



Quantification d'incertitude par réduction du modèle de dispersion atmosphérique appliquée à l'accident nucléaire de Fukushima

Ngoc Bao Tran Le

► To cite this version:

Ngoc Bao Tran Le. Quantification d'incertitude par réduction du modèle de dispersion atmosphérique appliquée à l'accident nucléaire de Fukushima. Modélisation et simulation. Sorbonne Université, 2021. Français. NNT : 2021SORUS044 . tel-03472277

HAL Id: tel-03472277

<https://theses.hal.science/tel-03472277>

Submitted on 9 Dec 2021

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

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

Thèse de doctorat de Sorbonne université

Présentée et soutenue publiquement le 10 juin 2021 par

Ngoc Bao Tran Le

pour l'obtention du diplôme de docteur de Sorbonne université

Spécialité : Mathématiques appliquées

Quantification d'incertitude par réduction du modèle de dispersion atmosphérique appliquée à l'accident nucléaire de Fukushima

préparée à l'IRSN–Institut de radioprotection et de sûreté nucléaire et l'INRIA
Paris–Institut national de recherche en informatique et en automatique

Jury :

<i>Directrice :</i>	Isabelle HERLIN	- INRIA Lille
<i>Co-directeurs :</i>	Irène KORSAKISSOK	- IRSN
	Vivien MALLET	- INRIA Paris
<i>Rapporteurs :</i>	Bertrand IOOSS	- EDF R&D
	Céline HELBERT	- École Centrale Lyon
<i>Examineurs :</i>	Sylvie THIRIA	- LOCEAN IPSL
	Matthias BEEKMANN	- LISA IPSL
	Pierre PINSON	- Technical University of Denmark

Remerciements

J'aimerais tout d'abord remercier mon directeur de thèse, M. Vivien Mallet, chargé de recherche à l'INRIA, pour m'avoir accueilli au sein de l'équipe CLIME puis ANGE de l'INRIA, où j'ai passé deux ans de ma thèse. Je lui suis également reconnaissante pour le temps conséquent qu'il m'a accordé, ses qualités pédagogiques et scientifiques, sa franchise et sa sympathie. J'ai beaucoup appris à ses côtés et je lui adresse ma gratitude pour tout cela.

J'adresse de chaleureux remerciements à ma co-encadrante de thèse, Mme Irène Korsakissok, chercheuse à l'IRSN, de m'avoir accueillie au sein de l'équipe BMCA de l'IRSN, où j'ai passé ma première année de thèse. Je lui adresse également les sincères remerciements pour son attention de tout instant sur mes travaux, pour ses conseils avisés et son écoute qui ont été prépondérants pour la bonne réussite de cette thèse. Son énergie et sa confiance ont été des éléments moteurs pour moi. J'ai pris de grand plaisir à travailler avec elle.

Je tiens à remercier Mme Céline Helbert et M. Bertrand Iooss pour le temps consacré à mon manuscrit de thèse.

Je tiens à remercier tous les membres du jury de la soutenance Mme Isabelle Herlin, Mme Céline Helbert, Mme Sylvie Thiria, M. Bertrand Iooss, M. Matthias Beekmann et M. Pierre Pinson pour leurs attentions.

Un grand merci à Mme Anne Mathieu pour son implication dans le projet, notamment sur la propagation d'incertitude par méthode de Monte Carlo. Elle m'a beaucoup appris, j'ai apprécié son enthousiasme et sa sympathie.

J'associe à ces remerciement M. Damien Didier, ancien chef du bureau BMCA, Mme Isabelle Herlin, ancien chef d'équipe CLIME de l'INRIA Paris, et M. Jacques Sainte-Marie, ancien chef d'équipe ANGE de l'INRIA Paris, pour m'avoir accueillie chaleureusement au sein de leur équipes pendant mes deux dernières années de thèse.

Bien sûr, atteindre ces objectifs n'aurait pas été possible sans l'aide des membres de SE-SUC/IRSN et ANGE/INRIA qui ont participé au projet.

Merci à Raphaël Périllat, Olivier Saunier, Arnaud Quérel et Denis Quélo qui m'ont donné des conseils avisés et ont été d'une grande ressource pour l'application du modèle ℓdX .

Je tiens à remercier tous mes collègues d'INRIA, Janelle Hammond, Frédéric Allaire, Antoine Lessieur, Raphaël Ventura, Ruiwei Chen pour leurs aides et leurs bonnes humeurs. Nous avons partagé de bons moments.

Résumé :

En cas d'accident nucléaire pouvant impliquer un rejet de radionucléides dans l'environnement, l'IRSN utilise des modèles de dispersion atmosphérique pour prédire les conséquences sanitaires et environnementales et préconiser des actions de protection de la population. En phase post-accidentelle (une fois les rejets effectués dans l'environnement), les simulations complètent les mesures de terrain pour l'évaluation des effets à court et à long terme sur l'environnement et la santé humaine. Les modèles de dispersion atmosphérique calculent, sur un domaine donné, les concentrations de radionucléides dans l'atmosphère, les quantités déposées au sol, ainsi que les débits de dose provenant du rayonnement gamma émis par les radionucléides présents dans l'air et au sol.

Les prévisions issues de ces simulations sont entachées de grandes incertitudes, en raison de la connaissance limitée des données d'entrée (prévisions météorologiques et terme source), des paramètres physiques des modèles de dispersion (vitesses de dépôt, coefficients de diffusion, coefficients de lessivage, etc.) et des approximations numériques de ces modèles. Le but de cette thèse est de mettre en place des méthodes permettant d'attribuer des densités de probabilité (PDF – probability density functions) aux erreurs inconnues sur les variables d'entrée et de calibrer ces PDF à l'aide d'observations de terrain. Ces méthodes ont été appliquées à la catastrophe nucléaire de Fukushima, en utilisant un modèle de dispersion opérationnel de l'IRSN et des mesures de radioactivité collectées pendant ou après la catastrophe. Dans les processus d'évaluation et d'étalonnage (calibration), nous avons utilisé des observations radiologiques de l'activité dans l'air, des dépôts et du débit de dose collectées au Japon.

Les méthodes de calibration nécessitent un grand nombre d'appels au modèle, de sorte que le temps de calcul est beaucoup trop élevé pour pouvoir utiliser le modèle physique complet. Pour contourner ce problème, nous avons construit un métamodèle, c'est-à-dire une approximation mathématique du modèle physique, fondé sur l'émulation statistique. Nous avons utilisé deux méthodes d'interpolation : les RBF — fonctions de base radiale — et le krigeage (processus gaussiens). L'une des difficultés a été de représenter l'incertitude liée aux variables météorologiques. Ces variables sont de grande dimension (2D ou 4D variables en temps), et leurs incertitudes sont dépendantes de l'espace et du temps, et corrélées entre elles. Pour les représenter, nous avons utilisé un ensemble de 50 prévisions météorologiques fournies par le centre européen de prévision météorologique (CEP), construites pour échantillonner les variables météorologiques incertaines. Nous avons paramétré les champs météorologiques comme une combinaison linéaire convexe de l'ensemble des prévisions, ce qui a soulevé un certain nombre de difficultés dans la conception d'un plan d'expérience, dans la métamodélisation et dans le processus de calibration. De même, neuf termes source, représentant des séquences de rejets dans l'atmosphère compatibles avec les observations dans l'environnement lors de l'accident de Fukushima, et issus de la littérature, ont été combinés avec des poids aléatoires. Une étude comparative a montré que le temps de construction et d'exécution de RBF était assez faible, avec une émulation acceptable des résultats issus du modèle complet. Le krigeage, plus lent, a donné de meilleurs résultats et a permis en outre d'obtenir un écart type de l'erreur de prédiction. Cet écart type a été évalué par comparaison à un ensemble de résultats de simulations généré spécifiquement pour la vérification (et non utilisé pour la construction de l'émulateur).

Les métamodèles ont ensuite été utilisés dans une procédure de calibration d'ensemble. Dans un premier temps, une optimisation déterministe a été réalisée. Le but était de trouver

le meilleur ensemble de paramètres qui minimise les scores du modèle par comparaison aux données, comme la RMSE - Root Mean Square Error. Ensuite, une optimisation probabiliste a été étudiée afin d'obtenir une distribution de probabilité pour chaque paramètre (poids des combinaisons convexes et autres paramètres scalaires du modèle). L'algorithme de Metropolis-Hastings a été appliqué pour échantillonner la distribution des paramètres. La vraisemblance tenait compte des incertitudes des données d'observation. Les échantillons obtenus par la méthode de Metropolis-Hastings étaient constitués de plusieurs chaînes de Markov exécutées en parallèle. Cette méthode a été tout d'abord appliquée à un modèle météorologique simple, calibré sur la base des observations météorologiques disponibles sur le sol japonais durant l'accident de Fukushima. Pour la calibration complète des incertitudes liées au modèle de dispersion, des difficultés ont découlé du nombre élevé de paramètres d'entrée (19), de la faible sensibilité aux poids de l'ensemble météorologique et de la combinaison convexe des prévisions météorologiques.

Mots clés :

Métamodèle,	calibration	statistique,	quantification	d'incertitude,	mo-
dèle	de	dispersion	atmosphérique,	MCMC,	Fukushima,
					crise.

Abstract :

In case of a nuclear accident that may imply a release of radionuclides into the environment, the French Institute of Radiation Protection and Nuclear Safety (IRSN) uses atmospheric dispersion models to infer mitigation actions and protect the population. In the post-accidental phase (after the release actually occurred), they also complement field measurements for the assessment of short- and long-term environmental and human health impacts. These atmospheric dispersion models compute, over a given domain, the concentrations of radionuclides in the atmosphere, the deposited quantities on the ground, as well as the gamma dose rates (radiation emitted by radionuclides present in the air and on the ground).

The forecasts made by these simulations involve large uncertainties, due to the limited knowledge on input data (weather forecasts and source term), physical parameters of the model (deposition velocities, diffusion coefficients, scavenging factors, etc.), and numerical approximations. The aim of our study was to implement methods to assign probability density functions (PDFs) to the unknown errors on the input variables and to calibrate those PDFs using field observations. This was applied to the Fukushima nuclear disaster, using IRSN's operational dispersion model and field measurements collected during or after the disaster. In the evaluation and calibration processes, we made use of radiological observations of activity concentration in the air, deposition and gamma dose rates collected in Japan.

Calibration methods require a large number of calls to the model, so that the computational time is far too high to use the physical model. To circumvent this issue, we built a metamodel, that is, a mathematical approximation of the physical model, based on statistical emulation. We used two interpolation methods : RBFs—Radial Basis Functions—and Kriging (Gaussian processes). One of the difficulties was to deal with the uncertain meteorological fields. Those are time-varying 2D or 3D variables whose uncertainties are time and space-dependent, as well as correlated between the different variables. To describe these uncertainties, we use an ensemble of 50 meteorological forecasts provided by the European Centre for Medium-Range Forecast (ECMWF) whose purpose is to sample the uncertain meteorological variables. We parameterized the set of possible meteorological fields as a convex linear combination of the ensemble of forecasts, which raised a number of difficulties in the design of experiment, in the metamodeling and the calibration. Similarly, nine credible release scenarios, compatible with the observations in the environment, were derived from the Fukushima literature and combined with random weights. A comparison study showed that the construction and execution time of RBF were rather low, with acceptable emulation of the results obtained with the physical model. Kriging was slower but showed better results, and provided in addition a standard deviation of the prediction error. This standard deviation was evaluated on an independent verification set of simulations (not used to build the emulator).

The metamodels were then used in an ensemble-based calibration procedure. Firstly, a deterministic optimization was carried out. The aim was to find the best set of parameters that would minimize model-to-data scores such as the RMSE—Root Mean Square Error. Then, a probabilistic calibration was implemented in order to obtain a probability distribution for each parameter (convex combination and other parameters). The Metropolis-Hastings algorithm was applied to sample the distribution of uncertain parameters. The likelihood took into account the uncertainties in observational data. Samples obtained by Metropolis-Hastings were made of several Markov chain run in parallel. This was first applied to a simple meteorological model,

using meteorological observations available on the Japanese territory during the Fukushima accident. For the full calibration of all uncertainties related to the dispersion model, difficulties arose from the high number of input parameters (19), the low sensitivity to meteorological ensemble weights, and the convex combination of the meteorological forecasts.

Keywords :

Metamodel, statistical calibration, uncertainty quantification, atmospheric dispersion model, MCMC, Fukushima, emergency.

Avant-propos

La modélisation de dispersion atmosphérique a pour objectif de simuler la concentration des polluants dans l'air dans un domaine spatial donné pendant une période de temps précise. Le résultat sert à estimer l'effet des polluants sur les écosystèmes. Lors de la modélisation de la dispersion atmosphérique, les principaux processus à prendre en compte sont : le transport, la diffusion, les réactions physico-chimiques et le dépôt au sol. Parallèlement, la dispersion des polluants dépend aussi de nombreux facteurs tels que les conditions de rejet (hauteur et quantité de rejet de chaque polluant, moment d'émission, etc.), les conditions météorologiques (pluie, vent, température) et orographiques (type de végétation, relief).

Les applications de la modélisation de la dispersion atmosphérique s'étendent à de nombreux domaines : prévision de la qualité de l'air, suivi d'installations industrielles, évaluation des risques chimiques ou radiologiques, etc. L'émission d'un produit dans l'environnement peut avoir une caractéristique :

- chronique, représentée par des rejets plus ou moins continus dans le temps,
- accidentelle, représentée par des échappements soudains de particules ou gaz. Une pollution accidentelle est, comme son nom l'indique, une pollution imprévue, et qui nécessite la mise en place de moyens de gestion pour en limiter les effets sur les populations et l'environnement.

À l'IRSN – Institut de Radioprotection et de Sécurité Nucléaire, des modèles de dispersion atmosphérique sont développés et intégrés dans la plateforme *C3X*. Ils sont utilisés afin d'estimer les conséquences sanitaires et environnementales suite à un rejet accidentel de radionucléides dans l'atmosphère. Pour simuler la dispersion à l'échelle locale, l'IRSN utilise le modèle *pX* lorsqu'on se trouve à courte distance de la source de rejet (à moins de 80 km). Pour une échelle plus grande, régionale ou continentale, un autre modèle à longue distance est utilisé, nommé *ldX*. Ces modèles permettent de simuler la concentration dans l'air, le dépôt au sol et le débit dose du rayonnement gamma émis par les radionucléides. Leurs résultats sont utilisés pour donner des recommandations aux autorités qui décident des actions de protection de la population (évacuation, mise à l'abri, prise d'iode stable) en fonction des prévisions de dose sur le territoire environnant le site accidenté.

L'accident nucléaire de Fukushima a eu lieu le 11 mars 2011 au Japon. Un gigantesque tsunami créé par un violent séisme de magnitude 9 a entraîné une panne des systèmes de refroidissement de secours des réacteurs nucléaires de la centrale de Fukushima-Daiichi. Cette catastrophe, de niveau 7 sur l'échelle INES–International Nuclear Event Scale, est la plus grave arrivée dans une centrale nucléaire depuis l'accident de Tchernobyl en 1986. C'est un accident combinant les conséquences d'un accident nucléaire et d'un tremblement de terre. Pour la sécurité des habitants suite aux fuites de radiation dans l'environnement, l'évacuation a été effectuée dans un rayon de 20 kilomètres, avec une extension à certaines zones situées au-delà en fonction des mesures de radioactivité dans l'environnement. Cette évacuation, justifiée par les normes de radioprotection, a eu des conséquences profondes sur l'économie, le bien-être et la santé des personnes évacuées (traumatisme psycho-social, décès de personnes malades ou âgées, etc.). Or,

dans une situation accidentelle telle que celle de Fukushima, les incertitudes sont par nature très importantes, notamment par manque de connaissance précise sur la situation réelle de l'installation lié à des défauts de matériels (sondes de mesure défectueuses, par exemple), des problèmes de communication, des incompréhensions, des erreurs humaines. Dans une telle situation, les décisions pour protéger les populations doivent être prises le plus rapidement possible sur la base d'informations partielles. Pour limiter les conséquences sanitaire, des marges assez grandes sont généralement prises pour englober ces incertitudes et éviter d'avoir des populations exposées. Une estimation plus précise des zones contaminées, incluant une gradation dans l'évaluation du risque, permettrait d'augmenter l'efficacité des actions de protection en ciblant des zones prioritaires et, potentiellement, de diminuer les dommages liés aux évacuations en limitant leur recours.

L'accident de Fukushima a conduit à de très nombreuses mesures de radioactivité dans l'environnement. Des équipes de recherche dans le monde entier ont cherché à reproduire la séquence accidentelle qui a abouti à la contamination observée dans l'environnement. Pour bien reproduire la radioactivité qui a été mesurée sur le territoire japonais et dans le monde, plusieurs données clés sont nécessaires. Tout d'abord, les champs météorologiques, donnant l'évolution de variables telles que le vent ou la pluie sur tout le territoire au cours de la période accidentelle. Ces données sont fournies par des services de prévision météorologique qui s'appuient sur des modèles (eux-mêmes entachés d'incertitudes) et des observations météorologiques (dont beaucoup sont manquantes dans les premiers jours de l'accident de Fukushima, les stations de mesure ayant été endommagées par le tsunami). Le deuxième élément clé est le terme source, c'est-à-dire la quantité et la composition du rejet au cours du temps. Ces données sont encore très incertaines plusieurs années après l'accident, notamment à cause des difficultés à reconstituer la séquence accidentelle par manque d'informations et de mesures fiables sur l'état du réacteur. De ce fait, les termes sources publiés dans la littérature sont tous issus d'une reconstitution sur la base de mesures radiologiques dans l'environnement, d'informations météorologiques et de modèles de dispersion (méthodes de modélisation inverse). Enfin, la dernière « brique » permettant de simuler correctement la dispersion du panache accidentel est le modèle de dispersion, lui aussi entaché d'incertitudes. Du fait des nombreuses observations radiologiques disponibles, la catastrophe de Fukushima est un cas d'étude privilégié pour la validation et l'amélioration des outils opérationnels de dispersion atmosphérique utilisés en cas d'accident. Toutefois, celle-ci ne peut se faire sans prendre en compte les nombreuses sources d'incertitudes. Cette étude vise à utiliser les mesures de radioactivité collectées sur le territoire japonais dans la calibration statistique du modèle de longue distance ℓdX de l'IRSN qui sera présenté dans la partie 1.3 du chapitre 1. L'objectif est de quantifier les incertitudes liées aux données d'entrée (terme source et météorologie) ainsi qu'au modèle lui-même.

Plan de la thèse

Le chapitre 1 a pour objectif d'introduire les notions générales de la modélisation de transport des polluants dans l'atmosphère, incluant la structure de l'atmosphère et les phénomènes impliqués dans le transport réactif. La première partie 1.3 de ce chapitre présente l'équation d'advection-diffusion sur la concentration de polluants dans l'air. Ensuite, nous parlerons des

sources d'incertitude de ce type de modèle numérique. La partie 1.5 présentera en détail ses principales caractéristiques, ainsi que la prise en compte des incertitudes de chaque élément. La partie 1.4 présentera les principaux événements du déroulement de l'accident de Fukushima. Enfin, nous introduirons la problématique stochastique associée et présenterons les algorithmes mathématiques qui permettent de traiter les incertitudes de modèle, dans la section 1.1, ainsi les méthodes de réduction du modèle par interpolation dans la section 1.2.

Le deuxième chapitre parlera de la mise en place d'un processus de propagation d'incertitude par méthode de Monte Carlo 2. Ce travail était l'objet principal de la première année de la thèse. Un ensemble de deux cent simulations du modèle ℓdX a été réalisé en faisant varier indépendamment des paramètres d'entrée suivant leurs plages de variation. Nous avons comparé ensuite le résultat de ces simulations aux observations radiologiques, collectées directement sur le territoire japonais par nos partenaires. De nombreux indicateurs et scores ont été utilisés pour évaluer la relation entre les mesures et les sorties du modèle numérique.

La deuxième et une partie de la troisième année de la thèse ont concerné la création de métamodèles. La difficulté principale de ce travail se trouve dans la construction du plan d'expérience, surtout la modélisation de l'ensemble de prévision météorologique et les termes sources. Cette étape a pour objectif de créer une imitation mathématique du modèle physique ℓdX afin de réduire les temps de calcul. Pour le cas de Fukushima, avec une résolution assez grossière, une simulation de ℓdX dure environ 10 minutes. Ce chiffre n'est pas énorme mais une étude de calibration statistique a besoin de plusieurs millions d'appels au modèle. Un calcul de 10 minutes devient problématique. Les métamodèles sont établis suivant des méthodes d'apprentissage statistique, qui donnent une prédiction proche du résultat du modèle physique mais avec un temps de calcul beaucoup plus court (de l'ordre de quelques secondes).

À ce stade, nous avons réuni tous les éléments nécessaires pour effectuer une étude de calibration statistique du modèle ℓdX . Cependant, avant de l'appliquer directement au modèle radiologique, qui est très complexe, l'algorithme d'échantillonnage a d'abord été testé dans un cas plus simple, consistant à calibrer un modèle météorologique statistique simplifié sur les observations météorologiques disponibles au Japon sur la période de l'accident 4, pour ensuite appliquer la méthode de calibration au modèle de dispersion atmosphérique avec les observations radiologiques de Fukushima 5.

Table des matières

1	Introduction	9
1.1	Approche stochastique pour un modèle numérique	9
1.1.1	Introduction de l'inférence bayésienne	11
1.1.2	Méthodes de Monte Carlo par chaîne de Markov (MCMC)	14
1.1.3	Diagnostic de convergence des chaînes de Markov	15
1.2	Métamodélisation : objectifs et méthodes	17
1.2.1	Plan d'expérience par hypercube latin	17
1.2.2	Construction du métamodèle	19
1.2.3	Interpolation par fonctions de base radiale-RBF	20
1.2.4	Processus gaussien, krigeage	21
1.3	Modèle de dispersion atmosphérique	23
1.3.1	Caractéristiques de l'atmosphère	23
1.3.2	Dispersion de polluants	26
1.3.3	Modélisation mathématique du phénomène de transport atmosphérique	27
1.3.4	Résolution numérique de l'équation advection-diffusion	28
1.4	Accident de la centrale nucléaire Fukushima-Daiichi	32
1.5	Principales sources d'incertitudes	35
1.5.1	Ensemble de prévisions météorologiques	35
1.5.2	Termes sources de l'accident de Fukushima	38
1.5.3	Paramètres incertains dans le modèle	39
2	Propagation d'incertitude par méthode de Monte Carlo	43
2.1	Introduction	43
2.1.1	Context	43
2.1.2	Uncertainty analysis	44
2.1.3	Objectives of the study	46
2.2	Analysis of a meteorological ensemble	46
2.2.1	Description of ensemble weather forecasts	46
2.2.2	Model-to-data comparison of the meteorological ensemble	49
2.3	Monte Carlo simulations	56
2.3.1	ℓdX description	56
2.3.2	Simulation setup	56
2.3.3	Description of the measurements	58
2.3.4	Model-to-data comparison for Monte Carlo simulations	60
2.4	Synthesis and perspectives	67
3	Métamodélisation du modèle de dispersion à longue distance ℓdX	71
3.1	Introduction	71
3.2	Creation of DOE and model's output	72
3.2.1	Input : Physical parameters	72

3.2.2	Input : Source term	73
3.2.3	Input : Weather forecast	74
3.2.4	Outputs' emulation	75
3.3	Construction of metamodels	75
3.3.1	RBFs—Radial Basis Functions	76
3.3.2	Kriging—Gaussian processes (GP) modeling	76
3.4	Metamodels' evaluation	77
3.4.1	RBFs vs Kriging	78
3.4.2	Studying error predicted by GP models	80
3.4.3	Comparing metamodels with different meteorological perturbations	82
3.5	Applications of metamodels	85
3.5.1	Sensitivity analysis	85
3.5.2	Minimization of the RMSE on concentration	85
3.6	Conclusions and perspectives	89
4	Calibration statistique du modèle météorologique	91
4.1	Metropolis-Hastings et diagnostic de convergence Gelman-Brooks	92
4.1.1	Algorithme Metropolis-Hastings	92
4.1.2	Vérification du critère Gelman-Brooks sur un cas simple	93
4.2	Calibration du modèle météorologique	94
4.2.1	Approche bayésienne	96
4.2.2	Construction de la loi <i>a posteriori</i>	97
4.2.3	Échantillonnage par Metropolis-Hastings	98
4.2.4	Vérification du résultat de calibration	100
4.3	Conclusions	100
5	Optimisation probabiliste des sorties radiologiques de ℓdX	103
5.1	Construction de la loi <i>a posteriori</i> par inférence bayésienne	103
5.1.1	Prise en compte de la corrélation des observations radiologiques dans la vraisemblance	103
5.1.2	Informations <i>a priori</i> sur les paramètres incertains	104
5.2	Cas synthétique : jeu d'entrée optimal connu	104
5.2.1	Calibration du dépôt avec observations corrélées	106
5.2.2	Calibration d'activité volumique intégrée sans corrélation entre les observations	110
5.3	Cas réel : observations lors de l'accident de Fukushima	111
5.3.1	Évaluation des métamodèles utilisés	113
5.3.2	Calibration avec le débit dose et une corrélation entre les observations	114
5.3.3	Calibration avec l'activité volumique et sans corrélation	115
5.3.4	Calibration avec le dépôt et avec corrélation	123
5.4	Conclusion	124
6	Conclusion	127

Annexes	131
A Annexe du chapitre d'introduction	133
A.1 Fonction de covariance	133
A.2 Déterminant d'une matrice carré	135
B Annexe d'article Monte Carlo	139
B.1 Interpreting the rank histogram	139
B.2 Illustration of subgrid effects	139
B.3 Figure of Merit in Space – FMS	139
B.4 Figure of Merit in Time – FMT	140
B.5 Root Mean Square Error – RMSE	140
C Annexe d'article de métamodélisation	141
C.1 LHS and complementary LHS	141
C.2 Dirichlet distribution	141
C.3 Reduction of weather ensemble size	143
Bibliographie	145

Introduction

Dans ce chapitre, nous exposons une démarche stochastique dans le prolongement de la modélisation numérique, en particulier avec des algorithmes de traitement des incertitudes utilisés dans la calibration statistique autour de modèles numériques. De plus, une partie centrale de cette thèse se concentre sur la métamodélisation, pour conduire une réduction de modèle numérique. Il en résulte un métamodèle qui est ensuite utilisé dans l'algorithme de calibration afin de réduire les temps de calcul. Tout ce processus, y compris la métamodélisation et la calibration statistique, est appliqué au problème de la dispersion atmosphérique sur le cas de l'accident nucléaire de Fukushima en 2011.

Contents

1.1	Approche stochastique pour un modèle numérique	9
1.1.1	Introduction de l'inférence bayésienne	11
1.1.2	Méthodes de Monte Carlo par chaîne de Markov (MCMC)	14
1.1.3	Diagnostic de convergence des chaînes de Markov	15
1.2	Métamodélisation : objectifs et méthodes	17
1.2.1	Plan d'expérience par hypercube latin	17
1.2.2	Construction du métamodèle	19
1.2.3	Interpolation par fonctions de base radiale-RBF	20
1.2.4	Processus gaussien, krigeage	21
1.3	Modèle de dispersion atmosphérique	23
1.3.1	Caractéristiques de l'atmosphère	23
1.3.2	Dispersion de polluants	26
1.3.3	Modélisation mathématique du phénomène de transport atmosphérique	27
1.3.4	Résolution numérique de l'équation advection-diffusion	28
1.4	Accident de la centrale nucléaire Fukushima-Daiichi	32
1.5	Principales sources d'incertitudes	35
1.5.1	Ensemble de prévisions météorologiques	35
1.5.2	Termes sources de l'accident de Fukushima	38
1.5.3	Paramètres incertains dans le modèle	39

1.1 Approche stochastique pour un modèle numérique

Considérons un modèle $\mathcal{G}$ qui transforme des entrées $\mathbf{x} \in \mathbb{R}^d$ en sorties $\mathbf{y} \in \mathbb{R}^n$. Par la suite, $\mathcal{G}$ représente un modèle numérique, que l'on utilise pour modéliser un phénomène physique, par

exemple pour calculer la concentration des particules dans l'air, et peut être noté comme suit

$$\mathbf{y} = \mathcal{G}(\mathbf{x}).$$

Nous appelons également $\mathbf{y}^{obs}$ le vecteur qui contient des mesures réelles de la sortie $\mathbf{y}$ du modèle $\mathcal{G}$. Si les valeurs obtenues de $\mathbf{y}^{obs}$ sont exactes, c'est-à-dire qu'il n'y a aucune différence entre la vraie valeur $\mathbf{y}^*$ et celle mesurée $\mathbf{y}^{obs}$, on dit que ces observations ne contiennent pas d'erreur, aussi appelée *erreur d'observation*. En réalité, il existe toujours une marge entre la vraie valeur $\mathbf{y}^*$ et ses valeurs mesurées. Prenons un exemple de mesure de température de l'atmosphère. Au même endroit et même moment, il est possible que deux appareils de même modèle fournissent deux valeurs distinctes, même si la différence est très faible. Des erreurs peuvent provenir de diverses sources : la mise en œuvre de la mesure, l'état de l'appareil, le lieu d'exposition (à l'abri ou exposé), etc. Nous notons cette erreur d'observation ε_{obs} , et alors

$$\mathbf{y}^* = \mathbf{y}^{obs} + \varepsilon_{obs}, \quad \varepsilon_{obs} \in \mathbb{R}^n.$$

D'autre part, le modèle $\mathcal{G}$, qui représente un phénomène physique, est sujet à des incertitudes, souvent fortes, que nous appelons ici *erreur du modèle*. Cette erreur peut être la conséquence de la mise en équation du phénomène physique, de la discrétisation dans la modélisation numérique ou des incertitudes des paramètres d'entrée $\mathbf{x}$. La relation entre la simulation de $\mathcal{G}$, les observations $\mathbf{y}^{obs}$ et la réalité $\mathbf{y}^*$ est introduite de manière suivante

$$\begin{aligned} \mathbf{y}^* &= \mathbf{y}^{obs} + \varepsilon_{obs} \\ &= \mathcal{G}(\mathbf{x}) + \varepsilon_{mod} \end{aligned} \tag{1.1}$$

où ε_{mod} représente l'erreur du modèle $\mathcal{G}$.

La vraie valeur $\mathbf{y}^*$ n'est pas connue. Dans certaines situations, elle est seulement estimée à partir des valeurs observées $\mathbf{y}^{obs}$ et des erreurs d'observation ε_{obs} évaluées par l'expérience. Pour cette raison, l'équation 1.1 est réduite à une relation entre les observations et l'estimation du modèle numérique $\mathcal{G}$:

$$\mathbf{y}^{obs} = \mathcal{G}(\mathbf{x}) + \varepsilon^* \tag{1.2}$$

où $\varepsilon^* = \varepsilon_{mod} - \varepsilon_{obs}$ est l'écart total entre les observations $\mathbf{y}^{obs}$ et le modèle $\mathcal{G}$. Cette relation modèle-mesure 1.2 sera utilisée pour former la fonction de vraisemblance, qui sera présentée dans la partie 1.1.1.

Dans plusieurs domaines, le modèle numérique $\mathcal{G}$ a été validé par des tests en laboratoire avec des conditions maîtrisées pour réduire l'erreur du modèle ε_{mod} . Dans ce cas, l'écart modèle-mesure se forme principalement à cause des incertitudes de l'entrée $\mathbf{x}$. Ces incertitudes sont souvent déterminées *a priori* suivant un jugement d'expert et la littérature du domaine étudié. L'objectif de cette thèse est d'appliquer les méthodes de calibration statistique pour retrouver la distribution de probabilité de $\mathbf{x}$, qui permet de réduire l'erreur modèle-mesure. Autrement dit, nous voulons utiliser les informations apportées par les observations $\mathbf{y}^{obs}$ pour réduire et estimer à la fois les incertitudes d'entrée. Cette partie propose de poser le problème mathématique de l'inférence bayésienne qui sera appliqué au cas de Fukushima dans les chapitres 4 et 5.

1.1.1 Introduction de l'inférence bayésienne

Rappels de probabilités

La démarche statistique classique repose essentiellement sur un principe de vraisemblance qui considère que ce qui a été observé rend compte de manière exhaustive du phénomène. Mais les observations ne fournissent qu'une image du phénomène et celle-ci peut être mauvaise. Un certain nombre de théorèmes permettent d'évaluer la qualité des estimateurs si le nombre d'observation est suffisant. L'analyse bayésienne des problèmes statistiques propose d'introduire dans la démarche d'inférence, l'information dont nous disposons *a priori*.

Définition 1 (Probabilité conditionnelle). *Soient A et B deux événements tels que $\mathbb{P}(B) \neq 0$, on appelle la probabilité conditionnelle de A sachant B la valeur suivante :*

$$\mathbb{P}(A|B) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

Théorème 1 (Probabilités totales). *Soient A et B deux événements tels que $\mathbb{P}(B) \neq 0$, alors :*

$$\mathbb{P}(B) = \mathbb{P}(B|A) \mathbb{P}(A) + \mathbb{P}(B|\bar{A}) \mathbb{P}(\bar{A})$$

Démonstration.

$$\begin{aligned} \mathbb{P}(B) &= \mathbb{P}(B \cap A) + \mathbb{P}(B \cap \bar{A}) \\ &= \mathbb{P}(B|A) \mathbb{P}(A) + \mathbb{P}(B|\bar{A}) \mathbb{P}(\bar{A}) \end{aligned}$$

□

Théorème 2 (Bayes). *Soient A et B deux événements tels que $\mathbb{P}(B) \neq 0$, alors le théorème de Bayes sur les ensembles discrets s'écrit comme suit :*

$$\begin{aligned} \mathbb{P}(A|B) &= \frac{\mathbb{P}(B|A) \mathbb{P}(A)}{\mathbb{P}(B)} \\ &= \frac{\mathbb{P}(B|A) \mathbb{P}(A)}{\mathbb{P}(B|A) \mathbb{P}(A) + \mathbb{P}(B|\bar{A}) \mathbb{P}(\bar{A})} \end{aligned}$$

Démonstration.

$$\mathbb{P}(A|B) \mathbb{P}(B) = \mathbb{P}(A \cap B) = \mathbb{P}(B|A) \mathbb{P}(A)$$

□

En réalité, il existe souvent plus que deux événements à consulter. L'extension de la probabilité conditionnelle 1 sur deux événements indépendants est introduite par la définition 2.

Définition 2 (Probabilité conditionnelle avec notion d'indépendance). *Soient A , B et C des événements tels que $\mathbb{P}(C) \neq 0$, alors A est indépendant de B conditionnellement à C si :*

$$\mathbb{P}(A \cap B|C) = \mathbb{P}(A|C) \mathbb{P}(B|C)$$

Dans cette étude, nous travaillons exclusivement sur les probabilités continues. La densité conditionnelle sera définie de manière suivante :

Définition 3 (Densité conditionnelle). *Soient X et Y deux variables aléatoires de la loi jointe $f(x, y)$ sous réserve de non nullité du dénominateur, on définit la densité conditionnelle :*

$$f(x|y) := \frac{f(x, y)}{\int f(x, y) dx}$$

Modèle statistique

On cherche à estimer une quantité d'intérêt de $\hat{x}$ à partir d'observations $\mathbf{y}^{obs} = (y_1^{obs}, \dots, y_n^{obs})$. Pour cela, on se donne un modèle statistique qui consiste à se donner $\mathbf{Y} = (Y_1, \dots, Y_n)$ un vecteur aléatoire continu à valeur dans $\mathbb{R}^n$ dont les composantes sont indépendantes et suivent la même loi. Cette loi dépend d'un paramètre $\mathbf{x} \in \Theta \subset \mathbb{R}^p$. On définit une manière de mesurer la qualité d'un paramètre donné pour un ensemble d'observations :

Définition 4. *Si $(Y_k)_{k \leq n}$ sont des variables continues iid (indépendantes et identiquement distribuées) dont la loi est une densité f_X dépendant d'un paramètre $\mathbf{X}$, on appelle la fonction de vraisemblance la fonction $\mathcal{L}$ définit par :*

$$\mathcal{L}(\mathbf{X}, \mathbf{Y} = \mathbf{y}^{obs}) = \prod_{k=1}^n f_X(Y_k)$$

La fonction de vraisemblance $\mathcal{L}(\mathbf{X}, \mathbf{Y} = \mathbf{y}^{obs}) := \mathbb{P}(\mathbf{Y} = \mathbf{y}^{obs} | \mathbf{X})$ n'est pas une probabilité de $\mathbf{X}$, mais c'est la probabilité des observations $\mathbf{y}^{obs}$ sachant $\mathbf{X}$. La maximisation donne $\mathbf{x}_{MLE}$, paramètre du maximum de la vraisemblance, qui est la valeur la plus vraisemblable de $\mathbf{X}$ dans Θ pour les observations $\mathbf{y}^{obs} = (y_1^{obs}, \dots, y_n^{obs})$. Cette fonction joue un rôle central dans la méthode bayésienne.

D'autre part, la relation entre les observations $\mathbf{y}^{obs}$ et les paramètres $\mathbf{x}$ a été introduite par l'équation 1.2. Autrement dit, la vraisemblance $\mathcal{L}(\mathbf{X}, \mathbf{y}^{obs})$ correspond à la distribution de l'erreur $\boldsymbol{\varepsilon}^*$. Dans la suite de cette étude, nous supposons que l'écart modèle-mesure $\boldsymbol{\varepsilon}^* = (\varepsilon_1, \dots, \varepsilon_n)$ suit une loi normale centrée multidimensionnelle de dimension n de covariance $\Sigma \in \mathbb{R}^{n \times n}$, notée comme suit

$$\boldsymbol{\varepsilon}^* \sim \mathcal{N}(\mathbf{0}, \Sigma)$$

où Σ est une matrice réelle définie comme suit

$$\Sigma = \begin{pmatrix} \sigma_1^2 & cov(\varepsilon_1, \varepsilon_2) & cov(\varepsilon_1, \varepsilon_3) & \cdots & cov(\varepsilon_1, \varepsilon_n) \\ cov(\varepsilon_2, \varepsilon_1) & \sigma_2^2 & cov(\varepsilon_2, \varepsilon_3) & \cdots & cov(\varepsilon_2, \varepsilon_n) \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ cov(\varepsilon_n, \varepsilon_1) & cov(\varepsilon_n, \varepsilon_2) & cov(\varepsilon_n, \varepsilon_3) & \cdots & \sigma_n^2 \end{pmatrix}$$

où σ_i est l'écart type de la loi marginale de la i ème composante de $\boldsymbol{\varepsilon}^*$ et $cov(a, b)$ est la fonction de covariance entre deux éléments a et b . Les caractéristiques de cette fonction seront détaillées dans l'annexe A.1. La matrice de covariance Σ est toujours symétrique grâce à la symétrie de

la fonction de covariance. Dans le cas d'indépendance entre les composantes de ε^* , la matrice de covariance Σ devient diagonale. En conséquence de la définition 4, la vraisemblance qui relie les entrées $\mathbf{x}$ et les observations $\mathbf{y}^{obs}$ s'écrit de manière suivante :

$$\begin{aligned}
 \mathcal{L}(\mathbf{X}, \mathbf{y}^{obs}) = f_{\varepsilon^*}(\mathbf{y}^{obs} - \mathcal{G}(\mathbf{X})) &= \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2} (\mathbf{y}^{obs} - \mathcal{G}(\mathbf{X}))^T \Sigma^{-1} (\mathbf{y}^{obs} - \mathcal{G}(\mathbf{X}))\right) \\
 &= \prod_{i=1}^n \frac{1}{\sigma_i \sqrt{2\pi}} \exp\left(-\frac{(y_i^{obs} - g_i(\mathbf{X}))^2}{2\sigma_i^2}\right) \\
 &= (2\pi)^{-n/2} \prod_{i=1}^n \exp(-\log(\sigma_i)) \cdot \exp\left(-\frac{(y_i^{obs} - g_i(\mathbf{X}))^2}{2\sigma_i^2}\right) \\
 &= (2\pi)^{-n/2} \exp\left[-\sum_{i=1}^n \log(\sigma_i) + \frac{(y_i^{obs} - g_i(\mathbf{X}))^2}{2\sigma_i^2}\right] \quad (1.3)
 \end{aligned}$$

où $\mathcal{G}(\mathbf{x}) = (g_1(\mathbf{x}), \dots, g_n(\mathbf{x}))$ et $|\Sigma|$ est le déterminant de la matrice Σ qui est détaillé dans l'annexe A.2. Une application concrète du développement de la vraisemblance 1.3 sera expliqué dans la partie sur la calibration d'un modèle météorologique simplifié 4.

Approche bayésienne

Nous appelons l'information *a priori* sur $\mathbf{X}$ toute information disponible sur ce paramètre en dehors de celle apportée par les observations. L'information *a priori* de $\mathbf{X}$ est entachée d'incertitude. Il est essentiel de modéliser cette information à travers une loi de probabilité, appelée loi *a priori*. Sa densité est notée $\pi(\mathbf{x})$. L'approche bayésienne consiste à utiliser cette loi *a priori* du paramètre $\mathbf{X}$ et la loi des observations. On appelle loi des observations, la loi conditionnelle de $\mathbf{Y}^{obs}$ sachant $\mathbf{X}$, notée $f(\mathbf{y}|\mathbf{x})$. Dans la plupart des cas, cette loi des observations est représentée par la vraisemblance, définie dans le paragraphe qui précède (1.1.1).

Alors que les méthodes statistiques classiques reposent sur la loi des observations, ou la fonction de vraisemblance, la méthode bayésienne repose sur la loi *a posteriori*, notée $\pi(\mathbf{x}|\mathbf{y})$. La loi *a posteriori* peut s'interpréter comme le résumé (dans le sens probabiliste) des informations disponibles sur $\mathbf{X}$, une fois que $\mathbf{y}^{obs}$ est observé. L'essentiel de la méthode bayésienne consiste, sur la base de la loi *a priori* et de l'observation effectuée, à déterminer la loi du paramètre conditionnellement aux observations, la loi *a posteriori* qui actualise les informations sur ce paramètre.

Suivant la figure 1.1, la loi *a posteriori* est la loi conditionnelle de $\mathbf{X}$ sachant $\mathbf{y}^{obs}$. En vertu de la formule de Bayes,

$$\pi(\mathbf{x}|\mathbf{Y} = \mathbf{y}^{obs}) = \frac{\mathcal{L}(\mathbf{x}, \mathbf{y}^{obs}) \pi(\mathbf{x})}{\int_{\Theta} \mathcal{L}(\mathbf{x}, \mathbf{y}^{obs}) \pi(\mathbf{x}) d\mathbf{x}}$$

où $\int_{\Theta} \mathcal{L}(\mathbf{x}, \mathbf{y}^{obs}) \pi(\mathbf{x}) d\mathbf{x}$ est une constante, souvent inconnue. On peut alors constater que la loi *a posteriori* est connue à un facteur près : elle est proportionnelle au produit de la

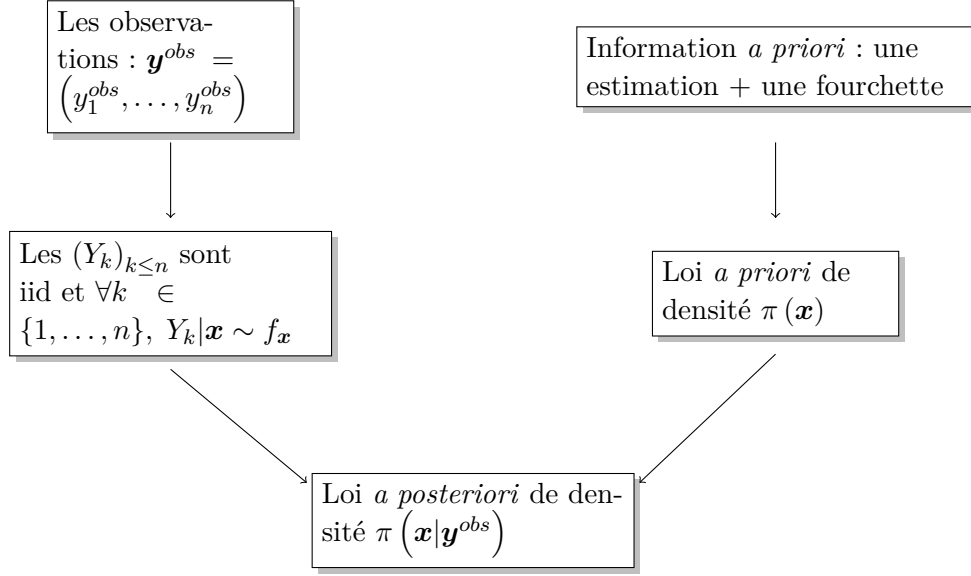


FIGURE 1.1 – Résumé de l’approche bayésienne dans le cadre de la statistique paramétrique inférentielle, source [Dupuis 2007].

vraisemblance et de la loi *a priori*, et donc $\pi(\mathbf{x} | \mathbf{y}^{obs}) \propto \mathcal{L}(\mathbf{x}, \mathbf{y}^{obs}) \pi(\mathbf{x})$. Toute la statistique bayésienne repose sur cette loi *a posteriori*. Les difficultés essentielles proviennent d’une part, de la détermination et du choix de la loi *a priori* et, d’autre part, du calcul explicite de la loi *a posteriori*.

1.1.2 Méthodes de Monte Carlo par chaîne de Markov (MCMC)

Le but des méthodes MCMC est d’approcher une loi cible g à l’aide d’une chaîne de Markov de mesure invariante g . Dans notre étude, nous souhaitons échantillonner la distribution *a posteriori* $\pi(\mathbf{X} | \mathbf{y}^{obs})$, qui remplacera la loi cible g . On peut ensuite utiliser cette approche pour fournir une estimation de la variable recherchée. En effet, si $Z_1, \dots, Z_m \sim \pi(\mathbf{x} | \mathbf{y}^{obs})$, alors on peut prendre comme estimateur de $\mathbf{X}$:

- $\hat{\mathbf{X}}_m := \frac{1}{m} \sum_{i=1}^m Z_i$ (estimateur de Monte Carlo) ;
- ou $\hat{\mathbf{X}}_m := \text{mediane}(Z_1, \dots, Z_m)$;
- ou $\hat{\mathbf{X}}_m := \text{argmax hist}(Z_1, \dots, Z_m)$.

L’idée générale des méthodes de MCMC est de considérer une chaîne de Markov qui produit des échantillons corrélés

$$\mathbf{x}^{(0)} \longrightarrow \mathbf{x}^{(1)} \longrightarrow \mathbf{x}^{(2)} \longrightarrow \dots$$

tels que pour i suffisamment grand, X_i suit à peu près la loi g . On note, par la suite, $\mathcal{X}$ l’espace d’état. Dans notre cas d’étude, $\mathcal{X}$ est fini et inclus dans $\mathbb{R}^p$. On peut définir la chaîne de Markov de la manière suivante :

Définition 5. Une chaîne de Markov $(X_0, x_1, \dots)$ avec $X_i \in \mathcal{X}$ est une suite de variables

aléatoires vérifiant

$$f(X_{i+1}|X_i, \dots, X_1) = K(X_{i+1}|X_i)$$

où $K : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}^+$, appelé *noyau de Markov* ou *noyau de transition*, vérifie : $\forall x \in \mathcal{X}$, $x' \mapsto K(x'|x)$ est une densité de probabilité (ou loi discrète).

Supposons que g s'écrit $g(\mathbf{x}) = \gamma(\mathbf{x})/Z$ avec Z une constante de normalisation. On a également $\mathcal{X} = \Theta$. On veut une chaîne de Markov qui admette g comme mesure invariante et telle que Z n'apparaisse pas dans le noyau de transition. Pour cela on se donne un noyau de Markov $q(\mathbf{x}'|\mathbf{x})$ et on utilise l'algorithme de Metropolis-Hastings (cf. algorithme 1).

Algorithm 1: Algorithme d'échantillonnage de Metropolis-Hastings.

Data:

- la densité cible g ;
- le noyau de transition q ;
- le nombre d'itérations N_{it} ;

Result: un échantillon de $\mathbf{X}$ suivant la distribution g ;

1 initialization : le point de départ $\mathbf{x}^{(0)}$;

2 for $i = 1, \dots, N_{it}$ do

1. générer une nouvelle valeur $\mathbf{x}^* \sim q(\cdot|\mathbf{x}^{(i-1)})$ (loi de proposition ou de transition),
2. calculer la probabilité d'acceptation :

$$\alpha(\mathbf{x}^*, \mathbf{x}^{(i-1)}) = \min \left[1, \frac{g(\mathbf{x}^*) \cdot q(\mathbf{x}^{(i-1)}|\mathbf{x}^*)}{g(\mathbf{x}^{(i-1)}) \cdot q(\mathbf{x}^*|\mathbf{x}^{(i-1)})} \right]$$

3. on prend :

$$\mathbf{x}_i = \begin{cases} \mathbf{x}^* & \text{avec une probabilité } \alpha, \\ \mathbf{x}^{(i-1)} & \text{avec une probabilité } 1 - \alpha \end{cases}$$

3 end

Dans la plupart des cas, la distribution cible g n'est connue qu'à une constante près, mais avec l'algorithme de Metropolis-Hastings, cet aspect n'influence pas le résultat de l'échantillonnage. Nous pouvons exploiter la loi *a posteriori*, $\pi(\mathbf{x}|\mathbf{y}^{obs})$, de l'inférence bayésienne à la place de g . L'objectif est d'obtenir un ensemble de valeurs vraisemblables du paramètre $\mathbf{X}$ sachant les observations $\mathbf{y}^{obs}$.

1.1.3 Diagnostic de convergence des chaînes de Markov

En théorie, la chaîne de Markov atteint la distribution de la loi cible à l'infini. Mais en pratique, avoir une infinité d'itérations est irréaliste. Quelle taille d'échantillon faut-il pour atteindre le régime stationnaire ? Quelques diagnostics de convergence ont été proposés dans les années 90 :

- [Geweke 1991] compare les moyennes calculées à partir de différents segments de la chaîne de Markov ;
- [Raftery 1992] calcule la longueur minimale de la chaîne de Markov pour atteindre un certain intervalle de confiance.

Ces diagnostics se basent sur l'utilisation d'une seule chaîne de Markov. En théorie, une chaîne de Markov atteint sûrement la loi cible, mais le temps d'exécution risque d'être très long dans certains cas. Un autre critère de convergence a été proposé par l'étude de [Brooks 1998] en utilisant plusieurs chaînes en parallèle afin d'accélérer l'obtention de la distribution représentant la loi cible. Cette méthode est une extension du critère proposé par [Gelman 1992]. Suivant ce travail, la convergence vers la distribution cible des chaînes de Markov peut être déterminée en étudiant certaines inférences de différentes chaînes lancées en parallèle. Cette méthode utilise plusieurs réplifications de chaînes pour déterminer si le régime stationnaire a été atteint sur la deuxième moitié de chaque chaîne. La méthode suppose que m chaînes de Markov sont lancées en parallèle en partant de m points distincts. Après $2n$ itérations, les n premiers échantillons sont rejetés car les chaînes n'ont pas encore convergé. Après cette étape, toutes les chaînes sont censées être équiprobables et leurs comportements statistiques sont similaires : la convergence est atteinte ou presque. Plus précisément, la méthode de [Gelman 1992] porte sur la comparaison des inférences (moyenne et variance) apportées par chaque chaîne aux inférences obtenues en mélangeant toutes les chaînes ensemble. Un contrôle de convergence est effectué toutes les $b \approx n/20$ itérations.

Dans le cas multidimensionnel, où les échantillons sont de taille p , [Brooks 1998] a proposé de remplacer les estimateurs scalaires par les $p \times p$ -matrices de covariance B et W , respectivement inter-chaîne et intra-chaîne. Soient $\mathbf{X} \in \mathbb{R}^p$ le vecteur aléatoire que l'on veut échantillonner suivant la distribution cible f , μ le vecteur d'espérance et Σ la matrice de variance-covariance de cette distribution. Appelons $\hat{\mu}$ l'estimateur de μ , $\mathbf{x}_{jt}$ le t ème tirage de la chaîne j . Nous avons :

$$\hat{\mu} = \bar{\mathbf{x}}_{..} = \frac{1}{m} \sum_{j=1}^m \bar{\mathbf{x}}_{j.} = \frac{1}{m} \sum_{j=1}^m \frac{1}{n} \sum_{t=1}^n \mathbf{x}_{jt}$$

L'estimateur de la matrice variance-covariance postérieure Σ est défini par :

$$\hat{V} = \frac{n-1}{n} W + \left(1 + \frac{1}{m}\right) B/n$$

où la variance intra-chaîne W est définie comme suit :

$$W = \frac{1}{m(n-1)} \sum_{j=1}^m \sum_{t=1}^n (\mathbf{x}_{jt} - \bar{\mathbf{x}}_{j.}) (\mathbf{x}_{jt} - \bar{\mathbf{x}}_{j.})'$$

et la variance inter-chaîne B/n est :

$$B/n = \frac{1}{m-1} \sum_{j=1}^m (\bar{\mathbf{x}}_{j.} - \bar{\mathbf{x}}_{..}) (\bar{\mathbf{x}}_{j.} - \bar{\mathbf{x}}_{..})'$$

Les matrices $\hat{V}$, W et B/n sont de taille $p \times p$. L'indice de convergence défini dans cette méthode est la distance entre deux matrices W et $\hat{V}$. Cette distance est résumée par un scalaire

qui devrait être proche de 1 quand l'échantillonnage atteint la convergence. L'une des manières pour mesurer cette distance est de calculer la valeur propre maximale de la matrice $W^{-1}V$. L'idée est qu'à chaque b itérations, on calcule

$$R_c = \max_{\alpha} \frac{\alpha' V \alpha}{\alpha' W \alpha} \quad (1.4)$$

avec $\alpha \in \mathbb{R}^p$. Quand R_c est vraiment proche de 1, au moins d'ordre 10^{-3} , on peut conclure que la distribution cible est atteinte.

Le paragraphe 4.1.2 détaille l'application du diagnostic de Gelman-Brooks dans un cas simple : échantillonner une loi gaussienne de dimension 20. Ce test vise à vérifier la précision de l'indicateur R_c sur la convergence vers la loi cible des deux chaînes de Markov lancées en parallèle. Le résultat montre qu'au bout de 400 000 itérations, l'indicateur R_c vaut 1.001 et l'erreur en norme L^2 entre la moyenne empirique et théorique est d'ordre 0.4 pour un espace de dimension 20. On peut conclure qu'à 400 000 itérations, les échantillons représentent déjà une distribution de la loi cible. L'obtention rapide de la loi cible vient de la simplicité de la loi à échantillonner. En réalité, certaines lois nécessitent des millions d'itérations pour converger, d'où le choix d'une méthode "multi-chaîne".

1.2 Métamodélisation : objectifs et méthodes

La section 1.1 a mis en évidence les problématiques stochastiques liées à un modèle de simulation numérique, ainsi que les méthodes de MCMC pour quantifier des incertitudes de simulation. Cependant, les algorithmes de MCMC requièrent toujours un très grand nombre de simulations, ou appels au modèle original. Ce chiffre peut atteindre des centaines de milliers ou des millions de simulations afin d'obtenir un régime stationnaire qui converge vers la distribution *a posteriori* de $\mathbf{x}$. Si le temps de calcul du modèle $\mathcal{G}$ est court, cela ne pose pas de problème. En revanche, il existe des modèles numériques dont les simulations coûtent cher en temps de calcul, par exemple un modèle eulérien qui résout l'équation advection-diffusion sur un maillage fin tel qu'utilisé ici et défini en partie 1.3.4. Dans certains cas, une simulation dure quelques minutes, et dans d'autres cas, elle peut même prendre jusqu'à quelques heures, voire une journée entière. Dans ce cas, l'échantillonnage par méthodes de MCMC devient excessivement coûteux et finalement impossible à réaliser.

Pour contourner ce problème, il existe des méthodes de réduction de modèle reposant sur un apprentissage à partir de simulations du modèle numérique original. Ces méthodes construisent une reproduction du modèle initial $\mathcal{G}$ à partir d'une base d'apprentissage, appelée plan d'expérience ou *Design Of Experiment*—DOE—en anglais. Le nouveau modèle obtenu par ce processus, appelé par la suite métamodèle, a pour objectif de diminuer le temps de calcul tout en fournissant une bonne approximation du modèle initial $\mathcal{G}$. Le processus pour obtenir un métamodèle est illustré par la figure 1.2.

1.2.1 Plan d'expérience par hypercube latin

La construction d'un métamodèle demande tout d'abord un plan d'expérience (DOE). Ce dernier sert à définir une série de jeux d'entrée du modèle $\mathcal{G}$ à partir de laquelle des simulations

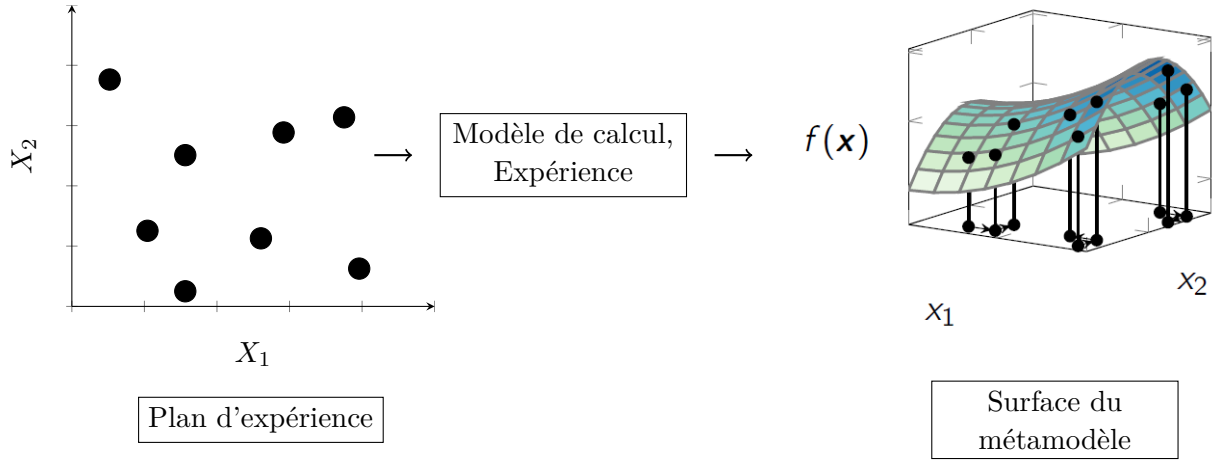


FIGURE 1.2 – Étapes de la métamodélisation

avec ce modèle seront effectuées. Les informations apportées par ces résultats seront ensuite exploitées par l'étape d'interpolation. Il est important de bien définir le DOE afin de récupérer le maximum d'information sur le modèle initial $\mathcal{G}$. Il existe des algorithmes qui permettent de répartir uniformément les points du DOE selon toutes les dimensions du domaine étudié pour $\mathbf{x}$.

La méthode d'échantillonnage par hypercube latin a été utilisée dans de nombreux domaines informatiques depuis 1975, suivant [McKay 1979]. Cet algorithme permet de générer un ensemble de jeux de données dans l'hypercube $[0, 1]^p$. Cet ensemble peut être ensuite projeté sur un autre domaine si besoin. Le N_s -échantillonnage par l'hypercube latin sélectionne N_s valeurs distinctes pour chacune des p variables $x_1, x_2, \dots, x_p$. L'intervalle de variation de chaque entrée x_i est découpé en N_s segments disjoints. La largeur de chaque segment dépend de la distribution *a priori* de chaque composante x_i . Plus concrètement, si x_i suit une loi normale d'espérance μ , les segments découpés seront plus larges aux extrémités de la plage de variation et plus étroits autour de la moyenne μ . Cette manière de répartir permet d'obtenir un échantillon de x_i en respectant la loi *a priori* de cette variable. L'étape suivante consiste à choisir aléatoirement une valeur de x_i dans chaque sous-intervalle. Effectivement, les N_s valeurs obtenues de x_1 sont couplées aléatoirement avec les N_s valeurs de x_2 . Ces N_s paires sont ensuite combinées avec les N_s valeurs de x_3 pour former N_s triplets, etc. jusqu'à ce que N_s p -uplets soient formés. Cette construction garantit que l'on n'a qu'une valeur dans chaque sous-intervalle de chaque direction de l'espace. Cette méthode sera résumée à nouveau dans le chapitre 3.

L'objectif de l'échantillonnage par hypercube latin est d'explorer le plus possible l'espace des entrées $\mathbf{x}$. La construction des échantillons suivant cette méthode permet de ne pas avoir deux échantillons dans le même intervalle de valeurs. Cependant, les plans d'expérience par hypercube latin (Latin Hypercube Sampling–LHS–en anglais) ne sont pas uniques. Pour certains plans, même s'ils respectent les règles de construction de l'hypercube latin, l'exploration du domaine n'est pas tout à fait optimale. Nous illustrons avec deux exemples dans l'annexe C.1.

Ainsi, pour pallier ce problème, des plans LHS optimisés ont été mis en place. Appelons

$P = (\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(N_s)})$ un LHS de N_s échantillons ; chaque échantillon $\mathbf{x}^{(i)}$ est un vecteur de l'espace $\llbracket 0, 1 \rrbracket^p$. Nous notons également $\mathcal{P}(N_s, \llbracket 0, 1 \rrbracket^p)$ l'espace des plans LHS de N_s échantillons dans $\llbracket 0, 1 \rrbracket^p$, alors $P \in \mathcal{P}(N_s, \llbracket 0, 1 \rrbracket^p)$. Plusieurs critères d'optimisation existent. Nous pouvons citer les deux plus connus qui se fondent sur la distance entre les points :

- *Maximin* : Notons $d_{ij} = \|\mathbf{x}^{(i)} - \mathbf{x}^{(j)}\|_{L^q}$ la distance entre le i ème et j ème point du plan LHS P ; avec $\|\cdot\|_{L^q}$ la norme de l'espace L^q . L'optimisation du LHS applique souvent la norme euclidienne, où $q = 2$. Un plan maximin est un plan qui maximise la distance minimale entre deux points distincts du plan. Le critère maximin peut s'écrire comme suit

$$P^* = \max_{P \in \mathcal{P}(N_s, \llbracket 0, 1 \rrbracket^p)} \left(\min_{i \neq j, \forall (i,j) \in \llbracket 1, N_s \rrbracket^2} d_{ij} \right). \quad (1.5)$$

- *Minimax* : Un plan minimax est un plan qui minimise la distance maximale entre un point du domaine hors de l'échantillon, noté α , et le point de l'échantillon le plus proche. Le critère minimax s'écrit de façon suivante

$$\min_{P \in \mathcal{P}} \left[\max_{\alpha \in D} \left(\min_{i \in \llbracket 1, N_s \rrbracket} d(\mathbf{x}^{(i)}, \alpha) \right) \right]$$

où $D \subset \llbracket 0, 1 \rrbracket^p$ est le domaine de variation de $\mathbf{x}$.

En grande dimension, c'est-à-dire p élevé, l'optimisation suivant les critères *minimax* et *maximin* devient de moins en moins performante. L'article [Damblin 2013] a présenté un autre critère qui utilise des suites à faible discrédance. Il existe plusieurs définitions de la discrédance, elle peut notamment s'écrire

$$D^*(P) = \sup_{\alpha \in \llbracket 0, 1 \rrbracket^p} \left| \frac{1}{N_s} \sum_{i=1}^{N_s} \mathbb{1}_{(\mathbf{x}^{(i)} \in [0, \alpha])} - \prod_{j=1}^p \alpha_j \right|.$$

Les suites à faible discrédance sont des suites quasi-aléatoires ayant une faible valeur de discrédance. Il existe plusieurs familles de suites à discrédance faible comme les suites de Sobol, suites de Halton, etc. Un échantillon formé par une suite à faible discrédance minimise la déviation de la répartition dans l'espace de cet échantillon et celle d'un échantillon de même taille tiré par la loi uniforme.

1.2.2 Construction du métamodèle

Rappelons les notations suivantes :

- $X_{DOE} = (\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(N_s)}) \in \mathcal{P}(N_s, \llbracket 0, 1 \rrbracket^p) \subset \mathbb{R}^{N_s \times p}$ le plan d'expérience de N_s points en dimension p ;
- $\mathbf{y}_{DOE} = (y_{DOE}^1, \dots, y_{DOE}^{N_s}) = \mathcal{G}(X_{DOE})$ le vecteur des résultats de simulation du modèle $\mathcal{G}$ sur le plan X_{DOE} .

Le développement de cette section sera repris de manière plus concise dans le chapitre 3 et en version anglaise.

L'idée essentielle du métamodèle est de créer une surface d'interpolation à partir des réponses $\mathbf{y}_{DOE}$ du modèle initial $\mathcal{G}$. Dans certaines approches, la métamodélisation d'un modèle numérique en point $\mathbf{x}$ peut se séparer en deux parties : la tendance linéaire, notée $f(\mathbf{x})$, et

l'interpolation, notée $z(\mathbf{x})$. Nous appelons $\hat{\mathcal{G}}(\mathbf{x})$ la prévision du métamodèle de $\mathcal{G}$ au point $\mathbf{x}$ dans l'espace des entrées. Il peut s'exprimer

$$\hat{\mathcal{G}}(\mathbf{x}) = f(\mathbf{x}) + z(\mathbf{x}) \quad (1.6)$$

On note $Y = f(\mathbf{X})$ la variable aléatoire réelle à expliquer et $\mathbf{X}$ un vecteur explicatif déterministe. Cette tendance linéaire revient à supposer qu'en moyenne, $\mathbb{E}(Y)$ est une fonction affine de $\mathbf{X}$ et s'écrit sous la forme

$$f(\mathbf{X}) = \beta_0 + \sum_{k=1}^p \beta_k X_k. \quad (1.7)$$

La méthode des moindres carrés ordinaires permet de minimiser l'écart entre la tendance $f(X_{DOE})$ et les réponses du plan d'expérience y_{DOE} dont la solution est :

$$\hat{\beta} = \left(X_{DOE}^T X_{DOE} \right)^{-1} X_{DOE}^T \mathbf{y}_{DOE}.$$

1.2.3 Interpolation par fonctions de base radiale—RBF

La construction du métamodèle par RBF procède en deux étapes : estimation de la tendance linéaire 1.7, puis estimation de $z(\mathbf{x})$. Après avoir obtenu l'estimateur des moindres carrés pour la partie linéaire, la déduction des valeurs de la fonction d'interpolation z sur le plan d'expérience X_{DOE} , présentée dans l'équation 1.6, s'écrit $z(X_{DOE}) = \mathbf{y}_{DOE} - f(X_{DOE})$. Nous notons également $z(X_{DOE}) = z_{DOE}$. Pour estimer cette partie du métamodèle, la première méthode appliquée dans cette étude est l'interpolation par fonctions de base radiale. Cette interpolation implique le passage de la surface de réponse par chaque valeur du plan d'expérience. Il existe plusieurs fonctions de base radiale, par exemple :

- Multiquadratique : $\Phi(r) = \sqrt{1 + \|r\|_2^2}$;
- Inverse : $\frac{1}{\sqrt{1 + \|r\|_2^2}}$;
- Gaussienne : $\exp(-\|r\|_2^2)$;
- Spline de plaque fine : $\|r\|_2^2 \cdot \log(\|r\|)$,

où $\|\cdot\|_2$ est la norme dans l'espace L^2 .

La prédiction du modèle d'interpolation par RBF $\hat{z}$ au point $\mathbf{x} \in \mathbb{R}^p$ s'écrit comme suit

$$\hat{z}(\mathbf{x}) = \sum_{i=1}^{N_s} \alpha_i \Phi(\mathbf{x} - \mathbf{x}^{(i)})$$

où $(\mathbf{x}^{(i)})_i$ sont des échantillons du DOE P et $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_{N_s}) \in \mathbb{R}^{N_s}$ est le vecteur de pondération à déterminer. Les composantes du vecteur $\boldsymbol{\alpha}$ doivent respecter les conditions suivantes

- $\forall i \in \llbracket 1, N_s \rrbracket, \quad \alpha_i \in [0, 1]$;
- $\sum_{i=1}^{N_s} \alpha_i = 1$.

En effet, le principe de l'interpolation par RBF consiste en la pondération de la distance entre le point à estimer $\mathbf{x}$ et les points dans le DOE X_{DOE} . Autrement dit, plus $\mathbf{x}$ se trouve proche d'un point de l'échantillon $\mathbf{x}^{(i)}$ du DOE X_{DOE} , plus α_i s'approche de 1 et les autres composantes

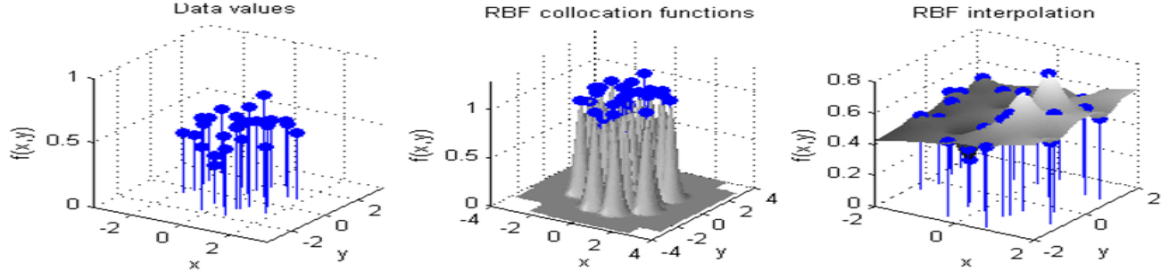


FIGURE 1.3 – Illustration de trois étapes principales de l'interpolation par RBF dans la présentation [Yue 2011]. La première figure correspond à la construction du DOE. Celle au milieu montre l'association des fonctions de base radiale à chaque point du DOE. Enfin, l'interpolation par RBF crée une surface d'apprentissage par méthode de pondération.

$\alpha_{j \neq i}$ s'annulent. D'après cette propriété, l'application du modèle interpolé $\hat{z}$ sur les points du DOE P s'écrit comme suit

$$\forall j \in \llbracket 1, N_s \rrbracket, \quad \hat{z}(\mathbf{x}^{(j)}) = z_{DOE,j} = \sum_{i=1}^{N_s} \alpha_i \Phi(\mathbf{x}^{(j)} - \mathbf{x}^{(i)}) \quad (1.8)$$

ou sous forme matricielle

$$\begin{aligned} \mathbf{z}_{DOE} = \begin{pmatrix} z_{DOE,1} \\ z_{DOE,2} \\ \vdots \\ z_{DOE,N_s} \end{pmatrix} &= \begin{pmatrix} \phi_{11} & \phi_{12} & \cdots & \phi_{1 N_s} \\ \phi_{21} & \phi_{22} & \cdots & \phi_{2 N_s} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{N_s 1} & \phi_{N_s 2} & \cdots & \phi_{N_s N_s} \end{pmatrix} \cdot \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_{N_s} \end{pmatrix} \\ &= \Phi \cdot \boldsymbol{\alpha} \end{aligned} \quad (1.9)$$

où $\phi_{ij} = \Phi(\mathbf{x}^{(i)} - \mathbf{x}^{(j)})$. L'estimation de $\boldsymbol{\alpha}$ revient à résoudre un système linéaire de rang N_s . La figure 1.3 permet de visualiser les différentes étapes de l'interpolation par RBF.

En conclusion, la prédiction du métamodèle $\hat{G}$ au point $\mathbf{x} \in \mathbb{R}^p$ devient

$$\hat{G}(\mathbf{x}) = \beta_0 + \hat{\beta} \mathbf{x} + \Phi(\mathbf{x}) \hat{\boldsymbol{\alpha}}$$

où $\hat{\boldsymbol{\alpha}}$ est la solution du système 1.9. L'avantage de l'interpolation par RBF est que le temps de construction d'un métamodèle est assez court car la solution analytique du problème des moindres carrés est simple. La deuxième méthode que nous allons présenter (processus gaussien ou krigeage) donne des prédictions de meilleure qualité mais elle nécessite un temps de construction beaucoup moins avantageux.

1.2.4 Processus gaussien, krigeage

L'application de cette méthode nécessite toujours le plan d'expérience X_{DOE} et les simulations du modèle initial $\mathbf{y}_{DOE}$. Suivant la théorie du krigeage, la sortie déterministe $\mathbf{y}_{DOE}$ est considérée comme la réalisation d'un champ aléatoire, noté $Y(\mathbf{x})$

$$Y(\mathbf{x}) = \mathcal{F}(\mathbf{x}) + \mathcal{Z}(\mathbf{x})$$

où $\mathcal{F}(\mathbf{x})$ est la fonction moyenne et $\mathcal{Z}(\mathbf{x})$ est un processus gaussien centré. En réalité, il existe plusieurs choix pour la fonction moyenne $\mathcal{F}$ comme :

- Fonction constante : $\mathcal{F}(\mathbf{x}) = c$;
- Polynôme de degré 1 : $\mathcal{F}(\mathbf{x}) = \beta_0 + \sum_{k=1}^p \beta_k x_k$;
- Polynôme de degré supérieur ;
- Modèle additif : $\mathcal{F}(\mathbf{x}) = \beta_0 + \sum_{k=1}^p \beta_k h_k(x_k) + \sum_{k,l=1, l \neq k}^p \beta_{kl} h_{kl}(x_k, x_l)$;
- etc.

En comparant cette explication avec le développement 1.6, le métamodèle $\hat{G}$ par méthode de krigeage correspond à un processus gaussien dont la moyenne est une tendance linéaire. La partie d'interpolation z suivra évidemment un processus gaussien centré de variance σ^2 et de fonction de covariance R

$$\text{cov}(z(\mathbf{x}), z(\mathbf{x}')) = R(\mathbf{x}, \mathbf{x}') = \sigma^2 r(\mathbf{x} - \mathbf{x}').$$

Suivant les propriétés d'un processus gaussien, la loi jointe des échantillons du DOE suit la loi normale

$$\begin{aligned} Y_{DOE} &\sim \mathcal{N}(\mathcal{F}_{DOE}, \Sigma_{DOE}) \\ \mathcal{F}_{DOE} &= \left(f(\mathbf{x}^{(1)}), \dots, f(\mathbf{x}^{(N_s)}) \right)^T \\ (\Sigma_{DOE})_{ij} &= R(\mathbf{x}^{(i)}, \mathbf{x}^{(j)}) = \sigma^2 r(\mathbf{x}^{(i)} - \mathbf{x}^{(j)}). \end{aligned}$$

Appelons $\mathbf{x}^*$ le nouveau point de prédiction non inclus dans P . La loi jointe des échantillons du DOE et $\mathbf{x}^*$ s'écrit comme suit

$$[Y_{DOE}, Y^*] \sim \mathcal{N} \left(\begin{pmatrix} \mathcal{F}_{DOE} \\ f^* \end{pmatrix}, \begin{pmatrix} \Sigma_{DOE} & \mathbf{R}^* \\ \mathbf{R}^{*T} & \sigma^2 \end{pmatrix} \right)$$

où Y^* est la variable aléatoire représentant la valeur de $\hat{G}(\mathbf{x}^*)$, $f^* = f(\mathbf{x}^*) = \beta_0 + \sum_{k=1}^p \beta_k x_k^*$ et $\mathbf{R}^* = \sigma^2 \left(r(\mathbf{x}^{(1)} - \mathbf{x}^*), \dots, r(\mathbf{x}^{(N_s)} - \mathbf{x}^*) \right)^T$. La prédiction au point $\mathbf{x}^*$ par modèle de krigeage suit la loi conditionnelle $Y^*|Y^P$ qui, par conséquence des processus gaussiens, est une loi normale

$$Y^*|Y_{DOE} \sim \mathcal{N}(\mu(\mathbf{x}^*), \tilde{\sigma}^2(\mathbf{x}^*)) \quad (1.10)$$

$$\begin{aligned} \mu(\mathbf{x}^*) &= E(Y^*|Y_{DOE}) = f(\mathbf{x}^*) + \mathbf{R}^{*T} (\Sigma_{DOE})^{-1} Z_{DOE} \\ \tilde{\sigma}^2(\mathbf{x}^*) &= \sigma^2 - \mathbf{R}^{*T} (\Sigma_{DOE})^{-1} \mathbf{R}^* \end{aligned} \quad (1.11)$$

avec $Z_{DOE} = \mathcal{Z}(X_{DOE})$.

La figure 1.4 représente en dimension 1 la prédiction du modèle krigeage (ligne rouge continue) du modèle initial (ligne pointillée), ainsi que l'erreur de prédiction (zone grise). En conclusion, le prédicteur du métamodèle par krigeage est donné par l'expression 1.10 ; c'est-à-dire la moyenne du processus gaussien. D'autre part, dans la loi gaussienne, une moyenne s'accompagne toujours d'une variance ou un écart type, qui représente l'erreur moyenne autour de son espérance. C'est pourquoi la variance du prédicteur 1.11 peut être considérée comme une estimation de l'erreur du prédicteur. Cela implique un avantage potentiel de cette méthode : le métamodèle

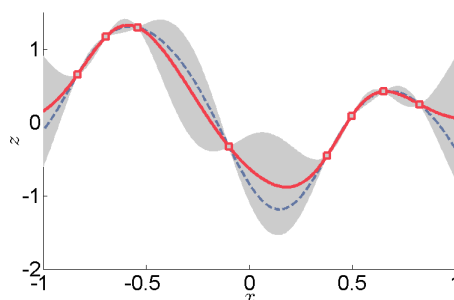


FIGURE 1.4 – Illustration d'un métamodèle par krigeage. La ligne rouge est le modèle initial émulé ; les points correspondent à la base d'apprentissage ; la ligne bleue en discontinu représente la prévision du krigeage et la bande grise est l'écart type des possibilités prévues par la méthode de krigeage.

prédit la valeur d'un point et potentiellement l'écart type de cette prédiction. Les métamodèles construits dans la section 3 ont pour noyau de corrélation "Matern 5/2" et l'estimation des hyper-paramètres de krigeage est faite via la méthode du maximum de vraisemblance.

Les deux méthodes d'interpolation, RBF et krigeage, seront utilisées durant cette thèse. L'étape de validation et la comparaison entre ces deux méthodes seront abordées dans le chapitre 3. Jusqu'ici, nous nous sommes concentrés sur les méthodes mathématiques du traitement d'incertitude de manière générale. La suite de ce chapitre présentera l'application principale de cette étude : la modélisation numérique de la dispersion des éléments radiologiques dans l'atmosphère.

1.3 Modèle de dispersion atmosphérique

La diffusion des particules dans l'air dépend de plusieurs facteurs :

- Les conditions de rejet (nature du nuage de produit, type d'émission) ;
- les conditions météorologiques (champ de vent, pluie et température, etc.) ;
- l'environnement (végétation, nature du sol, topographie, etc.).

1.3.1 Caractéristiques de l'atmosphère

Structure verticale de l'atmosphère terrestre

L'atmosphère est la couche gazeuse qui entoure la Terre. Sa structure est assez complexe, que ce soit en composition chimique, en stratification (température) ou suivant l'altitude. Si on utilise la température comme paramètre de stratification, l'atmosphère se divise verticalement (de plus forte à plus faible altitude) en cinq couches principales : l'exosphère, la thermosphère, la mésosphère, la stratosphère et la troposphère, suivant [Prieur 2014]. La figure 1.5 illustre l'organisation des couches atmosphériques.

- La thermosphère se situe en dessous de l'exosphère, entre 85 km et 600 km en altitude. Elle est caractérisée par une forte croissance de température avec l'altitude.

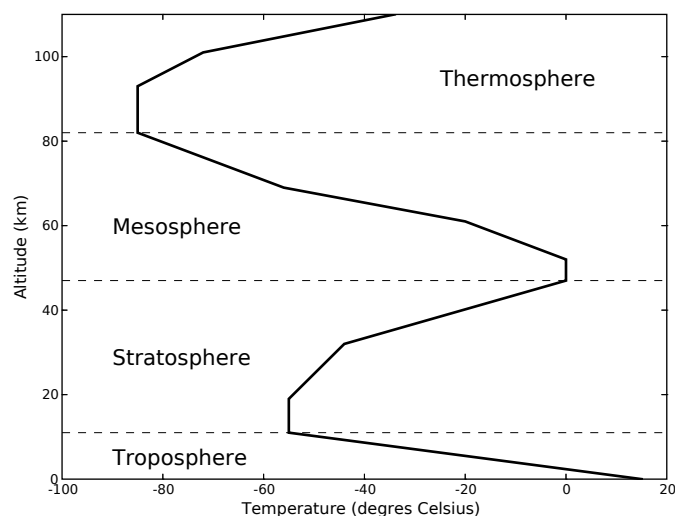


FIGURE 1.5 – Schéma de la structure verticale atmosphérique

- La mésosphère se situe entre 50 km et 90 km et est caractérisée par une décroissance de température avec l'altitude. C'est à sa limite supérieure qu'on trouve la température la plus basse de l'atmosphère terrestre : elle peut atteindre 200°K, soit -73°C . C'est une zone de transition entre la Terre et l'espace.
- La stratosphère est l'avant-dernière strate de l'atmosphère, entre 10 km et 50 km d'altitude. La température de l'air y est relativement constante jusqu'à 20 km et ensuite se met à augmenter régulièrement de 1°C jusqu'à 32 km (moyenne stratosphère) puis de 2.8°C par km pour atteindre la température de -3°C à 50 km (stratopause). À cause de l'augmentation de température avec l'altitude, il n'y a pas de convection, ce qui stabilise les conditions atmosphériques dans cette couche.
- La troposphère est la couche la plus basse de l'atmosphère terrestre. L'épaisseur de cette enveloppe varie entre 8 km dans les hautes latitudes jusqu'à 18 km au-dessus de l'équateur. Elle varie aussi en fonction des saisons (maximale en été et minimale en hiver). La troposphère représente la plus grande partie de la masse atmosphérique. C'est une zone dense et turbulente, dans laquelle se déroule la plupart des phénomènes météorologiques. Sa densité est caractéristique, tout comme son gradient thermique de $-6.5^{\circ}\text{C}/\text{km}$.

Mouvement de l'atmosphère

Cette enveloppe gazeuse n'est pas immobile ; elle circule toujours et présente des caractéristiques extrêmement variables suivant les échelles spatiales et temporelles considérées (petite brise/vent orageux, échelle locale/planétaire, etc.). Ces mouvements d'air se produisent à cause du chauffage différentiel de la surface terrestre. Le système Terre-atmosphère reçoit de l'énergie du rayonnement solaire. Quand les rayons solaires arrivent à la surface externe de l'atmosphère, 30% d'énergie des ondes électromagnétiques est réfléchi et repart dans l'espace. Le reste est

absorbé par le système et provoque le réchauffement de la surface terrestre. Le rayonnement solaire incident au sol atteint le maximum dans la région équatoriale et le minimum vers les pôles. Par conséquent, ce chauffage différentiel pôle/équateur crée un flux de chaleur qui est responsable des grands mouvements des masses d'air de la Terre. Sans ce flux d'énergie des zones tropicales vers les pôles, il y aurait aux pôles des températures beaucoup plus basses, et dans la région équatoriale des températures beaucoup plus élevées, que celles effectivement rencontrées.

La couche limite atmosphérique

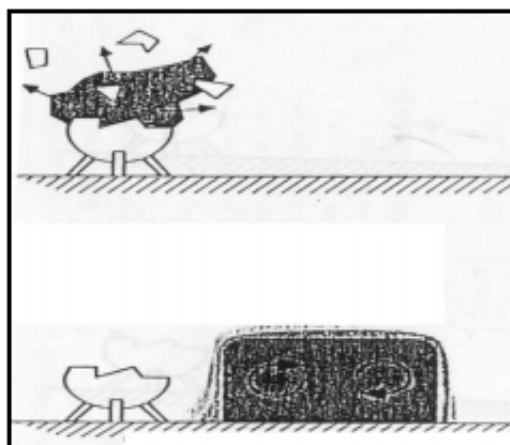
La dispersion atmosphérique se base sur les conditions météorologiques dans les basses couches de la troposphère. Les polluants émis sont plus ou moins rapidement dispersés selon deux phénomènes majeurs :

- le transport horizontal par le champ de vent, qui explique le déplacement des nuages de polluants sur de longue distance, suivant différentes directions ;
- la turbulence verticale liée aux effets de turbulence dans la couche limite atmosphérique (CLA).

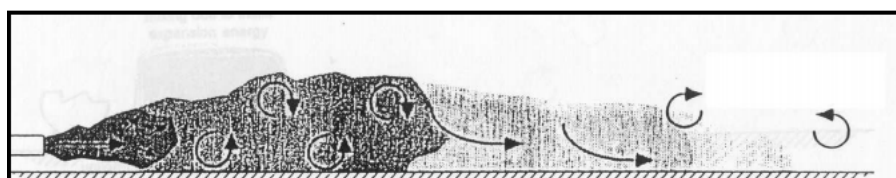
Le CLA est la partie de la troposphère directement influencée par la présence de la surface terrestre, caractérisée en général par une baisse de température en fonction de l'altitude (-0.65°C sur 100 m dans une atmosphère neutre. Les effets de la turbulence dans la couche limite ont principalement une double origine : dynamique avec le cisaillement vertical de vent induit par la condition d'adhérence au sol (vent nul), et thermique avec le gradient vertical de température (le sol est chauffé au cours de la journée puis refroidit pendant la nuit).

Un des paramètres qui influencent le plus le mélange de polluants est la hauteur de la couche limite, aussi appelée hauteur de mélange car elle fixe un volume de dilution pour la masse de polluants émis. En effet, le mélange de polluants dans la CLA se produit très rapidement sur la verticale (dans un temps caractéristique d'environ une heure), tandis que la hauteur de couche limite agit souvent comme un « couvercle » limitant la diffusion des polluants vers la troposphère « libre » (où l'écoulement, généralement laminaire, n'est plus directement influencé par le sol). Nous distinguons trois états principaux de stabilité de l'atmosphère :

- la CLA neutre ;
- la CLA instable, dans laquelle les polluants sont fortement brassés par la turbulence d'origine thermique. Sous l'effet du rayonnement solaire, la surface se chauffe plus rapidement que l'air. La chaleur se transfère de la Terre dans l'air, diminue la masse volumique de ce dernier. L'air chaud monte, rencontre un air plus froid en altitude, ce qui amplifie son mouvement ascendant. L'air froid, lui, redescend au niveau du sol. Cela crée un mélange convectif dans l'atmosphère, favorable à une diffusion verticale rapide des polluants. Cet état de stabilité est associé à une hauteur de CLA assez grande (de l'ordre de 1000 à 3000 mètres).
- la CLA stable (souvent la nuit) où les polluants restent accumulés au niveau du sol. La nuit, la Terre ne reçoit plus l'énergie thermique du soleil. Le sol refroidit plus rapidement que l'air, et la couche d'air au niveau du sol reste plus froide que les couches supérieures, ce qui a pour effet de freiner les échanges entre les couches. La turbulence atmosphérique, de nuit, est uniquement d'origine mécanique, et beaucoup plus faible que la turbulence



(a) Rejet instantané d'un produit



(b) Rejet continu d'un produit

FIGURE 1.6 – Deux types d'émissions de polluants dans l'atmosphère, source [INERIS 2002]

thermique qui a lieu dans la journée. La hauteur de couche limite stable est beaucoup plus faible (quelques centaines, voire seulement dizaines de mètres). Au-dessus de la couche limite stable se trouve la couche résiduelle, qui contient les polluants qui ont été mélangés dans la journée.

1.3.2 Dispersion de polluants

La dispersion atmosphérique comporte quatre processus qu'il importe de caractériser au mieux pour la simuler correctement :

- Émission : les rejets de polluants dans l'atmosphère peuvent se dérouler de façon instantanée ou continue, par une source d'émission ponctuelle ou diffuse. Les rejets instantanés sont caractérisés par une libération ponctuelle d'une quantité de polluants, appelée « bouffée » (cf. figure 1.6a). Une émission continue est caractérisée par la durée d'émission et le débit du rejet, et se traduit par un « panache » (cf. figure 1.6b). Dans le cadre d'un rejet par une cheminée, la source peut être considérée en première approximation comme ponctuelle (cette hypothèse est acceptable dès que la distance à la source est suffisamment grande). Les caractéristiques thermodynamiques du rejet (vitesse d'éjection et température) peuvent être prises en compte, mais sont rapidement « oubliées » par le nuage de polluants dès qu'il est suffisamment mélangé à l'air ambiant.
- Transport dans l'atmosphère : dans cette étape du processus, les deux facteurs qui conduisent à la dispersion du panache sont l'advection effectuée par le vent horizontal et

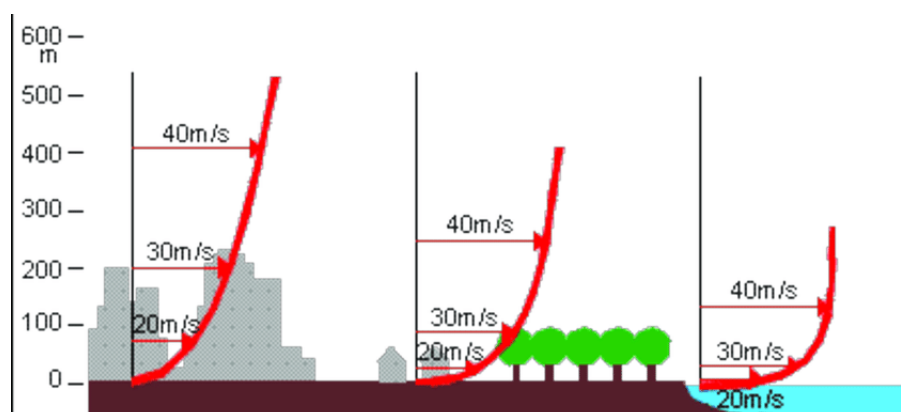


FIGURE 1.7 – Profil vertical du vent suivant la rugosité au sol, source [Dubois 2014]

la diffusion verticale par turbulence des masses d'air.

- Advection : au sein de la CLA, le vent vertical est quasiment nul (il excède rarement les 1 cm/s). La direction et la vitesse de transport du panache sont données par les vents horizontaux. Du fait du frottement de l'air sur le sol, le profil de vent adopte un caractère logarithmique, plus ou moins perturbé par les processus thermiques, cf. figure 1.7.
- Turbulence : les deux origines de la turbulence, thermique et mécanique, ont été décrites plus haut. La turbulence conduit à la diffusion (étalement) du panache dans toutes les directions.
- Dépôt au sol : durant le transport atmosphérique, une partie des polluants est déposée au sol sous certaines conditions de dépôt sec ou humide. Les polluants peuvent se déposer sur les surfaces bâties, la végétation, les eaux de surface, etc. Les processus fondamentaux du dépôt sec sont la sédimentation gravitationnelle, les impacts par inertie ou interception avec des obstacles, et la diffusion. Les précipitations (y compris pluie, neige, brouillard, etc.) sont responsables du dépôt humide. En présence de précipitations, le dépôt humide devient le processus prépondérant de rabattement des polluants au sol et sur les surfaces, notamment du fait que le nuage est lessivé sur toute la colonne verticale, alors que le dépôt sec concerne les molécules de polluants les plus proches du sol ou des obstacles.
- Réactions chimiques ou radiologiques : les polluants peuvent réagir chimiquement avec les autres substances présentes dans l'atmosphère, pour créer d'autres produits. Dans le cas de radionucléides, la décroissance radioactive fait disparaître les isotopes avec un temps caractéristique correspondant à leur demi-vie, tandis que des isotopes « fils » sont créés par filiation radioactive, figure 1.8.

1.3.3 Modélisation mathématique du phénomène de transport atmosphérique

Les modèles de dispersion atmosphérique sont des descriptions mathématiques du transport et des réactions chimiques des polluants (ou radiologiques dans le cas des radionucléides). Ils s'agit généralement de couplage entre des équations aux dérivées partielles (EDP) dont les

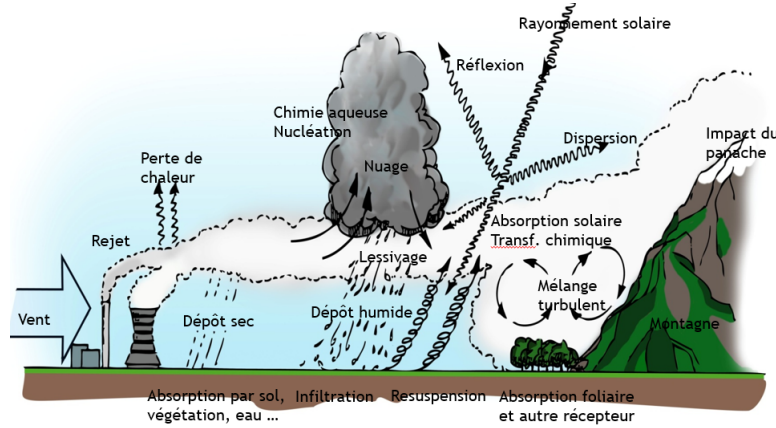


FIGURE 1.8 – Processus de dispersion des polluants dans l’air, source IRSN/SESUC

inconnues sont des concentrations de polluants.

Nous considérons l’équation d’advection-diffusion qui régit le transport d’un isotope dans un écoulement atmosphérique. La présentation de cette équation est indispensable non seulement pour comprendre les phénomènes physiques mais également pour identifier l’origine des erreurs rencontrées dans la modélisation de la dispersion. Dans les modèles de dispersion, l’advection par le vent et la diffusion turbulente sont pris en compte, ainsi que le terme source (émission) et les termes « puits » (dépôts sec et humide). Le principe de conservation appliqué à un radioisotope r permet d’écrire l’équation d’advection-diffusion de la concentration instantanée c comme suit :

$$\frac{\partial c_r}{\partial t} + \text{div}(\mathbf{w}c_r) = \text{div}\left(\rho \mathbf{K} \nabla \frac{c_r}{\rho}\right) - \mathbf{F} \mathbf{c} + E_r - \Lambda c_r \quad (1.12)$$

où c_r est la concentration dans l’air de l’espèce r ; $\mathbf{c}$ le vecteur de concentration de tous les radionucléides reliés par la matrice de filiation radioactive $\mathbf{F}$; $\mathbf{w} = (w_u, w_v, w_z)^T$ le vecteur de vitesse du vent en dimension trois ; ρ la densité de l’air. La matrice de diffusion par turbulence $\mathbf{K}$ est supposée être diagonale avec les coefficients diagonaux (K_u, K_v, K_z) . E_r est le terme d’émission, dû à la présence de sources qui rejettent l’espèce dans l’atmosphère, et enfin Λ est le coefficient de lessivage pour la perte par dépôt humide en cas de précipitation. La condition aux limites au niveau du sol s’écrit comme $\rho \mathbf{K} \nabla \frac{c_r}{\rho} \cdot \mathbf{n} = v_d c_r$, où $\mathbf{n}$ est le vecteur normal à la surface terrestre qui est orienté dans le sens des altitudes croissantes, et v_d est la vitesse de dépôt sec.

1.3.4 Résolution numérique de l’équation advection-diffusion

La résolution de l’équation d’advection-diffusion requiert en général, soit des hypothèses simplificatrices pour pouvoir la résoudre de façon analytique, soit de faire appel à des techniques numériques de discrétisation. Plusieurs types d’approche existante permettent de simuler la dispersion atmosphérique des polluants. Ils diffèrent l’un de l’autre par l’échelle de simulation (de quelques mètres jusqu’à l’échelle régionale ou continentale). Pour présenter ces approches, nous les avons classées en quatre catégories : modèle gaussien, modèle lagrangien, modèle eulérien et

modèle de mécanique des fluides. Chacun de ces modèles sera détaillé dans les paragraphes qui suivent.

Modèle à bouffées gaussiennes

Dans un modèle à bouffées, l'approche considère le rejet de polluants comme un enchaînement de rejets instantanés. Chacun de ces rejets est modélisé par une bouffée gaussienne qui diffuse au cours du temps. La distribution à l'instant t_i , qui correspond au i -ème rejet, de concentration induite par ce rejet instantané (à un instant t_0) et ponctuel (au point de coordonnées (x_0, y_0, z_0)) d'une masse M_i de polluants, est fournie par l'équation 1.13.

$$\bar{c}(x, y, z, t) = \sum_{i=1}^n \frac{M_i}{[2\pi]^{3/2} \sigma_{x_i} \sigma_{y_i} \sigma_{z_i}} \exp \left(-\frac{[x - x_0 - u(t - t_i)]^2}{2\sigma_{x_i}^2} - \frac{(y - y_0)^2}{2\sigma_{y_i}^2} \right) \left[\exp \left(-\frac{(z - z_0)^2}{2\sigma_{z_i}^2} \right) + \alpha \exp \left(-\frac{(z + z_0)^2}{2\sigma_{z_i}^2} \right) \right]$$

où :

- M_i est la masse du i ème rejet instantané $[kg]$;
- u est la vitesse du vent $[m/s]$;
- n est le nombre de rejets instantanés considérés ;
- t_i est instant de fin d'émission du i ème rejet $[s]$;
- t_{i-1} est instant de fin d'émission du $(i - 1)$ ème rejet et début d'émission du i ème rejet $[s]$;
- $\sigma_{x_i}, \sigma_{y_i}, \sigma_{z_i}$ sont les écarts types de la distribution gaussienne du i ème rejet instantané de masse M_i par rapport à sa localisation à l'instant t $[m]$;
- α est le coefficient de réflexion au sol.

Le centre de masse de chaque bouffée est transporté par le champ de vent, qui est supposé homogène au sein de chaque bouffée mais qui peut être hétérogène et évoluer en temps. À un instant t donné, la concentration en un point dans l'espace est obtenue en rassemblant la contribution de toutes les bouffées émises depuis le début du rejet, cf. figure 1.9.

Dans le cas d'un rejet continu à débit constant et d'un champ météorologique uniforme et homogène, cette équation peut encore se simplifier en supprimant le terme dépendant de x . La solution est alors un panache stationnaire (équation 1.13).

$$\bar{c}(x, y, z, t) = \frac{Q}{2\pi\sigma_y \sigma_z} \exp \left(-\frac{(y - y_0)^2}{1\sigma_y^2} \right) \left[\exp \left(-\frac{(z - z_0)^2}{2\sigma_z^2} \right) + \alpha \exp \left(-\frac{(z + z_0)^2}{2\sigma_z^2} \right) \right] \quad (1.13)$$

Cette équation peut être reliée à l'équation générale d'advection-diffusion 1.12, sous réserves des hypothèses simplificatrices suivantes :

1. La source est ponctuelle ;
2. Le rejet est continu ;
3. La météo est uniforme (en temps) et homogène (en espace) ;
4. La diffusion dans la direction du vent est négligeable par rapport à l'advection.

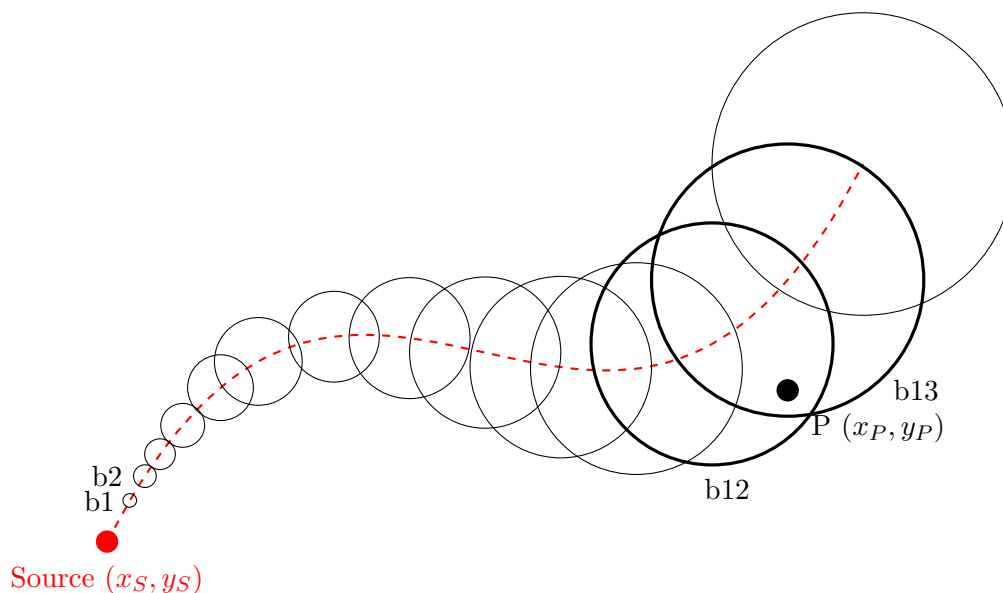


FIGURE 1.9 – Illustration de l’émission d’un modèle de dispersion par bouffées gaussiennes. Le point rouge représente la source de rejet, la ligne discontinue montre la trajectoire du centre de la bouffée. Les cercles illustrent les positions de différentes bouffées utilisées pour discrétiser le panache. La concentration du point P , qui se trouve dans la zone de fusion de deux bouffées b12 et b13, est la combinaison linéaire de ces deux bouffées.

Lorsque ces hypothèses sont vérifiées, l’équation 1.13 peut être retrouvée à partir de l’équation d’advection diffusion, avec la relation suivante entre l’écart type gaussien et le coefficient de diffusion turbulente : $\sigma_i(t) = \sqrt{2K_i(t - t_0)}$, avec $i \in \{u, x, z\}$. L’avantage le plus important des modèles gaussiens est le faible temps de calcul. C’est une caractéristique essentielle d’un outil opérationnel pour situation urgente. Le modèle gaussien de panache est le plus rapide. En revanche, il est limité à un écoulement uniforme et homogène. Le temps de calcul du modèle à bouffées gaussiennes dépend du nombre de bouffées considérées. Toutefois, il permet de traiter un rejet instantané ou continu variable en temps ainsi qu’un champ météorologique hétérogène et variable en temps. Sa limitation principale est l’hypothèse de météorologie homogène au sein de la bouffée, qui n’est plus valide à de trop grandes distances. Les modèles gaussiens sont adaptés à des rejets ponctuels (accidentels ou chroniques) de type source industrielle. Ils sont plus difficilement utilisables pour la modélisation de sources diffuses (pollution urbaine par exemple) car la multiplicité des sources augmente considérablement le temps de calcul.

Modèle lagrangien

Les modèles lagrangiens stochastiques sont basés sur l’observation du caractère aléatoire du mouvement de chaque particule à travers l’atmosphère (figure 1.10). Les particules se déplacent à la vitesse moyenne du vent et subissent en outre l’influence de la turbulence. L’effet de cette turbulence est modélisée en ajoutant une vitesse aléatoire additionnelle au mouvement moyen de chaque particule. Pour prévoir le comportement du nuage de polluant rejeté, il faut étudier les

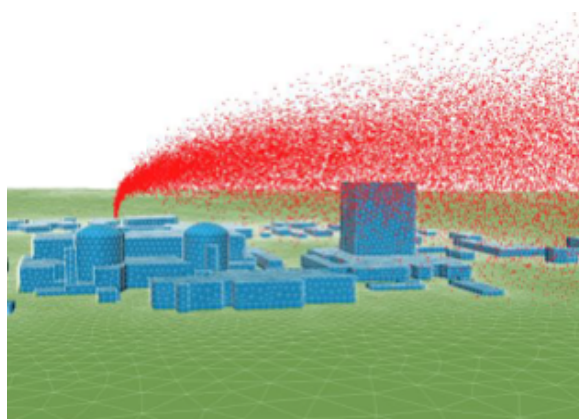


FIGURE 1.10 – Illustration du modèle lagrangien, source IRSN/ECL

trajectoires d'un grand nombre de particules passant par la source et en effectuer une moyenne. Les concentrations sont ensuite calculées aux points d'intérêt en comptant le nombre de particules présentes dans un volume de contrôle. Plus le point est éloigné, plus le nombre initial de particules rejetées devra être grand afin d'avoir toujours un nombre de particules suffisant dans le volume de contrôle pour avoir des statistiques significatives.

Le temps de calcul des modèles lagrangiens reste plus élevé que celui des modèles à bouffées gaussiennes, du fait du grand nombre de particules à considérer pour pouvoir obtenir des statistiques représentatives d'un grand nombre de réalisations. De la même façon que les modèles à bouffées gaussiennes, ils permettent de traiter des rejets instantanés ou continus, variables en temps, et pour n'importe quel type d'écoulement. Au contraire, dans le cas où plusieurs sources de rejet sont concernées, le temps de calcul sera très lourd car il est proportionnel au nombre de sources, contrairement à l'approche eulérienne du paragraphe 1.3.4. Ce type de modèle n'est en outre pas adapté pour modéliser des réactions chimiques.

Modèle eulérien

Un modèle eulérien résout l'équation 1.12 en utilisant des méthodes numériques de type différences finies, éléments finis ou volumes finis. Le principe est d'observer le mouvement des particules de polluants en se plaçant sur une grille de récepteurs fixes, souvent celle des données météorologiques.

La résolution d'une équation aux dérivées partielles, telle que l'équation d'advection-diffusion, commence par une association de cette équation à des conditions initiales en temps et des conditions aux limites du domaine spatial. Une fois qu'un maillage dans l'espace et une discrétisation en temps sont introduits, il est possible de remplacer le système précédent, composé de l'équation d'advection-diffusion avec des conditions initiales et aux limites, par un système discret pour lequel on cherche une solution sur les noeuds du maillage spatio-temporel.

Le modèle eulérien développé par l'IRSN, nommé *ldX*, est une version adaptée du modèle Polyphemus/Polair3D, utilisé dans la prédiction de la qualité de l'air. Ce modèle résout l'équation d'advection-diffusion 1.12 par le schéma numérique des volumes finis. Donc pour chaque maille de la discrétisation spatiale, il calcule la concentration moyenne dans toute la maille, cf.

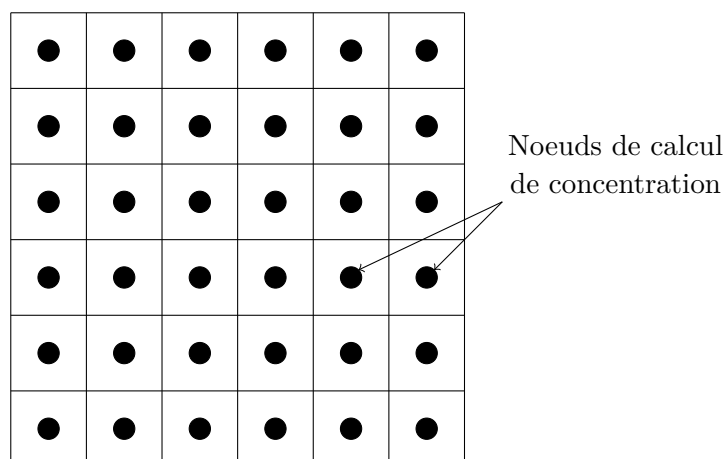


FIGURE 1.11 – Maillage de type volumes finis avec leurs centres où sont localisées les concentrations moyennes.

figure 1.11. À cause de cette caractéristique, nous appliquons ce modèle plutôt pour les estimations à grande échelle (régionale ou continentale). Quand le panache est proche de la source (à moins de quatre mailles), il n'est souvent pas assez dispersé pour que la concentration moyenne dans la maille soit représentative à la position du panache, voir annexe B.2.

Le modèle eulérien permet de tenir compte des champs météorologiques 3D et variables en temps et de la cinétique du rejet (figure 1.12). Contrairement au modèle lagrangien, le temps de calcul du modèle eulérien est pratiquement indépendant du nombre de sources d'émission mais fortement dépendant de la résolution du maillage. Il est plus facile d'intégrer des effets spécifiques dans le modèle eulérien comme des réactions physico-chimiques ou des dépôts. Cependant, ce type de modèle présente certains inconvénients. La principale limitation est que les phénomènes de taille caractéristique inférieure à celle de la maille ne sont pas représentés, puisque toutes les grandeurs (météorologiques ou liées au polluant) sont des moyennes sur les mailles. Ainsi, un maillage grossier introduit une forte erreur de dilution numérique du panache, ainsi que d'autres approximations (mauvaise prise en compte du relief ou des précipitations). De plus, l'équation d'advection-diffusion est résolue sur tout le maillage et non seulement aux points d'intérêt (points d'observation, par exemple). La qualité (exactitude) du résultat dépend donc fortement de la résolution du maillage et de la précision des données météorologiques.

1.4 Accident de la centrale nucléaire Fukushima-Daiichi

Le 11 mars 2011, un séisme de magnitude 9.0 s'est produit à 80 km au large de l'île d'Honshu au Japon [irsn.fr 2011]. Ce séisme a provoqué un tsunami qui a frappé le côte nord-est du Japon, ce qui a privé la centrale nucléaire de Fukushima-Daiichi de ses sources externes d'électricité et de ses moyens de refroidissement du coeur des réacteurs nucléaires. Quelques heures plus tard, de premiers rejets radioactifs étaient détectés et plus de 100 000 personnes évacuées.

La centrale exploitée par TEPCO comporte six réacteurs à eau bouillante (BWR–Boiling Water Reactor), de puissances respectives :

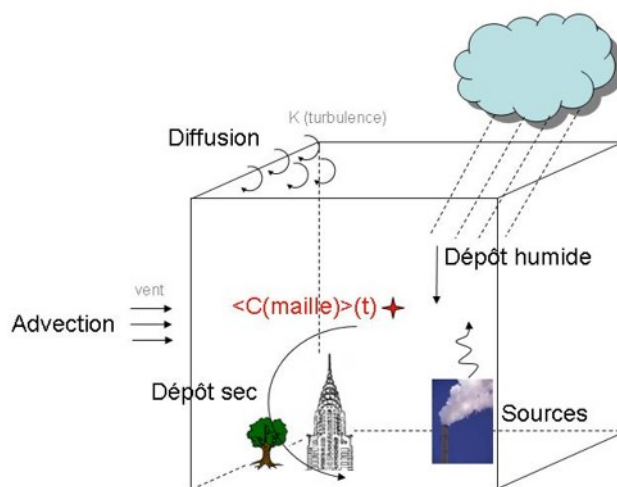


FIGURE 1.12 – Les interactions dans une maille du modèle eulérien, source IRSN/BMCA

- Réacteur 1 : 460 MWe
- Réacteur 2 à 5 : 784 MWe
- Réacteur 6 : 1100 MWe

Les réacteurs 1, 2 et 3 étaient en fonctionnement à pleine puissance au moment du séisme, les réacteurs 4, 5 et 6 étaient à l'arrêt pour maintenance. Le tremblement de terre a entraîné :

- Une perte des alimentations électriques externes des réacteurs et le démarrage des alimentations électriques internes (au site), c'est-à-dire des groupes électrogènes de secours à moteur diesel pour faire fonctionner des pompes de refroidissement ;
- l'arrêt automatique des réacteurs par insertion des grappes de commandes dans les coeurs, étouffant la réaction de fission par absorption des neutrons ;
- cinquante-cinq minutes plus tard, une vague de tsunami de 14 m de hauteur.

La vague a eu pour conséquence :

- Un endommagement des prises d'eau en mer conduisant à la perte de la source froide, nécessaire pour évacuer en permanence la chaleur résiduelle des réacteurs après l'arrêt d'urgence ;
- La perte de diesel de secours des réacteurs de 1 à 4.

À la suite de la perte des diesels, le refroidissement des coeurs a été temporairement assuré par des systèmes internes, qui ont été à leur tour perdus successivement sur les trois réacteurs, notamment par défaillance de batteries électriques. Les moyens de refroidissement de secours n'ont alors plus été opérationnels, ce qui a entraîné l'endommagement du combustible et la fusion des coeurs des réacteurs 1, 2 et 3. Durant trois semaines, des rejets successifs de radioactivité ont été détectés dans l'environnement, liés à différents événements sur l'installation. En particulier, des explosions liées à la formation d'hydrogène dans le bâtiment réacteur sont survenues successivement sur les trois réacteurs entre le 11 et le 15 mars. Jusqu'au 30 mars, différentes séquences de rejets dans l'atmosphère ont pu être reconstituées à partir de mesures de radioactivité dans l'environnement. Les rejets se sont poursuivis, dans des proportions beaucoup plus faibles, plusieurs semaines après.

Durant les trois semaines de rejets significatifs dans l’atmosphère, des modèles de dispersion atmosphérique ont été utilisés, au Japon comme à l’étranger, pour prévoir et anticiper les trajectoires possibles des panaches radioactifs. La plateforme WSPEEDI, développée par JAEA – Japanese Atomic Energy Agency, a fourni des prévisions de contamination au fur et à mesure de l’évolution de l’accident. Cependant, ces prévisions étaient tributaires de la qualité des données météorologiques disponibles, d’une part, et de la connaissance disponible sur l’état des réacteurs et les actions en cours par les opérateurs sur la centrale (par exemple, réalisation d’un éventage, c’est-à-dire l’action de relâcher volontairement de la vapeur radioactive dans l’atmosphère afin de diminuer la pression de l’enceinte du réacteur et d’éviter une explosion). L’accès à ces informations étant partiel, les prévisions de trajectoire des panaches radioactifs étaient elles-mêmes entachées d’incertitudes. En particulier, la trace de contamination du sol au nord-ouest de la centrale de Fukushima Daiichi, visible sur les relevés de mesures aéroportées effectuées dès le 17 mars par les ministères de l’environnement américain et japonais (DOE, 2011), n’avait été prédite par aucun modèle de dispersion [Mathieu 2012]. Cette contamination, reconstituée par différentes campagnes de mesures après les rejets, est présentée figure 1.13. Cette incapacité des modèles à prédire les dépôts correctement vient du fait que la contamination a eu lieu durant quelques heures, par la conjonction de rejets, de vents tournant vers le nord-ouest et d’un épisode de précipitations qui a rabattu les polluants au sol. L’incapacité des modèles de dispersion déterministes à prévoir cet épisode de contamination jusqu’à plusieurs semaines après l’accident illustre la nécessité de mieux prendre en compte les incertitudes, et d’intégrer la possibilité de prévoir des scénarios ayant une faible probabilité d’occurrence mais un fort impact sur l’environnement.

Après l’accident, malgré le grand nombre d’observations radiologiques dans l’environnement, les modèles de dispersion atmosphérique ont été utilisés en complément de celles-ci pour caractériser et comprendre les rejets et leur impact sur l’environnement [Mathieu 2018a]. En particulier, dans cette phase post-accidentelle, les modèles de dispersion permettent de :

- donner une vision spatiale tridimensionnelle de la contamination et son évolution dans le temps, alors que les mesures ; fournissent, soit un « cliché » instantané avec une bonne résolution spatiale mais cumulé dans le temps (c’est le cas des mesures de dépôt), soit une bonne évolution temporelle mais seulement en des points répartis de façon non homogène sur le territoire (mesures d’activité et de débit de dose) ;
- reconstituer les doses potentiellement reçues en incluant des radionucléides à vie courte, non disponibles dans les mesures faites après l’accident ;
- comprendre les phénomènes expliquant les différentes traces de contaminations (dépôt sec ou humide, lessivage d’un panache en hauteur...) ;
- reconstituer les événements de rejet, en conjonction avec les mesures disponibles et les données météorologiques.

Pour atteindre ces différents objectifs, il est nécessaire de quantifier au mieux les différentes sources d’incertitudes liées aux données et aux modèles de dispersion, afin d’évaluer la confiance que l’on peut avoir dans les résultats des simulations.

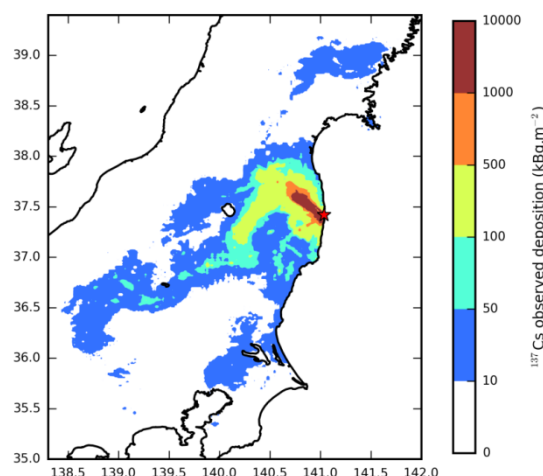


FIGURE 1.13 – Carte des dépôts de ^{137}Cs obtenue par mesures aéroportées et prélèvements de terrain durant différentes campagnes ([MEXT 2011b], [MEXT 2011a], [MEXT 2012])

1.5 Principales sources d'incertitudes

La dispersion atmosphérique des polluants (ici, il s'agit de radio-isotopes), dépend principalement des conditions météorologiques et des caractéristiques de la source de rejet. Ce sont toutes les deux des données complexes de grande dimension. Une prévision météorologique est un ensemble de variables (pluie, vent, température, pression, etc.) qui sont des champs de données en 2D ou 3D variant en temps. Le terme source contient des séries temporelles de quantités de rejet des radionucléides dans l'atmosphère. De plus, au sein du modèle, il existe des paramétrisations permettant de représenter différents processus physiques présentés au paragraphe 1.3.2, qui sont elles-mêmes sources d'erreurs. Le résultat des modèles de dispersion, plus précisément ℓdX dans cette étude, dépend fortement de ces trois éléments. Ces trois sources d'incertitudes sont détaillées dans [Mathieu 2018b]. L'objet de cette partie est de décrire la façon dont ces différentes sources d'incertitudes ont été prises en compte dans le cas de Fukushima. Les ensembles météorologiques utilisés pour représenter l'incertitude météorologique sont décrits dans la partie 1.5.1. L'ensemble de termes sources issus de la littérature pour le cas de Fukushima est brièvement décrit dans la partie 1.5.2. Enfin, les différents paramètres incertains du modèle de dispersion sont présentés en partie 1.5.3. Ces données et les incertitudes associées sont utilisées dans les études détaillées dans les chapitres suivants : la propagation par méthode Monte Carlo au chapitre 2, la construction des métamodèles au chapitre 3, et enfin la calibration au chapitre 5.

1.5.1 Ensemble de prévisions météorologiques

Les modèles numériques de prévision météorologique (NWP – Numerical Weather Prediction models) sont utilisés pour la prévision de l'état futur de l'atmosphère, en assimilant

des observations de l'état actuel de ce dernier. Il existe trois sources d'erreurs qui entachent ces prévisions. D'une part, les observations météorologiques disponibles ne permettent pas de représenter parfaitement l'état initial de l'atmosphère (appelé conditions initiales) du fait des imprécisions de mesure et de la couverture partielle des observations. Le processus consistant à intégrer l'information de ces observations aux données de modélisation 3D pour produire une représentation cohérente de l'état de l'atmosphère est appelé assimilation de données, et il est lui-même source d'erreurs et d'approximations. La seconde source d'erreur est la représentation des conditions aux limites, notamment les interactions entre l'atmosphère et l'océan. Enfin, la dernière source d'erreur est le modèle météorologique lui-même, qui est fondé sur une discrétisation numérique des équations de la mécanique des fluides sur un maillage. Les erreurs liées à cette représentation numérique et à l'utilisation de paramétrisations pour représenter les processus sous-maille ont été évoquées pour les modèles eulériens de dispersion atmosphérique au paragraphe 1.3.4 ; cela s'applique également aux modèles météorologiques. Pour prendre en compte ces différentes incertitudes, le Centre Européen de Prévision (CEP,) ou European Centre for Medium-Range Weather Forecasts (ECMWF) en anglais, a mis au point un système permettant de prédire la fiabilité des prévisions. Ce système, opérationnel au CEP depuis 1992, est le système de prévision d'ensemble, Ensemble Prediction System–EPS en anglais [Leutbecher 2008].

Le système de prévision d'ensemble du CEP représente les incertitudes liées aux conditions initiales en créant un ensemble de 50 états initiaux légèrement différents, créés en perturbant les conditions initiales en tenant compte de l'erreur d'observation (figure 1.14). Le système d'assimilation de données ajuste la trajectoire de chaque membre de l'ensemble pour les rapprocher des observations, ce qui crée un ensemble initial (« first guess »). L'atmosphère étant un système chaotique, des perturbations initiales assez faibles peuvent conduire, après propagation dans le temps, à des changements significatifs à grande échelle. Le taux de croissance de cette erreur dépend de la situation météorologique : certains jours, la prévision sera toujours plus ou moins similaire et les petites perturbations n'auront que peu d'impact sur la prévision. Dans d'autres cas, un petit changement de condition initiale pourra conduire à des prévisions radicalement différentes pour les jours suivants (figure 1.15). En plus des incertitudes sur les conditions initiales, les erreurs de modèle, liées notamment aux paramétrisations physiques, sont prises en compte en perturbant celles-ci de façon stochastique afin de tenir compte des incertitudes inhérentes aux processus sous-maille. En combinant ces deux méthodes, la dispersion de l'ensemble (souvent représenté par son écart type) est supposé représentatif de la plage d'incertitude des variables prédites sur une période de temps donnée. Les perturbations initiales sont ajustées de façon à ce que le taux de croissance de l'erreur soit représentatif de l'incertitude généralement observée par les comparaisons modèle-mesure au bout d'un certain temps. Pour les prévisions du CEP, l'ensemble est optimisé pour être représentatif de l'erreur à échéance de prévision de deux jours. Toutefois, il est utilisé pour toutes les échéances de prévision jusqu'à 15 jours.

Les ensembles météorologiques sont donc calibrés pour représenter les incertitudes de certaines variables météorologiques à une échéance de prévision donnée. Toutefois, les variables d'intérêt sur lesquelles ces erreurs sont calibrées (typiquement la température à 500 hPa) sont différentes des variables utilisées pour la dispersion atmosphérique. Dans le cas de la prévision météorologique, il s'agit en effet d'être représentatif de l'échelle synoptique, c'est-à-dire des

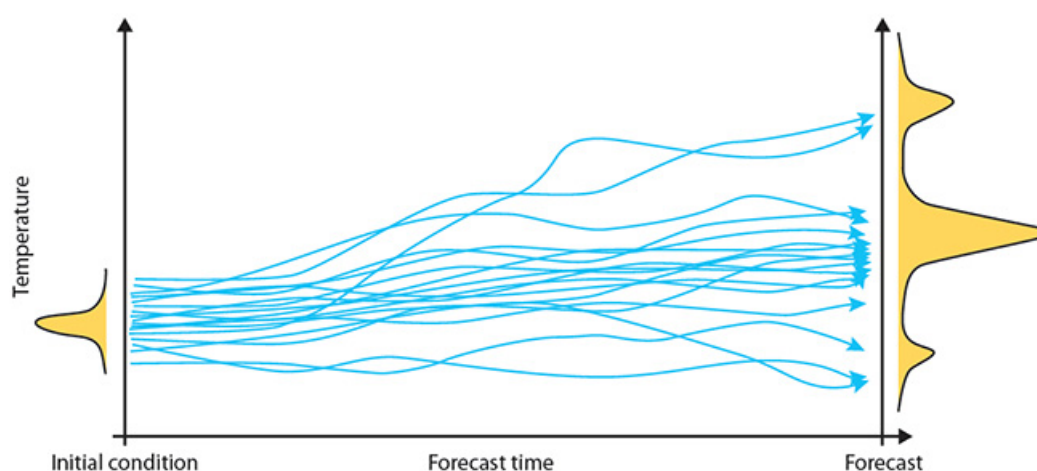


FIGURE 1.14 – Illustration de perturbations de l'état initial de l'atmosphère, [Leutbecher 2007]

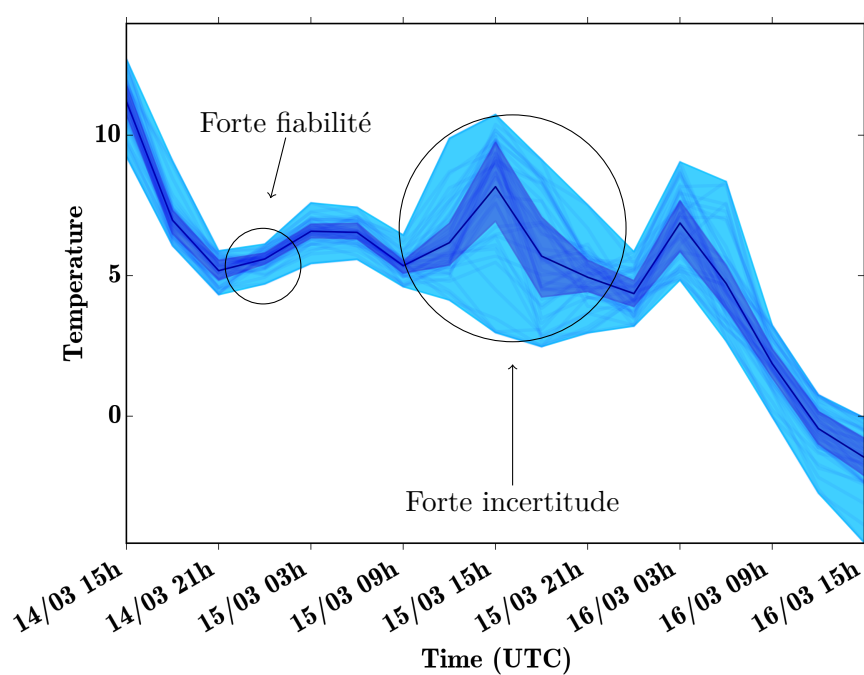


FIGURE 1.15 – Exemple de la prévision d'ensemble météorologique sur la température. La bande bleue claire contient tous les membres de l'ensemble, aussi appelée enveloppe ; la bande bleue foncée contient 50% des membres, soit 25 membres et la ligne bleu marine est la moyenne d'ensemble.

mouvements d'air à grande échelle dans la troposphère libre, hors de l'influence directe du sol. Les ensembles météorologiques sont également utilisés pour représenter des données utiles pour la dispersion telles que les précipitations ou le vent dans les basses couches de l'atmosphère, mais leur étalement n'est pas nécessairement représentatif des incertitudes liées à ces variables. L'un des objectifs de ce travail est d'étudier l'utilisation de ces ensembles météorologiques en entrée des modèles de dispersion ainsi que leurs limitations, pour fournir une estimation des incertitudes en dispersion.

1.5.2 Termes sources de l'accident de Fukushima

La première information à fournir à un modèle de dispersion atmosphérique pour décrire les conséquences d'un rejet accidentel est le terme source. Le terme source est une expression technique utilisée pour décrire le rejet accidentel de matières radioactives à partir d'une installation nucléaire dans l'environnement. Non seulement la quantité de radioactivité libérée est importante mais aussi sa répartition dans le temps, ainsi que les formes chimiques et physiques de ces matières. Autrement dit, un terme source contient des séries temporelles de quantité de rejet dans l'environnement des radionucléides sous différents états physiques (particule ou gaz) et isotopes.

L'estimation d'un terme source peut se faire par diverses manières. Dans l'étude de l'accident nucléaire à Fukushima, les termes source ont été reconstruits majoritairement *a posteriori*. Cela signifie que les informations sur l'émission sont calculées à partir des informations combinées des mesures de radioactivité dans l'air et au sol, à l'aide de modèles de dispersion et d'informations météorologiques, une fois que le rejet a eu lieu. Des méthodes de modélisation inverse plus ou moins sophistiquées ont été utilisées pour estimer les valeurs du terme source à partir des mesures du terrain. Pour modéliser les incertitudes de ce type de d'entrée, nous utilisons ici neuf termes sources dans la littérature de Fukushima : [Katata 2015], [Terada 2012], [Mathieu 2012], [Saunier 2013], [Saunier 2016], [Stohl 2012], [Winiarek 2012], [Winiarek 2014], [Yumimoto 2016].

Le terme source [Terada 2012] a été construit sur la base des mesures de concentration dans l'air et de modélisation. Le terme source de [Katata 2015] est une mise à jour du précédent. Il estime les émissions atmosphériques à partir de modélisations atmosphériques et océaniques comparées aux mesures de la concentration ou de débit de dose dans l'environnement. Le terme source [Yumimoto 2016] est encore une mise à jour des termes sources précédents, incluant davantage de mesures. Le terme source de [Mathieu 2012] a été estimé dans le mois qui a suivi l'accident, sur la base d'une combinaison entre les informations d'événements de rejet sur l'installation, les mesures de radioactivité (débit de dose) en champ proche, et les résultats des modèles de dispersion à courte et longue distance. Le terme source de [Saunier 2013] utilise des mesures de débit de dose dans le cadre d'une approche automatique de modélisation inverse basée sur une minimisation des écarts entre les résultats des simulations de dispersion et les mesures disponibles. [Saunier 2016] utilise la même approche avec des mesures d'activité de ^{137}Cs dans l'air. Tous les termes sources précédents se basent sur des mesures effectuées sur l'île de Honshu. Le terme source [Stohl 2012] utilise une combinaison de mesure d'activité volumique à l'échelle du globe de ^{137}Cs et ^{133}Xe ainsi que quelques mesures au Japon, et un *a priori*. Les deux termes sources [Winiarek 2012] et [Winiarek 2014] ont été établis respectivement sur des

mesures d'activité volumique et de dépôt de ^{137}Cs , en ajoutant une couche d'estimation des erreurs antérieures lors de l'assimilation des observations provenant de plusieurs ensembles de données.

1.5.3 Paramètres incertains dans le modèle

Les sources d'incertitudes liées aux modèles de dispersion sont introduites par plusieurs paramètres scalaires représentant différents processus dans les EDP, par exemple la vitesse de dépôt, lessivage, coefficient de diffusion, etc. Ces paramètres ont une influence non négligeable sur les résultats de prédiction du modèle. [Girard 2016] a étudié la sensibilité des sorties radioactives (concentration, dépôt et débit dose) par rapport aux variations des paramètres d'entrée, y compris des conditions météorologiques, le terme source utilisé et les paramètres de caractérisation de l'accident de Fukushima. Leurs résultats ont montré que certains de ces variables influencent considérablement les résultats du modèle.

Coefficient de lessivage ou dépôt humide

Le paramétrage du coefficient de lessivage est de la forme $\Lambda_r = ap_0^b$, où p_0 est l'intensité de la pluie. Nous appelons, par la suite, a le facteur de lessivage et b l'exposant de lessivage. La physique distingue différentes valeurs de a et b pour le lessivage ayant lieu en dessous du nuage (rabattement des particules radioactives par les gouttes de pluie ou les flocons de neige) et à l'intérieur des nuages (captation des particules par les gouttelettes de nuage), figure 1.16. Le choix de modèles exponentiels simples avec seulement deux paramètres pour représenter à la fois le lessivage dans le nuage et sous le nuage est déterminé par la faible connaissance des paramètres microphysiques, tels que la distribution de la taille des particules, requis par des modèles plus détaillés. Dans la littérature, les coefficients de lessivage utilisés par les modèles peuvent varier d'environ un facteur 10, ce qui peut avoir des répercussions significatives sur la prévision des zones contaminées par le dépôt humide ([Quérel 2015]; [Quérel 2016]).

Vitesse de dépôt sec

La plage de variation de la vitesse de dépôt sec est issue d'une revue de la littérature : [Maryon 1991], [Rodos 1999], [Baklanov 2001], [Brandt 2002]. On a supposé que toutes les espèces particulières avaient la même vitesse de dépôt sec, fondée sur un diamètre des particules de $1\mu\text{m}$, celui des aérosols atmosphériques. Cette hypothèse est motivée par le fait qu'à une distance suffisamment grande de la source, les particules les plus lourdes se sont déposées au sol et les plus petites se sont agglomérées ou ont été captées par les végétaux ou les gouttes (en cas de pluie).

Coefficient de diffusion horizontale

Les coefficients de diffusion horizontale, K_u et K_v (premier et deuxième élément de la diagonale de la matrice $\mathbf{K}$) sont constants et homogènes. Peu de références sur la détermination des coefficients de diffusion horizontale ont été trouvées dans la littérature. La plage de variation utilisée dans le chapitre de propagation par méthode Monte Carlo et dans la construction

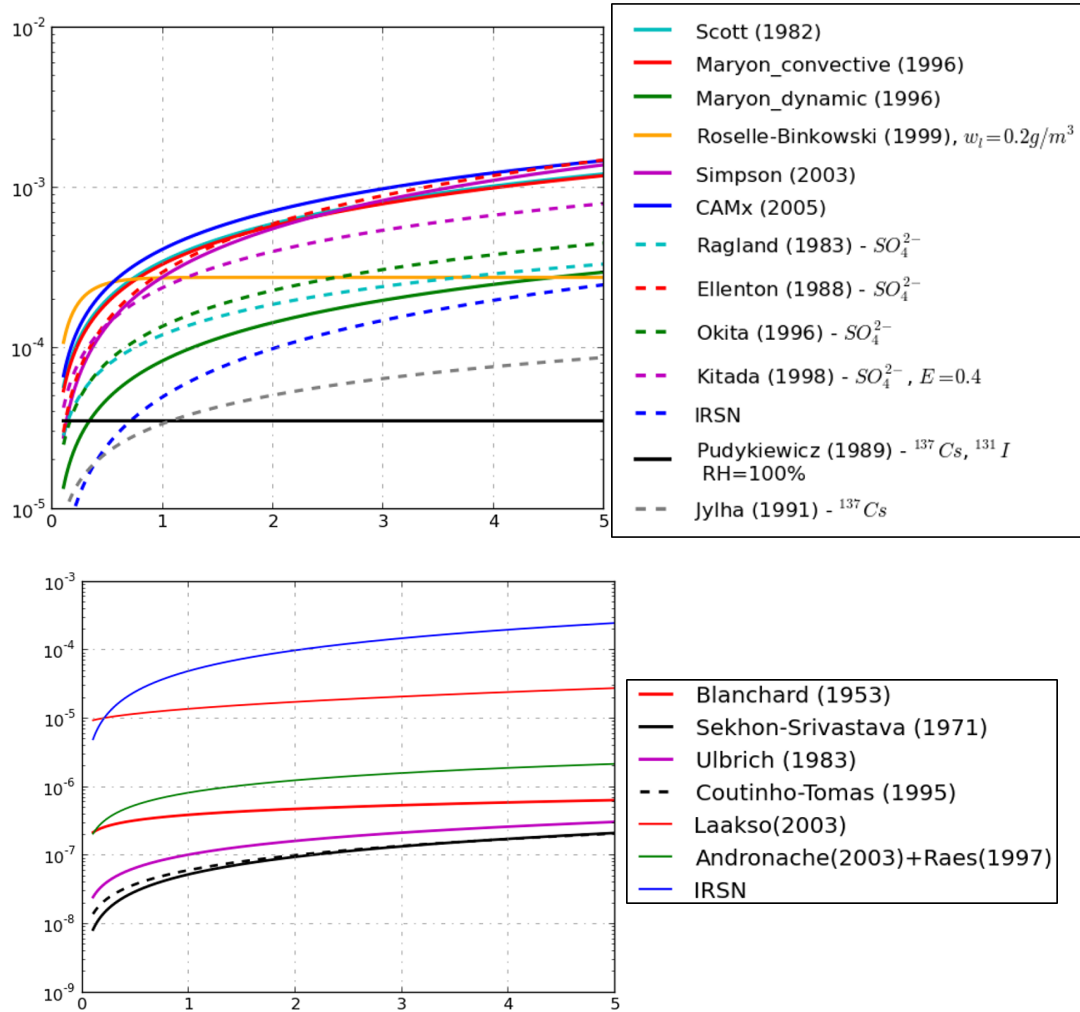


FIGURE 1.16 – Exemple de plage de variation des coefficients de lessivage dans le nuage (en haut), [Qu  rel 2015], et sous le nuage (en bas), [Duhanyan 2011], suivant diff  rents mod  les

des métamodèles est assez large et englobe les valeurs typiques trouvées dans la littérature ([Girard 2016], [Brandt 1998], [Ryall 1998], [Sørensen 1998], [Yamartino 2000]).

Coefficient de diffusion verticale

Le coefficient de diffusion verticale, K_z (troisième élément diagonal de la matrice $\mathbf{K}$), est calculé par la paramétrisation de [Louis 1979] au-dessus de la couche limite et à l’intérieur de la couche limite stable. Il est calculé avec le paramétrage de [Troen 1986] dans la couche limite instable. [Hanna 2001] suggère un facteur d’incertitude de 3 avec une hypothèse log-normale englobant 95% des valeurs possibles.

Propagation d'incertitude par méthode de Monte Carlo

Contents

2.1 Introduction	43
2.1.1 Context	43
2.1.2 Uncertainty analysis	44
2.1.3 Objectives of the study	46
2.2 Analysis of a meteorological ensemble	46
2.2.1 Description of ensemble weather forecasts	46
2.2.2 Model-to-data comparison of the meteorological ensemble	49
2.3 Monte Carlo simulations	56
2.3.1 ℓdX description	56
2.3.2 Simulation setup	56
2.3.3 Description of the measurements	58
2.3.4 Model-to-data comparison for Monte Carlo simulations	60
2.4 Synthesis and perspectives	67

2.1 Introduction

2.1.1 Context

Atmospheric dispersion modelling is the numerical implementation of physical models that describe the evolution of air pollutants in the atmosphere. At IRSN (French Institute of Radiation Protection and Nuclear Safety), atmospheric dispersion models are developed and used to determine the consequences of an accidental release of radionuclides in the atmosphere. The results provide an estimate of the surface location of the potentially contaminated areas and the corresponding impact on human health and environment. These estimations can be used as a tool for decision making. On their basis, countermeasures may be proposed to protect the population, such as sheltering, evacuation, thyroid protection by using potassium stable iodine, restriction of food intake or water supply.

At IRSN, the two dispersion models used are the short-range model, pX , used within a few tens of kilometers from the source, and the long-range model, ℓdX . The latter, used in this study, is an Eulerian transport model, which is based on the solution of a system of 3D advection-diffusion equations. ℓdX is an operational version of Polyphemus/Polair3D [Mallet 2007]. The

two models have been used during and after the Fukushima nuclear disaster in 2011, to carry out atmospheric dispersion simulations, both at local scale [Korsakissok 2013] with the pX model and at regional scale with the ℓdX model ([Mathieu 2012]; [Saunier 2013]). These studies have allowed a better understanding of the different contamination events, but also shed light on the uncertainties that still linger more than eight years after the accident.

Dealing with uncertain information is inherent to a crisis situation, where lack of information and incorrect information (e.g., human errors) are key issues, especially in the early stage of an accident. As the knowledge of the damaged installation and sequence of events improves with time, the confidence in the simulation results should increase accordingly. However, as proven by the Fukushima case, uncertainties can still be large at a post-accident stage, even when a large number of environmental observations are available. Besides, time is a key issue in the decision making process during an emergency. Currently, numerical simulations are carried out in a deterministic way, leading to a single result obtained by making conservative hypotheses in an attempt to encompass all sources of uncertainties. This practical approach has its limitations, because (a) it is very difficult to ensure that no possible contamination scenario has been left out, and (b) it may lead to unnecessary protective actions that can have large human and economic consequences. Therefore, a better understanding and quantification of all sources of uncertainties related to atmospheric dispersion simulations is crucial.

2.1.2 Uncertainty analysis

The stages of uncertainty analysis of numerical models are :

1. Identify the main sources of uncertainty and the parameters that significantly represent them,
2. For each parameter, construct a probability distribution that represents our *a priori* knowledge on this variable,
3. Determine the output variables that are sensitive to these input variables,
4. Propagate the uncertainties through the model,
5. Use the observations to improve our knowledge on these parameters (calibration).

Input uncertainties

Uncertainties in atmospheric dispersion simulations primarily stem from input data, namely meteorological fields and source term. The former comprise two or three-dimensional fields varying in time for several variables such as wind, rain, or atmospheric stability. The latter consist in the time-varying release rate for several emitted species, as well as emission location and release height (including possible buoyancy effects).

As far as meteorological fields are concerned, the timing and location of rain events, particularly convective showers, are especially difficult to forecast. In the case of low wind speed, the wind direction is not well established and uncertainties can be very large. Therefore, meteorological uncertainties depend on the weather situation as well as on the forecast lead (i.e. how far in the future the weather forecast is used) and the meteorological model characteristics, in particular, its spatial and temporal resolution. In the case of the Fukushima accident, several

studies highlighted the uncertainty of the meteorological inputs, and their effect on deposition patterns ([Draxler 2015]; [Arnold 2015]; [Leadbetter 2015]).

The *a priori* knowledge of a source term, in the pre or early-release stage, is derived from the available knowledge of the state of the reactor and on-site operators' actions. In that case, uncertainties can be huge but very dependent on the accidental scenario. There are *deep uncertainties* that may not be quantified, and are not part of the present study. In the post-accident stage, i.e. when radiological observations are available, *a posteriori* source terms can be constructed by taking into account these measurements. In the case of the Fukushima accident, all source terms in the literature were derived from environmental measurements, but there is still a large variability in these estimations, both on the total emitted quantities for the different radionuclides (a factor 2 or more) and the emission rate as a function of time [Mathieu 2018a].

Aside from these two sources of uncertainties, physical processes such as dry deposition, wet scavenging, or turbulent diffusion, are represented with physical parameterizations that are also subject to errors. Finally, the dispersion model in itself leads to biases due to numerical and mathematical errors. [Mathieu 2018b] conducted a literature review of the range of variation for the most common physical parameters used in atmospheric dispersion models. Concerning the Fukushima case, the sensitivity to wet deposition schemes was investigated by [Arnold 2015], [Quérel 2015] and [Leadbetter 2015].

Sensitivity and uncertainty propagation

The aforementioned studies conducted after the Fukushima accident are all *local* sensitivity analyses, devoted to investigate the influence of one particular kind of variable on a given output. While these studies allow analyzing in detail the effect of a particular physical process, they do not provide a global view on the joint effect of all uncertain variables together. A *global* sensitivity analysis based on Polair3D was carried out by [Girard 2014, Girard 2016] on the Fukushima case. These papers studied the relative influence of a set of uncertain inputs on several outputs. They concluded that the highest uncertainties in nuclear dispersion simulations for the Fukushima accident were due to the weather forecasts and the source term, and quantified the part of output variance related to each input variable. These studies were carried out by using large ranges of variation in an attempt to broadly encompass the input uncertainties, without trying to determine realistic input distributions. In particular, constant, homogeneous perturbations were applied to the meteorological fields, while meteorological uncertainties are inter-dependent and highly heterogeneous. To better represent these uncertainties, meteorological ensembles, constructed by weather offices to be representative of weather forecast uncertainties, have been used in the case of Fukushima ([Sørensen 2016]; [Kajino 2019]) as well as in hypothetical accident scenarios ([Sørensen 2016]; [Korsakissok 2019]). These studies accounted for meteorological uncertainties only, or, in some cases, both meteorological and source term uncertainties in a systematic approach (that is, all combinations of meteorology and source terms were simulated), which implies a significant computational burden.

To properly conduct an uncertainty analysis, prior input distributions have to be determined for all uncertain input parameters; these distributions are then sampled and propagated within the atmospheric dispersion model, typically with a Monte Carlo method. This approach, combined with the use of weather forecast ensembles, was carried out with the short-range

model pX on the Fukushima case and compared to the systematic, cross-simulations approach mentioned earlier ([Périllat 2016]). Simulations were compared to radiological observations of dose rate and airborne deposition measurements collected in Japan.

2.1.3 Objectives of the study

The aim of this study is twofold : (a) to investigate the use of a coarse-resolution weather forecast ensemble for propagation of meteorological uncertainty, and (b) to propagate the uncertainties in the Fukushima case at continental scale (up to a few hundreds of km), using the ℓdX model.

[Périllat 2016] showed that using a high-resolution meteorological ensemble may not be significantly beneficial for our purpose, and a coarse-resolution, operational ensemble may be sufficient to encompass most uncertainties. The first part of this paper aims at analyzing more thoroughly an operational weather forecast ensemble provided by the European Center for Medium-range Weather Forecasts (ECMWF), as could be used in case of an emergency (Section 2.2). In particular, the effects of grid resolution (Section 2.2.2) and meteorological observation errors (Section 2.2.2) are investigated.

The second part of this paper (Section 2.3) presents Monte Carlo simulations for the Fukushima disaster with ℓdX model. In this study, we used all members of the weather ensemble, six *a posteriori* source terms from the literature, and random perturbations on the other inputs. The simulations are compared to radiological observations of activity concentration in air, gamma dose rate and airborne deposition measurements collected in Japan. The principal objectives are the validation of the *a priori* distributions of the uncertain inputs and the analysis of sensitive outputs.

2.2 Analysis of a meteorological ensemble

2.2.1 Description of ensemble weather forecasts

Principle of ensemble forecast

Numerical weather prediction (NWP) models are used predict the evolution of the state of the atmosphere. The first step is to evaluate the current state of the atmosphere (called the analysis). To achieve this goal, millions of observations are available everyday. The various observations and *a priori* simulated estimations are combined to make a consistent 3-D multivariate representation of the atmosphere, which is a process called data assimilation. The resulting analysis is subject to uncertainties coming from the observations partial coverage, observational errors as well as errors introduced by the data assimilation process itself. The analysis is used as initial condition to the subsequent forecasts made by NWP models, which rely on a discretization of the mathematical equations governing the physical processes occurring in the atmosphere. Therefore, further uncertainties are introduced in the forecasts by the physical and numerical approximations made by the model. Boundary conditions, such as interactions of the atmosphere with the land or ocean, are imperfectly taken into account, leading to an additional source of uncertainties. These three sources of uncertainty or error can all be reasons for a forecast to differ from what actually occurs. Due to the chaotic nature of the atmosphere, small initial

errors can grow during the forecast and lead to substantial changes in large-scale patterns. The rate at which these inevitable errors grow vary from day to day (it is flow-dependent). Sometimes, small errors in the initial conditions may result in dramatically different outcomes, while on other occasions, the forecasts will be more or less the same regardless the slight perturbations in the initial conditions.

An ensemble is a set of forecasts run from different initial conditions to account for initial uncertainties, and possibly including NWP model errors as well. The ensemble of forecasts provides a range of possible future scenarios describing how the atmospheric flow may evolve, in accordance with our knowledge of its initial state and of the model capabilities. The ECMWF-ENS operational ensemble prediction system (EPS), as described in [Leutbecher 2014], comprises fifty simulations representing slightly different starting conditions, plus one "control" member (without perturbation). Each simulation is called a member. The initial perturbations are generated from an ensemble data assimilation (EDA) process [Buizza 2008]. It consists in adjusting the recent ensemble forecast (providing a "first guess") to available observations using 4D-var techniques. Ideally, the EDA would be sufficient to represent the initial uncertainties. As this step does not provide enough spread, this approach is completed by another one called Singular Vector-SV [Leutbecher 2014]. This mathematical technique generates perturbations to an initial state that will grow the fastest over a given time interval (here, 48 hours). Using these perturbations together with the EDA helps to compensate for having a relatively small number of ensemble members, by ensuring to produce fast-growing forecast spread.

In addition, the ECMWF-ENS forecast takes account of model uncertainties. Physical parametrizations are used to represent the effect of processes that are unresolved (or partially resolved) by the model. In particular, subgrid-scale phenomena, occurring at a scale smaller than that of a grid box, are parametrized. Uncertainties arise from the parametrized processes due to two aspects :

- Using simplified representations of processes not fully resolved : the impact on the resolved flow is represented via some bulk/averaged quantity (e.g., surface drag, convection rates, phase transitions, radiation transfers),
- Representing processes that are poorly constrained due to lack of observational coverage, or not fully understood (e.g., vertical cloud-overlap, composition of the atmosphere, non-orographic drag).

ECMWF uses a combination of two methods : SPPT (Stochastically Perturbed Parametrization Tendencies) and SKEB (Stochastic Kinetic Energy Backscatter scheme) to account for these different aspects of the model uncertainty [Palmer 2009]. The SPPT scheme introduces a random element to the parametrized variables, in order to account for the subgrid-scale uncertainty. The SKEB scheme accounts for energy transfers between subgrid scales and larger scales. In the ensemble construction, each initial (perturbed) state is then used with a different version of the numerical model, modified by random perturbations from SPPT and SKEB schemes. This inclusion of model uncertainties generates a larger ensemble's spread, which improves the forecasting skill by making the spread more representative of the overall forecast error. In practice, there are limitations to these schemes in terms of physical consistency of the forecast and partial representation of the model uncertainties. A better representation of model uncertainties is currently under development at ECMWF [Ollinaho 2017].

Application of ensemble weather forecast to the Fukushima study

The ECMWF-ENS operational ensemble used in this study was retrieved from the ECMWF archive in 2017. It contains 50 members with 0.25° of horizontal resolution and three hour-time step, the domain spans 40×40 grid boxes that cover a great part of Honshu island. The vertical resolution of the data is coarse, with only five pressure levels to cover the boundary layer height : 50000, 70000, 85000, 92500, 100000 Pa. For dispersion purposes, the meteorological variables are then interpolated on a finer grid adopted to our needs (see section 2.3.2). The ECMWF operational ensemble forecast is representative of the kind of data that could be retrieved and used on a relatively short term for emergency response. Ensembles of finer resolution have been made operational for weather forecast at national level as computer efficiency increases, but for a given computational burden, there is a balance between the respective benefit of increasing resolution compared to using more ensemble members to represent uncertainty [Raynaud 2017]). The huge amount of data for high-resolution ensembles currently poses a challenge for downloading, processing, using and storing the data and results. On the other hand, previous studies have shown that coarse-resolution ensembles may be sufficient to provide a first, broad view of the main uncertainties related to dispersion [Périllat 2016]. One of the main objectives of this study is to further investigate to what extent an ensemble of this (spatial and temporal) resolution may be appropriate to account for uncertainties in the first stage of emergency response.

Another concern is that such an ensemble is designed to capture the spread of meteorological variables on synoptic scales, while the domain and scales in this study are much smaller. Weather forecast ensembles, however, are also used in practice to predict local variables within the boundary layer, such as rain showers or wind at 10-m level. It is acknowledged that, generally, these ensembles are under-dispersive when compared to boundary layer variables, since they fail to take into account all sources of uncertainties at this scale. Before using them for dispersion within the boundary layer, a first step is to perform a model-to-data comparison to such local-scale variables in order to have a better view of this under-dispersion (section 2.2.2).

A third challenge in this study is posed by the length of the accident (three weeks). Usually, weather forecasts are made up to about five days. When the accidental scenario spans a longer time period, it raises the issue of how to properly combine different ensemble forecasts made from different initial times to cover the whole simulation period. To deal with this issue, this study uses a meteorological ensemble made of forecasts that started from 12 hours to 36 hours before the current time. For day D at 00 :00, the ensemble of forecasts was simulated starting from analyzed states (i.e., initial states merging simulations and observations) for 12 :00 of the day $D - 1$. We make use of the subsequent forecasts until the end of day D . For day $D + 1$, we switch to the better forecasts that originate from day D at 12 :00. Depending on the construction of the set, the disturbances of the initial state are independent of the day to the other. This means that the simulations of a member M of the day D and $D + 1$ do not come from the same perturbation. This leads to assimilation jumps when we go from one day to the next. Figure 2.1 illustrates the cycle, using forecasts with a time step of three hours.

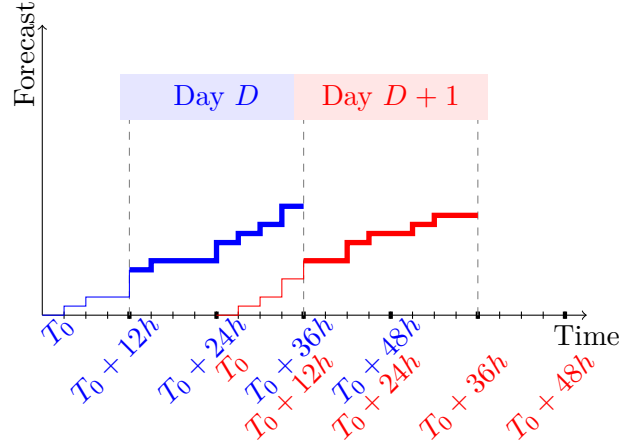


FIGURE 2.1 – Illustration of which meteorological forecasts are used (in bold) during the day. The forecasts originate from 12 :00 on one day and are used the next day only.

This process leaves clear room for improvement and research on methods to insure a better temporal continuity in the ensemble members for long simulation periods. This ensemble, however, is representative of what could be constructed in a real emergency case, in the sense that 24 hours is the typical forecast period that would be used in order to anticipate the consequences of potential releases. The Fukushima accident is a succession of independent releases coming from different units and installation events [Mathieu 2018a]. It is probable that this kind of situation would be dealt with by repeatedly using 24-hour forecasts used to infer or update recommendations for the protection of the population. The ensemble constructed here is made in this perspective. The first twelve hours of the forecast are not used in order to let the perturbations grow, then the next 24 hours are used by the simulation.

2.2.2 Model-to-data comparison of the meteorological ensemble

The Automated Meteorological Data Acquisition System (AMeDAS) in Japan provides observations of wind direction and velocity, precipitation and/or temperature on more than a thousand measurement stations, from which 653 are within our simulation domain. Stations within the same grid cell are averaged for comparison purposes, and there are 354 cells containing at least one meteorological measurement point. The comparisons are made from March 12th to March 30th 2011 ; some observations are given every 10 minutes, others every 1 hour. As our forecast time step is 3 hours, the observations are taken from 1 hour-time step observations (or cumulated for the rain) to correspond to the model time resolution. Therefore, our comparisons are made on 354 points and 150 time steps. Finally, among the selected points and time steps, the observed values of wind and temperature are unavailable about 20 % of the time. Indeed, many stations were out of order during the first days following the tsunami that triggered the Fukushima accident.

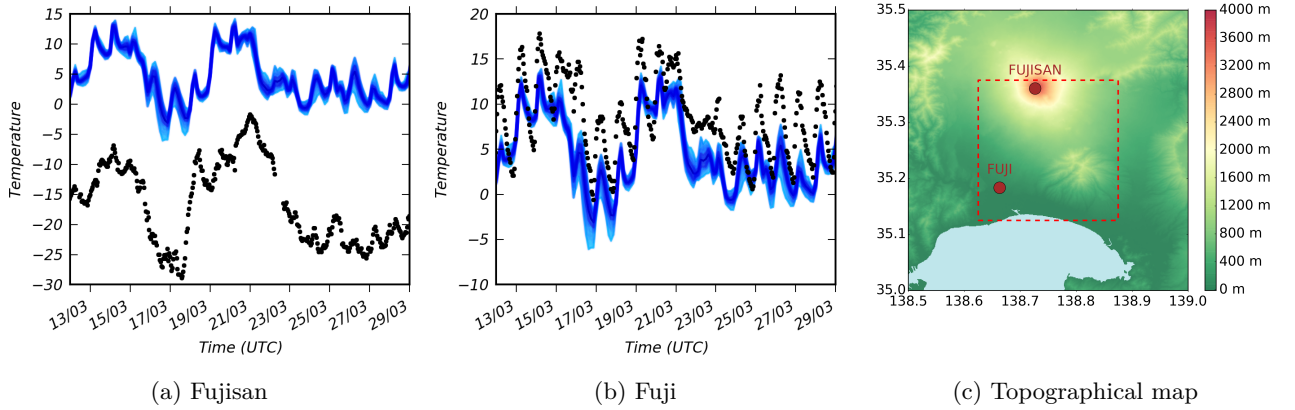


FIGURE 2.2 – Temperature comparisons at two stations in the same grid cell (Fujisan and Fuji), and topographical map of this cell. The weather forecast ensemble, which contains 50 members, corresponds to the blue band and the black dots are measured values.

Subgrid-scale variability

While ground observations are representative of one given location, a forecast gives the average of variables over a three-dimensional grid cell. In addition, the horizontal resolution of this ensemble is 0.25° , i.e about $22 \text{ km} \times 28 \text{ km}$ at the center of the simulation domain. A cell can be spread over different topographies. Therefore, the comparison to observations from a given monitoring stations can be biased. For example, the temperature predicted at Fujisan is much higher than the observation ; but the model-to-data difference is constant in time, about 20°C . This phenomenon can be explained as follow :

- The altitude of Fujisan station is about 3775 m above sea level. The temperature observed at this station is obviously lower than that at a lower altitude ;
- With a large resolution, the cell containing the Fujisan station has a large variation in altitude. Weather forecast takes into account this variation and compute the average across the cell, figure 2.2.

Stations like Fujisan create a bias on the evaluation of weather ensemble. If we compare the forecast interpolated at each station to the measurements, the bias can be important. To circumvent this issue, it is preferable to compare the simulations to averages of all stations in each grid cell. However, this approach is limited by the small number of stations in one grid cell.

Probabilistic indicators for ensemble evaluation

There are several output products for ensemble weather forecast systems, such as meteorograms, which display the time evolution of the distribution of meteorological parameters from the ensemble at a given location, as illustrated by Figure 2.2. To evaluate the quality of an ensemble by comparison to observations, it is necessary to derive probabilistic scores that will provide a view on the capability of the ensemble to correctly represent the uncertainty for a given meteorological variable. At ECMWF, the overall performance of the forecasts is

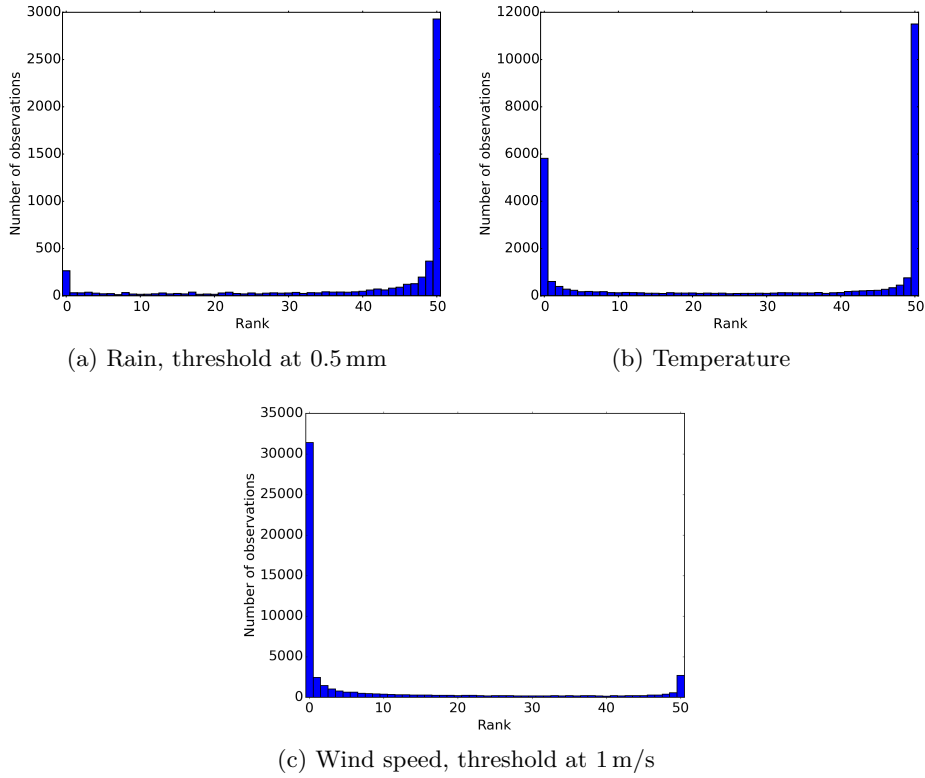


FIGURE 2.3 – Rank histograms of ensembles.

evaluated by using a set of headline scores which summarize different aspects of the forecast. Commonly used verification metrics are the rank histogram, spread-skill diagrams, Brier score and Continuous Ranked Probability Score (CRPS) [Hudson 2017].

While the ECMWF products are routinely verified against observations using these scores [Haiden 2019], the variables of interest are those of synoptic scale, as discuss in Section 2.2.1, whilst we focus here on boundary layer variables, which are much more difficult to represent but are of prime interest for dispersion. Besides, routine verification is often made for a particular forecast lead time, whereas in this case different lead times are used (from 12 to 36 hours after analysis time), to better represent what will actually be used for dispersion. Therefore, the verification conducted here is not representative of the ensemble quality for weather forecast, but rather, of how it may be suitable for use for atmospheric dispersion purposes.

Rank histograms are a good way to measure how well the ensemble is representative of the observed uncertainty. The rank of an observation is determined by counting how many members of the ensemble are below this observation. A "perfect" ensemble would have a flat histogram (see appendix B.1).

Figure 2.3 shows the rank histograms for precipitation, temperature, wind speed and wind direction of original and perturbed ensembles generated with all measurement stations. For the wind speed, we removed the values below 1 m s^{-1} because the measurements of low wind speeds

are highly uncertain. The ensemble does not spread out enough, hence its rank histograms show high external bars. There is a tendency to overestimate the wind speed (high left bar). This bias is partially explained by all the coastal cells that contain a large proportion of sea area, where the wind speed is generally higher than over land. The histogram for precipitation shows, with a high bar on the right, that the weather ensemble underestimates the observations. The chosen threshold was 0.5 mm h^{-1} , the rain gauge threshold, in order to avoid a bias due to the number of occurrences with no rain, where all members and observations are equal to zero. Therefore, only periods where either an observation or a member of the ensemble features a rain episode are taken into account. This leads to a limited number of observations for the construction of the rank diagram. There are only 9.7% rain observations whose values are higher than 0.5 mm h^{-1} . As discussed in Section 2.2.1, the operational ensemble forecasts are not designed to capture the full uncertainty of boundary layer parameters. This error is still greater due to the spatial and temporal scales of the ensemble. To accurately represent precipitation, the spatial and temporal scales are crucial. The subgrid effect described in Section 2.2.2 can lead to significant bias, due to the discrepancy between the cell average and the observation on station. Besides, the three-hour time step may lead to a delay in the beginning or end of the rain episode, since rainfall rates are interpolated between two time steps. These representativeness errors are not well taken into account by the ensemble approach, and are the main limitation of this dataset.

Accounting for meteorological observation errors in the comparison

The comparison of forecast members to the observational cell averages show that the ensemble's spread is not sufficient to encompass the observations. There are biases on all variables. The aim of this part is to take into account observational uncertainties in the model-to-observation comparisons. This allows to infer whether observation errors on boundary layer variables can explain a good part of the ensemble under-dispersion, or whether efforts should be put on other causes when designing meteorological ensembles for dispersion purposes. Another objective of this section is to illustrate how to take into account measurement errors during ensemble verification, which should ideally be applied in the future when comparing dispersion ensembles to radiological measurements (although not done in this study).

According to [Bowler 2009], taking into account observational errors can not be carried out on the same way for all variables. In the generation of the rank histograms, we propose to perturb each meteorological member by adding a random value drawn from a Gaussian density with mean zero, and a standard deviation that depends on each variable. After this perturbation, an ensemble member and the observation should follow the same distribution again, and the generation of the rank histogram is hopefully more consistent. This perturbation is not applied to the 3-D meteorological fields but only to the data used for model-to-observation comparisons. We can summarize that in the equation 2.1 below, where Y is the variable that we want to perturb given by an ensemble member :

$$\hat{Y} = Y + \varepsilon, \quad \varepsilon \sim \mathcal{N}(0, \sigma_Y) \text{ or } \mathcal{U}(a, b) \quad (2.1)$$

For the precipitation, the data is stored in cumulative form. [Bowler 2009] perturbed the precipitations by adding an error to the logarithm of the six-hours accumulated precipitation. It amounts to a multiplicative error of 10% on the accumulation. This perturbation must be

applied on the values higher than 0.1 mm/h. However, there are differences of accuracy between the rain gauges in Japan and in Europe. Indeed, Japanese rain gauges are less accurate than European rain gauges, 0.5 mm of accumulated rain per hour in Japan and 0.1 mm/h in Europe. The perturbation given by [Bowler 2009] is therefore not sufficient to improve the weather ensemble compared to the AMEDAS observations. For this reason, we chose to use instead a uniform distribution for the precipitation, $\mathcal{U}(-0.5, 0.5)$. Another important variable that was not presented in [Bowler 2009] is the wind direction. Like the rain, a perturbation with uniform distribution was applied for this variable. The Table 2.1 shows the estimated observational errors for all variables.

Variable (Y)	Distribution	σ_Y
Temperature [K]	Normal	1.5
Wind speed [m/s]	Normal	1.7
Precipitation [mm/h]	Uniform	$[-0.5, +0.5]$
Wind direction [$^\circ$]	Uniform	$[-10, +10]$

TABLE 2.1 – Observational errors, according to [Bowler 2009] or as perturbed in this study.

The figure 2.4 illustrates the model-to-data comparison between the original ensemble and its perturbation following table 2.1. These graphs show the rain, temperature and wind speed in a grid cell, that contains the meteorological stations Naganuma, Shirakawa and Kooriyama. This cell is chosen because there are measurement stations of air concentration (Sukagawa and Yabuki) and of gamma dose rates (Sukagawa, Yabuki, Shirakawa and Nishigo). The light-blue band extends across the weather ensemble envelop. The dark-blue band is delimited by the first and the third quartiles of the ensemble. The green lines are the observations of the stations in this cell. The red points are the observations averages. Another way to graphically represent the observational errors would have been to show error bars on the observation points. Here, we chose to apply the perturbations to the ensemble, but it is just a way to visualize to what extent the measurement error would influence the ensemble-to-observations comparison.

The previous curves bring a local view on the model-to-data comparison at a few measurement stations. However, evaluating an ensemble requires a global vision at all stations, and statistical tools can be used for this purpose. An ensemble is well designed if its spread is properly related to its observed errors. For a global interpretation, we used rank histograms.

Figure 2.5 shows the rank histograms including observation errors for precipitation, temperature, wind speed and wind direction ensembles generated with all measurement stations. For the wind speed, we removed the values below 1 m s^{-1} because the measurements of low wind speeds are highly uncertain. Compared to the original ensemble shown in Figure 2.3, the perturbation clearly improves the rank histograms for wind speed and temperature, with much higher external bars (the scales on the left and right figures are not the same). It does not correct, however, a tendency to overestimate the wind speed (high left bar). For rain, the bias

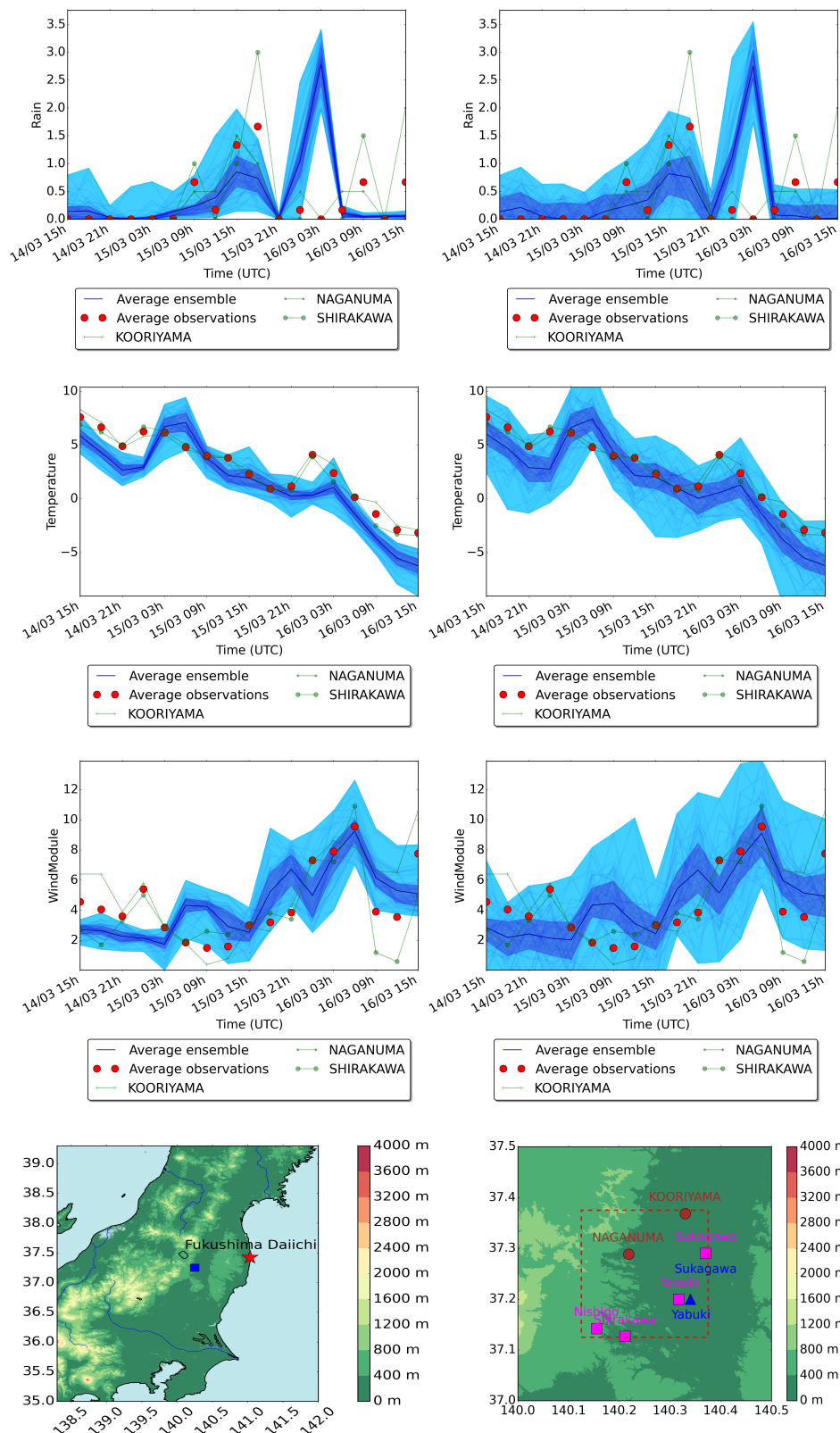


FIGURE 2.4 – Rain [mm], temperature [°C] and wind speed [m/s] in a grid cell on March 14-16. The left column shows the original ensemble and the right column shows the perturbed ensemble. In the last line, the map on the left shows the location of the cell. The map on the right is a zoom on this cell. The red points are the meteorological stations, the pink points are the concentration stations and the blue points are the dose-rate stations. Note that Shirakawa is both a meteorological station and a dose-rate station.

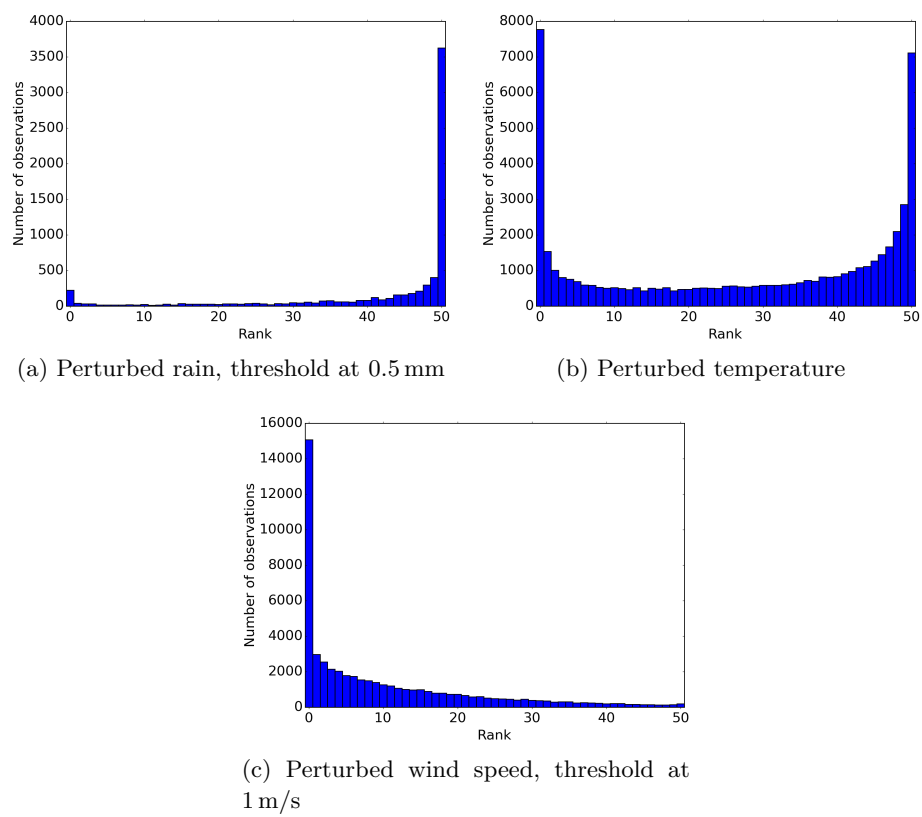


FIGURE 2.5 – Rank histograms of perturbed ensembles.

toward underestimation shown in Figure 2.3 is still higher with the perturbed ensemble, whose rank diagram is not improved.

Clearly, taking into account the observation errors described in this section is not sufficient to make a large difference on the ensemble spread for boundary layer variables. In the remaining part of this study, the original ensemble is used, without any kind of perturbation. While the ensemble described in this section clearly shows its limitations as far as the representation of boundary layer meteorological variables is concerned, it is representative of what could be used in case of emergency with regard to operational constraints. The goal of the next sections is to propagate these uncertainties, along with others, through an atmospheric dispersion model and determine whether this might nonetheless provide a first assessment of uncertainties for operational purposes. A possibility is that the uncertainties on plume trajectories may accumulate over time and/or space, and this, combined with source term and dispersion model uncertainties, might compensate for the limitations of the meteorological ensemble itself.

2.3 Monte Carlo simulations

2.3.1 ℓdX description

ℓdX is an Eulerian transport model that solves advection-diffusion equations over a fixed 3D-grid. The original model comes from Polyphemus/Polair3D which is an air quality modeling system. IRSN developed this model by adding criteria to take into account radionuclide characteristics. The advection-diffusion equation of this model for a precise radionuclide, noted by index r is written as follows :

$$\frac{\partial c_r}{\partial t} + \text{div}(\mathbf{w}c_r) = \text{div}\left(\rho\mathbf{K}\nabla\frac{c_r}{\rho}\right) - \mathbf{F}\mathbf{c} + E_r - \Lambda c_r \quad (2.2)$$

where c_r is the concentration in the air of radionuclide r , $\mathbf{c}$ the vector of concentrations of all radionuclides, $\mathbf{F}$ the matrix of decay coefficients linking these radionuclides, $\mathbf{w} = (w_u, w_v, w_z)$ the wind velocity in dimension three, ρ the air density, $\mathbf{K}$ the turbulent diffusion matrix, this matrix is diagonal with diagonal (K_u, K_v, K_z) , E_r the source term and Λ the scavenging coefficient.

At each simulation, ℓdX uses a weather forecast and a source term to calculate the concentration in the air of radionuclides. At ground level, the deposition is also calculated using boundary conditions. The deposition is the amount of radionuclides deposited on the ground, speed and amount of deposition depend on weather condition, the type of soil and vegetation, etc. From these two elements, concentration in the air and deposition on the ground, we can calculate the dose rate of gamma radiation. This measure makes it possible to estimate the radioactive consequences on human health and environment.

2.3.2 Simulation setup

Once all sources of uncertainties have been identified, we sample them according to given probability distributions. The Monte Carlo ensemble contains 200 simulations and is carried out on ten parameters, detailed in Table 2.2. For each simulation, a meteorological member and a source term are randomly chosen in the weather ensemble and the six source terms,

with a uniform distribution. For the other parameters, we have reused the same variations as proposed by [Girard 2014]. The parameters are not all perturbed in the same way, some are added a random value, some are multiplied by a random value, and others are replaced by a random value. All parameters are perturbed independently. The six source terms used are from the literature. Almost ten years after the accident, no reliable source term has been produced by relying only on the reactor physics and on-site events. All source terms in the literature are based on more or less sophisticated inverse or reverse modeling techniques, which combine environmental observation data, meteorological and dispersion modeling. An account of the different source terms available in the literature and the various techniques and observation datasets used to obtain them is available in [Mathieu 2018a]. Here, six source terms have been retained. Some source terms were not selected, either because they did not significantly differ from another source term (to avoid introducing bias in the ensemble) or because they were not deemed relevant due to the spatial and temporal scales of interest in the study, or because insufficient information was available. Although they are all *a posteriori* source terms based on observations, the variability in the overall released quantities is important. The total release of ^{137}Cs over the whole period ranges from 8 PBq to 20.6 PBq, and the differences are even greater when looking at particular release events, due to large differences in the release kinetics [Mathieu 2018a].

As this variability is representative of the *a posteriori* uncertainty, additional perturbations have been applied on the released quantities (called *emission factor* in Table 2.2) and on the emission release time (called *emission delay* in Table 2.2). The values used for these perturbations come from [Girard 2014]. They are crude perturbations, in the sense that the uncertainty on the released time and quantity is dependent on the release event, and therefore should be time-dependent. Besides, to be representative of *a priori* uncertainties, it would be better to use source terms derived only from installation events, without incorporating information from the observations, if available in the future.

Similarly to [Girard 2014], we made use of *ldX* which is an Eulerian model for radionuclide dispersion described in Section 2.3.1. The simulation domain and grid resolution used by the dispersion model coincide on the horizontal with the discretization of the weather forecasts, that is, a grid cell size of 0.25° horizontal resolution and a domain that spans 40×40 grid boxes to cover a great part of Honshu island. As the vertical resolution of the NWP ensemble is very coarse, the meteorological fields are interpolated on a vertical grid used for dispersion by the *ldX* model. The center altitudes are 20 m, 100 m, 220 m, 340 m, 500 m, 700 m, 1000 m, 1500 m, 2200 m, 3000 m, 3850 m and 4650 m. The model computes every hour and in each grid cell the air activity concentration and deposition values for all radionuclides released, taking into account dry deposition, scavenging processes due to precipitation, radioactive decay and progeny. This yields three-dimensional fields (resp. two-dimensional for deposition) varying in time, during the three weeks simulated. From these variables, the ambient gamma dose rate can be derived.

Variable	Distribution/Method	Variation space
ECMWF ensemble member	Discrete/replace	[member 1, ..., member 50]
Source term	Discrete replace	[[Katata 2015], [Terada 2012], [Mathieu 2012], [Saunier 2013], [Saunier 2016], [Winiarek 2012]]
Emission factor f_e ¹	Log-normal/multiplication	[1/3, 3]
Source elevation [m]	Discrete/replace	[0-40, 40-160, 160-280, 280-400]
Emission delay [hours]	Truncated normal/addition	[-6, +6]
Dry deposition velocity v_d [m/s]	Uniform/replace	[$5 \cdot 10^{-4}$, $5 \cdot 10^{-3}$]
Scavenging factor a [$h s^{-1} m m^{-1}$] ²	Uniform/replace	[10^{-7} , 10^{-4}]
Scavenging exponent b^2	Uniform/replace	[0.6, 1]
Horizontal diffusion K_h [$m^2 s^{-1}$] ³	Uniform/replace	[0, 1.5×10^4]
Factor of vertical diffusion f_{Kz} ⁴	Uniform/multiplication	[1/3, 3]

TABLE 2.2 – Variation space of ℓdX input variables.

2.3.3 Description of the measurements

A review of the radiological observations available for the Fukushima case, with their advantages and drawbacks, was made by [Mathieu 2018a]. Here, we briefly describe the datasets that are used for our comparison purposes.

Air activity concentration

In this paper, we make use of the hourly concentrations of ^{137}Cs retrieved by [Oura 2015]. The air activity concentration is given in Bq m^{-3} . The data was obtained from the air quality automated monitoring network measuring suspended particulate matter (SPM) on filter tapes. They were measured too late to detect short-lived radionuclides, but they give crucial information on the temporal variation of ^{137}Cs concentration close to ground level, therefore indicating the passage of different plumes. Figure 2.6 shows the spatial distribution of the 108 stations; unfortunately the repartition is not uniform, especially over the mostly contaminated area (in

-
1. $E_r^{\text{perturb}} = f_e \cdot E_r$
 2. Scavenging coefficient : $\Lambda_r = a p_0^b$, where p_0 is the rain intensity [$mm.h^{-1}$]
 3. K_h is computed following K_u and K_v .
 4. $K_z^{\text{perturb}} = f_{Kz} \cdot K_z$.

grey).

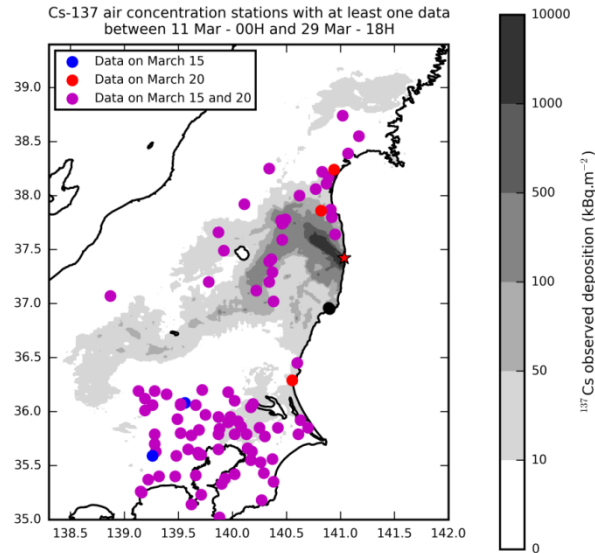


FIGURE 2.6 – Position of activity monitoring stations for ^{137}Cs and their availability periods. From [Quérel 2016]. In grey, the area where ^{137}Cs deposition is higher than 10 kBq m^{-2} .

Deposition

Among the long-lived species, ^{137}Cs contaminated abundantly the Japanese territory and was measured during several airborne and ground measurement campaigns. The airborne measurements obtained from the various surveys conducted jointly by Japanese authorities (MEXT) and the U.S. Department Of Energy (DOE) were calibrated against the ground readings to map the total ^{137}Cs deposit over Honshu island [Sanada 2014]. The values were corrected from the radioactive decay to represent the deposition on April 1st 2011, after the end of all significant releases. While this kind of measurements gives a time-integrated information, there is a very fine spatial coverage (a few hundreds of meters). Here, the observations were averaged over a mesh of resolution 1 km, which amounts to more than 30000 points. Figure 2.7 shows the corresponding deposition map (values below 10 kBq m^{-2} were not measured).

Gamma dose rate

The gamma dose rate measurements were provided by automated stations as early as March 11st, with a 10-minutes time step. The majority of dose rate measurements are available in the IAEA database [IAEA 2012]. There are 88 stations spread over Japan, although the spatial coverage is heterogeneous (Figure 2.8). The dose rate readings have the advantage of measuring the contribution from all radionuclides, including the short-lived species that could not be detected by the other monitoring systems. It is composed of two parts : the direct plume contribution (“cloud-shine”) and the gamma-ray emitted by radionuclides deposited on the ground (“ground-shine”). Hence, the cloud-shine is usually responsible for peak values observed

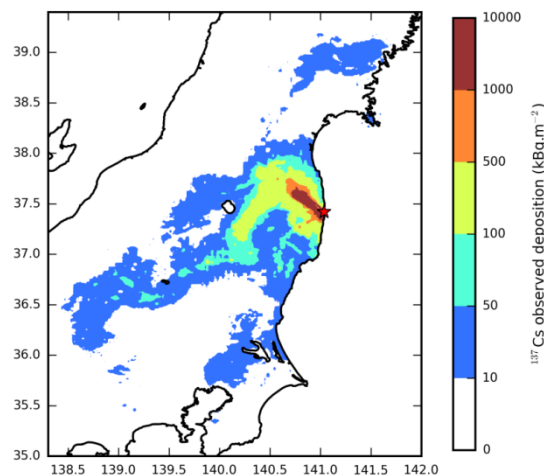


FIGURE 2.7 – Observed deposition of ^{137}Cs established for April 1st 2011, [Quérel 2016].

during the plume passage, whereas ground-shine corresponds to a lasting contribution that lingers after the plume has left the area, and decreases because of the radioactive decay.

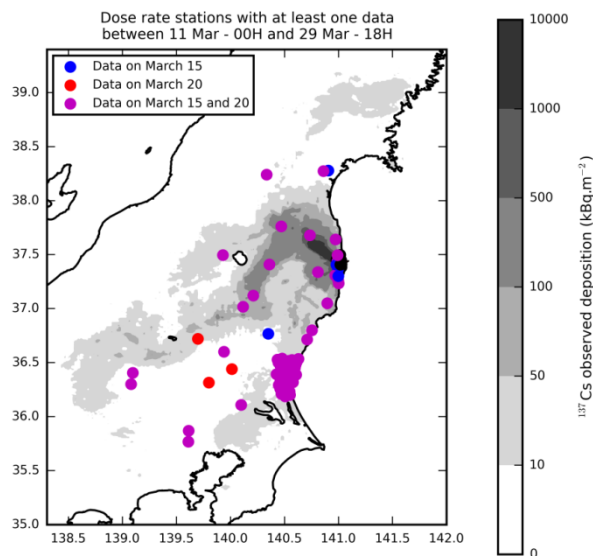


FIGURE 2.8 – Position of dose rate monitoring stations for ^{137}Cs and their availability periods. See [Quérel 2016]. In grey, the area where ^{137}Cs deposition is higher than 10 kBq m^{-2} .

2.3.4 Model-to-data comparison for Monte Carlo simulations

The 200 simulations from the Monte Carlo sampling described in Section 2.3.2 are now compared to the observation data detailed in Section 2.3.3. For this purpose, the 3-D ^{137}Cs air concentration (resp. gamma dose rate) fields given by each member of the ensemble were

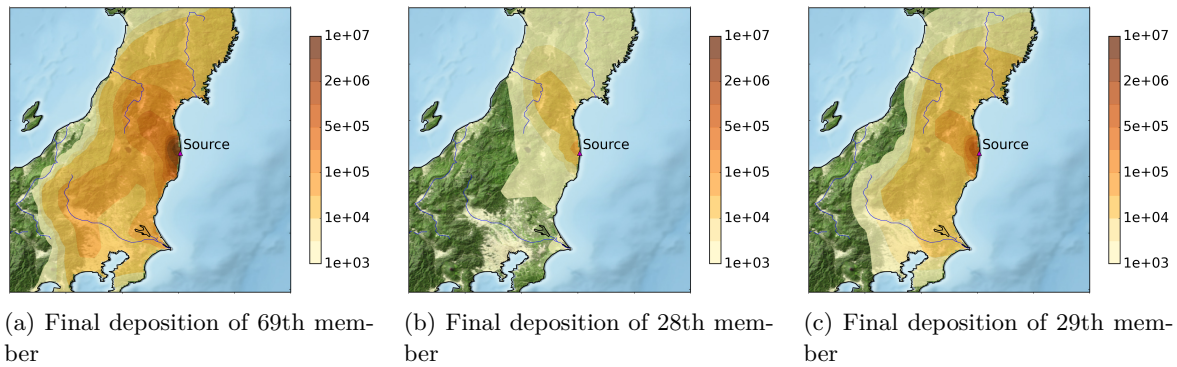


FIGURE 2.9 – Illustration of the final deposition of ^{137}Cs , computed by three ensemble members, in Bq m^{-2} .

interpolated on the corresponding station locations. As far as deposition was concerned, the 2-D ^{137}Cs deposition maps at the final time step of the simulation were compared to the observation map.

Deposition

Pollutants can be removed from the air during the transport by various processes. Deposition concerns the transfer from airborne materials, gases or particles, to the soil, water and vegetation. When precipitation is absent, this process is called dry deposition, otherwise, it is called wet deposition or scavenging. During the Fukushima accident, significant deposition of ^{137}Cs was measured in Honshu island. A substantial proportion comes from wet deposition, especially during the two contamination episodes of March 14-16 and March 20-22 [Korsakissok 2013]. In the Monte Carlo ensemble, some simulations provide a deposition map comparable to observations, as illustrated by Figure 2.9a. Other simulations show a lower deposition, either because the simulated plume was carried away over the ocean or since precipitation did not occur during the passage of the plume, as in Figures 2.9b and 2.9c.

Rank histogram The rank histograms are built without observed deposition values below the threshold 5 kBq m^{-2} . According to figure 2.10a, the ensemble shows a bias as it tends to underestimate the observations. The rank map 2.10b indicates the overestimated (blue points) and underestimated (red points) areas. According to [Mathieu 2018a], the long red branch, first north-west and then south-west, was mainly contaminated during March 14-16. The presence of precipitation and fog, together with a concentrated plume, caused a strong deposition that most simulations failed to reproduce. In particular, fog and light rains are not well predicted by weather forecasts and these uncertainties may not be well represented by the meteorological ensemble. Besides, overestimation in the north-west (the blue area) and underestimation in the south-west (the little red area) can be due to the subgrid effect. The plume came into a valley and was wedged there but our ensemble can not reproduce this because of the coarse resolution.

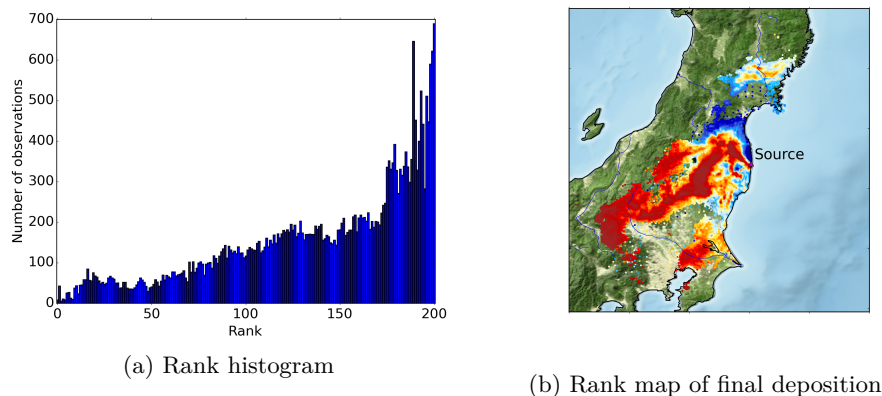


FIGURE 2.10 – Rank study of final deposition of ^{137}Cs over all points with a threshold at 5 kBq m^{-3} . (a) Rank histogram. (b) rank map : the color of each point corresponds to the rank of the observation. The dark blue points are overestimated, the dark red are underestimated.

RMSE The computation of the RMSE on deposition was carried out with all Monte Carlo members without using a threshold. Firstly, it was calculated over all measurement points. The value obtained is much higher than the observational average, as shown in the first line of table 2.3. Secondly, it was calculated over the cell average, i.e. model-to-data error is then the difference between the prediction in a cell and the observational mean in this cell. The fraction “Min/Obs mean” decreases from 3.24 to 1.42 because there are not subgrid effects in the second case. The RMSEs of the third line use only the measurement points whose distances from the source are higher than 80 km. The fraction “Min/Obs mean” is even better because the short-range error is eliminated. In the last line, both subgrid effects and short-range discrepancies are filtered out.

	Min	Mean	Median	Obs mean [kBq/m ²]	Min / Obs mean
RMSE all points	263.27	269.05	282.46	81.06	3.24
RMSE all cells (cell average)	67.18	92.36	91.5	47.44	1.42
RMSE points > 80 km	16.69	22.11	21.63	22.15	0.75
RMSE cells > 80 km	10.56	17.49	14.72	15.72	0.67

TABLE 2.3 – RMSE on the final deposition.

FMS—figure of merit in space Figures of merit, like RMSE, are analytic parameters used for evaluating the prediction accuracy in relation to observations. FMS is calculated at a fixed threshold, and it is defined as the amount of overlapping between the observational and predicted areas where values are greater than the threshold. It is given by the ratio between the intersection

of the areas and their union (see Appendix B.3). The value 100 % in FMS represents a perfect cover of prediction over observation. In table 2.4, when the threshold increases, the performance of the model decreases because there are less measurement points. A small shift in deposited direction changes immediately FMS values. Following [Korsakissok 2013], a forecast whose FMS value reaches about 50 % is satisfactory, so our simulation ensemble is acceptable according to that statistical indicator.

Thres [kBq/m^2]	Max	Mean	Median
5	99.16	93.08	82.07
37	57.64	48.20	42.39

TABLE 2.4 – FMS [%] for final deposition at all stations.

Gamma dose rate

Figure 2.11 presents the gamma dose rate measured at two stations, Shirakawa and Kita Ibaraki, as well as the ℓdX predictions by Monte Carlo method in these stations. Shirakawa is a dose rate measurement station in a cell that also includes weather and concentration stations. The abrupt increase informs on the plume arrival. After the peak, the plume departure drives to a decrease of dose rate, but it stays high because of wet deposition. Then, its decreases is due to radioactivity as shown in figure 2.11a. In case of sole dry deposition, there are less radionuclides deposited, so gamma dose rate will decrease brutally after the plume departure, as in figure 2.11b. Shirakawa is a typical example for wet deposition and Kita Ibaraki for dry deposition. Kita Ibaraki is a station on the south-west coast as shown in figures 2.11c and 2.11d.

Rank histogram The triangular shape of the rank histogram of gamma dose rates shows that the ensemble spreads out largely in comparison to the observations. A threshold of $0.1 \mu Sv h^{-1}$ was applied, to account for background noise. Contrary to the rank histogram of deposition, the rank histogram of gamma dose rate has low bars on the left, as shown Figure 2.12. This shape suggests that the ensemble is over-dispersed by comparison to this output variable. Each gamma dose rate station provides many measurements, but they mostly bring similar, correlated information. After the plume has left the area, the only difference between two consecutive measurement points comes from the radioactive decay of the deposited radionuclides. Therefore, in that case, most observations get similar ranks, as it is the case at Shirakawa station (see Figure 2.11a). This issue creates peaks in the rank histogram.

RMSE The RMSE of the gamma dose rate is much better than the one of air activity as shown in table 2.5. The subgrid effect and the delay in time affect less substantially the RMSE for dose rate than for air concentration. Indeed, concerning the subgrid effect, the measurement stations of gamma dose rate can detect gamma radiation within a few kilometers, while a measurement of a concentration represents the concentration levels in the close vicinity of the station. Concerning the time delay, wet deposition took place at most of stations and the dose rate decreases slowly after wet deposition. Thus, the time delay after the passage of the plume is not visible. For stations at more than 80 km from the source, the ensemble minimum, average,

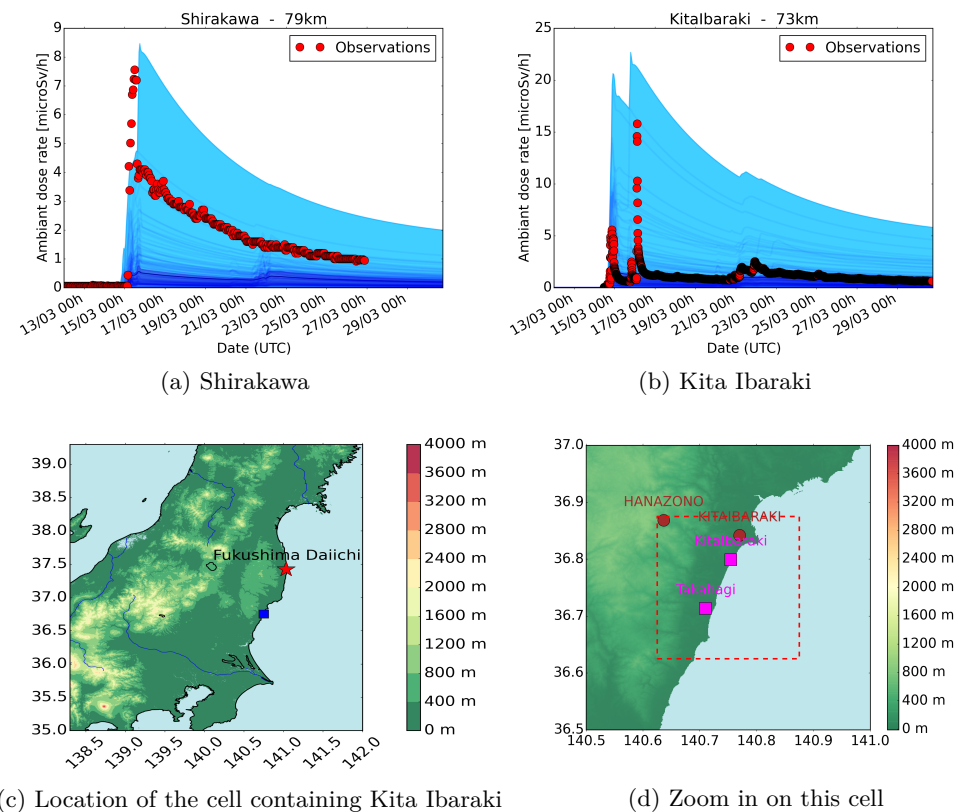


FIGURE 2.11 – Illustration for gamma dose rates in Shirakawa (wet deposition) and Kita Ibaraki (dry deposition). The blue strip indicates the stretch of ℓdX simulations, the red points correspond to measurement of gamma dose rate. Two bottom figures show the location of the cell where these two stations are located, and include a zoom on this cell.

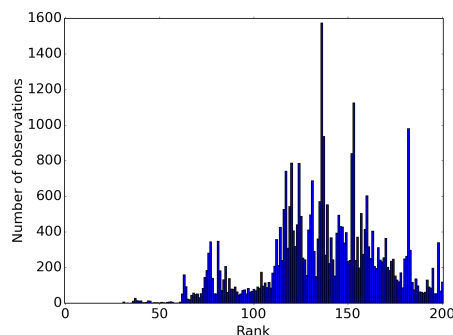


FIGURE 2.12 – Rank histogram of gamma dose rate over all measurement stations with threshold at $0.1\mu Sv/h$.

median and the observational mean are similar. The relational fraction “Min / Obs mean” decreases significantly when we remove the stations at less than 80 km from the source, where the numerical diffusion due to the grid resolution is too high.

	Min	Mean	Median	Obs mean	Min / Obs mean
All stations	1.85	2.86	2.19	0.68	2.72
Stations > 80 km	0.33	0.34	0.39	0.34	0.97
Stations < 80 km	4.43	6.97	5.3	2.4	1.84

TABLE 2.5 – RMSE for gamma dose rates.

FMT—figure of merit in time In the same way as FMS, the computation of FMT is based on a threshold. It represents the ratio of overlap between prediction and observation over time (see the definition in appendix B.4). The table 2.6 confirms that the ℓdX results over the area beyond 80 km are better than those over entire Japanese territory or closer to the source. The FMT value of the ensemble average is similar to the maximum FMT, and 58% is the result of most ensemble members. Following [Korsakissok 2013], this is a satisfactory value.

	Max	Mean	Median
All stations	41.1	33.09	28.22
Stations > 80 km	58.83	58.61	42.30
Stations < 80 km	35.74	23.40	21.17

TABLE 2.6 – FMT [%] for gamma dose rate.

Air activity concentration

Generally, two principal episodes of contamination were observed and will be studied here : March 14-16 and March 20-22 [Mathieu 2018a]. During the first period, the observational amplitude was correctly represented by the ensemble but at some stations, a delay in time and a discrepancy in the duration was observed. The highest values of most stations are included in the ensemble envelop. Some members largely overestimate the observations. However, at 57 out of 108 stations, the plumes simulated by all members do not stay long enough to cover all observations, as illustrated by Figure 2.13. All plumes modeled by the ensemble arrive too late at 22 stations. For the second episode, simulations tend to underestimate observations and a delay in time unfolded for several stations. At 39 stations out of 108, the ensemble envelop does not reach the highest observed values. The plumes arrive late at 32 stations.

In Figure 2.13, two measurement stations of air concentration are shown : Sukagawa and Yabuki. These stations are located in the same cell, about 70 km in the south-west of the Fukushima plant (Figure 2.4). This figure is a good illustration of the subgrid effect. ℓdX computes the cell average of air concentrations. Hence, the simulations give the same result at these two stations. However, the mesh used is coarse, and the plume does not fill entirely the cell, as illustrated by Figure B.1 in Appendix B.2. During the first episode, Yabuki station

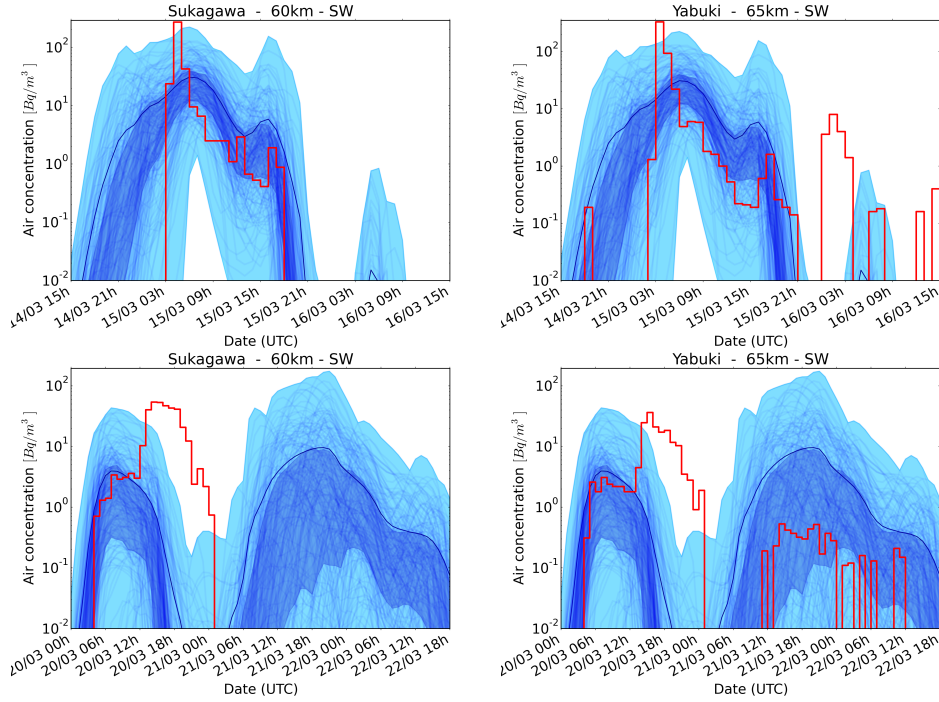


FIGURE 2.13 – Model-to-data comparison of air concentrations at two stations located in the same cell : Sukagawa (left) and Yabuki (right), during two episodes : 14-16 March 2011 (top) and 20-22 March 2011 (bottom). The red line is the observation. The simulations are represented by the light-blue lines. The dark blue line is the average of the simulations.

detected several plumes whereas Sukagawa station observed only one plume from 03 :00 to 21 :00 on March 15th. The same phenomenon took place during the second episode. Therefore, stations inside the plume give measurements that differ from those given by stations outside the plume, whereas the simulations are identical. This can produce a significant model-to-data discrepancy.

Rank histogram Figure 2.14 presents a rank histogram for air concentrations. This histogram does not use all measurements. We selected stations whose distances from the source are greater than 80 km because ℓdX is a long-range model. The resolution used is about $25 \text{ km} \times 25 \text{ km}$ and the model results are reliable only beyond about three cells from the source, to account for unrealistic numerical diffusion close to the source. We also put a threshold at 1 Bq m^{-3} in order to discriminate stations with a significant signal. This rank histogram is U-shaped with two external bars that are very high. The left bar indicates some overestimation of the ensemble, probably because of the subgrid effect. The right bar corresponds to some underestimation, probably because of the delay described in the previous paragraph. The concentration measurements are very difficult to reproduce because of the subgrid effect, which leads to a rather bad rank histogram.

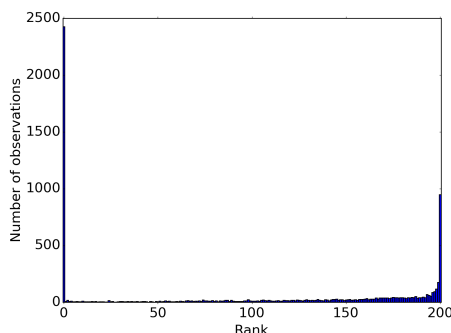


FIGURE 2.14 – Rank histogram of air concentrations in March 2011.

RMSE—Root mean square error The RMSE is a statistical indicator widely for model-to-data comparison (see the definition in Section B.5). Table 2.7 presents for air concentrations : the best RMSE among all ensemble members, the RMSE of the ensemble average and that of the median. Overall, model performance is acceptable when RMSE and observational mean have the same order of magnitude. Compared to the observational average, these RMSE are too high, which may be explained by the delay and the subgrid effect presented in the previous paragraphs.

Thres.	Min	Mean	Median	Obs. mean	Min / Obs mean
0.1 Bq/m^3	13.43	18.76	14.03	3.49	3.84
1 Bq/m^3	15.18	21.21	15.86	4.38	3.47
10 Bq/m^3	20.87	29.24	21.83	7.14	2.92

TABLE 2.7 – RMSEs on air activity concentrations of ^{137}Cs during entire March 2011, for the best RMSE among all ensemble members (Min), the RMSE of the ensemble average (Mean) and that of the median.

2.4 Synthesis and perspectives

Uncertainties carried by the weather forecasts influence very substantially the results of atmospheric dispersion models : air concentration activity, deposition and gamma dose rate. Among all meteorological variables, wind direction and speed, and precipitations are obviously two major factors that influence the plume trajectory and subsequent contamination patterns. The wind direction determines direction and spread of plumes, and precipitation has a strong influence on the deposited quantities. The model-to-data comparison to meteorological observations provided a first insight into the suitability of weather ensemble forecasts as a descriptor of uncertainties related to these boundary layer variables. In this study, the operational ECMWF-ENS forecast was retrieved and used for the three weeks of the Fukushima disaster (12-30 March 2011). This raised several issues. First, the horizontal and vertical resolution of this ensemble, as stored by ECMWF, is quite coarse, with a 0.25°

horizontal grid resolution and only five pressure levels within the boundary layer. This in turn raises the concern of subgrid variability which will create discrepancies with local variables at meteorological stations. Another issue was the fact that the ECMWF ensemble prediction system is tuned to be representative of large-scale errors. This spread is not sufficient to represent the uncertainty within the boundary layer height, as illustrated in this paper by the rank diagrams for meteorological variables. Finally, another challenge was to combine different ensemble forecasts from various initial times to span the three-week period of the accident. This was done by simply juxtaposing successive 24-hour forecasts. This may be viewed as representative of what could be used operationally in case of an emergency, as successive 24-hour forecasts would be used to infer or complement recommendations for the protection of the population, as the accidental sequence develops. However, this leaves clear room for future research in order to ensure a smoother transition between different ensembles. The ensemble performance shown in this study is therefore representative of the quality of an ensemble that could realistically be used for operational purposes. This paper also proposed a simple way to account for observation errors in the ensemble evaluation. Perturbing the weather forecast ensemble following given probability densities allowed to take into account the observational errors for some important variables. Although wind speed and temperature perturbations allowed to significantly improve the weather ensemble rank diagram, the results for precipitation and wind direction still showed significant under-dispersion. These results highlight the benefit of using an appropriate estimation of the observation errors. Clearly, errors for precipitation still have to be refined.

The other question addressed in this study was then to what extent this ensemble may be used, despite its limitations, to evaluate the uncertainties in the dose evaluations made by atmospheric dispersion models. The study of [Girard 2014] ascertained that three uncertainty sources are meteorological forecasts, source term and some parameters used inside the model. The perturbation by Monte Carlo was carried out over ten input variables of ℓdX . Among those, the weather forecasts and the source term are uniformly selected in discrete sets. Other variables follow different probability distributions. Two hundred perturbed data samples were generated, and as many dispersion simulations were computed. Observations of air concentrations, deposited quantities and gamma dose rates allowed us to evaluate this simulation ensemble with several statistical indicators : rank histogram, RMSE, FMS and FMT. The removal of the area within 80 km of the source confirms the ℓdX effectiveness for long distances.

In conclusion, the aim of this study was to implement a method of perturbation on inputs and analyze the uncertainty propagation in ℓdX . In the weather ensemble assessment, we understood that the coarse resolution of the forecasts was responsible for a significant part of the model-to-data discrepancy, because of the subgrid effect. With the Monte Carlo perturbations, the simulations' output spread was much larger than what was observed in the meteorological ensemble. Although the ensemble did not prove sufficient to represent uncertainties for air activity concentrations, the results encompassed reasonably well the deposition and gamma dose rate observations despite the resolution. One of the questions when dealing with ensemble simulations, especially in an emergency context, is the computational time. Therefore, the use of a coarse-resolution meteorological ensemble may be an acceptable choice, compared to a high-resolution ensemble which may improve the accuracy but would significantly increase the

computational burden. To take into account the subgrid effect, suitable perturbations may be studied in the future to compensate for this uncertainty not taken into account in our ensemble.

A next step could be to improve the perturbations on the key inputs so that the ensemble better fits the observations. The comparison should rely on probabilistic scores, like the Brier score or the rank histogram. Several statistical methods are available to carry out an ensemble calibration. Among them are Bayesian inference schemes, for instance, based on Monte Carlo Markov chain (MCMC) algorithms. This approach makes use of observations to calibrate the model inputs and samples from the a posteriori probability distribution of each input variable. However, this approach requires a huge number of simulations, of the order of several millions. It is not reachable with current models. A future study should therefore focus on meta-models' construction. A meta-model is a mathematical approximation of some model but with a much lower computational time. The results of the meta-model must be similar to those of the model, so that they can replace the model in the MCMC algorithms.

Taking into account the errors of radiological observations for the ensemble evaluation would still be necessary, and they may be taken into account in the same way as for the weather observations in section 2.2.2. A more complex issue, that we raised in this study, is the redundancy in the observational data. When two stations are close, they may essentially bring the same information. This may also happen at one given station because of high correlation in time, such as gamma dose rate stations when they are measuring the radioactive decay of the surrounding deposition.

Métamodélisation du modèle de dispersion à longue distance ldX

Contents

3.1	Introduction	71
3.2	Creation of DOE and model's output	72
3.2.1	Input : Physical parameters	72
3.2.2	Input : Source term	73
3.2.3	Input : Weather forecast	74
3.2.4	Outputs' emulation	75
3.3	Construction of metamodels	75
3.3.1	RBFs—Radial Basis Functions	76
3.3.2	Kriging—Gaussian processes (GP) modeling	76
3.4	Metamodels' evaluation	77
3.4.1	RBFs vs Kriging	78
3.4.2	Studying error predicted by GP models	80
3.4.3	Comparing metamodels with different meteorological perturbations	82
3.5	Applications of metamodels	85
3.5.1	Sensitivity analysis	85
3.5.2	Minimization of the RMSE on concentration	85
3.6	Conclusions and perspectives	89

3.1 Introduction

Dispersion models are evolving to meet new requirements in the context of nuclear risk assessment and emergency response systems. In case of unintended emission of radionuclides, these models are currently utilized within the French Institute for Radiation Protection and Nuclear Safety (IRSN), to track the transport of radioactive plumes and to prepare protection actions. The models are subject to large uncertainties that require specific assessment, propagation and possibly reduction using observational data.

In March 2011, a tsunami arrived on the East coast of Japan and created a nuclear disaster in Fukushima-Daiichi. To date, many simulations were carried out by worldwide establishments, but they were subject to errors, sometimes during critical phases of the plume dispersion. The simulation approach was deterministic, but the considerable uncertainties would point

toward more stochastic approaches. There are indeed three main sources of uncertainties in atmospheric dispersion modeling : weather forecasts, emission source terms and shortcomings in the numerical model. Taking into account these uncertainties is crucial in order to increase the reliability of decisions.

Propagating the uncertainties in a dispersion model is costly because of the complexity and computation cost of a single numerical simulation. Classical approaches relies on multiple simulations, which may be intractable in practice. Ideally, the aim is to determine the PDFs of the inputs (through a calibration) and compute the PDFs of the outputs (uncertainty propagation). Calibration methods require very large number of simulations (several thousands or a few millions). For the Fukushima accident, current dispersion models spend about 7 minutes for a single run. To circumvent this computational cost, we built an approximation of the numerical model, called metamodel. The metamodeling relies on emulation through interpolation in high dimension, either using Radial Basis Functions (RBFs) or Kriging (Gaussian processes).

The dispersion model used in this study is ℓdX , which is part of the platform *C3X* developed by IRSN for emergency cases. This paper is structured into five parts : part 3.1 explains the context and subject of the study ; part 3.2 shows how to model the input uncertainties to create a design of experiment (DOE) ; the construction and evaluation of metamodels will be respectively explained in details in parts 3.3 and 3.4 ; part 3.5 will talk about an application of metamodels in our uncertainty study ; and the last part 3.6 will summarize and give some perspectives.

3.2 Creation of DOE and model's output

The construction of the metamodel requires a training set of several simulations of the physical model. These simulations must be carried out cleverly so that the maximum amount of information on the model's behavior can be retrieved with as few simulations as possible. In other words, the input data-sets must be organized to best cover the model's inputs space. These data-sets form a series of experiments, called design of experiment (DOE), and relevant information for the training will be found at the output of each simulation. This section focuses on creating a design of experiment for uncertain parameters of ℓdX . The modeling of scalar parameters followed a Latin hyper-cube sampling (LHS). The inputs were perturbed in the same way as [Le 2020], except that the variation interval of the scavenging factor was modified. For the high dimension inputs (source term and weather forecast), we did not perturb in the same way as [Le 2020] ; the uncertainty modeling of these inputs will be detailed in sections 3.2.2 and 3.2.3.

3.2.1 Input : Physical parameters

Besides the source term and the meteorology, there are eight uncertain scalar parameters in the model ; see the list in table 3.1. The sampling space of LHS is therefore a eight-dimensional hyper-cube, cf figure C.1. The parameters are considered independent. There are different methods to perturb the parameters. A few parameters were perturbed by adding a random value ; others were multiplied by a random value, and the others were entirely replaced by a random

value. The variation ranges were determined by using typical variation ranges, either used in practice of simulations, or found in bibliographical review :

- Dry deposition velocity : [Maryon 1991], [Rodos 1999], [Baklanov 2001], [Brandt 2002];
- Horizontal diffusion : [Girard 2016], [Brandt 1998], [Ryall 1998], [Sørensen 1998], [Yamartino 2000];
- Vertical diffusion : [Louis 1979], [Troen 1986], [Hanna 2001].

Variable	Method	Variation range
Emission delay [<i>hours</i>]	addition	$[-6, +6]$
Source elevation [<i>m</i>]	replace	$[0 - 400]$
Dry deposition velocity [<i>m/s</i>]	replace	$[5.10^{-4}, 5.10^{-3}]$
Scavenging factor [$hs^{-1}mm^{-1}$]	replace	$[10^{-7}, 10^{-4}]$
Scavenging exponent	replace	$[0.6, 1]$
Horizontal diffusion [m^2/s]	replace	$[0, 1.5] \times 10^{-4}$
Vertical diffusion	multiplication	$[1/3, 3]$
Release factor	multiplication	$[1/3, 3]$

TABLE 3.1 – Variation ranges of physical parameters.

3.2.2 Input : Source term

The second source of uncertainties is the source term, which consists of time series of released quantities, used as input data. For the Fukushima nuclear disaster, several establishments built their source terms using inverse methods. The release time, type of radionuclides and quantity were determined by using the radioactive measurements collected on the Japanese territory. In this study, we use nine principal source terms, which come from the Fukushima literature : [Katata 2015], [Terada 2012], [Mathieu 2012], [Saunier 2016], [Saunier 2013], [Stohl 2012], [Winiarek 2012], [Winiarek 2014] and [Yumimoto 2016].

Each source term has their own strengths and limitations. For instance, a source term inferred from deposition measurements gives a deposition result better than another inferred from concentration observations. In the study of uncertainty propagation [Le 2020], for each simulation, we randomly chose a source term among the nine source terms. In the present paper, the construction of the metamodel is based on the interpolation between the points of DOE. Hence we do not select a single source term, but prefer to combine them linearly with nine weights that are positive and sum up to 1. In other words, a new source term is a convex combination of the original source terms, with weights $(\omega_1, \dots, \omega_9)$ so that

$$\forall i \in \{1, \dots, 9\}, 0 < \omega_i < 1 \text{ and } \sum_{i=1}^9 \omega_i = 1. \quad (3.1)$$

To produce these coefficients, we used the Dirichlet distribution; see appendix C.2. This distribution is configured by a vector of nine positive numbers. The parameter $\alpha = (1, 1, \dots, 1)$ means that the generated points are uniformly shared out in space. High values of α correspond to heavy concentration of points in the center of the space, low values of α mean that the

distribution has higher probability density in the vicinity of the corners and boundaries of the space.

3.2.3 Input : Weather forecast

Weather forecast is the third source of uncertainties we considered. A meteorological data fields consist of several variables in 2D or 3D varying in time. Several studies proved that some meteorological variables have a significant impact on model outputs [[Draxler 2015]]. For instance, wind decides the direction of plumes, and rain determines the quantity of ground deposition. So, modeling uncertainties of weather forecasts is essential to improve the trustworthiness of our dispersion simulations. Weather services produce not only one simulation but an ensemble of forecasts, which is supposed to be representative of the uncertainties. The European Centre for Medium-Range Weather Forecasts (ECMWF) provides an ensemble of fifty member that we exploited for the Fukushima nuclear disaster. The horizontal resolution is 0.25° (about 25 km) with a three-hour time step. These weather forecasts were carried out on a 40×40 horizontal field with 12 vertical levels, whose center altitudes are : 20 m, 100 m, 220 m, 340 m, 500 m, 700 m, 1000 m, 1500 m, 2200 m, 3000 m, 3850 m and 4650 m.

For the same reason as for the source term, the weather forecasts were not randomly selected, but instead linearly combined with random weights. Two perturbation schemes may be employed : linear convex combination (as for the source term in the paragraph 3.2.2) and perturbation of the ensemble spread around the mean. In the first option, if all prediction members are used (fifty members), the input space is too large, and the learning process will require a huge number of experimental simulations. To circumvent this issue, we selected a few ensemble members such that the standard deviation of the reduced ensemble and the entire ensemble remain similar (e.g., 5% of original deviation), cf C.3. Twenty members among the fifty members were chosen, and the creation of the DOE for the meteorological part took then place in the same way as for the source term. For each value in space and time, and for each meteorological variable, the same convex coefficients were applied.

This paragraph details the second suggestion. In this case, only wind and rain were perturbed, other variables were taken as the ensemble mean. The new wind (respectively rain) forecast was determined as follows :

$$m_{new} = \mu_{ens} + k\sigma_{ens}$$

where μ_{ens} and σ_{ens} are respectively the mean and standard deviation of the ensemble (over the fifty members). k is the perturbation factor that will change and be included in the metamodel input space. The wind prediction consists of zonal and meridional winds ; the same perturbation factor was applied to these two components. So, instead of twenty correlated parameters as with the first perturbation scheme, the meteorological part is now perturbed with only two independent parameters (k^{wind} and k^{rain}). The DOE sampling could be carried out at the same time with the physical parameters, i.e., with the LHS. However, a variation space is required in the LHS algorithm. The distribution of k is chosen to be uniform and centered at 0. The extreme values k^* lead to the values $m_{new}^* = \mu_{ens} \pm k^* \sigma_{ens}$. These extreme values are determined such that m_{new}^* covers as much as possible the original ensemble. In other words, we choose a value of k that minimizes the error between the two ensemble envelopes. We found 3 and 4 as the extreme values for the wind and rain respectively : $k^{wind} \sim \mathcal{U}(-3, 3)$ and $k^{rain} \sim \mathcal{U}(-4, 4)$.

This paper compares metamodels built using the two perturbation schemes of the weather forecasts.

3.2.4 Outputs' emulation

The ldX outputs are all high-dimensional data fields. Activity concentration and gamma dose rate are 3D fields varying in time. Final deposition is computed over a 2D grid. After simulation, these outputs are then interpolated to obtain the estimates at measurement stations. In addition, the observations available for the Fukushima case are rather abundant. A description of these was presented in [Le 2020]. We have access to 108 measurement stations for activity concentration and 88 stations for gamma dose rate. For deposition, the airborne measurement was carried out on Honshu Island with a one-kilometer resolution grid; see table 3.2.

However, emulation methods only apply to scalar outputs. Creating a metamodel for each ldX -output element would not be effective. To circumvent this issue, we emulated values integrated in time at measurement stations (total dose rate, cumulative concentration, etc.) or model-to-observation scores (RMSE, FMT, FMS, etc.). Emulators for the model-to-observation scores can be used for an optimization procedure, as explained in section 3.5. The aim could then be to determine an optimal set of input parameters that minimizes the model-to-observation discrepancy.

Output	Dimension	Measurement information
Concentration	3D $\times$ time	108 stations $\times$ 451 one-hour time steps
Gamma dose rate	3D $\times$ time	88 stations $\times$ 451 one-hour time steps
Deposition	2D	> 30 000 measurement cells of a 0.1° spatial resolution

TABLE 3.2 – ldX outputs' information

3.3 Construction of metamodels

RBFs and Kriging are two metamodeling methods used in this study. They will be detailed in parts 3.3.1 and 3.3.2. Let us denote y the model that we want to emulate. The aim is to construct a mathematical approximation $\hat{y}$, which allows to obtain similar results but with much lower computational time. Let n be the number of simulations of the learning base and also the number of corresponding points in the DOE (here, $n = 10000$); d be the dimension of these points ($d = 33$ or 19 depending on the weather perturbation scheme). We can summarize the learning base as follows :

- The design of experiment is $X_{DOE} = [\mathbf{x}_1^T, \dots, \mathbf{x}_n^T]^T$, $\forall i \in \{1, \dots, n\}$, $\mathbf{x}_i \in \mathbb{R}^d$, then $X_{DOE} \in \mathbb{R}^{n \times d}$.
- The learning set is $\mathbf{y}_{DOE} = [y(\mathbf{x}_1), \dots, y(\mathbf{x}_n)]^T = [y_1, \dots, y_n]^T \in \mathbb{R}^n$, where the $(y_i)_i$ are the output to emulate, hence model-to-measurement scores or scalar values integrated over time.

Emulating y consists in two parts, as shown in equation 3.3. A linear trend, noted f , is first implemented with coefficients that are the solution of a pseudo-linear system. Then, an interpo-

lation, noted z , is carried out on the residual of the interpolation f . This means that z is trained with the residuals $y(\mathbf{x}) - f(\mathbf{x})$. In the following equation 3.3, the β coefficients are computed by the "least squares" approximation method.

$$\hat{y}(\mathbf{x}) = \hat{y}(x^1, \dots, x^d) = f(\mathbf{x}) + z(\mathbf{x}) \quad (3.2)$$

$$= \beta^0 + \sum_{j=1}^d \beta^j x^j + z(\mathbf{x}) \quad (3.3)$$

3.3.1 RBFs—Radial Basis Functions

This method is based on the distance weighting between DOE points and the target point, noted $\mathbf{x}_*$, where we want to estimate the output value z with $\hat{z}$. RBF interpolation seeks an approximation of z in the form :

$$z(\mathbf{x}_*) = \sum_{i=1}^n \omega_i \Phi(\mathbf{x}_* - \mathbf{x}_i) \quad (3.4)$$

where $\Phi(\mathbf{r}) = \phi(\|\mathbf{r}\|)$ is a radial function. There are several types of radial function. In this study, we use the multiquadric function $\Phi(\mathbf{r}) = \sqrt{1 + \|\mathbf{r}\|_2^2}$. Let $\mathbf{z}_{DOE} = [z(\mathbf{x}_1), \dots, z(\mathbf{x}_n)]^T$ be the learning base of residuals, then :

$$\forall k \in \{1, \dots, n\}, \quad z_{DOE(k)} = \sum_{i=1}^n \omega_i \Phi(\mathbf{x}_k - \mathbf{x}_i) \quad (3.5)$$

Since $\mathbf{z}_{DOE}$ is known and we expect $\hat{z} = z$ at the DOE points, the coefficients $(\omega_i)_i$ are determined by solving the following linear system :

$$A\boldsymbol{\omega} = \mathbf{z}_{DOE} \quad (3.6)$$

The matrix A has elements $A_{lk} = \Phi(\mathbf{x}_l - \mathbf{x}_k)$ and is symmetric because Φ is a radial function. The construction of RBFs metamodels relied on the Python package SCIPY. The building time is quite fast : only few seconds to create an emulator.

3.3.2 Kriging—Gaussian processes (GP) modeling

The RBF method applies only to the stochastic part z of the metamodel y . The regression part f is calculated separately. In the Kriging approach, the interpolation part z is seen as a stationary centered Gaussian process, whose covariance between two input points $\mathbf{x}$ and $\mathbf{x}_*$ is written as follows : $Cov(z(\mathbf{x}), z(\mathbf{x}_*)) = \sigma^2 r(\mathbf{x} - \mathbf{x}_*)$, according to [Marrel 2009]. In the following, let Y be the random variable of the function y , the value of interest ; and Z the random variable of the stochastic part z . Also, σ^2 indicates the variance of y and r represents the correlation function between the outputs at $\mathbf{x}$ and $\mathbf{x}_*$.

We define R_{DOE} as the matrix of correlation between the training data points X_{DOE} :

$$R_{DOE(lk)} = r(\mathbf{x}_l - \mathbf{x}_k), \quad \forall l, k \in \{1, \dots, n\} \quad (3.7)$$

and $\mathbf{f}_{DOE} = f(X_{DOE}) = [f(\mathbf{x}_1), \dots, f(\mathbf{x}_n)]^T$ as the regression part on DOE. Under the assumption of a GP model, the response sample $\hat{\mathbf{y}}_{DOE}$ follows a multivariate normal distribution :

$$\hat{\mathbf{y}}_{DOE} \sim \mathcal{N}(\mathbf{f}_{DOE}, R_{DOE}) \quad (3.8)$$

For a new point $\mathbf{x}_*$ outside of the learning base, $\hat{y}_*$ is considered as the estimate of the model y at $\mathbf{x}_*$. The correlation vector on z between DOE points and $\mathbf{x}_*$ is :

$$\mathbf{r}_* = [r(\mathbf{x}_1 - \mathbf{x}_*), \dots, r(\mathbf{x}_n - \mathbf{x}_*)]^T \quad (3.9)$$

The joint probability distribution of output between DOE and the new point is :

$$\begin{pmatrix} \hat{\mathbf{y}}_{DOE} \\ \hat{y}_* \end{pmatrix} = \mathcal{N} \left(\begin{pmatrix} \mathbf{f}_{DOE} \\ f_* \end{pmatrix}, \begin{pmatrix} R_{DOE} & \mathbf{r}_* \\ \mathbf{r}_*^T & \sigma^2 \end{pmatrix} \right) \quad (3.10)$$

By conditioning this joint distribution on the learning sample, the prediction of GP model on the point $\mathbf{x}_*$ is characterized by the following conditional distribution :

$$\hat{y}_{*|\hat{\mathbf{y}}_{DOE}} \sim \mathcal{N} \left(E[\hat{y}_{*|\hat{\mathbf{y}}_{DOE}}], Var[\hat{y}_{*|\hat{\mathbf{y}}_{DOE}}] \right) \quad (3.11)$$

where $E[\hat{y}_{*|\hat{\mathbf{y}}_{DOE}}] = f_* + \mathbf{r}_*^T R_{DOE}^{-1} (\mathbf{y}_{DOE} - \mathbf{f}_{DOE})$ is considered as the prediction value of GP model on $\mathbf{x}_*$, and $Var[\hat{y}_{*|\hat{\mathbf{y}}_{DOE}}] = \sigma^2 - \mathbf{r}_*^T R_{DOE}^{-1} \mathbf{r}_*$ approximates the prediction error.

The metamodels by GP model was established using R package DICEKRIGING. Building a GP model takes about 3 hours but the execution time is around 12 ms to make a call to the metamodel (while a simulation of the physical model takes 7-8 minutes). Now, we will evaluate how the metamodels can replace the original model.

3.4 Metamodels' evaluation

In intensive uncertainty studies, metamodels are created to replace the physical model in order to reduce the computational time. They must satisfy some criteria. A metamodel is considered of quality if it reproduces all main variations of the physical model.

To evaluate the metamodels, several predictions are made at points that are not used during the metamodel fitting process. A comparison model-metamodel score is used to evaluate the difference between the physical model response and the metamodel predictions. Cross validation is the most popular method. An advantage of this method is that we do not need to perform new simulations. A metamodel is built on each subset of the original DOE and the remaining points are used to evaluate the prediction of the metamodel. So for physical models that are very expensive in computational time, this method is a good option. But in our case, the construction time of metamodel is also long (for Kriging) because of the large size of the input space. Moreover, the removal of points from the learning sample can break the optimal structure of LHS and can cause a bad prediction.

In this paper, we produce a new sample set for testing purposes. New simulations are required but only one metamodel is built. The points in the testing set are distributed in the input space in order to fill the areas not covered by the original DOE. This method has been proposed by [Iooss 2010] and this sample set is called complementary LHS, cf figure C.2.

3.4.1 RBFs vs Kriging

At present, we perform the evaluation of metamodels as well as the comparison of the performance between two metamodeling methods. The quality of the metamodel predictor is measured by the *standardized mean squared error* between the original model response and the emulator prediction :

$$SMSE = \frac{\overline{(Y_* - \hat{Y}(X_*))^2}}{Var(Y_*)} \quad (3.12)$$

where X_* is a test sample in the complementary LHS, Y_* and $\hat{Y}(X_*)$ are respectively the response of ℓdX and the metamodel prediction at this point. The discrepancy is averaged over all testing points. The *mean squared errors* are positive measures of the quality of an estimator and a value close to zero informs of a good surrogate. Dividing by the variance of the physical response gives us relative information between the model-metamodel error and the output unit.

This part will not only talk about the metamodel evaluation but also the comparison of two metamodeling methods : RBFs and Kriging. The SMSE score allows to assess the quality of emulators ; see first and second columns of table 3.3. In general, Kriging emulators approximate the physical model better than RBFs ones. This result has been confirmed by a graphical plot

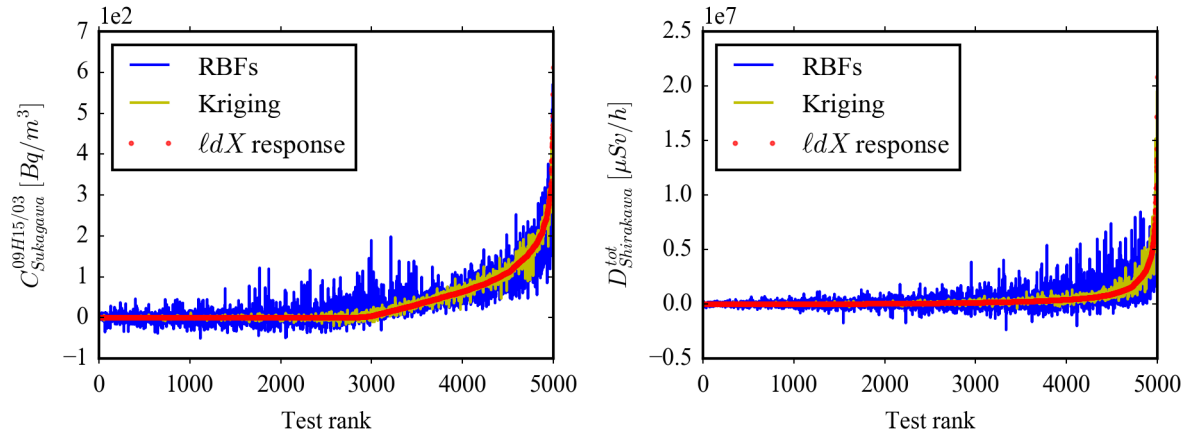
Output	Model-metamodel error		Model-to-measurement error
	RBFs	Kriging	
Concentration	0.155	0.014	3.84
Gamma dose rate	0.677	0.066	2.72
Mean deposition	0.014	0.001	3.24

TABLE 3.3 – Comparison of performance between RBFs and Kriging emulators with SMSE scores.

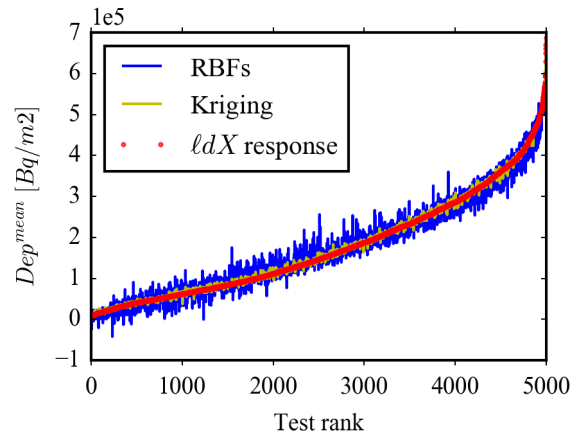
in figure 3.1. It represents the comparison of predictions by RBFs (blue line) and Kriging (red line) to the ℓdX responses (red points). The test is performed on a complementary DOE of five thousand points and the result has been arranged in the increasing order of the ℓdX responses for clarity. Several values of interest were emulated (concentration integrated over time, mean deposition, RMSE of gamma dose, etc.) and the yellow line of Kriging is always closer to ℓdX responses than the blue line of RBFs.

However, the model-metamodel discrepancy is often negligible compared to the model-to-measurement error (third column of table 3.4). In the worst cases, the ratio between the two errors is 24% for RBFs and 2.4% for Kriging. Note that the model-to-measurement deviation was further studied in the article [Le 2020], where an uncertainty propagation was carried out by a Monte Carlo method. In conclusion, the metamodels are generally considered as good approximations of the physical model, especially with Kriging, except emulation with RBFs for Gamma dose rate.

Moreover, Kriging not only emulates well the physical model, it also calculates the prediction error that is given by $Var[\hat{Y}_{*|\hat{Y}_{DOE}}]$. The latter allows to estimate the quality of a prediction result without the exact output of the original model. It will be detailed in part 3.4.2. For RBFs emulators, the prediction is not as good as Kriging but the construction and execution



(a) Concentration at Sukagawa, a measurement station, (b) Gamma dose rate integrated over time at Shirakawa at a given time



(c) Average deposition in space

FIGURE 3.1 – Comparison to ldX responses of RBFs and Kriging metamodels.

time are much shorter. The advantages and drawbacks of these two methods are summarized in table 3.4. In addition, for RBFs, once an emulator is built, we make use of some parameters, A matrix, to accelerate the construction of other emulators, while this will not work with Kriging. In Kriging process, the kernel is optimized by maximum likelihood estimation. Because of the high dimension of the DOE, this optimization could be very long. For the applications found in section 3.5, we use RBF emulation to minimize a model-to-measurement score. Since we need to call the metamodel many times, RBF which has a short execution time is more suitable.

	RBFs	Kriging
Construction time for one emulator	A few seconds	About 3 hours
Execution time for one thousand calls	2.5 seconds	12 seconds
Quality	Acceptable	Good
Reusing some construction parameters	Yes	No
Computing the prediction error	No	Yes

TABLE 3.4 – Advantages and disadvantages of RBFs and Kriging emulations.

3.4.2 Studying error predicted by GP models

According to the paragraph 3.3.2, the GP model provides, in addition to the predictor, the prediction error, also called the MSE variance. This variance does not depend on the estimated output but on the covariance function r and the distribution of the DOE samples. It makes it possible to visualize the areas where the estimation is inaccurate and thus to estimate the quality of the GP model.

This section focuses on a relationship analysis between the prediction error estimated by Kriging, called $\sigma_{Kriging}$, and the model-metamodel discrepancy, $\varepsilon_{\ell dX-Kriging}$. The objective is to evaluate the reliability of this prediction error. Following the construction of the GP model, this prediction error must follow a centered normal distribution with the variance $\sigma_{Kriging}^2$, hence $\varepsilon_{\ell dX-Kriging} \sim \mathcal{N}(0, \sigma_{Kriging}^2)$. Thus, it can be deduced that the fraction $\frac{\varepsilon_{\ell dX-Kriging}}{\sigma_{Kriging}}$ theoretically follows the standard normal distribution, $\frac{\varepsilon_{\ell dX-Kriging}}{\sigma_{Kriging}} \sim \mathcal{N}(0, 1)$.

The figure 3.2 illustrates the relation between the error predicted by Kriging and the error committed by this metamodel according to three kinds of ℓdX outputs : concentration, gamma dose rate and deposition. In the same way as for the metamodels' evaluation, prediction tests were done on the complementary LHS whose samples are well distinguished from those of the DOE. The tests are arranged in ascending order according to the MSE, and the figures are showed on a logarithmic scale to facilitate observation. The red line represents the prediction error estimated by Kriging. The blue line is the deviation between metamodel predictions and the response of the physical model. In general, the error predicted by Kriging is greater than the model-metamodel discrepancy. To confirm this interpretation, we have also plotted the histograms of the ratio $\frac{\varepsilon_{\ell dX-Kriging}}{\sigma_{Kriging}}$ to check whether it follows a standard normal distribution. The figure 3.3 shows the histograms of this ratio by comparing with the probability density function (PDF) of the standard normal distribution (continuous black line) and the centered normal distribution whose deviation is 0.5. It can be concluded that the error committed by

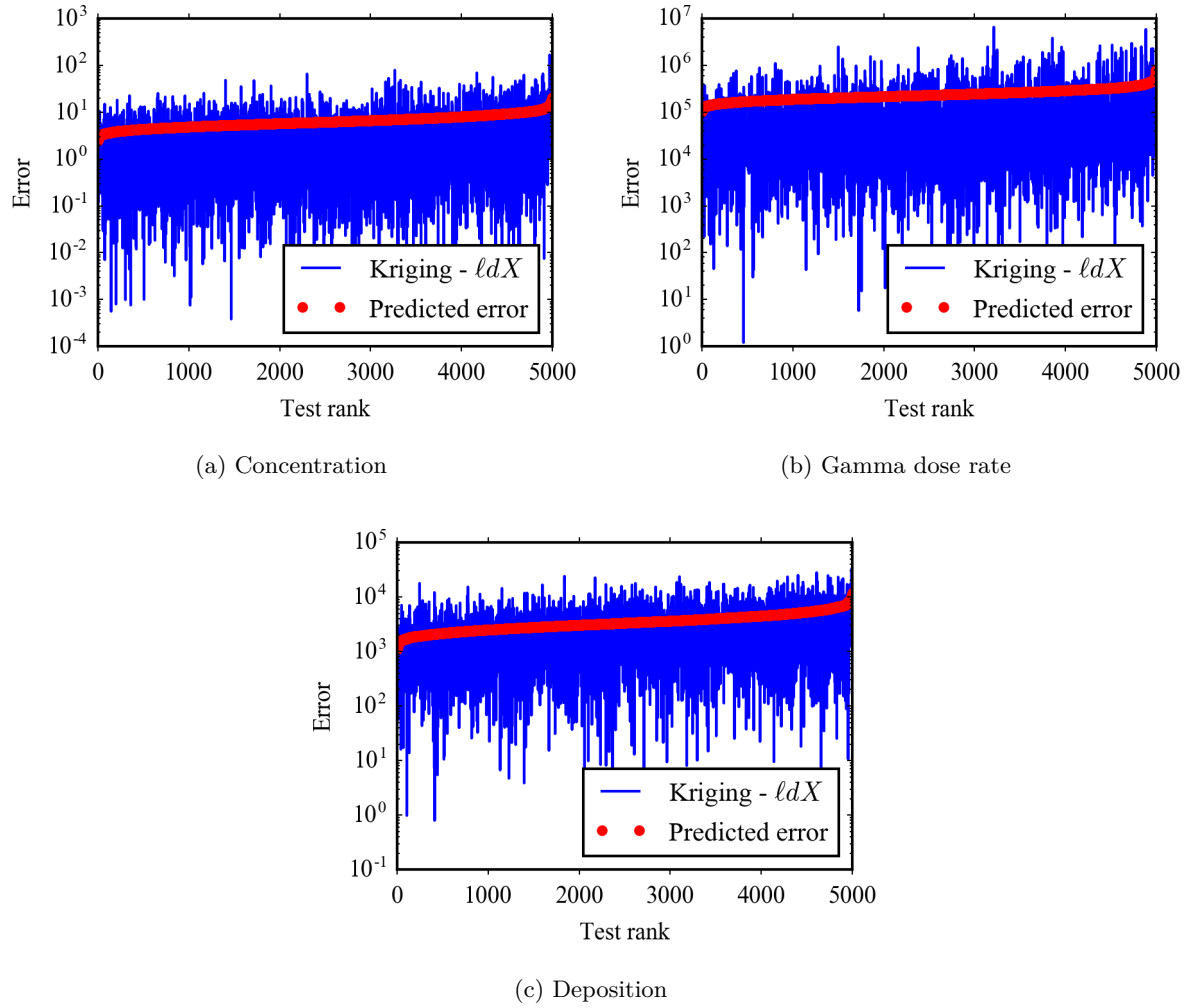


FIGURE 3.2 – Comparison between error predicted by Kriging and model-metamodel discrepancy.

Kriging is lower than the error predicted by this type of metamodel in most cases.

3.4.3 Comparing metamodels with different meteorological perturbations

The section 3.2.3 introduced two perturbation schemes for taking into account the uncertainty of the weather-forecast ensemble : by linear unit combination and by dispersion factor. The first method assumes that the weather forecast used as input to the ℓdX model should be a convex linear combination of the weather ensemble. This allows get information from each forecast without breaking the linear physical constraints. Note that the approach allows to reproduce the operational approach where the forecaster chooses a single meteorological simulation, hence putting a unit weight on this simulation, and no weight on the other simulations.

The metamodels will be used later in a deterministic optimization or even further for a statistical calibration. When the meteorological perturbations already bring 20 parameters to optimize, the dimension of the input space can lead to difficulties. The second perturbation scheme softens the problem. Only wind and rain are perturbed in this case. The amplitude of the factor k has been determined such that the upper and lower edges of new ensemble are as close as possible to the original ensemble. We note k_{max}^{wind} , k_{min}^{wind} , k_{max}^{rain} and k_{min}^{rain} respectively the higher and lower perturbation boundaries of wind and rain. These values are the results of the minimization of the RMSE between original and new envelopes. With this method, we have reduced the dimension of the weather inputs from twenty to two. We are interested here in the quality of emulators built on these two perturbation schemes : what is the best method to take into account the uncertainties of the weather forecasts? Did the dimension reduction decrease the quality of prediction ?

The SMSE score shows that, in general, perturbing with a convex combination produces better emulators, regardless of the use of either RBFs or Kriging ; see table 3.5. This result is confirmed by the figure 3.4. Despite the poorer quality of the metamodels, perturbing with a dispersion factor still builds emulators with reasonable errors, compared to the model-to-measurement error in table 3.3. In addition, this method greatly simplifies thereafter the calibration process because it makes use of fewer parameters.

Metamodel	Emulated value	Convex combination	Dispersion factor
RBFs	Concentration	0.06	0.155
	Gamma dose rate	0.027	0.677
	Distance of deposition	0.09	0.243
Kriging	Concentration	0.017	0.014
	Gamma dose rate	0.016	0.066
	Distance of deposition	0.05	0.168

TABLE 3.5 – Comparison between two perturbation schemes for the weather part, using SMSE scores.

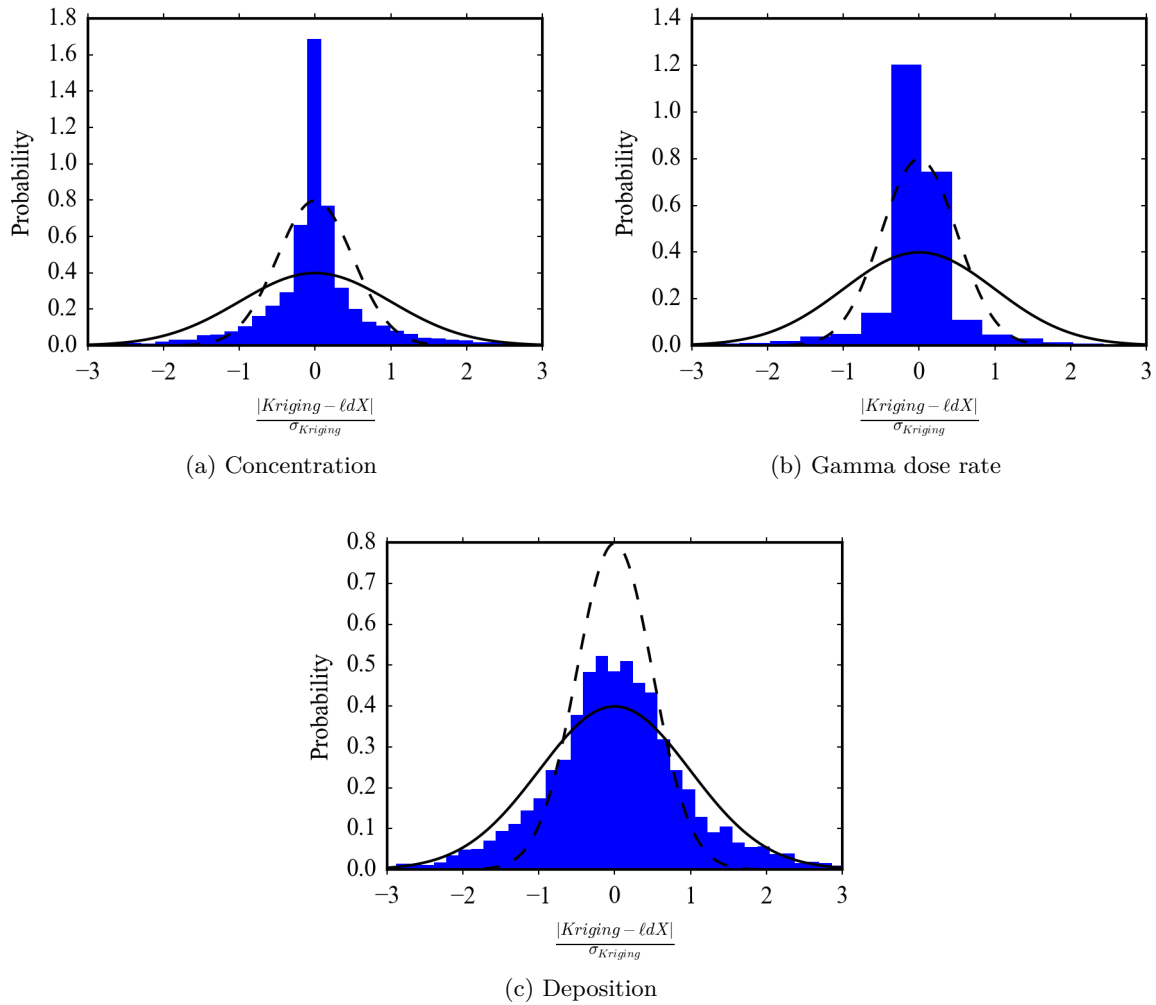
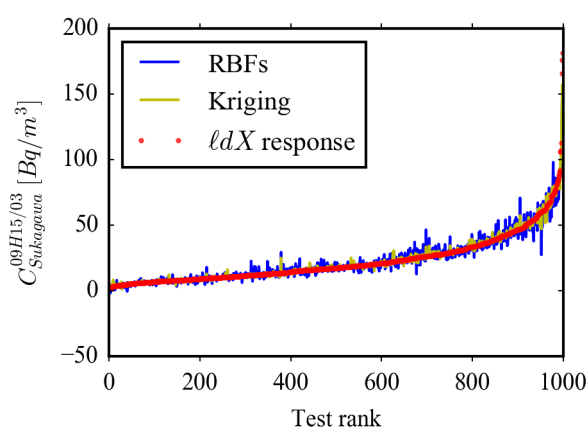
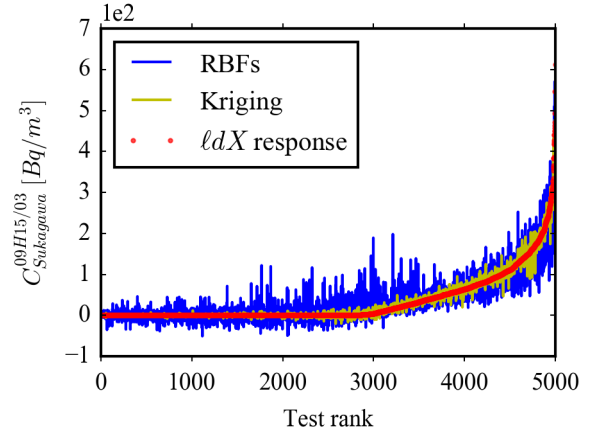


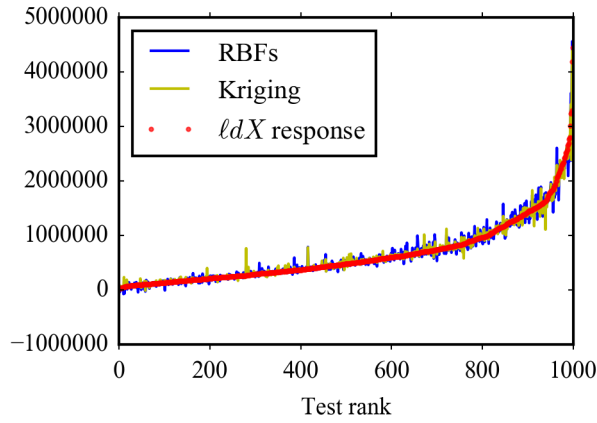
FIGURE 3.3 – Histograms of $\frac{\varepsilon_{\ell dX - Kriging}}{\sigma_{Kriging}}$ following the ℓdX outputs. The continuous curve corresponds to the probability density function (PDF) of the standard distribution. The dashed line represents the PDF of the centered normal distribution with 0.5 deviation.



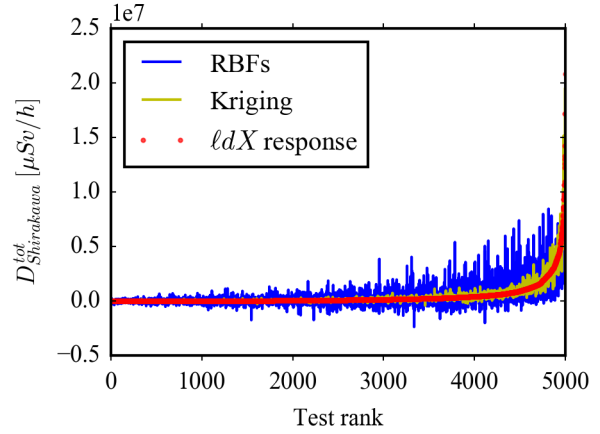
(a) Concentration at Sukagawa with convex combination



(b) Concentration at Sukagawa with dispersion factor



(c) Total dose rate in Shirakawa with convex combination



(d) Total dose rate in Shirakawa with dispersion factor

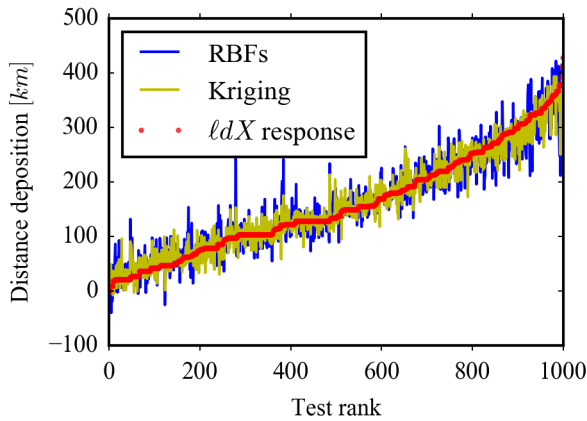
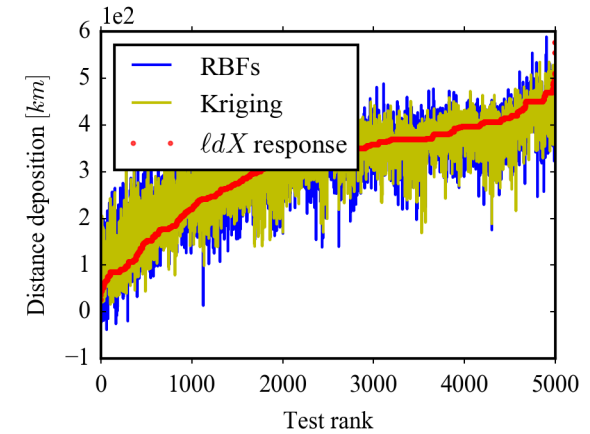

(e) Distance from the source of deposition higher than 37 kBq/m^3 with convex combination

(f) Distance of deposition higher than 37 kBq/m^3 with dispersion factor

FIGURE 3.4 – Comparison between two perturbation schemes for the meteorological part.

3.5 Applications of metamodels

3.5.1 Sensitivity analysis

One application of metamodeling is to study the sensitivity of the model to changes in the inputs. For example, one can plot cross sections of some model outputs against different inputs. The figure 3.5 represents the variation of the RMSE—*Root Mean Squared Error*—score on activity concentrations against three parameter pairs. Obviously, a RMSE value close to zero indicates an excellent simulation which almost coincides with the observations. This corresponds to the green areas in the figures. This clearly shows that some parameters have a greater influence on the performance than the others; for example, the activity concentration is more sensitive to the emission delay than the vertical diffusion; see figure 3.5c. Some parameters have complex effects.

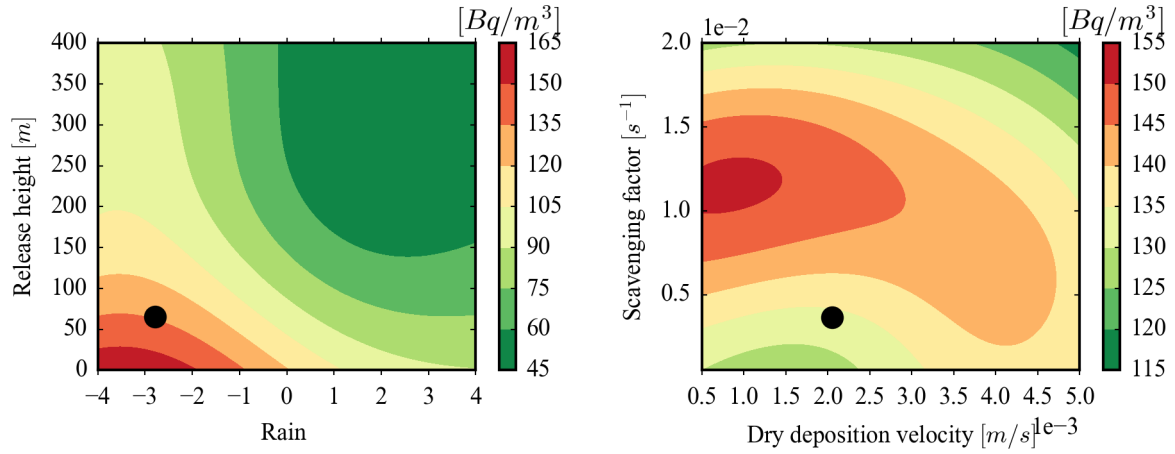
Cross sections with perturbations on the source terms are harder to plot because the coefficients of the convex linear combination are correlated to each other. We have chosen to set all weights to zero except for triplets of source terms. The figure 3.6 represents the FMT score of the concentration against variation in the weights of the following triplets of source terms : [Yumimoto 2016]—[Saunier 2016]—[Saunier 2013], [Yumimoto 2016]—[Katata 2015]—[Mathieu 2012] and [Winiarek 2014]—[Stohl 2012]—[Katata 2015]. Some source terms provide a lower RMSE value than the others. But the dimension of the inputs is high, and the RMSE function is likely to have plenty of local minima in other parts of the space.

3.5.2 Minimization of the RMSE on concentration

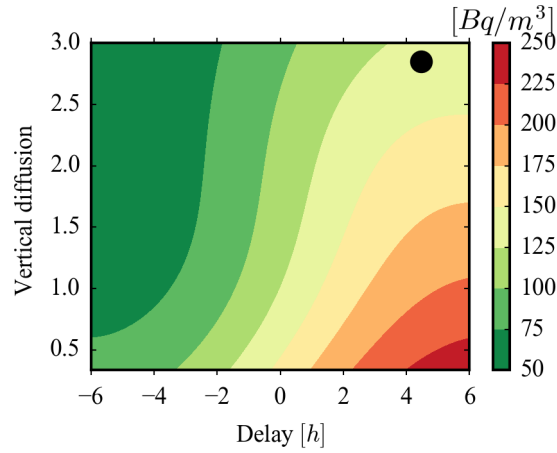
Since dispersion models are used in emergency cases, it is essential to find the optimal input parameters that allow outputs to get close to reality. A first optimization was carried out with the objective of finding the best input values that minimize the model-to-data discrepancy. Here, we applied the minimization of the RMSE on the activity concentration, while the weather forecasts were perturbed with the second perturbation scheme (see 3.2.3). This way is chosen in order to reduce the dimension of the input space. When the linear convex combination on the source terms is fixed, the minimization always finds a unique global minimum. However, when the the convex combination of the source terms is not fixed, many local minima appear. According to the figures 3.5 and 3.6, it seems that the RMSE function does contain several local minima. It is likely that a large number of such minima, spread all over the input space, make it difficult to converge to the global minimum. To verify this hypothesis, we studied the behavior of the RMSE on the DOE.

RMSE at selected DOE points against the distance between the selected points and the best point of the DOE (i.e., the point where the RMSE is minimal).

The table 3.6 shows the five lowest RMSE values of the DOE as well as the Euclidean distance between these points. Their RMSE scores are quite similar but they are quite distant in the input space. Indeed the average distance between DOE points is 0.656, and the selected points are farther from each other. By checking the RMSE against the Euclidean distance between the best point and the other points in figure 3.7, we see that very distant points in space can have a similar RMSE. This explains why the minimization process has trouble finding the global minimum. Next, using the metamodel, we have estimated the concentration RMSE on a much

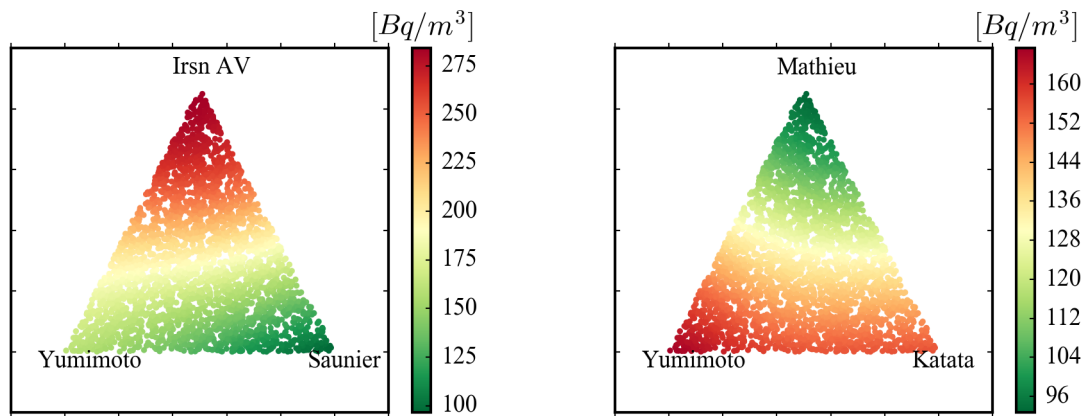


(a) Against the rain perturbation factor and the release height (b) Against dry deposition velocity and scavenging factor



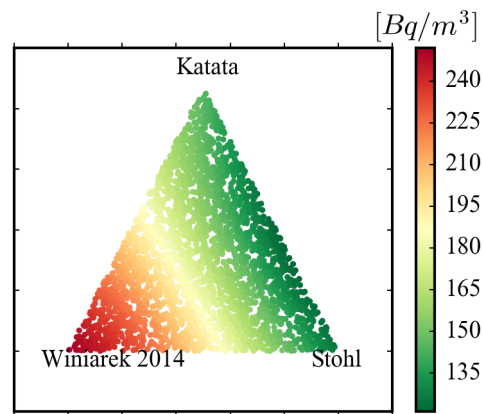
(c) Against the delay on the source term and the perturbation on the vertical diffusion

FIGURE 3.5 – Illustration of cross sections of the metamodel that emulates the RMSE score on activity concentrations. The black point indicates a point in the DOE where the output value is exact.



(a) Yumimoto - Saunier - IRSN inverse by concentration

(b) Yumimoto - Katata - Mathieu



(c) Winiarek 2014 - Stohl - Katata

FIGURE 3.6 – RMSE on activity concentration following several source terms.

	Best point 31.596	Point 2 31.672	Point 3 31.909	Point 4 31.926	Point 5 32.013
Best point	0	1.029	1.088	1.117	0.814
Point 2	-	0	0.973	1.228	1.139
Point 3	-	-	0	0.856	0.754
Point 4	-	-	-	0	1.275

TABLE 3.6 – Euclidean distance between the best points of the DOE.

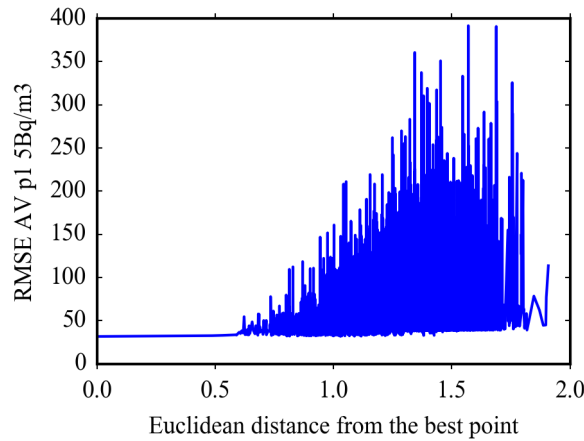


FIGURE 3.7 – RMSE values of DOE points against the distance from the best point.

denser set of points (1 000 000 samples) and then selected the best point of this database. A simulation with the original model ℓdX at that point allowed us to evaluate the optimization.

The best RMSE value on the concentration of the learning base is 31.596 Bq/m^3 . The Kriging and RBF emulators respectively find 10.962 Bq/m^3 and 2.105 Bq/m^3 as the best RMSE among the 1 000 000 points of the dense DOE. However, simulation with ℓdX on these points gives higher values because of the metamodel error. However, the best point of the DOE and the Kriging estimate agree that the source term [Saunier 2013] should have the largest weight. This source term was constructed after the Fukushima accident using observations of the activity concentration. This result is potentially correct.

	Surogate	ℓdX	Main source term
Best of DOE	-	31.596	[Saunier 2013]
Best of denser DOE by Kriging	10.962	39.035	[Saunier 2013]
Best of denser DOE by RBFs	2.105	39.819	[Mathieu 2012]

TABLE 3.7 – Comparing the RMSE of the concentration predicted by metamodels and ℓdX response.

3.6 Conclusions and perspectives

On the one hand, the constructed metamodels reproduce several outputs of the physical model ℓdX on the Fukushima disaster. These are local outputs like the concentration at a given moment and at a given measurement station, or global outputs like the average deposition on the entire Japanese territory. Despite the satisfactory quality of the emulators, these mathematical approximations could be improved. One of the proposals is to increase the performance of the DOE because combining LHS with simplex combinations may leave unsampled areas in the input space.

On the other hand, the radioactive measurements collected during the Fukushima nuclear disaster make it possible to optimize the dispersion models and then to better prepare for emergency cases. The first improvement of the ℓdX model is through deterministic optimization which aims to find the best input parameters that minimize the model-to-measure gap. This process does not converge because of the complexity of the simplex combination. But the quantification of uncertainty aims to find the probability distribution of uncertain parameters, instead of a single deterministic values. Markov Chain Monte Carlo (MCMC) sampling may be applied in order to sample the uncertain parameters in place of the non-convergent deterministic optimization.

In statistical calibration, especially in nuclear field, using several types of measurements may be useful because each type carries different information. For example, the deposition used in this study is a very dense measurement (in space), reconstructed after the accident, but it has no notion of time. This means that the simulations can contribute to the deposition at any time, as long as the final quantity is consistent with the measurements. Meanwhile, concentrations and gamma dose rates are time series measurements. To properly calibrate a simulation with these measurements, the calibration should bring some consistency with the temporal events. One difficulty is that our model is based on a coarse spatial grid, which can cause sub-grid effects to give poor results when comparing with the observations. The mixture of measurements allows us to have more constraints and to refine the obtained distributions.

Calibration statistique du modèle météorologique

La calibration statistique utilise les observations pour améliorer la connaissance sur les paramètres incertains du modèle. Contrairement à l'optimisation déterministe, ce processus a pour objectif de trouver non pas la valeur la plus vraisemblable pour chaque paramètre mais une distribution de probabilité de ces paramètres. Ce résultat est très important dans la situation de crise car le niveau de fiabilité des prévisions est toujours une information importante. Les modèles de dispersion atmosphérique sont très complexes ; ils utilisent des données d'entrée de grande dimension et plusieurs paramètres scalaires. Afin d'évaluer la méthode de calibration, un premier travail a été appliqué à la prévision météorologique en utilisant les observations du vent. Dans ce chapitre, il s'agit de présenter un modèle statistique simplifié pour prévoir le vent, de le calibrer et de vérifier la convergence des résultats.

Contents

4.1	Metropolis-Hastings et diagnostic de convergence Gelman-Brooks . . .	92
4.1.1	Algorithme Metropolis-Hastings	92
4.1.2	Vérification du critère Gelman-Brooks sur un cas simple	93
4.2	Calibration du modèle météorologique	94
4.2.1	Approche bayésienne	96
4.2.2	Construction de la loi <i>a posteriori</i>	97
4.2.3	Échantillonnage par Metropolis-Hastings	98
4.2.4	Vérification du résultat de calibration	100
4.3	Conclusions	100

4.1 Metropolis-Hastings et diagnostic de convergence Gelman-Brooks

Dans cette étude, la calibration statistique du modèle de prévision météorologique utilise un algorithme de la famille des méthodes MCMC, Metropolis-Hastings, qui a été présenté dans le chapitre 1. La théorie des chaînes de Markov est introduite dans le premier chapitre de cette thèse. Ce dernier a également mentionné les critères de [Brooks 1998], qui permettent de vérifier la convergence des chaînes de Markov issues de l'algorithme de MCMC. Nous allons appliquer, par la suite, ces critères dans l'ensemble de nos résultats de calibration. Ci-dessous, nous détaillons d'abord les étapes du processus d'échantillonnage de Metropolis-Hastings.

4.1.1 Algorithme Metropolis-Hastings

Pour obtenir des échantillons d'une variable aléatoire X , nous avons besoin de la fonction de répartition ou bien de la densité de probabilité de cette variable. Dans certains cas, ces fonctions sont complexes et difficiles à calculer. Dans un cadre bayésien, la densité de probabilité n'est souvent connue qu'à une constante de normalisation près (cf. section 4.2.1). Les méthodes de Monte Carlo par chaîne de Markov (MCMC) sont une solution à ce problème. Le but de ces méthodes est d'échantillonner une distribution cible f à l'aide d'une chaîne de Markov de mesure invariante f . Un des algorithmes les plus importants de MCMC est Metropolis-Hastings [Hastings 1970].

Le principe de l'échantillonnage par Metropolis-Hastings est de produire une série d'échantillons corrélés. La génération de l'état suivant dépend de l'état qui précède par une loi conditionnelle $q(x|y)$, appelée loi instrumentale ou loi candidate. Le choix de cette loi peut changer la vitesse de convergence de la chaîne de Markov vers la distribution cible f . Par exemple, si la loi instrumentale a un écart type trop large, la chaîne explore vite l'espace de la variable X mais l'algorithme ne sélectionne que peu de nouveaux tirages dans la ou les zones de concentration de la distribution f . En revanche, si cet écart type est trop faible, la chaîne accepte facilement de nouveaux candidats, mais la convergence peut être longue car la chaîne se déplace lentement. L'échantillonnage par Metropolis-Hastings se déroule comme suit :

- Initialisation : choix de l'état initial $\theta^{(0)}$ et du nombre d'itérations N ,
- Pour $t = 1, \dots, N$:
 1. Générer $y \sim q(\cdot | \theta^{(t-1)})$,
 2. Acceptation-rejet :

$$\theta^{(t)} = \begin{cases} y & \text{avec probabilité } \rho(\theta^{(t)}, y) \\ \theta^{t-1} & \text{avec probabilité } 1 - \rho(\theta^{(t)}, y) \end{cases}$$

où

$$\rho(a, b) = \min \left[\frac{f(b)}{f(a)} \cdot \frac{q(a|b)}{q(b|a)}, 1 \right] \quad (4.1)$$

Remarque

- La probabilité d'acceptation $\rho(a, b)$, équation 4.1, explique pourquoi la constante de normalisation n'est pas nécessaire. La division de $f(b)$ par $f(a)$ fait disparaître cet élément.
- De plus, si la loi instrumentale est symétrique, c'est-à-dire $q(a|b) = q(b|a)$, la probabilité d'acceptation sera simplifiée :

$$\rho(a, b) = \min \left[\frac{f(b)}{f(a)}, 1 \right]$$

L'algorithme doit générer des tirages suffisamment longtemps pour atteindre le régime stationnaire de la distribution cible. Dans le cas où le régime stationnaire est atteint, l'histogramme des tirages obtenus par Metropolis-Hastings représente la densité de probabilité de la distribution f . Cependant, les premiers tirages sont souvent ignorés à cause d'une période d'initialisation (burn-in). Pour expliquer ce phénomène, il faut revenir sur les caractéristiques des chaînes de Markov. Une chaîne de Markov n'a pas de mémoire, son comportement futur ne dépend que de son emplacement actuel et pas de la manière dont il s'y est rendu. Donc le fait d'enlever les premiers tirages rend la chaîne indépendante de sa valeur initiale. Dans la plupart des cas, la première moitié de la chaîne sera éliminée.

4.1.2 Vérification du critère Gelman-Brooks sur un cas simple

Avant de tester la convergence d'un échantillon obtenu avec l'algorithme de Metropolis-Hastings, le critère de Gelman-Brooks [Brooks 1998] sera utilisé dans la suite. Ci-dessous, ce critère est testé sur un cas simple. Le test est réalisé avec l'échantillonnage d'une loi cible connue (densité gaussienne de dimension $p = 20$). Le but est de vérifier si une valeur de R_c , équation 1.4, proche de 1 correspond bien à la convergence des chaînes de Markov. Dans ce test, la distribution cible est la loi jointe de 20 variables aléatoires gaussiennes, dont la densité s'écrit comme suit :

$$f(\boldsymbol{\theta}) = \frac{1}{(2\pi)^{p/2} |\Sigma|^{1/2}} \exp \left[-\frac{1}{2} (\boldsymbol{\theta} - \boldsymbol{\mu})^T \Sigma^{-1} (\boldsymbol{\theta} - \boldsymbol{\mu}) \right]$$

où $\boldsymbol{\mu}$ est le vecteur d'espérance de taille p et Σ est la matrice de covariance de taille $p \times p$. Ici, les composantes de $\boldsymbol{\mu}$ sont fixées à des valeurs différentes :

$$\boldsymbol{\mu} = (-12, -11, 3, -15, 7, -12, 11, -12, 9, -11, -11, 9, -9, -6, 8, -8, -20, 10, 11, 10)^T$$

La matrice Σ est quasi-diagonale, c'est-à-dire que les composantes extra-diagonales sont beaucoup plus faibles que celles sur la diagonale. Cela implique que les variables aléatoires sont presque indépendantes entre elles et que la loi jointe n'est pas loin du produit des lois marginales. Dans le cas d'existence de corrélations, déterminer les lois marginales serait plus délicat avec la densité gaussienne, car on doit assurer que la matrice de covariance Σ est inversible.

La loi instrumentale $q(\cdot|\boldsymbol{\theta})$ est la loi uniforme en dimension 20 : $\mathcal{U}(\boldsymbol{\theta} - 1, \boldsymbol{\theta} + 1)$. Avec l'algorithme de Metropolis-Hastings, deux chaînes de Markov de taille 1 000 000 chacune ont été établies parallèlement. Les 500 000 derniers échantillons de chaque chaîne ont été sauvegardés pour appliquer le diagnostic de Gelman-Brooks. Ce critère identifie le régime stationnaire au cours des itérations, alors la baisse de l'indicateur R_c doit bien correspondre à l'approche des inférences empiriques ($\hat{\mu}$ et $\hat{\Sigma}$) vers la cible ($\boldsymbol{\mu}$ et Σ) ; cf. figure 4.1. Autrement dit, nous voulons

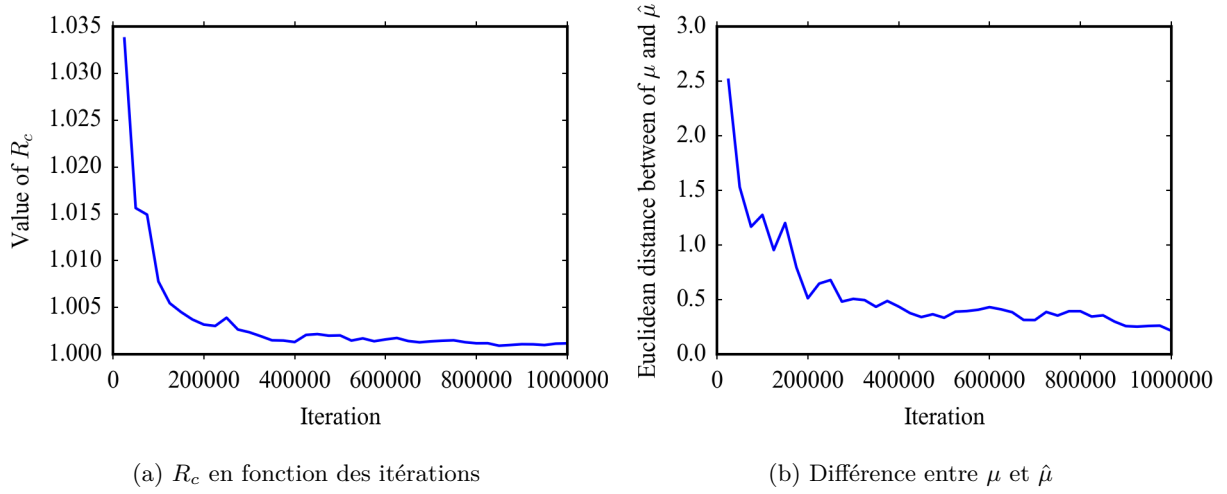


FIGURE 4.1 – Indicateurs de convergence

vérifier si R_c est bien un indicateur de convergence des chaînes de Markov. De plus, le vecteur $\hat{\mu}$ obtenu par ces deux chaînes de Markov est

$$\hat{\mu} = (-12.01, -10.95, 2.97, -15, 6.97, -12, 10.95, -11.98, 8.95, -11.02, -11.02, 9.03, \\ -9.14, -6.01, 7.94, -8.02, -20.06, 9.96, 11.01, 9.93)^T$$

Le vecteur final $\hat{\mu}$ est assez proche de μ . La distance euclidienne entre μ et $\hat{\mu}$ baisse fortement au cours des itérations, comme le montre la figure 4.1b. Donc la baisse vers 1 de R_c correspond bien à la convergence en moyenne des variables aléatoires. De plus, comme la loi cible est connue, il suffit de comparer les histogrammes obtenus avec les itérations à la PDF de la loi cible (*Probability Density Function*), comme à la figure 4.2. À partir des figures 4.1 et 4.2, on peut conclure que les chaînes de Markov dans ce test convergent bien vers la loi cible et que l'indicateur de [Brooks 1998] signale bien la convergence de ces chaînes.

4.2 Calibration du modèle météorologique

Nous disposons des observations météorologiques pour plusieurs variables (température, vent, pluie, etc.), qui permettent d'évaluer un ensemble de simulations météorologiques correspondant. Les observations météorologiques proviennent du réseau AMeDAS du Japon, et ont été détaillées dans la partie 2.2.2. La localisation des stations d'observation AMeDAS est représentée dans la figure 4.3.

Dans ce chapitre, nous nous intéressons à l'utilisation de ces observations pour trouver la distribution de probabilité de certains paramètres du modèle météorologique en appliquant l'algorithme de Metropolis-Hastings. Cependant, nous disposons seulement d'ensemble de prévisions de l'ECMWF, qui est le résultat d'un modèle météorologique. Cet ensemble de prévisions se compose de 50 simulations ou membres qui représentent chacun une possibilité d'évolution de l'atmosphère dans le futur. À partir de cet ensemble de 50 membres, nous formulons un modèle

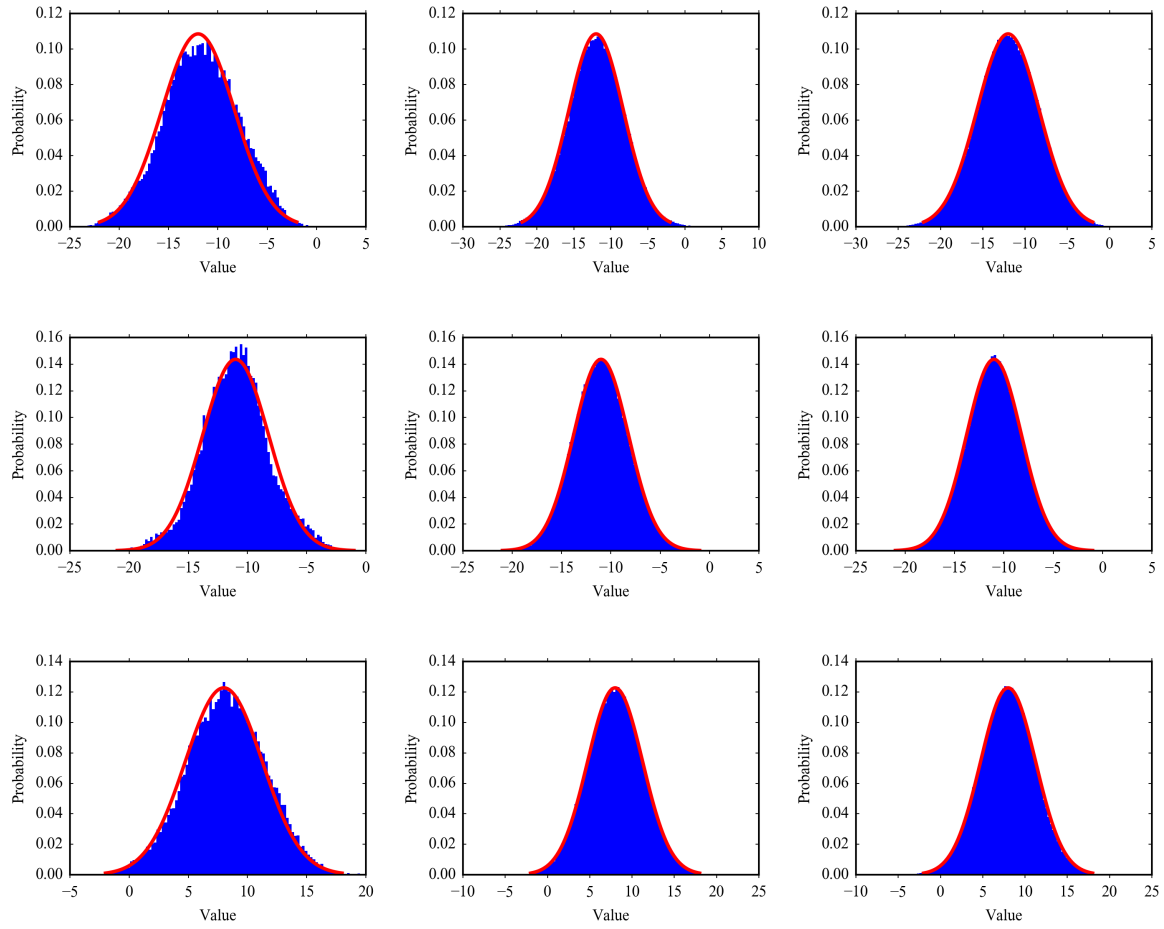


FIGURE 4.2 – Illustration d’histogrammes du premier paramètre (en haut), du 10e paramètre (au milieu) et du 15e paramètre (en bas) au bout de 50 000 itérations (à gauche), 500 000 itérations (au milieu) et 1 000 000 itérations (à droite). La ligne rouge représente la loi cible marginale de chaque paramètre.

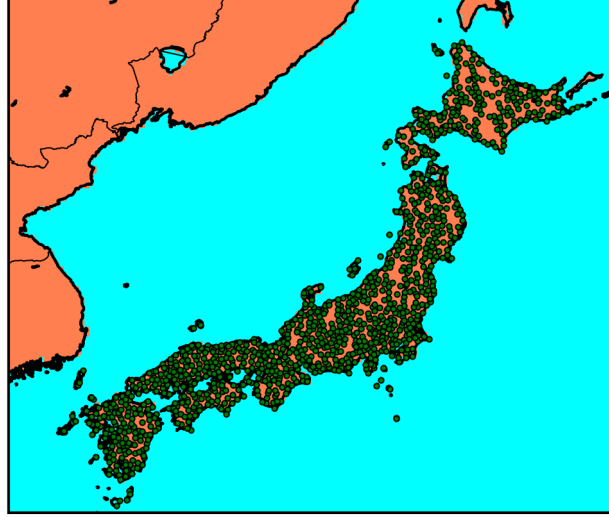


FIGURE 4.3 – Carte de position des stations d’observation météorologiques. Chaque point vert correspond à une station de mesure du réseau AMeDAS.

probabiliste. Pour chaque variable météorologique (température, vent, pluie, etc.), ce modèle s’écrit

$$\mathcal{M}(p) = \mathcal{M}(\alpha, \beta) = \alpha\mu_{ens} + \beta\sigma_{ens} \in \mathbb{R}^n \quad (4.2)$$

où μ_{ens} et σ_{ens} sont respectivement la moyenne et l’écart type de ces 50 prévisions, n est le nombre d’observations (nombre de stations $\times$ nombre de pas de temps) et $p = (\alpha, \beta)$ est le vecteur des paramètres à calibrer. Le modèle est construit en appliquant le facteur α au biais empirique de l’ensemble, et le facteur β à l’écart type empirique de l’ensemble. Grâce à β , on espère corriger une sous-estimation ou une sur-estimation systématique de l’incertitude de prévision sur chaque variable considérée. L’algorithme de Metropolis-Hastings demande que la loi cible des paramètres incertains soit connue à un facteur près. Cette densité peut être déterminée à l’aide de l’inférence bayésienne.

4.2.1 Approche bayésienne

L’inférence bayésienne est une démarche logique permettant de calculer ou réviser la probabilité d’une hypothèse. Elle fait intervenir à la fois les quantités connues (observations, par exemple) et les quantités inconnues (les paramètres qui n’ont pas été observés). Différente de l’inférence classique, l’inférence bayésienne introduit une information préalable (avant les observations) disponible sur les variables d’intérêt $\mathbf{p}$, appelée information *a priori*. Dans l’inférence bayésienne, le paramètre $\mathbf{p}$ est probabilisé, et on lui associe une loi de probabilité dite loi *a priori*, notée $\pi(\mathbf{p})$. Appelons $\mathbf{y} = (y_1, \dots, y_n)$ les observations sur lesquelles l’analyse statistique se fonde. L’approche statistique bayésienne calcule une loi conditionnelle sur $\mathbf{y}$, actualisation de la loi *a priori* $\pi(\mathbf{p})$ au vu de l’observation. Cette loi conditionnelle s’appelle loi *a posteriori*,

notée $\pi(\mathbf{p}|\mathbf{y})$. Suivant le théorème de Bayes en version continue, la loi *a posteriori* s'écrit

$$\pi(\mathbf{p}|\mathbf{y}) = \frac{f(\mathbf{y}|\mathbf{p})\pi(\mathbf{p})}{f(\mathbf{y})}$$

où $f(\mathbf{y}|\mathbf{p})$ désigne la vraisemblance et $f(\mathbf{y})$ est la loi marginale ou prédictive

$$f(\mathbf{y}) = \int f(\mathbf{y}|\mathbf{p})\pi(\mathbf{p})d\mathbf{p} \quad (4.3)$$

4.2.2 Construction de la loi *a posteriori*

La calibration a été effectuée sur la vitesse du vent. Appelons

- $\mathbf{y} = (y_1, \dots, y_n)$ le vecteur des observations du module du vent,
- $\mathbf{x} = \mathcal{M}(\alpha, \beta) \in \mathbb{R}^n$ le vecteur de sortie du modèle 4.2

L'erreur modèle-mesure $\boldsymbol{\varepsilon} \in \mathbb{R}^n$ est supposée suivre une loi normale centrée : $\varepsilon_i \sim \mathcal{N}(0, \sigma^2) \quad \forall i \in \{1, \dots, n\}$. Dans ce test, nous supposons que le modèle 4.2 est parfait, alors le problème peut être résumé de manière suivante :

$$\forall i \in \{1, \dots, n\}, \quad \varepsilon_i = y_i - x_i \quad (4.4)$$

$$= y_i - \mathcal{M}_i(\alpha, \beta) \sim \mathcal{N}(0, \sigma^2) \quad (4.5)$$

Cependant, l'écart type σ de cette erreur n'est pas connu, il peut être introduit dans la calibration. Les paramètres à calibrer deviennent donc $\boldsymbol{\theta} = (\alpha, \beta, \sigma)$, et le modèle statistique sera $\widetilde{\mathcal{M}}(\boldsymbol{\theta}) = \mathcal{M}(\alpha, \beta) + \mathcal{N}(0, \sigma^2)$. La vraisemblance est la densité de probabilité de x (ou de $\boldsymbol{\theta}$) à condition que y se produise. En supposant que les ε_i sont indépendants entre eux, la vraisemblance de $\boldsymbol{\theta}$ sachant que y est le produit des lois gaussiennes de ε_i , ce qui donne la relation entre α, β, σ et y :

$$f(\mathbf{y}|\boldsymbol{\theta}) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y_i - \mathcal{M}_i(\alpha, \beta))^2}{2\sigma^2}\right) \quad (4.6)$$

$$\Rightarrow \log[f(\mathbf{y}|\boldsymbol{\theta})] = -\frac{n}{2} \log(2\pi) - \sum_{i=1}^n \left(\log(\sigma) + \frac{(y_i - \mathcal{M}_i(\alpha, \beta))^2}{2\sigma^2} \right) \quad (4.7)$$

Quand le nombre d'observations est abondant, le terme sous l'exponentiel risque d'être très négatif et cela peut engendrer des erreurs d'approximation dans les calculs numériques. Pour travailler avec la vraisemblance, il est conseillé de passer par le logarithme de cette fonction pour éviter les problèmes numériques. Une fois que la vraisemblance est établie, et avec les lois *a priori*, on peut appliquer l'inférence bayésienne pour obtenir la loi *a posteriori*. Le terme de normalisation 4.3 n'est souvent pas connu, nous connaissons donc la distribution *a posteriori* à une constante de proportionnalité près :

$$\pi(\boldsymbol{\theta}|\mathbf{y}) \propto f(\mathbf{y}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})$$

Le tableau 4.1 représente les distributions *a priori* qui ont été choisies pour les trois variables de $\boldsymbol{\theta}$. Les lois *a priori* ont été choisies telles que le support de l'ensemble des paramètres d'entrée soit le plus large possible afin d'atteindre toutes les possibilités. Ensuite, l'étape d'acceptation de Metropolis-Hastings va pondérer les zones les plus probables.

Variable	Distribution <i>a priori</i>
α	$\mathcal{N}(1, 10^2)$ tronquée sur $[0, +\infty[$
β	$\mathcal{N}(0, 5^2)$
σ	$\mathcal{N}(0.2, 1^2)$ tronquée sur $[0, +\infty[$

TABLE 4.1 – Distribution *a priori* de $\theta = (\alpha, \beta, \sigma)$

4.2.3 Échantillonnage par Metropolis-Hastings

Pour obtenir la distribution de θ , six chaînes de Markov ont été lancées en partant de différents points sur 1 000 000 itérations. Les points de départ sont donnés par l'égalité 4.8. Les 500 000 premières itérations ont été enlevées pour éviter le burn-in. À chaque itération, la loi instrumentale de θ est donnée par le tableau 4.2. Les histogrammes de ces trois paramètres, représentés dans la figure 4.4, visualisent la formation des distributions

$$\Theta = [(\alpha_i, \beta_i, \sigma_i)]_{i=1, \dots, 6} = \begin{pmatrix} (1, 5, 0.1) \\ (10, 5, 5) \\ (20, 0.1, 10) \\ (5, -6, 6) \\ (2, 10, 20) \\ (15, -2, 2) \end{pmatrix} \quad (4.8)$$

Variable	Loi instrumentale
α	$\mathcal{N}(\alpha^{(t)}, 0.1^2)$ tronquée sur $[0, +\infty[$
β	$\mathcal{N}(\beta^{(t)}, 0.1^2)$
σ	$\mathcal{N}(\sigma^{(t)}, 0.1^2)$ tronquée sur $[0, +\infty[$

TABLE 4.2 – Distribution instrumentale de θ

Au cours des itérations, les histogrammes évoluent et quand la convergence s'enclenche, une forme de cloche apparaît qui indique la distribution des paramètres de θ . L'indicateur de convergence R_c a baissé de 1.45 (après 25 000 itérations) à 1.005 (après 1 000 000 itérations), cf. figure 4.5. Le vecteur des moyennes et la matrice de covariance empirique sont

$$\hat{\mu}_{(\alpha, \beta, \sigma)} = (0.37, 0.22, 1.76), \hat{\Sigma} = \begin{pmatrix} 0.14 & 0.08 & 0.65 \\ 0.08 & 0.05 & 0.38 \\ 0.65 & 0.38 & 3.1 \end{pmatrix}$$

D'autre part, la valeur la plus vraisemblable de σ est 1.76 qui correspond bien à l'erreur d'observation de la vitesse du vent présentée dans l'article [Bowler 2009]. Suivant l'algorithme de [Brooks 1998], on peut conclure que l'ensemble de ces six chaînes convergent vers la distribution cible. Une fois la calibration terminée, une vérification par des simulations directes permet de valider le résultat de l'échantillonnage.

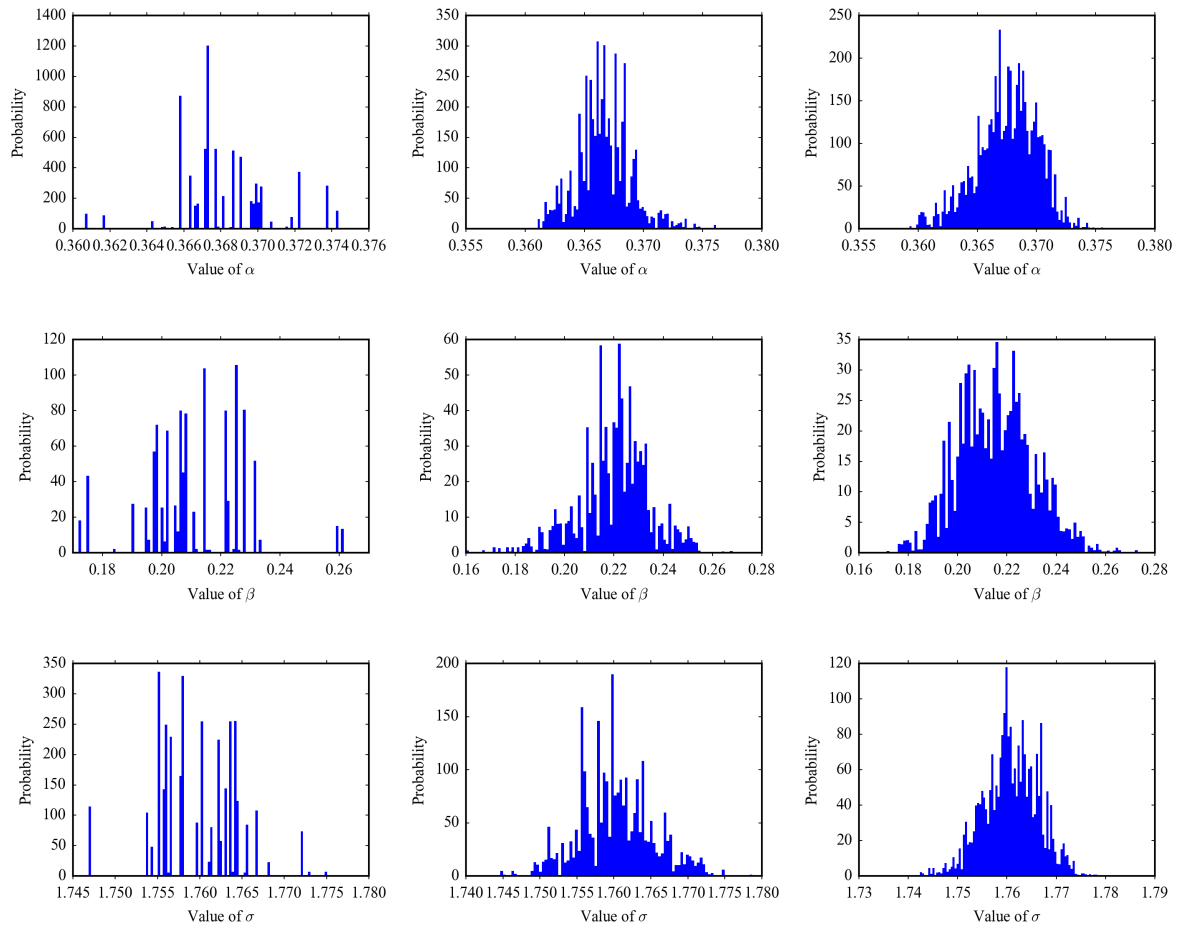


FIGURE 4.4 – Histogramme de θ au bout de 25 000 (à gauche), 500 000 (au milieu) et 1 000 000 itérations (à droite).

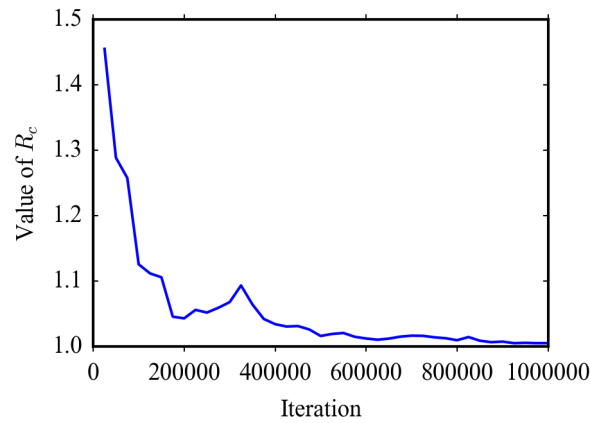


FIGURE 4.5 – Indication de convergence R_c en fonction du nombre d'itérations.

4.2.4 Vérification du résultat de calibration

Pour évaluer les échantillons obtenus après la calibration, 50 tirages de θ ont été choisis parmi les dernières moitiés des six chaînes de Markov. Dans la partie 4.2.2, l'erreur modèle-mesure ε est supposée suivre une loi normale centrée d'écart type σ . Cela signifie qu'une fois appliquées les valeurs de α et β dans l'équation 4.2, il faut ajouter la partie stochastique ε à chaque prévision :

$$Y(\alpha, \beta, \sigma) = \mathcal{N}(\mathcal{M}(\alpha, \beta), \sigma^2 \mathbb{I})$$

où $\mathbb{I}$ est la matrice d'identité de taille $n \times n$. Ce nouvel ensemble a été comparé aux observations afin d'évaluer la calibration par Metropolis-Hastings, ainsi que le critère de convergence de [Brooks 1998]. La figure 4.6 permet de visualiser l'évolution entre l'ensemble météorologique original et le résultat de la calibration. La bande verte représente l'ensemble de prévision météorologique original de l'ECMWF. La bande bleue est l'ensemble des 50 nouveaux membres échantillonnés après la calibration. Les points rouges sont les observations obtenues en moyennant les stations dans la même maille. La bande bleue se rapproche beaucoup plus des observations que la bande verte et l'ajout de la partie stochastique ε rend ce nouvel ensemble plus dispersé et permet d'englober presque toutes les observations.

Une autre façon de vérifier la qualité de cet ensemble est d'utiliser le diagramme de rang. La figure 4.7 montre le diagramme de rang de ces deux ensembles. La construction des diagrammes de rang est effectuée sur un seuil de 1 m/s sur les observations, i.e., l'observation est incluse seulement si elle est supérieure au seuil. Cette figure montre que l'ensemble fourni par l'ECMWF est très peu dispersé et il surestime souvent les observations. Au contraire, le nouvel ensemble donne un diagramme de rang beaucoup plus plat. Cependant, il a tendance à sous-estimer légèrement les observations.

4.3 Conclusions

La méthode de MCMC a permis de trouver la distribution de probabilité des paramètres α et β , ainsi que l'écart type σ de l'erreur du modèle stochastique. Dans ce cas test, les chaînes de Markov atteignent la convergence assez rapidement et la vérification par simulation directe a prouvé que les échantillons obtenus génèrent une distribution cohérente avec les observations. Ce résultat permet également d'accéder au niveau de fiabilité de la prévision à grâce aux distributions des sorties.

L'objectif de ce cas test simplifié est de vérifier la mise en place de l'algorithme de MCMC et le diagnostic de convergence de Brooks. Dans la prochaine étape, il s'agit d'appliquer ces deux méthodes à la calibration statistique de ℓdX en utilisant les mesures radiologiques.

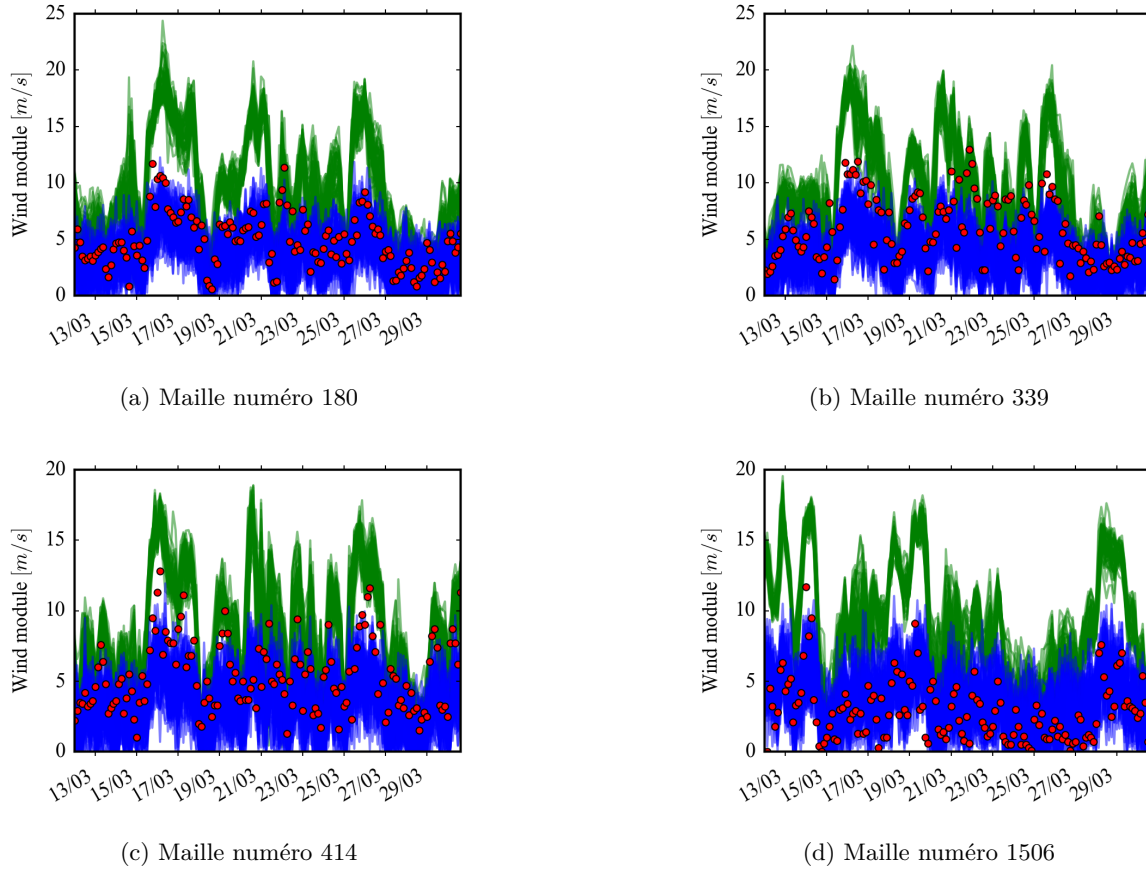


FIGURE 4.6 – Comparaison du résultat de calibration (bleu) et l'ensemble météorologique original (vert). Les points rouges sont des observations utilisées dans la calibration.

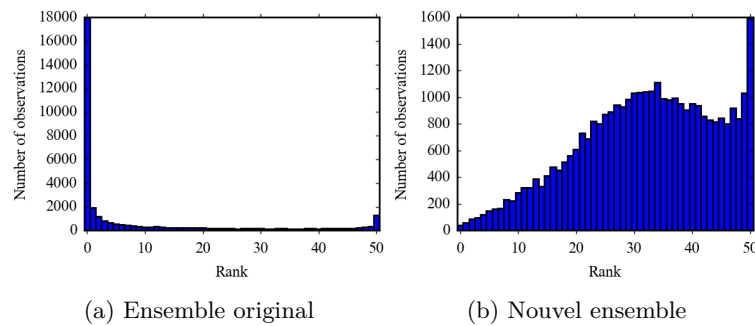


FIGURE 4.7 – Diagrammes de rang pour la vitesse du vent de l'ensemble original provenant de l'ECMWF et l'ensemble calibré par Metropolis-Hastings.

Optimisation probabiliste des sorties radiologiques de ℓdX

L'objectif final de ce travail est de retrouver les distributions de probabilité des paramètres d'entrée incertains du modèle opérationnel ℓdX . Le chapitre précédent joue le rôle d'un test préliminaire de l'algorithme d'échantillonnage Metropolis-Hastings sur un modèle météorologique simplifié. Nous allons ensuite appliquer cette méthode dans l'étude d'incertitude du modèle radiologique ℓdX en utilisant les trois types de sorties présentés dans le chapitre de propagation par méthode de Monte Carlo 2.3.3.

Contents

5.1	Construction de la loi <i>a posteriori</i> par inférence bayésienne	103
5.1.1	Prise en compte de la corrélation des observations radiologiques dans la vraisemblance	103
5.1.2	Informations <i>a priori</i> sur les paramètres incertains	104
5.2	Cas synthétique : jeu d'entrée optimal connu	104
5.2.1	Calibration du dépôt avec observations corrélées	106
5.2.2	Calibration d'activité volumique intégrée sans corrélation entre les observations	110
5.3	Cas réel : observations lors de l'accident de Fukushima	111
5.3.1	Évaluation des métamodèles utilisés	113
5.3.2	Calibration avec le débit dose et une corrélation entre les observations . . .	114
5.3.3	Calibration avec l'activité volumique et sans corrélation	115
5.3.4	Calibration avec le dépôt et avec corrélation	123
5.4	Conclusion	124

5.1 Construction de la loi *a posteriori* par inférence bayésienne

5.1.1 Prise en compte de la corrélation des observations radiologiques dans la vraisemblance

La calibration du modèle météorologique simplifié supposait que les observations étaient toutes indépendantes entre elles. Cela signifie que chaque observation apporte une information unique qui permet de localiser la valeur optimale des paramètres incertains. Cependant, les mesures radiologiques apportent souvent des informations redondantes. Nous avons donc supposé qu'il existe une corrélation entre les observations (spatiale pour le dépôt ; et temporelle pour

l'activité volumique et le débit de dose). Dans ce cas, la relation entre les observations et la sortie de ℓdX ne s'écrit plus comme l'équation 4.5 mais de la manière suivante :

$$\varepsilon = \mathbf{y} - \mathcal{M}(\mathbf{p}) \sim \mathcal{N}(\mathbf{0}, \Sigma) \quad (5.1)$$

où $\mathbf{y}$ et $\mathcal{M}(\mathbf{p})$ sont respectivement les observations radiologiques (concentration, dépôt ou débit dose) et les sorties du modèle ℓdX avec le jeu de données d'entrée $\mathbf{p}$; tous les deux sont de taille n . $\mathbf{0}$ est le vecteur zéro de taille n ; Σ est la matrice de covariance représentant la dépendance entre les observations; $|\Sigma|$ est le déterminant de cette matrice. La seule condition à respecter ici est que la matrice Σ doit être une matrice carrée et symétrique définie positive, donc inversible. Les propriétés de cette matrice ont été abordées dans section 1.1.1. Donc, la vraisemblance et les logarithmes des équations 4.6 et 4.7 se transforment en

$$f(\mathbf{y}|\mathbf{p}) = (2\pi)^{-n/2} |\Sigma|^{-1} \exp\left(-\frac{1}{2} (\mathbf{y} - \mathcal{M}(\mathbf{p}))^T \Sigma^{-1} (\mathbf{y} - \mathcal{M}(\mathbf{p}))\right) \quad (5.2)$$

$$\begin{aligned} \log[f(\mathbf{y}|\mathbf{p})] &= -\frac{n}{2} \log(2\pi) - \log(|\Sigma|) \\ &\quad - \frac{1}{2} (\mathbf{y} - \mathcal{M}(\mathbf{p}))^T \Sigma^{-1} (\mathbf{y} - \mathcal{M}(\mathbf{p})) \end{aligned} \quad (5.3)$$

Par la suite, la fonction de vraisemblance sera utilisée dans l'algorithme de Metropolis-Hastings via un métamodèle. Autrement dit, nous émuloons la vraisemblance ou son logarithme sur la base d'un tirage de simulations issu d'un LHS, ainsi qu'il est présenté dans le chapitre sur la méta-modélisation.

5.1.2 Informations *a priori* sur les paramètres incertains

Pour construire la loi *a posteriori*, nous avons besoin non seulement de la vraisemblance mais aussi de la loi *a priori*. La loi *a priori* correspond à nos connaissances préalables. Les jugements d'expert fournissent des plages de variation de ces paramètres, et des distributions uniformes ont été appliquées sur ces intervalles. Cf. tableau 5.1. La densité de la loi *a priori* est constante et sera éliminée par la probabilité d'acceptation de l'algorithme de Métropolis-Hastings 4.1. Les distributions uniformes servent juste à limiter l'intervalle de variation des paramètres.

Les distributions des paramètres sont supposées indépendantes entre elles, alors la loi jointe *a priori* devient la composition directe des lois marginales. Autrement dit, la distribution *a priori* que nous allons utiliser sera le produit de toutes les distributions uniformes du tableau 5.1.

5.2 Cas synthétique : jeu d'entrée optimal connu

L'objectif de la thèse est d'utiliser les mesures radiologiques pour retrouver une distribution de probabilité des paramètres incertains de ℓdX , présentés dans le paragraphe précédent 5.1.2. Mais avant d'appliquer les méthodes de calibration au cas avec des mesures réelles de l'accident de Fukushima, nous avons effectué une calibration en utilisant des observations synthétiques afin de vérifier l'exactitude de l'algorithme de Metropolis-Hastings, ainsi que d'évaluer l'influence de l'émulation dans ce processus. Ce cas test est appelé synthétique car un jeu de données d'entrée a été choisi pour créer des observations « artificielles », sur lesquelles nous allons construire la

Variable [Notation/unité]	Distribution <i>a priori</i>
Facteur de dispersion du vent [k_{wind}]	$\mathcal{U}[-3, 3]$
Facteur de dispersion de la pluie [k_{rain}]	$\mathcal{U}[-4, 4]$
Terme source	$\mathcal{D}(\alpha = (1, \dots, 1)^T)$
Décalage en temps [dt/h]	$\mathcal{U}[-6, +6]$
Hauteur d'émission [z/m]	$\mathcal{U}[0, 400]$
Vitesse de dépôt sec [vd/ms^{-1}]	$\mathcal{U}[5.10^{-4}, 5.10^{-3}]$
Facteur de lessivage [$a/hs^{-1}mm^{-1}$]	$\mathcal{U}[5.10^{-4}, 2.10^{-2}]$
Exposant de lessivage [b]	$\mathcal{U}[0.6, 1]$
Diffusion horizontale [$K_z/\times 10^4 m^2 s^{-1}$]	$\mathcal{U}[0., 1.5]$
Facteur de diffusion verticale [α_{K_z}]	$\mathcal{U}[1/3, 3]$
Amplitude de rejet [Amp]	$\mathcal{U}[1/3, 3]$

TABLE 5.1 – Distribution *a priori* des paramètres d'entrée du métamodèle

vraisemblance. Comme nous connaissons déjà le point de référence, le but de cette étude est de vérifier si les chaînes de Markov arrivent à trouver le bon jeu de données d'entrée de départ.

Dans ce même cas synthétique, nous nous concentrons sur deux types de sortie : la concentration (activité volumique) aux stations de mesure et le dépôt final, qui ont des caractéristiques complémentaires. L'activité volumique est une série temporelle qui représente le passage des panaches à un endroit précis. D'après le chapitre de propagation par méthode de Monte Carlo, l'activité volumique est une sortie très difficile à englober par l'ensemble de simulations, comme le montraient les diagrammes de rang en forme de U, avec de grandes barres aux extrémités. L'accident de Fukushima s'étale sur un mois entier et se compose de plusieurs événements de rejet, suivant [Mathieu 2018a]. La calibration de l'activité volumique sur tout le mois de mars 2011 pourra être très compliquée à atteindre. Pour cette raison, nous avons calibré l'activité volumique intégrée en temps sur seulement trois jours (du 14 au 16 mars). Le fait d'utiliser des valeurs intégrées en temps nous permet d'éliminer les erreurs dues à des décalages temporels de quelques heures, et donc rend un taux d'acceptation d'échantillonnage plus élevé. De plus, l'activité volumique intégrée sur une période de 24 ou 48 heures est proportionnelle à la dose qui serait calculée au point d'observation pour évaluer l'impact radiologique du rejet. Il s'agit donc d'une variable d'intérêt pour l'évaluation des conséquences radiologiques.

Le dépôt final ne donne pas d'information temporelle sur la contamination car il est déjà intégré dans le temps, mais c'est une mesure à résolution spatiale très fine. La calibration du dépôt peut faire sortir plusieurs scénarios de cet événement. Pour cette sortie, nous ne pouvons pas utiliser toutes les mesures (plus de 30 000 points) car ce sont des mesures aéroportées effectuées à une haute résolution spatiale. Cela explique la corrélation entre les points d'observation. Une solution est de moyenner ces points d'observation sur la grille de simulation, ce qui conduit finalement à 134 mesures de dépôt moyennées en espace.

L'accident de Fukushima est fortement complexe, avec un nombre important de mesures. Cela crée des contraintes très fortes dans l'échantillonnage. Il est probable que la distribution finale soit fortement concentrée à un endroit dans l'espace de très grande dimension (dimension

19 dans ce cas). Pour contourner ce problème, nous avons testé en parallèle, pour chaque type de sortie, deux situations : des observations indépendantes, équations 4.6 et 4.7, et de observations corrélées, équations 5.2 et 5.3. Parmi ces équations, les variables σ_i représentent l'erreur d'observation de chaque mesure. La matrice de covariance Σ définie positive a comme diagonale le vecteur $\sigma^2 = (\sigma_1^2, \dots, \sigma_n^2)^T$. Les termes extra-diagonaux de Σ s'écrivent

$$s_{ij} = \sigma_i \sigma_j \left(1 + \frac{d}{L}\right) \exp\left(-\frac{d}{L}\right) \quad (5.4)$$

où d est la distance entre deux observations i et j (cette distance peut être l'écart temporel ou spatial ou une combinaison des deux), L est la distance de référence que nous avons choisie. Pour ces deux mesures précédentes, L et d correspondent à la distance spatiale entre deux observations, avec $L = 10$ km.

Concernant les métamodèles utilisés dans la calibration du cas synthétique, nous avons appliqué les métamodèles par krigeage qui ont une meilleure précision par rapport à RBF. Ces métamodèles interpolent les réponses de ℓdX sur un espace de dimension 19, dont deux coefficients pour la prévision météorologique, neuf termes sources et huit paramètres scalaires.

5.2.1 Calibration du dépôt avec observations corrélées

Nous choisissons le dépôt comme première application de l'échantillonnage par MCMC aux mesures radiologiques. Cette variable représente l'accumulation des radionucléides déposés au sol pendant l'intégralité de la durée de l'accident. Elle ne porte pas d'information chronologique. Ainsi plusieurs scénarios peuvent produire une bonne carte de dépôt final. Cependant, elle a été effectuée sur un maillage très fin et peut apporter beaucoup de contraintes. La distribution a posteriori selon la procédure bayésienne risque alors d'être très concentrée, et la convergence des chaînes de Markov pourrait être très longue. Pour cette raison, un calcul pour moyenner les observations sur le maillage de simulation a été mis en place. L'explication du calcul de la moyenne a été détaillé dans le paragraphe précédent 5.2. De plus, nous avons en mis place une corrélation entre les observations, avec une distance de référence L de 10 km. Nous rappelons que le maillage de simulation fait $0.25^\circ \times 0.25^\circ$, soit environ $25 \text{ km} \times 25 \text{ km}$. Le choix de distance de référence à 10 km signifie que deux mailles adjacentes sont fortement corrélées. La corrélation entre les observations joue le rôle de grouper les points de mesures qui sont proches et donc de diminuer la redondance d'information portée par les observations.

Dans ce cas test, cinq chaînes de Markov ont été lancées en parallèle en partant de cinq points de départ distincts. Après deux millions d'itérations, les cinq chaînes se regroupent toutes dans une zone. La convergence a été mesurée par [Brooks 1998], cf. figure 5.1. Le rapport R_c entre deux matrices de covariance intra-chaîne et inter-chaîne est proche de 1 : cela signifie que les cinq chaînes sont d'accord sur la zone optimale de l'espace des entrées.

Une fois la convergence de l'échantillonnage vérifiée, nous souhaitons évaluer si les chaînes de Markov arrivent à retrouver le point de référence. Les histogrammes de l'échantillonnage 5.2 affichent les valeurs les plus vraisemblables de chaque paramètre, et la ligne rouge représente la valeur de référence, qui est utilisée comme observation. Pour certains paramètres, les chaînes de Markov ont réussi à localiser plus ou moins la valeur de référence, comme le facteur de lessivage ou les termes sources. Cependant, trois parmi dix neuf paramètres ont une distribution

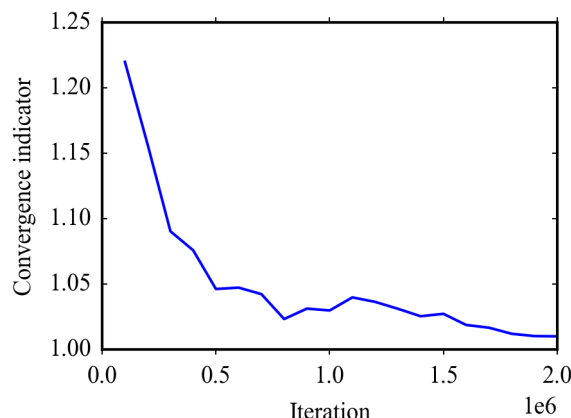


FIGURE 5.1 – Indicateur de convergence [Brooks 1998] sur la calibration du dépôt avec une corrélation entre les observations

complètement distincte du point optimal : c'est le cas de la pluie, du décalage en temps et de la vitesse de dépôt sec. Nous avons ensuite fixé ces trois paramètres aux valeurs de référence et fait varier les autres. Alors, d'autres paramètres qui étaient correctement retrouvés sont devenus mal identifiés. Ce problème est dû aux approximations du métamodèle, ce que nous montrons ci-dessous.

L'échantillonnage par Metropolis-Hastings est fondé sur la loi *a posteriori*, et la distribution obtenue est proportionnelle au produit de la vraisemblance et de la distribution *a priori*. La figure 5.3 présente les coupes du logarithme de la vraisemblance suivant les trois variables : décalage en temps, facteur de dispersion de la pluie et vitesse de dépôt sec. Ce sont les variables dont les distributions *a posteriori* ne retrouvent pas la valeur de référence qui est la droite rouge dans la figure 5.2. Théoriquement, les distributions issues de Metropolis-Hastings doivent retrouver exactement les valeurs de référence qui ont été utilisées pour créer les observations synthétiques. L'idée de réaliser des coupes de la vraisemblance a pour objectif d'identifier l'origine du problème : l'algorithme d'échantillonnage ou l'erreur du métamodèle. Si le maximum de la vraisemblance se trouve au sommet des distributions, cela signifie que l'algorithme fonctionne bien et que l'écart entre les distributions et leurs valeurs de référence vient de l'erreur du métamodèle. Nous nous intéressons seulement à la vraisemblance et pas au produit entre la vraisemblance et l'*a priori* car nous avons supposé que les distributions *a priori* sont uniformes et donc vont se simplifier dans le calcul de la probabilité d'acceptation de Metropolis-Hastings. Ici, nous affichons le logarithme car les valeurs de la vraisemblance se trouvent dans l'intervalle $[0, 1]$. Donc, avec la plage de variation des paramètres, les courbes peuvent être fortement aplaties, provoquant des difficultés d'interprétation. Selon ces coupes, la valeur de référence, les traits rouges dans la figure 5.2, ne se trouve pas au sommet de la courbe. En même temps, les histogrammes ont la même forme que les courbes, ce qui signifie que l'algorithme de Metropolis-Hastings a bien convergé. L'écart de l'échantillonnage vient de l'erreur du métamodèle.

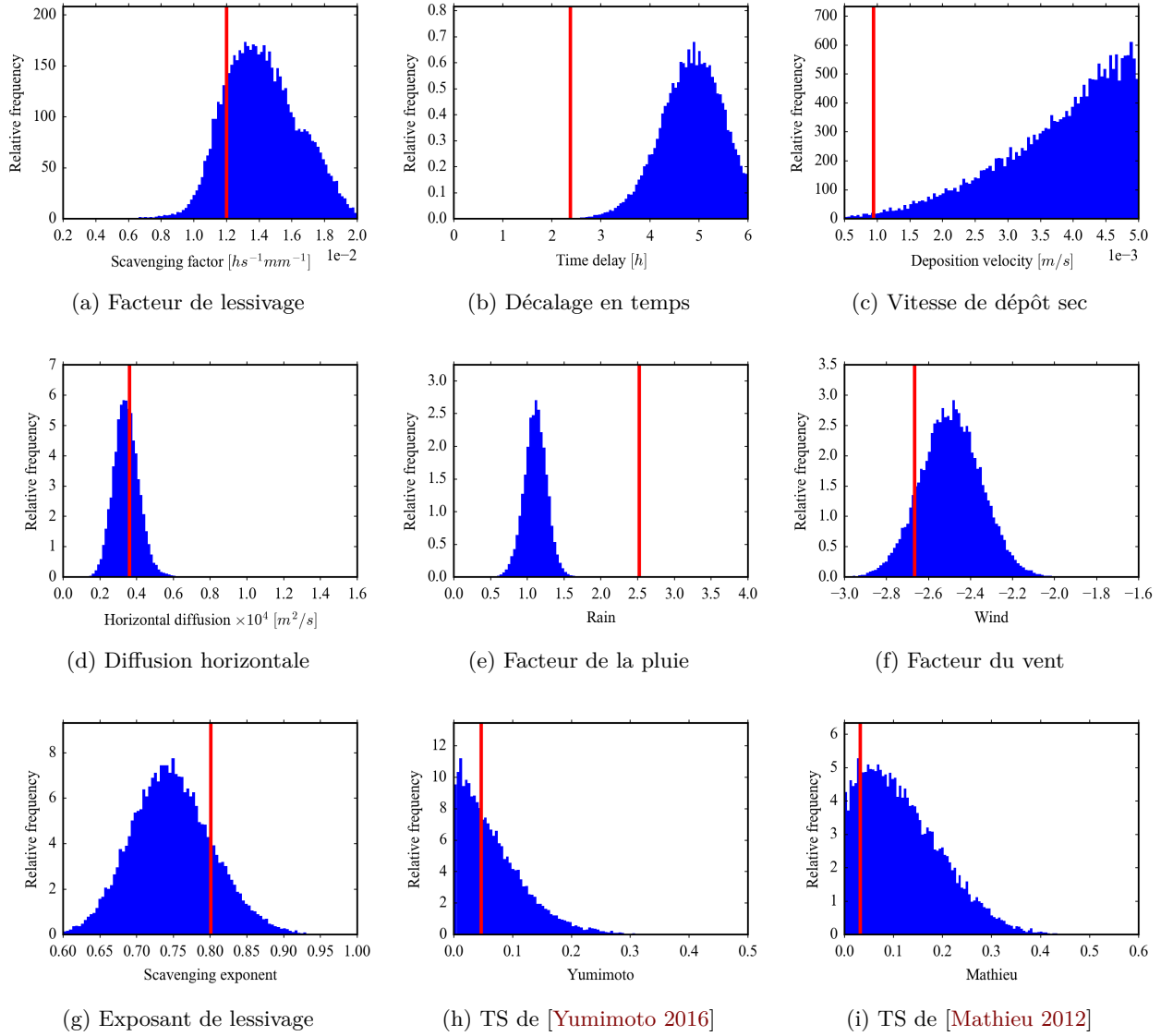


FIGURE 5.2 – Histogramme par échantillonnage MCMC de tous les paramètres. Les lignes rouges sont les valeurs optimales utilisées au début pour créer les observations synthétiques.

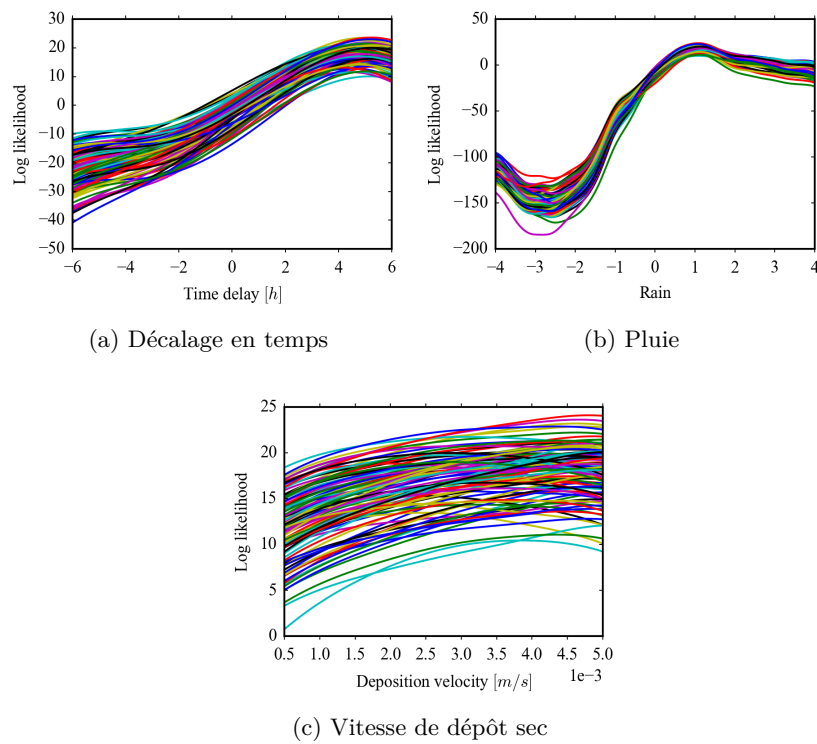


FIGURE 5.3 – Coupes du logarithme de la vraisemblance suivant les variables pour lesquelles l'approche MCMC n'a pas convergé vers les valeurs optimales

5.2.2 Calibration d'activité volumique intégrée sans corrélation entre les observations

Malgré la mise en place de corrélation entre les points d'observation, la calibration du dépôt, section 5.2.1, montre que les chaînes de Markov convergent difficilement dans certains cas. En réalité, il existe toujours une corrélation entre les observations. Dans le cas du dépôt final, la corrélation est systématiquement conseillée car les points de mesure sont vraiment très proches l'un de l'autre. Ils sont de plus collectés par le même appareil de mesure. Pour l'activité volumique, nous ne disposons que d'une centaine de stations d'observation, qui sont dispersées sur toute l'île de Honshu avec une distance moyenne de quelques kilomètres (cf. figure 2.6). C'est pourquoi nous avons supposé que la corrélation entre observations n'est pas nécessaire dans ce cas.

Puisque il était difficile d'obtenir la convergence dans le cas du dépôt final, nous avons d'abord commencé par la calibration en 1D afin de détecter où se trouve le problème de convergence des chaînes de Markov. Cela signifie que tous les paramètres d'entrée ont été fixés sauf un seul. Le cas d'étude suivant présente la calibration du facteur de dispersion du vent basée sur les mesures de concentration. Nous rappelons que le facteur de dispersion du vent dans ce chapitre correspond au facteur β de l'équation 4.2 dans la calibration du modèle météorologique simplifié.

Cas 1D : facteur de dispersion du vent

Dans ce cas en dimension 1, les chaînes de Markov s'accordent bien sur la localisation de la valeur de référence. La convergence est atteinte assez rapidement. Le sommet de la distribution ne tombe pas exactement sur la valeur attendue (figure 5.4) à cause de l'erreur du métamodèle. Cependant, la distance entre la moyenne des échantillons et la valeur de référence reste faible :

$$\frac{|x_{opt} - x_{mean}|}{\sigma_{post}} = 0.5214 \quad (5.5)$$

L'équation 5.5 implique bien que la moyenne *a posteriori* et la valeur de référence ne sont pas confondues. La vraie valeur optimale se trouve quand même dans la zone de forte probabilité, à moins d'une fois l'écart type de la moyenne obtenue. De plus, la loi *a priori* du vent est une loi uniforme d'écart type 3. Avec la méthode d'échantillonnage par MCMC, nous avons réussi à construire une distribution de probabilité *a posteriori* avec un écart type beaucoup plus faible : 0.82. Cela signifie qu'une partie des incertitudes des paramètres d'entrée a été absorbée par le processus d'échantillonnage.

Cas 19D : tous les paramètres d'entrée

La calibration de l'activité volumique intégrée a fonctionné pour le cas d'un seul paramètre d'entrée. Nous nous intéressons maintenant à l'application sur tout l'espace des entrées, de dimension 19. De la même façon, cinq chaînes de Markov ont été lancées en parallèle et arrêtées au bout d'un million itérations. L'indicateur de [Brooks 1998] signale bien le début de la convergence des chaînes, cf. figure 5.5.

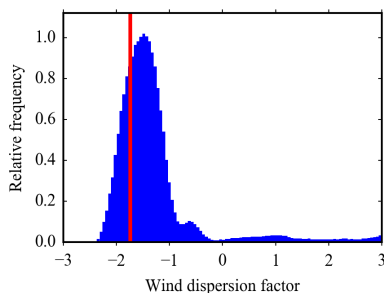


FIGURE 5.4 – Histogramme de MCMC en 1D basé sur la concentration intégrée. La ligne rouge indique la valeur optimale que l'on cherche à approcher.

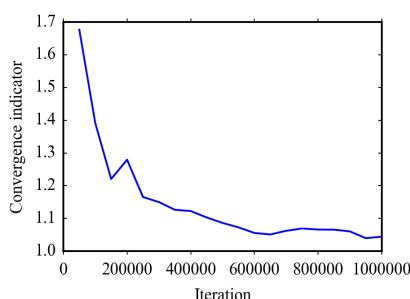


FIGURE 5.5 – Indicateur de convergence de [Brooks 1998]

Les histogrammes de probabilité ont été également tracés. Cependant, dans ce cas test, certains paramètres adoptent un histogramme avec un maximum bien marqué (facteur de dispersion du vent ou exposant de lessivage). D'autres paramètres ont une distribution beaucoup plus plate (facteur de lessivage et décalage en temps). Cf. figure 5.6. En général, la combinaison convexe du terme source est bien retrouvée. Une explication probable pour ces distributions aplaties est que ces paramètres sont peu sensibles à la vraisemblance de ce type de sortie.

5.3 Cas réel : observations lors de l'accident de Fukushima

Les tests avec des mesures synthétiques permettent de vérifier non seulement le fonctionnement de l'algorithme de Metropolis-Hastings, mais aussi l'impact des erreurs de métamodèle dans le résultat de la calibration. Dans certains cas, la méthode d'émulation par krigeage avait des difficultés à trouver la matrice de covariance optimale pour construire le métamodèle. Ce phénomène peut très bien se reproduire dans l'étude suivante, avec des mesures réelles de Fukushima. Donc, les problèmes rencontrés dans la section précédente doivent être pris en compte afin d'éviter des difficultés similaires.

Chaque procédure de calibration par Metropolis-Hastings utilise une fonction de vraisemblance différente. Cela signifie qu'un métamodèle doit être créé pour chaque cas, et l'évaluation de métamodèle est indispensable. Ici encore, nous nous concentrons sur deux types de sorties : l'activité volumique et le dépôt final. L'espace des entrées est de dimension 19 : deux facteurs

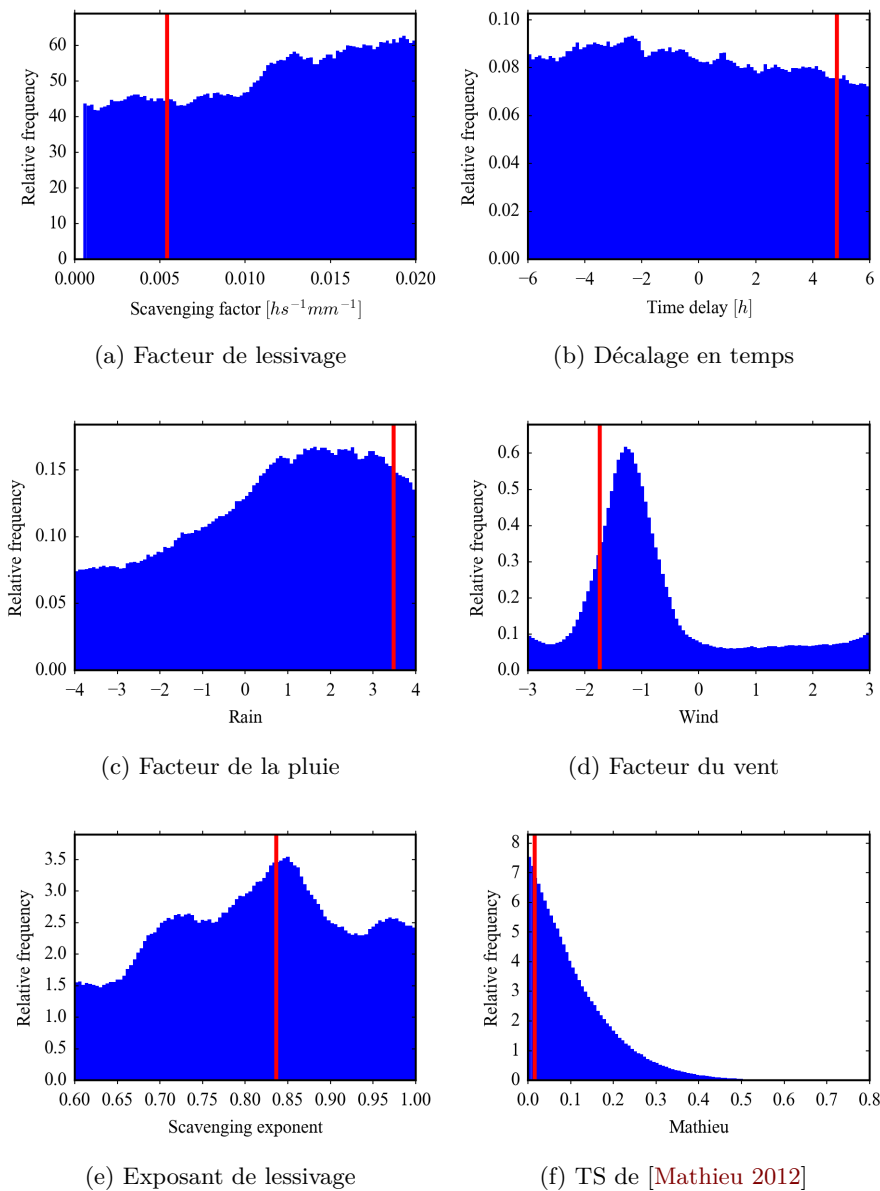


FIGURE 5.6 – Résultat du MCMC sur la calibration de l'activité volumique intégrée en dimension 19. Les lignes rouges sont les valeurs optimales utilisées au début pour créer les observations synthétiques.

météorologiques, neuf coefficients de terme source et huit paramètres scalaires.

5.3.1 Évaluation des métamodèles utilisés

La procédure a été construite en parallèle sur le débit dose, l'activité volumique et le dépôt final. La calibration du débit dose a été effectuée en premier avec l'interpolation par RBF, alors que l'activité volumique et dépôt final ont été exploités avec un métamodèle construit par krigeage. La construction de métamodèle par RBF est beaucoup moins coûteuse : environ deux à trois minutes sont nécessaires pour RBF contre trois heures pour un métamodèle par krigeage. Cependant, RBF est moins précis que le krigeage dans la prédiction des sorties du modèle. En général, il faut compter un facteur 3 sur les scores d'erreur modèle-métamodèle entre RBF et le krigeage. Dans ce paragraphe, nous présentons l'évaluation des métamodèles par krigeage pour les trois types de sorties.

Dans la vraisemblance, nous avons considéré que les observations de débit de dose sont indépendantes entre elles, et qu'il n'y a donc pas de corrélation entre les mesures pour la calibration de ces sorties. Pour la concentration intégrée, il était impossible d'appliquer une corrélation entre les observations, sous la forme de 5.4, dans la fonction de vraisemblance. En effet, peu importe la valeur de distance de référence L , la matrice de covariance Σ des mesures n'est jamais définie positive. Le déterminant de son inverse est souvent à l'infini et donc, le calcul de la vraisemblance sur le DOE est incorrect. Nous allons présenter uniquement le cas d'indépendance entre les observations pour la concentration intégrée.

Le dépôt final n'a pas de problèmes avec la corrélation entre les mesures. Mais avec un ensemble d'observations indépendantes, la calibration du dépôt n'a pas convergé. Les chaînes de Markov restent bloquées dans des minima locaux. Chaque observation indépendante porte une information qui sera prise en compte durant le déroulement de l'algorithme MCMC. L'accident de Fukushima est un cas d'étude complexe dont les mesures contiennent également des incertitudes. Comme dans la calibration du cas synthétique, le maillage des mesures aéroportées du dépôt final est vraiment très fin (environ 1 km). L'appareil utilisé a survolé l'île de Honshu et capté des valeurs dans la zone programmée (cf. figure 2.7). Si nous utilisons toutes ces mesures, cela entraînera une forte contrainte et les distributions *a posteriori* risquent d'être très concentrées, et donc très difficiles à échantillonner. La calibration utilisant des valeurs moyennées sur le maillage avec une corrélation d'observation facilite la convergence des chaînes de Markov dans la calibration statistique.

En conclusion, nous avons appliqué une vraisemblance sans corrélation pour la calibration de l'activité volumique intégrée et du débit de dose, et une vraisemblance avec corrélation pour le dépôt final. Pour évaluer le métamodèle, nous utilisons le score SMSE comme dans les parties précédentes. De plus, afin d'adopter une vision globale sur la qualité des métamodèles, nous proposons d'interpréter un deuxième score, noté NSMSE—Normalized Standardized Mean Squared Error, dont la formule est

$$NSMSE = \frac{1}{n} \sum_{i=1}^n \left(\frac{y_i - f(p_i)}{y_i} \right)^2 \quad (5.6)$$

Comme tous les scores de la famille *Mean Squared Error*, la NSMSE ne contient que des valeurs positives ou nulles. Pour interpréter ce score, une valeur proche de 0 indique une prévision

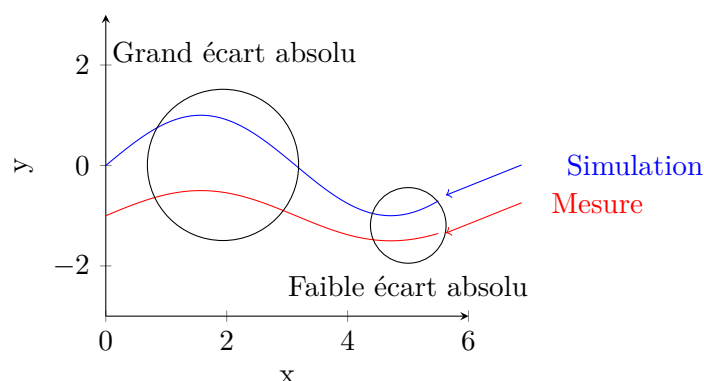


FIGURE 5.7 – Illustration de l'écart modèle-mesure

de bonne qualité, c'est-à-dire qu'elle est proche des observations. Plus la NSMSE est grande, plus la prévision s'écarte des mesures. L'interprétation de la NSMSE ne contredit pas celle de la SMSE mais elle joue plutôt un rôle complémentaire. La différence dans la formule entre la SMSE 3.12 et NSMSE 5.6 indique que la SMSE a tendance à privilégier la différence modèle-mesure absolue, tandis que la NSMSE se concentre sur les écarts relatifs. En effet, les grands écarts modèle-mesure sont plus importants dans le score SMSE. Souvent, les grands écarts modèle-mesure sont liés aux fortes valeurs ; nous avons alors tendance à oublier les petites différences, qui possèdent souvent une erreur relative non négligeable. Pour la NSMSE, ces différences modèle-mesure sont normalisées avant d'être moyennées. Cela permet de rattraper la différence en valeur absolue des écarts modèle-mesure. La figure 5.7 illustre les deux types d'écart modèle-mesure abordés auparavant.

Dans le tableau 5.2, les métamodèles sont de bonne qualité selon le SMSE mais moins avec le NSMSE. La figure 5.8 représente la prédiction du métamodèle en comparant à la réponse du modèle physique sur la fonction de vraisemblance. Les métamodèles créés émulent assez bien la fonction d'intérêt, la vraisemblance.

	Concentration intégrée	Dépôt final de seuil 10 kBq/m ³
SMSE	0.0669	0.1464
NSMSE	1.2579	6.0519

TABLE 5.2 – Évaluation des métamodèles de la vraisemblance avec des mesures réelles de Fukushima.

5.3.2 Calibration avec le débit dose et une corrélation entre les observations

Nous avons commencé la calibration des mesures radiologiques de Fukushima par les stations de débit dose. Une corrélation spatiale est appliquée entre les stations, selon la formule 5.4, car les appareils de mesure captent le rayonnement gamma dans un voisinage de quelques kilomètres. Le métamodèle utilisé dans la calibration des mesures de débit de dose repose sur une interpolation par RBF (et non un krigeage). Cette méthode a l'avantage que le temps de construction des émulateurs est assez court. Un émulateur peut être construit en quelques minutes. Bien que

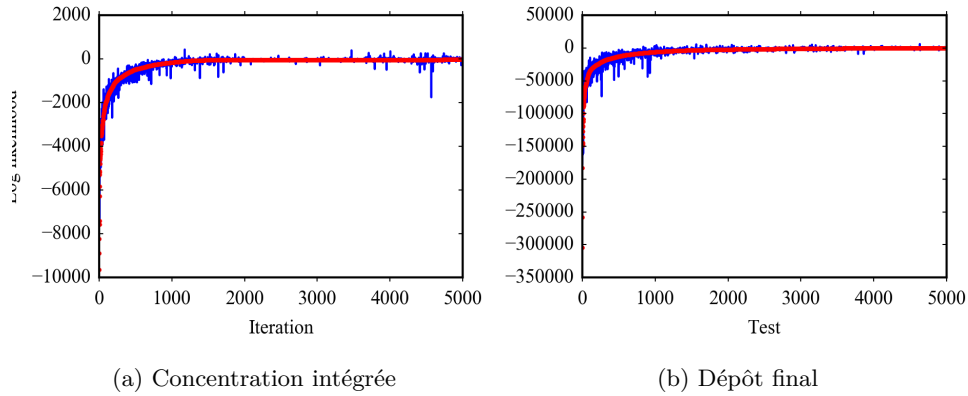


FIGURE 5.8 – Comparaison de la prédiction du métamodèle de la vraisemblance à la réponse de ldX sur LHS complémentaire. La ligne bleue est la prédiction du métamodèle et la rouge est le résultat des tests de LHS complémentaire.

la précision de ces métamodèles soit inférieure à celle avec krigeage, cette étape nous permet d’avoir une vision rapide sur le résultat de la calibration par MCMC.

Les distributions résultant de la calibration du débit dose se trouvent à la figure 5.9. L’algorithme de Metropolis-Hastings a réussi à trouver une valeur optimale pour certains paramètres, comme les composantes du coefficient de lessivage ou la diffusion verticale K_z . Les termes sources les plus probables, suivant les tirages MCMC, sont [Saunier 2013], [Saunier 2016] et [Winiarek 2012]. En revanche, il existe des paramètres dont l’histogramme des tirages est complètement écrasé, comme l’amplitude de rejet, décalage en temps ou les facteurs météorologiques. En réalité, il n’est pas possible de prendre en compte toutes les incertitudes de modélisation physique. L’histogramme de l’amplitude de rejet, du décalage en temps et les paramètres météorologiques confirment cette tendance. Une partie des incertitudes ne sont pas prises en compte, et l’algorithme d’échantillonnage doit augmenter artificiellement la dispersion sur les paramètres, ici considérés incertains, pour pouvoir augmenter l’écart modèle - mesure. D’autre part, le fait d’utiliser un métamodèle ajoute d’incertitude sur les simulations. C’est pourquoi il est important d’assurer que le métamodèle imite bien le modèle physique.

5.3.3 Calibration avec l’activité volumique et sans corrélation

Comme nous avons vu dans le chapitre précédent, plusieurs problèmes peuvent apparaître, et l’obtention d’une convergence sur tous les paramètres est très difficile. Nous avons découpé ce problème en petites étapes. D’abord, nous faisons varier seulement le facteur du vent et toutes les autres variables sont fixées—cf. section 5.3.3. Deuxièmement, les paramètres liés à la météorologie et tous les paramètres scalaires sont fixés ; seule la combinaison des termes sources varie. Ensuite, nous calibrons la météorologie et le terme source ; et puis, nous avançons vers la calibration totale avec la variation de tous les paramètres d’entrée. Pour chaque test, les paramètres invariants sont fixés à la valeur du centre de leur plage de variation.

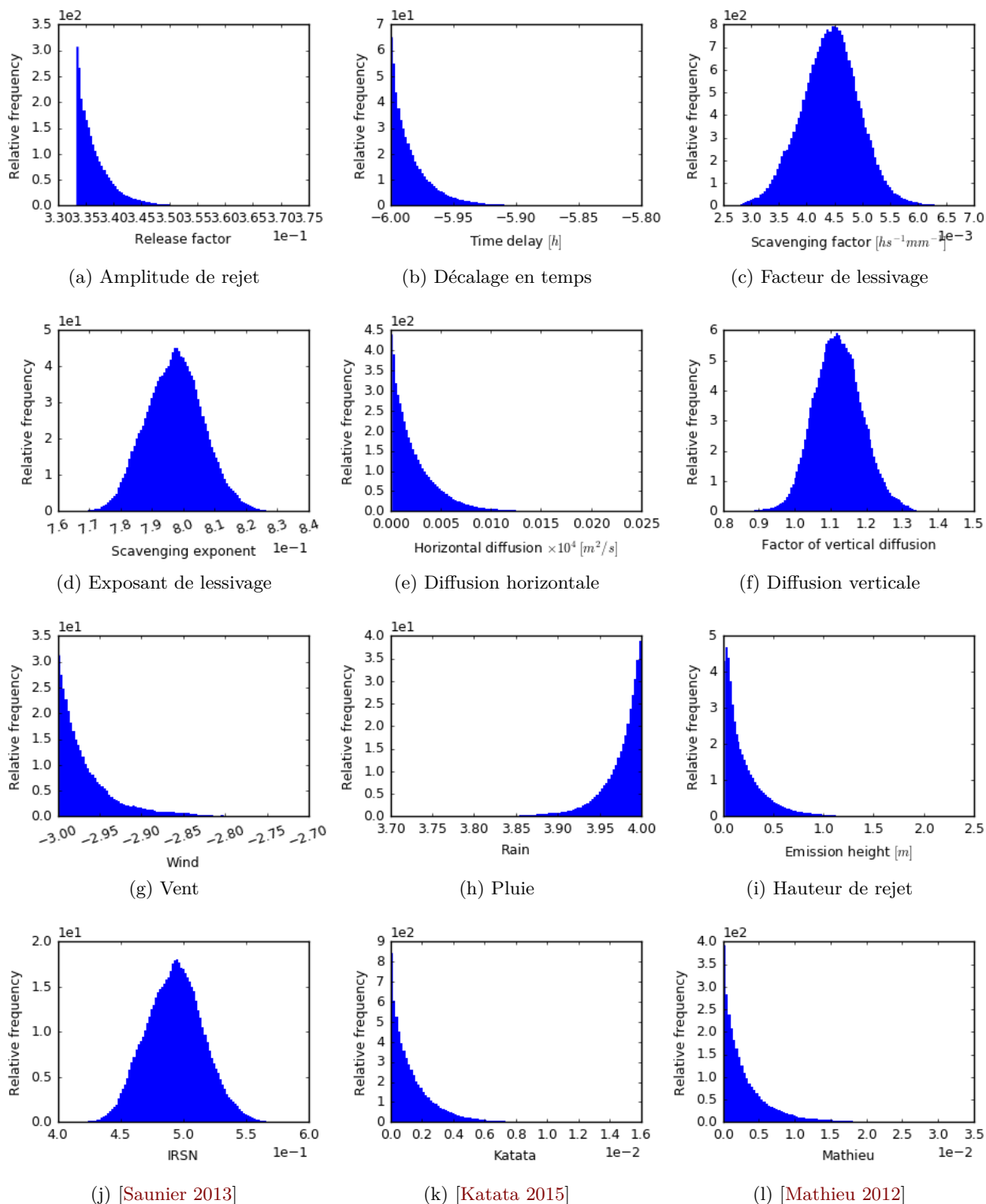


FIGURE 5.9 – Résultat de calibration de débit dose sur toutes les stations pendant la période entière

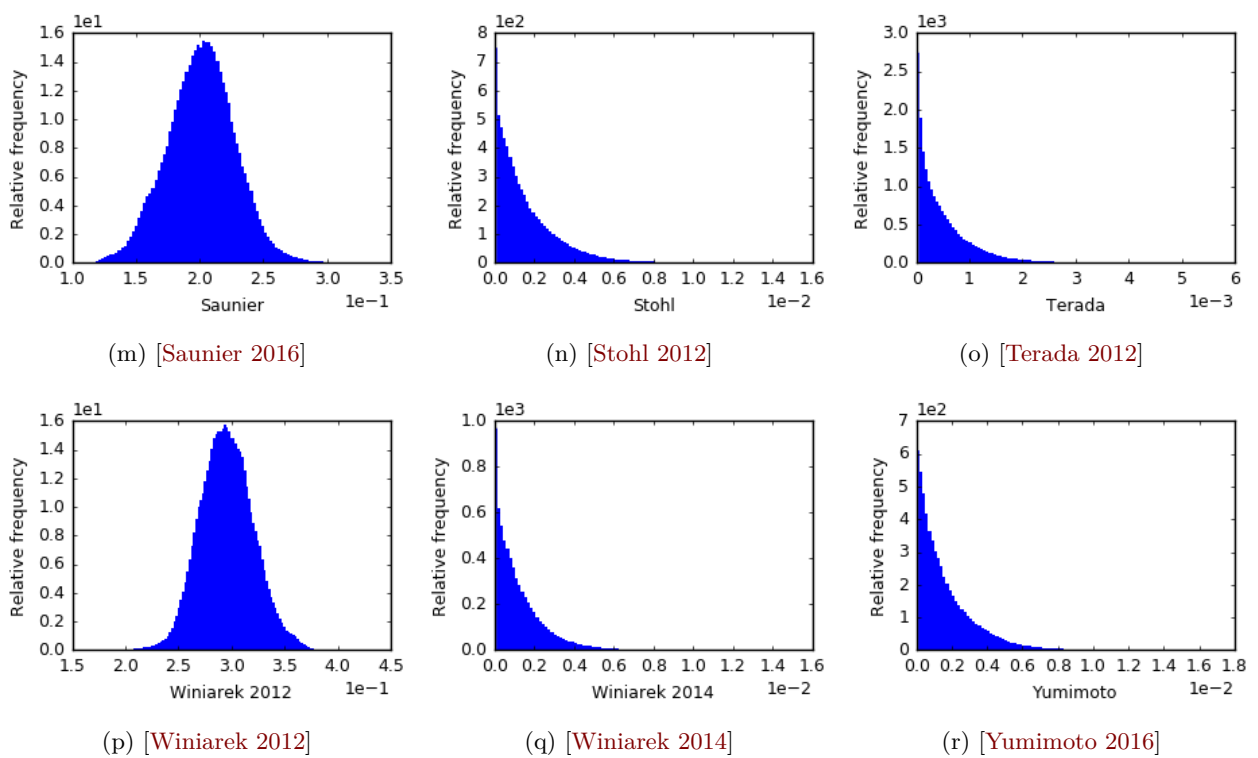


FIGURE 5.9 – Résultat de calibration de débit dose sur toutes les stations pendant la période entière

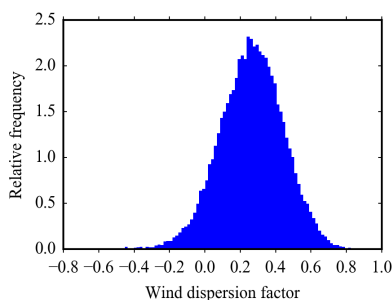


FIGURE 5.10 – Résultat de la calibration de l’activité volumique en faisant varier seulement le facteur du vent

Facteur de dispersion du vent

La calibration d’un seul paramètre se déroule avec succès. En dimension 1, la vraisemblance adopte un maximum bien net. Les chaînes de Markov convergent rapidement. Le sommet de la distribution présentée à la figure 5.10 se trouve à 0.3; cela signifie que le vent optimal qui minimise l’erreur modèle-mesure se trouve au centre de l’ensemble de prévisions météorologiques.

Terme source

Pour augmenter le niveau de complexité, cinq chaînes de Markov ont été lancées en faisant varier seulement la combinaison des termes sources. Les autres paramètres sont fixés au centre de plage de variation. La figure 5.11 montre que deux termes sources ont un coefficient nettement plus élevé que les autres. Cela signifie que le rejet correspondant au panache du 14-16 mars peut être représenté par une combinaison des termes sources [Winiarek 2012] et [Yumimoto 2016]. Les autres termes sources, ayant un poids beaucoup plus faible, pourront être prépondérants lors de la calibration sur une autre période.

Combinaison de la météorologie et du terme source

Maintenant, nous combinons les paramètres de la météorologie et du terme source. Les chaînes ont des difficultés de convergence : chaque chaîne reste logée dans un minimum local, comme l’illustre la figure 5.12. Cette figure présente la valeur des échantillons de deux variables acceptés par l’algorithme de MCMC : le facteur sur les précipitations et le poids donné au terme source [Stohl 2012]. Quand on augmente le nombre de variables incertaines, l’écart type de la distribution recherchée peut être faible, voire négligeable, selon certaines directions (liées aux paramètres ou à une combinaison de paramètres). Cela implique une distribution très pointue et donc difficile à échantillonner.

Pour expliquer l’absence de convergence des chaînes de Markov dans la calibration des mesures de l’activité volumique, une hypothèse possible est que les chaînes sont entraînées dans des minima locaux, qui possèdent le même poids dans la distribution *a posteriori*. Autrement dit, il est possible que l’erreur modèle-mesure de ces minima locaux soit équivalente. Les itérations d’une chaîne de Markov ont tendance à converger vers un jeu de paramètres qui sont un

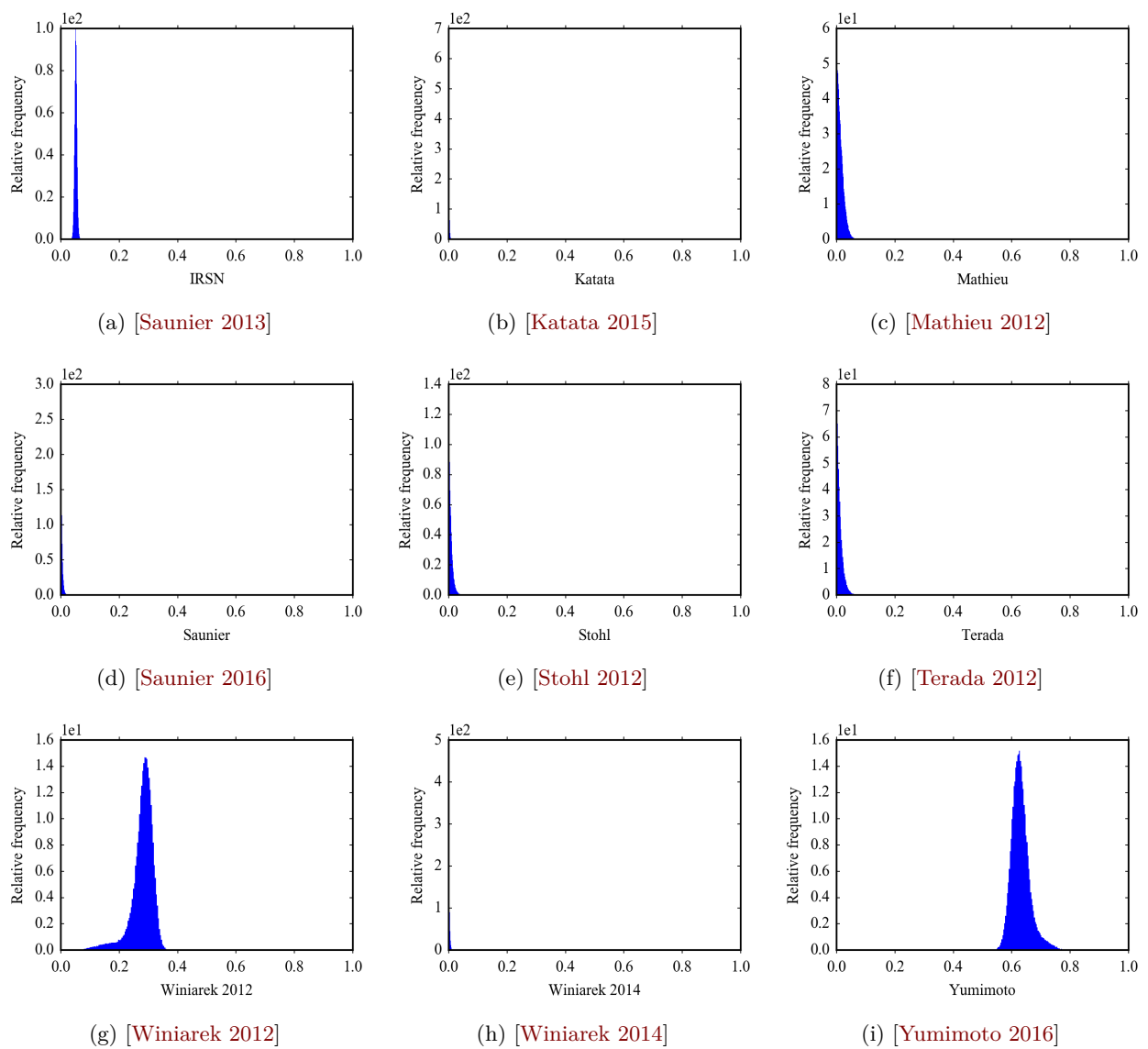


FIGURE 5.11 – Résultat de calibration pour le terme source sur la concentration intégrée du 14 au 16 mars

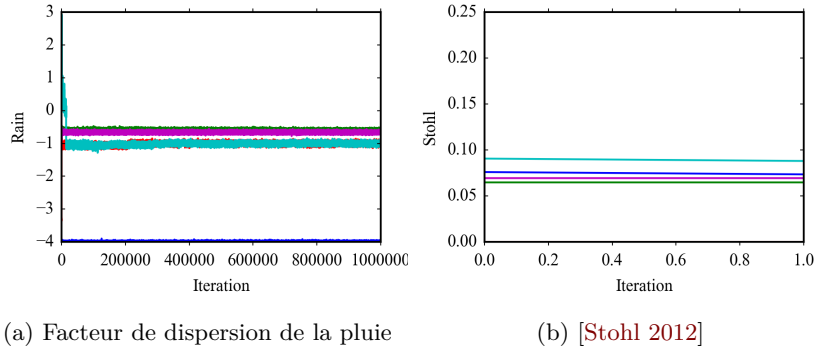


FIGURE 5.12 – Résultat de MCMC pour les paramètres météorologiques et le terme source sur la concentration intégrée

maximum de la distribution *a posteriori*. Si cette dernière possède un grand nombre de maxima locaux de même poids, par exemple avec une distribution plate, les chaînes de Markov vont avoir des difficultés à identifier un maximum global. Pour vérifier cette hypothèse, deux cents simulations directes ont été réalisées. Les jeux de paramètres d'entrée de ces simulations ont été choisis parmi les échantillons issus de l'algorithme Metropolis-Hastings. Nous avons reporté dans le tableau 5.3 la RMSE modèle-mesure de ces deux cents simulations directes. On peut comparer avec le tableau 2.7 de la partie de propagation par Monte Carlo.

L'algorithme de MCMC a réussi à réduire l'incertitude de l'activité volumique, car l'écart entre les valeurs minimales, moyenne et médiane est plus faible. Nous trouvons également que la RMSE moyenne pour les seuils de 0.1 Bq/m^3 et 1 Bq/m^3 est plus basse que celle des simulations Monte Carlo. Cela signifie que les simulations se rapprochent des observations. En revanche, le seuil 10 Bq/m^3 n'est peut-être pas le bon seuil pour vérifier cette hypothèse. Avec un seuil élevé, il y a un risque de filtrer un grand nombre de mesures, et avec le décalage en temps modèle-mesure, les scores comme la RMSE peuvent augmenter très rapidement. La figure 5.13 visualise, en échelle logarithmique, l'activité volumique de ces deux cents simulations sur deux stations de mesure (Sukagawa et Yabuki), en correspondance de la comparaison modèle-mesure effectuée dans la partie 2. Les stations Yabuki et Sukagawa, qui se trouvent tous les deux dans une maille de simulation ℓdX , ont été choisies pour démontrer l'effet sous maille, qui pourrait provoquer un résultat faussé dans la comparaison modèle-mesure. D'autre part, nous avons mis la station Asahi, à 80 km à l'est de Tokyo ou 200 km au sud de la centrale Fukushima-Daiichi. La figure 5.14 montre les simulations *a posteriori* de l'activité volumique à la station Asahi sur les deux périodes importantes de l'accident : du 14 au 16 mars sur la colonne gauche et du 20 au 22 mars sur la colonne droite. Enfin, la figure 5.15 donne un exemple de carte de dépôt final d'une de ces deux cents simulations *a posteriori*. Ici, nous avons calibré le modèle ℓdX suivant les observations de l'activité volumique et non du dépôt final. La réalisation de cette carte a pour l'objectif de vérifier la relation entre deux types de sorties.

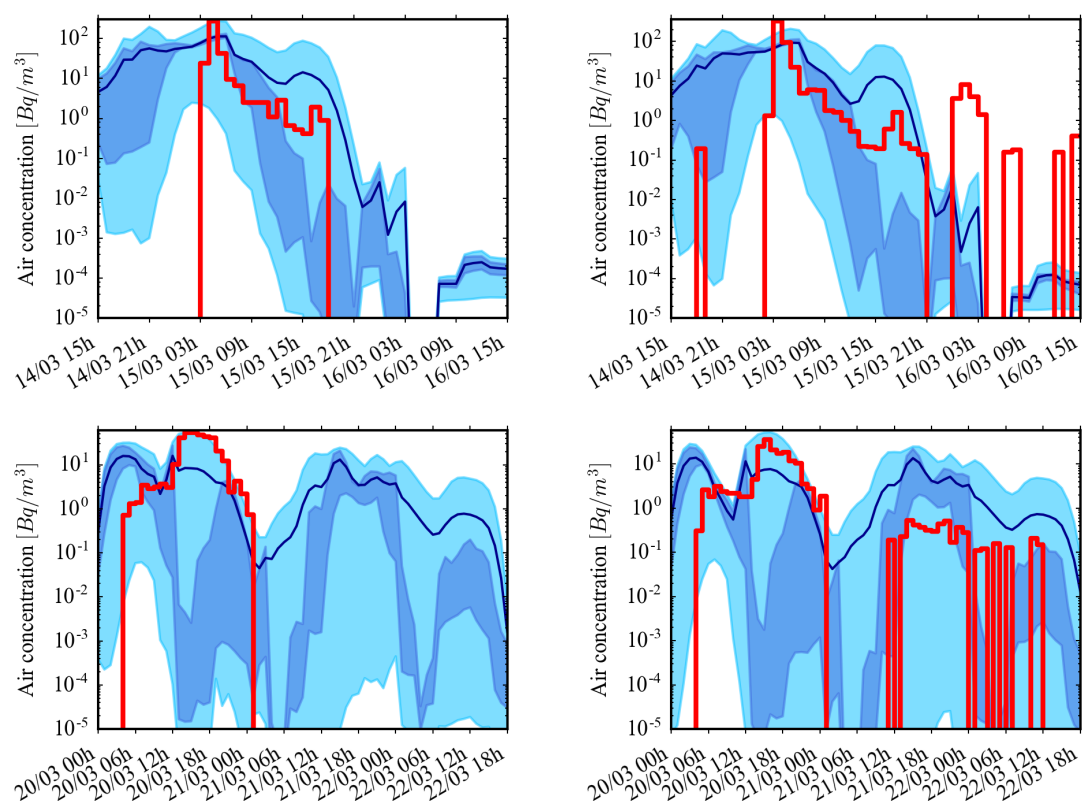


FIGURE 5.13 – Activité volumique aux stations Sukagawa et Yabuki après la calibration de MCMC sur la période du 14 au 16 mars. La ligne rouge représente les mesures de l'activité volumique. La bande bleue claire est l'enveloppe des deux cents simulations, la bande bleue foncée contient 50% des simulations et la ligne bleue marine est la moyenne.

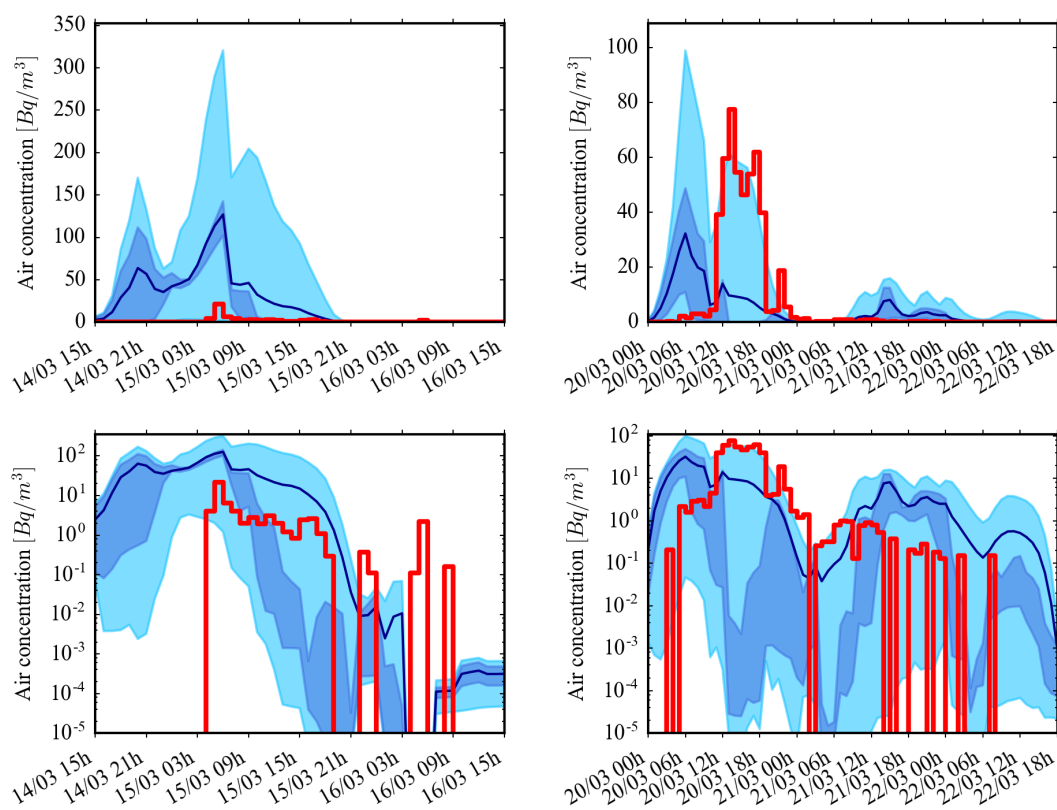


FIGURE 5.14 – Autre illustration de l'activité volumique à la station Asahi en échelle normale (en haut) et échelle logarithmique (en bas)

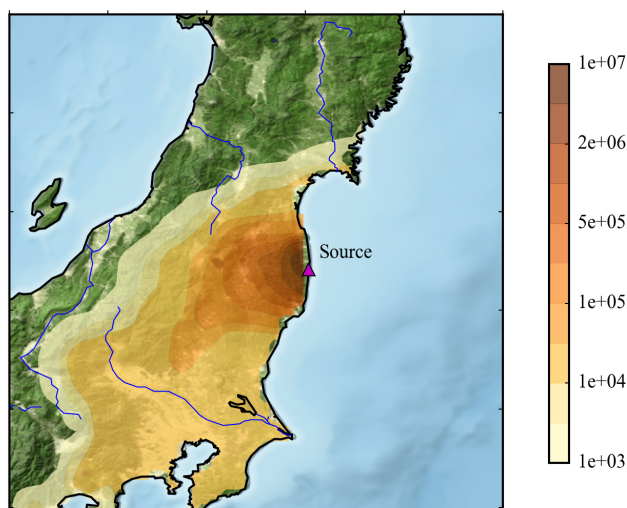


FIGURE 5.15 – Une carte de dépôt final issue des échantillons par MCMC

Seuil	Min	Moyenne	Médiane	Seuil	Min	Moyenne	Médiane
0.1 Bq/m^3	13.43	18.76	14.03	0.1 Bq/m^3	13.3	14.16	13.64
1 Bq/m^3	15.18	21.21	15.86	1 Bq/m^3	18.03	18.73	18.24
10 Bq/m^3	20.87	29.24	21.83	10 Bq/m^3	37.8	45.24	46.34

TABLE 5.3 – RSME modèle-mesure des résultats de propagation par Monte Carlo (gauche) et de MCMC (droite) sur l'activité volumique du 14 au 16 mars avec la combinaison des paramètres météorologiques et termes sources

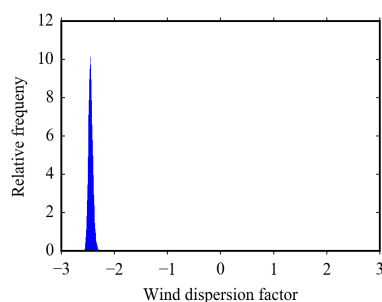


FIGURE 5.16 – Distribution obtenue après calibration du vent avec observation du dépôt final

5.3.4 Calibration avec le dépôt et avec corrélation

De la même façon, nous avons effectué la calibration du dépôt final en plusieurs étapes : seulement le vent, figure 5.16; ensuite le terme source. En exploitant le dépôt final, les chaînes de Markov ne convergent pas au niveau des termes sources, comme l'illustre la figure 5.17. Comme le dépôt est une mesure qui n'a pas de notion temporelle, plusieurs scénarios peuvent donner une bonne carte de dépôt final mais avec des événements intermédiaires complètement différents.

Facteur de dispersion du vent

Terme source

Dans les figures 5.12 et 5.17, les chaînes de Markov se logent dans certains coins de l'espace d'entrée de grande dimension. Cela signifie que la fonction coût, induite par la loi *a posteriori*, possède un certain nombre de minima locaux. Malgré la non convergence des chaînes de Markov, nous avons réalisé les simulations directes avec les échantillons obtenus par l'algorithme de Metropolis-Hastings. La RMSE modèle-mesure a été calculée sur deux cents simulations, dont les paramètres d'entrée ont été prélevés aléatoirement dans les chaînes de Markov—cf. tableau 5.4. En comparant avec le résultat de Monte Carlo du tableau 2.3, le résultat de la calibration paraît moins intéressant. Une hypothèse proposée pour expliquer ce phénomène est que parmi les paramètres d'entrée, certains ont des influences négligeables sur la sortie du modèle, ou autrement dit sur la fonction coût. La perturbation de ces variables n'apporte pas assez d'informations pour identifier les distributions optimales de l'ensemble des paramètres.

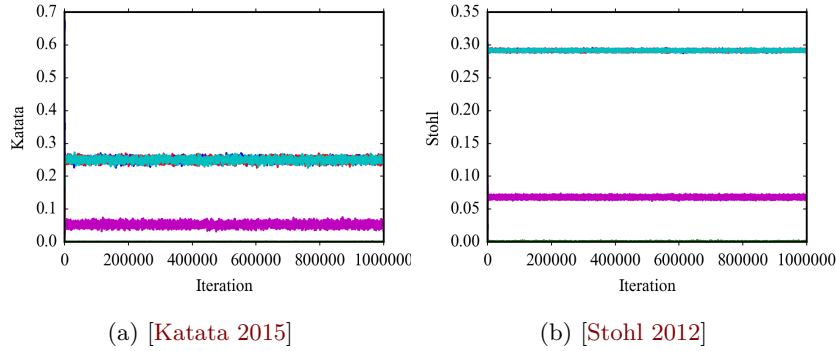


FIGURE 5.17 – Résultat de calibration de terme source avec observation du dépôt final : valeurs de la pondération pour deux termes sources, en fonction du temps

Une suggestion pour de futures études est de fixer certaines variables, comme les coefficients de lessivage, qui ont une faible répercussion sur les sorties du modèle numérique étudiées ici.

	Min	Moyenne	Médiane
RMSE sur toutes les cellules	92.01	99.66	103.16
RMSE > 80 km	42.66	58	66.73

TABLE 5.4 – RMSE modèle-mesure suivant le résultat de MCMC sur le dépôt final. Tous les paramètres sont fixés sauf la combinaison des termes sources.

5.4 Conclusion

Ce chapitre se concentre surtout sur la calibration de la première période de l'activité volumique. Malgré l'absence de convergence des chaînes de Markov avec la combinaison des paramètres météorologiques et le terme source, les simulations *a posteriori* semblent montrer une réduction d'incertitude des paramètres d'entrée du modèle ℓdX . Il serait intéressant de réaliser la même étude sur la deuxième période de l'accident, du 20 au 22 mars, et ensuite de comparer les résultats obtenus entre ces deux périodes. Plus précisément, il est important de vérifier si les valeurs privilégiées dans les distributions *a posteriori* sont les mêmes pour chaque paramètre. Par exemple, nous pourrions voir s'il s'agit de la même combinaison de termes sources, ou les mêmes facteurs météorologiques liés à la dispersion.

D'autre part, la calibration réelle du dépôt final a posé plus de difficultés que dans le cas synthétique. Avec les études précédentes, par exemple [Mathieu 2018a], aucune simulation n'a réussi à fournir une carte de dépôt final de manière satisfaisante. Le point le plus difficile de ce type de sortie est l'absence de contrainte temporelle. Cela signifie que plusieurs simulations peuvent donner des cartes de dépôt similaires, avec des valeurs des paramètres d'entrée parfois irréalistes. Pour améliorer la calibration, nous proposons de fixer certains paramètres non essentiels et d'éliminer des termes sources peu probables, afin de rendre la distribution *a posteriori*

plus facile à échantillonner.

Pour le débit dose, une première version de la calibration a été faite en utilisant le méta-modèle par RBF. Cependant, ce métamodèle possède une forte erreur par rapport au modèle physique. Par la suite, il sera intéressant de construire un métamodèle sur ce type de sortie en exploitant le krigeage, et en espérant ainsi obtenir un métamodèle de meilleur qualité.

Conclusion

Cette thèse réalise une application des méthodes d'optimisation et de calibration statistique afin de quantifier les incertitudes dans des modèles de dispersion atmosphérique. Les principaux processus peuvent être classés comme suit :

- Comprendre le fonctionnement des modèles de dispersion et identifier les principales sources d'incertitude dans ce type de modèle numérique ;
- Étudier la propagation d'incertitude via le modèle de dispersion en perturbant les variables d'entrée sur des plages de variation conseillées par des experts ;
- Déterminer la méthode la plus adaptée pour quantifier les incertitudes. Ici, nous souhaitons construire une distribution de probabilité pour chaque paramètre d'entrée grâce à une calibration statistique adaptée ;
- Créer des outils nécessaires pour pouvoir appliquer la calibration statistique : construction du métamodèle, donc création d'un plan d'expérience et implantation de RBF et krigeage ;
- Appliquer des algorithmes de MCMC afin de retrouver la distribution des paramètres d'entrée en utilisant les mesures réelles.

Ces étapes produisent un processus continu dont l'objectif final est d'améliorer nos connaissances sur les incertitudes du modèle de dispersion atmosphérique ℓdX . Ce travail permet d'adopter une première vision sur le profil des distributions de probabilité des paramètres d'entrée de ce modèle. En situation de crise nucléaire, les modèles de dispersion seront utilisés afin d'estimer l'évolution des panaches radioactifs, ainsi que les conséquences environnementales à anticiper. Il est essentiel de mettre en place un outil opérationnel qui permet d'automatiser tout ce processus (construction du métamodèle, optimisation déterministe et calibration statistique).

Propagation d'incertitude par méthode de Monte Carlo

Bilan

Les modèles de dispersion présentés dans la partie 1.3 ont pour l'objectif de résoudre l'équation d'advection-diffusion sur l'évolution de la concentration des polluants dans l'atmosphère. Ici nous nous sommes concentrés sur la concentration des radionucléides au cours du temps. Chaque modèle apporte ses propres avantages et inconvénients. Le modèle à bouffées gaussien, pX , n'est pas très coûteux en temps de calcul. Il fournit une meilleure estimation à l'échelle locale, tandis que le modèle eulérien ℓdX repose sur la discrétisation de l'équation d'advection-diffusion en temps et en espace sur un maillage. Ce modèle est couplé à un modèle eulérien qui détermine les champs de vent et de turbulence sur le même maillage. Le modèle eulérien s'utilise plutôt pour les simulations à l'échelle régionale ou continentale. La résolution des

modèles eulériens est complexe car elle nécessite de résoudre un système d'équations dont les variables sont les concentrations des polluants dans chaque maille. Les modèles de dispersion sont intégrés dans la plateforme opérationnelle *c3X*, développée par l'IRSN. Les performances obtenues sont très satisfaisantes lors de la comparaison avec d'autres modèles opérationnels en cours utilisés. Trois principales sources d'incertitude ont été identifiées dans ces modèles, parmi lesquelles les données météorologiques et le terme source sont les plus essentielles dans les résultats du modèle. Certains paramètres dans la configuration du phénomène jouent également un rôle indispensable mais nous les avons perturbés de manière indépendante l'un de l'autre. Par exemple, les deux composantes du coefficient de lessivage, a et b dans la formule $\Lambda = ap_0^b$, sont considérées indépendantes entre elles, tandis qu'en réalité, il existe certainement une relation de corrélation entre ces deux paramètres. Les mesures collectées autour de la catastrophe de Fukushima (concentration, débit de dose et dépôt de ^{137}Cs) servent à vérifier et renforcer les informations *a priori* des entrées du modèle, dans un cadre déterministe ou probabiliste.

Résultats

L'objectif du chapitre 2 est de donner une vision préliminaire sur l'utilisation des mesures dans l'évaluation d'un ensemble de prévisions, via le diagramme de rang. Ce dernier représente la répartition des observations vis-à-vis des modèles de l'ensemble. Cet outil a été utilisé à la fois pour évaluer l'ensemble météorologique, qui sera une entrée de notre modèle *ldX*, et ensuite, évaluer l'ensemble des sorties radiologiques. La comparaison aux observations de l'ensemble météorologique montre qu'il existe d'importantes incertitudes dans ces données d'entrées. Cela a été vérifié par les diagrammes de rang en forme de U avec de très hautes barres aux extrémités. Malgré ce défaut dans l'ensemble météorologique utilisé, le dépôt et débit dose à la sortie semblent assez bien englober les mesures. Le dépôt est une sortie qui n'est pas très difficile à obtenir car il ne présente pas de dimension temporelle, alors que l'évolution temporelle du rejet est très incertaine dans le cas de Fukushima. En revanche, l'activité volumique est une sortie qui se caractérise par une série temporelle. Cela signifie qu'une bonne évaluation de l'activité volumique nécessite une bonne précision en terme de valeur et de dynamique. Ici, on peut conclure qu'il est essentiel de mettre en place la calibration statistique afin de réduire les incertitudes existantes dans les paramètres précédents. Une perspective de ce travail sera d'affiner les conditions d'optimisation : construire des métamodèles seulement sur certains paramètres pour une calibration avec moins d'erreurs.

Construction du métamodèle

L'objectif d'un métamodèle est de créer une imitation du modèle physique en gardant une bonne précision de la sortie et en réduisant le temps de calcul de manière non-négligeable. Les étapes principales pour obtenir un métamodèle sont :

- Créer un plan d'expérience,
- Déterminer et calculer les valeurs d'intérêt,
- Choisir la méthode d'interpolation,
- Évaluer le métamodèle obtenu.

Établissement du plan d'expérience

Notre cas d'étude s'effectue dans un espace d'entrée de dimension assez importante avec certaines variables corrélées entre elles. Le plan d'expérience dans le chapitre 3 se construit de manière séparée pour des variables indépendantes, via l'algorithme de LHS, et pour celles avec une corrélation. Cela a créé un ensemble d'entrée hybride, dont la répartition dans l'espace peut être améliorée. Il serait intéressant de mettre en place un algorithme qui crée un équivalent du LHS en tenant compte des corrélations entre les variables.

Évaluation du métamodèle et comparaison entre deux méthodes de métamodélisation

Avoir deux méthodes d'interpolation nous laisse le choix pour adapter au mieux nos critères. En général, les métamodèles émulent assez bien le modèle physique, sauf RBF pour le débit dose qui a une erreur non négligeable. Les métamodèles avec RBF représentent un peu moins bien les sorties du modèle physique que le krigeage, mais le coût de construction est nettement moins élevé. Le krigeage fournit toujours des métamodèles de meilleure qualité accompagnés de prédictions d'erreur. Par ailleurs, la comparaison entre deux manières de perturber des données météorologiques montre que, malgré la difficulté de convergence, les nouvelles données météorologiques perturbées par la combinaison convexe restent plus proches des données initiales, et possiblement plus réalistes.

Calibration des mesures radioactives

La calibration statistique des mesures radiologiques s'effectue principalement sur le dépôt final et l'activité volumique. Ces deux mesures portent des propriétés spécifiques et complémentaires. L'accident de Fukushima est caractérisé par plusieurs phases de rejets qui créent une vraie difficulté dans l'analyse. Le dépôt final ne fournit pas d'information temporelle, alors que plusieurs scénarios peuvent donner le même dépôt final. Cela explique pourquoi les distributions dans la calibration du dépôt sont aussi plates. En revanche, l'activité volumique est particulièrement caractérisée par une série temporelle. Malgré le découpage en temps (optimisation sur des tranches de deux jours), l'algorithme de MCMC n'arrive pas à converger. Il existe plusieurs possibilités pour expliquer ce phénomène : erreur du métamodèle trop forte, algorithme de MCMC pas assez robuste, etc. Dans la partie d'évaluation des métamodèles, nous avons vu que les métamodèles reposant sur la combinaison convexe de la météorologie donnent des erreurs plus faibles. Malgré l'absence de convergence dans l'optimisation déterministe de ces métamodèles, il est possible d'obtenir des résultats plus performants dans la calibration statistique.

Travail réalisé

Malgré les difficultés rencontrées avec la convergence des chaînes de Markov dans l'algorithme de Metropolis-Hastings, les tâches suivantes ont été achevées :

- Étude de la propagation d'incertitude par méthode de Monte Carlo ;
- Mise en place de plusieurs méthodes et scores afin d'évaluer les simulations : RMSE, FMS, FMT, diagramme de rang, etc. ;

- Construction et test de plusieurs métamodèles en utilisant deux méthodes d'interpolation : par RBF et avec le krigeage ;
- Calibration réussie d'un modèle météorologique simplifié ;
- Mise en place et vérification du fonctionnement d'un diagnostic de convergence des chaînes de Markov, [Brooks 1998] ;
- Calibration du modèle ℓdX dans un cas synthétique pour évaluer l'efficacité de l'algorithme de Metropolis-Hastings ;
- Calibration en utilisant les mesures réelles radiologiques de l'accident de Fukushima sur l'activité volumique et le dépôt final.

Le dernier chapitre 5 nous a montré que la calibration statistique a réussi à réduire l'incertitude des entrées du modèle ℓdX . Les simulations *a posteriori*, générées à partir des échantillons choisis par l'algorithme Metropolis-Hastings, donnent des résultats plus proches des observations. Cette étude de calibration était réalisable grâce aux métamodèles, qui permettent de réduire le temps de calcul pendant l'échantillonnage. Cette thèse a mis en place également une première version de métamodélisation en utilisant un ensemble de prévisions météorologiques et de termes sources. Celui-ci a nécessité l'interpolation de résultats de simulation dans un espace de grande dimension, ce qui a été le plus difficile dans la métamodélisation. Il était cependant essentiel de construire notre approche dans ce cadre car nos modèles de dispersion reposent, en cas de crise, sur le type de données d'entrée exploité ici.

Améliorations possibles

Dans une suite possible de ces travaux, on peut suggérer les pistes suivantes :

- Fixer les paramètres non essentiels, et peut-être éliminer certains termes sources peu probables, pour améliorer la calibration du dépôt final ;
- Construire le métamodèle et mettre en place la calibration de la deuxième période de l'accident sur l'activité volumique ;
- Construire le métamodèle et mettre en place la calibration du débit dose ;
- Mettre en place une calibration croisée en combinant plusieurs sorties de ℓdX .

Annexes

Annexe du chapitre d'introduction

A.1 Fonction de covariance

Définition 6. Soient $\mathbf{x} = (x_1, x_2, \dots, x_n)$ et $\mathbf{y} = (y_1, y_2, \dots, y_n)$ deux échantillons de moyennes respectives $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ et $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$. On appelle la covariance empirique de $\mathbf{x}$ et $\mathbf{y}$ et on note $\text{cov}(\mathbf{x}, \mathbf{y})$ le nombre

$$\text{cov}(\mathbf{x}, \mathbf{y}) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}).$$

On appelle la variance empirique de $\mathbf{x}$ et on note $\text{var}(\mathbf{x})$ le nombre

$$\text{var}(\mathbf{x}) = \text{cov}(\mathbf{x}, \mathbf{x}) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

et $\sqrt{\text{var}(\mathbf{x})}$ s'appelle l'écart type de l'échantillon $\mathbf{x}$. On appelle corrélation de $\mathbf{x}$ avec $\mathbf{y}$ et on note $\text{cor}(\mathbf{x}, \mathbf{y})$ le nombre

$$\rho = \text{cor}(\mathbf{x}, \mathbf{y}) = \frac{\text{cov}(\mathbf{x}, \mathbf{y})}{\sqrt{\text{var}(\mathbf{x})}\sqrt{\text{var}(\mathbf{y})}}$$

Proposition 1. Soient $\mathbf{x}, \mathbf{x}', \mathbf{y}$ et $\mathbf{y}'$ des échantillons de même taille n . Notons $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ et $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ la moyenne de $\mathbf{x}$ et $\mathbf{y}$. Soient λ et λ' des réels dans $\mathbb{R}$. Les relations suivantes sont satisfaites :

1. Bilinéarité à gauche : $\text{cov}(\lambda\mathbf{x} + \lambda'\mathbf{x}', \mathbf{y}) = \lambda\text{cov}(\mathbf{x}, \mathbf{y}) + \lambda'\text{cov}(\mathbf{x}', \mathbf{y})$
2. Bilinéarité à droite : $\text{cov}(\mathbf{x}, \lambda\mathbf{y} + \lambda'\mathbf{y}') = \lambda\text{cov}(\mathbf{x}, \mathbf{y}) + \lambda'\text{cov}(\mathbf{x}, \mathbf{y}')$
3. Symétrie : $\text{cov}(\mathbf{x}, \mathbf{y}) = \text{cov}(\mathbf{y}, \mathbf{x})$
4. Homogénéité de degré deux : $\text{var}(\lambda\mathbf{x}) = \lambda^2\text{var}(\mathbf{x}) \geq 0$
5. Formules de Huggens :

$$\text{cov}(\mathbf{x}, \mathbf{y}) = \frac{1}{n} \sum_{i=1}^n x_i y_i - \bar{x}\bar{y}, \text{ et } \text{var}(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2$$

6. Inégalité de Cauchy-Schwarz : $(\text{cov}(\mathbf{x}, \mathbf{y}))^2 \leq \text{var}(\mathbf{x})\text{var}(\mathbf{y})$ et donc $\rho^2 \leq 1$.

Démonstration. La démonstration des propriétés de la fonction variance - covariance 1 se déroule comme suit :

1. Bilinéarité à gauche : La moyenne de la combinaison $\lambda \mathbf{x} + \lambda' \mathbf{x}'$ s'écrit

$$\begin{aligned}
 \overline{\lambda \mathbf{x} + \lambda' \mathbf{x}'} &= \frac{1}{n} \sum_{i=1}^n \lambda x_i + \lambda' x'_i \\
 &= \frac{1}{n} \sum_{i=1}^n \lambda x_i + \frac{1}{n} \sum_{i=1}^n \lambda' x'_i \\
 &= \frac{\lambda}{n} \sum_{i=1}^n x_i + \frac{\lambda'}{n} \sum_{i=1}^n x'_i \\
 &= \lambda \bar{\mathbf{x}} + \lambda' \bar{\mathbf{x}}'.
 \end{aligned} \tag{A.1}$$

Nous pouvons ensuite remplacer l'expression A.1 dans le calcul de covariance

$$\begin{aligned}
 cov(\lambda \mathbf{x} + \lambda' \mathbf{x}', \mathbf{y}) &= \frac{1}{n} \sum_{i=1}^n (\lambda x_i + \lambda' x'_i - \overline{\lambda \mathbf{x} + \lambda' \mathbf{x}'}) (y_i - \bar{\mathbf{y}}) \\
 &= \frac{1}{n} \sum_{i=1}^n (\lambda x_i - \lambda \bar{\mathbf{x}} + \lambda' x'_i - \lambda' \bar{\mathbf{x}}') (y_i - \bar{\mathbf{y}}) \\
 &= \frac{\lambda}{n} \sum_{i=1}^n (x_i - \bar{\mathbf{x}}) (y_i - \bar{\mathbf{y}}) + \frac{\lambda'}{n} \sum_{i=1}^n (x'_i - \bar{\mathbf{x}}') (y_i - \bar{\mathbf{y}}) \\
 &= \lambda cov(\mathbf{x}, \mathbf{y}) + \lambda' cov(\mathbf{x}', \mathbf{y}).
 \end{aligned}$$

2. La démonstration de la bilinéarité à droite se déroule de la même façon.
3. La symétrie est une déduction directe de la formule de la définition.
4. Homogénéité de degré deux :

$$\begin{aligned}
 var(\lambda \mathbf{x}) &= \frac{1}{n} \sum_{i=1}^n (\lambda x_i - \overline{\lambda \mathbf{x}})^2 \\
 &= \frac{1}{n} \sum_{i=1}^n (\lambda x_i - \lambda \bar{\mathbf{x}})^2 \\
 &= \frac{\lambda^2}{n} \sum_{i=1}^n (x_i - \bar{\mathbf{x}})^2 \\
 &= \lambda^2 var(\mathbf{x})
 \end{aligned}$$

5. Formules de Huygens :

$$\begin{aligned}
 cov(\mathbf{x}, \mathbf{y}) &= \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \\
 &= \frac{1}{n} \sum_{i=1}^n (x_i - x_i \bar{y} - y_i \bar{x} + \bar{x} \bar{y}) \\
 &= \frac{1}{n} \sum_{i=1}^n x_i y_i - \frac{\bar{y}}{n} \sum_{i=1}^n x_i - \frac{\bar{x}}{n} \sum_{i=1}^n y_i + \frac{n}{n} \bar{x} \bar{y} \\
 &\quad (\text{les moyennes } \bar{x} \text{ et } \bar{y} \text{ ne dépendent plus d'indices}) \\
 &= \frac{1}{n} \sum_{i=1}^n x_i y_i - \bar{x} \bar{y} - \bar{x} \bar{y} + \bar{x} \bar{y} \\
 &= \frac{1}{n} \sum_{i=1}^n x_i y_i - \bar{x} \bar{y}
 \end{aligned}$$

6. Inégalité de Cauchy-Schwarz : Considérons la variance de $\lambda \mathbf{x} + \mathbf{y}$ pour tout λ réel

$$\begin{aligned}
 var(\lambda \mathbf{x} + \mathbf{y}) &= \frac{1}{n} \sum_{i=1}^n (\lambda x_i + y_i - \overline{\lambda \mathbf{x} + \mathbf{y}})^2 \\
 &= \frac{1}{n} \sum_{i=1}^n (\lambda x_i + y_i - \lambda \bar{x} - \bar{y})^2 \\
 &= \frac{1}{n} \sum_{i=1}^n (\lambda^2 (x_i - \bar{x})^2 + 2\lambda (x_i - \bar{x})(y_i - \bar{y}) + (y_i - \bar{y})^2) \\
 &= \frac{\lambda^2}{n} \sum_{i=1}^n (x_i - \bar{x})^2 + \frac{2\lambda}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) + \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2 \\
 &= \lambda^2 var(\mathbf{x}) + 2\lambda cov(\mathbf{x}, \mathbf{y}) + var(\mathbf{y}) \tag{A.2}
 \end{aligned}$$

De plus, grâce à la définition, la fonction de variance est toujours positive ou nulle pour tout λ réel. Si on considère l'expression A.2 comme un polynôme de degré deux de λ alors le discriminant A.3 d'un polynôme positif ou nul doit être négatif ou nul

$$\begin{aligned}
 \Delta &= 4(cov(\mathbf{x}, \mathbf{y}))^2 - 4var(\mathbf{x})var(\mathbf{y}) \leq 0 \\
 &\Leftrightarrow (cov(\mathbf{x}, \mathbf{y}))^2 \leq var(\mathbf{x})var(\mathbf{y})
 \end{aligned} \tag{A.3}$$

d'où le résultat. Remarquons que le discriminant est nul si et seulement si $cov(\lambda \mathbf{x} + \mathbf{y})$ peut s'annuler, c'est-à-dire s'il existe un réel λ tel que $\lambda \mathbf{x} + \mathbf{y}$ soit constant. Donc l'inégalité de Cauchy-Schwarz ne peut être une égalité que s'il existe une relation affine entre $\mathbf{x}$ et $\mathbf{y}$. Une fois que l'inégalité de Cauchy-Schwarz est prouvée, il est naturel de diviser la covariance par le produit des écarts types pour obtenir la dernière expression $\rho^2 \leq 1$. $\square$

A.2 Déterminant d'une matrice carré

On note $\mathcal{M}_n(\mathbb{R})$ l'ensemble des matrices carrées de taille $n \times n$ dont les composantes sont réelles. Si A est une matrice de l'ensemble $\mathcal{M}_n(\mathbb{R})$, et si i et j sont des entiers, tels que $1 \leq i \leq n$

et $1 \leq j \leq n$, on notera A_{ij} la matrice de $\mathcal{M}_{n-1}(\mathbb{R})$ obtenue en enlevant la i ème ligne et la j ème colonne de A .

Soit A dans $\mathcal{M}_n(\mathbb{R})$. On définit son déterminant $\det(A) \in \mathbb{R}$ par récurrence sur n . Pour $n = 1$, si $A = (a)$, on pose simplement $\det(A) = a$. Supposons maintenant $n > 1$, et supposons que l'on a déjà défini le déterminant des matrices de $\mathcal{M}_{n-1}(\mathbb{R})$. Alors on pose la définition suivante.

Définition 7. Soit $A \in \mathcal{M}_n(\mathbb{R})$, de coefficients a_{ij} . On définit :

$$\begin{aligned} \det(A) &= a_{11}\det(A_{11}) - a_{12}\det(A_{12}) + \dots + (-1)^{1+n}a_{1n}\det(A_{1n}) \\ &= \sum_{j=1}^n (-1)^{1+j}a_{1j}\det(A_{1j}) \end{aligned}$$

On note aussi le déterminant de A par le tableau de ses coefficients entre deux barres :

$$\det(A) = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$$

Pour $n = 2$, on a $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$. Pour $n = 3$, on a $\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = aei - afh - bdi + bfg + cdh - ceg$.

À partir de la définition, nous pouvons en déduire que :

- Le déterminant de la matrice nulle $\mathbf{0}_n$ vaut 0 ;
- Le déterminant de la matrice identité $\mathbf{I}_n = \begin{pmatrix} 1 & \dots & 1 \\ \vdots & 1 & \vdots \\ 1 \dots & 1 \end{pmatrix}$ vaut 1.

Donnons maintenant quelques propriétés importantes du déterminant : comment se comporte le déterminant face aux opérations élémentaires sur les colonnes ?

Proposition 2. Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice ayant des colonnes $c_1, \dots, c_n$. On note A' la matrice obtenue par une des opérations élémentaires sur les colonnes, qui sont :

1. $c_i \rightarrow \lambda c_i$ avec $\lambda \neq 0$: A' est obtenue en multipliant une colonne de A par un scalaire non nul. Alors $\det(A') = \lambda \det(A)$.
2. $c_i \rightarrow c_i + \lambda c_j$ avec $\lambda \in \mathbb{R}$ et $i \neq j$: A' est obtenue en ajoutant à une colonne de A un multiple d'une autre colonne de A . Alors $\det(A') = \det(A)$.
3. $c_i \leftrightarrow c_j$: A' est obtenue en échangeant deux colonnes distinctes de A . Alors $\det(A') = -\det(A)$.

Plus généralement pour 2 : l'opération $c_i \rightarrow c_i + \sum_{j=1, j \neq i}^n \lambda_j c_j$ d'ajouter une combinaison linéaire des autres colonnes conserve le déterminant.

Démonstration. 1. La première propriété découle de la linéarité du déterminant par rapport à chaque colonne de la matrice A , les autres étant fixées.

2. Soit $A = (c_1 \dots c_i \dots c_j \dots c_n)$ une matrice représentée par ses vecteurs colonnes c_k . L'opération $c_i \rightarrow c_i + \lambda c_j$ transforme la matrice A en la matrice $A' = (c_1 \dots c_i + \lambda c_j \dots c_j \dots c_n)$. Par linéarité par rapport à la colonne i , on sait que :

$$\det(A') = \det(A) + \lambda \det(c_1 \dots c_j \dots c_j \dots c_n)$$

Or les colonnes de la matrice $(c_1 \dots c_j \dots c_j \dots c_n)$ sont identiques, donc son déterminant est nul.

3. Si on échange les colonnes i et j de $A = (c_1 \dots c_i \dots c_j \dots c_n)$ on obtient la matrice $A' = (c_1 \dots c_j \dots c_i \dots c_n)$, où le vecteur c_j se trouve en colonne i et le vecteur c_i se trouve en colonne j . Introduisons alors une troisième matrice $B = (c_1 \dots c_i + c_j \dots c_j + c_i \dots c_n)$. Cette matrice a deux colonnes distinctes égales, donc $\det(B) = 0$.

D'autre part, nous pouvons développer ce déterminant en utilisant la multilinéarité, c'est-à-dire linéarité par rapport à chaque colonne. Ceci donne :

$$\begin{aligned} 0 = \det(B) &= \det(c_1 \dots c_i + c_j \dots c_j + c_i \dots c_n) \\ &= \det(c_1 \dots c_i \dots c_j + c_i \dots c_n) \\ &\quad + \det(c_1 \dots c_j \dots c_j + c_i \dots c_n) \\ &= \det(c_1 \dots c_i \dots c_j \dots c_n) \\ &\quad + \det(c_1 \dots c_i \dots c_i \dots c_n) \\ &\quad + \det(c_1 \dots c_j \dots c_j \dots c_n) \\ &\quad + \det(c_1 \dots c_j \dots c_i \dots c_n) \\ &= \det(A) + 0 + 0 + \det(A') \end{aligned}$$

□

Corollaire 1. Si une colonne c_i de la matrice A est combinaison linéaire des autres colonnes, alors $\det(A) = 0$.

Dans la formule de la vraisemblance, la matrice de covariance doit être inversible. Le théorème suivant met en relation l'inversibilité d'une matrice et son déterminant.

Théorème 3. Une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est inversible si et seulement si $\det(A) \neq 0$.

Ainsi, une famille de vecteurs $(C_1, \dots, C_n)$ est une base de $\mathbb{R}^n$ si et seulement si $\det(A) \neq 0$, où (C_i) est la i ème colonne de la matrice A .

Démonstration. 1. Soit A non-inversible, donc avec une colonne C_j combinaison linéaire des autres colonnes. En enlevant cette combinaison linéaire de C_j , ce qui ne change pas le déterminant, nous obtenons une colonne nulle, donc $\det(A) = 0$.

2. Soit A inversible, donc avec les colonnes $(C_1, \dots, C_n)$ qui forment une base de $\mathbb{R}^n$. Supposons par absurde que $\det(A) = 0$. Soit $(e_1, \dots, e_n)$ la base canonique de $\mathbb{R}^n$; écrivons $e_j = \sum_i b_{ij} C_i$. Alors :

$$\det(e_1, \dots, e_n) = \det\left(\sum_{i_1} b_{i_1 1} C_{i_1}, \dots, \sum_{i_n} b_{i_n n} C_{i_n}\right) = \sum_{i_1, \dots, i_n} b_{i_1 1} b_{i_n n} \det(C_{i_1}, \dots, C_{i_n}) = 0$$

Ceci est contradictoire avec l'hypothèse de départ.

□

Annexe d'article Monte Carlo

B.1 Interpreting the rank histogram

The shape analysis of a rank histogram allows to conclude on potential biases in the ensemble as well as over- or under-spread. In the rank histogram, the height of the bar i is the number of observations that are higher than exactly i ensemble members. If an ensemble is reliable, the observations are indistinguishable from any other ensemble member, and the rank histogram is flat. If an ensemble does not spread out enough, the external bars will be higher than the internal bars. If an ensemble spreads out too much, the observations' dispersion is smaller than the ensemble's spread, and the rank histogram is shaped like a triangle. In case the ensemble has a bias, e.g., overestimates (resp. underestimates) the variable, the histogram will show higher bars on the left (resp. on the right) than on the other side. In practice, the rank histograms are seldom flat because all uncertainties are not taken into account in the generation of the ensemble.

B.2 Illustration of subgrid effects

B.3 Figure of Merit in Space – FMS

The FMS formula is

$$FMS = \frac{A_m \cap A_p}{A_m \cup A_p} 100$$

where A_m is the observed area above the threshold and A_p is the predicted area above the same threshold.

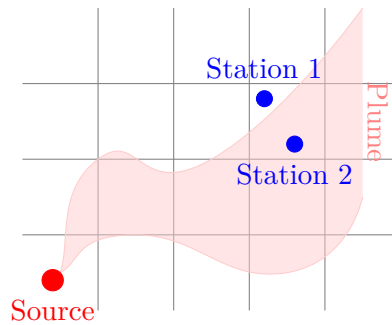


FIGURE B.1 – Subgrid effect description.

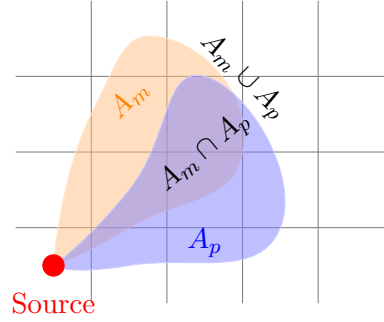


FIGURE B.2 – Illustration of FMS.

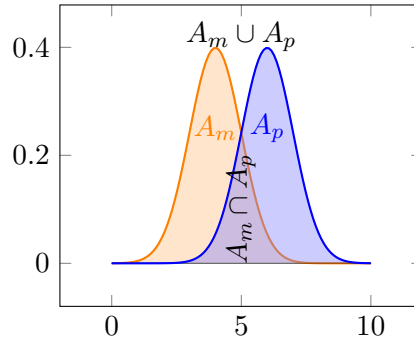


FIGURE B.3 – Illustration of FMT.

B.4 Figure of Merit in Time – FMT

The FMT formula is the same as the FMS formula. The figure B.3 shows how to find the formula components. The intersection of A_m and A_p is related to the minimum between the curves and the union is related to the the maximum.

B.5 Root Mean Square Error – RMSE

Indicator RMSE is the standard deviation of model-to-data error, it aggregates the magnitudes of the difference between simulation and observation in space for various times into a single value. It allows to compare the accuracies of different simulations. RMSE is always positive and the value 0 represents a perfect prediction. In practice, this value is never reached because of various reasons : model error, observational error, etc. Let $Y = (y_1, y_2, \dots, y_n)$ be the observed vector and $\hat{Y} = (\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n)$ the predicted vector of variable y . The RMSE formula is written as

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}}$$

Annexe d'article de métamodélisation

C.1 LHS and complementary LHS

The construction of the complementary LHS is based on the initial DOE. Its points are chosen such that the merger of these two series forms another LHS. To obtain this result, [Damblin 2013] introduces several optimization methods in order to maximize the space filling after the merger.

C.2 Dirichlet distribution

The n -dimensional Dirichlet distribution is a distribution on probability vectors, $\nu = (\nu_1, \dots, \nu_n)$, whose elements are non-negative and sum up to 1. The n -dimensional Dirichlet distribution is parameterized by a n -dimensional vector α of positive reals. In 2-dimension, the Dirichlet distribution is essentially the Beta distribution. In other words, the Dirichlet distribution is the multinomial generalization of the Beta distribution. The main features of a Dirichlet distribution $Dir(\alpha = (\alpha_1, \dots, \alpha_n))$ can be summarized as in table C.1. The figure C.3 illustrates the probability density function (PDF) of some three-dimensional Dirichlet distributions with different α parameters. When $\alpha = (1, \dots, 1)$, all points are equally likely or the PDF is similar everywhere in the triangle. The blue area represents a low probability of appearance, and the red one means a strong probability. The white line on the color bar indicates the single probability density of the uniform PDF $Dir(\alpha = (1, \dots, 1))$. The high values of α lead to a low standard deviation, with a concentration of the density around the so-called mode point.

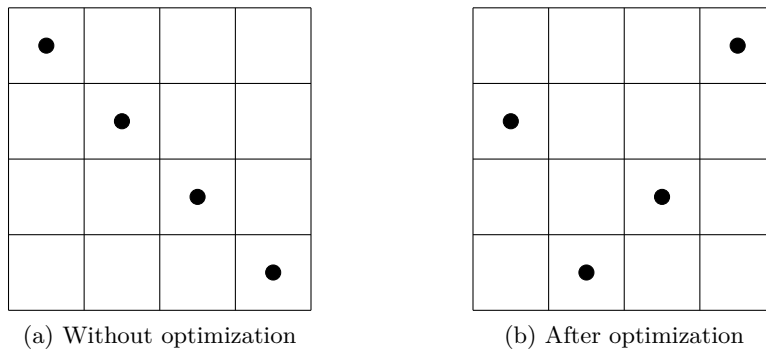


FIGURE C.1 – Illustration of Latin Hyper-cube Sampling

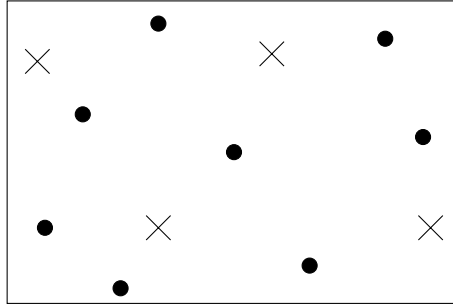


FIGURE C.2 – Illustration of complementary Latin Hyper-cube Sampling. The round points represent the initial LHS and the crosses are the complementary LHS sample.

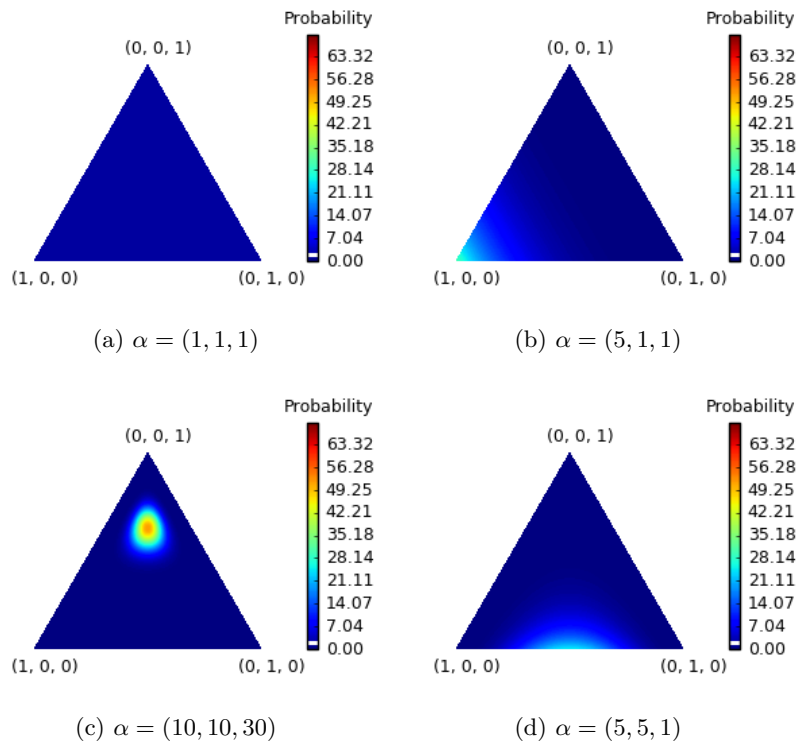


FIGURE C.3 – Illustrations of Dirichlet distributions in dimension three

Support	$v = (v_1, \dots, v_n), \quad (v_i)_i \in [0, 1] \text{ and } \sum_{i=1}^n v_i = 1$
PDF	$f_\alpha(v) = \frac{1}{B(\alpha)} \prod_{i=1}^n v_i^{\alpha_i} - 1$
Mode	$x_i = \frac{\alpha_i - 1}{\sum_{i=1}^n \alpha_i - n} \quad \alpha_i > 1$
Mean	$E[V_i] = \frac{\alpha_i}{\alpha_0}$
Variance	$Var(V_i) = \frac{\alpha_i(\alpha_0 - \alpha_i)}{\alpha_0^2(\alpha_0 + 1)}$
Covariance	$Cov(V_i, V_j) = \frac{-\alpha_i \alpha_j}{\alpha_0(\alpha_0 + 1)}, \quad (i \neq j)$

TABLE C.1 – Description of Dirichlet distribution, where $B(\alpha) = \frac{\prod_{i=1}^n \Gamma(\alpha_i)}{\Gamma(\sum_{i=1}^n \alpha_i)}$ and $\alpha_0 = \sum_{i=1}^n \alpha_i$.

C.3 Reduction of weather ensemble size

Algorithm 2: Reduce weather ensemble size

Data:

- N_{new} : The size of new ensemble ;
- ε : Error rate ;
- $M = (m_1, \dots, m_N)$: the original ensemble, $m_i \in \mathbb{R}^{N_t}$ where N_t is the number of time steps.

Result: $M_{new} = (m_1, \dots, m_{N_{new}})$: the vector contains indices of selected members.

1 Initialization :

- Compute $\sigma_{old} = \frac{1}{N_t} \sum_{j=1}^{N_t} \sqrt{Var(M)} ;$
- Generate $M_{new} = (m_1, \dots, m_{i_{N_{new}}})$ a subensemble of M ;
- Compute $\sigma_{new} = \frac{1}{N_t} \sqrt{Var(M_{new})}$

while $\frac{\sigma_{new}}{\sigma_{old}} \cdot 100 > \varepsilon$ **do**

- Compute $\sigma_{new} = \frac{1}{N_t} \sqrt{Var(M_{new}^j)}$ where $M_{new}^j = (m_{i_1}^j, \dots, m_{i_{N_{new}}}^j)$ and $(i_1, \dots, i_{N_{new}})$ is a random draw without replacement in $\{1, \dots, N\}$.

end

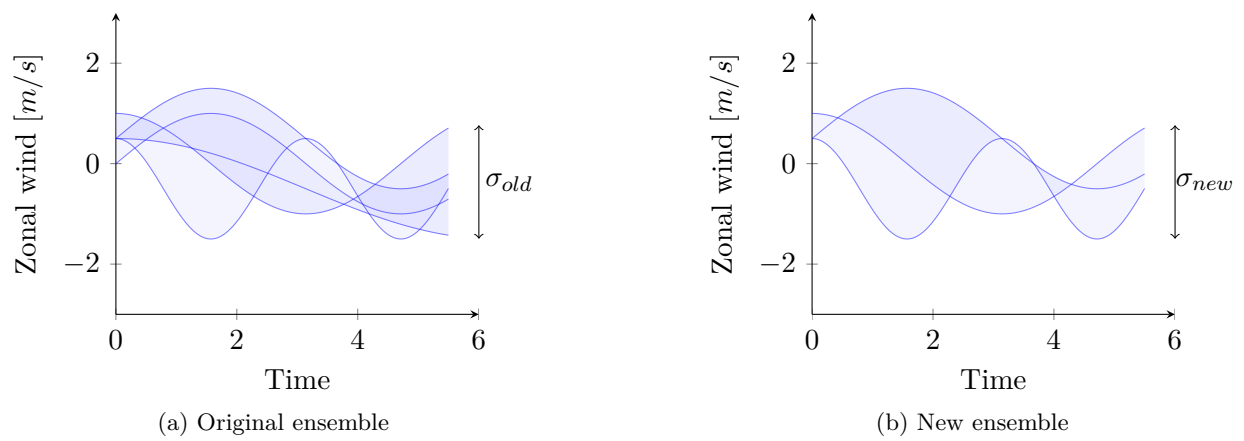


FIGURE C.4 – Illustration of the reduction of ensemble size

Bibliographie

- [Arnold 2015] D. Arnold, C. Maurer, G. Wotawa, R. Draxler, K. Saito et P. Seibert. *Influence of the meteorological input on the atmospheric transport modelling with FLEXPART of radionuclides from the Fukushima Daiichi nuclear accident*. Journal of Environmental Radioactivity, vol. 139, pages pp.212 – 225, 2015. (Cité en page 45.)
- [Baklanov 2001] A. Baklanov et J.H. Sørensen. *Parameterisation of radionuclide deposition in atmospheric long-range transport modelling*. Physics and Chemistry of the Earth, Part B : Hydrology, Oceans and Atmosphere, vol. 26, no. 10, pages 787 – 799, 2001. (Cité en pages 39 et 73.)
- [Bowler 2009] N. E. Bowler, A. Arribas, S. E. Beare, K. R. Mylne et G. J. Shutts. *The local ETKF and SKEB : Upgrades to the MOGREPS short-range ensemble prediction system*. Quarterly Journal of the Royal Meteorological Society, vol. Soc. 135, page pp.767–776, 2009. (Cité en pages 52, 53 et 98.)
- [Brandt 1998] Jørgen Brandt, Annemarie Bastrup-birk, Jesper H. Christensen, Torben Mikkelsen, Søren Thykier-Nielsen et Zahari Zlatev. *Testing the importance of accurate meteorological input fields and parameterizations in atmospheric transport modelling using dream - validation against ETEX-1*. Atmospheric Environment, vol. 32, no. 24, pages 4167 – 4186, 1998. (Cité en pages 41 et 73.)
- [Brandt 2002] J. Brandt, J. H. Christensen et L. M. Frohn. *Modelling transport and deposition of caesium and iodine from the Chernobyl accident using the DREAM model*. Atmospheric Chemistry and Physics, vol. 2, no. 5, pages 397–417, 2002. (Cité en pages 39 et 73.)
- [Brooks 1998] S. P. Brooks et A. Gelman. *General Methods for Monitoring Convergence of Iterative Simulations*. Journal of Computational and Graphical Statistics, vol. 7, no. 4, pages 434–455, 1998. (Cité en pages 16, 92, 93, 94, 98, 100, 106, 107, 110, 111 et 130.)
- [Buizza 2008] Roberto Buizza, Martin Leutbecher et Lars Isaksen. *Potential use of an ensemble of analyses in the ECMWF Ensemble Prediction System*. Quarterly Journal of the Royal Meteorological Society, vol. 134, no. 637, pages 2051–2066, 2008. (Cité en page 47.)
- [Damblin 2013] G. Damblin, M. Couplet et B. Iooss. *Numerical studies of space-filling designs : optimization of Latin Hypercube Samples and subprojection properties*. Journal of Simulation, vol. 7, no. 4, pages 276–289, 2013. (Cité en pages 19 et 141.)
- [Draxler 2015] R. Draxler, D. Arnold, M. Chino, S. Galmarini, M. Hort, A. Jones, S. Leadbetter, A. Malo, C. Maurer, G. Rolph, K. Saito, R. Servranckx, T. Shimbori, E. Solazzo et G. Wotawa. *World Meteorological Organization’s model simulations of the radionuclide dispersion and deposition from the Fukushima Daiichi nuclear power plant accident*. Journal of Environmental Radioactivity, vol. 139, pages pp.172 – 184, 2015. (Cité en pages 45 et 74.)
- [Dubois 2014] Catherine Dubois. *Adapter les quartiers et les bâtiments au réchauffement climatique ; une feuille de route pour les architectes et les designers urbains québécois*. PhD thesis, 12 2014. (Cité en page 27.)

- [Duhanyan 2011] N. Duhanyan et Y. Roustan. *Below-cloud scavenging by rain of atmospheric gases and particulates*. Atmospheric Environment, vol. 45, no. 39, pages 7201 – 7217, 2011. (Cité en page 40.)
- [Dupuis 2007] J. Dupuis. *Statistique bayésienne et algorithmes MCMC*, 09 2007. (Cité en page 14.)
- [Gelman 1992] Andrew Gelman et Donald B. Rubin. *Inference from Iterative Simulation Using Multiple Sequences*. Statist. Sci., vol. 7, no. 4, pages 457–472, 11 1992. (Cité en page 16.)
- [Geweke 1991] John F. Geweke. *Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments*. Staff Report 148, Federal Reserve Bank of Minneapolis, 1991. (Cité en page 16.)
- [Girard 2014] S. Girard, I. Korsakissok et V. Mallet. *Screening sensitivity analysis of a radio-nuclides atmospheric dispersion model applied to the Fukushima disaster*. Atmospheric environment, Elsevier, vol. 95, pages pp.490–500, 2014. (Cité en pages 45, 57 et 68.)
- [Girard 2016] S. Girard, V. Mallet, I. Korsakissok et A. Mathieu. *Emulation and Sobol’ sensitivity analysis of an atmospheric dispersion model applied to the Fukushima nuclear accident*. Journal of Geophysical Research : Atmospheres, American Geophysical Union, vol. Journal of Geophysical Research : Atmospheres, American Geophysical Union, pages pp.3484 – 3496, 2016. (Cité en pages 39, 41, 45 et 73.)
- [Haiden 2019] Thomas Haiden, Martin Janousek, Frédéric Vitart, Laura Ferranti et Fernando Prates. *Evaluation of ECMWF forecasts, including the 2019 upgrade*. no. 853, 11 2019. (Cité en page 51.)
- [Hanna 2001] Steven R Hanna, Zhigang Lu, H Christopher Frey, Neil Wheeler, Jeffrey Vukovich, Saravanan Arunachalam, Mark Fernau et D Alan Hansen. *Uncertainties in predicted ozone concentrations due to input uncertainties for the UAM-V photochemical grid model applied to the July 1995 OTAG domain*. Atmospheric Environment, vol. 35, no. 5, pages 891 – 903, 2001. (Cité en pages 41 et 73.)
- [Hastings 1970] W. K. Hastings. *Monte Carlo Sampling Methods Using Markov Chains and Their Applications*. Biometrika, vol. 57, no. 1, pages 97–109, 1970. (Cité en page 92.)
- [Hudson 2017] Debbie Hudson et Beth Ebert. *Ensemble verification metrics*. In ECMWF Annual Seminar, 2017. (Cité en page 51.)
- [IAEA 2012] IAEA. *Fukushima Monitoring Data Base*, 2012. (Cité en page 59.)
- [INERIS 2002] INERIS. *Méthodes pour l’évaluation et la prévention des risques accidentels (DRA-006)*, 02 2002. (Cité en page 26.)
- [Iooss 2010] B. Iooss, L. Boussouf, V. Feuillard et A. Marrel. *Numerical studies of the metamodel fitting and validation processes*. International Journal of Advances in Systems and Measurements, vol. 3, pages pp.11–21, 01 2010. (Cité en page 77.)
- [irsn.fr 2011] irsn.fr. *L’impact du séisme sur les installations nucléaires de Fukushima-Daiichi*, 2011. (Cité en page 32.)
- [Kajino 2019] M. Kajino, T.T Sekiyama, Y. Igarashi, G. Katata, M. Sawada, K. Adachi, Y. Zai-zen, H. Tsuruta et T. Nakajima. *Deposition and Dispersion of Radio-Cesium Released*

- Due to the Fukushima Nuclear Accident : Sensitivity to Meteorological Models and Physical Modules.* Journal of Geophysical Research : Atmospheres, vol. 124, no. 3, pages pp.1823–1845, 2019. (Cité en page 45.)
- [Katata 2015] G. Katata, M. Chino, T. Kobayashi, H. Terada, M. Ota, H. Nagai, M. Kajino, R. Draxler, M. C. Hort, A. Malo, T. Torii et Y. Sanada. *Detailed source term estimation of the atmospheric release for the Fukushima Daiichi Nuclear Power Station accident by coupling simulations of an atmospheric dispersion model with an improved deposition scheme and oceanic dispersion model.* Atmos. Chem. Phys., vol. 15, pages pp.1029–1070, 2015. (Cité en pages 38, 58, 73, 85, 116, 119 et 124.)
- [Korsakissok 2013] I. Korsakissok, A. Mathieu et D. Didier. *Atmospheric dispersion and ground deposition induced by the Fukushima Nuclear Power Plant accident : A local-scale simulation and sensitivity study.* Atmospheric Environment, vol. 70, pages pp.267–279, 2013. (Cité en pages 44, 61, 63 et 65.)
- [Korsakissok 2019] I. Korsakissok, S. Andronopoulos, P. Astrup, P. Bedwell, K. Chevalier-Jabet, H. De Vries, G. Geertsema, F. Gering, T. Hamburger, H. Klein, S. Leadbetter, A. Mathieu, T. Pazmandi, R. Périllat, C. Rudas, A. Sogachev, P. Szanto, J. Tomas, C. Twenhöfel et J. Wellings. *Comparison of ensembles of atmospheric dispersion simulations : lessons learnt from the confidence project about uncertainty quantification.* In 19th international conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, Bruges, Belgium, 2019. (Cité en page 45.)
- [Le 2020] N.B.T Le, I. Korsakissok, A. Mathieu, V. Mallet, R. Périllat et D. Didier. *Uncertainty study on atmospheric dispersion model by using Monte Carlo perturbation, and applied to the Fukushima nuclear accident : Part II - long-range model.* Atmospheric Environment, 2020. Submitted. (Cité en pages 72, 73, 75 et 78.)
- [Leadbetter 2015] S.J. Leadbetter, M.C. Hort, A.R. Jones, H.N. Webster et R.R. Draxler. *Sensitivity of the modelled deposition of Caesium-137 from the Fukushima Dai-ichi nuclear power plant to the wet deposition parameterisation in NAME.* Journal of Environmental Radioactivity, vol. 139, pages pp.200 – 211, 2015. (Cité en page 45.)
- [Leutbecher 2007] M. Leutbecher et T.N. Palmer. *Ensemble forecasting.* no. 514, page 31, 02 2007. (Cité en page 37.)
- [Leutbecher 2008] Martin Leutbecher et Tim N. Palmer. *Ensemble forecasting.* J. Comput. Physics, vol. 227, pages 3515–3539, 2008. (Cité en page 36.)
- [Leutbecher 2014] M. Leutbecher et S. T. K. Lang. *On the reliability of ensemble variance in subspaces defined by singular vectors.* Quarterly Journal of the Royal Meteorological Society, vol. 140, no. 682, pages 1453–1466, 2014. (Cité en page 47.)
- [Louis 1979] J.F. Louis. *A parametric model of vertical eddy fluxes in the atmosphere.* Boundary-Layer Meteorology, vol. 17, pages pp. 187–202, 1979. (Cité en pages 41 et 73.)
- [Mallet 2007] V. Mallet, D. Quélo, B. Sportisse, M. Ahmed de Biasi, É. Debry, I. Korsakissok, L. Wu, Y. Roustan, K. Sartelet, M. Tombette et H. Foudhi. *Technical Note : The air quality modeling system Polyphemus.* Atmospheric Chemistry and Physics, vol. 7(20), pages pp.5479–5487, 2007. (Cité en page 43.)

- [Marrel 2009] A. Marrel, B. Iooss, B. Laurent et O. Roustant. *Calculations of Sobol indices for the Gaussian process metamodel*. Reliability Engineering and System Safety, vol. 94, pages pp.742–751, 2009. (Cité en page 76.)
- [Maryon 1991] R.H. Maryon, F.B. Smith, B.J. Conway et D.M. Goddard. *The U.K. nuclear accident model*. Progress in Nuclear Energy, vol. 26, no. 2, pages 85 – 104, 1991. (Cité en pages 39 et 73.)
- [Mathieu 2012] A. Mathieu, I. Korsakissok, D. Quélo, J. Groëll, M. Tombette, D. Didier, E. Quentric, O. Saunier, J.-P. Benoit et O. Isnard. *Atmospheric dispersion and deposition of ra-dionuclides from the Fukushima Daiichi nuclear power plant accident*. Elements, vol. 8, pages pp.195–200, 2012. (Cité en pages 34, 38, 44, 58, 73, 85, 88, 108, 112, 116 et 119.)
- [Mathieu 2018a] A. Mathieu, M. Kajino, I. Korsakissok, R. Périllat, D. Quélo, A. Quérel, O. Saunier, T. T. Sekiyama, Y. Igarashi et D. Didier. *Fukushima Daiichi-derived radionuclides in the atmosphere, transport and deposition in Japan : A review*. Applied Geochemistry, vol. 91, pages pp.122–139, 2018. (Cité en pages 34, 45, 49, 57, 58, 61, 65, 105 et 124.)
- [Mathieu 2018b] A. Mathieu, I. Korsakissok, S. Leadbetter, S. Andronopoulos, P. Bedwell, J. Wellings, G. Geertsema, A. Jones, J. Tomas, H. de Vries, H. Klein, R. Périllat, K. Chevalier-Jabet, S. Stéphani, S. Fougerolle, V. Créach, E. Cogez, T. Hamburger, T. Pazmandi, C. Rudas, Sogachev A. et P. Szanto. *D9.1 Guidelines ranking uncertainties for atmospheric dispersion*, 02 2018. (Cité en pages 35 et 45.)
- [McKay 1979] M. D. McKay, R. J. Beckman et W. J. Conover. *A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code*. Technometrics, vol. 21, no. 2, pages 239–245, 1979. (Cité en page 18.)
- [MEXT 2011a] MEXT. *Results of Airborne Monitoring Survey by MEXT and Ibaraki Prefecture*. Rapport technique, Ministry of Education, Culture, Sports, Science and Technology of Japan, 2011. (Cité en page 35.)
- [MEXT 2011b] MEXT. *Results of the Airborne Monitoring Survey by MEXT and Gunma Prefecture*. Rapport technique, Ministry of Education, Culture, Sports, Science and Technology of Japan, 2011. (Cité en page 35.)
- [MEXT 2012] MEXT. *Results of the fifth airborne monitoring survey and airborne monitoring survey outside 80 km from the Fukushima Dai-ichi NPP*. Rapport technique, Ministry of Education, Culture, Sports, Science and Technology of Japan, 2012. (Cité en page 35.)
- [Ollinaho 2017] Pirkka Ollinaho, Sarah-Jane Lock, Martin Leutbecher, Peter Bechtold, Anton Beljaars, Alessio Bozzo, Richard M. Forbes, Thomas Haiden, Robin J. Hogan et Irina Sandu. *Towards process-level representation of model uncertainties : stochastically perturbed parametrizations in the ECMWF ensemble*. Quarterly Journal of the Royal Meteorological Society, vol. 143, no. 702, pages 408–422, 2017. (Cité en page 47.)
- [Oura 2015] Y. Oura, M. Ebihara, H. Tsuruta, T. Nakajima, T. Ohara, M. Ishimoto, H. Sawahata et W. Nitta. *A database of hourly atmospheric concentrations of radiocesium (^{134}Cs and ^{137}Cs) in suspended particulate matter collected in March 2011 at 99 air*

- pollution monitoring stations in Eastern Japan*. J. Nucl. Radiochem. Sci., vol. 15, pages pp.1–12, 2015. (Cité en page 58.)
- [Palmer 2009] T.N. Palmer, Roberto Buizza, F. Doblas-Reyes, T. Jung, Martin Leutbecher, G.J. Shutts, M. Steinheimer et Antje Weisheimer. *Stochastic parametrization and model uncertainty*. no. 598, page 42, 10 2009. (Cité en page 47.)
- [Prieur 2014] Jean-Louis Prieur. *Contribution to Signal, Image & Instrumentation in Astronomy*. Habilitation à diriger des recherches, Université Paul Sabatier (Toulouse), Septembre 2014. (Cité en page 23.)
- [Périllat 2016] R. Périllat, I. Korsakissok, V. Mallet, A. Mathieu, T. Sekiyama, M. Kajino, K. Adachi, Y. Igarashi, T. Maki et D. Didier. *Using meteorological ensembles for atmospheric dispersion modelling of the Fukushima nuclear accident*. In 17th international conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, Budapest, Hungary, 2016. (Cité en pages 46 et 48.)
- [Quérel 2015] A. Quérel, Y. Roustan, D. Quélo et J.P. Benoit. *Hints to discriminate the choice of wet deposition models applied to an accidental radioactive release*. International Journal of Environment and Pollution, vol. 58, no. 4, pages 268–279, 2015. (Cité en pages 39, 40 et 45.)
- [Quérel 2016] A. Quérel, D. Quélo, Y. Roustan, A. Mathieu, M. Kajino, T. Sekiyama, K. Adachi, D. Didier, Y. Igarashi et T. Maki. *Fukushima : Lessons Learned on Wet Deposition from a Combined Analysis of Radiation Dose Rate and Volume Activity Measurements of Cesium-137*. In Goldschmidt Conferences, Yokohama, Japon, 2016. (Cité en pages 39, 59 et 60.)
- [Raftery 1992] Adrian E. Raftery et Steven M. Lewis. *[Practical Markov Chain Monte Carlo] : Comment : One Long Run with Diagnostics : Implementation Strategies for Markov Chain Monte Carlo*. Statistical Science, vol. 7, no. 4, pages 493–497, 1992. (Cité en page 16.)
- [Raynaud 2017] Laure Raynaud et François Bouttier. *The impact of horizontal resolution and ensemble size for convective-scale probabilistic forecasts*. Quarterly Journal of the Royal Meteorological Society, vol. 143, no. 709, pages 3037–3047, 2017. (Cité en page 48.)
- [Rodas 1999] Rodas, Thykier-Nielsen S., Deme S. et Mikkelsen T. *Description of the atmospheric dispersion module RIMPUFF*. 01 1999. (Cité en pages 39 et 73.)
- [Ryall 1998] D.B. Ryall et R.H. Maryon. *Validation of the UK Met. Office's name model against the ETEX dataset*. Atmospheric Environment, vol. 32, no. 24, pages 4265 – 4276, 1998. (Cité en pages 41 et 73.)
- [Sanada 2014] Y. Sanada, T. Sugita, Y. Nishizawa, A. Kondo et T. Torii. *The aerial radiation monitoring in Japan after the Fukushima Daiichi nuclear power plant accident*. Progress in Nuclear Science and Technology, vol. 4, pages pp.76–80, 2014. (Cité en page 59.)
- [Saunier 2013] O. Saunier, A. Mathieu, D. Didier, M. Tombette, D. Quélo, V. Winiarek, et M. Bocquet. *An inverse modeling method to assess the source term of the Fukushima Nuclear Power Plant accident using gamma dose rate observations*. Atmos. Chem. Phys., vol. 13, pages pp.11403–11421, 2013. (Cité en pages 38, 44, 58, 73, 85, 88, 115, 116 et 119.)

- [Saunier 2016] O. Saunier, A. Mathieu, T.T. Sekiyama, M. Kajino, K. Adachi, M. Bocquet, T. Maki, Higarashi et D. Didier. *A new perspective on the Fukushima releases brought by newly available ^{137}Cs air concentration observations and reliable meteorological fields*. 2016. Budapest, Hungary. (Cité en pages 38, 58, 73, 85, 115, 117 et 119.)
- [Stohl 2012] A. Stohl, P. Seibert, G. Wotawa, D. Arnold, J. F. Burkhart, S. Eckhardt, C. Tapia, A. Vargas et T. J. Yasunari. *Xenon-133 and caesium-137 releases into the atmosphere from the Fukushima Dai-ichi nuclear power plant : determination of the source term, atmospheric dispersion, and deposition*. *Atmospheric Chemistry and Physics*, vol. 12, no. 5, pages pp.2313–2343, 2012. (Cité en pages 38, 73, 85, 117, 118, 119, 120 et 124.)
- [Sørensen 1998] Jens Havskov Sørensen. *Sensitivity of the DERMA long-range Gaussian dispersion model to meteorological input and diffusion parameters*. *Atmospheric Environment*, vol. 32, no. 24, pages 4195 – 4206, 1998. (Cité en pages 41 et 73.)
- [Sørensen 2016] J.H. Sørensen, B. Amstrup, H. Feddersen, J. Bartnicki, H. Klein, M. Simonsen, B. Lauritzen, S.C. Hoe, C. Israelson et J. Lindgren. *Fukushima Accident : UNcertainty of Atmospheric dispersion modelling (FAUNA)*. Rapport technique, Nordic Nuclear Safety Research (NKS), 2016. Rapport NKS-360. (Cité en page 45.)
- [Terada 2012] H. Terada, G. Katata, M. Chino et H. Nagai. *Atmospheric discharge and dispersion of radionuclides during the Fukushima Dai-ichi Nuclear Power Plant accident. Part II : verification of the source term and analysis of regional-scale atmospheric dispersion*. *Journal of Environmental Radioactivity*, vol. 112, pages pp.141–154, 2012. (Cité en pages 38, 58, 73, 117 et 119.)
- [Troen 1986] I Troen et Larry Mahrt. *A simple model of the atmospheric boundary layer : Sensitivity to surface evaporation*. *Boundary-Layer Meteorology*, vol. 37, pages 129–148, 01 1986. (Cité en pages 41 et 73.)
- [Winiarek 2012] V. Winiarek, M. Bocquet, O. Saunier et A. Mathieu. *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*, vol. 117, no. D5, 2012. (Cité en pages 38, 58, 73, 115, 117, 118 et 119.)
- [Winiarek 2014] V. Winiarek, M. Bocquet, N. Duhanyan, Y. Roustan, O. Saunier et A. Mathieu. *Estimation of the caesium-137 source term from the Fukushima Daiichi nuclear power plant using a consistent joint assimilation of air concentration and deposition observations*. *Atmospheric Environment*, vol. 82, pages pp.268–279, 2014. (Cité en pages 38, 73, 85, 117 et 119.)
- [Yamartino 2000] R. Yamartino. *Refinement of horizontal diffusion in photochemical grid models*. In 11th conference on the application of air pollution meteorology, AMS/AWMA, Long Beach, California, USA, 2000. (Cité en pages 41 et 73.)
- [Yue 2011] Thomas C.L. Yue. *A mesh-free numerical method for three-dimensional nonlinear Schrödinger equation*. In Department of Computer Science and Information Systems, Birkbeck, University of London, 2011. (Cité en page 21.)
- [Yumimoto 2016] Keiya Yumimoto, Yu Morino, Toshimasa Ohara, Yasuji Oura, Mitsuru Ebihara, Haruo Tsuruta et Teruyuki Nakajima. *Inverse modeling of the ^{137}Cs source term*

of the Fukushima Dai-ichi Nuclear Power Plant accident constrained by a deposition map monitored by aircraft. Journal of Environmental Radioactivity, vol. 164, 2016. (Cité en pages 38, 73, 85, 108, 117, 118 et 119.)