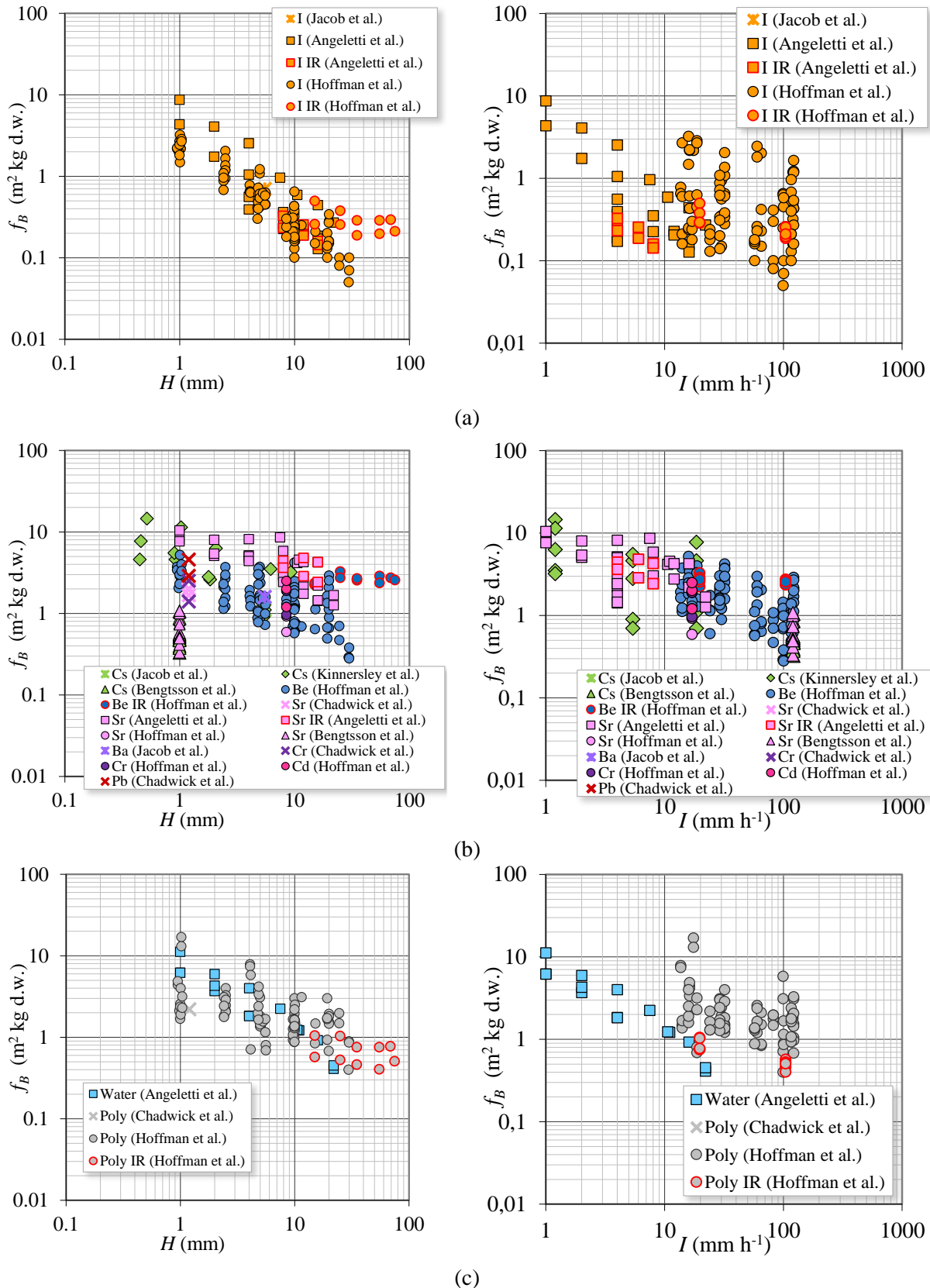
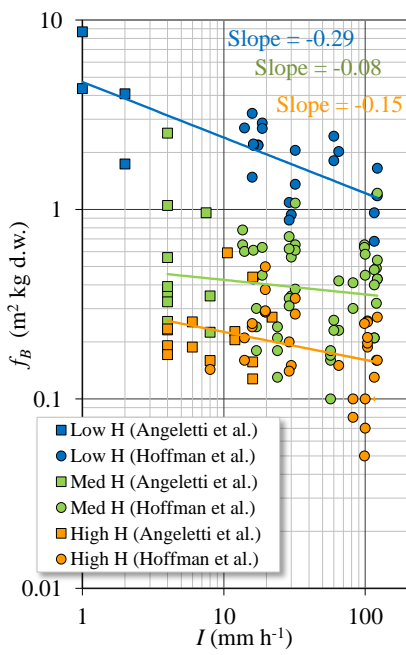
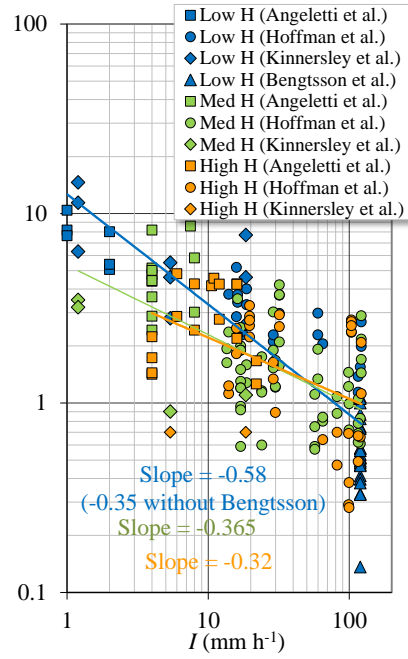


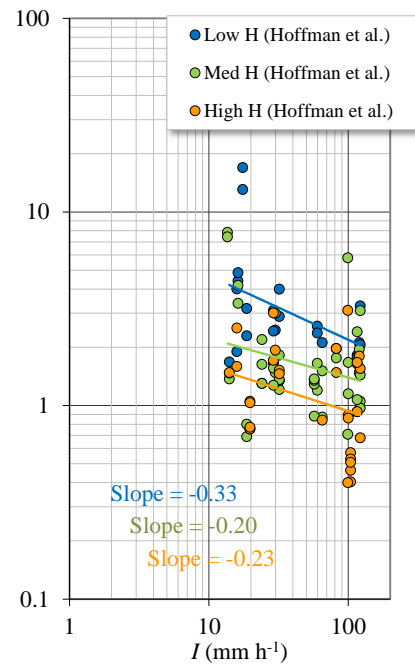
Figure 1. Mass interception fraction f_B (m² kg⁻¹ dry weight) as a function of rainfall height H and rainfall intensity I , for: (a) Anion, (b) Cations and (c) water and polystyrene particles. Experiments performed with successive rainfall events (i.e. intermittent rain) are denoted IR in the legend. See main text and Tab. 1 for a detailed description of the various data sets.



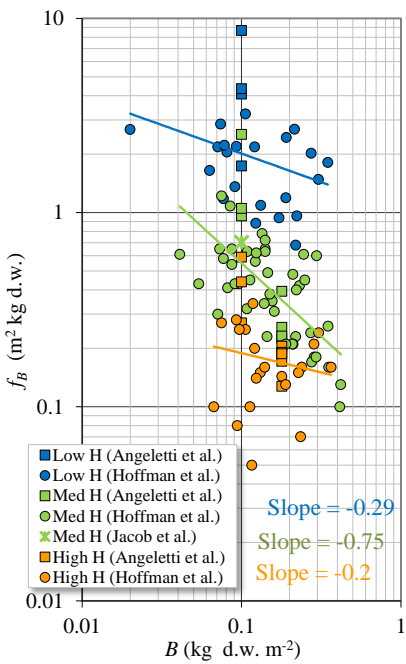
(a) Anion (I)



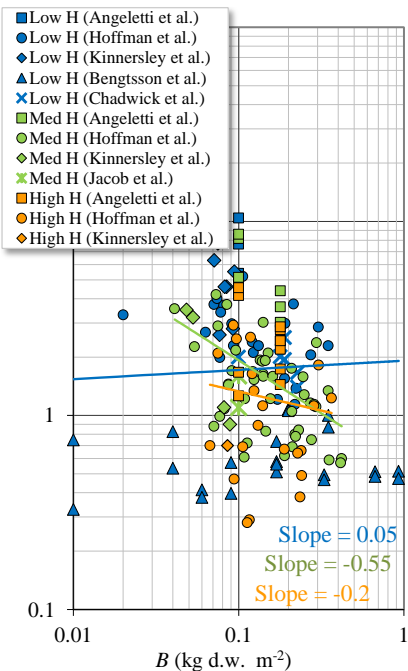
(b) Cations (Be, Sr, Cs, Cr, Cd)



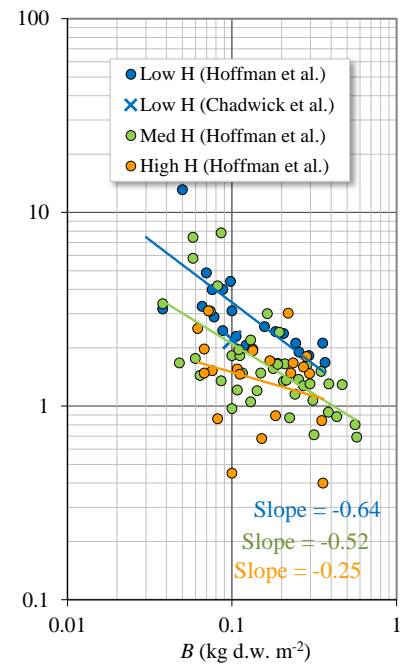
(c) Polystyrene particles



(d) Anion (I)

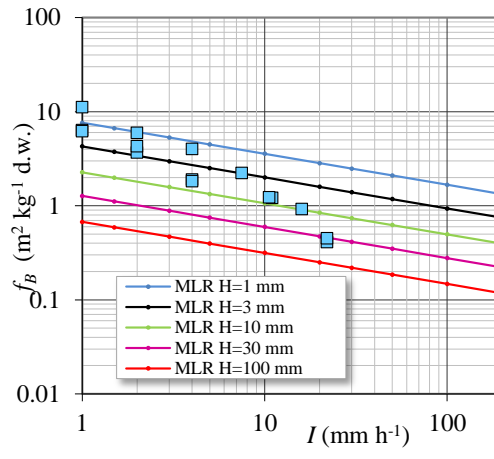
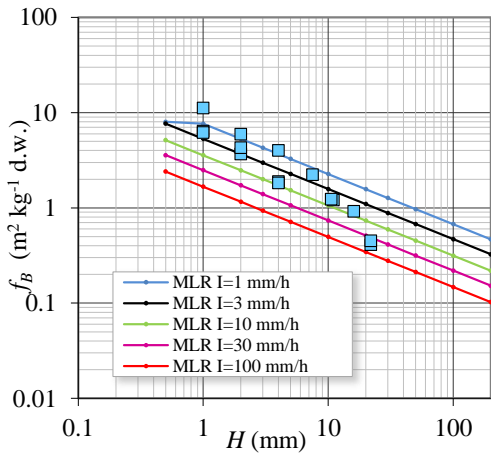


(e) Cations (Be, Sr, Cs, Cr, Cd)

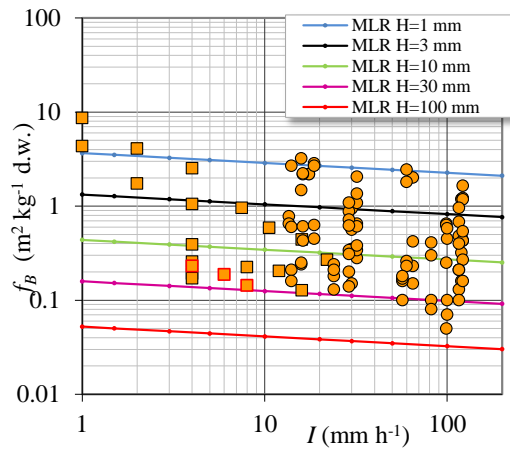
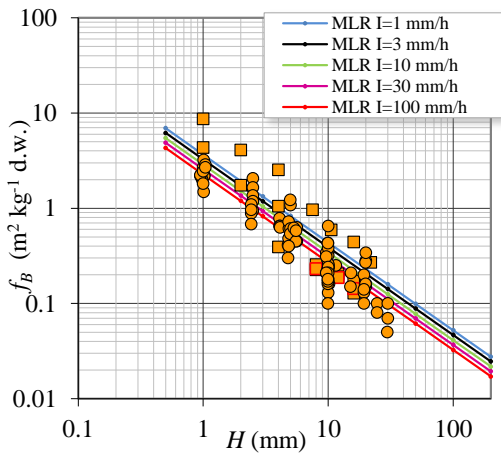


(f) Polystyrene particles

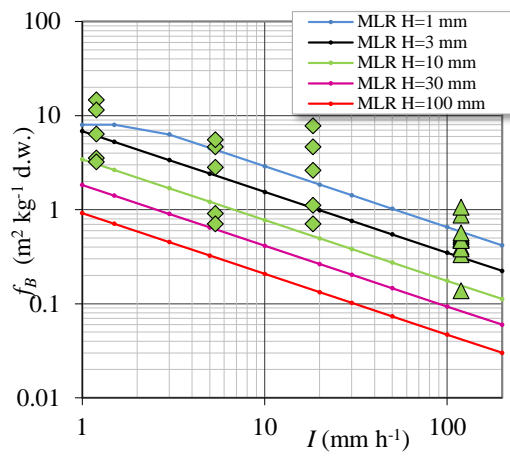
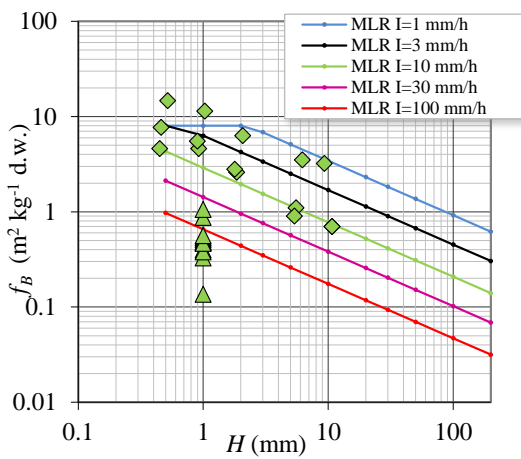
Figure 2. Mass interception fraction f_B ($\text{m}^2 \text{kg}^{-1}$ dry weight) as a function of rainfall intensity I (mm h^{-1}) and dry biomass B ($\text{kg dry weight m}^{-2}$) for: (a,d) anion, (b,e) cations and (c,f) polystyrene particles. For each family of substances, data are further classified in 3 different groups of rainfall heights: Low H (0.4-3 mm), Med H (3-12 mm) and High H (12-75 mm). Regression lines are superimposed for each group of data. See main text and Tab. 1 for a detailed description of data sets.



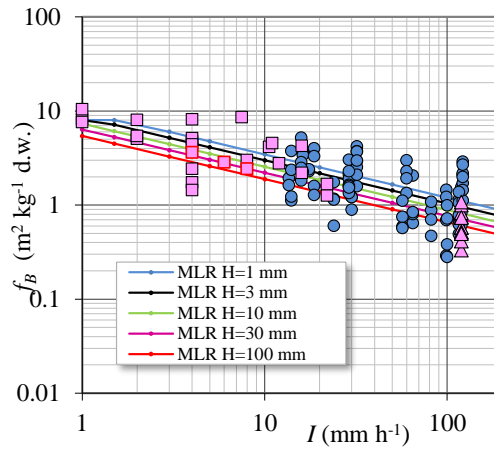
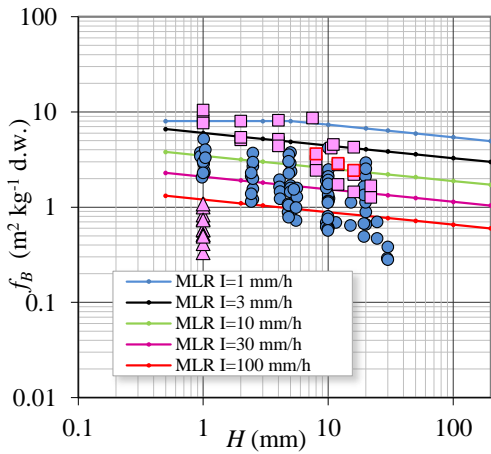
(a)



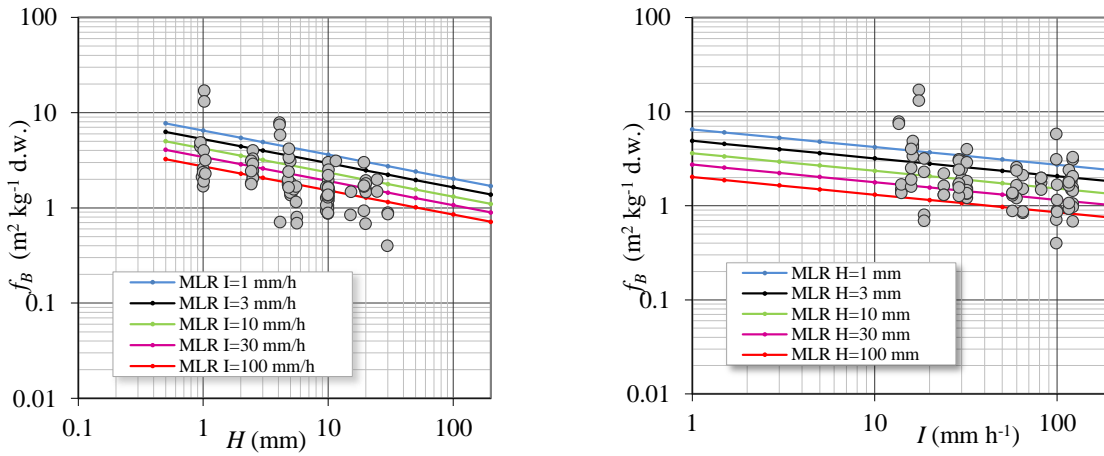
(b)



(c)



(d)



(e)

Figure 3. Best estimate (50th percentile) of the mass interception fraction f_B ($\text{m}^2 \text{kg}^{-1}$ dry weight) by the HBI model for: (a) water, (b) anion, (c) Cesium, (d) divalent cations and (e) polystyrene particles, for a mean biomass of $0.125 \text{ kg dry weight m}^{-2}$. Observations for which H , B and I were all available are displayed as colored points (same legend as in Figure 1).

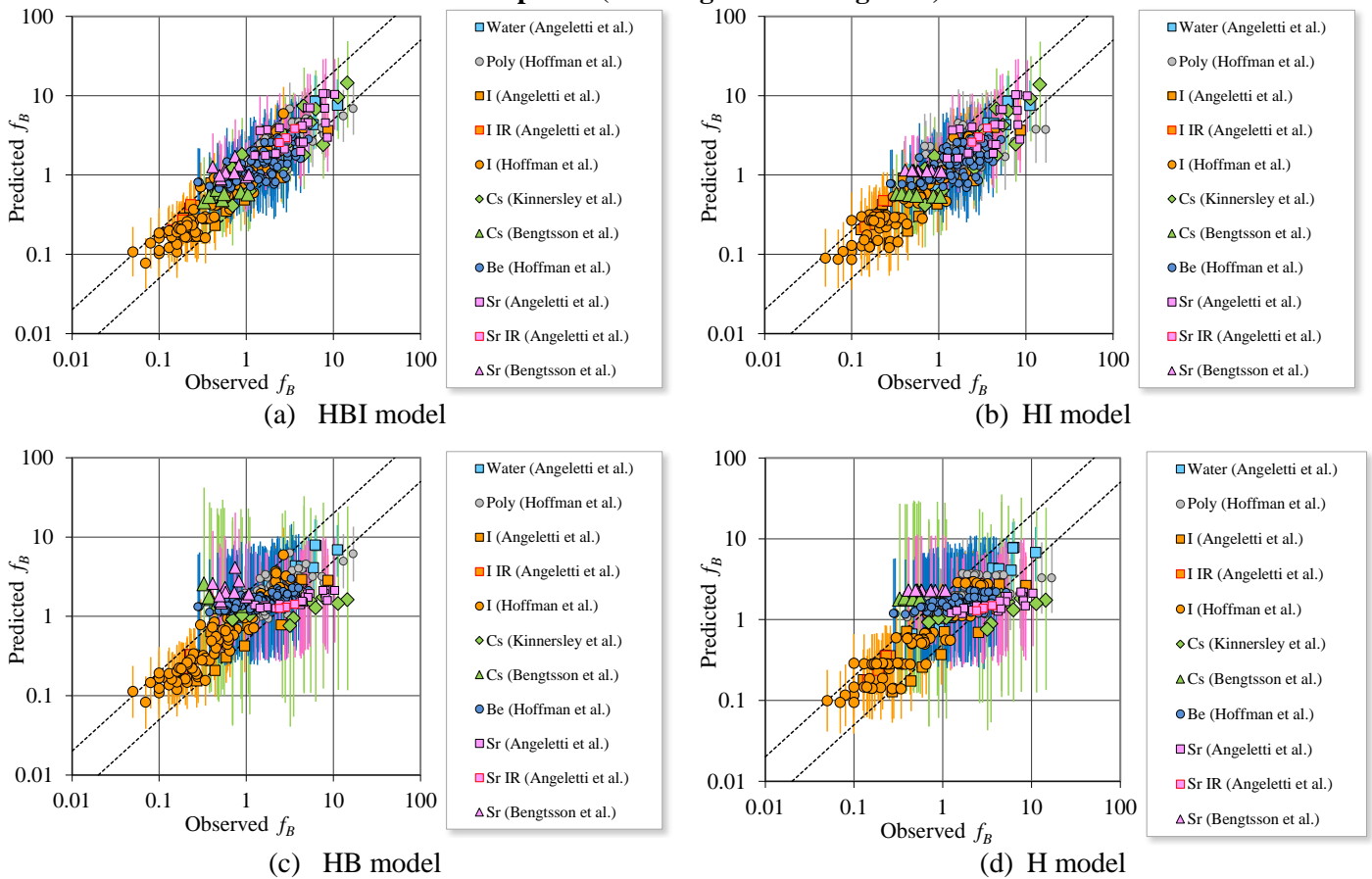


Figure 4. Predicted versus observed mass interception fraction f_B ($\text{m}^2 \text{kg}^{-1}$ dry weight) for all observations where H , B and I were available, using the following models: (a) HBI, (b) HI, (c) HB and (d) H. Error bars quantifies the 95% credible interval.

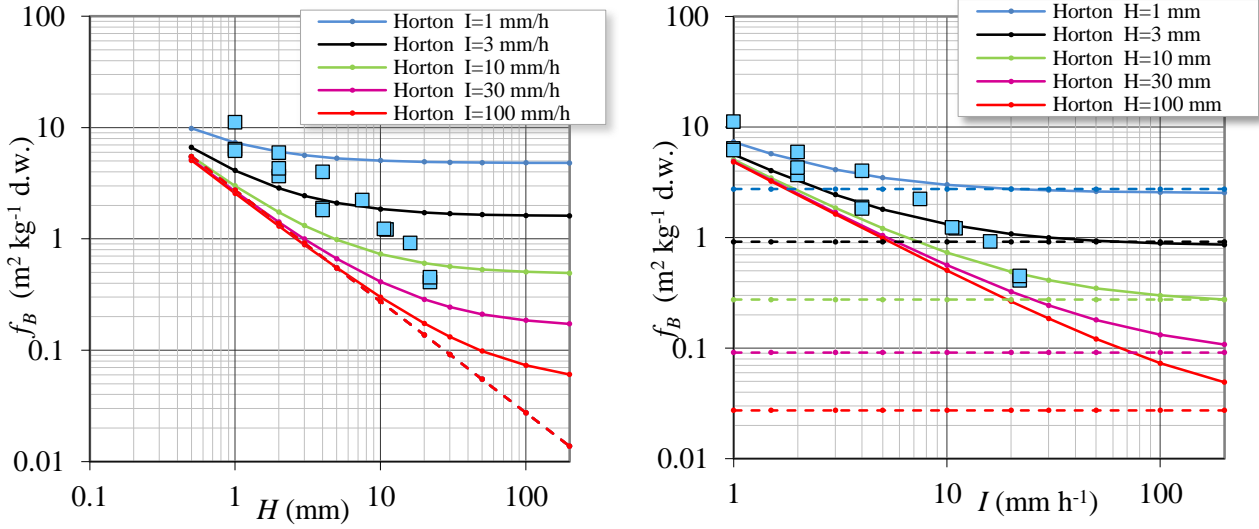


Figure 5. Best estimate (50th percentile) of the mass interception fraction f_B ($\text{m}^2 \text{kg}^{-1}$ dry weight) of water by the Horton's model, for a mean biomass of $0.125 \text{ kg dry weight m}^{-2}$. Dotted and solid lines correspond to $E_0=0 \text{ mm h}^{-1}$ and $E_0=0.6 \text{ mm h}^{-1}$, respectively. Observations for which H , B and I were all available are displayed as colored points (same legend as in Figure 1).

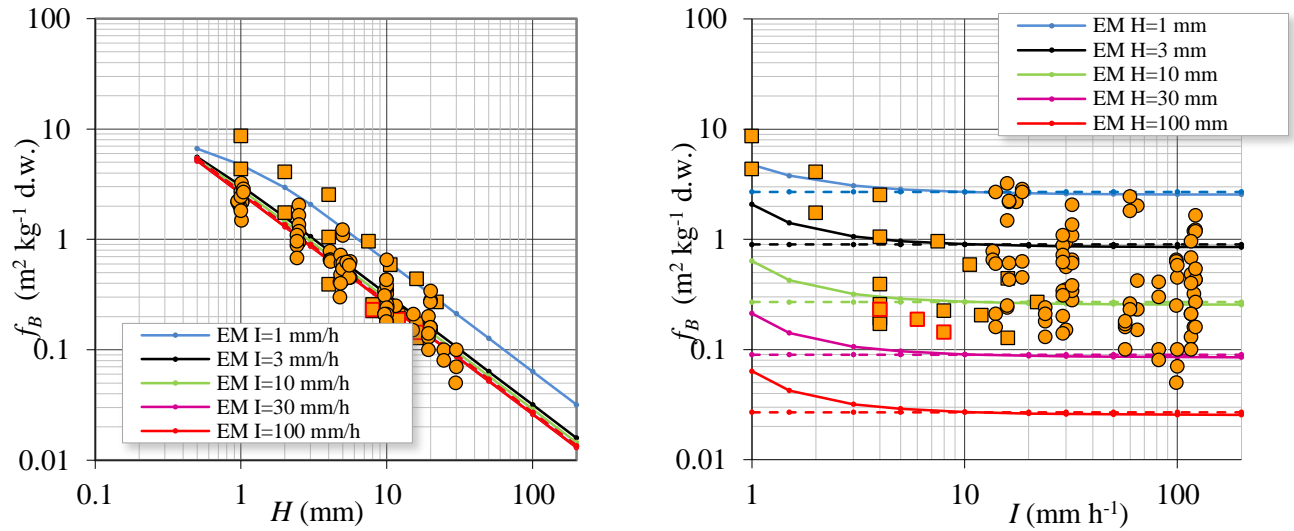


Figure 6. Best estimate (50th percentile) of the mass interception fraction f_B ($\text{m}^2 \text{kg}^{-1}$ dry weight) of iodine by the equilibrium model (EM) or kinetic model (KM), for a mean biomass of $0.125 \text{ kg dry weight m}^{-2}$. Dotted and solid lines correspond to $E_0=0 \text{ mm h}^{-1}$ and $E_0=0.6 \text{ mm h}^{-1}$, respectively. Predictions by the kinetic model are identical. Observations for which H , B and I were all available are displayed as colored points (same legend as in Figure 1).

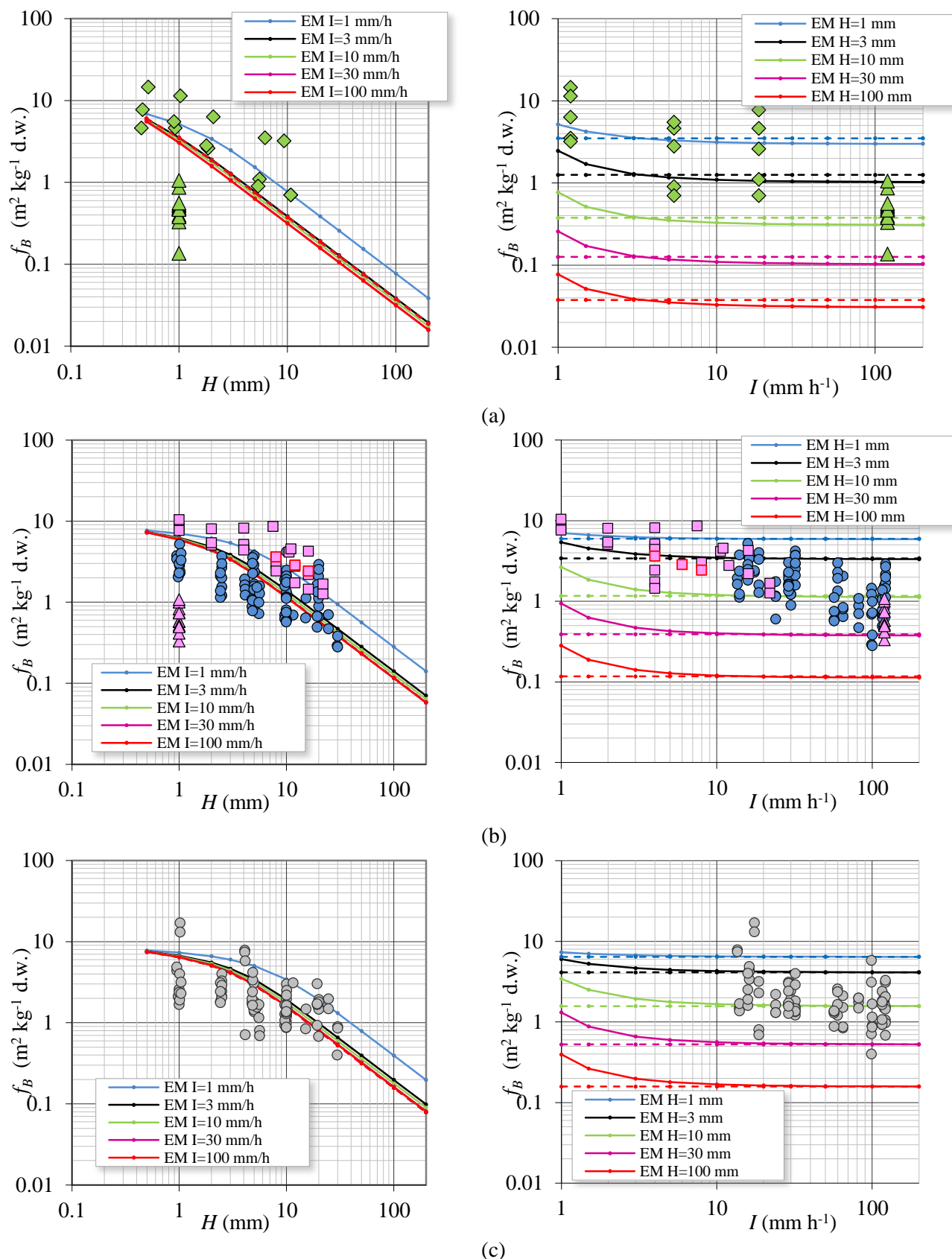


Figure 7. Best estimate (50th percentile) of the mass interception fraction f_B ($\text{m}^2 \text{kg}^{-1}$ dry weight) by the equilibrium model (EM) for: (a) Cesium, (b) divalent cations and (c) polystyrene particles, for a mean biomass of $0.125 \text{ kg dry weight m}^{-2}$. Dotted and solid lines correspond to $E_0=0 \text{ mm h}^{-1}$ and $E_0=0.6 \text{ mm h}^{-1}$, respectively. Observations for which H , B and I were all available are displayed as colored points (same legend as in Figure 1).

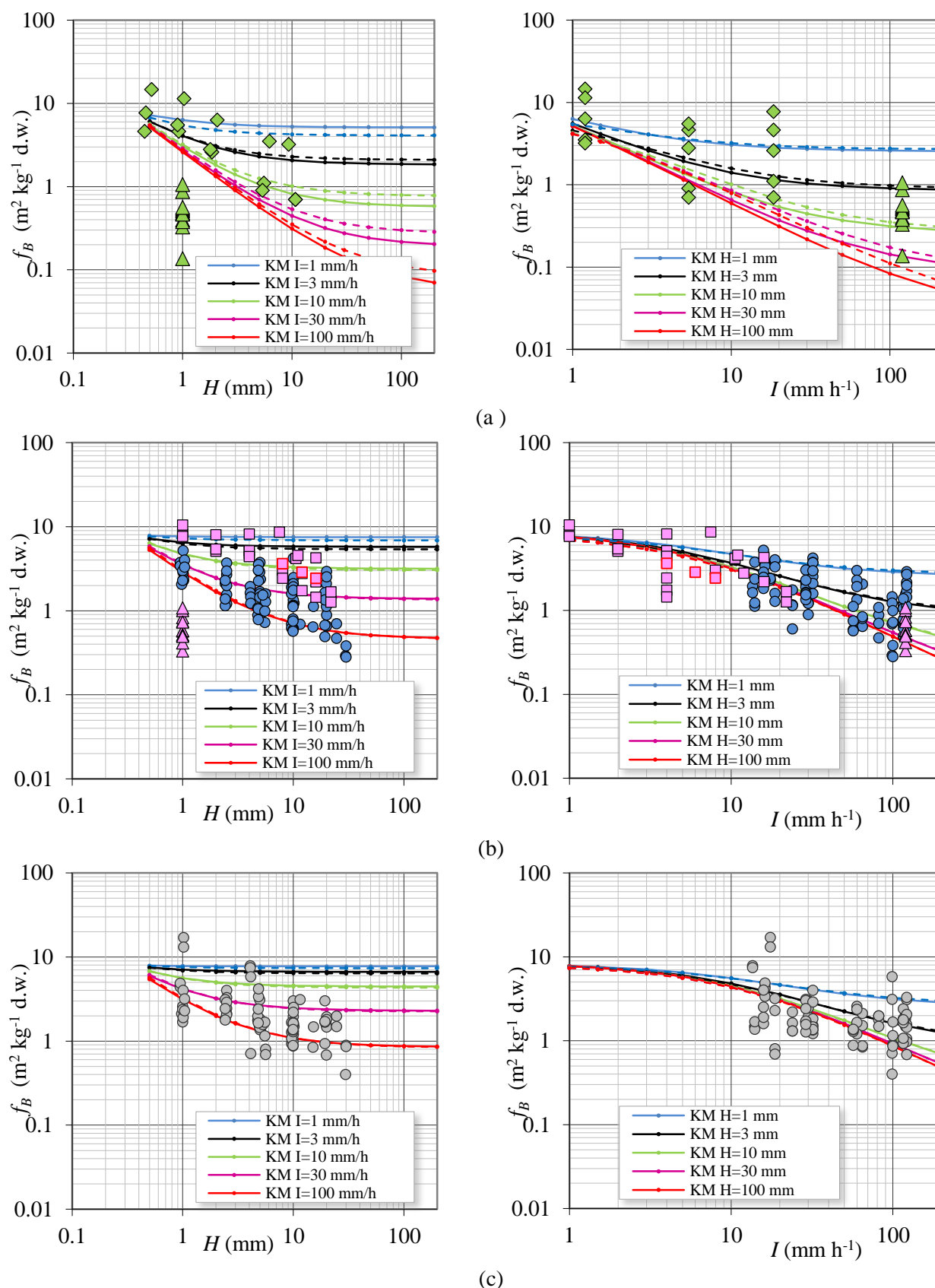


Figure 8. Best estimate (50th percentile) of the mass interception fraction f_B (m² kg⁻¹ dry weight) by the kinetic model (KM) for: (a) Cesium, (b) divalent cations and (c) polystyrene particles, for a mean biomass of 0.125 kg dry weight m⁻². Dotted and solid lines correspond to $E_0=0$ mm h⁻¹ and $E_0=0.6$ mm h⁻¹, respectively. Observations for which H , B and I were all available are displayed as colored points (same legend as in Figure 1).

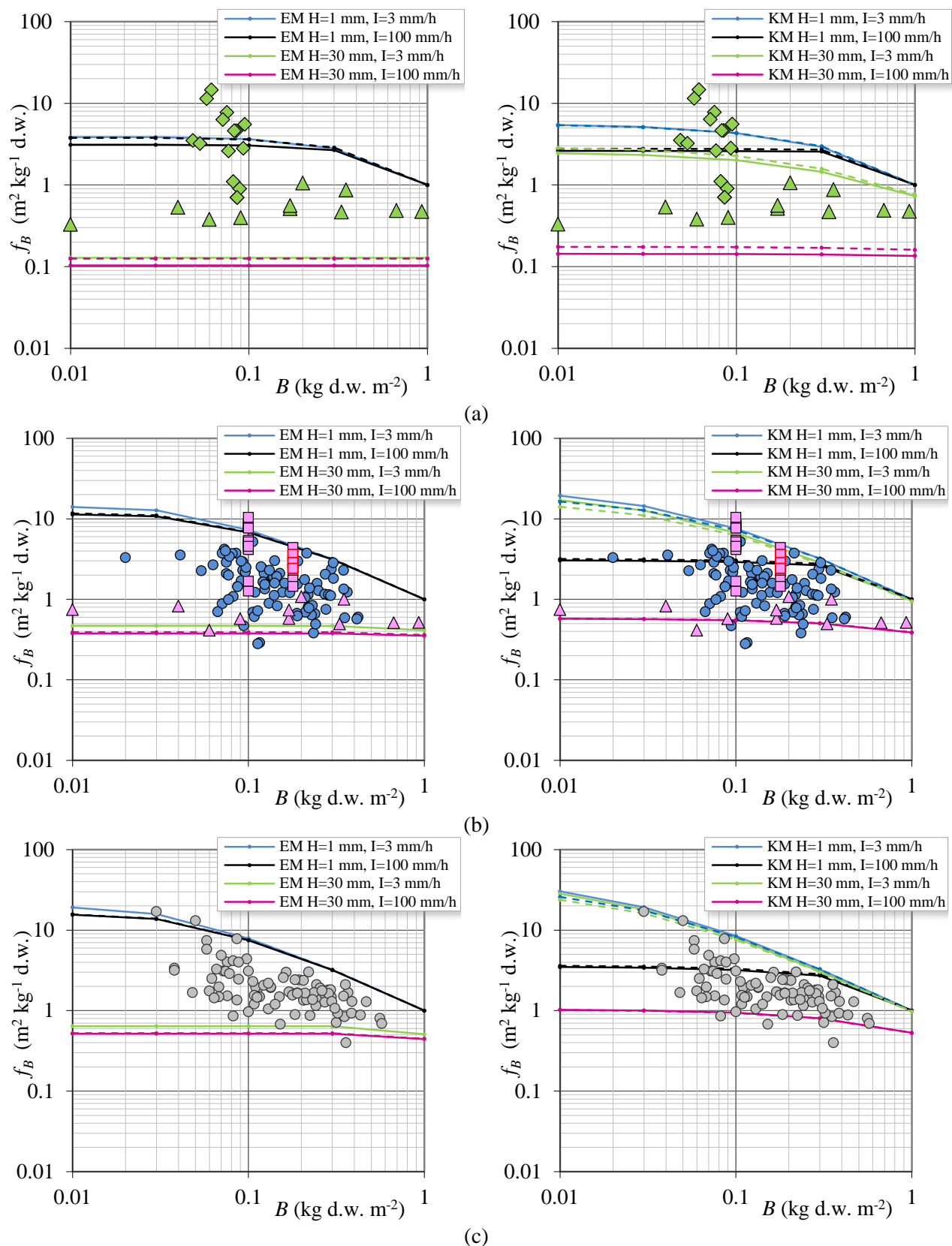


Figure 9. Best estimate (50th percentile) of the mass interception fraction f_B ($\text{m}^2 \text{kg}^{-1}$ dry weight) by the equilibrium model (EM) and kinetic model (KM) for: (a) Cesium, (b) divalent cations and (c) polystyrene particles, with varying rainfall characteristics. Dotted and solid lines correspond to $E_0=0 \text{ mm h}^{-1}$ and $E_0=0.6 \text{ mm h}^{-1}$, respectively. Observations for which H , B and I were all available are displayed as colored points (same legend as in Figure 1).

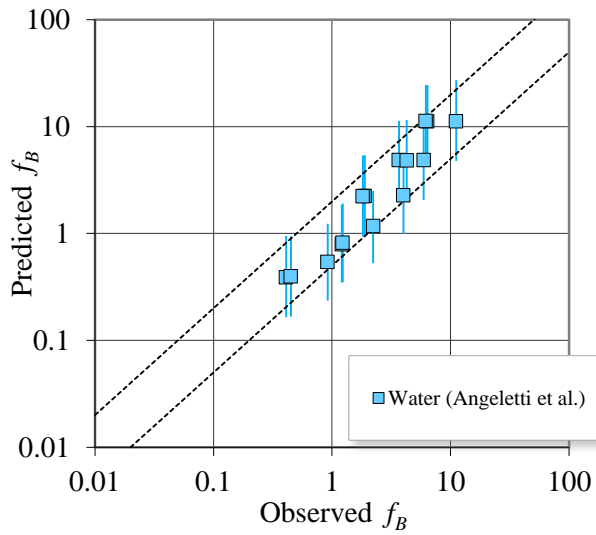


Figure 10. Predicted versus observed mass interception fraction f_B ($\text{m}^2 \text{kg}^{-1}$ dry weight) for water using Horton's type model with $E_0=0.6 \text{ mm h}^{-1}$.

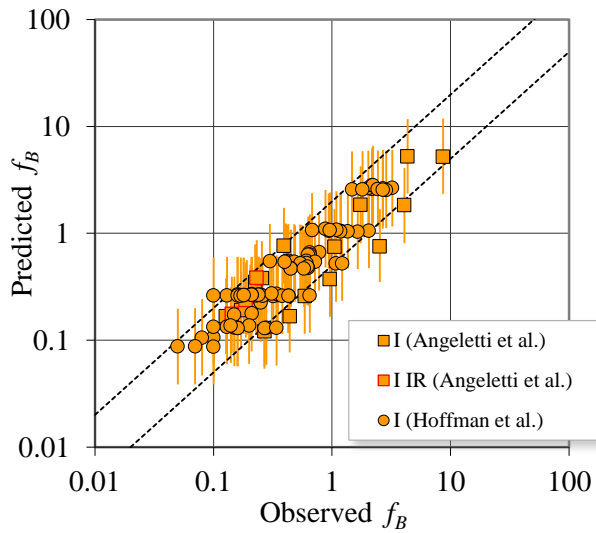


Figure 11. Predicted versus observed mass interception fraction f_B ($\text{m}^2 \text{kg}^{-1}$ dry weight) for Iodine using the EM/KM models with $E_0=0.6 \text{ mm h}^{-1}$.

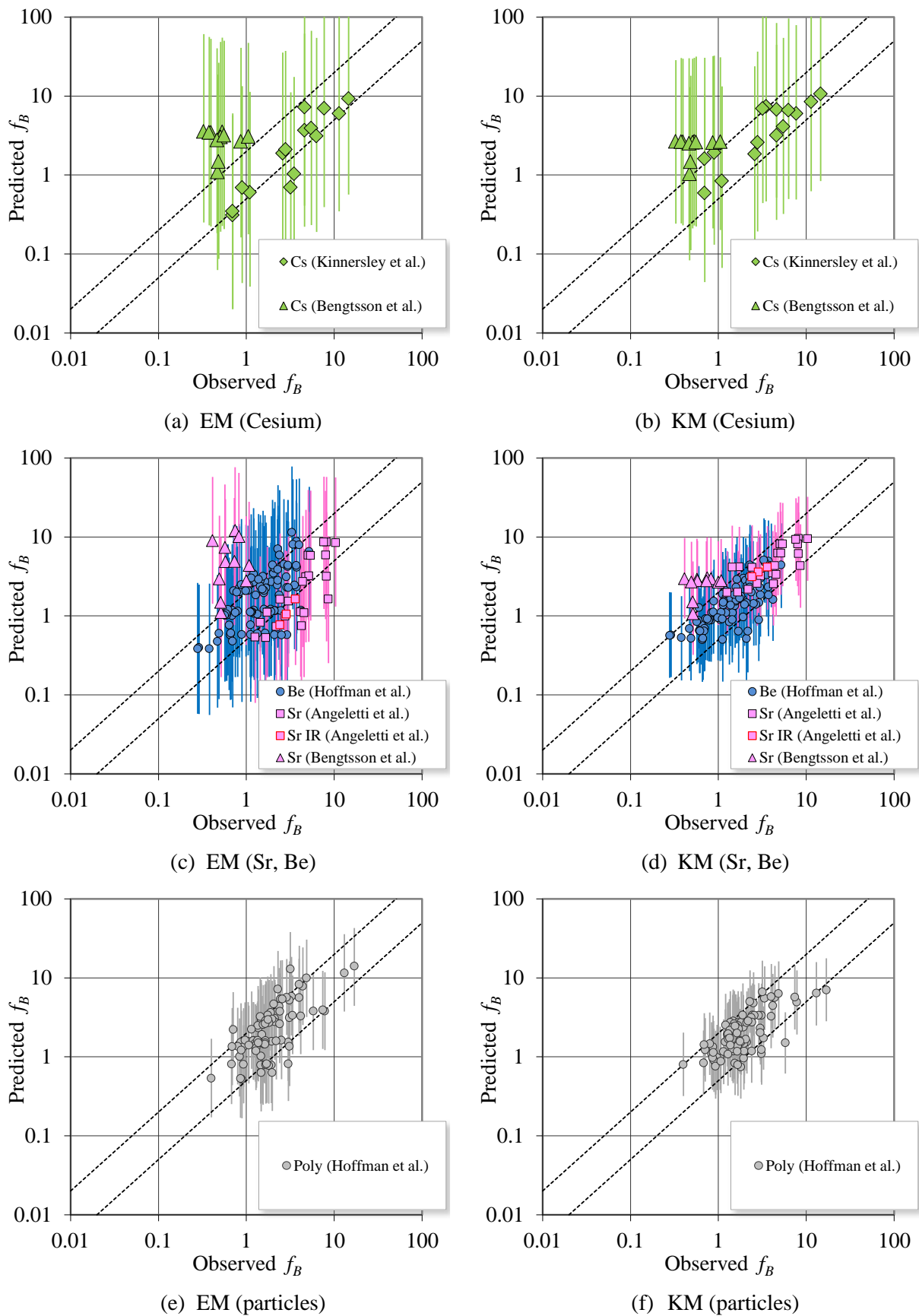
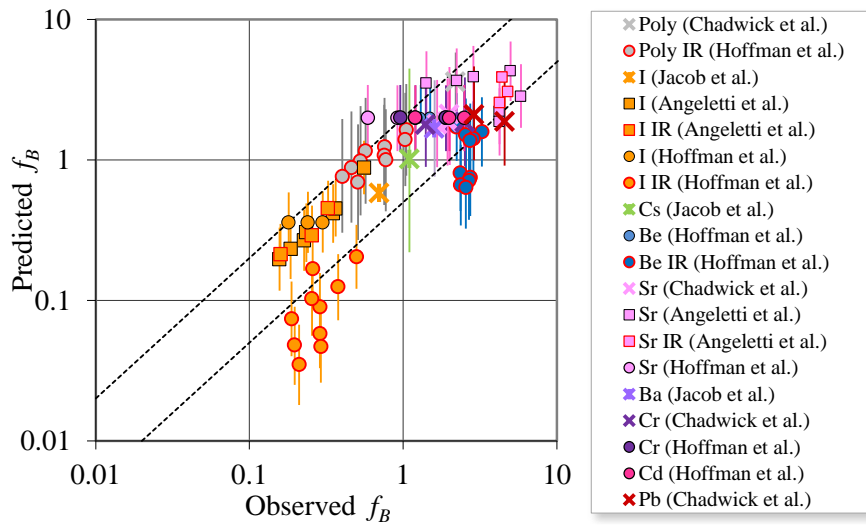
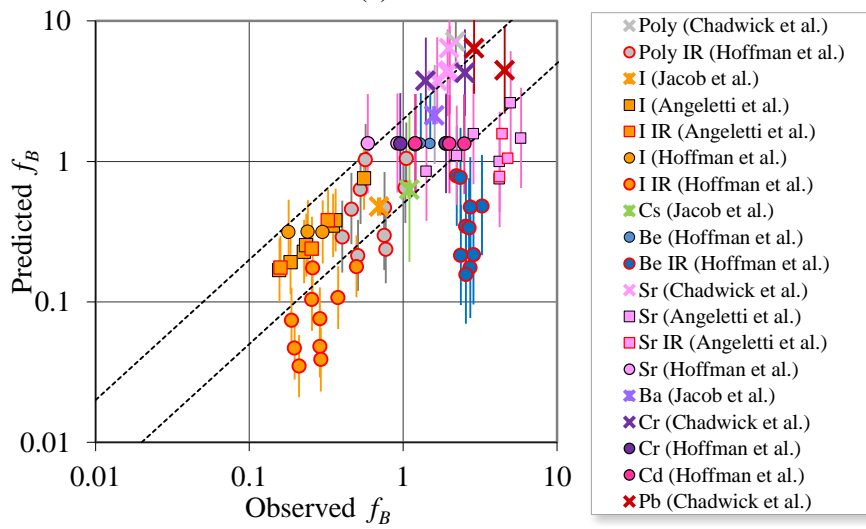


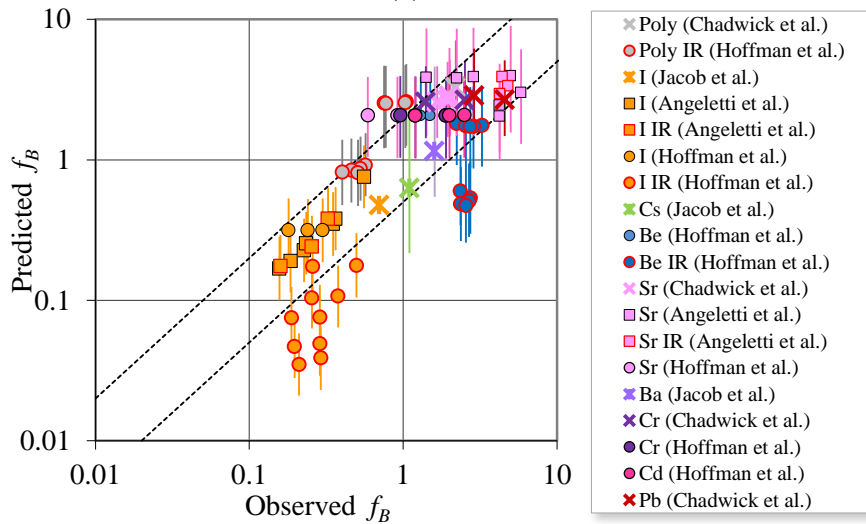
Figure 12. Predicted versus observed mass interception fraction f_B ($\text{m}^2 \text{kg}^{-1}$ dry weight) using the equilibrium model (EM) and kinetic model (KM) with $E_0=0.6 \text{ mm h}^{-1}$ for: (a,b) Cesium, (c,d) divalent cations and (e,f) polystyrene particles.



(a) HI/HB models



(b) EM model



(c) KM model

Figure 13. Predicted versus observed mass interception fraction f_B ($\text{m}^2 \text{kg}^{-1}$ dry weight) for all observations where either B or I was not available, using: (a) HI or HB model, (b) EM and (c) KM models, with $E_0=0.6 \text{ mm h}^{-1}$.

SUPPLEMENTARY INFORMATION

S.1 Brief description of the experiments considered in this study

Experiments by Chadwick and Chamberlain (1970) were done on an area of old grassland consisting of grasses with a little clover and a considerable amount of weeds, especially plantain and dandelion. Solutions of ^{85}Sr (strontium chloride, SrCl_2), ^{51}Cr (Sodium chromate, Na_2CrO_4), ^{210}Pb (lead nitrate, $\text{Pb}(\text{NO}_3)_2$) and a suspension of insoluble $1\ \mu\text{m}$ polystyrene particles were sprayed onto grassland and the initial retention by the herbage measured. The spray droplets had a diameter of $300\ \mu\text{m}$, and the volume of water used was equivalent to a precipitation of 1.2 mm. The equivalent precipitation intensity was not indicated but it is reasonable to assume that spraying took some time (i.e. a few minutes to a few tens of minutes) and implied a rather moderate intensity (i.e. a few millimeters to a few centimeters per hour). For each of the four experiments done (2 in winter, 2 in summer), 6 plots of $1.2\ \text{m}^2$ (for a total sampling area of about $7.2\ \text{m}^2$) were sampled and enabled to characterize a standard error associated to the variability between the 6 plots. Two hours were elapsed between deposition and the sampling of plants.

Angeletti and Levi (1977a, 1977b) conducted experiments on the interception of solutions of ^{85}Sr (strontium chloride, SrCl_2) and ^{131}I (sodium iodide, NaI) by ray-grass (*Lolium italicum*), clover (*Trifolium album*) or a mixture of them (grass contributing to about 75% of the total biomass) that were cultivated in greenhouses. The characteristic heights of grass and clover were respectively 15 cm and 10 cm, but biomass was not reported for pure clover plots. For each type of plant and rainfall intensity, multiple pots were sampled in order to assess the variability between (i.e. 15 and 5 pot samples in the experiments conducted by Angeletti et al., 1977a, and Angeletti et al., 1977b respectively within a total area of 0.215 and $0.134\ \text{m}^2$ respectively). Plant sampling occurred 24 hours after radioactive contamination. Another series of experiments were conducted with two consecutive deposition events, of one hour each, with 24 hours of interval in between.

Hoffman et al. (1989, 1992) conducted an extensive series of field experiments on one-meter-square stands of clover (*Trifolium repens*) or fescue (*Festuca arundinacae*) and mixed stands of old field vegetation composed of fescue grasses, other grass species and weeds. All experiments were conducted in summer season. Solutions of ^{131}I (iodide, I^- , or periodate, IO_4^-), ^7Be (Beryllium dichloride, BeCl_2) and a suspension of insoluble polystyrene microspheres of 3, 9 and $25\ \mu\text{m}$ diameter were applied with a rainfall simulator with droplet sizes ranging from 1.7 to 2.5 mm. Only diameter of $3\ \mu\text{m}$ will be considered in this study. The initial interception and retention of rain-deposited substances were investigated by sampling pasture vegetation in each plot within a $0.33\ \text{m}^2$ ring, without any replicate. Sampling began 1 hour after application. In a separate experiment, simulated rain was applied intermittently to fescue grass with approximately 30 min between two applications (apart from one where 12 hours were elapsed), up to 5 consecutive applications. Unfortunately, no biomass measurements were made in these latter experiments. Our study includes further data acquired by Hoffman et al. (1995) for 3 different types of plants: fescue (*Festuca arundinacae*), bush clover (*Lespedeza bicolor*) and golden rod (*Solidago altissima*) and a wider range of contaminants: soluble ^{51}Cr (as CrO_4^{2-}), ^{85}Sr (as Sr^{2+}), ^{109}Cd (as Cd^{2+}), ^7Be (as Be^{2+}) and ^{131}I (as I^-). It was shown that Cr added to an artificial rain as chromate (CrO_4^{2-}) had retention values similar to the other cations, because CrO_4^{2-} in aqueous solution was likely to be readily reduced to a hydrated Cr^{3+} ion with a net positive charge. Plants of each species except the fescue were selected from a nearby field site and cultivated in $176\ \text{cm}^2$ pots while for fescue subplots of $100\ \text{cm}^2$ each were taken from grass plots grown in greenhouse. Three replicates were made for each type of plants that result in total sampling areas of approximately $530\ \text{cm}^2$ and $300\ \text{cm}^2$ for bush clover and golden rod and

fescue respectively. The characteristic heights of grass, clover and weed were respectively 10, 15 and 45 cm (3 replicates were made for each plant species). Application of artificial rain was carried out in an environmental growth chamber; average drop size was 1.7 mm. Following the application of the artificial rain the plants were allowed to dry for 1 hour without disturbance.

Direct estimates of wet interception of Chernobyl fallouts by grass were included in present study. Jacob et al. (1993) estimated interception fractions for ^{131}I , ^{137}Cs and ^{140}Ba through the radioactivity measurement in samples collected after the first rainfall event (on May 1, 1986) following the Chernobyl accident in Munich. The biomass of grass and the characteristic rainfall for this event were not measured but estimated.

Kinnersley et al. (1997) went on to perform a series of experiments to quantify the interception of rain-borne Cesium (as CsCl) by grass (*Lolium perenne*) in controlled conditions for rainfall composition and intensities typical of those encountered in the U.K. Vegetation was grown outdoors in tubes, and trimmed regularly to a height of 10 cm; 6 samples were collected at different time intervals and for different intensity values, each within a square of approximately 324 cm² (resulting in a total sampling area of about 0.2 m²), to estimate the variability.

The interception and retention of wet deposited ^{134}Cs and ^{85}Sr (in the form of chlorides) by a ley pasture were recently investigated by Bengtsson et al. (2014) in a field trial using a rainfall simulator. The experimental ley was a mixture of clover and grass, which composed of: 60% timothy (*Phleum pratense*), 30% meadow fescue (*Festuca pratensis*) and 10% clover species (*Trifolium pratens*, *T. repens*). The sampling of contaminated above-ground biomass was done within a frame of 625 cm² in the middle of each one-meter-square parcel, with 3 replicates for each of the different growth stages investigated.

S.2 Choice of the prior distributions for inferred models parameters

The regression coefficients of the MLR model were assigned very large prior distributions (i.e. $N(0,100)$) which carry very little intrinsic information.

Regarding EM/KM models, the prior probability density function for L was considered uniform and ranged between 0.1 and 0.5 mm. This range is slightly larger than the values reported for grass vegetation, between 0.2 and 0.4 mm, by Clark (1940) and Merriam (1961). For some of the plant plots considered in this study (Angeletti et al., 1977a; Hoffman et al., 1995; Kinnersley et al., 1997; Bengtsson et al., 2014), the specific leaf area was reported or could be calculated as LAI/B . The 31 values of SLA ranged from 1.64 to 60 m² kg⁻¹ dw, with a median value of 25.5 m² kg⁻¹ dw and 4 outsider values smaller than 3 m² kg⁻¹ dw (from Bengtsson et al., 2014). If we exclude outsiders, the remaining 27 values are consistent with the values reported for temperate weed species by Storkey et al. (2004, 2005, 2006) that ranged between 15 and 60 m² kg⁻¹ dw. They are also consistent with the values reported for temperate grasses by Cruz et al. (2002) and implemented in the PASIM model for predicting dynamics of biomass and structure of herbage in managed permanent pastures by Jouven et al. (2006a, 2006b): these values ranged from 19 to 33 m² kg⁻¹ dw. It was thus decided to start calculations with a uniform prior distribution between 10 and 60 m² kg⁻¹ dw. Regarding substance-dependent parameters – concentration ratio CR or absorption velocity J/K –, large a prior distribution were imposed as no information exists in the literature. Values ranged from 0 to 100 mm or 100 mm h⁻¹, for a non-interacting and strongly fixed substance, respectively.

7.3 Résultats, limites et perspectives

L'un des premiers défis de cette étude a consisté à collecter une base de données la plus étendue possible de facteurs d'interception (massique) humide renseignant en sus de ce facteur la hauteur de pluie et une information minimale sur les caractéristiques physico-chimiques des substances considérées. Une partie de cette base (363 données) a permis d'estimer les paramètres des différents modèles bayésiens hiérarchiques (empiriques et mécanistes) et le reste, composé de jeux de données incomplets (77 données pour lesquelles soit la biomasse, soit l'intensité de la pluie est absente) de tester les capacités prédictives des modèles. Les comparaisons faites entre les différents modèles (empiriques et basés sur les mécanismes) ont permis de souligner quelques points intéressants.

- La normalisation du facteur d'interception humide pour obtenir un facteur d'interception massique réduit la dépendance à la biomasse mais ne l'annule pas.
- Pour toutes les substances considérées, une dépendance entre le facteur d'interception et la hauteur de pluie est globalement observée.
- L'intensité de la pluie est un facteur important à prendre en compte dans la modélisation de l'interception humide. Cette dépendance est peu discutée dans la littérature. Une décroissance du facteur d'interception massique en fonction de l'intensité est observée.
- Le processus d'interception varie suivant la substance considérée ; ce qui souligne d'une part que les particules cations et les anions doivent être traités séparément. D'autre part, cela montre que la charge électrique semble influencer sur le comportement des cations dissous (monovalents comme le césium versus cations divalents comme le béryllium et le strontium). Cette différence, mise en évidence par le modèle empirique complet (prenant en compte la biomasse, la hauteur et l'intensité de pluie et séparant toutes les substances), est expliquée par le processus d'absorption modélisée suivant les deux approches décrites plus haut.

Les modèles empiriques dépendant uniquement de la hauteur de pluie, ou de la hauteur de pluie combinée à la biomasse ou encore de la hauteur et de l'intensité de la pluie produisent des prédictions moins bonnes que celles du modèle empirique complet, mais celles-ci sont plutôt satisfaisantes compte tenu de l'information disponible (biomasse et/ou intensité de la pluie sont souvent inconnues). En situation accidentelle (*e.g.* accident de Fukushima) de tels modèles emboîtés sont particulièrement utiles, en raison du peu d'informations disponibles à

court et moyen terme (pendant les rejets et quelques temps après leur arrêt). Une posture prudente est néanmoins requise dans l'analyse de leurs résultats.

L'une des nouveautés apportées par ce travail réside dans le développement de modèles basés sur les mécanismes en action lors du processus d'interception humide. On peut, cependant noter que le modèle adoptant une approche à l'équilibre du processus d'absorption reproduit moins bien les observations que le modèle basé sur une absorption à cinétique lente (et moins bien que le modèle empirique complet). La prise en compte de la variabilité spécifique à l'espèce végétale sur les paramètres SLA et L désignant respectivement la surface foliaire spécifique (exprimée en $m^2 \cdot kg^{-1}$ poids sec) et l'épaisseur caractéristique de rétention d'eau (exprimée en mm) peut constituer un point d'amélioration.

En vue de comparer les deux approches mécanistes (d'absorption à l'équilibre et d'absorption à cinétique lente), les paramètres SLA et L ont été estimés de manière commune *i.e.* dans la même inférence bayésienne. Cette approche a consisté à superposer deux versions dupliquées du jeu de données ($363 \times 2 = 726$ données) et à faire correspondre à l'aide de fonctions indicatrices chaque modèle mécaniste à un duplicata des données. Le graphe acyclique direct est illustré par [Fig. 7.1-\(a\)](#). Rappelons que seuls les paramètres SLA et L sont communs, dans cette configuration, aux deux modèles. Cette approche peut être discutée car, même si on travaille sur les mêmes données, elle pose un problème de cohérence vis-à-vis de l'utilisation de l'information disponible. On peut alors penser que les paramètres SLA et L sont estimés deux fois sur la base de deux modèles d'erreurs différents ayant néanmoins les mêmes structures (modèles d'erreur log-normaux). De plus, les deux modèles développés reflètent deux comportements un peu extrêmes (l'équilibre versus une cinétique lente). L'approche de modélisation commune aux paramètres SLA et L évoquée ci-dessus, combinée au fait que la réalité pourrait se situer entre les deux modèles mécanistes soutient l'idée de développer un modèle alternatif entre l'approche à l'équilibre et l'approche cinétique. Ainsi, un mélange de lois log-normales où le paramètre de mélange symboliserait ce compromis comme illustré par [Fig. 7.1-\(b\)](#), est fortement envisageable. Une telle approche permettrait également de s'affranchir du problème de cohérence lié à l'utilisation de l'information disponible. Dans ce cas un seul modèle d'erreur est défini et les données ne sont pas dupliquées. Des approches telles que la *Bayesian Model Averaging* (BMA) permettant de combiner différents modèles pourraient également être une alternative intéressante (voir *e.g.* [Hoeting et al. 1999](#)). De telles approches ont notamment été utilisées pour modéliser l'incertitude de modèle ([Madigan and Raftery, 1994](#); [Raftery, 1995](#); [Draper, 1995](#)).

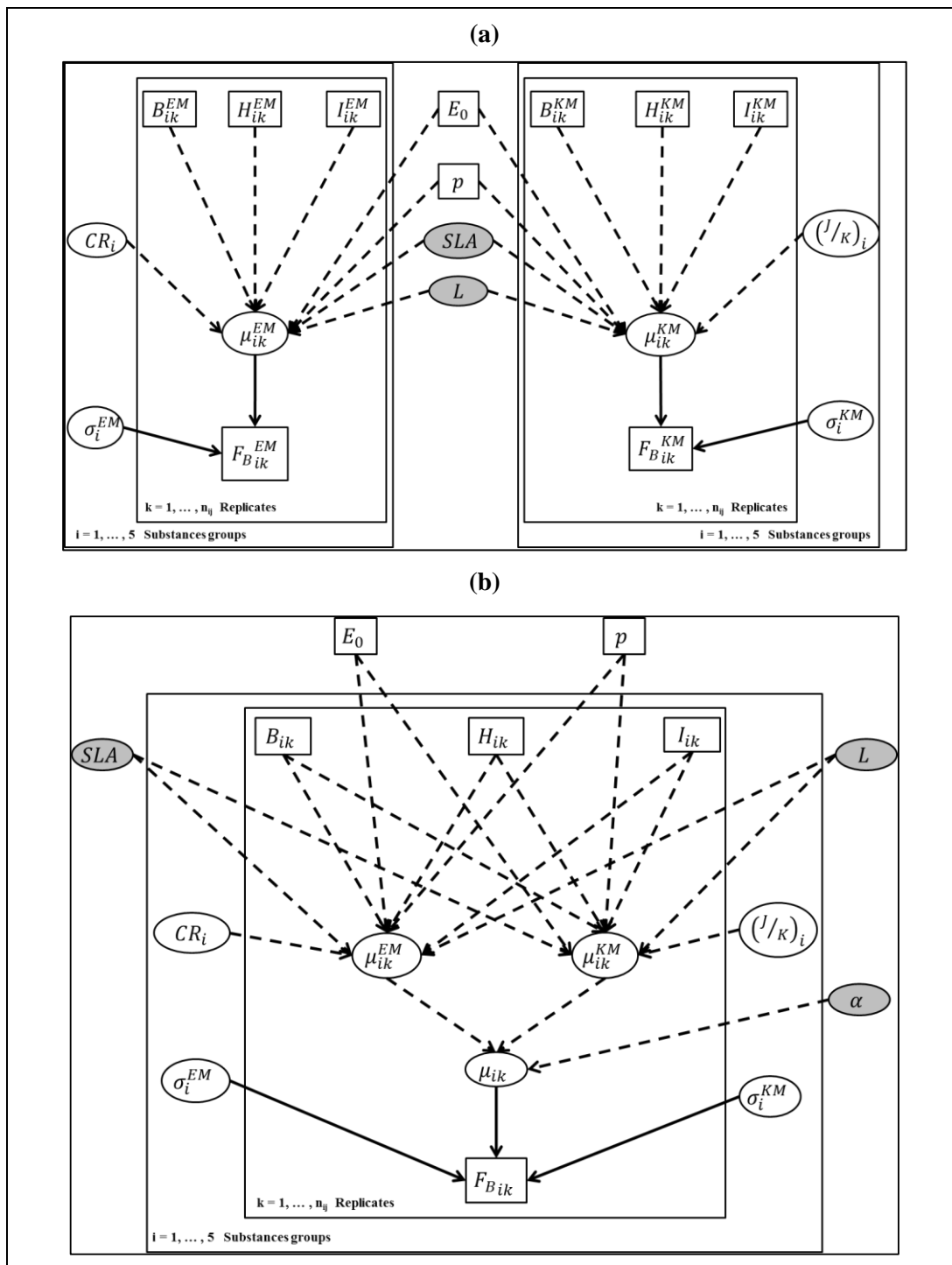


Fig. 7.1 : (a) DAG des modèles mécanistes utilisant un jeu de données dupliqué pour l'estimation commune des paramètres SLA et L . (b) DAG du modèle de mélange basé sur les deux modèles mécanistes. Le nouveau paramètre α est le paramètre de mélange. Les carrés à fond blanc correspondent aux données fixes (observations et paramètres fixes), les ellipses à fond gris sont les paramètres et les ellipses à fond blanc les variables latentes. Les flèches pleines désignent des relations stochastiques et celles en pointillées des relations déterministes.

8. Décroissance biomécanique

Ce chapitre est consacré à l'étude de la décroissance biomécanique (introduite à la [Section 3.1.2.2](#)) de la concentration de radionucléides déposés sur le couvert végétal (*e.g.* herbe de pâturage, herbacées, légumes feuilles, autres types de légumes) suite à un dépôt survenu en conditions sèches et/ou humides. Dans ce chapitre nous nous intéressons également à la propagation des incertitudes sur le facteur d'interception sèche et la constante de décroissance biomécanique dans le cadre d'une évaluation de l'impact dosimétrique à la suite d'un dépôt standardisé de $1 \text{ kBq}\cdot\text{m}^{-2}$ de césium (Cs-134/-137), d'iode (I-131) et de strontium (Sr-90). Les travaux présentés dans la suite s'inscrivent dans le cadre des actions menées dans le groupe de travail « Human Food Chain » du projet COMET. Ces travaux ont fait l'objet d'un article rédigé en collaboration avec Ali Hosseini, Astrid Liland et Håvard Thorring du NRPA et Marie Simon-Cornu de l'IRSN. L'article qui a été soumis le 28 Janvier 2016 dans la revue *Journal of Environmental Radioactivity* est reproduit dans la [Section 8.2](#).

8.1 Objectifs

8.1.1 Inférence bayésienne sur la constante de décroissance biomécanique

Dans la lignée des deux précédents [Chapitres \(6 et 7\)](#), l'un des principaux objectifs de cette étude est d'améliorer la connaissance sur le processus de décroissance biomécanique ou sur les incertitudes associées grâce à la collecte de données. Cette étude vise également à exploiter cette information dans un cadre de modélisation permettant de caractériser l'incertitude et la variabilité spécifique aux espèces végétales et aux radionucléides considérés. Ainsi, trois bases de données étendues ont été constituées, la première pour l'iode, la deuxième pour des radionucléides sous forme particulaire émis en extérieur (accidents ou expérimentations de terrain), la troisième pour des formes particulières émises en conditions contrôlées (laboratoire). Les données collectées sont issues d'études couvrant plusieurs

décennies allant des années 1960 à 2014. Ces bases de données sont notamment constituées de mesures réalisées suite à l'accident de Tchernobyl.

L'un des objectifs de ce travail était de développer, dans un cadre de modélisation bayésienne, des modèles hiérarchiques offrant la possibilité de caractériser l'incertitude sur les constantes de décroissance et d'étudier, dans ce cadre, les variations potentiellement induites par les caractéristiques physiques des radionucléides et/ou par les différentes espèces végétales étudiées (sachant que nous considérons des légumes feuilles, de l'herbe de pâturage et d'autres espèces herbacées). Cette étude a été l'occasion de montrer davantage l'intérêt d'utiliser l'approche bayésienne (hiérarchique) pour la quantification de l'incertitude et de la variabilité en radioécologie et, en particulier, dans la modélisation des processus environnementaux de transfert de radionucléides. En effet, très peu d'études utilisant ce type de modélisation existent en radioécologie. On peut citer par exemple [Hosseini *et al.* \(2013\)](#) parmi les rares. Dans ce travail, nous introduisons brièvement l'approche bayésienne hiérarchique en construisant progressivement un modèle hiérarchique appliqué à la constante de décroissance biomécanique suivant différentes hypothèses visant à caractériser les différentes sources de variabilité. Une inférence bayésienne est ensuite utilisée pour inférer les paramètres de ces modèles hiérarchiques.

8.1.2 Propagation des incertitudes

Ce travail a également visé à conduire une analyse d'incertitude et de sensibilité dans le cadre des travaux engagés par le groupe de travail "Human Food Chain" du projet COMET. Les actions définies dans ce groupe de travail ont pour objectif premier d'améliorer la qualité et la robustesse des paramètres radioécologiques clés pour l'exposition humaine par voie alimentaire par l'acquisition de « nouvelles » données et leur analyse suivant des approches pertinentes. Dans ce contexte, un scénario de calcul d'impact dosimétrique a été défini en collaboration avec les différentes parties engagées dans le groupe de travail "Human Food Chain" en vue d'évaluer l'impact, sur les modèles d'évaluation de risque, des différentes sources d'incertitude et de variabilité caractérisées pour les paramètres radioécologiques d'intérêt. Ce scénario consiste à considérer, à la date du 1^{er} Août 2013, un dépôt sec standardisé (unitaire) de Cs-134, Cs-137, I-131 et Sr-90 (1 kBq.m^{-2} pour chaque isotope) puis d'évaluer le transfert aux produits alimentaires (lait de brebis, de chèvre et de vache; viande de bœuf, de mouton et de porc; farine de blé et légumes feuilles) et l'impact

dosimétrique pour différentes classes d'âge au cours d'une période de 5 ans. Dans l'article ci-dessous, nous présentons uniquement les résultats relatifs au lait de vache, à la viande de bœuf et aux légumes feuilles pour une période d'un an.

8.2 Article: Uncertainty and variability in food chain exposure by radioactive fallouts. What can Hierarchical Bayesian modelling bring to radioecology community?

Mouhamadou M. Sy^{1*}, Ali Hosseini², Astrid Liland², Marie Simon-Cornu¹, Håvard Thørring²

¹ Institut de Radioprotection et de Sûreté Nucléaire (IRSN), Laboratoire de Modélisation pour l'Expertise Environnementale (LM2E), Cadarache, Bâtiment 159, St Paul-lez-Durance, 13115 France.

² Norwegian Radiation Protection Authority, Grini næringspark 13, Post Box 55, NO-1332 Østerås, Norway.

*Corresponding author. Email address: mouhamadou.sy87@hotmail.fr ; moustapha.sy@irsn.fr

Abstract

Better characterization of the key processes in environmental transfer of radionuclides has been pointed out as a main research focus by the radiological community. The present study was in two parts: (a) Development of hierarchical Bayesian models of weathering loss, i.e. the reduction of radionuclides concentrations in plant foliage due to wind, rain, biological dilution due to growth etc.; (b) Test scenario that incorporates results from part (a) and results from earlier studies on dry interception.

In part (a), three datasets of weathering loss values were built based on published literature covering a period of approximately 6 decades (from 1960s to 2014). For each dataset, several hierarchical Bayesian models were constructed, and models providing best fit were selected upon the application of Deviance Information Criterion (DIC). Differences between iodine, field particles and lab particles have been shown regarding expected value (95% credible intervals in brackets) of weathering rate estimated at 0.076 (0.006,0.66), 0.058 (0.028,0.11), and 0.04 (0.029,0.054) day⁻¹, respectively. The developed models demonstrated the influence on the weathering loss of plant species for iodine, plant species and radionuclides for field particles and radionuclides for lab particles.

In part (b), we assessed the impact of the uncertainty/variability about weathering loss and dry interception of radioactive fallout on the assessment of effective doses. Probabilistic simulations were performed within a defined propagation scenario (dry deposition of 1kBq.m⁻² of Cs-134, Cs-137, I-131 and Sr-90) to predict the concentrations of these radionuclides in 9 foodstuffs ((fresh and canned) leafy vegetables; flour; beef, pork and sheep meat; (fresh and canned) cow milk, and goat milk) and the effective ingestion doses received by an adult (≥17 years old) and a young child (2-7 years old) for a period of 1 year after the deposition. Sensitivity analyses using Spearman correlation coefficients were also implemented.

Keywords: Hierarchical Bayesian, Meta-analysis, Weathering loss, Uncertainty, Variability, Radioecology

1 Introduction

In the Strategic Research Agenda of radioecology (Hinton *et al.* 2013), three challenges were defined which aims at providing a quantitative and integrative assessment of radionuclide impacts on humans and wildlife for a broad range of exposure scenarios. The first one of these three challenges is “To predict human and wildlife exposure in a robust way by quantifying key processes that influence radionuclide transfers and exposure”. Four research lines (RLs) linked to this challenge were outlined (Hinton *et al.* 2013):

RL1: “Identify and mathematically represent key processes that make significant contributions to the environmental transfers of radionuclides and resultant exposures of humans and wildlife”;

RL2: “Acquire the data necessary for parameterization of the key processes controlling the transfer of radionuclides”;

RL3: “Develop transfer and exposure models that incorporate physical, chemical and biological interactions, and enable predictions to be made spatially and temporally”; and

RL4: “Represent radionuclide transfer and exposure at a landscape or global environmental level with an indication of the associated uncertainty”.

Parameterization of processes controlling the transfer of radionuclides and assessing associated parametric uncertainty in exposure models is thus a key aspect of the suggested research activities for the field of radioecology.

The scarcity of data is one of the major sources of uncertainty (Hinton *et al.* 2013) and, as identified by RL2, there is a need to both acquire more data, and also consider alternative modelling approaches. Extrapolation approaches are experiencing significant developments in radioecology since they allow to “making the best use of the available knowledge” (Beresford *et al.* 2015; Brown *et al.* 2013). Another alternative approach is the Hierarchical Bayesian Meta-Analysis (abbreviated HBMA), i.e. combining expert knowledge and data from different sources under the Bayesian paradigm by considering them in a single hierarchical modelling framework. This may allow a low number of empirical observations to be supported by inferences from more comprehensive, larger datasets (Hinton *et al.* 2013). Collating a wide range of data may lead to heterogeneous and potentially not directly comparable information. Sy *et al.* (2015) identified parameterization and measurement heterogeneity as sources of data heterogeneity. The main difficulty lies on answering how heterogeneous our data are and on the definition of a modelling framework which enables to take into account such data heterogeneity. This introduces the

concept of (data and parameters) exchangeability which characterizes the mathematical judgment of the similarity between data collated from different sources (Bernardo, 1996; Jackman, 2009; Sy *et al.*, 2015). HBMA provides well-suited features to combine data from multiple sources in a flexible modelling structure that also allows for characterization of uncertainty and variability within the same modelling structure. The Bayesian paradigm makes use of the Bayes rule (Bayes and Price, 1763) and the concept of conditional probabilities. When the purpose of analysis is to estimate model parameters and make statistical inferences, the Bayesian approach provides a cohesive framework by combining information drawn from prior knowledge (e.g. historical related data and/or expert advice) and data (Carlin *et al.* 2006; Sy *et al.*, 2015). Upon such combination a new state-of-knowledge characterized by a probability distribution called *posterior* is then obtained (Parent and Bernier, 2007; Sy *et al.*, 2015). The power of HBMA comes from its ability to accommodate complexity. By applying conditional distributions and establishing conditional relationship between various components of models (processes, parameters and data) it allows inferences to be made regarding errors in variables, random effects, hidden variables, and multiple data sets at different scales Clark (2005). HBMA modelling is now getting broad scientific recognition with increasing application in different fields within environmental sciences (Carlin *et al.* 2006).

In radioecology, this approach has been applied to model (i) transfers in forest ecosystems by Linkov *et al.* (1999), (ii) transfer to wild plants and animals (Hosseini *et al.* 2013), (iii) sorption-desorption in soils by Nicoulaud-Gouin *et al.* (2014; in revision), and (iv) foliar interception in dry conditions (Sy *et al.* 2015). The present study is a continuation of the latter one applied to the foliar pathway, which is crucial in post-accidental conditions. Following the initial deposition of radionuclides onto plant surfaces, environmental removal processes combined with radioactive decay will reduce the quantity of initial contamination on the vegetation surface (Chamberlain, 1970). The term “field loss” was originally introduced by Russell (1965) referring to the different environmental processes and the mechanical actions leading to the loss of radionuclides from the external and the internal compartments of plants. This “field loss” process is generally referred as to weathering in the literature. With respect to TRS 472 (IAEA, 2009), the “field loss” term has been rephrased as “weathering loss”

in this study. Factors such as the growth of plants and the weathering of deposited materials due to wind, rain and dieback of leaves (or their surface layers) are identified as influencing the rate at which radioactive substances are removed from plants surfaces. Indeed it was suggested by Miller and Hoffman (1983) that the wind action may lead to a rapid initial loss of material immediately after deposition since particles may be shaken off of vegetation by wind. In addition, precipitation, fog, dew and mist are considered to remove radionuclides from the surface and interior of the plant by wash-off and leaching processes, respectively. Moreover, the growth of the plant which can be represented, up to maturity, by an exponential process is assumed to induce an exponential decrease of plant contamination representing a biological dilution effect (Miller and Hoffman, 1983).

The present study aims at exploiting the features of HBMA to characterize uncertainty about this weathering loss parameter. The use of the uncertainty distributions, obtained by this approach as well as those resulting from earlier work on foliar interception (Sy *et al.* 2015), is illustrated using a case study in which human exposure through the food chain is assessed for the first year after two theoretical accidental scenarios defined in the EC project COMET (Liland *et al.* 2013).

2 Hierarchical Bayesian modelling of the weathering loss parameter

2.1 Weathering loss data

The weathering parameter is generally expressed as a weathering rate described as follows

$$\lambda^W = \frac{\ln 2}{T^W} \quad (1)$$

T^W is the weathering half-life – i.e. the time necessary for one-half of the radioactivity to be removed by environmental processes (as described in the previous section).

In the literature, λ^W is estimated by regression of empirical datasets in two ways: (i) Concentrations per unit mass of vegetation (Bq.kg^{-1}) are plotted as a function of time (in days). In such a case the biological dilution is implicitly included in the weathering loss value, and the weathering loss parameter (λ^W) is denoted WL in this article (i). Alternatively the concentrations per unit of ground area (Bq.m^{-2}) are plotted as a function of time (in days), and in such a case the biological dilution is excluded from the “weathering loss” and the

weathering loss parameter (λ^W) is denoted WL^* in this article (ii).

For the present work, the weathering half-life database compiled in connection with the NKS-PARDNOR project (Andersson et al., 2011), has been expanded to include more publications from the 1960s till 2014 considering weathering loss of any radionuclide in the following crop species: herbaceous vegetation, leafy and non-leafy vegetables and pasture grass. The major part of these data was obtained in natural field conditions including mostly post-Chernobyl and nuclear weapons testing measurements, and the rest of the data were from laboratory studies. Thus, 317 weathering half-life values weathering loss were collated. This set of values was divided in subsets according to radionuclide (iodine versus other nuclides) and to the source of data (laboratory or field). Table 1 summarizes the main features of the compiled data. The detailed databases are supplied as Excel spreadsheets in supplementary materials (S-1, S-2 and S-3).

2.2 Data modelling

The weathering loss values acquired in the database were analyzed with different Bayesian models. In a first approach, one single distribution is assumed for all weathering rates λ^W (i.e. excluding the biological dilution component) whatever the radionuclides and plants and another distribution is assumed for biological dilution of each plant (whatever the radionuclide). This approach is called “simple” as opposed to the “hierarchical” approach in which a new level of complexity is introduced in simple Bayesian models, to account for between-plants and/or between-radionuclides variability.

2.2.1 Simple Bayesian modelling

Let WL_{ijk} (in day^{-1}) a weathering loss value including the biological dilution (i.e. on the basis of mass of vegetation) observed for the plant species $i = 1, \dots, I$, the radionuclide $j = 1, \dots, J$ and the replicate $k = 1, \dots, n_{ij}$. A Log-normal error model was defined so that the logarithm of the weathering loss values are assumed to be identically and independently distributed (i.i.d) realizations of a Gaussian distribution with unknown variance and mean parameter.

$$\log WL_{ijk} = \mu_{ijk} + \varepsilon_{ijk} \quad (2)$$

$$\mu_{ijk} = \log (\lambda^W + I(\text{Measure Basis} (WL_{ijk}) = \text{per unit mass}) \times \lambda_i^{BD}) \quad (3)$$

$$\varepsilon_{ijk} \sim \text{i.i.d } N(0, \sigma^2) \quad (4)$$

Similarly, let WL_{ijk}^* (in day^{-1}) a weathering loss value excluding the biological dilution (i.e. on the basis of area) observed for the plant species $i = 1, \dots, I$, the radionuclide $j = 1, \dots, J$ and the replicate $k = 1, \dots, n_{ij}$.

$$\log WL_{ijk}^* = \mu_{ijk}^* + \varepsilon_{ijk}^* \quad (5)$$

$$\mu_{ijk}^* = \log \lambda^W \quad (6)$$

$$\varepsilon_{ijk}^* \sim \text{i.i.d } N(0, \sigma^{*2}) \quad (7)$$

The expression $I(\text{Measure Basis} (i) = \text{per unit mass})$ denotes the indicator function which specifies the measurement basis used to determine the weathering loss value i.e. the value is equal to 1 if the weathering loss value was derived on a per unit mass basis and 0 otherwise. The unknown quantities μ_{ijk} and μ_{ijk}^* define the expected values of $\log WL_{ijk}$ and $\log WL_{ijk}^*$. The parameter λ^W represents the average rate (in day^{-1}) at which the deposited radioactive materials are removed from plants surface due to the combination of environmental processes such as wind, rain and dieback of plants surface. This parameter is here assumed to be similar for all considered plant species and radionuclides whatever their characteristics, describing the uncertainty induced by an average behavior which is common to all plants and all particles. The unknown quantity λ_i^{BD} defines the rate (expressed also in day^{-1}) at which the concentration of materials in a plant is reduced due to its biological growth over time. Herein, this parameter refers to an environmental process which is assumed to be plant species-specific and totally independent of the characteristics of the particles. Given the two types of measurements, λ_i^{BD} is a latent variable, only considered in modelling the weathering loss data obtained in field conditions to take into account the biological dilution effect and its uncertainty. The quantities ε_{ijk} and ε_{ijk}^* represent stochastic variables which denote the discrepancy between the logarithm of the expected weathering loss value and the logarithm of the observed weathering loss measured in field and lab conditions. This discrepancy may arise, for instance, from the variability that can be observed between two plants of the same species, from two aerosols with different sizes particles and

from the uncertainty linked to modelling errors. The parameters σ^2 and σ^{*2} denote the unexplained residual variability. The two simple Bayesian structures result in three models denoted MI_I for iodine, MPI_I (field conditions) and MPL_I (laboratory conditions) for the particles.

The Bayesian approach, within this parametric framework, consists in inferring the posterior probability distribution of the vector of unknown parameters (λ^W , λ_i^{BD} , σ^2 and σ^{*2}), namely θ given the observed data $p(\theta|WL)$ (or $p(\theta|WL^*)$) by the means of the Bayes' rule which combines the likelihood $p(WL|\theta)$ (or $p(WL^*|\theta)$) and the joint prior distribution $p(\theta)$. The likelihood function expresses, how probable the current data are under the assumption that they are generated by a model containing parameters θ (Clark 2005), whereas the prior and the posterior describes the current and the updated (given the observations) state-of-knowledge on θ , respectively. Thus, the Bayesian modelling framework, through employing conditional relationships between observed and unobserved stochastic quantities, can be seen as a two-entity-modelling structure providing a powerful tool for updating data and describing uncertainty:

$$p(\text{parameter}|\text{data}) \propto p(\text{data}|\text{parameter}) \times p(\text{parameter})$$

This formularization means that the expressions on both sides of the symbol \propto are proportional.

2.2.2 From simple to hierarchical Bayesian structures

The essence of the hierarchical modelling is based on the development of a multilevel stochastic specification which captures the important components to explain the behavior of a complex process and thus accommodates complexity via the dissection of the problem into levels (Carlin *et al.* 2004; Clarke, 2005). Berliner (1996) proposed an extended decomposition of the problem in three entities including stochastic elements such as parameters and latent variables as expressed below.

$$\begin{aligned} p(\text{parameter}|\text{process}, \text{data}) & \\ & \propto p(\text{data}|\text{process}, \text{parameter}) \\ & \times p(\text{process}|\text{parameter}) \\ & \times p(\text{parameter}) \end{aligned}$$

The notion of 'hierarchy' then arises because the process model describing the processes sits "above" the model for the data (Jackman, 2009), and

consequently constitutes the second entity of the hierarchical models as identified by Berliner (1996). In this study, the concept of 'hierarchy' is introduced by the considerations related to the variations in the average weathering loss values and their dispersion which may be expected depending on the different plant species, the different radionuclides or the physical forms of the considered substances. We propose the following hierarchical model which captures simultaneously the species-specific variability and the variability induced by the different radionuclides (or types of radionuclides) of concern.

$$\log WL_{ijk} = \mu_{ijk} + \varepsilon_{ijk} \quad (8)$$

$$\mu_{ijk} = \log (\lambda_{ij}^W + I(\text{Measure Basis } (WL_{ijk}) = \text{per unit mass}) \times \lambda_i^{BD}) \quad (9)$$

$$\log WL_{ijk}^* = \mu_{ijk}^* + \varepsilon_{ijk}^* \quad (10)$$

$$\mu_{ijk}^* = \log \lambda_{ij}^W \quad (11)$$

$$\lambda_{ij}^W \sim \text{i.i.d. } \text{LogN}(\mu_{\lambda^w}, \sigma_{\lambda^w}^2) \quad (12)$$

$$\varepsilon_{ijk} \sim \text{i.i.d. } N(0, \sigma^2); \varepsilon_{ijk}^* \sim \text{i.i.d. } N(0, \sigma^{*2}) \quad (13)$$

Definitions associated to the unknowns μ_{ijk} , μ_{ijk}^* , λ_i^{BD} , ε_{ijk} , ε_{ijk}^* , σ^2 and σ^{*2} were provided in the previous sections. The latent variable λ_{ij}^W still denotes the average weathering rate at which the deposited radioactive material $j = 1, \dots, J$ is removed from the surface of the plant species $i = 1, \dots, I$. It was assumed a simultaneous dependence of λ_{ij}^W on the plant species and to the radionuclides (or the physical form of the radionuclide for the radioiodine case). Each combination between one plant species and one radioactive material is then supposed to have its own weathering loss value which describes the specific interaction or affinity between the plant surface characteristics and the nature of the considered element. However, this interaction is assumed to be a common feature of all the plant-radionuclide pairs considered and hence the weathering rate is assumed to stem from a common log-normal distribution with parameters μ_{λ^w} and σ_{λ^w} . Thus, variations between plant species - radionuclides couples are given by σ_{λ^w} .

Making use of this hierarchical structure, three main hypotheses were investigated with respect to the presence and absence of interactions between radionuclides and plant species and their impacts on weathering rate. On one hand the environmental loss constant is assumed to depend only on the plant species peculiarities and on the other hand it is assumed to vary depending only on the

characteristics of the radionuclides of interest (or physical form for radioiodine). Each of these hypotheses resulted in three competing hierarchical models for each dataset: MI_2 to MI_4 for radioiodine, MPF_2 to MPF_4 and MPL_2 to MPL_4 for particulate radionuclide studied in field and in lab conditions respectively (see Table 2).

The hierarchical Bayesian approach, in its global application in many fields and specifically in this context of radioecology, represents an intermediate strategy between (simple Bayesian models) that do not take into account the assumption of variability among defined groups, and highly parameterized models that generally assign an estimate of parameters to each plant species and/or radionuclides (or type of radionuclides) (Jiao *et al.* 2009). This approach presents a strategy to avoid under-fitting and / over-fitting data that we are trying to model by enabling to accommodate enough parameters to fit the data properly while making use of a population distribution to structure dependencies into the parameters. It can be seen as a nested hierarchy of stochastic relations for which the challenging issue is to estimate the latent variables located in the middle layers of the structure.

2.2.3 Bayesian inference and posterior computations

Parameters of the twelve resulting models (MI_m, MPF_m and $MPL_m : m = 1, \dots, 4$) were inferred under the Bayesian framework. Simple and hierarchical Bayesian models are generally seen as networks in which the unknown quantities are represented by nodes and the conditional distributions define the relationships between the nodes: terminal nodes are called parameters and nodes with parents are called latent variables (Sy *et al.* 2015).

One of the requirements in Bayesian inference concerns the definition of prior distributions linked to the parameters. The choice of the prior distributions is very vital in Bayesian analysis as posterior distribution is seen as a compromise between the prior and the observed data (Gelman *et al.* 2004). In addition, this choice is delicate since the prior distributions express the modeler's uncertainty on the unknown quantities and enable to combine multiple sources of information. The main difficulty is then based on the definition of probability distributions which do not carry high intrinsic information within the analysis. In this study, the adopted strategy was mainly to assign probability distributions with low intrinsic

information (the so-called weakly informative priors) to the models parameters. Indeed, inverse gamma distributions with both scale and intensity parameters fixed at 0.001 ($IG(0.001, 0.001)$) were assigned to the variance parameters ($\sigma_{\lambda^w}^2, \sigma^2$ and σ^{*2}) of the twelve models (MI_m, MPF_m and $MPL_m : m = 1, \dots, 4$) following (Gelman 2006). Log-normal distributions ($LN(0, 10^4)$) which are supposed to carry low intrinsic information were defined for models MI_1, MPF_1 and MPL_1 . A normal distribution with low intrinsic information with mean equal to 0 and high variance (10^4) was assigned to the parameter μ_{λ^w} for all concerned models (MI_m, MPF_m and $MPL_m : m = 2, \dots, 4$). Different flat uniform prior distributions were assigned to the biological dilution rates (unknown quantities $\lambda_i^{BD} : i = 1, \dots, I$) for each associated plant species. The lower and upper bounds of these uniform prior distributions were defined following the values suggested by (Calmet *et al.* 2001). In contrast to all other selected priors, which generally provide very little information on the considered subject, the latter ones (i.e. uniform priors) are more informative and reflect the current available knowledge regarding the biological dilution rates.

The computation of the marginal posterior distributions was performed within the environment of R 3.1.1 (R Core Team, 2014) by using Markov Chain Monte Carlo simulation method as developed in the software JAGS 3.3.0 (Plummer, 2003). Three independent chains starting from three random positions were computed for each unknown parameter for each model. A burn-in-period of $5 \cdot 10^4$ iterations was defined to obtain stationary chains and 10^5 additional runs were performed after the burn-in-period. A thin was fixed at 20 to reduce autocorrelation in the three chains meaning that only one run was kept every 20 iterations. The convergence of the chains was checked by graphically visualizing whether the three chains had mixed well, and by computing the Gelman-Rubin convergence diagnostic for each parameter. For each parameter, this convergence diagnostic is based on the assessment of the within-chain and the between-chains variances (Brooks and Gelman, 1998). For each model, inferences and empirical quantities (such as the posterior mean, median, standard deviation and percentiles) associated to each unknown quantity were derived from the posterior samples which are based on the pool of three thinned samples for a total sample size $N = 1.5 \cdot 10^4$

(5000 x 3 chains). The posterior results are displayed in Table 6.

Regarding the three datasets of concern, the corresponding models were compared on the basis of the Deviance Information Criterion (DIC) values calculated for each competing model to assess which model fits the weathering loss data best. DIC is a penalized likelihood criteria used in Bayesian analysis which had been introduced by Spiegelhalter *et al.* (2002) to compare models by checking their goodness-of-fit (characterized by the posterior mean of the deviance) and their parsimonious characteristics, i.e. inclusion of few parameters, (described by an estimate of the effective number of parameters, pD). DIC is given by the expression $DIC = Dbar + pD$. Based on the DIC rule the model with the smallest DIC value is the best with respect to the fitting performances (Table 5).

3 Propagation of uncertainty in post-accidental scenarios

3.1 Scenarios definition and specification of parametric uncertainty

A “standardized dry deposition scenario” was specified in the COMET project (Liland *et al.* 2013): the deposition is modelled to have occurred on August 1st 2013, assuming 1 kBq.m⁻² of each of four radionuclides: Cs-134, Cs-137, Sr-90 and I-131. Calculations were performed with the modeling platform SYMBIOSE (Gonze *et al.* 2011, Simon-Cornu *et al.* 2015), for the first year after the accident. Transfer to 9 foodstuffs, (fresh and canned) leafy vegetables; flour; beef, pork and sheep meat; (fresh and canned) cow milk, and goat milk were considered. Effective doses (abbreviated ED hereafter) for the 4 radionuclides taken separately were also calculated for 2 age groups of great interest in risk assessment; Adult (≥ 17 years) and Young Child (2-7years). ED values are calculated following the equation below.

$$ED_r^i(t) = ED_r^i(t - 1) + \sum_{j=1} C_r^j(t) \times IR_j^i \times DC_r^i \quad (14)$$

Where $ED_r^i(t)$ is the effective dose (cumulated over the time and expressed in Sievert (Sv)) of the radioactive substance $r \in \{Cs-134, Cs-137, I-131 \text{ and } Sr-90\}$ received at the date t by an individual belonging to the age group $i = 1, 2$. The quantity $C_r^j(t)$ is the

concentration (in Bq/kg fw) of the radioactive substance r predicted in the food item $j = 1, \dots, 9$. The quantities IR_j^i and DC_r^i denote the food ingestion rate (the quantity of food consumed per unit time usually expressed in kg per day) and the effective dose coefficient (Sv/Bq) which converts the intake activity to the effective dose, respectively. The values of both quantities were defined according to SYMBIOSE (see above) default values which are based on French food consumption standard for IR_j^i (Favier *et al.* 1995; Bertrand, 1993) and on International consensus for DC_r^i (ICRP, 1996).

The fixed values and the probability distribution functions of the considered parameters are defined in Tables 3, 4 and Table S-4 of the supplementary materials. The uncertainty distributions for the weathering loss parameter were derived from the previous part of this study. The uncertainty distributions assigned to the dry interception fractions were derived following the HBMA model developed by Sy *et al.* (2015) to produce posterior distributions of dry interception fractions of Cs-134, Cs-137, Sr-90 and I-131 by leafy vegetables and pasture grass (used for domestic animals feeding). The considered HBMA model is described in annex A. The uncertainty distributions assigned to root transfer parameters and to the parameters relative to the transfer to animal products were defined by Simon-Cornu *et al.* (2015) on the basis of the TRS 472 values (IAEA, 2009).

3.2 Numerical simulations

Ordinary differential equations were solved with a time step of 6 hours (Euler solver). The uncertainty of the outputs (concentrations in the foodstuffs and effective doses) was characterized by Monte Carlo simulation (1000 iterations). A second Monte Carlo simulation (1000 iterations as well) was run to check convergence.

Uncertainty in the outputs (concentration values and the effective doses) was quantified using the uncertainty ratio (UR), i.e. the ratio between the 95th percentile and the median of the distribution of concentration values and of the distribution of effective doses. Spearman correlation coefficients were calculated to assess sensitivity of model outputs to changes in the input parameters.

4 Results and discussion

4.1 Posterior results of the HBMA on the weathering loss parameter

The posterior distributions obtained by performing the Bayesian inference of the unknown parameters of the models MI_1 to MI_4 , MPF_1 to MPF_4 and MPL_1 to MPL_4 are summarized in Table 6, Table 7 and Table 8, respectively.

4.1.1 Simple structures

The posterior mean of the average weathering rate is estimated to be about 0.07 day^{-1} with a 95% credible interval (CI) of (0.062 ; 0.078) for model MI_1 and at approximately 0.06 day^{-1} (0.05, 0.07) and 0.04 (0.035, 0.045) day^{-1} for model MPF_1 and MPL_1 , respectively. These 95% CIs correspond to environmental half-life ranging from approximately 9 to 11 days for iodine, from 10 to 14 days for “field” particles and from 15 to 20 days for “laboratory” particles. These differences suggest that iodine in field conditions is removed from plants foliage faster than particles in such conditions. This observation is consistent with the conclusions drawn from the earlier reviews (Kirchner, 1994; Miller and Hoffman, 1983; Pröhl and Hoffman, 1996). However, these ranges are very similar since the 95% CIs overlap and do not demonstrate a large variability between these elements. In addition, dissimilarities are observed between “field” particles and laboratory particles as particles in laboratory conditions are removed slower than particles in natural field conditions. This shows the difference between the field and the laboratory conditions and the impact of the environmental conditions on the fate of elements in the ecosystems.

The posterior mean values of the biological dilution rate obtained for herbaceous vegetation are generally higher than the posterior means estimated for pasture grass for both iodine and particles studied in field conditions. Indeed, the mean values are estimated for iodine at approximately 0.04 day^{-1} (0.016, 0.086) for herbaceous vegetation and 0.026 day^{-1} (0.015, 0.048) for pasture and for particles at approximately 0.024 day^{-1} (0.015, 0.049) for herbaceous vegetation and 0.019 day^{-1} (0.015, 0.03) for pasture grass. The differences between the values which relate to herbaceous vegetation and to pasture grass reflect the variability, independently from the considered elements, which can be expected in natural field conditions between plant species such as herbaceous vegetation and pasture grass. Such variability is also observed regarding the uncertainty on the biological dilution rates, which higher for herbaceous vegetation.

The posterior means related to the unexplained residual variability σ (or σ^* for MPL_1) are relatively

high for the three models MI_1 , MPF_1 and MPL_1 and are estimated at approximately 0.56 day^{-1} (0.5, 0.64), 0.68 day^{-1} (0.61, 0.77) and 0.51 day^{-1} (0.43, 0.61), respectively. These values are not comparable since the models were fitted to different datasets (see Section 2) but a certain consistency is observed when comparing the three models. Posterior values obtained for laboratory particles are lower; this can be explained by the fact that MPL_1 is less affected than MI_1 and MPF_1 by the uncertainty associated with the biological dilution rate.

Results drawn from the inference of the simple Bayesian model structures give preliminary insights on the differences that can be observed between iodine and the other particles, between particles studied in natural field conditions and those studied in laboratory conditions and between plant species with respect to the biological growth of plants.

4.1.2 Hierarchical models

The hierarchical framework incorporates a level of complexity which enables studying the specific variations across elements, their physical forms and the different considered plant species.

For iodine, the posterior means and the posterior medians estimated for the average weathering rates for models MI_2 to MI_4 , which range from 0.05 to 0.075 day^{-1} , are generally similar to the posterior mean and median estimated for MI_1 (approximately 0.07 day^{-1}). This corresponds to environmental half-life values varying between 9 and 14 days, which are consistent with values generally found in the literature (IAEA, 1996; Miller and Hoffman, 1983). The average weathering rates estimated for particulate iodine are found to be slightly higher, but very similar, to those estimated for gaseous iodine, suggesting similar removal behavior between both physical forms. The “unknown” physical form shows lower values of environmental half-life, approximately 9 days. The unknown iodine values correspond to environmental measurements performed after Chernobyl fallout (see Table S-1 in supplementary materials) whereas the values of particulate and gaseous iodine mainly correspond to studies implemented in controlled conditions (see Tables S-1 and S-2 in supplementary materials). The environmental conditions occurring during the experimentations might be an explanation of such slower removal but there is no clear evidence supporting this. The posterior means of the average weathering rates of pasture grass are slightly higher than values obtained for herbaceous vegetation. The removal behavior can be considered similar between

pasture grass and herbaceous vegetation, so the observed difference might be mainly due to the measurements obtained after the Chernobyl fallouts. The variability which is induced by the combined effect of plant species and the physical forms of iodine are found to be not significant. The posterior distributions obtained for the parameters and the latent variables of model MI₄ are rather similar between the different sub-groups.

For particles studied in laboratory conditions consistency is found between the models MPL₁ to MPL₄ regarding the posterior means and medians obtained for the average weathering rates oscillating around 0.04 day⁻¹. The variability induced by the effect of the interactions between plant species and the elements does not appear significant. However, dissimilarities are found between the considered radionuclides; the posterior means of the average weathering rates are estimated at approximately 0.03 day⁻¹ for cesium and ruthenium, 0.04 day⁻¹ for cerium, cobalt and manganese and at 0.054 day⁻¹ for strontium. One could speculate whether differences between radionuclides could be due to element specific solubility properties and / or different degrees of absorption onto plant surfaces (negatively charged) or possibly foliar uptake (IAEA, 2009). Slight dissimilarities are found between plant species with posterior means of the average weathering rates estimated at 0.04 day⁻¹. For particles studied in field conditions, dissimilarities exist between MPF₁ and models MPF₂ to MPF₄. These dissimilarities are mainly due to the data related to rubidium in non-leafy vegetables for which weathering rate values range from approximately 0.09 to 0.42 day⁻¹. The posterior means of the average weathering rate estimated for leafy vegetables are generally higher than those estimated for herbaceous vegetation and pasture grass suggesting the importance of the species specific-variability on the environmental removal process. Similarly for the models MPL₂ to MPL₄, an element specific-variability characterized by differences of posterior weathering rate values is observed. The interactions between the plant species and the elements do not appear to be significant since, for a given element, only slight differences are observed between the different plant species.

The posterior results associated to biological dilution rates obtained for models MI₂ to MI₄ and MPF₂ to MPF₄ are very similar to those obtained for MI₁ and MPF₁ respectively. Higher values of biological dilution rates are found for herbaceous vegetation compared to values obtained for pasture

grass reflecting the species specific-variability in the environment.

The posterior means related to the unexplained residual variability σ for models MI₂ to MI₄ are similar to those obtained for MI₁ and are estimated to be about 0.56. The posterior means obtained for models MPF₂ to MPF₄ range from 0.58 to 0.62. These values are lower than the posterior mean estimated for MPF₁ suggesting a better specification of the overall variability. The posterior mean of the unexplained residual variability estimated for MPL₃ is similar to the posterior mean estimated for MPL₁ (approximately 0.51) whereas the values obtained for MPL₂ and MPL₄ are lower, approximately 0.44. It seems that MPL₂ and MPL₄ models which differ from MPL₁ and MPL₃ due to the characterization of the element-specific variability, provide a better specification of the overall variability.

4.2 Model Selection

The models MI₁ to MI₄, MPF₁ to MPF₄ and MPL₁ to MPL₄ were compared based on the DIC values summarized in Table 5. The posterior results presented in the previous section have highlighted the influence of the plant species on the environmental weathering rate and on the biological dilution rate for iodine. In addition, both for particles studied in field conditions and those studied under laboratory conditions, the element-specific variability on the environmental removal process of such particulate substances is underlined. These preliminary observations are confirmed by the DIC values. Indeed, models MI₃, MPF₂ and MPL₂ offers the best fitting performances compared to the other models in each group. For iodine, this means that the characterization of the plant species-specific variability when modelling the environmental removal process should be taken into account as an important source of variations. In contrast, the characterization of the variability induced by the various physical forms of iodine is of lesser interest. Some precautions are required since the difference between the DIC values of models MI₁ to MI₄ is not significant; values varying from -402.7 to -401.5. Nevertheless, the values of the posterior distribution of the parameter $\sigma_{\lambda w}$, characterizing inter-species variability, tend to confirm that species-specific variability is more influential on the variations of the weathering rate and by definition on the variations of the weathering loss. For field particles and particles studied in laboratory conditions, the DIC values and the posterior distributions of the

parameter $\sigma_{\lambda w}$ estimated for MPF₂ to MPF₄ and MPL₂ to MPL₄ confirm that element-specific variability is of greater interest than the species-specific variability on the environmental removal process.

Figures 1-3 illustrate the prior and the posterior distributions of the parameters of the models MI₃, MPF₂ and MPL₂ respectively. The considered weathering loss data contribute to strongly update the prior belief on the parameters of MI₃, MPF₂ and MPL₂. For the parameters $\mu_{\lambda w}$, $\sigma_{\lambda w}$ and σ of these three models, the strategy was to assign prior distribution functions which carry very low intrinsic information since no precise information is available in the literature about the range of plausible values for such parameters. For the biological dilution rate parameters λ_1^{BD} and λ_2^{BD} of models MI₃ and MPF₂, the up-to-date uncertainty characterized by the posterior shows the impact of the collated data on the prior belief. Indeed, the central tendency of the prior distribution is located at approximately 0.058 day⁻¹ whereas the central tendencies of the posteriors are estimated at approximately 0.04 and 0.03 day⁻¹ for MI₃ and approximately 0.024 and 0.019 day⁻¹ for MPF₂.

In the following section posterior distribution of plausible weathering loss values are predicted using the models MI₃ for ¹³¹I and MPF₂ for cesium isotopes (¹³⁴Cs and ¹³⁷Cs) and ⁹⁰Sr.

4.3 Propagation of uncertainty: standardized scenario

In the next sections the results obtained from the Monte Carlo simulations concerning the contamination of 9 food items by Cs-134/-137, I-131 and Sr-90 following a standardized dry deposition scenario are presented and discussed. For sake of clarity, in regard to the large amounts of results, only the concentrations of Cs-134/-137, I-131 and Sr-90 in beef meat, fresh cow milk and leafy vegetables are discussed in detail.

4.3.1 Uncertainty analysis

Table 9 summarizes the concentrations of Cs-134/-137, I-131 and Sr-90 predicted in beef meat, fresh cow milk and leafy vegetables at approximately 1 month, 3 months and 1 year after the deposition date and Figure 4 illustrates the dynamics of Cs-134/-137, I-131 and Sr-90 in beef meat, in fresh cow milk and in leafy vegetables. Approximately one month after the deposition date, the highest concentrations of Cs-134/-137 are found in beef meat, whereas I-

131 and Sr-90 are mainly present in leafy vegetables. Three months after the deposition date, the highest values of cesium, iodine and strontium concentrations are observed in fresh cow milk whereas one year after highest concentrations of cesium and strontium are observed in beef meat. The uncertainty on the concentrations in leafy vegetables is somewhat constant across the considered radionuclides the first month after the deposition with URs varying from 3.3 to 3.6. The uncertainty about the concentrations in beef meat and cow milk varies more at this period across the radionuclides with URs ranging from 4.4 to 6.25 for beef meat and from 5.4 to 9.75 for fresh cow milk. This is mainly explained by the way each element behaves within beef and cow metabolism (IAEA, 2009). Over time, the uncertainty about the concentration in beef meat does not vary much, whereas the uncertainty about the concentration in cow milk has decreased three months after the deposition, and has increased one year after the deposition. These fluctuations, explained by seasonal variations of feeding conditions (i.e. indoors with contaminated or not contaminated forage, outdoors with less and less contaminated grass.), are more immediate for milk and smoother for meat (integration over time). For leafy vegetables, the uncertainty about the concentrations of cesium, iodine and strontium has increased strongly due to the uncertainty on the slope of the concentrations curves (the weathering loss parameter).

Table 10 summarizes the distribution of the effective doses received by an adult and a young child predicted for cesium, iodine and strontium. The dynamics of the effective dose of ¹³⁷Cs, ¹³¹I and ⁹⁰Sr received by an adult and a young child are illustrated in Figure 5. During the first month following the deposition, the effective doses of cesium isotopes received by an adult are found to be higher than the doses received by a young child, whereas the effective doses of ¹³¹I and ⁹⁰Sr are higher for young child. The food dietary habits (and specifically the different ingestion rates), and life stage as young child has higher growth rate can explain such differences. However, UR values calculated for adult and young child are found to be very similar for each given radionuclide. This reflects that the uncertainty on the concentration values is not significant to explain the inter-individual variability of ingested doses since the uncertainty and the variability on parameters relating to food dietary habits are not taken into account (fixed values are used for these parameters in this study see Table S-4 of supplementary materials). The uncertainty on the

effective doses of cesium and strontium predicted during this period is found to be very slightly variable across these radionuclides with UR values estimated at approximately 2.2 for cesium isotopes and at approximately 1.9 for strontium. For iodine a difference is observed with URs at approximately 3.3. The uncertainty on the effective doses is found to be rather constant over the time since one year after the deposition the variation of UR values are not very high: variations of approximately 13%, 3% and 4% are observed on UR values calculated for cesium, iodine and strontium one year after the deposition date. Such observation is mainly explained by the fact that most of the effective dose is received during the first month following the deposition date.

One important point in food risk analysis is to assess the influence of each food item or combination of food items on the human population exposure to derive typical food exposure patterns (Kroes *et al.* 2002; Sy *et al.* 2013). Figure 6 gives an illustration of the contribution of each food item on the predicted average effective dose of ^{137}Cs , ^{131}I and ^{90}Sr received by an adult and a young child. The food items with contribution equal or close to zero are not visible. The fresh cow milk, beef meat and leafy vegetables are the food items which contribute the most on effective doses of ^{137}Cs received by an adult and a young child one and three months after the deposition date. One year after the deposition date, processed products such as canned leafy vegetables and canned milk become more influential on the effective dose of ^{137}Cs as expected. The assessment of the contribution of each food item gives a first vision on the influential food pattern and indirectly the potential influence of the environmental process on the dose received by human population. In the next section, the impact of the radioecological processes is assessed more precisely.

4.3.2 Sensitivity analysis

Tables 11 and 12 display the Spearman correlation coefficients (SCC) calculated between the concentrations of cesium, iodine and strontium in the food items and the considered uncertain input parameters and between the effective doses and these uncertain input parameters respectively (see Table 3). Only the most influential parameters with respect to SCC values are presented.

One month after the deposition the most influential parameters on the uncertainty about the concentration of cesium, iodine and strontium in

beef meat and cow milk are the foliar biomass of grass, the equilibrium transfer factor from feed to product (milk/meat), the dry interception fraction of grass and the weathering loss rate associated to grass. The concentrations of the different elements in leafy vegetables are mainly influenced by the weathering loss parameter and the dry interception associated to leafy vegetables. Knowing that the foliar biomass of leafy vegetables is included, by definition within the platform SYMBIOSE (in the parameterization of the dry mass interception fraction) this parameter is also influential on the uncertainty about the concentrations in leafy vegetables. Three months and 1 year after the deposition, the dry interception fraction and the foliar biomass of grass and leafy vegetables tend to be less influential; this reflects the influence of the foliar pathway during the early phase following the deposition (i.e. the first few months). The influence of the weathering loss parameter for both leafy vegetables and animal products (beef meat and cow milk), and the influence of the equilibrium transfer factor on the animal product concentrations, are found to be constant during all the period of interest (i.e. the year following the deposition).

For the effective doses, the foliar biomass of grass and leafy vegetables, the dry interception, the weathering loss parameter and the equilibrium transfer factor to animal products are the most influential parameters. These observations confirm the influence of the foliar pathway on the effective doses as previously suggested by SCC values calculated for the concentrations of cesium, iodine and strontium in food items and by the contribution of each food item on the predicted effective doses. Indeed, as previously mentioned, most of the effective dose coincides with the dose received during the first month following the deposition date. In addition, as the uncertainty on the food ingestion rates and on the other parameters driving the food dietary habits of each individual is not taken into account (fixed values are used see Table S-4 of the supplementary materials) the uncertainty found about the food concentrations is directly reported on the uncertainty on the effective doses. This propagation of the standardized dry scenario enables us to have a clear perception of the different radioecological parameters influencing on the levels and the (temporal) variations of cesium, iodine and strontium in food items and their subsequent influence on the dose received by individuals belonging to different age groups after the ingestion of contaminated foods.

5 Conclusions

In line with the research strategy suggested by the radioecological community (Hinton et al. 2013), the main objective of the present work was to contribute in improving risk assessment models using in radioecology. This was done by applying Hierarchical modelling under the Bayesian framework.

By using weathering loss as an example, this study has shown the advantage of adopting Bayesian Hierarchical Modelling to produce more robust estimates for model parameters. Indeed, the HBMA approach, applied on the three datasets of weathering loss values (for iodine, particles studied under field and lab conditions), has provided the possibility to define a unique framework which enables to simultaneously characterize the uncertainty on weathering loss and its connection with the variations between plant species, between radionuclides, and the variations between plant species - radionuclides couples. In addition, this approach enables taking into account, in an explicit way, the effect of the biological dilution. To compare models and assessing their fit, DIC method was applied to the constructed HBMA models. Furthermore, selected models were used to show differences between iodine, field particles and lab particles. The impact of plant species for iodine, plant species and radionuclides for field particles and radionuclides for lab particles (because no dilution due to biological growth was considered) on the weathering loss was demonstrated.

Making use of the HBMA models developed for the weathering loss in this study and for the dry interception fraction developed by Sy *et al.* (2015), the impact of the uncertainty/variability about these radioecological parameters on effective dose predictions was evaluated within a defined propagation scenario (dry deposition of Cs-134, Cs-137, Sr-90 and I-131, 1 kBq.m⁻²). Both parameters - weathering loss and dry interception fraction – were found to be influential on the effective doses the first year following the deposition for both adult and young child.

With respect to the research lines 3 and 4 outlined by (Hinton et al. 2013) and listed in the introduction, HBMA can be an interesting modelling approach to derive robust radioecological (and also dietary-related) parameters values depending on the different spatial (regional) and temporal (seasonal) patterns. Such regionalization of radioecological parameters can be, for instance, implemented by

incorporating the spatial variability of some variables (e.g. the aboveground biomass values) as extra levels in hierarchical models. As an example, this alternative can be implemented in a straightforward way using the HBMA models developed by Sy *et al.* (2015) for the dry interception fraction.

Acknowledgements

This work was partially financed by the COMET project through the European Commission (Grant agreement n° 604974) and the Research Council of Norway (project number 230295/E40). The Region PACA (France) is acknowledged for financially supporting the PhD study of xxx. xxx is acknowledged for financially supporting the development of SYMBIOSE.

Annex

A Addressing uncertainty/variability on the dry interception fractions using HBMA model

The interception process following the initial deposition of (radioactive) substances on plant canopies represents the entrance of such substances in the food chain and is generally expressed as an interception fraction defining the proportion of deposited materials that is retained by vegetation (Pröhl, 2009). Two different parameterizations of the interception fraction are mainly used in the literature:

- The ratio of the radioactivity intercepted by plant leaves to the total radioactivity deposited on both ground and plant leaves (both expressed in Becquerel per square meter). This ratio, denoted IF in this article and expressed in (Bq.m⁻² per Bq.m⁻²), takes its value over the open interval (0,1).
- The normalization of IF by the standing aboveground biomass of plants denoted MIF (for mass interception fraction) and expressed in (kilogram of fresh weight per square meter of ground).

Let DIF_{ijk} be the dry interception fraction considered for the plant species $i = 1, \dots, I$, the radionuclide $j = 1, \dots, J$ and the replicate $k = 1, \dots, n_{ij}$. Let DIF_{ijk}^* the logit-transformation of the dry interception fraction DIF_{ijk} . The quantity DIF_{ijk}^* was assumed by (Sy *et al.* 2015) to be Gaussian distributed as described below.

$$DIF_{ijk}^* \sim i.i.d N(\mu_{ijk}^{DIF}, (\sigma^{DIF})^2) \quad (A-1)$$

$$\mu_{ijk}^{DIF*} = \text{logit} \left(A \times (1 - \exp^{-G_j \times B_{ijk}}) \right) \quad (A-2)$$

$$N_{LV}^{out} \sim \text{Poisson}(5); N_{LV}^{in} \sim \text{Poisson}(12) \quad (A-5)$$

$$q \sim \text{Beta}(20, 2) \quad (A-6)$$

$$W_{LV} \sim \text{Normal}(400, 50) \quad (A-7)$$

The quantity μ_{ijk}^{DIF*} denotes the expected value of the logit-transformed dry interception fraction. The parameter A corresponds to the asymptotic dry interception fraction which describes a saturation phenomenon related to high values of plants biomass. Sy *et al.* (2015) demonstrated that this asymptotic interception can be assumed to be similar to all plant species since plants peculiarities are not significant when plant leaves covers entirely the ground surface. The parameter G_j is the rate of convergence at which the expected value μ_{ijk}^{DIF*} reaches the dry asymptotic interception A . It was demonstrated that this rate of convergence is specific to the physical characteristics of the radionuclides of concern. More specifically, aerosols and gaseous iodine were found to have different rates of convergence.

The parameter σ^{DIF} represents the unknown variances describing the discrepancy between the observed values of the logit-transformed dry interception fraction and the expected values μ_{ijk}^{DIF*} . This discrepancy was assumed by (Sy *et al.* 2015) to be independent to the type of radionuclides of concern.

The quantity B_{ijk} is a covariate representing the values of the aboveground standing fresh biomass of the plant species $i = 1, \dots, I$ expressed in ($\text{kg} \cdot \text{m}^{-2}$, fw). For pasture grass, the dry aboveground biomass values were assumed to vary uniformly between 0.02 and 0.14 $\text{kg} \cdot \text{dw} \cdot \text{m}^{-2}$. This range of values was defined for an illustrative purpose with, in addition, the background idea to be roughly representative of a broad range of pastures usually seen in temperate environments (Peeters, 1993; Smit *et al.* 2008). Considering that the dry matter content values stem from a uniform distribution with bounds fixed at 0.1 and 0.33 (IAEA, 2010) the corresponding fresh biomass values vary between 0.06 and 1.37 $\text{kg} \cdot \text{m}^{-2}$, fw. Regarding leafy vegetables, the aboveground biomass values were derived using the following Poisson mixture model which was originally developed by Sy *et al.* (2015) to predict lettuce fresh biomass values.

$$B_{LV} = \frac{(N_{LV} \times W_{LV})}{1000} \quad (A-3)$$

$$N_{LV} = q \times N_{LV}^{out} + (1 - q) \times N_{LV}^{in} \quad (A-4)$$

Let B_{LV} be the aboveground fresh biomass calculated for leafy vegetables and expressed in $\text{kg} \cdot \text{fw} \cdot \text{m}^{-2}$. The parameters N_{LV} and W_{LV} denote the number of plants per square meter of ground and the weight of one plant (in grams fw) respectively. A Gaussian distribution with mean fixed at 400 grams and standard deviation at 50 grams was considered for the weight of one plant at harvest according to Sy *et al.* (2015). The number of plants per m^2 is assumed to stem from the mixture of two Poisson distributions characterizing the number of plants cropped outdoor and indoor with parameters fixed at 5 and 12 respectively. A beta distribution with shape parameters defined at 20 and 2 was assigned to the mixture ratio to reproduce preferentially the biomass values corresponding to outdoor cultures.

Within the propagation scenario defined in this article the posterior distributions obtained by Sy *et al.* (2015) using HBMA and assigned to the unknown parameters (A, G_j and σ^{DIF}) and the values and/or probability distributions assigned to the covariate (B_{ijk}) were combined to reproduce plausible values of dry interception fractions. The fixed values and the probability distributions functions are detailed in Table 3.

B Supplementary Materials

The detailed databases corresponding to Iodine, particles studied under field conditions and particles studied under laboratory conditions are supplied as the Excel spreadsheets S-1, S-2 and S-3 respectively. Tables S-4 and S-5 display the values of the parameters which relate to human food consumption (daily ingestion rate of food, food stocking delay and the proportion of food origin) and of the ingestion (effective) dose coefficients considered in this study respectively. Supplementary materials related to this article can be found at xxx.

References

Andersson, K.G., Nielsen, S.P., Thørring, H., Hansen, H.S., Joensen, H.P., Isaksson, M., Kostianen, E., Suolanen, V., Pálsson, S.E., 2011. Revision of deposition and weathering parameters for the ingestion dose module (ECOSYS) of the ARGOS and RODOS decision support systems. *J. Environ. Radioact.* 102, 1024-1031.

- Bayes, M., Price, M., 1763. An Essay towards Solving a Problem in the Doctrine of Chances. By the Late Rev. Mr. Bayes, F. R. S. Communicated by Mr. Price, in a Letter to John Canton, A. M. F. R. S. Philosophical Transactions 53, 370-418. 10.1098/rstl.1763.0053.
- Bengtsson S.B., Gårdenäs A.I., Eriksson J., Vinichuk M., Rosén K., 2014. Interception and retention of wet-deposited radiocesium and radiostrontium on a ley mixture of grass and clover. Science of the Total Environment, 497-498: 412-419.
- Bereford, N.A., Wood, M.D., Vives i Batlle, J., Yankovich, T.L., Bradshaw, C., Willey, N., 2015. Making the most of what we have: application of extrapolation approaches in radioecological wildlife transfer models. J. Environ. Radioact. 151, 373-386.
- Berliner, M.L., 1996. Hierarchical Bayesian time series models. In: Maximum Entropy and Bayesian Methods. Proceedings of the Fifteenth International Workshop on Maximum Entropy and Bayesian Methods. 79, 15-22.
- Bernardo, J.M., 1996. The concept of exchangeability and its applications. Far East J. Math. Sci. 4, 111-121.
- Bertrand M. (1993). Consommation et lieux d'achats des produits alimentaires en 1991, collection INSEE Résultats, Consommation-Modes de vie n°54-55.
- Black, S.C. Barth, D.S., 1976. Radioiodine prediction model for nuclear tests. US-EPA Environmental Monitoring and Support Laboratory. EPA-600/4-76-027.
- Brooks, S.P., Gelman, A., 1998. General methods for monitoring convergence of iterative simulations. J. Comput. Graph. Stat. 7 (4), 434-455.
- Brown, J.E., Beresford, N.A., Hosseini, A., 2013. Approaches to providing missing transfer parameter values in the ERICA Tool – How well do they work? J. Environ. Radioact. 126, 399-411.
- Calmet, D., Crabol, B., Germain, P., Fiévet, B., Boust, D., Bailly du Bois, P., Louvat, D., Renaud, Ph., Rommens, C., Vray, F., 2001. Identification et estimation des incertitudes associées aux évaluations de doses obtenues par le modèle mathématique du Groupe Radioécologie Nord Cotentin. IRSN Report DPRE/SERNAT/2001-13. Available at: http://www.gep-nucleaire.org/norcot/gepnc/sections/travauxgep/duxieme_mission/volume_2_distribution/view
- Carlin, B.P., Clark, J.S., Gelfand, A.E., 2006. Elements of hierarchical Bayesian inference. In: Clark, J.S., Gelfand, A.E., 2006. Hierarchical modelling for the environmental sciences, Oxford University Press Inc., New York.
- Clark, J.S., 2005. Why environmental scientists are becoming Bayesians. Ecology Letters, 8, 2-14.
- Chamberlain, A.C., 1970. Interception and retention of radioactive aerosols by vegetation. Atmospheric Environment. 4, 57-78.
- Chamberlain, A.C., 1987. Environmental impact of particles emitted from Windscale piles, 1954-7. Science of the Total Environ. 63, 139-160.
- Favier J.-C., Ireland-Ripert J., Toque C., Feinberg M., (1995). Répertoire général des aliments – table de composition, 2ème édition, INRA éditions.
- Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., Rubin, D.B., 2004. Bayesian Data analysis, Third edition. Chapman & Hall-CRC Texts in Statistical Science, 676 pages.
- Gelman, A., 2006. Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). Bayesian anal. 1 (3), 515-534.
- Gonze, M.-A., Mourlon, C., Garcia-Sanchez, L., Le Dizès, S., Nicoulaud, V., Métivier, J.-M., Simon-Cornu, M., Gerber P.-P., Vermorel, F., 2011. SYMBIOSE: A Simulation Platform for Performing Radiological Risk Assessments, ICRER 2011, 19-24 June, Hamilton (Canada), site (consulted on 03/12/2013): http://www.irsn.fr/EN/Research/Scientific-tools/Computer-codes/SYMBIOSE/Documents/ICRER2011_SymbiosePlatform_Poster02.pdf
- Heinemann, K., Vogt, K.J., Angeletti, L., 1974. Deposition and biological half-life of elemental iodine on grass and clover. In Proceedings Atmosphere-Surface Exchange of Particulate and Gaseous Pollutants—1974 Symposium. 136-152.
- Hinton, T.G., Garnier-Laplace, J., Vandenhove, H., Dowdall, M., Adam-Guillermin, C., Alonzo, F., Barnett, C., Beaugelin-Seiller, K., Beresford, N.A., Bradshaw, C., Brown, J., Eyrolle, F., Fevrier, L., Gariel, J.C., Gilbin, R., Hertel-Aas, T., Horemans, N., Howard, B.J., Ikäheimonen, T., Mora, J.C., Oughton, D., Real, A., Salbu, B., Simon-Cornu, M., Steiner, M., Sweeck, L., Vives i Batlle, J., 2013. An invitation to contribute to a strategic research agenda in radioecology. J. Environ. Radioact. 115, 73-82.
- Hoffman, F.O., Garten, C.T., Jr., Huckabee, J.W., Lucas, D.M., 1982. Interception and retention of

- technetium by vegetation and soil. *J. Environ. Qual.* 11, 134-141.
- Hosseini, A., Stenberg, K., Avila, R., Beresford, N.A., Brown, J.E., 2013. Application of the Bayesian approach for derivation of PDFs for concentration ratio values. *J. Environ. Radioact.* 126, 376-387.
- IAEA, 1996. Modelling of Radionuclide Interception and Loss Process in Vegetation and of Transfer in Semi-natural Ecosystems: Second Report of the VAMP Terrestrial Working Group. IAEA-TECDOC-857. International Atomic Energy Agency, Vienna, Austria.
- IAEA, 2009. Quantification of Radionuclide Transfer in Terrestrial and Freshwater Environments for radiological Assessments. IAEA-TECDOC-1616. International Atomic Energy Agency, Vienna, Austria.
- IAEA, 2010. Handbook of Parameter Values for the Prediction of Radionuclide Transfer in Terrestrial and Freshwater Environments. IAEA-TRS-472. IAEA, Vienna.
- International Commission on Radiological Protection (1996). Age-dependent Dose to Members of the Public from Intake of Radionuclides : Part 5 Compilation of Ingestion and Inhalation Dose Coefficients, Publication 72 (Annals of the ICRP Vol. 26, No 1), Pergamon Press, Oxford.
- Jackman, S., 2009. Bayesian Analysis for the Social Sciences. John Wiley & Sons, Ltd, Chichester.
- Jiao, Y., Hayes, C., Cortés, E., 2009. Hierarchical Bayesian approach for population dynamics modelling of fish complexes without species-specific data. – *ICES Journal of Marine Science*, 66: 367–377.
- Kirchmann, R., Fagnart, E., Van Puymbroeck, S., 1966. Studies on the foliar contamination by radiocaesium and radiostrontium. In: Symposium Radioecological concentration processes. Stockholm, 25-29 April 1966, 475-483.
- Kirchner, G., 1994. Transport of iodine and cesium via the grass-cow-milk pathway after the Chernobyl accident. *Health Phys.* 66 (6), 653-665.
- Köhler, H., Peterson, S.R., Hoffman, F.O., 1991. BIOMOV5 Technical Report 13 Part 1. Scenario A4: Multiple Model Testing using Chernobyl Fallout Data of I-131 in Forage and Milk and Cs-137 in Forage, Milk, Beef and Grain. ISSN 1100-0392.
- Kroes, R., Müller, D., Lambe, J., Löwik, M.R.H., van Klaveren, J., Kleiner, J., Massey, R., Mayer, S., Urieta, I., Verger, P., Visconti, A., 2002. Assessment of intake from the diet. *Food and Chem. Toxicol.* 40, 327-385.
- Liland, A., Barnett, C.L., Beresford, N.A., Bradshaw, C., Brown, J., Calmon, Ph., Mora, J.C., Chmielewska, I., Duffa, C., Février, L., Gaschak, S., Iospje, M., Kasparov, V., Lahtinen, J., Michalik, B., Nanba, K., Outola, L., Real, A., Robles, B., Salbu, B., Skipperud, L., Simon-Cornu, M., Stark, K., Steiner, M., Thorrying, H., Vanhoudt, N., Vives i Battle, J., Vandenhove, H., 2013. Detailed plan for the COMET WP3 Initial Research Activity – list of research projects and goals, participants and timing. DELIVERABLE (D-N°3.1). available at https://wiki.ceh.ac.uk/display/COM/Publications?atl_token=c4dbabba10d86d0584fc9e3ad5bc2416e46c9528
- Linkov, I., Burmistrov, D., Kandjekar, M., Schell, W.R., 1999. Reducing uncertainty in the radionuclide transport modeling for the Chernobyl forests using Bayesian updating. In: Contaminated forests. Volume 58 NATOScience Series, 143-150.
- Martin, C.J., Heaton, B., Robb, J.D., 1988. Studies of ¹³¹I, ¹³⁷Cs and ¹⁰³Ru in milk, meat and vegetables in North East Scotland following the Chernobyl Accident. *J. of Environ. Radioact.* 6, 247-259.
- Milbourn, G.M., Taylor, R., 1965. The contamination of Grassland with radioactive strontium-I Initial retention and loss. *Radiation Botany.* 5, 337-347.
- Millard, G.C., Fraley, L., Markham, O.D., 1983. Deposition and retention of ¹⁴¹Ce and ¹³⁴Cs aerosols on cool desert vegetation. *Health Phys.* 44 (4), 349-357.
- Miller, C.W., Hoffman, F.O. 1982. An analysis of reported values of the environmental half-time for radionuclides deposited on the surfaces of vegetation. Proceedings of an International Symposium on migration in the terrestrial environment of long-lived radionuclides from the nuclear fuel cycle organized by the international atomic energy agency, the commission of the european communities and the oecd nuclear energy agency and held in Knoxville, USA, 27 - 31 July, 1981. IAEA-SM-257/63.
- Miller, C.W., Hoffman, F.O. 1983. An examination of the environmental half-time for radionuclides deposited on vegetation. *Health Phys.* 45 (3), 731-744.
- Monte, L., 1991. Analysis of radiocontamination data, collected in Italy following the Chernobyl accident, for the evaluation of transfer parameters

- of radionuclides in the Deposition-Vegetation-Cow-Milk Pathway. *J. Environ. Radioact.* 14, 317-329.
- Mück, K., Roth, K., Gerzabek, M.H., Oberländer, H.-E., 1994. Effective half-lives of I- and Cs-Isotopes in Grassland shortly after fallout. *J. Environ. Radioact.* 24, 127-143.
- Nicoulaud-Gouin, V., Garcia-Sanchez, L., Giacalone, M., Attard, J.C., Martin-Garin, A., Bois, F.Y., (In revision). Identifiability of sorption parameters in stirred flow-through reactor experiments and their identification with a Bayesian approach. *J. Environ. Radioact.*
- Parent, E., Bernier, J., 2007. *Le raisonnement bayésien: modélisation et inférence*, Springer, Paris.
- Peeters, A., 1993. Extensification of grasslands in Europe. Proceedings of the 18th International Grassland Conference, Canada, Paper 15, 81-83.
- Plummer, M., 2003. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003). Vienna. March, 20-22.
- Pröhl, G., Hoffman, F.O., 1996. Radionuclide interception and loss processes in vegetation. Modelling of radionuclide interception and loss processes in vegetation and of transfer in semi-natural ecosystems: Second report of the VAMP Terrestrial Working Group. IAEA. Vienna, 9-47.
- Pröhl, G., 2009. Interception of dry and wet deposited radionuclides by vegetation. *J. Environ. Radioact.* 100 (9), 675-682.
- R Core Team (2014). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
- Russell, R.S., 1965. An introductory review. Interception and retention of airborne material on plants. *Health Physics.* 11, 1305-1315.
- Simon-Cornu, M., Beaugelin-Seiller, K., Boyer, P., Calmon, P., Garcia-Sanchez, L., Murlon, C., Nicoulaud, V., Sy, M.M., Gonze, M.A., 2015. Evaluating variability and uncertainty in radiological impact assessment with the probabilistic database of SYMBIOSE. *J. Environ. Radioact.* 139, 91-102.
- Smit, H.J., Metzger, M.J., Ewert, F., 2008. Spatial distribution of grassland productivity and land use in Europe. *Agricultural Systems.* 98, 208-219.
- Spiegelhalter, D.J., Best, N.G., Carlin, B.P., Linde, A.V.D., 2002. Bayesian Measures of model complexity and fit. *J. Roy. Stat. Soc. Series B (Statistical Methodology)* 64 (4), 583-639.
- Sy, M.M., Feinberg, M., Verger, P., Barré, T., Cléménçon, S., Crépet, A., 2013. New approach for the assessment of cluster diets. *Food and Cham. Toxicol.* 52, 180-187.
- Sy, M.M., Ancelet, S., Henner, P., Hurtevent, P., Simon-Cornu, M., 2015. Foliar interception of radionuclides in dry conditions: a meta-analysis using a Bayesian modeling approach. *J. Environ. Radioact.* 147, 63-75.
- International Commission on Radiological Protection (1996b). Age-dependent Dose to Members of the Public from Intake of Radionuclides : Part 5 Compilation of Ingestion and Inhalation Dose Coefficients, Publication 72 (Annals of the ICRP Vol. 26, No 1), Pergamon Press, Oxford.

1 Tables

2 **Table 1: Datasets of weathering rate (or weathering loss) values collated and considered in the HBMA.**

Database	Element	Radionuclide Physical Form	Plant Species	Min	Max	Unit	Experimental Conditions	Measurement Basis	Reference*	Number of data per reference
I	Iodine (115)	Gaseous (34) Particle (25) Unknown (56)	Herbaceous Vegetation (9) Pasture Grass (106)	0.013	0.244	day-1	Field	per unit area (23) per unit mass (92)	Heinemann <i>et al.</i> 1974	4
									Black and Barth 1976	19
									Miller and Hoffman 1982	2
									Miller and Hoffman 1983	30
									Chamberlain 1987	2
									Köhler <i>et al.</i> 1991	13
									Monte 1991	3
									Kirchner 1994	32
									Muck <i>et al.</i> 1994	3
									IAEA 1996	6
Total = 115									Pröhl and Hoffman 1996	1
F	Cerium (2) Cesium (81) Rubidium (8) Ruthenium (2) Strontium (34) Technetium (8) Zirconium (1)	Particle (136)	Herbaceous Vegetation (16) Leafy Vegetables (3) Non-leafy Vegetables (6) Pasture Grass (111)	0.014	0.416	day-1	Field	per unit area (17) per unit mass (119)	Milbourn and Taylor 1965**	2
									Kirchman <i>et al.</i> 1966	8
									Hoffman <i>et al.</i> 1982	8
									Millard <i>et al.</i> 1983	4
									Miller and Hoffman 1983	7
									Martin <i>et al.</i> 1988	1
									Köhler <i>et al.</i> 1991	8
									Monte 1991	3
									Kirchner 1994	21
									Muck <i>et al.</i> 1994	7
Total = 136									IAEA 1996	12
									Pröhl and Hoffman 1996	9
									Bengtsson <i>et al.</i> 2014	46
L	Cerium (1) Cesium (17)	Particle (66)	Leafy Vegetables (25) Non-leafy Vegetables (37)	0.012	0.158	day-1		per unit area (66) per unit mass (0)	Milbourn and Taylor 1965**	1
									Lab	Miller and Hoffman 1983

Cobalt (10)
Manganese
(10)
Ruthenium (10)
Strontium (18)

Pasture Grass (4)

IAEA 2009

62

Total = 66

3 * As often as possible, the source is the original one, i.e. the publication in which the dataset was first presented. Otherwise, the source is the review from which the dataset
4 was extracted.

5 ** The weathering loss value was estimated by non-linear regression based on one exponential function applied to the published kinetics.

6

7 **Table 2: Prior distributions defined for the parameters of the considered models for iodine and the other particles. G = Gamma; LN = LogNormal; N**
8 **= Normal and U = Uniform.**

Model	Model Formulation	Parameter	Plant Species	Unit	Prior distribution
MI1	$\log WL_{ijk} = \mu_{ijk} + \varepsilon_{ijk}$ $\mu_{ijk} = \log (\lambda^W + I(\text{Measure Basis } (WL_{ijk})$ $= \text{per unit mass}) \times \lambda_i^{BD})$	λ_1^{BD}	Herbaceous Vegetation	day ⁻¹	U (0.015, 0.1)
		λ_2^{BD}	Pasture Grass	day ⁻¹	U (0.015, 0.1)
		λ^W	-	day ⁻¹	LN (0, 10 ⁴)
		σ^2	-	day ⁻¹	IG(0.001,0.001)
MI2	$\log WL_{ijk} = \mu_{ijk} + \varepsilon_{ijk}$ $\mu_{ijk} = \log (\lambda_{i,j}^W + I(\text{Measure Basis } (WL_{ijk})$ $= \text{per unit mass}) \times \lambda_i^{BD})$	λ_1^{BD}	Herbaceous Vegetation	day ⁻¹	U (0.015, 0.1)
		λ_2^{BD}	Pasture Grass	day ⁻¹	U (0.015, 0.1)
		μ_{λ^W}	-	day ⁻¹	N (0,10 ⁴)
		$\sigma_{\lambda^W}^2$	-	day ⁻¹	IG(0.001,0.001)
MI3	$\log WL_{ijk} = \mu_{ijk} + \varepsilon_{ijk}$ $\mu_{ijk} = \log (\lambda_{i,j}^W + I(\text{Measure Basis } (WL_{ijk})$ $= \text{per unit mass}) \times \lambda_i^{BD})$	λ_1^{BD}	Herbaceous Vegetation	day ⁻¹	U (0.015, 0.1)
		λ_2^{BD}	Pasture Grass	day ⁻¹	U (0.015, 0.1)
		μ_{λ^W}	-	day ⁻¹	N (0,10 ⁴)
		$\sigma_{\lambda^W}^2$	-	day ⁻¹	IG(0.001,0.001)
MI4	$\log WL_{ijk} = \mu_{ijk} + \varepsilon_{ijk}$	λ_1^{BD}	Herbaceous Vegetation	day ⁻¹	U (0.015, 0.1)
		λ_2^{BD}	Pasture Grass	day ⁻¹	U (0.015, 0.1)
		μ_{λ^W}	-	day ⁻¹	N (0,10 ⁴)
		σ^2	-	day ⁻¹	IG(0.001,0.001)

	$\mu_{ijk} = \log (\lambda_{i,j}^W + I(\text{Measure Basis } (WL_{ijk})$ $= \text{per unit mass}) \times \lambda_i^{BD})$	$\sigma_{\lambda^W}^2$	-	day ⁻¹	IG(0.001,0.001)
		σ^2	-	day ⁻¹	IG(0.001,0.001)
MPF1	$\log WL_{ijk} = \mu_{ijk} + \varepsilon_{ijk}$	λ_1^{BD}	Herbaceous Vegetation	day ⁻¹	U (0.015, 0.1)
		λ_2^{BD}	Leafy Vegetable	day ⁻¹	U (0.03, 0.14)
		λ_3^{BD}	Non-leafy Vegetables	day ⁻¹	U (0.03, 0.14)
		λ_4^{BD}	Pasture Grass	day ⁻¹	U (0.015, 0.1)
	$\mu_{ijk} = \log (\lambda_{i,j}^W + I(\text{Measure Basis } (WL_{ijk})$ $= \text{per unit mass}) \times \lambda_i^{BD})$	λ^W	-	day ⁻¹	LN (0, 10 ⁴)
		σ^2	-	day ⁻¹	IG(0.001,0.001)
MPF2	$\log WL_{ijk} = \mu_{ijk} + \varepsilon_{ijk}$	λ_1^{BD}	Herbaceous Vegetation	day ⁻¹	U (0.015, 0.1)
		λ_2^{BD}	Leafy Vegetable	day ⁻¹	U (0.03, 0.14)
		λ_3^{BD}	Non-leafy Vegetables	day ⁻¹	U (0.03, 0.14)
		λ_4^{BD}	Pasture Grass	day ⁻¹	U (0.015, 0.1)
	$\mu_{ijk} = \log (\lambda_{i,j}^W + I(\text{Measure Basis } (WL_{ijk})$ $= \text{per unit mass}) \times \lambda_i^{BD})$	μ_{λ^W}	-	day ⁻¹	N (0,10 ⁴)
		$\sigma_{\lambda^W}^2$	-	day ⁻¹	IG(0.001,0.001)
		σ^2	-	day ⁻¹	IG(0.001,0.001)
MPF3	$\log WL_{ijk} = \mu_{ijk} + \varepsilon_{ijk}$	λ_1^{BD}	Herbaceous Vegetation	day ⁻¹	U (0.015, 0.1)
		λ_2^{BD}	Leafy Vegetable	day ⁻¹	U (0.03, 0.14)
		λ_3^{BD}	Non-leafy Vegetables	day ⁻¹	U (0.03, 0.14)
		λ_4^{BD}	Pasture Grass	day ⁻¹	U (0.015, 0.1)
	$\mu_{ijk} = \log (\lambda_{i,j}^W + I(\text{Measure Basis } (WL_{ijk})$ $= \text{per unit mass}) \times \lambda_i^{BD})$	μ_{λ^W}	-	day ⁻¹	N (0,10 ⁴)
		$\sigma_{\lambda^W}^2$	-	day ⁻¹	IG(0.001,0.001)
		σ^2	-	day ⁻¹	IG(0.001,0.001)
MPF4	$\log WL_{ijk} = \mu_{ijk} + \varepsilon_{ijk}$	λ_1^{BD}	Herbaceous Vegetation	day ⁻¹	U (0.015, 0.1)
		λ_2^{BD}	Leafy Vegetable	day ⁻¹	U (0.03, 0.14)
		λ_3^{BD}	Non-leafy Vegetables	day ⁻¹	U (0.03, 0.14)
		λ_4^{BD}	Pasture Grass	day ⁻¹	U (0.015, 0.1)
	$\mu_{ijk} = \log (\lambda_{i,j}^W + I(\text{Measure Basis } (WL_{ijk})$ $= \text{per unit mass}) \times \lambda_i^{BD})$	μ_{λ^W}	-	day ⁻¹	N (0,10 ⁴)
		$\sigma_{\lambda^W}^2$	-	day ⁻¹	IG(0.001,0.001)
		σ^2	-	day ⁻¹	IG(0.001,0.001)
MPL1	$WL_{ijk}^* = \mu_{ijk}^* + \varepsilon_{ijk}^*$	λ^W	-	day ⁻¹	LN (0, 10 ⁴)

	$\mu_{ijk}^* = \log \lambda^W$	σ^{*2}	-	day ⁻¹	IG(0.001,0.001)
MPL2	$WL_{ijk}^* = \mu_{ijk}^* + \varepsilon_{ijk}^*$	μ_{λ^W}	-	day ⁻¹	N (0,10 ⁴)
	$\mu_{ijk}^* = \log \lambda_{i,j}^W$	$\sigma_{\lambda^W}^2$	-	day ⁻¹	IG(0.001,0.001)
MPL3	$WL_{ijk}^* = \mu_{ijk}^* + \varepsilon_{ijk}^*$	μ_{λ^W}	-	day ⁻¹	N (0,10 ⁴)
	$\mu_{ijk}^* = \log \lambda_{i,j}^W$	$\sigma_{\lambda^W}^2$	-	day ⁻¹	IG(0.001,0.001)
MPL4	$WL_{ijk}^* = \mu_{ijk}^* + \varepsilon_{ijk}^*$	μ_{λ^W}	-	day ⁻¹	N (0,10 ⁴)
	$\mu_{ijk}^* = \log \lambda_{i,j}^W$	$\sigma_{\lambda^W}^2$	-	day ⁻¹	IG(0.001,0.001)

9 **Table 3: Input parameter values and probability distribution functions considered in the propagation of uncertainty scenario. Only the parameters**
10 **relating to the foliar and the root uptake are considered. LN = LogNormal; N = Normal and U = Uniform.**

Food Chain Pathway	Input Parameter / Covariates	Element / Isotopes	Plant Species	Soil	Unit	Value / Distribution	Source
	Total Deposits	Cesium / Iodine / Strontium	-	-	Bq.m ⁻²	1000	This article (Section 3.1)
Scenario setting parameters	Fresh Biomass <i>B</i>	-	Grass	-	kg.m ⁻²	LN(-1.05 , 0.61)	This article (Annex A)
		-	Leafy Vegetables	-	kg.m ⁻²	N(1.14 , 0.42)	
	Radioactive Decay λ^{Cs}	Cs-134	-	-	day ⁻¹	9.19 10 ⁻⁴	Nucleonica
		Cs-137	-	-	day ⁻¹	6.32 10 ⁻⁵	
		I-131	-	-	day ⁻¹	8.64 10 ⁻²	
		Sr-90	-	-	day ⁻¹	6.59 10 ⁻⁵	
Foliar Pathway	Dry Interception Fraction DIF (or DMIF)	Cesium / Strontium	Grass	-	-	N(0.61 , 0.15)	This article (Annex A)
			Leafy Vegetables	-	m ² .kg ⁻¹	LN(-1.37 , 0.33)	
	Iodine	Grass	-	-	N(0.42 , 0.16)		
		Leafy Vegetables	-	m ² .kg ⁻¹	LN(-1.50 , 0.29)		

		Iodine	Grass	-	day ⁻¹	LN(-2.07 , 0.60)	
			Leafy Vegetables	-	day ⁻¹	LN(-2.85 , 0.60)	
	Weathering loss parameter λ^{WL}	Cesium	Grass	-	day ⁻¹	LN(-2.54 , 0.59)	Models MI ₃ and MPF ₂ of this article (See Section 2)
			Leafy Vegetables	-	day ⁻¹	LN(-2.80 , 0.59)	
		Strontium	Grass	-	day ⁻¹	LN(-2.60 , 0.59)	
			Leafy Vegetables	-	day ⁻¹	LN(-2.89 , 0.59)	
	Dry Matter Content f^{dw}	-	Grass	-	-	U(0.1 , 0.33)	IAEA (2010)
			-	Leafy Vegetables	-	-	
Root Uptake	Soil-to-Plant Transfer Factor TF^{Up}	Cesium	Grass	Clay	-	LN(-1.39 , 1.44)	Simon-Cornu <i>et al.</i> 2015
			Grass	Sandy	-	LN(-1.39 , 1.44)	
		Iodine	Grass	Clay	-	LN(-5.6 , 1.828)	
			Grass	Sandy	-	LN(-5.6 , 1.828)	
		Strontium	Grass	Clay	-	LN(0.26 , 0.805)	
			Grass	Sandy	-	LN(0.26 , 0.805)	
		Cesium	Leafy Vegetables	Clay	-	LN(-2.79 , 1.828)	
			Leafy Vegetables	Sandy	-	LN(-2.81 , 1.828)	
		Iodine	Leafy Vegetables	Clay	-	LN(-5.04 , 1.34)	
			Leafy Vegetables	Sandy	-	LN(-5.04 , 1.34)	
		Strontium	Leafy Vegetables	Clay	-	LN(-0.28 , 1.83)	
			Leafy Vegetables	Sandy	-	LN(-0.27 , 1.83)	
	Surface Layer Height h_{SL}	-	-	m	0.002	SYMBIOSE default	
	Soil Distribution Coefficient Kd	Cesium	-	Clay	L.kg ⁻¹	LN(7.09 , 1.986)	Simon-Cornu <i>et al.</i> 2015
			-	Sandy	L.kg ⁻¹	LN(7.09 , 1.986)	

	I-131	-	Clay	L.kg ⁻¹	LN(1.93 , 1.72)	
		-	Sandy	L.kg ⁻¹	LN(1.93 , 1.72)	
	Sr-90	-	Clay	L.kg ⁻¹	LN(3.95 , 1.81)	
		-	Sandy	L.kg ⁻¹	LN(3.95 , 1.81)	
Soil Density ρ	-	-	Clay	kg.m ⁻³	1300	SYMBIOSE default
	-	-	Sandy	kg.m ⁻³	1550	
Root Layer water content θ	-	-	-	L. m ⁻³	0.4	SYMBIOSE default
Root Layer height h_{RL}	-	-	-	m	0.15	SYMBIOSE default
Coefficient of the Migration RL - UL α	Cesium / Iodine / Strontium	-	Clay	-	10	SYMBIOSE default
		-	Sandy	-	10	SYMBIOSE default

11 **Table 4: Probability distribution functions and values considered in the propagation of**
 12 **uncertainty scenario for the equilibrium transfer (to animal products) factor. LN = LogNormal.**

Parameter	Substance	Product	Unit	Value / Distribution	Source
Equilibrium Transfer Factor	Cesium	Beef Meat	day.kg ⁻¹	LN(-3.82 , 0.89)	Simon-Cornu <i>et al.</i> 2015
		Cow Milk	day.kg ⁻¹	LN(-5.38 , 0.71)	
		Goat Milk	day.kg ⁻¹	LN(-2.21 , 0.81)	
		Pork Meat	day.kg ⁻¹	LN(-1.61 , 0.41)	
		Sheep Meat	day.kg ⁻¹	LN(-1.66 , 0.80)	
	Iodine	Beef Meat	day.kg ⁻¹	LN(-4.74 , 0.75)	
		Cow Milk	day.kg ⁻¹	LN(-5.22 , 0.89)	
		Goat Milk	day.kg ⁻¹	LN(-1.51 , 1.09)	
		Pork Meat	day.kg ⁻¹	LN(-3.46 , 0.38)	
		Sheep Meat	day.kg ⁻¹	0.03	
	Strontium	Beef Meat	day.kg ⁻¹	LN(-6.65 , 1.09)	
		Cow Milk	day.kg ⁻¹	LN(-6.64 , 0.54)	
		Goat Milk	day.kg ⁻¹	LN(4.14 , 0.71)	
		Pork Meat	day.kg ⁻¹	LN(-5.99 , 1.01)	
		Sheep Meat	day.kg ⁻¹	LN(-6.50 , 0.54)	

13
 14 **Table 5: Deviance information criterion (DIC) calculated for the competing models for iodine**
 15 **and the other particles.**

Models	Dbar	pD	DIC = Dbar + pD
MI1	-404.50	3.00	-401.50
MI2	-407.40	4.90	-402.50
MI3	-406.80	4.10	-402.70
MI4	-407.60	5.30	-402.30
MPF1	-474.90	2.20	-472.70
MPF2	-517.70	7.20	-510.50
MPF3	-501.60	5.20	-496.40
MPF4	-518.80	10.60	-508.20
MPL1	-327.60	2.10	-325.50
MPL2	-347.80	6.10	-341.70
MPL3	-327.00	2.80	-324.20
MPL4	-348.30	9.50	-338.80

16

17 **Table 6: Summary statistics (i.e. posterior means, standard deviation (Sd), medians and 95%CI) of the simulated posterior distributions of the**
 18 **parameters and latent variables of the models MI₁ to MI₄.**

Model	Parameter / Latent Variable	Symbols	Physical Form Iodine	Plant Species	Unit	Post. Mean	Post. Sd	Post. Median	95%CI
MI1	Parameters	λ_1^{BD}	-	Herbaceous Vegetation	day ⁻¹	0.041	0.019	0.038	[0.016, 0.086]
		λ_2^{BD}	-	Pasture Grass	day ⁻¹	0.026	0.009	0.024	[0.015, 0.048]
		λ^W	-	-	day ⁻¹	0.070	0.004	0.070	[0.062, 0.078]
		σ	-	-	day ⁻¹	0.564	0.038	0.562	[0.495, 0.644]
MI2	Parameters	λ_1^{BD}	-	Herbaceous Vegetation	day ⁻¹	0.045	0.019	0.043	[0.017, 0.089]
		λ_2^{BD}	-	Pasture Grass	day ⁻¹	0.030	0.011	0.028	[0.016, 0.055]
		μ_{λ^W}	-	-	day ⁻¹	-2.714	0.276	-2.699	[-3.098, -2.37]
		σ_{λ^W}	-	-	day ⁻¹	0.229	0.430	0.146	[0.031, 0.903]
		σ	-	-	day ⁻¹	0.557	0.038	0.555	[0.49, 0.638]
	Latent Variables	$\lambda_{,1}^W$	Particle	-	day ⁻¹	0.065	0.006	0.065	[0.053, 0.078]
		$\lambda_{,2}^W$	Unknown	-	day ⁻¹	0.075	0.006	0.075	[0.065, 0.087]
		$\lambda_{,3}^W$	Vapor	-	day ⁻¹	0.061	0.007	0.062	[0.047, 0.075]
MI3	Parameters	λ_1^{BD}	-	Herbaceous Vegetation	day ⁻¹	0.052	0.021	0.051	[0.018, 0.094]
		λ_2^{BD}	-	Pasture Grass	day ⁻¹	0.026	0.009	0.024	[0.015, 0.047]
		μ_{λ^W}	-	-	day ⁻¹	-2.573	3.652	-2.764	[-5.05, 0.417]
		σ_{λ^W}	-	-	day ⁻¹	1.686	10.251	0.354	[0.036, 10.503]
		σ	-	-	day ⁻¹	0.559	0.038	0.556	[0.491, 0.639]
	Latent Variables	$\lambda_{1,}^W$	-	Herbaceous Vegetation	day ⁻¹	0.051	0.014	0.051	[0.027, 0.077]
		$\lambda_{2,3}^W$	-	Pasture Grass	day ⁻¹	0.071	0.004	0.071	[0.063, 0.079]
MI4	Parameters	λ_1^{BD}	-	Herbaceous Vegetation	day ⁻¹	0.047	0.020	0.045	[0.017, 0.091]
		λ_2^{BD}	-	Pasture Grass	day ⁻¹	0.029	0.010	0.027	[0.016, 0.054]
		μ_{λ^W}	-	-	day ⁻¹	-2.753	0.174	-2.729	[-3.121, -2.522]
		σ_{λ^W}	-	-	day ⁻¹	0.192	0.208	0.139	[0.03, 0.665]
		σ	-	-	day ⁻¹	0.557	0.038	0.555	[0.488, 0.638]
	Latent Variables	$\lambda_{1,1}^W$	Particle	Herbaceous Vegetation	day ⁻¹	0.062	0.011	0.063	[0.038, 0.082]
		$\lambda_{1,3}^W$	Vapor	Herbaceous Vegetation	day ⁻¹	0.059	0.014	0.061	[0.027, 0.081]

$\lambda_{2,1}^W$	Particle	Pasture Grass	day ⁻¹	0.066	0.007	0.066	[0.054, 0.08]
$\lambda_{2,2}^W$	Unknown	Pasture Grass	day ⁻¹	0.075	0.006	0.074	[0.065, 0.086]
$\lambda_{2,3}^W$	Vapor	Pasture Grass	day ⁻¹	0.063	0.007	0.063	[0.048, 0.076]

19
20
21
22

Table 7: Summary statistics (i.e. posterior means, standard deviation (Sd), medians and 95%CI) of the simulated posterior distributions of the parameters and latent variables of the models MPF₁ to MPF₄.

Model	Parameter / Latent Variable	Symbols	Element	Plant Species	Unit	Post. Mean	Post. Sd	Post. Median	95%CI
MPF1	Parameters	λ_1^{BD}	-	Herbaceous Vegetation	day ⁻¹	0.024	0.009	0.021	[0.015, 0.049]
		λ_2^{BD*}	-	Leafy Vegetables	day ⁻¹	n/a	n/a	n/a	n/a
		λ_3^{BD*}	-	Non-leafy Vegetables	day ⁻¹	n/a	n/a	n/a	n/a
		λ_4^{BD}	-	Pasture Grass	day ⁻¹	0.019	0.004	0.018	[0.015, 0.03]
		λ^W	-	-	day ⁻¹	0.061	0.004	0.061	[0.054, 0.069]
		σ	-	-	day ⁻¹	0.683	0.043	0.682	[0.606, 0.772]
MPF2	Parameters	λ_1^{BD}	-	Herbaceous Vegetation	day ⁻¹	0.022	0.007	0.020	[0.015, 0.04]
		λ_2^{BD*}	-	Leafy Vegetables	day ⁻¹	n/a	n/a	n/a	n/a
		λ_3^{BD*}	-	Non-leafy Vegetables	day ⁻¹	n/a	n/a	n/a	n/a
		λ_4^{BD}	-	Pasture Grass	day ⁻¹	0.018	0.003	0.017	[0.015, 0.026]
		μ_{λ^W}	-	-	day ⁻¹	-2.851	0.351	-2.839	[-3.572, -2.19]
		σ_{λ^W}	-	-	day ⁻¹	0.775	0.342	0.702	[0.367, 1.618]
		σ	-	-	day ⁻¹	0.584	0.037	0.582	[0.517, 0.662]
	Latent Variables	$\lambda_{.,1}^W$	Cerium	-	day ⁻¹	0.053	0.020	0.049	[0.024, 0.099]
		$\lambda_{.,2}^W$	Cesium	-	day ⁻¹	0.061	0.004	0.060	[0.053, 0.069]
		$\lambda_{.,3}^W$	Rubidium	-	day ⁻¹	0.207	0.045	0.202	[0.132, 0.306]
		$\lambda_{.,4}^W$	Ruthenium	-	day ⁻¹	0.059	0.023	0.055	[0.025, 0.116]
$\lambda_{.,5}^W$		Strontium	-	day ⁻¹	0.056	0.006	0.055	[0.045, 0.067]	

		$\lambda_{,6}^W$	Technetium	-	day ⁻¹	0.033	0.008	0.032	[0.02, 0.05]	
		$\lambda_{,7}^W$	Zirconium	-	day ⁻¹	0.044	0.026	0.040	[0.008, 0.107]	
MPF3	Parameters	λ_1^{BD}	-	Herbaceous Vegetation	day ⁻¹	0.022	0.008	0.020	[0.015, 0.043]	
		λ_2^{BD*}	-	Leafy Vegetables	day ⁻¹	n/a	n/a	n/a	n/a	
		λ_3^{BD*}	-	Non-leafy Vegetables	day ⁻¹	n/a	n/a	n/a	n/a	
		λ_4^{BD}	-	Pasture Grass	day ⁻¹	0.018	0.003	0.017	[0.015, 0.028]	
		μ_{λ^W}	-	-	day ⁻¹	-2.510	0.586	-2.525	[-3.569, -1.403]	
		σ_{λ^W}	-	-	day ⁻¹	0.912	0.696	0.741	[0.314, 2.542]	
		σ	-	-	day ⁻¹	0.619	0.039	0.618	[0.549, 0.702]	
	Latent Variables	$\lambda_{1,.}^W$	-	Herbaceous Vegetation	day ⁻¹	0.047	0.008	0.046	[0.034, 0.064]	
		$\lambda_{2,.}^W$	-	Leafy Vegetables	day ⁻¹	0.084	0.028	0.079	[0.042, 0.151]	
		$\lambda_{3,.}^W$	-	Non-leafy Vegetables	day ⁻¹	0.205	0.055	0.198	[0.116, 0.33]	
		$\lambda_{4,.}^W$	-	Pasture Grass	day ⁻¹	0.059	0.004	0.059	[0.053, 0.067]	
	MPF4	Parameters	λ_1^{BD}	-	Herbaceous Vegetation	day ⁻¹	0.022	0.007	0.020	[0.015, 0.04]
			λ_2^{BD*}	-	Leafy Vegetables	day ⁻¹	n/a	n/a	n/a	n/a
λ_3^{BD*}			-	Non-leafy Vegetables	day ⁻¹	n/a	n/a	n/a	n/a	
λ_4^{BD}			-	Pasture Grass	day ⁻¹	0.018	0.003	0.017	[0.015, 0.027]	
μ_{λ^W}			-	-	day ⁻¹	-2.804	0.254	-2.799	[-3.329, -2.307]	
σ_{λ^W}			-	-	day ⁻¹	0.731	0.232	0.694	[0.389, 1.294]	
σ			-	-	day ⁻¹	0.582	0.038	0.580	[0.514, 0.662]	
Latent Variables		$\lambda_{1,1}^W$	Cerium	Herbaceous Vegetation	day ⁻¹	0.053	0.020	0.050	[0.025, 0.1]	
		$\lambda_{1,2}^W$	Cesium	Herbaceous Vegetation	day ⁻¹	0.059	0.022	0.056	[0.028, 0.113]	
		$\lambda_{2,2}^W$	Cesium	Leafy Vegetables	day ⁻¹	0.075	0.028	0.070	[0.035, 0.141]	
	$\lambda_{4,2}^W$	Cesium	Pasture Grass	day ⁻¹	0.060	0.004	0.060	[0.053, 0.069]		
	$\lambda_{1,3}^W$	Rubidium	Herbaceous Vegetation	day ⁻¹	0.177	0.072	0.163	[0.078, 0.356]		

$\lambda_{3,3}^W$	Rubidium	Non-leafy Vegetables	day ⁻¹	0.201	0.048	0.195	[0.121, 0.311]
$\lambda_{2,4}^W$	Ruthenium	Leafy Vegetables	day ⁻¹	0.086	0.043	0.077	[0.032, 0.193]
$\lambda_{4,4}^W$	Ruthenium	Pasture Grass	day ⁻¹	0.042	0.024	0.037	[0.009, 0.101]
$\lambda_{1,5}^W$	Strontium	Herbaceous Vegetation	day ⁻¹	0.035	0.013	0.033	[0.015, 0.066]
$\lambda_{4,5}^W$	Strontium	Pasture Grass	day ⁻¹	0.059	0.006	0.058	[0.048, 0.072]
$\lambda_{1,6}^W$	Technetium	Herbaceous Vegetation	day ⁻¹	0.033	0.008	0.032	[0.02, 0.05]
$\lambda_{4,7}^W$	Zirconium	Pasture Grass	day ⁻¹	0.046	0.026	0.041	[0.01, 0.109]

23 *The posterior distributions of λ_2^{BD} and λ_3^{BD} were not provided since all the weathering loss values which relate to leafy and non-leafy vegetables were measured on a per unit
24 area basis.
25 n/a means not applicable
26

27 **Table 8: Summary statistics (i.e. posterior means, standard deviation (Sd), medians and 95%CI) of the approximate posterior distributions of the**
28 **parameters and latent variables of the models MPL₁ to MPL₄.**

Model	Parameter / Latent Variable	Symbols	Element	Plant Species	Unit	Post. Mean	Post. Sd	Post. Median	95%CI
MPL1	Parameters	λ^W	-	-	day ⁻¹	0.040	0.003	0.040	[0.035, 0.045]
		σ	-	-	day ⁻¹	0.510	0.046	0.507	[0.43, 0.609]
MPL2	Parameters	μ_{λ^W}	-	-	day ⁻¹	-3.230	0.153	-3.231	[-3.534, -2.925]
		σ_{λ^W}	-	-	day ⁻¹	0.300	0.151	0.268	[0.116, 0.671]
		σ	-	-	day ⁻¹	0.437	0.041	0.434	[0.367, 0.525]
	Latent Variables	$\lambda_{,1}^W$	Cerium	-	day ⁻¹	0.041	0.011	0.040	[0.024, 0.066]
		$\lambda_{,2}^W$	Cobalt	-	day ⁻¹	0.044	0.006	0.044	[0.034, 0.056]
		$\lambda_{,3}^W$	Cesium	-	day ⁻¹	0.030	0.003	0.030	[0.024, 0.037]
		$\lambda_{,4}^W$	Manganese	-	day ⁻¹	0.042	0.005	0.042	[0.033, 0.053]
		$\lambda_{,5}^W$	Ruthenium	-	day ⁻¹	0.033	0.004	0.032	[0.025, 0.042]
		$\lambda_{,6}^W$	Strontium	-	day ⁻¹	0.054	0.006	0.054	[0.044, 0.066]
		MPL3	Parameters	μ_{λ^W}	-	-	day ⁻¹	-3.220	0.346
σ_{λ^W}	-			-	day ⁻¹	0.146	1.303	0.083	[0.024, 0.566]
σ	-			-	day ⁻¹	0.512	0.046	0.509	[0.432, 0.611]

	Latent Variables	$\lambda_{1,}^W$	-	Leafy Vegetables	day ⁻¹	0.039	0.003	0.039	[0.033, 0.046]
		$\lambda_{2,}^W$	-	Non-leafy Vegetables	day ⁻¹	0.040	0.003	0.040	[0.035, 0.047]
		$\lambda_{3,}^W$	-	Pasture Grass	day ⁻¹	0.040	0.006	0.040	[0.031, 0.053]
MPL4	Parameters	μ_{λ^W}	-	-	day ⁻¹	-3.233	0.093	-3.233	[-3.418, -3.048]
		σ_{λ^W}	-	-	day ⁻¹	0.244	0.084	0.234	[0.025, 0.042]
		σ	-	-	day ⁻¹	0.436	0.043	0.432	[0.367, 0.525]
	Latent Variables	$\lambda_{3,1}^W$	Cerium	Pasture Grass	day ⁻¹	0.041	0.009	0.040	[0.026, 0.062]
		$\lambda_{1,2}^W$	Cobalt	Leafy Vegetables	day ⁻¹	0.046	0.007	0.046	[0.034, 0.063]
		$\lambda_{2,2}^W$	Cobalt	Non-leafy Vegetables	day ⁻¹	0.041	0.006	0.040	[0.03, 0.054]
		$\lambda_{1,3}^W$	Cesium	Leafy Vegetables	day ⁻¹	0.033	0.005	0.033	[0.023, 0.044]
		$\lambda_{2,3}^W$	Cesium	Non-leafy Vegetables	day ⁻¹	0.031	0.004	0.030	[0.024, 0.039]
		$\lambda_{3,3}^W$	Cesium	Pasture Grass	day ⁻¹	0.041	0.009	0.040	[0.025, 0.062]
		$\lambda_{1,4}^W$	Manganese	Leafy Vegetables	day ⁻¹	0.043	0.007	0.043	[0.032, 0.058]
		$\lambda_{2,4}^W$	Manganese	Non-leafy Vegetables	day ⁻¹	0.040	0.006	0.040	[0.029, 0.054]
		$\lambda_{1,5}^W$	Ruthenium	Leafy Vegetables	day ⁻¹	0.034	0.005	0.033	[0.024, 0.045]
		$\lambda_{2,5}^W$	Ruthenium	Non-leafy Vegetables	day ⁻¹	0.035	0.005	0.034	[0.025, 0.046]
		$\lambda_{1,6}^W$	Strontium	Leafy Vegetables	day ⁻¹	0.043	0.007	0.042	[0.032, 0.057]
		$\lambda_{2,6}^W$	Strontium	Non-leafy vegetables	day ⁻¹	0.059	0.008	0.059	[0.044, 0.076]
		$\lambda_{3,6}^W$	Strontium	Pasture Grass	day ⁻¹	0.041	0.008	0.040	[0.027, 0.059]

30
31
32
33
34

Table 9: Distributions of concentrations of Cs-134/-137, I-131 and Sr-90 (Bq/kg) predicted in beef meat, cow milk and leafy vegetables at approximately 1 month, 3 months and 12 months after the deposition event. Brackets represent the 5th, the 50th and the 95th percentiles of the distributions of concentrations values (i.e. (5th, 50th, 95th)).

Isotope	Food Item	04/09/2013	01/11/2013	01/08/2014
Cs-134	Beef Meat	(9.1, 47.1 , 205.5)	(4.4, 23.7 , 113.6)	(7.1, 25.4 , 84.4)
	Cow Milk	(1.2, 8.4 , 57.1)	(32.1, 87 , 240.2)	(0.035, 0.11 , 0.4)
	Leafy Vegetables	(2.6, 29.1 , 104.7)	(0.001, 0.8 , 23.4)	(0, 0.001 , 0.054)
Cs-137	Beef Meat	(9.4, 48.5 , 211.6)	(4.8, 25.7 , 122.9)	(9.6, 34.8 , 115.4)
	Cow Milk	(1.2, 8.7 , 58.7)	(34.7, 94.1 , 259.9)	(0.05, 0.2 , 0.5)
	Leafy Vegetables	(2.7, 29.9 , 107.8)	(0.002, 0.881 , 25.3)	(0, 0.001 , 0.07)
I-131	Beef Meat	(0.08, 0.4 , 2.5)	(0, 0.003 , 0.02)	insignificant
	Cow Milk	(0.02, 0.4 , 3.9)	(0.01, 0.04 , 0.2)	insignificant
	Leafy Vegetables	(0.13, 1.5 , 4.9)	(0, 0 , 0.008)	insignificant
Sr-90	Beef Meat	(0.6, 3.6 , 22.5)	(0.23, 1.45 , 10.4)	(0.5, 2.6 , 13)
	Cow Milk	(0.6, 4.7 , 25.3)	(12.1, 25.26 , 55.03)	(0.5, 1 , 2.2)
	Leafy Vegetables	(2.4, 29.9 , 105.7)	(0.02, 1.15 , 24.5)	(0.004, 0.08 , 1.3)

35
36
37
38
39
40

Table 10: Distributions of the effective doses of Cs-134/-137, I-131 and Sr-90 (µSv) received by ingestion by an adult (≥ 17 years old) and a young child (2 – 7 years old). The effective doses are predicted at approximately 1 month, 3 months, and 12 months after the deposition event. Brackets represent the 5th, the 50th and the 95th percentiles of the distributions of the effective doses (i.e. (5th, 50th, 95th)).

Isotope	Age Group	04/09/2013	01/11/2013	01/08/2014	01/09/2014
Cs-134	Adult	(4.3,8.9,19.7)	(7.6, 16.3 , 36.3)	(38.4, 78.2 , 152)	(39, 79.3 , 153.2)
	Young Child	(2.1, 4.3 , 8.9)	(3.5, 7.7 , 16.4)	(16.9, 34 , 65.1)	(17.2, 34.4 , 65.9)
Cs-137	Adult	(3, 6.2 , 13.6)	(5.4, 11.5 , 25.7)	(30.3, 62 , 120.5)	(30.8, 62.8 , 122.2)
	Young Child	(1.6, 3.2 , 6.6)	(2.7, 5.8 , 12.5)	(14.3, 28.9 , 55.4)	(14.5, 29.4 , 56.2)
I-131	Adult	(1.7, 3.6 , 12)	(1.7, 3.6 , 12.2)	(1.7, 3.6 , 12.2)	(1.7, 3.6 , 12.2)
	Young Child	(5.6, 11.3 , 34.3)	(5.7, 11.5 , 34.8)	(5.7, 11.5 , 34.9)	(5.7, 11.5 , 34.9)
Sr-90	Adult	(4.1, 8.3 , 15.4)	(5.4, 11.9 , 25)	(18.5, 32.3 , 57.4)	(18.9, 32.8 , 58)
	Young Child	(5.1, 10.3 , 18.8)	(6.6, 14.2 , 31.3)	(20.4, 35.5 , 63.3)	(20.8, 36 , 63.8)

41

42 **Table 11: Spearman correlation coefficients calculated (for both dry and wet scenarios) between**
 43 **concentrations of Cs-134/-137, I-131 and Sr-90 in beef meat, cow milk and leafy vegetables. Only**
 44 **parameters with highest impact on the distribution of concentrations values are displayed.**

Isotope	Food Item	Input Parameter ^(a)	04/09/2013	01/11/2013	01/08/2014
Cs-134/-137	Beef Meat	Foliar Biomass – Grass	-0.478	-0.453	0.035
		ETF	0.744	0.696	1.000
		Dry Interception – Grass	0.161	0.154	0.004
		Weathering loss – Grass	-0.237	-0.369	0.079
	Cow Milk	Foliar Biomass – Grass	-0.409	-0.001	-0.195
		ETF	0.484	1.000	0.815
		Weathering loss – Grass	-0.710	0.023	-0.361
	Leafy Vegetables	Dry Interception	0.230	0.054	0.030
		Weathering loss	-0.951	-0.992	-0.384
	I-131	Beef Meat	Foliar Biomass – Grass	-0.484	-0.462
ETF			0.834	0.801	0.998
Dry Interception – Grass			0.197	0.190	0.028
Weathering loss – Grass			-0.196	-0.333	-0.035
Cow Milk		Foliar Biomass – Grass	-0.506	-0.307	-0.383
		ETF	0.787	0.469	0.598
		Dry Interception – Grass	0.222	0.137	0.165
		Weathering loss – Grass	-0.115	-0.785	-0.632
Leafy Vegetables		Dry Interception	0.268	0.122	0.046
		Weathering loss	-0.968	-0.995	-0.412
Sr-90	Beef Meat	Foliar Biomass – Grass	-0.413	-0.392	-0.007
		Weathering loss – Grass	-0.206	-0.348	-0.043
		Dry Interception – Grass	0.155	0.151	0.020
		ETF	0.839	0.801	0.997
	Cow Milk	Weathering loss – Grass	-0.775	-0.029	-0.112
		Foliar Biomass – Grass	-0.409	0.027	-0.052
		Dry Interception – Grass	0.155	0.021	0.045
		ETF	0.395	1.000	0.981
	Leafy Vegetables	Weathering loss	-0.953	-0.947	-0.074
		Dry Interception	0.306	0.135	0.067

45 ^(a) ETF means Equilibrium Transfer Factor

46 **Table 12: Spearman correlation coefficients calculated (for both dry and wet scenarios) between**
 47 **the effective ingested dose of Cs-134/-137, I-131 and Sr-90 received by an adult (≥ 17 years old)**
 48 **and a young child (2 – 7 years old). Only parameters with highest impact on the distribution of**
 49 **effective doses are displayed.**

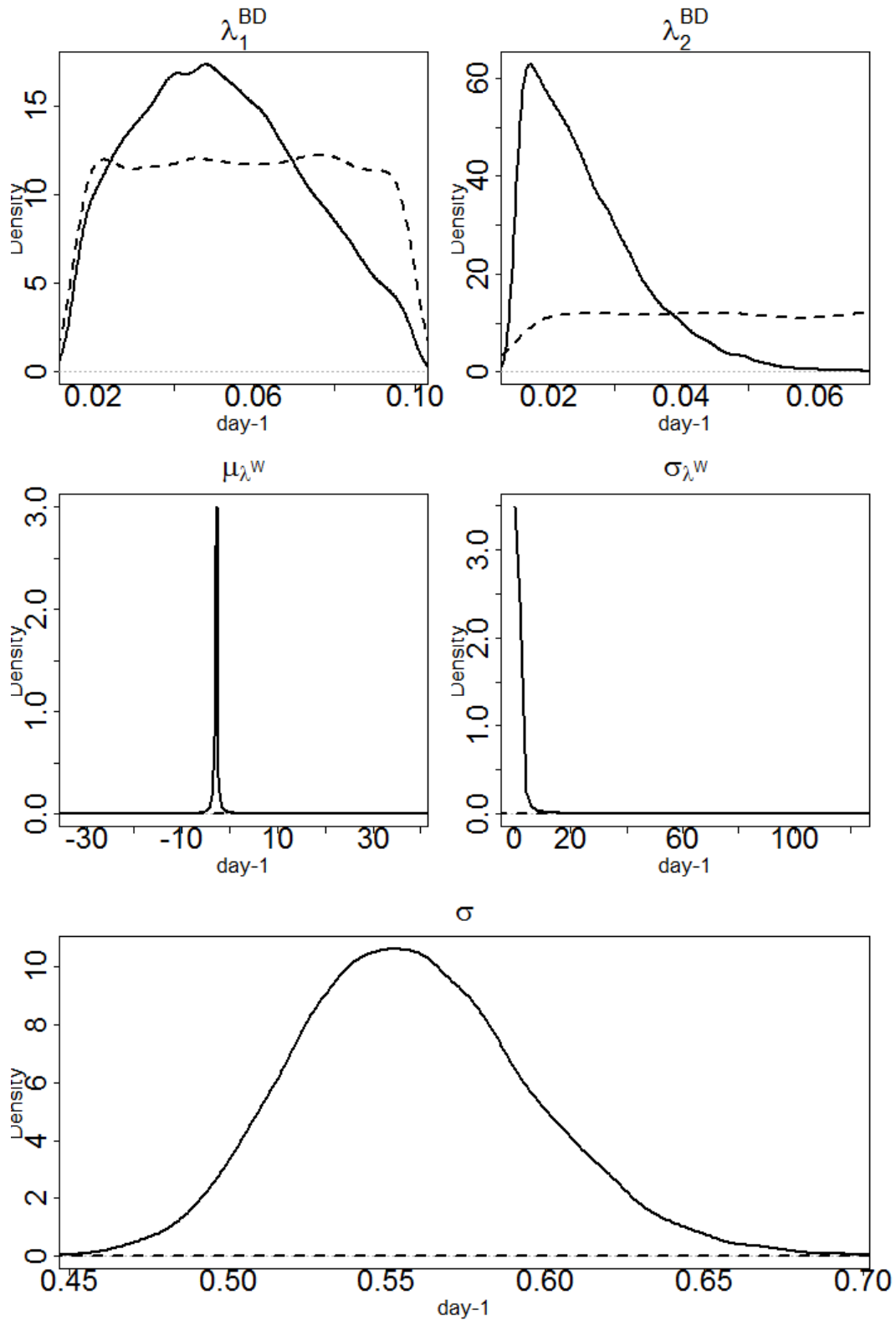
Isotope	Age Group	Input parameter ^(a)	04/09/2013	01/11/2013	01/08/2014
Cs-134/-137	Adult	Foliar Biomass – Grass	-0.544	-0.510	-0.134
		Weathering loss – Grass	-0.307	-0.335	-0.045
		Weathering loss - Leafy Vegetables	-0.307	-0.315	-0.061
		Dry Interception - Leafy Vegetables	0.313	0.201	0.054
		ETF - Cow Milk	0.463	0.454	0.705
	Young Child	Foliar Biomass – Grass	-0.514	-0.491	-0.134
		Weathering loss - Leafy Vegetables	-0.362	-0.375	-0.088
		Weathering loss – Grass	-0.285	-0.319	-0.044
		Dry Interception - Leafy Vegetables	0.357	0.232	0.067
		ETF - Cow Milk	0.431	0.418	0.685
I-131	Adult	Foliar Biomass – Grass	-0.466	-0.466	-0.466
		Weathering loss – Grass	-0.208	-0.213	-0.213
		Weathering loss - Leafy Vegetables	-0.186	-0.195	-0.195
		Dry Interception - Leafy Vegetables	0.194	0.193	0.193
		Dry Interception – Grass	0.209	0.209	0.209
		ETF - Cow Milk	0.732	0.727	0.728
	Young Child	Foliar Biomass – Grass	-0.452	-0.452	-0.452
		Weathering loss - Leafy Vegetables	-0.224	-0.234	-0.234
		Weathering loss – Grass	-0.207	-0.212	-0.212
		Dry Interception – Grass	0.204	0.204	0.204
		Dry Interception - Leafy Vegetables	0.237	0.236	0.236
		ETF - Cow Milk	0.713	0.709	0.709
Sr-90	Adult	Weathering loss - Leafy Vegetables	-0.627	-0.724	-0.435
		Foliar Biomass – Grass	-0.323	-0.284	-0.116
		Weathering loss – Grass	-0.164	-0.202	-0.117
		ETF - Cow Milk	0.298	0.316	0.757
		Dry Interception - Leafy Vegetables	0.559	0.422	0.263
	Young Child	Weathering loss - Leafy Vegetables	-0.663	-0.765	-0.505
		Foliar Biomass – Grass	-0.272	-0.242	-0.108
		ETF - Cow Milk	0.257	0.271	0.701
		Dry Interception - Leafy Vegetables	0.586	0.443	0.300

50

^(a) ETF means Equilibrium Transfer Factor

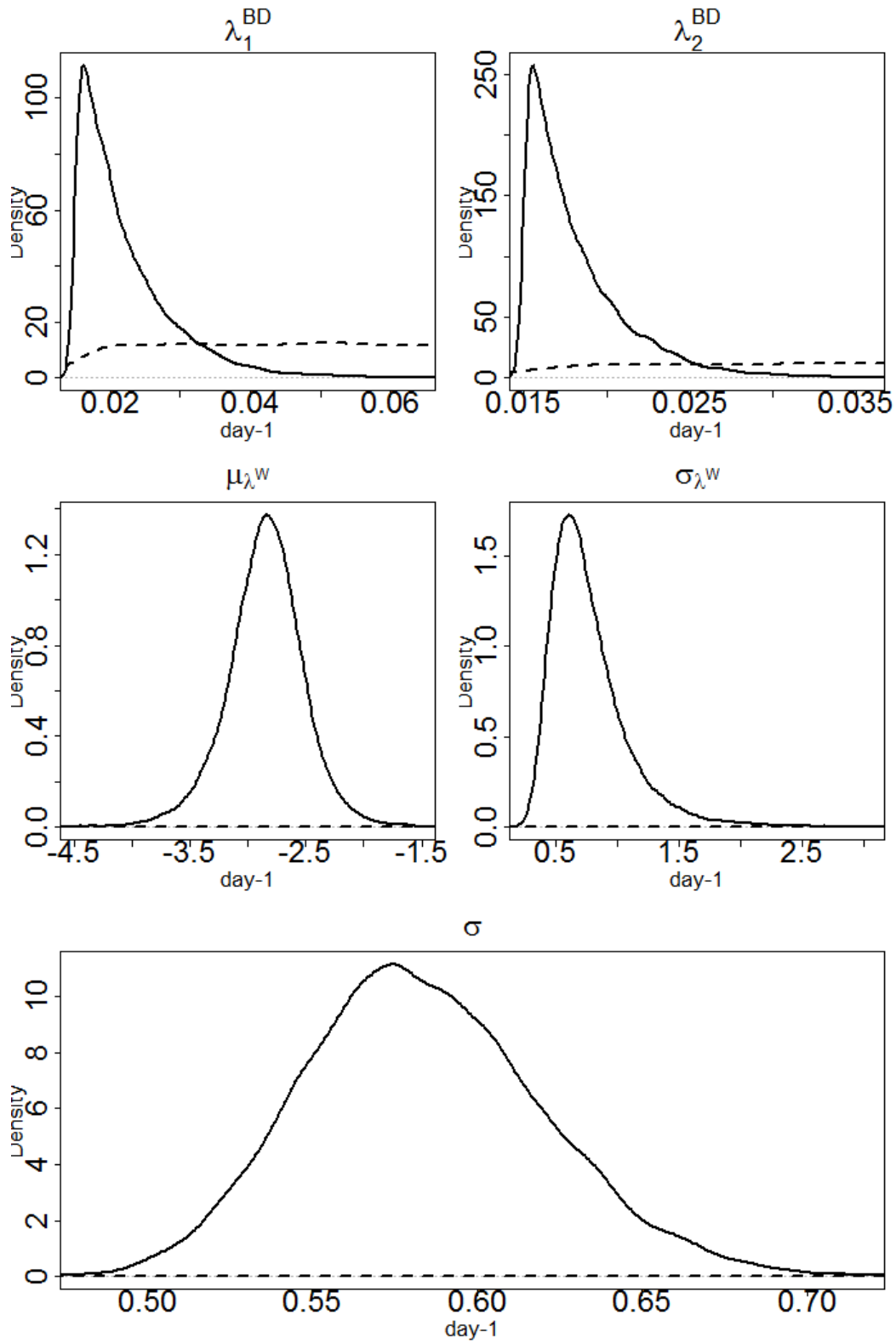
51 **Figures**

52 **Figure 1: Prior and posterior probability distributions of the parameters**
53 **$(\lambda_1^{BD}, \lambda_2^{BD}, \mu_{\lambda^W}, \sigma_{\lambda^W}$ and $\sigma)$ of the selected model MI₃. Dashed lines correspond to the priors and**
54 **the solid lines to the posteriors.**



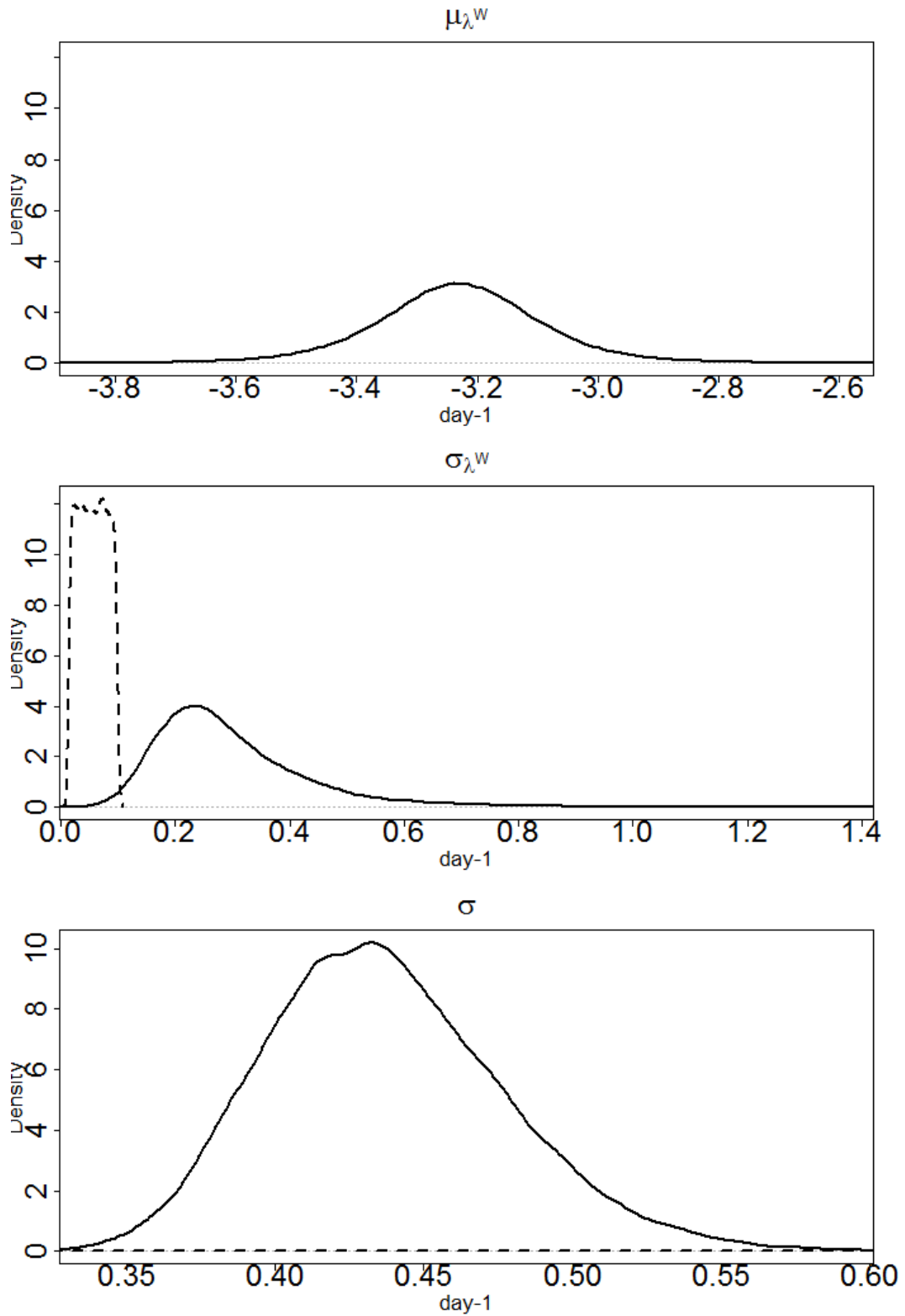
55

56 **Figure 2: Prior and posterior probability distributions of the parameters**
 57 **$(\lambda_1^{BD}, \lambda_2^{BD}, \mu_{\lambda^W}, \sigma_{\lambda^W}$ and σ) of the selected model MPF_2 . Dashed lines correspond to the priors**
 58 **and the solid lines to the posteriors.**



59

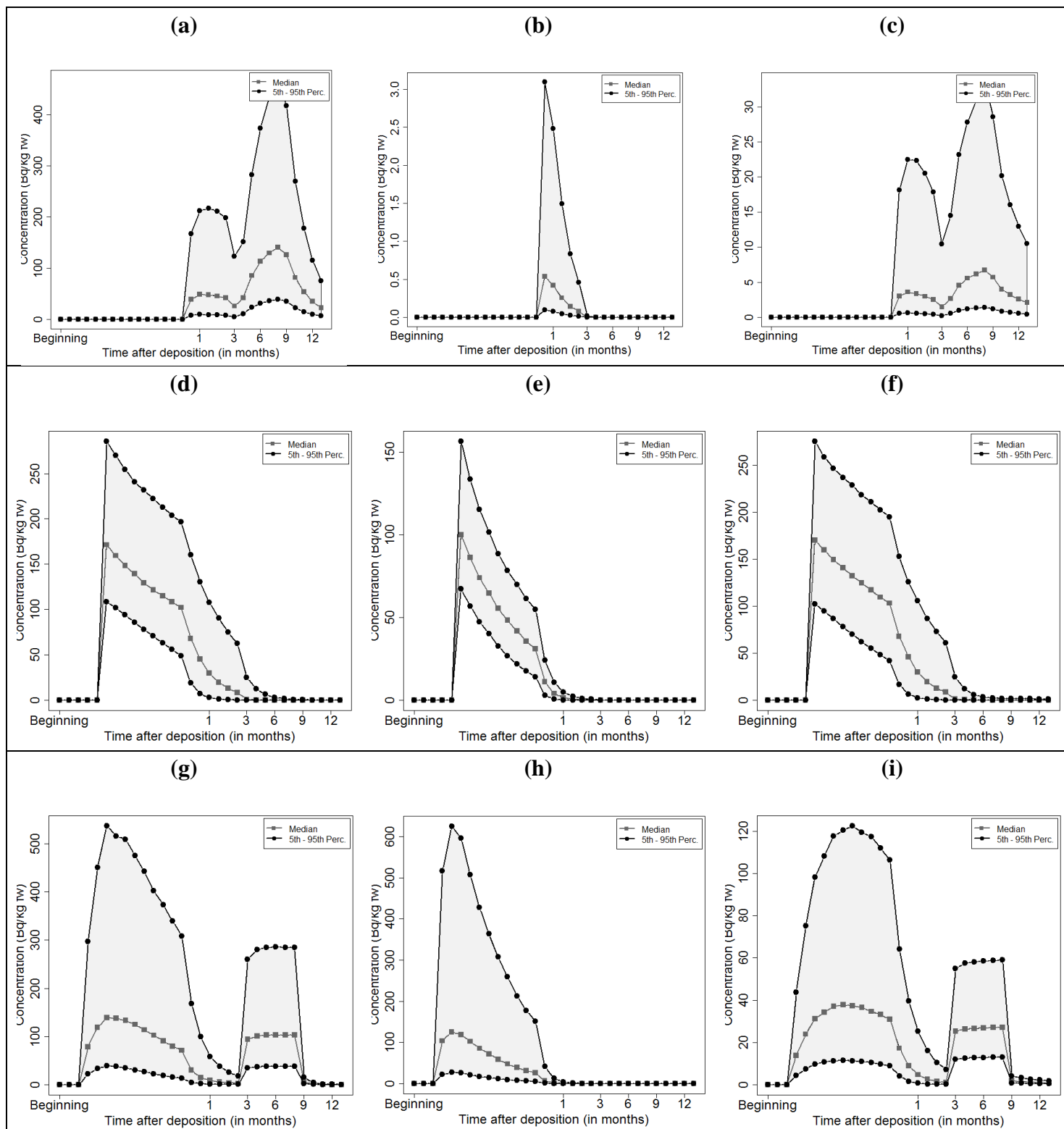
60 **Figure 3: Prior and posterior probability distributions of the parameters (μ_{λ^w} , σ_{λ^w} and σ) of the**
61 **selected model MPL₂. Dashed lines correspond to the priors and the solid lines to the posteriors.**



62

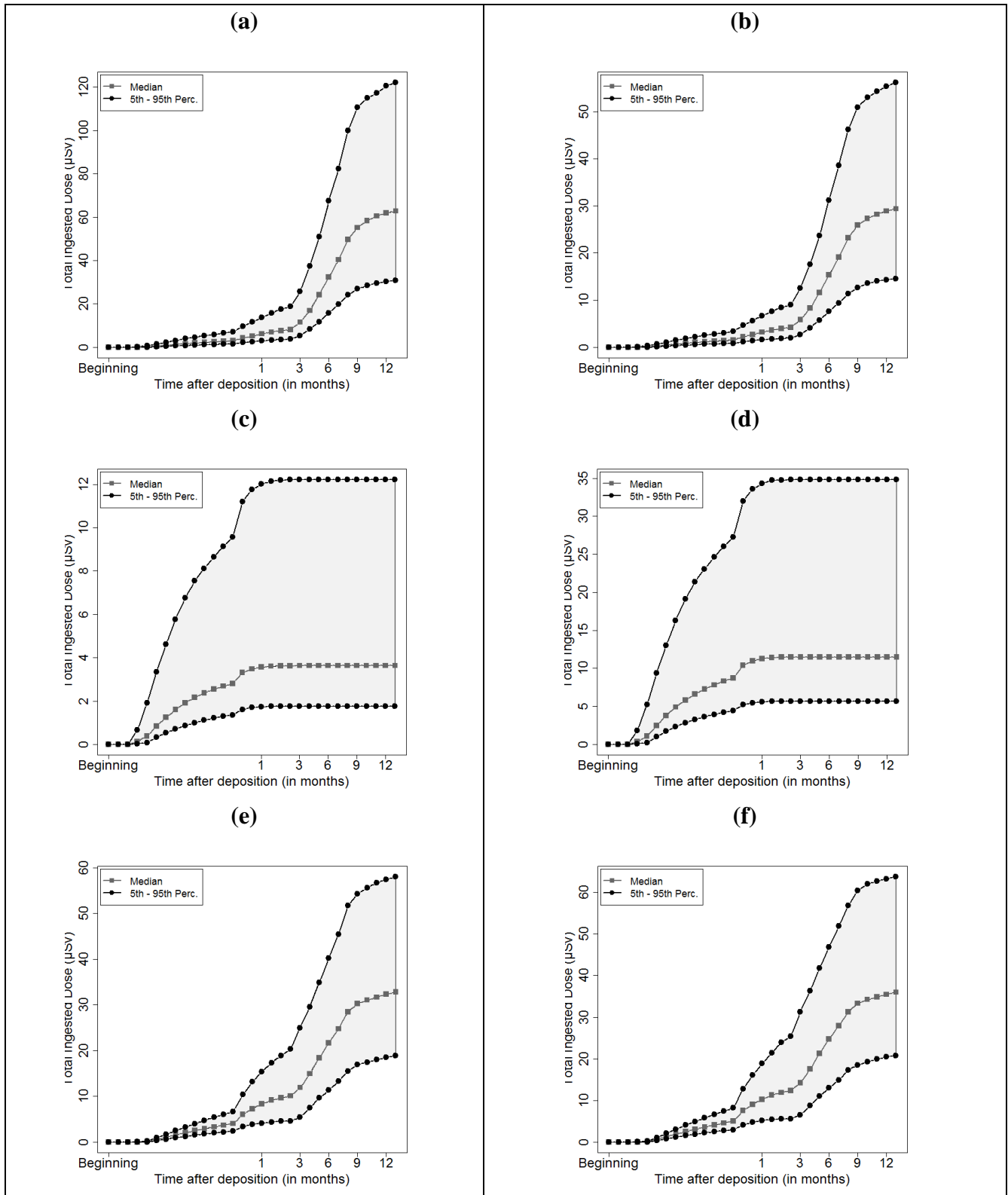
63
64
65
66
67

Figure 4: Evolution of the distribution of the concentrations (in Bq.kg⁻¹ fw) of ¹³⁷Cs, ¹³¹I and ⁹⁰Sr in beef meat (a-b-c), in leafy vegetables (d-e-f) and in cow milk (g-h-i) from August 1st, 2013 to September 1st, 2014 (13 months after deposition). The 5th and 95th percentiles are represented by black curves (with black points) and the median by the grey curve (with grey squares). The grey bandwidth illustrates the dynamics of the uncertainty (90%CI) about the predicted concentration values.



68

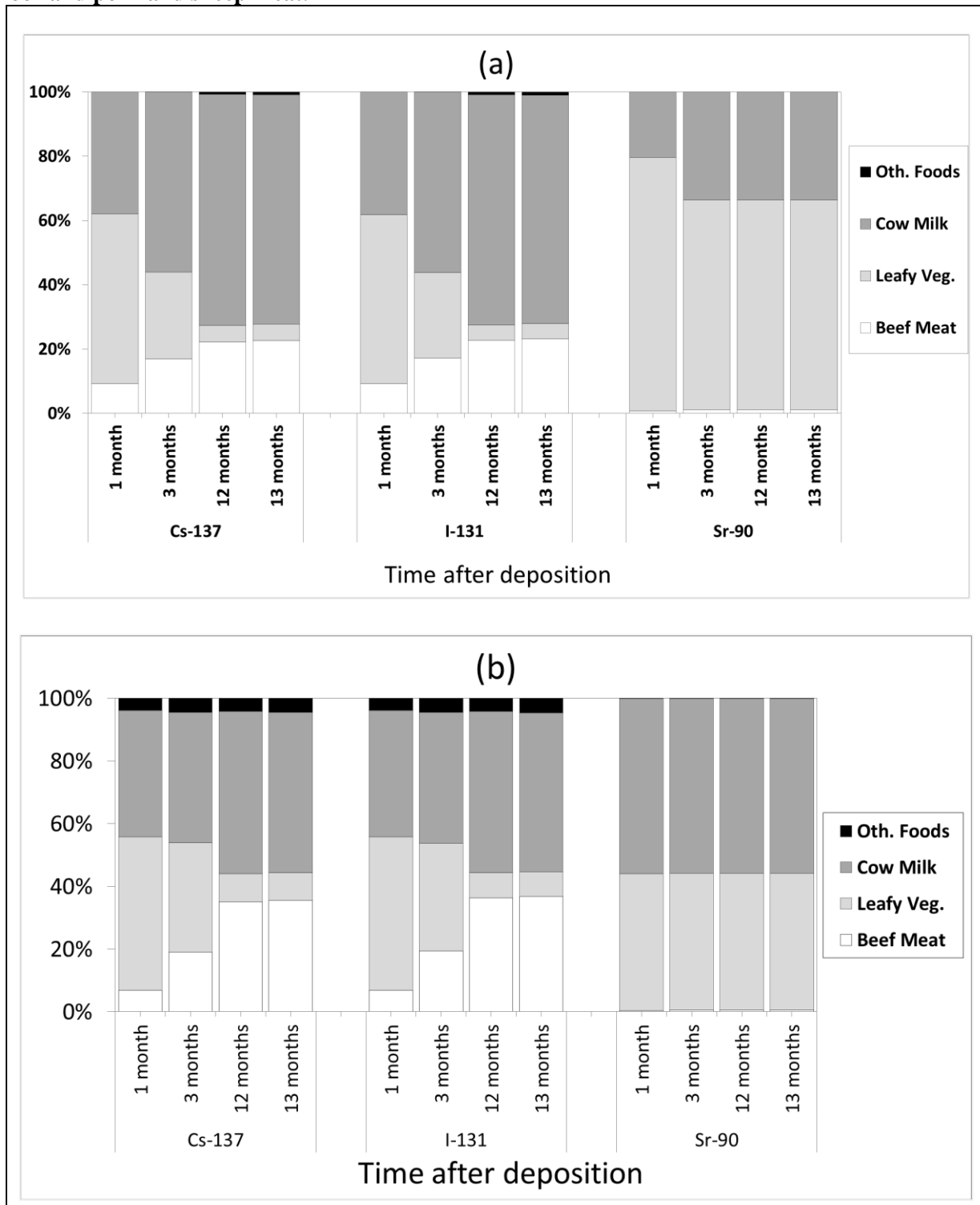
69 **Figure 5: Evolution of the effective doses (in μSv) of ^{137}Cs (a-b), ^{131}I (c-d) and ^{90}Sr (e-f) received by (a-**
 70 **c-e) an adult and (b-d-f) a young child predicted from August 1st, 2013 to September 1st, 2014 (13**
 71 **months after deposition). The 5th and 95th percentiles are represented by black curves (with black**
 72 **points) and the median by the grey curve (with grey squares). The grey bandwidth illustrates the**
 73 **dynamics of the uncertainty (90%CI) about the predicted effective doses.**



74

75
76
77
78

Figure 6: Contribution (expressed as a proportion) of food items on the average effective doses of ^{137}Cs , ^{131}I and ^{90}Sr received by (a) an adult and (b) a young child at four different dates after deposition. Cow Milk includes fresh and canned cow milk. Leafy Veg. includes fresh and canned leafy vegetables. Oth. Foods includes goat milk, flour and pork and sheep meat.



79

8.3 Résultats, limites et perspectives

Les analyses préliminaires et la méta-analyse conduites dans cette étude ont permis de mettre en évidence des différences de comportement entre l'iode (sous forme particulaire ou gazeuse) et les autres radionucléides, rejetés essentiellement sous la forme de particules. En effet, on observe une décroissance biomécanique plus rapide pour l'iode. Cette différence entre l'iode et les autres particules pourrait s'expliquer par une capacité moindre de l'iode à se fixer à la surface des plantes ; l'anion iodure étant repoussé par la charge négative de la surface des feuilles (cf. article reproduit à la [Section 7.2](#) du [Chapitre 7](#)), mais ceci reste à préciser.

Dans cette étude, il a également été observé une différence entre les particules étudiées en extérieur et soumises à des conditions naturelles et les particules étudiées en laboratoire. Cet écart sur la décroissance biomécanique s'explique non seulement par la prise en compte dans les observations de terrain de la dilution biologique (qui contribue à réduire l'activité massique), mais également de l'intervention d'autres mécanismes contribuant à accélérer la décroissance biomécanique, comme l'intervention du vent et de la pluie.

L'une des contributions intéressantes apportées par ce travail concerne la prise en compte dans les modèles hiérarchiques de l'incertitude et de la variabilité inter-espèces végétales sur la constante de décroissance par dilution biologique. La prise en compte explicite de ce paramètre apporte une réelle nouveauté car dans les différentes études considérées, sa valeur est simplement « diluée » dans la constante de décroissance biomécanique.

L'analyse d'incertitude et l'analyse de sensibilité basée sur la régression linéaire et le calcul de coefficients de corrélation de Spearman entre les sorties des modèles et les entrées incertaines ont mis en évidence, comme attendu, l'influence dominante du transfert foliaire (interception sèche et décroissance biomécanique) dans les premiers jours, voire mois, suivant le dépôt. L'influence des paramètres dictant la mobilité des radionucléides dans le sol et leur transfert *via* le système racinaire des plantes ne se manifeste que plusieurs mois après le dépôt. Les distributions de probabilité relatives à l'incertitude sur ces paramètres sont issues de [Simon-Cornu et al. \(2015\)](#) ; le transfert racinaire n'étant pas au cœur de ce travail de recherche. Cependant, nous avons conscience que pour une évaluation de l'exposition sur le long terme, les questions relatives à l'incertitude et la variabilité sur les processus de transfert racinaire d'une part, et à l'incertitude de modèle générée par la modélisation de la distribution

des radionucléides dans le sol d'autre part, restent ouvertes et nécessitent plus de développements.

Concernant le transfert au lait de vache et à la viande de bœuf, des facteurs de transfert sont utilisés et les distributions de probabilité sur ces facteurs sont principalement issues de [Simon-Cornu *et al.* \(2015\)](#). Cette approche, utilisant des facteurs de transfert, peut être discutée car elle ne permet pas de prendre en compte de façon explicite, mais elle le fait implicitement les caractéristiques physiologiques et métaboliques des animaux ; ce qui représente une source supplémentaire d'incertitudes de modèle et paramétrique. Ces questions sont discutées plus en détail dans le [Chapitre 10](#).

L'une des limites dans cette étude concerne la caractérisation de l'incertitude et de la variabilité interindividuelle sur les paramètres (consommation journalière, délai entre récolte et consommation, part d'autoconsommation, *etc.*) reflétant le comportement alimentaire de chaque classe d'âge. En effet, pour chaque produit alimentaire, une quantité moyenne (proposée par défaut dans la plateforme SYMBIOSE et basée sur des régimes alimentaires français) de produit alimentaire consommée chaque jour a été fixée. Cela constitue une limite suggérant des investigations complémentaires pour plus de réalisme. Dans le cadre des actions définies dans le groupe de travail "Human Food Chain", la définition de paramètres radioécologiques et de paramètres contextuels (régimes alimentaires, pratiques agricoles) caractéristiques des contextes spatio-temporels des différents pays (Espagne, Finlande, France, Norvège, Ukraine) constitue l'un des axes de recherche prioritaires.

Partie III : Propagation des incertitudes

9. Application à l'accident de Fukushima

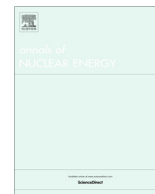
Dans ce chapitre, nous présentons une analyse d'incertitude et de sensibilité relative à la contamination des parcelles agricoles suite à un accident réel : l'accident de Fukushima. Une première analyse d'incertitude et de sensibilité a été introduite dans le [Chapitre 8](#). Dans cette étude de cas nous nous sommes principalement concentrés sur la contamination des légumes feuilles par les isotopes 134 et 137 du césium. Les objectifs principaux de cette étude étaient d'étudier l'impact de l'incertitude sur l'ensemble des paramètres radioécologiques décrivant les transferts foliaire et racinaire sur les niveaux de concentration de césium dans les légumes feuilles et d'identifier les paramètres les plus influents. Les travaux présentés dans la suite ont fait l'objet d'un article rédigé en collaboration avec Marc-André Gonze, Jean-Michel Métivier, Valérie Nicoulaud-Gouin et Marie Simon-Cornu. L'article qui a été accepté le 17 Décembre 2015 dans la revue *Annals of Nuclear Energy* est reproduit dans la [Section 9.1](#).

9.1 Article: Uncertainty analysis in post-accidental risk assessment models: An application to the Fukushima accident



Contents lists available at ScienceDirect

Annals of Nuclear Energy

journal homepage: www.elsevier.com/locate/anucene

Uncertainty analysis in post-accidental risk assessment models: An application to the Fukushima accident

Mouhamadou Moustapha Sy*, Marc-André Gonze, Jean-Michel Métivier, Valérie Nicoulaud-Gouin, Marie Simon-Cornu

Institut de Radioprotection et de Sûreté Nucléaire (IRSN), Laboratoire de Modélisation pour l'Expertise Environnementale (LM2E), Cadarache, Bâtiment 159, St Paul-lez-Durance 13115, France

ARTICLE INFO

Article history:

Received 31 July 2015

Received in revised form 9 December 2015

Accepted 17 December 2015

Available online xxx

Keywords:

Uncertainty

Sensitivity analysis

SYMBIOSE

Radioecology

Fukushima accident

ABSTRACT

Environmental contamination subsequent to the atmospheric releases during the Fukushima accident resulted in high radioactive concentrations in feed and foodstuffs. Producing a realistic health risk assessment after severe nuclear accidents, and developing a sufficient understanding of environmental transfer and exposure processes, appears to be a research priority. Specifically, the characterization of uncertainties in the human ingestion pathway, as outlined by the radioecological community, is of great interest. The present work aims to (i) characterize spatial variability and parametric uncertainties raised by the processes involved in the transfer of radionuclides (^{134}Cs and ^{137}Cs) after atmospheric releases during the Fukushima accident into the terrestrial ecosystems, and (ii) study the impact of these variability and uncertainties on radioactive contamination of leafy vegetables. The implemented approach quantified uncertainties under a probabilistic modelling framework. This resulted in probability distributions derived mainly from Bayesian inference and by performing transfer calculations in the modelling platform SYMBIOSE.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

The atmospheric releases following severe nuclear disasters, such as those at Chernobyl and Fukushima, may potentially generate huge environmental contamination that could subsequently contaminate feed and foodstuffs at high concentrations. Providing reliable risk assessments in accidental situations is of great importance. However, it is a complex issue for crisis managers since the transfer of radionuclides from the atmosphere to terrestrial ecosystems is governed by several miscellaneous processes which involve complex physical, chemical and biological mechanisms. Indeed, while released in the atmosphere either in a gaseous form or as particles, the radionuclides are deposited onto plant canopies as a result of gravitation, meteorological, and aerodynamic conditions. Subsequently, the fraction of radionuclides initially retained at the surface of foliage (agricultural crops and forest canopies) and the fraction of radionuclides on the ground represent a source of food chain contamination and a potential risk for human populations. Developing a sufficient understanding of the environmental transfer and the processes leading to exposure of human popula-

tions in a (post-) accidental context appears a priority for decision-makers during crisis management (Hinton et al., 2013). Specifically, the characterization of the uncertainties associated with these processes is one of the key research areas underlined by the radioecology community in order to enhance risk assessment models when studying accidental atmospheric emissions of radioactive elements and their impact on the human food chain. *Uncertainty* is defined in this article as the lack of perfect knowledge about the adequacy of a defined model to reflect the situation of concern for a given impact assessment (Cullen and Frey, 1999). *Uncertainty* is reducible if additional knowledge becomes available and it includes *model uncertainty* and *input uncertainty*. The latter, also called *parametric uncertainty* describes the lack of perfect knowledge about the values of parameters and variables of one model equations (Cullen and Frey, 1999). *Variability* is defined herein as the true heterogeneity inherent to the physical, chemical or biological processes of interest or the diversity within the scenario of concern (Simon-Cornu et al., 2015). *Variability* is not reducible with the supply of additional information.

International organizations such as the World Health Organization (WHO) and the Food and Agriculture Organization of the United Nations (FAO) have recommended quantifying uncertainty to obtain more realistic food risk assessments for many kinds of hazards (Sy et al., 2015). Within the context of environmental

* Corresponding author.

E-mail addresses: moustapha.sy@irsn.fr, mouhamadou.sy87@hotmail.fr (M.M. Sy).

transfer of radionuclides, the International Atomic Energy Agency (IAEA, 1989) provided similar recommendations and a great number of developments were devoted to providing reliable knowledge and a good representation of the *parametric uncertainty* of processes such as radionuclides (foliar) interception, “field loss” to ground or resuspension from ground (IAEA, 2010).

To characterize *parametric uncertainty*, probabilistic modelling is one of the broadly used approaches. The probabilistic framework offers theoretical bases with convenient and flexible tools for characterizing *parametric uncertainty*. However, even though some applications of the probabilistic modelling exist in the field of radioecology the use of probability distributions in dosimetric impact assessments is relatively new. As far as we know, in the few probabilistic food risk assessments that were published after the Fukushima accident (Evangelidou et al., 2014; Koizumi et al., 2012; UNSCEAR, 2013; Yamaguchi, 2012), uncertainties in the radioecological transfers was not the main focus.

The present study aimed to characterize uncertainties when modelling the human exposure to radiocesium isotopes (^{134}Cs and ^{137}Cs) through ingestion of leafy vegetables during an accidental context. As the leaves of leafy vegetables directly intercept the radioactive deposits we focused on the activity concentration of radiocesium isotopes (^{134}Cs and ^{137}Cs) in such vegetables/plants. This work mainly consisted of (i) implementing quantitative approaches to characterize the different parametric uncertainties raised by the physical, chemical and biological processes underlying the transfer of radionuclides into the terrestrial ecosystems after severe accidents and (ii) to apply those approaches in a real severe accident scenario: the environmental contamination subsequent to the Fukushima accident. To address the first aim (i), a probabilistic modelling framework was adopted by using Bayesian inference and this resulted in the derivation of probability distributions. Uncertainty and sensitivity analyses were applied to the environmental impact calculations by using Monte Carlo simulations in SYMBIOSE (which is a modelling platform developed by the French Institute of Radiation Protection and Nuclear Safety, IRSN) in response to the second objective (ii).

2. Material and methods

In this section, the modelling platform SYMBIOSE is presented and a description of the models which are used to predict the fate of radioactive cesium in agricultural lands and specifically to calculate the concentrations of radiocesium in leafy vegetables is provided.

2.1. SYMBIOSE modelling platform

SYMBIOSE is a simulation platform for assessing the fate and transport of radioactive pollutants in ecosystems, and their impact on humans (Gonze et al., 2011; Simon-Cornu et al., 2015). Environmental models implemented in SYMBIOSE address media, such as atmospheric, terrestrial, freshwater and marine systems, as well as the major transfer processes at their interfaces. Hundreds of components and interactions are accounted for in the system, most of which are modelled using a dynamical approach. Modelled exposure pathways include: external exposure (cloud shine and ground shine); internal exposure by inhalation (plume and resuspension from the ground), and by ingestion (agricultural products, freshwater fishes, seafood). This platform is flexible enough to deal with a wide range of situations, including post-accidental crisis predictions. Calculations can be performed for various spatial situations, from simple and generic to complex and site-specific landscape-level situations. SYMBIOSE also enables accounting for uncertainty in radioecological models in a probabilistic modelling framework

by using either Monte Carlo simulations or runs driven by Latin Hypercube sampling. Parametric uncertainty is defined via probability density functions (PDFs).

2.2. Modelling cesium transfer in leafy vegetables using SYMBIOSE: application to the Fukushima accident

Within the SYMBIOSE modelling platform, different modules are devoted to the transfer of radionuclides to plant canopies specific to different types of vegetation (annual crops, pasture grass and vegetable products). Specifically, the transfer of radionuclides to vegetable products such as leafy vegetables is governed by the combination of several processes that mainly include: the atmospheric deposition of radionuclides onto plant and ground surfaces; the interception of the deposited radionuclides by the leaves of plants; the field losses induced by the effect of wind, rain and the growth of plants; and the migration from ground to plant leaves. The different processes mentioned are further detailed in the sections that follow. Fig. 1 gives a conceptual illustration of the transfer of radioelements to agricultural lands, with emphasis on leafy vegetables.

2.2.1. Atmospheric deposition scenario

The spatial dispersion of ^{134}Cs and ^{137}Cs deposition and the estimation of dry deposit proportions within 80 km around the Fukushima Daiichi nuclear power plant (FDNPP) were deduced from the methodology proposed by Gonze et al. (2014). This method is based on the analysis of two kinds of field-based measurements of radiocesium deposits performed by the Japanese authorities:

- In situ spectrometry measures of bare soils in inhabited areas.
- Airborne gamma-ray spectrometry measures, which integrate radiation from the underlying land surfaces.

The methodology, its assumptions, and the conclusions drawn are further discussed by Gonze et al. (2014, 2015). These inputs, D^{Cs} (in $\text{Bq}\cdot\text{m}^{-2}\cdot\text{day}^{-1}$) the flux of radiocesium deposited on agricultural lands (soil + canopy), and K^{Cs} , the proportion of radiocesium deposited in dry conditions, i.e., the ratio between dry and wet deposits, are illustrated in Fig. 2.

2.2.2. Transfer models

Let $[\text{Cs}]_{\text{LV}}^{\text{Fol}}(t_H)$ and $[\text{Cs}]_{\text{LV}}^{\text{Root}}(t_H)$ denote the concentrations (expressed in kilobecquerels per kilogram of plant fresh weight, $\text{kBq}\cdot\text{kg}^{-1}$ fw) of radiocesium in the foliar and the root compartments of leafy vegetables at harvest (t_H). The contamination of leafy vegetables by radiocesium at harvest $[\text{Cs}]_{\text{LV}}(t_H)$ was modelled, accounting for both direct contamination through the foliar pathway and indirect contamination via root uptake as expressed below (Eq. (1)).

$$[\text{Cs}]_{\text{LV}}(t_H) = [\text{Cs}]_{\text{LV}}^{\text{Fol}}(t_H) + [\text{Cs}]_{\text{LV}}^{\text{Root}}(t_H) \quad (1)$$

2.2.2.1. Foliar pathway. The concentration of radiocesium in leafy vegetables induced by foliar uptake is given by the mass conservation equation taking into account the amount of cesium deposited and initially retained on the plant surface, the radioactive decay and the potential “field losses” due to the combined action of wind, rain and the growth of plants. This mass balance is given by the first-order differential equation below (2).

$$\frac{d}{dt}[\text{Cs}]_{\text{LV}}^{\text{Fol}} = D^{\text{Cs}} \times \left(K^{\text{Cs}} \times IF_d + (1 - K^{\text{Cs}}) \times IF_w(H^{\text{rain}}) \right) - (\lambda^{\text{FL}} + \lambda^{\text{CS}}) \times [\text{Cs}]_{\text{LV}}^{\text{Fol}} \quad (2)$$

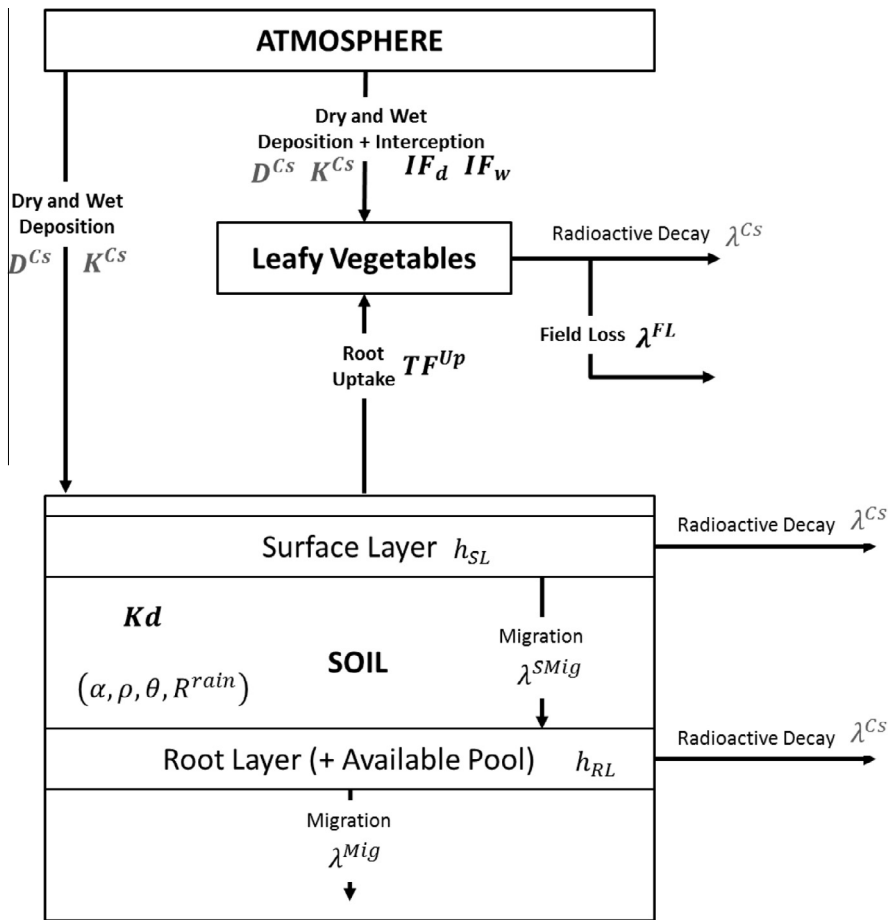


Fig. 1. Schematic illustration of the risk assessment model considered in this study with an emphasis to the transfer of radiocesium to leafy vegetables in an accidental context. The contextual parameters are represented in grey and the other parameters in black. Parameters for which probability distributions are defined are printed in bold (see Table 1).

$$IF_w(H^{rain}) = (IF_w(5) - IF_d) \times \frac{H^{rain}}{5} + IF_d \quad \text{if } 0 < H^{rain} < 5 \quad (3)$$

The parameter λ^{FL} corresponds to “field loss” due to the weathering effect of wind and rain and to the growth of plants (day^{-1}) and λ^{Cs} is the radioactive decay (in day^{-1}) specific to each considered isotope. IF_d and IF_w (expressed in $\text{m}^2 \cdot \text{kg}^{-1}$) are the mass interception fractions which depend on the total precipitation H^{rain} , occurring during the deposition events. H^{rain} is a meteorological input expressed in (mm). IF_d (dry fraction) corresponds to $H^{rain} = 0$ mm, whereas IF_w (wet fraction) is tabulated for H^{rain} values ranging from 5 mm to more than 15 mm, as described in Table 1, and interpolated between these values (Eq. (3)).

2.2.2.2. Root uptake. Let the soil be a superposition of two layers which both interact following migration mechanisms:

- The “surface layer” (SL), which is the interface between the atmosphere and the lower layer, and
- The “root layer” (RL) in which the roots of the plant get the nutrients necessary for plant growth. RL also includes an “available pool” (AP) representing the part of radiocesium available for the root uptake.

The concentration of radiocesium in the root compartment of leafy vegetables at harvest is derived from the concentration of radiocesium in the available pool as described below (Eq. (4)).

$$[Cs]_{LV}^{Root}(t_H) = f^{dw} \times TF^{Up} \times [Cs]_{AP}(t_H) \quad (4)$$

TF^{Up} which is defined by the IAEA (2010) as the ratio of the dry weight concentration in plants to the dry weight concentration in the specified soil layer (in kg dw of soil per kg dw of aboveground plant biomass) denotes the soil-to-plant transfer factor considered at equilibrium, and f^{dw} is the dry matter content of the plant (expressed in kg of dry weight aboveground biomass per kg of fresh weight aboveground biomass). This formulation postulates that the contamination is transferred “in one time” specifically at the harvest date and that an equilibrium is reached between the concentration in the leafy vegetables and the concentration in the available pool of the soil $[Cs]_{AP}(t_H)$. The migration of radiocesium through soil layers to the aerial part of plants is described in Annex A.

2.3. Uncertainty and sensitivity analyses using SYMBIOSE

Uncertainty analysis and sensitivity analysis are stepwise procedures which basically aim at characterizing the uncertainty on the considered outputs and at prioritizing the uncertain inputs respectively. Further details on both approaches are provided in Annexes B and C. One crucial step within both procedures refers to the definition of the probability distribution functions characterizing the input parameters’ uncertainty. The foliar pathway and the root uptake models presented above involve on one hand inputs that are “contextual”, i.e. specific to the accidental scenario (D^{Cs} and K^{Cs}) and on the other hand “radioecological” inputs that could be considered as common to any accidental scenario (IF_d , IF_w , λ^{FL} , λ^{Cs} , K_d , TF^{Up} , etc.).

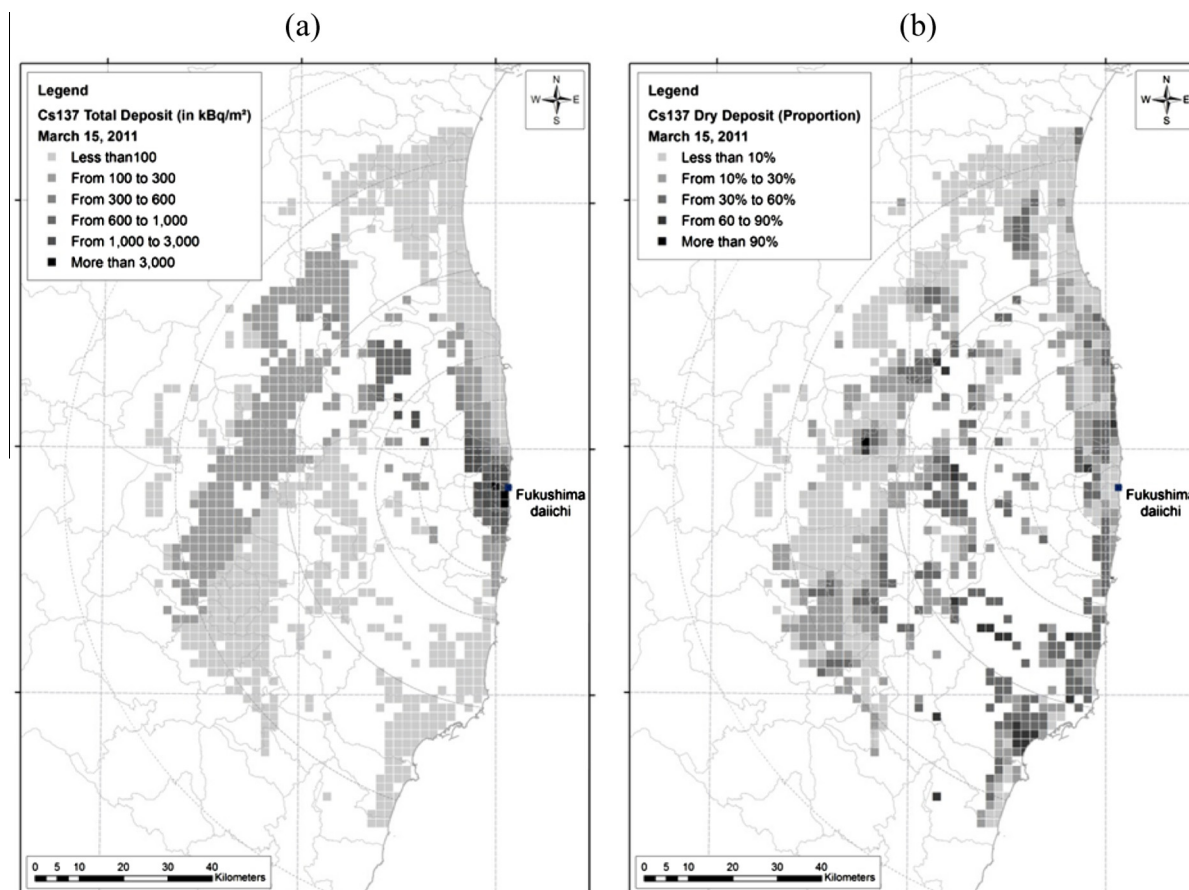


Fig. 2. Maps of (a) the total ^{137}Cs deposit on the agricultural areas expressed in $\text{kBq}\cdot\text{m}^{-2}$ within the 80 km around the FDNPP and (b) the dry deposit fraction (expressed in %).

2.3.1. Parametric uncertainty in models inputs

Addressing the uncertainty of the contextual inputs mentioned above is a complex issue since it requires accurate information (on climatic conditions during the deposition, land cover, agricultural practices, etc.) about the accident event. The collection of relevant data during the emergency phase is a cumbersome task due to the different constraints faced by the crisis managers.

Herein, we focus on the description of the approaches used to derive the uncertainty probability distributions assigned to the uncertain radioecological parameters introduced in Section 2.2.2: dry and wet interception fractions (IF_d and IF_w , respectively), the field loss rates (λ^{FL}), the soil distribution coefficient (Kd) and the soil-to-plant transfer factor (TF^{Up}).

Bayesian modelling was the approach implemented to define PDFs for dry and wet interception fractions. The Bayesian modelling is fundamentally based on the conditional probability theory and the Bayes Theorem (Bayes and Price, 1763). This approach can be used as a statistical estimation method. It combines prior knowledge (e.g. drawn from expert advice or historical data) and observed information (likelihood of the observed data given the model) on the parameter(s) of interest to produce an up-to-date knowledge of the considered parameter(s), which corresponds to a certain probability distribution that is called posterior (Parent and Bernier, 2007). Bayesian modelling was applied on data that were collected from both the published scientific literature and personal communications, and resulted in 145 data for dry interception and 335 data for wet interception. These databases include covariates such as aboveground fresh biomass values, precipitation amount (for wet interception), etc. Bayesian meta-analyses were then carried out on these two datasets separately. Performing a

meta-analysis under the Bayesian framework offers the possibility to derive PDFs, called posterior probability distributions, reflecting the uncertainty about unknown parameters of interest given the different combined datasets (Sy et al., 2015). Thus, posterior probability distributions of interception fractions were produced for leafy vegetables satisfying the specifications in SYMBIOSE. Indeed, in SYMBIOSE the (dry and wet) interception fractions for leafy vegetables are expressed as the mass interception fraction. As a consequence, log-normal distributions were assigned to the dry and wet mass interception fractions on the basis of the works presented by Sy et al. (2015) and Gonze and Sy (in preparation), respectively (Table 1).

For the field loss parameter λ^{FL} , triangular distributions were fitted for leafy vegetables on the basis of data collected from the literature. These triangular distributions were derived on the basis of field loss data associated to different leafy vegetables species (broccoli, cabbage, Chinese cabbage and lettuce) and published by IAEA (1996, 2009) and Miller and Hoffman (1983). In most data, the field loss parameter is expressed as field loss half-life (expressed in days); then a conversion to the field loss rate was applied (expressed in day^{-1}). Thus, we obtained a triangular T (min = 0.02, max = 0.09, mode = 0.04).

The probability density functions allocated to the parameters Kd and TF^{Up} were essentially derived from the probabilistic database of radioecological parameters defined in SYMBIOSE that is described in details by Simon-Cornu et al. (2015). This database is mainly drawn from the IAEA TRS472 technical report (IAEA, 2010) in which a set of radioecological parameter values (and if available some related probability distributions) integrating miscellaneous characteristics of soil, type of vegetation, etc. are

Table 1

Definition of the models input parameters (related to ^{137}Cs). Specifications for the Morris method are defined. The uncertainty characterization for the (considered) uncertain input parameters is specified by probability density functions and the references are provided.

Model	Input variable/parameter	Symbol	Rain amount (mm)	Unit	Value/probability distribution	Morris range	Source
Contextual parameters	^{137}Cs radioactive decay	λ^{Cs}	–	day^{-1}	6.32E–05	–	Nucleonica GmbH (2014)
	Total ^{137}Cs deposit	D^{Cs}	–	$\text{kBq}\cdot\text{m}^{-2}$	See Fig. 2(a)	–	Gonze et al. (2014)
	^{137}Cs dry deposit proportion	K^{Cs}	–	–	See Fig. 2(b)	–	Gonze et al. (2014)
	Rainfall height (during deposition)	H^{rain}	–	mm	2.45	–	JMA (2013)
	Effective rainfall (excluding deposition)	R^{rain}	–	$\text{mm}\cdot\text{day}^{-1}$	Based on the average daily rainfall heights	–	JMA (2013)
Foliar pathway	Dry interception fraction	IF_d	0	$\text{m}^2\cdot\text{kg}^{-1}$	$\text{LogN}(-1.282, 0.349)$	[0.14; 0.55]	Sy et al. (2015)
	Wet interception fraction	IF_w	5	$\text{m}^2\cdot\text{kg}^{-1}$	$\text{LogN}(-2.995, 0.875)$	[0.009; 0.278]	Gonze and Sy (In preparation) and Kinnersley et al. (1997)
			10	$\text{m}^2\cdot\text{kg}^{-1}$	$\text{LogN}(-3.136, 0.863)$	[0.008; 0.236]	
			15	$\text{m}^2\cdot\text{kg}^{-1}$	$\text{LogN}(-3.252, 0.872)$	[0.007; 0.214]	
Field loss parameter	λ^{FL}	–	day^{-1}	$T(0.02, 0.04, 0.09)$	[0.02; 0.09]	IAEA (1996, 2009), Miller and Hoffman (1983)	
Root uptake	Dry matter content	f^{dw}	–	–	0.09	–	IAEA (2010), Calmon (2012)
	Soil-to-plant transfer factor	TF^{LP}	–	–	$\text{LogN}(-2.804, 1.823)$	$[1.7 \cdot 10^{-3}; 2.16]$	Simon-Cornu et al. (2015)
	Surface layer height	h_{SL}	–	m	0.002	–	Expert Advice
	Soil distribution coefficient	Kd	–	$\text{L}\cdot\text{kg}^{-1}$	$\text{LogN}(7.09, 1.986)$	$[24.5; 5.88 \cdot 10^5]$	Simon-Cornu et al. (2015)
	Soil density	ρ	–	$\text{kg}\cdot\text{m}^{-3}$	1550	–	Expert Advice
	Root layer water content	θ	–	$\text{L}\cdot\text{m}^{-3}$	0.4	–	Beaugelin-Seiller and Mourlon (1999), Calmon and Mourlon (2009)
	Root layer height	h_{RL}	–	m	0.15	–	Calmon and Mourlon (2009)
	Proportionality coefficient for the RL – UL migration	α	–	–	10	–	Expert advice

compiled. Table 1 lists all the probability distributions related to the uncertainty of the considered radioecological parameters.

2.3.2. Numerical simulations and uncertainty propagation

Different numerical simulations using the Monte Carlo sampling method (with samples size set to 1000) were performed in the SYMBIOSE platform, in which ordinary differential equations are solved, using the 4th order Runge–Kutta method.

Deposits were modelled from March 15th, 2011, till April 1st, 2011, by assuming a constant daily deposit amount during this period. Transfers were modelled from March 15th, 2011, till March 15th, 2012. Parallel probabilistic simulation was performed using the same sample size in order to check the convergence of the simulation.

Uncertainty analysis calculations applied to the whole landscape were performed using the spatial ground deposits D^{Cs} illustrated in Fig. 1 and the spatial proportion of dry deposits K^{Cs} was assessed using the methodology developed by Gonze et al. (2014). The probability distributions allocated to the uncertain parameters and described in Section 2.3.1 and the deterministic values defined for the other parameters were also considered (Table 1). The sensitivity analyses were performed by computing Spearman correlation coefficients (SCC) and Morris indices for different dates and different meshes.

3. Results

3.1. Uncertainty analysis

In this section, we present the analysis of the uncertainty characterization regarding the levels of ^{137}Cs in leafy vegetables by providing a three-dimensional reading of the results (time, space, and parametric uncertainty). Fig. 3(a) and (b) respectively provide the 5th and 95th percentile of the spatial distribution of ^{137}Cs concentration values predicted on April 1st, 2011. Thus, each map considered separately gives an insight on the spatial variation of ^{137}Cs

concentration within the considered area for a given level of uncertainty. For both maps and then for both levels of uncertainty, the maximum values within the spatial distribution of ^{137}Cs concentrations predicted in leafy vegetables since April 1st, 2011, coincided to the last modelled fallout.

Table 2 summarizes the distributions of ^{137}Cs concentration values predicted at five different dates for three different meshes corresponding to the 5th, 50th and 95th percentiles of spatial variability at each of the five dates. These meshes are respectively located at the southern (Iwaki), south-eastern (Izumizaki village) and north-eastern (Iitate) parts of the considered area. A large variation of ^{137}Cs concentration values is observed between the three different meshes. Spatial variability appears right-skewed and constant over time.

Consider the variability ratio (VR), the uncertainty ratio (UR) and the overall uncertainty ratio (OvUR) as the measures of either variability or uncertainty combined. These ratios were introduced by Özkaynak et al. (2009) to characterize the overall variability and overall uncertainty combined about the outputs derived from a 2-dimensional Monte Carlo simulation. Herein, VR is the ratio of the 95th percentile of spatial variability to the central tendency (i.e., 50th percentile of uncertainty for the 50th percentile of spatial variability), UR corresponds to the ratio of the 95th percentile of uncertainty for the 50th percentile of spatial variability to the central tendency, and OvUR the ratio of the 95th percentile of uncertainty for the 95th percentile of spatial variability to the central tendency. The measures VR, UR and OvUR were calculated for five selected dates and are represented in Table 3.

The values of VR are estimated at approximately 12 for the five different dates which confirms that the spatial variability of concentrations in leafy vegetables is simply explained by the initial spatial distribution of the ^{137}Cs deposits, themselves explained by the meteorological conditions (wind, precipitation, etc.) occurring during the deposition events.

Fig. 4 illustrates the evolution of the ^{137}Cs concentrations predicted in leafy vegetables over the considered time giving an additional insight on the related uncertainty at a local scale with an

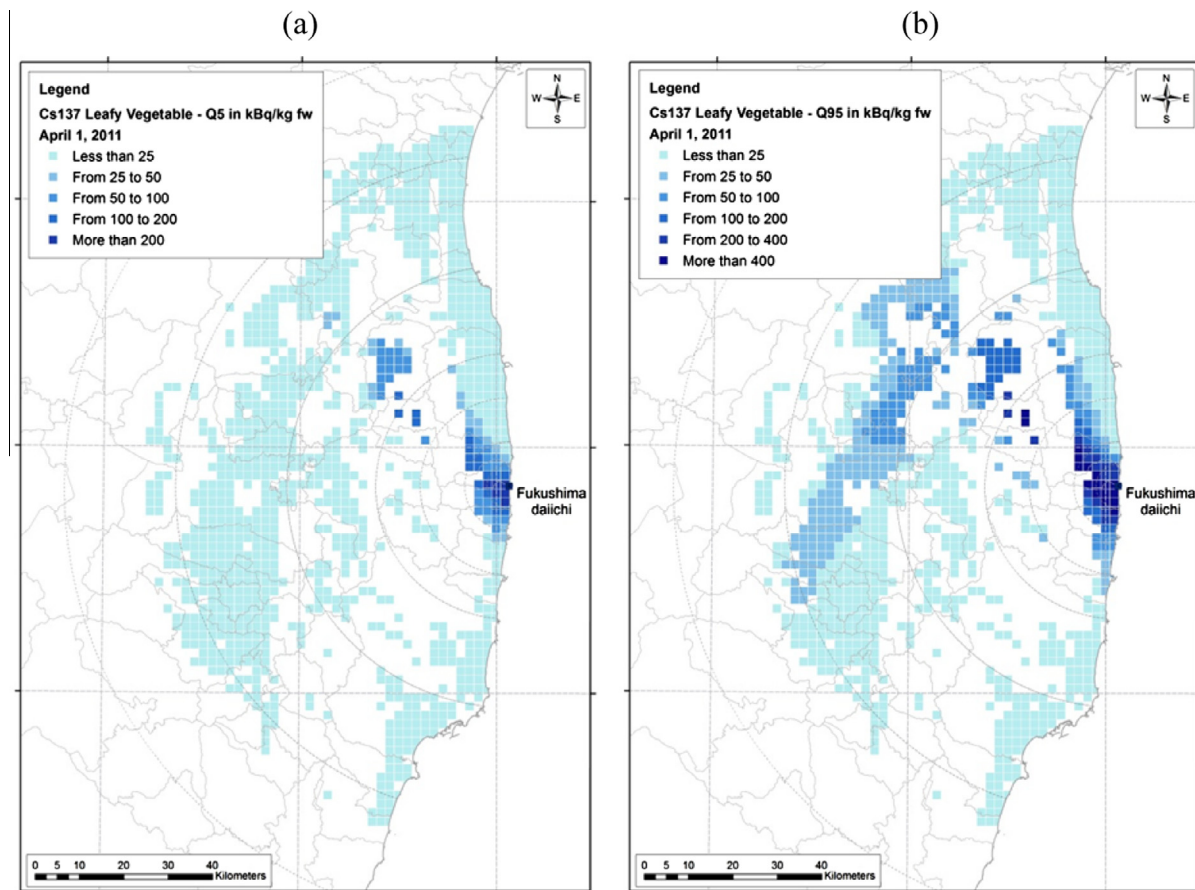


Fig. 3. Maps of (a) the 5th percentile of ^{137}Cs concentrations predicted in leafy vegetables on April 1st, 2011 within the 80 km around the FDNPP expressed in $\text{kBq}\cdot\text{kg}^{-1}$ fw and (b) the 95th percentile of ^{137}Cs concentrations predicted in leafy vegetables on April 1st, 2011.

Table 2

Uncertainty distributions of ^{137}Cs concentrations in leafy vegetables calculated at five different dates for the 5th, the 50th and the 95th percentiles of spatial variability.

Mesh ^a	Dates	5th Perc.	Median	95th Perc.
Min (F.Iwaki.12.34)	01/04/2011	1.36	2.38	3.79
	15/04/2011	0.58	1.21	2.26
	15/05/2011	0.07	0.29	0.86
	15/09/2011	$1.8 \cdot 10^{-5}$	$9.8 \cdot 10^{-4}$	$2.4 \cdot 10^{-2}$
	15/03/2012	10^{-6}	$3.1 \cdot 10^{-5}$	$9.5 \cdot 10^{-4}$
Med (F.Izumizaki.23.8)	01/04/2011	4.63	7.96	12.76
	15/04/2011	2	4.10	7.67
	15/05/2011	0.21	0.97	2.92
	15/09/2011	$5.6 \cdot 10^{-5}$	$3.3 \cdot 10^{-3}$	$8 \cdot 10^{-2}$
	15/03/2012	$2 \cdot 10^{-6}$	$9.2 \cdot 10^{-5}$	$3 \cdot 10^{-3}$
Max (F.litate.54.26)	01/04/2011	53.84	93.29	149.12
	15/04/2011	23.08	47.92	89.37
	15/05/2011	2.55	11.29	34.43
	15/09/2011	$6.9 \cdot 10^{-4}$	$3.9 \cdot 10^{-2}$	$9.1 \cdot 10^{-1}$
	15/03/2012	$2.6 \cdot 10^{-5}$	$1.1 \cdot 10^{-3}$	$3.7 \cdot 10^{-2}$

^a Min, Med and Max = 5th, 50th and 95th percentiles of (spatial) variability.

emphasis on the mesh corresponding to the 95th percentile of spatial variability which is located in Iitate. This area is of interest as high values of ^{137}Cs concentrations are predicted and it was included in the second evacuation plan managed by the Japanese authorities, which lasted approximately two months after the deposition started on March 2011. Specifically, interest lies in the fact that human populations living in this area during these two months were likely to have ingested highly contaminated leafy vegetables. A median concentration at, approximately

Table 3

Measures of the predicted variability and uncertainty calculated at five different dates.

Dates	Variability ratio ^a	Uncertainty ratio ^b	Overall uncertainty ratio ^c
01/04/2011	11.72	1.60	18.73
15/04/2011	11.69	1.87	21.79
15/05/2011	11.64	3.01	35.43
15/09/2011	11.81	24.08	276.60
15/03/2012	11.96	32.58	396.68

^a Variability ratio = ratio of 95th percentile of spatial variability to central tendency (50th percentile of uncertainty of the 50th percentile of spatial variability).

^b Uncertainty ratio = ratio of the 95th percentile of uncertainty of the 50th percentile of spatial variability to central tendency.

^c Overall uncertainty ratio = ratio of the 95th percentile of uncertainty of the 95th percentile of spatial variability to central tendency.

93 $\text{kBq}\cdot\text{kg}^{-1}$ fw with a 90% confidence interval (CI) (54, 150) is predicted. One month and two months after the deposition started the concentrations predicted had decreased with median concentrations at approximately 50 and 11 $\text{kBq}\cdot\text{kg}^{-1}$ fw with 90% CI (23, 90) and (2, 35), respectively. One year after, the predictions are considerably lower, around 1 $\text{Bq}\cdot\text{kg}^{-1}$ fw. Between April 1st and May 15th 2011, UR and OvUR values are doubled and vary between 1.6–3 and 18.73–35.43 reflecting a rather small uncertainty. One year after (around April 15th 2012) UR and OvUR values reach approximately 33–400 describing a kind of uncertainty “increase” knowing that the variability remains stable over the time. This uncertainty “increase” is explained by the uncertainty about the slope (field loss values).

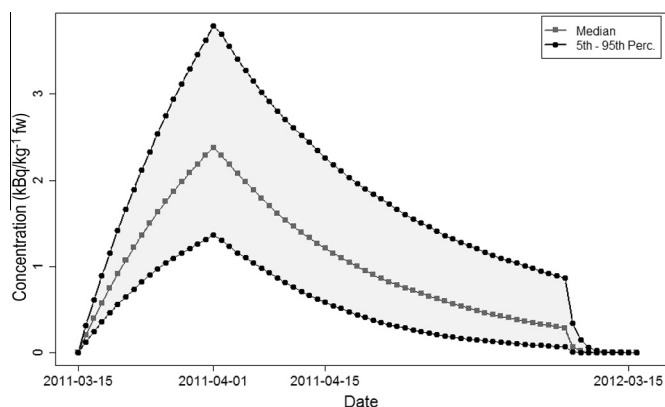


Fig. 4. Evolution of the distribution of ^{137}Cs concentrations (in $\text{kBq}\cdot\text{kg}^{-1}\text{fw}$) predicted in leafy vegetables in one mesh located in the city of Iitate from March 15th, 2011 to March 15th, 2012. The 5th and 95th percentiles are represented by black curves and the median by the grey curve. The grey bandwidth illustrates the evolution over the time of the uncertainty (90%CI) about the ^{137}Cs concentration values.

3.2. Sensitivity analysis

The SCC between the uncertain input parameters and the ^{137}Cs concentration values in leafy vegetables and the Morris indices (μ^* , σ) associated to each input parameter were calculated for five different dates and for all the meshes within the considered area. The SCC values and Morris indices calculated for the three meshes corresponding to the 5th, 50th and 95th percentiles of spatial variability are presented in Tables 4 and 5, respectively.

Two weeks after the deposition events started, the dry interception fraction, the wet interception fraction at a rain height of 5 mm, and the field loss parameter are the most influential input parameters on uncertainty with regards to the SCC values and Morris indices calculated on April 1st, 2011. This influence is confirmed since the sum of the square of SCC values is equal to about 0.94 meaning that approximately 94% of the variance of ^{137}Cs concentration values predicted on April 1st, 2011 is explained by these three parameters. On September, 2011 the field loss parameter and both interception factors are still the input parameters with the most influence on the uncertainty of ^{137}Cs concentrations in leafy vegetables. The SCC value calculated for the field loss parameter is negative and reflects the fact that the higher the field loss value the lower the concentration of ^{137}Cs in leafy vegetables. The positive SCC values obtained for the dry and the wet interception parameters say the higher the interception values the higher the concentrations of ^{137}Cs in leafy vegetables. In addition, the impact of the field loss parameter grows increasingly over time whereas the influence of the interception fractions decreases in time, as confirmed by SCC values and Morris indices calculated for March 15th, 2012. Indeed, the field loss parameter is still influential whereas the SCC values and the Morris indices of the (dry and wet) interception fractions are close to zero. In addition, parameters related to the transfer of radioactive materials via the root uptake, such as the soil distribution coefficient K_d and the soil-to-plant transfer factor TF^{Up} are the most influential parameters on the level of ^{137}Cs in leafy vegetables a few months after the accident. This confirms the known influence of the foliar pathway during the early phase following an accidental release of radionuclides in the atmosphere and the impact of the root transfer parameters in long-term contamination. Specifically, the dry and wet interception fractions and the field loss parameter have an important role during the early period following the accident by defining the initial contamination levels. The root transfer parameters (and also the field loss parameter) become as expected the

main route explaining the concentrations of ^{137}Cs in leafy vegetables grown during the post-accidental phase.

The SCC values and Morris indices for the interception fractions, which vary between the three selected meshes, show the spatial variability of the influence of such parameters within the 80 km around FDNPP, as also illustrated by Figs. 5 and 6. Indeed, the SCC values calculated for the dry interception fractions and the wet interception on April 1st, 2011, range from 0.81 to 0.94 and from 0.03 to 0.43, respectively. This variation is also observed one and two months after the deposition started with values ranging from 0.57 to 0.67 and from 0.28 to 0.34 for the dry interception and from 0.02 to 0.31 and 0.003 to 0.15 for the wet interception. However, one year after the accident the spatial variability associated to SCC values and Morris indices of the dry interception fraction decreases significantly, whereas the spatial variability associated to SCC values and Morris indices of the wet interception fraction remains similar.

The spatial variation of SCC values and Morris indices associated to both dry and wet interception fractions describe how ^{137}Cs concentrations depend on local meteorological conditions during deposition events, which is specifically characterized by the spatially variable proportion of dry deposits observed at a local scale. The interpretation of these results require some precautions because in meshes where the proportion of dry deposits is null (i.e., meshes where ^{137}Cs particles are deposited entirely by wet deposition, e.g. 50th and 95th percentiles of spatial variability) the dry interception fraction appears paradoxically as the more influential parameter. Considering a rain height equal to 2.45 mm for such meshes, the influential parameter is the wet interception fraction resulting from the interpolation between the dry ($H^{rain} = 0$ mm) and wet interception fraction tabulated for a rain height of 5 mm (see Section 2.2.2)

In contrast to the dry and wet interception fractions, SCC values and Morris indices calculated for the field loss parameter, the soil distribution coefficient and the soil-to-plant transfer factor show low variation over the 80 km around the FDNPP. For instance, on April 1st, 2011, SCC values of the parameters λ^{Fl} , K_d and TF^{Up} range from -0.32 to -0.30 , -0.06 to -0.04 and 0.01 to 0.02 , respectively. Likewise, the impact of these parameters on ^{137}Cs levels in leafy vegetables is spatially homogeneous one, two, and six months, and one year after the deposition started.

4. Discussion

The work presented here is integrated in the framework of guidelines initiated by the radioecological community with regards to improving risk assessment models in accidental contexts (Hinton et al., 2013). Results have provided a simultaneous insight on not only the spatial and temporal variations of ^{137}Cs concentration levels in leafy vegetables within the area of concern but also an idea of the uncertainty, i.e., the range of radiocesium concentration values that could be expected for a given date and location during the post-accidental phase. Sensitivity analyses have illustrated the influence of the foliar pathway in the first two months. Consistency between Spearman correlation and the Morris method was verified.

The results obtained may give insight on the key features that describe the transfer of radionuclides in the terrestrial ecosystems following atmospheric releases and should guide the uncertainty characterization efforts of radiation protection and nuclear safety scientific communities. Further investigations in some key points are needed to improve assessment models and tools dedicated to nuclear accidents and their impact on the environment.

- The modelling assumption about the deposits dynamic (constant flux over two weeks) can be further investigated since defining punctual deposition events with variable deposit

Table 4
Spearman correlation coefficients between uncertain input parameters and cesium concentrations calculated at five different dates for the 5th the 50th and the 95th percentiles of spatial variability.

Mesh ^a	Uncertain parameters	Dates				
		01/04/2011	15/04/2011	15/05/2011	15/09/2011	15/03/2012
Min (F.Iwaki.12.34)	Dry interception	0.807	0.566	0.284	0.068	0.032
	Wet interception 5 mm	0.43	0.311	0.155	0.024	0.020
	Field-loss parameter	-0.311	-0.712	-0.927	-0.961	-0.305
	<i>Kd</i>	-0.060	-0.045	-0.035	-0.128	-0.649
	<i>TF^{Up}</i>	0.009	0.018	0.024	0.120	0.604
Med (F.Izumizaki.23.8)	Dry interception	0.865	0.602	0.301	0.073	0.033
	Wet interception 5 mm	0.314	0.224	0.110	0.011	0.017
	Field-loss parameter	-0.316	-0.719	-0.929	-0.964	-0.315
	<i>Kd</i>	-0.056	-0.042	-0.033	-0.121	-0.645
	<i>TF^{Up}</i>	0.011	0.021	0.025	0.115	0.600
Max (F.litate.54.26)	Dry interception	0.837	0.584	0.292	0.071	0.033
	Wet interception 5 mm	0.375	0.269	0.133	0.018	0.018
	Field-loss parameter	-0.314	-0.717	-0.928	-0.962	-0.310
	<i>Kd</i>	-0.059	-0.044	-0.034	-0.125	-0.648
	<i>TF^{Up}</i>	0.010	0.020	0.025	0.118	0.602

^a Min, Med and Max = 5th, 50th and 95th percentiles of (spatial) variability.

Table 5
Morris indices (μ^* , σ) of the uncertain input parameters and cesium concentrations calculated at five different dates for the 5th the 50th and the 95th percentiles of spatial variability.

Mesh ^a	Uncertain parameters	Dates									
		01/04/2011		15/04/2011		15/05/2011		15/09/2011		15/03/2012	
		μ^*	σ	μ^*	σ	μ^*	σ	μ^*	σ	μ^*	σ
Min (F.Iwaki.12.34)	Dry interception	0.45	0.09	0.29	0.16	0.17	0.20	0.13	0.22	0.05	0.08
	Wet interception 5 mm	0.29	0.06	0.21	0.11	0.14	0.14	0.11	0.16	0.04	0.06
	Field-loss parameter	0.27	0.10	0.49	0.21	0.57	0.36	0.52	0.49	0.19	0.19
	<i>Kd</i>	$4.5 \cdot 10^{-5}$	$5.5 \cdot 10^{-5}$	$1.7 \cdot 10^{-4}$	$2.1 \cdot 10^{-4}$	$6.7 \cdot 10^{-4}$	$8.1 \cdot 10^{-4}$	0.02	0.03	0.35	0.42
	<i>TF^{Up}</i>	$2.4 \cdot 10^{-5}$	$4.2 \cdot 10^{-5}$	$9.2 \cdot 10^{-5}$	$1.6 \cdot 10^{-4}$	$3.5 \cdot 10^{-4}$	$6.3 \cdot 10^{-4}$	0.01	0.02	0.18	0.32
Med (F.Izumizaki.23.8)	Dry interception	0.52	0.10	0.33	0.18	0.20	0.23	0.14	0.25	0.06	0.10
	Wet interception 5 mm	0.22	0.05	0.15	0.09	0.10	0.11	0.09	0.12	0.03	0.05
	Field-loss parameter	0.28	0.11	0.50	0.23	0.58	0.37	0.53	0.51	0.20	0.20
	<i>Kd</i>	$4.3 \cdot 10^{-5}$	$5.2 \cdot 10^{-5}$	$1.6 \cdot 10^{-4}$	$2.0 \cdot 10^{-4}$	$6.3 \cdot 10^{-4}$	$7.6 \cdot 10^{-4}$	0.02	0.03	0.34	0.41
	<i>TF^{Up}</i>	$2.3 \cdot 10^{-5}$	$4.0 \cdot 10^{-5}$	$8.7 \cdot 10^{-5}$	$1.5 \cdot 10^{-4}$	$3.3 \cdot 10^{-4}$	$5.9 \cdot 10^{-4}$	0.01	0.02	0.18	0.32
Max (F.litate.54.26)	Dry interception	0.48	0.09	0.31	0.17	0.19	0.21	0.13	0.23	0.05	0.09
	Wet interception 5 mm	0.26	0.05	0.18	0.10	0.12	0.13	0.10	0.14	0.04	0.05
	Field-loss parameter	0.27	0.11	0.50	0.22	0.57	0.37	0.52	0.50	0.20	0.19
	<i>Kd</i>	$4.4 \cdot 10^{-5}$	$5.3 \cdot 10^{-5}$	$1.7 \cdot 10^{-4}$	$2.0 \cdot 10^{-4}$	$6.5 \cdot 10^{-4}$	$7.9 \cdot 10^{-4}$	0.02	0.03	0.34	0.41
	<i>TF^{Up}</i>	$2.3 \cdot 10^{-5}$	$4.1 \cdot 10^{-5}$	$8.9 \cdot 10^{-5}$	$1.6 \cdot 10^{-4}$	$3.4 \cdot 10^{-4}$	$6.1 \cdot 10^{-4}$	0.01	0.02	0.18	0.32

^a Min, Med and Max = 5th, 50th and 95th percentiles of (spatial) variability.

amounts could lead to higher concentration in leafy vegetables. This question is not easy to tackle and remains broadly open to further developments. Readers interested in uncertainty and sensitivity analyses applied to atmospheric dispersion following Fukushima releases are invited to refer to [Evangelidou et al. \(2014\)](#) and [Girard et al. \(2014\)](#).

- The uncertainty about meteorological conditions during the deposition events, and specifically the rainfall height and its spatial variability, is still a delicate point ([Champion et al., 2013](#)). This lack of relevant information leads to uncertainty in modelling the nature of deposits (dry versus wet contributions) and its spatial variability. [Gonze et al. \(2015\)](#) have discussed the interest of developing operational methods, which improve knowledge on the contextual inputs (such as the nature of deposits), to enhance risk assessment during crisis management. In this study, uncertainty about all contextual parameters, including deposits and their nature (dry versus wet), was not modelled.
- The way the wet interception was modelled, with an interpolation between two values of H^{rain} , including 0 mm (and then the so-called “dry” interception factor), is to be improved. The capacity of plant canopies to retain the water, the interactions between the physical/chemical characteristics of the deposited materials and the peculiarities of plant foliage might also be important sources of uncertainty/variability ([Pröhl, 2009](#)). Models combining the water foliage capacities to retain the water and the absorption kinetics of radionuclides with different chemical characteristics need to be developed to better characterize interactions during the wet interception process, as discussed by [Gonze and Sy \(in preparation\)](#).
- Parameters such as the dry matter content of plants are also of great interest since it could strongly depend on the seasonality and on the characteristics of the plants of concern ([Müller and Pröhl, 1993](#)). The impact of uncertainty on this parameter was discussed by [Sy et al. \(2015\)](#). One way to account for the related uncertainty was proposed on the basis of values drawn from the TRS472 ([IAEA, 2010](#)), but this was not accounted for in this article.
- The fate of the foliar contamination after the interception is mostly governed by the field loss parameter, for which uncertainty characterization is still in progress. Moreover, the model chosen in this study neglects the fact that after a vegetative period (e.g. 60 days), all leafy vegetables that had intercepted the plume have been harvested, which interrupts the foliar pathway. Thus our model may overestimate the influence of the foliar pathway on the concentration values at medium term.

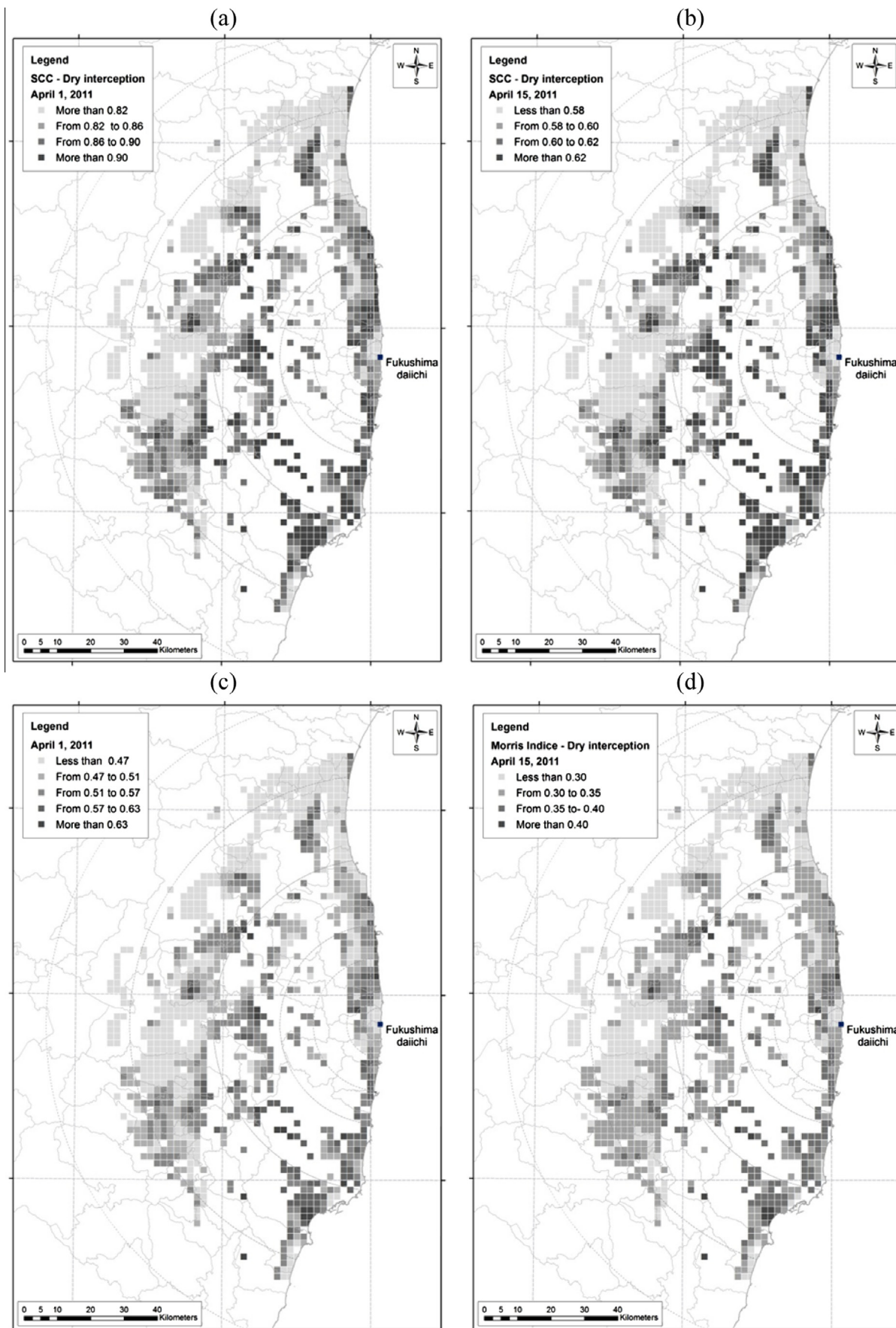


Fig. 5. Maps of the Spearman correlation coefficients ((a) and (b)) and maps of the Morris indices ((c) and (d)) calculated for the dry interception parameter on (a, c) April 1st, 2011 and (b, d) April 15th, 2011.

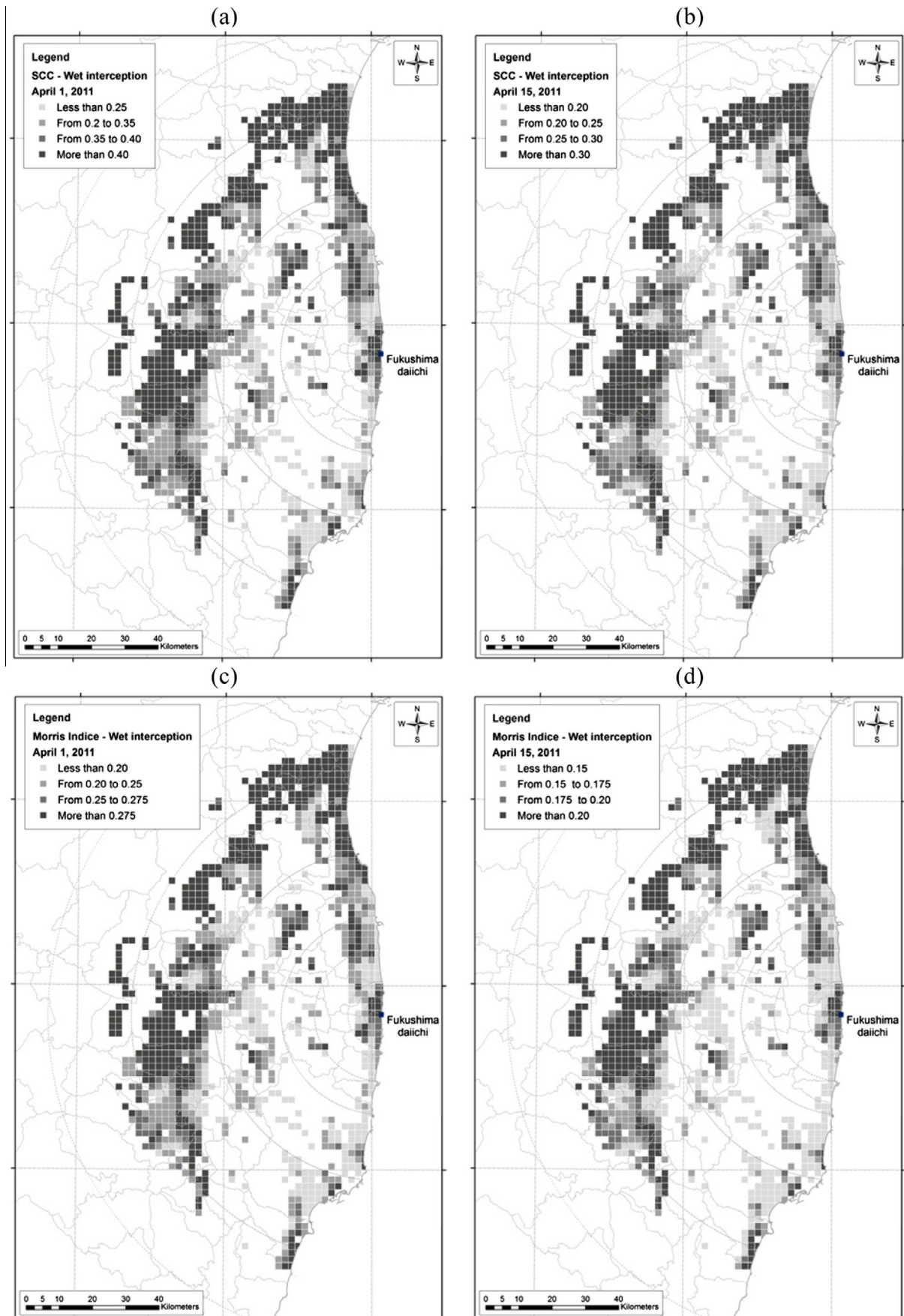


Fig. 6. Maps of the Spearman correlation coefficients ((a) and (b)) and maps of the Morris indices ((c) and (d)) calculated for the wet interception parameter (5 mm) on (a, c) April 1st, 2011 and (b, d) April 15th, 2011.

- Lastly, “contextual parameters” including, for example, the sheltering effect of greenhouse crops, agricultural practices, and the food diets of the human population, in the context of severe accidents were considered fixed, even if the obvious scenario uncertainty is still to be characterized.

Modelling and statistical methods, as used in this study, are proposed to be key in improving our predicting capacity. In parallel, evaluating the discrepancies between environmental monitoring data and risk assessment model outputs, e.g. concentrations in foodstuffs during the post-accidental phase, is also of great interest since it may give a strong insight on how far our model can identify the necessary developments.

Acknowledgements

The Region PACA (France) is acknowledged for financially supporting the PhD study of the first author. EDF (French electricity operator) is acknowledged for financially supporting the development of SYMBIOSE. The European commission is acknowledged for financially supporting the COMET project in which uncertainty analysis in post-accidental risk assessment models is discussed.

Annex

A. Migration of radiocesium through soil layers

Evaluating the concentration of radiocesium in leafy vegetables induced by the root uptake requires studying the characteristics of the soil on which those leafy vegetables are growing and the exchange mechanisms occurring within the soil. The mechanisms describing the migration of radiocesium between the surface layer and the root layer are defined by the following (first-order differential) equations.

$$[Cs]_{SL} = \frac{S_{SL}}{\rho \times h_{SL}} \quad (A-1)$$

$$[Cs]_{AP} = \frac{S_{AP}}{\rho \times h_{RL}} \quad (A-2)$$

$$\lambda^{Mig} = \frac{1}{h_{RL}} \times \frac{R^{rain}}{\theta + \rho \times Kd} \quad (A-3)$$

$$\lambda^{SMig} = \alpha \times \lambda^{Mig} \quad (A-4)$$

$$\frac{d}{dt} S_{SL} = D^{Cs} - \lambda^{SMig} \times S_{SL} - \lambda^{Cs} \times S_{SL} \quad (A-5)$$

$$\frac{d}{dt} S_{AP} = \lambda^{SMig} \times S_{SL} - \lambda^{Mig} \times S_{AP} - \lambda^{Cs} \times S_{AP} \quad (A-6)$$

S_{SL} and S_{AP} define the stocks in the surface layer and in the available pool of the root layer respectively (in mol.m^{-2}), ρ is the density of the considered soil (in kg.m^{-3}), h_{SL} and h_{RL} correspond to the height of SL and RL respectively, θ (L.m^{-3}) expresses the water content within RL, R^{rain} is a meteorological input (expressed in meters per time unit) describing the effective rainfall that reaches the ground, Kd is the soil distribution coefficient (in L.kg^{-1}), λ^{SMig} represents the migration rate from SL to RL (in day^{-1}) and λ^{Mig} is the migration rate (in day^{-1}) from the RL to the underlying layers (UL) that is assumed to be proportional to R^{rain} .

B. Uncertainty analysis

According to Monod et al. (2006), uncertainty analysis (UA) aims to answer the question: What is the level of uncertainty of the output(s) $Y = M(X)$ that is induced by the uncertainty of input (s) X ? The main objectives of UA, within the framework of risk assessment models, are then to identify and to characterize the various sources of uncertainties associated with the input parameters involved in the assessment models and to assess their impact on the output(s) by quantifying them, or if not possible by qualifying them. Probabilistic modelling is one appropriate approach since it offers the theoretical framework and the technical instruments for performing uncertainty analysis. Within a probabilistic framework, UA consists in determining the probability distribution of the output Y that is induced by the probability distributions of the input parameters. Monod et al. (2006) identified 4 steps as the classical way of implementing UA procedure when analytical calculation of the output probability distribution is not feasible. We consider below a three-step procedure as the 2nd and the 3rd step of Monod et al. (2006) can be merged.

- 1st step: Uncertainty quantification. This step is used to define the probability distribution functions associated with the input parameters. Despite the broad range of tools at our disposal, this is a fastidious point since it may include subjectivity (Monod et al., 2006). Indeed, the choice (or the definition) of the adequate probability distribution strongly depends on the available knowledge and on the judgment one have on the input parameters. Basically, this knowledge could be picked out from “historical-related” data or expert advice (Oakley and O’Hagan, 2007).
- 2nd step: Uncertainty propagation. Herein, the purpose is to generate for each input parameter random samples from the corresponding probability distribution function and to calculate the model outputs. Different sampling methods exist such as simple random sampling, Latin Hypercube sampling or space filling design methods. See IAEA (1989), Monod et al. (2006) and Pronzato and Müller (2012) for more details. The main difficulty of this step lies in the calculations step that could be time-consuming depending on model complexity and samples size.
- 3rd step: Output uncertainty characterization. This step aims to give either mathematical or graphical representations of the output uncertainty by generally computing descriptive statistics such as means, variance or the percentiles of the probability distributions.

C. Sensitivity analysis

The main objective of sensitivity analysis is to identify the input parameters which impact more the uncertainty on the output(s) (Saltelli et al., 2004). The general procedure of SA is similar to UA procedure but differs in the way that the 3rd step is dedicated in SA to the computation of “sensitivity indices” which measure the influence of uncertain parameters and enable to classify the influence of those parameters (Monod et al., 2006).

Several methods of SA exist and can be classified into two main approaches: local and global sensitivity analysis methods. Local sensitivity analysis consists in studying the variation in the output values that is induced by a perturbation of the inputs in the vicinity of a reference value (Girard et al., 2014; Monod et al., 2006). Local sensitivity analysis methods are based on the calculations of the derivatives of the model function in the neighborhood of the inputs X and enable to evaluate the behavior of the model around one or several reference values. Global sensitivity analysis aims, in contrast to local SA, at estimating the relative importance of the inputs

over their whole domain of variation (Girard et al., 2014; Saltelli, 2002). Hereafter, we focus on two types of global sensitivity analysis techniques: the Spearman correlation coefficient calculation that is a regression-based method included in the range of SA method based on the computation of a measure of importance and the Morris method which is a screening-based SA method.

C.1. Spearman correlation coefficient

Let Y and X_i ($i = 1, \dots, I$) be two vectors with the same length describing the output values and the values of the i th input parameter respectively. The Spearman correlation coefficient (SCC) between Y and X_i is described by the equation below.

$$\text{SCC}(Y, X_i) = \frac{\text{Cov}(R_Y, R_{X_i})}{\sqrt{\text{Var}(R_Y) * \text{Var}(R_{X_i})}} \quad (\text{C-1})$$

R_Y and R_{X_i} are the rank vectors associated to Y and to X_i respectively, cov and var are the covariance and the variance functions respectively. If the Spearman coefficient is equal or close to +1 (or -1) it means that there is a strong positive (or negative) relationship between the output and the input parameters. This reflects in both cases a certain sensitivity of the uncertainty of the output Y to the uncertainty of the considered input X_i . If the Spearman coefficient is zero or close to zero it means that there is a null (or weak) relationship between the output and the input. However some precautions are required when the results are analyzed since the Spearman correlation coefficient does not give insight on any cause-to-effect relationship. Sensitivity analyses are illustrated by “Tornado charts” which represent mainly horizontal bar plots of the Spearman correlation coefficients calculated between the output and input parameters of concern.

C.2. Morris method

Morris method aims to examine the changes in an output that are unambiguously attributed to the changes in individual inputs (Morris, 1991). The method is based on computing for each input a number of incremental ratios, called “elementary effects” by Morris (1991), which are then averaged to assess the overall importance of the input (Campolongo et al., 2007). Thus, the Morris method consists first on defining a design of experiment that is commonly a “one-at-a-time” (OAT) design in which each parameter varies once over its “discretized” domain of variation. Thereafter, the next stage consists in computing the elementary effects for each parameter by repeating r times OAT designs by defining the starting point of each OAT design randomly. This “randomization” was suggested by Morris (1991) in order to take into account, in the analysis, interactions which are effects that appear when two or more inputs vary simultaneously (Girard et al., 2014). Let $X = (X_1, \dots, X_K)$ be a set of K -dimensional vectors of the input parameters values, $M(X_1, \dots, X_K)$ be the response value and r be the number of random repetitions. The elementary effect related to the k th input parameter and the i th repetition, given an input variation δ , expresses as follows.

$$EE_k^i = \frac{M(X_1^i, \dots, X_k^i + \delta, \dots, X_K^i) - M(X_1^i, \dots, X_k^i, \dots, X_K^i)}{\delta} \quad (\text{C-2})$$

Once the elementary effects are calculated for each input parameter, the sensitivity analysis is then based on some descriptive statistics estimated from the distribution of elementary effect values. Morris (1991) proposed, as the two sensitivity indices defining the decision rule, the mean μ_k which assesses the overall influence of the parameter k on the output and the standard deviation σ_k that estimates the totality of high-order effects (i.e. interactions with other parameters and non-linearity)

(Nicoulaud-Gouin et al., 2014). Consequently given the values of μ_k and σ_k ($k = 1, \dots, K$) Morris method enables to identify three groups of input parameters: (i) parameters with negligible effect characterized by low μ_k value, (ii) parameters with linear effect without interactions having low σ_k and high μ_k values and (iii) parameters with non-linear and/or interactions effects corresponding to high μ_k and σ_k values. Campolongo et al. (2007) introduced μ_k^* (the average of the absolute value of the elementary effects) as an alternative refinement when the response is not monotone (Girard et al., 2014). A simultaneous examination of μ_k and μ_k^* is then required since the sign (positive or negative) of the effect is lost while μ_k^* is estimated (Campolongo et al., 2007).

Glossary

AP	available pool
CI	confidence interval
FAO	Food and Agriculture Organization of the United Nations
FDNPP	Fukushima Dai-Ichi Nuclear Power Plant
IAEA	International Atomic Energy Agency
IF	interception fraction
IRSN	French Institute of Radiation Protection and Nuclear Safety
OAT	one-at-a-time
OvUR	overall uncertainty ratio
PDF	probability distribution function
RL	root layer
SCC	Spearman correlation coefficients
SA	sensitivity analysis
SL	surface layer
TF	transfer factor
UA	uncertainty analysis
UL	underlying layers
UNSCEAR	United Nations Scientific Committee on the Effects of Atomic Radiation
UR	uncertainty ratio
VR	variability ratio
WHO	World Health Organization

References

- Bayes, M., Price, M., 1763. An essay towards solving a problem in the doctrine of chances. By the late Rev. Mr. Bayes, F.R.S. Communicated by Mr. Price, in a letter to John Canton, A.M.F.R.S. *Philos. Trans.* 53, 370–418. <http://dx.doi.org/10.1098/rstl.1763.0053>.
- Beaugelin-Seiller, K., Mourlon, C., 1999. Guide d'évaluation de l'impact de rejets radioactifs atmosphériques- édition corrigée et complétée. IPSN Report DPRE/SERLAB/99-011.
- Calmon, Ph., 2012. Modification des valeurs des paramètres de transferts racinaire et foliaire de SYMBOISE. IRSN Report PRP-ENV/SESURE/2012-10.
- Calmon, Ph., Mourlon, C., 2009. Equations et paramètres du logiciel ASTRAL v 2.3. IRSN Report DEI/SECURE/09-06.
- Campolongo, F., Cariboni, J., Saltelli, A., 2007. An effective screening design for sensitivity analysis of large models. *Environ. Modell. Software* 22, 1509–1518.
- Champion, D., Korsakissok, I., Didier, D., Mathieu, A., Quélo, D., Groell, J., Quentric, E., Tombette, M., Benoit, J.-P., Saunier, O., Parache, V., Simon-Cornu, M., Gonze, M. A., Renaud, Ph., Cessac, B., Navarro, E., Servant-Perrier, A.-C., 2013. The IRSN's earliest assessments of the Fukushima accident's consequences for the terrestrial environment in Japan. *Radioprotection* 48 (1), 11–37. <http://dx.doi.org/10.1051/radiopro/2012052>.
- Cullen, A.C., Frey, H.C., 1999. *Probabilistic Techniques in Exposure Assessment: A Handbook for Dealing with Variability and Uncertainty in Models and Inputs*. Plenum Press, New York, 336 pages.
- Evangelio, N., Balkanski, Y., Cozic, A., Moller, A.P., 2014. How “lucky” we are that the Fukushima disaster occurred in early spring predictions on the contamination levels from various fission products released from the accident and updates on the risk assessment for solid and thyroid cancers. *Sci. Total Environ.* 500–501, 155–172.

- Girard, S., Korsakissok, I., Mallet, V., 2014. Screening sensitivity analysis of a radionuclides atmospheric dispersion model applied to the Fukushima disaster. *Atmos. Environ.* 95, 490–500.
- Gonze, M.A., Sy, M.M., in preparation. Wet captation of airborne radionuclides by herbaceous plants: data review and modelling.
- Gonze, M.A., Mourlon, C., Garcia-Sanchez, L., Le Dizès, S., Nicoulaud, V., Gerber, P.P., Vermorel, F., 2011. SYMBIOSE: a simulation platform for conducting radiological risk assessments. In: International Conference on Radioecology and Environmental Radioactivity (ICRER). 19–24 June 2011, Hamilton, Ontario, Canada.
- Gonze, M.A., Renaud, Ph., Korsakissok, I., Kato, H., Hinton, T.G., Mourlon, C., Simon-Cornu, M., 2014. Assessment of dry and wet atmospheric deposits of radioactive aerosols: application to Fukushima Radiocaesium fallout. *Environ. Sci. Technol.* 48, 11268–11276.
- Gonze, M.A., Mourlon, C., Calmon, Ph., Manach, E., Debayle, C., Baccou, J., 2015. Modelling the dynamics of ambient dose rates induced by radiocaesium in the Fukushima terrestrial environment. *J. Environ. Radioact.* <http://dx.doi.org/10.1016/j.jenvrad.2015.06.003>.
- Hinton, T.G., Garnier-Laplace, J., Vandenhoove, H., Dowdall, M., Adam-Guillermin, C., Alonzo, F., Barnett, C., Beaugelin-Seiller, K., Beresford, N.A., Bradshaw, C., Brown, J., Eyrolle, F., Fevrier, L., Gariel, J.C., Gilbin, R., Hertel-Aas, T., Horemans, N., Howard, B.J., Ikäheimonen, T., Mora, J.C., Oughton, D., Real, A., Salbu, B., Simon-Cornu, M., Steiner, M., Sweeck, L., Vives, I., Battle, J., 2013. An invitation to contribute to a strategic research agenda in radioecology. *J. Environ. Radioact.* 115, 73–82.
- IAEA, 1989. Evaluating the reliability of predictions made using environmental transfer models Vienna. IAEA Safety Series 100, 106.
- IAEA, 1996. Modelling of Radionuclide Interception and Loss Process in Vegetation and of Transfer in Semi-natural Ecosystems: Second Report of the VAMP Terrestrial Working Group. IAEA-TECDOC-857. International Atomic Energy Agency, Vienna, Austria.
- IAEA, 2009. Quantification of Radionuclide Transfer in Terrestrial and Freshwater Environments for Radiological Assessments. IAEA-TECDOC-1616. International Atomic Energy Agency, Vienna, Austria.
- IAEA, 2010. Handbook of Parameter Values for the Prediction of Radionuclide Transfer in Terrestrial and Freshwater Environments. IAEA-TRS-472. International Atomic Energy Agency, Vienna, Austria.
- JMA, 2013. Japan meteorological agency: statistics weather information, <<http://www.data.jma.go.jp/obd/stats/etrn/index.php>>.
- Kinnersley, R.P., Goddard, A.J.H., Minski, M.J., Shaw, G., 1997. Interception of caesium-contaminated rain by vegetation. *Atmos. Environ.* 31 (8), 1137–1145.
- Koizumi, A., Harada, K.H., Niisoe, T., Adachi, A., Fujii, Y., Hitomi, T., Kobayashi, H., Wada, Y., Watanabe, T., Ishikawa, H., 2012. Preliminary assessment of ecological exposure of adult residents in Fukushima Prefecture to radioactive cesium through ingestion and inhalation. *Environ. Health Preventative Med.* 17, 292–298.
- Miller, C.W., Hoffman, F.O., 1983. An examination of the environmental half-time for radionuclides deposited on vegetation. *Health Phys.* 45 (3), 731–744.
- Monod, H., Naud, C., Makowski, D., 2006. Uncertainty and sensitivity analysis for crop models. In: Working with dynamic crop models: Evaluation, analysis, parameterization and applications. Elsevier Science, Amsterdam, The Netherlands, pp. 55–100.
- Morris, M.D., 1991. Factorial sampling plans for preliminary computational experiments. *Technometrics* 33 (2), 161–174.
- Müller, H., Pröhl, G., 1993. ECOSYS-87: a dynamic model for assessing radiological consequences of nuclear accidents. *Health Phys.* 64 (3), 232–252.
- Nicoulaud-Gouin, V., Garcia-Sanchez, L., Métivier, J.M., Gonze, M.A., 2014. Sensitivity analysis of a radionuclide transfer model describing contaminated vegetation in Fukushima prefecture, using Morris and Sobol' indices. In: 3rd International Conference on Radioecology & Environmental Radioactivity (ICRER) 7–12 September 2014, Barcelona, Spain.
- Nucleonica GmbH, 2014. Nucleonica Nuclear Science Portal (<<http://www.nucleonica.com/>>), version 3.0.49, Karlsruhe, Germany.
- Oakley, J.E., O'Hagan, A., 2007. Uncertainty in prior elicitation: a nonparametric approach. *Biometrika* 94, 427–441.
- Özkaynak, H., Frey, H.C., Burke, J., Pinder, R.W., 2009. Analysis of coupled model uncertainties in source-to-dose modelling of human exposures to ambient air pollution: a PM2.5 case study. *Atmos. Environ.* 43 (9), 1641–1649.
- Parent, E., Bernier, J., 2007. Le Raisonnement Bayésien: Modélisation et Inférence. Springer, Paris.
- Pröhl, G., 2009. Interception of dry and wet deposited radionuclides by vegetation. *J. of Environ. Radioact.* 100, 675–682.
- Pronzato, L., Müller, W.G., 2012. Design of computer experiments: space filling and beyond. *Stat. Comput.* 22, 681–701. <http://dx.doi.org/10.1007/s11222-011-9242-3>.
- Saltelli, A., 2002. Making best use of model evaluations to compute sensitivity indices. *Comput. Phys. Commun.* 145, 280–297.
- Saltelli, A., Tarantola, S., Campolongo, F., Ratto, M., 2004. Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models. John Wiley & Sons, Ltd., Chichester, 232 pages.
- Simon-Cornu, M., Beaugelin-Seiller, K., Boyer, P., Calmon, P., Garcia-Sanchez, L., Mourlon, C., Nicoulaud, V., Sy, M.M., Gonze, M.A., 2015. Evaluating variability and uncertainty in radiological impact assessment with the probabilistic database of SYMBIOSE. *J. Environ. Radioact.* 139, 91–102.
- Sy, M.M., Ancelet, S., Henner, P., Hurtevent, P., Simon-Cornu, M., 2015. Foliar interception of radionuclides in dry conditions: a meta-analysis using a Bayesian modelling approach. *J. Environ. Radioact.* 147, 63–75.
- UNSCEAR, 2013. Sources, effects and risks of ionizing radiation: volume I report to the general assembly and scientific annex A. <http://www.unscear.org/unscear/en/publications/2013_1.html>.
- Yamaguchi, I., 2012. Radioactive concentration of food caused by Fukushima nuclear power plant disaster and new radiological standards for foodstuffs in the existing exposure situation in Japan after a severe nuclear accident. *Jpn. J. Health Phys.* 47, 144–147.

9.2 Résultats, limites et perspectives

Dans cette étude, nous nous sommes intéressés à l'évaluation de l'exposition par consommation de légumes feuilles dans un périmètre de 80 km autour de la centrale nucléaire de Fukushima Dai-Ichi dans une période d'un an après l'accident survenu sur cette installation. Cette analyse d'incertitude a principalement porté sur la prédiction des niveaux de concentrations en césium dans les légumes feuilles. Il est important de noter que l'incertitude et la variabilité sur le facteur d'interception sèche ont été caractérisées à partir du modèle bayésien hiérarchique présenté au Chapitre 6 et que l'incertitude sur le facteur d'interception humide et sur la constante de décroissance biomécanique ont été déduites à partir des bases de données présentées dans les Sections 7.2 et 8.2.

Dans ce travail, la variabilité spatiale des dépôts et de la proportion de dépôt sec de césium dans les 80 km autour de la centrale de Fukushima est caractérisée. Les valeurs utilisées sont issues des estimations réalisées par (Gonze *et al.* 2014) basées sur des mesures aéroportées et au sol produites par les autorités japonaises quelques mois après l'accident. La caractérisation de cette variabilité spatiale nous a permis d'étudier, en différents points de la zone considérée et au cours du temps sur une période d'un an, l'influence combinée des paramètres radioécologiques et de la nature des dépôts. L'utilisation des méthodes d'analyse de sensibilité de Morris et les coefficients de corrélation de Spearman ont notamment permis d'identifier en chaque point de la zone d'intérêt les paramètres radioécologiques les plus influents (facteurs d'interception sèche et humide, constante de décroissance biomécanique, facteur de transfert racinaire, coefficient de distribution dans le sol) sur la prévision des niveaux de césium dans les légumes feuilles. Cependant, l'incertitude sur la quantité de césium déposée en chaque point n'a pas été quantifiée. La caractérisation de l'incertitude sur les dépôts constitue une des limites de ce travail car des variations importantes des valeurs de dépôt génèrent des variations importantes des niveaux de césium dans les légumes feuilles ; ce qui peut altérer considérablement la qualité des évaluations de l'exposition.

Cette étude met également en évidence l'influence des paramètres contextuels tels que la nature des dépôts (sec ou humide), les conditions météorologiques et la présence de serres sur les évaluations de risque et la nécessité de caractériser de manière robuste la connaissance ou l'absence de connaissance sur de tels paramètres. Nous nous sommes limités à la prédiction

des niveaux de concentrations de césium dans les légumes feuilles car peu de connaissances sur les comportements alimentaires des populations locales sont disponibles.

Partie IV: Conclusion Générale

10. Conclusion Générale

L'analyse des risques liés à l'exposition de l'homme aux polluants, et en particulier aux substances radioactives émises en situation accidentelle dans l'atmosphère, est généralement entachée d'*incertitude* et de *variabilité* qui peuvent altérer la qualité des prédictions des modèles d'évaluation de risque et, par conséquent, affecter la fiabilité du processus de prise de décisions. Quelques années après l'accident de Fukushima, les questions relatives à l'analyse des incertitudes et leur prise en compte dans le contexte de l'évaluation post-accidentelle des risques alimentaires sont plus que jamais d'actualité. A travers l'agenda de recherche stratégique (SRA pour Strategic Research Agenda en anglais), la communauté scientifique a identifié la modélisation des sources d'*incertitude* et de *variabilité* comme l'un des axes prioritaires de recherche en radioécologie pour les vingt prochaines années ([Hinton et al. 2013](#)). De manière plus spécifique, cette volonté globale, affichée par la communauté en radioécologie, d'améliorer les connaissances sur les processus de transfert de radionucléides dans l'environnement en réduisant les sources d'incertitude et en caractérisant la variabilité environnementale, se manifeste par exemple à travers les attentes relatives à l'exposition humaine par ingestion de denrées alimentaires contaminées, soulevées dans le cadre des actions définies par les groupes de travail du projet COMET ([Liland et al. 2013](#)).

L'objectif premier de mon projet de recherche était de développer une approche méthodologique pour la prise en compte des sources d'*incertitude* et de *variabilité* dans les modèles d'évaluation de risque alimentaire suite à un accident nucléaire majeur en vue d'améliorer les outils d'aide à la décision en situation accidentelle (*e.g.* la plateforme SYMBIOSE développée par l'IRSN). Ainsi, ce travail est une contribution permettant de lever quelques verrous sur la manière d'appréhender et de prendre en compte l'incertitude et la variabilité dans les modèles d'évaluation de risque en nous appuyant sur des méthodes statistiques robustes que nous jugeons appropriées.

10.1 Pluridisciplinarité et choix méthodologiques

La problématique dans ce projet de thèse est marquée par une transversalité qui s'explique globalement par :

- les différents compartiments de la biosphère impliqués à savoir les systèmes atmosphérique et agricole (sols agricoles, cultures, élevage d'animaux) et l'homme en tant que notre cible finale dans la biosphère.
- les différentes questions théoriques et pratiques soulevées par la modélisation statistique de l'*incertitude* et de la *variabilité* appliquée aux différents compartiments de la biosphère et aux différents types de processus en jeu. On peut alors noter une certaine pluridisciplinarité car ces processus couvrent de nombreux domaines scientifiques (*e.g.* transport de radionucléides dans l'atmosphère, mécanismes de dépôt sur le sol et sur le couvert végétal, transfert sol-plantes, étude des pratiques agricoles, physiologie et métabolisme des animaux domestiques dictant le transfert aux produits animaux, sociologie des comportements alimentaires des populations humaines, mathématiques, implémentation numérique de méthodes statistiques, *etc.*).

La modélisation de l'évaluation de l'exposition humaine, depuis le rejet de substances radioactives dans l'atmosphère à la consommation de denrées contaminées, peut être vue comme la construction d'une chaîne dans laquelle chaque maillon correspond à un modèle impliquant un ensemble de processus incertains et variables. Face à cette pluridisciplinarité et à la multiplicité des modèles il était important de définir en amont les contours du champ d'application, *i.e.* notre périmètre de travail. Ce questionnement préliminaire avait pour vocation d'orienter la réflexion et les développements méthodologiques vers les processus clés en situation accidentelle et post-accidentelle. Ainsi, nous nous sommes intéressés en priorité aux processus dictant l'entrée et le devenir des radionucléides dans la chaîne alimentaire en considérant le flux de dépôt comme le point de départ (donnée d'entrée) de la modélisation et la quantité de radionucléides ingérée par l'homme comme la sortie principale. Ce choix implique que l'incertitude et la variabilité intervenant dans les processus de dispersion atmosphérique et dans les habitudes alimentaires humaines n'ont pas été caractérisées explicitement même si nous avons conscience de leur impact potentiel sur les évaluations de risque. De plus, notre intérêt s'est porté principalement sur le risque d'exposition du fait de l'ingestion de légumes feuilles, de lait (vache, brebis et chèvre) et de viande (bœuf, mouton et porc). Le transfert sol-plante étant, quelques semaines à quelques mois après la fin des rejets, une voie secondaire de contamination des denrées agricoles nous avons considéré des modèles simplifiés à 2 compartiments (cf. [Section 3.2.1.2](#)).

Le positionnement a été, tout au long de ce projet, d'identifier les processus d'intérêt de la chaîne alimentaire d'exposition en situation accidentelle et post-accidentelle sur la base des besoins opérationnels et scientifiques, et d'étudier par la suite les questions théoriques et pratiques qui en découlent. Dans les sections suivantes nous dressons un bilan des travaux effectués, et discutons les limites et proposons quelques directions de recherche en perspectives.

10.2 Bilan des travaux

Dans le cadre des travaux engagés dans ce projet, nous nous sommes essentiellement intéressés à la modélisation de l'*incertitude* et de la *variabilité* autour des processus d'interception foliaire (sèche ou humide) de radionucléides par les plantes (herbe, légumes feuilles et autres légumes) et de décroissance biomécanique de la concentration dans les plantes. Ces différents processus sont influencés par un ensemble de facteurs desquels on a une connaissance plus ou moins précise, comme les conditions de dépôt, les caractéristiques physico-chimiques des radionucléides déposés ou l'état de développement des plantes parmi ces facteurs influents (cf. [Section 3.2.1.1](#)).

Données

Une étape prépondérante dans la modélisation de ces processus a consisté à répertorier les sources potentielles d'incertitude et de variabilité et à collecter des données afférentes nous avons alors privilégié les données de la littérature présentées sous une forme aussi brute que possible *i.e.* des estimations des facteurs d'interception (sèche ou humide) issues de mesures directes de concentrations dans les plantes ou encore des estimations des constantes de décroissance directement obtenues à partir des chroniques de concentrations. Les données sont principalement issues :

- de documents et rapports techniques de l'AIEA ([IAEA, 1996, 2009, 2010](#)),
- de revues antérieures de la littérature,
- de la « littérature grise » concernant des expérimentations menées à l'IRSN et au CEA,
- des échanges de connaissance *via* les groupes de travail du projet COMET et à travers le réseau Européen d'excellence STAR.

Lorsque les données brutes n'étaient pas directement disponibles dans les revues, nous sommes reportés en priorité aux publications d'origine pour avoir une information non-

résumée. Avec les apports de ces différentes sources nous avons compilé, pour les facteurs d'interception sèche et humide et pour la constante de décroissance biomécanique, les bases de données les plus riches possible, contenant notamment des informations additionnelles relatives au stade de développement des plantes, à la nature des dépôts, aux caractéristiques des radionucléides. De cette collecte de données émerge une question sous-jacente relative à la probable hétérogénéité entre données. Par exemple, les concentrations massiques de radionucléides dans les végétaux peuvent être exprimées relativement à la masse sèche ou à la masse fraîche ; ce qui peut générer une source supplémentaire d'incertitude/variabilité car la teneur en eau des plantes peut varier d'une variété à l'autre ou d'une espèce à l'autre en fonction des conditions climatiques elles-mêmes variant selon les saisons et les régions d'intérêt. Cette question relative à l'hétérogénéité des données a notamment été discutée dans les [Chapitres 6](#) et [8](#) sur l'interception sèche et sur la décroissance biomécanique respectivement.

Modélisation

Sur la base des données compilées, des modèles hiérarchiques ont été construits et calibrés pour le facteur d'interception sèche (cf. [Chapitre 6](#)), le facteur d'interception humide (cf. [Chapitre 7](#)) et la constante de décroissance biomécanique (cf. [Chapitre 8](#)) en adoptant le paradigme bayésien. L'utilisation de la modélisation bayésienne hiérarchique a permis dans chaque application de synthétiser toute l'information disponible et de caractériser l'incertitude relative au phénomène ainsi que l'effet de la variabilité spécifique aux espèces végétales et aux radionucléides considérés. Dans ce contexte, la construction des couches latentes a permis de mieux saisir la variabilité résultant des processus (très souvent complexes à spécifier) dictant la variabilité entre espèces végétales ou entre radionucléides ayant des caractéristiques différentes (*e.g.* radionucléides sous forme particulaire ou gazeuse, anions versus cations). Le choix de l'approche bayésienne repose sur ses nombreux avantages présentés ci-dessous.

- Une séparation entre incertitude et variabilité. La structure hiérarchique des modèles implémentés permet en effet cette séparation, contrairement à d'autres approches.
- Les techniques d'inférence par l'utilisation d'algorithmes de type MCMC sont particulièrement adaptées à la calibration de modèles complexes ayant un grand nombre de paramètres ou construits avec une structure hiérarchique.
- La caractérisation de la loi jointe des paramètres propose une représentation de l'incertitude globale sur l'ensemble des paramètres. Cette caractérisation est alors particulièrement appropriée (car la loi jointe tient compte des dépendances entre

paramètres) lorsque l'on souhaite propager l'incertitude sur ces paramètres ou lorsque l'on souhaite utiliser le modèle à des fins prédictives.

- Contrairement aux méthodes statistiques classiques reposant sur une hypothèse asymptotique difficilement vérifiable, l'inférence bayésienne permet la représentation directe de l'incertitude sur l'estimation des paramètres avec la production des intervalles de crédibilité (*e.g.* un intervalle de crédibilité à 90% correspond à l'ensemble des valeurs comprises entre les 5^{ème} et 95^{ème} percentiles de la distribution *a posteriori* du paramètre).

L'une des principales critiques généralement émises à l'encontre de l'approche bayésienne concerne la prise en compte d'une information *a priori* additionnelle dans l'analyse et par conséquent le choix des distributions *a priori* sur les paramètres car ce choix peut nous conduire à incorporer une quantité d'information (subjective) importante dans l'analyse. Pour chaque paramètre, la stratégie a globalement été de considérer soit son sens physique, soit de des considérations mathématiques telles que son support (*e.g.* dans le [Chapitre 6](#) le facteur d'interception sèche représente une proportion car il est défini comme le ratio entre la concentration déposée à la surface des plantes sur la concentration totale déposée à la surface du sol et des plantes ; ce qui implique que les valeurs prises par ce facteur sont comprises entre 0 et 1). Ainsi, des distributions *a priori* portant relativement très peu d'information intrinsèque (*e.g.* loi normale avec une espérance à 0 et une variance de l'ordre de 10000) ont été assignées aux paramètres pour lesquels notre connaissance était plutôt vague. Chaque fois que nous avons défini une distribution *a priori* plus informative nous nous sommes basés sur la littérature. Ce fut le cas pour le ratio poids sec poids frais ([Chapitre 6](#)), pour la surface foliaire spécifique ([Chapitre 7](#)), pour l'épaisseur caractéristique de rétention d'eau ([Chapitre 7](#)) et pour la constante de dilution biologique ([Chapitre 8](#)).

Une dernière étape du travail a consisté à évaluer l'impact des sources d'incertitude et de variabilité identifiées et caractérisées (dans les étapes précédentes) sur les sorties des modèles radioécologiques d'évaluation de risque par le biais d'une analyse probabiliste des incertitudes. Ce travail a également permis d'évaluer l'influence des différents paramètres incertains et/ou variables sur les résultats de modélisation de la chaîne alimentaire (cf. [Chapitres 8 et 9](#)) à travers des analyses de sensibilité. Deux études de cas ont été traitées.

- La première correspond au dépôt standardisé de quatre isotopes radioactifs (Cs-134, Cs-137, I-131 et Sr-90) en conditions sèches. Une valeur de 1 kBq.m⁻² a été retenue pour chaque isotope (voir [Chapitre 8](#)).

- La seconde porte sur la contamination des légumes feuilles par le césium suite à l'accident de Fukushima (voir [Chapitre 9](#)).

Ces deux applications nous ont permis d'analyser l'influence des paramètres radioécologiques sur les niveaux de concentrations de radionucléides dans différentes denrées alimentaires et notamment d'évaluer la distribution temporelle et spatiale de cette influence.

10.3 Limites et perspectives

10.3.1 Régionalisation des paramètres radioécologiques

Dans le cadre du projet COMET (cf. [Chapitre 8](#)), des actions ont été définies par le groupe de travail « Human Food Chain » dans l'optique d'améliorer la caractérisation de l'incertitude et de la variabilité des paramètres radioécologiques et des paramètres anthropiques (habitudes alimentaires, pratiques agricoles *etc.*). En effet, les modèles de transfert de radionucléides dans les écosystèmes terrestres considérés dans les outils d'aide à la décision ARGOS et RODOS utilisés dans des pays comme la Norvège et la Finlande ont été paramétrés sur la base du modèle ECOSYS développé par [Müller et Pröhl \(1993\)](#) avec, pour la plupart des paramètres, des valeurs par défaut caractéristiques des écosystèmes allemands. L'un des objectifs principaux était donc d'enrichir ces outils d'aide à la décision en proposant des paramètres radioécologiques dont les valeurs sont caractéristiques des différentes régions européennes considérées (essentiellement Scandinavie et pourtour méditerranéen).

L'interception foliaire a été discutée dans les [Chapitres 6](#) et [7](#). Dans les deux cas le rendement (biomasse aérienne) apparaît comme une variable explicative importante présentant une variabilité spatio-temporelle. Plusieurs études ont permis de mettre en place une cartographie des rendements potentiels de l'herbe de pâturage (exprimés en poids sec) à l'échelle Européenne ([Corrall, 1988](#) ; [Lee, 1983](#) ; [Peeters, 1993](#) ; [Smit *et al.* 2008](#)). La figure ci-après ([Fig. 10.1](#)) illustre une spatialisation des biomasses en Europe. L'application des modèles hiérarchiques développés dans les [Chapitres 6](#) et [7](#) à ces données permettrait par exemple de simuler pour chaque classe de valeurs de biomasse et, par conséquent, pour chaque pays (voire région), des valeurs plausibles de facteurs d'interception sèche ou humide. De telles données resteront cependant entachées d'incertitude. Sur le même principe, des

méthodes permettant de caractériser la variabilité temporelle des valeurs de biomasse sèche de l'herbe de prairie permettraient par exemple de caractériser relativement simplement la variabilité temporelle des facteurs d'interception sèche ou humide par application de ces modèles hiérarchiques.

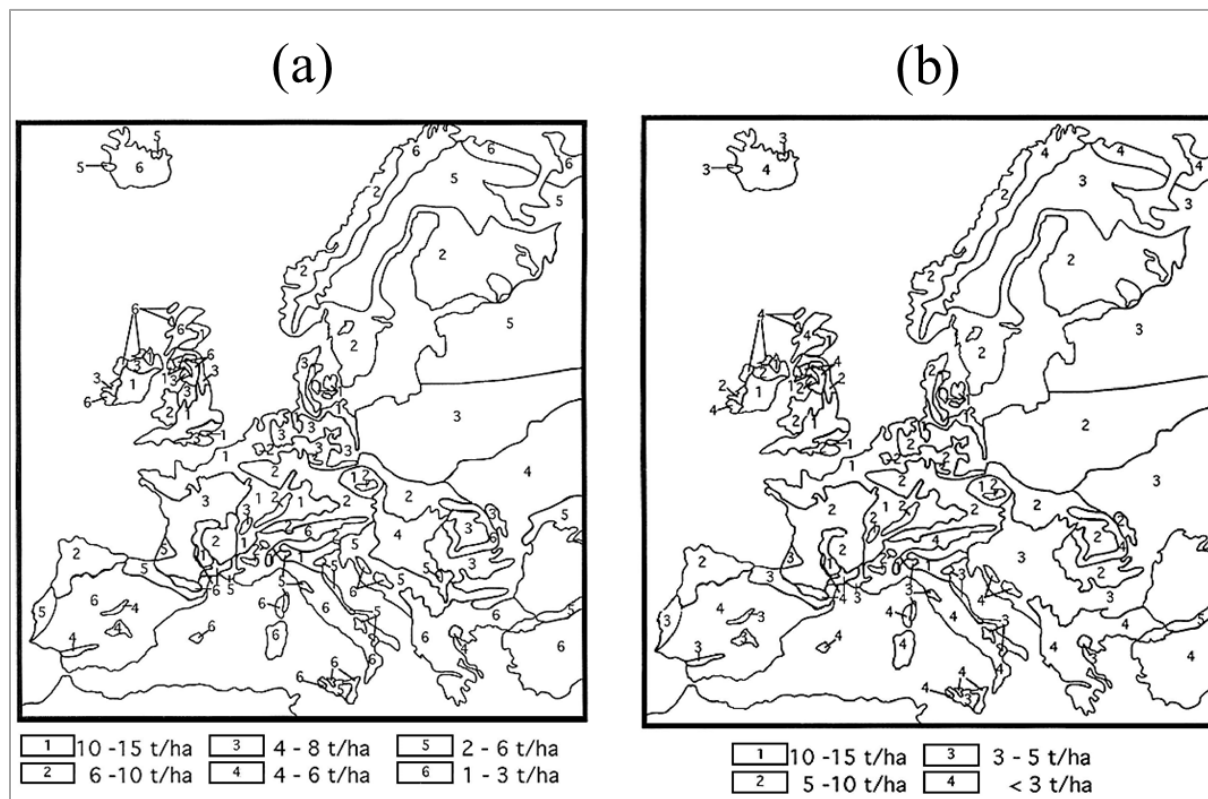


Fig. 10.1 : Cartes du rendement de l'herbe de pâturage en Europe. Le rendement est exprimé en tonnes (de matière sèche) par hectare ce qui équivaut à des kg (matière sèche) par m². (a) illustre des classes qui se chevauchent et (b) des classes disjointes. Illustration (Source : Peeters (1993) adapté de (Lee, 1983)).

10.3.2 Modélisation du transfert aux produits animaux

Le transfert aux produits animaux est généralement quantifié à l'aide d'un facteur de transfert à l'équilibre (cf. Section 3.2.2). L'AIEA (IAEA, 2009, 2010) a récemment compilé à travers ses réseaux d'expert une base de données conséquente de facteurs de transfert pour différents produits animaux (lait de brebis, de chèvre, de vache ; viande de bœuf, de mouton, porc ; œufs). Les données compilées par l'AIEA sont restituées sous une forme résumée (*i.e.* moyennes arithmétique et géométrique, écart-types standard et géométrique et nombre de données utilisées *etc.*). L'incertitude sur ces facteurs de transfert a été quantifiée par Simon-

Cornu *et al.* (2015). Ainsi, la stratégie suivante a été appliquée par Simon-Cornu *et al.* (2015) :

- lorsque la moyenne et l'écart-type géométriques (GM et GSD) sont renseignés sur la base de $n > 10$ données brutes, une loi log-normale tronquée de paramètres (min, max et p) est définie en posant $p = 0.95$, $min = \exp(\ln(GM) - 2\ln(GSD))$ et $max = \exp(\ln(GM) + 2\ln(GSD))$,
- lorsqu'entre 2 et 9 valeurs sont renseignées et/ou si la moyenne et l'écart-type géométriques sont remplacées par des valeurs min et max alors une loi log-uniforme variant entre min et max est définie,
- lorsqu'au plus une valeur est renseignée aucune loi de distribution n'est définie.

Les lois de probabilité obtenues à partir de cette stratégie ont permis d'alimenter la base de données probabiliste de la plateforme SYMBIOSE et ont servi à caractériser l'incertitude sur les facteurs de transfert au lait de vache et à la viande de bœuf dans les calculs réalisés dans le [Chapitre 8](#).

Le modèle décrit en [Section 3.2.2](#) et utilisé au [Chapitre 8](#) reproduit la cinétique du transfert aux produits animaux caractérisée non seulement par le facteur de transfert à l'équilibre mais aussi par une ou deux constante(s) de décroissance propres à un ou deux pool(s) métabolique(s). Au [Chapitre 8](#), des valeurs fixes (sans incertitude paramétrique) ont été utilisées pour les constantes de décroissance métabolique, ce qui est une limite de cette approche. Les travaux réalisés durant le stage de Rym Allouche-Chérif ([Allouche-Chérif, 2015](#)) que j'ai co-encadré, proposaient d'aller plus loin par la mise à jour de ces paramètres, en utilisant des mesures réalisées en France après l'accident de Fukushima ([Parache *et al.* 2011](#)) correspondant à une contamination chronique des animaux, et les mesures issues de l'étude menée par ([Beresford *et al.* 1990](#)) correspondant à une prise unique de contamination. La prise en compte de l'incertitude sur ces paramètres constitue une perspective de recherche.

La modélisation du transfert de radionucléides dans les produits animaux proposée dans la plateforme SYMBIOSE ne tient compte ni de l'âge, ni de la durée de vie des animaux domestiques et ne permet donc pas de suivre l'évolution de la contamination pour des cohortes. De plus, la date de naissance des animaux est considérée comme antérieure ou égale à la date de démarrage du calcul ; ce qui peut être maximisant du fait de la comptabilisation pour la cohorte d'intérêt des flux ingérés avant sa naissance. Il est important de préciser

également que l'inhalation et l'ingestion involontaire de sol ne sont pas explicitement modélisées comme des voies d'exposition interne pour les animaux domestiques.

Il pourrait être intéressant d'aborder ces questions à travers une modélisation plus exhaustive, voire plus mécaniste, c'est-à-dire prenant en compte les caractéristiques physiologiques et métaboliques des animaux. Une approche basée sur des modèles de type PBTK/TD (pour *physiologically based toxicokinetic / toxicodynamic*) peut par exemple être utilisée. En considérant l'organisme des animaux comme un ensemble de compartiments (*e.g.* les organes ou les systèmes cardiaque, respiratoire et digestif, *etc.*), les modèles PBTK/TD permettent de modéliser l'absorption, la distribution, la métabolisation et l'excrétion de polluants en décrivant les échanges de flux sanguins entre les différents compartiments considérés. Plusieurs applications reposant sur l'approche bayésienne pour la prise en compte de la variabilité interindividuelle (pour l'homme ou pour les animaux domestiques) ont été publiées en toxicologie des contaminants chimiques (*e.g.* [Barton et al. 2007](#) ; [Béchaux et al. 2014b](#) ; [Bernillon and Bois, 2001](#) ; [Bessems et al. 2014](#); [Bois et al. 2010](#)).

Cette approche a été envisagée lors du stage de Rym Allouche-Chérif. Cependant, la mise en œuvre d'une telle approche nécessite des jeux de données suffisamment riches pour permettre la calibration des différents paramètres et surtout la caractérisation de l'incertitude et de la variabilité interindividuelle (entre les différents animaux d'une cohorte, ce qui est peu répandu).

L'implication de l'IRSN dans des projets européens (avec la possibilité d'échanges des connaissances) (<https://wiki.ceh.ac.uk/x/YoFsD>) pourrait permettre de mettre en œuvre les perspectives identifiées plus haut.

10.3.3 Incertitude et Variabilité sur les paramètres contextuels

Globalement, l'incertitude et la variabilité sur les paramètres contextuels n'ont pas été abordées de manière approfondie à l'instar de celles caractérisant les paramètres radioécologiques. Les paramètres contextuels correspondent à l'ensemble des paramètres dont les valeurs sont spécifiques au scénario étudié, *i.e.* une valeur valable pour un scénario A ne l'est pas nécessairement pour un scénario B. Ainsi la caractérisation de l'incertitude et de la variabilité sur de tels paramètres nécessite des informations précises en lien direct avec le

scénario accidentel. Généralement très peu d'informations sont disponibles a minima à très court terme sur les conditions climatiques durant l'accident (direction du vent, hauteur et intensité des pluies, type de précipitations, *etc.*), et même à plus long terme pour les pratiques agricoles (culture sous serres, épandage d'engrais, *etc.*) et sur les comportements alimentaires des populations (types d'aliments habituellement consommés, pratiques d'hygiène, réception et respect des contre-indications gouvernementales, *etc.*). Notre démarche a donc consisté, dans certains cas, à considérer des hypothèses conservatives (*e.g.* pas de cultures sous serre, pas de réduction de la concentration du fait du lavage des légumes, des délais de stockage des produits alimentaires très courts, une part de denrées alimentaires produites localement fixée à 100%, *etc.*) et, dans d'autres cas, à adopter une approche semi-quantitative (*e.g.* en considérant deux scénarios de dépôt, l'un sec et l'autre humide). Cette stratégie est discutable et des améliorations peuvent être envisagées. Quelques pistes sont données dans la suite.

10.3.3.1 Dépôts atmosphériques

Le flux de dépôt correspond dans le cadre de ces travaux à la donnée d'entrée. Dans la première étude de cas, les quantités de Cs-134, Cs-137, I-131 et Sr-90 déposés sont fixées à 1 kBq.m⁻² (cf. [Chapitre 8](#)). Dans la seconde étude de cas, différentes valeurs sont considérées dans un rayon de 80 km autour de la centrale de Fukushima (cf. [Chapitre 9](#)). Ces valeurs sont issues des mesures réalisées par les autorités japonaises après l'accident. Ainsi, il est possible de caractériser la variabilité spatiale des dépôts en considérant en chaque point de l'espace la concentration. Cependant, le manque de connaissances ne permet pas de prendre en compte l'incertitude sur ces concentrations massiques.

Il pourrait par exemple être intéressant de proposer un modèle spatial permettant de caractériser la variabilité spatiale et de renseigner, en chaque point, de l'espace l'incertitude sur la quantité de radionucléides déposés. Des modèles basées sur les champs aléatoires ([Saint-Geours, 2012](#)) et combinant champs aléatoires et modélisation bayésienne hiérarchique ([Ancelet, 2008](#) ; [Chagneau, 2009](#) ; [Sall, 2009](#)) peuvent être une approche particulièrement intéressante pour l'analyse de données spatiales. Dans le cadre du projet EDOFU ([Gonze et al. 2013](#)), des travaux ont été développés permettant de simuler, dans la plateforme SYMBIOSE, les concentrations de radionucléides en considérant différentes cartes de dépôt reproduisant la variabilité spatiale et l'incertitude sur les dépôts ([Mac, 2015](#)). Ces

développements offrent donc la possibilité de considérer simultanément la variabilité spatiale sur les dépôts et la plage de variation en chaque point de l'espace.

Il est néanmoins important de préciser que l'utilisation de ces méthodes statistiques apporte une contribution intéressante uniquement sur la propagation simultanée de l'incertitude et de la variabilité spatiale ; car seuls la méthode et les outils utilisés sont exportables dans le cas d'un autre scénario. Dans le cadre de la quantification de l'incertitude et de la variabilité spatiale, des informations précises sur le scénario d'intérêt sont nécessaires à la pertinence et à la fiabilité des prédictions obtenues à partir de ces modèles dans la phase de propagation des sources d'incertitude et de variabilité spatiale.

10.3.3.2 Les approches alternatives comme une solution crédible

Nous distinguons dans ce travail l'*incertitude* et la *variabilité*, l'*incertitude* étant définie comme un manque de connaissances, une imprécision ou le caractère incomplet des informations disponibles et pouvant être de trois types : incertitude de modèle, incertitude de scénario ou incertitude paramétrique. L'application des approches alternatives dans le contexte des évaluations de risque alimentaire suite à un accident nucléaire offre des perspectives intéressantes concernant principalement la différenciation entre paramètres contextuels tels que la nature (sèche ou humide) des dépôts de radionucléides et paramètres faiblement scénario-dépendants. L'influence de ces paramètres dans les évaluations de risque a notamment été mise en évidence dans le cas d'étude relatif à la contamination des légumes feuilles par le césium suite à l'accident de Fukushima (cf. [Chapitre 9](#)). L'*incertitude* sur ces deux « types » de paramètres caractérisant un manque de connaissances est très souvent prise en compte dans les évaluations de risque comme une incertitude paramétrique ; ce qui ne permet, pas par exemple, de modéliser l'*incertitude* relative à l'occurrence de plusieurs scénarios de contamination dans une même analyse d'incertitude. Dans l'intégralité des applications utilisant ces approches alternatives, la distinction est plutôt faite entre *incertitude* et *variabilité* (définie comme une *incertitude aléatoire* par les différents auteurs). L'application des approches alternatives permettrait donc de différencier ces deux types d'incertitude (*incertitude de scénario* et *incertitude paramétrique*) en plus de la distinction faite généralement entre *incertitude* et *variabilité*. Si l'on suppose que la principale source d'information pour réduire ou caractériser l'incertitude de scénario sur des paramètres

contextuels provient d'avis d'expert (ex : pour l'accident de Fukushima, une information pertinente sur les habitudes agricoles ou alimentaires des populations locales est censée provenir d'un « expert » ou de toute autre personne au fait des réalités locales dans la préfecture de Fukushima), ces approches constituent, dans le cadre des évaluations de risque, une alternative intéressante à l'approche probabiliste classique. (cf. [Annexe D](#)).

Des réserves peuvent néanmoins être émises sur la pertinence et surtout la faisabilité des approches probabilistes non classiques ou des approches « extra-probabilistes » du fait de la difficulté des analystes et des décideurs à communiquer sur les incertitudes sur la base de formalismes sortant du cadre habituel des analyses de risque probabilistes ou encore du fait des potentielles difficultés techniques à mettre en œuvre ces méthodes. La citation suivante illustre bien ce scepticisme :

« Les paradigmes « extra-probabilistes », comme la théorie des possibilités ou la théorie Dempster-Shafer (DST) présentent une grande souplesse pour aller au-delà des limites du calcul probabiliste quand l'information est très limitée et hétérogène (Dempster, 1967, Dubois and Prade, 1988). Néanmoins, ils conduisent souvent à des besoins de calculs beaucoup trop lourds : ils cumulent typiquement des calculs d'optimisation déterministe à de la simulation probabiliste, ce qui multiplie immensément le nombre d'appels au code physique quand x est de dimension importante. Par ailleurs, les mesures d'importance pour hiérarchiser les sources d'incertitudes y sont encore peu développées. Enfin les résultats paraissent plus délicats à communiquer à des autorités réglementaires. » (De Rocquigny, 2006a).

10.3.4 Concilier modèle et mesures environnementales

Des résultats préliminaires obtenus lors d'une comparaison "modèle-mesure", notamment les écarts observés entre les prédictions des modèles et les mesures réalisées par les autorités Japonaises suite à l'accident de Fukushima ([Renaud et al. 2013 a, b](#)), ont soulevé le besoin de développer une approche méthodologique permettant de combiner les mesures issues de la surveillance environnementale et les prédictions obtenues par modélisation. Pareille approche permettrait, par exemple, de réduire les sources d'incertitude, de corriger quelques biais de modélisation et de limiter l'impact d'hypothèses un peu trop conservatives dans les

évaluations de risque nucléaire. Cette problématique est aujourd'hui d'un intérêt scientifique pour l'IRSN.

Selon le contexte, les doses sont évaluées soit par une approche partant de la connaissance radiologique des milieux (*i.e.* généralement les activités volumiques, surfaciques ou massiques des radionucléides dans les différents compartiments de l'environnement) et de quelques hypothèses calculatoires simples, soit par une approche de modélisation partant d'une donnée d'entrée (*e.g.* le rejet, ou le flux de dépôt) et utilisant des modèles physiques, biologiques ou chimiques de transferts et d'expositions (SYMBIOSE (cf. [Annexe B](#)), CROM (IAEA, 2001), ACADIE (GRNC, 1999), *etc.*). Ces deux approches sont développées relativement indépendamment l'une de l'autre, et rarement confrontées l'une à l'autre. Cette distinction, voire indépendance, entre approche essentiellement orientée vers la modélisation et approche reposant majoritairement sur des données observées dans l'environnement n'est pas spécifique à l'évaluation des impacts dosimétriques. Dans le domaine du risque alimentaire au sens large, on observe également une opposition entre, d'une part, des approches de modélisation « de la fourche à la fourchette » (*from fark to fork*) et, d'autre part, des approches d'acquisition de données au plus proche de l'exposition humaine, c'est-à-dire la contamination « dans l'assiette », dans les études totales de l'alimentation (*Total Diet Studies*, [Moy and Vannoort, 2013](#)). Différents développements ont été réalisés récemment dans le cadre de l'évaluation des risques alimentaires liés aux produits chimiques en vue de proposer des méthodes, permettant de confronter modèles et mesures issues de la surveillance environnementale ([Béchaux, 2014](#) ; [Béchaux et al. 2014b](#) ; [Bois and Boudet, 2006](#)). Plus généralement, pour l'exposition toutes voies confondues (incluant l'inhalation) à une pollution environnementale, on retrouve encore ([Ciffroy et al. soumis](#)) l'opposition entre calculs simples utilisant des mesures environnementales au plus proche de l'exposition versus modélisation multi-milieux (avec l'exemple récent du modèle MERLIN-EXPO développé dans le cadre des projets européens 2-FUN puis 4-FUN <http://merlin-expo.eu/>). Dans le domaine de la radioprotection et de la gestion des accidents nucléaires, de nombreuses applications combinant modèle et mesures de terrain par l'intermédiaire de la modélisation inverse et des techniques d'assimilation de données (particulièrement répandues en météorologie (voir [Blayo et al. 2012](#) ; [Kalnay, 2003](#)) sont également de plus en plus répandues concernant par exemple l'estimation du terme source suite à un accident nucléaire ([Bocquet, 2013](#) ; [Bocquet and Sportisse, 2007](#) ; [Mallet et al. 2007](#) ; [Mathieu et al. 2012](#) ; [Saunier et al. 2013](#) ; [Winiarek et al. 2012](#)). Dans le cadre de ce projet de recherche,

l'approche visant à confronter modèles et mesures apporterait une contribution intéressante à la caractérisation des sources d'*incertitude* et de *variabilité* dans la modélisation du transfert de radionucléides aux produits animaux. Comme évoqué dans la [Section 10.3.2](#), les mesures réalisées sur le territoire français suite à l'accident de Fukushima par [Parache *et al.* \(2011\)](#) et par [Saey *et al.* \(2013\)](#) constituent dans ce contexte une base de connaissances intéressante dans le cadre du transfert de l'iode et du césium au lait de vache ou à la viande de bœuf.

De manière formelle, ce sujet de recherche suscite un intérêt manifeste au sein du pôle de radioprotection de l'IRSN et fait l'objet du lancement d'études futures visant la construction d'une démarche méthodologique d'exploitation combinée des données de surveillance et de la modélisation pour évaluer notamment l'exposition de populations humaines à la radioactivité présente dans l'environnement et l'*incertitude* et la *variabilité* associées.

Références bibliographiques

A

- Adams, H.D., Williams, A.P., Xu, C., Rauscher, S.A., Jiang, X., McDowell, N.G., 2013. Empirical and process-based approaches to climate-induced forest mortality models. *Frontiers in Plant Science*, 4, 438.
- Allouche-Chérif, R., 2015. Etude du transfert des radionucléides vers les produits animaux suite à un accident nucléaire. Rapport Master 2 Aix-Marseille Université. 30 pages.
- Ancelet, S., 2008. Exploiter l'approche hiérarchique bayésienne pour la modélisation statistique de structures spatiales. Thèse de Doctorat Biostatistique. AgroParistech.
- Anderson, T.W., Darling, D.A., 1952. Asymptotic Theory of Certain "Goodness of Fit" Criteria Based on Stochastic Processes. *The Annals of Mathematical Statistics*. 23, 193-212.
- Anderson, T.W., Darling, D.A., 1954. A Test of Goodness of Fit. *Journal of the American Statistical Association*. 49, 765-769.
- Anspaugh, L.R., 1987. Retention by vegetation of radionuclides deposited in rainfall – A literature summary. UCRL-53810.
- Arunraj, N.S., Mandal, S., Maiti, J., 2013. Modeling uncertainty in risk assessment: An integrated approach with fuzzy set theory and Monte Carlo simulation. *Accident Analysis and Prevention*. 55, 242-255.

B

- Baccou, J., Chojnacki, E., Mercat-Rommens, C., Baudrit, C., 2008. Extending Monte Carlo simulations to represent and propagate uncertainties in presence of incomplete knowledge : application to the transfer of a radionuclide in the environment. *Journal of Environmental Engineering*. 134, 362-368.
- Bannör, K.F., Scherer, M., 2014. Model Risk and Uncertainty – Illustrated with examples from mathematical finance. In: *Risk – A Multidisciplinary introduction*, 279-306. Springer International Publishing, Switzerland.
- Baraldi, P., Pedroni, N., Zio, E., Ferrario, E., Pasanisi, A., Couplet, M., 2011. Monte Carlo and fuzzy interval propagation of hybrid uncertainties on a risk model for the design of a flood protection dike. In: C. Bérenguer, A. Grall & C. Guedes Soares (Eds.), *Advances in Safety, Reliability and Risk Management - Proceedings of the European Safety and*

- RELIability (ESREL) 2011 Conference, Troyes, France, 18-23 September 2011, pp. 2167-2175, Taylor & Francis Group, London, UK, 2012, ISBN 978-0-415-68379-1.
- Barton, H.A., Chiu, W.A., Setzer, W., Andersen, M.E., Bailer, A.J., Bois, F.Y., DeWoskin, R.S., Hays, S., Johanson, G., Jones, N., Loizou, G., MacPhail, R.C., Portier, C.J., Spendiff, M., Tan, Y.-M., 2007. Characterizing uncertainty and variability in physiologically based pharmacokinetic models: state of the science and needs for research and implementation. *Toxicological Sciences*. 99, 395-402.
- Baudrit, C., Dubois, D., Guyonnet, D., Fargier, H., 2004. Joint treatment of imprecision and randomness in uncertainty propagation. In : *Information Processing and Management of Uncertainty in Knowledge-based Systems*, Perugia, Italy, 4-9 July 2004.
- Baudrit C., 2005. Représentation et propagation de connaissances imprécises et incertaines: Application à l'évaluation des risques liés aux sites et sols pollués. Thèse de Doctorat de l'université Paul Sabatier, Toulouse III, 198 p.
- Baudrit, C., Guyonnet, D., Dubois, D., 2005. Post-processing the hybrid approach for addressing uncertainty in risk assessments. *Journal of Environmental Engineering*. 131, 1750-1754.
- Bayes, M., Price, M., 1763. An Essay towards Solving a Problem in the Doctrine of Chances. By the Late Rev. Mr. Bayes, F. R. S. Communicated by Mr. Price, in a Letter to John Canton, A. M. F. R. S. *Philosophical Transactions* 53, 370-418. 10.1098/rstl.1763.0053.
- Béchaux, C., 2014. Modélisation dynamique de l'exposition intégrant la toxicocinétique pour améliorer l'évaluation du risque à des substances chimiques. Thèse de Doctorat Santé Publique – Mathématiques Appliquées. AgroParistech.
- Béchaux, C., Bodin, L., Cléménçon, S., Crépet, A., 2014a. PBPK and population modelling to interpret urine cadmium concentrations of the French population. *Toxicology and applied Pharmacology*. 279, 364-372.
- Béchaux, C., Crépet, A., Cléménçon, S., 2014b. Improving Dietary Exposure Models by Imputing Biomonitoring Data through ABC Methods. *The international journal of biostatistics*. 10, 277-287.
- Bedwell, P., Mortimer, K., Wellings, J., Sherwood, J., Leadbetter, S.J., Haywood, S.M., Charnock, T., Jones, A.R., Hort, M.C., 2015. An assessment of the doses received by members of the public in Japan following the nuclear accident at Fukushima Daiichi nuclear power plant. *Journal of Radiological Protection*. 35, 869–890.
- Bengtsson S.B., Gårdenäs A.I., Eriksson J., Vinichuk M., Rosén K., 2014. Interception and retention of wet-deposited radiocesium and radiostrontium on a ley mixture of grass and clover. *Science of the Total Environment*, 497-498: 412-419.

- Beresford, N.A., Howard, B.J., Mayes, R.W., Lamb, C.S. & Barnett, C.L. 1990. Dynamics of radionuclides in sheep tissues, final Report. TFS Project T07051d1, 200pp. Central Electricity Generating Board & Ministry of Agriculture, Fisheries and Food.
- Beresford, N.A., Wood, M.D., Vives i Batlle, J., Yankovich, T.L., Bradshaw, C., Willey, N., 2015. Making the most of what we have: application of extrapolation approaches in radioecological wildlife transfer models. *Journal of Environmental Radioactivity*. 151, 373-386.
- Berliner, M.L., 1996. Hierarchical Bayesian time series models. In: *Maximum Entropy and Bayesian Methods. Proceedings of the Fifteenth International Workshop on Maximum Entropy and Bayesian Methods*. 79, 15-22.
- Bernillon, P., Bois, F.Y., Statistical issues in toxicokinetic modeling: a Bayesian perspective. *Environmental Health Perspective*. 108, 883 - 893.
- Bertrand M. (1993). Consommation et lieux d'achats des produits alimentaires en 1991, collection INSEE Résultats, Consommation-Modes de vie n°54-55.
- Bessems, J.G., *et al.* 2014. PBTK modelling platforms and parameter estimation tools to enable animal-free risk assessment. Recommendations from a joint EPAA – EURL ECVAM ADME workshop. *Regulatory Toxicology and Pharmacology*. 68, 119-139.
- Blayo, E., Bocquet, M., Cosme, E., Cugliandolo, L., eds., 2012. *Advanced Data Assimilation for Geosciences: Lecture Notes of the Les Houches School of Physics. Special Issue, June 2012 (Lecture Notes of the Les Houches Summer School)*.
- Bocquet, M., Sportisse, B., 2007. Modélisation inverse pour la qualité de l'air : éléments de méthodologie et exemples. *Pollution Atmosphérique*. 196, 395-404.
- Bocquet, M., 2013. Modélisation numérique de la dispersion atmosphérique accidentelle des radionucléides : l'état de l'art de la recherche. *Revue du Centre de Défense NBC, Ministère de la Défense*. 47-49.
- Bois, F.Y., Boudet, C., 2006. Integrating Environmental Monitoring and Human Biomarker Data in Risk Assessment: The Role of Modeling. *International Society for Environmental Epidemiology and International Society of Exposure Analysis Conference, ISEE/ISEA Conference Paris*.
- Bois, F.Y., Jamei, M., Clewell, H.J., 2010. PBPK modelling of inter-individual variability in the pharmacokinetics of environmental chemicals. *Toxicology*. 278, 256-267.
- Boreux, J.-J., Parent, E., Bernier, J., 2010. *Pratique du calcul bayésien*. Springer, Paris.

Brevault, L., Belesdent, M., Bérend, N., Le Riche, R., 2013. Comparison of different global sensitivity analysis methods for aerospace vehicle optimal design. 10th World Congress on Structural and Multidisciplinary Optimization. Orlando, Florida, May 19-24, 2013.

Brown, J.E., Beresford, N.A., Hosseini, A., 2013. Approaches to providing missing transfer parameter values in the ERICA Tool – How well do they work? *Journal of Environmental Radioactivity*. 126, 399-411.

C

Calec, N., 2013. Dépôt sec des aérosols submicroniques sur une surface liquide en mouvement. Thèse de Doctorat Mécanique, Physique, Micro et Nanoélectronique. Aix-Marseille Université.

Calmon, Ph., 2009. Equations et paramètres du logiciel ASTRAL V2.3. Rapport IRSN N° PRP-ENV/SESURE/LERCM/09-06 (révision de DEI/SECRE/06-47).

Campolongo, F., Cariboni, J., Saltelli, A., 2007. An effective screening design for sensitivity analysis of large models. *Environmental Modelling and Software*. 22, 1509-1518.

Caniou, Y., 2012. Analyse de sensibilité globale pour les modèles de simulation imbriqués et multi-échelles. Thèse de Doctorat Génie Mécanique. Université Blaise Pascal – Clermont II.

Cariboni, J., Gatelli, D., Liska, R., Saltelli, A., 2007. The role of sensitivity analysis in ecological modelling. *Ecological Modelling*. 203, 167-182.

Carlin, B.P., Clark, J.S., Gelfand, A.E., 2006. Elements of hierarchical Bayesian inference. In: Clark, J.S., Gelfand, A.E., 2006. *Hierarchical modelling for the environmental sciences*, Oxford University Press Inc., New York.

Casella, G., George, E.I., 1992. Explaining the Gibbs Sampler. *The American Statistician*. 46, 167-174.

Chadwick, R.C., Chamberlain, A.C., 1970. Field loss of radionuclides from grass. *Atmospheric Environment*. (1967). 4, 51-56.

Chagneau, P., 2009. Modélisation bayésienne hiérarchique pour la prédiction multivariée de processus spatiaux non gaussiens et processus ponctuels hétérogènes d'intensité liée à une variable prédite. Thèse de Doctorat Biostatistique. Université Montpellier 2.

Chamberlain, A.C., Garland, J.A., 1991. Interception of radioactive fallout by vegetation. Harwell, Atomic Energy Research Establishment. [AERE-R--13826](#).

Ciffroy P., *et al.* (Soumis). Development of a standard documentation protocol for communicating exposure models. *Science of the Total Environment*. En cours de révision.

- Ciric, C., 2012. Conception et développement d'un nouveau modèle d'écosystème aquatique adapté pour décrire la dynamique des espèces dans des mésocosmes lotiques. Thèse de Doctorat Université de Lyon.
- Clark, J.S., 2005. Why environmental scientists are becoming Bayesians. *Ecology Letters*, 8, 2-14.
- Codex Alimentarius. 2007. Working Principles for Risk Analysis for Food Safety for Application by Governments. Codex Alimentarius.
- CODIRPA, 2012. Eléments de doctrine pour la gestion post-accidentelle d'un accident nucléaire. Version Finale. 5 Octobre 2012. <http://www.asn.fr/Informer/Actualites/Gestion-post-accidentelle-d-un-accident-nucleaire>.
- Commeau, N., 2012. Modélisation de la contamination par *Listeria monocytogenes* pour l'amélioration de la surveillance dans les industries agro-alimentaires. Thèse de Doctorat Biostatistique. AgroParistech.
- Comte de Buffon (Georges Louis Leclerc). 1777. Essai d'arithmétique morale. 103 pages.
- Corrall, A.J., 1988. Prediction of production from grassland, in Information Bulletin of the FAO European Research Co-operative Network on Pastures and Field Crops, Herba 1, pp. 25– 28, Agron. Inst. of the Univ. of Florence, Florence, Italy.
- Coughtrey, P.J., Thorne, M.C., 1983. Radionuclides distribution and transport in terrestrial and aquatic ecosystems. A critical review of data. Volume I. ISBN 90-6191-278-4.
- Crépet, A., 2007. Statistique bayésienne et Monte Carlo de second ordre pour l'évaluation des risques microbiologiques. Thèse de Doctorat Mathématiques Appliquées et Applications des Mathématiques. AgroParistech.
- Cukier, H., Levine, R.I., Shuler, K., 1978. Nonlinear sensitivity analysis of multiparameter model systems. *Journal of Computational Physics*, 26, 1-42.
- Cuddington, K., Fortin, M.-J., Gerber, L.R., Hastings, A., Liebhold, A., O'Connor, M., Ray, C., 2013. Process-based models are required to manage ecological systems in a changing world. *Ecosphere*. 4, 20 pages.
- Cullen, A.C., Frey, H. C., 1999. Probabilistic techniques in exposure assessment: A handbook for dealing with variability and uncertainty in models and inputs, Springer.

D

- Da Veiga, S., 2007. Analyse d'incertitudes et de sensibilité. Application aux modèles de cinétique chimique. Thèse de Doctorat Statistiques. Université de Toulouse.

- Da Veiga, S., Wahl, F., Gamboa, F., 2009. Local polynomial estimation for sensitivity analysis on models with correlated inputs. *Technometrics*. 51, 425-463.
- Davison, A.C., Hinkley, D.V., 1997. *Bootstrap methods and their application*. Cambridge Press, UK. 582 pages.
- Delahaye, J-P., 2015. Modélisation mathématique. *Encyclopædia Universalis* [en ligne], URL : <http://www.universalis.fr/encyclopedie/modelisation-mathematique/>
- Delignette-Muller, M.L., Dutang, C., 2012. Fitting parametric univariate distributions to non censored or censored data using the R *fitdistrplus* package. Available at: http://114.207.245.114/moniwiki/pds/_ed_99_95_eb_a5_a0_eb_b6_84_ed_8f_ac_ec_a0_81_ed_95_a9_ec_84_b1/intro2fitdistrplus.pdf
- Dempster, A.P., 1967. Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics*. 38, 325-339.
- Depuydt, G., 2013. Etude expérimentale *in situ* du potentiel de lessivage de l'aérosol atmosphérique par les précipitations. Thèse de Doctorat Océan, Atmosphère et surfaces continentales. Université de Toulouse.
- De Rocquigny, E., 2006a. La maîtrise des incertitudes dans un contexte industriel. 1^{ère} Partie : Une approche méthodologique globale basée sur des exemples. *Journal de la Société Française de Statistique*, 147, 33-71.
- De Rocquigny, E., 2006b. La maîtrise des incertitudes dans un contexte industriel. 2^{ème} Partie : Revue des méthodes de modélisation statistique physique et numérique. *Journal de la Société Française de Statistique*, 147, 73-106.
- De Rocquigny, E., Devictor, N., Tarantola, S., 2008. *Uncertainty in industrial practice*. Wiley. 364 pages.
- De Rocquigny, E., 2009. Quantifying uncertainty in an industrial approach: an emerging consensus in an old epistemological debate. *S.A.P.I.EN.S.* 2(1) URL : <http://sapiens.revues.org/782>
- Destercke, S., 2008. Représentation et combinaison d'informations incertaines: nouveaux résultats avec applications aux études de sûreté nucléaire. Thèse de Doctorat Information. Université de Toulouse.
- Dobre, S., 2010. Analyses de sensibilité et d'identifiabilité globales. Application à l'estimation de paramètres photophysiques en thérapie photodynamique. Thèse de Doctorat Automatique / Robotique. Université Henri Poincaré – Nancy I.
- Draper, D., 1995. Assessment and Propagation of Model Uncertainty. *Journal of the Royal Statistical Society, Series B (Methodological)*. 57, 45–97.

- Draper, D., Pereira, A., Prado, P., Saltelli, A., Cheal, R., Eguilior, S. Mendes, B. Tarantola, S., 1999. Scenario and parametric uncertainty in GESAMAC: A methodological study in nuclear waste disposal risk assessment. *Computer Physics Communications*. 117, 142-155.
- Dubois, D., Prade, H., 1985. Evidence measures based on fuzzy information. *Automatica*, 21, 547-562.
- Dubois, D., Prade, H., 1988. *Possibility Theory: an approach to computerized processing of uncertainty*, Plenum Press, New York.
- Dubois, D., 2006. Possibility theory and statistical reasoning, *Computational Statistics & Data Analysis*. 51 (1) 47-69.
- Dubois, D., Guyonnet, D., 2011. Risk-informed decision-making in the presence of epistemic uncertainty. *International Journal of General Systems*. 40, 145-167.
- Duc, M., 2002. Contribution à l'étude des mécanismes de sorption aux interfaces solide-liquide: application aux cas des apatites et des oxy-hydroxydes. Thèse de Doctorat Material chemistry. Ecole Centrale de Paris, 2002.

E

- Efron, B., 1979. Bootstrap methods. Another look at the Jackknife. *Annals of Statistics*.
- Efron, B., Tibshirani, R.J., 1993. *An introduction to the Bootstrap*. Chapman and Hall/CRC, Florida, USA.
- EFSA, 2006. Guidance of the Scientific Committee on a request from EFSA related to Uncertainties in Dietary Exposure Assessment. *The EFSA Journal* 438, 1-54.
- EFSA, 2008. Guidance Document for the use of the Concise European Food Consumption Database in Exposure Assessment. EFSA/DATEX/2008/01. Parma, Italy.
- EFSA, 2011. Guidance on the environmental risk assessment of plant pests. EFSA Panel on Plant Health. *EFSA Journal* 2011; 9, 2460.
- EFSA, 2012. Scientific Opinion on the development of a risk ranking framework on biological hazards. EFSA Panel on Biological Hazards. *EFSA Journal* 2012; 10, 2724.
- EFSA, 2015. Guidance on Uncertainty in 1 EFSA Scientific Assessment. <http://www.efsa.europa.eu/fr/consultations/call/150618>. Parma, Italy.
- Ellouze, M., 2009. Evaluation des intégrateurs temps température en tant qu'outils de gestion des risques sanitaires et d'altération des aliments. Thèse de Doctorat AgroParistech.

F

- Faivre, R., Iooss, B., Mahévas, S., Makowski, D., Monod, H., 2013. Analyse de sensibilité et exploration de modèles. Application aux sciences de la nature et de l'environnement. Quae, France. ISBN 978-2-7592-1906-3.
- FAO/WHO, 2011. FAO/WHO guide for application of risk analysis principles and procedures during food safety emergencies. Rome, Food and Agriculture Organization of the United Nations and World Health Organization.
- Favier J.-C., Ireland-Ripert J., Toque C., Feinberg M., 1995. Répertoire général des aliments – table de composition, 2ème édition, INRA éditions.
- Ferson, S., Ginzburg, L., Kreinovich, V., Myers, D.M., Sentz, K., 2003. Construction Probability Boxes and Dempster-Shafer structures. Sandia National Laboratories, Technical report SANDD2002-4015.
- Frey, H.C., 1993. Separating variability and uncertainty in exposure assessment: Motivations and method. Paper No. 93-79.01. Proceedings of the 86th Annual Meeting Air and Waste Management Association Pittsburgh, Pennsylvania June 1993.
- Frey H.C., Rhodes, D.S., 1998. Characterization and Simulation of Uncertain Frequency Distributions: Effects of Distribution Choice, Variability, Uncertainty, and Parameter Dependence, Human and Ecological Risk Assessment: An International Journal. 4, 423-468.
- Frey, H.C., Patil, S.R. 2002. Identification and review of sensitivity analysis methods. Risk Analysis 22, 553-578.

G

- Garger, E.K., Hoffman, F.O., Thiessen, K.M., Galeriu, D., Kryshev, A.I., Lev, T., Miller, C.W., Nair, S.K., Talerko, N., Watkins, B., 1999. Test of existing mathematical models for atmospheric resuspension of radionuclides, Journal of Environmental Radioactivity. 42, 157-175.
- Gelfand, A.E., Smith, A.F., 1990. Sampling-based approaches to calculating marginal densities. Journal of the American Statistical Association. 85, 398-409.
- Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., Rubin, D.B., 2004. Bayesian Data analysis, Third edition. Chapman & Hall-CRC Texts in Statistical Science, 676 pages.
- Geman, S., Geman, D., 1984. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. IEEE Transactions on Pattern Analysis and Machine Intelligence. 6, 721-741.

- Girard, S., Korsakissok, I., Mallet, V., 2014. Screening sensitivity analysis of a radionuclides atmospheric dispersion model applied to the Fukushima disaster. *Atmospheric Environment*. 95, 490-500.
- Goguen, J.A., 1967. L-Fuzzy sets. *J. Mathematical Analysis and applications*. 18, 145-174.
- Gonze, M.-A., Mourlon, C., Garcia-Sanchez, L., Le Dizès, S., Nicoulaud, V., Métivier, J.-M., Simon-Cornu, M., Gerber P.-P., Vermorel, F., 2011. SYMBIOSE: A Simulation Platform for Performing Radiological Risk Assessments, ICRER 2011, 19-24 June, Hamilton (Canada), site (consulted on 03/12/2013): http://www.irsn.fr/EN/Research/Scientific-tools/Computer-codes/SYMBIOSE/Documents/ICRER2011_SymbiosePlatform_Poster02.pdf
- Gonze, M.-A., Mourlon, C., Nicoulaud-Gouin, V., Métivier, J.M., Debayle, C., Manach, E., Calmon, Ph., Gurriaran, R., Baccou, J., Courbet, C., Simonucci, C., Crosland, E., 2013. Etude d'évaluation Dosimétrique Externe dans la région de Fukushima. Le projet EDOFU. Rapport IRSN PRP-ENV/SERIS/2013-036.
- Gonze, M.A., Renaud, Ph., Korsakissok, I., Kato, H., Hinton, T.G., Mourlon, C., Simon-Cornu, M., 2014. Assessment of dry and wet atmospheric deposits of radioactive aerosols: application to Fukushima Radiocaesium fallout. *Environmental Science and Technology*. 48, 11268-11276.
- Gonze M.-A., Sy M.M., (Soumis le 22 Janvier 2016). Interception of Wet Deposited Radionuclides by Herbaceous Vegetation: Data Review and Modelling. *Science of the Total Environment*.
- Groupe Radioécologie Nord-Cotentin (GRNC), 1999. Rapport de synthèse : Estimation des niveaux d'exposition aux rayonnements ionisants et des risques de leucémies associés de populations du Nord Cotentin. http://www.irsn.fr/FR/base_de_connaissances/Environnement/surveillance-environnement/GRNC/Documents/irsn_gnrc_rapport-synthese-fr.pdf
- 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*. 126, 68-78, 2003.
- Guyonnet, D., Ménard, Y., Baudrit, C., Dubois, D., 2005. HyRisk – Traitement Hybride des incertitudes en évaluation des risques. Rapport BRGM/RP 53714.

H

- Haas, C.N., Rose, J.B., Gerba, Ch. Regli, S., 1993. Risk assessment of virus in drinking water. *Risk Analysis*. 13, 545-552.

- Hastings, W.K., 1970. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*. 57, 97-109.
- Hinton, T.G., Garnier-Laplace, J., Vandenhove, H., Dowdall, M., Adam-Guillermin, C., Alonzo, F., Barnett, C., Beaugelin-Seiller, K., Beresford, N.A., Bradshaw, C., Brown, J., Eyrolle, F., Fevrier, L., Gariel, J.C., Gilbin, R., Hertel-Aas, T., Horemans, N., Howard, B.J., Ikäheimonen, T., Mora, J.C., Oughton, D., Real, A., Salbu, B., Simon-Cornu, M., Steiner, M., Sweeck, L., Vives i Batlle, J., 2013. An invitation to contribute to a strategic research agenda in radioecology. *Journal of Environmental Radioactivity*. 115, 73-82.
- Hoeffding, W., 1948. A class of statistics with asymptotically normal distributions. *Annals of Mathematical Statistics*. 19, 293-325.
- Hoeting, J.A., Madigan, D., Raftery, A.E., Volinsky, C.T., 1999. Bayesian Model Averaging: A tutorial. *Statistical Science*. 14, 382-417.
- Hoffman, F.O., Frank, M.L., Blaylock, B.G., von Bernuth, R.D., Deming, E.J., Graham, R.V., Mohrbacher, D.A., Waters, A.E., 1989. Pasture grass interception and retention of ^{131}I , ^7Be and insoluble microspheres deposited in rain. Oak Ridge National Laboratory report ORNL-6542. Environmental Sciences Division Publication No. 3247.
- Hoffman, F.O., Thiessen, K.M., Frank, M.L., Blaylock, B.G., 1992. Quantification of the interception and initial retention of radioactive contaminants deposited on pasture grass by simulated rain. *Atmospheric Environment*, 26, 3313-3321.
- Hoffman, F.O., Hammonds, J.S., 1994. Propagation of uncertainty in risk assessments: The need to distinguish between uncertainty due to lack of knowledge and uncertainty due to variability. *Risk Analysis*. 14, 707-712.
- Hoffman, F.O., Thiessen, K.M., 1996. The use of Chernobyl data to test model predictions for interindividual variability of Cs-137 concentrations in humans. *Reliability Engineering & System Safety*. 54, 197-202.
- Homma, T., Saltelli, A., 1996. Importance measures in global sensitivity analysis of nonlinear models. *Reliability Engineering and System Safety*. 52, 1-17.
- Hosseini, A., Stenberg, K., Avila, R., Beresford, N.A., Brown, J.E., 2013. Application of the Bayesian approach for derivation of PDFs for concentration ratio values. *Journal of Environmental Radioactivity*. 126, 376-387.
- Hurtevent, P., Thiry, Y., Levchuk, S., Yoschenko, V., Henner, P., Madoz-Escande, C., Leclerc, E., Colle, C., Kashparov, V., 2013. Translocation of ^{125}I , ^{75}Se and ^{36}Cl to Wheat edible parts following wet foliar contamination under field conditions, *Journal of Environmental Radioactivity*. 121, 43-54.

- IAEA, 1989. Evaluating the reliability of predictions made using environmental transfer models. Vienna. IAEA Safety Series no. 100: 106.
- IAEA, 1992. Modelling of Resuspension, seasonality and losses during food processing. First report of the VAMP Terrestrial Working Group. IAEA-TECDOC-647. International Atomic Energy Agency, Vienna, Austria.
- IAEA, 1994. Handbook of parameter values for the prediction of radionuclide transfer in temperate environments. IAEA-TRS-364. International Atomic Energy Agency, Vienna, Austria.
- IAEA, 1996. Modelling of Radionuclide Interception and Loss Process in Vegetation and of Transfer in Semi-natural Ecosystems: Second Report of the VAMP Terrestrial Working Group. IAEA-TECDOC-857. International Atomic Energy Agency, Vienna, Austria.
- IAEA, 2001. Generic Models for Use in Assessing the Impact of Discharges of Radioactive Substances to the Environment. Safety Reports Series. No.19. Vienna.
- IAEA, 2009. Quantification of Radionuclide Transfer in Terrestrial and Freshwater Environments for radiological Assessments. IAEA-TECDOC-1616. International Atomic Energy Agency, Vienna, Austria.
- IAEA, 2010. Handbook of Parameter Values for the Prediction of Radionuclide Transfer in Terrestrial and Freshwater Environments. IAEA-TRS-472. IAEA, Vienna.
- ICRP, 1996. Age-dependent Dose to Members of the Public from Intake of Radionuclides : Part 5 Compilation of Ingestion and Inhalation Dose Coefficients, Publication 72 (Annals of the ICRP Vol. 26, No 1), Pergamon Press, Oxford.
- Iooss, B., 2009. Contributions to the uncertainty management in numerical modelling: wave propagation in random media and analysis of computer experiments. HDR Mathematics. Université Toulouse.
- Iooss, B., 2011. Revue sur l'analyse de sensibilité globale de modèles numériques. Journal de la Société Française de Statistique. 152, 3-25.
- Iooss, B., Mahévas, S., 2013. Revue et objectifs des méthodes d'analyse de sensibilité globale. In: Faivre, R., Iooss, B., Mahévas, S., Makowski, D., Monod, H., 2013. Analyse de sensibilité et exploration de modèles. Application aux sciences de la nature et de l'environnement. Quae, Paris.
- Iooss, B., Lemaitre, P., 2014. A review of global sensitivity analysis methods. In: Meloni, C., Dellino, G., Uncertainty management in Simulation-Optimization of Complex Systems : Algorithms and Applications, Springer, 2015.

- IPCS, 2004. Risk Assessment Terminology. IPCS. WHO, Geneva.
<http://www.who.int/ipcs/methods/harmonization/areas/ipcsterminologyparts1and2.pdf>
- IPCS, 2009. Risk Assessment and its Role in Risk Analysis. In: Principles and Methods for the Risk Assessment of Chemicals in Food. Environmental Health Criteria 240. ISBN 978 92 4 157240 8.
- IRSN, 2015a. Le projet AMORAD. Site internet de l'Institut de Radioprotection et de la Sécurité Nucléaire.<http://www.irsn.fr/FR/Larecherche/Organisation/Programmes/Amorad/Pages/projet-Amorad.aspx#.Vo4UZ0roiUk>
- IRSN, 2015b. <http://www.irsn.fr/FR/Pages/Home.aspx>.
- IUR, 1992. Protocol developed by the Working Group on Soil to Plant Transfer, 1982–1992, IUR, Saint-Paul-lez-Durance, France.
- Jackman, S., 2009. Bayesian Analysis for the Social Sciences. John Wiley & Sons, Ltd, Chichester.
- Jang, H.-K., Kim, J.-Y., Lee, J.-K., 2009. Radiological risk assessment for field radiography based on two dimensional Monte Carlo analysis. Applied Radiation and Isotopes. 67, 1521-1525.
- Jiao, Y., Hayes, C., Cortés, E., 2009. Hierarchical Bayesian approach for population dynamics modelling of fish complexes without species-specific data. – ICES Journal of Marine Science. 66, 367–377.

K-L

- Kalnay E., 2003. Atmospheric modeling, data assimilation and predictability. Cambridge University Press, Cambridge. 369 pages.
- Kann, A., Weyant, J.P., 2000. Approaches for performing uncertainty analysis in large-scale energy/economic policy models. Environmental Modelling and Assessment. 5, 29-46.
- Kaufman, A., 1973. Introduction à la théorie des sous—ensembles flous. Vol 1: Eléments théoriques de base. Masson, Paris. 432 pages.
- Kendall, M., 1955. Rank Correlation Methods. Hafner Publishing Co.
- Kentel, E., Aral, M.M., 2005. 2D Monte Carlo versus 2D Fuzzy Monte Carlo health risk assessment. Stochastic Environmental Research and Risk Assessment. 19, 86-96.
- Kettler, S., Kennedy, M., McNamara, C., Oberdörfer, R., O'Mahony, C., Schnabel, J., Smith, B., Sprong, C., Faludi, R., 2015. Assessing and reporting uncertainties in dietary exposure

- analysis. Mapping of uncertainties in a tiered approach. *Food and Chemical Toxicology*. 82, 79-95.
- King, D.M., Perera, B.J.C., 2013. Morris method of sensitivity analysis applied to assess the importance of input variables on urban water supply yield – A case study. *Journal of Hydrology*. 477, 17-32.
- Kinnersley, R.P., Goddard, A.J.H., Minski, M.J., Shaw, G., 1997. Interception of Caesium-contaminated rain by vegetation. *Atmospheric Environment*. 31, 1137-1145.
- Kirchner, G., Steiner, M., 2008. Uncertainties in radioecological assessment models—Their nature and approaches to reduce them. *Applied Radiation and Isotopes*. 66, 1750-1753.
- Lee, J., 1983. The spatial pattern of grassland production in Europe. Proc. 9th General Meeting of the European Grassland Federation. 11-20.
- Lelieveld, J., Kunkel, D., Lawrence, M.G., 2012. Global risk of radioactive fallout after major nuclear reactor accidents. *Atmospheric Chemistry and Physics*. 12, 4245-4258.
- Liland, A., Barnett, C.L., Beresford, N.A., Bradshaw, C., Brown, J., Calmon, Ph., Mora, J.C., Chmielewska, I., Duffa, C., Février, L., Gaschak, S., Iospje, M., Kasparov, V., Lahtinen, J., Michalik, B., Nanba, K., Outola, L., Real, A., Robles, B., Salbu, B., Skipperud, L., Simon-Cornu, M., Stark, K., Steiner, M., Thorrying, H., Vanhoudt, N., Vives i Battle, J., Vandenhove, H., 2013. Detailed plan for the COMET WP3 Initial Research Activity – list of research projects and goals, participants and timing. DELIVERABLE (D-N°3.1). available at https://wiki.ceh.ac.uk/display/COM/Publications?atl_token=c4dbabba10d86d0584fc9e3ad5bc2416e46c9528
- Linkov, I., Burmistrov, D., 2003. Model Uncertainty and Choices Made by Modelers: Lessons Learned from the International Atomic Energy Agency Model Intercomparisons. *Risk Analysis*. 23, 1297-1308.

M-N

- Mac, T.P.A., 2015. Study of spatial uncertainties by geostatistical simulations: application to the estimation of external doses in Fukushima with SYMBIOSE platform. Rapport de stage IRSN.
- Madigan, D., Raftery, A.E., 1994. Model Selection and Accounting for Model Uncertainty in Graphical Models Using Occam's Window. *Journal of the American Statistical Association*. 89, 1535-1546.
- Mallet, Q., Quélo, D., Sportisse, B., Korsakissok, I., 2007. Polyphemus : une plate-forme multimodèles pour la pollution atmosphérique et l'évaluation des risques. *Pollution Atmosphérique*, APPA. 423-431.

- Mara, T.A., Tarantola, S., Annoni, P., 2015. Non-parametric methods for global sensitivity analysis of model output with dependent inputs. *Environmental Modelling & Software*. 72, 173-183.
- Mathieu, A., Korsakissok, I., Quélo, D., Groëll, J., Tombette, M., Didier, D., Quentric, E., Saunier, O., Benoit, J.-P., Isnard, O. 2012. Atmospheric dispersion and deposition of radionuclides from the Fukushima-Daiichi nuclear power plant accident. *Elements*. 8, 195-200.
- Mayes, R.W., Lamb, C.S., 1989. A possible method for estimating the true absorption coefficient for radiocaesium in ruminants. *Science of the Total Environment*. 85, 263-266.
- Mayes, R.W., Beresford, N.A., Howard, B.J., Vandecasteele, C.M., Stakelum, G., 1996. Use of true absorption coefficient as a measure of bioavailability of radiocaesium in ruminants. *Journal of Radiation and Environmental Biophysics*. 35, 101-109.
- Metropolis, N., Ulam, S., 1949. The Monte Carlo Method. *Journal of the American Statistical Association*. 44, 335-341.
- Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E., 1953. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*. 21, 1087.
- Miehle, P., Battaglia, M., Sands, P.J., Forrester, D.I., Feikema, P.M., Livesley, S.J., Morris, J.D., Arndt, S.K., 2009. A comparison of four process-based models and a statistical regression model to predict growth of *Eucalyptus globulus* plantations, *Ecological Modelling*. 220, 734-746.
- Miller, C.W., Hoffman, F.O. 1983. An examination of the environmental half-time for radionuclides deposited on vegetation. *Health Physics*. 45, 731-744.
- Monod, H., Naud, C., Makowski, D., 2006. Uncertainty and sensitivity analysis for crop models. In Wallach, D., Makowski, D., Jones, J.W., Working with dynamic crop models: Evaluation, analysis, parameterization and applications, Elsevier, 55-100.
- Morris, M.D., 1991. Factorial Sampling Plans for Preliminary Computational Experiments. *Technometrics*. 33, 161-174.
- Moy G.G., Vannoort R.W., 2013. Total Diet Studies. Springer-Verlag, New York. 550 pages.
- Müller, H., Pröhl, G., 1993. ECOSYS-87: A dynamic model for assessing radiological consequences of nuclear accidents. *Health Physics*. 64, 232-252.
- Nauta, M.J., 2000. Separation of uncertainty and variability in quantitative microbial risk assessment models. *International Journal of Food Microbiology*. 57, 9-18.

Nicoulaud-Gouin, V., Garcia-Sanchez, L., Métivier, J.M., Gonze, M.A., 2014. Sensitivity analysis of a radionuclide transfer model describing contaminated vegetation in Fukushima prefecture, using Morris and Sobol' indices. 3rd International Conference on Radioecology & Environmental Radioactivity (ICRER). 7-12 September 2014, Barcelona, Spain.

NRC, 2009. Science and decisions: advancing risk assessment. Committee on Improving Risk Analysis Approaches used by the US EPA. National Academies Press, Washington DC.

Nucleonica GmbH, 2014. Nucleonica Nuclear Science Portal (<http://www.nucleonica.com/>), version 3.0.49, Karlsruhe, Germany.

O

Oakley, J.E., O'Hagan, A., 2007. Uncertainty in prior elicitation: A nonparametric approach. *Biometrika*. 94, 427-441.

Oberkampf, W.L., Helton, J.C., Joslyn, C.A., Wojtkiewicz, S.F., Ferson, S., 2004. Challenge problems: uncertainty in system response given uncertain parameters. *Reliability Engineering and System Safety*. 85, 11-19.

OCDE, 2010. Evaluation de risques d'accidents nucléaires comparés à ceux d'autres filières énergétiques. Rapport AEN N°6862. ISBN 978-92-64-99123-1.

Odum E.P., 1955. Consideration of the total environment in power reactor waste disposal. International Conference on the Peaceful Uses of Atomic Energy, Geneva, Switzerland.

Özkaynak, H., Frey, H.C., Burke, J., Pinder, R.W., 2009. Analysis of coupled model uncertainties in source-to-dose modelling of human exposures to ambient air pollution: A PM2.5 case study. *Atmospheric Environment*. 43, 1641-1649.

P-Q

Parache, V., Pourcelot, L., Roussel-Debet, S., Orjollet, D., Leblanc, F., Soria, C., Gurriaran, R., 2011. Transfer of ¹³¹I from Fukushima to the Vegetation and Milk in France, *Environmental Science & Technology*. 45, 9998-10003.

Parent, E., Bernier, J., 2007. *Le raisonnement bayésien: modélisation et inférence*, Springer, Paris.

Pasanisi, A., 2004. Aide à la décision dans la gestion des parcs de compteurs d'eau potable. Thèse de Doctorat Sciences of the Universe. ENGREF (AgroParistech).

Pasanisi, A., 2014. Analyse d'incertitudes et aide à la décision : contributions méthodologiques, techniques et managériales aux études d'ingénierie et de R&D. HDR Université de Technologie de Compiègne.

- Peeters, A., 1993. Extensification of grasslands in Europe. Proceedings of the 18th International Grassland Conference, Canada, Paper 15, 81-83.
- Petroff, A., 2005. Etude mécaniste du dépôt sec d'aérosols sur les couverts végétaux. Thèse de Doctorat Mécanique des Fluides. Aix-Marseille Université.
- Pouillot, R., Albert, I. Cornu, M., Denis, J.-B., 2003. Estimation of uncertainty and variability in bacterial growth using Bayesian inference. Application to *Listeria monocytogenes*. International Journal of Food Microbiology. 81, 87-104.
- Pouillot, R., Miconnet, N., Afchain, A.-L., Delignette-Muller, M. L. Beaufort, A., Rosso, L., Denis, J.-B., Cornu, M., 2007. Quantitative Risk Assessment of *Listeria monocytogenes* in French Cold-Smoked Salmon: I. Quantitative Exposure Assessment. Risk Analysis. 27, 683-700.
- Pouillot, R., Delignette-Muller, M.L., 2010. Evaluating variability and uncertainty separately in microbial quantitative risk assessment using two R packages. International Journal of Food Microbiology. 142, 330-340.
- Pröhl, G., Hoffman, F.O., 1996. Radionuclide Interception and Loss Processes in Vegetation. In Modelling of Radionuclide Interception and Loss Processes in Vegetation and of Transfer in Semi-natural Ecosystems: Second Report of the VAMP Terrestrial Working Group. IAEA, Vienna, 9-47.
- Pröhl, G., 2009. Interception of dry and wet deposited radionuclides by vegetation. Journal Environmental Radioactivity. 100, 675-682.

R

- R CoreTeam, 2015. R: A language and environment for statistical computing. Vienna, R Foundation for Statistical Computing.
- Raftery, A.E., 1995. Bayesian Model Selection in Social Research. Sociological Methodology. 25, 111-163.
- Renaud, Ph., Simon-Cornu, M., Gonze, M.-A., Mourlon, C., Parache, V., Korsakissok, I., Navarro, E., Cessac, B., Rannou, A., Champion, D., 2013a. Contamination of Japanese foodstuffs of terrestrial origin after the Fukushima nuclear accident and related dose assessments. Part 1: foodstuff contamination. Radioprotection. 48, 493-509.
- Renaud, Ph., Simon-Cornu, M., Gonze, M.-A., Mourlon, C., Parache, V., Korsakissok, I., Navarro, E., Cessac, B., Rannou, A., Champion, D., 2013b. Contamination of Japanese foodstuffs of terrestrial origin after the Fukushima nuclear accident and related dose assessment. Part 2: ingestion dose assessments for the first year after the accident. Radioprotection. 49, 17-22

- Rigaux, C., 2013. Méthodes de Monte Carlo du second ordre et d'inférence bayésienne pour l'évaluation des risques microbiologiques et des bénéfices nutritionnels dans la transformation des légumes. Thèse de Doctorat Biostatistique. AgroParistech.
- Rigaux, C., Denis, J.-B., Albert, I., Carlin, F., 2013. A meta-analysis accounting for sources of variability to estimate heat resistance reference parameters of bacteria using hierarchical Bayesian modeling: Estimation of D at 121.1 °C and pH 7, zT and zpH of *Geobacillus stearothermophilus*. *International Journal of Food Microbiology*. 161, 112-120.
- Robert, C.P., and Casella, G., 2011. Méthodes de Monte-Carlo avec R, Springer, France SARL.
- Roupsard, P., 2013. Etude phénoménologique du dépôt sec d'aérosols en milieu urbain : Influence des propriétés des surfaces, de la turbulence et des conditions météorologiques. Thèse de Doctorat INSA Rouen.
- Roustan, Y., 2005. Modélisation de la dispersion atmosphérique du mercure, du plomb et du cadmium à l'échelle européenne. Thèse de Doctorat Ocean, Atmosphere. Ecole des Ponts Paristech.
- Ruano, M.V., Ribes, J., Seco, A., Ferrer, J., 2012. An improved sampling strategy based on trajectory design for application of the Morris method to systems with many input factors. *Environmental Modelling & software*. 37, 103-109.

S-T

- Saey, L., Pourcelot, L., Parache, V., Roussel-Debet, S., Gurriaran, R., Orjollet, D., Leblanc, F., Renaud, Ph., 2013. Facteurs de transfert du césium au lait et à la viande déterminés à partir de mesures faites en France en 2011 à la suite de l'accident de Fukushima-Daichi. *Radioprotection*. 48, 341-354.
- Saint-Geours, N., 2012. Analyse de sensibilité de modèles spatialisés. Application à l'analyse coût-bénéfice de projets de prévention des inondations. Thèse de Doctorat Biostatistiques. Université de Montpellier 2.
- Sall, C.E., 2009. Modélisation spatiale hiérarchique bayésienne de l'apparentement génétique et de l'héritabilité en milieu naturel à l'aide de marqueurs moléculaires. Thèse de Doctorat Biostatistique. Université Montpellier 2.
- Saltelli, A., Tarantola, S., Chan, K.P.-S., 1999. A Quantitative model-independent method for global sensitivity analysis of model output. *American Statistical Association and the American Society for Quality*. 41, 39-56.
- Saltelli, A., 2002. Making best use of model evaluations to compute sensitivity indices. *Computer Physics Communication*. 145, 280-297.

- Saltelli, A., Tarantola, S., 2002. On the relative importance of input factors in mathematical models. Safety assessment for nuclear waste disposal. *Journal of American Statistical Association*. 97, 702-709.
- Saltelli, A., Tarantola, S., Campolongo, F. and Ratto, M., 2004, *Sensitivity Analysis in Practice*, Chichester : Wiley.
- Saporta, G., 2006. *Probabilités, analyse des données et statistique*, Editions Technip.
- Saunier, O., Mathieu, A., Didier, D., Tombette, M., Quélo, D., Winiarek, V., Bocquet, M., 2013. An inverse modeling method to assess the source term of the Fukushima Nuclear Power Plant accident using gamma dose rate observations. *Atmospheric Chemistry and Physics*, European Geosciences Union. 13, 11403-11421.
- SCHER, SCENIHR, SCCS, 2013. Making risk assessment more relevant for risk management. Scientific Committee on Health and Environmental Risks, Scientific Committee on Emerging and Newly Identified Health Risks, and Scientific Committee on Consumer Safety. March 2013.
- Sehmel, G.A., 1980. Particle and gas dry deposition: A review. *Atmospheric Environment*. 1967 14, 983-1011.
- SGDSN, 2014a. Plan National de réponse à un accident nucléaire ou radiologique majeur. NUMÉRO 200/SGDSN/PSE/PSN - ÉDITION FÉVRIER 2014.
- SGDSN. 2014b. Synthèse du plan de réponse à un accident nucléaire majeur. Service d'information du Gouvernement.
- Shafer, G., 1976. *A Mathematical Theory of Evidence*, Princeton University Press, Princeton.
- Simon-Cornu M., Beaugelin-Seiller K., Boyer P., Calmon P., Garcia-Sanchez L., Mourlon C., Nicoulaud V., Sy M. M., Gonze M-A., 2015. Evaluating variability and uncertainty in radiological impact assessment with the probabilistic database of SYMBIOSE. *Environmental Radioactivity*. 139, 91-102.
- Simpkins, A.A., 2005. Method for estimating ingestion doses to the public near the Savannah River site following an accidental atmospheric release. *Health Physics*. 88, 133-138.
- Smit, H.J., Metzger, M.J., Ewert, F., 2008. Spatial distribution of grassland productivity and land use in Europe. *Agricultural Systems*. 98, 208-219.
- Sobol, I.M., 1993. Sensitivity estimates for nonlinear mathematical models. *Mathematical Modelling and Computational Experiments*. 1, 407- 414.
- Spiegelhalter, D., Dawid, A., Lauritzen, S., and Cowell, R. (1993). Bayesian analysis in expert systems. *Statistical Science*. 8, 219–283.

- Steinhauser, G., Brandl, A., Johnson, T.E., Comparison of the Chernobyl and Fukushima nuclear accidents: A review of the environmental impacts, *Science of The Total Environment*, 470–471, 800-817.
- Sudret, B., 2007. Uncertainty propagation and sensitivity analysis in mechanical models. Contributions to structural reliability and stochastic spectral methods. HDR Université Blaise Pascal.
- Sy M.M., Ancelet S., Henner P., Hurtevent P., Simon-Cornu M., 2015. Foliar interception of radionuclides in dry conditions: A meta-analysis using a Bayesian modeling approach. *J. Environmental Radioactivity*. 147, pp 63-75.
- Sy M.M., Gonze M.A., Nicoulaud-Gouin V., Métivier J.M., Simon-Cornu M., 2016. Uncertainty analysis in post-accidental risk assessment models: an application to the Fukushima accident. *Annals of Nuclear Energy*. In press.
- Sy M.M., Hosseini, A., Liland, A., Simon-Cornu, M., Thorrying, H., (Soumis le 28 Janvier 2016). Uncertainty and variability in food chain exposure by radioactive fallouts. What can Hierarchical Bayesian modelling bring to radioecology community? *Journal of Environmental Radioactivity*.
- Thiessen, K.M., Thorne, M.C., Maul, P.R., Pröhl, G., Wheeler, H.S., 1999. Modelling radionuclide distribution and transport in the environment. *Environmental Pollution*. 100, 151-177.
- Tschiersch, J., Shinonaga, T., Heuberger, H., 2009. Dry deposition of gaseous radioiodine and particulate radiocaesium onto leafy vegetables. *Science of the Total Environment*. 407, 5685-5693.
- Tuffin, B., 2010. La simulation de Monte Carlo, Editions Lavoisier. Paris, France.

U-V

- UNSCEAR, 2013. Sources, Effects and Risks of Ionizing Radiation .Volume I: Report to the General Assembly. Scientific Annex A: Levels and effects of radiation exposure due to the nuclear accident after the 2011 great east-Japan earthquake and tsunami. UNSCEAR REPORT 2013. e-ISBN: 978-92-1-056501-1.
- Vandecasteele, C.M., Baker, S., Förstel, H., Muzinsky, M., Millan, R., Madoz-Escande, C., Tormos, J., Sauras, T., Schulte, E., Colle C., 2001. Interception, retention and translocation under greenhouse conditions of radiocaesium and radiostrontium from a simulated accidental source. *Science of the Total Environment*. 278, 199-214.
- Vandecasteele, C.M., 2004. Environmental monitoring and radioecology: A necessary synergy. *Journal of Environmental Radioactivity*. 72, 17-23.

- Van Leeuwen, C.J., Vermeire, T.G., 2007. Risk assessment of chemicals. An Introduction. 2nd Edition. Springer, Netherlands. 705 pages. ISBN 978-1-4020-6102-8.
- Villon, J., 2004. Théorie du Regret, Incertitude et Pessimisme. Thèse de Doctorat Sciences Economiques. Université de Caen.
- Vose, D. (2008). Risk analysis: a quantitative guide, Wiley.

W-Z

- Ward, G.M., Johnson, J.E., Stewart, H.F., 1965. Cesium-137 passage from precipitation to milk. Proceedings of the second conference on radioactive fallout from nuclear weapons tests, (Klement, A.W., Ed.). National Technical Information Service, Springfield (1965) 703-710.
- WHO/IPCS, 2008. Uncertainty and data quality in exposure assessment. Part 1: guidance on characterizing and communicating uncertainty in exposure assessment. Available online at: <http://www.who.int/ipcs/methods/harmonization/areas/uncertainty%20.pdf>
- WHO, 2012. Health risk assessment from the nuclear accident after the 2011 Great East Japan earthquake and tsunami, based on a preliminary dose estimation. WHO Report. ISBN 978 92 4 150513 0.
- Winiarek, V., Bocquet, M., Saunier, O., Mathieu, A., 2012. Estimation of errors in the inverse modeling of accidental release of atmospheric pollutant: Application to the reconstruction of the cesium-137 and iodine-131 source terms from the Fukushima Daiichi power plant. Journal of Geophysical Research: Atmospheres, American Geophysical Union. 117 (D05122).
- Winkler, R.L., Hays, W.L., 1975. Statistics. Holt, Rinehart and Winston, New York.
- Zadeh, L.A., 1965. Fuzzy Sets. Information and Control. 8, 338-353.
- Zadeh, LA., 1978. Fuzzy Sets as a basis for a theory of possibility. Fuzzy Sets and Syst.1, 3-28.
- Zio, E., Pedroni, N., 2013. Literature review of methods for representing uncertainty, numéro 2013-03 des Cahiers de la Sécurité Industrielle, Fondation pour une Culture de Sécurité Industrielle, Toulouse, France (ISSN 2100-3874). Disponible à l'adresse <http://www.FonCSI.org/fr/>.

Annexes

A Accidents de Tchernobyl et de Fukushima

A.1 Classement des accidents nucléaires

Après l'accident de la centrale de Tchernobyl (Ukraine, 1986) et afin d'aider la population et les médias à comprendre immédiatement la gravité d'un incident ou d'un accident dans le domaine nucléaire, une échelle de gravité semblable à l'échelle de Richter a été créée (IRSN, 2015b). L'échelle INES (International Nuclear Events Scale) s'applique à tout événement se produisant dans les installations nucléaires de base (INB) civiles et militaires, ainsi que lors du transport des matières nucléaires (IRSN, 2015b). L'application de l'échelle INES aux INB se fonde sur trois critères de classement : les conséquences à l'extérieur des installations, les conséquences de l'événement à l'intérieur du site et la dégradation des lignes de défense du site (IRSN, 2015b). A ce jour seuls deux événements ont été classés au plus haut niveau de l'échelle INES.

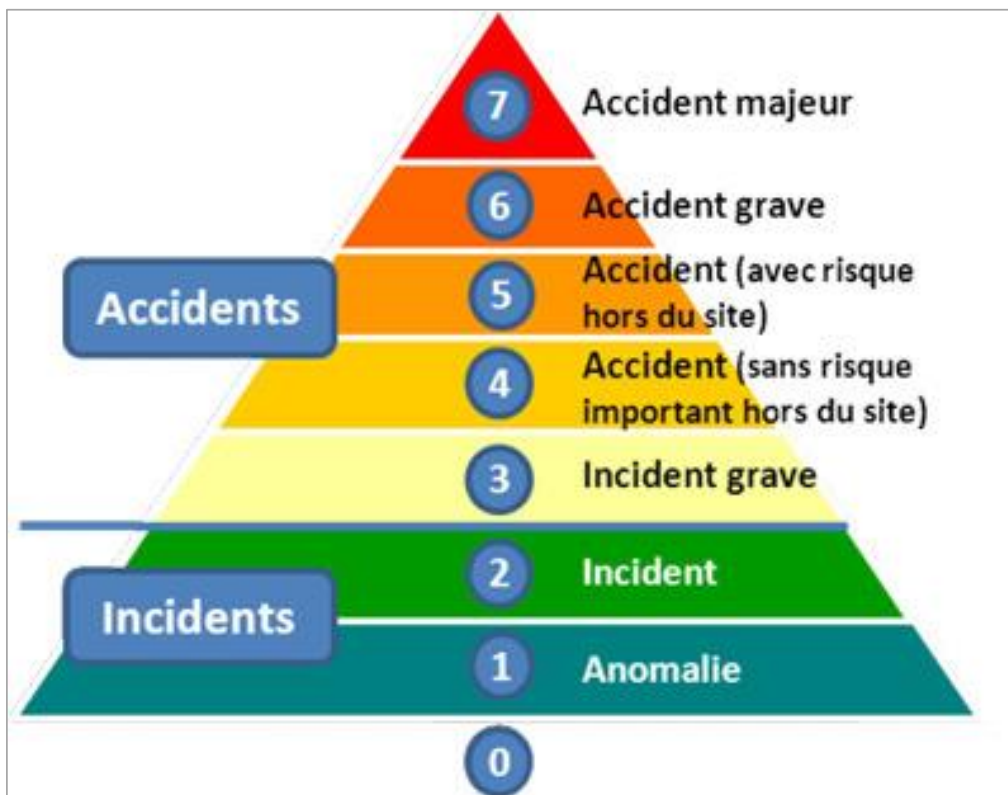


Fig. A. 1 : Echelle INES des accidents nucléaires

A.2 Dépôts de césium

La figure ci-dessous illustre les dépôts de césium ($Cs-134 + Cs-137$) lors des accidents de Tchernobyl en 1986 et de Fukushima en 2011.

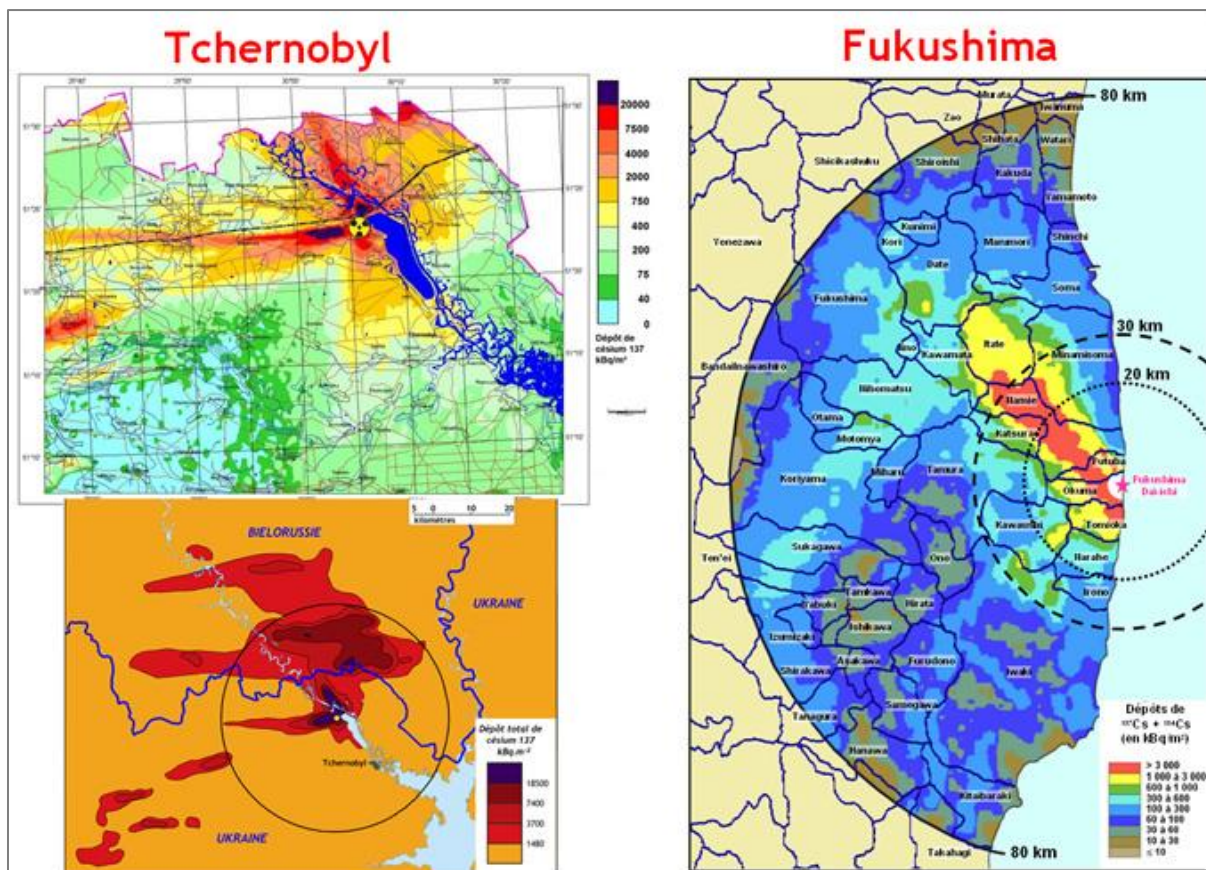


Fig. A. 2 : Cartes des dépôts de césium lors des accidents de Fukushima (à gauche) et de Tchernobyl (Source Médiathèque IRSN:

<http://www.irsn.fr/FR/connaissances/Environnement/expertises-accidents/comparaison-tchernobyl-fukushima/Pages/sommaire.aspx#.VqiNY0roiUk>).

A.3 Plans d'évacuation Fukushima

La figure ci-dessous (Fig. A. 3) propose à titre illustratif le plan d'évacuation des habitations organisé par le gouvernement Japonais sous l'égide du Ministère de l'Economie, du Commerce et de l'Industrie. Ce plan d'évacuation ne fait pas référence aux actions entreprises concernant la consommation de denrées alimentaires contaminées. Les décisions liées à la consommation de produits alimentaires ont été prises sous l'égide du Ministère Japonais de l'Education, de la Culture, du Sport, des Sciences et de la Technologie.

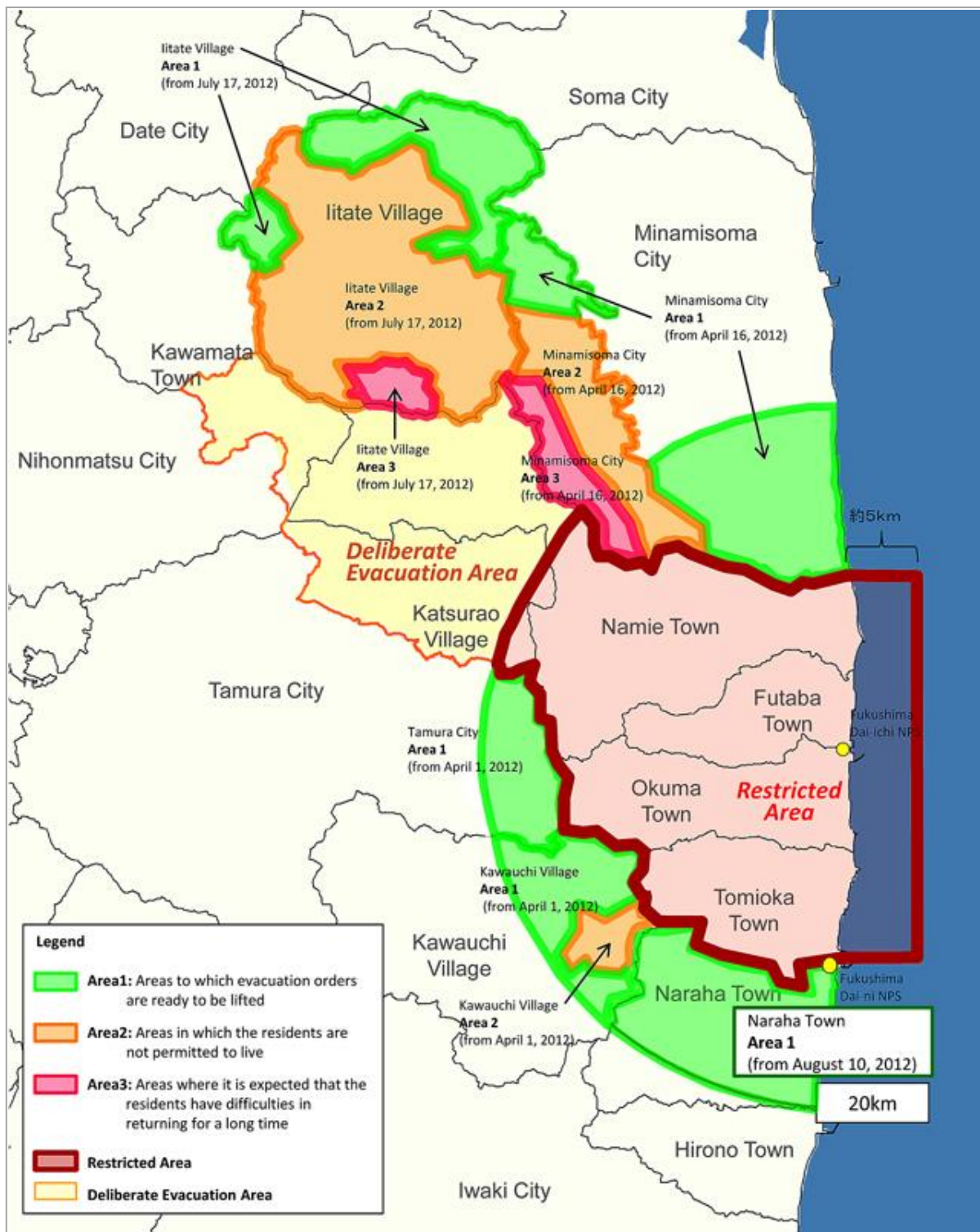


Fig. A. 3 : Plan d'évacuation au mois de juillet 2012 (Source : <http://www.meti.go.jp/english/>)

B SYMBIOSE et le modèle ASTRAL

Dans cette partie j'introduis deux outils développés par l'IRSN apportant une assistance technique dans le cadre de la modélisation des transferts de radionucléides dans les écosystèmes et vers l'homme.

B.1 ASTRAL

ASTRAL (acronyme pour ASsistance Technique en Radioprotection post-AccidenteLle) est un logiciel développé dans le cadre d'un projet cofinancé par l'IRSN et EDF offrant entre autres la possibilité, en cas de rejets accidentels, de tester différents scénarios de gestion de cette situation accidentelle. Le point de départ des estimations est le dépôt sur le sol des radionucléides. La phase de dispersion atmosphérique et les conséquences de l'exposition au nuage ne sont pas traitées par ce code. Les calculs effectués concernent à la fois l'évolution dans le temps des concentrations des radionucléides dans les éléments de la chaîne alimentaire (écosystèmes agricole et forestier) et l'évaluation des doses dues à l'exposition externe et interne. Parmi les principales fonctionnalités offertes par ASTRAL, on peut citer :

- le calcul de concentrations dans différents compartiments de la chaîne alimentaire,
- le calcul de doses reçues par des individus sur une période donnée *via* trois voies de contamination (exposition aux dépôts, inhalation de particules du dépôt remises en suspension, ingestion de denrées contaminées),
- le calcul d'indices de concentration, rapport entre les concentrations et la limite commerciale d'un produit alimentaire, permettant l'évaluation d'un impact radio-agronomique,
- la prise en compte dynamique de contre-mesures.

Les modèles implémentés dans ASTRAL sont des équations analytiques, en cohérence avec les modèles introduits au [Chapitre 3](#). Pour une présentation plus détaillée des paramètres et des équations d'ASTRAL je vous renvoie à ([Calmon, 2009](#)).

B.2 La plateforme de modélisation SYMBIOSE

SYMBIOSE est une plate-forme de modélisation du transfert de radionucléides dans les écosystèmes et de calcul de l'impact dosimétrique à l'homme développée par l'IRSN et cofinancée par Electricité De France (EDF) (Gonze *et al.* 2011). Cette plateforme permet de modéliser l'évolution dans le temps et dans l'espace du comportement des radionucléides au sein des principaux systèmes constitutifs d'un environnement continental. SYMBIOSE permet également de considérer différents types de rejets (atmosphériques, fluviaux ou marins) et d'évaluer les doses et débits de dose induits pour les populations humaines. Différentes voies d'exposition sont prises en compte : l'exposition externe lors du passage du panache et hors panache, l'exposition interne par inhalation, par transfert transcutané de tritium, par ingestion accidentelle de sable ou d'eau de mer, par ingestion de denrées alimentaires contaminées (ex : eau potable, lait, légumes, viande, poisson, œufs *etc.*). SYMBIOSE est un outil d'aide à la décision qui a pour vocation d'accompagner les études d'évaluation du risque radiologique pour une large gamme de situations : fonctionnement normal ou accidentel des installations nucléaires, démantèlement des installations, évaluation de sûreté des installations de stockage de déchets.

Les modèles implémentés dans SYMBIOSE sont des systèmes d'équations différentielles ordinaires, introduits au [Chapitre 3](#).

Les calculs SYMBIOSE pouvant être déterministes ou probabilistes mobilisent un ensemble de variables d'entrée et de paramètres.

- Le terme source et les conditions de rejet. Les données correspondant à cette variable doivent être renseignées par l'utilisateur.
- Un ensemble de variables contextuelles comprenant par exemple les variables météorologiques, les pratiques agricoles et les habitudes alimentaires des populations, les budgets temps de ces populations *etc.* Des valeurs par défaut ont été adoptées sur la base des régimes existant en France. Ces valeurs peuvent être modifiées sans trop de difficultés par les utilisateurs.
- Des paramètres caractérisant les processus physiques, chimiques et biologiques et les différents mécanismes radioécologiques. Les valeurs de ces paramètres sont principalement issues des rapports techniques de l'AIEA (IAEA, 1994, 2010) produits sur fond de consensus international par les experts impliqués dans les différents groupes de

travail de l'AIEA. Ces valeurs peuvent également être modifiées sans trop de difficultés par les utilisateurs.

Ainsi, la base de données de SYMBIOSE couvre 25 denrées alimentaires issues de plusieurs types de cultures (légumes feuilles, légumes racine, légumineuse et environ 10 cultures céréalières) et de plusieurs animaux (bœuf, mouton, poulet, cochon pour la viande ; vache, brebis, chèvre pour le lait ; poule pour les œufs ; différents poissons de rivière et poissons marins) et environ 277 isotopes de 64 éléments au total.

C Lois de probabilités usuelles

Dans cette section nous présentons quelques distributions de probabilité usuelles

C.1 Lois discrètes

▪ Loi de Bernoulli

La loi de Bernoulli permet de caractériser la probabilité qu'un événement se réalise ou non. Elle caractérise deux états disjoints tels que *réussite* et *échec*, *oui* et *non* ou *vrai* et *faux*. Par exemple la probabilité de tomber sur pile ou sur face lorsque l'on jette une pièce. Soit $p \in [0,1]$ et X une variable aléatoire suivant une loi de Bernoulli $\mathfrak{B}(1, p)$. Alors X vérifie :

$$P(X = k) = (1 - p) \times \delta_0(k) + p \times \delta_1(k) \quad (\text{C-1})$$

▪ Loi Binomiale

La loi binomiale est une généralisation de la loi de Bernoulli car elle s'applique à la répétition d'expériences aléatoires un nombre fini de fois. Elle permet par exemple de caractériser le nombre de fois que la face « pile » apparaît lorsque que l'on jette une pièce de monnaie un nombre n de fois. Soit $p \in [0,1]$ et X une variable aléatoire suivant une loi Binomiale $\mathfrak{B}(n, p)$. Alors X vérifie :

$$P(X = k) = \begin{cases} \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} & \text{si } k \in \{0, 1, 2, \dots, n\} \\ 0 & \text{sinon} \end{cases} \quad (\text{C-2})$$

$$\begin{cases} \mathbb{E}(X) = n \times p \\ \text{Var}(X) = n \times p \times (1-p) \end{cases} \quad (\text{C-3})$$

▪ Loi Binomiale Négative

La loi binomiale décrit la situation suivante : une expérience consiste en une série de tirages indépendants, donnant un "succès" avec probabilité $p \in [0,1]$ et un "échec" avec une probabilité complémentaire $(1-p) \in [0,1]$. Soit n le nombre de succès. La variable aléatoire représentant le nombre d'échecs (avant l'obtention du nombre donné n de succès) suit alors une loi binomiale négative. Soit X une variable aléatoire binomiale négative $\text{Neg}\mathfrak{B}(n, p)$. Alors X vérifie :

$$P(X = k) = \begin{cases} \frac{(n+k-1)!}{k!(n+k-1)!} p^n (1-p)^k & \text{si } k \in \{0, 1, 2, \dots, n\} \\ 0 & \text{sinon} \end{cases} \quad (\text{C-4})$$

$$\begin{cases} \mathbb{E}(X) = \frac{n \times (1-p)}{p} \\ \text{Var}(X) = \frac{n \times (1-p)}{p^2} \end{cases} \quad (\text{C-5})$$

▪ Loi Hypergéométrique

La loi hypergéométrique permet généralement de modéliser l'expérience suivante : tirer k boules blanches parmi n tirages sans remise dans une urne contenant N boules au total dont m boules blanches et $N - m$ boules d'une autre couleur. Soit X une variable aléatoire suivant une loi hypergéométrique $\mathcal{H}(N, m, n)$. Alors X vérifie :

$$P(X = k) = \begin{cases} \frac{\binom{k}{m} \binom{m-k}{N-m}}{\binom{n}{N}} \\ 0 & \text{sinon} \end{cases} \quad (\text{C-6})$$

$$\begin{cases} \mathbb{E}(X) = n \times p \\ \text{Var}(X) = n \times p \times (1-p) \frac{N-n}{N-1} \end{cases} \quad (\text{C-7})$$

▪ Loi Multinomiale

La loi multinomiale généralise la loi binomiale dans le sens où l'on considère plusieurs expériences différentes caractérisées par les variables aléatoires X_1, \dots, X_m , répétées un chacune un nombre fini de fois $n_i : i = 1, \dots, m$ et $p_i \in [0, 1] : i = 1 \dots, m$. Soit n_1, \dots, n_m tel que $\sum_{i=1}^m n_i = n$. On appelle loi multinomiale $M(n, p_1, \dots, p_m)$, la loi vérifiant :

$$P((X_1, \dots, X_m) = (n_1, \dots, n_m)) = \begin{cases} \frac{n!}{n_1! \dots n_m!} p_1^{n_1} \dots p_m^{n_m} & (n_1, \dots, n_m) \in \mathbb{N}^m \\ 0 & \text{sinon} \end{cases} \quad (\text{C-8})$$

$$\begin{cases} \mathbb{E}(X_i) = n \times p_i \\ \text{Var}(X_i) = n \times p_i \times (1 - p_i) \\ \text{Cov}(X_i, X_j) = -n \times p_i \times p_j & \text{avec } i \neq j \end{cases} \quad (\text{C-9})$$

▪ Loi de Poisson

La loi de poisson est généralement utilisée pour modéliser un comptage. Par exemple, le nombre de personnes attendant par jour à une file d'attente à un guichet de poste. Soit X une variable aléatoire de paramètre λ . Alors x vérifie :

$$P(X = k) = \begin{cases} e^{-\lambda} \frac{\lambda^k}{k!} & k \in \mathbb{N} \\ 0 & \text{sinon} \end{cases} \quad (\text{C-10})$$

$$\begin{cases} \mathbb{E}(X) = \lambda \\ \text{Var}(X) = \lambda \end{cases} \quad (\text{C-11})$$

C.2 Lois continues

▪ Loi uniforme

Soit $a, b \in \mathbb{R}$ tel que $a < b$. La loi uniforme sur $[a, b]$ caractérise l'équiprobabilité sur l'intervalle $[a, b]$. Soit X la variable aléatoire suivant une loi uniforme $U(a, b)$. Alors X vérifie :

$$f_X(x) = \begin{cases} \frac{1}{b-a} & x \in [a, b] \\ 0 & \text{sinon} \end{cases} \quad (\text{C-12})$$

$$\begin{cases} \mathbb{E}(X) = \frac{a+b}{2} \\ \text{Var}(X) = \frac{(b-a)^2}{12} \end{cases} \quad (\text{C-13})$$

▪ Loi Normale

La loi normale est aussi appelée loi de Gauss ou loi de Laplace-Gauss. Elles sont omniprésentes dans la modélisation de tout phénomène physique, chimique ou biologique. Cette omniprésence est liée au théorème central limite postulant qu'une somme infinie de variable aléatoire de moment d'ordre 2 fini est une loi normale. Ainsi, la loi normale est parfaitement adaptée à la modélisation de phénomènes résultante de la somme de plusieurs autres phénomènes. Soit X une variable aléatoire suivant une loi normale $N(\mu, \sigma)$. Alors X vérifie :

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right); \quad x \in \mathbb{R} \quad (\text{C-14})$$

$$\begin{cases} \mathbb{E}(X) = \mu \\ \text{Var}(X) = \sigma \end{cases} \quad (\text{C-15})$$

▪ Loi Log-Normale

Une variable aléatoire suit une loi log-normale si sa transformée logarithmique $\log(X)$ suit une loi normale. La loi log-normale est généralement utilisée pour modéliser des phénomènes résultant de la multiplication d'autres phénomènes. Soit X une variable aléatoire suivant une loi log-normale. Alors X vérifie :

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2x}} \exp\left(-\frac{(\log x - \mu)^2}{2\sigma^2}\right); x \in \mathbb{R} \quad (\text{C-16})$$

$$\begin{cases} \mathbb{E}(X) = e^{\mu + \frac{\sigma^2}{2}} \\ \text{Var}(X) = (e^{\sigma^2} - 1) \times e^{2\mu + \sigma^2} \end{cases} \quad (\text{C-17})$$

▪ Loi Beta

La loi Beta permet de modéliser des phénomènes dont les valeurs des réalisations sont comprises entre 0 et 1. Elle présente l'avantage d'être particulièrement flexible, les paramètres lui permettant de prendre plusieurs formes. Soit X une variable aléatoire suivant une loi $Beta(\alpha, \beta)$. Alors X vérifie :

$$f_X(x) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} x^{\alpha-1}(1-x)^{\beta-1}; x \in [0,1] \quad (\text{C-18})$$

$$\begin{cases} \mathbb{E}(X) = \frac{\alpha}{\alpha+\beta} \\ \text{Var}(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \end{cases} \quad (\text{C-19})$$

▪ Loi Exponentielle

La loi exponentielle permet par exemple de caractériser la durée de vie d'une ampoule. Soit X une variable aléatoire $Exp(\lambda)$. Alors X vérifie :

$$f_X(x) = \lambda \times e^{-\lambda x}; x \in \mathbb{R}_+^* \quad (\text{C-20})$$

$$\begin{cases} \mathbb{E}(X) = \frac{1}{\lambda} \\ \text{Var}(X) = \frac{1}{\lambda^2} \end{cases} \quad (\text{C-21})$$

▪ Gamma

La loi gamma est une généralisation de la loi exponentielle. Elle permet de modéliser une grande variété de phénomènes, et tout particulièrement les phénomènes se déroulant au cours du temps où par essence, le temps écoulé est une grandeur réelle positive ; c'est le cas par exemple dans l'analyse de survie. Soit X une variable aléatoire suivant une loi Gamma $\Gamma(\alpha, \beta)$. Alors X vérifie :

$$f_X(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} \times x^{\alpha-1} e^{-\beta x}; x \in \mathbb{R}_+^* \quad (\text{C-22})$$

$$\begin{cases} \mathbb{E}(X) = \frac{\alpha}{\beta} \\ \text{Var}(X) = \frac{\alpha}{\beta^2} \end{cases} \quad (\text{C-23})$$

▪ Lois des événements rares

Les lois de Gumbel, Pareto et Weibull sont généralement utilisées pour modéliser des événements rares. Par exemple en assurances elles peuvent être utilisées pour la modélisation du risque de crues ou d'autres phénomènes climatiques (catastrophes naturelles) dont les occurrences sont rares mais les effets dévastateurs. En finance, ces lois peuvent également être utilisées dans la prévention des crashes boursiers.

• Loi de Gumbel

Soit X une variable aléatoire suivant une loi de Gumbel de paramètres μ et $\beta > 0$. Alors X vérifie :

$$f_X(x) = \frac{\exp(-z) \times z}{\beta} \text{ avec } z = \exp\left(-\frac{x-\mu}{\beta}\right) \quad x \in \mathbb{R}_+^* \quad (\text{C-24})$$

$$\begin{cases} \mathbb{E}(X) = \mu + \beta \times \gamma \\ \text{Var}(X) = \frac{\pi^2}{6} \beta^2 \end{cases} \quad (\text{C-25})$$

• Loi de Pareto

Soit X une variable aléatoire suivant une loi de Pareto de paramètres x_m et k . Alors X vérifie :

$$f_X(x) = \frac{k \times x_m^k}{x^{k+1}} \text{ avec } ; x \in [x_m, +\infty[, x_m, k \in \mathbb{R}_+^* \quad (\text{C-26})$$

$$\begin{cases} \mathbb{E}(X) = \frac{k \times x_m}{k-1}, k > 1 \\ \text{Var}(X) = \frac{k \times x_m^2}{(k-1)^2(k-2)}, k > 2 \end{cases} \quad (\text{C-27})$$

• Loi de Weibull

Soit X une variable aléatoire suivant une loi de Weibull λ et k strictement positifs. Alors X vérifie :

$$f_X(x) = \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} e^{-\left(\frac{x}{\lambda}\right)^k} \text{ avec } \lambda, k > 0 ; x \in \mathbb{R}_+ \quad (\text{C-28})$$

$$\begin{cases} \mathbb{E}(X) = \lambda \times \Gamma\left(1 + \frac{1}{k}\right) \\ \text{Var}(X) = \lambda^2 \times \Gamma\left(1 + \frac{2}{k}\right) - \mu^2 \end{cases} \quad (\text{C-29})$$

D Méthodes de Monte Carlo et méthode Hybride

Dans cette section, nous présentons un exemple illustrant la propagation des incertitudes d'une part dans le cadre probabiliste et d'autre part suivant la méthode hybride de propagation des incertitudes introduite dans la [Section 5.3.3](#). Les paramètres d'entrée dits variables sont représentés par des distributions de probabilités. Les paramètres d'entrée dits incertains sont représentés soit par des distributions de probabilité (propagation Monte Carlo) soit par des distributions de possibilité pour l'approche alternative (propagation hybride). Cet exemple est proposé uniquement à titre illustratif et n'a pas vocation à comparer les approches probabiliste et alternative. Cependant, il permet d'exhiber le dilemme en face duquel le modélisateur ou les décideurs peuvent se retrouver lorsqu'il faut choisir une approche et l'appliquer à un problème ou à un scénario donné.

D.1 Définition du cas d'étude

Dans cet exemple, nous nous intéressons à la prédiction de la concentration initiale de Cs-137 dans les légumes feuilles suite à un dépôt atmosphérique accidentel de césium sur des parcelles agricoles. Nous considérons dans nos calculs la présence de serres. Le modèle permettant le calcul de la concentration initiale est donnée ci-dessous :

$$[Cs](t_0) = \frac{Dep}{B \times DMC} \times (K_s \times PS_s \times IF_s + (1 - K_s) \times PS_h \times IF_h) \quad (D-1)$$

La quantité $[Cs](t_0)$ désigne la sortie du modèle, *i.e.* la concentration initiale prédite dans les légumes feuilles exprimée en Bq/kg de poids frais. Les paramètres B et DMC représentent respectivement la biomasse des légumes feuilles (en $kg \cdot m^{-2}$) exprimée en poids sec et le ratio poids sec poids frais pour passer aux poids frais. Le paramètre K_s désigne la proportion de dépôt sec, les paramètres IF_s et IF_h correspondent respectivement aux facteurs d'interception sèche et d'interception humide et les paramètres PS_s et PS_h désignent des facteurs d'abris associés au dépôt sec et au dépôt humide respectivement et quantifient l'effet de protection due à la présence de serres. Ces facteurs de protection correspondent à des proportions et sont donc comprises entre 0 et 1.

Dans cette étude de cas, nous définissons principalement trois scénarios de calcul.

- Le scénario A définit un calcul de type MC1D où l'*incertitude* et la *variabilité* sont quantifiées à l'aide de distributions de probabilité. Des distributions de probabilité sont assignées à tous les paramètres.
- Le scénario B correspond à un calcul de type MC2D dans lequel l'*incertitude* et la *variabilité* sur les paramètres d'entrée sont traitées séparément. Des distributions de probabilité sont assignées aux paramètres incertains et variables.
- Dans le scénario C, nous séparons l'*incertitude* et la *variabilité* en représentant les paramètres incertains par des distributions de possibilité et les paramètres variables par des distributions de probabilité.

Les distributions de probabilité utilisées dans les scénarios A, B et C et les distributions de possibilité utilisées dans le scénario C sont définies dans le tableau [Tab. D. 1](#) et illustrées par les figures [D.1](#) et [D.2](#).

Tab. D. 1 : Définition des distributions de probabilité et de possibilité associées aux paramètres incertains et variables. N = Loi normale ; U = Loi uniforme ; Tr = Loi de probabilité triangulaire/ distribution de possibilité triangulaire.

Type Paramètre	Paramètre	Distribution Probabilité	Scenarios	Distribution Possibilité	Scenarios
	<i>Dep</i>	$N(10^4, 10^3)$	A-B-C	-	-
	<i>B</i>	$Tr(1.5, 3.5, 6)$	A-B-C	-	-
Variabilité	<i>DMC</i>	$U(0.1, 0.33)$	A-B-C	-	-
	<i>IF_s</i>	$U(0.6, 0.9)$	A-B-C	-	-
	<i>IF_h</i>	$U(0.25, 0.55)$	A-B-C	-	-
	<i>K_s</i>	$Tr(0.25, 0.45, 0.65)$	A-B	$Tr(0.25, 0.45, 0.65)$	C
Incertitude	<i>PS_s</i>	$Tr(0.1, 0.2, 0.4)$	A-B	$Tr(0.1, 0.2, 0.4)$	C
	<i>PS_h</i>	$Tr(0.2, 0.4, 0.6)$	A-B	$Tr(0.2, 0.4, 0.6)$	C

Pour quantifier les résultats des simulations nous calculerons trois indices, II/I, III/I et IV/I, permettant d'évaluer respectivement la variabilité globale, l'incertitude globale et la combinaison de l'incertitude et de la variabilité sur les concentrations de Cs-137 prédites dans

les légumes feuilles. Ces indices ont été introduits par [Özkaynak *et al.* \(2009\)](#) pour évaluer l'impact de l'incertitude et de la variabilité des entrées sur l'incertitude et la variabilité combinée sur la sortie du modèle. Ici l'indice II/I désigne le ratio entre le 97,5^{ème} percentile du 50^{ème} percentile d'incertitude et la tendance centrale (*i.e.* la médiane du 50^{ème} percentile d'incertitude), l'indice III/I désigne le ratio entre la médiane du 97,5^{ème} percentile d'incertitude sur la tendance centrale et l'indice IV/I correspond au ratio entre le 97,5^{ème} percentile de variabilité du 97,5^{ème} percentile d'incertitude et la tendance centrale. Pour la sortie de la méthode hybride nous avons calculé pour chaque percentile de variabilité le milieu de l'intervalle $[Bel(x), Pl(x)]$ défini par $(Bel(x) + Pl(x))/2$. Pour rappel, les termes *Bel* et *Pl* correspondent aux fonctions de croyance de Dempster-Shafer et désignent, respectivement, les mesures de crédibilité et de plausibilité définies dans la [Section 5.2.2.2](#). Les différents indices d'Özkaynak sont représentés dans le tableau [Tab. D. 3](#).

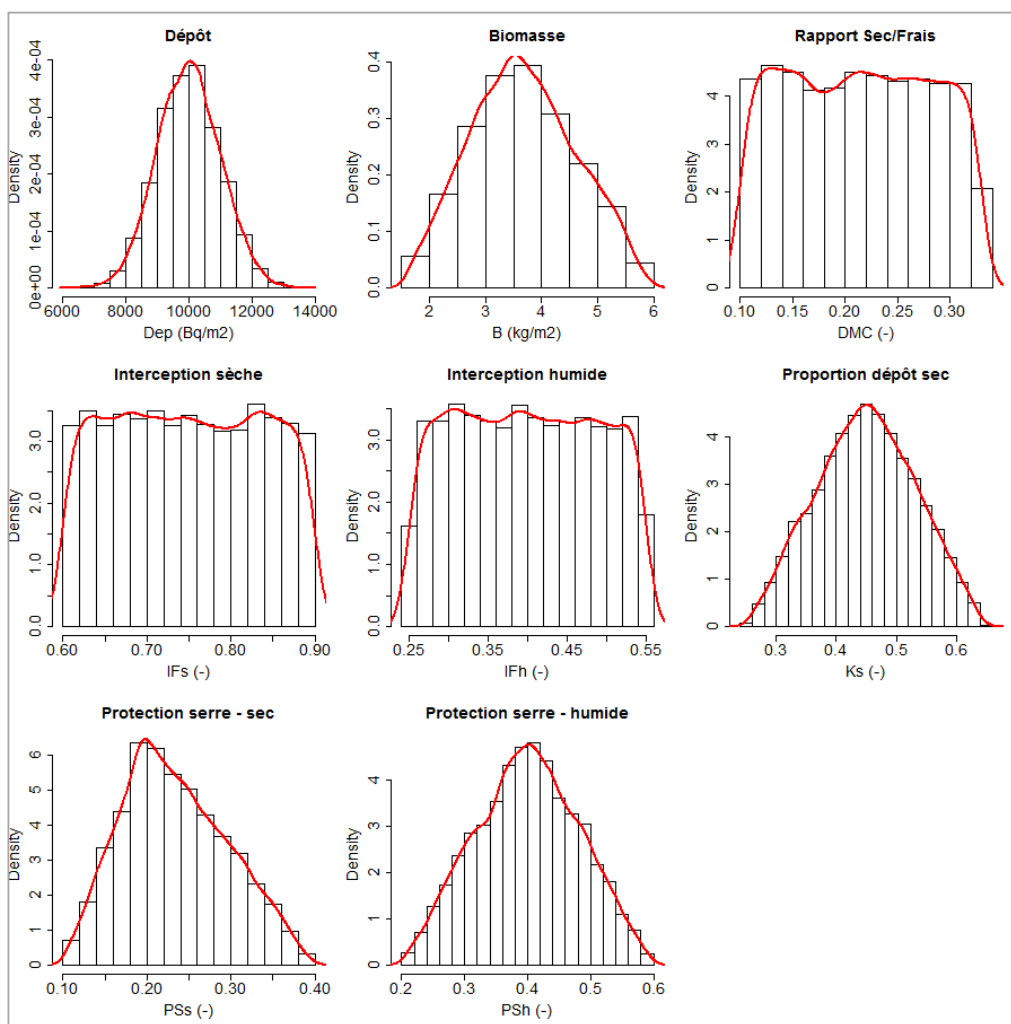


Fig. D. 1 : Histogrammes des distributions de probabilités des paramètres incertains et variables utilisés dans les scénarios A, B ou C. Les courbes rouges représentent les densités des distributions de probabilité.

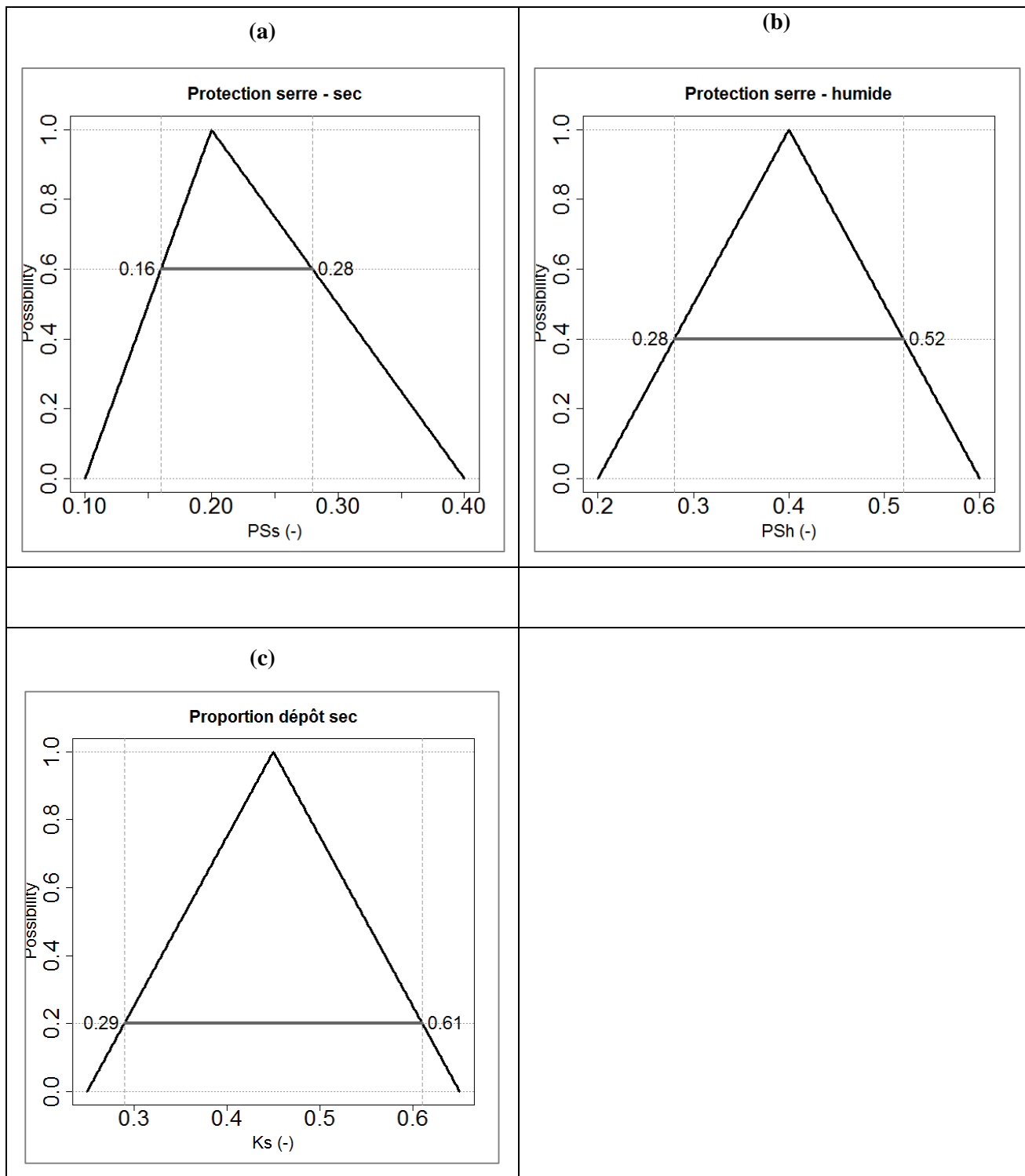


Fig. D. 2: Distributions de possibilité des paramètres incertains ((a) PS_s , (b) PS_h et (c) K_s) utilisées dans le scénario C. Les segments désignent les α -coupes avec α égal à 0.6, 0.4 et 0.2 dans les figures (a), (b) et (c) respectivement.

Tab. D. 2 : Résultats des simulations des méthodes MC1D, MC2D et Hybride. Les concentrations sont exprimées en Bq.kg⁻¹ de poids frais.

(^a) la colonne (Bel+Pl)/2 correspond au milieu de l'intervalle pour chaque percentile de variabilité (*i.e.* chaque ligne).

Les exposants I, II, III et IV correspondent aux quantités utilisées pour calculer les indices d'Özkaynak *et al.* (2009).

Inc. → Var. ↓	MC1D	MC2D							Hybride ^(a)		
	Min	2.5 ^e Perc.	5 ^e Perc.	50 ^e Perc.	95 ^e Perc.	97.5 ^e Perc.	Max	Bel	$\frac{Bel + Pl}{2}$	Pl	
2.5 ^e Perc.	918	513	689	728	1000	1300	1349	1533	601	903	1205
5 ^e Perc.	1029	574	768	818	1130	1464	1516	1729	684	1023	1361
50 ^e Perc.	2181 ^{I, III}	1104	1492	1582	2170 ^I	2801	2906 ^{III}	3333	1501	2180 ^I	2858 ^{III}
95 ^e Perc.	5204	2436	3298	3504	4795	6169	6388	7268	3626	5170	6714
97.5 ^e Perc.	6090 ^{II, IV}	2949	4018	4280	5847 ^{II}	7480	7758 ^{IV}	8854	4266	6057 ^{II}	7848 ^{IV}

Tab. D. 3 : Indices d'Özkaynak *et al.* (2009) calculés pour les méthodes MC1D, MC2D et Hybride.

Méthode	II/I (Var.)	III/I (Inc.)	IV/I (Inc. & Var.)
MC1D	2,79	1	2,79
MC2D	2,67	1,34	3,58
Hybride	2,78	1,31	3,6

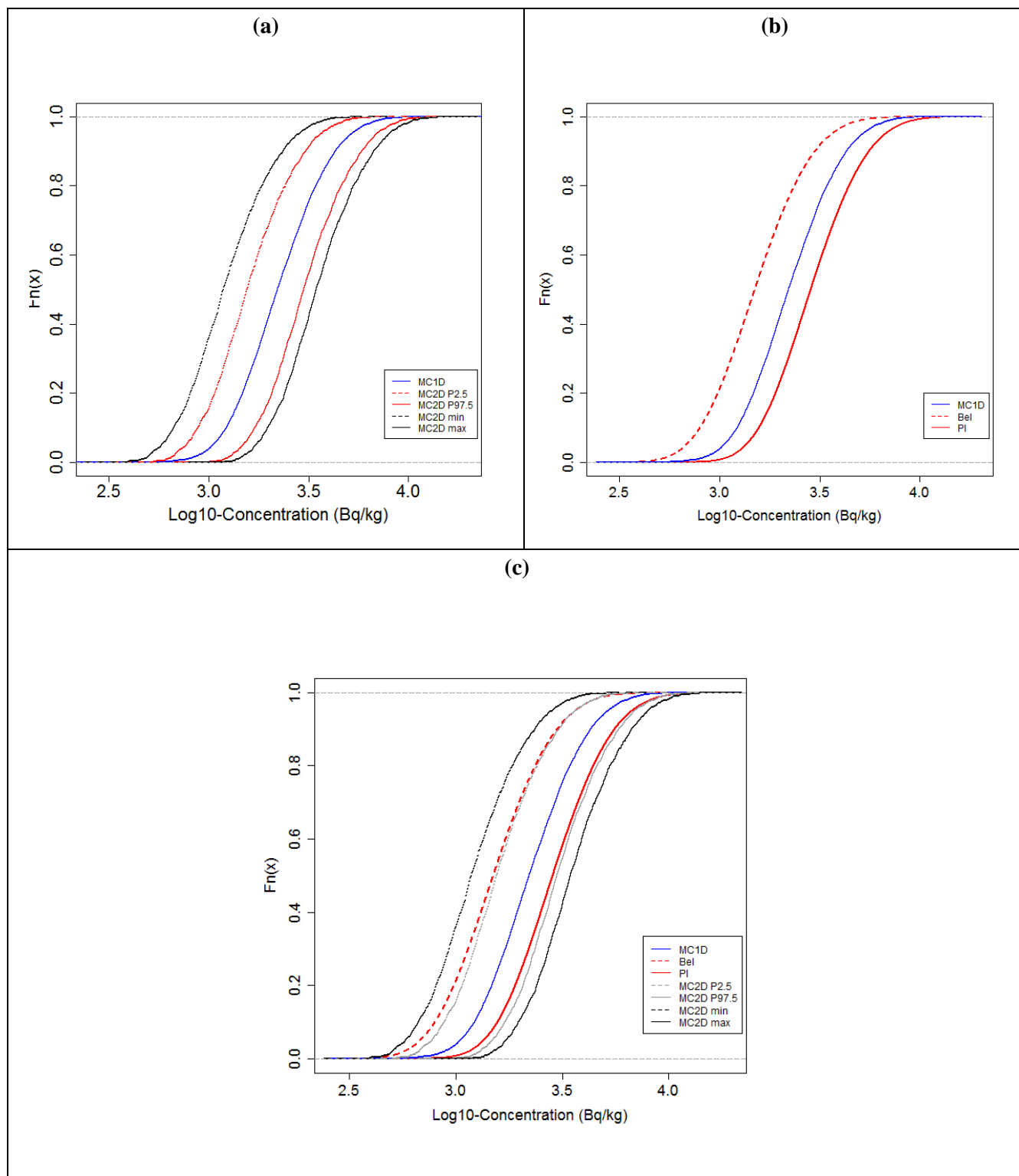


Fig. D. 3 : Fonctions de répartition empiriques des logarithmes en base 10 des concentrations de Cs-137 dans les légumes feuilles. (a) illustre les résultats issus des scénarios A et B, (b) les résultats des scénarios A et C et (c) les résultats des scénarios A, B et C.

D.2 Résultats

Les résultats des simulations relatives aux scénarios A, B et C sont résumés dans le tableau [Tab. D. 2](#) et illustrés par [Fig. D. 3](#). Globalement les résultats sont similaires entre les différentes méthodes. Les valeurs des percentiles de la distribution de probabilité obtenue par la méthode MC1D sont très proches de celles des percentiles calculés pour $\frac{Bel+Pl}{2}$ par la méthode hybride. Cette proximité est également observée lorsque l'on s'intéresse aux valeurs de l'indice II/I. Comme attendu, l'indice III/I est égal à 1 pour la méthode MC1D car l'incertitude et la variabilité sont confondues. Les résultats obtenus pour les scénarios B et C reflètent la combinaison de l'incertitude et de la variabilité sur les concentrations de Cs-137 dans les légumes feuilles. Les valeurs des indices III/I et IV/I sont similaires entre la méthode MC2D et la méthode hybride ; ce qui implique que l'incertitude et la variabilité relatives aux variations sur les concentrations de Cs-137 dans les légumes feuilles sont reproduites de manière similaire. Nous observons que les bornes *min* et *max* de la dimension d'incertitude obtenues par la méthode MC2D dessinent une enveloppe englobant les résultats de la méthode Hybride (cf. [Fig. D. 3-\(c\)](#)). La famille de distributions de probabilité définie par les fonctions de croyance *Bel* et *Pl* est donc entièrement contenue dans la famille de probabilité définie par les bornes *min* et *max* des distributions issues de la méthode MC2D. Cet écart observé entre les deux méthodes sur les distributions extrêmes d'incertitude pose la question de la gestion des extrêmes. En effet, considérons les deux cas suivants :

- (i.) Les résultats issus de la méthode MC2D traduisent réalistement les extrêmes
- (ii.) Les résultats issus de la méthode hybride traduisent réalistement les extrêmes

Dans le premier cas (i.) la méthode hybride semble mener à une surévaluation des valeurs minimales et à une sous-évaluation des valeurs minimales de concentrations ; ce qui peut mener à une évaluation des risques pas assez conservative car on peut rater des valeurs. Dans le second cas (ii.) la méthode MC2D semble mener à une surévaluation des valeurs maximales et à une sous-évaluation des valeurs minimales ; ce qui peut mener à une évaluation des risques trop conservative.

D.3 Conclusion

Comme déjà évoqué précédemment (cf. [Section 5.3.1.2](#)), la méthode MC1D ne permet pas de séparer *incertitude* et *variabilité* ce qui ne permet pas d'évaluer tous les scénarios possibles et peut conduire à une sous-évaluation ou à une surévaluation des doses. La méthode MC2D sépare les paramètres incertains et les paramètres variables en propageant indépendamment l'*incertitude* et la *variabilité* ; ce type d'approches peut notamment générer une surévaluation de l'amplitude des variations de la sortie et conduire à des décisions trop conservatives. La méthode hybride permet de séparer *incertitude* et *variabilité* par simulation Monte Carlo sur les paramètres variables et calcul d'intervalles sur les paramètres incertains. Comme la méthode MC2D, cette méthode permet d'évaluer les différents scénarios possibles (ici des scénarios majoritairement humide et sec).

Le nombre de paramètres incertains combiné à la complexité du modèle pourrait être rédhibitoire dans l'utilisation de la méthode hybride. En effet, si nous considérons K paramètres incertains représentés par des distributions de possibilité, la méthode hybride nécessite d'évaluer pour chaque niveau de possibilité α considéré 2^K combinaisons des bornes des α -coupes ; ce qui pose de sérieux problèmes d'optimisation lorsque le modèle est non-monotone.

Résumé. Les accidents nucléaires majeurs de Tchernobyl (Avril 1986) et de Fukushima (Mars 2011) ont généré une importante contamination environnementale avec des quantités conséquentes de substances radioactives libérées dans l'atmosphère. L'évaluation des risques, en situation d'urgence nucléaire, est entachée d'incertitudes sur le transfert de substances radioactives dans les écosystèmes terrestres et vers l'homme à travers la chaîne alimentaire, pouvant altérer la fiabilité des décisions. L'ampleur des répercussions des accidents de Tchernobyl et de Fukushima a mis en évidence la difficulté de gérer les conséquences post-accidentelles de tels événements et surtout d'appréhender les différentes sources d'incertitudes dans les prises de décision. L'objectif de cette thèse est de développer une méthodologie pour la prise en compte des incertitudes dans les modèles d'évaluation de risque environnemental et alimentaire en vue d'améliorer les outils d'aide à la décision en situation accidentelle. Dans cette optique, différents modèles bayésiens hiérarchiques visant à saisir, dans un cadre unique de modélisation, l'incertitude et la variabilité sur des paramètres radioécologiques d'intérêt en situation post-accidentelle (facteurs d'interception sèche et d'interception humide et constante de décroissance biomécanique) ont été développés. Les paramètres de ces modèles ont été estimés par inférence bayésienne sur des données collectées à partir d'une revue étendue de la littérature. L'influence sur les modèles d'évaluation de risque, de l'incertitude et de la variabilité autour de ces paramètres radioécologiques a ensuite été évaluée par simulations probabilistes et analyses de sensibilité, appliquées à deux études de cas : un accident hypothétique simulant un dépôt standardisé de radionucléides et l'accident survenu à la centrale nucléaire de Fukushima. Les travaux réalisés dans cette thèse contribuent à améliorer la connaissance autour des processus clés de transfert de radionucléides dans l'environnement et la paramétrisation des modèles radioécologiques d'évaluation de risque en ligne avec les axes de recherche prioritaires définis par la communauté scientifique en radioécologie.

Mots-clés : *Incertitude/Variabilité, Modélisation bayésienne, Accidents nucléaires majeurs, Radioécologie, Evaluation du risque alimentaire.*

Abstract. Major nuclear accidents of Chernobyl (April, 1986) and Fukushima (March, 2011) have led to a huge environmental contamination with important amounts of radionuclides released in the atmosphere. Risk assessment, in case of nuclear emergency, is confronted to uncertainties on the transfer of radioactive substances in terrestrial ecosystems and to human population through the food chain, which could affect the reliability of decisions. The extent of the repercussions of Chernobyl and Fukushima accidents highlighted the difficulty of managing the consequences of such disasters and specifically to accommodate the different sources of uncertainty within decision-making processes. The objective of this research project is to develop a methodology to take into account uncertainties within environmental and food risk assessment models in order to improve decision support tools used for accidental situations. In this regard, different hierarchical Bayesian models aiming at capturing, within a unique modelling framework, uncertainty and variability about radioecological parameters of great important for accidental situation (dry and wet interception fractions and weathering loss parameter) were developed. Models parameters were estimated by Bayesian inference applied on databases obtained by an extended literature review. The impact on the risk assessment models of uncertainty and variability about these radioecological parameters was then assessed by stochastic simulations and sensitivity analyses applied on two case-studies: a hypothetical accident simulating a standardized deposition of radionuclides and the accident of Fukushima nuclear power plant. The works developed in this project contribute to enhance knowledge on key processes governing environmental transfer of radionuclides and to improve the parameterization of the radioecological risk assessment models with respect to the research lines outlined by the scientific community in radioecology.

Keywords: *Uncertainty/Variability, Bayesian modelling, Major nuclear accidents, Radioecology, Food risk assessment.*
