



HAL
open science

Propagation et réduction des incertitudes dans les modèles de changement d'occupation des sols

Ahlem Ferchichi

► **To cite this version:**

Ahlem Ferchichi. Propagation et réduction des incertitudes dans les modèles de changement d'occupation des sols . Intelligence artificielle [cs.AI]. Université de la Manouba 2017. Français. NNT: . tel-01667936

HAL Id: tel-01667936

<https://theses.hal.science/tel-01667936v1>

Submitted on 2 Jan 2018

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.

MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE
SCIENTIFIQUE

UNIVERSITÉ DE LA MANOUBA
ÉCOLE NATIONALE DES SCIENCES DE L'INFORMATIQUE



THÈSE

Présentée en vue de l'obtention du diplôme de

DOCTEUR EN INFORMATIQUE

par

Ahlem Ferchichi

Titre

**Propagation et réduction des incertitudes dans les modèles de
changement d'occupation des sols**

Réalisée au sein du laboratoire RIADI

**Laboratoire de Recherche en génie logiciel, Applications distribuées, systèmes
Décisionnels et Imagerie intelligente**

Soutenue le 14 décembre 2017 devant le jury composé de

Président : M. Moncef Tagina, Professeur, ENSI, Université de la Manouba

Rapporteur : M. Pierre Gançarski, Professeur, Université de Strasbourg (France)

Rapporteur : M. Ezzeddine Zagrouba, ISI, Professeur, Université Tunis El Manar

Examineur : M. Ouajdi Korbaa, Professeur, ISITCom, Université de Sousse

Directeur de thèse : M. Imed Riadh Farah, Professeur, ISAMM, Université de la Manouba

Résumé

Cette thèse présente une approche de propagation et de réduction des incertitudes liées au processus de changement d'occupation des sols (COS). Les modèles de COS ont, généralement, deux sources d'incertitude qui sont l'incertitude liée aux paramètres du modèle et l'incertitude liée à la structure du modèle. En outre, les incertitudes peuvent être subdivisées en deux types : aléatoire et épistémique. Le premier provient de la variabilité naturelle, alors que, le deuxième provient d'un manque de connaissance. Le problème principal est que la théorie de probabilités classique ne fait pas une distinction claire entre les incertitudes aléatoires et épistémiques de la manière dont elles sont représentées, c'est-à-dire les deux types sont décrits avec une distribution de probabilités. L'objectif de cette thèse est de propager et de réduire les incertitudes liées aux modèles de COS en utilisant la théorie des fonctions de croyance. Plus précisément, l'approche proposée est divisée en six étapes principales : 1) une étape d'identification des incertitudes de paramètres permettant d'identifier les paramètres d'entrée incertains, leurs types d'incertitude et leurs corrélations, 2) une étape de modélisation des paramètres permettant de modéliser et transformer tous les paramètres incertains dans le cadre de la théorie des fonctions de croyance, 3) une étape de propagation des incertitudes de paramètres permettant de propager les incertitudes aléatoires et épistémiques de paramètres tout en considérant les corrélations, 4) une étape de modélisation et de propagation des incertitudes liées à la structure du modèle permettant de modéliser et propager les incertitudes associées à la structure du modèle, 5) une étape d'analyse de sensibilité permettant de tester l'effet des sources d'incertitudes, et 6) une étape d'estimation basée sur les limites de confiance de Kolmogorov-Smirnov permettant d'optimiser les sources sélectionnées. Dans cette thèse, l'approche proposée est appliquée sur un scénario d'urbanisation. Les changements actuels et futurs d'une étude de cas ont été étudiés en se basant sur des images satellitaires Landsat multitemporelles. Ces données sont utilisées pour la préparation de la carte de prédiction de l'année 2025. Les résultats montrent que l'approche proposée basée sur la théorie des fonctions de croyance a le potentiel de réduire les incertitudes dans les modèles de COS.

Mots-clefs : Modèles de COS, propagation des incertitudes, incertitude de paramètres, incertitude de la structure du modèle, incertitude aléatoire, incertitude épistémique, analyse de corrélation, théorie des fonctions de croyance, analyse de sensibilité.

Abstract

This thesis presents an approach for propagating and reducing uncertainty related to the process of land cover change (LCC). LCC models have, almost, two sources of uncertainty which are the uncertainty related to model parameters and the uncertainty related to model structure. Furthermore, uncertainties can be subdivided into two types: aleatory and epistemic. The former comes from the natural variability, while the latter represents a lack of knowledge. The main problem is that classical probability theory does not make a clear distinction between aleatory and epistemic uncertainties in the way they are represented, i.e., both of them are described with a probability distribution. The aim of this thesis is to propagate and reduce the uncertainty related to LCC models using belief function theory. More precisely, the proposed approach is divided into six main steps: 1) a step for parameter uncertainty identification used to identify the uncertain input parameters, their types of uncertainty and their correlations, 2) a step for parameter modeling used to model and to transform all uncertain parameters in belief functions framework, 3) a step for parameter uncertainty propagation used to propagate aleatory and epistemic uncertainty of input parameters with considering correlations, 4) a step for model structure modeling and propagating used to model and propagate the uncertainty associated to model structure, 5) a step for sensitivity analysis performed to test the effect of uncertainty sources, and 6) a step for estimation based on confidence limits of Kolmogorov-Smirnov used to optimize the selected ones. In this thesis, the proposed approach is applied to an urbanization scenario. The present and future changes of a case study were studied using multi-temporal Landsat satellite images. These data are used for the preparation of a prediction map of the year 2025. Results show that the proposed approach based on the belief function theory has a potential for reducing uncertainty in LCC modeling.

Keywords : LCC models, uncertainty propagation, parameter uncertainty, model structure uncertainty, aleatory uncertainty, epistemic uncertainty, correlation analysis, belief function theory, sensitivity analysis.

Dédicaces

*Je voudrais tout d'abord remercier infiniment Dieu
je dédie mes travaux de thèse
à mes chers parents Mohamed et Latifa,
à mes soeurs Hela, Olfa, Sonia et Aya,
à mon frère Malek,
à toute ma famille,
à tous mes amis
et à tous les gens que j'aime*

Remerciements

Je tiens tout d'abord à adresser mes vifs remerciements et ma profonde reconnaissance à Monsieur Imed Riadh FARAH, Professeur à l'Université de la Manouba, pour la confiance, pour avoir dirigé mon travail, pour son assistance et sa disponibilité, pour tous les encouragements, les remarques, les suggestions et les soutiens continus et pour l'intérêt qu'il a manifesté pour mon travail.

Je tiens à remercier Monsieur Moncef TAGINA, Professeur à l'Université de la Manouba, pour l'honneur qu'il me fait en présidant ce jury.

Je remercie également Messieurs Pierre GANCARSKI, Professeur à l'Université de Strasbourg et Ezzeddine ZAGROUBA, Professeur à l'Université Tunis El Manar, d'avoir accepté la fastidieuse tâche de rapporter ce travail. Je les remercie pour les remarques et les critiques qu'ils m'ont prodiguées et qui ont permis de perfectionner le présent manuscrit.

Mes remerciements vont aussi à Monsieur Ouajdi KORBAA, Professeur à l'Université de Sousse, d'avoir bien voulu participer à ce jury et juger mon travail.

Merci à Monsieur Wadii BOULILA, pour son dévouement, ses conseils bienveillants, ses critiques et son aide précieuse.

Merci à tous les membres du laboratoire RIADI pour ces années passées ensemble et pour tout ce que vous avez apporté pour la finalisation de ce travail.

Table des matières

Introduction générale	1
I Les modèles de changement d'occupation des sols	4
1 Introduction	5
2 Définitions	5
2.1 Changement d'occupation et d'utilisation des sols	5
2.2 La télédétection et les COS	7
3 Modèles de changement d'occupation des sols	7
3.1 Introduction aux modèles de COS	7
3.2 Approches de modélisation du COS	9
3.3 Composants des modèles de COS	21
4 Conclusion	23
II Etude des incertitudes et de leur propagation dans les modèles de COS	24
1 Introduction	25
2 Types d'incertitudes	25
2.1 Incertitude aléatoire	26
2.2 Incertitude épistémique	26
2.3 Aléatoire ou épistémique ?	27
3 Classification des sources d'incertitudes	28
3.1 Incertitudes sur les paramètres	28
3.2 Incertitudes sur la structure des modèles de COS	31
4 Propagation des incertitudes dans les modèles de COS	32
4.1 Les méthodes probabilistes	32
4.2 Les méthodes non probabilistes	38

4.3 Synthèse et analyses comparatives	43
5 Conclusion	45
III Approche proposée pour la propagation des incertitudes dans les modèles de COS	46
1 Introduction	47
2 Approche de propagation des incertitudes	47
2.1 Analyse d'incertitude de paramètres du modèle	47
2.2 Analyse d'incertitude de la structure du modèle	58
3 Conclusion	64
IV Approche proposée pour la réduction des incertitudes dans les modèles de COS	65
1 Introduction	66
2 Approche de réduction des incertitudes	66
2.1 Analyse de sensibilité	67
2.2 Estimation	73
2.3 Evaluation de la performance des modèles de COS	75
3 Conclusion	76
V Expérimentation	78
1 Introduction	79
2 Application de l'approche proposée	79
2.1 Description du scénario d'application	79
2.2 Présentation de la base d'images et de la région d'étude	79
2.3 Présentation de modèles de COS et de paramètres utilisés	82
2.4 Résultats d'application de l'approche proposée de la propagation des incertitudes	82

2.5	Résultats d'application de l'approche proposée de la réduction des incertitudes	90
2.6	Résultats des cartes de prédiction de COS	94
3	Evaluation de l'approche proposée	99
4	Conclusion	100
	Conclusion générale et perspectives	101
	Liste des publications	104
	Bibliographie	106

Table des figures

1	Les modèles d'automates cellulaires.	10
2	Structure de modèle SLEUTH [Chaudhuri et Clarke 2013].	11
3	Structure du modèle CLUE [Veldkamp et Fresco 1996].	12
4	Structure des modèles basés sur les agents.	13
5	Structure de modèle de réseau de neurones artificiels.	14
6	Structure de modèle des machines à vecteurs de support.	15
7	Structure de modèle des algorithmes génétiques.	15
8	Structure de modèle CART.	16
9	Structure des modèles de Markov caché.	17
10	Structure de modèle DINAMICA.	19
11	Structure de modèle CA-MARKOV.	20
12	Structure de modèle LCM.	21
13	Modèle conceptuel d'un modèle de simulation [Verstegen et al. 2016].	22
14	Illustration du principe de propagation des incertitudes.	32
15	Distributions statistiques.	34
16	Intervalles focaux $[a_i, b_i]$ et les masses de croyance m_i associées.	41
17	Fonctions de croyance et plausibilité cumulatives.	42
18	Architecture de l'approche proposée pour la propagation des incertitudes.	48
19	Représentation de la fonction de densité Gaussienne de probabilités.	50
20	Structure évidentielle.	51
21	Relation entre deux paramètres dans le cadre probabiliste.	51
22	Relation entre deux paramètres représentés par des intervalles.	53
23	Convolution entre les structures évidentielles sans considération de la corrélation entre les deux paramètres A_e et B_a	56
24	Modèle ellipsoïdal basé sur le cadre de discernement conjoint.	57
25	Convolution entre les structures de Dempster-Shafer par le modèle ellipsoïdal avec considération de la corrélation entre les deux paramètres A_e et B_a	58
26	Représentation de l'incertitude de la structure du modèle (H_i sont les hypothèses et f_i sont les différentes structures de modèle).	59
27	Structure évidentielle pour chaque modèle f_1 et f_2	63

28	Structure évidentielle de la variable de sortie Y avec propagation de l'incertitude relative aux paramètres et à la structure du modèle.	63
29	Architecture de l'approche de réduction des incertitudes.	67
30	Représentation graphique de classement d'importance des cinq paramètres SI, BI, A, NDVI et Tem à travers deux structures f_1 et f_2 du modèle DINAMICA.	71
31	Comparaison entre la sensibilité de paramètres SI, BI, A, NDVI et Tem avec et sans considération de la corrélation pour les deux structures f_1 et f_2	72
32	Structures évidentielles pour la surface, le NDVI et la température.	73
33	Les structures évidentielles estimées pour la surface, le NDVI et la température.	74
34	Structure évidentielle de la variable de sortie Y avec réduction de l'incertitude de paramètres et de la structure du modèle.	75
35	Scénario d'application de l'approche proposée.	80
36	Localisation de la zone d'étude.	81
37	Cartes d'occupation des sols.	81
38	Valeurs de paramètres dans chaque type d'occupation des sols.	83
39	Sorties de quatre modèles de COS représentées par des distributions de croyance et de plausibilité avec seulement une propagation des incertitudes aléatoires de paramètres.	86
40	Sorties de quatre modèles de COS représentées par des distributions de croyance et de plausibilité avec une propagation des incertitudes aléatoires et épistémiques de paramètres.	87
41	Sorties de quatre modèles de COS représentées par des distributions de croyance et de plausibilité avec une propagation des incertitudes tout en considérant les corrélations entre les paramètres en comparaison aux celles où la corrélation n'est pas prise en compte.	88
42	Sorties de trois structures différentes des quatre modèles de COS représentées par des distributions de croyance et de plausibilité.	89
43	Sorties des quatre modèles de COS représentées par des distributions de croyance et de plausibilité avec la propagation des incertitudes de paramètres et de la structure du modèle en comparaison à celles causées par la propagation de paramètres seulement.	90
44	Classement de l'importance de tous les paramètres d'entrée à travers trois structures différentes pour les quatre modèles de COS.	92
45	Structures évidentielles estimées des paramètres les plus influents en comparaison avec celles non estimées.	94

46	Sorties des quatre modèles de COS représentées par des distributions de croyance et de plausibilité avec l'application de l'approche de la réduction des incertitudes.	95
47	Sorties des quatre modèles de COS en comparaison aux changements réels entre 1987 et 2014 de la région du Caire : (a) résultats sans application de l'approche proposée et (b) résultats avec application de l'approche proposée.	96
48	Comparaison entre les quatre modèles de COS pour les critères de performance R^2 , NSE et $PBIAS$ dans les deux périodes de validation. . . .	97
49	Croissance urbaine dans la région du Caire à travers les modèles DYNAMICA, SLEUTH, CA MARKOV et LCM.	98

Liste des tableaux

1	Sources d'incertitude de paramètres d'entrée des modèles de COS.	29
2	Approches de réduction d'incertitude liée aux paramètres des modèles de COS.	30
3	Comparaison des études de propagation des incertitudes dans les modèles de COS.	44
4	Description, formule et gamme de paramètres d'entrée du modèle.	49
5	Exemple	52
6	L'ensemble de BPA pour la sortie $Y = A_e + B_a$ obtenu au moyen de produit cartésien entre les paires interval-masse $\{([1, 6], 0, 25), ([5, 9], 0, 5), ([7, 10], 0, 25)\}$ de A_e et $\{([2, 4], 0, 5), ([3, 5], 0, 5)\}$ de B_a sous l'hypothèse de l'indépendance.	55
7	Collection de paires intervalle-BPA de cinq paramètres (indice de forme, indice de luminosité, surface, NDVI et température).	62
8	Mesure d'importance des cinq paramètres SI, BI, A, NDVI et Tem à travers deux structures f_1 et f_2 du modèle DINAMICA.	71
9	Mesure d'importance des cinq paramètres SI, BI, A, NDVI et Tem avec et sans considération de la corrélation à travers deux structures f_1 et f_2	72
10	Collection de paires d'intervalles-masses de surface, NDVI et température.	73
11	Les limites de confiances du Kolmogorov-Smirnov pour la surface, le NDVI et la température.	74
12	Valeurs incertaines de paramètres aléatoires et épistémiques pour chaque type d'occupation des sols (Urb : urbain, Agc : agriculture, Wat : eau, et Dsr : désert).	84
13	Coefficients de corrélations entre tous les paramètres.	85
14	Top huit paramètres les plus importants identifiés par les structures originales de modèles de COS DINAMICA, SLEUTH, CA-MARKOV et LCM.	93
15	Comparaison entre les changements réels, les changements obtenus par l'application de l'approche proposée à travers quatre modèles de COS et ceux obtenus par les approches présentées dans [Ferchichi <i>et al.</i> 2016] et [Boulila <i>et al.</i> 2014].	100

Introduction générale

Contexte et problématique de la thèse

La surface de la terre a été considérablement changée au cours des dernières décennies par les humains en raison de l'urbanisation, de la déforestation et de l'agriculture. Ainsi, le changement d'occupation des sols (COS) est apparu comme un facteur fondamental du changement environnemental mondial et a donc reçu une attention croissante des scientifiques et des décideurs. Un certain nombre de modèles ont été établis pour simuler les tendances de COS tels que DINAMICA [Soares-Filho *et al.* 2002, De Almeida *et al.* 2005], SLEUTH (Slope, Land Use, Exclusion, Urban, Transportation, and Hill Shading) [Jantz *et al.* 2010, Hua *et al.* 2014], CA-MARKOV (Cellular Automata-MARKOV) [Arsanjani *et al.* 2011, Gong *et al.* 2015], LCM (Land Change Modeler) [Tewolde et Cabral 2011, Mishra *et al.* 2014], etc. Ces modèles qui sont développés pour mieux comprendre, évaluer et prédire le changement futur d'occupation des sols, contiennent des incertitudes dans leurs composants principaux tels que les paramètres d'entrée et la structure du modèle. Ces incertitudes se propagent à travers le modèle, ce qui génère des incertitudes dans l'état de système modélisé, c'est-à-dire dans les prédictions d'occupation des sols [Krüger et Lakes 2015]. Les prédictions modélisées basées sur des résultats incertains peuvent conduire à un biais ou à une fausse précision dans les prédictions issues des modèles de COS. Tenir compte des incertitudes dans les modèles de COS est devenu alors une étape évidente, voire nécessaire, pour améliorer les prédictions et la robustesse de ces modèles [Verburg *et al.* 2013, Versteegen *et al.* 2016].

Malgré les récents développements significatifs en puissance de calcul et de modélisation de COS, la question de la façon d'aborder correctement les incertitudes associées aux prédictions des modèles de COS reste difficile [van der Kwast *et al.* 2012]. Cette question doit être dûment prise en compte pour la modélisation de COS pour réaliser son potentiel maximum pratique dans les processus décisionnels environnementaux. De toute évidence, la clé pour traiter correctement les incertitudes dans les modèles de COS est de comprendre, de propager et de réduire les incertitudes inhérentes à la modélisation des changements d'une manière cohérente et systématique [Versteegen *et al.* 2016].

Pour aborder adéquatement les incertitudes dans la modélisation de COS, il y a trois aspects distincts mais connexes à considérer : la compréhension, la propagation et la réduction des incertitudes. De toute évidence, la compréhension des incertitudes fait partie intégrante de toute application de la propagation et/ou de la réduction des incertitudes.

D'autre part, pour gérer ces incertitudes de nombreuses méthodes d'analyse d'incertitude ont été introduites telles que la théorie des fonctions de croyance, qui peut être utilisée pour mettre en oeuvre des procédures et une meilleure compréhension de réduire les incertitudes [Lindley 2013].

De toute évidence, sans une bonne compréhension préalable de tous les types et les sources des incertitudes et les relations entre eux, il est difficile de propager et de réduire les incertitudes de manière significative. Ceci est dû au fait que différents types d'incertitudes peuvent introduire des caractéristiques d'erreur significativement différentes qui nécessitent des techniques différentes pour les traiter ; et l'absence de sources importantes des incertitudes peuvent conduire à des prédictions trompeuses dans les résultats

des modèles de COS. À ce jour, notre compréhension sur les incertitudes des modèles de COS est encore loin d'être terminée et il ya beaucoup de place pour de nouveaux efforts dans la recherche de moyens cohérents et systématiques pour aborder ce sujet.

Dans la littérature, plusieurs sources et types des incertitudes sont traités. Deux grandes sources se démarquent : l'incertitude de paramètres et l'incertitude de la structure du modèle. La première est l'incertitude des paramètres entrants d'un modèle. En revanche, la deuxième source est l'erreur introduite par la façon dont le modèle est fait. Ces incertitudes peuvent avoir deux origines : le premier provient du caractère aléatoire de l'information du à une variabilité naturelle résultant de phénomènes stochastiques. On parle alors d'incertitude de variabilité ou d'incertitude stochastique. Le second est lié au caractère imprécis de l'information lié à un manque de connaissance et qui résulte par exemple d'erreurs systématiques lors de mesures ou d'avis d'experts. On parle alors d'incertitude épistémique. Dans le calcul de risque, ces deux notions sont souvent confondues alors qu'elles devraient être traitées de manière différente. En outre, l'analyse de corrélation entre les paramètres d'entrée du modèle est souvent négligée dans le processus de traitement des incertitudes. Cela affecte la fiabilité du modèle et rend plus difficile de mieux déterminer l'incertitude liée aux paramètres du modèle.

Dans cette thèse, il y a plusieurs défis scientifiques à relever : (1) définir un modèle de COS permettant de prédire, d'une manière fidèle, le changement futur d'occupation des sols en présence de plusieurs sources et types des incertitudes et en présence des informations corrélées ; (2) analyser et étudier, d'une manière intelligente, les informations incertaines et corrélées en développant des mécanismes de modélisation et de propagation dans le cadre de la théorie des fonctions de croyance ; et (3) proposer des techniques de réduction afin de rendre les modèles de COS plus performants et faire face aux différents types et sources des incertitudes.

Objectif de la thèse

L'objectif principal de cette thèse est de proposer une approche de propagation et de réduction des incertitudes affectant les résultats des modèles de COS.

Les objectifs de cette thèse sont :

- Proposer une modélisation des incertitudes aléatoires et épistémiques liées aux paramètres et à la structure du modèle dans le cadre de la théorie des fonctions de croyance.
- Etablir une étude des corrélations entre les paramètres d'entrée incertains des modèles de COS basée sur le coefficient de corrélation.
- Suggérer une propagation des incertitudes de paramètres et de la structure du modèle basée sur la théorie des fonctions de croyance. Ceci permet de donner des informations sur l'incertitude associée à la sortie des modèles de COS calculée au regard des incertitudes sur les variables critiques.
- Faire une réduction des incertitudes basée sur l'analyse de sensibilité. Cette analyse tente de trouver les variables les plus importantes pour la calibration des modèles de COS afin de réduire leurs incertitudes.

Structure de la thèse

Ce manuscrit est organisé en deux parties :

-
- La première partie est dédiée à l'état de l'art qui est composé de deux chapitres.

Le chapitre I présente dans sa première partie une étude bibliographique sur les principaux modèles de COS. Nous présentons un aperçu général sur les changements d'occupation et d'utilisation des sols. Ainsi, nous montrons l'importance de données de télédétection pour le suivie. La deuxième partie de ce chapitre est réservée pour la présentation des différentes approches de modélisation de COS et ses structures. Enfin, nous donnons un aperçu sur les principaux composants d'un modèle de simulation.

Le chapitre II passe en revue les différents types et sources des incertitudes qui sont rencontrés lors de l'analyse d'un modèle de COS. Ainsi, nous décrivons les deux grandes familles d'approches de la propagation des incertitudes définies dans les applications des modèles de COS.

- Dans la deuxième partie, nous décrivons nos contributions. Cette partie est structurée en trois chapitres.

Dans le chapitre III, nous présentons l'approche proposée de la propagation des incertitudes dans les modèles de COS. Le processus de cette approche consiste en trois étapes principales à savoir la modélisation des incertitudes aléatoires et épistémiques, l'étude de corrélation et la propagation des incertitudes de paramètres et de la structure du modèle.

Dans le chapitre IV, nous présentons l'approche proposée de la réduction des incertitudes dans les modèles de COS. Le processus de cette approche consiste en trois étapes principales à savoir l'analyse de sensibilité, l'estimation et l'évaluation de performance.

Le chapitre V détaille l'expérimentation menée pour valider les contributions proposées. Nous présentons, en premier lieu, les principaux résultats obtenus suite à l'application de l'approche sur un scénario d'urbanisation. En second lieu, nous consacrons à l'évaluation des performances de cette approche.

Enfin, un bilan global de la thèse est fourni en conclusion générale où nous rappelons nos différentes contributions. Un ensemble de perspectives pour la recherche future est également discuté dans ce bilan.

Chapitre I

Les modèles de changement d'occupation des sols

Sommaire

1	Introduction	5
2	Définitions	5
2.1	Changement d'occupation et d'utilisation des sols	5
2.2	La télédétection et les COS	7
3	Modèles de changement d'occupation des sols	7
3.1	Introduction aux modèles de COS	7
3.2	Approches de modélisation du COS	9
3.2.1	Modèles basés sur les automates cellulaires	9
3.2.2	Modèles basés sur la régression	11
3.2.3	Modèles basés sur les agents	13
3.2.4	Modèles basés sur l'apprentissage automatique	14
3.2.5	Modèles basés sur les chaines de Markov	17
3.2.6	Modèles hybrides	18
3.3	Composants des modèles de COS	21
4	Conclusion	23

1 Introduction

Ce premier chapitre présente une étude bibliographique, non exhaustif, sur les principaux modèles de COS. À cet effet, nous présentons les différentes approches de modélisation de COS qui sont discutées dans la littérature. Après avoir présenté le principe et les caractéristiques de chaque approche, les structures et les composants de ces modèles sont explorés en profondeur. Dans ce chapitre, nous présentons d'abord un aperçu général sur les changements d'occupation et d'utilisation des sols. Ensuite, nous montrons l'importance de données de télédétection sur les COS.

2 Définitions

Dans cette section, nous présentons plusieurs aspects liés au COS qui constitue notre domaine d'intérêt dans cette thèse. Ces aspects concernent : le changement d'occupation des sols, le changement d'utilisation des sols, l'urbanisation, l'intensification agricole, la déforestation et le rôle de la télédétection.

2.1 Changement d'occupation et d'utilisation des sols

Dans le but de prédire les COS avec un minimum d'impacts négatifs, une bonne compréhension du système terrestre est nécessaire. Les changements d'occupation et d'utilisation des sols ont un large éventail de facteurs qui peuvent être naturels ou anthropiques. Les changements dans l'utilisation des sols (*land use change*, en anglais) sont définis comme toutes les activités humaines sur la terre. Les changements dans l'occupation des sols (*land cover change*, en anglais) correspondent à une classification physique du terrain. Les COS peuvent être naturels, par exemple, la succession écologique ou les incendies de forêt naturels [Cihlar et Jansen 2001], mais peuvent être également induits par l'homme, par exemple, un changement des zones cultivées en zones urbaines. Dans le dernier cas, le COS est causé par un changement d'utilisation des sols, pour l'exemple donné, un changement de l'utilisation agricole à l'utilisation résidentielle. Dans le débat sur la durabilité, l'intérêt porté aux facteurs de changement d'occupation et d'utilisation des sols est axé sur les facteurs naturels et humains, car ceux-ci sont ceux que nous pouvons essayer d'orienter pour établir des changements plus durables. Pour cette raison, l'accent est mis dans cette thèse sur les changements induits par la nature et l'homme.

L'identification, la délimitation et la cartographie des COS sont considérées comme des activités importantes à l'appui de la gestion durable des ressources naturelles. Selon Giri *et al.* [Giri *et al.* 2003], l'analyse des séries chronologiques du COS et l'identification des forces motrices responsables de ces changements sont nécessaires pour la gestion durable des ressources naturelles et à la prédiction des trajectoires d'occupation du sol. Dans [Lambin *et al.* 2003], Lambin *et al.* ont affirmé que les informations sur le COS sont nécessaires en ce qui concerne les changements qui surviennent, où et quand ils se produisent, les taux auxquels ils se produisent et les forces sociales et physiques qui entraînent ces changements. Dans [Sala *et al.* 2000], Sala *et al.* estiment qu'au 21^e siècle, les COS ont le plus grand impact global, environ 1,5 fois plus grand que le changement climatique et plus de deux fois plus important que d'autres facteurs tels

que les changements dans le CO_2 atmosphérique et le dépôt d'azote.

La conversion et la modification sont les deux formes de COS [Turner *et al.* 1995, Lambin *et al.* 2003] où la première est un remplacement complet d'une catégorie d'occupation du sol par une autre (par exemple, le gain urbain, la déforestation ou la perte agricole). Le dernier est, cependant, un changement dans une catégorie d'occupation des sols. L'intensification agricole est un exemple de modification d'occupation des sols qui augmente la production alimentaire [Tilman 1999]. La détection de la modification d'occupation des sols nécessite des représentations d'occupation des sols, qui varient progressivement dans l'espace et dans le temps. Par conséquent, le COS s'effectue par de nombreuses voies telles que l'urbanisation, la déforestation, l'intensification agricole, la désertification, etc. En ce qui suit, nous allons présenter brièvement les trois termes : l'urbanisation, l'intensification agricole et la déforestation.

Urbanisation

L'urbanisation est une expression désignant le phénomène de développement des surfaces urbanisées en périphérie des villes. Le facteur le plus apparent du COS est la croissance de la population mondiale [Ramankutty et Foley 1999, Agarwal *et al.* 2002, Alexander *et al.* 2015], car les gens ont besoin d'espace pour vivre. L'augmentation de la croissance de la population chaque année est toujours suivie par la demande croissante des terres, en particulier les besoins de la production, de logement, d'agriculture et d'autres. Récemment, l'espèce humaine a franchi un seuil important puisque plus de la moitié d'entre nous vivent maintenant dans les villes. L'empreinte urbaine double aujourd'hui tous les 30 ans. À ce rythme, l'utilisation des terres urbaines devrait atteindre 10% de la superficie du sol d'ici 2070.

Intensification agricole

L'expansion des terres agricoles est un phénomène fondé sur un accroissement de la production agricole optimisé par rapport à la disponibilité des facteurs de production. Ce phénomène est également une tendance majeure du COS mondial. À l'heure actuelle, les terres agricoles couvrent environ un tiers de la surface terrestre mondiale et se sont développées dans les forêts, mais aussi dans les steppes et les savanes, pour répondre à la demande croissante de produits alimentaires [Lambin *et al.* 2003]. Ce processus d'expansion agricole a été progressivement remplacé dans les années 1960 par un processus d'intensification dans certaines régions du monde, à savoir une augmentation de la production alimentaire par hectare, principalement grâce au labour mécanisé, à l'utilisation d'engrais et à l'irrigation.

Déforestation

La déforestation est le phénomène de régression des surfaces couvertes de forêts. Elle est considérée aussi comme la principale cause de dégradation des sols [Allen *et al.* 2013]. Au cours des deux dernières décennies, moins de 60 000 arbres forestiers ont été enlevés et remplacés par des arbres fruitiers, des maisons ou des bâtiments [Manson 2007]. Les décideurs, les gestionnaires des ressources naturelles et les décideurs sont intéressés à préserver les écosystèmes forestiers pour la conservation de la biodiversité [Kirschbaum 2000], à soutenir les produits du bois [Wernick 2007], à réduire les émissions dues au déforestation et à la dégradation et à diminuer les impacts sur le cycle de l'eau [Peng *et al.* 2002, Ray *et al.* 2010]. Les gens dans le monde sont aujourd'hui plus dépendants des ressources forestières pour répondre aux besoins essentiels que jamais auparavant (par

exemple l'oxygène et les aliments) [FAO 2009, Pijanowski *et al.* 2010].

2.2 La télédétection et les COS

Allen *et al.* [Allen *et al.* 2013] ont montré que les méthodes traditionnelles de collecte de données telles que les données démographiques et les cartes d'échantillons n'étaient pas satisfaisantes dans le but de gérer les COS. Des informations exactes sur les COS sont donc très essentielles pour l'analyse de ces changements. Pour obtenir cette information, les données de télédétection peuvent être utilisées car elles fournissent des informations sur l'occupation du sol. La télédétection se réfère à la science ou à l'art d'acquérir des informations d'un objet ou d'un phénomène sur la surface terrestre sans contact physique avec celle-ci. Cela peut se faire en détectant et en enregistrant l'énergie réfléchie ou émise et les informations traitées, analysées et appliquées à un problème donné [Allen *et al.* 2013].

La télédétection peut fournir des ensembles de données spatialement cohérentes qui couvrent de vastes zones avec des détails élevés et une fréquence temporelle élevée, y compris des séries chronologiques historiques. Cela a été de plus en plus efficace avec les capteurs émergents qui fournissent des données à haute résolution spectrale, spatiale et temporelle aux chercheurs. Les informations dérivées de la télédétection peuvent aider à estimer les niveaux et les taux de déforestation, d'urbanisation, de dégradation des zones humides et de nombreux autres phénomènes au niveau du sol, ceci permet une meilleure compréhension pour la planification et la gestion des terres [Longley *et al.* 2001].

La télédétection peut être utilisée pour évaluer spatialement les COS. Mais, les changements futurs ne sont pas simplement une extrapolation des changements passés. De ce fait, des modèles de simulation sont utilisés pour projeter les changements futurs de l'occupation des sols [Veldkamp *et al.* 2001]. Les impacts du COS peuvent varier considérablement dans l'espace. Par conséquent, des modèles de COS spatialement explicites sont appropriés pour étudier les impacts attendus. Nous trouverons ci-après une étude détaillée sur les modèles de COS qui sont utilisés pour produire des prédictions de COS et les composants de ces modèles.

3 Modèles de changement d'occupation des sols

Cette section commence à fournir les notions de base qui sous-tendent les modèles de simulation du COS. Les approches de modélisation sont ensuite mises en évidence.

3.1 Introduction aux modèles de COS

Les modèles de COS commencent par une situation d'occupation du sol initiale pour une zone de cas d'étude donnée. Ils utilisent une fonction de transition inférée, représentant les processus de changement, pour simuler l'expansion et la contraction d'un ensemble prédéfini de types d'occupation des sols sur une période donnée. Ces modèles contribuent à améliorer notre compréhension du système terrestre en établissant

des relations cause-effet et en les testant sur des données historiques. De cette façon, ils aident à identifier les facteurs de COS et leur importance relative. De plus, les modèles de COS peuvent être utilisés pour explorer les changements futurs d'occupation des sols pour différents scénarios [Lambin *et al.* 2000, Magliocca *et al.* 2015].

Le développement du modèle doit tenir compte de la calibration et de la validation. La calibration est le réglage des paramètres d'entrée afin d'assurer la meilleure qualité d'ajustement entre la sortie du modèle et les données observées. La validation, en revanche, montre que le modèle est précis en fonction de l'utilisation prévue du modèle [Rykiel 1996]. Il existe deux façons communes de séparer les processus de calibration et de validation les uns des autres [Pontius *et al.* 2004] : le temps et l'espace.

La séparation par le temps est la méthode la plus courante utilisée pour la modélisation du COS. Un sous-ensemble de l'ensemble de données entre les temps t_1 et t_2 est choisi aléatoirement pour former le modèle et ensuite l'information entière au temps t_1 est utilisée pour simuler le changement de t_1 à un point ultérieur t_2 . Ensuite, la carte simulée de t_2 est comparée à une carte réelle de t_2 . La validation est réalisée pour évaluer comment le modèle peut prédire une troisième étape de temps au futur [Pontius *et al.* 2004]. Toutes les informations qui sont acquises à l'instant t_2 sont utilisées pour prédire le changement de t_2 à une date ultérieure t_3 . La carte prédite de t_3 est généralement comparée à une carte observée de t_3 pour évaluer le niveau d'accord entre les deux.

La séparation par l'espace est une autre manière commune. Le modèle utilise les données de la première région pour adapter les paramètres du calibrage. Ainsi, un sous-ensemble de l'ensemble de données de la première région a été sélectionné au hasard pour former le modèle. Puis, l'ensemble de l'information provenant de la première région a été utilisé pour simuler les COS. Ensuite, la carte simulée est comparée à une carte réelle pour la première région. Le modèle ajusté pour la première région est appliqué à l'ensemble de l'information provenant de la deuxième région pour prédire les COS pour la validation. Après, la carte simulée est comparée à une carte réelle pour la deuxième région. La plupart des modèles de COS utilisent au moins deux cartes de séries temporelles pour la calibration [Veldkamp et Fresco 1996].

Ces modèles peuvent utiliser un large éventail d'informations (images de télédétection classifiées, variables biophysiques et socioéconomiques, scénarios et indicateurs économiques, enquêtes sur le terrain, etc.). En général, le COS est modélisé en utilisant les COS passés ou la distribution spatiale de l'occupation des sols. Le principe est de développer un modèle mathématique qui estime le potentiel de changement en fonction d'un ensemble de variables spatiales explicatives [Veldkamp et Lambin 2001, Paegelow et Camacho Olmedo 2008]. Ce modèle mathématique peut être également basé sur un ensemble des hypothèses théoriques. Il compare habituellement deux cartes d'occupation des sols à partir de deux dates différentes pour estimer les modèles et les processus de changement (types de transitions et taux de changement) pour le paramétrage du modèle. L'analyse des changements passés ou des distributions d'occupation des sols permet d'évaluer dans quelle mesure les lieux pourraient probablement changer dans le futur [Kolb *et al.* 2013]. Ce potentiel de changement est également appelé propension, probabilité, susceptibilité et adéquation dans la littérature. Ces cartes par elles-mêmes peuvent être considérées comme un premier produit généré dans le processus de modélisation.

3.2 Approches de modélisation du COS

Diverses approches de modélisation du COS existent dans la littérature. Ces approches peuvent être classées de différentes façons [Parker *et al.* 2003, Gaucherel et Houet 2009], sur la base : des objectifs de modélisation, des théories appliquées, des types d'occupation des sols, des types de données utilisés, des perspectives disciplinaires et des niveaux spatiaux et temporels d'analyse. Plusieurs travaux récents ont distingué six catégories principales de modèles [Falahatkar *et al.* 2011, Arsanjani *et al.* 2013, Tayyebi *et al.* 2014, Parsa *et al.* 2016, Samardzic-Petrovic *et al.* 2016] : i) modèles basés sur les automates cellulaires, ii) modèles basés sur la régression ou les systèmes statistiques, iii) modèles basés sur les agents ou les systèmes évolutifs, iv) modèles basés sur l'apprentissage automatique, v) modèles basés sur les chaînes de Markov, et vi) modèles hybrides. Les sous sections suivantes représentent une description de chacune de ces approches tout en résumant sa structure et ses applications.

3.2.1 Modèles basés sur les automates cellulaires

La méthode de modélisation de COS la plus utilisée jusqu'à présent est les automates cellulaires. En 1940, Ulam et Von Neumann ont développé les automates cellulaires pour simuler des modèles complexes [Neumann et Burks 1966]. Les automates cellulaires, qui peuvent capturer une grande variété de comportements locaux et de modèles globaux [Wolfram 1984]. Ils comprennent cinq composantes de base (Figure 1) : (1) espace de grille qui peut être représenté sous forme de cellules régulières ou irrégulières, (2) chaque cellule a un statut qui peut changer par les attributs de collection de cellules dans ses voisins, (3) les règles de transition qui sont utilisées pour classer les données, (4) un voisinage qui définit l'étendue de l'influence des cellules qui entourent la cellule centrale, (5) une étape de temps.

L'objectif de la calibration des automates cellulaires est de trouver la meilleure combinaison de règles de transition pour modéliser les COS [Batty et Xie 1994a, Batty et Xie 1994b, Landis et Zhang 1998, Batty *et al.* 1999]. Les automates cellulaires peuvent délibérément articuler des modèles globaux à travers des processus locaux [Batty et Xie 1994a, Batty et Xie 1994b]. Les automates cellulaires ont été utilisés pour simuler divers phénomènes spatio-temporels, y compris les COS [Ménard et Marceau 2007, Xua *et al.* 2016], la simulation du changement urbain [Batty *et al.* 1999, Deep et Saklani 2014, Aburas *et al.* 2016], la propagation du feu [Jellouli *et al.* 2016] et la compétition d'espèces [Chopard et Lagrava 2006].

Il est difficile de définir des meilleures combinaisons de règles de transitions lorsqu'il y a beaucoup de variables, car les modèles de COS sont complexes [Batty et Xie 1994a, Li et Yeh 2000]. Les variations sont dues aux nombreuses façons possibles de définir les règles de transition. Il y a deux manières pour effectuer le calibrage des automates cellulaires. La première, connu comme un modèle hybride, consiste à l'utilisation d'autres méthodes pour calibrer les modèles des automates cellulaires [Wu 2002]. La deuxième manière, appelée approche essai-erreur, ne nécessite pas l'utilisation d'autres approches. Les résultats de simulation obtenus à partir de différentes combinaisons de paramètres sont comparés dans [Clarke *et al.* 1997]. White *et al.* [White *et al.* 1997] ont proposé une méthode intuitive basée sur l'approche essai-erreur pour obtenir une matrice de

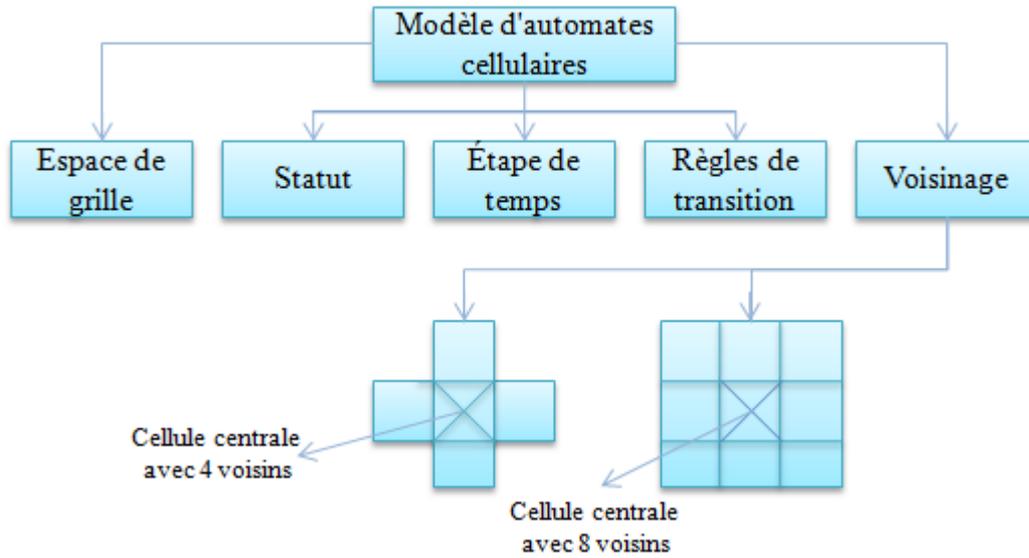


FIGURE 1 – Les modèles d'automates cellulaires.

paramètres pour la simulation urbaine, mais ils ont montré que cette approche est très coûteuse en temps.

Le modèle SLEUTH

SLEUTH (*Slope, Land Use, Exclusion, Urban, Transportation, and Hill Shading*) est l'un des modèles de COS de la première génération qui utilise les automates cellulaires pour simuler l'urbanisation et qui a été appliqué pour la première fois dans la région de la baie de San Francisco [Clarke *et al.* 1997]. Il utilise des données spatiales et temporelles de deux temps ou plus pour simuler le gain urbain dans le futur ou la perte urbaine dans le passé [Clarke *et al.* 1997, Clarke et Gaydos 1998, Yang et Lo 2003, Rienow et Goetzke 2014]. La version originale de SLEUTH a été modifiée pour une variété d'applications. SLEUTH a été appliqué aux grandes villes aux États-Unis, y compris Detroit, Chicago, New York, Washington, San Francisco et Albuquerque, et d'autres pays tels que les Pays-Bas, le Portugal et l'Australie.

Le modèle SLEUTH a la capacité de simuler quatre formes d'expansion urbaine (Figure 2) : expansion par diffusion, en continuité de l'urbain existant, le long des routes et par création de nouveaux centres. Ce modèle permet aux utilisateurs : (i) une maîtrise de la quantité de changement de manière exogène au modèle et indépendante des tendances passées ; (ii) une maîtrise des formes urbaines en offrant à l'utilisateur la possibilité de spécifier les pourcentages contributifs de chacune des formes d'urbanisation adaptées à chaque scénario ; et (iii) l'intégration d'un facteur additionnel de localisation des changements, en l'occurrence l'attractivité du territoire ou le prix des loyers. De plus, SLEUTH est conçu pour exécuter de façon automatique la spatialisation dynamique d'un scénario prospectif décliné en sous-périodes présentant des dynamiques spatiales variables au cours du temps [Doukari *et al.* 2016].

Pour calibrer SLEUTH, les valeurs initiales des paramètres de l'utilisateur sont requises et le modèle utilise Monte Carlo comme une approche itérative pour vérifier la combinaison des ensembles de paramètres. Monte Carlo trouve les paramètres à travers trois étapes itératives (par exemple grand, moyen et petit) afin de minimiser la diffé-

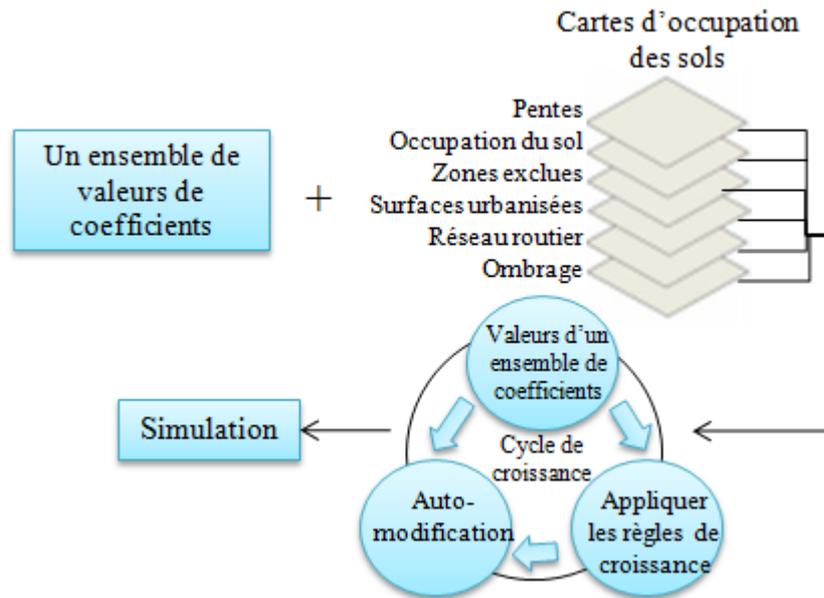


FIGURE 2 – Structure de modèle SLEUTH [Chaudhuri et Clarke 2013].

rence entre la carte simulée et la carte de référence. Plusieurs études ont été consacrées à la calibration de SLEUTH [Jantz *et al.* 2003, Nigussie et Altunkaynak 2016]. Les paramètres et la carte d'occupation des sols sont la sortie de l'étape de calibration [Clarke et Gaydos 1998]. SLEUTH utilise également une fonction d'auto-modification pour une simulation réaliste [Clarke *et al.* 1997]. Cette dernière modifie les valeurs des coefficients au fur et à mesure de l'itération du modèle. Sans auto-modification, SLEUTH simule un taux de croissance linéaire, produisant le même nombre de nouvelles cellules urbaines. Les détails sur le modèle SLEUTH et sa mise en oeuvre sont disponibles sur le site Web du Projet Gigalopolis (<http://www.ncgia.ucsb.edu/projects/gig/>).

3.2.2 Modèles basés sur la régression

Les entrées de la régression logistique sont des valeurs des variables prédictives à l'instant t_1 tandis que la sortie est un changement binaire entre t_1 et t_2 . La sortie de la régression logistique est égale à 1 (indique qu'il y a un changement) ou 0 (indique pas de changement). La fonction de régression logistique donnée par He et Lo [He et Lo 2007] est bornée entre 0 et 1. La sortie de régression logistique donne la probabilité de COS pour chaque cellule en fonction des variables prédictives spatiales. L'interception et les coefficients doivent être estimés, respectivement, comme un paramètre fixe et pour chaque prédicteur spatial, respectivement. La régression logistique a une forme non linéaire, mais peut être transformée en une forme linéaire avec une transformation simple [Pontius et Schneider 2001, Tayyebi *et al.* 2010]. Ensuite, les coefficients des variables prédictives spatiales sont estimés en utilisant la fonction transformée. Au lieu d'adapter une régression logistique avec l'aptitude de COS dans chaque cellule comme le résultat, le logarithme de la probabilité est considéré comme le résultat. Si un coefficient de régression particulier est nul, alors la variable explicative correspondante n'est pas associée à l'occurrence de la réponse.

Le modèle CLUE

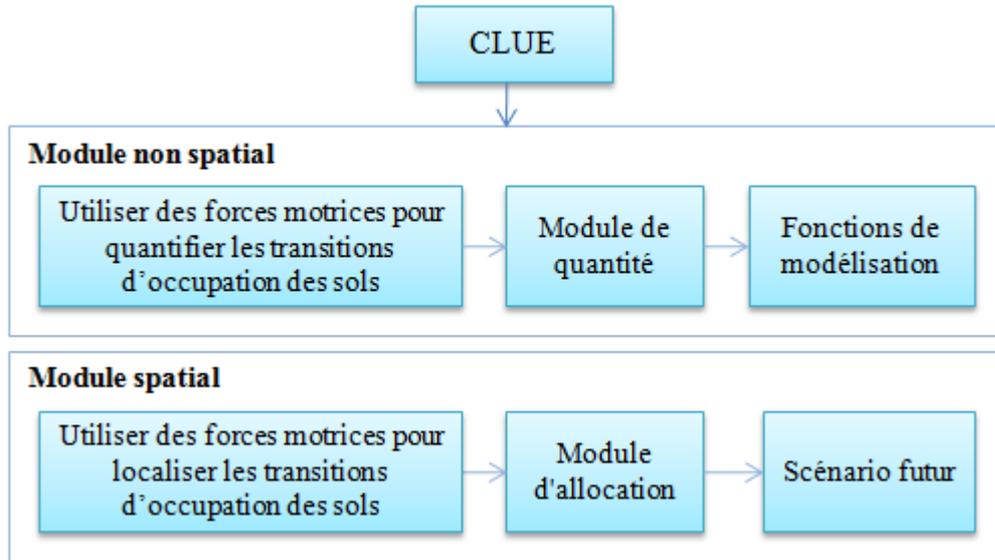


FIGURE 3 – Structure du modèle CLUE [Veldkamp et Fresco 1996].

Le modèle CLUE est l'un des modèles de base qui utilise la régression logistique pour simuler les COS. Il utilise des relations empiriques quantifiées entre l'occupation des sols et ses facteurs en combinaison avec une modélisation dynamique de la compétition entre les classes d'occupation des sols [Veldkamp et Fresco 1996, Verburg *et al.* 1999]. La deuxième version du modèle a été développée pour une application régionale et a été appelée CLUE-S (*Conversion of Land Use and its Effects at Small regional extent*) [Verburg *et al.* 2002, Verburg et Veldkamp 2004]. CLUE a été appliqué à l'échelle nationale pour l'Equateur, la Chine et l'Indonésie. Le modèle CLUE comporte deux modules distincts, appelés un module de quantité et un module d'allocation (Figure 3). Le module de quantité calcule le nombre de transitions pour chaque classe d'occupation des sols ; tandis que le module d'allocation localise la quantité donnée de COS à des emplacements différents. Le module de quantité en CLUE peut exécuter des modèles différents allant des modèles simples aux modèles complexes. Les résultats du module de quantité sont une entrée directe pour le module d'allocation.

Le modèle CLUE intègre quatre composantes principales : -1) option politique met en évidence les zones de la carte où les COS sont restreints ou peuvent impliquer des arrangements de stimulation pour une certaine utilisation du sol sur un emplacement. CLUE utilise une matrice de conversion de COS pour afficher les transitions qui sont interdites par une certaine stratégie. -2) Les paramètres de conversion et les séquences de transition d'occupation des sols sont deux ensembles de paramètres pour caractériser les catégories d'occupation des sols. Les paramètres de conversion sont liés à la réversibilité de COS allant de 0 (conversion facile) à 1 (changement irréversible). Les séquences de transition d'occupation des sols montrent les paramètres de conversion et les caractéristiques temporelles qui peuvent être définies à l'aide d'une matrice de conversion. -3) Les exigences d'occupation des sols contrôlent la simulation en définissant les changements requis de l'occupation des sols (quantité de COS). L'extrapolation des tendances dans les données historiques de COS dans le futur proche est une technique courante pour calculer les exigences d'occupation des sols. -4) Les conversions d'occupation des sols devraient avoir lieu dans les endroits ayant la plus grande aptitude (emplacement

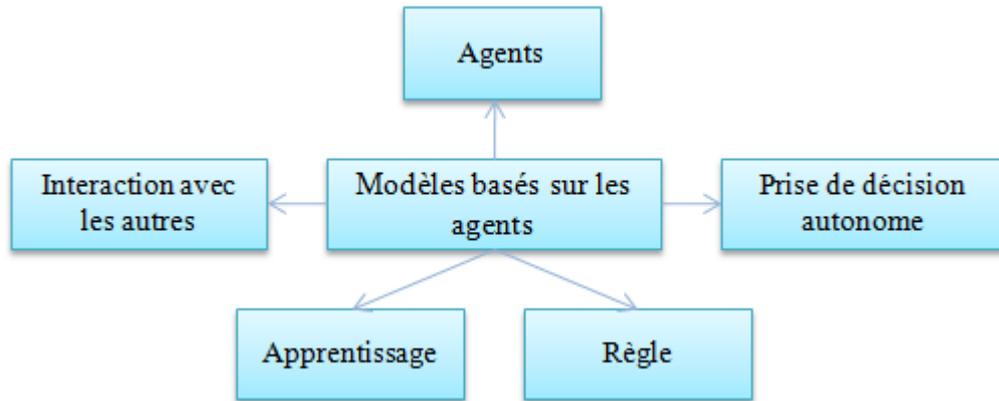


FIGURE 4 – Structure des modèles basés sur les agents.

de COS). Une approche statistique (par exemple une régression logistique) est généralement utilisée pour quantifier les relations entre les emplacements de COS et un ensemble de facteurs indépendants. Pour exécuter le modèle, il est nécessaire d’avoir des données spatiales pour un temps au minimum ; Cependant, pour réaliser la calibration et la validation du modèle, il est nécessaire d’avoir des données correspondant à plusieurs dates.

3.2.3 Modèles basés sur les agents

Différentes études ont montré que les modèles basés sur les agents sont utiles pour explorer les processus de COS [Parker *et al.* 2003, Verburg 2006]. Ces modèles ont été utilisés plus particulièrement pour simuler les COS à l’échelle locale en raison de sa nature complexe [Acosta-Michlik et Espaldon 2008]. Les applications de ces modèles ont été plus limitées dans la planification et la prise de décision. Les calculs des modèles basés sur les agents sont extensifs. Ils intègrent généralement des modèles empiriques pour le paramétrage [Valbuena *et al.* 2008] ou ils sont couplés avec d’autres modèles de COS affectivement pour rendre la simulation plus rapide (par exemple, les modèles des automates cellulaires et les modèles de chaîne de Markov) [Parker *et al.* 2008].

Les modèles basés sur les agents simulent les COS comme un résultat de l’interaction entre les agents individuels (par exemple, les décisions concernant la politique, la planification et la gestion) [Parker *et al.* 2003, Parker *et al.* 2008]. Dans un modèle de COS (Figure 4), les agents sont des unités de prise de décision dans les modèles basés sur les agents et ils agissent de façon autonome dans l’environnement.

Les agents doivent partager l’espace spatial pour interagir les uns avec les autres et répondre à l’environnement. Les règles sont le résultat de l’interaction entre les agents et déterminent le comportement futur des actions. Le comportement des agents peut changer de règles simples à des décisions complexes. Les agents ont généralement des caractéristiques uniques (par exemple l’ADN dans le corps humain ou l’empreinte digitale) qui les permettent d’être identifiés par d’autres (Figure 1-4). Les agents changent leurs comportements pour satisfaire leur objectif souhaité en comparant le résultat de leur comportement par rapport à ses objectifs.

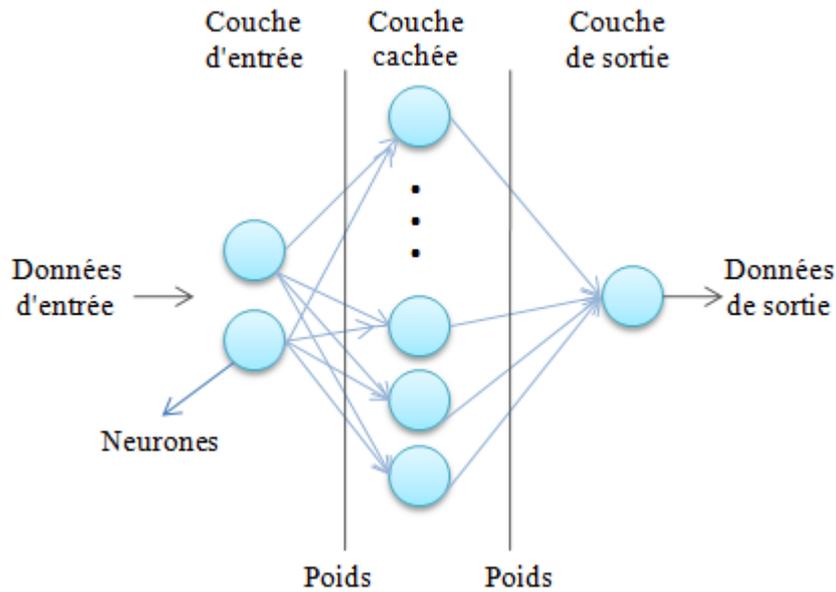


FIGURE 5 – Structure de modèle de réseau de neurones artificiels.

3.2.4 Modèles basés sur l'apprentissage automatique

L'apprentissage automatique est un noyau de l'intelligence artificielle. Il permet d'étudier des algorithmes informatiques pour l'apprentissage pour accomplir une tâche. L'accent mis sur l'apprentissage automatique est de concevoir des algorithmes d'apprentissage qui font automatiquement l'apprentissage sans l'aide humaine. Les chercheurs dans le domaine de l'apprentissage automatique sont familiers avec la plupart des modèles statistiques. Les réseaux de neurones artificiels, les machines à vecteurs de support, les algorithmes génétiques et les arbres de décisions appartiennent à ce type de modèles.

Réseaux de neurones artificiels

De nombreux chercheurs ont constaté que les réseaux de neurones artificiels peuvent résoudre les problèmes de classification avec précision. Les réseaux de neurones artificiels sont constitués de neurones (par exemple la structure de cerveaux humains) dans des couches qui simulent les COS (Figure 5).

Machines à Vecteurs de Support

Comme les réseaux de neurones artificiels, les machines à vecteurs de support peuvent être utilisées comme un algorithme d'apprentissage automatique pour détecter des tendances non linéaires dans les données. Les machines à vecteurs de support projettent des données à une dimension supérieure (appelée espace de Hilbert, Figure 6) pour contraindre un hyperplan de classification optimal. Dans l'espace de Hilbert, les machines à vecteurs de support peuvent classer linéairement les tendances de COS en utilisant la minimisation des risques structurels et la maximisation des marges [Vapnik 1998]. La fonction de l'hyperplan optimal de séparation est développée à l'aide de la fonction du noyau et des vecteurs de support. Contrairement aux réseaux de neurones artificiels, les machines à vecteurs de support fournissent un hyperplan optimal unique et global.

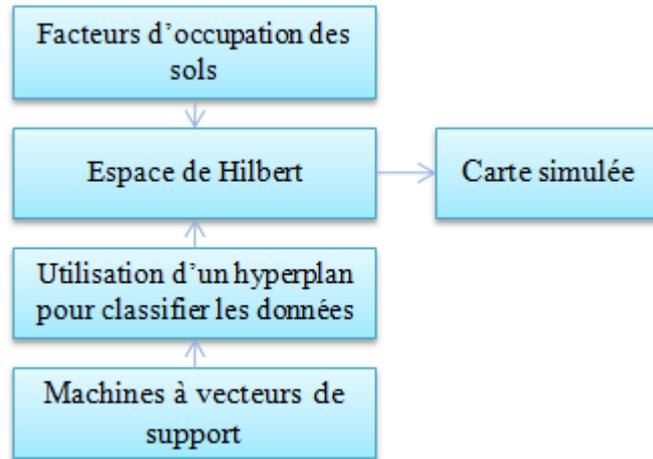


FIGURE 6 – Structure de modèle des machines à vecteurs de support.

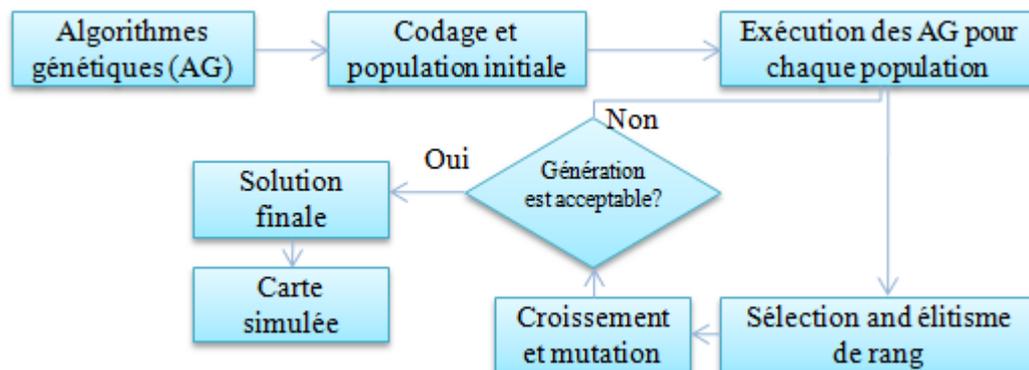


FIGURE 7 – Structure de modèle des algorithmes génétiques.

Algorithmes génétiques

Les algorithmes génétiques, issus de la théorie de la sélection naturelle en biologie, effectuent une recherche dans l'espace pour trouver la solution optimale [Holland 1975, Goldberg 1989]. Les algorithmes génétiques peuvent estimer le minimum ou le maximum global en utilisant des fonctions de fitness (Figure 7).

L'approche de calibration des algorithmes génétiques comprend les étapes suivantes [Shan *et al.* 2008] : premièrement, l'algorithme génétique génère aléatoirement la population initiale de solutions et est codé en chaînes binaires (encodage et population initiale). Chaque chaîne en style binaire codé correspond à une solution. Deuxièmement, l'algorithme génétique s'exécute ensuite pour chaque chaîne de la population avec la carte de COS de référence. Une carte de référence est utilisée pour classer les solutions initiales sur la base d'une fonction de fitness. Cette étape est responsable de sélectionner les chaînes pour la prochaine génération. Pour la sélection de rang, toutes les chaînes sont ordonnées en fonction de leur fonction de fitness dans l'ordre croissant et enfin, la dernière étape (croisement et mutation) est de produire la prochaine génération de solutions. L'opération de croisement produit la prochaine génération dans chaque exécution qui sont censés d'avoir de même ou de meilleure qualité que leurs parents. La mutation empêche la solution de se verrouiller sur les minima locaux. L'apprentissage

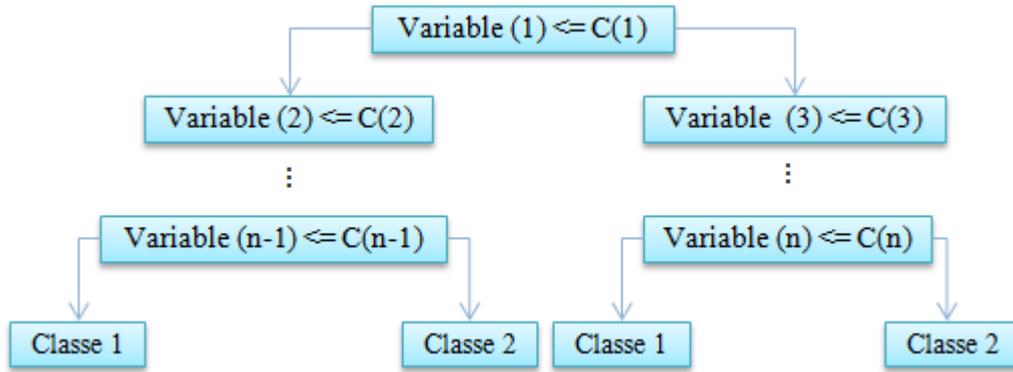


FIGURE 8 – Structure de modèle CART.

de l'algorithme génétique se poursuit pendant un certain nombre de générations et la génération avec les fonctions de fitness minimales choisit comme solution finale.

Arbres de décisions

Les arbres de décision font partie des méthodes les plus populaires en apprentissage automatique. En effet, ils constituent une technique puissante de fouille de données, souvent utilisée pour les tâches de classification et de prédiction. Le processus de construction de l'arbre de décision est présenté dans [Quinlan 1993]. La structure de base de l'arbre de décision est constituée d'un nud racine, d'un nombre de nuds internes et enfin d'un ensemble de nuds terminaux. Les données sont divisées de manière récursive dans l'arbre de décision en fonction du cadre de classification défini. A chaque nud, une règle de décision est requise et ceci peut être implémenté en utilisant un test de partitionnement.

CART (*Classification and Regression Tree*) est l'un des modèles courants d'arbres de décision qui classent les données hiérarchiquement (Figure 8). CART produit un modèle avec une structure utilisant une série de règles si-alors-sinon [Breiman *et al.* 1984]. CART est responsable d'identifier des partitionnements à chaque nud afin de mieux diviser les données. Les nuds de l'arbre sont représentatifs de chaque variable dans CART. L'emplacement des nuds au niveau hiérarchique montre la contribution de chaque variable. Les nuds en haut (respectivement en bas) ont une contribution plus élevée (respectivement plus basse) pour la modélisation. L'indice de Gini utilise une fonction d'impureté pour trouver la meilleure répartition [Breiman *et al.* 1984] entre toutes les valeurs uniques parmi les prédicteurs pour fragmenter les données.

En général, les avantages de l'arbre de décision sont : la construction rapide de l'arbre et l'interprétation facile des règles de décision générées à partir de ce dernier du type si-alors. C'est pourquoi, les problèmes de prédiction dans plusieurs domaines [Levashenko et Zaitseva 2012, Ebadi *et al.* 2012], notamment en télédétection, sont souvent résolus par cette technique [Boulila *et al.* 2011, Li 2011]. Cependant, un de ses inconvénients, est l'utilisation fréquente de variables moins pertinentes pour l'étape de construction de l'arbre (sur-apprentissage). Ce problème est résolu par l'étape d'élagage qui consiste à supprimer les sous-arbres superflus ou trop liés aux données, et ceci afin d'améliorer l'aspect prédictif de l'arbre d'une part, et réduire sa complexité d'autre part.

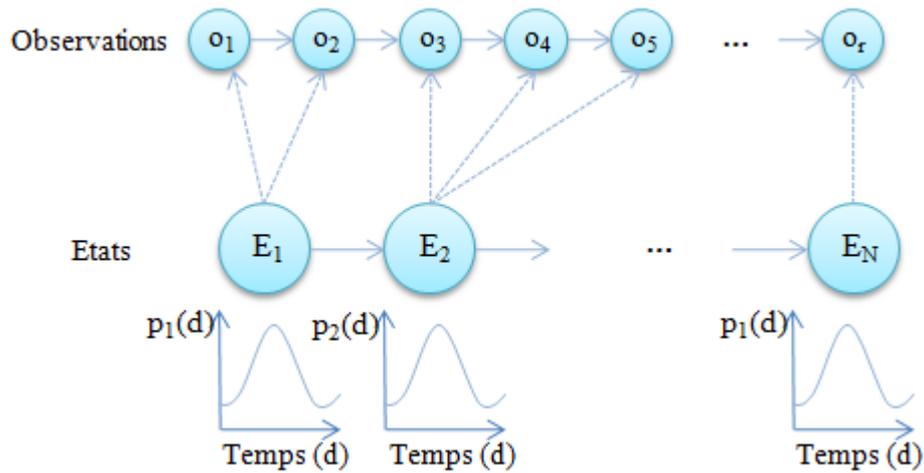


FIGURE 9 – Structure des modèles de Markov caché.

3.2.5 Modèles basés sur les chaînes de Markov

Les chaînes de Markov est un modèle de processus stochastique qui décrit la probabilité de changement d'un état à un autre, c'est-à-dire d'un type d'occupation du sol à un autre, en utilisant une matrice de probabilité de transition [Logsdon *et al.* 1996]. La probabilité de transition serait la probabilité qu'un type d'occupation du sol (pixels) à l'instant t_0 change à un autre type d'occupation du sol au temps t_1 . Par conséquent, les COS entre les dates ont été utilisés pour développer une matrice de probabilité de transition et de prédire ensuite les occupations des sols pour un temps futur. Cette matrice est le résultat du croisement entre les images en établissant une erreur proportionnelle. Les cartes de probabilité de transition sont élaborées à partir des trois occupations des sols qui sont générées par le processus de classification. Elles sont utilisées pour calculer les projections des changements possibles entre les périodes spécifiées. Le modèle de Markov fournit une méthodologie simple et par lequel un système dynamique pourrait être disséqué et examiné [Muller et Middleton 1994, Logsdon *et al.* 1996, Dongjie *et al.* 2008, Dadhich et Hanaoka 2010, Zhang *et al.* 2011]. De nombreuses recherches ont témoigné la précision du modèle de Markov [Zhang *et al.* 2011].

Les modèles de Markov caché dérivent d'autres modèles comme le modèle de Markov hiérarchique, le modèle de Markov gaussien ou encore appelé filtre, le modèle de Markov couplé, le modèle de Markov triplet et les modèles de Markov à temps continu. Les modèles de Markov caché sont généralement utilisés pour résoudre trois types de problèmes (Figure 9) : (1) le calcul de la probabilité d'une séquence donnée, (2) le calcul du chemin le plus probable, c'est-à-dire la suite d'états optimale, connaissant une suite d'observations, et (3) la mise au point des paramètres du modèle. En général, les observations sont les valeurs de caractéristiques des images satellites et les états correspondent implicitement aux phases phénologiques des types d'occupation des sols.

Les travaux à base des modèles de Markov caché se prêtent d'avantage à l'utilisation des outils faisant appel aux mathématiques du calcul des probabilités, à la théorie des graphes (graphes orientés) et à l'algèbre linéaire (matrice de transition). Par conséquent, il permet une estimation rapide et une prédiction du futur. Or, l'exploitation des modèles de Markov caché est souvent limitée au traitement d'images telles que la seg-

mentation. Néanmoins des travaux récents sont menés dans la direction de la prédiction des événements spatiotemporels.

3.2.6 Modèles hybrides

Les modèles hybrides qui combinent plusieurs modèles par exemple les modèles cellulaires/les modèles statistiques (ex. DINAMICA), les modèles cellulaires/les modèles de Markov (ex. CA-MARKOV), les modèles de Markov/les modèles statistiques (ex. LCM), etc. Les modèles originaux ont été considérablement modifiés afin d'améliorer les résultats de simulation de COS. L'utilisation de plusieurs variables avec des gammes plus larges dans les règles de transition rend la modélisation de l'occupation des sols plus sophistiqué.

Le modèle DINAMICA

DINAMICA est un modèle d'automates cellulaires qui présente des fonctions transitoires à base de proximité multi-échelle. Il intègre l'approche spatiale de rétroaction à un moteur de simulation multi-étapes stochastique, et applique la régression logistique pour calculer les probabilités de transition dynamiques spatiales. Selon White *et al.* [White *et al.* 2000], le modèle d'automate cellulaire classique se compose de : (1) un espace euclidien divisé en un ensemble de cellules identiques, (2) un voisinage cellulaire d'une taille et d'une forme définies, (3) un ensemble d'états de cellule discrète, (4) un ensemble de règles de transition qui déterminent l'état d'une cellule en fonction des états des cellules dans un voisinage, et (5) des étapes de temps discrètes avec tous les états de cellule mis à jour simultanément. Généralement conçu pour simuler les COS, DINAMICA incarne les concepts ci-dessus en utilisant une séquence d'algorithmes spécialement développés, qui sont présentés dans (Figure 10).

Pour exécuter les simulations, le modèle fonctionne en phases, chacune ayant ses propres paramètres : (1) le nombre d'incrément de temps, (2) la matrice de transition, étant les taux fixés à l'intérieur de la phase, (3) une valeur de saturation éventuelle pour chaque type d'occupation des sols, (4) Le temps de séjour minimum pour chaque type de transition avant qu'une cellule change son état, (5) les coefficients du modèle logistique ont appliqué pour calculer chaque probabilité de transition spatiale, et (6) le pourcentage de transitions exécutées par chaque fonction de transition ainsi que les paramètres signifient la moyenne et la variance de la taille du patch de chaque type d'occupation des sols. Chaque phase a des paramètres fixes. Par conséquent, le modèle peut être exécuté en utilisant plusieurs phases, chacune consistant en plusieurs étapes de temps.

Un sous-ensemble de variables utilisées dans la simulation est statique. Ces variables doivent être calculées une seule fois avant le début de la simulation. Puis elles sont stockées dans un cube de données raster. DINAMICA peut également tenir compte des hypothèses concernant les taux de transition distincts, en fonction de l'arrangement et la partition spatiale, en utilisant une carte de partition spatiale comme variable statique. Un autre sous-ensemble de variables est dynamique, ce qui signifie qu'elles doivent être calculées avant chaque itération du modèle. Plus d'informations sur le développement et les différentes hypothèses proposées de modèle DINAMICA peuvent être trouvées dans [Soares-Filho *et al.* 2002].

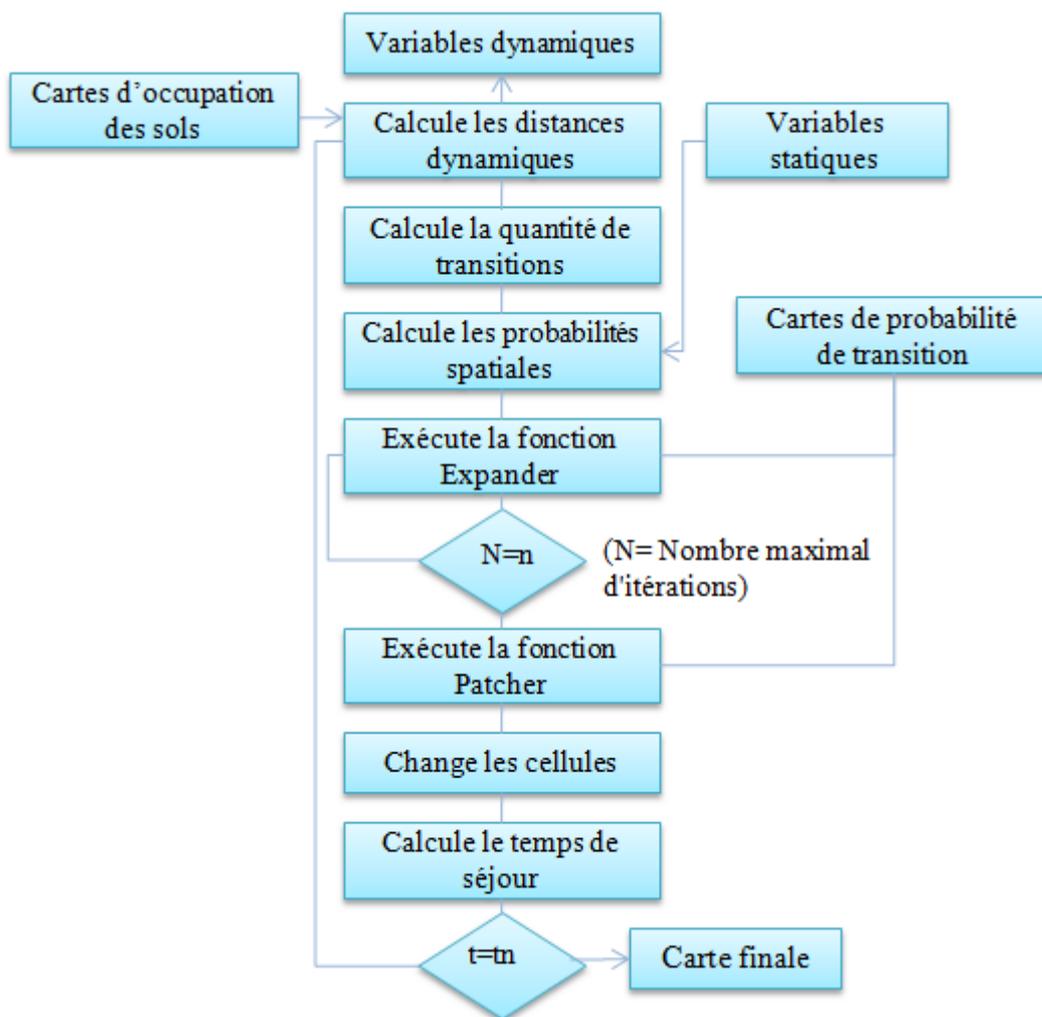


FIGURE 10 – Structure de modèle DINAMICA.

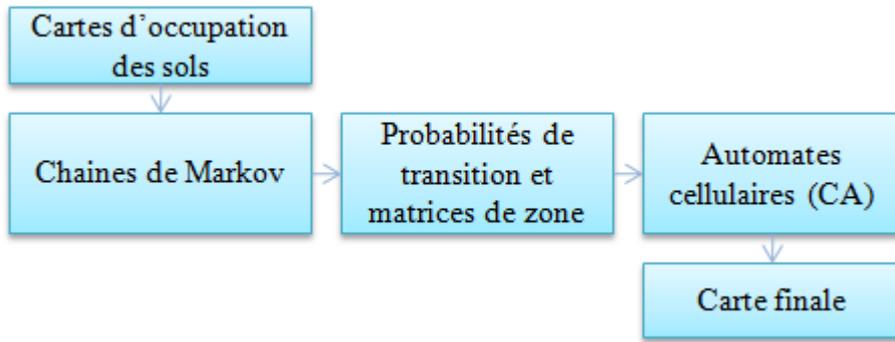


FIGURE 11 – Structure de modèle CA-MARKOV.

DINAMICA a été appliqué à diverses études, telles que la modélisation de la croissance urbaine [Almeida *et al.* 2003, Thapa et Murayama 2011], la déforestation tropicale [Soares-Filho *et al.* 2002], les régimes d'incendie [Silvestrini *et al.* 2011, Soares-Filho *et al.* 2012], l'évaluation des coûts et des avantages de la réduction des émissions de carbone dues à la déforestation et à la dégradation des forêts [Merry *et al.* 2009, Nepstad *et al.* 2009] et l'analyse de ROC [Mas *et al.* 2013].

Le modèle CA-MARKOV

Le modèle hybride de Markov et d'Automates Cellulaires (CA-MARKOV) utilise des matrices de chaîne de Markov pour déterminer la quantité de changements. Il utilise, aussi, des cartes d'aptitude et des automates cellulaires pour répartir spatialement ces changements (Figure 11). Le processus de Markov contrôle la dynamique temporelle entre les types d'occupation par l'utilisation de probabilités de transition. Pour la dynamique spatiale, elle est contrôlée par des règles locales à travers un mécanisme des automates cellulaires (CA) en considérant soit la configuration de voisinage, soit les probabilités de transition.

La combinaison de chaînes de Markov et d'Automates Cellulaires permet de simuler l'évolution de l'occupation des sols représentée par les pixels. Chaque pixel peut prendre une valeur d'un ensemble fini d'états. Tous les pixels sont affectés par une fonction de transition qui prend comme arguments les valeurs mesurées et les valeurs des pixels voisins en fonction du temps. Pour une zone d'étude donnée, la fonction de transition a été déterminée en fonction de la différence entre deux dates t_1 et t_2 . CA-MARKOV a ensuite utilisé cette fonction de transition pour prédire l'occupation du sol pour des dates futures. En d'autres termes, la matrice de probabilité de transition, créée à partir des changements observés entre t_1 et t_2 , les cartes de probabilité de transition à l'instant t_2 et chaque scénario ont été utilisés pour produire des cartes d'occupation des sols pour une date future. Dans un processus itératif, CA-MARKOV utilise les cartes de probabilité de transition de chaque type d'occupation du sol pour établir l'aptitude inhérente de chaque pixel pour passer d'un type d'occupation du sol à un autre. Une application détaillée de ce modèle est illustrée dans [Paegelow et Camacho Olmedo 2005]. Des exemples d'applications peuvent être trouvés dans [Houet et Hubert-Moy 2006, Subedi *et al.* 2013, Nouri *et al.* 2014, Hyandye et Martz 2017].

Le modèle LCM

Le LCM (*Land Change Modeler*) est une suite d'outils avec lesquels l'analyse et la

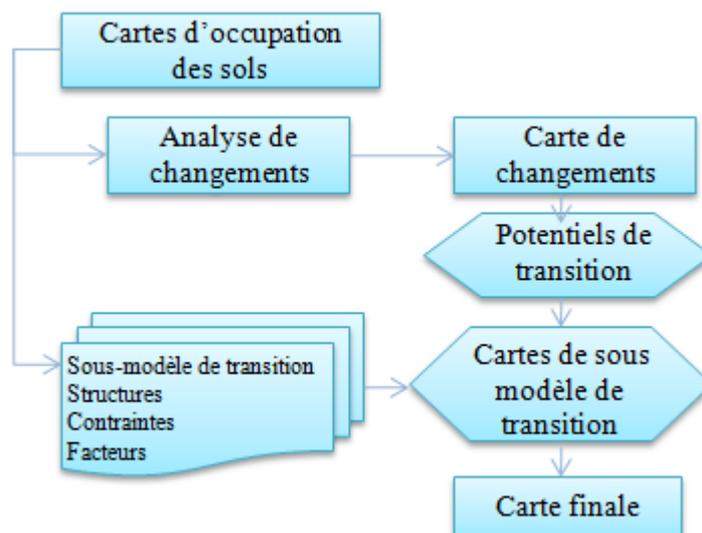


FIGURE 12 – Structure de modèle LCM.

modélisation de COS peuvent être combinés avec des évaluations de la biodiversité et des émissions de gaz à effet de serre. Le module de modélisation des changements est basé sur des matrices de chaînes de Markov et des cartes de susceptibilité de transition obtenues par la régression logistique ou par des machines d'apprentissage [Eastman 2009, Johnson 2009] (Figure 12).

Le LCM offre la possibilité d'estimer la quantité de changements futurs pour chaque transition. A partir des données d'entrée, les probabilités de transition et la quantité de changement sont estimées par extrapolation des tendances passées à l'aide d'un module probabiliste composé de chaînes de Markov. Le processus markovien suppose que l'état d'un système au temps t_2 peut être prédit en se basant sur son évolution entre ses états à t_0 et t_1 . Cette quantité de changement peut également être définie manuellement ou provenir d'un modèle externe. Au final, le couplage des chaînes de Markov et des cartes de probabilités de transition au sein d'un module d'allocation spatiale des changements tenant compte des différentes transitions permet de bénéficier des avantages comparatifs des deux approches.

Le LCM a été appliqué pour identifier les tendances de COS [Václavík et Rogan 2009], la déforestation tropicale [Koi et Murayama 2010], la croissance urbaine [Aguejdad et Houet 2008, Mishra *et al.* 2014], l'érosion sous différents scénarios de conservation [Gaspari *et al.* 2009] et la modélisation de l'habitat [Gontier *et al.* 2009].

3.3 Composants des modèles de COS

La simulation numérique est un outil permettant de simuler des phénomènes réels. La simulation sert à étudier le fonctionnement et les propriétés du système modélisé ainsi qu'à prévoir son évolution. Tous les modèles de simulation sont constitués de différents composants du modèle [Kruger et Lakes 2015, Verstegen *et al.* 2016]. Les entrées du modèle constituent la première composante. Ils s'agissent dans le modèle de COS, des demandes pour tous les types d'occupation des sols sur une période de temps donnée et un certain nombre d'attributs spatiaux qui influent sur les emplacements des COS.

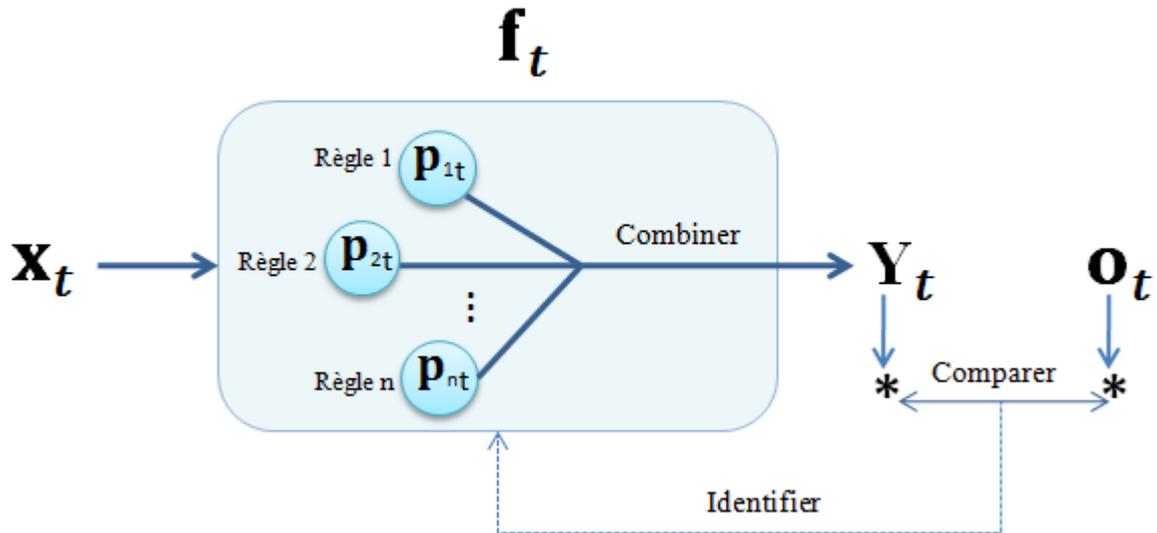


FIGURE 13 – Modèle conceptuel d'un modèle de simulation [Verstegen *et al.* 2016].

L'ensemble des entrées à l'instant t est noté X_t (Figure 13). Le deuxième composant est l'état du système (la sortie du modèle) à l'instant t , noté Y_t (Figure 13), par exemple le COS. L'état système à l'instant $t - 1$ est également l'une des entrées du calcul de l'état du système à l'instant t car, dans presque tous les systèmes d'occupation des sols, l'état actuel du système dépend de l'état du système passé. Le noyau d'un modèle de simulation est l'ensemble des règles de transition qui représentent les processus qui conduisent au changement dans l'état du système dans le temps. Dans le cas d'un modèle de COS, une règle de transition est une fonction qui calcule l'aptitude de chaque emplacement pour un type particulier d'occupation des sols, en respectant l'attribut spatial qui influe sur l'emplacement de ce type d'occupation des sols. Le troisième composant englobant la technique de sélection et de combinaison des règles de transition, est appelée structure de modèle f_t (Figure 13). Les règles de transition peuvent contenir des paramètres p_t définissant les caractéristiques du processus représenté par la règle de transition. Un modèle de COS peut avoir, par exemple, un paramètre qui spécifie si la relation entre les valeurs de l'attribut et les valeurs d'aptitude est linéaire ou exponentielle. Les paramètres sont le quatrième composant du modèle.

Si nous voulons utiliser un modèle d'occupation des sols pour prédire les changements futurs d'occupation des sols, il est essentiel de trouver la structure du modèle et les valeurs des paramètres qui donnent une représentation optimale du système terrestre étudié. L'identification de la structure du modèle et des valeurs des paramètres peut être réalisée en comparant les changements modélisés et les changements réels O_t pour la même étape à l'instant t (Figure 13). Ensuite, il faut sélectionner les valeurs des paramètres et la structure du modèle qui minimisent la différence entre les changements modélisés et réels (observés). Les observations ou les changements réels peuvent être des cartes d'occupation des sols, mais aussi des mesures agrégées ou des métriques spatiales de la configuration (une mesure des modèles spatiaux) ou de la composition (une mesure des proportions de différentes classes) [Csillag *et Boots* 2005].

4 Conclusion

Dans ce chapitre, nous avons présenté une brève introduction sur des terminologies de base dans l'étude de COS, ce qui met en évidence l'importance de données de télédétection. Les approches fondamentales de modélisation du COS ont été introduites par la suite. Les composants de ces modèles de simulation ont été également abordés.

Aussi, il est essentiel d'évaluer la solidité des prédictions de COS. En effet, des décisions basées sur des prédictions erronées peuvent être coûteuses, d'un point de vue environnemental ou économique. En d'autres termes, il est nécessaire de reconnaître l'incertitude dans ces prédictions. Connaître la source de l'incertitude offre également des moyens de la réduire. Dans le chapitre suivant, nous allons présenter les sources, les types d'incertitude associés aux modèles de COS et les méthodes existantes pour les modéliser.

Chapitre II

Etude des incertitudes et de leur propagation dans les modèles de COS

Sommaire

1	Introduction	25
2	Types d'incertitudes	25
2.1	Incertitude aléatoire	26
2.2	Incertitude épistémique	26
2.3	Aléatoire ou épistémique?	27
3	Classification des sources d'incertitudes	28
3.1	Incertitudes sur les paramètres	28
3.2	Incertitudes sur la structure des modèles de COS	31
4	Propagation des incertitudes dans les modèles de COS	32
4.1	Les méthodes probabilistes	32
4.1.1	Généralités	33
4.1.2	Méthodes probabilistes de propagation des incertitudes	35
4.2	Les méthodes non probabilistes	38
4.2.1	La théorie des intervalles	38
4.2.2	La théorie des ensembles flous	38
4.2.3	La théorie des possibilités	39
4.2.4	La théorie des probabilités imprécises	40
4.2.5	La théorie des fonctions de croyance	40
4.3	Synthèse et analyses comparatives	43
5	Conclusion	45

1 Introduction

La construction d'un modèle numérique de COS n'est pas simple. Les processus dynamiques dans le système terrestre sont complexes et difficiles à comprendre. Un modèle numérique ne pourra jamais restituer la réponse réelle de la structure qu'il représente, aussi avancé et précis qu'il soit. Pour cela, les ingénieurs doivent prendre des décisions ayant des conséquences sur la fiabilité et la performance de ces modèles. Généralement, ceci consiste à identifier les sources d'incertitudes, distinguer leurs types et les modéliser afin d'améliorer la prise de décision concernant le COS.

Dans ce deuxième chapitre, nous présentons, dans la première partie, une étude bibliographique sur les types et les sources d'incertitudes rencontrés lors de l'analyse d'un modèle de COS. Cette étude permettra de montrer qu'il est nécessaire d'incorporer aux modèles numériques une stratégie de propagation et réduction des incertitudes. Ainsi, ces modèles deviendront plus robustes aux incertitudes et fourniront une meilleure prise de décision. Dans la deuxième partie de ce chapitre, un état de l'art, non exhaustif, des principales méthodes de propagation des incertitudes dans les modèles de COS seront dressé. Des analyses comparatives de plusieurs méthodes permettent d'évaluer leur capacité de prédiction en termes de précision.

2 Types d'incertitudes

L'incertitude est un concept très vaste et qui a plusieurs définitions. Selon Haimès [Haimès 2009], l'incertitude est l'incapacité à connaître le vrai état d'un système. Ayyub et Klir dans [Ayyub et Klir 2006], quant à eux, l'interprètent comme l'incomplétude de la connaissance en général due à des difficultés intrinsèques à acquérir de la connaissance. Les deux visions sont compatibles puisque l'incapacité à déterminer un état mène nécessairement à une connaissance incomplète. De nombreux auteurs ont abordé la classification et la terminologie de l'incertitude [Crosetto *et al.* 2001, Oijen et Thomson 2010, Boulila *et al.* 2014, Cockx *et al.* 2014]. Bien qu'ils ont défini l'incertitude dans une variété de façons, la plupart distinguent deux types d'incertitudes selon la quantité d'information disponible sur les fluctuations de la grandeur d'intérêt : aléatoire lorsqu'elle est due à une variabilité naturelle, et épistémique si elle est due à un manque de connaissance [Hoffman et Hammonds 1994, Ferson et Ginzburg 1996, Oberkampf *et al.* 2002, Helton *et al.* 2004, Der Kiureghian et Ditlevsen 2009]. Dans [Oberkampf *et al.* 2002], nous distinguons de plus l'erreur, une lacune identifiable dans la phase de modélisation ou d'expérimentation qui n'est pas due à un manque de connaissance.

Il existe néanmoins d'autres manières de classer les incertitudes. Camerer et Weber [Camerer et Weber 1992] emploient le terme ambiguïté pour parler du manque d'information. Ce terme est défini comme l'incertitude par rapport à la probabilité provenant du manque d'informations importantes qui pourraient être connues. Dubois [Dubois 2010] prétend que toutes les incertitudes sont de nature épistémique. Mais il reconnaît que l'incertitude peut avoir plusieurs origines, à savoir être due à la variabilité naturelle d'un phénomène, à des informations incomplètes ou contradictoires. Blockley [Blockley 2013] propose une définition de l'incertitude dans un espace orthogonal tridimensionnel : l'axe de l'aléatoire (*randomness*, en anglais) qui signifie un manque d'un motif ou

schéma spécifique, l'axe de l'incomplétude (*incompleteness*, en anglais) qui se réfère à ce qu'on ne connaît pas, et l'axe du flou (*fuzziness*, en anglais) qui indique une imprécision ou ambiguïté.

2.1 Incertitude aléatoire

Le mot aléatoire dérive du terme latin *alea* qui signifie " lancer les dés ". L'incertitude est considérée aléatoire si elle est irréductible provenant d'une variabilité naturelle. Une incertitude est qualifiée comme aléatoire si on suppose qu'elle est intrinsèque à un phénomène aléatoire, et apparaît généralement entre les membres d'une même population présentant des variations temporelles ou spatiales. Par exemple, la variation des valeurs de NDVI (*Normalized Difference Vegetation Index*) d'une année à une autre. Dans la littérature, nous trouvons d'autres références à ce type d'incertitude, comme l'incertitude irréductible, la variabilité, l'incertitude intrinsèque, l'incertitude objective et l'incertitude stochastique.

Pour modèles de COS, nous considérons que les incertitudes sont aléatoires si elles sont liées par exemple [Peters 1995, Hulley *et al.* 2012]

- aux caractéristiques spectrales (les valeurs moyennes des bandes spectrales pour chaque objet d'images, etc.) ;
- aux caractéristiques de végétation ;
- aux caractéristiques climatiques (température, humidité, etc) ;

L'utilisation d'un modèle probabiliste s'avère appropriée pour représenter et propager les incertitudes aléatoires. D'un point de vue mathématique, ce modèle est souvent une densité de probabilité (notée PDF pour *Probability Density Function*) ou une fonction de répartition (notée CDF pour *Cumulative Density Function*).

2.2 Incertitude épistémique

Le mot épistémique dérive du terme grec *επιστημη* (*episteme*) qui signifie connaissance. Dans la littérature, nous trouvons d'autres références à ce type d'incertitude, comme l'incertitude réductible (elle est réductible par l'acquisition de nouvelles connaissances ou des avis de spécialistes ou d'améliorer la précision de mesure), l'état de connaissance, l'incertitude subjective et l'incertitude cognitive. Ainsi, une incertitude est considérée épistémique si elle provient d'un manque d'information, de méconnaissance, d'imprécision, d'ignorance ou des erreurs humaines. Elle apparaît généralement lors de la modélisation de systèmes complexes, lorsque des phénomènes physiques sont négligés ou des représentations sont simplifiées. L'incertitude épistémique se manifeste également quand nous disposons de peu ou aucune de données expérimentales pour un paramètre physique, et lorsque nous avons une compréhension limitée des phénomènes physiques complexes étudiés. Pour modèles de COS, nous considérons que les incertitudes sont épistémiques si elles sont liées, par exemple, aux caractéristiques de forme ou de texture [Peters 1995]. A titre d'exemple, lors de l'identification des objets issus des images satellitaires, les experts ont indiqué que la zone de végétation est entre 30.7 et 31 *ha*. Dans ce cas, l'incertitude sur la surface peut être considérée comme épistémique.

Le choix du cadre probabiliste est justifié pour les incertitudes aléatoires pour lesquelles nous disposons de données suffisantes permettant l'ajustement des fonctions de densité de probabilité. Le choix peut être remis en cause pour les incertitudes épistémiques dont la caractérisation s'appuie uniquement sur l'avis d'expert. Ainsi, il est commun de représenter les incertitudes épistémiques dans des intervalles de confiance [Moore 1966]. Il est difficile de quantifier ou modéliser ces incertitudes, mais nous pouvons les réduire en rassemblant davantage d'informations (résultats expérimentaux, opinion d'expert, etc.) ou en améliorant la définition des modèles, ce qui nécessite du temps et des ressources. Habituellement, l'incertitude épistémique est représentée par un intervalle sans une PDF associée, ou bien par une PDF décrivant le niveau de croyance et plausibilité de l'analyste [Agarwal *et al.* 2004, Park et Grandhi 2012].

2.3 Aléatoire ou épistémique ?

La différence entre l'incertitude épistémique et aléatoire n'est pas toujours facile à faire lors de l'analyse d'un système terrestre. La distinction entre les incertitudes aléatoires et épistémiques est importante car les décisions prises en conséquence sont différentes. Dans le cas de l'incertitude aléatoire (due à une variabilité), une action concrète doit être prise pour éviter les effets dangereux potentiels d'une telle variabilité. Alors que, pour le cas de l'incertitude épistémique (due à des informations incomplètes), la meilleure décision serait probablement de chercher à la réduire en recueillant des informations complémentaires. Dans [Morgan et Henrion 1992], Morgan et Henrion ont affirmé qu'il est important de distinguer les types d'incertitude, dont la variété peut engendrer de la confusion. Dans [Thompson 2002], Thompson a déclaré que la distinction entre l'incertitude aléatoire et l'incertitude épistémique est très importante dans la gestion des risques. En effet, la connaissance de l'incertitude aléatoire conduit à une meilleure compréhension de la répartition des risques, et connaître l'incertitude épistémique aide les décideurs à chercher des informations appropriées et à améliorer les systèmes.

De plus, il est essentiel de pouvoir distinguer les types d'incertitudes car les représentations mathématiques et les méthodes de propagation sont différentes [Ferson et Ginzburg 1996]. La prise en compte de l'incertitude aléatoire et l'incertitude épistémique avec une seule distribution de probabilités peut entraîner une perte d'information en raison du caractère différent des incertitudes [Vose 2008]. Par conséquent, une identification et une hiérarchisation plus détaillées des sources d'incertitude sont utiles pour choisir les méthodologies les plus pertinentes pour propager les différents types d'incertitude [De Rocquigny *et al.* 2008].

Néanmoins, même si certains auteurs prétendent qu'une distinction franche et nette entre ces deux types est possible, il est souvent très difficile de déterminer, lors de la phase de modélisation de COS, si une incertitude particulière devrait être considérée aléatoire ou épistémique. Le rôle du constructeur du modèle de COS est d'effectuer cette distinction selon les circonstances et les données disponibles. Ses choix personnels de modélisation sont conditionnés par la connaissance scientifique et les données disponibles, mais surtout par le besoin pratique d'avoir un modèle fin pour améliorer la prise de décision à partir des calculs prédictifs [Kiureghian et Ditlevsen 2009].

3 Classification des sources d'incertitudes

Au cours des années passées, un intérêt considérable s'est manifesté sur la recherche des sources d'incertitudes, surtout celles introduites dans les modèles mathématiques et les outils de simulation numérique. Les ingénieurs, les scientifiques et les décideurs ont commencé à distinguer et caractériser les différentes sources responsables de l'écart observé entre les prédictions numériques et le comportement de la structure réelle. Si la modélisation est comprise comme un processus comportant différentes étapes de modélisation, un certain degré d'incertitude est associé à chaque étape de modélisation [Walker *et al.* 2003]. La quantité d'incertitude dans l'une des étapes dépend de l'incertitude des autres étapes [Krayner von Krauss *et al.* 2006]. En essayant de réduire l'incertitude dans le modèle, il est donc utile de couvrir les dépendances entre ces étapes. Dans le cas contraire, la réduction d'une seule source pourrait conduire à une augmentation de l'incertitude dans une seconde source. L'incertitude qui se pose dans le processus de modélisation se propage et regroupe l'incertitude totale des résultats du modèle [Golde-wijk et Verburg 2013]. L'importance et la complexité de l'interaction entre les différentes sources d'incertitude sont connues dans la communauté des changements de terres, les approches qui prennent systématiquement en compte ces problèmes sont rares [Golde-wijk et Verburg 2013, Tayyebi *et al.* 2014].

Généralement, les incertitudes dans les modèles de COS proviennent principalement de deux grandes sources : les incertitudes de paramètres du modèle et les incertitudes de la structure du modèle dans les phases de conception et modélisation. Il est très important de prendre en compte chacune de ces sources de manière appropriée selon la quantité d'information disponible afin d'apporter les améliorations nécessaires à la prise de décision concernant les COS. Une description complète de ces sources est effectuée dans [Oberkampf et Roy 2010].

3.1 Incertitudes sur les paramètres

Ce sont des incertitudes liées aux paramètres d'entrée du modèle de COS, tels que ceux relatifs aux descripteurs des objets d'images satellitaires. Ces paramètres sont estimés à partir de données issues du retour d'expérience. En théorie, il faut disposer d'un nombre infini de données observées pour que l'estimation puisse être considérée comme dénuée d'erreur. Vu que ceci est impossible en pratique, il en résulte la présence d'incertitudes sur ces paramètres. Ces incertitudes se propagent dans le modèle et engendrent des incertitudes dans le résultat final. Dans les modèles de COS, les paramètres d'entrée d'un modèle sont généralement constitués de :

- paramètres décrivant la composition spectrale d'une région donnée (paramètres spectraux), comme les valeurs moyennes des bandes spectrales, l'indice de luminosité, etc. ;
- paramètres définissant l'arrangement structural des surfaces et leurs relations avec leur environnement direct (paramètres de texture), comme le contraste, l'entropie, l'énergie, etc. ;
- paramètres décrivant la géométrie d'un objet satellite (paramètres de forme), comme l'indice de forme, la densité, la surface, etc.
- paramètres isolant les zones de végétation (paramètres de végétation), comme

Paramètres d'entrée	Sources d'incertitude	Références
Paramètres spectraux	La réflectance spectrale de la surface, l'étalonnage du capteur, le bruit du capteur, l'effet de pixels mixtes, les nuages de sous-pixels, effet d'un changement dans l'emplacement de canal, l'enregistrement de pixels entre plusieurs canaux spectraux, le profil de la température et l'humidité atmosphériques, les conditions atmosphériques, les conditions de fonctionnement des instruments, les nuages, la topographie, la géométrie de visée, les particules de brume.	[Peters 1995]
Paramètres de texture	L'interaction spatiale entre la taille de l'objet dans la scène et la résolution spatiale du capteur, l'effet de frontières, l'ambiguïté dans la distinction objet / arrière-plan.	
Paramètres de forme	La comptabilisation de la position saisonnière du Soleil par rapport à la Terre, les conditions dans lesquelles l'image à été acquise, les changements dans l'illumination de la scène, les conditions atmosphériques, la géométrie d'observation.	[Peters 1995]
Paramètres de végétation	La variation dans la luminosité de fond de sol, les bandes rouge et NIR, les perturbations atmosphériques, les perturbations atmosphériques en fonction de l'état de l'atmosphère et de la surface du sol au moment des deux acquisitions, la variabilité dans la structure de sous-pixel de la végétation.	
Paramètres climatiques	La correction atmosphérique, le bruit du capteur, l'émissivité de la surface de la terre, les absorbeurs de gaz, les effets angulaires, l'imperfection de longueur d'onde.	[Hulley et al. 2012]

TABLE 1 – Sources d'incertitude de paramètres d'entrée des modèles de COS.

NDVI (*Normalized Difference Vegetation Index*), etc. ;

- paramètres liés aux conditions atmosphériques (paramètres climatiques), comme la température, l'humidité, etc.

Les incertitudes de paramètres sont souvent considérées de nature aléatoire. Néanmoins, dans un modèle de COS, il existe des paramètres de nature épistémique car leur incertitude est due au manque de connaissance dans un modèle physique. Généralement, le type d'incertitude de chaque paramètre dépend de sources de leur incertitude. Par conséquent, il est nécessaire d'identifier les sources d'incertitudes liées aux paramètres.

Le Tableau 1 présente les différents paramètres d'entrée pour les modèles de COS et leurs sources d'incertitude.

Dans la littérature, plusieurs études ont proposé des approches de réduction des incertitudes aléatoires et épistémiques de paramètres des modèles de COS. Le Tableau 2 présente ces approches de réduction. Selon ces études, nous pouvons améliorer la qualité de paramètres avant d'être utilisés dans l'étape de propagation.

En outre, l'incertitude dans les distributions de paramètres [Heijungs et Frischknecht 2005] et le manque d'information sur les corrélations entre les paramètres d'entrée restent un défi dans la plupart des études de COS [Gao et Dennis 2001, Pea 2008]. Généralement,

Paramètres d'entrée	Facteurs de réduction d'incertitude de paramètres	Références
Paramètres spectraux	exigences strictes pour la conception de l'instrument, choix des procédures appropriées pour le calibrage de bord, choix des algorithmes appropriés pour la correction radiométrique et atmosphérique, réduction de la plage de longueur d'onde de rayonnement ou de mesure spectrale, réduction des ombres des nuages et des effets de contamination des nuages et réduction des erreurs du système de capteurs lui-même.	[Peters 1995]
Paramètres de texture	utilisation d'une résolution spatiale élevée et choix des méthodes appropriées pour la segmentation.	[Miura <i>et al.</i> 2000, Pesaresi et Benediktsson 2001]
Paramètres de forme	amélioration de la stabilité des plates-formes et de la vitesse du transporteur, amélioration de la technologique des capteurs eux-mêmes, réduction des effets des conditions atmosphériques, amélioration de la qualité globale de la segmentation, réduction de nombre de mauvais pixels et de la taille des mauvaises zones, et amélioration de l'incertitude de la réponse des pixels.	[Peters 1995]
Paramètres de végétation	choix des algorithmes appropriés pour la correction atmosphérique, réduction des erreurs dans les mesures de surface pour les bandes rouges et NIR, réduction des effets des variations temporelles dans le zénith du soleil et les angles d'azimut, et réduction des effets d'angle du soleil et de la contamination du bruit.	[Jimenez-Munoz et Sobrino 2006, Hird et McDermid 2009]
Paramètres climatiques	Choix des algorithmes appropriés pour la correction atmosphérique, réduction des erreurs du système de capteurs lui-même, réduction des variations d'émissivité, réduction des effets d'angle du soleil, réduction des erreurs de calibrage du radiomètre et des erreurs de rayonnement, et réduction des erreurs de variabilité spatio-temporelle des nuages.	[Jimenez-Munoz et Sobrino 2006, Hulley <i>et al.</i> 2012]

TABLE 2 – Approches de réduction d'incertitude liée aux paramètres des modèles de COS.

les modèles de COS contiennent un grand nombre de paramètres corrélés conduisant à des problèmes non-identifiables.

3.2 Incertitudes sur la structure des modèles de COS

Ces incertitudes résultent de simplifications et d'hypothèses relatives à un manque de connaissance sur un système ou sur le comportement d'une structure dans des conditions variées [Ditlevsen 1983]. Un modèle de COS est une représentation parcimonieuse d'un système terrestre dans une forme mathématique, abstraite, numérique, ou expérimentale. Aucun modèle ne peut représenter complètement une situation ou un processus réel en raison des hypothèses formulées au cours du processus de modélisation. En dehors de l'hypothèse simplificatrice, les modèles peuvent également varier en fonction des décisions prises au cours du processus de modélisation à l'égard de la préférence du modéleur et les exigences de l'utilisateur du modèle. Par exemple, un modéleur peut construire une variété de modèles d'éléments finis pour analyser un système d'ingénierie en faisant varier les types d'éléments. Pour résumer, différents modèles peuvent être construits pour un certain système physique, mais aucun d'entre eux est une représentation complète du comportement physique. Compte tenu des modèles d'approximation pour un système terrestre physique, le problème peut se produire s'il est possible de sélectionner un modèle physique qui représente le système avec la plus haute fidélité de l'ensemble des modèles considérés. En effet, il est au-delà de notre capacité à sélectionner la meilleure approximation d'un ensemble de modèles. Ceci parce que nous ne pouvons pas mesurer les degrés relatifs de différence entre un système physique et un ensemble de modèles se rapprochant du système. Ainsi, l'incertitude est impliquée dans le choix du meilleur modèle à partir d'un ensemble de modèles. Dans cette thèse, cette incertitude est appelée incertitude de la structure du modèle. L'incertitude de la structure du modèle est principalement épistémique car elle vient de notre manque de connaissances.

En général, le processus de modélisation induit des erreurs de modélisation, liées au niveau de détail du modèle, qui empêchent le modèle d'être complètement représentatif du système réel. Parmi ces erreurs, nous citons :

- les erreurs numériques qui incluent l'erreur de discrétisation, l'erreur de convergence, et les erreurs faites lors de l'implémentation et la programmation (parfois qualifiées d'incertitudes numériques) ;
- la simplification, la négligence ou l'ignorance des dépendances structurelles entre les différentes variables du modèle mathématique ou du modèle probabiliste ;
- les hypothèses simplificatrices au niveau des lois de comportement ;
- les formulations mathématiques approchées et les approximations effectuées dans la modélisation de certaines parties du système ;

Une description complète de différentes sources d'incertitude de la structure du modèle est effectuée dans [Morgan et Henrion 1992, Haimes 2009]. Ces incertitudes se propagent à travers le modèle COS, ce qui génère une incertitude dans l'état du système modélisé, c'est-à-dire dans les prédictions d'occupation des sols. L'incertitude de la structure du modèle devrait être prise en considération lors de la prédiction des COS. Ignorant l'incertitude de la structure du modèle peut conduire à sous-estimer la variabilité des prédictions ou générer des prédictions erronées [Draper 1995, Raftery 1996]. Afin

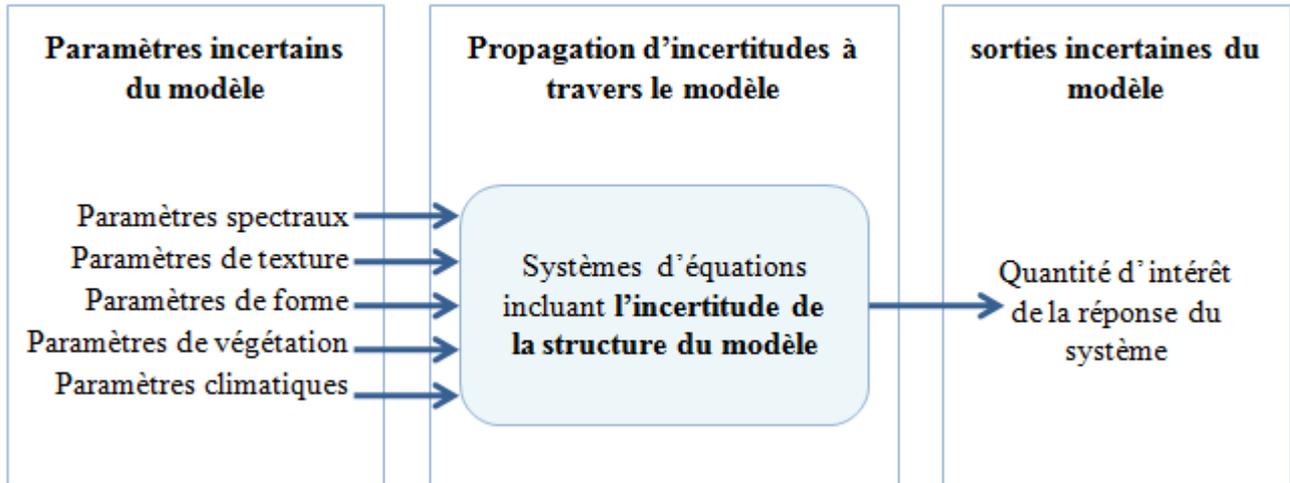


FIGURE 14 – Illustration du principe de propagation des incertitudes.

de réduire les incertitudes de la structure du modèle, il suffit de construire le modèle pour qu'il soit le plus proche possible de la structure réelle.

4 Propagation des incertitudes dans les modèles de COS

Après avoir détaillé les types et sources d'incertitude dans les modèles de COS, nous allons décrire les différentes méthodes de propagation d'incertitude pour les modèles de COS.

Les modèles de COS sont de plus en plus complexes ayant des caractéristiques non-linéaires. En analysant numériquement ces modèles, les ingénieurs tendent généralement à considérer que leurs modèles mathématiques et leurs paramètres sont déterministes. Cependant, le retour au physique montre que ces éléments sont souvent incertains, à cause d'une variabilité naturelle (incertitude aléatoire) ou d'un manque de connaissance (incertitude épistémique). Selon cette distinction de deux types d'incertitude, les experts classent les méthodes de propagation d'incertitude dans les modèles numériques en deux grandes familles (Figure 14) :

- les méthodes probabilistes pour traiter les incertitudes aléatoires [Shi et Ehlers 1996, Crosettoa *et al.* 2001, Boulila *et al.* 2014, Sexton *et al.* 2015];
- les méthodes non probabilistes pour traiter les incertitudes épistémiques [Zhao *et al.* 2009, Jacquin 2010, Abdallah *et al.* 2013, Ferchichi *et al.* 2015];

Cette section brosse un état de l'art, non exhaustif, de ces méthodes qui sont utilisées dans plusieurs domaines de la télédétection.

4.1 Les méthodes probabilistes

Les méthodes probabilistes sont les méthodes non déterministes les plus employées pour le traitement des incertitudes car elles sont basées sur l'usage des probabilités.

Leur objectif est d'étudier la variabilité des grandeurs de sortie du modèle causée par des incertitudes aléatoires liées aux paramètres d'entrée.

Dans ce qui suit, nous rappelons quelques notions de base de la théorie des probabilités.

4.1.1 Généralités

La notion de probabilité est liée à celle d'expérience aléatoire. Une expérience est aléatoire si nous ne pouvons pas prédire avec certitude son résultat. Le résultat d'une expérience aléatoire est appelé réalisation, notée θ . Tous les résultats possibles d'une réalisation forment un espace appelé *l'espace des réalisations*, noté Θ . Un événement B est défini comme un sous-ensemble de Θ contenant des réalisations $\theta \in \Theta$. Ainsi, l'ensemble certain est tout l'espace Θ et l'événement impossible est l'ensemble vide \emptyset . L'intersection (respectivement l'union) de deux événements B et C est notée $B \cap C$ (respectivement $B \cup C$). En considérant que Θ est dénombrable, nous notons dans ce cas $\mathcal{F} = \mathcal{P}(\Theta)$ l'ensemble des sous ensembles de Θ . Nous disons alors que B est un événement si $B \in \mathcal{F}$.

L'objectif de la théorie des probabilités est d'associer des chiffres à des événements, à savoir leur probabilité de réalisation. Une probabilité (ou mesure de probabilité) est une fonction \mathbb{P} de l'ensemble des événements \mathcal{F} vers $[0, 1]$ telle que :

- $\mathbb{P}(\Theta) = 1$ et $\mathbb{P}(\emptyset) = 0$.
- Si $(B_i, i \in I)$ est une collection finie ou dénombrable d'événements ($B_i \in \mathcal{F}, \forall i \in I$) disjoints deux à deux ($B_i \cap B_j = \emptyset, i \neq j, B_i, B_j \in \mathcal{F}, i, j \in I$), alors :

$$\mathbb{P}\left(\bigcup_{i \in I} B_i\right) = \sum_{i \in I} \mathbb{P}(B_i) \quad (1)$$

Le triplet $(\Theta, \mathcal{F}, \mathbb{P})$ est appelé espace de probabilités. Une variable aléatoire réelle X est une application mesurable $X : \theta \in \Theta \rightarrow X(\theta) \in D_X \subset \mathbb{R}$. Nous préférons la notation $X(\theta)$ quand nous voudrions souligner le caractère aléatoire de X . Si D_X est un ensemble discret, la variable aléatoire est dite discrète, sinon elle est dite continue. La variable aléatoire X est complètement définie par sa fonction de répartition F_X prenant des valeurs entre 0 et 1 :

$$F_X(x) = \mathbb{P}(X \leq x), x \in X \quad (2)$$

La variable aléatoire X peut être également définie par sa densité de probabilité f_X qui n'est autre que la fonction dérivée de $F_X : F_X(x) = \int_{-\infty}^x f_X(u) du$ pour tout $x \in \mathbb{R}$.

La notion d'espérance est la formalisation du concept de "en moyenne". L'espérance mathématique $\mathbb{E}[\bullet]$ d'une variable aléatoire réelle correspond à la valeur que nous s'attendons à trouver, en moyenne, si nous répétons un grand nombre de fois la même expérience aléatoire. Dans le cas de variables aléatoires continues, la moyenne μ , la variance $Var[\bullet]$, l'écart-type σ et le moment d'ordre n s'écrivent respectivement :

$$\mu_X \equiv \mathbb{E}[X] = \int_{D_X} x f_X(x) dx \quad (3)$$

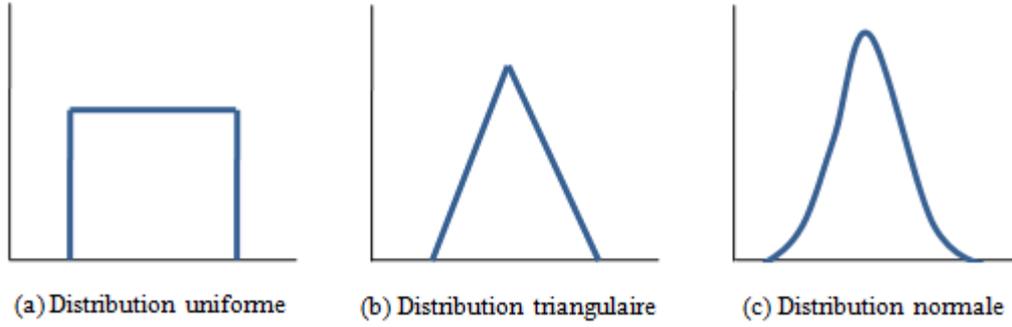


FIGURE 15 – Distributions statistiques.

$$Var[X] = \sigma^2 = \mathbb{E}[(X - \mu_X)^2] = \int_{D_X} (x - \mu_X)^2 f_X(x) dx \quad (4)$$

$$\mathbb{E}[X^n] = \int_{D_X} x^n f_X(x) dx \quad (5)$$

Un vecteur aléatoire est un vecteur dont les composantes sont des variables aléatoires. Connaissant la densité $f_{(X,Y)}$ du vecteur aléatoire continu (X, Y) , la covariance des variables X et Y est définie par :

$$Cov[X, Y] = \mathbb{E}[(X - \mu_X)(Y - \mu_Y)] = \int_{D_X} \int_{D_Y} (x - \mu_X)(y - \mu_Y) f_{(X,Y)}(x, y) dx dy \quad (6)$$

Un champs aléatoire est une fonction à valeurs dans \mathbb{R} , qui dépendent de l'aléa θ et de l'espace x . Il est généralement caractérisé par la donnée d'une moyenne et d'une variance qui dépendent de l'espace, et d'une fonction de covariance entre deux points de l'espace.

Parmi les lois à densité d'usage fréquent en modèles probabilistes de COS, nous trouvons :

- La loi uniforme : la distribution uniforme (Figure 15a) est l'une des distributions continues les plus simples, définie uniquement par des valeurs minimales et maximales. Elle est généralement utilisée lorsque la forme de la distribution est inconnue ; ainsi, elle suppose que chaque valeur possible a une probabilité égale.
- La loi triangulaire : la distribution triangulaire (Figure 15b), contrairement à la distribution uniforme, elle permet d'exprimer une probabilité plus grande pour les valeurs centrales sans toutefois connaître la vraie distribution. Ses paramètres sont la moyenne et la distance entre celle-ci et le minimum ou le maximum. Il est également possible de se servir d'une distribution triangulaire asymétrique. Dans ce cas, les paramètres sont le minimum, le maximum et le mode.
- La loi gaussienne ou normale : il s'agit de la distribution la plus connue et la plus répandue en statistique classique, puisque, selon le théorème central limite, la somme d'une infinité de variables aléatoires indépendantes et identiquement distribuées est normalement distribuée. Cette distribution est symétrique (Figure 15c) et ses paramètres sont la moyenne et l'écart-type. Ainsi, la distribution normale peut être utilisée pour modéliser la variabilité naturelle lorsque la moyenne et l'écart-type sont connus [Smith *et al.* 1992].

La loi normale a une importance capitale en calcul des probabilités et en statistique. Par exemple, elle apparaît naturellement comme la loi d'erreurs de mesures, d'après

le théorème central limite. Elle permet de construire des modèles robustes pour lesquels des calculs explicites peuvent être faits. Voici une synthèse critique des principales distributions [Morgan et Henrion 1992, Heijungs et Frischknecht 2005].

4.1.2 Méthodes probabilistes de propagation des incertitudes

Les méthodes probabilistes sont généralement utilisées pour modéliser les incertitudes provenant d'une variabilité naturelle. Elles permettent la propagation d'incertitudes des données ou des paramètres caractérisant un modèle mathématique donné de COS. En littérature, les méthodes qui se basent sur la théorie des probabilités sont les plus utilisées vu que les incertitudes sont généralement considérées aléatoires. Les méthodes probabilistes peuvent être réparties en cinq catégories :

- La première catégorie se constitue des méthodes d'échantillonnage telles que la simulation de Monte Carlo (*Monte Carlo Simulation*, en anglais) [Fishman 1996], l'échantillonnage préférentiel (*Importance Sampling*, en anglais) [Melchers 1989] et l'échantillonnage d'hypercube latin (*Latin Hypercube Sampling*, en anglais) [Helton et Davis 2003]. Les méthodes d'échantillonnage permettent d'obtenir un nombre des réalisations générés au hasard pour les variables d'entrée aléatoires afin d'obtenir des informations stochastiques sur les réponses du système. La procédure de mise en oeuvre de **la simulation de Monte Carlo** consiste à spécifier des distributions de probabilités représentant des variables d'entrée aléatoires, la génération d'ensembles d'échantillonnage à partir des distributions, la simulation des sorties de réponse en utilisant les ensembles d'échantillonnage générés et le calcul d'informations probabilistes sur les réponses. Cette méthode est avantageuse car elle peut s'appliquer à n'importe quel modèle physique indépendamment de sa complexité, en plus d'être simple à implémenter sans modification du modèle mathématique. Cependant, son utilisation s'avère prohibitive car la vitesse de convergence est en σ/\sqrt{n} , où σ est l'écart type de la distribution résultante et n est le nombre d'itérations. L'obtention d'une valeur stable et précise de la réponse d'un modèle d'une structure réelle exige alors un grand nombre d'échantillons, ce qui rend le temps de calcul très élevé. **L'échantillonnage préférentiel** a été développé pour surmonter les limites de la méthode de Monte Carlo. La méthode de l'échantillonnage préférentiel est une technique de réduction de variance qui utilise un ensemble d'échantillonnage généré à partir d'une distribution biaisée qui se concentre sur les régions importantes d'une variable d'entrée. Cependant, pour que ces techniques restent avantageuses, le nombre de variables aléatoires indépendantes doit rester faible. **L'échantillonnage d'hypercube latin** (LHS) est une autre technique de réduction de variance qui génère une distribution de collections plausibles de valeurs de variables d'entrée à partir d'une distribution multidimensionnelle pour réduire le nombre de simulations pour des problèmes de dimension plus élevée. Les limites du LHS sont discutées dans [Huntington et Lyrintzis 1998], en présentant deux techniques permettant d'améliorer sa performance. De plus, pour réduire davantage les temps de calcul, des variantes sont proposées dans [Olsson *et al.* 2003] avec une étude comparative sur trois problèmes de fiabilité. Un avantage fondamental des méthodes d'échantillonnage est leur utilisation directe de sorties simulées pour obtenir des informations probabilistes sur les réponses du système qui ne peuvent être résolues facilement par des procédures connues. L'inconvé-

nient des méthodes d'échantillonnage est que l'analyse d'incertitude utilisant des techniques d'échantillonnage nécessiterait un très grand nombre de réalisations, ce qui rend pratiquement impossible d'utiliser des méthodes pour des problèmes d'ingénierie à grande échelle.

- La seconde catégorie se constitue des méthodes basées sur l'expansion locale, telles que l'expansion de la série de Taylor [Heuvelink *et al.* 1989], la méthode de perturbation [Sudret 2002] et la méthode de décomposition de Neumann [Wang *et al.* 2013]. L'idée des méthodes d'expansion est d'estimer les moments statistiques des réponses du système avec une petite perturbation dans les variables d'entrée aléatoires. **L'expansion de la série de Taylor et la méthode de perturbation** impliquent l'expansion d'une réponse du système comme une série en termes de petite perturbation aléatoire dans les variables d'entrée et la solution subséquente des coefficients de série. L'expansion de la série de Taylor et la méthode de perturbation exigent des dérivées partielles d'ordre élevé pour maintenir une bonne précision. **La méthode de décomposition de Neumann** consiste à trouver l'inverse des matrices aléatoires par l'expansion de la série de Neumann des matrices. Les méthodes d'expansion nécessitent moins de calculs et conviennent à de petits niveaux d'incertitude. Les méthodes basées sur l'expansion locale ont été largement utilisées pour calculer les deux premiers moments statistiques des réponses du système. Le principal inconvénient de ces méthodes est qu'elles pourraient devenir inefficaces ou inexactes lorsque le nombre ou le degré d'incertitude d'entrée est important. En outre, ces méthodes peuvent ne pas être pratiques pour des applications d'ingénierie à grande échelle en raison de l'exigence de dérivées partielles d'ordre élevé des réponses du système.
- La troisième catégorie se constitue des méthodes basées sur les points les plus probables (*most probable point (MPP) based methods*, en anglais) [Choi *et al.* 2006]. **La méthode de fiabilité du premier ordre (FORM) et les méthodes de fiabilité du second ordre (SORM)** font partie de cette catégorie et constituent deux méthodes d'approximation populaires qui estiment la probabilité de défaillance associée à une fonction d'état limite. La distinction entre les deux méthodes dépend de l'utilisation d'une expansion de la série de Taylor du premier ordre ou du second ordre pour approximer une fonction d'état limite au MPP. L'indice de sûreté de rotation symétrique est introduit par la transformation de Hasofer et Lind (HL) [Hasofer et Lind 1974] pour les deux méthodes. Si l'approximation d'une fonction d'état limite au MPP est exacte, les méthodes basées sur MPP peuvent produire des résultats satisfaisants. En outre, ces méthodes peuvent calculer l'effet relatif de chaque variable d'entrée (facteur de sensibilité) sur la probabilité de défaillance sans coût de calcul supplémentaire. Cependant, la FORM et la SORM peuvent entraîner de grandes erreurs lorsque la fonction d'état limite est fortement non linéaire. De plus, ces méthodes nécessitent le calcul des dérivées du premier ordre ou du second ordre des réponses du système qui sont extrêmement intensive pour une simulation multi-physique avec un grand nombre de variables aléatoires. Récemment, des améliorations des méthodes basées sur MPP ont été développées telles que la méthode FOSPA (*First-Order Saddlepoint Approximation method*) [Du et Sudjianto 2004] et la méthode MVFOSA (*Mean-Value First Order Saddlepoint Approximation*) [Huang et Du 2008] pour rendre le calcul de la fonction de défaillance plus efficace, plus précis et plus robuste.

- La quatrième catégorie se constitue des méthodes basées sur l’expansion stochastique. **La méthode spectrale des éléments finis stochastiques** (*Spectral Stochastic Finite Element Method*, en anglais) [Ghanem et Spanos 1991] et **la méthode de surface de réponse stochastique** (*Stochastic Response Surface Method*, en anglais) [Isukapalli *et al.* 1998] peuvent être classées dans ce groupe. En 1990, Ghanem et Spanos [Ghanem et Spanos 1991] ont proposé la SSFEM (*Spectral Stochastic Finite Element Method*) dans lequel la méthode des éléments finis est étendue pour traiter des systèmes incertains gouvernés par des équations aux dérivées partielles stochastiques. SSFEM est une approche bien adaptée pour une analyse impliquant des champs aléatoires. Dans cette méthode, l’extension de Karhunen-Loève (KL) [Karhunen 1947, Loève 1977] est effectivement utilisée pour représenter des champs aléatoires. L’idée de base du SSFEM est de représenter une réponse du système par une combinaison linéaire de polynômes orthogonaux de variables aléatoires standard. SRSM (*Stochastic Response Surface Method*) utilise également des polynômes orthogonaux pour représenter des réponses d’un système stochastique, mais ne nécessite aucune modification des codes FE (*Finite Element*) déjà construits contrairement au SSFEM. En raison des procédures de formulation non intrusives, SRSM fournit une approche pratique pour l’analyse stochastique des problèmes d’ingénierie à grande échelle. Les méthodes basées sur l’expansion stochastique peuvent fournir des informations sur les statistiques d’ordre supérieur des comportements du système stochastique, même pour de grands niveaux d’incertitude d’entrée impliqués. Cependant, ces méthodes peuvent être inappropriées pour des problèmes impliquant un haut non linéarité et une discontinuité élevée dans lesquelles une divergence potentielle dans des moments d’ordre supérieur peut être observée.
- La dernière catégorie comporte les méthodes basées sur l’intégration numérique. **Les méthodes d’intégration numérique multidimensionnelle** font partie de cette famille et sont caractérisées par l’utilisation de règles en quadrature basées sur des moments [Evans 1967]. La fonction de réponse est évaluée à un ensemble de points d’échantillonnage bien conçus, souvent appelés points de quadrature, pour l’intégration numérique. Une méthode nouvellement développée et appartenant à cette catégorie est **la méthode de réduction de dimension** (DR) [Rahman et Xu 2004, Youn *et al.* 2006]. La méthode DR se rapproche de la réponse du système par des intégrations multidimensionnelles basées sur une décomposition additive d’une fonction de réponse. Ceci rend cette méthode sensiblement plus simple et plus efficace que les méthodes traditionnelles d’intégration multidimensionnelle. Cependant, la méthode DR peut produire une erreur relativement importante pour le second ordre. En effet, les moments supérieurs de réponses d’un système impliquent une grande variation aléatoire et/ou une non-linéarité élevée. Les méthodes basées sur l’intégration numérique ne nécessitent pas de calculer les dérivées des réponses, mais peuvent prédire avec précision les moments de réponse. L’inconvénient des méthodes basées sur le moment est leur inefficacité numérique. Les méthodes basées sur le moment tendent à devenir inefficaces quand le nombre de variables aléatoires devient grand.

Pour plus du détail sur des travaux récents se rapportant à l’application des méthodes probabilistes de propagation des incertitudes dans les modèles de COS, on pourra consulter [Boulila *et al.* 2014, Krüger et Lakes 2015, Versteegen *et al.* 2016]. Pour les

modèles de COS, nous pouvons constater que l'ensemble des méthodes décrites précédemment reste encore trop coûteux. De plus, ne disposant pas de méthode prenant en compte différents types et sources d'incertitudes de manière simple et peu intrusive.

4.2 Les méthodes non probabilistes

Comme nous l'avons précisé dans le paragraphe 2.2, les incertitudes peuvent être épistémiques, dues à des informations manquantes, imprécises ou incomplètes. Une modélisation probabiliste est alors injustifiée, insuffisante et entraîne une déformation du problème traité. En effet, l'ignorance de la cause de variation d'une grandeur ne rend pas cette variation aléatoire. Afin de traiter ce type d'incertitude dans la modélisation et les analyses numériques non déterministes, des méthodes non probabilistes ont émergé. Une synthèse générale de ces approches non probabilistes est présentée dans [Kang *et al.* 2016], et les activités de recherche et applications récentes sont décrites dans [Chakraborty *et al.* 2013, Jadidi *et al.* 2014].

4.2.1 La théorie des intervalles

La théorie des intervalles (*Interval Theory*, en anglais) [Moore 1966] fait partie des premiers développements dans ce domaine. Elle permet de calculer les bornes de l'intervalle dans lequel varie la grandeur d'intérêt sans connaissance préalable sur la dispersion des paramètres d'entrée incertains. Les règles de cette théorie sont élaborées par Moore [Moore 1966]. L'objectif est de calculer les valeurs les plus petites et les plus grandes de la production dans le cadre des opérations arithmétiques. Si un intervalle est défini par sa limite inférieure \underline{a} et supérieure \bar{a} , alors nous pouvons effectuer les opérations suivantes :

$$[\underline{a}, \bar{a}] + [\underline{b}, \bar{b}] = [\underline{a} + \underline{b}, \bar{a} + \bar{b}] \quad (7)$$

$$[\underline{a}, \bar{a}] - [\underline{b}, \bar{b}] = [\underline{a} - \bar{b}, \bar{a} - \underline{b}] \quad (8)$$

$$[\underline{a}, \bar{a}] \times [\underline{b}, \bar{b}] = [\min(\underline{a}\underline{b}, \underline{a}\bar{b}, \bar{a}\underline{b}, \bar{a}\bar{b}), \max(\underline{a}\underline{b}, \underline{a}\bar{b}, \bar{a}\underline{b}, \bar{a}\bar{b})] \quad (9)$$

$$[\underline{a}, \bar{a}] \div [\underline{b}, \bar{b}] = [\underline{a}, \bar{a}] \times [1/\bar{b}, 1/\underline{b}], \text{ si } 0 \notin [\underline{b}, \bar{b}] \quad (10)$$

En utilisant ces opérations, la propagation des intervalles montre un phénomène pessimiste de surestimation : l'incertitude de la somme de deux intervalles est supérieure à la somme des incertitudes de chaque intervalle. Des améliorations, telles que le développement de l'arithmétique affine [De Figueiredo et Stolfi 2004], ont été effectuées afin de réduire la croissance de l'incertitude lorsqu'elle se propage dans un modèle. Des applications de la théorie des intervalles ont été effectuées sur des problèmes de télédétection [Sander *et al.* 2006].

4.2.2 La théorie des ensembles flous

La théorie des ensembles flous (*Fuzzy Set Theory*, en anglais) [Zadeh 1999] utilise le concept de logique floue (*fuzzy logic*, en anglais) basé plutôt sur le raisonnement humain. Les ensembles flous sont des ensembles dont les éléments ont des degrés d'appartenance décrits par une fonction dite d'appartenance évaluée dans l'intervalle d'unité

réelle $[0, 1]$. En particulier, nous pouvons souligner le développement d'une méthode des Éléments Finis flous (*Fuzzy FEM*, en anglais) [Rao et Sawyer 1995], où les dispersions des paramètres incertains sont modélisées par des intervalles flous et la réponse est calculée à l'aide de l'arithmétique des intervalles. Les intervalles flous ont été utilisés pour représenter l'incertitude due à l'information imprécise dans les modèles d'évaluation environnementale [Mertens et Huwe 2002, Assaghir 2010].

Un avantage de la théorie des ensembles flous est qu'elle fournit des modèles d'inférence. Ces modèles peuvent traiter l'imprécision intrinsèquement présente dans la connaissance humaine avec la fonction d'appartenance caractérisée par la transition progressive de l'adhésion à la non-adhésion. De plus, Leyk et Zimmermann [Leyk et Zimmermann 2007] ont conclu que cette approche a fourni une représentation plus précise des variables incertaines non-aléatoires dans l'analyse de COS et plus de flexibilité pour les ensembles de données dont les distributions de probabilités étaient difficiles à déterminer. Ainsi, la propagation de l'incertitude par cette approche était plus efficace du point de vue du calcul que les simulations probabilistes.

Cependant, il est difficile de fournir des fonctions d'appartenance optimales pour le système physique analysé car les fonctions d'appartenance sont conçues par des experts ayant une connaissance partielle du système donné. Un autre inconvénient de la théorie des ensembles flous est qu'il existe de nombreuses façons d'interpréter les règles floues et de combiner les sorties de plusieurs règles floues. Dans [Real et al. 2010], Real et al. ont montré que même si les intervalles flous étaient efficaces du point de vue du calcul et fournissaient des résultats comparables à ceux d'une approche probabiliste, ils offraient des informations insuffisantes sur l'incertitude de la sortie (c'est-à-dire un intervalle de confiance approximatif). Ceci pourrait être moins utile dans les applications de COS. Ainsi, l'interprétation des résultats basés sur des ensembles flous pour la prise de décision semble être un défi [Real et al. 2010]. De plus, son application pratique est encore limitée en raison de la complexité du modèle étudié. Par exemple, Sarmiento et al. [Sarmiento et al. 2013] ont illustré la faisabilité de la propagation floue dans un modèle de COS. Mais, ils ont observé une faible efficacité lorsque le modèle n'était pas monotone.

4.2.3 La théorie des possibilités

La théorie des possibilités (*Possibility Theory*, en anglais), introduite par Dubois et Prade [Dubois et Prade 1998], est considérée comme une extension de la théorie des ensembles flous. Cette théorie peut fidèlement représenter des états d'ignorance partielle par l'utilisation de la distribution de possibilités. Les mesures de possibilité et de nécessité définies par une distribution de possibilités sont utilisées pour représenter la confiance d'un agent quant à la probabilité que des événements connexes se produisent. La théorie de la possibilité a les mêmes avantages et inconvénients à ceux de la théorie des ensembles flous mentionnés ci-dessus puisque cette dernière fournit une base naturelle pour la première. La théorie des possibilités a été largement utilisée dans plusieurs domaines d'ingénieries [André et Lopes 2012], mais semble moins employée en analyse de COS pour le traitement des incertitudes épistémiques [Ferchichi et al. 2015].

4.2.4 La théorie des probabilités imprécises

La théorie des probabilités imprécises (*Imprecise Probabilities*, en anglais), développée par Walley [Walley 1991, Walley 2000], est une théorie mathématique constituée le cadre le plus général de représentation des incertitudes. Les probabilités imprécises sont définies comme un encadrement d'une distribution de probabilité par deux distributions haute et basse. Cette approche est intéressante lorsqu'on sait le type de modèle aléatoire avec une connaissance imprécise sur les paramètres, ou lorsque notre connaissance est limitée à certaines valeurs remarquables de la distribution. Une nouvelle forme de probabilités imprécises, basée sur les intervalles généralisés, a été proposée récemment dans [Wang 2010] pour améliorer l'évaluation de la fiabilité des systèmes. Les analystes du risque débattent de l'utilisation de la probabilité imprécise dans l'analyse de l'incertitude. Certains croient que la probabilité imprécise pourrait confondre les gens en raison de son interprétation des résultats imprécis [Lindley 2010]. De plus, l'approche subjective (ou bayésienne) permet à un décideur de classer différentes sources d'opinion avec une seule estimation de la probabilité. Cependant, d'autres affirment que la prise de décisions dans l'analyse des risques fondée sur des résultats imprécis ne devrait pas suivre la vue subjective (ou bayésienne) de la probabilité si le jugement d'expert ne peut pas fournir une expression plus précise, parce que l'approche subjective ignore l'imprécision [Ferson et Tucker 2006, Aven et Zio 2011, Rinderknecht *et al.* 2012].

Dans ce qui suit, nous allons présenter en détail la théorie des fonctions de croyance, puisque c'est la méthode de base dans cette thèse.

4.2.5 La théorie des fonctions de croyance

La théorie des fonctions de croyance (*Theory of Belief Functions* ou *Evidence Theory*, en anglais), aussi connue sous le nom de la théorie de Dempster-Shafer (*Dempster-Shafer Theory*, en anglais), initialement développé par Dempster [Dempster 1968] et formalisé par Shafer [Shafer 1976]. Elle est basée sur un fondement mathématique robuste qui permet de représenter l'incertitude que nous pouvons confronter sur une information ou un phénomène à traiter. Elle fournit une méthode pour combiner des corps distincts de preuves collectées à partir de différentes sources. En plus, cette théorie a été largement considérée comme une généralisation de l'inférence Bayésienne [Dempster 1968, Dempster 2008], ou de manière équivalente, que la théorie des probabilités est un cas particulier de la théorie mathématique d'évidence lors de la manipulation de la fonction de masse élémentaire bayésienne. Donc, cette théorie a le potentiel de combiner la représentation probabiliste et la représentation d'intervalle dans une représentation unifiée.

La théorie des fonctions de croyance est capable de gérer l'incertitude à partir de deux fonctions : la fonction de croyance (*Bel*) et la fonction de plausibilité (*Pl*) qui sont dérivées de la fonction de masse élémentaire (BPA, aussi appelée fonction de masse). La BPA pour un ensemble donné A (noté $m(A)$) indique la proportion de toutes les preuves disponibles qui supporte la supposition qu'un élément particulier de x appartient à l'ensemble A . Elle a des axiomes tels que :

$$m : 2^\Omega \rightarrow [0, 1] \quad (11)$$

$$m(\emptyset) = 0 \quad (12)$$

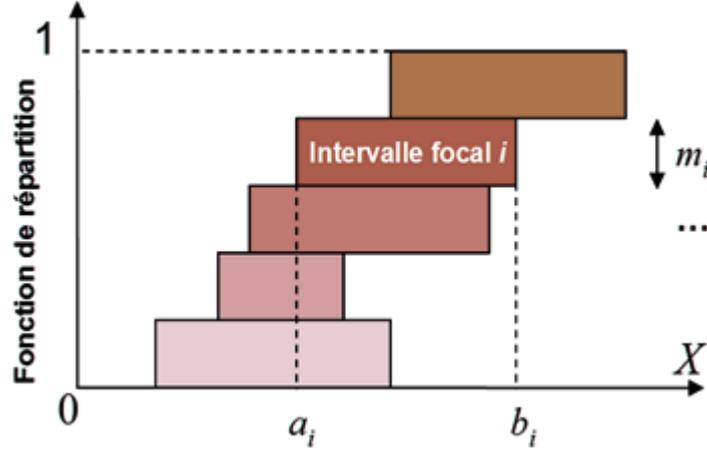


FIGURE 16 – Intervalles focaux $[a_i, b_i]$ et les masses de croyance m_i associées.

$$\sum_{A \in 2^\Omega} m(A) = 1 \quad (13)$$

où 2^Ω est l'ensemble de puissance qui comprend tous les sous-ensembles possibles, y compris l'ensemble vide \emptyset . A est un sous-ensemble (appelé "élément focal") de 2^Ω . Les fonctions de croyance et de plausibilité sont définies à partir du BPA. La fonction de croyance de A est la somme de la BPA de tous les sous-ensembles (B) de A ($B \subseteq A$) :

$$Bel(A) = \sum_{B \subseteq A} m(B) \quad \forall B \in 2^\Omega \quad et B \neq \emptyset \quad (14)$$

La fonction de plausibilité de A est la somme de la BPA d'un sous-ensemble (C) de 2^Ω , avec la condition que l'intersection de C et A est un ensemble non vide ($C \cap A \neq \emptyset$) :

$$Pl(A) = \sum_{C \cap A \neq \emptyset} m(C) \quad \forall C \in 2^\Omega \quad et C \neq \emptyset \quad (15)$$

Les trois concepts peuvent également être utilisés dans des distributions de probabilité continues où chaque élément focal du paramètre incertain X peut être exprimé comme un intervalle ($[a_i, b_i]$) avec m_i (ou BPA_i) (où $a_i \leq b_i$ pour tout i). Comme nous pouvons le voir dans la Figure 16, l'incertitude associée au paramètre incertain X peut être représentée par ce que l'on appelle une structure d'intervalles focaux de Dempster-Shafer [Dempster 1968] comme suit :

$$\{([a_1, b_1], m_1), ([a_2, b_2], m_2), \dots, ([a_n, b_n], m_n)\} \\ \text{où } a_i \leq b_i, \sum_{i=1}^n m_i([a_i, b_i]) = 1 \quad et \quad [a_i, b_i] \subseteq \mathbb{R}, \quad 1 \leq i \leq n. \quad (16)$$

Du point de vue calculatoire, cette représentation est pratique pour la propagation des incertitudes à travers un modèle par la simulation [Le Duy *et al.* 2010].

Ainsi, l'ensemble de X est la collection de ces intervalles avec leurs masses m_i associées, et la somme de m_i est égale à 1. Donc, les équations 14 et 15 peuvent être transformées comme suit :

$$Bel(X \in \mathbb{R}) = \sum_{[a_i, b_i] \subseteq x} m([a_i, b_i]) \quad (17)$$

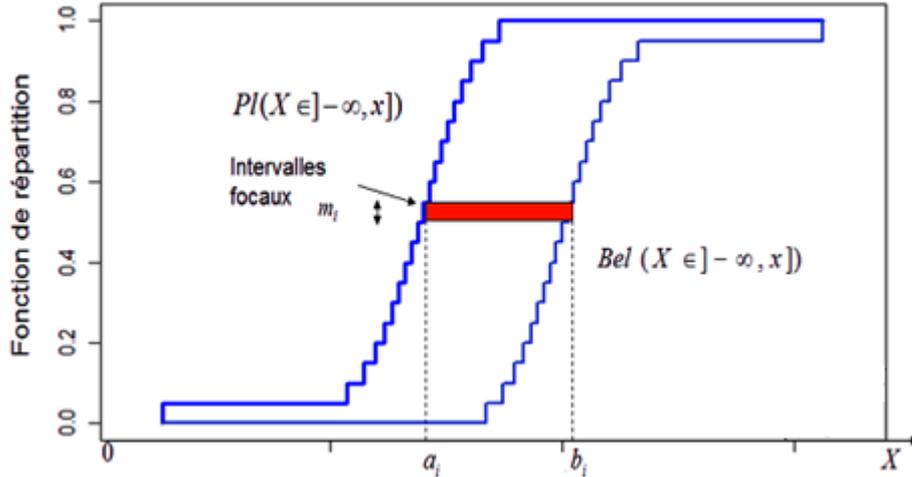


FIGURE 17 – Fonctions de croyance et plausibilité cumulatives.

$$Pl(X \in R) = \sum_{[a_i, b_i] \cap x \neq \emptyset} m([a_i, b_i]) \quad (18)$$

Comme dans la théorie des probabilités, où la fonction de répartition $F(x)$ (*Cumulative Distribution Function*, en anglais) est définie sur les intervalles semi-fermés $] -\infty, x] \subseteq \mathbb{R}$ avec $F(x) = P(X \in] -\infty, x])$, les fonctions $Bel(X \in] -\infty, x])$ et $Pl(X \in] -\infty, x])$ dans la théorie des fonctions de croyance sont respectivement appelées fonction de croyance cumulative et fonction de plausibilité cumulative (*belief/plausibility cumulative functions*, en anglais) [Yager 2004]. Ces deux fonctions peuvent être obtenues comme suit (Figure 17) :

$$Bel(X \in] -\infty, x]) = \sum_{b_i \leq x} m([a_i, b_i]) \quad (19)$$

$$Pl(X \in] -\infty, x]) = \sum_{a_i \leq x} m([a_i, b_i]) \quad (20)$$

Les fonctions de croyance et de plausibilité peuvent être considérées comme des fonctions de probabilité inférieure et supérieure de X avec une valeur donnée x [Ferson *et al.* 2003], respectivement. La vraie fonction de probabilité de X ($Pr(x)$) se trouve à l'intérieur de ces deux fonctions, est interprétée comme suit :

$$Bel(x) \leq P(x) \leq Pl(x) \quad (21)$$

avec un intervalle de la moyenne de X

$$\sum m_i a_i \leq \mathbb{E}(X) \leq \sum m_i b_i \quad (22)$$

L'avantage d'utiliser cette théorie réside dans le fait qu'elle peut réussir à quantifier le degré d'incertitude quand la quantité d'information disponible est faible. Elle fournit deux mesures incertaines, à savoir la croyance et la plausibilité. Les fonctions et la règle de combinaison de la théorie représentent de manière appropriée les preuves fournies par le système de connaissances expert et son agrégation. Un autre avantage est que cette théorie a le potentiel de quantifier des incertitudes aléatoires et épistémiques [Chen *et*

[Corson 2014], et de résoudre des problèmes multidimensionnels [Ferson *et al.* 2004]. Cependant, lorsqu'un conflit important se produit entre deux éléments d'évidence, la règle de combinaison de Dempster donne un résultat contraire à l'intuition. En outre, il existe un obstacle à l'application de la théorie des fonctions de croyance que la quantité des efforts de calcul requise pour la combinaison des évidences augmente de façon exponentielle avec le nombre d'événements probables considérés. Des algorithmes de réduction du coût de calcul sont proposés dans [Rana *et al.* 2012] afin d'améliorer l'emploi pratique de cette théorie comme outil général d'analyse des incertitudes sur des cartes d'occupation des sols de grande taille. On trouve dans [Yang *et al.* 2016] une application de la théorie des fonctions de croyance dans le cadre de la classification des couvertures complexes de terrains urbains contenant des incertitudes. Récemment, des méthodologies, utilisant les structures mathématiques de la théorie des fonctions de croyance, ont été développées pour propager à la fois les incertitudes de paramètres et les incertitudes de la structure du modèle dans les applications nucléaires [Le Duy *et al.* 2013] et les structures mécaniques [Park et Grandhi 2012].

4.3 Synthèse et analyses comparatives

Les méthodes de propagation mentionnées au-dessus ont été largement appliquées dans divers problèmes de télédétection, tels que la fusion d'images [Saeidi *et al.* 2014], la détection de changement [Rana *et al.* 2012, Liu *et al.* 2012], la prédiction de changement [Al-sharif et Pradhan 2015], la segmentation d'images [Fouque *et al.* 2001] et la classification [Yang *et al.* 2016]. Cependant, quelques travaux ont utilisé ces méthodes pour la propagation d'incertitude dans les modèles de COS [van der Kwast *et al.* 2012, Verburg *et al.* 2013, Tayyebi *et al.* 2014, Boulila *et al.* 2014, Krüger et Lakes 2015, Verstegen *et al.* 2016, Ferchichi *et al.* 2016].

Le Tableau 3 présente chacun de ces travaux tout en précisant le modèle de COS utilisé, le type de l'incertitude (d'où vient l'incertitude), la source d'incertitude traitée, les méthodes de propagation et de réduction des incertitudes choisies et le domaine d'application.

Nous remarquons que les approches proposées pour la propagation des incertitudes dans les modèles de COS ont modélisé un seul type d'incertitude. Alors que, comme nous avons montré dans la section 2, plusieurs types d'incertitudes peuvent accompagner les données dans les modèles de COS. Généralement, le choix des méthodes de propagation dépend de la façon dont l'incertitude est représentée [Lloyd et Ries 2007]. Plusieurs méthodes hybrides ont été proposées dans la littérature pour remédier à ce problème [Helton *et al.* 2004, Arunraj *et al.* 2013]. Dans [Helton *et al.* 2004], Helton *et al.* ont comparé des méthodes pour représenter l'incertitude. Ils ont discuté des moyens d'agrèger leur influence sur les prédictions du modèle. Helton *et al.* ont proposé différentes façons de représenter l'incertitude des variables d'entrée en structures probabilistes (CDF ou CDF imprécise) et les ont propagées à travers des modèles avec la méthode Monte Carlo. Cette propagation est similaire à celle de la méthode Monte Carlo classique, sauf que le processus d'échantillonnage génère un intervalle aléatoire au lieu d'une seule valeur aléatoire. Ensuite, chaque itération est calculée avec l'arithmétique des intervalles. La combinaison de la méthode Monte Carlo et la théorie des intervalles permet à l'incertitude stochastique et à l'information imprécise d'être considérées simultanément dans le

Auteurs/date	Modèle(s) de changement	Type(s) d'incertitude	Source(s) d'incertitude	Méthode(s) de propagation	Méthode(s) de réduction	Domaine d'application
Van der Kwast <i>et al.</i> [Van der Kwast <i>et al.</i> 2012]	Automate cellulaire	Incertitude aléatoire	Incertitude de paramètres	Méthode probabiliste (Monte Carlo)	Techniques de filtrage des particules et d'assimilation des données	Changement des types d'occupation des sols
Verbürg <i>et al.</i> [Verbürg <i>et al.</i> 2013]	Régression logistique (le modèle CLUE)	Non spécifié	Incertitude de paramètres et de la structure du modèle	Méthode probabiliste	Analyse de sensibilité	Suivi de l'agriculture
Tayyebi <i>et al.</i> [Tayyebi <i>et al.</i> 2014]	Apprentissage automatique et régression logistique	Non spécifié	Incertitude de paramètres	Méthode probabiliste	Analyse de sensibilité	Urbanisation
Boulija <i>et al.</i> [Boulija <i>et al.</i> 2014]	Apprentissage automatique	Incertitude aléatoire	Incertitude de paramètres	Méthode probabiliste (Expansion stochastique)	Analyse de sensibilité (Indices de Sobol)	Changement des types d'occupation des sols
Krger <i>et al.</i> [Krger <i>et al.</i> 2015]	Modèle Bayésien de réseaux de croyance	Non spécifié	Incertitude de la structure du modèle	Méthode probabiliste (Monte Carlo)	Analyse de sensibilité	Déforestation
Verstegen <i>et al.</i> [Verstegen <i>et al.</i> 2016]	Automate cellulaire	Non spécifié	Incertitude de paramètres et de la structure du modèle	Méthode probabiliste (Monte Carlo)	Techniques de filtrage des particules et d'assimilation des données	Changement des types d'occupation des sols
Ferchichi <i>et al.</i> [Ferchichi <i>et al.</i> 2016]	Apprentissage automatique	Incertitude épistémique	Incertitude de paramètres et de la structure du modèle	La théorie de possibilités	Analyse de sensibilité (Mesure de Hartley-like)	Changement des types d'occupation des sols

TABLE 3 – Comparaison des études de propagation des incertitudes dans les modèles de COS.

modèle lorsque les deux types existent. D'autres travaux ont développé des approches hybrides pour la propagation conjointe des distributions de probabilité et des intervalles flous dans l'évaluation des risques [Guyonnet *et al.* 2003, Baudrit *et al.* 2005, Baudrit *et al.* 2006, Baraldi et Zio 2008, Arunraj *et al.* 2013]. Les résultats ont été présentés sous la forme d'une distribution de probabilité bornée qui distingue l'incertitude due au caractère aléatoire de celle due à une information incomplète. L'application de ces méthodes hybrides est cependant limitée en raison de l'intensité de calcul de la méthode de Monte Carlo et les théories des intervalles. Par ailleurs, les corrélations entre les paramètres aléatoires et épistémiques restent à explorer [Ferson *et al.* 2004].

Une méthode rapide et simple prenant en compte différents types et sources d'incertitudes serait souhaitable. La théorie des fonctions de croyance, comme nous présentons dans la section 4.2.5, fournit des règles de combinaison qui peuvent fusionner des informations dans un modèle de simulation de différentes sources dans le même cadre. Cependant, l'interprétation des résultats basée sur cette théorie semble complexe et plus difficile à comprendre qu'une seule distribution probabiliste. Lindley [Lindley 2013] a soutenu une représentation simplifiée de l'incertitude. Il a affirmé que les structures d'intervalles focaux de Dempster-Shafer [Dempster 1968] ont l'avantage de considérer différents types d'incertitude pour l'analyse d'incertitude de paramètres et de structure des modèles numériques et fournir des estimations plus réalistes. Ces structures peuvent être aussi utilisées pour résoudre des problèmes multidimensionnels [Ferson *et al.* 2004]. Sur la base de ces facteurs, la théorie des fonctions de croyance connait un succès croissant dans plusieurs applications d'ingénieries [Zargar *et al.* 2012, Park et Grandhi 2012, Abdallah *et al.* 2013, Jiang *et al.* 2013, Le Duy *et al.* 2013, Chen et Corson 2014] et, en particulier, de télédétection [Fouque *et al.* 2001, Rana *et al.* 2012, Liu *et al.* 2012, Saeidi *et al.* 2014, Al-sharif et Pradhan 2015, Yang *et al.* 2016].

5 Conclusion

Dans ce chapitre, nous avons mené une étude sur les types et les sources des incertitudes affectant les modèles de COS. Puis, nous avons dressé un état de l'art, non exhaustif, des principales méthodes de propagation des incertitudes. Nous remarquons, suite à cette étude de l'état de l'art, que peu des travaux pour la propagation des incertitudes dans les modèles de COS ont été proposés. En plus, nous remarquons qu'il n'existe actuellement aucune approche qu'a pris en considération à la fois, les incertitudes aléatoires/épistémiques, les incertitudes représentant les paramètres/la structure du modèle et les corrélations entre les paramètres. D'où l'idée de proposer des méthodes de propagation et de réduction des incertitudes tout en considérant ces facteurs en se basant sur la théorie des fonctions de croyance pour améliorer la qualité des résultats des modèles de COS. Pour résoudre ces problèmes, les deux chapitres suivants présentent l'approche proposée.

Chapitre III

Approche proposée pour la propagation des incertitudes dans les modèles de COS

Sommaire

1	Introduction	47
2	Approche de propagation des incertitudes	47
2.1	Analyse d'incertitude de paramètres du modèle	47
2.1.1	Modélisation de l'incertitude de paramètres	48
2.1.2	Etude de la corrélation entre les paramètres	50
2.1.3	Propagation de l'incertitude de paramètres	54
2.2	Analyse d'incertitude de la structure du modèle	58
2.2.1	Caractérisation d'incertitude de la structure du modèle	59
2.2.2	Modélisation de l'incertitude de la structure du modèle	61
2.2.3	Propagation de l'incertitude de la structure du modèle	61
3	Conclusion	64

1 Introduction

Dans ce chapitre, nous présentons notre approche de propagation des incertitudes dans les modèles de COS [Ferchichi *et al.* 2017a]. L'objectif de cette approche est de propager les incertitudes liées aux paramètres et à la structure d'un modèle de COS en utilisant la théorie des fonctions de croyance. L'approche proposée prend les avantages de la théorie des fonctions de croyance pour combiner différents types d'incertitudes (aléatoires et/ou épistémiques) d'une manière très naturelle. La théorie des fonctions de croyance permet aussi de prendre en considération la corrélation entre les paramètres dans le processus de COS.

Dans ce chapitre, nous détaillons les étapes principales du processus de propagation des incertitudes à savoir la modélisation des incertitudes aléatoires et épistémiques, l'étude de corrélation et la propagation des incertitudes de paramètres et de la structure du modèle.

2 Approche de propagation des incertitudes

Dans cette section, nous intéressons en premier lieu à la représentation des incertitudes aléatoires et épistémiques de paramètres par la théorie de fonction de croyance. La modélisation des incertitudes aléatoires dans la pratique traditionnelle repose sur l'hypothèse d'une distribution probabiliste alors que la modélisation des incertitudes épistémiques dans le cadre de la théorie des fonctions de croyance ne repose pas sur cette hypothèse. Pour cela, nous avons proposé deux différents types de modélisation pour ces deux types d'incertitudes. Ensuite, nous étudions les corrélations entre les paramètres d'entrée des modèles de COS. La propagation des incertitudes de paramètres dans le cadre de la théorie des fonctions de croyance est présentée dans la section suivante. Dans la deuxième partie de ce chapitre, nous allons étudier les incertitudes représentant la structure du modèle. Pour se faire, nous commençons par la caractérisation et la modélisation de ces incertitudes. Puis, nous présentons le principe de la propagation. Nous illustrons chaque procédure du processus proposé de propagation par un exemple simple. La Figure 18 présente l'architecture générale de l'approche proposée de la propagation des incertitudes.

2.1 Analyse d'incertitude de paramètres du modèle

Les paramètres d'entrée des modèles de COS sont généralement les caractéristiques des objets contenus dans les images satellitaires. Dans cette étude, nous avons choisi d'utiliser 5 paramètres spectraux, 3 paramètres de texture, 5 paramètres de forme, 1 paramètre de végétation et 2 paramètres climatiques (voir Tableau 4) [Boulila *et al.* 2011, Pu *et al.* 2011, Laliberte *et al.* 2012, Puertas *et al.* 2014]. Les paramètres spectraux sont : l'indice de luminosité (BI) et les valeurs moyennes de la bande bleue (MB), la bande rouge (MR), la bande verte (MG) et la bande NIR (MN) pour chaque objet d'images. Les paramètres de texture sont : le contraste (Ctr), l'entropie (Ent) et le second moment angulaire (ASM) générés à partir de GLCM (*Gray Level Co-occurrence Matrix*). Les paramètres de forme ou géométriques sont : l'ajustement rectangulaire

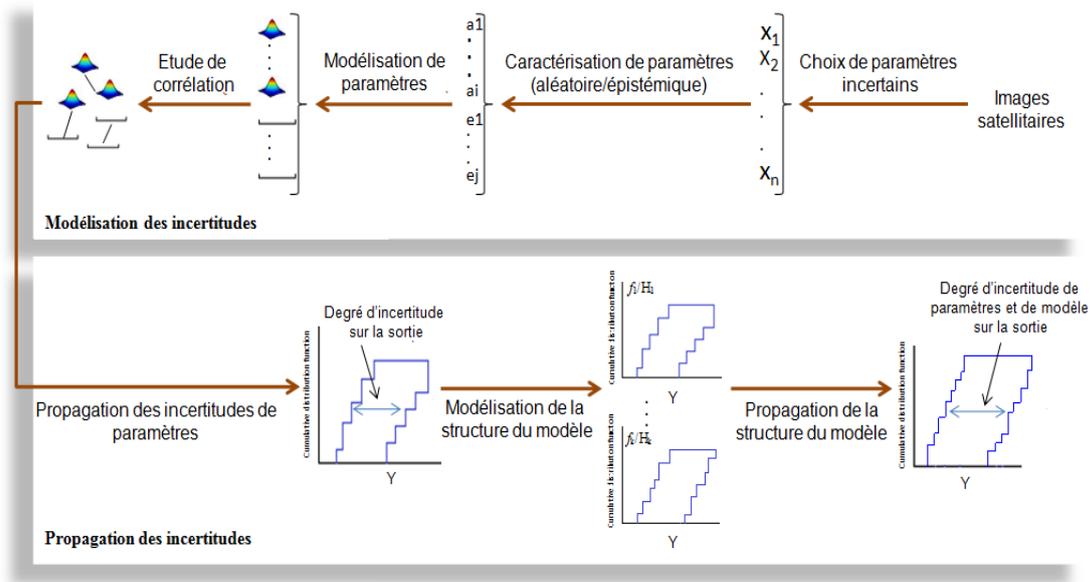


FIGURE 18 – Architecture de l'approche proposée pour la propagation des incertitudes.

(RF), l'ajustement elliptique (EF), l'indice de forme (SI), la densité (D) et la surface (A). Le paramètre de végétation est : le NDVI (*Normalized Difference Vegetation Index*) qui est le rapport de la différence entre la bande NIR et la réflectance rouge. Enfin, les paramètres climatiques sont : la température (Tem) et l'humidité (Hum).

Pour l'approche proposée, nous devons d'abord identifier le type d'incertitude lié à chaque paramètre. Pour expliquer la méthode proposée, considérons une fonction de transfert simple qui présente des paramètres incertains de types différents

$$Y = f(a, e) \quad (23)$$

où a représente les paramètres de type aléatoires, e représente les paramètres de type épistémiques et Y est la réponse du modèle de COS.

2.1.1 Modélisation de l'incertitude de paramètres

Modélisation de paramètres aléatoires

Pour la modélisation de paramètres de types aléatoires, les distributions de probabilités restent les méthodes les plus solides et ceci a été montré et confirmé par des vastes approches développées dans la littérature (voir chapitre 2). Plusieurs types de distributions ont été proposés tels que triangulaires, normales, uniformes, etc. Dans notre cas, nous avons choisi d'utiliser une distribution de type normale pour modéliser les paramètres aléatoires a .

Nous disons que X_a suit une loi normale ou Gaussienne de paramètre aléatoire $\mu \in R$ et $\sigma \in R^{+*}$ (voir Figure 19), noté $X_a \mathcal{N}(\mu, \sigma^2)$, si elle admet une densité de probabilités (pdf) de la forme :

$$f_{\mathcal{N}}(\mu, \sigma^2)(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-1/2\left(\frac{x-\mu}{\sigma}\right)^2\right), \quad \forall x \in \mathbb{R} \quad (24)$$

Paramètres	Description	Formule	Gamme
BI	La luminosité est la moyenne des valeurs de gris de toutes les bandes pour chaque pixel/segment.	$\frac{TM1+\dots+TM7}{7}$	$[C_k^{min}, C_k^{max}]$, C_k^{min} : la plus faible valeur d'intensité possible et C_k^{max} : la valeur d'intensité la plus brillante possible
MB	La valeur moyenne de la bande spectrale bleue.	-	$[C_k^{min}, C_k^{max}]$
MR	La valeur moyenne de la bande spectrale rouge.	-	$[C_k^{min}, C_k^{max}]$
MG	La valeur moyenne de la bande spectrale verte.	-	$[C_k^{min}, C_k^{max}]$
MN	La valeur moyenne de la bande spectrale proche infrarouge.	-	$[C_k^{min}, C_k^{max}]$
Ctr	Le contraste est la somme de la variance des carrés.	$\sum_{i,j=0}^{N-1} P_{i,j}(i-j)^2$, $P_{i,j}$: la matrice de co-occurrence, i et j : les niveaux de gris, N : le nombre de niveaux de gris	$[0, 90]$
Ent	Quantifie le caractère aléatoire de la répartition de l'intensité du niveau de gris .	$\sum_{i,j=0}^{N-1} (-\ln P_{i,j})$	$[0, 90]$
ASM	Mesure l'uniformité de la texture.	$\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} P(i,j)^2$	$[0, 90]$
RF	Mesure l'écartement des objets	$a/l_{max} \times l_{min}$, a : zone du polygone, l_{max} : longueur du grand axe, and l_{min} : longueur du petit axe	$[0, 1]$
EF	Explique comment un objet correspond à une ellipse.	-	$[0, 1]$
SI	L'indice de forme décrit la lisibilité de la limite d'un objet.	$p/(4 \times \sqrt{a})$, p : périmètre du polygone	$[1, \infty]$
D	La densité est définie comme le rapport entre la surface de l'objet et son rayon [Navulur 2007].	$\sqrt{P_v}/(1 + \sqrt{\sigma_x^2 + \sigma_y^2})$, $\sqrt{P_v}$: diamètre d'un objet carré avec P_v pixels, et $\sqrt{\sigma_x^2 + \sigma_y^2}$: diamètre de l'ellipse ajustée sur le segment	$[0, \text{Selon la forme de l'objet de l'image}]$
A	La zone de l'objet identifié.	-	$[0, \text{taille de la scène}]$
NDVI	Le NDVI est basé sur le principe de la différence spectrale qui est basée sur une forte absorbance de la végétation dans la réflectance rouge et proche infrarouge du spectre [Bonn et Rochon 1996].	$\frac{TM4-TM3}{TM4+TM3}$	$[-1, 1]$
Tem	Décrit le degré de chaleur ou de froid d'un objet.	-	\mathbb{R}
Hum	L'humidité est définie comme le déficit de pression de vapeur.	$\frac{e}{e_s} \times 100\%$, e : La pression partielle de la vapeur d'eau réelle, et e_s : La pression de la vapeur d'eau saturée	\mathbb{R}

TABLE 4 – Description, formule et gamme de paramètres d'entrée du modèle.

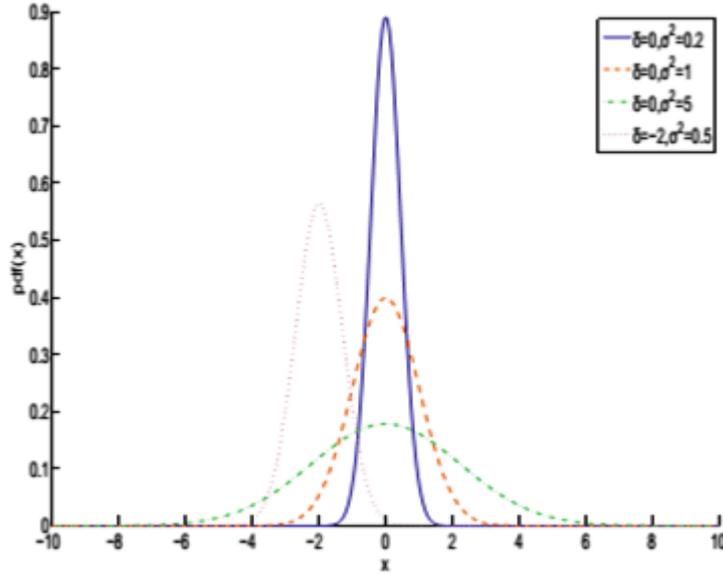


FIGURE 19 – Représentation de la fonction de densité Gaussienne de probabilités.

La variable μ est appelé la moyenne et la variable σ est l'écart-type (nous pouvons parler de variance avec σ^2).

Modélisation de paramètres épistémiques

Dans le cadre de la théorie des fonctions de croyance, chaque paramètre de type épistémique e est généralement modélisé par un ensemble des intervalles de confiance associé à une structure des fonctions de masse ou BPA (voir chapitre 2) comme suit :

$$\{[e_1^L, e_1^U]/(m(1), [e_2^L, e_2^U]/(m(1), \dots, [e_k^L, e_k^U]/(m(k), \dots | k \in (1, 2, \dots, P))\} \quad (25)$$

où P est le nombre total de sous-intervalles de e et $m(k)$ représente la valeur de BPA associée à le k -ième sous-intervalle $[E_K^L, E_K^U]$. Quand il ya des structures de BPA différentes, nous pouvons utiliser la règle de combinaison pour les intégrer dans une structure BPA conjointe comme $e_j/m(e_j)$ ($j \in [1, 2, \dots, P]$), où e_j est aussi un intervalle comme $[e_K^L, e_K^U]$ et $m(e_j)$ est la valeur de BPA associée à l'intervalle e_j .

Graphiquement, chaque paire peut être représentée par un rectangle dont la base correspond à l'intervalle X_K situé sur l'axe horizontal et dont la hauteur, sur l'axe des ordonnées, est égale à la masse ou BPA associée (Figure 20).

2.1.2 Etude de la corrélation entre les paramètres

La corrélation est une mesure de l'association entre deux paramètres. L'un des concepts les plus fondamentaux dans de nombreuses applications est le concept de corrélation. Si deux paramètres sont corrélés, cela signifie qu'il vous pouvez utiliser des informations sur un paramètre pour prédire les valeurs de l'autre paramètre. En fait, une forte corrélation entre deux paramètres suggère essentiellement que les deux paramètres peuvent capturer des informations similaires : les considérer simultanément peut aider à améliorer la décision finale.

La corrélation définie sur des ensembles conventionnels précis a été largement discu-

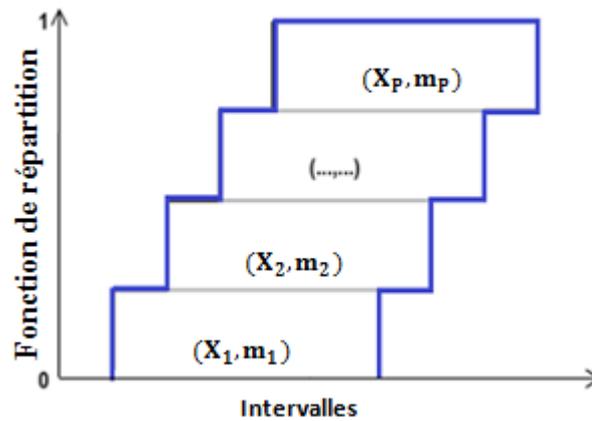


FIGURE 20 – Structure évidentielle.

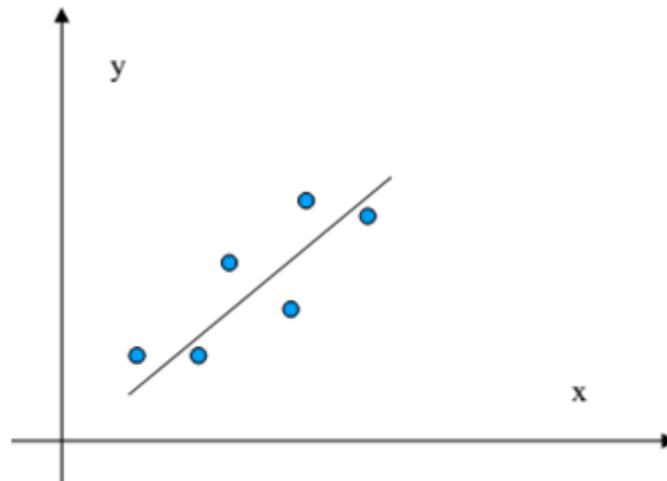


FIGURE 21 – Relation entre deux paramètres dans le cadre probabiliste.

tée dans les statistiques classiques. Cependant, le problème de la corrélation en présence des données incertaines n'est pas bien étudié dans la littérature. Dans ce qui suit, nous allons détailler le processus qui traite la corrélation entre les paramètres aléatoires et épistémiques.

Corrélation entre les paramètres aléatoires

Pour calculer le degré de corrélation entre deux paramètres aléatoires qui sont représentés par des distributions normales de probabilités, nous pouvons utiliser le coefficient de corrélation (Figure 21).

Le coefficient de corrélation représente une interrelation entre deux paramètres aléatoires, $a_1(X)$ et $a_2(Y)$, et généralement noté par ρ dans les statistiques traditionnelles. Dans les applications pratiques, le coefficient de corrélation ρ est généralement estimé par un coefficient de corrélation d'échantillon r qui est donné par :

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} \quad (26)$$

$f(x, y)$	1	2	3	$f_X(x)$
1	0.25	0.25	0	0.5
2	0	0.25	0.25	0.5
$f_Y(y)$	0.25	0.5	0.25	1

TABLE 5 – Exemple

où

$$\bar{x} = \mu = \sum_{i=1}^n x_i/n \quad \text{et} \quad \bar{y} = \mu = \sum_{i=1}^n y_i/n \quad (27)$$

- $-1 \leq r \leq 1$,
- si $r > 0$, X et Y sont positivement corrélés,
- si $r < 0$, sont négativement corrélés, et
- si $r = 0$, X et Y ne sont pas corrélés.

Exemple 1. Supposons que X et Y sont deux paramètres aléatoires qui ont des fonctions de distributions de probabilités normales conjointes (voir Tableau 5).

Alors, $\mu_X = 3/2$, $\mu_Y = 2$, $\sigma_X = 1/2$, et $\sigma_Y = \sqrt{1/2}$.

Quel est le coefficient de corrélation de X et Y ? Avant de répondre à cette question nous allons d'abord calculer la covariance entre les deux paramètres aléatoires X et Y . Si X et Y sont des paramètres aléatoires avec un support conjoint S , alors la covariance de X et Y est définie par :

$$\begin{aligned} cov(X, Y) &= \sigma_{XY} = \sum_{(x,y) \in S} (x - \mu_x)(y - \mu_y)f(x, y) \\ &= (1 - 3/2)(1 - 2)(1/4) + (1 - 3/2)(2 - 2)(1/4) + (1 - 3/2)(3 - 2)(0) \\ &\quad + (2 - 3/2)(1 - 2)(0) + (2 - 3/2)(2 - 2)(1/4) + (2 - 3/2)(3 - 2)(1/4) \\ &= 1/8 + 0 + 0 + 0 + 0 + 1/8 = 1/4 \end{aligned} \quad (28)$$

Le coefficient de corrélation est alors égale à :

$$r = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{1/4}{(1/2)(\sqrt{1/2})} = 0.71 \quad (29)$$

D'après ce résultat, nous pouvons dire que les deux paramètres aléatoires X et Y ont une corrélation positive forte.

Corrélation entre paramètres épistémiques

Comme nous l'avons déjà mentionné, les paramètres épistémiques sont généralement modélisés par des intervalles de confiances. Supposons $IX_i = [x_i^L, x_i^U]$ les intervalles observés pour le paramètre épistémique $e_1(X)$, où x_i^L et x_i^U sont respectivement la borne inférieure et la borne supérieure pour X_i . $IY_i = [y_i^L, y_i^U]$ les intervalles observés pour le paramètre épistémique $e_2(Y)$, où y_i^L et y_i^U sont respectivement la borne inférieure et la borne supérieure pour Y_i . Dans ce qui suit, nous allons définir les critères pour évaluer le coefficient de corrélation pour des intervalles (voir Figure 22).

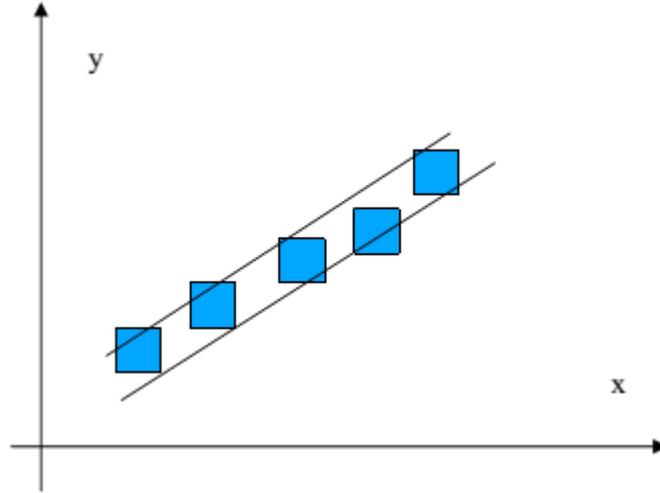


FIGURE 22 – Relation entre deux paramètres représentés par des intervalles.

Définition : Soit r_{Intj} le coefficient de corrélation entre des intervalles (IX, IY) . Etant donné $j = 1, 2, 3, \dots$ soit $\{(x_{ij}, y_{ij}) | x_{ij} \in IX_i, y_{ij} \in IY_i \text{ et } i = 1, 2, \dots, n\}$ un ensemble de données d'échantillon de (IX, IY) . Le coefficient de corrélation entre les intervalles (IX, IY) est donné par

$$r_{Intj} = \frac{\sum_{i=1}^n (x_{ij} - \bar{x}_j)(y_{ij} - \bar{y}_j)}{\sqrt{\sum_{i=1}^n (x_{ij} - \bar{x}_j)^2} \sqrt{\sum_{i=1}^n (y_{ij} - \bar{y}_j)^2}} \quad (30)$$

où

$$\bar{x}_j = \sum_{i=1}^n x_{ij}/n \quad \text{et} \quad \bar{y}_j = \sum_{i=1}^n y_{ij}/n \quad (31)$$

En effet, le coefficient de corrélation entre des intervalles doit être un intervalle. Donc, soient $r^L = \inf_j \{r_{Intj} | j = 1, 2, 3, \dots\}$ et $r^U = \sup_j \{r_{Intj} | j = 1, 2, 3, \dots\}$, alors le coefficient de corrélation des intervalles de (IX, IY) est donné par $r = [r^L, r^U]$. D'après l'inégalité de Cauchy-Schwarz [Dunn 2005], $r = [r^L, r^U]$ est un sous-intervalle de $[-1, 1]$ donné par l'équation (32)

$$\left(\sum_{i=1}^n (x_{ij} - \bar{x}_j)(y_{ij} - \bar{y}_j) \right)^2 \leq \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2 \sum_{i=1}^n (y_{ij} - \bar{y}_j)^2 \quad (32)$$

pour $x_{ij} \in IX_i, y_{ij} \in IY_i, i = 1, 2, \dots, n$ et $j = 1, 2, 3, \dots$

$$r_{Intj}^2 = \frac{(\sum_{i=1}^n (x_{ij} - \bar{x}_j)(y_{ij} - \bar{y}_j))^2}{\sum_{i=1}^n (x_{ij} - \bar{x}_j)^2 \sum_{i=1}^n (y_{ij} - \bar{y}_j)^2} \leq 1 \quad (33)$$

alors $-1 \leq r_{Intj}^2 \leq 1$ pour chaque j et $-1 \leq r^L \leq r^U \leq 1$. Par conséquent, $r = [r^L, r^U] \subseteq [-1, 1]$. Selon les opérations des intervalles pour l'inégalité précédente, nous avons les définitions suivantes. Soit $r = [r^L, r^U]$ un coefficient de corrélation des intervalles

- si $r^U > r^L \geq 0$, alors il y a une corrélation positive entre IX et IY
- si $r^U < r^L \leq 0$, alors il y a une corrélation négative entre IX et IY
- si $r^U = r^L = 0$, alors il n'a pas une corrélation entre IX et IY

– si $r^L < 0 < r^U$, alors nous ne pouvons pas juger si IX et IY ont des corrélations

Dans le cadre de la théorie des fonctions de croyance, les paramètres épistémiques sont modélisés par des intervalles avec des fonctions de masse associées. Denoeux et Masson [Masson *et al.* 2003, Hebert *et al.* 2003, Denoeux *et al.* 2005] ont proposé des méthodes permettant d'étendre la notion de corrélation à des paramètres de types épistémiques. Ils ont étendu le coefficient de corrélation au cas des nombres flous (*Fuzzy correlation coefficient*, en anglais). La corrélation entre des paramètres représentés par des ensembles flous a été largement étudiée dans la littérature [Chiang et Lin 1991, Gershenkorn et Manko 1991, Yu 1993, Bustince et Burillo 1995, Hong et Hwang 1995, Hsu et Wu 2010]. Dans le cadre possibiliste, Carlsson et Fullér [Carlsson et Fullér 2001] et Carlsson *et al.* [Carlsson *et al.* 2005] ont proposé un coefficient de corrélation (*Possibilistic correlation coefficient*, en anglais) basé sur le nombre flou. La section 2.1.3 présentera la méthode proposée pour calculer le coefficient de corrélation entre les intervalles dans le cadre de la théorie des fonctions de croyance.

2.1.3 Propagation de l'incertitude de paramètres

La propagation de l'incertitude est un moyen d'estimer l'incertitude dans les résultats finaux d'un modèle de COS due à l'incertitude des paramètres d'entrée de ce modèle. Étant donné un modèle de changement (f) dont les paramètres sont incertains de types aléatoires notés a et de types épistémiques notés e , la sortie du modèle peut être obtenue comme suit :

$$Y = f(a, e) \quad (34)$$

Comme nous avons indiqué dans la section 2.1.1, chaque paramètre de type épistémique e , est exprimé par des intervalles avec des fonctions de masse ou BPA associées $\{[e_1^L, e_1^U]/(m(1)), [e_2^L, e_2^U]/(m(2)), \dots, [e_k^L, e_k^U]/(m(k)), \dots | k \in (1, 2, \dots, P)\}$. Chaque paramètre de type aléatoire a est représenté par une distribution normale de probabilités a (μ, σ). Alors que, la propagation des incertitudes via la théorie des fonctions de croyance, peuvent prendre seulement comme entrées des intervalles avec BPA associées (c.à.d. une structure d'intervalles focaux) parce qu'elle a proposé spécifiquement pour modéliser des incertitudes épistémiques. Pour cela, il est important de mettre l'ensemble de paramètres du modèle dans un cadre commun. Pour résoudre ce problème, nous avons proposé de transformer les distributions normales de probabilités de paramètres aléatoires en des intervalles avec des fonctions de masse associées.

Dans la littérature, plusieurs auteurs ont proposé des approches pour transformer des fonctions de probabilité en des fonctions de croyance [Fiche *et al.* 2003, Caron *et al.* 2008, Doré *et al.* 2011, Hu *et al.* 2011]. Sur la base de l'approche de Hu *et al.* [Hu *et al.* 2011], une densité de probabilité normale de moyenne μ et d'écart-type σ peut être tronquée approximativement à $[\mu - \xi\sigma, \mu + \xi\sigma]$. Ainsi, nous pouvons discrétiser l'intervalle approximatif en N sous-intervalles $[a_i^L, a_i^U]$, $i \in [1, 2, \dots, N]$, et pour chaque sous-intervalle la valeur de fonction de masse est définie par l'équation suivante :

$$m(a_i) = \int_{a_i^L}^{a_i^U} f(x) dx, i \in [1, 2, \dots, N]. \quad (35)$$

où X_{a_i} est défini comme $\{a_i | x \in [a_i^L, a_i^U]\}$ et $f(x)$ est la fonction de distribution de densité de probabilité de x .

$Y = A_e + B_a$	[2,4], 0.5	[3,5], 0.5
[1,6], 0.25	[3,10], 0.125	[4,11], 0.125
[5,9], 0.5	[7,13], 0.25	[8,14], 0.25
[7,10], 0.25	[9,14], 0.125	[10,15], 0.125

TABLE 6 – L'ensemble de BPA pour la sortie $Y = A_e + B_a$ obtenu au moyen de produit cartésien entre les paires interval-masse $\{([1, 6], 0, 25), ([5, 9], 0, 5), ([7, 10], 0, 25)\}$ de A_e et $\{([2, 4], 0, 5), ([3, 5], 0, 5)\}$ de B_a sous l'hypothèse de l'indépendance.

Pour un paramètre aléatoire, les valeurs de BPA équivalentes dans des intervalles spécifiés sont égales à la zone sous la pdf. Après avoir obtenu les structures de BPA de tous les paramètres incertains (aléatoires et épistémiques), nous pouvons les propager à travers un modèle. La propagation de l'incertitude est effectuée par un produit cartésien [Balch 2012]. En général, la méthode de propagation du produit cartésien pour des structures évidentielles donne des limites correctes sur la probabilité lors de la propagation de deux variables avec une relation complètement inconnue. Pour définir la structure de BPA conjointe, cette méthode est synthétisée comme suit :

$$C = a \times e = \{c_{ij} = a_i \times e_j\} \quad (36)$$

où C désigne l'ensemble cartésien de tous les paramètres incertains et c_{ij} est l'élément de C .

Cas d'indépendance de paramètres

Lorsque deux paramètres incertains, e et a , sont non corrélés (indépendants), la BPA conjointe pour c_{ij} est définie en multipliant la BPA de a_i au BPA de e_j .

$$m(c_{ij}) = m(a_i) \times m(e_j) \quad (37)$$

L'élément focal c_{ij} est inclus par le cadre de discernement conjoint, et sa BPA est égale à la multiplication des BPA marginales correspondantes.

Exemple 2. Supposons que nous avons un modèle simple à deux paramètres incertains i.e. $Y = f(X)$, $X = \{A_e, B_a\}$. A_e représente l'indice de forme d'un objet urbain dans une zone donnée qui est considérée comme épistémique et B_a représente l'indice de luminosité de même objet qui est représenté comme aléatoire. Supposons que A_e est caractérisé par la structure évidentielle suivante $\{([1, 6], 0, 25), ([5, 9], 0, 5), ([7, 10], 0, 25)\}$ et que B_a est caractérisé, après transformation de sa distribution de probabilité, par la structure évidentielle suivante $\{([2, 4], 0, 5), ([3, 5], 0, 5)\}$. Les fonctions de plausibilité et de croyance cumulatives pour ces deux entrées sont présentées dans les graphiques de la Figure 23. Le Tableau 6 rapporte l'ensemble de BPA de la sortie $Y = A_e + B_a$ construite (sous l'hypothèse de l'indépendance entre A_e et B_a) au moyen du produit cartésien. Les fonctions de plausibilité et de croyance cumulatives pour la sortie sont affichées sur le graphique à droite de la Figure 23.

Cas de dépendance de paramètres

En présence de paramètres corrélés dans un modèle, plusieurs travaux ont été développés pour déterminer la BPA conjointe pour un élément focal c_{ij} [Ferson *et al.* 2004]. Dans notre cas, nous allons proposer un nouveau modèle de la théorie des fonctions de croyance qui prend en compte la corrélation entre les paramètres basé sur le modèle

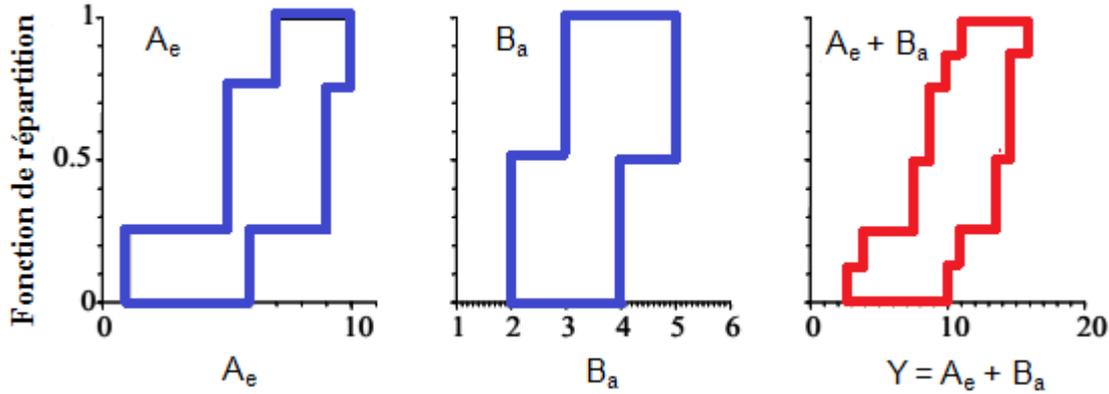


FIGURE 23 – Convolution entre les structures évidentielles sans considération de la corrélation entre les deux paramètres A_e et B_a .

ellipsoïdal [Luo *et al.* 2008, Jiang *et al.* 2013]. Ce modèle est proposé initialement pour l'analyse de l'incertitude non probabiliste. Dans cette thèse, le modèle ellipsoïdal est étendu pour traiter des paramètres évidentiels corrélés. A cet effet, un ellipsoïde multi-dimensionnel est construit en faisant en sorte que toutes les réalisations possibles des paramètres évidentiels inter-corrélés N -dimensionnels tombent dans un cadre de discernement (*Frame of discernment*, en anglais) conjoint comme le montre l'équation suivante (voir Algorithme 1) :

$$\Omega = \{\mathbf{X} | (\mathbf{X} - \mathbf{X}^c)^T \mathbf{G} (\mathbf{X} - \mathbf{X}^c) \leq 1\} \quad (38)$$

où le centre ellipsoïdal X^c est obtenu par les cadres de discernement marginaux comme suit :

$$X_m^c = \frac{X_m^L + X_m^R}{2}, m = 1, 2, \dots, N \quad (39)$$

où $X_m \in c_{ij}$ sont les paramètres évidentiels (paramètres aléatoires et épistémiques). La matrice de caractéristique symétrique \mathbf{G} détermine la taille et l'orientation de l'ellipsoïde, reflétant le degré et la manière de corrélation entre les paramètres évidentiels. Évidemment, nous n'assignons des probabilités de croyance qu'aux éléments c_{ij} qui tombent partiellement ou totalement dans le modèle ellipsoïde. Ainsi, une BPA conjointe est formulée comme suit :

$$m(c_{ij} \cap \Omega) = \frac{m(a_i) \times m(e_j)}{S}, c_{ij} \cap \Omega \neq \emptyset \quad (40)$$

où S est un facteur de normalisation pour que les BPA totales de m sont égales à 1, ce qui est donné par

$$S = \sum_{c_{ij} \cap \Omega \neq \emptyset} m(c_{ij}) \quad (41)$$

Algorithme 1. Construction du modèle ellipsoïdal

Soit t échantillons expérimentaux $X^{(r)}, r = 1, 2, \dots, t$ pour les N paramètres évidentiels. Chaque échantillon est un vecteur N -dimensionnel.

Le modèle ellipsoïdal est établi comme suit :

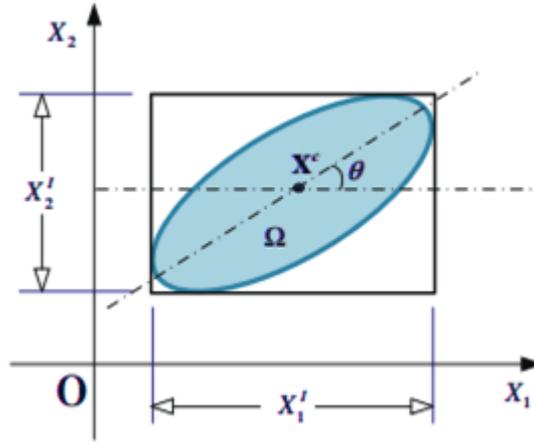


FIGURE 24 – Modèle ellipsoïdal basé sur le cadre de discernement conjoint.

1. Prendre une paire de paramètres évidentiels X_m et X_n ($m \neq n$) à la fois à partir de l'ensemble de paramètres incertains.
2. Extraire les valeurs de X_m et X_n à partir de t échantillons expérimentaux et construire un ensemble d'échantillons bivariant correspondant $\left(X_m^{(r)}, X_n^{(r)}\right)$, $r = 1, 2, \dots, t$.
3. Créer une ellipse minimale enveloppant les échantillons bivariants obtenus et obtenir l'angle de rotation θ correspondant comme représenté sur la Figure 24.
4. Calculer la covariance (Cov) et le coefficient de corrélation (r) des deux paramètres incertains X_m et X_n en fonction de la valeur de θ :
 - $Cov(X_m, X_n) = \frac{\tan(\theta)}{1-\tan^2(\theta)}((X_m^w)^2 - (X_n^w)^2)$ où $X_m^w = \frac{X_m^L + X_m^R}{2}$ et $X_n^w = \frac{X_n^L + X_n^R}{2}$ représentent les rayons de X_m et X_n , respectivement.
 - $r_{X_m X_n} = \frac{Cov(X_m, X_n)}{X_m^w X_n^w}$, $-1 \leq r_{X_m X_n} \leq 1$.
5. Répéter le processus ci-dessus pour toutes les paires de paramètres incertains jusqu'à obtenir un total de $N(N-1)/2$ covariances et coefficients de corrélation pour tous les paramètres.
6. Créer une matrice de covariance C sur la base des covariances calculées.
7. Le modèle ellipsoïdal peut être obtenu par :

$$\Omega = \{\mathbf{X} | (\mathbf{X} - \mathbf{X}^c)^T \mathbf{C}^{-1} (\mathbf{X} - \mathbf{X}^c) \leq 1\} \quad (42)$$

Exemple 3. Prenons les mêmes données que l'exemple 1, le paramètre épistémique 'indice de forme' A_e avec la structure de Dempster-Shafer suivante $\{([1, 6], 0, 25), ([5, 9], 0, 5), ([7, 10], 0, 25)\}$ et le paramètre aléatoire 'indice de luminosité' B_a avec la structure de Dempster-Shafer suivante $\{([2, 4], 0, 5), ([3, 5], 0, 5)\}$, et leur coefficient de corrélation est de 0,2. Supposons que nous voulons utiliser le modèle ellipsoïdal pour déterminer l'ensemble de BPA de la sortie $Y = A_e + B_a$. Les fonctions de plausibilité et de crédibilité cumulatives pour les deux entrées et la réponse du modèle Y tout en considérant les corrélations sont présentées dans les graphiques de la Figure 25.

Dans cet exemple, pour le modèle $Y = A_e + B_a$, nous n'avons pas obtenu des résultats de propagation très différents par rapport aux résultats obtenus dans l'exemple 1, parce que les deux paramètres A_e et B_a ont une faible corrélation ($r_{A_e B_a} = 0.2$).

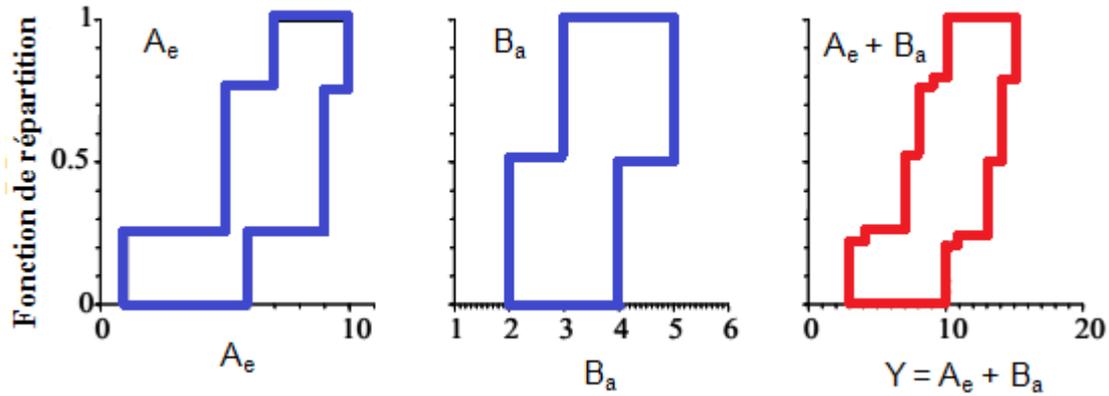


FIGURE 25 – Convolution entre les structures de Dempster-Shafer par le modèle ellipsoïdal avec considération de la corrélation entre les deux paramètres A_e et B_a .

Une fois que la structure de BPA conjointe de tous les paramètres incertains est définie, la propagation des incertitudes de paramètres à travers le modèle $Y = f(a, e)$ est effectuée et l'intervalle focal de la variable de sortie Y est calculé par :

$$[Y_{min}, Y_{max}] = [\min_{\mathbf{x} \in c_{ij}} f(\mathbf{X}), \max_{\mathbf{x} \in c_{ij}} f(\mathbf{X})] \quad (43)$$

La croyance Bel est obtenue en additionnant tous les BPA conjointes de l'ensemble d'éléments X , dans laquelle la réponse maximale Y_{max} est inférieure à la valeur limite (Y_{limite}). La plausibilité Pl est obtenue en additionnant les BPA conjointes de ces ensembles d'éléments, dans lequel Y_{limite} est supérieure à la réponse minimale Y_{min} .

2.2 Analyse d'incertitude de la structure du modèle

Le processus décrit dans la section 2.1 est proposé pour le traitement des incertitudes de paramètres en supposant que le modèle f est déterministe et bien établi. Dans notre cas, ce modèle reflète la structure d'un modèle de COS. Toutefois, dans certains cas, le modèle f peut être sujet à des incertitudes du fait des hypothèses ou des simplifications qui sont faites par des analystes au moment de la modélisation des systèmes terrestres. Ces incertitudes induites par les erreurs de modélisation qui viennent s'ajouter aux incertitudes sur les paramètres et résultent en un modèle non représentatif du système réel. Si les incertitudes de la structure du modèle ne sont pas considérées lors de l'optimisation, la solution obtenue pourra se révéler en pratique bien moins performante que prévue, voire ne pas répondre aux performances imposées. À l'heure actuelle, les incertitudes qui affectent la structure du modèle sont la plupart du temps négligées dans le processus de traitement des incertitudes, bien que selon plusieurs travaux [Droguett et Mosleh 2012] leur impact sur les résultats peuvent être important et parfois même plus important que l'impact des incertitudes de paramètres. Cette source d'incertitude devrait donc être prise en compte dans le processus décisionnel final. La définition des incertitudes de la structure du modèle a été donnée dans le chapitre 2.

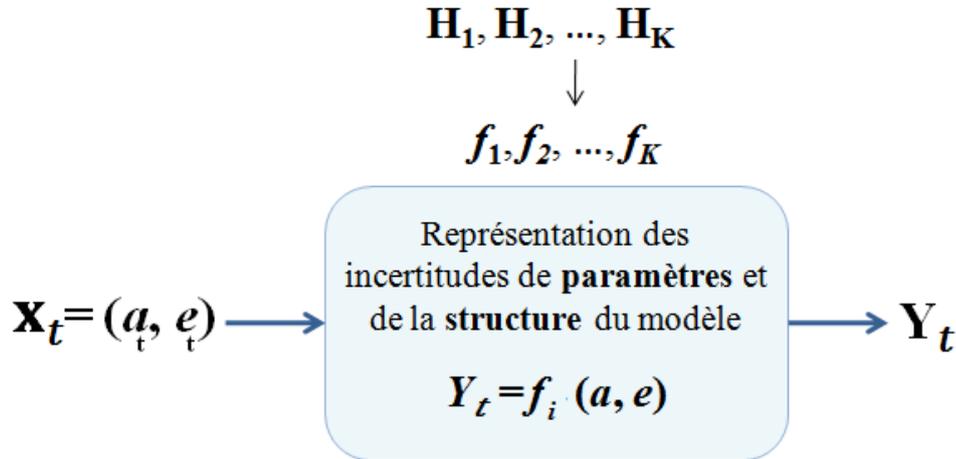


FIGURE 26 – Représentation de l'incertitude de la structure du modèle (H_i sont les hypothèses et f_i sont les différentes structures de modèle).

2.2.1 Caractérisation d'incertitude de la structure du modèle

En général, chaque structure d'un modèle de COS est représentée par un certain nombre d'hypothèses qui sont des décisions ou des jugements considérés par les analystes. D'une structure à une autre, les représentations sont différentes. Supposons qu'un ensemble de K hypothèses $\{H_i\}$, $1 \leq i \leq K$ sont données par K experts différents, alors pour chaque hypothèse H_i , un modèle de COS (noté f) est construit. Nous avons alors un ensemble de K structures différentes de ce modèle $\{f_i\}$, $1 \leq i \leq K$ (voir Figure 26).

Dans un modèle de COS, quand deux hypothèses H_1, H_2 sont données par deux experts différents, nous obtenons alors deux structures différentes f_1 et f_2 pour ce modèle. Pour chaque structure de modèle, les résultats de ce modèle Y peuvent être différents. Par exemple, la construction de la structure du modèle DINAMICA de COS est basée sur un ensemble d'étapes [Soares-Filho *et al.* 2002]. Chacune ayant ses propres paramètres :

- le nombre de pas de temps ;
- la matrice de transition, avec les taux fixés à l'intérieur de la phase. DINAMICA peut également tenir compte des hypothèses concernant les taux de transition distincts, en fonction de la partition et l'arrangement spatial spécifiques, en utilisant une carte de partition spatiale comme variable statique.
- une valeur de saturation éventuelle pour chaque type d'occupation des sols. Cette valeur est calculée en utilisant l'équation suivante :

$$Rate'(ij) = Rate(ij) \times (M - v)/(M + v) \quad (44)$$

où M est un pourcentage de type d'occupation des sols pris à un temps fixé, et v la superficie minimale dans laquelle un type d'occupation des sols demeurera.

- le temps de séjour minimum pour chaque type de transition avant qu'une cellule change son état ;
- les coefficients du modèle logistique sont appliqués pour calculer chaque probabilité de transition spatiale P_{ij} . Les probabilités de transition spatiale sont calculées pour chaque cellule dans la carte d'occupation des sols et pour chaque transition spécifiée. La méthode choisie utilise un modèle logistique sous la forme

de :

$$P\left(i \Rightarrow \frac{j(x, y)}{V}\right) = \frac{e^{G_{i \Rightarrow j}(V)}}{1 + \sum_i^j e^{G_{i \Rightarrow j}(V)}} \quad (45)$$

où V est un vecteur de k variables spatiales, mesuré à l'emplacement x, y et représenté par $V_{1xy}, V_{2xy}, \dots, V_{kxy}$. Chaque $G_{ij}(V)$ est calculée par l'équation suivante :

$$G_{ij}(V) = \beta_{0,ij} + \beta_{1,ij}V_{1xy} + \dots + \beta_{k,ij}V_{kxy} \quad (46)$$

- le **pourcentage de transitions** exécuté par chaque fonction de transition ainsi que les paramètres 'moyenne et variance' de la taille du patch de chaque type d'occupation des sols. Le modèle DINAMICA utilise deux fonctions de transition : Fonction *Expander* et Fonction *Patcher*. La combinaison de ces deux fonctions est montrée dans l'équation suivante :

$$Q_{ij} = r \times (\text{ExpanderFunction}) + s \times (\text{PatcherFunction}) \quad (47)$$

où Q_{ij} représente la quantité totale de transitions de type ij spécifiée par l'étape de simulation et r et s sont respectivement le pourcentage de transitions exécuté par chaque fonction, avec $r + s = 1$. Les deux fonctions de transition utilisent un mécanisme de sélection stochastique. L'algorithme appliqué consiste à lire d'abord la carte d'occupation des sols pour capter les cellules avec des probabilités plus élevées et ensuite les disposer dans un tableau de données. La sélection des cellules se fait de façon aléatoire (le mécanisme de choix stochastique interne peut être desserré ou serré en fonction du degré de randomisation désiré). L'algorithme de la fonction *Expander* est exprimé par l'équation suivante :

$$\begin{aligned} \text{If } n_j > 3 \text{ alors } P'(ij)(xy) &= P(ij)(xy) \\ \text{sinon } P'(ij)(xy) &= P(ij)(xy) \times (n_j)/4 \end{aligned} \quad (48)$$

où n_j correspond au nombre de cellules de type j se produisant dans une fenêtre 3×3 . Cette méthode garantit que la $P'(ij)$ maximale sera la $P(ij)$ originale, quand qu'un type de cellule i est entouré d'au moins 50% de cellules voisines de type j . La fonction *Patcher* utilise un dispositif qui recherche des cellules autour d'un emplacement choisi pour une transition combinée. Ceci se fait d'abord en choisissant la cellule centrale du nouveau patch et en sélectionnant un nombre spécifique de cellules autour de la cellule centrale, en fonction de leurs probabilités de transition P_{ij} . La taille des nouveaux patches et des marges d'expansion est établie selon une distribution de probabilité lognormal. Donc, il est nécessaire de spécifier les paramètres de cette distribution (la moyenne et la variance).

Le modèle DINAMICA est écrit en langage C^{++} orienté objet (accessible via Internet sur : www.csr.ufmg.br/dinamica/). Les étapes principales de ce modèle mentionnées ci-dessus sont basées sur un nombre d'hypothèses. Ces hypothèses peuvent être changées d'un expert à un autre ou d'une application à une autre, ce qui peut produire des incertitudes au niveau des résultats des COS. En plus, les simplifications et les équations mathématiques utilisées lors de la construction de ce modèle peuvent aussi conduire à des incertitudes.

2.2.2 Modélisation de l'incertitude de la structure du modèle

Dans la plupart des cas, l'incertitude représentant la structure d'un modèle de COS est une forme d'incertitude épistémique parce que nous ne savons pas si leurs constructions sont raisonnables et complètes. Elle serait une incertitude aléatoire seulement si la structure du modèle lui-même doit être changée dans le temps, dans l'espace ou entre les composants de certaines populations.

Dans la littérature [Alvin *et al.* 1998], les incertitudes de la structure du modèle sont généralement traitées par l'approche par pondération ou l'approche Bayésienne du modèle moyenné (*Bayesian Model Averaging*, en anglais) dans le cadre probabiliste. Dans ces approches, un ensemble de modèles plausibles est sélectionné selon les hypothèses alternatives données par des avis d'experts ou par des caractéristiques physiques alternatives du système. Des probabilités subjectives (i.e. des poids) représentant le niveau de croyance des analystes sont ensuite assignés aux modèles alternatifs. Pour chaque modèle, la propagation des incertitudes paramétriques associées à ce modèle est effectuée. Étant donné les poids, ces modèles sont agrégés pour calculer la distribution probabiliste à la sortie.

Dans [Soize 2010], Soize décrit une méthodologie, appelée approche probabiliste généralisée des incertitudes, permettant de prendre en compte de manière séparée les incertitudes de la structure du modèle et les incertitudes liées aux paramètres du système. Après la construction d'un modèle éléments finis nominal, un modèle numérique stochastique, tenant compte à la fois les deux sources d'incertitudes, est élaboré dans un contexte probabiliste.

Ladevèze *et al.* [Ladevèze *et al.* 2006] ont développé la théorie des méconnaissances (*Lack-Of-Knowledge (LOK) Theory*, en anglais) pour modéliser des incertitudes de types épistémiques représentant la structure d'un modèle. Cette théorie cherche à quantifier l'incertitude à partir de la différence entre le modèle numérique et la structure réelle. L'idée de cette théorie est de globaliser les incertitudes à l'échelle des sous-structures constitutives de la structure étudiée en les définissant comme des variables scalaires situées dans un intervalle dont les bornes sont des variables aléatoires.

Dans cette thèse, puisque les incertitudes de paramètres sont modélisées par la théorie des fonctions de croyance, nous utilisons cette théorie pour propager les incertitudes liées à la structure du modèle.

2.2.3 Propagation de l'incertitude de la structure du modèle

La propagation de l'incertitude de la structure du modèle est mise en oeuvre en combinaison avec la propagation de l'incertitude de paramètres. Dans cette section, comme les incertitudes de paramètres d'entrée sont propagées par la théorie des fonctions de croyance, nous utilisons cette technique dans ce cadre.

Supposons qu'un ensemble de structures du modèle $f_k, 1 \leq k \leq K$ représente une facette du phénomène physique réel. Pour chaque structure du modèle f_k , l'incertitude de paramètres se propage à travers ce modèle. En conséquence, nous avons un ensemble de représentations pour la variable de sortie Y . La différence entre ces représentations reflète la variation liée aux incertitudes de la structure du modèle.

Indice de forme		Indice de luminosité		Surface		NDVI		Température	
<i>Intervalle</i>	<i>BPA</i>	<i>Intervalle</i>	<i>BPA</i>	<i>Intervalle</i>	<i>BPA</i>	<i>Intervalle</i>	<i>BPA</i>	<i>Intervalle</i>	<i>BPA</i>
[1, 6]	0,25	[2, 4]	0.50	[0, 1200]	0.04	[-1,-0.5]	0.50	[0,10]	0.05
[5, 9]	0.50	[3, 5]	0.50	[1200, 2400]	0.09	[-0.5,0]	0.35	[10,20]	0.25
[7, 10]	0,25	–	–	[2400, 3600]	0.22	[0,0.5]	0.1	[20,30]	0.5
–	–	–	–	[3600, 4800]	0.65	[0.5,1]	0.05	[30,40]	0.20
Somme	1	Somme	1	Somme	1	Somme	1	Somme	1

TABLE 7 – Collection de paires intervalle-BPA de cinq paramètres (indice de forme, indice de luminosité, surface, NDVI et température).

Dans [Alvin *et al.* 1998, Ladevèze *et al.* 2006, Soize 2010], chaque représentation de la variable de sortie Y est caractérisée par une distribution de probabilité. Cependant, dans le cadre de la théorie des fonctions de croyance, elle est modélisée par une paire de fonctions de croyance et plausibilité. Ainsi, pour toutes les structures du modèle f_k ($1 \leq k \leq K$), nous avons un ensemble de paires de fonctions de croyance et de plausibilité de la variable de sortie Y , noté $\{[Bel_1(Y), Pl_1(Y)], [Bel_2(Y), Pl_2(Y)], \dots, [Bel_K(Y), Pl_K(Y)]\}$. La différence entre ces représentations reflète la variation associée à l'incertitude de la structure du modèle. Ces différentes représentations $[Bel_i(Y), Pl_i(Y)]$, $1 \leq i \leq K$ peuvent être combinées en une seule représentation. Par conséquent, la représentation de l'incertitude de la variable de sortie Y peut être obtenue par les formules suivantes :

$$Bel^*(Y) = \min(Bel_1(Y), Bel_2(Y), \dots, Bel_K(Y)) \quad (49)$$

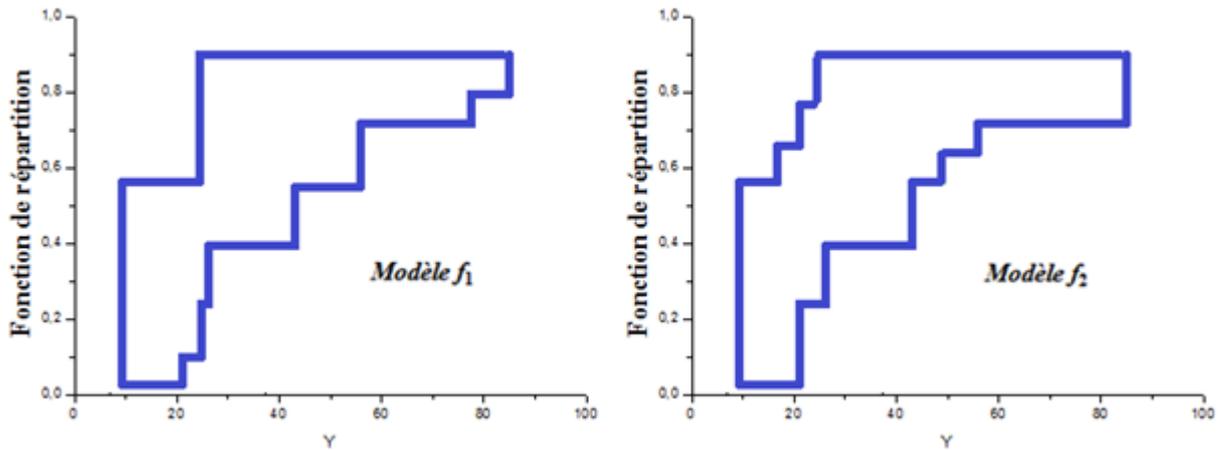
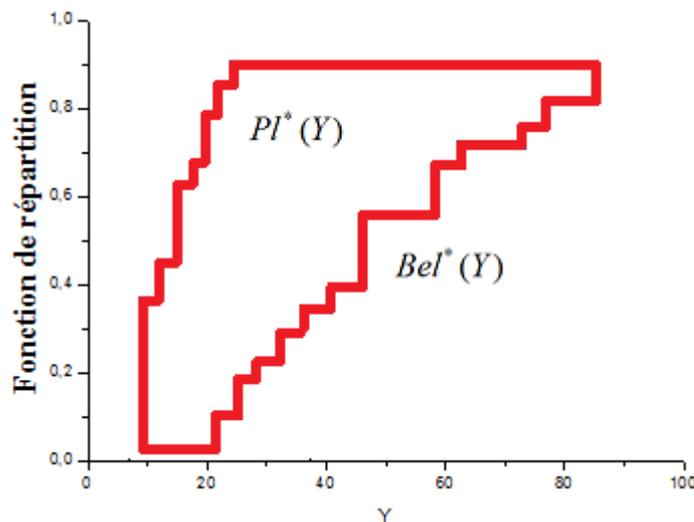
$$Pl^*(Y) = \max(Pl_1(Y), Pl_2(Y), \dots, Pl_K(Y)) \quad (50)$$

Les fonctions de croyance et de plausibilité $[Bel^*(Y), Pl^*(Y)]$ prennent en compte à la fois les incertitudes de paramètres et de la structure du modèle dans le résultat de sortie finale.

Exemple 4. Pour illustrer la propagation des incertitudes de la structure du modèle avec l'approche proposée, nous prenons les deux paramètres incertains étudiés dans les exemples 2 et 3 qui sont l'indice de forme et l'indice de luminosité, et en ajoutant 3 autres paramètres incertains tels que : la surface, le NDVI et la température. Le Tableau 7 rapporte l'ensemble de paramètres avec leurs intervalles et les BPA associées.

Dans cet exemple, nous proposons d'utiliser le modèle de COS DINAMICA en tenant compte des hypothèses modifiées du taux de transition de la structure originale du modèle étudié dans la section 2.2.1 sont modifiées. Ainsi, selon les hypothèses originales et les hypothèses modifiées nous avons deux structures différentes du modèle DINAMICA. Notons f_1 (respectivement f_2) le modèle correspondant à la première hypothèse H_1 (respectivement à la deuxième hypothèse H_2). Pour prendre en compte l'incertitude de la structure du modèle dans le résultat final, pour chaque modèle f_1 ou f_2 , les incertitudes associées aux cinq paramètres (indice de forme, indice de luminosité, surface, NDVI, et température) sont d'abord propagées.

La Figure 27 présente les fonctions de croyance et plausibilité cumulatives correspondant à chaque modèle. Ces fonctions sont ensuite propagées en utilisant la théorie des

FIGURE 27 – Structure évidentielle pour chaque modèle f_1 et f_2 .FIGURE 28 – Structure évidentielle de la variable de sortie Y avec propagation de l'incertitude relative aux paramètres et à la structure du modèle.

fonctions de croyance. La Figure 28 représente le résultat de la propagation des incertitudes relative à la structure du modèle en combinaison avec les incertitudes propagées de paramètres.

Figure 28 montre un écart important entre les fonctions de croyance et de plausibilité avec les résultats présentés dans la Figure 27. Ceci implique que l'effet combiné de l'incertitude de la structure du modèle et de paramètres conduit à une incertitude plus large à la sortie du modèle.

La distance entre les fonctions de croyance et de plausibilité $Bel^*(Y)$ et $Pl^*(Y)$ représente le degré ou l'effet de l'incertitude des paramètres et de la structure du modèle de COS. Dans cette étude, cette distance peut être calculée par la distance de Kolmogorov-Smirnov (KS) [Bain et Engelhardt 1991]. La distance de KS est une mesure utilisée dans le test statistique [Bain et Engelhardt 1991]. Elle est définie comme la différence maximale entre la CDF d'échantillonnage et la CDF hypothétisée. Cette distance mesure comment rapprocher la CDF d'échantillonnage à la CDF hypothétisée.

Dans l'approche proposée, nous adoptons la même idée pour mesurer la différence entre les distributions $Bel^*(Y)$ et $Pl^*(Y)$. La distance de KS entre ces deux distributions est calculée comme suit :

$$d_{KS} = \max_Y [Pl^*(Y)] - Bel^*(Y) \quad (51)$$

L'équation implique que la distance de KS est la différence maximale entre les deux courbes de $Bel^*(Y)$ et $Pl^*(Y)$.

3 Conclusion

Dans ce chapitre, nous avons proposé une démarche pour la représentation et la propagation des incertitudes de paramètres et de la structure du modèle dans le cadre de la théorie des fonctions de croyance pour les applications des modèles de COS.

Nous avons proposé dans notre démarche de modéliser les incertitudes aléatoires de paramètres par des distributions normales de probabilités et les incertitudes de types épistémiques par des intervalles avec des fonctions de masse associées (intervalles focaux). Les corrélations entre ces différents paramètres sont également étudiées. En ce qui concerne la propagation des incertitudes de paramètres, la procédure a été présentée et illustrée par un exemple simple. La propagation des incertitudes dans ce cas consiste à propager des intervalles focaux de paramètres aléatoires et épistémiques tout en considérant la corrélation qui existe entre ces paramètres. La méthode proposée pour ce faire est la théorie des fonctions de croyance sur la base d'un modèle ellipsoïdal.

Après l'étude de la propagation des incertitudes de paramètres, nous nous sommes intéressés à la prise en compte des incertitudes de la structure du modèle. Nous avons commencé cette étude par la caractérisation des incertitudes de modèle selon le nombre de sources d'incertitude, dans les applications des modèles de COS. La propagation des incertitudes de la structure du modèle, quant à elle, est effectuée d'abord par la propagation des incertitudes de paramètres étant donné une structure du modèle. A l'issue de cette étape, nous obtenons un ensemble de représentations par des fonctions de croyance et de plausibilité pour la variable de sortie. La différence entre ces représentations reflète les impacts des incertitudes de la structure du modèle. Les incertitudes de la structure du modèle sont ensuite incluses dans le résultat final par la combinaison de ces représentations avec la théorie des fonctions de croyance.

Dans ce chapitre, nous avons détaillé le processus de la propagation des incertitudes de paramètres et de la structure du modèle. Les résultats obtenus à l'issue de ce processus sont présentés pour les étapes du processus de la réduction des incertitudes qui seront étudiées dans le chapitre suivant.

Chapitre IV

Approche proposée pour la réduction des incertitudes dans les modèles de COS

Sommaire

1	Introduction	66
2	Approche de réduction des incertitudes	66
2.1	Analyse de sensibilité	67
2.1.1	Choix de la méthode d'analyse de sensibilité	68
2.1.2	Quantification des incertitudes de paramètres	69
2.1.3	Quantification des incertitudes de la structure du modèle	70
2.2	Estimation	73
2.3	Evaluation de la performance des modèles de COS	75
3	Conclusion	76

1 Introduction

Dans le chapitre précédent, nous avons étudié les étapes du processus de propagation des incertitudes, à savoir la modélisation et la propagation des incertitudes avec la théorie des fonctions de croyance. Dans ce chapitre, nous allons détailler le processus de réduction des incertitudes dans les modèles de COS.

Ainsi, dans ce chapitre, nous présentons les étapes principales du processus proposé de réduction des incertitudes à savoir l'analyse de sensibilité, l'estimation et l'évaluation de performance [Ferchichi *et al.* 2017b].

2 Approche de réduction des incertitudes

Comme nous l'avons indiqué dans le Tableau 3 de chapitre 2, de nombreuses approches ont été proposées pour réduire l'incertitude dans les modèles de COS. Parmi ces approches de réduction, plusieurs méthodes d'analyse de sensibilité ont été développées pour comprendre l'impact des sources d'incertitude sur les sorties du modèle [Verburg *et al.* 2013, Boulila *et al.* 2014, Tayyebi *et al.* 2014, Mondal *et al.* 2015, Krüger et Lakes 2015, Ferchichi *et al.* 2016]. La compréhension des effets des incertitudes sur les résultats des modèles de simulation est essentielle pour une bonne utilisation de ces modèles. D'autre part, la tâche de l'optimisation en présence de grand nombre de paramètres incertains est complexe et prend beaucoup de temps. A ce niveau, l'analyse de sensibilité, qui implique la variation des paramètres du modèle et l'évaluation de la façon dont les résultats de prédiction changent avec ces variations, s'est révélée efficace et robuste pour sélectionner des variables qui ont un effet important sur la sortie des modèles de COS [Boulila *et al.* 2014, Ferchichi *et al.* 2016] et a été appliquée dans de nombreuses études [Mishra et Susaki 2014, Mousivanda *et al.* 2014, Elhag 2014]. Des travaux antérieurs ont montré que l'analyse de sensibilité peut réduire l'incertitude du modèle [Hamby 1994], éviter le sur-paramétrage [Du *et al.* 2013], aider à comprendre la structure du modèle [Jakeman *et al.* 2006, Munoz-Carpena *et al.* 2007], et générer des résultats de modèles plus précis. Dans [Saltelli *et al.* 2004], Saltelli *et al.* ont affirmé qu'il serait difficile de concevoir ou d'améliorer un modèle sans faire l'analyse de sensibilité.

Dans cette thèse, l'application des méthodes d'analyse de sensibilité tente de trouver les paramètres et les structures les plus importants pour la calibration de modèle du COS. Ceux qui ne sont pas importants peuvent être préfixés en tant que valeur constante (voir Figure 29).

L'analyse de sensibilité de paramètres du modèle permet de tester les effets des paramètres sur les résultats, tandis que l'analyse de sensibilité de la structure du modèle permet de modifier la structure et d'identifier les facteurs des structures possibles qui affectent la robustesse des résultats (varier la structure du modèle et voir l'impact sur les résultats et les compromis entre les choix) [Ferchichi *et al.* 2016]. La connaissance des variables les plus pertinentes peut augmenter les chances de trouver un modèle satisfaisant.

Après avoir effectué l'analyse de sensibilité, l'estimation des paramètres et des structures du modèle les plus influents est souhaitée afin d'optimiser leurs valeurs qui abou-

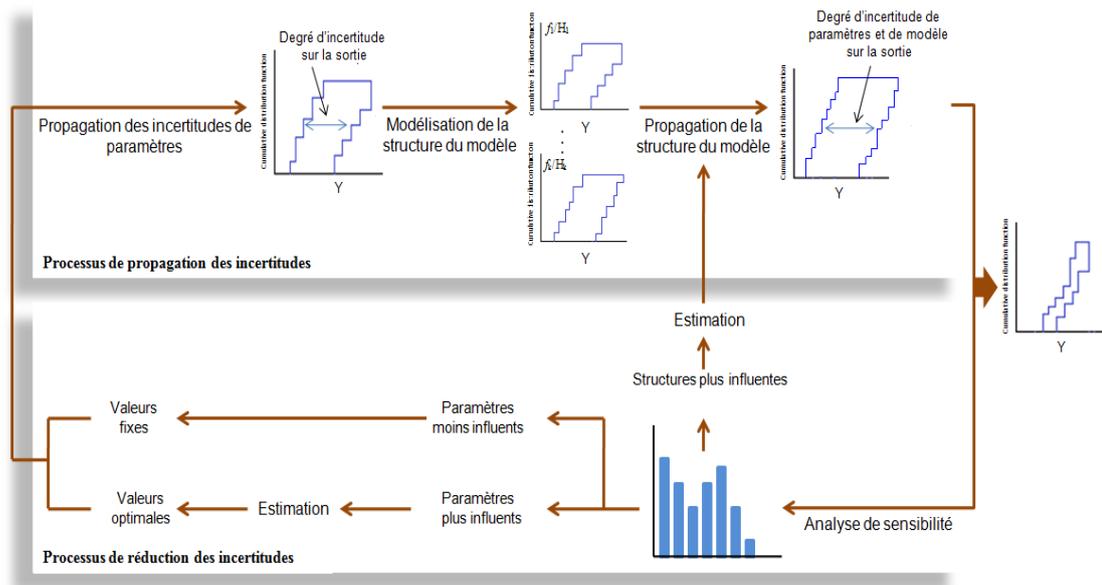


FIGURE 29 – Architecture de l'approche de réduction des incertitudes.

tissent à une représentation optimale du modèle du système d'occupation du sol étudié (voir Figure 4-1).

Dans ce qui suit, nous allons détailler le processus de l'approche proposée de réduction des incertitudes dans les modèles de COS [Ferchichi *et al.* 2017b]. Dans la première partie, nous commençons par présenter l'approche proposée de l'analyse de sensibilité pour savoir quelles sont les sources d'incertitude qui contribuent le plus à l'incertitude de la sortie du modèle. Cette analyse permet de hiérarchiser les sources d'incertitude de l'étude, c'est à dire de cibler les paramètres et les structures du modèle les plus influents sur la sortie et nécessitant le plus de précision. Dans la deuxième partie, une méthode d'estimation est proposée pour estimer davantage les sources sélectionnées à partir de données et les principaux critères d'évaluation que nous pouvons utiliser pour évaluer la performance des modèles de COS. Nous illustrons chaque procédure du processus de réduction par un exemple simple.

2.1 Analyse de sensibilité

L'analyse de sensibilité permet de comprendre le comportement du système à modéliser, et de déterminer les sources d'incertitude du modèle qui nécessiteraient une plus grande précision, ou une amélioration dans leur traitement [Saltelli *et al.* 2004]. Dans ce travail, l'analyse de sensibilité est utilisée pour identifier les paramètres incertains et les structures les plus importants dans les modèles de COS. L'analyse de sensibilité est utilisée ici comme un moyen efficace pour réduire les incertitudes. Comme nous avons déjà présenté dans le chapitre précédent de propagation de l'incertitude, les incertitudes de paramètres et de la structure du modèle sont propagés à travers le modèle f à la sortie de model Y par la théorie des fonctions de croyance. Les résultats de la propagation sont des fonctions de croyance et de plausibilité (Bel^* et Pl^*). L'écart entre ces deux distributions indique l'effet de l'incertitude de deux sources. Si l'écart est trop grand, il sera difficile de prendre des décisions.

2.1.1 Choix de la méthode d'analyse de sensibilité

Plusieurs méthodes d'analyse de sensibilité ont été développées dans la littérature [Pianosi *et al.* 2016], y compris les méthodes de criblage [Sanchez-Canales *et al.* 2012], les méthodes d'analyse différentielle [Bettemier 2010, Haihua *et al.* 2015], les méthodes basées sur la variance [Zielinskaa et Sunb 2010], les méthodes basées sur l'échantillonnage [Helton *et al.* 2006] et les méthodes basées sur l'entropie [Li *et al.* 2014]. Parmi celles-ci, les méthodes basées sur la variance sont les plus populaires. Ces dernières proviennent de la décomposition de la variance de la sortie du modèle. La méthode FAST (*Fourier Amplitude Sensitivity Test*) [Xiao *et al.* 2013, Wei et Hua 2013], les mesures d'importance [Homma et Saltelli 1996], la méthode de Saltellis [Song *et al.* 2012] et les indices de Sobol [Gerardino-Neira *et al.* 2008] appartiennent à ce type de méthodes. Généralement, ces méthodes ont été développées pour quantifier des incertitudes aléatoires qui sont principalement représentées par des distributions probabilistes. L'analyse de sensibilité avec une représentation probabiliste est généralement nommée analyse de sensibilité probabiliste.

Cependant, dans le cadre de notre travail, nous sommes intéressés par la théorie des fonctions de croyance. Dans la littérature, des méthodes non probabilistes d'analyse de sensibilité ont été proposées [Hagashi et Klir 1983, Klir et Wiermann 1998, Sengupta et Pal 2000, Ferson et Tucker 2006, Ivanov et Tokmakian 2011, Datta 2011, Chutta *et al.* 2013, Ferchichi *et al.* 2016]. Dans [Hagashi et Klir 1983, Klir et Wiermann 1998, Datta 2011, Ivanov et Tokmakian 2011, Chutta *et al.* 2013], les auteurs ont utilisé la mesure de Hartley-like d'analyse de sensibilité pour quantifier des incertitudes représentées dans le cadre de la théorie des ensembles flous. Dans [Ferchichi *et al.* 2016], Ferchichi *et al.* ont proposé une méthode basée sur la mesure de Hartley-like pour tester l'impact des incertitudes dans le cas où l'incertitude de la sortie du modèle est représentée par une distribution possibiliste. De nombreuses études ont confirmé la robustesse, l'efficacité et la facilité de l'utilisation de la stratégie du pincement (*pinching strategy*, en anglais) pour générer une analyse de sensibilité dans le cadre de la théorie des fonctions de croyance et ceci dans plusieurs domaines [Ferson et Tucker 2006, Helton *et al.* 2006, Hall 2006, Oberguggenberger *et al.* 2009, Ali *et al.* 2012]. La stratégie du pincement est proposée par Ferson et Tucker [Ferson et Tucker 2006] et elle est analogue à la variance telle qu'elle est utilisée dans les méthodes basées sur la variance [Zielinskaa et Sunb 2010].

Le choix de la stratégie du pincement est argumenté par le fait que cette méthode a le potentiel de quantifier des incertitudes représentées par des fonctions de croyance et de plausibilité en présence des incertitudes aléatoires/épistémiques et des corrélations entre les paramètres. La stratégie du pincement utilise des outils numériques pour concevoir des solutions globales d'amélioration. L'objectif de cette stratégie consiste à estimer la valeur de l'information empirique supplémentaire. La valeur des informations supplémentaires requises peut être prédite en comparant l'incertitude du modèle avant et après une opération appelée 'pincement'.

Ainsi, la quantification des incertitudes à travers la stratégie du pincement sera traitée à la fois pour les paramètres et la structure du modèle dans les sections suivantes.

2.1.2 Quantification des incertitudes de paramètres

Comme nous l'avons vu dans le chapitre 3, la propagation des incertitudes de paramètres est effectuée par les fonctions de croyance (*Bel* et *Pl*). L'écart entre ces deux fonctions reflète alors le degré ou l'effet d'incertitude de paramètres. Il est donc important d'étudier les données d'entrée qui contribuent à cet écart. La réduction de cet écart implique la nécessité de recueillir davantage d'informations sur les paramètres d'entrée en tenant compte des coûts et des efforts associés. L'incertitude dans la sortie du modèle Y peut être mesurée en utilisant la fonction $unc()$ comme suit

$$unc(Y) := \|Bel - Pl\|_1 = \int_{-\infty}^{\infty} |Bel - Pl| dx \quad (52)$$

La fonction $unc(Y)$ peut être considérée comme l'aire du domaine compris entre la fonction de croyance (*Bel*) et la fonction de plausibilité (*Pl*). Par conséquent, cette fonction $unc(Y)$ reflète le degré d'incertitude de la variable de sortie Y . En plus, la fonction $unc(Y)$ est égale à zéro lorsque les intervalles focaux sont réduits à des singletons. Dans ce cas, les fonctions de croyance et de plausibilité (*Bel* et *Pl*) coïncident avec la CDF de la théorie des probabilités.

Ainsi, l'étude de sensibilité de la variable de sortie Y peut s'effectuer en mesurant l'effet de l'écart de la variation de l'incertitude sur un paramètre d'entrée. Le degré d'incertitude de la variable de sortie, $unc(Y)$, est comparé avec le degré d'incertitude $unc(Y_i)$ où Y_i est la variable de sortie obtenue après modification de l'incertitude du $i^{\text{ème}}$ paramètre d'entrée. La mesure de la sensibilité du $i^{\text{ème}}$ paramètre d'entrée à l'incertitude de la variable de sortie Y est calculée en utilisant l'équation suivante :

$$S_i = 1 - \frac{unc(Y_i)}{unc(Y)} \quad (53)$$

où $unc(Y_i)$ et $unc(Y)$ sont respectivement la quantité d'incertitude mesurée par rapport à la sortie avec et sans pincement des paramètres d'entrée du modèle. Plus la valeur S_i est grande, plus l'incertitude du $i^{\text{ème}}$ paramètre d'entrée à l'incertitude de la variable de sortie Y est importante. Cette étude de sensibilité est effectuée pour quantifier l'effet de chaque paramètre aléatoire et épistémique. L'impact d'incertitude sur chaque paramètre est évalué un par un. La variation de l'incertitude du $i^{\text{ème}}$ paramètre d'entrée peut être effectuée de plusieurs manières. Le $i^{\text{ème}}$ paramètre peut être remplacé par une valeur sans incertitude. Généralement, nous ne connaissons pas la valeur correcte avec certitude ; donc ce remplacement doit être de nature conjecturale. Les stratégies suivantes sont également suivies pour estimer les sensibilités à partir des évaluations des fonctions de croyance (*Bel*) et de plausibilité (*Pl*) :

- Remplacer le $i^{\text{ème}}$ paramètre par une valeur point,
- Remplacer le $i^{\text{ème}}$ paramètre par une fonction de distribution précise, ou
- Remplacer le $i^{\text{ème}}$ paramètre par un intervalle de zéro variance.

Le remplacement des distributions de *Bel* et de *Pl* du $i^{\text{ème}}$ paramètre par une valeur point serait de sélectionner aléatoirement une valeur dans l'intervalle de ce paramètre. La mesure de la sensibilité S_i est ensuite moyennée quand le paramètre prend un certain nombre de valeurs possibles. Cette stratégie est souvent utilisée dans le cas où les informations sur les paramètres du modèle sont données tout simplement sous forme

d'intervalle sans avoir une connaissance sur la valeur la plus représentative. Une limitation rencontrée dans cette stratégie est qu'en échantillonnant les valeurs possibles de chaque paramètre, le temps du calcul de S_i pourrait être très long pour un modèle complexe et comportant un nombre important de paramètres incertains. Le remplacement des distributions de Bel et de Pl par une distribution de probabilité précise, serait de pincer l'incertitude sur la distribution alors que le remplacement de ces distributions par une valeur ponctuelle serait de pincer à la fois l'incertitude et la variabilité de la quantité.

Pour les entrées qui sont connues pour être variables (variance supérieure à zéro), un tel pincement est contrefactuel, mais il peut néanmoins être informatif. En particulier, il peut être utile pour planifier des stratégies d'assainissement. Dans certaines situations, il peut être raisonnable de remplacer les distributions de Bel et de Pl par un intervalle mais prescrit pour avoir une variance de zéro. Cela aurait pour effet de pincer la variabilité mais de laisser l'incertitude.

Dans notre approche, nous proposons d'utiliser la valeur nominale (la valeur la plus représentative) de chaque paramètre obtenue à partir de retour d'expérience ou par avis d'expert. Cette manière de faire est compatible, d'une part, avec le calcul probabiliste de référence du modèle de COS (sans incertitude) qui est effectué avec les valeurs nominales et d'autre part, permet de réduire significativement le temps de quantification des incertitudes dans ce modèle.

Notons que cette méthode d'analyse de sensibilité peut pincer plusieurs paramètres d'entrée (groupe de paramètres) en même temps, ce qui serait utile pour étudier les corrélations entre les paramètres aléatoires et/ou épistémiques.

2.1.3 Quantification des incertitudes de la structure du modèle

L'incertitude de la structure du modèle est l'une des sources les plus importantes d'incertitude dans la modélisation de COS et qui est difficile à quantifier. La réduction de cette incertitude nécessite d'identifier la structure du modèle de COS la plus appropriée pour un système terrestre donné. Pour évaluer la cohérence des modèles de COS, nous pouvons utiliser des techniques d'évaluation du modèle pour tester leur capacité à reproduire le comportement du système terrestre.

Dans cette thèse, nous proposons d'utiliser la mesure d'importance précédemment discutée, de la même manière, pour mesurer l'impact des sources d'incertitude liées à la structure du modèle sur les résultats à la sortie.

Exemple 1. Pour illustrer la mesure de l'analyse de sensibilité étudiée, prenons l'exemple 4 dans le chapitre 3. Le Tableau 8 et la Figure 30 montrent les résultats de la mesure du pincement pour les cinq paramètres : indice de forme (SI), indice de luminosité (BI), la surface (A), le NDVI et la température (Tem). Ces paramètres sont obtenus à travers les deux différentes structures f_1 et f_2 du modèle DINAMICA.

La mesure d'importance S_i est calculée à partir des fonctions de croyance et de plausibilité de la variable de sortie du modèle Y prenant en compte des incertitudes de cinq paramètres pour chacune des deux structures. Plus la valeur S_i est grande, plus l'incertitude de paramètre est importante. La structure du modèle la plus influente montre généralement une plus grande sensibilité des paramètres.

f_1		f_2	
Paramètre	S_i	Paramètre	S_i
SI	0.193	SI	0.246
BI	0.206	BI	0.281
A	0.614	A	0.693
NDVI	0.751	NDVI	0.822
Tem	0.558	Tem	0.575

TABLE 8 – Mesure d'importance des cinq paramètres SI, BI, A, NDVI et Tem à travers deux structures f_1 et f_2 du modèle DINAMICA.

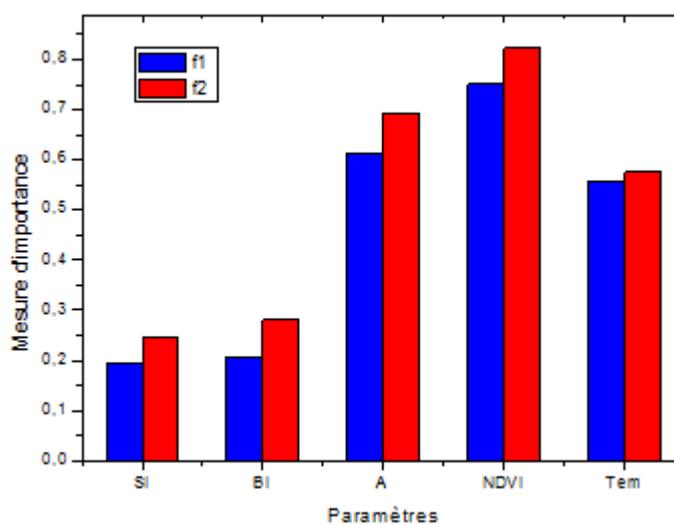


FIGURE 30 – Représentation graphique de classement d'importance des cinq paramètres SI, BI, A, NDVI et Tem à travers deux structures f_1 et f_2 du modèle DINAMICA.

Paramètre	f_1		Paramètre	f_2	
	S_i avec corrélation	S_i sans corrélation		S_i avec corrélation	S_i sans corrélation
SI	0.193	0.213	SI	0.246	0.301
BI	0.206	0.237	BI	0.281	0.377
A	0.614	0.721	A	0.693	0.699
NDVI	0.751	0.798	NDVI	0.822	0.842
Tem	0.558	0.569	Tem	0.575	0.587

TABLE 9 – Mesure d’importance des cinq paramètres SI, BI, A, NDVI et Tem avec et sans considération de la corrélation à travers deux structures f_1 et f_2 .

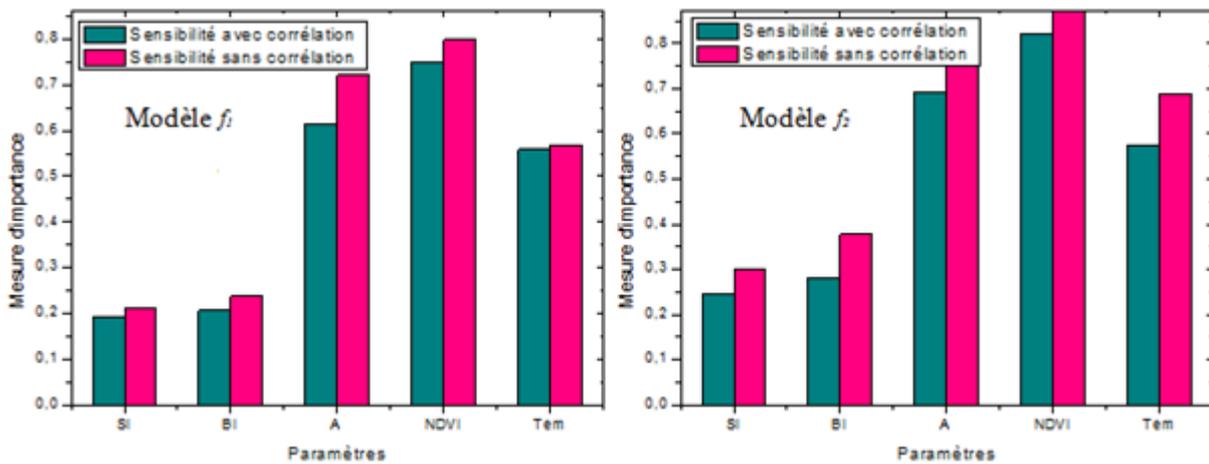


FIGURE 31 – Comparaison entre la sensibilité de paramètres SI, BI, A, NDVI et Tem avec et sans considération de la corrélation pour les deux structures f_1 et f_2 .

Comme le montre la Figure 30, les paramètres A, NDVI et Tem dans les deux structures f_1 et f_2 contribuent la plus à l’incertitude sur la variable de sortie Y et ont l’impact le plus important sur la décision. Nous remarquons aussi que les paramètres de f_2 sont relativement sensibles que ceux de f_1 .

Il est important de tester la différence de l’analyse de sensibilité avec et sans considération de la corrélation entre les paramètres des modèles de COS. Ainsi, il est important de distinguer la contribution de l’incertitude pour un paramètre individuel dans les deux cas. Par conséquent, les résultats obtenus de cinq paramètres SI, BI, A, NDVI et Tem dans les deux structures f_1 et f_2 sont montrés dans le Tableau 9 et présentés à la Figure 31.

Les résultats obtenus dans la Figure 31 montrent que tous les paramètres dans les deux structures f_1 et f_2 sont plus sensibles dans le cas où les corrélations ne sont pas prises en considération. Ainsi, il est important de distinguer les différences entre les deux cas. Malgré ces différents résultats, nous avons obtenu la même classification d’importance de paramètres. Les paramètres A, NDVI et Tem dans les deux structures f_1 et f_2 contribuent la plus à l’incertitude sur la variable de sortie Y et ont l’impact le plus important sur la décision.

Surface			NDVI			Température		
<i>Intervalle</i>	<i>BPA</i>	<i>F(x)</i>	<i>Intervalle</i>	<i>BPA</i>	<i>F(x)</i>	<i>Intervalle</i>	<i>BPA</i>	<i>F(x)</i>
[0, 1200]	0.04	0.04	[-1,-0.5]	0.5	0.5	[0,10]	0.05	0.05
[1200, 2400]	0.09	0.13	[-0.5,0]	0.35	0.85	[10,20]	0.25	0.3
[2400, 3600]	0.22	0.35	[0,0.5]	0.1	0.95	[20,30]	0.5	0.8
[3600, 4800]	0.65	1	[0.5,1]	0.05	1	[30,40]	0.20	1
Somme	1		Somme	1		Somme	1	

TABLE 10 – Collection de paires d’intervalles-masses de surface, NDVI et température.

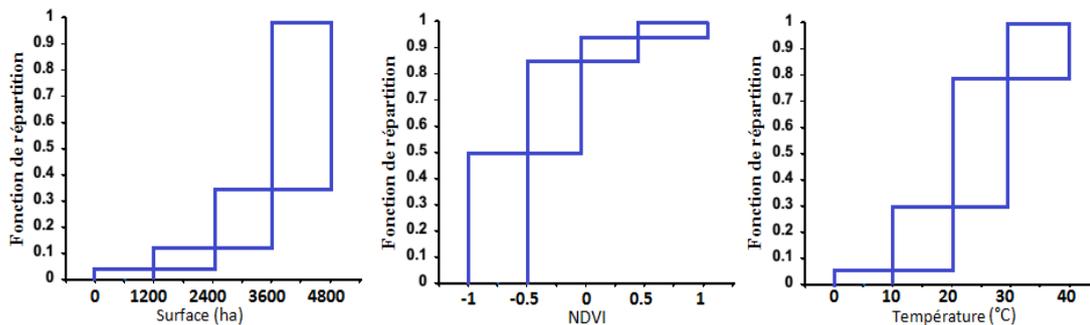


FIGURE 32 – Structures évidentielles pour la surface, le NDVI et la température.

2.2 Estimation

Après avoir sélectionné les sources d’incertitude les plus importantes, le but de cette section est d’estimer les sources sélectionnées à partir des données et de fixer les autres à des valeurs nominales afin d’obtenir des valeurs optimales. Comme nous avons déjà mentionné, tous les paramètres et les structures incertains des modèles de COS sont modélisés par des intervalles avec des fonctions de masse associées (intervalles focaux).

Pour estimer les paramètres les plus influents, nous proposons d’utiliser les limites de confiance du Kolmogorov-Smirnov afin de donner un certain niveau de confiance pour chaque paramètre. L’estimation des intervalles de confiance du Kolmogorov-Smirnov d’une structure évidentielle est donnée par l’expression proposée par Ferson *et al.* [Ferson *et al.* 2003]

$$\text{Min}(1, \max(0, F(x) \pm D(\alpha, n))) \quad (54)$$

où $F(x)$ est une fonction de distribution, α est le niveau de confiance, et $D(\alpha, n)$ est la statistique de Kolmogorov pour un niveau de confiance $100(1 - \alpha)\%$ et un nombre n des intervalles. Les valeurs de $D(\alpha, n)$ ont été proposées par Miller [Miller 1956] et ont été modifiées par Lilliefors [Lilliefors 1967].

Exemple 2. Pour illustrer la procédure d’estimation des paramètres les plus influents, nous prenons les résultats obtenus dans l’exemple 1. Nous allons estimer les valeurs des paramètres suivants : surface, NDVI et température. Chaque paramètre va être traité par des intervalles avec des fonctions de masse associées comme montre le Tableau 10. Les représentations graphiques des structures évidentielles obtenues de paramètres suivants : surface, NDVI, et température sont schématisées dans la Figure 32.

Surface				NDVI				Température			
Intervalle	$F(x)$	Min	Max	Intervalle	$F(x)$	Min	Max	Intervalle	$F(x)$	Min	Max
[0, 1200]	0.04	0	0.421	[-1,-0.5]	0.5	0.119	0.881	[0,10]	0.05	0	0.431
[1200, 2400]	0.13	0	0.511	[-0.5,0]	0.85	0.469	1.000	[10,20]	0.3	0	0.681
[2400, 3600]	0.35	0	0.731	[0,0.5]	0.95	0.569	1.000	[20,30]	0.8	0.419	1.000
[3600, 4800]	1	0.619	1.000	[0.5,1]	1	0.619	1.000	[30,40]	1	0.619	1.000

TABLE 11 – Les limites de confiances du Kolmogorov-Smirnov pour la surface, le NDVI et la température.

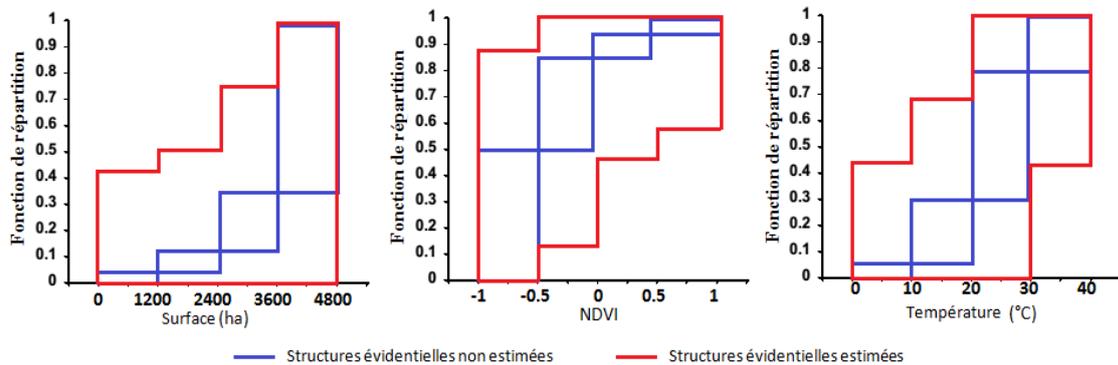


FIGURE 33 – Les structures évidentielles estimées pour la surface, le NDVI et la température.

Selon l'expression d'estimation proposée par Ferson *et al.* [Ferson *et al.* 2003] (voir équation (54)), les limites de confiance du Kolmogorov-Smirnov tout en utilisant les valeurs modifiées pour $D(\alpha, n)$ de ces structures évidentielles pour un intervalle de confiance de 95% ($\alpha = 95\%$), pour $D(95\%, 4) = 0.381$ sont représentées dans le Tableau 11.

La Figure 33 représente les fonctions de croyance et de plausibilité en comparaison avec les structures évidentielles sans estimation. Il est à noter que les limites de confiance du Kolmogorov-Smirnov précédemment présentées peuvent être, de la même manière, utilisées pour estimer les structures des modèles de COS.

Après avoir appliqué les méthodes de réduction, la propagation des incertitudes par la théorie des fonctions de croyance est appliquée à nouveau en réduisant les incertitudes liées aux paramètres et à la structure du modèle de COS et ceci pour générer les nouvelles plages d'incertitude de prédiction.

Exemple 3. Pour illustrer l'importance de l'approche proposée de réduction des incertitudes dans les modèles de COS, nous introduisons les valeurs optimales obtenues de l'exemple 2 pour les paramètres et la structure les plus influents à nouveau à l'étape de propagation.

La Figure 34 présente les fonctions de croyance et plausibilité cumulatives de la réponse finale du modèle Y .

Les résultats obtenus dans cet exemple montrent que les plages d'incertitudes de la

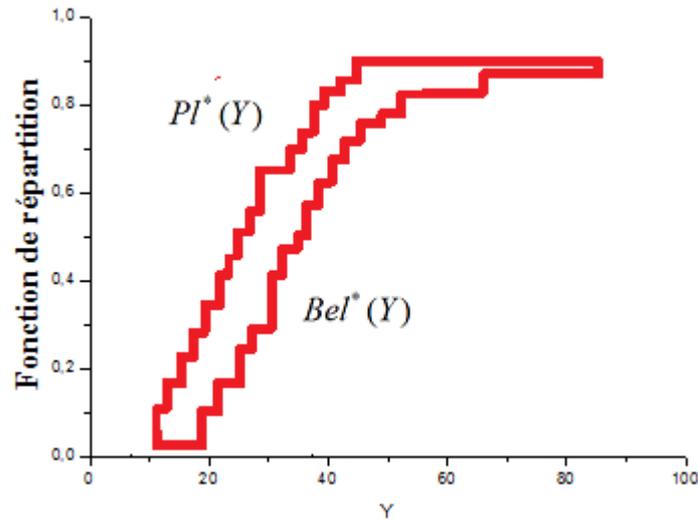


FIGURE 34 – Structure évidentielle de la variable de sortie Y avec réduction de l'incertitude de paramètres et de la structure du modèle.

sortie du modèle DINAMICA avant l'implémentation de l'approche de réduction des incertitudes sont relativement importantes (voir Figure 28 dans le chapitre 3), alors que ces plages sont relativement faibles avec l'application de l'approche proposée (voir Figure 34).

2.3 Evaluation de la performance des modèles de COS

Une fois qu'un modèle de COS est calibré et appliqué, une validation de la sortie du modèle avec des données de référence est complétée dans la majorité des études. Ceci est réalisé en résumant les pixels bien et mal classifiés dans une matrice de confusion [Comber *et al.* 2012]. Une variété de mesures de précision, telles que le désaccord général (*the overall disagreement*, en anglais), la précision de l'utilisateur (*the user' accuracy*, en anglais), l'indice de Kappa (*Kappa index*, en anglais), l'indice de Kappa flou (*the Fuzzy Kappa index*, en anglais) ou la mesure ROC (*the relative operating characteristic*, en anglais), sont basées sur cette matrice [Cohen 1960, Mas *et al.* 2013, van Vliet *et al.* 2013, Olofsson *et al.* 2014, Pontius et Parmentier 2014]. En outre, le désaccord de répartition et le désaccord de la quantité sont des mesures largement utilisées [Pontius et Millones 2011]. Ils séparent l'erreur totale en deux parties. La première partie fait référence au désaccord de la quantité des différentes catégories de changement des sols dans la sortie du modèle et les données de référence, tandis que la seconde quantifie le décalage spatial.

Dans cette thèse, pour évaluer la performance des modèles de COS et comparer les résultats simulés par rapport aux résultats observés, 3 mesures statistiques sont utilisées : le coefficient de détermination (R^2), le critère de Nash-Sutcliffe (NSE) et le pourcentage de biais ($PBIAS$). Plusieurs travaux récents ont montré l'efficacité des ces critères d'évaluation pour les modèles de COS [El-Sadek et Irvem 2014, Nigussie et Altunkaynak 2016].

- R^2 décrit le degré de colinéarité entre les cartes d'occupation des sols simulées et

mesurées. Il décrit la proportion de la variance dans les données mesurées expliquées par le modèle de COS. La valeur de R^2 varie de 0 à 1, avec des valeurs plus élevées indiquant moins de variance d'erreur.

$$R^2 = \frac{\sum_{i=1}^n (Q_{oi} - Q_o)(Q_{pi} - Q_p)}{\sqrt{\sum_{i=1}^n (Q_{oi} - Q_o)^2} \sqrt{\sum_{i=1}^n (Q_{pi} - Q_p)^2}} \quad (55)$$

où n est le nombre total d'observations, Q_{oi} , Q_{pi} , Q_o et Q_p représentent, respectivement, les valeurs observées, simulées, observées moyennes et simulées moyennes.

- NSE , qui est compris entre $-\infty$ et 1, mesure le pouvoir prédictif du modèle de COS. Les résultats d'un modèle de COS seront d'autant meilleurs que le NSE se rapproche de 1. Par comparaison avec d'autres critères, celui-ci a l'avantage d'être d'interprétation simple. En effet, s'il est positif, le modèle donne de meilleurs résultats qu'un modèle naïf donnant à chaque pas de temps un débit constant égal à la moyenne des débits observés ; s'il est négatif, c'est l'inverse.

$$NSE = \frac{\sum_{i=1}^n (Q_{oi} - Q_o)^2 - \sum_{i=1}^n (Q_{pi} - Q_{oi})^2}{\sum_{i=1}^n (Q_{oi} - Q_o)^2} \quad (56)$$

- $PBIAS$ mesure le pourcentage d'écart entre les cartes d'occupation des sols simulées et celles observées. Un $PBIAS$ négatif indique que les valeurs simulées sont plus élevées que celles observées (surestimation du biais du modèle de COS), et un $PBIAS$ positif indique que les valeurs simulées sont inférieures à celles observées (sous-estimation du biais).

$$PBIAS = \frac{\sum_{i=1}^n (Q_{pi} - Q_{oi})}{\sum_{i=1}^n (Q_{oi})} \quad (57)$$

Les critères mentionnés au dessus sont utilisés pour comparer quantitativement les résultats de COS obtenus en se basant sur des prédictions du modèle de COS et des données d'observation de la même période.

3 Conclusion

Dans ce chapitre, nous avons étudié les étapes du processus de la réduction des incertitudes dans le cadre de la théorie des fonctions de croyance pour les applications des modèles de COS.

Afin de donner aux décideurs des outils permettant d'identifier les paramètres et les structure du modèle de COS qui ont un impact important sur l'incertitude de la sortie, une stratégie basée sur le pincement est proposée. Cette stratégie permet de pincer des paramètres de différents types (aléatoires et/ou épistémiques). La stratégie du pincement permet également d'étudier plusieurs paramètres d'entrée (groupe de paramètres) en même temps, ce qui serait utile d'étudier les corrélations entre les paramètres des modèles de COS.

Après avoir obtenu les sources d'incertitude les plus importantes sur les résultats des modèles de COS, les limites de confiance de Kolmogorov-Smirnov sont utilisées pour les estimer afin d'obtenir des valeurs optimales. Après l'étude du processus de la

réduction des incertitudes, nous nous sommes intéressés à l'évaluation de la performance des modèles de COS. Pour ce faire, nous avons présenté trois mesures statistiques à savoir le coefficient de détermination (R^2), le critère de Nash-Sutcliffe (NSE) et le pourcentage de biais ($PBIAS$).

Nous présentons dans le chapitre suivant l'implémentation et la validation de l'approche proposée pour la propagation et la réduction des incertitudes dans les modèles de COS. Ensuite, nous montrons les résultats des différentes expérimentations que nous avons menées et qui nous ont permis d'évaluer et de valider l'approche proposée.

Chapitre V

Expérimentation

Sommaire

1	Introduction	79
2	Application de l'approche proposée	79
2.1	Description du scénario d'application	79
2.2	Présentation de la base d'images et de la région d'étude	79
2.2.1	Présentation de la zone d'étude	79
2.2.2	Présentation de la base d'images	80
2.3	Présentation de modèles de COS et de paramètres utilisés	82
2.4	Résultats d'application de l'approche proposée de la propagation des incertitudes	82
2.4.1	Etape 1 : Modélisation de paramètres incertains	82
2.4.2	Etape 2 : Etude de corrélation	83
2.4.3	Etape 3 : Propagation des incertitudes de paramètres	85
2.4.4	Etape 4 : Propagation des incertitudes de la structure de modèle	89
2.5	Résultats d'application de l'approche proposée de la réduction des incertitudes	90
2.5.1	Etape 5 : Etude de sensibilité	91
2.5.2	Etape 6 : Estimation	93
2.6	Résultats des cartes de prédiction de COS	94
3	Evaluation de l'approche proposée	99
4	Conclusion	100

1 Introduction

Dans les deux chapitres précédents, nous avons présenté notre approche de propagation et de réduction d'incertitude basée sur la théorie des fonctions de croyance dans les modèles de COS. Nous nous intéressons dans ce chapitre à présenter les principaux résultats obtenus suite à l'application de cette approche sur un scénario d'urbanisation.

Des analyses comparatives avec la théorie de possibilités permettent d'évaluer la capacité de la théorie des fonctions de croyance ainsi étendue à assurer une conception robuste en présence de différents types et sources des incertitudes.

2 Application de l'approche proposée

Dans cette section, nous décrivons d'abord le scénario d'application. Ensuite, nous présentons la zone d'étude et la base d'images utilisées dans le scénario d'application. Puis, nous dressons les modèles de COS et les paramètres choisis pour ce scénario. Enfin, Nous détaillons les résultats obtenus.

2.1 Description du scénario d'application

Le but du scénario est d'appliquer notre approche pour la prédiction des changements d'une zone urbaine. Le scénario d'application est divisé en deux parties principales. Chacune ayant ses propres étapes :

- Propagation des incertitudes dans les modèles de COS
 - Etape 1 : Modélisation des incertitudes
 - Etape 2 : Etude de corrélation
 - Etape 3 : Propagation des incertitudes de paramètres
 - Etape 4 : Propagation des incertitudes de la structure du modèle
- Réduction des incertitudes dans les modèles de COS
 - Etape 5 : Etude de sensibilité
 - Etape 6 : Estimation

La Figure 35 illustre les six étapes de l'approche proposée.

2.2 Présentation de la base d'images et de la région d'étude

2.2.1 Présentation de la zone d'étude

La zone d'étude pour le scénario d'application considéré est la zone du Caire. Le Caire, la capitale de l'Egypte, est la ville la plus peuplée en Egypte (Figure 36) et est considéré comme une méga-ville mondiale. La cartographie de COS est importante pour comprendre et analyser les relations entre la géomorphologie (Highlands et déserts), les ressources naturelles (terres agricoles et la rivière du Nil) et les activités humaines. Les terres agricoles autour du Caire ont connu des pratiques d'empiètement graves en raison d'une croissance démographique rapide. Ainsi, les plaines désertiques adjacentes

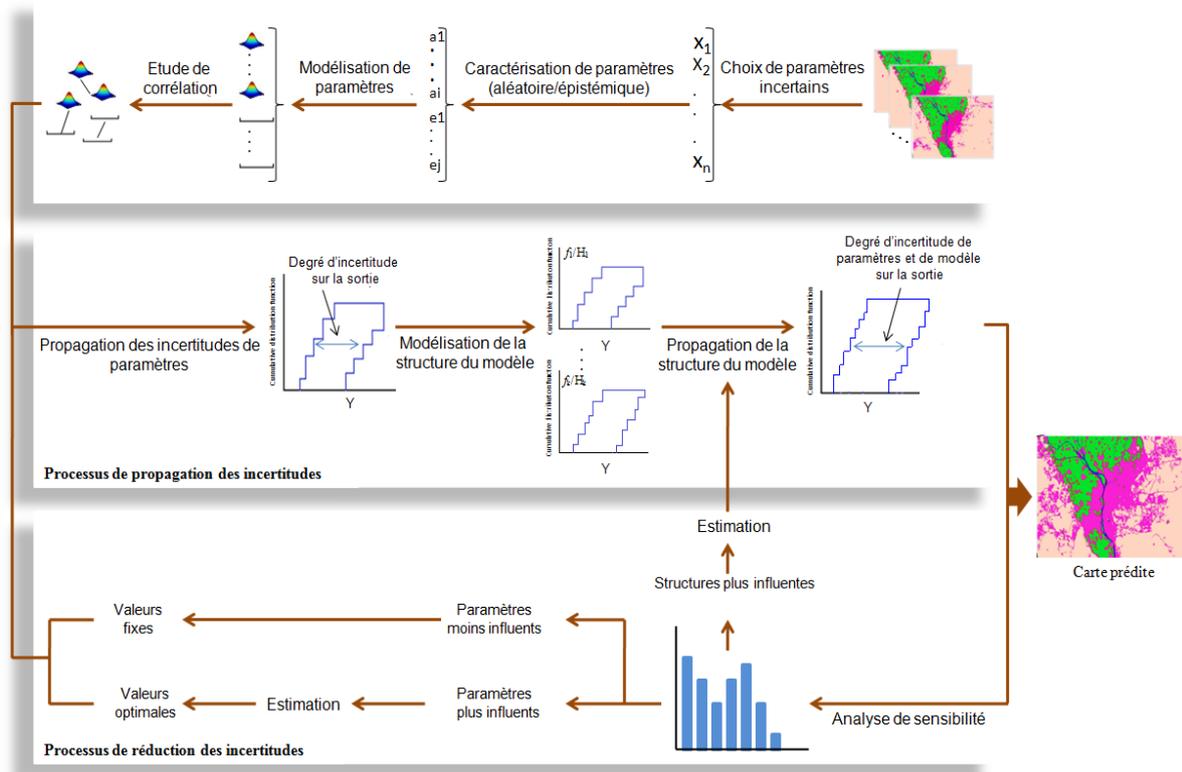


FIGURE 35 – Scénario d'application de l'approche proposée.

ont également assisté à des pratiques d'urbanisation pour englober la croissance de la population intensive.

Différentes études ont déjà été réalisées pour la détection et la modélisation de COS dans la zone du Caire [Sutton et Fahmi 2001, Yin *et al.* 2005, de Noronha *et al.* 2011, Mohamed 2012]. La population du Caire (ville du Caire et Gizeh) est passée d'environ 6,4 millions en 1976 [Sutton et Fahmi 2001] à environ 12,5 millions en 2006 selon l'Agence centrale de la Mobilisation publique et des Statistiques d'Egypte (CAPMAS). L'importance du Caire découle de son emplacement à mi-chemin entre la vallée du Nil et le delta. En plus, les principales installations et services du gouvernement se trouvent au Caire. Ainsi, le besoin d'étudier les COS de la zone du Caire est important.

2.2.2 Présentation de la base d'images

Afin de prédire le COS de la zone du Caire, nous avons utilisé des images provenant de la base de données en ligne de la USGS¹ (*United States Geological Survey*). Cette base offre des séries de données satellitaires prétraitées, multitemporelles et multi-capteurs. Dans cette étude, trois images Landsat TM5 sont obtenues à partir de la base USGS. Ces trois images acquises en 6 Avril 1987, 8 Août 2003 et 15 Mars 2014, respectivement, sont classifiées en quatre types d'occupation des sols qui sont urbain, agriculture, désert et eau pour produire des cartes de COS (Figure 37).

Au cours de cette période, la population du Caire a augmenté d'environ 7 millions

1. www.usgs.gov

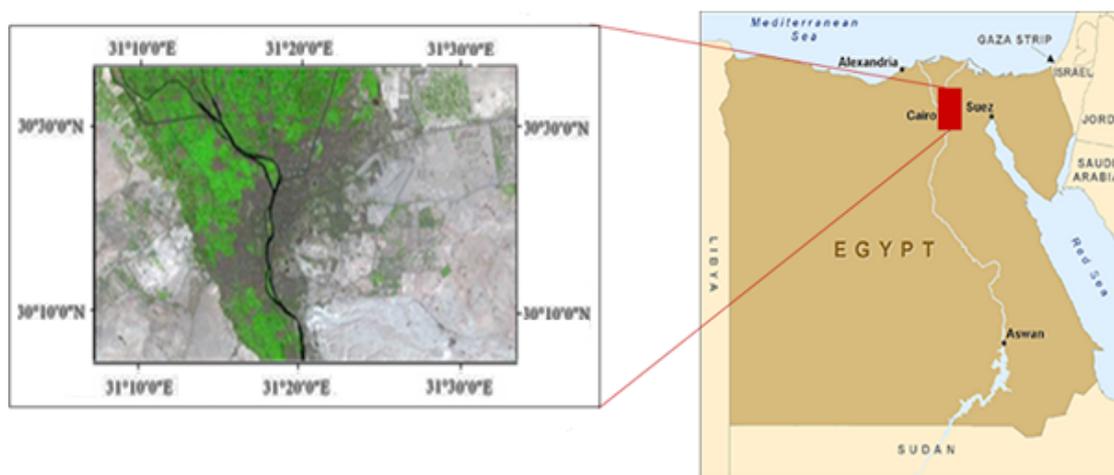


FIGURE 36 – Localisation de la zone d'étude.

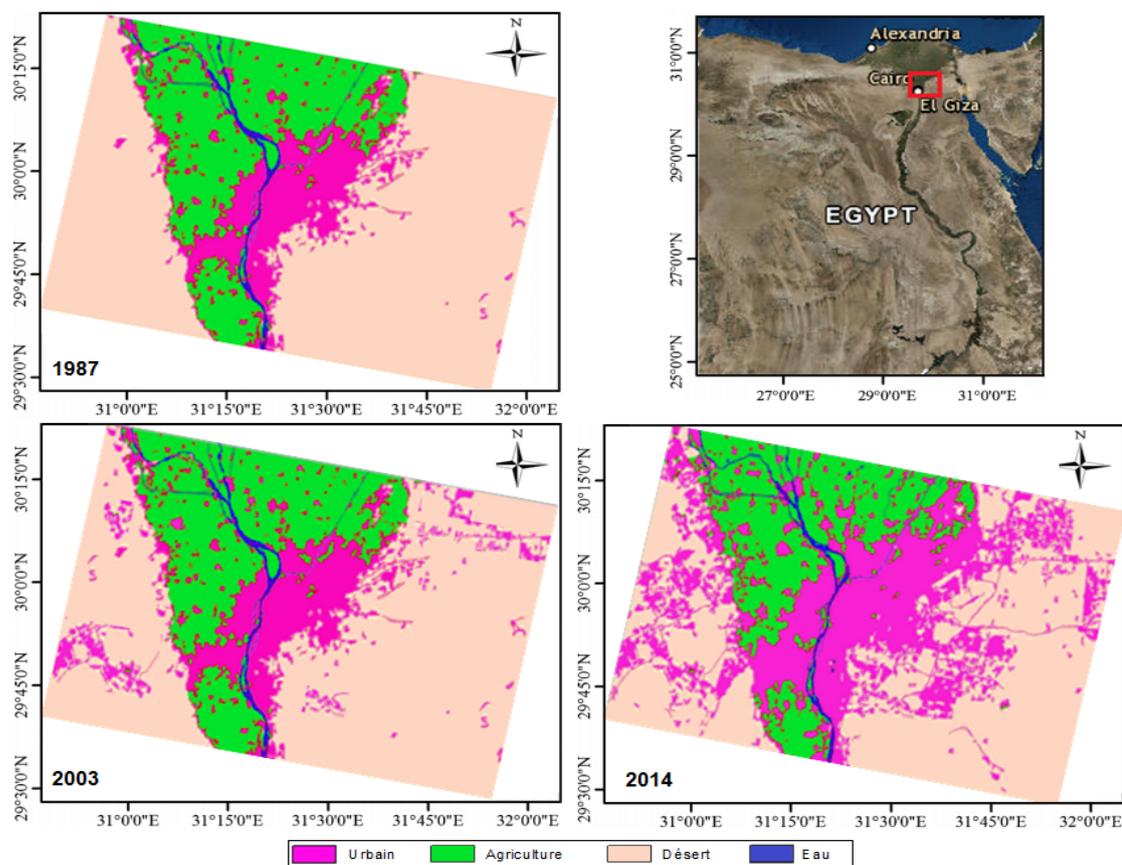


FIGURE 37 – Cartes d'occupation des sols.

en 1987 à plus de 15 millions en 2014. La croissance récente de la population a causé la ville et ses zones urbaines associées à s'étendre dans le désert. Aussi, ces trois images satellites montrent également une augmentation globale des terres urbaines développées (magenta) par rapport aux terres agricoles précédentes (vert).

Dans cette étude, les données satellitaires ont été classifiées après corrections initiales et traitement afin de pouvoir extraire des informations utiles de ces images. Des corrections spectrales, géométriques et atmosphériques des images ont été réalisées pour rendre les caractéristiques manifestes, améliorer la qualité des images et éliminer les effets néfastes de la lumière et de l'atmosphère.

2.3 Présentation de modèles de COS et de paramètres utilisés

Comme nous l'avons évoqué dans le chapitre 1, plusieurs modèles ont été proposés pour analyser, détecter et prédire les COS. Généralement, la plupart des études ont été limitées à quantifier l'incertitude à travers un seul modèle de COS. Seulement quelques études dans la littérature ont examiné l'effet de choix du modèle sur les résultats des prédictions et ont étudié la complexité de ces modèles [Pontius et Malanson 2005, Mas et al. 2014].

Dans cette étude, l'approche proposée est appliquée en utilisant trois modèles hybrides de COS à savoir DINAMICA, CA-MARKOV et LCM, et un modèle d'urbanisation SLEUTH. Ces modèles sont différents dans la façon dont ils simulent les COS. Cette application peut aider les utilisateurs de modèles de COS à comprendre la précision de ces modèles et les variables pertinentes affectant la robustesse des sorties de chaque modèle. Chacun des quatre modèles contient un certain nombre de paramètres, décrivant les caractéristiques des objets extraits des images satellitaires. Ces paramètres présentent intrinsèquement une variabilité et une manque des connaissances, ce qui conduit à l'ignorance de leurs vraies valeurs et des inexactitudes dans leur estimation. Toute méthodologie d'analyse qui néglige l'incertitude liée à ses paramètres peut conduire à des solutions non optimales et ne fournit pas une évaluation précise. Dans cette étude, nous considérons 16 paramètres d'entrée pour les quatre modèles choisis (voir chapitre 3) : cinq paramètres spectraux (BI, MB, MR, MG et MN), trois paramètres de texture (Ctr, Ent et ASM), cinq paramètres de forme (RF, EF, SI, D et A), un paramètre de végétation (NDVI) et deux paramètres climatiques (Tem et Hum). Ces paramètres sont utilisés dans cette étude pour classifier les quatre types d'occupation des sols de la zone du Caire à savoir urbain, agriculture, désert et eau.

Les résultats présentés dans les sections suivantes sont effectués à travers le logiciel de statistique R [R Development 2012].

2.4 Résultats d'application de l'approche proposée de la propagation des incertitudes

2.4.1 Etape 1 : Modélisation de paramètres incertains

Afin de modéliser les paramètres incertains, la Figure 38 montre les valeurs des paramètres d'entrée dans chaque type d'occupation des sols selon la zone d'étude.

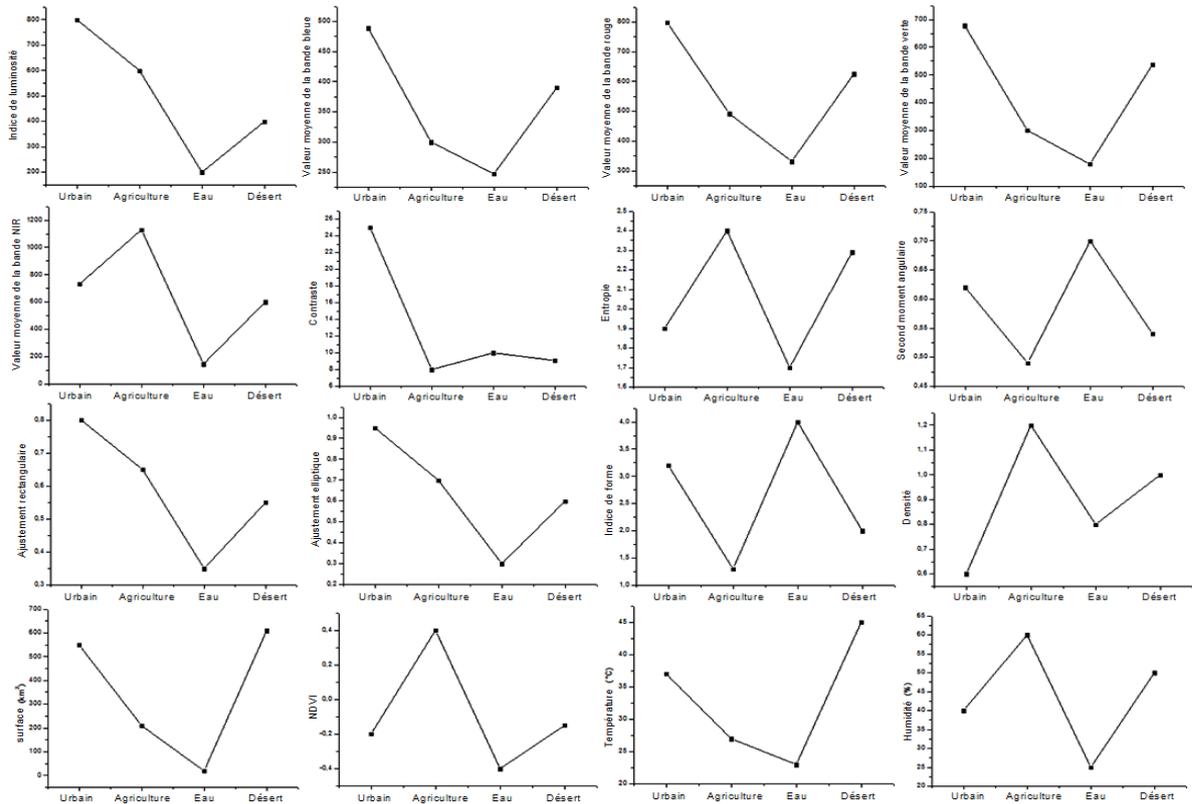


FIGURE 38 – Valeurs de paramètres dans chaque type d'occupation des sols.

Idéalement, chaque valeur de ces paramètres doit être accompagnée d'une description incertaine. Comme nous l'avons mentionné, les incertitudes liées aux paramètres des modèles de COS sont deux types : aléatoire et/ou épistémique. Dans cette étude, les paramètres spectraux, de végétation et climatiques sont considérés comme aléatoires parce que leurs incertitudes sont généralement dues à la variabilité naturelle. Ces paramètres sont décrits par des distributions normales de probabilité (moyenne et écart type). Les paramètres de texture et de forme sont considérés comme épistémiques parce que leurs incertitudes sont dues au manque de connaissances, et ils sont représentés par des intervalles avec des limites spécifiées. Le Tableau 12 présente les valeurs incertaines de tous les paramètres aléatoires et épistémiques dans chaque type d'occupation des sols.

2.4.2 Etape 2 : Etude de corrélation

Dans cette étude, l'analyse de corrélation est utilisée pour identifier les relations entre les paramètres aléatoires et épistémiques des modèles de COS. Les valeurs de corrélation entre ces paramètres sont présentées dans le Tableau 13 (voir la section 2.1.4 de chapitre 4 pour le calcul de corrélations). Les corrélations les plus élevées entre les paramètres sont clairement illustrées par des cellules grises plus foncées et les paramètres moins corrélés par des cellules grises claires dans le Tableau 13. Les paramètres peu ou pas corrélés sont présentés sans couleur. Des paramètres fortement corrélés sont suggérés pour réduire la dimensionnalité de l'espace de paramètres. Les valeurs positives indiquent que les paramètres reliés ont le même effet sur la sortie d'un modèle de COS, tandis que

Paramètres	Moyenne				écart type				Intervalle			
	<i>Urb</i>	<i>Agc</i>	<i>Wat</i>	<i>Dsr</i>	<i>Urb</i>	<i>Agc</i>	<i>Wat</i>	<i>Dsr</i>	<i>Urb</i>	<i>Agc</i>	<i>Wat</i>	<i>Dsr</i>
BI	792	602	200	398	170	214	34	72	-	-	-	-
MB	489	301	247	390	173	21	16	70	-	-	-	-
MR	798	492	330	626	317	49	37	158	-	-	-	-
MG	678	302	180	538	274	48	33	181	-	-	-	-
MN	733	1132	147	602	282	117	22	214	-	-	-	-
Ctr	-	-	-	-	-	-	-	-	[24,26]	[7.5,9]	[9.5,10.5]	[8,8.5]
Ent	-	-	-	-	-	-	-	-	[1.7,2]	[2.2,2.5]	[1.5,2]	[2.25,2.3]
ASM	-	-	-	-	-	-	-	-	[0.6,0.65]	[0.47,0.5]	[0.7,0.75]	[0.5,0.55]
RF	-	-	-	-	-	-	-	-	[0.7,0.8]	[0.6,0.7]	[0.3,0.4]	[0.5,0.6]
EF	-	-	-	-	-	-	-	-	[0.9,1]	[0.6,0.7]	[0.2,0.3]	[0.4,0.5]
SI	-	-	-	-	-	-	-	-	[3,3.5]	[1,1.5]	[3.8,4.2]	[2,2.5]
D	-	-	-	-	-	-	-	-	[0.5,0.8]	[1,1.4]	[0.7,0.9]	[0.9,1.2]
A	-	-	-	-	-	-	-	-	[500,600]	[200,300]	[20,30]	[600,700]
NDVI	-0.19	0.43	-0.2	0.28	0.13	0.26	0.25	0.20	-	-	-	-
Tem	36.48	28.4	26.5	42.8	1.9	0.8	1.7	0.83	-	-	-	-
Hum	39.18	60	25.5	49	1.02	1.9	1.7	0.8	-	-	-	-

TABLE 12 – Valeurs incertaines de paramètres aléatoires et épistémiques pour chaque type d'occupation des sols (*Urb* : urbain, *Agc* : agriculture, *Wat* : eau, et *Dsr* : désert).

	BI	MB	MR	MG	MN	Ctr	Ent	ASM	RF	EF	SI	D	A	NDVI	Tem	Hum
BI	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
MB	0.98	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-
MR	0.99	0.96	1	-	-	-	-	-	-	-	-	-	-	-	-	-
MG	0.98	0.94	0.95	1	-	-	-	-	-	-	-	-	-	-	-	-
MN	0.98	0.95	0.95	0.98	1	-	-	-	-	-	-	-	-	-	-	-
Ctr	0.25	0.85	0.83	0.81	0.79	1	-	-	-	-	-	-	-	-	-	-
Ent	0.11	0.76	0.72	0.81	0.72	0.89	1	-	-	-	-	-	-	-	-	-
ASM	0.29	0.85	0.83	0.87	0.79	0.97	-0.98	1	-	-	-	-	-	-	-	-
RF	0.12	0	0	0	0	0	0	0	1	-	-	-	-	-	-	-
EF	0.18	0	0	0	0	0	0	0	0.99	1	-	-	-	-	-	-
SI	0.26	0	0	0	0	0	0	0	0.89	0.90	1	-	-	-	-	-
D	0.06	0	0	0	0	0	0	0	0.98	0.99	0.98	1	-	-	-	-
A	0	0	0	0	0	0	0	0	0.76	0.76	0.78	0.76	1	-	-	-
NDVI	0.82	0.42	0.98	0.41	0.98	0.32	0.28	0.32	0.01	0.01	0.01	0.11	0.21	1	-	-
Tem	0.01	0	0	0	0	0	0	0	0	0	0	0.21	0	0.78	1	-
Hum	0.08	0	0	0	0	0	0	0	0	0	0	0.02	0	0.81	-0.80	1

TABLE 13 – Coefficients de corrélations entre tous les paramètres.

les valeurs négatives désignent des effets opposés de deux paramètres reliés sur la sortie de ce modèle.

Nous remarquons que les paramètres MB/BI, MR/BI, MR/MB, etc, montrent les coefficients de corrélation les plus élevés, qui sont extrêmement proches de 1. D'autres paramètres comme Ctr/BI, ASM/BI, SI/BI, NDVI/MB, NDVI/MG, NDVI/Ctr, NDVI/Ent, NDVI/ASM, NDVI/A et Tem/D avec des coefficients de corrélation de 0,25, 0,29, 0,26, 0,42, 0,41, 0,32, 0,28, 0,32, 0,21 et 0,21, respectivement, sont aussi relativement corrélés.

La capture de la vraie relation entre les paramètres peut être cruciale pour un calcul précis dans les prédictions du modèle.

2.4.3 Etape 3 : Propagation des incertitudes de paramètres

Comme nous l'avons indiqué, les paramètres de modèles de COS sont entachés par des incertitudes aléatoires et épistémiques. Ces incertitudes se propagent à travers le modèle, ce qui génère des incertitudes dans le processus de modélisation. Ignorer cette distinction peut affecter les résultats de la propagation de l'incertitude.

Dans cette étude, la théorie des fonctions de croyance est utilisée pour propager les incertitudes représentant les paramètres à travers les quatre modèles de COS : DYNAMICA, SLEUTH, CA-MARKOV et LCM. Pour illustrer l'importance de la propagation de deux types d'incertitude à travers ces modèles de COS, l'analyse avec une hypothèse d'incertitude aléatoire pure est menée où tous les 16 paramètres incertains sont traités comme aléatoires avec des distributions normales de probabilités. Dans ce cas, la fonction de distribution cumulative (CDF) de la sortie de chaque modèle représentant uniquement l'incertitude dans les paramètres est obtenue par la théorie des fonctions de

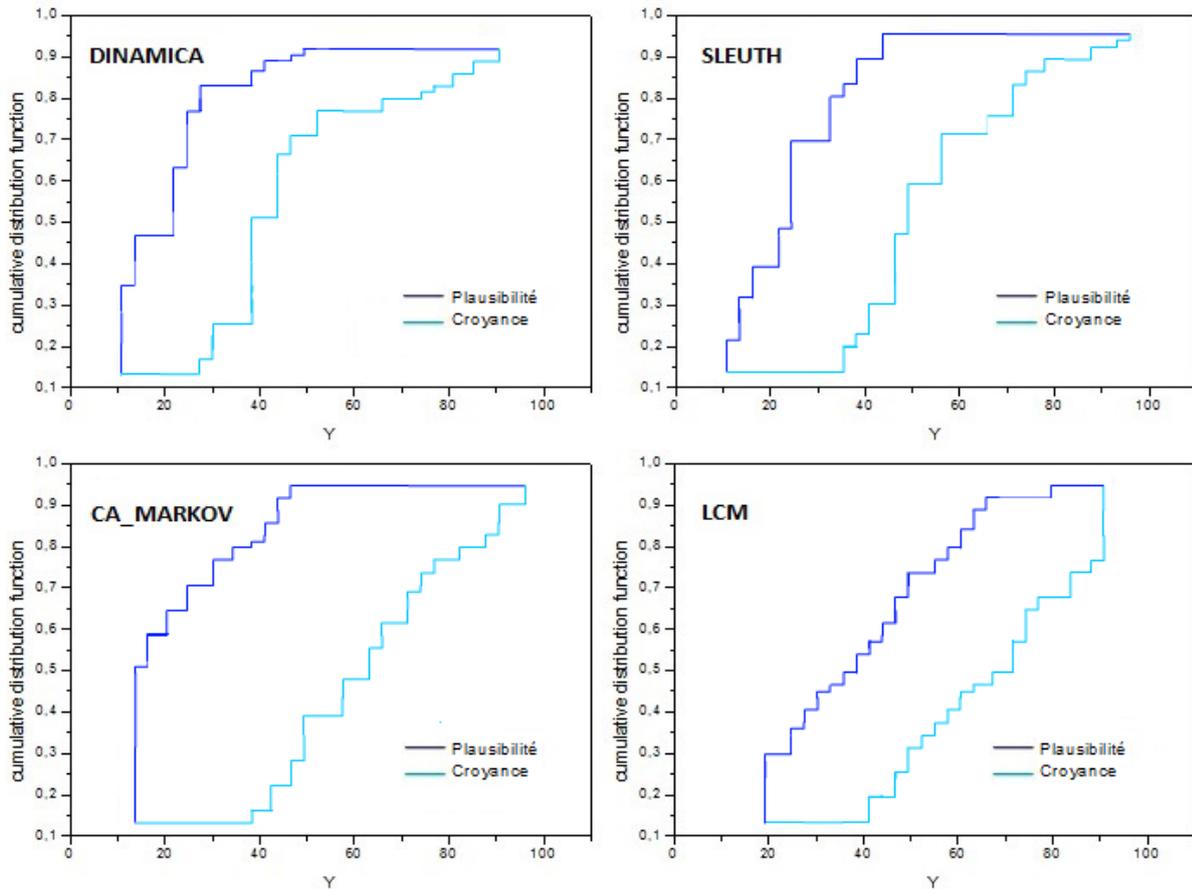


FIGURE 39 – Sorties de quatre modèles de COS représentées par des distributions de croyance et de plausibilité avec seulement une propagation des incertitudes aléatoires de paramètres.

croyance. La Figure 39 montre ces distributions basées sur 10.000 échantillons.

Maintenant, pour la propagation des incertitudes aléatoires et épistémiques de paramètres dans les modèles de COS, une approche basée sur les intervalles est utilisée pour modéliser les paramètres épistémiques, et les distributions normales de probabilités sont utilisées pour modéliser les paramètres aléatoires. 10.000 échantillons sont choisis pour les deux cas aléatoires et épistémiques. Les CDF produites de quatre modèles de COS sont illustrées sur la Figure 40. Les résultats obtenus dans cette figure indiquent des incertitudes importantes dans les quatre modèles. En outre, nous pouvons constater que l'incertitude des paramètres conduit à une plus grande incertitude dans le résultat du modèle CA-MARKOV en comparaison à celle causée par le reste des modèles. Cependant, DINAMICA conduit à plus faible degré d'incertitude. Par conséquent, l'incertitude liée aux paramètres du modèle devrait être plus soigneusement simulée tout en réduisant son occurrence et accordant plus d'attention à ses effets sur la prise de décision des modèles de COS.

Comme nous l'avons également évoqué, les paramètres des modèles de COS sont fortement corrélés. Alors, il est important d'étudier l'effet de ces paramètres corrélés sur la variation de la réponse de sortie et sur les résultats de propagation des incertitudes.

Selon le coefficient de corrélation, par exemple, nous avons trouvé des corrélations

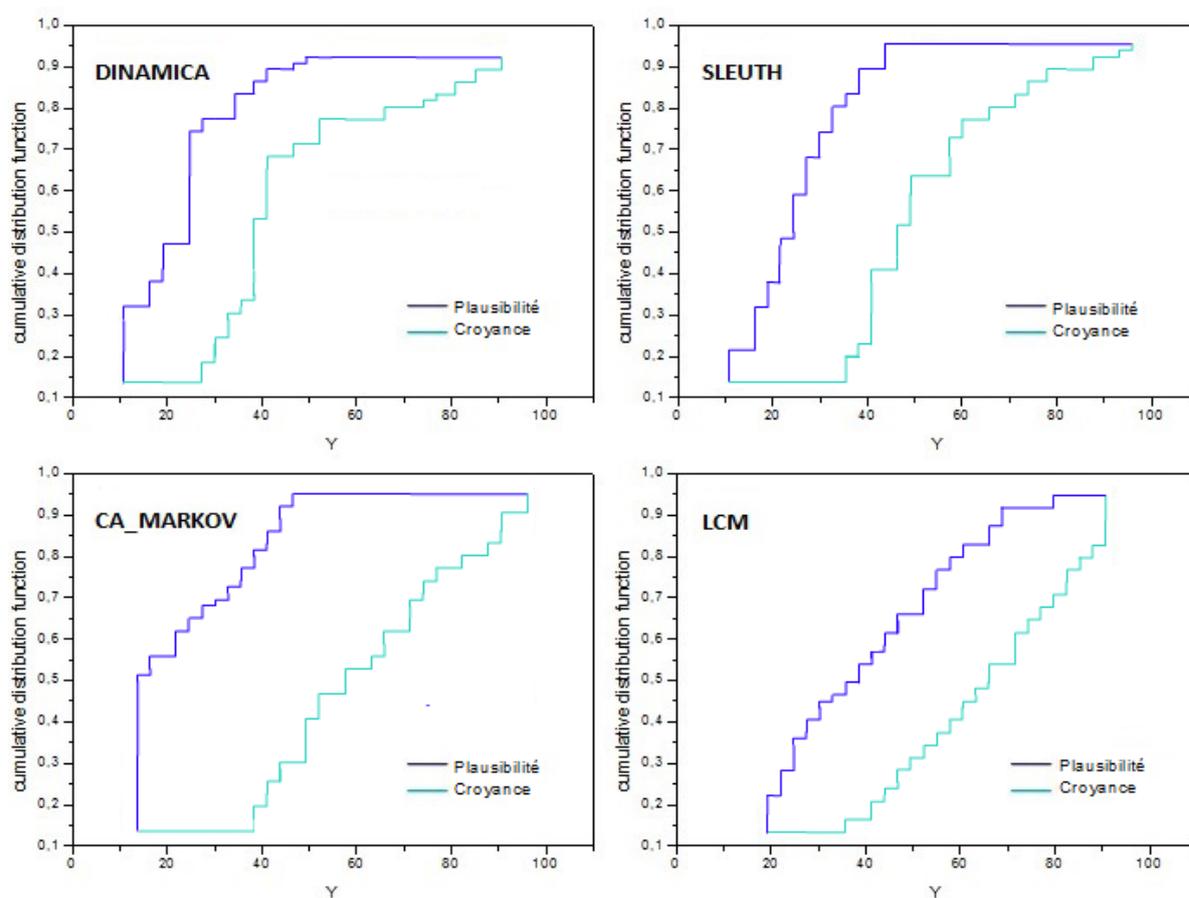


FIGURE 40 – Sorties de quatre modèles de COS représentées par des distributions de croyance et de plausibilité avec une propagation des incertitudes aléatoires et épistémiques de paramètres.

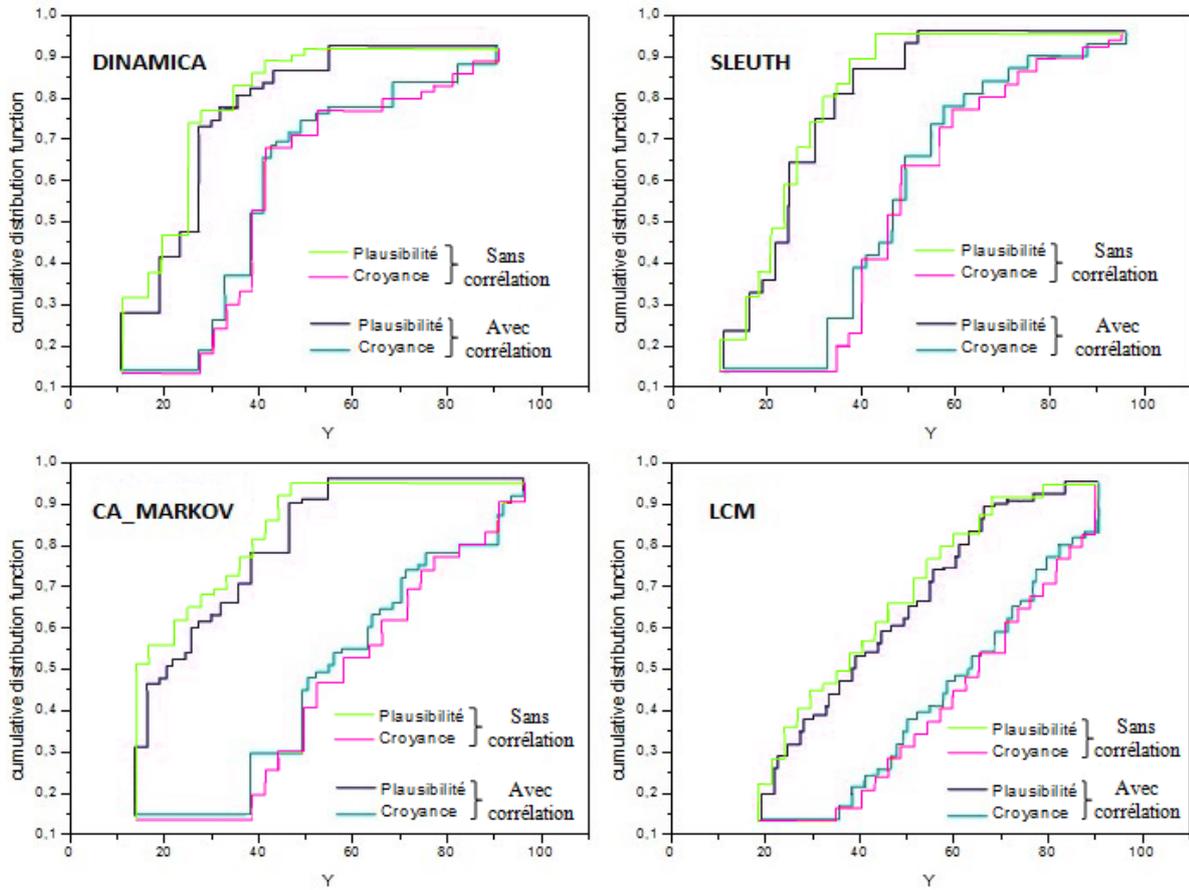


FIGURE 41 – Sorties de quatre modèles de COS représentées par des distributions de croyance et de plausibilité avec une propagation des incertitudes tout en considérant les corrélations entre les paramètres en comparaison aux celles où la corrélation n'est pas prise en compte.

statistiquement significatives entre les paramètres de forme (voir Tableau 13). De manière générale, l'indice de forme (SI) est significativement corrélé avec la surface (A), la densité (D), l'ajustement rectangulaire (RF) et l'ajustement elliptique (EF).

De plus, nous avons trouvé des corrélations statistiquement significatives entre le paramètre NDVI et les paramètres climatiques (voir Tableau 13). Le paramètre NDVI a une plus forte corrélation avec la température et l'humidité. Aussi, nous notons que les paramètres spectraux sont corrélés avec les paramètres de texture.

La Figure 41 montre une comparaison de la distribution de la sortie lorsque les corrélations entre les paramètres sont considérées par rapport à la distribution de la sortie lorsque la corrélation entre les paramètres n'est pas prise en compte dans les modèles de COS. Sur la base de l'approche proposée de l'étude de corrélation (voir chapitre 3), les résultats montrent qu'il n'y a pas une grande différence entre la propagation de l'incertitude avec et sans prise en compte de la corrélation entre les paramètres. Cependant, il est important de distinguer la contribution corrélée et non corrélée de l'incertitude pour un paramètre individuel, ce qui aiderait les ingénieurs à décider s'ils devaient se concentrer sur les variations d'un ensemble de paramètres ou de chaque paramètre à part.

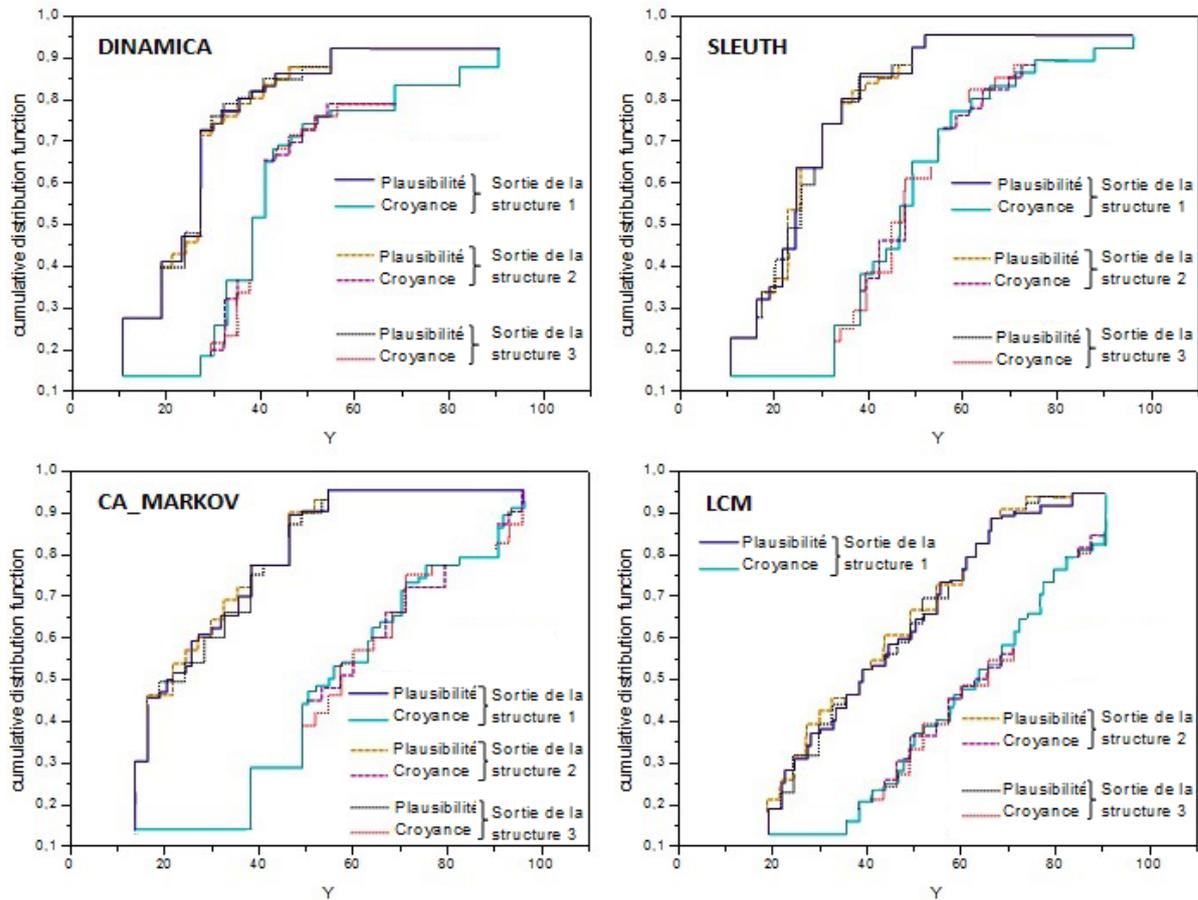


FIGURE 42 – Sorties de trois structures différentes des quatre modèles de COS représentées par des distributions de croyance et de plausibilité.

2.4.4 Etape 4 : Propagation des incertitudes de la structure de modèle

La compréhension de la façon dont les différentes sources d'incertitude influencent les résultats des modèles est importante. Afin d'illustrer l'importance de l'incertitude de la structure du modèle, nous utilisons trois structures différentes pour chaque modèle de COS utilisé. En conséquence, nous obtenons trois structures différentes avec des hypothèses différentes pour chaque modèle. Pour tenir compte des incertitudes liées à la structure du modèle dans le résultat final, les incertitudes liées aux paramètres sont d'abord propagées à travers chaque structure du modèle et ceci pour chaque modèle choisi. Ensuite, nous obtenons trois représentations incertaines de COS dans chaque modèle, qui sont présentées dans la Figure 42. La différence entre ces trois représentations illustre l'impact de l'incertitude de la structure du modèle. Par rapport aux résultats des structures originales des modèles, nous pouvons constater que cette différence est importante.

Nous remarquons que les résultats obtenus pour les modèles DINAMICA, SLEUTH, CA-MARKOV et LCM sont très similaires car nous n'avons pas utilisé des hypothèses très différentes par rapport à celles d'origine.

La propagation des incertitudes des différentes structures du modèle donne des indications sur la contribution de ces structures à la variable de sortie.

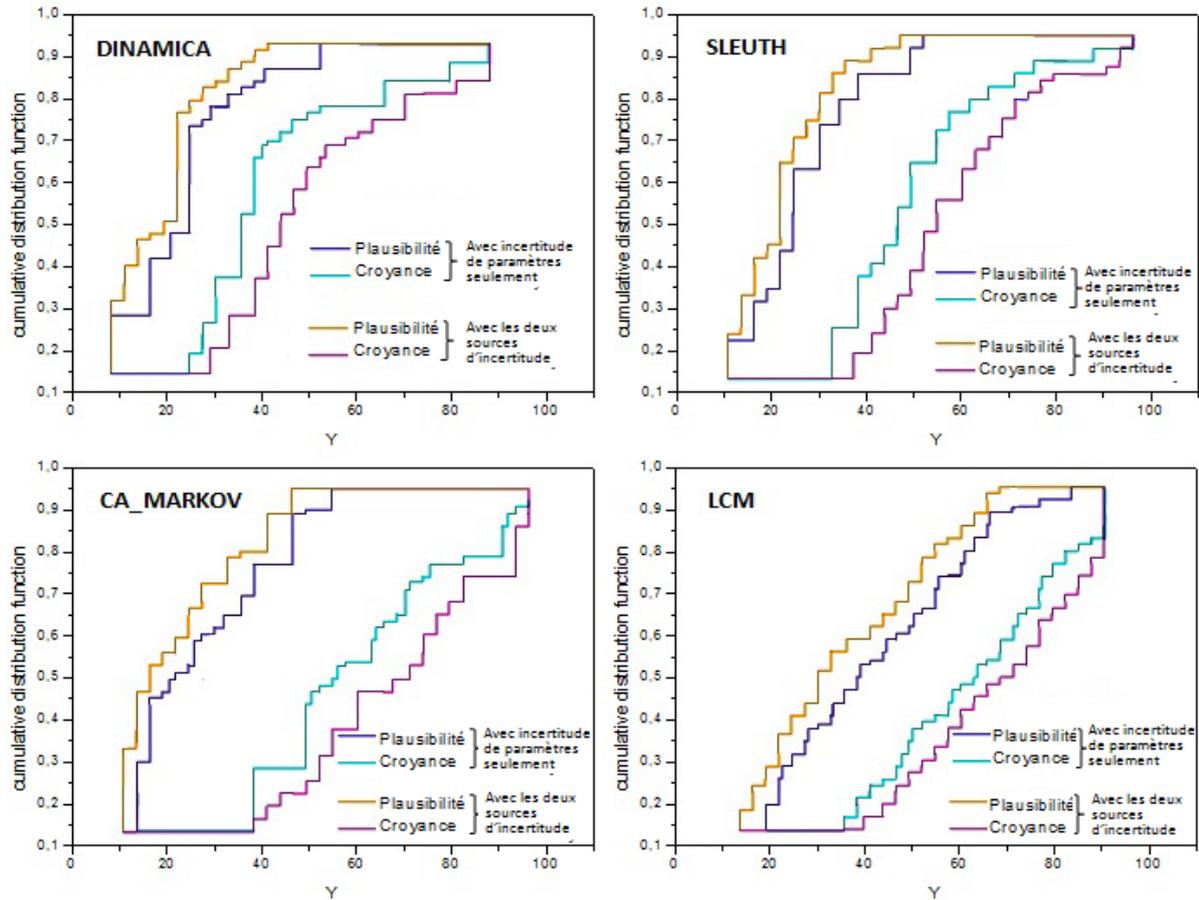


FIGURE 43 – Sorties des quatre modèles de COS représentées par des distributions de croyance et de plausibilité avec la propagation des incertitudes de paramètres et de la structure du modèle en comparaison à celles causées par la propagation de paramètres seulement.

La Figure 43 montre une comparaison des distributions de croyance et de plausibilité des quatre modèles DINAMICA, SLEUTH, CA-MARKOV et LCM tout en considérant à la fois les incertitudes liées aux paramètres et à la structure du modèle, et en considèrent uniquement les incertitudes de paramètres. Pour tous les modèles, nous remarquons que l'effet intégré de l'incertitude représentant à la fois les paramètres et la structure du modèle conduit à une plus grande incertitude liée au COS en comparaison au cas de la propagation des incertitudes de paramètres seulement.

Dans notre étude, nous remarquons une plus grande incertitude de paramètres entraîne également une plus grande incertitude de la structure du modèle.

Nous concluons que la propagation des incertitudes démontrée dans cette thèse aide significativement les utilisateurs à choisir le modèle approprié de COS.

2.5 Résultats d'application de l'approche proposée de la réduction des incertitudes

Une fois que les incertitudes de sortie sont connues, des méthodes sont appliquées pour réduire ces incertitudes dans les modèles de COS. Le processus de réduction des

incertitudes comporte deux étapes principales : l'analyse de sensibilité par la stratégie du pincement et l'estimation par les limites de confiance de Kolmogorov-Smirnov.

2.5.1 Etape 5 : Etude de sensibilité

L'identification des variables les plus influentes qui affectent la stabilité de sortie des modèles de COS en utilisant une analyse de sensibilité facilitera un meilleur fonctionnement et un meilleur contrôle avec moins de surveillance par les opérateurs de système.

Dans cette thèse, l'analyse de sensibilité par la stratégie du pincement est appliquée pour tester l'effet de 16 paramètres incertains et des trois structures différentes à travers les modèles DINAMICA, SLEUTH, CA-MARKOV et LCM.

La Figure 44 illustre les résultats de l'analyse de sensibilité. Les différentes hauteurs des barres révèlent les différents niveaux de sensibilité. Une barre longue indique une plus grande sensibilité de paramètre. Les variations des paramètres sont représentées individuellement pour chacune des trois structures pour chaque modèle de COS. La structure du modèle la plus influente montre généralement une plus grande sensibilité de paramètres. Les structures 1 et 2 de modèles DINAMICA, SLEUTH et CA-MARKOV ont donné presque les mêmes résultats. En revanche, les paramètres de la structure 3 sont relativement sensibles par rapport aux structures 1 et 2. Dans le modèle LCM, les structures 1 et 3 ont donné presque les mêmes résultats, tandis que les paramètres de la structure 2 sont relativement sensibles en comparaison aux structures 1 et 3.

Les résultats indiquent que les incertitudes représentant la structure du modèle jouent un rôle important pour la réalisation de l'analyse de sensibilité et pour la réduction des incertitudes dans les modèles de COS. La contribution globale de paramètres BI, MR, MN, RF, EF, SI, D et NDVI, qui sont les plus influents et indiquent une grande sensibilité avec une gamme de 0.2708-0.5593, 0.2997-0.5999, 0.3818-0.6919 et 0.2908-0.5801, respectivement, aux modèles originaux DINAMICA, SLEUTH, CA-MARKOV et LCM. Les paramètres RF, EF, SI et D présentent des tendances de variation similaires entre eux car ce sont des paramètres corrélés pour la représentation spatiale des objets. En général, ces paramètres jouent un rôle de plus en plus important dans la performance du modèle. Les paramètres MB, MG, Ctr, Ent, ASM, A, Tem et Hum présentent un effet relativement important avec une gamme de 0.019-0.090, 0.025-0.100, 0.100-0.260 et 0.008-0.120 aux modèles originaux DINAMICA, SLEUTH, CA-MARKOV et LCM, respectivement. Ces paramètres avec peu d'effet sur l'incertitude des sorties du modèle (par exemple, les valeurs de sensibilité des paramètres sont moins de 0.260) sont considérés comme non influents.

Malgré ces différents résultats obtenus par les quatre modèles de COS, nous avons obtenu la même classification d'importance de paramètres. Dans la structure originale du modèle DINAMICA, les valeurs de sensibilité des huit paramètres les plus sensibles étudiés, du plus élevée au plus faible, sont 0.5593, 0.4997, 0.4811, 0.4807, 0.4501, 0.4329, 0.2812 et 0.2708 pour BI, RF, D, NDVI, MN, SI, EF et MR, respectivement. À l'instar des résultats présentés par la structure originale du modèle SLEUTH, les paramètres BI, RF, D, NDVI, MN, SI, EF et MR montrent la plus grande sensibilité à la performance du modèle, avec une valeur minimale supérieure à 0.29, représentée par le paramètre MR. Selon la structure original du modèle CA-MARKOV, les valeurs de sensibilité, du plus

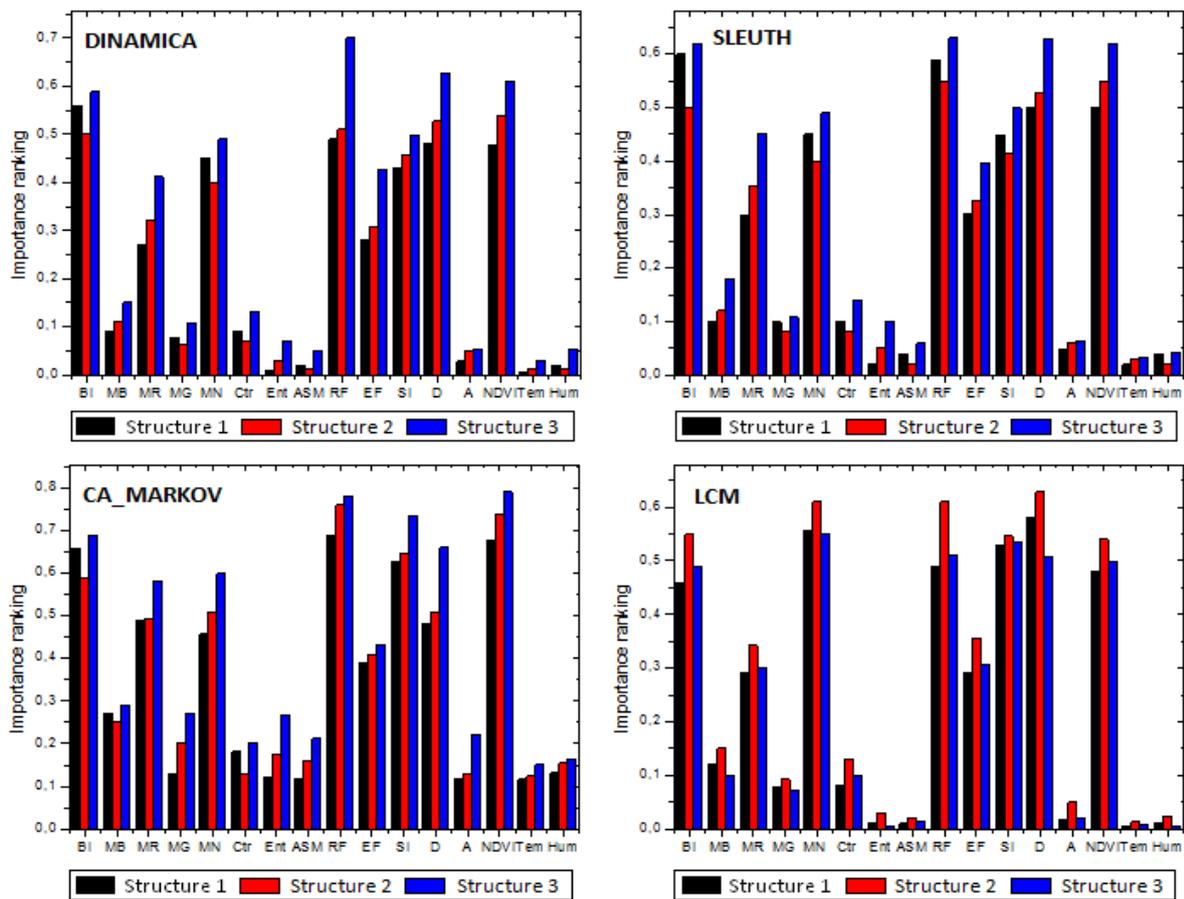


FIGURE 44 – Classement de l'importance de tous les paramètres d'entrée à travers trois structures différentes pour les quatre modèles de COS.

Rank	DINAMICA	SLEUTH	CA-MARKOV	LCM
1	BI	BI	RF	D
2	RF	RF	NDVI	MN
3	D	D	BI	SI
4	NDVI	NDVI	SI	RF
5	MN	MN	MR	NDVI
6	SI	SI	D	BI
7	EF	EF	MN	EF
8	MR	MR	EF	MR

TABLE 14 – Top huit paramètres les plus importants identifiés par les structures originales de modèles de COS DINAMICA, SLEUTH, CA-MARKOV et LCM.

élevée au plus faible, sont 0.6919, 0.6875, 0.6571, 0.6189, 0.4729, 0.4511, 0.4462 et 0.3818 pour RF, NDVI, BI, SI, MR, D, MN et EF, respectivement. En outre, ces paramètres présentent également une grande sensibilité dans la structure originale du modèle LCM car ils présentent une sensibilité élevée supérieure à 0.55, représentée par la densité D, dans la période de simulation. Selon les résultats obtenus, une forte augmentation des valeurs de sensibilité est observée à partir du modèle CA-MARKOV.

Les huit paramètres les plus importants qui sont identifiés par les modèles DINAMICA, SLEUTH, CA-MARKOV et LCM sont présentés dans le Tableau 14.

Les modèles DINAMICA et SLEUTH ont donné le même classement pour les paramètres les plus influents (les huit paramètres ont le même ordre pour les deux modèles). Les modèles CA-MARKOV et LCM donnent les mêmes huit paramètres les plus influents, mais avec quelques différences de positionnement. BI est le paramètre le plus influent par DINAMICA et SLEUTH. CA-MARKOV surestime RF comme 1^{er} et NDVI comme 2^{me} paramètres les plus influents et ces paramètres sont classés comme 2^{me} et 4^{me} respectivement pour les modèles DINAMICA et SLEUTH. Le modèle LCM identifie les deux premiers paramètres comme D et MN, respectivement, et les 3^{me} et 4^{me} paramètres comme SI et RF, respectivement. Le classement prioritaire de tous les paramètres incertains a été évalué à travers les quatre structures de modèles utilisés. Il est clairement évident que tous les modèles ne produisent pas les mêmes niveaux d'importance pour les différents paramètres, mais que les paramètres les plus influents peuvent être identifiés à l'aide de ces différents modèles. Les résultats montrent également que la stratégie du pincement d'analyse de sensibilité est plus adaptée, offrant un bon équilibre entre précision et efficacité.

Après avoir appliqué le processus d'analyse de sensibilité, nous considérerons seulement les structures et les paramètres (BI, MR, MN, RF, EF, SI, D et NDVI) du modèle les plus influents pour une estimation optimale des sources d'incertitude.

2.5.2 Etape 6 : Estimation

Après avoir identifié les variables les plus influentes sur les résultats des modèles de COS, les limites de confiance de Kolmogorov-Smirnov sont utilisées pour les estimer afin d'obtenir des valeurs optimales.

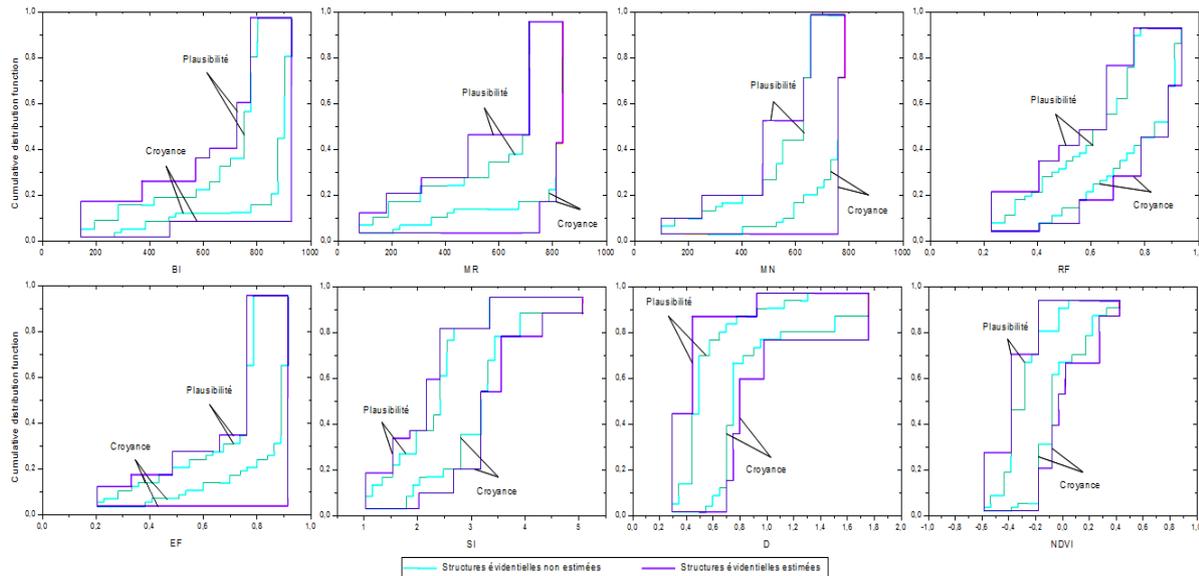


FIGURE 45 – Structures évidentielles estimées des paramètres les plus influents en comparaison avec celles non estimées.

Les résultats de l'estimation des paramètres sont présentés dans la Figure 45.

Les limites de confiance sont également utilisées pour estimer les structures du modèle les plus influentes.

Après avoir appliqué les méthodes de la réduction, la propagation des incertitudes par la théorie des fonctions de croyance est appliquée à nouveau pour exécuter les quatre modèles de COS pour générer les nouvelles plages d'incertitude de prédiction. La Figure 46 montre les résultats finaux. Les plages d'incertitude de sortie des quatre modèles avant l'implémentation de l'approche proposée sont relativement importantes (Figure 43), alors que ces plages sont relativement faibles par l'application de l'approche proposée (Figure 46).

Les résultats obtenus dans cette thèse montrent que les quatre modèles de COS ont une bonne précision et peuvent bien représenter les changements réels avec des erreurs minimales. De plus, l'approche proposée à travers le modèle DINAMICA entraîne une réduction globale importante de l'incertitude en comparaison aux autres modèles.

Enfin, nous pouvons conclure que l'incertitude est inhérente pour les deux composants principaux de la modélisation de COS. Les travaux futurs devraient tenir compte plus explicitement de ces sources d'incertitude pour améliorer la robustesse des résultats des modèles de COS.

2.6 Résultats des cartes de prédiction de COS

La validation des modèles de prédiction de COS est généralement effectuée en comparant le résultat simulé à la carte empirique pour déterminer la capacité de prédiction des modèles de COS. Dans cette thèse, pour prédire et analyser la croissance actuelle et future de la région du Caire, de nombreuses images satellitaires Landsat sont utilisées. Ces images sont obtenues à des différentes dates (1987, 2003 et 2014) et sont classi-

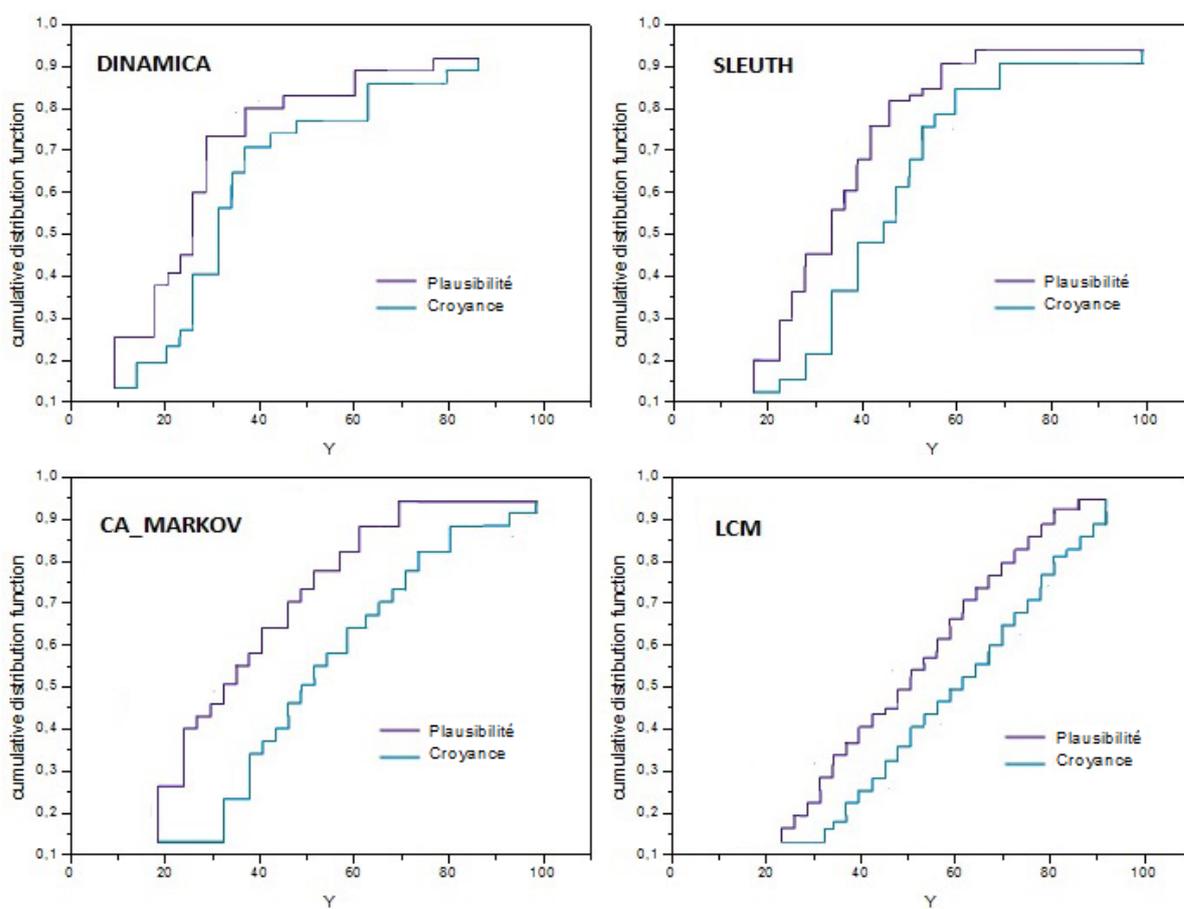


FIGURE 46 – Sorties des quatre modèles de COS représentées par des distributions de croyance et de plausibilité avec l'application de l'approche de la réduction des incertitudes.

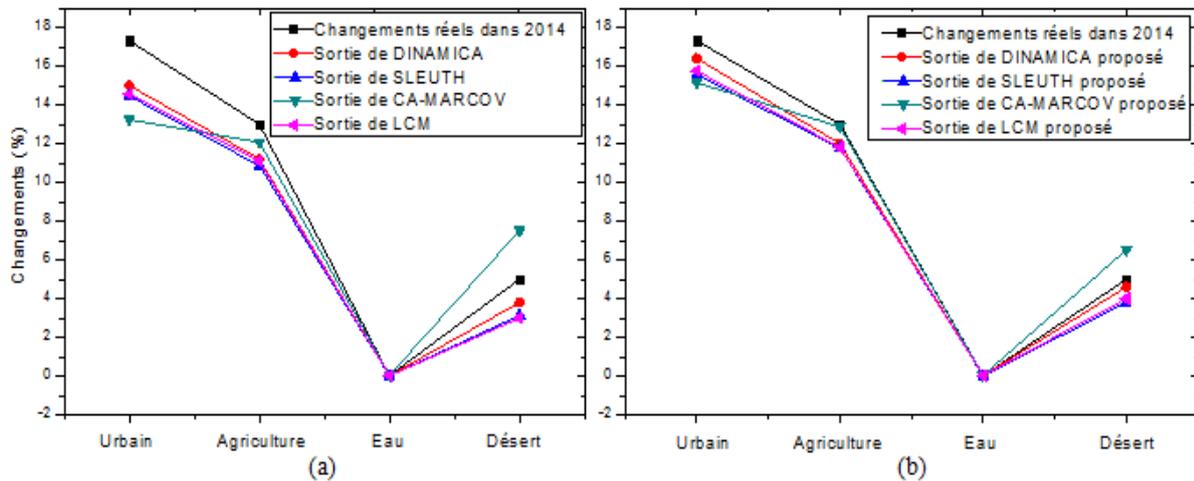


FIGURE 47 – Sorties des quatre modèles de COS en comparaison aux changements réels entre 1987 et 2014 de la région du Caire : (a) résultats sans application de l'approche proposée et (b) résultats avec application de l'approche proposée.

fiées pour produire des cartes de COS. Pour valider, l'analyse de détection de COS est utilisée pour analyser les changements au fil du temps. Ces changements sont enregistrés dans une matrice de transition qui représente l'entrée aux deux étapes : calibration et modélisation des changements. Après cela, les changements entre 1987 et 2003 sont simulés en utilisant les modèles DINAMICA, SLEUTH, CA-MARKOV et LCM tout en appliquant l'approche proposée de la propagation et la réduction des incertitudes pour produire une carte prédite de 2014. Les cartes obtenues sont comparées à la carte réelle de changement pour la même année, dérivée d'images satellitaires, afin de valider l'approche proposée. L'étape suivante consiste à simuler les changements en 2025, en supposant que les changements entre 2003 et 2014 se poursuivront au cours des 11 prochaines années.

Afin d'évaluer la précision de l'approche proposée, la Figure 47(a) indique une comparaison entre les changements simulés des modèles DINAMICA, SLEUTH, CA-MARKOV et LCM et les changements réels pour les différents types d'occupation des sols entre 1987 et 2014 sans application de l'approche proposée. Les comparaisons entre les prédictions simulées à travers les quatre modèles choisis et les observations avec l'application de l'approche proposée sont présentées dans la Figure 47(b). Comme nous l'avons noté, les prédictions des modèles SLEUTH et LCM ne présentent pas de différences significatives. En plus, les résultats montrent que l'approche proposée à travers le modèle DINAMICA peut prédire le changement de la zone du Caire mieux que les modèles SLEUTH, CA-MARKOV et LCM. Alors que, CA-MARKOV donne des meilleurs résultats pour les changements des terres agricoles.

La comparaison entre les changements réels et simulés est basée sur les critères statistiques R^2 , NSE et $PBIAS$. La Figure 48 montre la performance de l'approche proposée de la propagation et de la réduction des incertitudes à travers les modèles DINAMICA, SLEUTH, CA-MARKOV et LCM pour l'ensemble de critères définis dans les deux périodes de validation (1987-2003 et 2003-2014).

DINAMICA, avec l'application de l'approche proposée, donne les meilleurs résultats pour simuler les COS avec les valeurs de R^2 , NSE et $PBIAS$ sont 0,93, 0,89 et 0,25,

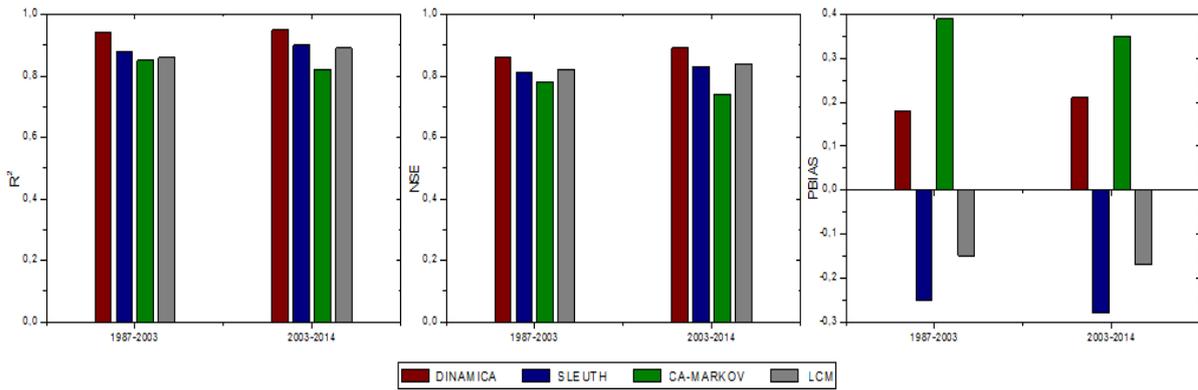


FIGURE 48 – Comparaison entre les quatre modèles de COS pour les critères de performance R^2 , NSE et $PBIAS$ dans les deux périodes de validation.

respectivement, pendant la période de validation 2003-2014 dans la région du Caire. En outre, il n'y a pas de différences évidentes dans la performance de modèle DINAMICA avec l'application de l'approche proposée et l'application de cette dernière à travers les autres modèles pour prédire les changements de la zone étudiée. Les résultats ont montré que les modèles SLEUTH, CA-MARKOV et LCM avec l'application de l'approche proposée sont bons selon les critères définis. En plus, les résultats montrent un bon accord entre les cartes d'occupation des sols réelles et simulées, et cela montre que les méthodes proposées sont fiables pour la région du Caire. Ainsi, ces méthodes peuvent être utilisées pour prédire des changements futurs d'occupation des sols. Selon les résultats obtenus, les tendances les plus importantes dans les deux périodes (1987-2003 et 2003-2014) sont les transitions de l'agriculture et de désert vers des zones urbaines.

Plus de 16 ans, de 1987 à 2003, l'agriculture a perdu 12,7% vers les zones urbaines et a perdu presque le même pourcentage (12%) dans 11 ans seulement, de 2003 à 2014. Cela implique la transformation des terres agricoles au Caire en zones urbaines. En outre, 2,8% des zones désertiques sont devenues urbaines entre 1987 et 2003. Ce pourcentage a été augmenté à 5% entre 2003 et 2014, résultant de l'application de stratégies de reconstruction du désert à la construction de nouvelles colonies urbaines en dehors de la vallée du Nil.

Après cela, l'étape suivante consiste à simuler les changements en 2025, en supposant que les changements entre 1987 et 2014 continueront au cours des 11 prochaines années. 13% de l'agriculture et 3,8% du désert en 2014 devraient se transformer en zones urbaines en 2025. La Figure 49 montre l'étalement urbain au Caire en 1987, 2003, 2014 et les zones urbaines prédites en 2025 à travers les modèles DINAMICA, SLEUTH, CA-MARKOV et LCM avec l'application de l'approche proposée.

La construction de terrains résidentiels a connu une évolution de plus de 27 ans, de 1987 à 2014. Cela pourrait être attribué à l'augmentation de la population en raison de la demande accrue de terres urbaines. Selon les résultats obtenus à travers le modèle DINAMICA, les zones urbaines de la région du Caire devraient augmenter à environ 22% de la zone en 2025. Cette vaste croissance de l'urbanisation non planifiée représente une grave menace pour les systèmes socio-écologiques. Ceci implique qu'il existe un risque important de perdre les régions agricoles ce qui conduit à l'augmentation de la menace de l'insécurité alimentaire. Cela indique un risque actuel et à venir pour le patrimoine

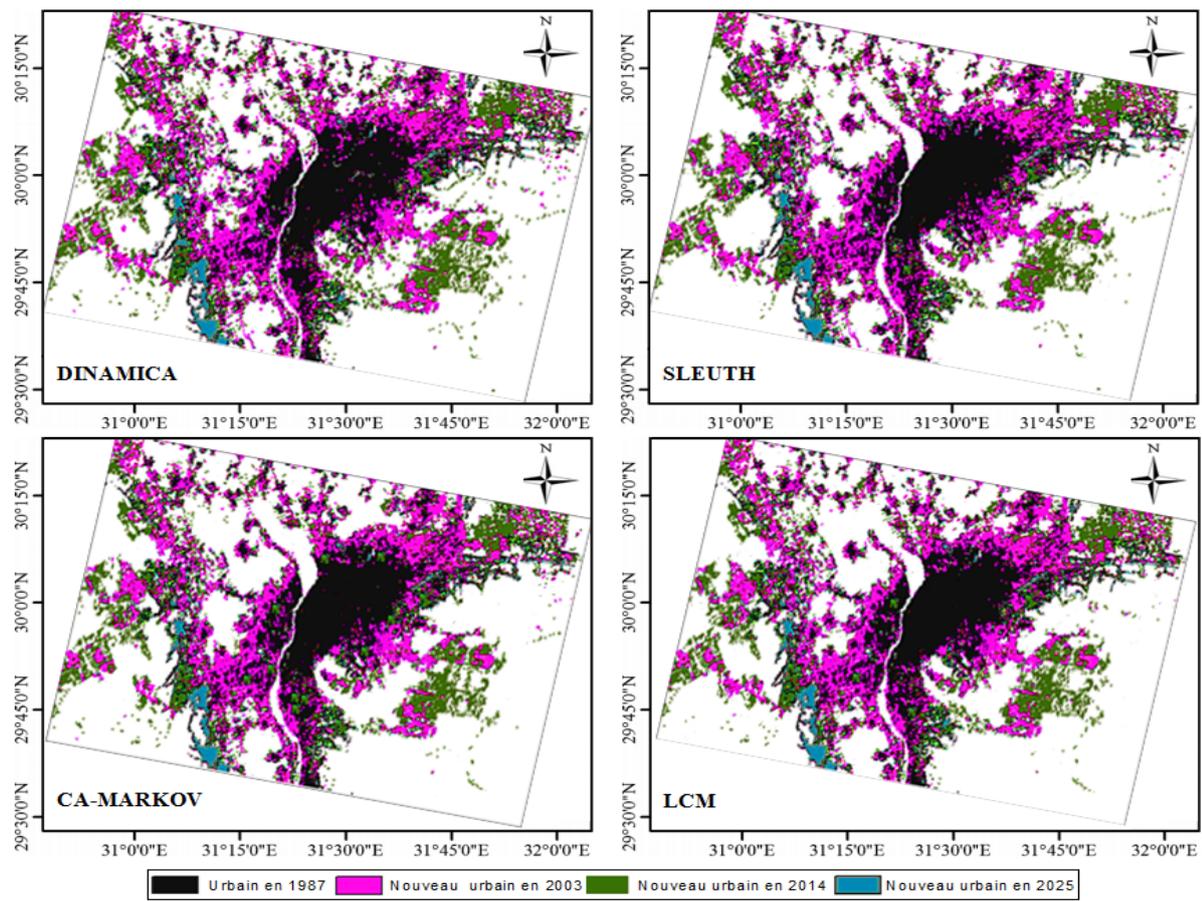


FIGURE 49 – Croissance urbaine dans la région du Caire à travers les modèles DINAMICA, SLEUTH, CA MARKOV et LCM.

culturel de la région du Caire, en particulier avec la croissance de la population humaine dans la capitale égyptienne.

3 Evaluation de l'approche proposée

Pour montrer la performance de l'approche proposée, il est important de la comparer avec d'autres approches probabilistes ou non probabilistes.

Dans cette thèse, nous comparons l'approche proposée de la propagation et de la réduction des incertitudes avec deux approches existantes [Boulila *et al.* 2014, Ferchichi *et al.* 2016]. L'approche proposée dans [Ferchichi *et al.* 2016] permet de propager et de réduire les incertitudes liées aux paramètres et à la structure des modèles de COS par la théorie de possibilités. Le modèle de COS, qui est utilisé dans cette approche, a été développé par Boulila *et al.* [Boulila *et al.* 2011]. Ce modèle est basé sur une approche d'apprentissage automatique, l'arbre de décision, pour construire des prédictions de COS dans la région du Caire. L'approche présentée dans [Ferchichi *et al.* 2016] consiste en quatre étapes principales :

- Modélisation de paramètres : cette étape permet de modéliser tous les paramètres choisis par des distributions de possibilités tout en considérant que tous les paramètres sont de types épistémiques.
- Propagation des incertitudes de paramètres : cette étape permet de propager les incertitudes liées aux paramètres par la théorie de possibilités sans tenir en compte les corrélations entre les paramètres.
- Propagation des incertitudes de la structure de modèle : cette étape permet de propager les incertitudes liées à la structure de modèle de COS par la théorie de possibilités.
- Etude de sensibilité : cette étape permet de performer une analyse de sensibilité par la mesure de Hartley-like pour identifier les sources des incertitudes les plus importantes.

L'approche proposée dans [Boulila *et al.* 2014] permet de propager et de réduire les incertitudes liées seulement aux paramètres d'entrée des modèles de COS par une méthode probabiliste basée sur l'expansion stochastique. Cette approche a également utilisé le modèle de COS présenté dans [Boulila *et al.* 2011] pour simuler des prédictions de COS. L'approche présentée dans [Boulila *et al.* 2014] consiste en trois étapes principales :

- Modélisation de paramètres : cette étape permet de modéliser tous les paramètres choisis par des distributions de probabilités tout en considérant que tous les paramètres sont de types aléatoires.
- Propagation des incertitudes de paramètres : cette étape permet de propager les incertitudes liées aux paramètres par la méthode probabiliste de collocation sans tenir en compte les corrélations entre les paramètres.
- Etude de sensibilité : cette étape permet de performer une analyse de sensibilité par les indices de Sobol pour identifier les paramètres les plus importants.

Le Tableau 15 représente les pourcentages de changement des quatre types d'occupation des sols (urbain, agriculture, eau et désert). Ce Tableau montre la différence entre

	Urbain(%)	Agriculture(%)	Eau (%)	Désert(%)
Changements réels en 2014	17.32	13.00	0.02	5.00
Approche présentée dans [Ferchichi <i>et al.</i> 2016]	15.63	13.80	0.01	4.03
Approche présentée dans [Boulila <i>et al.</i> 2014]	15.21	14.19	0.01	6.37
Notre approche proposée appliquée à DINAMICA	16.58	12.36	0.02	4.29
Notre approche proposée appliquée à SLEUTH	16.02	12.11	0.02	4.11
Notre approche proposée appliquée à CA-MARKOV	15.91	12.84	0.02	6.18
Notre approche proposée appliquée à LCM	16.27	12.19	0.02	4.21

TABLE 15 – Comparaison entre les changements réels, les changements obtenus par l’application de l’approche proposée à travers quatre modèles de COS et ceux obtenus par les approches présentées dans [Ferchichi *et al.* 2016] et [Boulila *et al.* 2014].

les changements réels, les changements de notre approche proposée à travers les quatre modèles de COS (DINAMICA, SLEUTH, CA-MARKOV et LCM) et ceux obtenus par les approches présentées dans [Ferchichi *et al.* 2016] et [Boulila *et al.* 2014].

Nous remarquons que les résultats obtenus dans le Tableau 5-4 indiquent que l’approche proposée produit de meilleurs résultats dans les quatre modèles de COS en comparaison à ceux obtenus par les approches des références. Ceci montre d’une part l’importance de la théorie des fonctions de croyance, mais d’autre part l’importance de l’intégration des incertitudes de paramètres et de la structure du modèle, de la distinction entre les incertitudes aléatoires et épistémiques, et de la considération de la corrélation entre les paramètres.

4 Conclusion

Dans ce chapitre, nous avons présenté un scénario d’application de l’approche proposée de la propagation et de la réduction des incertitudes dans les modèles de COS. Ce scénario est divisé en six étapes : modélisation des incertitudes, étude de corrélation, propagation des incertitudes de paramètres, propagation des incertitudes de la structure du modèle, étude de sensibilité et estimation.

Le scénario d’application a pour objectif la prédiction des changements de la zone urbaine de la région de Cairo entre les dates 1987 et 2014. Les résultats obtenus montrent que les zones d’agriculture et désertiques sont les zones les plus concernées par l’étalement urbain.

Afin d’évaluer les performances de notre approche, nous avons choisi une approche de propagation et de réduction des incertitudes basée sur la théorie de possibilités. L’évaluation a montré les bonnes performances de notre approche.

Conclusion générale et perspectives

L'objectif de cette thèse est de proposer une démarche de propagation et de réduction des incertitudes pour les modèles de COS. L'approche proposée permet d'améliorer les décisions concernant les COS et ceci par la modélisation des incertitudes qui accompagnent le processus de prédiction de changements.

Dans ce qui suit, nous synthétisons nos travaux réalisés au cours de cette thèse. Les perspectives de recherche qui pourraient être envisagées suite à ces travaux sont également discutées.

Synthèse

Les travaux de cette thèse constituent à la modélisation des incertitudes qui accompagnent le processus de COS. Les incertitudes sont de plusieurs types et sources. Peu de travaux dans la littérature se sont concentrés sur la propagation et la réduction des incertitudes liées aux modèles de COS. L'ensemble de ces travaux se fondent sous les méthodes probabilistes.

Dans cette thèse, dans un premier temps, nous avons étudié les différents types et sources d'incertitudes rencontrées lors de l'analyse d'un modèle de COS. Généralement, les incertitudes sont dues à une variabilité naturelle ou à un manque de connaissance, et sont liées aux paramètres d'entrée du modèle et à la structure du modèle. Par la suite, nous nous sommes intéressés à la problématique de modélisation et de propagation de ces types et sources des incertitudes dans les modèles de COS. Nous avons proposé une approche de propagation des incertitudes basée sur la théorie des fonctions de croyance. Nous commençons par modéliser les incertitudes liées aux paramètres et à la structure du modèle de COS. Ensuite, les incertitudes (aléatoires/épistémiques) liées aux paramètres sont tout d'abord propagées tout en considérant la corrélation entre les paramètres à travers les modèles de COS. Après cela, les incertitudes liées à la structure du modèle sont également propagées par la théorie des fonctions de croyance. La propagation des incertitudes liées à la structure du modèle, quant à elle, est effectuée d'abord par la propagation des incertitudes de paramètres.

Dans un second temps, nous avons proposé une approche de réduction des incertitudes dans les modèles de COS composée (i) premièrement d'une analyse de sensibilité basée sur la stratégie du pincement pour savoir quelles sont les sources d'incertitude qui contribuent le plus à l'incertitude de la sortie des modèles de COS. Cette analyse permet de hiérarchiser les sources d'incertitude de l'étude, c'est à dire de cibler les paramètres et les structures du modèle les plus influents sur la sortie et nécessitant le plus de précision ; (ii) deuxièmement, d'une méthode d'estimation basée sur les limites de confiance du Kolmogorov-Smirnov pour optimiser les sources sélectionnées ; et (iii) troisièmement, des critères d'évaluation tels que R^2 , NSE et $PBIAS$ pour évaluer la performance des modèles de COS.

Par la suite, notre approche a été expérimentée pour valider et évaluer leur efficacité et leur pertinence. Pour ce faire, nous avons choisi comme site d'étude la zone du Caire et nous avons utilisé des images satellitaires provenant de la base de données en ligne de la USGS. A travers les expérimentations réalisées, nous avons montré que les méthodes de propagation et de réduction des incertitudes sont des éléments nécessaires pour le paramétrage et le calibrage des modèles de COS et pour une prédiction de la

performance dans un intervalle de confiance donné. L'approche proposée devrait offrir une évaluation raisonnable et réaliste des systèmes terrestres étudiés face à divers types et sources des incertitudes. L'existence d'incertitude, qui est inévitable, ne devrait pas restreindre la recherche et la prise de décision, mais au contraire devrait aider les utilisateurs de modèles de COS à compléter leurs connaissances par rapport au système étudié, en identifiant les informations clés nécessaires. Les résultats obtenus dans cette thèse montrent que l'application de la théorie des fonctions de croyance pour la propagation des incertitudes permet d'améliorer la précision des modèles de COS. Nous avons également montré l'importance de considérer la corrélation entre les paramètres d'entrée des modèles de COS. Le degré d'incertitude varie d'un modèle à un autre et le choix de la structure du modèle de COS est encore plus important. Ainsi, l'effet de l'incertitude ne doit pas être négligé lors de la modélisation des prédictions de COS. Ainsi, notre objectif est de déterminer un bon compromis entre la précision et le coût du calcul des modèles de COS. Ainsi, l'analyse de sensibilité peut simplifier ce problème et produire plus de précision sur le COS. L'analyse de sensibilité s'est avérée efficace et robuste pour identifier les paramètres et les structures qui ont un effet important sur la sortie des modèles de COS. Dans cette thèse, les résultats de l'analyse de sensibilité montrent généralement que les incertitudes des paramètres et des structures considérés ont un impact significatif sur les décisions des modèles de COS. Le degré d'influence varie d'un paramètre à un autre et d'une structure à une autre. Mais, en général, l'ignorance ou l'évaluation inadéquate de l'incertitude dans les modèles de COS a un coût quantifiable. En plus, l'incorporation de corrélations entre les paramètres dans l'analyse de sensibilité fournit des prédictions plus réalistes sur le COS. L'analyse de sensibilité basée sur la stratégie du pincement, pour les modèles de COS est également très importante. Par conséquent, l'analyse faite dans cette thèse peut être très utile pour identifier les sources d'incertitudes les plus pertinentes qui nécessitent une attention particulière dans les pratiques de planification.

Perspectives

Pour améliorer l'approche proposée, nous pouvons évoquer des perspectives et des pistes préliminaires quant à de futures recherches. Parmi celles-ci, nous pouvons citer :

- Amélioration de l'approche. Dans cette thèse, nous avons appliqué l'approche proposée sur quatre modèles différents de COS à savoir DINAMICA, SLEUTH, CA-MARKOV et LCM afin de monter la performance de notre travail. Ces modèles sont basés sur les automates cellulaires, la régression logistique et les chaînes de Markov. Les travaux ultérieurs devraient également appliquer l'approche proposée sur d'autres modèles très populaires de COS tels que les modèles basés sur l'apprentissage automatique.
- La mise en oeuvre des méthodes basées sur la construction du méta-modèle pour améliorer le temps de calcul de traitement des incertitudes. Pour réduire les temps de calculs, nous pouvons utiliser un méta-modèle afin d'approcher les résultats de simulation mis en oeuvre avec l'optimisation.
- Mettre en ligne un outil de propagation et de réduction des incertitudes basé sur la théorie des fonctions de croyance. Cet outil permettant aux chercheurs de fournir des données, dans un format particulier, et d'obtenir une réponse. Ceci aidera les chercheurs à améliorer les performances de leurs modèles de COS.

- Utilisation plus d'études de cas. En particulier, il serait bon d'avoir une configuration complètement différente. Dans cette thèse, pour valider l'approche proposée, nous avons choisi la région du Caire d'Égypte dans le but de suivre l'étalement urbain. Une alternative serait d'appliquer l'approche proposée sur d'autres scénarios comme la déforestation et l'intensification agricole. En effet, plusieurs travaux dans la littérature ont montré que les scénarios d'occupation des sols choisis sont l'une de principales sources des incertitudes. Ainsi, nous pouvons proposer des mécanismes pour propager et réduire cette source d'incertitude.
- Dans cette thèse, l'approche proposée est réalisée pour propager et réduire les incertitudes associées aux prédictions de COS. Ainsi, nous comptons appliquer notre approche sur d'autres domaines tels que la classification, la segmentation, le suivi, etc.
- Nous comptons modéliser les différents types d'incertitudes et étudier les corrélations dans le cadre de la théorie de possibilités. Ceci peut montrer l'efficacité de l'approche proposée.

Liste des publications

Liste des journaux

- Propagating aleatory and epistemic uncertainty in land cover change prediction process, *Ecological Informatics* 37 : 24-37 (2017).
- Reducing uncertainties in land cover change prediction models using sensitivity analysis, *Knowledge and Information Systems* 1-22 (2017).
- Towards an uncertainty reduction framework for land-cover change prediction using possibility theory, *Vietnam Journal of Computer Science* 4(3) 195-209 (2016).

Liste des conférences

- Improvement of LCC Prediction Modeling Based on Correlated Parameters and Model Structure Uncertainty Propagation. *IJCCI 2015* : 270-290.
- Using Evidence Theory in Land Cover Change Prediction to Model Imperfection Propagation with Correlated Inputs Parameters. *IJCCI (FCTA) 2015* : 47-56.
- An Intelligent Possibilistic Approach to Reduce the Effect of the Imperfection Propagation on Land Cover Change Prediction. *ICCCI (2) 2015* : 520-529.
- Parameter and structural model imperfection propagation using evidence theory in land cover change prediction. *IPAS 2014* : 1-6.
- An approach based on adaptive decision tree for land cover change prediction in satellite images. *KDIR-KMIS 2013* : 82-90.

Liste des présentations

- Propagation d'imperfection de paramètres et de modèle : application à la prédiction de changements en imagerie satellitaire. *ATSIP 2015*, Monastir, Tunisie.
- Reducing Imperfection System in Land Cover Change Prediction Process. *TAIMA 2015*, Hammamet, Tunisie.
- Parameter and structural model imperfection propagation using evidence theory in land cover change prediction. *IPAS 2014*, Hammamet, Tunisie.
- Improving land cover change prediction : Adaptive Decision Tree Approach. *BFTA 2013*, Carthage, Tunisie.

Références

- [Abdallah *et al.* 2013]N.B. Abdallah, N. Mouhous-Voyneau et T. Denoeux (2013). Using Dempster-Shafer theory to model uncertainty in climate change and environmental impact assessments. *Int. Conf. Inf. Fusion* 2117–2124.
- [Aburas *et al.* 2016]M.M. Aburas, Y.M. Ho, M.F. Ramli et Z.H. Ash'aari (2016). The simulation and prediction of spatio-temporal urban growth trends using cellular automata models : A review. *International Journal of Applied Earth Observation and Geoinformation* 52 : 380–389.
- [Acosta-Michlik et Espaldon 2008]L. Acosta-Michlik et V. Espaldon (2008). Assessing vulnerability of selected farming communities in the Philippines based on a behavioural model of agent's adaptation to global environmental change. *Glob Environ Change* 18 : 554–563.
- [Agarwal *et al.* 2002]C. Agarwal, G.M. Green, J.M. Grove, T.P. Evans, C.M. Schweik, (2002). A Review and Assessment of Land-Use Change Models : Dynamics of Space, Time, and Human Choice. *US Department of Agriculture, Forest Service, 1*.
- [Agarwal et al. 2004]H. Agarwal, J.E. Renaud, E.L. Preston et D. Padmanabhan (2004). Uncertainty quantification using evidence theory in multidisciplinary design optimization. *Reliability Engineering and System Safety* 85(1-3) : 281–294.
- [Aguejdad et Houet 2008]R. Aguejdad et T. Houet (2008). Modeling of urban sprawl using the land change modeler on a French metropolitan Area (Rennes) : foresee the unpredictable. *In : Spatial Landscape Modelling : From Dynamic Approaches to Functional Evaluations*.
- [Alexander *et al.* 2015]P. Alexander, M.D.A. Rounsevell, C. Dislich, J.R. Dodson, K. Engström, D. Moran, (2015). Drivers for global agricultural land use change : The nexus of diet, population, yield and bioenergy. *Global Environmental Change* 35 : 138–147.
- [Ali *et al.* 2012]T. Ali, H. Boruah et P. Dutta (2012). Sensitivity Analysis in Radiological Risk Assessment using Probability Bounds Analysis. *International Journal of Computer Applications* 44(17) : 1–5.
- [Allen *et al.* 2013]T.R. Allen, Y. Wang et T.W. Crawford, (2013). Remote Sensing of Land Cover Dynamics. *Treatise on Geomorphology* 3 : 80–102.
- [Almeida *et al.* 2003]C.M. Almeida, M. Batty, A.M.V. Monteiro, G. Câmara, B.S. Soares-Filho, G.C. Cerqueira (2003). Stochastic cellular automata modeling of urban land use dynamics : Empirical development and estimation. *Computers, Environment and Urban Systems* 27 : 481–509.
- [Al-sharif et Pradhan 2015]A.A.A. Al-sharif et B. Pradhan (2015). Spatio-temporal prediction of urban expansion using bivariate statistical models : assessment of the efficacy of evidential belief functions and frequency ratio models. *Appl. Spat. Anal. Policy* 1–19.

- [Alvin *et al.* 1998]K. Alvin, W. Oberkampf, K. Diegertan et B. (1998). Rutherford Uncertainty quantification in computational structural dynamics : a new paradigm for model validation. *Proceedings of the 16th International Modal Analysis Conference* 3243 : 1191–1198.
- [André et Lopes 2012]J.C.S. André et D.R. Lopes 2012. On the use of possibility theory in uncertainty analysis of life cycle inventory. *The International Journal of Life Cycle Assessment* 17(3) : 350–361.
- [Arsanjani *et al.* 2013]J.J. Arsanjani, M. Helbich, W. Kainz et A.D. Bolorani (2013). Integration of logistic regression, Markov chain and cellular automata models to simulate urban expansion. *International Journal of Applied Earth Observation and Geoinformation* 21 : 265–275.
- [Arunraj *et al.* 2013]N.S. Arunraj, S. Mandal et J. Maiti (2013). Modeling uncertainty in risk assessment : An integrated approach with fuzzy set theory and Monte Carlo simulation. *Accident Analysis and Prevention* 55 :242–255.
- [Assaghir 2010]Z. Assaghir (2010) Analyse formelle de concepts et fusion d’informations : application à l’estimation et au contrôle d’incertitude des indicateurs agri-environnementaux. *Institut National Polytechnique de LorraineINPL*.
- [Aven et Zio 2011]T. Aven et E. Zio (2011) Some considerations on the treatment of uncertainties in risk assessment for practical decision making. *Reliability Engineering and System Safety* 96 (1) :64–74.
- [Ayyub et Klir 2006]B.M. Ayyub et G.J. Klir (2006). Uncertainty modeling and analysis in engineering and the sciences. CRC Press.
- [Bain et Engelhardt 1991]L. Bain et M. Engelhardt (1991). Introduction to Probability and Mathematical Statistic. 2nd ed., CA, Duxbury, Pacific Grove.
- [Balch 2012]M.S. Balch (2012). Mathematical foundations for a theory of confidence structures. *Int. J. Approx. Reason.* 53 (7) : 1003–1019.
- [Baraldi et Zio 2008]P. Baraldi et E. Zio (2008). A combined Monte Carlo and possibilistic approach to uncertainty propagation in event tree analysis. *Risk Analysis* 28 (5) :1309–1326.
- [Batty et Xie 1994a]M. Batty et Y. Xie (1994a). From cells to cities. *Environment and Planning B* 21 : 31–48.
- [Batty et Xie 1994b]M. Batty et Y. Xie (1994b). Urban analysis in a GIS environment : population density modeling using ARC/INFO. *In A.S. Fotheringham, P.A. Rogerson, Spatial analysis and GIS (pp. 189-219). London : Taylor and Francis.*
- [Batty *et al.* 1999][58] M. Batty, Y. Xie et Z. Sun (1999). Modeling urban dynamics through GIS-based cellular automata. *Computers, Environment and Urban Systems*23 : 205–233.
- [Baudrit *et al.* 2005]C. Baudrit, D. Guyonnet et D. Dubois (2005). Post-processing the hybrid method for addressing uncertainty in risk assessments. *Journal of Environmental Engineering-ASCE* 131 (12) :1750–1754.

- [Baudrit *et al.* 2006]C. Baudrit, D. Dubois et D. Guyonnet (2006). Joint propagation of probabilistic and possibilistic information in risk assessment. *IEEE Transactions on Fuzzy Systems* 14 (5) :593–608.
- [Bettemier 2010]O.H. Bettemier (2010). Error Estimation of Orthorectification of Small Satellite Images by Differential Sensitivity Analysis. *Journal of Aeronautics and Space Technologies* 4(4) : 65–74.
- [Blockley 2013]D. Blockley (2013). Analysing uncertainties : Towards comparing Bayesian and interval probabilities'. *Mechanical Systems and Signal Processing* 37(1-2) : 30–42.
- [Bonn et Rochon 1996]F. Bonn et G. Rochon (1996). Accurate Remote Sensing, Thematic Application. Quebec University, Quebec.
- [Boulila *et al.* 2011]W. Boulila, I.R. Farah, K.S. Ettabaa, B. Solaiman, H. Ben Ghézala (2011). A data mining based approach to predict spatiotemporal changes in satellite image. *International Journal of Applied Earth Observation and Geoinformation* 13 : 386–395.
- [Boulila *et al.* 2014]W. Boulila, A. Bouatay et I.R. Farah (2014). A probabilistic collocation method for the imperfection propagation : application to land cover change prediction. *J. Multimedia Process. Technol.* 5(1) : 12–32.
- [Breiman *et al.* 1984]L. Breiman, J.H. Friedman, R.A. Olshen et C.J. Stone (1984). Classification and Regression Trees. Belmont, CA : Wadsworth.
- [Bustince et Burillo 1995]H. Bustince et P. Burillo (1995). Correlation of interval-valued intuitionistic fuzzy sets. *Fuzzy Sets and Systems* 74 : 237–244.
- [Camerer et Weber 1992]C. Camerer et M. Weber (1992). Recent developments in modeling preferences : Uncertainty and ambiguity. *Journal of Risk and Uncertainty* 5(4) : 325–370.
- [Carlsson et Fullér 2001]C. Carlsson et R. Fullér (2001). On possibilistic mean value and variance of fuzzy numbers. *Fuzzy Sets and Systems* 122 : 315–326.
- [Carlsson *et al.* 2005]C. Carlsson, R. Fullér et P. Majlender (2005). On possibilistic correlation. *Fuzzy Sets and Systems* 155 : 425–445.
- [Caron *et al.* 2008]F. Caron, B. Ristic, E. Duflos et P. Vanheeghe (2008). Least committed basic belief density induced by a multivariate gaussian : Formulation with applications. *International Journal of Approximate Reasoning* 48(2) :419–436.
- [Chakraborty *et al.* 2013]M. Chakraborty, A. Skowron, M. Maiti et S. Kar (2013). Facets of Uncertainties and Applications. ICFUA, Kolkata, India.
- [Chaudhuri et Clarke 2013]G. Chaudhuri, K.C. Clarke (2013). The SLEUTH Land Use Change Model : A Review. *The International Journal of Environmental Resources Research* 1(1) : 88–104.
- [Chen et Corson 2014]X. Chen et M. Corson (2014). Application of Dempster-Shafer theory to integrate methods to propagate variability and epistemic uncertainty in agricultural LCA. *Proceedings of the 9th International Conference on Life Cycle Assessment in the Agri-Food Sector*, pp.216–225

- [Chiang et Lin 1991]D.A. Chiang et N.P. Lin (1991). Correlation offuzzy sets. *Fuzzy Sets and Systems* 102 : 221–226.
- [Choi *et al.* 2006]S.K. Choi, R.V. Grandhi et R.A. Canfield (2006). Reliability-based Structural Design, Springer, London.
- [Chopard et Lagrava 2006]B. Chopard et D. Lagrava (2006). A Cellular Automata Model for Species Competition and Evolution. In : El Yacoubi S., Chopard B., Bandini S. (eds) Cellular Automata. ACRI 2006. Vol. 4173 : 277-286.
- [Chutta *et al.* 2013]R. Chutta, S. Mahanta et D. Datta (2013). Sensitivity analysis of atmospheric dispersion model-Rimpuff using Hartley-like measure. *J.Appl.Math.Inform.* 31(1-2) : 99–110.
- [Cihlar et Jansen 2001]J. Cihlar et L.J.M. Jansen (2001). From land cover to land usa : A methodology for efficient land use mapping over large areas. *Professional Geographer* 53(2) :275–289.
- [Clarke *et al.* 1997]K.C. Clarke, S. Hoppen et L. Gaydos (1997). A Self-Modifying Cellular Automaton Model of Historical Urbanization in the San Francisco Bay Area. *Environmental Planning B : Planning and Design* 24 : 247–261.
- [Clarke et Gaydos 1998]K.C. Clarke et L. Gaydos (1998). Loose-coupling a cellular automaton model and GIS : Long-term urban growth prediction for San Francisco and Washington/Baltimore. *International Journal of Geographic Information Science* 12 : 699–714.
- [Cockx *et al.* 2014]K. Cockx, T. de Voorde et F. Canters (2014). Quantifying uncertainty in remote sensing-based urban land-use mapping. *Int. J. Appl. Earth Obs. Geoinf.* 31 : 154–166.
- [Cohen 1960]J. Cohen (1960). A coefficient of agreement for nominal scales. *Educational and psychological measurement* 20 (1) : 37–46.
- [Comber *et al.* 2012]A. Comber (2012). Spatial analysis of remote sensing image classification accuracy. *Remote Sensing of Environment* 127 : 237–246.
- [Crosettoa *et al.* 2001]M. Crosettoa, J.A.M Ruiz et B. Crippac (2001). Uncertainty propagation in models driven by remotely sensed data. *Remote Sens. Environ.* 76 (3) : 373–385.
- [Csillag et Boots 2005]F. Csillag et B. Boots (2005). Toward Comparing Maps as Spatial Processes, in Fisher, P. (Ed.), *Developments in Spatial Data Handling ; 11th International Symposium on Spatial Data Handling*. Springer Berlin Heidelberg, pp. 641–652.
- [Dadhich et Hanaoka 2010]P.N. Dadhich et S. Hanaoka (2010). Remote sensing, GIS and Markov’s method for land use change detection and prediction of Jaipur district. *J Geomatics.* 4 : 9–15.
- [Datta 2011]D. Datta (2011). Measures of sensitivity and uncertainty with the fuzzy model of solute transport through groundwater. In : *International Conference on Communication and Industrial Application*, pp. 1–6.
- [De Figueiredo et Stolfi 2004]L. H. De Figueiredo et J. Stolfi (2004). Affine arithmetic : Concepts and applications. *Numerical Algorithms* 37(1) :147–158.

- [Deep et Saklani 2014]S. Deep et A. Saklani (2014). Urban sprawl modeling using cellular automata. *The Egyptian Journal of Remote Sensing and Space Science* 17(2) : 179–187.
- [Dempster 1968]A.P. Dempster (1968). A generalization of Bayesian inference. *Journal of the Royal Statistical Society*. 30 :205–247.
- [Dempster 2008]A. Dempster (2008). The Dempster-Shafer calculus for statisticians. *IJAR48* :365–377.
- [Denoeux *et al.* 2005]T. Denoeux, M. Masson et P.A. Hébert (2005). Nonparametric rank-based statistics and significance tests for fuzzy data. *Fuzzy sets and systems* 153(1) :1–28.
- [de Noronha *et al.* 2011]V.E. de Noronha, M. Caetano et P. Nijkamp (2011). Trapped between antiquity and urbanism-a multi-criteria assessment model of the greater cairo metropolitan area. *J. Land. Use. Sci.* 6 : 283–299.
- [De Rocquigny *et al.* 2008]E. De Rocquigny, N. Devictor et S. Tarantola (2008). Uncertainty in industrial practice : a guide to quantitative uncertainty management. Wiley, Chichester, UK.
- [Der Kiureghian et Ditlevsen 2009]A. Der Kiureghian et O. Ditlevsen (2009). Aleatory or epistemic ? does it matter ? *Structural Safety* 31(2) : 105–112.
- [Ditlevsen 1983]O. Ditlevsen (1983). Model uncertainty in structural reliability. *Structural safety* 1(1) :73–86.
- [Dongjie *et al.* 2008]G. Dongjie, G. Weijun, W. Kazuyuki et F. Hidetoshi (2008). Land use change of Kitakyushu based on landscape ecology and Markov model. *J Geogr Sci.* 18 : 455–468.
- [Doré *et al.* 2011]P.E. Doré, A. Martin, I. Abi-Zeid, A.L. Jusselme et P. Maupin (2011). Belief functions induced by multimodal probability density functions, an application to the search and rescue problem, *RAIRO Operations Research* 44(4) : 323–343.
- [Doukari *et al.* 2016]O. Doukari, R. Aguejdad, T.Houet (2016). *SLEUTH** : An urban growth model for a scenario-driven approach, *Revue internationale de géomatique* 26(1) : 7–32.
- [Draper 1995]D. Draper (1995). Assessment and Propagation of Model Uncertainty. *Journal of the Royal Statistical Society series B* 57(1) : 45–97.
- [Droguett et Mosleh 2012]E.L. Droguett et A. Mosleh (2012). Integrated treatment of model and parameter uncertainties through a Bayesian approach. *Journal of Risk and Reliability* 227(1) : 41–54.
- [Du *et al.* 2013]J. Du, H. Rui, T. Zuo, Q. Li, D. Zheng, A. Chen, Y. Xu et C.Y. Xu (2013). Hydrological Simulation by SWAT Model with Fixed and Varied Parameterization Approaches Under Land Use Change. *Water Resour Manage* 27 :2823–2838.
- [Dubois et Prade 1998]D. Dubois et H. Prade (1998). Possibility theory : qualitative and quantitative aspects. In *Quantified representation of uncertainty and imprecision*, pp. 169–226.

- [Dubois 2010]D. Dubois (2010). Representation, propagation, and decision issues in risk analysis under incomplete probabilistic information. *Risk Analysis* 30(3) : 361–368.
- [Du et Sudjianto 2004]X. Du et A. Sudjianto (2004). First order saddlepoint approximation for reliability analysis. *AIAA journal* 42(6) :1199–1207.
- [Dunn 2005]W.M. Dunn III (2005). A quick proof that the least squares formulas give a local minimum. *College Math. J.* 36 : 64–65.
- [Eastman 2009]R. Eastman (2009). Idrisi Taiga, Guide to GIS and Image Processing, Manual Version 16.02. Clark University, p. 342.
- [Ebadi *et al.* 2012]M. Ebadi, M.A. Ahmadi, K.F. Hikoei, N. Rashidinia (2012). Application of Fuzzy Decision Tree Analysis for Prediction Asphaltene Precipitation Due Natural Depletion : Case Study, *Australian Journal of Basic and Applied Sciences* 6 : 190–197.
- [Elhag 2014]M. Elhag (2014). Sensitivity analysis assessment of remotely based vegetation indices to improve water resources management. *Environment, Development and Sustainability* 16(6) : 1209–1222.
- [El-Sadek et Irvem 2014]A. El-Sadek et A. Irvem (2014). Evaluating the impact of land use uncertainty on the simulated streamflow and sediment yield of the Seyhan River basin using the SWAT model. *Turk. J. Agric. For.* 38 : 515–530.
- [Evans 1967]D.H. Evans (1967). An Application of Numerical Integration Techniques to Statistical Tolerancing. *Technometrics* 9(3) :441–456.
- [Falihatkar *et al.* 2011]S. Falahatkar, A.R. Soffianian, S.J. Khajeddin, H.R. Ziaee et M.A. Nadoushan (2011). Integration of remote sensing data and GIS for prediction of land cover map. *Intern. J. Geom. and Geosci.* 1(4).
- [FAO 2009]FAO (2009). Seed security for food security in the light of climate change and soaring food prices : challenges and opportunities. *Rome, FAO.*
- [Ferchichi *et al.* 2015]A. Ferchichi, W. Boulila et I.R. Farah (2015). An intelligent possibilistic approach to reduce the effect of the imperfection propagation on land cover change prediction. *International Conference on Computational Collective Intelligence* 9330 : 520–529.
- [Ferchichi *et al.* 2016]A. Ferchichi, W. Boulila, I.R. Farah (2016). Towards an uncertainty reduction framework for land-cover change prediction using possibility theory. *Vietnam J Comput Sci*, pp. 1–15.
- [Ferchichi *et al.* 2017a]A. Ferchichi, W. Boulila et I.R. Farah (2017). Propagating aleatory and epistemic uncertainty in land cover change prediction process. *Ecological Informatics* 37 : 24–37.
- [Ferchichi *et al.* 2017b]A. Ferchichi, W. Boulila et I.R. Farah (2017). Reducing uncertainties in land cover change prediction models using sensitivity analysis. *Knowledge And Information Systems*. In review.
- [Ferson et Ginzburg 1996]S. Ferson et L.R. Ginzburg (1996). Different methods are needed to propagate ignorance and variability. *Reliability Engineering and System Safety* 54(2-3) : 133–144.

- [Ferson *et al.* 2003]S. Ferson, L. Ginzburg, V. Kreinovich, D. Myers et K. Sentz (2003). Construction of probability boxes and Dempster-Shafer structures. Note technique, Sandia National Laboratories.
- [Ferson *et al.* 2004]S. Ferson, J. Hajagos, D. Berleant, J. Zhang, W.T.Tucker, L. Ginzburg et W. Oberkampf (2004). Dependence in Dempster-Shafer theory and probability bounds analysis, SAND.
- [Ferson et Tucker 2006]S. Ferson et W.T. Tucker (2006). Sensitivity in risk analyses with uncertain numbers. Sandia National Laboratories.
- [Fiche *et al.* 2003]A. Fiche, A. Martin, J.C. Cexus, A. Khenchaf, Continuous belief functions and α -stable distributions.
- [Fishman 1996]G.S. Fishman (1996). Monte Carlo : Concepts, Algorithms and Applications. Springer-Verlag.
- [Fouque *et al.* 2001]L. Fouque, A. Bendjebbour, Y. Delignon, V.Samson et W. Pieczynski (2001). Multisensor image segmentation using Dempster-Shafer fusion in Markov fields context. *IEEE Trans. Geosci. Remote Sens.* 39 (8) : 1789–1798.
- [Gao et Dennis 2001]Z.Q. Gao et O. Dennis (2001). The temporal and spatial relationship between NDVI and climatological parameters in Colorado. *J. Geogr. Sci.* 11 (4) : 411–419.
- [Gaucherel et Houet 2009]C. Gaucherel et T. Houet (2009). Preface to the selected papers on spatially explicit landscape modelling : current practices and challenges. *Ecol. Model.* 220(24) : 2477–3480.
- [Gerardino-Neira *et al.* 2008]C. Gerardino-Neira, J. Goodman, M. Velez-Reyes et W. Rivera (2008). Sensitivity Analysis of a Hyperspectral Inversion Model for Remote Sensing of Shallow Coastal Ecosystems. *IGARSS I-98-I-101*.
- [Gerstenkorn et Manko 1991]T. Gerstenkorn et J. Manko (1991). Correlation of intuitionistic fuzzy sets. *Fuzzy Sets and Systems* 44 : 39–43.
- [Ghanem et Spanos 1991]R.G. Ghanem et P.D. Spanos (1991). Stochastic Finite Elements : A Spectral Approach, Springer-Verlag, New York.
- [Gaspari *et al.* 2009]F.J. Gaspari, M.I. Delgado et G.E. Sensiterria (2009). Simulación espacio-temporal de la erosión hídrica superficial en una cuenca serrana bonaerense. Argentina. *GeoFocus* 9 : 67–82.
- [Giri *et al.* 2003]C. Giri, P. Defourny et S. Shrestha (2003). Land cover characterization and mapping of continental Southwest Asia multi-resolution satellite sensor data. *International Journal of Remote Sensing* 24(21) : 4181–4196.
- [Goldberg 1989]D.E. Goldberg (1989). Genetic Algorithms in Search, Optimization, and Machine Learning, Addison-Wesley, Massachusetts.
- [Goldewijk et Verburg 2013]K.K. Goldewijk et P.H. Verburg (2013). Uncertainties in global-scale reconstructions of historical land use : an illustration using the HYDE data set. *Landscape Ecology* 28(5) : 861-877.
- [Gontier *et al.* 2009]M. Gontier, U. Mortberg et B. Balfors (2009). Comparing GIS-based habitat models for applications in EIA and SEA. *Environ. Impact Assess. Rev.* 30(1) : 8–18.

- [Guyonnet *et al.* 2003]D. Guyonnet, B. Bourguine, D. Dubois, H. Fargier, B. Come et J.P. Chiles (2003). Hybrid approach for addressing uncertainty in risk assessments. *Journal of Environmental Engineering-ASCE* 129 (1) :68–78.
- [Hagashi et Klir 1983]M. Hagashi et G.J. Klir (1983). Measure of uncertainty and information based on possibility distribution. *Int. J. Gen. Syst.*9 : 43–58.
- [Haihua *et al.* 2015]X. Haihua, Y. Xianchuan, H. Dan et D. Sha (2015). Sensitivity Analysis of Hierarchical Hybrid Fuzzy-Neural Network. *International Journal on Smart Sensing and Intelligent Systems* 8(3) : 1837–1854.
- [Haines 2009]Y.Y. Haines (2009). Risk Modeling, Assessment, and Management. John Wiley and Sons.
- [Hall 2006]J.W. Hall (2006). Uncertainty-based sensitivity indices for imprecise probability distributions. *Reliab. Eng. Syst. Saf.* 91(1011) : 1443–1451.
- [Hamby 1994]D.M. Hamby (1994). A review of techniques for parameter sensitivity analysis of environmental models. *Environ Monit Assess* 32 : 135–154.
- [Hasofer et Lind 1974]A.M. Hasofer, N. Lind (1974). Exact and Invariant Second Moment Code Format. *Journal of the Engineering Mechanics Division* 100(1) : 111-121.
- [He et Lo 2007]Z. He et C. Lo (2007). Modelling urban growth in Atlanta using logistic regression. *Computer Environment Urban System* 31(6) : 667–688.
- [Hebert *et al.* 2003]P.A. Hebert, M. Masson et T. Denoeux (2003). Fuzzy rank correlation between fuzzy numbers. 10th IFSA World congress, pages 224-227, 29 Juin-2 Juillet, Istanbul, Turquie.
- [Heijungs et Frischknecht 2005]R. Heijungs et R. Frischknecht (2005). Representing statistical distributions for uncertain parameters in LCA. *International Journal of Life Cycle Assessment* 10 (4) : 248–254.
- [Helton et Davis 2003]J.C. Helton et F.J. Davis (2003). Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliability Engineering and System Safety* 81(1) :23–69.
- [Helton *et al.* 2004]J.C. Helton, J.D. Johnson et W.L. Oberkampf (2004). An exploration of alternative approaches to the representation of uncertainty in model predictions. *Reliability Engineering and System Safety* 85 (1) :39–71.
- [Helton *et al.* 2006]J. Helton, J. Johnson, W. Oberkampf et C. Sallaberry (2006). Sensitivity analysis in conjunction with evidence theory representations of epistemic uncertainty. *Reliab. Eng. Syst. Saf.* 91 : 1414–1434.
- [Helton *et al.* 2006]J.C. Helton, J.D. Johnson, C.J. Sallaberry, C.B. Storlie (2006). Survey of sampling-based methods for uncertainty and sensitivity analysis. *Reliability Engineering and System Safety* 91(10-11) : 1175–1209.
- [Heuvelink *et al.* 1989]G.B.M. Heuvelink, P.A. Burrough et A. Stein (1989). Propagation of errors in spatial modelling with GIS. *International Journal of Geographical Information Science* 3(4) : 303 –322.
- [Hird et McDermid 2009]J.N. Hird et G.J. McDermid (2009). Noise reduction of NDVI time series : An empirical comparison of selected techniques. *Remote Sensing of Environment* 113(1) : 248–258.

- [Hoffman et Hammonds 1994]F.O. Hoffman et J.S. Hammonds (1994). Propagation of uncertainty in risk assessment : the need to distinguish between uncertainty due to lack of knowledge and uncertainty due to variability. *Risk Anal.* 14(5) : 707–712.
- [Holland 1975]J.H. Holland (1975). *Adaptation in Natural and Artificial Systems*, University of Michigan Press, Ann Arbor, Michigan.
- [Homma et Saltelli 1996]T. Homma et A. Saltelli (1996). Importance measures in global sensitivity analysis of nonlinear models. *Reliability Engineering and System Safety* 52(1) : 1–17.
- [Hong et Hwang 1995]D.H. Hong et S.Y. Hwang (1995). Correlation of intuitionistic fuzzy sets in probability space. *Fuzzy Sets and Systems* 75 :77–81.
- [Houet et Hubert-Moy 2006]T. Houet et L. Hubert-Moy (2006). Modelling and projecting land-use and land-cover changes with a cellular automaton considering landscape trajectories : an improvement for simulation of plausible future states. In : *EARSeL eProceedings,N5*, pp. 63–76.
- [Hsu et Wu 2010]H.L. Hsu et B. Wu (2010). AN Innovative approach on fuzzy correlation coefficient with interval data. *International Journal of Innovative Computing, Information and Control* 63(A) : 1049–1058.
- [Hu et al. 2011]J. Hu, F. Wei, R. Ge et T. Liang (2011). Propagation of Parameter Uncertainty Based on Evidence Theory, IEEE.
- [Huang et Du 2008]B. Huang et X. Du (2008). Probabilistic uncertainty analysis by mean-value first order saddlepoint approximation. *Reliability Engineering and System safety* 93(2) :325–336.
- [Hulley et al. 2012]G.C. Hulley, C.G. Hughes et S.J. Hook (2012). Quantifying uncertainties in land surface temperature and emissivity retrievals from ASTER and MODIS thermal infrared data. *Journal of Geophysical Research* 117(D23) : 1–18.
- [Huntington et Lyrantzis 1998]D. Huntington et C. Lyrantzis (1998). Improvements to and limitations of latin hypercube sampling. *Probabilistic Engineering Mechanics* 13(4) :245–253.
- [Hyandye et Martz 2017]C. Hyandye et L.W. Martz, A Markovian and cellular automata land-use change predictive model of the Usangu Catchment. *International Journal of Remote Sensing* 38(1).
- [Isukapalli et al. 1998]S.S. Isukapalli, A. Roy et P.G. Georgopoulos (1998). Stochastic Response Surface Methods (SRSMs) for Uncertainty Propagation : Application to Environmental and Biological Systems. *Risk Analysis* 18(3) : 351–363.
- [Ivanov et Tokmakian 2011]L.M. Ivanov et R.T. Tokmakian (2011). Sensitivity analysis of nonlinear models to parameter perturbations for small size ensembles of model outputs. *Int. J. Bifurc. Chaos* 21(12) : 3589–3609.
- [Jacquin 2010]A.P. Jacquin (2010). Possibilistic uncertainty analysis of a conceptual model of snowmelt runoff. *Hydrol. Earth. Syst. Sci.* 14 : 1681–1695.

- [Jadidi *et al.* 2014]A. Jadidi, M.A. Mostafavi, Y. Bédard et K. Shahriari (2014). Spatial Representation of Coastal Risk : A Fuzzy Approach to Deal with Uncertainty. *ISPRS Int. J. Geo-Inf.* 3(3) : 1077–1100.
- [Jakeman *et al.* 2006]A.J. Jakeman, R.A. Letcher, J.P. Norton (2006). Ten iterative steps in development and evaluation of environmental models. *Environmental Modelling and Software* 21 : 602–614.
- [Jantz *et al.* 2003]C.A. Jantz, S.J. Goetz et M.K. Shelley (2003). Using the SLEUTH urban growth model to simulate the impacts of future policy scenarios on urban land use in the Baltimore/Washington metropolitan area. *Environment and Planning B* 31 : 251–271.
- [Jellouli *et al.* 2016]O. Jellouli, A. Bernoussi, M. Mâatouk et M. Amharref (2016). Forest fire modelling using cellular automata : application to the watershed Oued Laou (Morocco). *Methods, Tools and Applications in Engineering and Related Sciences* 22(5) : 493-507.
- [Jiang *et al.* 2013]C. Jiang, B. Wang, Z.R. Li, X. Han et D.J. Yu (2013). An evidence-theory model considering dependence among parameters and its application in structural reliability analysis. *Eng. Struct.* 57 : 12–22.
- [Jimenez-Munoz et Sobrino 2006]J.C. Jimenez-Munoz, J.A. Sobrino (2006). Error sources on the land surface temperature retrieved from thermal infrared single channel remote sensing data. *International Journal of Remote Sensing* 27(5) : 999–1014.
- [Johnson 2009]S. Johnson (2009). An Evaluation of Land Change Modeler for ArcGIS for Ecological Analysis of Landscape Composition. Master of Science (Department of Geography)d(GIScience Specialization).
- [Kang *et al.* 2016]R. Kang, Q. Zhang, Z. Zeng, E. Zio, X. Li (2016). Measuring reliability under epistemic uncertainty : Review on non-probabilistic reliability metrics. *Chinese Journal of Aeronautics* 29(3) : 571-579.
- [Karhunen 1947]K. Karhunen (1947). Uber lineare Methoden in der Wahrscheinlichkeitsrechnung, *Ann. Acad. Sci. Fennicae.* " Ser. A. I. Math.-Phys. 79(37).
- [Kirschbaum 2000]M.U.F. Kirschbaum (2000). Forest growth and species distribution in a changing climate. *Tree Physiology* 20 : 309–322.
- [Kiureghian et Ditlevsen 2009]A. Der Kiureghian et O. Ditlevsen (2009). Aleatory or epistemic? does it matter? *Structural Safety* 31(2) :105–112.
- [Klir et Wiermann 1998]G.J. Klir et M.J. Wiermann (1998). Uncertainty Based Information. Elements of Generalised Information Theory. Physica-Verlg, Heidelberg.
- [Koi et Murayama 2010]D.D. Koi et Y. Murayama (2010). Forecasting areas vulnerable to forest conversion in the Tam Dao National Park Region. *Vietnam. Remote Sens.* 2 : 1249–1272.
- [Kolb *et al.* 2013]M. Kolb, J.F. Mas et L. Galicia 2013. Evaluating drivers and transition potential models in a complex landscape in southern Mexico. *Int. J. Geogr. Inf. Sci.*

- [Krayser von Krauss *et al.* 2006]M.P. Krayser von Krauss, W.E. Walker, J.P. van der Sluijs, P. Janssen, M.B.A. van Asselt et J. Rotmans (2006). Response to "To what extent, and how, might uncertainty be defined" by Norton, Brown, and Mysiak. *Journal-Bridging Sciences and Policy* 6 (1) : 89–94.
- [Krüger et Lakes 2015]C. Krüger et T. Lakes (2015). Bayesian belief networks as a versatile method for assessing uncertainty in land-change modeling. *International Journal of Geographical Information Science* 29(1) : 111–131.
- [Ladevèze *et al.* 2006]P. Ladevèze, G. Puel et T. Romeuf (2006). Lack of knowledge in structural model validation. *Computer Methods in Applied Mechanics and Engineering* 195 :4697–4710.
- [Laliberte *et al.* 2012]A.S. Laliberte, D.M. Browning et A. Rango (2012). A comparison of three feature selection methods for object-based classification of sub-decimeter resolution UltraCam-L imagery. *International Journal of Applied Earth Observation and Geoinformation* 15 : 70–78.
- [Lambin *et al.* 2003]E.F.Lambin, H.J. Geist et E. Lepers (2003). Dynamics of land-use and land-cover change in tropical regions. *Annual Review of Environment and Resources* 28 : 205–241.
- [Lambin *et al.* 2000]E.F. Lambin, M.D.A. Rounsevell et H.J. Geist (2000). Are agricultural land-use models able to predict changes in land-use intensity Agriculture. *Ecosystems and Environment* 82 : 321–331.
- [Landis et Zhang 1998]J. Landis et M. Zhang (1998). The second generation of the California urban futures model. *Part 1 : model logic and theory. Environment and Planning B : Planning and Design* 30 : 657–666.
- [Le Duy *et al.* 2010]T. Le Duy, D. Vasseur, L. Dieulle, C. Bérenguer et M. Couplet (2010). Uncertainty analysis by Dempster-Shafer theory in probabilistic risk assessment. *In Proceeding of the ESREL Conference*.
- [Le Duy *et al.* 2013]T. Le Duy, L. Dieulle, D. Vasseur, C. Bérenguer et M. Couplet (2013). An alternative comprehensive framework using belief functions for parameter and model uncertainty analysis in nuclear probabilistic risk assessment applications. *Proc IMechE Part O : J Risk and Reliability* 0(0) : 1–20.
- [Levashenko et Zaitseva 2012]V. Levashenko et E. Zaitseva (2012). Fuzzy Decision Trees in Medical Decision Making Support System, *Proceedings of the Federated Conference on Computer Science and Information Systems*, pp.213–219.
- [Leyk et Zimmermann 2007]S. Leyk et N.E. Zimmermann 2007. Improving land change detection based on uncertain survey maps using fuzzy sets. *Landscape Ecology* 22(2) : 257–272.
- [Li 2011]W. Li (2011). Dynamic monitoring of land use/cover changes based on CBERS imagery. *International Conference on Geoinformatics*, pp.1–5.
- [Li *et al.* 2014]C. Li, W. Wang, J. Xiong et P. Chen (2014). Sensitivity Analysis for Urban Drainage Modeling Using Mutual Information. *Entropy* 16 : 5738–5752.

- [Li et Yeh 2000]X. Li et A.G.O. Yeh (2000). Modelling sustainable urban development by the integration of constrained cellular automata and GIS. *International Journal of Geographical Information Science* 14 : 131–152.
- [Lilliefors 1967]H. Lilliefors (1967). On the KolmogorovSmirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association* 62 : 399–402.
- [Lindley 2010]D.V. Lindley (2013). Understanding uncertainty. John Wiley and Sons.
- [Liu et al. 2012]Z. Liu, J. Dezert, G. Mercier et Q. Pan (2012). Dynamic evidential reasoning for change detection in remote sensing images. *IEEE Trans. Geosci. Remote Sens.* 50 (5) : 1955–1967.
- [Lloyd et Ries 2007]S.M. Lloyd et R. Ries (2007) Characterizing, propagating, and analyzing uncertainty in life-cycle assessment - A survey of quantitative approaches. *Journal of Industrial Ecology* 11 (1) :161–179.
- [Loève 1977]M. Loève (1977). Probability theory, Springer-Verlag, New York-Heidelberg.
- [Logsdon et al. 1996]M.G. Logsdon, J.E.Bell et V.F. Westerlund (1996). Probability mapping of land use change : a GIS interface for visualizing transition probability. *Comput Environ Urban Syst.* 20 : 389–398.
- [Longley et al. 2001]P.A. Longley, M.J. Barnsley et J.P. Donnay (2001). Remote sensing and urban analysis : a research agenda. In : *J.P. Donnay, M. J. Barnsley, and P.A. Longley (Eds.), Remote Sensing and Urban Analysis. London and New York. Taylor and Francis* : 245–258.
- [Luo et al. 2008]Y.J. Luo, Z. Kang, Z. Luo et L. Alex (2008). Continuum topology optimization with nonprobabilistic reliability constraints based on multi-ellipsoid convex model. *J. Struct. Multidiscip. Optimization* 39(3) : 297–310.
- [Magliocca et al. 2015]N.R. Magliocca, J. van Vliet, C. Brown, T.P. Evans, T. Houet, P. Messerli, J.P. Messina, K.A. Nicholas, C. Ornetsmüller, J. Sagebiel, V. Schweizer, P.H. Verburg et Q. Yu (2015). From meta-studies to modeling : Using synthesis knowledge to build broadly applicable process-based land change models. *Environmental Modelling and Software* 72 : 10–20.
- [Manson 2007]S.M. Manson (2007). Challenges in evaluating models of geographic complexity. *Environment and Planning B : Planning and Design* 34 : 245–260.
- [Mas et al. 2013]J.F. Mas, B. Soares Filho, R.G. Pontius, M. Farfán Gutiérrez et H. Rodrigues (2013). A suite of tools for ROC analysis of spatial models. *ISPRS Int. J. Geo-inf.* 2(3).
- [Mas et al. 2014]J.F. Mas, M. Kolb, M. Paegelow, M.T.C. Olmedo et T. Houet (2014). Inductive pattern-based land use/cover change models : A comparison of four software packages. *Environmental Modelling and Software* 51 : 94–111.
- [Masson et al. 2003]M. Masson, P.A. Hebert et T. Denoeux (2003). Corrélation de rang entre nombres flous. Actes des Rencontres Francophones sur la Logique Floue et ses Applications (LFA'03), 137-142, Tours, France.

- [Matsinos et Troumbis 2002]Y.G. Matsinos et A.Y. Troumbis (2002). Modeling competition, dispersal and effects of disturbance in the dynamics of a grassland community using a cellular automaton model. *Ecological Modelling* 149 : 71–83.
- [Melchers 1989]R. Melchers (1989). Importance sampling in structural systems. *Structural Safety*. 6(1) :3–10.
- [Ménard et Marceau 2007]A. Ménard et D.J. Marceau (2007). Simulating the impact of forest management scenarios in an agricultural landscape of southern Quebec, Canada, using a geographic cellular automata. *Landscape and Urban Planning* 79(3-4) : 253–265.
- [Merry *et al.* 2009]F. Merry, B.S. Soares Filho, D. Nepstad, G. Amacher et H. Rodrigues (2009). Balancing conservation and economic sustainability : the future of the Amazon timber industry. *Environ. Manag.* 44(3) : 395–407.
- [Mertens et Huwe 2002]M. Mertens et B. Huwe (2002). FuN-Balance : a fuzzy balance approach for the calculation of nitrate leaching with incorporation of data imprecision. *Geoderma* 109(3) :269–287.
- [Miller 1956]L.H. Miller (1956). Table of percentage points of Kolmogorov statistics. *Journal of the American Statistical Association* 51 : 111–121.
- [Mishra et Susaki 2014]B. Mishra et J. Susaki (2014). Sensitivity Analysis for L-Band Polarimetric Descriptors and Fusion for Urban Land Cover Change Detection. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 7(10) : 4231–4242.
- [Mishra *et al.* 2014]V.N. Mishra, P.K. Rai et K. Mohan (2014). Prediction of land use changes based on land change modeler (LCM) using remote sensing : a case study of Muzaffarpur (Bihar), India. *J. Geogr. Inst. Cvijic.* 64(1) : 111–127.
- [Miura *et al.* 2000]T. Miura, A.R. Huete et H. Yoshioka (2000). Evaluation of Sensor Calibration Uncertainties on Vegetation Indices for MODIS. *IEEE Transactions on Geoscience and Remote Sensing* 38(3) : 1399–1409.
- [Mohamed 2012]E. Mohamed (2012). Analysis of urban growth at Cairo, Egypt using remote sensing and GIS. *Nat. Sci.* 4.
- [Mondal *et al.* 2015]M.S. Mondal, P.K. Garg, N. Sharma et M. Kappas (2015). Cellular automata (ca) markov modeling of LULC change and sensitivity analysis to identify sensitive parameter(s). In : Proceedings of the 27th International Cartographic Conference 38 :(818).
- [Moore 1966]R.M. Moore (1966). Interval Analysis. Prentice-Hall.
- [Morgan et Henrion 1992]M.G. Morgan et M. Henrion (1992). Uncertainty : a guide to dealing with uncertainty in quantitative risk and policy analysis. Cambridge University Press, New York, USA.
- [Mousivanda *et al.* 2014]A. Mousivanda, M. Menentia, B. Gorte et W. Verhoef (2014). Global sensitivity analysis of the spectral radiance of a soilvegetation system. *Remote Sensing of Environment* 145 :131–144.

- [Muller et Middleton 1994]R.M. Muller et J. Middleton (1994). A Markov model of land-use change dynamics in the Niagara region, Ontario, Canada. *Landscape Ecol.* 9 : 151–157.
- [Munoz-Carpena *et al.* 2007]R. Munoz-Carpena, Z. Zajac et Y.M. Kuo (2007). Global sensitivity and uncertainty analyses of the water quality model VFSMOD-W. *Transactions of the ASABE* 50(5) :1719–1732.
- [Navulur 2007]K. Navulur (2007). Multispectral image analysis using the object-oriented paradigm.. CRC Press, Taylor and Francis Group.
- [Nepstad *et al.* 2009]D. Nepstad, B.S. Soares-Filho, F. Merry, A. Lima, P. Moutinho, J. Carter, M. Bowman, A. Cattaneo, H. Rodrigues, S. Schwartzman, D. McGrath, C. Stickler, R. Lubowski, P. Piris-Cabezaz, S. Rivero, A. Alencar, O. Almeida et O. Stella (2009). The end of deforestation in the Brazilian Amazon. *Science* 326(5958) : 1350–1351.
- [Neumann et Burks 1966]J. Von Neumann et A.W. Burks (1966). Theory of self reproducing automata. *Urbana, Illinois : University of Illinois Press.*
- [Nigussie et Altunkaynak 2016]T.A. Nigussie et A. Altunkaynak (2016). Assessing the Hydrological Response of Ayamama Watershed from Urbanization Predicted under Various Landuse Policy Scenarios, *Water Resour Manage* 30 : 3427–3441.
- [Nouri *et al.* 2014]J. Nouri, A. Gharagozlou, R. Arjmandi, S. Faryadi et M. Adl (2014). Predicting Urban Land Use Changes Using a CA-Markov Model. *Arab J Sci Eng.*
- [Oberguggenberger *et al.* 2009]M. Oberguggenberger, J. King et B. Schmelzer (2009). Classical and imprecise probability methods for sensitivity analysis in engineering : a case study. *Int.J. Approx. Reason.* 50 : 680–693.
- [Oberkampf *et al.* 2002]W.L. Oberkampf, S.M. DeLand, B. M. Rutherford, K.V. Diegert et K.F. Alvin (2002). Error and uncertainty in modeling and simulation. *Reliability Engineering and System Safety* 75(3) : 333–357.
- [Oberkampf et Roy 2010]W.L. Oberkampf et C.J. Roy (2010). Verification and validation in scientific computing. Cambridge University Press.
- [Oijen et Thomson 2010]M.V. Oijen et A. Thomson (2010). Toward Bayesian uncertainty quantification for forestry models used in the United Kingdom Greenhouse Gas Inventory for land use, land use change, and forestry. *Clim. Change* 103 : 55–67.
- [Olofsson *et al.* 2014]P. Olofsson (2014). Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment* 148 : 42–57.
- [Olsson *et al.* 2003]A. Olsson, G. Sandberg et O. Dahlblom (2003). On latin hypercube sampling for structural reliability analysis. *Structural safety* 25(1) :47–68.
- [Paegelow et Camacho Olmedo 2005]M. Paegelow et M.T. Camacho Olmedo (2005). Possibilities and limits of prospective GIS land cover modelling - a compared case study : Garrotxes (France) and Alta Alpujarra Granadina (Spain). *Int. J. Geogr. Inf. Sci.* 19(6) : 697–722.

- [Paegelow et Camacho Olmedo 2008]M. Paegelow et M.T. Camacho Olmedo 2008. Modelling Environmental Dynamics. Advances in Geomatic Simulations. *Series Environmental Science. Springer Verlag, Berlin, Heidelberg, New York.*
- [Park et Grandhi 2012]I. Park et R.V. Grandhi (2012). Quantification of model-form and parametric uncertainty using evidence theory. *Structural Safety* 39 : 44–51.
- [Parker *et al.* 2003]D.C. Parker, S.M. Manson, M.A. Janssen, M.J. Hoffmann et P. Deadman (2003). Multi-agent systems for the simulation of land-use and land-cover change : a review. *Ann. Assoc. Am. Geogr.* 93(2) : 314–337.
- [Parker *et al.* 2008]D.C. Parker, D. Brown, J. Polhill, P.J. Deadman, S.M. Manson (2008). Illustrating a new 'conceptual design pattern' for agent-based models of land use via five case studies-the Mr. Potatohead framework. *In : Lopez Paredes A, Hernandez Iglesias C (eds) Agent-based modelling in natural resource management. Pearson Education, Upper Saddle River*, pp 23–51.
- [Parsa *et al.* 2016]V.A. Parsa, A. Yavari et A. Nejadi (2016). Spatio-temporal analysis of land use/land cover pattern changes in Arasbaran Biosphere Reserve : Iran. *Model. Earth Syst. Environ.* 2 :178.
- [Pea 2008]M.A. Pea (2008). Relationships between remotely sensed surface parameters associated with the urban heat sink formation in Santiago, Chile. *Int. J. Remote Sens.* 29 (15) : 4385–4404.
- [Peng *et al.* 2002]Y. Peng, U. Lohmann, R. Leaitch, C. Banic et M. Couture, (2002). The cloud albedocloud droplet effective radius relationship for clean and polluted clouds from RACE and FIRE. *ACE. J. Geophys. Res.* 107(D11).
- [Pesaresi et Benediktsson 2001]M. Pesaresi et J.A. Benediktsson 2002. A New Approach for the Morphological Segmentation of High-Resolution Satellite Imagery. *IEEE Transactions on Geoscience and Remote Sensing* 39(2) : 309–320.
- [Peters 1995]R.A. Peters (1995). A new algorithm for image noise reduction using mathematical morphology. *IEEE Transactions on Image Processing* 4(5) : 554–568.
- [Pianosi *et al.* 2016]F. Pianosi, K. Beven, J. Freer, J.W. Hall, J. Rougier, D.B. Stephenson et T. Wagener (2016). Sensitivity analysis of environmental models : A systematic review with practical workflow. *Environmental Modelling and Software* 79 : 214–232.
- [Pijanowski *et al.* 2010]B. Pijanowski, L. Iverson, C. Drew, H. Bulley, J. Rhemtulla, M. Wimberly, A. Bartsch et J. Peng (2010). Addressing the Interplay of Poverty and the Ecology of Landscapes : A Grand Challenge Topic for Landscape Ecologists. *Landscape Ecology*.
- [Pontius et Schneider 2001]R.G. Pontius et L. Schneider (2001). Land-use change model validation by a ROC method for the Ipswich watershed, Massachusetts, USA. *Agriculture, Ecosystems and Environment* 85(1-3) : 239–248.
- [Pontius *et al.* 2004]R.G. Pontius, D. Huffaker et K. Denman (2004). Useful techniques of validation for spatially explicit land-change models. *Ecological Modelling* 179(4) : 445–61.

- [Pontius et Malanson 2005]G.R. Pontius et J. Malanson (2005). Comparison of the structure and accuracy of two land change models. *International Journal of Geographical Information Science* 19(2) : 243–265.
- [Pontius et Parmentier 2014]R.G. Pontius et B. Parmentier (2014). Recommendations for using the relative operating characteristic (ROC). *Landscape Ecology* 29 (3) : 367–382.
- [Pontius et Millones 2011]R.G. Pontius et M. Millones (2011). Death to Kappa : birth of quantity disagreement and allocation disagreement for accuracy assessment. *International Journal of Remote Sensing. International Journal of Remote Sensing* 32 (15) : 4407–4429.
- [Pu et al. 2011]R. Pu, S. Landry et Q. Yu (2011). Object-based urban detailed land cover classification with high spatial resolution IKONOS imagery. *International Journal of Remote Sensing* 32(12) : 3285–3308.
- [Puertas et al. 2014]O.L. Puertas, C. Henríquez et F.J. Meza (2014). Assessing spatial dynamics of urban growth using an integrated land use model. Application in Santiago Metropolitan Area, 2010-2045. *Land Use Policy* 38 : 415–425.
- [Quinlan 1993]R. Quinlan (1993). *Programs for Machine Learning*. Morgan Kaufman, San Mateo.
- [R Development 2012]R Development Core Team (2012). *R : a Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria .
- [Raftery 1996]A.E. Raftery (1996). Approximate Bayes Factors and Accounting for Model Uncertainty in Generalized Linear Models. *Biometrika* 83(2) 251–266.
- [Rahman et Xu 2004]S. Rahman et H. Xu (2004). A Univariate Dimension-reduction Method for MultiDimensional Integration in Stochastic Mechanics. *Probabilistic Engineering Mechanics* 19(4) :393–408.
- [Ramankutty et Foley 1999]N. Ramankutty et J.A. Foley (1999). Estimating Historical Changes in Global Land Cover : Croplands From 1700 to 1992. *Global Biogeochemical Cycles* 13 : 997–1027.
- [Rana et al. 2012]Y.H. Rana, X. Lia, L. Lua et Z.Y. Lib (2012). Large-scale land cover mapping with the integration of multi-source information based on the Dempster-Shafer theory. *International Journal of Geographical Information Science* 26(1) : 169–191.
- [Rao et Sawyer 1995]S.S. Rao et J.P. Sawyer (1995). Fuzzy finite element approach for analysis of imprecisely defined systems. *AIAA Journal* 33(12) :2364–2370.
- [Ray et al. 2010]D.K. Ray, J.M. Duckles et B.C. Pijanowski (2010). The Impact of Future Land Use Scenarios on Runoff Volumes in the Muskegon River Watershed. *Environmental Management*.
- [Real et al. 2010]R. Real, A.L. Márquez, J. Olivero et A. Estrada, Species distribution models in climate change scenarios are still not useful for informing policy planning : an uncertainty assessment using fuzzy logic 2010. *Ecography*33, (2) :304–314.

- [Rienow et Goetzke 2014]A. Rienow et R. Goetzke (2012). Supporting SLEUTH—Enhancing a cellular automaton with support vector machines for urban growth modeling. *Computers Environment and Urban Systems* 49 :1–16.
- [Rinderknecht et al. 2012]S.L. Rinderknecht, M.E. Borsuk et P. Reichert (2012). Bridging uncertain and ambiguous knowledge with imprecise probabilities. *Environmental Modelling and Software* 36 :122–130.
- [Rykiel 1996]E.J. Rykiel (1996). Testing ecological models : the meaning of validation. *Ecol. Model.* 90 : 229–244.
- [Saeidi et al. 2014]V. Saeidi, B. Pradhan, M.O. Idrees et Z. Abd Latif (2014). Fusion of airborne LiDAR with multispectral SPOT 5 image for enhancement of feature extraction using Dempster-Shafer theory. *IEEE Trans. Geosci. Remote Sens.* 52 (10) : 6017–6025.
- [Sala et al. 2000]O.E. Sala, F.S. Chapin III, J.J. Armesto, E. Berlow, J. Bloomfield, R. Dirzo, E. Huber-Sanwald, L.F. Huenneke, R.B. Jackson, A. Kinzig, R. Leemans, D.M. Lodge, H.A. Mooney, M. Oesterheld, N.L. Poff, M.T. Sykes, B.H. Walker, M. Walker, D.H. Wall, (2000). Global biodiversity scenarios for the year 2100. *Science* 287(5459) : 1770–1774.
- [Saltelli et al. 2004]A. Saltelli, S. Tarantola, F. Campolongo et M. Ratto (2004). Sensitivity Analysis in Practice : A Guide to Assessing Scientific Models. Joint Research Centre of the European Commission, Ispra, Italy.
- [Samardzic-Petrovic et al. 2016]M. Samardzic-Petrovic, S. Dragicevic, M. Kovacevic et B. Bajat (2016). Modeling Urban Land Use Changes Using Support Vector Machines. *Transactions in GIS* 20(5) : 718–734.
- [Sanchez-Canales et al. 2012]M. Sanchez-Canales, A.L. Benito, A. Passuello, M. Terrado, G. Ziv, V. Acuna, M. Schuhmacher et F.J. Elorza (2012). Sensitivity analysis of ecosystem service valuation in a Mediterranean watershed. *Science of The Total Environment* 440 : 140–153.
- [Sander et al. 2006]P. Sander, B. Bergbäck et T. Öberg (2006). Uncertain Numbers and Uncertainty in the Selection of Input Distributions-Consequences for a Probabilistic Risk Assessment of Contaminated Land. *Risk Analysis* 26(5) : 1363–1375.
- [Sarmiento et al. 2013]P. Sarmiento, H. Carrão, M. Caetano, C.C. Fonte et N. Cortês (2013). Incorporating Uncertainty in the Accuracy Assessment of Land Cover Maps Using Fuzzy Numbers and Fuzzy Arithmetic.
- [Sengupta et Pal 2000]A. Sengupta et T.K. Pal (2000). Theory and methodology : on comparing interval numbers. *Eur. J. Oper. Res.* 127 : 28–43.
- [Sexton et al. 2015]J.O. Sexton, P. Noojipady, A. Anand, X.P. Song, S. McMahon, C. Huang, M. Feng, S. Channan et J.R. Townshend (2015). A model for the propagation of uncertainty from continuous estimates of tree cover to categorical forest cover and change. *Remote Sens. Environ.* 156 : 418–425.
- [Shafer 1976]G. Shafer (1976). A Mathematical Theory of Evidence, volume 1. Princeton University Press.

- [Shan *et al.* 2008]J. Shan, S. Alkheder et J. Wang (2008). Genetic Algorithms for the Calibration of Cellular Automata Urban Growth Modeling. *Photogrammetric Engineering and Remote Sensing* 74(10) : 1267–1277.
- [Shi et Ehlers 1996]W.Z. Shi et M. Ehlers (1996). Determining uncertainties and their propagation in dynamic change detection based on classified remotely-sensed images. *Int. J. Remote Sens.* 17(14) : 2729–2741.
- [Silvestrini *et al.* 2011]R.A. Silvestrini, B.S. Soares-Filho, D. Nepstad, M. Coe, H.O. Rodrigues et R. Assunção (2011). Simulating fire regimes in the Amazon in response to climate change and deforestation. *Ecol. Appl.* 21(5) : 1573–1590.
- [Smith *et al.* 1992]A. Smith, P. Ryan et J. Evans (1992). The effect of neglecting correlations when propagating uncertainty and estimating the population distribution of risk. *Risk Analysis.*
- [Soares-Filho *et al.* 2002]B.S. Soares-Filho, G.C. Cerqueira, C.L. Pennachin (2002). DINAMICA-a stochastic cellular automata model designed to simulate the landscape dynamics in an Amazonian colonization frontier. *Ecological Modelling* 154 : 217–235.
- [Soares-Filho *et al.* 2012]B. Soares-Filho, R. Silvestrini, D. Nepstad, P. Brando, H. Rodrigues, A. Alencar, M. Coe, C. Locks, L. Lima, L. Hissa et C. Stickler (2012). Forest fragmentation, climate change and understory fire regimes on the Amazonian landscapes of the Xingu headwaters. *Landsc. Ecol.* 27(4) : 585–598.
- [Soize 2010]C. Soize (2010). Generalized probabilistic approach of uncertainties in computational dynamics using random matrices and polynomial chaos decompositions. *International Journal for Numerical Methods in Engineering* 81(8) : 939–970.
- [Song *et al.* 2012]X. Song, B.A. Bryan, K.I. Paul et G. Zhao (2012). Variance-based sensitivity analysis for a forest growth model. *Ecological Modelling* 247 : 135–143.
- [Subedi *et al.* 2013]P. Subedi, K. Subedi et B. Thapa (2013). Application of a Hybrid Cellular Automaton - Markov (CA-Markov) Model in Land-Use Change Prediction : A Case Study of Saddle Creek Drainage Basin, Florida. *Applied Ecology and Environmental Sciences* 1(6) : 126–132.
- [Sudret 2002]B. Sudret (2002). Comparison of the spectral stochastic finite element method with the perturbation method for second moment analysis. In Proc. 1st Int. ASRANet Colloquium, Glasgow, United Kingdom.
- [Sutton et Fahmi 2001]K. Sutton et W. Fahmi (2001). Cairo's urban growth and strategic master plans in the light of Egypt's 1996 population census results. *Cities* 18 (3) : 135–149.
- [Tayyebi *et al.* 2010]A. Tayyebi, M.R. Delavar, B.C. Pijanowski, M.J. Yazdanpanah, S. Saeedi et A.H. Tayyebi (2010). A Spatial Logistic Regression Model for Simulating Land Use Patterns, A Case Study of the Shiraz Metropolitan Area of Iran. *Advances in Earth Observation of Global Change, Edited by Emilio Chuvieco, Jonathan Li and Xiaojun Yang.*

- [Tayyebi *et al.* 2014]A.H. Tayyebi, A. Tayyebi et N. Khanna (2014). Assessing uncertainty dimensions in land-use change models : using swap and multiplicative error models for injecting attribute and positional errors in spatial data. *International Journal of Remote Sensing* 35(1) : 149–170.
- [Thapa et Murayama 2011]Rr.B. Thapa et Y. Murayama (2011). Urban growth modeling of Kathmandu metropolitan region, Nepal. *Comput. Environ. Urban Syst.* 35(1) : 25–34.
- [Thompson 2002]K.M. Thompson (2002). Variability and uncertainty meet risk management and risk communication. *Risk Analysis* 22(3) : 647–654.
- [Tilman 1999]D. Tilman (1999). Global environmental impacts of agricultural expansion : the need for sustainable and efficient practices. *Proc. Natl. Acad. Sci. USA* 96(11) : 5995–6000.
- [Turner *et al.* 1995]B.L.I. Turner, D. Skole, S. Sanderson, G. Fischer, L. Fresco et R. Leemans (1995). Land-Use and Land-Cover Change. *IGBP Report, N° 35, HDP Report N° 7. Stockholm and Geneva, IGBP and HDP.*
- [Václavík et Rogan 2009]T. Václavík et J. Rogan (2009). Identifying trends in land use/land cover changes in the context of post-socialist transformation in central Europe : a case study of the greater Olomouc region, Czech Republic. *GISci. Remote Sens.* 46(1) : 54–76.
- [Valbuena *et al.* 2008]D. Valbuena, P.H. Verburg et A.K. Bregt (2008). A method to define a typology for agent-based analysis in regional land-use research. *Agric Ecosyst Environ* 128 : 27–36.
- [van der Kwast *et al.* 2012]J. van der Kwast, L. Poelmans, T. Van de Voorde, K. de Jong, I. Uljee, D. Karssenbergh, F. Canters et G. Engelen (2012). Uncertainty analysis and data-assimilation of remote sensing data for the calibration of cellular automata based land-use models. *International Environmental Modelling and Software Society* , pp. 997–1004.
- [Vapnik 1998]V. Vapnik (1998). *Statistical Learning Theory*. Wiley, New York, 768.
- [van Vliet *et al.* 2013]J. van Vliet (2013). A fuzzy set approach to assess the predictive accuracy of land use simulations. *Ecological Modelling* 261 : 32–42.
- [Veldkamp et Fresco 1996]A. Veldkamp et L.O. Fresco (1996). CLUE-CR : an integrated multi-scale model to simulate land use change scenarios in Costa Rica. *Ecological modelling* 91 : 231–248.
- [Veldkamp et Lambin 2001]A. Veldkamp et E.F. Lambin (2001). Predicting land-use change. *Agriculture, Ecosystems and Environment* 85 : 1–6.
- [Verburg *et al.* 1999]P.H. Verburg, A. Veldkamp, G.H.J. de Koning, K. Kok et J. Bouma (1999). A spatial explicit allocation procedure for modelling the pattern of land use change based upon actual land use. *Ecological modelling* 116 : 45–61.
- [Verburg *et al.* 2002]P.H. Verburg, W. Soepboer, A. Veldkamp, R. Limpiada et V. Espaldon (2002). Modeling the spatial dynamics of regional land use : the CLUE-S model. *Environmental Management* 30(3) : 391–405.

- [Verburg et Veldkamp 2004]P.H. Verburg et A. Veldkamp (2004). Projecting land use transitions at forest fringes in the Philippines at two spatial scales. *Landscape ecology* 19(1) : 77–98.
- [Verburg 2006]P.H. Verburg (2006). Simulating feedbacks in land use and land cover change models. *Landscape Ecol* 21 : 1171–1183.
- [Verburg et al. 2013]P.H. Verburg, A. Tabeau et E. Hatna (2013). Assessing spatial uncertainties of land allocation using a scenario approach and sensitivity analysis : A study for land use in Europe. *Journal of Environmental Management* 127 : S132–S144.
- [Verstegen et al. 2016]J.A. Verstegen, D. Karssenbergh, F. van der Hilst et A.P.C. Faaij (2016). Detecting systemic change in a land use system by Bayesian data assimilation. *Environmental Modelling and Software* 75 : 424–438.
- [Vose 2008]D. Vose (2008). Risk analysis : a quantitative guide. 3rd edn. Wiley, Chichester, West Sussex, England.
- [Walker et al. 2003]W.E. Walker, P. Harremoes, J. Rotmans, J.P. Van der Sluijs, M.B.A. Van Asselt, P. Janssen et M.P. Kraayer von Krauss (2003). Defining Uncertainty : A Conceptual Basis for Uncertainty Management in Model-Based Decision Support. *Integrated Assessment* 4(1) : 5–17.
- [Walley 1991]P. Walley (1991). Statistical reasoning with imprecise probabilities, volume 42. Peter Walley.
- [Walley 2000]P. Walley (2000). Towards a unified theory of imprecise probability. *International Journal of Approximate Reasoning* 24(2-3) :125–148.
- [Wang 2010]Y. Wang (2010). Imprecise probabilities based on generalised intervals for system reliability assessment. *International Journal of Reliability and Safety* 4(4) :319–342.
- [Wang et al. 2013]X. Wang, S. Cen et C. Li (2013). Generalized Neumann Expansion and Its Application in Stochastic Finite Element Methods. *Mathematical Problems in Engineering* 2013 : 1–13.
- [Wei et Hua 2013]H. Wei et Y. Hua (2013). EFAST Method for Global Sensitivity Analysis of Remote Sensing Models Parameters. *Remote Sensing Technology and Application* 28(5) : 836–843.
- [Wernick 2007]I.K. Wernick (2007). Global Warming and the Industrial System. *International Relations and Security Network (ISN)*, Zurich, Switzerland.
- [White et al. 1997]R. White, G. Engelen et I. Uijee (1997). The use of constrained cellular automata for high-resolution modelling of urban land use dynamics. *Environment and Planning B : Planning and Design* 24 : 323–343.
- [White et al. 2000]R. White, G. Engelen, I. Uijee, C. Lavalley et D. Ehrlich (2000). Developing an urban land use simulator for European cities. In : Fullerton, K. (Ed.). *Proceedings of the Fifth EC GIS Workshop : GIS of Tomorrow*. European Commission Joint Research Centre, pp. 179–190.
- [Wolfram 1984]S. Wolfram (1984). Computer Software in Science and Mathematics. *Scientific American* pp. 188–203.

- [Wu 2002]F. Wu (2002). Calibration of stochastic cellular automata : The application to rural-urban land conversions. *International Journal of Geographical Information Science* 16(8) : 795–818.
- [Xiao *et al.* 2013]Y. Xiao, W. Zhao, D. Zhou et H. Gong (2013). Sensitivity Analysis of Vegetation Reflectance to Biochemical and Biophysical Variables at Leaf, Canopy, and Regional Scales. *IEEE Transactions on Geoscience and Remote Sensing* 1–11.
- [Xua *et al.* 2016]X. Xua, Z. Dua et H. Zhang (2016). Integrating the system dynamic and cellular automata models to predict land use and land cover change. *International Journal of Applied Earth Observation and Geoinformation* 52 : 568–579.
- [Yager 2004]R. Yager (2004). Cumulative distribution functions from Dempster-Shafer belief structures. *IEEE Trans Syst Man Cybern* 34(5).
- [Yang et Lo 2003]X. Yang et C.P. Lo (2003). Modelling urban growth and landscape change in the Atlanta metropolitan area. *International Journal of Geographical Information Science* 17 : 463–488.
- [Yang *et al.* 2016]Y. Yang, Y. Wang, K. Wu et X. Yu (2016). Classification of Complex Urban Fringe Land Cover Using Evidential Reasoning Based on Fuzzy Rough Set : A Case Study of Wuhan City. *Remote Sens.* 8(4) : 304–321.
- [Yin *et al.* 2005]Z.Y. Yin, D.J. Stewart, S. Bullard, J.T. Maclachlan (2005). Changes in urban built-up surface and population distribution patterns during 1986-1999 : a case study of Cairo. Egypt. *Comput. Environ. Urban Syst.* 29 : 595–616.
- [Youn *et al.* 2006]B.D. Youn, Z. Xi, L.J. Wells et P. Wang (2006). Enhanced Dimension Reduction (eDR) Method for Sensitivity-free Uncertainty Quantification. Proceedings of 11th AIAA/ISSMO multidisciplinary analysis and optimization conference, Portsmouth, VA.
- [Yu 1993]C. Yu (1993). Correlation of fuzzy numbers. *Fuzzy Sets and Systems* 55 : 303–307.
- [Zadeh 1999]L. Zadeh (1999). Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems* 100 :9–34.
- [Zargar *et al.* 2012]A. Zargar, R. Sadiq, G. Naser, A. ASCE, F.Khan et N. Neumann (2012). Dempster-Shafer theory for handling conflict in hydrological data : case of snow water equivalent. *J. Comput. Civ. Eng.* 26 (3) : 434–447.
- [Zhang *et al.* 2011]R. Zhang, C. Tang, S. Ma, H. Yuan, L. Gao et W. Fan (2011). Using Markov chains to analyze changes in wetland trends in arid Yinchuan plain, China. *Math Comput Modell.* 54 : 924–930.
- [Zhao *et al.* 2009]X. Zhao, A. Stein et X. Chen (2009). Application of random sets to model uncertainties of natural entities extracted from remote sensing images. *Stoch. Env. Res. Risk A* 24(5) : 713–723.
- [Zielinskaa et Sunb 2010]A.L. Zielinskaa et L. Sunb (2010). Applying time-dependent variance-based global sensitivity analysis to represent the dynamics of an agent-based model of land use change. *International Journal of Geographical Information Science* 24(12) : 1829–1850.

