



**Géosimulation multi-niveau de phénomènes complexes basés
sur les multiples interactions spatio-temporelles de nombreux
acteurs: développement d'un outil générique d'aide à la
décision pour la propagation des zoonoses.**

Thèse

Mondher Bouden

Doctorat en informatique
Philosophiæ Doctor (Ph.D.)

Québec, Canada

© Mondher Bouden, 2013

Résumé

Nous proposons dans cette thèse une nouvelle approche de géosimulation multi-niveau permettant de simuler la propagation d'une zoonose (maladie infectieuse qui se transmet des animaux aux humains) à différents niveaux de granularité. Cette approche est caractérisée entre autres par l'utilisation d'un modèle théorique original que nous avons nommé *MASTIM* (*Multi-Actor Spatio-Temporal Interaction Model*) permettant de simuler des populations contenant un nombre considérable d'individus en utilisant des modèles compartimentaux enrichis. *MASTIM* permet de spécifier non seulement l'évolution de ces populations, mais également les aspects relatifs aux interactions spatio-temporelles de ces populations incluant leurs déplacements dans l'environnement de simulation géoréférencé. Notre approche de géosimulation multi-niveau est caractérisée également par l'utilisation d'un environnement géographique virtuel informé (*IVGE*) qui est composé d'un ensemble de cellules élémentaires dans lesquelles les transitions des différents stades biologiques des populations concernées, ainsi que leurs interactions peuvent être plausiblement simulées. Par ailleurs, nous avons appliqué nos travaux de recherche au développement d'outils d'aide à la décision. Nous avons acquis une première expérience avec le développement d'un outil (*WNV-MAGS*) dont l'objectif principal est de simuler les comportements des populations de moustiques (*Culex*) et des oiseaux (corneilles) qui sont impliquées dans la propagation du Virus du Nil Occidental (VNO). Nous avons par la suite participé au développement d'un outil générique (*Zoonosis-MAGS*) qui peut être utilisé pour simuler la propagation d'une variété de zoonoses telles que la maladie de Lyme et le VNO. Ces outils pourraient fournir des informations utiles aux décideurs de la santé publique et les aider à prendre des décisions informées. En outre, nous pensons que nos travaux de recherche peuvent être appliqués non seulement au phénomène de la propagation des zoonoses, mais également à d'autres phénomènes faisant intervenir des interactions spatio-temporelles entre différents acteurs de plusieurs types.

Abstract

We propose in this thesis a new multi-level geosimulation approach to simulate the spread of a zoonosis (infectious disease transmitted from animals to humans) at different levels of granularity. This approach is characterized by using an original theoretical model named MASTIM (Multi-Actor Spatio-Temporal Interaction Model) which can be applied to simulate populations containing a huge number of individuals using extended compartmental models. MASTIM may specify not only the evolution of these populations, but also the aspects related to their spatio-temporal interactions, including their movements in the simulated georeferenced environment. Our multi-level geosimulation approach take advantage of an informed virtual geographic environment (*IVGE*) composed of a set of elementary cells in which the transitions of the different biological stages of the involved populations, as well as their interactions can be simulated plausibly. Furthermore, this approach has been applied to develop decision support tools. We got a first experience with the development of WNV-MAGS, a tool whose main purpose is to simulate the populations' behavior of mosquitoes (*Culex*) and birds (crows), which are involved in the spread of West Nile Virus (WNV). We subsequently participated in the development of a generic tool (Zoonosis-MAGS) that can be used to simulate the spread of a variety of zoonoses such as Lyme disease and WNV. These tools may provide useful information to help public health officers to make informed decisions. Besides, we believe that this research can be applied not only to the spread of zoonoses, but also to other phenomena involving spatio-temporal interactions between different actors of different types.

Résumé long

Bien que l'étude des systèmes complexes ait significativement progressé ces dernières années, elle reste un grand défi pour la science aussi bien sur le plan théorique que pratique. La notion de complexité repose sur l'idée fondamentale selon laquelle un système articulant des éléments divers constitue un tout qui est différent de la somme de ses parties. De ce fait, plusieurs approches de plus en plus sophistiquées ont été conçues pour modéliser et simuler ces systèmes complexes. Cependant, ces approches comme les modèles mathématiques, les automates cellulaires et les systèmes multi-agents classiques présentent encore certaines lacunes qui limitent leurs utilisations pour produire des simulations plausibles. Dans ce contexte, nous proposons dans cette thèse une nouvelle approche de géosimulation multi-niveau permettant de remédier aux lacunes des méthodes actuelles quand elles sont utilisées seules et en bénéficiant de leurs avantages quand elles sont utilisées ensemble, et ce, à différents niveaux de granularité. En outre, nous sommes particulièrement intéressés à appliquer notre approche dans le domaine de la santé publique, afin d'aider les décideurs à surtout mieux comprendre un phénomène aussi complexe que la propagation d'une zoonose (maladie infectieuse qui se transmet des animaux aux humains). Néanmoins, les multiples acteurs ("hôte", "vecteur", "réservoir"), incluant ceux représentant des populations contenant un nombre considérable d'individus, présentent une problématique de modélisation puisque les approches existantes comme les systèmes multi-agent ne sont pas capables de les simuler (il est difficile de représenter chaque individu appartenant à des populations énormes par un agent). Pour surmonter ces difficultés, nous proposons un modèle théorique original que nous avons nommé *MASTIM* (*Multi-Actor Spatio-Temporal Interaction Model*). Ce modèle permet de simuler des populations contenant un nombre d'individus très élevé en utilisant des modèles compartimentaux enrichis pour spécifier non seulement l'évolution de ces populations, mais également les aspects relatifs aux interactions spatio-temporelles de ces populations incluant leurs déplacements dans l'environnement de simulation. Par ailleurs, nous proposons une nouvelle approche permettant de générer un environnement géographique virtuel informé (*IVGE*) dans lequel se déroule la géosimulation. Cet environnement est composé d'un ensemble de cellules élémentaires (correspondant à des polygones qui sont

bien définis à l'aide d'un Système d'Information Géographique - SIG) dans lesquelles les transitions des différents stades biologiques des populations concernées, ainsi que leurs interactions peuvent être plausiblement simulées. Nous avons appliqué nos travaux de recherche au développement d'outils d'aide à la décision. Nous avons acquis une première expérience avec le développement d'un outil (*WNV-MAGS*) dont l'objectif principal est de simuler les comportements des populations de moustiques (*Culex*) et des oiseaux (corneilles) qui sont impliqués dans la propagation du Virus du Nil Occidental (VNO). Nous avons par la suite participé au développement d'un outil générique (*Zoonosis-MAGS*) permettant non seulement de simuler la propagation du VNO, mais aussi la propagation d'autres maladies comme la maladie de Lyme. Ainsi, nos contributions de recherche sont essentiellement: (1) l'approche de géosimulation multi-niveau permettant de modéliser un système complexe à plusieurs niveaux de granularité surtout que les résultats de la simulation d'une zoonose peuvent être différents en fonction du niveau de granularité choisi, (2) le modèle *MASTIM* qui formalise entre autres les interactions spatio-temporelle des populations contenant un grand nombre d'individus (p.ex. moustiques pour le VNO et tiques pour la maladie de Lyme), (3) l'*IVGE* permettant d'intégrer entre autres les caractéristiques biologiques (p.ex. habitat favorable) et comportementales (p.ex. distribution à travers des trajectoires) des espèces étudiées lors de la génération de l'environnement de simulation représentant des larges territoires. L'*IVGE* permet aussi de minimiser le nombre de cellules à traiter en appliquant un processus de fusion répétitif basé sur des seuils de favorabilité et (4) le développement des outils d'aide à la décision permettant d'évaluer les résultats de plusieurs scénarios alternatifs d'intervention dans le but de comprendre et estimer l'ampleur de l'évolution d'une zoonose. Ces outils pourraient donc aider les décideurs de la santé publique à intervenir sur le terrain après avoir pris des décisions informées. Finalement, nous croyons que nos travaux de recherche peuvent être appliqués non seulement au phénomène de la propagation des zoonoses, mais également à d'autres phénomènes faisant intervenir des interactions spatio-temporelles entre différents acteurs de plusieurs types (p.ex. la propagation d'autres types d'épidémies comme le SRAS).

Extented abstract

Although the study of complex systems has increased significantly in recent years, it remains a great challenge for science, both theoretically and practically. The complexity is based on the fundamental idea that a system is different from the sum of its parts. Therefore, several approaches increasingly sophisticated have been proposed to model and simulate these complex systems. However, these approaches such as mathematical models, cellular automata and traditional multi-agent systems have some weaknesses that limit their use to produce credible simulations. In this context, we propose in this thesis a new multi-level geosimulation approach which aims at overcoming the drawbacks of existing methods when used alone and benefiting from their advantages when used together, and at different levels of granularity. Besides, we are particularly interested in applying our approach to the field of public health to help decision makers better understand a complex phenomenon such as the spread of a zoonosis (infectious disease transmitted from animals to humans). But, a modeling problem arises because of the presence of multiple actors ("host", "vector", "reservoir") including those representing populations containing a huge number of individuals. This problem is not properly tackled by existing approaches such as multi-agent systems (it is difficult to represent each individual of huge populations by an agent). To overcome these difficulties, we propose an original theoretical model named MASTIM (Multi-Actor Spatio-Temporal Interaction Model). This model can simulate populations containing a huge number of individuals using extended compartmental models which may specify not only the evolution of these populations, but also the aspects related to their spatio-temporal interactions, including their movements in the simulation environment. Furthermore, we propose a new approach to create an informed virtual geographic environment (IVGE) in which the geosimulation unfolds. This environment is composed of a set of elementary cells (corresponding to polygons which are well defined using Geographic Information System - GIS) in which the transitions of the different biological stages of the involved populations, as well as their interactions can be simulated plausibly. We use the proposed model and approaches to develop decision support tools. We got a first experience with the development of a tool (WNV-MAGS) whose main purpose is to simulate the populations' behavior of mosquitoes (*Culex*) and birds (crows), which are

involved in the spread of West Nile Virus (WNV). We subsequently participated in the development of a generic tool (Zoonosis-MAGS) that can be used to simulate the spread of a variety of zoonoses such as Lyme disease and WNV. Thus, our research contributions are essentially: (1) the multi-level geosimulation approach which allows for modeling a complex system at several levels of granularity since the results of the simulation of a zoonosis may be different depending on the chosen level of granularity, (2) the MASTIM model that can formalize spatio-temporal interactions of populations containing a large number of individuals (e.g. mosquitoes for WNV and ticks for Lyme disease), (3) the IVGE which integrates the biological (e.g. suitable habitat) and behavioral (e.g. distribution through trajectories) characteristics of the studied species during the generation of the simulation environment which represents a large territory. The IVGE can also reduce the number of cells to be computed by applying a progressive merging process based on thresholds of suitability and (4) the development of decision support tools which allows the assessment of the results of alternative intervention scenarios to understand and estimate the magnitude of the evolution of a zoonosis. These tools may therefore help the public health decision makers to intervene on the ground after making informed decisions. Finally, we believe that this research can be applied not only to the spread of zoonoses, but also to other phenomena involving spatio-temporal interactions between different actors of different types (e.g. propagation of other types of epidemics such as SARS).

Avant-propos

Au début, j'aimerais adresser mes remerciements à mon directeur de recherche, le Dr. Bernard Moulin, pour sa confiance, ses conseils et son soutien financier et surtout moral. Tout au long de mon cheminement, il a su me guider tout en me laissant la liberté suffisante pour prendre mes propres décisions. Il était toujours compréhensif et toujours disponible.

Je remercie sincèrement l'examineur externe de ma thèse le Dr. Jérôme Gensel (professeur à l'Université Pierre Mendès de Grenoble, France) ainsi que le Dr. Nadir Belkhiter (vice-doyen aux études à la Faculté des sciences et de génie) et le Dr. Luc Lamontagne (professeur au département d'informatique et de génie logiciel) pour avoir accepté d'être les membres de mon jury.

Je remercie aussi l'INSPQ (Institut National de Santé Publique de Québec) ainsi que GEOIDE (le réseau canadien de centres d'excellence en géomatique) pour avoir financé mes recherches. Ma sincère reconnaissance va surtout au Dr. Pierre Gosselin de l'INSPQ qui m'a généreusement transmis une grande partie de son expertise dans le domaine de la propagation des maladies infectieuses.

Un gros merci à tous mes amis et collègues avec qui j'ai partagé mon quotidien durant des années. Je les remercie tous pour leurs soutiens et les beaux moments que nous avons passés ensemble, mais je ne vais pas pouvoir tous les citer.

Je tiens également à témoigner ma gratitude à mes parents à qui je dois beaucoup. Sans leur aide, je n'aurais pas pu arriver au bout du chemin. Merci également à mes 5 sœurs, 6 neveux et 7 nièces et à toute la famille en Tunisie pour tout l'amour que nous partageons ensemble et pour m'avoir soutenu toutes ces années.

Finalement et non la moindre, je tiens à remercier ma femme qui a été toujours à mes côtés et qui m'a énormément soutenu. Évidemment, je ne peux pas oublier mon petit ange Sharaf que j'aime énormément et qui m'apporte du bonheur jour après jour ☺ الحمد لله

*À mon père Abderrazek et à ma mère Zakia,
À mes soeurs Amel, Nedia, Basma, Fatma et Lamia,
À ma femme Habiba et à mon garçon Sharaf,
Pour leur amour et soutien.*

Table des matières

Résumé.....	iii
Abstract.....	v
Résumé long	vii
Extended abstract.....	ix
Avant-propos	xi
Table des matières	xv
Liste des tableaux.....	xxi
Liste des figures.....	xxiii
Chapitre I: Introduction générale	1
1 Contexte de la recherche.....	1
1.1 Modélisation et simulation d'un phénomène complexe	1
1.2 Propagation d'une maladie infectieuse.....	4
2 Problématique de la recherche.....	9
3 Hypothèse de la recherche.....	11
4 Objectifs de la recherche.....	13
5 Structure de la thèse.....	14
Chapitre II: État de l'art	16
1 Introduction.....	16
2 Présentation des systèmes complexes.....	16
2.1 Définition et caractéristiques	17
2.2 Notions de complexité	19
2.3 Exemples de systèmes complexes	19
3 Modélisation et simulation des systèmes complexes.....	21
3.1 Systèmes continus non-linéaires.....	23
3.1.1 Simulation à base d'équations mathématiques	24
3.1.2 Dynamique des systèmes.....	24
3.2 Automates cellulaires.....	28

3.3	Systèmes multi-agents.....	30
3.4	Systèmes hybrides.....	32
3.5	Géosimulation multi-agent.....	35
3.5.1	Généralités.....	35
3.5.2	Plateforme de simulation <i>MAGS</i>	36
3.6	Systèmes multi-niveaux.....	41
3.7	Systèmes holoniques.....	48
4	Conclusion.....	51
	Chapitre III: Domaine d'application: Propagation des zoonoses.....	52
1	Introduction.....	52
2	Présentation des maladies infectieuses.....	52
2.1	Définition et caractéristiques.....	52
2.2	Exemples de zoonoses.....	55
2.2.1	Virus du Nil Occidental (VNO).....	55
2.2.1.1	Cycle de transmission du VNO.....	56
2.2.1.2	Populations de moustiques et leur rôle vectoriel.....	57
2.2.1.3	Populations d'oiseaux et leur rôle amplificateur.....	59
2.2.1.4	Interactions entre les moustiques et les corvidés.....	60
2.2.2	Maladie de Lyme.....	62
3	Modélisation et simulation de la propagation des maladies infectieuses.....	65
3.1	Modèles mathématiques.....	65
3.1.1	Modèles compartimentaux.....	66
3.1.2	Modèles à <i>patches</i> (<i>patchy models</i>).....	68
3.1.3	Modèles de métapopulations.....	69
3.2	Autres approches de simulation.....	70
3.3	Outils informatiques.....	73
4	Conclusion.....	75
	Chapitre IV: Approche de géosimulation multi-niveau.....	76
1	Introduction.....	76

2	Vue d'ensemble de notre approche.....	76
2.1	Utilité d'un outil d'aide à la décision.....	76
2.2	Vue globale des différents aspects qui nous semblent importants.....	80
2.3	Esquisse des différents aspects qui nous intéressent	85
3	Modélisation des différents niveaux de granularité.....	85
3.1	Modélisation des échelles spatio-temporelles.....	89
3.2	Modélisation des différentes catégories d'acteurs participant dans la propagation des zoonoses	92
3.2.1	Modélisation des individus et des groupes d'individus.....	95
3.2.2	Modélisation des populations	96
3.2.3	Modélisation des interactions entre les différents acteurs	98
3.3	Niveau de détail des modèles utilisés pour la propagation des zoonoses.....	101
4	Discussion et conclusion.....	103
	Chapitre V: Modèle théorique <i>MASTIM</i>	105
1	Introduction.....	105
2	Présentation de <i>MASTIM</i>	105
3	Formalisme de <i>MASTIM</i>	106
3.1	Le système écologique.....	107
3.2	Les cellules	108
3.3	Les régions.....	109
3.4	Les zones sélectionnées	110
3.5	Les espèces	111
3.6	Les compartiments.....	112
3.7	Les attributs.....	113
3.8	Les transitions.....	115
3.8.1	Les transitions d'évolution.....	116
3.8.2	Les transitions de mortalité.....	118
3.8.3	Les transitions de transfert.....	118
3.9	Les événements et les conditions.....	119

3.10	Les interactions	120
3.11	Les processus.....	123
4	Applications de <i>MASTIM</i>	126
4.1	Un modèle pour le VNO	126
4.2	Un modèle pour la maladie de Lyme	133
4.2.1	Interaction entre les tiques et les rongeurs	136
4.2.2	Interaction entre les tiques et les oiseaux	137
4.2.3	Interaction entre les tiques et les chevreuils.....	137
5	Discussion et conclusion	138
Chapitre VI: Approche de génération de l'IVGE		141
1	Introduction	141
2	Présentation de l'IVGE.....	141
3	Création de l'IVGE.....	143
3.1	Vue d'ensemble de l'approche proposée	144
3.2	Utilisation des habitats favorables comme critère de fusion.....	147
3.3	Présentation et prétraitement des données SIG	149
3.4	Fusion progressive des cellules	156
3.5	Enrichissement par l'ajout des données dans les cellules	160
3.5.1	Données biologiques et climatiques.....	160
3.5.2	Identification des voisins et leurs orientations	160
3.5.3	Informations de couverture du sol.....	163
3.5.4	Informations sur la forme des cellules	164
3.5.5	Informations sur les trajectoires	165
4	Application de l'IVGE: Distribution spatiale des flots d'individus	165
4.1	Formalisme du transfert spatial des individus.....	166
4.1.1	Le patron de mobilité	166
4.1.2	La trajectoire	168
4.1.3	Le réseau de mobilité	171
4.1.4	Le flot d'individus.....	172

4.2	Implémentation de la distribution d'un flot d'individus.....	173
4.2.1	Présentation de l'approche de distribution.....	173
4.2.2	Étapes d'exécution de l'application de distribution.....	179
5	Discussion et conclusion.....	188
	Chapitre VII: Développement d'outils d'aide à la décision.....	191
1	Introduction.....	191
2	WNV-MAGS.....	191
2.1	Échelle spatiale Macro (Sud de Québec).....	195
2.1.1	Préparation des données et création de l' <i>IVGE</i>	195
2.1.2	Description du modèle mathématique à base de compartiments.....	198
2.1.3	Geosimulation des populations.....	200
2.1.4	Application de différents scénarios et calibration du système.....	203
2.2	Échelle spatiale Méso (Ottawa).....	206
2.2.1	Changement de l'échelle spatiale.....	207
2.2.2	Préparation des données et création de l' <i>IVGE</i>	208
2.2.3	Résultat de la simulation et validation du modèle.....	211
2.3	Discussion.....	214
3	Zoonosis-MAGS.....	215
3.1	Architecture conceptuelle.....	215
3.2	Formalisme de la simulation.....	219
3.2.1	L'état du système.....	219
3.2.2	Le scénario de la simulation.....	221
3.2.3	La simulation d'une zoonose.....	223
3.3	Géosimulation de la maladie de Lyme.....	227
3.3.1	Architecture technique.....	227
3.3.2	Spécification du modèle de simulation et du scénario.....	230
3.3.3	Résultats de la simulation.....	236
3.4	Discussion.....	242
4	Conclusion.....	244

Chapitre VIII : Conclusion générale	245
1 Synthèse	245
2 Contributions	247
3 Limites et travaux futurs	254
Bibliographie	261
Webliographie	272
ANNEXES	274

Liste des tableaux

Tableau 1: Une hiérarchie hypothétique de la complexité du monde réel [Boulding, 1956].	20
Tableau 2: Degrés de favorabilité de base (BSD) des habitats pour les tiques en fonction des types de couverture du sol de Géobase.....	148
Tableau 3: Description des principaux types de couverture du sol présents dans Géobase.	151
Tableau 4: Calcul du nouveau degré de favorabilité.	153
Tableau 5: Les principaux attributs attachés aux cellules.....	177
Tableau 6 : Spécifications des paramètres de la simulation.	192

Liste des figures

Figure 1: Modélisation d'un phénomène complexe dynamique.....	2
Figure 2: Modélisation d'un phénomène complexe dynamique avec prise en compte de l'aspect décisionnel.....	3
Figure 3: Principales approches utilisées pour la propagation des maladies infectieuses.	5
Figure 4: Liens entre les objectifs de la thèse et ses chapitres.....	15
Figure 5: Modèle conceptuel typique d'un système complexe [Parrott, 2002].	18
Figure 6: Approches de simulation sur plusieurs niveaux d'abstraction [Borshchev et Filippov, 2004].....	22
Figure 7: Boucle de rétroaction (<i>Feedback loop</i>) [Karsky, 2004].....	25
Figure 8: Exemple d'une variable d'état et de deux variables de flux [Karsky, 2004].	27
Figure 9: Passage du diagramme causal (a) au modèle dynamique (b) pour l'exemple de la gestion de production d'un produit [Karsky, 2004].....	27
Figure 10: Exemple d'automates cellulaires tridimensionnels [Site Web: <i>Wikipedia</i>].	28
Figure 11: Exemple de simulation de la propagation des feux de forêt en utilisant un diagramme de Voronoi [Carvalho <i>et al.</i> , 2006].	29
Figure 12: Exemple d'agent à comportement de type cognitif [Mandiau <i>et al.</i> , 2002].....	31
Figure 13: Système multi-agent selon différents niveaux de détail [Mandiau <i>et al.</i> , 2002].	31
Figure 14: Niveaux macro et micro de la chaîne logistique [Schieritz et Größler, 2003]. ...	34
Figure 15 : Le simulateur <i>MAGS</i> en 2D (a) et en 3D (b) [Bouden, 2004].....	37
Figure 16 : Les différentes composantes du système <i>MAGS</i>	38
Figure 17 : Simulation de la propagation d'un gaz (a) et du vent (b) [Bouden, 2004].....	39
Figure 18: Le modèle conceptuel <i>IMSHED</i> [An <i>et al.</i> , 2005].	43
Figure 19: Approche de simulation multi-niveau [White <i>et al.</i> , 2000].....	46
Figure 20: Modélisation à 3 niveaux d'abstraction [Mathieu <i>et al.</i> , 2007].	47
Figure 21: La notion de Holon dans un système multi-agent [Cossentino <i>et al.</i> , 2007].....	49
Figure 22: Exemple de la structure holonique d'une université [Cossentino <i>et al.</i> , 2007]...	50
Figure 23: Principaux types de maladies infectieuses [McMichael et Woodruff, 2008].	53
Figure 24 : Cycle de transmission du virus du Nil occidental [Site Web : <i>ASPC</i>].	56

Figure 25 : Cycle biologique des moustiques [Site Web : <i>ASPC</i>].	57
Figure 26: Diagramme de la vue d'ensemble des phénomènes qui nous intéressent [Bouden <i>et al.</i> , 2008a].	61
Figure 27: Stades (a) et cycle de développement des tiques (b) [Site Web: <i>CDCP</i>].	64
Figure 28: Modèle mathématique de VNO à 8 équations différentielles [Wonham <i>et al.</i> , 2004].	67
Figure 29: Équations faisant évoluer les populations de moustiques et d'oiseaux pour chacun des patches subdivisant la région étudiée [Gourley <i>et al.</i> , 2008].	68
Figure 30: Modèle épidémique [Hannon et Ruth, 2009].	70
Figure 31: Modélisation du cycle biologique des tiques en utilisant la dynamique des systèmes [Hannon et Ruth, 2009].	71
Figure 32: Dynamique des populations de tiques [Ogden <i>et al.</i> , 2005].	72
Figure 33: Modèle DS (a) et modèle SMA (b) en utilisant <i>Anylogic</i> [Rahmandad et Sterman, 2006].	74
Figure 34: Système de prise de décision et d'intervention.	77
Figure 35: L'utilisation d'un outil d'aide à la décision par les décideurs afin de prendre des décisions informées.	78
Figure 36: Les différents niveaux de granularité (Macro, Méso et/ou Micro) qui sont utilisés pour établir les différents plans d'intervention (stratégique, tactique et/ou opérationnel).	79
Figure 37: Les différents aspects qui nous semblent importants.	81
Figure 38: Liens entre les objectifs spécifiques de la thèse et la solution proposée.	85
Figure 39: Principaux facteurs influençant le choix des niveaux de granularité.	86
Figure 40: Les différents types de niveaux de granularité.	87
Figure 41: Les différentes combinaisons des niveaux de granularité.	89
Figure 42: Les différentes échelles spatiales pour représenter l'environnement géographique virtuel.	91
Figure 43: Spécification des acteurs impliqués dans la propagation d'une maladie infectieuse.	92

Figure 44: La représentation multi-niveau des différentes catégories d'acteurs.	94
Figure 45: Une holarchie représentant les différentes catégories d'acteurs.	95
Figure 46: Modélisation différente selon la catégorie de population.	97
Figure 47: Interactions des différents acteurs.	100
Figure 48: Système multi-modèle à différents niveaux de granularité.	102
Figure 49: Exemples de zones sélectionnées incluant (a) ou non (b) la totalité des cellules qui touchent les frontières.	110
Figure 50: Les différents types de processus qui nous intéressent.	124
Figure 51: Cycle biologique des <i>Culex</i> en tenant compte de l'infection.	128
Figure 52: Cycle biologique des corneilles en tenant compte de l'infection.	129
Figure 53: Les différentes interactions entre les <i>Culex</i> et les corneilles.	131
Figure 54: Le modèle biologique des tiques sans maladie.	133
Figure 55: Le modèle biologique des tiques sans maladie.	134
Figure 56: Le modèle biologique des tiques avec maladie.	135
Figure 57: Les interactions entre les tiques et les rongeurs.	136
Figure 58: Les interactions entre les tiques et les oiseaux.	137
Figure 59: Les interactions entre les tiques et les chevreuils.	138
Figure 60: Vue d'ensemble de l'approche permettant de générer l' <i>IVGE</i>	144
Figure 61: Les 6 régions SNRC qui nous intéressent.	150
Figure 62: Deux polygones complexes représentant une zone urbaine (à gauche) et une forêt (à droite).	150
Figure 63: Région 31H avec une représentation de certains types de couverture du sol.	152
Figure 64: Création d'un polygone plein à partir d'un polygone contenant 20 trous et 15 sous-trous.	152
Figure 65: Une fenêtre de l'application <i>Visual Basic</i> permettant de faire les requêtes spatiales.	154
Figure 66: Le résultat de la suppression des trous des polygones de la région 31H.	155
Figure 67: Le polygone (en rouge) contenant le seul trou non supprimé (en vert).	155
Figure 68: Fusion avec les cellules voisines en utilisant un seuil de favorabilité.	156

Figure 69: Première itération du processus de fusion progressive.....	157
Figure 70: Pourcentages des cellules supprimées après les différentes itérations de fusion.	159
Figure 71: Résultat final de la fusion progressive des polygones de la région 31H.	159
Figure 72: Information quantitative (a) et qualitative (b) utilisée pour stocker l'orientation des voisins.	162
Figure 73: Enrichissement de l' <i>IVGE</i> par l'ajout des données dans les cellules.	163
Figure 74: Calcul des distances pour traverser une cellule dans les différentes directions de ses voisins.....	164
Figure 75: Un exemple de trajectoire avec son extension.....	169
Figure 76: Un réseau de mobilité montrant deux trajectoires.	171
Figure 77: Application de distribution des individus à travers les cellules.....	176
Figure 78: Quelques attributs des cellules C_1 (a) et C_6 (b).....	178
Figure 79: Présentation de l'orientation des différentes cellules.	179
Figure 80: Premier déplacement des individus et prochaine cellule à traiter.....	180
Figure 81: Tris des cellules voisines de C_1 et qui ne sont pas encore traitées.	182
Figure 82: Deuxième distribution des individus et prochaines cellules à traiter.	184
Figure 83: Tris des cellules voisines de C_5 et qui ne sont pas encore traitées.	185
Figure 84: Troisième distribution des individus et prochaines cellules à traiter.....	186
Figure 85: Fin du processus de distribution.	187
Figure 86: Modèle conceptuel représentant une vue dynamique du système.	193
Figure 87: La zone géographique qui nous intéresse.	196
Figure 88: Répartition des populations initiales de <i>Culex</i> (les municipalités plus foncées contiennent plus de moustiques).	197
Figure 89: Répartition des degrés jours cumulés (p.ex. les régions en rouge sont plus chaudes que les régions en bleue).	198
Figure 90: (a) modèle mathématique de Wonham <i>et al.</i> (2004) vs (b) modèle amélioré qui a été utilisé pour le système <i>WNV-MAGS</i>	199
Figure 91: Utilisation d'une carte de densité intelligente pour représenter les <i>Culex</i>	202

Figure 92: Utilisation des dortoirs (systèmes de particules) pour représenter les corneilles.	202
Figure 93: Utilisation des scénarios alternatifs d'intervention.....	204
Figure 94: Subdivision de la région d'Ottawa selon les secteurs de recensement (SR).....	207
Figure 95: Zone géographique qui nous intéresse pour la région d'Ottawa.....	209
Figure 96: Population initiale de <i>Culex</i> selon les secteurs de recensement.....	210
Figure 97: Capture d'écran de <i>WNV-MAGS</i> avant la simulation pour la région d'Ottawa.	211
Figure 98: Simulation du VNO pour la région d'Ottawa.....	212
Figure 99: Stations de capture de <i>Culex</i> pour la région d'Ottawa (le secteur 5050011.04 contient 4 stations).....	213
Figure 100: Évolution de la population de <i>Culex</i> (données du modèle vs données de terrain) après application de larvicides pour le secteur de recensement 5050011.04.....	214
Figure 101: Architecture conceptuelle de <i>Zoonosis-MAGS</i> (basée sur la méthode EPAS).	217
Figure 102: Passage du système <i>Zoonosis-MAGS</i> d'un état à un autre.....	220
Figure 103: Ordre des étapes au cours d'un pas de simulation.....	224
Figure 104: Exemples pour illustrer la stratégie à deux passes.....	226
Figure 105: Les plus importants modules de l'architecture technique de <i>Zoonosis-MAGS</i>	228
Figure 106: Les plugins du modèle de simulation.....	229
Figure 107: (a) Connexion à la base de données, (b) Sélection des plugings.....	230
Figure 108: Spécification des hiérarchies des cellules et des données de populations.....	231
Figure 109: Utilisation de l' <i>IVGE</i> pour la simulation de la maladie de Lyme.....	232
Figure 110: Fenêtre de commandes de la simulation.....	233
Figure 111: Spécification du modèle à base de compartiments pour la maladie de Lyme.	233
Figure 112: Spécification des équations mathématiques.....	234
Figure 113: Spécification du modèle de simulation en utilisant Visio.....	235
Figure 114: Spécification d'un scénario.....	235
Figure 115: Spécification des événements dans un scénario.....	236

Figure 116: Les options de sauvegarde avant de démarrer la simulation.	237
Figure 117: Fin de la simulation en affichant le temps d'exécution.	238
Figure 118: Comparaison entre les résultats de <i>MATLAB</i> et ceux de <i>Zoonosis-MAGS</i>	238
Figure 119: Évolution des différents compartiments des tiques sur deux ans.	239
Figure 120: Évolution de la population totale des tiques sur 40 ans en utilisant une cellule au Sud dont la température est supérieure de 5% qu'une cellule au Nord.....	240
Figure 121: Augmentation du temps d'exécution de la simulation en fonction du nombre de cellules dans l' <i>IVGE</i>	242
Figure 122: Nos principales contributions de recherche.	248

Chapitre I: Introduction générale

Dans ce chapitre, nous présentons le contexte de notre recherche qui est la géosimulation multi-niveau des zoonoses. Nous identifions également la problématique et les hypothèses de recherche que nous adressons dans cette thèse. Nous mentionnons par la suite nos objectifs de recherche. Finalement, nous présentons la structure de la thèse en introduisant ses différents chapitres.

1 Contexte de la recherche

Le monde réel nous offre une multitude de phénomènes complexes allant de l'infiniment petit à l'infiniment grand. Il est même souvent ardu de se rendre compte des différents niveaux de complexité d'un unique phénomène. L'observateur du monde réel a besoin de trouver les moyens pour comprendre un phénomène aussi complexe, surtout s'il faut intervenir afin de changer ses conséquences sur la société et/ou sur l'environnement (p.ex. comprendre comment une bactérie se transmet d'un individu à un autre individu avant de trouver les moyens pour limiter la transmission de cette bactérie). Nous présentons dans ce qui suit une série de figures synthétiques que nous avons élaborées pour préciser notre contexte de recherche. Nous commençons en présentant le contexte général se rapportant à la modélisation et à la simulation d'un phénomène complexe. Nous présentons par la suite le contexte spécifique se rapportant à notre domaine d'application. Il s'agit de la propagation des maladies infectieuses.

1.1 Modélisation et simulation d'un phénomène complexe

Les concepteurs utilisent la modélisation et la simulation informatique pour reproduire virtuellement un ou plusieurs phénomènes sur un ordinateur afin de les analyser. Un phénomène réel peut être représenté par plusieurs systèmes complexes, étant donné qu'il peut souvent être modélisé suivant plusieurs points de vue, selon la vision de l'observateur. Il peut être également modélisé en utilisant des techniques et/ou des approches différentes (voir Figure 1).

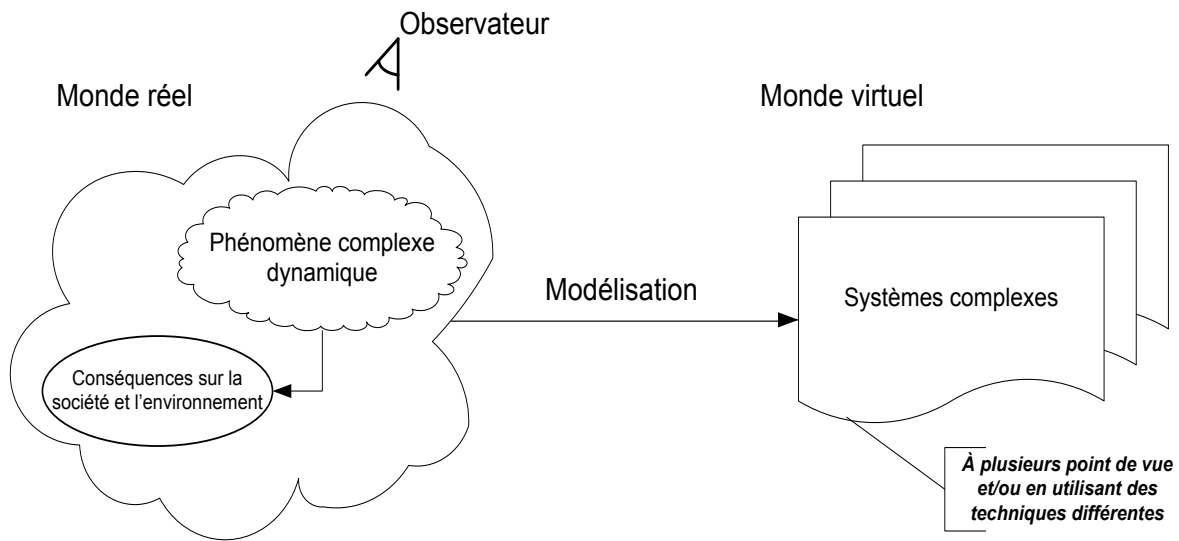


Figure 1: Modélisation d'un phénomène complexe dynamique.

La modélisation d'un phénomène complexe est une des premières étapes permettant de le simuler sur un ordinateur afin de pouvoir bien le comprendre et l'analyser. Fishwick [Fishwick, 1995] présente la simulation informatique comme une discipline qui se compose de la conception de modèles, l'exécution de ces modèles et finalement l'analyse des résultats de cette exécution. D'ailleurs, la simulation des systèmes complexes constitue actuellement un enjeu majeur dans de nombreux domaines tels que les neuro-sciences, la biologie, la physiologie, l'écologie, l'économie ou encore la sociologie. De plus, la simulation permet d'étudier des systèmes qui ne peuvent pas être directement observés ou mesurés. Elle est devenue donc un outil scientifique à part entière permettant d'étudier notamment l'évolution des phénomènes météorologiques, l'évolution des marchés financiers, l'évaluation de l'impact d'une explosion nucléaire ou encore la propagation d'une épidémie [Sichman *et al.*, 1998].

En outre, la simulation peut être utilisée afin de développer des outils d'aide à la décision permettant d'apporter des réponses pertinentes à des problèmes divers. En effet, ces outils permettent de faciliter la prise de décision stratégique, tactique et/ou opérationnelle dans des situations imprécises et/ou incertaines. D'ailleurs, les décideurs de divers secteurs (gouvernemental, militaire, industriel, médical, social) ont besoin de tels outils pour établir des plans d'intervention et de ce fait, influencer les conséquences d'un phénomène réel sur

la société et/ou l'environnement. Ces décideurs espèrent maintenir un certain contrôle sur le phénomène en agissant sur lui par l'intermédiaire de ce que nous appelons des "acteurs d'intervention". Ces derniers vont entrer en interaction avec les acteurs du phénomène et jouer un rôle important dans la limitation ou l'expansion des conséquences du phénomène. Cependant et pour pouvoir élaborer et comparer des scénarios alternatifs d'intervention, les décideurs ont besoin d'une compréhension globale du phénomène étudié et de moyens pour suivre son évolution, surtout lorsqu'ils doivent faire face à un phénomène qui implique un grand nombre d'acteurs de différents types (humains, animaux, statiques ou mobiles, ordinateurs, etc.) agissant dans des espaces géographiques très variés. Nous illustrons ces éléments dans la Figure 2.

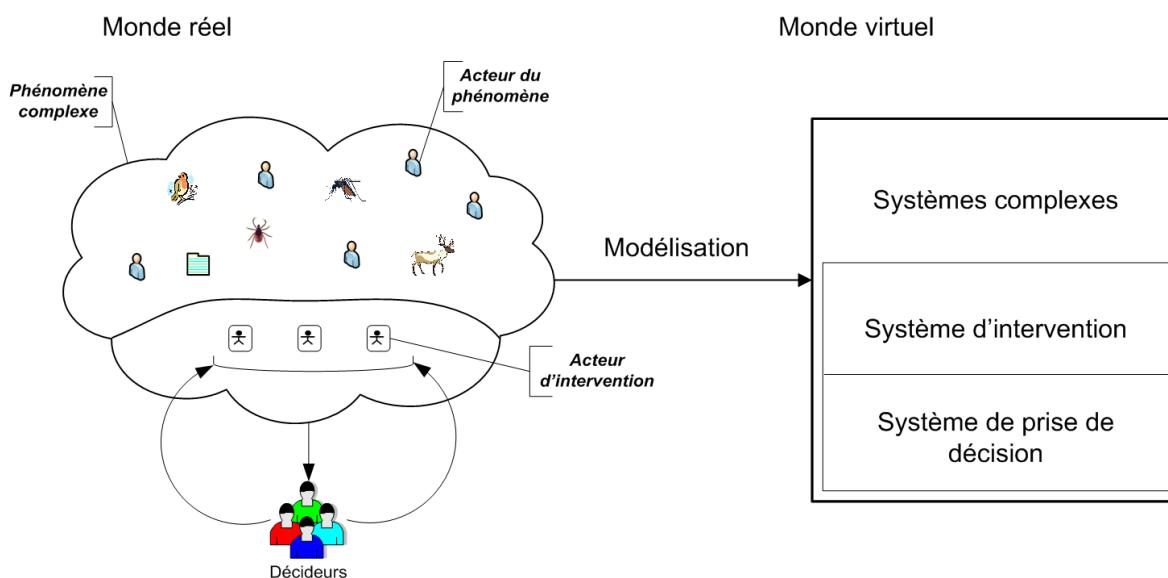


Figure 2: Modélisation d'un phénomène complexe dynamique avec prise en compte de l'aspect décisionnel.

D'un autre côté, nous avons mentionné qu'un phénomène réel peut être représenté par plusieurs systèmes complexes. Un système complexe peut être défini comme un système contenant un grand nombre de composants en interaction et dont le comportement global n'est pas dérivable de la somme des comportements de ses composants individuels (interactions non-linéaires). Typiquement, de tels systèmes exhibent une structure hiérarchique et des processus d'auto-organisation donnant souvent naissance à des propriétés émergentes [Drogoul et Ferber, 1994]. Parmi ces systèmes, nous distinguons le système de prise de décision et le système d'intervention. Ces deux sous-systèmes sont

normalement dépendants l'un de l'autre puisqu'une intervention sur le terrain est le résultat d'une prise de décision (voir Figure 2).

1.2 Propagation d'une maladie infectieuse

Nous sommes particulièrement intéressés à utiliser la modélisation et la simulation informatique dans le domaine de la santé publique, afin d'aider les décideurs à intervenir en temps voulu et de manière adéquate¹ en vue de contrôler la propagation de maladies infectieuses (voir Section 2 du chapitre III). Ces maladies sont le résultat de la transmission d'un micro-organisme, représentant l'agent pathogène, d'un individu infecté ("hôte": homme ou animal) à un individu sain. On parle d'une épizootie quand la transmission se fait entre animaux. L'épizootie peut se transformer en zoonose si elle se transmet à l'humain. Le contrôle d'une zoonose constitue aujourd'hui un enjeu majeur pour la santé publique. Ce contrôle passe par la compréhension des mécanismes de transmission de la maladie, qui sont généralement complexes du fait du mode de transmission indirect des maladies faisant intervenir de nombreux acteurs ("hôte", "vecteur"², "réservoir"³). Ce phénomène complexe mérite d'être modélisé et simulé si nous voulons aider les décideurs du domaine de la santé publique à limiter l'extension des zoonoses.

Par ailleurs, la propagation de ces maladies et leurs effets néfastes sur la santé humaine ont amené les responsables de la santé publique à développer des systèmes de surveillance permettant de mieux connaître le niveau de risque qu'elles peuvent représenter pour la population humaine. C'est le cas par exemple du système intégré des données de vigie sanitaire (SIDVS) qui a été développé en 2003 par l'Institut national de santé publique du Québec (INSPQ) [Gosselin *et al.*, 2005]. Cet outil vise notamment à effectuer la représentation cartographique en ligne, via un réseau Extranet, des principales données recueillies sur le terrain concernant le virus du Nil Occidental (VNO).

¹ Nous utilisons ici le terme « adéquate » pour désigner la satisfaction des décideurs par rapport à leurs attentes.

² Les vecteurs sont responsables de la propagation d'une maladie qui se transmet d'une façon indirecte (p.ex. moustiques, tiques, etc.).

³ Les réservoirs assurent le maintien de l'agent pathogène dans la nature (p.ex. oiseaux, rongeurs, etc.).

Néanmoins, ces programmes de surveillance ne permettent pas de faire des prévisions de la propagation probable de ces maladies. Une telle prévision, si elle s'avérait fiable, permettrait aux autorités de santé publique de déclencher des actions préventives au bon moment et en fonction du niveau de risque attendu. D'ailleurs, l'identification des zones vulnérables au moment opportun reste un grand défi, étant donné la complexité des phénomènes liés à la transmission des maladies infectieuses. De ce fait, plusieurs approches ont vu le jour ces dernières années visant à simuler, et par la même occasion tenter de prédire, le comportement de phénomènes complexes, spécialement ceux se rapportant à la propagation des maladies infectieuses [Hsieh *et al.*, 2006].

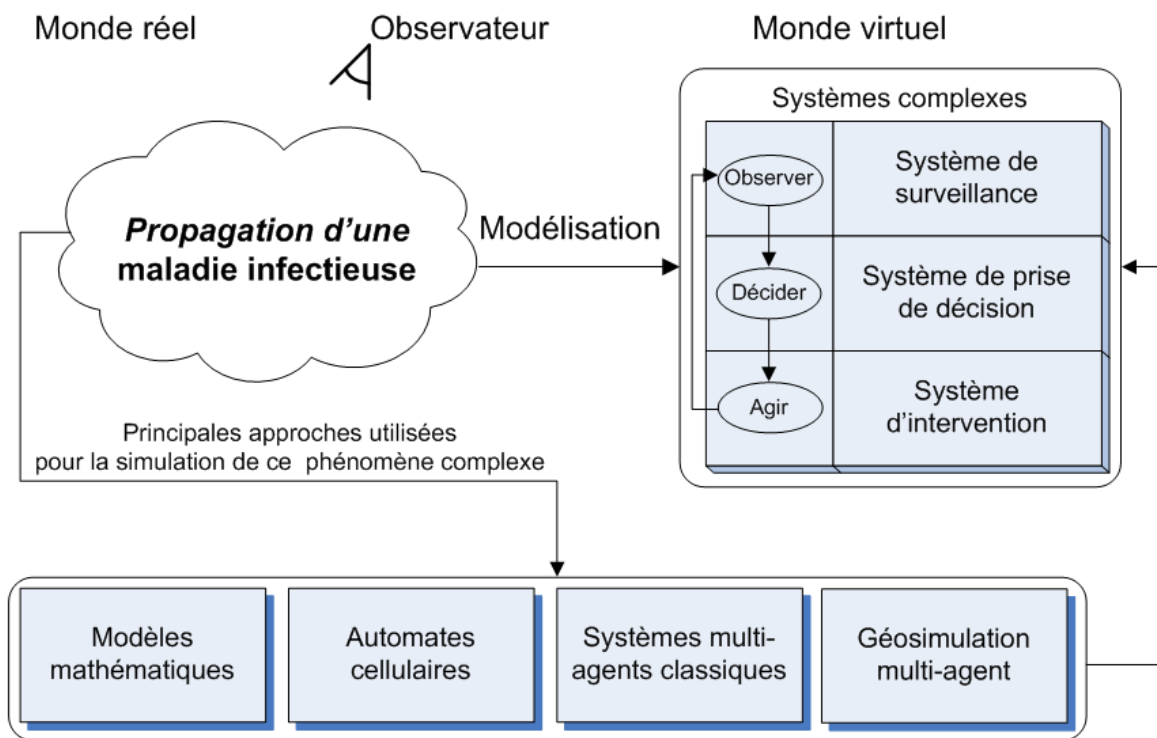


Figure 3: Principales approches utilisées pour la propagation des maladies infectieuses.

Dans la Figure 3, nous présentons justement les principales approches qui sont utilisées pour la simulation de la propagation d'une maladie infectieuse. Nous distinguons premièrement la simulation à base de modèles mathématiques qui utilise généralement des équations différentielles [Bowman *et al.*, 2005] (voir Section 3.1 du Chapitre III). Ces équations permettent de faire évoluer et interagir les populations impliquées dans la propagation d'une maladie. Cependant, une telle approche à elle seule présente certaines

lacunes. En effet, elle ne permet pas de prendre en considération les caractéristiques de l'espace géographique dans lequel les populations évoluent, sauf dans certains cas où quelques modèles utilisent un espace très agrégé, comme par exemple les *patchy models* [Liu *et al.*, 2006a], [Gourley *et al.*, 2008] (voir Section 3.1.2 du Chapitre III).

De plus, la simulation à base de modèles mathématiques, approche souvent qualifiée de macroscopique, apparaît comme plus apte à simuler un phénomène évoluant à grande échelle [Van Dyke Parunak *et al.*, 1998]. En revanche, elle s'avère inadéquate, voire incapable de modéliser l'impact d'événements localisés tels que l'application d'un traitement permettant de réduire ou stopper l'évolution d'une population responsable de la propagation d'une maladie et se trouvant dans une région bien déterminée (p.ex. application de larvicides et son effet sur la population larvaire de moustiques dans les puisards d'une municipalité). Par conséquent, les décideurs de la santé publique ne peuvent pas seulement s'appuyer sur les approches mathématiques lorsqu'ils doivent tenir compte de phénomènes reliés à des événements se produisant dans des zones précises de l'environnement géographique.

Nous distinguons deuxièmement la simulation à base d'automates cellulaires [Fu et Milne, 2003], [Beauchemin *et al.*, 2005], [Liu *et al.*, 2006b], [White *et al.*, 2009] (voir Section 3.2 du chapitre II). Cette approche présente également certaines limites, étant donné qu'elle modélise les caractéristiques spatiales des zones géographiques contenant des populations infectées par la maladie, mais elle ne permet pas de représenter les individus (ou groupe d'individus) ainsi que leur mobilité. De plus, le découpage de l'espace géographique en un ensemble de cellules de forme régulière est considéré comme artificiel et simpliste. Ce découpage ne reflète pas la "réalité du terrain" qui intéresse les décideurs (p.ex. régions administratives) afin d'intervenir dans les régions touchées par la maladie [Badariotti et Weber, 2002].

Nous distinguons troisièmement la simulation à base d'un système multi-agent classique [Emrich *et al.*, 2007], [Deng *et al.*, 2008], [Liu *et al.*, 2008], [Bauer *et al.*, 2009] (voir Section 3.3 du Chapitre II). Cette approche présente aussi d'autres limites puisqu'elle simule les vecteurs transmettant la maladie infectieuse comme des agents, mais

généralement elle ne profite pas des données fournies par un Système d'Information Géographique (SIG⁴ ou *GIS: Geographic information system*) afin de bien repérer ces agents dans l'environnement de simulation. De plus, cette approche n'est pas adaptée à la modélisation des populations contenant un nombre d'individus considérable. En effet, et selon le niveau de granularité recherché, il est souvent inutile et généralement impossible de représenter une population au niveau des individus ou même des groupes d'individus appartenant à cette population. De toute façon, les limites de ressources de calcul nous empêcheraient de représenter chaque individu par un agent surtout si la population que nous voulons simuler contient des millions, voire des milliards d'individus.

D'un autre côté, les trois approches que nous venons de mentionner ne modélisent pas tous les aspects concernant les interactions spatio-temporelles des populations impliquées dans la propagation de la maladie surtout que ces interactions se produisent sur un large territoire. Il s'agit d'une limitation importante, étant donné que ces interactions sont considérées comme des éléments essentiels pour créer des simulations crédibles. En outre et pour être utile pour la prise de décision, un système de simulation d'une zoonose devrait fournir à l'utilisateur la possibilité de spécifier des scénarios dans le cadre d'une analyse "*what-if*" [Haddad et Moulin, 2008] afin d'explorer, par exemple, l'influence de changements climatiques et les différentes stratégies d'intervention. D'ailleurs, les approches classiques de simulation utilisent peu souvent les données climatiques qui jouent un rôle important dans la transmission de l'agent pathogène responsable de l'infection. D'un autre côté, pour que la simulation soit plausible, il faut que les modèles utilisés soient bien calibrés, en utilisant notamment les données fournies par les systèmes de surveillance.

Nous pensons d'ailleurs qu'il est souhaitable qu'une approche de simulation de la propagation des maladies infectieuses soit capable entre autres de modéliser: (1) les acteurs de différentes espèces qui sont impliqués dans la transmission de l'infection en tenant compte de leurs cycles biologiques, (2) la localisation de ces acteurs dans l'espace en se

⁴ Un SIG est un système d'information permettant, à partir de diverses sources, de rassembler et d'organiser, de gérer, d'analyser et de combiner, d'élaborer et de présenter des informations spatialement référencées, autrement dit géoréférencées, ainsi que de produire des plans et des cartes. Ses usages couvrent les activités géomatiques de traitement, de partage et de diffusion de l'information géographique.
<http://www.cartographie.ird.fr/publi/documents/sig1.pdf>

basant sur un environnement géographique virtuel géoréférencé⁵ utilisant des données SIG et (3) les interactions de ces acteurs dans l'espace et dans le temps. De plus, de telles simulations doivent également permettre la modélisation de populations contenant un nombre considérable d'individus.

La géosimulation multi-agent (voir Section 3.5 du Chapitre II) pourrait éventuellement répondre à ces besoins [Benenson et Torrens, 2004], [Hu *et al.*, 2008]. D'ailleurs, le Groupe de Recherche en Informatique Cognitive (GRIC) de l'Université Laval dirigé par le Dr. Bernard Moulin a développé le simulateur *MAGS (Multi-Agent Geo-Simulation)* [Moulin *et al.*, 2003] (voir Section 3.5.2 du Chapitre II). Le développement de ce simulateur a été une étape importante montrant la faisabilité technique de création d'outils d'aide à la décision en se basant sur un puissant système de geosimulation multi-agent. Toutefois, cette approche présente elle aussi certaines limites puisqu'elle n'intègre pas les différents niveaux de granularité⁶ auxquels le phénomène peut être observé par les décideurs. Ces niveaux de granularité permettent de: (1) définir les échelles spatio-temporelles qu'il faut utiliser pour visualiser la progression de la maladie, (2) spécifier les niveaux de détail qu'il faut utiliser pour modéliser les différentes catégories d'acteurs (p.ex. individu, groupe et population) et (3) choisir la précision des modèles de simulation. En effet, un système multi-niveau (voir Section 3.6 du Chapitre II) peut aider l'utilisateur à élargir son spectre de compréhension d'un phénomène complexe. De nouvelles propriétés de ce phénomène peuvent apparaître en changeant le niveau de granularité surtout si les données sont disponibles pour pouvoir le faire. Le choix et la spécification de ces niveaux de granularité ont donc une incidence sur les résultats de la simulation [An *et al.*, 2005].

Par ailleurs, notre travail a fait partie du projet *MUSCAMAGS (MULTiSCALE Multi-Agent Geo-Simulation)* [Moulin, 2012] [Site Web: *MUSCAMAGS*] qui avait pour objectif général de développer des méthodes et des plateformes logicielles permettant des geosimulations

⁵ Un environnement géographique virtuel géoréférencé est une représentation virtuelle du territoire étudié. Il joue le rôle d'un support permettant par exemple de localiser, en utilisant des données SIG, les individus impliqués dans le phénomène simulé.

⁶ Un niveau de granularité est le degré de fragmentation d'un système en plus petites unités. Ce système ne peut plus être découpé en éléments plus petits quand on arrive à son niveau de granularité le plus fin (p.ex. les individus correspondent au niveau le plus fin pour une population).

multi-agents et multi-niveaux pour supporter l'aide à la décision dans des systèmes comportant des situations spatiales dynamiques faisant intervenir une multitude d'acteurs à plusieurs niveaux de granularité. De plus, nous sommes affiliés au projet *CODIGEOSIM* (*Geosimulation tools for simulating spatial-temporal spread patterns and evaluating health outcomes of communicable diseases*). Ce projet visait, entre autres, à offrir aux organismes gouvernementaux de meilleures stratégies pour prévenir la propagation de maladies infectieuses [Site Web: *CODIGEOSIM*]. Ces deux projets ont été supportés par GEOIDE, le réseau canadien de centres d'excellence en géomatique.

C'est dans ce contexte que notre recherche se situe. Nous essayons de contribuer à l'élaboration d'approches, techniques et modèles permettant la géosimulation multi-niveau de la propagation de zoonoses. Le fait que cette tendance soit très récente, qu'elle possède les avantages cités et qu'elle contribue clairement au développement d'outils d'aide à la décision, sont les raisons qui nous ont motivés à choisir ce contexte de recherche.

2 Problématique de la recherche

En se situant dans le contexte cité auparavant, la modélisation et la simulation d'un phénomène complexe comme celui de la propagation d'une zoonose reste un grand défi aussi bien sur le plan théorique que pratique. En effet, le phénomène qui nous intéresse résulte des interactions entre plusieurs individus de populations différentes. Ces individus évoluent dans l'espace et dans le temps en fonction de leurs besoins, mais également en fonction d'autres facteurs environnementaux et climatiques qui les influencent. De ce fait, il est difficile de modéliser tous ces aspects et d'ailleurs les approches existantes, que nous avons présentées dans la section précédente, présentent encore certaines lacunes qui limitent leurs utilisations pour produire des simulations pouvant satisfaire les attentes des décideurs de la santé publique.

Notre problématique générale peut donc s'énoncer comme suit :

Il y a un manque d'approches et d'outils génériques permettant de simuler la propagation d'une zoonose se produisant sur un large territoire (représenté par un environnement géographique virtuel géoréférencé) et qui peuvent supporter les aspects suivants :

- ✓ La simulation d'un grand nombre d'acteurs de différents types (individus, groupes ou populations d'humains et d'animaux), y compris ceux qui représentent des populations contenant un nombre considérable d'individus.
- ✓ La simulation des interactions spatio-temporelles entre ces multiples acteurs en tenant compte de leur localisation dans l'environnement en se basant sur des données géoréférencées.
- ✓ L'application et l'analyse des résultats de divers scénarios d'intervention dans un objectif d'aide à la décision.
- ✓ La simulation de tous ces aspects en utilisant des modèles à différents niveaux de granularité selon les échelles spatio-temporelles et selon les différentes catégories d'acteurs (population, groupe et individu).

Par ailleurs et pour apporter quelques éléments de réponse à cette problématique générale, nous visons à traiter les problématiques spécifiques suivantes:

Le niveau de granularité et l'aspect multi-modèle: Les approches actuelles de modélisation et de simulation de la propagation d'une zoonose n'intègrent pas les différents niveaux de granularité auxquels le phénomène peut être observé par les décideurs. De plus, il est difficile de choisir et de bien spécifier ces différents niveaux de granularité puisque c'est une notion qualitative et elle peut varier en fonction du problème ou du domaine traité, étant donné son caractère relatif (p.ex. la simulation à l'échelle d'un pays peut être considérée par un observateur comme étant à un niveau Macro, par contre un autre observateur peut la considérer à un niveau Méso s'il a une vision plus globale du phénomène étudié). D'un autre côté, les travaux actuels simulent la propagation des maladies infectieuses le plus souvent en utilisant une approche mono-modèle et ne profitent donc pas de la synergie que peut produire l'aspect multi-modèle à plusieurs niveaux d'abstraction.

Les interactions entre des acteurs contenant un nombre considérable d'individus: Les approches existantes ne modélisent pas les aspects importants concernant les interactions spatio-temporelles entre les différents types d'acteurs impliqués dans la propagation de la maladie et qui se produisent sur un large territoire. De plus, ces approches, comme les

systèmes multi-agents, se prêtent mal à la modélisation des populations contenant un nombre énorme d'individus. Il faut donc trouver d'autres alternatives de modélisation mieux adaptées à ce type de populations.

L'environnement géographique virtuel géoréférencé: Les approches actuelles n'intègrent pas dans l'environnement de simulation les caractéristiques géographiques qui influencent les cycles biologiques des espèces concernées (p.ex. zones qui sont favorables à la survie et à la prolifération des vecteurs propageant la maladie) et leurs comportements (p.ex. trajectoires de déplacements, corridors de migrations, etc.). D'ailleurs, ces approches ne permettent pas l'enrichissement de cet environnement (p.ex. avec des données de populations et des données climatiques). Cet enrichissement, qui donne à l'environnement son caractère informé⁷, permettrait entre autres d'améliorer l'efficacité de la simulation et de faciliter la modélisation de certains processus influençant la propagation d'une zoonose comme la migration des oiseaux.

Un outil générique de la simulation de la propagation des zoonoses: Il n'existe pas actuellement un outil permettant de simuler d'une façon générique la propagation des zoonoses. Cet outil pourrait être très utile aux décideurs dans le domaine de la santé publique et leur permettre d'appliquer plusieurs scénarios d'intervention dans le but de comprendre et estimer l'ampleur de l'évolution d'une zoonose. Cet outil pourrait éventuellement aider ces décideurs à intervenir sur le terrain après avoir pris des décisions éclairées⁸.

3 Hypothèse de la recherche

Nous avons mentionné que les approches comme les modèles mathématiques, les automates cellulaires, les systèmes multi-agents classiques et la géosimulation multi-agent présentent encore certaines lacunes qui limitent leurs utilisations pour produire des

⁷ Un environnement de simulation a le caractère informé si les données géométriques le caractérisant sont enrichies avec d'autres types de données pertinentes permettant de faciliter la simulation du phénomène étudié telles que des données sémantiques [Farenc *et al.*, 1999], biologiques, climatiques ou encore topologiques.

⁸ Une décision est éclairée si elle se base sur une étude préalable afin de bien comprendre et analyser le phénomène qui intéresse le décideur.

simulations plausibles⁹. Nous proposons donc une nouvelle approche, que nous nommons la géosimulation multi-niveau et nous formulons l'hypothèse que cette approche permet de remédier aux lacunes des méthodes actuelles quand elles sont utilisées seules et en bénéficiant de leurs avantages quand elles sont utilisées ensemble, et ce, à différents niveaux de granularité. Cette approche pourrait être utilisée pour simuler un phénomène aussi complexe que la propagation d'une zoonose. Nous pensons que cette approche peut être utilisée pour produire des simulations plausibles pour les raisons suivantes :

- ✓ Elle tient compte des différents niveaux de granularité auxquels le phénomène peut être observé par les décideurs. L'utilisateur peut choisir plusieurs combinaisons de niveaux de granularité afin de se rendre compte par exemple de l'influence du changement d'une échelle spatiale (p.ex. une simulation méso sur une ville par rapport à une simulation macro sur une province), d'une échelle temporelle (p.ex. un pas de simulation d'une journée par rapport à un pas de simulation d'une semaine), du niveau de catégorie d'acteurs (p.ex. simuler l'évolution d'un groupe d'individus par rapport à simuler l'évolution de toute une population) et du niveau de précision du ou des modèles de simulation utilisés (p.ex. simuler la mobilité de chaque individu par rapport à simuler les flux de déplacements de tous les individus).
- ✓ Elle peut être utilisée pour suivre dans l'espace et dans le temps la propagation d'une zoonose en simulant l'évolution des acteurs (incluant des populations contenant un nombre considérable d'individus) participant à cette propagation ainsi que les interactions qui peuvent se produire entre ces mêmes acteurs dans un environnement géographique virtuel informé représentant un large territoire.
- ✓ Elle se distingue des autres approches de simulation par sa prise en compte de l'espace en se basant sur des données géoréférencées. De ce fait, elle fournit une représentation analogique suffisamment fidèle à la réalité géographique. Les décideurs ont besoin justement de ce type de représentation afin d'avoir un repère par rapport à ce qui se passe sur le terrain.

⁹ Une simulation est plausible si elle donne un résultat permettant de satisfaire au moins une partie des attentes de l'utilisateur.

- ✓ Elle peut être utilisée pour développer un outil générique simulant la propagation d'une zoonose contenant un système de spécification de scénarios pouvant être utile pour les décideurs de la santé publique. Il faut signaler que notre laboratoire de recherche (GRIC) a déjà montré la faisabilité technique de création d'outils de géosimulation pour l'aide à la décision dans plusieurs projets. Ces projets ont donné à ce jour de bons résultats. Nous pouvons citer par exemple les travaux sur la simulation des foules [Moulin *et al.*, 2003], [Moulin *et al.*, 2004], [Bouden *et al.*, 2007], la propagation des feux de forêt [Sahli et Moulin, 2007], l'étude de comportements des clients dans un centre d'achat [Ali et Moulin, 2007], la planification des suites d'actions dans le domaine militaire [Haddad et Moulin, 2008], le trafic routier [Chaker et Moulin, 2008] ou encore l'identification des risques d'accident pour les trains [Mekni *et al.*, 2008].

4 Objectifs de la recherche

Étant donné la problématique et l'hypothèse de recherche énoncées dans les sections précédentes, notre objectif général est de concevoir de nouvelles approches et/ou de nouveaux modèles théoriques permettant de développer plus facilement un outil générique d'aide à la décision qui serait capable de simuler la propagation d'une zoonose, et ce, selon les attentes des responsables de la santé publique.

Pour atteindre ce but, nous visons les objectifs spécifiques suivants :

1. Proposer **une approche de géosimulation multi-niveau permettant l'intégration des différents niveaux de granularité** auxquels une zoonose peut être observée et comprise par les décideurs de la santé publique et selon la disponibilité des données qui alimentent les modèles de simulation. Cette nouvelle approche devrait remédier aux lacunes des approches actuelles (p.ex les modèles mathématiques, les automates cellulaires et les systèmes multi-agents classiques) quand elles sont utilisées seules en bénéficiant de leurs avantages quand elles sont utilisées ensemble, et ce, à différents niveaux d'abstraction.
2. Proposer **un modèle théorique permettant de modéliser et simuler les interactions spatio-temporelles d'un grand nombre d'acteurs** de différents types (individus, groupes ou populations d'humains et d'animaux), y compris ceux qui représentent des populations contenant un nombre considérable d'individus.

3. Proposer **une approche permettant de générer un environnement géographique virtuel informé** dans lequel se déroule la géosimulation en utilisant des données géoréférencées. Cet environnement devrait inclure les caractéristiques géographiques du paysage (p.ex. l'habitat favorable) pouvant influencer la propagation de la zoonose, mais également enrichi avec des données pertinentes (p.ex. biologiques, topologiques, climatiques, etc.) pouvant aider à modéliser certains processus importants pour le phénomène étudié (p.ex. la migration des oiseaux).

4. Développer **des prototypes permettant d'appliquer notre approche de géosimulation multi-niveau à au moins deux cas de zoonoses** afin de montrer sa généralité. Ces outils devraient permettre aux décideurs dans le domaine de la santé publique d'appliquer plusieurs scénarios d'intervention dans le but de comprendre et estimer l'ampleur de l'évolution d'une zoonose. Ces outils pourraient par la suite éventuellement aider ces décideurs à intervenir sur le terrain après avoir pris des décisions éclairées.

5 Structure de la thèse

Cette thèse est composée des huit chapitres suivants :

- ✓ Le chapitre I (le présent chapitre) est une introduction générale.
- ✓ Le chapitre II présente l'état de l'art relativement à la modélisation et la simulation des phénomènes complexes.
- ✓ Le chapitre III donne un aperçu du domaine qui nous intéresse qui est la propagation des maladies infectieuses.
- ✓ Le chapitre IV présente une vue d'ensemble de l'approche de géosimulation multi-niveau que nous proposons.
- ✓ Le chapitre V présente le modèle théorique *MASTIM* qui est une partie intégrante de l'approche proposée.
- ✓ Le chapitre VI présente la démarche proposée pour générer l'environnement géographique virtuel informé en utilisant les concepts définis dans le modèle *MATSIM*.
- ✓ Le chapitre VII présente les prototypes développés permettant de géosimuler la propagation de deux exemples de zoonoses.
- ✓ Le chapitre VIII présente une synthèse des contributions et les limites de cette thèse ainsi que les perspectives d'avenir.

Nous présentons d'ailleurs dans la Figure 4 le lien entre les objectifs de la thèse et ces chapitres. En effet le chapitre IV répond au premier objectif, le chapitre V répond au

deuxième objectif, le chapitre VI répond au troisième objectif et finalement le chapitre VII répond au quatrième objectif.

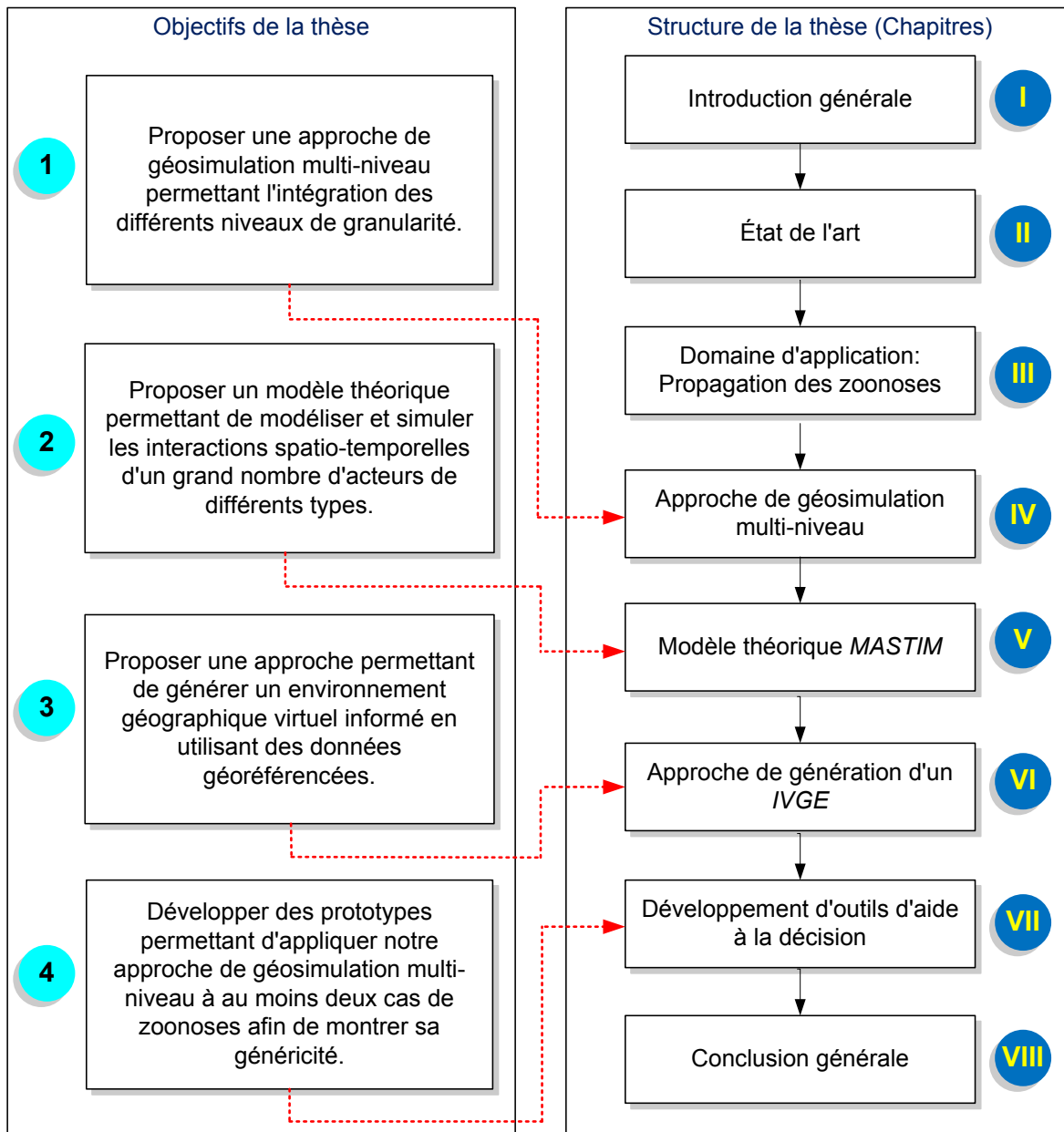


Figure 4: Liens entre les objectifs de la thèse et ses chapitres.

Chapitre II: État de l'art

1 Introduction

Étant donné notre problématique de recherche présentée dans le chapitre précédent, nous avons commencé par une étape d'exploration et d'analyse des travaux existants. Dans ce chapitre nous faisons une revue de littérature pour comprendre et caractériser d'une façon générale et brève les systèmes complexes. Nous essayons d'ailleurs d'expliquer la notion de complexité qui est une source de controverse chez les scientifiques. Nous présentons par la suite les principales approches et méthodes utilisées pour modéliser et simuler les systèmes complexes.

2 Présentation des systèmes complexes

Dans le monde réel, on observe de nombreux phénomènes de plus en plus complexes. D'ailleurs, l'accroissement rapide des possibilités techniques, la modification du contexte économique, l'irruption des machines intelligentes dans les organisations en tant qu'acteurs à part entière aux côtés de l'homme, nous incite à réaliser des systèmes sophistiqués. Les architectures des solutions informatiques s'éloignent d'une structure centralisée et tendent vers des architectures décentralisées afin de mieux s'intégrer dans une réalité de plus en plus complexe [Site Web: *Complex Systems*], [Axelrod et Cohen, 2000].

Par ailleurs, l'étude des systèmes complexes est très vaste. En fait, une simple recherche sur le Web, via un moteur de recherche, nous propose plus de 150 millions d'entrées dont plusieurs centaines de centres de recherches universitaires et presque autant de revues spécialisées. La plupart des études se présentent sous la forme de créations et d'applications de "méthodes de résolution de problèmes". Ces méthodes sont de plus en plus informatisées, mais laissent souvent encore à la charge des utilisateurs la tâche de formulation de leur problème dans des termes compatibles avec le langage informatique [Abbott, 2007]. De ce fait, les travaux sur les systèmes complexes constituent désormais un champ disciplinaire à part entière avec, par exemple, la conférence européenne dédiée

justement à l'étude de ces systèmes (*European Conference on Complex Systems*) [Site Web: ECSS].

2.1 Définition et caractéristiques

Nous n'avons pas trouvé une définition définitive des systèmes complexes étant donné que les auteurs ne s'entendent pas complètement sur la notion de complexité. Par contre, ils semblent être de même avis quand il s'agit de faire la différence entre un système complexe et un système compliqué. En fait, ce n'est pas parce qu'on ne comprend pas les processus ou les facteurs qui sont impliqués dans un système qu'il est nécessairement complexe. Il peut être donc tout simplement compliqué en fonction du degré de compréhension de l'observateur ou de l'utilisateur du système. Étymologiquement parlant, le mot compliqué (du latin *cum pliare*, empiler avec) signifie qu'il faut du temps et du talent pour comprendre l'objet d'étude alors que le mot complexe (du latin *cum plexus*, attaché avec) signifie qu'il y a beaucoup d'intrications, que "tout est lié" et que l'on ne peut étudier une partie du système de façon isolée [Site Web: *Wikipedia*]. De ce fait, les systèmes complexes sont généralement compliqués, mais le contraire n'est pas forcément vrai. Si on se base par exemple sur la revue *Science*, il est sorti en 1999 un volume dédié aux systèmes complexes [Site Web: *Science*]. On peut d'ailleurs retrouver plusieurs définitions dont notamment celle de Goldenfeld, Kadanoff et Brian Arthur. Ces derniers expliquent qu'un système complexe est un système très structuré qui montre une très grande variation au niveau de ses structures. Par contre, Whitesides et Ismagilov mentionnent dans la même revue qu'un système complexe est un système en évolution qui est très sensible à ses conditions initiales ou à une très petite perturbation. Il contient un nombre assez élevé de composantes qui interagissent. Il peut avoir également plusieurs branchements possibles qu'il peut prendre. D'un autre côté, Weng, Bhalla et Iyengar indiquent qu'un système complexe est un système dont les fonctions sont difficiles à vérifier [Site Web: *Complex Systems*].

Il n'existe donc pas une définition formelle largement acceptée de ce qu'est un système complexe. Cependant, certains auteurs peuvent qualifier un système de complexe par les trois propriétés suivantes : (1) S'il a beaucoup de composants, (2) si son comportement n'est pas immédiatement prévisible, et (3) s'il en émerge des propriétés d'auto-organisation [Murray, 1995]. De ce fait, un système complexe présente plusieurs caractéristiques. Peut-

être parmi les plus importantes sont l'auto-organisation et l'émergence de structures cohérentes comme l'apparition de certains motifs à un niveau supérieur (Figure 5) [Parrott, 2002]. Afin de mieux comprendre ce phénomène, nous pouvons citer l'exemple d'un bocal rempli de riz et de raisins. Si on agite ce contenant, nous pouvons facilement remarquer que les raisins vont se rassembler pour former un certain regroupement au-dessus du riz. De ce fait, ce regroupement émerge suite aux interactions des éléments du système. Cependant, cette notion d'émergence est encore à l'état de discussion, et un sujet de controverses [Arthur, 1993], [Sterman, 2000], [Sawyer, 2005].

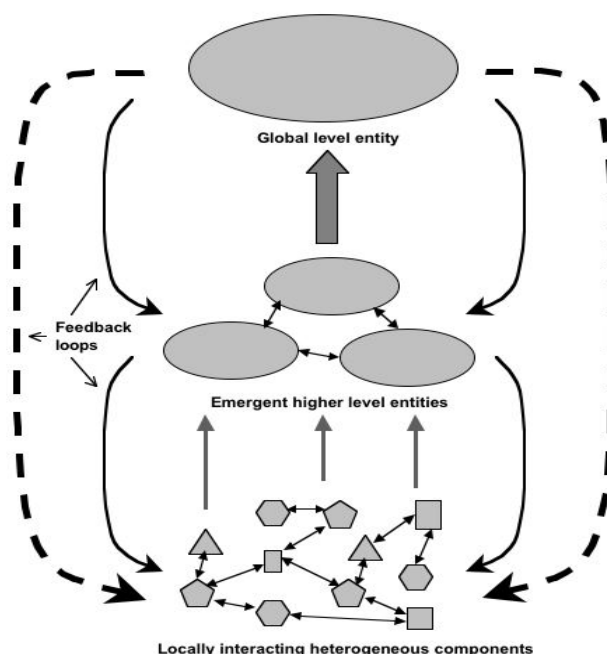


Figure 5: Modèle conceptuel typique d'un système complexe [Parrott, 2002].

De plus, un système complexe est caractérisé par le fait que la connaissance d'une partie du système ne permet pas d'affirmer que le reste du système est en moyenne dans le même état. En effet, plusieurs comportements possibles du système sont en compétition. Certains sont simples, d'autres chaotiques (désordonnés). Le système est souvent à la frontière entre les deux et alterne entre ces deux types de comportement. En outre, les interactions entre les composantes des systèmes ne sont pas triviales. En effet, ce n'est pas simplement toutes les composantes qui interagissent avec toutes les autres. Il y a au moins des liens privilégiés entre elles. Un système complexe peut être également décrit d'une façon plus globale par des boucles de rétroaction (*Feedback loops*) (Figure 5). Dans ce cas, l'état d'une entité a une

influence sur son état futur en lien avec l'état d'autres entités. Ces entités peuvent être elles-mêmes des systèmes complexes (par exemple, une société est composée d'humains eux-mêmes composés de cellules) [Weisbuch, 1989], [Brodu, 2007].

2.2 Notions de complexité

Le terme "complexité" est souvent interprété de façons très diverses selon le domaine d'étude. Chacune de ces interprétations est claire dans son domaine propre, mais c'est l'accumulation de toutes ces visions de la complexité qui crée la confusion. D'une façon générale, le concept de complexité s'attache davantage à la nature même du phénomène observé plutôt qu'aux modalités de résolution des problèmes [Volle, 2003].

Nous devons préciser que nous ne parlons pas dans ce contexte de la "classique" complexité algorithmique, mais plutôt de la complexité du monde réel. Prenons l'exemple des deux nombres binaires suivants : 101010101010101010 et 11010011000101101100. Ces deux nombres sont de la même longueur, mais le premier est moins complexe, car on peut le "résumer" à dix fois le motif '10' alors que l'autre ne peut que s'énumérer. C'est d'ailleurs l'idée intuitive derrière la mesure de complexité de Kolmogorov. Ce dernier avait défini la complexité d'un objet fini par la taille du plus petit programme informatique (au sens théorique) qui permet de produire cet objet [Bally, 2006]. D'un autre côté, la complexité d'après Rosnay [Rosnay, 1975] est attribuable aux facteurs suivants: (1) une grande variété des composants possédant des fonctions spécialisées; (2) des éléments organisés en niveaux hiérarchiques internes; (3) des interactions non-linéaires; (4) une difficulté voire une impossibilité de dénombrer de façon exhaustive les éléments qui constituent le système; et (5) une grande variété des liaisons possibles ne sont pas pour la plupart linéaires [Rosnay, 1975].

2.3 Exemples de systèmes complexes

Il existe bien sûr une infinité d'exemples de systèmes complexes. Les auteurs citent souvent le cas d'une colonie de fourmis qui échangent des phéromones et bâtissent une fourmilière, mais aucune fourmi n'a conscience de la fourmilière [Benkirane, 2006]. Un autre exemple serait un réseau de gènes qui interagissent par activations et inhibitions. Un ensemble de gènes activés définit un tissu. D'ailleurs, les gènes activés et inhibés ne sont pas les mêmes

dans les cellules de la peau ou dans celles d'un muscle. Un autre exemple serait la dynamique d'une cellule qui est constituée de protéines en réactions chimiques, son évolution permet une adaptation au milieu. Citons également le cas du cerveau humain qui est composé d'un très grand nombre de cellules, les neurones, de l'ordre de dix milliards. Ces cellules échangent des signaux électriques par l'intermédiaire de leurs synapses. Même si les divers types neuronaux sont en nombre réduit, c'est par leur structure de connexions qu'ils diffèrent. Un autre exemple intéressant est celui d'un tas de sable qui provoque des collisions entre les grains qui font naître des avalanches. Citons encore un vol d'étourneaux ou un troupeau de moutons, la propagation d'une épidémie, d'une rumeur ou du bouche-à-oreille sur un nouveau produit, des robots modulaires, des réseaux de criminalité, le développement d'un embryon, etc. [Site Web: CSCS], [Benkirane, 2006].

Tableau 1: Une hiérarchie hypothétique de la complexité du monde réel [Boulding, 1956].

Niveau		Caractéristiques	Exemples	Disciplines concernées
1.	Structure charpente	Statistique, Passivité	Le cristal, l'atome, la molécule	Représentations verbales iconiques/disciplines
2.	Objet actif	Structure dynamique, mouvements déterminés	L'horloge, le système solaire, la machine conventionnelle	Physique, sciences naturelles, lois de la mécanique
3.	Objet régulé	Homéostasie, rétroaction, régulation	Le thermostat, le régulateur à boules	Cybernétique, théorie de rétroaction
4.	Système ouvert	Structure auto-adaptative	La cellule, la flamme, les organismes en général	Théorie du métabolisme, information, ADN
5.	Organismes primaires	Spécialisation des fonctions, division du travail	La plante, les organismes de type végétal	Botanique
6.	Animal	Émergence de la mémoire, information/représentation	L'animal	Zoologie, théorie des automates
7.	Humain	Émergence de la métacognition, information/décision	L'être humain	Biologie, psychologie, symbolisme, épistémologie
8.	Organisation sociale	Rôles, communication, conscience sociale, populations	La famille, les associations diverses	Lois statistiques, histoire
9.	Systèmes transcendants	L'inaccessible, question sans réponses	L'idée de Dieu	

De plus, la réaction d'une foule suite à un événement inattendu est très complexe puisqu'elle fait intervenir la psychologie des individus comme leurs identités sociales. Citons finalement un exemple de système artificiel. Il s'agit d'un système informatique faisant intervenir de nombreux composants électroniques différents, que ce soit au niveau

des transistors ou des portes logiques. En fait, la grande majorité des systèmes naturels ou artificiels sont de nature complexe [Site Web: *SFT*]. Boulding [Boulding, 1956] a présenté d'ailleurs une classification des systèmes complexes selon leur degré de complexité (Tableau 1).

3 Modélisation et simulation des systèmes complexes

Dans diverses disciplines, on cherche, de différentes manières, des schémas de raisonnement permettant de concevoir un phénomène donné comme un "tout" et d'expliquer pourquoi et comment le comportement de l'ensemble des éléments qui constituent ce phénomène est autre que celui de chacun de ces éléments, pris séparément. Dans un certain nombre de domaines, il devient de plus en plus clair qu'analyser un phénomène donné en termes de ses composants est tout aussi inefficace sur le plan conceptuel que sur le plan opérationnel. Le processus traditionnel d'analyse se révèle inadéquat à la fois pour comprendre et pour agir. C'est ainsi que la décomposition des organismes vivants en cellules, réflexes ou comportements, ou encore en particules de matière ou en micro-unités sociales, ne paraît plus adéquate pour traiter les données expérimentales. De nouvelles théories se sont développées dans les domaines les plus divers (comme notamment la théorie de l'information, la théorie des jeux et des décisions, l'Intelligence artificielle). Le dénominateur commun de ces théories est le fait que l'on distingue toujours, sur le plan conceptuel, le tout de l'ensemble de ses parties. La Théorie Générale des Systèmes (qui matérialise ce "dénominateur commun" sur le plan théorique) remonte à 1937, avec l'œuvre du biologiste Ludwig Von Bertalanffy [Pouvreau, 2013].

Cette théorie a abouti à la systémique qui vise à modéliser des systèmes quelle que soit leur nature, qu'ils soient existants ou que l'on veuille les créer. Les techniques actuelles de maquettage ou de prototypage rapide permettent de simuler la dynamique de fonctionnement de divers types de systèmes [Site Web: *SFT*]. En effet, la simulation informatique peut être définie comme étant une technique permettant d'imiter le comportement d'un système. C'est un processus qui est composé de trois principales étapes inter-reliées. La première étape consiste en la conception ou le choix de modèles permettant de représenter le phénomène étudié. Il s'agit entre autres de rassembler les données et/ou les connaissances sur le système à simuler. Certaines d'entre elles sont typiquement

obtenues en utilisant des senseurs physiques ou humains. Par contre, d'autres sont obtenues en interviewant des experts du domaine ou en appliquant des techniques d'acquisition de la connaissance. La conception des modèles se fait donc à partir de ces données et/ou des connaissances acquises à partir des expériences antérieures sur des systèmes similaires. La deuxième étape consiste en l'exécution de ces modèles sur un ordinateur. Finalement, la dernière étape consiste en l'analyse des résultats de cette exécution. Des tests sont faits sur les données générées à partir des modèles mentionnés auparavant en faisant appel par exemple à des analyses statistiques. L'analyse la plus basique consisterait tout simplement à "regarder" les données et à en déduire des conclusions [Fishwick, 1995].

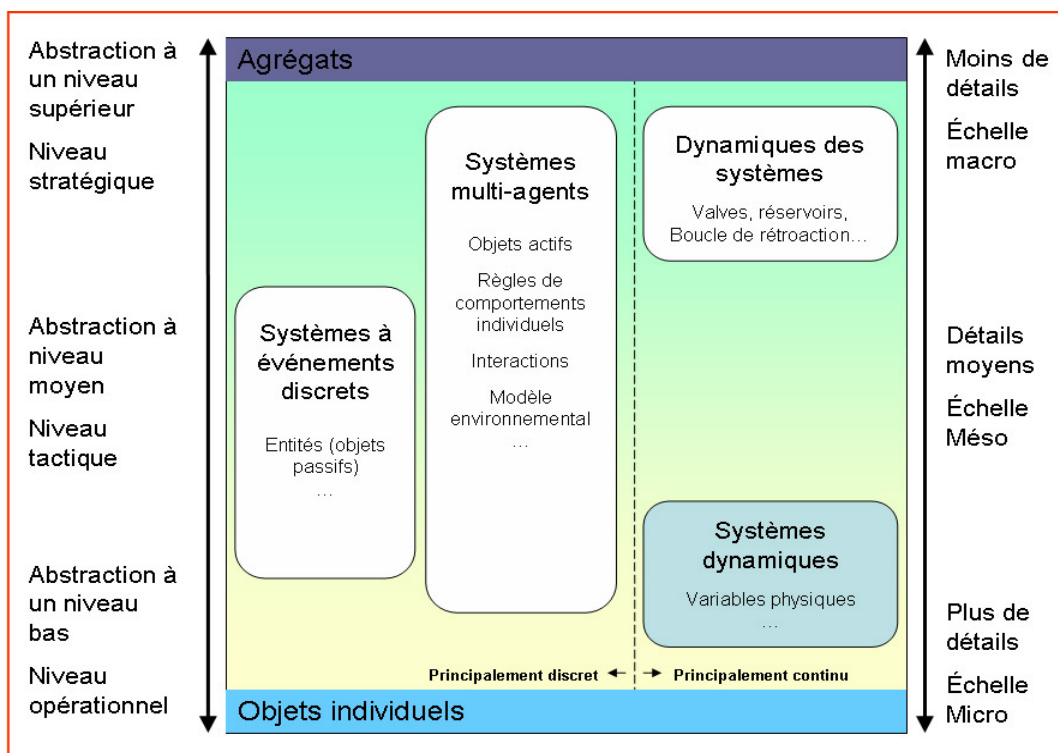


Figure 6: Approches de simulation sur plusieurs niveaux d'abstraction [Borshchev et Filippov, 2004].

Plusieurs approches ont été proposées pour simuler des phénomènes complexes, parmi lesquelles on peut distinguer deux catégories: (1) les approches qui traitent la plupart du temps des processus continus et (2) les approches qui utilisent un temps de simulation discret. La classification peut être également faite en fonction du niveau d'abstraction utilisé. Par exemple, la dynamique des systèmes (voir Section 3.1.2) s'intéresse aux agrégats et elle est située au niveau le plus élevé d'abstraction (échelle macro, avec moins

de détails). Par contre, les systèmes multi-agents (voir section 3.3) sont généralement utilisés à un niveau plus bas (échelle micro, avec plus de détails). Cependant, ces systèmes peuvent être également employés à travers tous les niveaux d'abstraction (Figure 6) [Borshchev et Filippov, 2004].

Nous présentons dans cette section les principales approches et techniques qui sont utilisées pour modéliser et simuler les systèmes complexes d'une façon générale. Nous commençons par les systèmes continus non-linéaires en présentant les systèmes à base d'équations mathématiques et la dynamique des systèmes. Nous présentons ensuite les systèmes à base d'automates cellulaires qui ont l'avantage d'ajouter une composante spatiale à la simulation, mais qui ne gèrent pas la mobilité des acteurs impliqués dans le phénomène étudié. Les systèmes multi-agents sont alors présentés puisqu'ils apportent une solution à cette lacune relativement à la mobilité des agents. Cependant, ces systèmes ne présentent pas l'espace d'une façon réaliste puisqu'ils utilisent rarement ou inadéquatement les données géographiques. De ce fait, nous présentons la géosimulation multi-agent pouvant résoudre ce problème, mais qui ne prend pas en considération les différents niveaux de granularité auxquels le phénomène peut être observé et qui semblent être un facteur très important afin de produire des simulations plausibles. Nous présentons alors les systèmes multi-niveaux qui essaient de simuler justement le phénomène étudié selon plusieurs niveaux de granularité. Nous terminons cette section en présentant les systèmes holoniques qui se caractérisent par la prise en compte des notions de hiérarchies et de structures ordonnées. Ces systèmes pourraient éventuellement servir à établir certains éléments de réponse à notre problématique.

3.1 Systèmes continus non-linéaires

Un système est continu, par opposition à un système discret, lorsque des variations des grandeurs physiques de sortie qui le caractérisent sont des fonctions du temps continues et que l'on peut donc définir ces grandeurs à chaque instant. Un système continu est un système mettant en jeu des signaux continus. Ils sont donc constitués d'éléments caractérisés par une mesure qui peut prendre une infinité de valeurs : température d'une pièce ou d'un objet, vitesse d'un mobile, niveau dans un réservoir, etc. La gestion de ces systèmes fait appel à des outils mathématiques aptes à la représentation de la dynamique

continue comme les équations différentielles assorties de diverses transformations. Nous nous intéressons bien sûr aux systèmes non linéaires, étant donné qu'ils peuvent être utilisés pour représenter les systèmes complexes [Wolfram, 2002].

3.1.1 Simulation à base d'équations mathématiques

On peut faire remonter l'origine de la simulation numérique intégrant des équations mathématiques aux travaux concernant la méthode des différences finies. Cette méthode permet de rechercher la solution d'une équation aux dérivées partielles. Une des principales applications de cette méthode est la résolution numérique des équations de la dynamique atmosphérique afin de prédire le temps. La simulation à base d'équations mathématiques se déroule schématiquement suivant 5 étapes. (1) On commence par définir le problème physique que l'on souhaite simuler et on précise si possible les questions auxquelles le modèle numérique doit tenter d'apporter une réponse. (2) On décrit par la suite ce problème à l'aide d'un système d'équations et de conditions aux limites convenablement choisies. (3) On remplace les équations différentielles par des équations algébriques. La résolution numérique de ces équations permet d'atteindre des solutions qui décrivent adéquatement la réalité physique du système. (4) On résout les systèmes d'équations algébriques à l'aide d'algorithmes numériques choisis en fonction de leur rapidité de calcul, de leurs besoins en taille mémoire et de la précision désirée. (5) On finit par tester le modèle numérique afin de vérifier que les algorithmes employés convergent vers une solution satisfaisante [Farge, 1988]. Nous présentons d'ailleurs, dans le chapitre III (section 3.1), quelques modèles mathématiques qui nous intéressent et qui sont utilisés pour simuler la propagation des maladies infectieuses.

3.1.2 Dynamique des systèmes

La dynamique des systèmes (DS) a été développée vers la fin des années 50 par Jay W. Forrester, travaillant au *MIT Sloan School of Management* [Site Web: MIT] qui a lancé une nouvelle discipline qu'il a appelée "dynamique industrielle" [Forrester, 1958]. Son but principal était d'étudier les caractéristiques de la rétro-information¹⁰ (*information-feedback*)

¹⁰ La rétro-information consiste à communiquer aux personnels la réaction des clients vis-à-vis de leur travail. Elle permet à l'employé de savoir ce qu'il fait correctement, quels points il devra améliorer et comment il peut le faire [Kirkwood, 1998].

des activités industrielles pour montrer comment la structure d'organisation, l'amplification (dans les politiques), et les délais (dans les décisions et les actions) agissent l'un sur l'autre pour influencer le succès d'une entreprise. Depuis cette époque, la méthode a été appliquée à une grande variété de systèmes sociaux [Kirkwood, 1998], [Schieritz et Milling, 2003], [Site Web: SDS].

En outre, la DS s'adresse à quiconque veut analyser et comprendre le comportement dynamique de structures complexes, quelle que soit leur nature. Il faut d'ailleurs mentionner que le terme « dynamique » sous-entend explicitement le changement. Ce concept de dynamique implique l'analyse des facteurs qui créent le changement, ou qui s'y opposent. De ce fait, la DS est destinée à étudier le comportement de systèmes dont la structure et les frontières, bien que déterminées, peuvent être changeantes, donc pas forcément connues d'avance [Kirkwood, 1998]. Les systèmes complexes analysés sont présentés en utilisant principalement des boucles de rétroaction représentant l'action en retour d'un effet sur sa propre cause. En effet, le comportement des structures d'un système résulte d'un ensemble de facteurs, de poussées ou d'inerties sans cesse évolutives. L'effet de ces facteurs, responsables du changement, modifie l'état du système. Cependant, ces facteurs ne sont pas figés, ils évoluent souvent de manière structurelle, c'est-à-dire en fonction de l'état de ce système qu'ils ont eux-mêmes contribué à changer. D'où justement la présence d'une boucle de rétroaction (Figure 7) [Karsky, 2004].

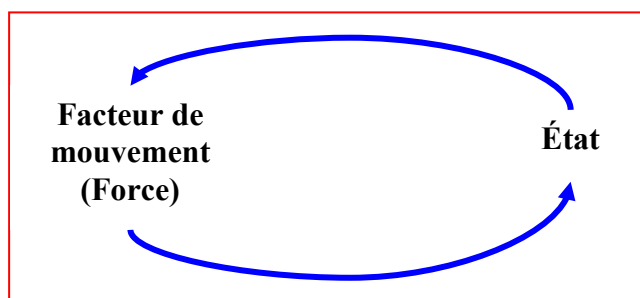


Figure 7: Boucle de rétroaction (*Feedback loop*) [Karsky, 2004].

On distingue deux types de boucles. Il y a les boucles négatives qui sont appelées également des boucles régulatrices ou encore stabilisatrices. Dans ce cas, une variation (accroissement ou baisse) cause un effet en sens inverse. Par exemple, un accroissement du prix d'un produit aura un effet contraire sur sa consommation. Il y a également les boucles

positives qui sont appelées également des boucles explosives. Par exemple, la publicité ou le bouche à oreille augmente la vente d'un produit et plus un produit est vendu plus il sera populaire, ce qui pourra encore augmenter sa vente [Kirkwood, 1998], [Karsky, 2004].

Par ailleurs, les boucles régulatrices ou explosives peuvent comporter un ou plusieurs délais entre variables ce qui représente la non-instantanéité des relations de cause à effet. De plus, les relations causales n'ont pas toutes les mêmes significations. Par exemple, la relation entre la production d'un produit et son stock de réserve est une relation physique puisqu'une quantité produite remplit automatiquement le stock d'un magasin. Par contre, la relation entre stock et production est une relation à la fois d'information et de décision. On peut donc décider de continuer à produire quel que soit le niveau des stocks. Il y a aussi toujours un certain délai entre l'information et la décision qui doit en résulter. L'interconnexion de plusieurs boucles peut justement engendrer un système complexe. D'ailleurs, le comportement temporel d'une structure composée de plusieurs boucles ne peut se déduire aisément de l'analyse du comportement individuel de chaque boucle. De plus, la plupart des relations d'influence entre les variables comportent un élément de non-linéarité entre cause et effet. Par exemple, l'augmentation du prix d'un produit entraîne la diminution de sa demande. Cette diminution est justement non linéaire. Elle ne forme pas une droite, mais plutôt une courbe [Karsky, 2004], [Sterman, 2000].

L'utilisation de la DS afin d'analyser un système complexe consiste à suivre certains processus et étapes de modélisation. En effet, on commence par mettre en évidence l'ensemble des éléments du problème représentant les variables du système. C'est l'étape d'analyse causale ou d'analyse structurelle permettant d'obtenir un modèle qualitatif simple représentant quelques boucles de rétroaction. On n'est pas obligé d'être exhaustif au départ, car les variables manquantes peuvent être ajoutées plus tard. Il faut par la suite identifier parmi les variables du système, celles qui paraissent être des variables d'accumulation. Ce sont les variables d'état qu'on appelle également "niveaux" par référence au niveau du liquide dans un récipient. Il faut identifier également les flux qui remplissent ou qui vident ces variables de niveau. De plus, il faut identifier les variables qui ont une influence sur ces flux qui sont normalement des variables d'information ou de décision. Par exemple, le

stock représentant un niveau (réservoir) est rempli par le flux de production (valve) et vidé par la livraison du produit correspondant (Figure 8 et Figure 9) [Karsky, 2004].

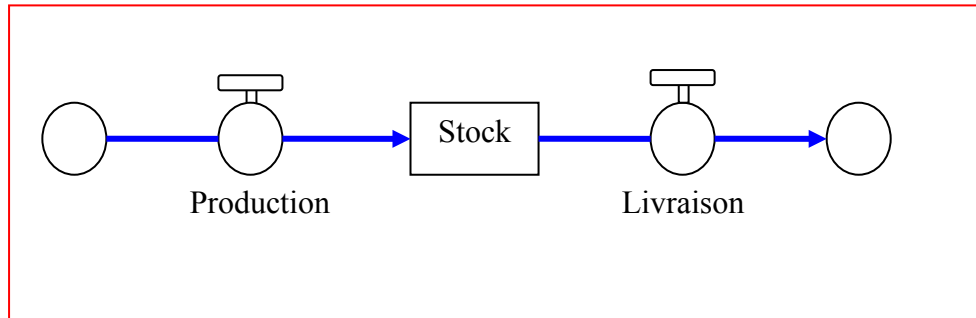


Figure 8: Exemple d'une variable d'état et de deux variables de flux [Karsky, 2004].

On passe ensuite par une étape de formalisation et de quantification en utilisant les équations différentielles pouvant représenter un système en mouvement continu. Finalement, la simulation nous permettra entre autres de tester le modèle en introduisant explicitement la variable temps. Il ne faut pas bien sûr oublier de terminer par les phases de calibration et de validation du modèle [Kirkwood, 1998].

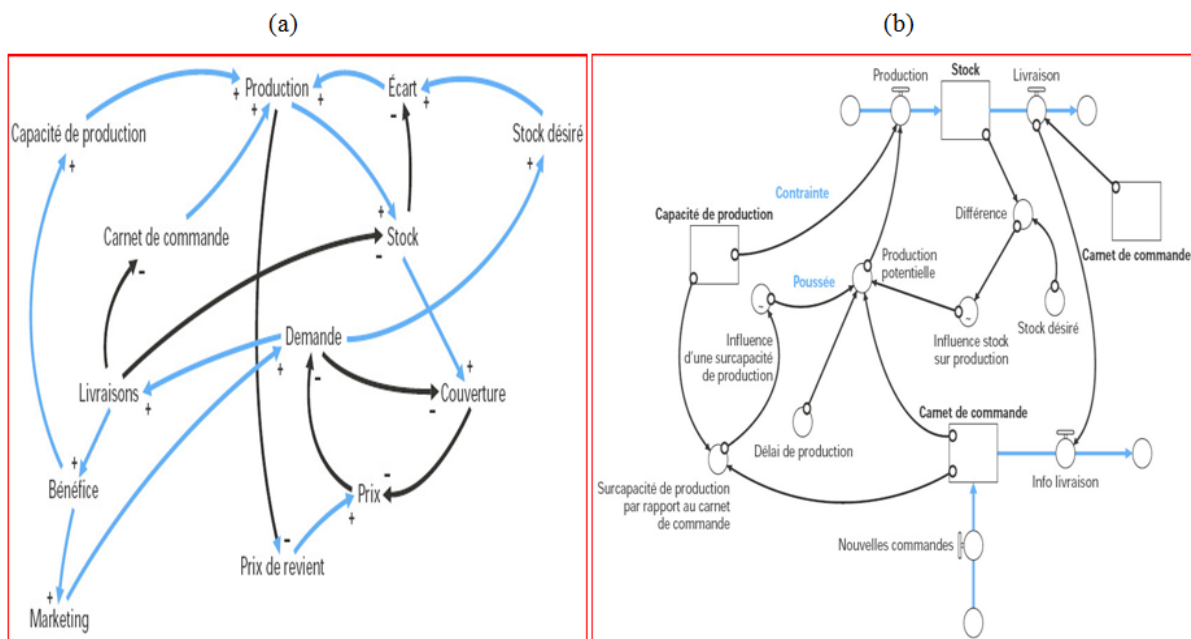


Figure 9: Passage du diagramme causal (a) au modèle dynamique (b) pour l'exemple de la gestion de production d'un produit [Karsky, 2004].

3.2 Automates cellulaires

Les automates cellulaires ont été inventés par Stanislaw Ulam et John Von Neumann à la fin des années 40 au *Los Alamos National Laboratory* [Site Web: *LANL*]. Plus tard, l'idée a été reprise dans des domaines autres que l'informatique et implémentée dans un grand nombre d'applications [Delorme, 1998].

Nous pouvons définir un automate cellulaire à l'aide de deux types de caractéristiques: structurelles et fonctionnelles. Les premières concernent l'aspect topologique du réseau cellulaire, les secondes concernent l'aspect dynamique de l'évolution du réseau au cours du temps. En effet, un automate cellulaire consiste en une grille régulière de «cellules» pouvant chacune prendre à un instant donné un «état» parmi un ensemble fini. L'état d'une cellule au temps $t + 1$ est fonction de l'état au temps t d'un nombre fini de cellules appelé son «voisinage». À chaque nouvelle unité de temps, les mêmes règles sont appliquées simultanément à toutes les cellules de la grille, produisant une nouvelle «génération» de cellules dépendant entièrement de la génération précédente [Weisbuch, 1989]. Le modèle des automates cellulaires est remarquable par l'écart entre la simplicité de sa définition et la complexité que peuvent atteindre certains comportements macroscopiques (l'évolution dans le temps de l'ensemble des cellules ne se réduit pas, simplement, à la règle locale qui définit le système). À ce titre, il constitue un des modèles standards pour l'étude des systèmes complexes. En effet, il peut entre autres faire émerger de nouveaux phénomènes [Wolfram, 1984].

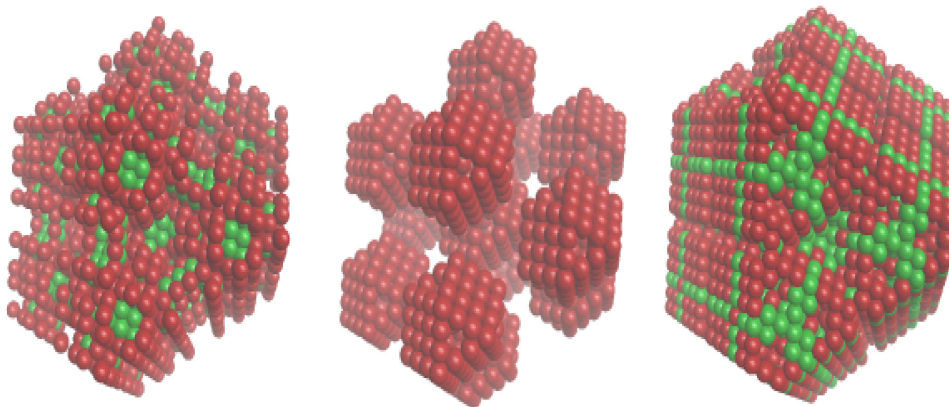


Figure 10: Exemple d'automates cellulaires tridimensionnels [Site Web: *Wikipedia*].

En outre, un réseau d'automates est constitué d'un ensemble d'automates reliés entre eux de telle sorte que les sorties des uns soient les entrées des autres. C'est donc un graphe orienté dont les nœuds sont les automates et les arêtes sont les connexions dirigées de la sortie d'un automate vers l'entrée d'un autre. Le formalisme des réseaux d'automates permet donc une description semi-quantitative des systèmes complexes naturels en utilisant par exemple des automates à plusieurs dimensions (Figure 10).

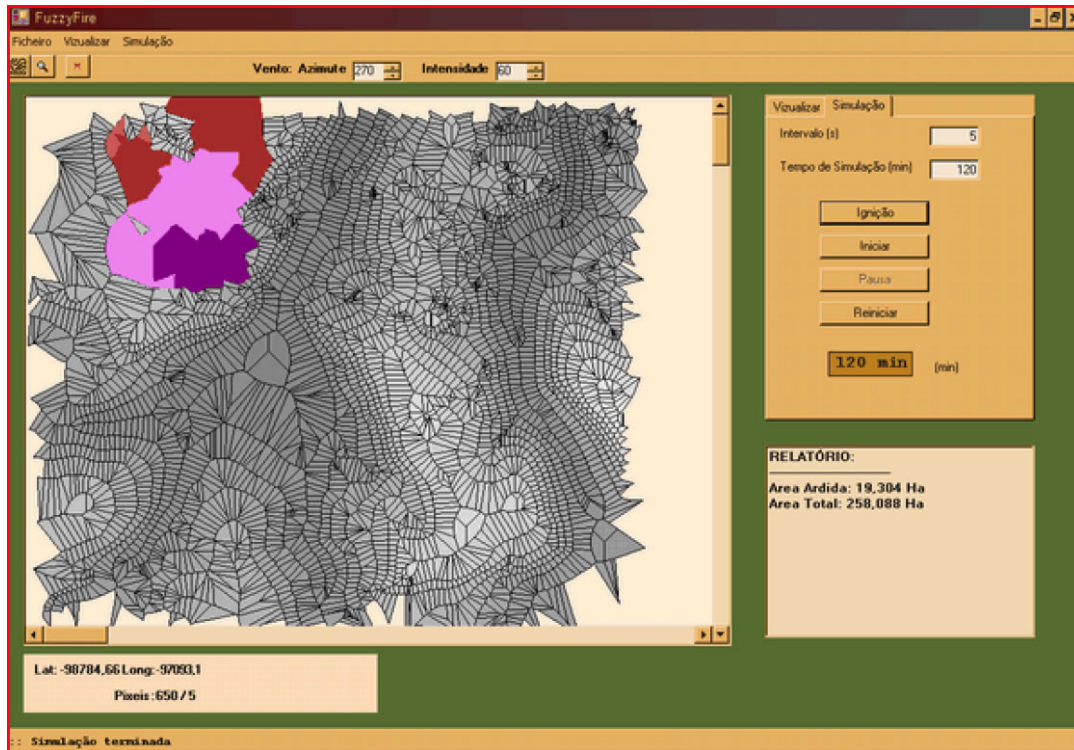


Figure 11: Exemple de simulation de la propagation des feux de forêt en utilisant un diagramme de Voronoi [Carvalho *et al.*, 2006].

Malgré tous les avantages qu'offre l'approche à base d'automates cellulaires, il paraît qu'elle fait face à plusieurs sérieux défis, surtout lorsqu'elle essaye de sortir des expérimentations académiques théoriques vers des applications plus réelles telles que la simulation urbaine [O'Sullivan et Torrens, 2000], [Benenson et Torrens, 2004]. Dans ce cas, nous pouvons d'ailleurs mentionner deux limites à l'utilisation de ces automates cellulaires pour la simulation dans des environnements géographiques. En effet, le découpage de l'espace géographique en un ensemble de cellules de forme régulière est une première limite de ces automates puisque la grille utilisée est considérée très superficielle et simpliste. Cette limite a été contournée par l'implémentation des automates cellulaires sous forme de réseaux irréguliers ou des partitions en utilisant un diagramme de Voronoi [Shi et

Pang, 2000]. Ce diagramme consiste à décomposer l'espace en se basant sur les distances séparant un ensemble discret d'objets ou de points bien particuliers appartenant à cet espace. Contrairement au découpage cellulaire, cette nouvelle décomposition permet d'avoir des relations de voisinage plus variées. C'est le cas par exemple lors de son utilisation pour simuler la propagation des feux de forêt (Figure 11) [Carvalho *et al.*, 2006]. La deuxième limite des automates cellulaires est qu'ils ne permettent pas de gérer les objets mobiles de l'environnement. Cette contrainte semble être majeure, étant donné que l'aspect dynamique des acteurs impliqués dans le phénomène qu'on veut simuler peut être très important.

3.3 Systèmes multi-agents

Les systèmes multi-agents (SMA) sont apparus à la fin des années 1980 et ils ont pris depuis lors une place de plus en plus importante en informatique et spécialement en intelligence artificielle distribuée. En effet, les problèmes complexes sont souvent physiquement distribués (p.ex. un réseau de transport dans lequel chaque voiture ou chaque avion est une entité autonome qui doit parvenir à sa destination). Nous pouvons d'ailleurs définir un système multi-agent comme étant un système informatique composé d'un nombre souvent assez élevé d'agents autonomes qui interagissent en vue de coopérer, cohabiter ou même compétitionner [Ferber, 1995], [Site Web: MAS].

En outre, un agent est une entité agissant dans un environnement, et elle est capable généralement de le percevoir, d'agir sur celui-ci et d'interagir avec les différents composants l'entourant. Cependant, on distingue deux catégories d'agents: (1) Les agents réactifs qui sont basés sur le principe de stimulus / réponse. Ils se contentent donc de réagir à des changements qui se produisent dans le système, et ceci par le déclenchement de règles préétablies. (2) Les agents cognitifs, par contre, possèdent une plus grande capacité de raisonnement, de traitement d'information et d'interaction. Ils sont caractérisés par certaines fonctions telles que la perception des signaux de l'environnement, ainsi que la possibilité de les mémoriser, la communication avec les autres agents, le contrôle de leurs actions et réactions, la coopération et également l'apprentissage [Wooldridge, 2002].

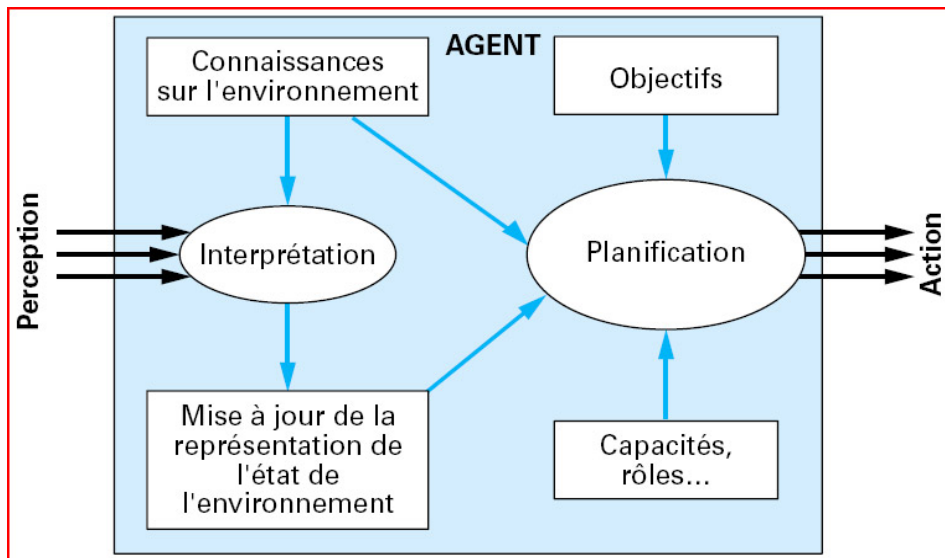


Figure 12: Exemple d'agent à comportement de type cognitif [Mandiau *et al.*, 2002].

Par ailleurs, les comportements de l'agent sont de trois ordres qui répondent aux trois phases générales de réalisation d'une tâche: (1) Une phase de perception (l'agent perçoit son environnement). (2) Une phase de cognition (l'agent détermine ce qui est à faire et décide quand et comment le faire). (3) Une phase d'action (la réalisation effective des actions qui ont été décidées) (Figure 12).

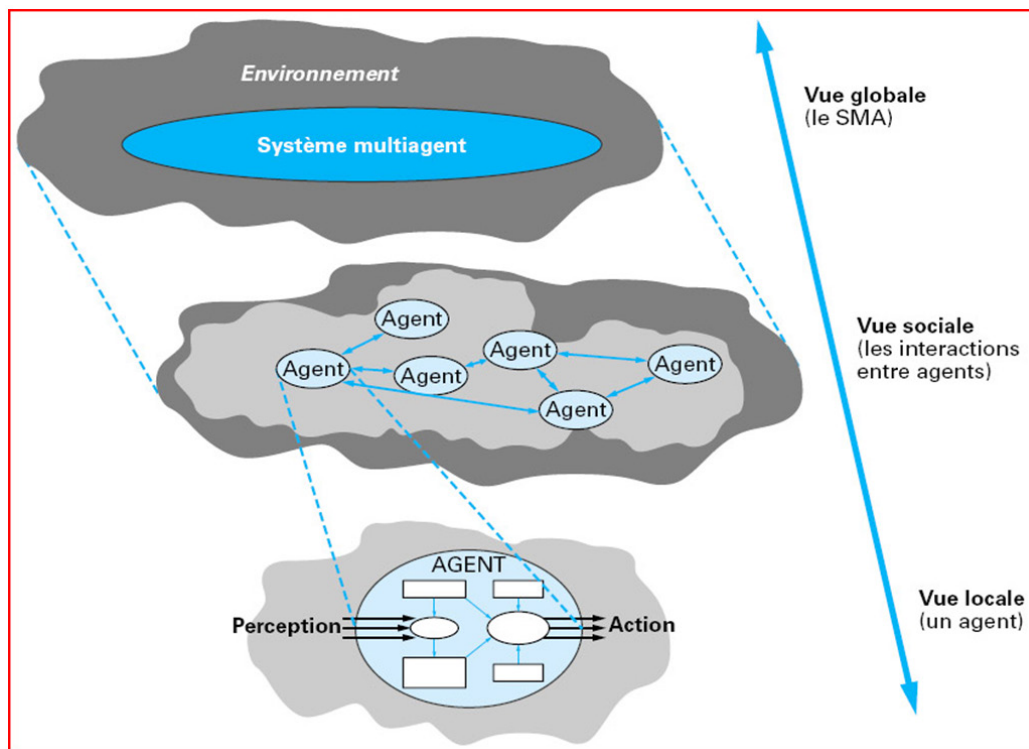


Figure 13: Système multi-agent selon différents niveaux de détail [Mandiau *et al.*, 2002].

Cette description d'un agent donne une vision locale du système. Par contre, la description du système multi-agent donne une vue globale et externe sur le système. Pour obtenir une vision complète, il faut ajouter la description intermédiaire qui concerne l'aspect « social » qui est interne au système et qui représente les interactions entre les agents. Les comportements individuels interfèrent et composent le comportement global du système par l'intermédiaire de cette couche d'interactions. De plus, le comportement global influence les comportements individuels via cette couche sociale (Figure 13) [Mandiau *et al.*, 2002].

De ce fait, l'avantage des SMA par rapport aux automates cellulaires est de prendre en compte explicitement un espace différencié et hiérarchisé, ainsi que les trajectoires de chaque agent ou groupe d'agents. Ils permettent de modéliser le comportement d'un ensemble d'entités plus ou moins organisé, selon un ensemble de lois ou de règles. Ces systèmes facilitent l'utilisation conjointe de règles qualitatives, d'équations quantitatives et de propriétés inhérentes aux configurations spatiales. De plus, les agents sont plus libres de naviguer et d'explorer leur environnement spatial, ce qui est impossible pour les machines d'état individuelles des automates cellulaires. Ceci est dû au fait que le comportement spatial des agents n'est pas contraint à une grille de cellules et que les interactions peuvent avoir lieu au-delà d'un voisinage immédiat. Le potentiel d'interaction est ainsi beaucoup plus grand avec les SMA qu'avec les automates cellulaires [Badariotti et Weber, 2002].

3.4 Systèmes hybrides

Comme nous venons de voir, plusieurs approches ont vu le jour pour simuler certains phénomènes complexes. Chacune de ces approches présente des avantages et des inconvénients. Plusieurs auteurs [Dong-Hwan et Jae-Ho, 1997], [Scholl, 2001], [Schieritz et Größler, 2003] ont justement pensé au couplage de certaines de ces approches pour créer une synergie entre elles, mais surtout pour combiner leurs forces afin d'arriver à produire un meilleur résultat lors de la simulation. D'ailleurs, le fait de pouvoir représenter une réalité la plus «complète» possible face à des besoins difficiles à cerner ne peut s'établir que dans la recherche d'une complémentarité des méthodes de conception autour d'un objectif commun [Schieritz et Milling, 2003].

Par ailleurs, la combinaison des automates cellulaires et des systèmes multi-agents est une technique qui est utilisée dans plusieurs travaux. Nous pouvons citer par exemple les travaux de Badariotti et Weber [Badariotti et Weber, 2002] qui ont modélisé la mobilité résidentielle et ses effets sur la structure urbaine de la ville de Bogota. En effet, un automate cellulaire est utilisé pour représenter les stocks de logements, à pas de temps non uniforme. Les zones où l'état d'occupation du stock de logements est inférieur à un moment donné à son optimum agissent comme des attracteurs. Les groupes d'individus ayant un comportement homogène et issus des cellules denses s'adaptent à cette situation par un mouvement de mobilité qui se reporte sur les zones attractrices. Le déplacement s'établit par contiguïté (voisinage). D'un autre côté, un modèle multi-agent introduit les individus (ou groupe d'individus) caractérisés par des comportements particuliers et des critères de choix résidentiels différents [Badariotti et Weber, 2002].

En outre, un autre exemple des systèmes hybrides est l'intégration d'un système comportant des évolutions continues et des phénomènes discrets qui leur sont liés. Il n'existe pas pour l'instant de théorie globale pour l'étude de ces systèmes, mais plutôt des approches basées sur l'extension de méthodes classiques issues des systèmes continus ou discrets, en vue de couvrir une gamme plus étendue d'applications [Branicky et Mattson, 1997]. Dans cet ordre d'idée, certains auteurs ont essayé de comparer la DS et les SMA dans le but de faire bénéficier mutuellement les avantages de ces deux approches l'une par rapport à l'autre [Stemate et Taylor, 2007]. Nous pouvons d'ailleurs citer un exemple d'intégration qui nous semble intéressant. Il s'agit des travaux de Schieritz et Größler [Schieritz et Größler, 2003] qui ont présenté une méthode utilisant les concepts de modélisation de la DS et des SMA afin de réduire la complexité dans la gestion de la chaîne logistique (c.-à-d. *supply chain management*). En effet, dans une entreprise de production, le temps de réalisation d'un produit est fortement conditionné par l'approvisionnement en matières premières, en éléments d'assemblage ou en pièces détachées à tous les niveaux de la chaîne de fabrication. On appelle ainsi "chaîne logistique" l'ensemble des maillons relatifs à la logistique d'approvisionnement (achats, approvisionnement, gestion des stocks, transport, manutention, etc.). Cette chaîne a été modélisée à deux niveaux d'agrégation (Figure 14). Le niveau macro montre un réseau d'agents qui sont potentiellement des participants dans la chaîne logistique. Chaque lien entre deux agents peut être interprété comme une relation entre le fournisseur et le client. La structure de la chaîne logistique est

un phénomène émergent résultant des décisions prises par les agents individuels. Au niveau micro, la structure interne de l'agent contient à la fois un module de commande et des modules d'évaluation pour chacun de ses fournisseurs. Ces modules sont représentés en utilisant la DS. Les auteurs ont déduit de leurs travaux que leur approche hybride a permis à la chaîne logistique d'être plus flexible en combinant la DS et les SMA [Schieritz et Größler, 2003].

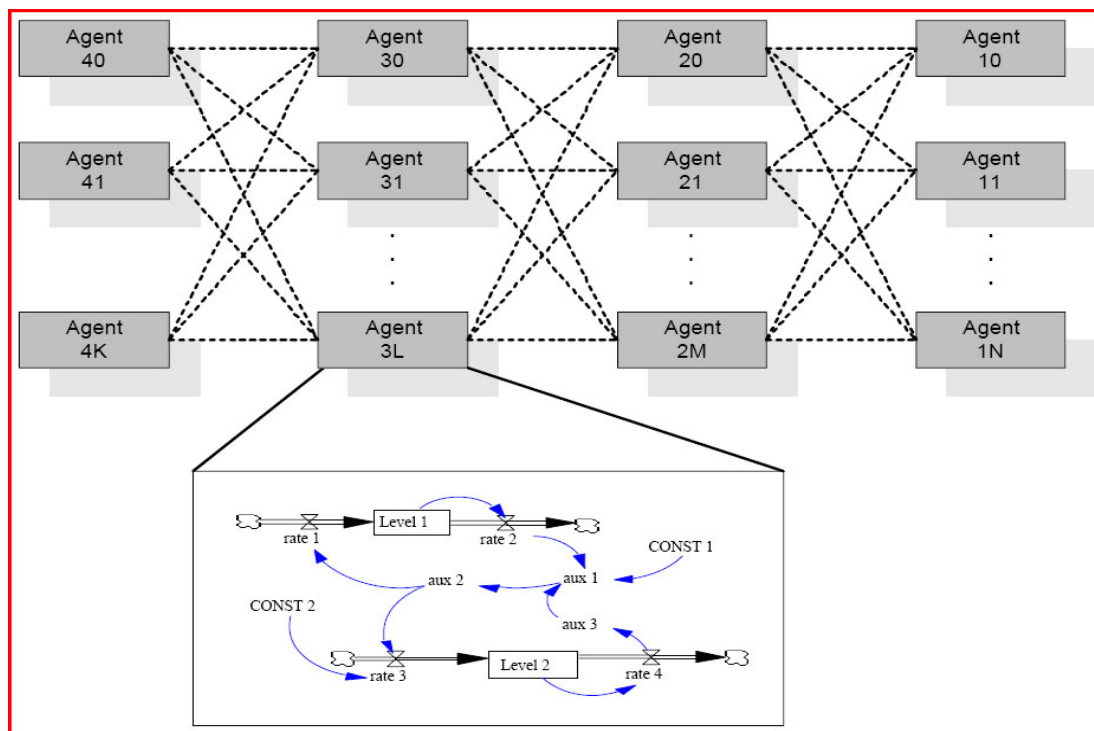


Figure 14: Niveaux macro et micro de la chaîne logistique [Schieritz et Größler, 2003].

Une autre étude intéressante dans laquelle Mabrouk et ses collaborateurs [Mabrouk *et al.*, 2007] ont développé un modèle micro de la dynamique de bactéries afin de mieux comprendre leurs interactions qui leur permettent de construire des structures collectives. Les auteurs ont utilisé des modèles macro et méso déjà existants pour les comparer au modèle micro développé. En effet, les bactéries sont utilisées dans la production de plusieurs produits agro-alimentaires, étant donné qu'elles peuvent dégrader et également synthétiser une large gamme de molécules. Ces bactéries doivent être cultivées dans des bioréacteurs afin de s'assurer d'une fonction métabolique suffisante. Elles vont former des systèmes très complexes dont les propriétés individuelles ainsi que les interactions microscopiques peuvent influencer les dynamiques macroscopiques du bioréacteur. Il faut donc tenir compte du niveau micro afin d'améliorer par exemple le processus de

fermentation de certains produits. Au niveau macro, la dynamique des bactéries est calculée en utilisant des équations différentielles permettant d'évaluer des variables comme la concentration en nutriments et de prédire leur évolution quantitative dans le temps. Au niveau méso, le processus de floculation est pris en considération et surtout la distribution de la taille des flocons de bactéries formées. Les auteurs utilisent un modèle hybride pour le niveau méso. En effet, un modèle discret est utilisé pour les flocons et un modèle continu est utilisé pour les nutriments. Au niveau micro, les cellules des bactéries sont modélisées par un système multi-agent. Cette étude a montré que les résultats du modèle micro sont statistiquement équivalents à ceux fournis par le modèle macro et méso [Mabrouk *et al.*, 2007].

3.5 Géosimulation multi-agent

3.5.1 Généralités

La géosimulation décrit une nouvelle forme de simulation basée sur une combinaison de diverses techniques et de théories (automates cellulaires, système multi-agent, etc.), offrant ainsi une perspective unique qui fait défaut chez la simulation traditionnelle. Par ailleurs, le concept de la géosimulation multi-agent [Moulin *et al.*, 2003] a émergé suite aux lacunes que présentent les systèmes multi-agent classiques. En effet, ces derniers n'utilisent pas un environnement de simulation géoréférencé. Par contre, la géosimulation possède l'avantage de structurer les connaissances spatiales de l'environnement en utilisant les données SIG. De plus, la recherche en géosimulation s'appuie sur plusieurs techniques pour améliorer la technologie de la simulation spatiale. Nous pouvons citer par exemple la dérivation de nouveaux algorithmes pour le calcul spatial, les nouvelles méthodologies pour concevoir des entités spatiales et les relations qui les relient, l'application des modèles de simulation aux problèmes du monde réel et les nouveaux logiciels pour expérimenter avec les systèmes géographiques [Schüle *et al.*, 2004] [Torrens, 2006].

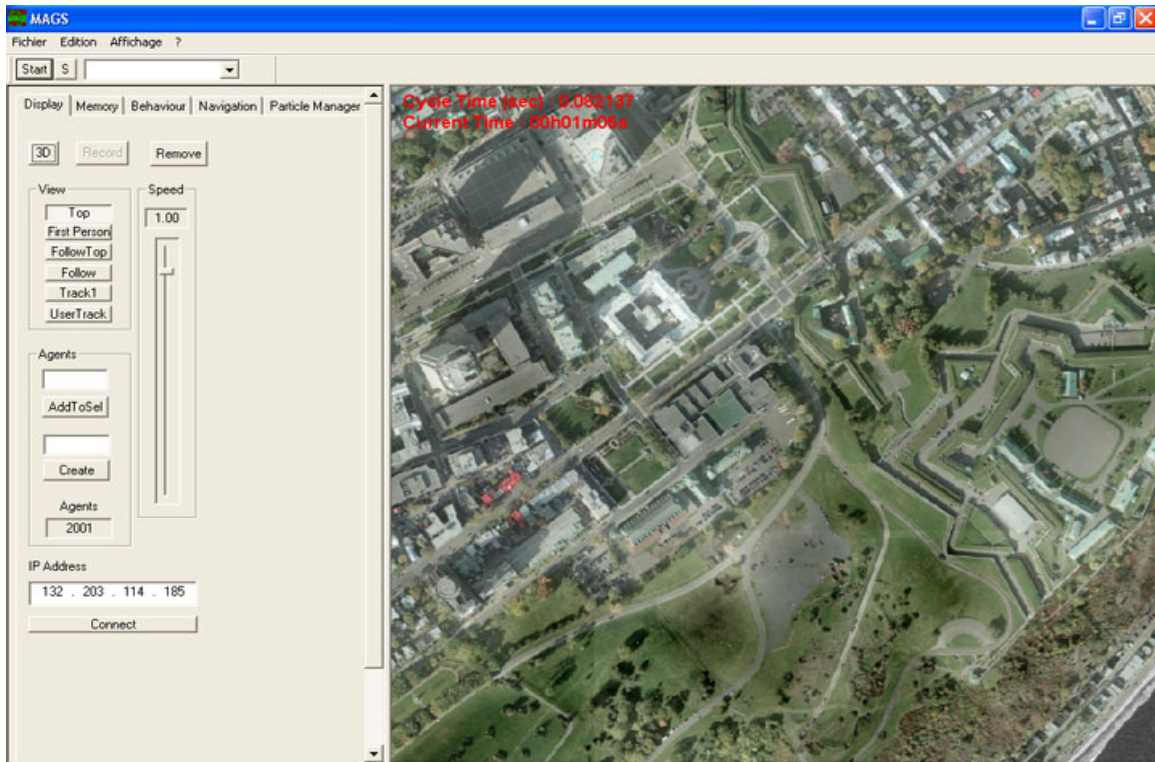
En outre, la géosimulation se distingue des approches traditionnelles de modélisation par le fait qu'elle met l'accent sur le comportement interactif des objets géographiques élémentaires selon une large variété de moyens. Les interactions peuvent être aussi bien des flux que d'autres relations spatiales telles que la diffusion, l'agrégation, l'action à distance, etc. De plus, la géosimulation s'écarte de l'objectif par défaut de la simulation

traditionnelle qui se limite à la prédiction. Elle tend, à travers les projets réalisés à ce jour, vers des simulations permettant d'inclure et de modifier des scénarios événementiels. La géosimulation est devenue donc un outil d'aide à la décision [Moulin *et al.*, 2003] [Benenson et Torrens, 2004]. D'un autre côté, le terme géosimulation a été initialement utilisé pour simuler des phénomènes urbains y compris le comportement des foules. Ce genre de simulation est basée sur l'évolution d'individus et d'autres unités du monde réel, et ce, dans un temps approchant le temps réel [Moulin *et al.*, 2003] [Benenson et Torrens, 2004] [Moulin *et al.*, 2004]. Cependant, la géosimulation a été généralisée pour étudier une multitude de phénomènes dans plusieurs domaines d'application. Nous avons d'ailleurs déjà cité les travaux de [Sahli et Moulin, 2007], [Ali et Moulin, 2007], [Haddad et Moulin, 2008], [Chaker et Moulin, 2008] et [Mekni *et al.*, 2008]. Par ailleurs, quelques plateformes de géosimulation basées sur les agents existent déjà, comme c'est le cas par exemple de *Swarm* [Minar *et al.*, 1996]. Cependant, nous aimerions présenter dans ce qui suit la plateforme de simulation *MAGS*, étant donné que c'est elle que nous avons choisie pour nos premières expérimentations afin de trouver des solutions à notre problématique.

3.5.2 Plateforme de simulation *MAGS*

Le Groupe de Recherche en Informatique Cognitive (GRIC) dirigé par le Dr. Bernard Moulin a été le pionnier dans l'utilisation de l'approche de géosimulation multi-agent. En effet, cette équipe a développé la plate-forme générique de simulation *MAGS* (c.-à-d. *Multi-Agent Geo-Simulation*). Cette plate-forme permet de simuler des milliers d'agents logiciels interagissant dans des environnements géographiques virtuels en 2D ou en 3D (Figure 15). Ces agents sont dotés de capacités spatiales et cognitives telles que la perception, la mémorisation et la navigation. En effet, les agents de *MAGS* sont capables de percevoir plusieurs objets de natures différentes. Un agent peut par exemple percevoir les caractéristiques du terrain de l'environnement de simulation (élevations, pentes, etc.), les autres agents logiciels qui sont dans son champ visuel ou encore les événements spéciaux comme une explosion. Des algorithmes efficaces et optimisés ont été implémentés afin de simuler la perception de milliers d'agents en temps réel [Moulin *et al.*, 2003].

(a)



(b)

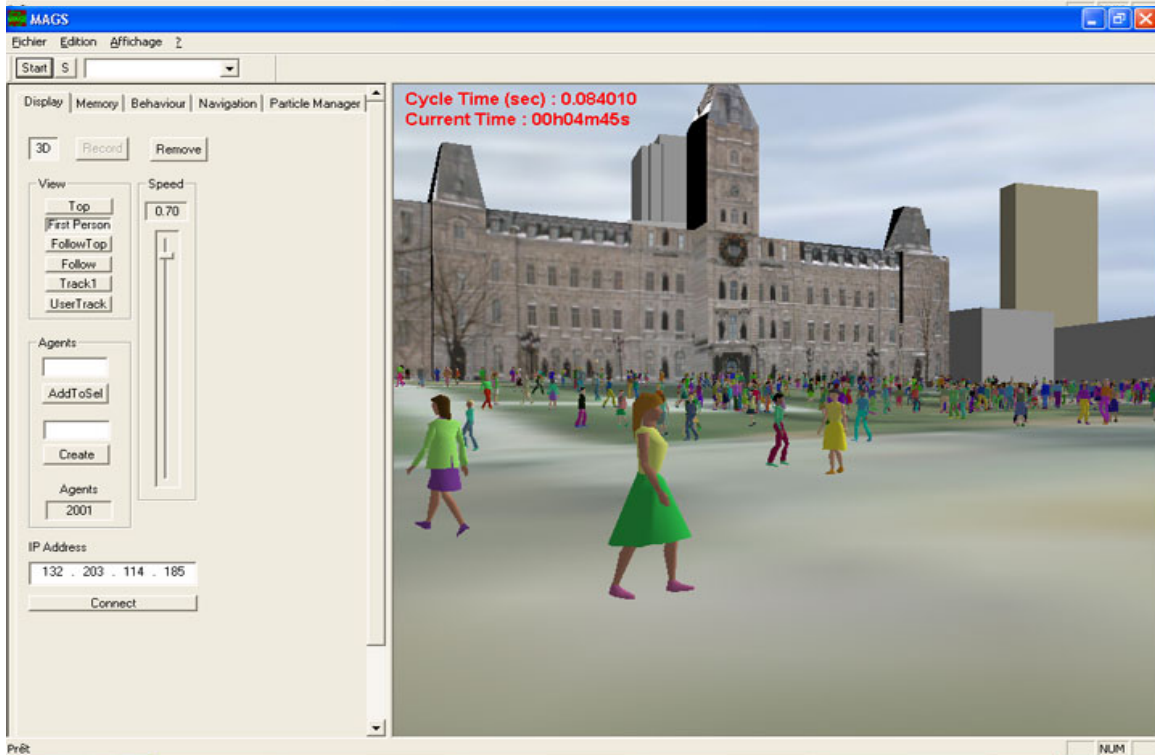


Figure 15 : Le simulateur *MAGS* en 2D (a) et en 3D (b) [Bouden, 2004].

De plus, les agents de *MAGS* se déplacent dans l'environnement de simulation de façon autonome. Ils peuvent suivre des chemins spécifiques comme les routes ou les trottoirs. Ils peuvent également naviguer à travers des espaces ouverts en évitant les autres agents ou les

objets de l'environnement. D'un autre côté, les agents sont capables d'interagir avec l'environnement de simulation et prendre des décisions qui respectent leurs états internes et leurs préférences. Ces agents ont également des objectifs à satisfaire. Ces objectifs sont organisés de façons hiérarchiques et sont associés à des actions que l'agent peut accomplir. Un objectif est relié à des règles qui contiennent des contraintes sur les actions et sur l'activation de l'objectif [Moulin *et al.*, 2003].

Par ailleurs, la plateforme de simulation *MAGS* comprend plusieurs modules lui permettant d'effectuer les tâches demandées (Figure 16). Il comprend un module pour l'environnement, un module pour les agents, une interface (*User Interface*), un contrôleur (*Processes Module*), un moteur d'affichage (*Display module*), un module comportemental (*Behaviour Module*), un module d'états (*States Module*) et un module de communication réseau (*Socket Controller Module*).

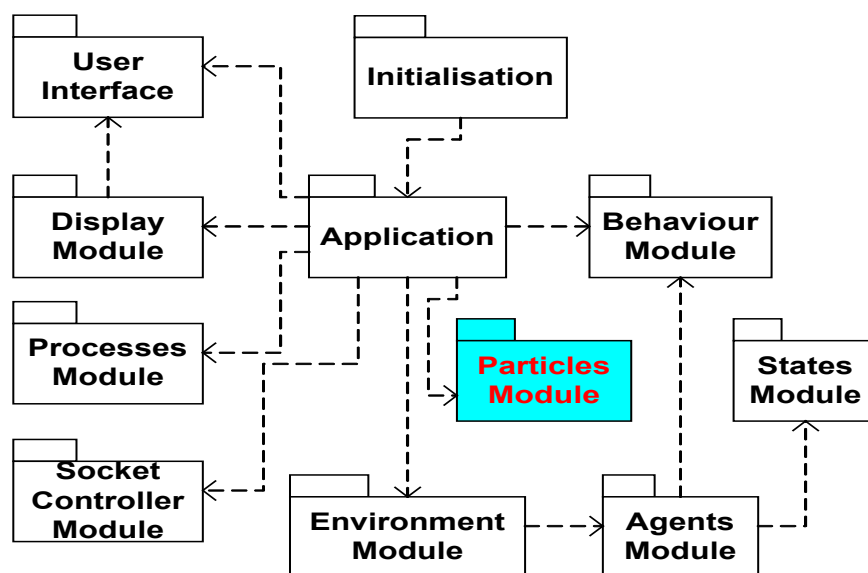
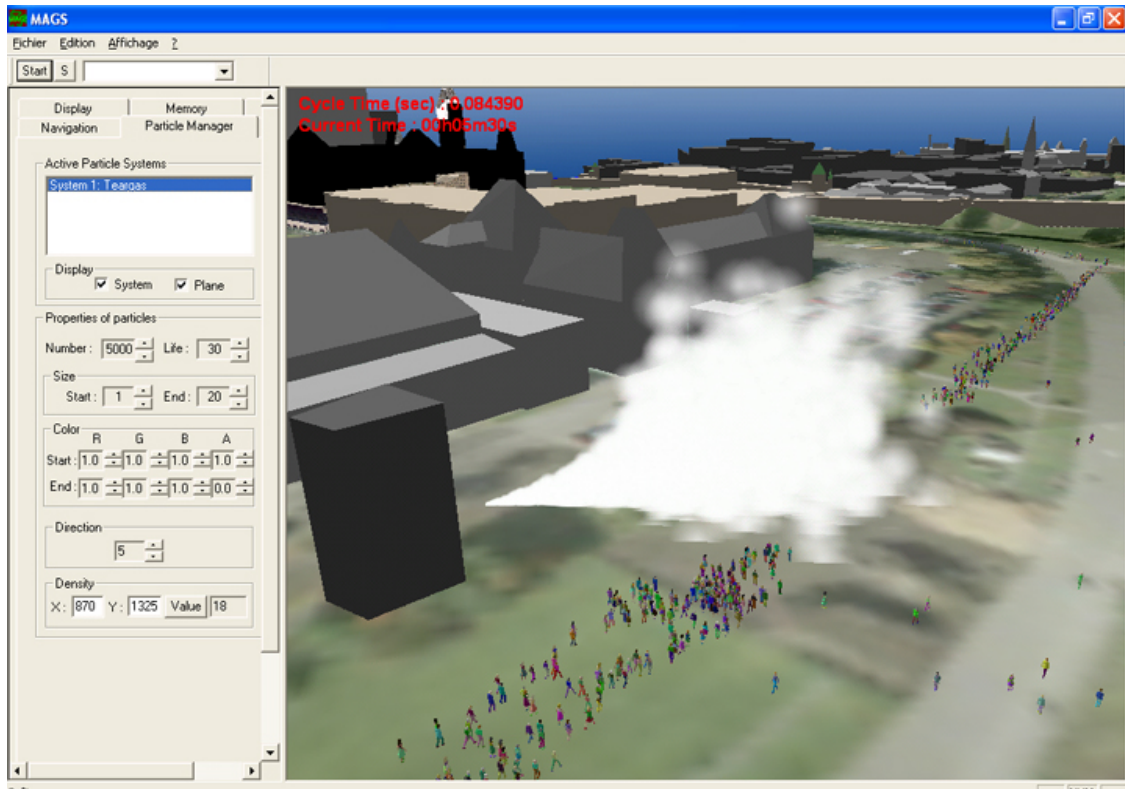


Figure 16 : Les différentes composantes du système *MAGS*.

En plus de ces modules, *MAGS* renferme un module de particules (*Particles Module*) qui a été ajouté dans le cadre de notre travail de maîtrise [Bouden, 2004] afin de simuler des formes irrégulières comme la fumée ou un gaz. En effet, nous avons conçu un gestionnaire de systèmes de particules qui nous a permis de faire réagir les agents, présents dans le monde virtuel, à un gaz ou à la fumée qui se répand dans l'espace. De plus, nous avons pu simuler le vent et son influence sur ces phénomènes (Figure 17).

(a)



(b)

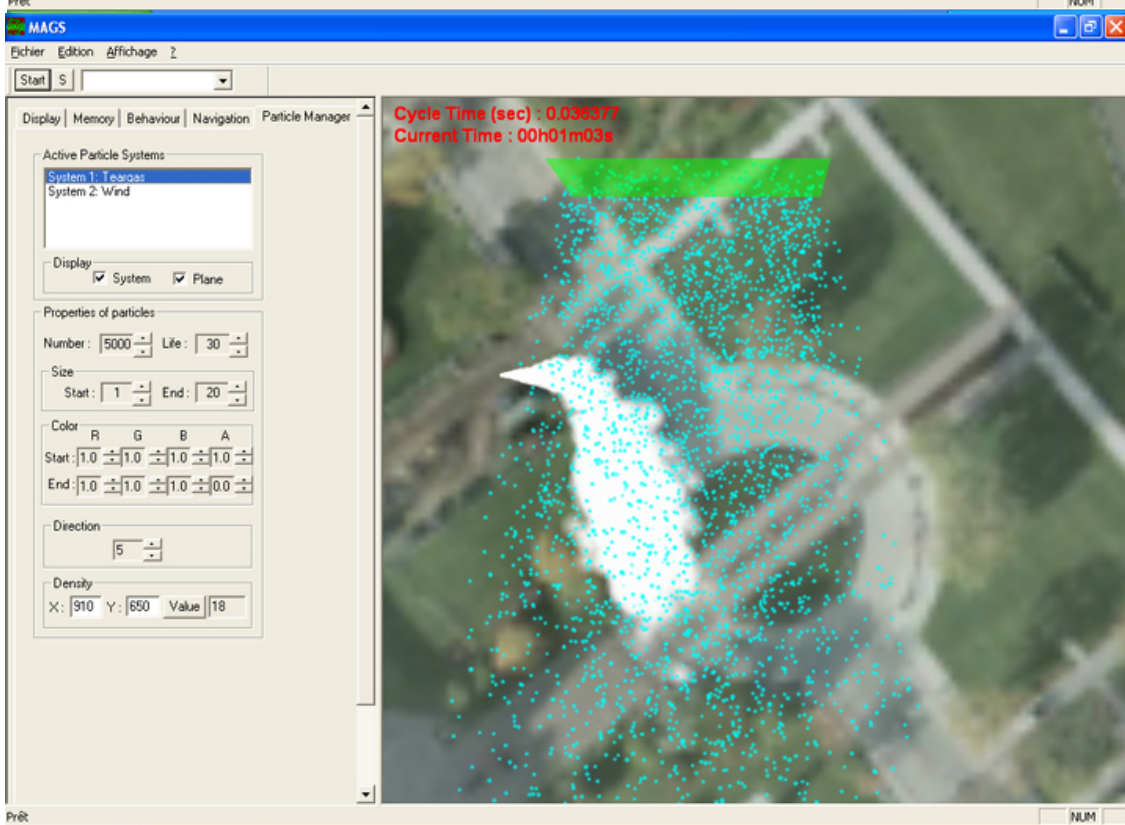


Figure 17 : Simulation de la propagation d'un gaz (a) et du vent (b) [Bouden, 2004].

Un système de particules, tel que décrit par Reeves [Reeves, 1983], représente un objet quelconque sous la forme d'un ensemble de particules en évolution dans un espace continu. C'est une méthode de modélisation qui utilise des entités prédéfinies (particules) comme primitives. L'intérêt principal du système de particules est sa grande flexibilité et la facilité avec laquelle on peut modéliser et animer des phénomènes complexes. En effet, la simplicité d'une particule nous permet d'en utiliser de grandes quantités, ce qui donne une impression visuelle de complexité. D'ailleurs, les systèmes de particules ont été initialement conçus pour simuler dynamiquement la morphologie irrégulière d'éléments naturels tels que les nuages, l'eau, le feu ou encore la fumée. Pour cette catégorie d'objets flous (fluides), il n'existe pas de frontière précise entre l'intérieur et l'extérieur de l'objet : il y a diffusion à travers tout l'espace. Il est ainsi impossible de définir, comme à l'ordinaire, leur surface sous la forme d'un maillage de facettes polygonales [Reeves, 1983].

Cependant, le champ d'application des systèmes de particules s'est considérablement étendu, notamment pour simuler le processus d'évacuation d'une foule dans un stade (chaque particule représentant un individu) ou encore pour reproduire le déploiement de coussins pneumatiques de protection (airbags). La modélisation de végétaux et de leur croissance a également été réalisée à l'aide de systèmes où chaque particule représente "une brique de base" de l'organisme vivant [Gareau, 1997]. Par ailleurs, Reynolds [Reynolds, 1987] a modélisé le vol d'oiseaux, les troupeaux d'animaux et les bancs de poissons en utilisant un système de particules dans lequel les particules (les animaux) peuvent connaître la position des particules voisines et modifier leur mouvement afin de rester près d'elles. Une particule a une perception de l'environnement limitée à son entourage immédiat. Elle ne perçoit donc pas tous les membres du troupeau, mais seulement ceux qui se trouvent près d'elle. Les règles d'attroupement, qui peuvent être considérées comme une simulation du comportement animal, sont : éviter une collision avec les voisins; aller à la même vitesse que les voisins; se maintenir au centre des particules voisines. Les particules peuvent être groupées en hiérarchies, de sorte qu'il est possible de diriger le troupeau en entier en modifiant manuellement les paramètres du chef du troupeau [Reynolds, 1987].

C'est pour toutes ces raisons que nous avons pensé utiliser *MAGS*, y compris son module de particules, afin de faire nos premières expérimentations concernant la simulation de la

40

propagation des maladies infectieuses. Nous avons d'ailleurs utilisé ces systèmes de particules pour simuler le comportement des groupes de corneilles impliquées dans la transmission du virus du Nil occidental. Nous allons y revenir dans le chapitre VII à la section concernant le développement du simulateur *WNV-MAGS* (voir Section 2 du Chapitre VII).

3.6 Systèmes multi-niveaux

Les systèmes multi-niveaux (SMN) permettent de simuler le phénomène étudié à plusieurs niveaux de granularité. En effet, la variation du niveau apporte à chaque fois un supplément d'information à la simulation et une vision différente pour interpréter les résultats escomptés et par conséquent, une meilleure compréhension du système étudié. D'ailleurs, la majorité des approches utilisées actuellement ignorent souvent l'aspect multi-niveau des données et ainsi, dans beaucoup de situations, ont pour résultat une grande perte dans leur puissance prédictive ou explicative [An *et al.*, 2005].

Par ailleurs, la notion de granularité définit la taille du plus petit élément d'un système. De ce fait, le système ne peut plus être découpé en éléments plus petits quand on arrive à son niveau de granularité le plus fin. Par exemple, la granularité la plus fine de la matière a longtemps été l'atome. Cependant, la découverte des constituants de l'atome (protons, neutrons et électrons) a changé ce niveau de granularité. D'autre part, cette notion de niveau de granularité est qualitative et elle peut varier en fonction du problème ou du domaine traité, étant donné son caractère relatif [Pereira, 2002] [Burger et Todd, 2006]. Par exemple, une plante est immense si on la compare à un atome d'une de ses cellules végétales, mais elle est ridiculement petite si on la compare à une galaxie.

En outre, la simulation multi-niveau permet de remédier aux limitations de ressources, car les limites de l'ordinateur d'une part et les contraintes financières et temporelles des projets d'autre part, rendent impossible une simulation très détaillée de l'ensemble du phénomène étudié. La simulation multi-niveau offre ainsi une approche évolutive, dans le sens où les modèles peuvent être raffinés et implémentés au fur et à mesure que les données sont disponibles. Par conséquent, elle permet de détailler de plus en plus certains aspects du phénomène étudié en offrant des solutions logicielles (performances des algorithmes de

simulation) et matérielles (rapidité de calcul et mémoire des ordinateurs) qui s'améliorent toujours plus [Chaker et Moulin, 2008].

D'un autre côté, la simulation reproduit, en fonction des attentes de leurs utilisateurs, des phénomènes avec le plus souvent trois niveaux de détail différents [Jakovljevic et Basch, 2004]. Il s'agit d'une macro, méso ou microsimulation. Les modèles macroscopiques représentent les phénomènes avec l'abstraction la plus élevée. Ils sont souvent employés pour évaluer les dynamiques globales des agrégats d'individus situés sur un vaste territoire. Les modèles mésoscopiques sont utilisés pour simuler des groupes d'individus par l'intermédiaire de comportements simples. Pour leur part, les modèles microscopiques sont utilisés pour simuler les individus caractérisés par un comportement relativement complexe et qui ont une capacité d'interaction et de perception de l'environnement. Les travaux de Mabrouk et ses collaborateurs [Mabrouk *et al.*, 2007] expliquent très bien ces notions qui sont appliquées aux populations des bactéries. Par contre, la simulation peut utiliser, dans un cadre restreint, des modèles picoscopiques ou même nanoscopiques permettant de représenter de manière très fine certains phénomènes.

Nous savons également que le niveau de granularité est une notion plutôt vaste [Pereira, 2002]. En effet, elle peut être reliée aux différentes échelles spatiales afin de simuler par exemple un phénomène à des niveaux Macro, Méso et/ou Micro. Cependant, elle peut être également reliée à la variation de l'échelle temporelle (par exemple, la variation du lieu de résidence s'observe sur des années alors que la variation du trafic routier sur un tronçon de route s'observe sur des minutes). D'ailleurs, les termes « grain » et « étendue » ont été souvent employés pour caractériser les différentes échelles spatio-temporelles. Le grain est l'unité fondamentale par laquelle un phénomène est mesuré ou décrit, alors que l'étendue dénote la zone spatiale ou la durée temporelle de ce phénomène [Burger et Todd, 2006]. D'un autre côté, la notion de niveau peut être également reliée aux différentes catégories de regroupements des acteurs du système étudié. Par exemple, les individus, les groupes et les populations ne peuvent pas être modélisés à un même niveau de granularité puisque le choix de ce niveau dépend de la disponibilité des modèles utilisés, ainsi que des données permettant de les alimenter. Ce choix dépend également des besoins de l'utilisateur. Il est donc primordial de bien choisir les différents niveaux de granularité, surtout qu'un mauvais

choix pourrait rendre le passage d'un niveau à un autre plus ou moins ardu lors de la visualisation des résultats ou encore de l'analyse du phénomène étudié.

Nous présentons d'ailleurs dans cette partie quatre projets afin d'illustrer l'utilisation des SMN par des cas pratiques dans des domaines différents. En effet, An et ses collaborateurs [An *et al.*, 2005] ont intégré des échelles spatio-temporelles dans un modèle multi-agent afin de simuler l'impact de l'expansion de la population humaine rurale sur les forêts ainsi que sur les habitats des pandas géants (espèce en voie de disparition). Les auteurs se sont intéressés au contexte de la réserve naturelle *Wolong* de Chine dans laquelle habitent des paysans qui utilisent le bois comme source d'énergie. Cette consommation du bois a entraîné une dégradation notable des habitats des pandas suite au déboisement des forêts. De ce fait, ces auteurs ont pu développer un modèle conceptuel nommé *IMSHED* (*Integrative Model for Simulating Household and Ecosystem Dynamics*) qui utilise plusieurs facteurs socio-économiques et démographiques afin de suivre l'évolution dynamique des foyers ruraux. Ce modèle renferme trois composantes : (1) le développement des foyers, (2) la demande en bois de chauffage et (3) la disponibilité et la collecte de ce même bois de chauffage (Figure 18).

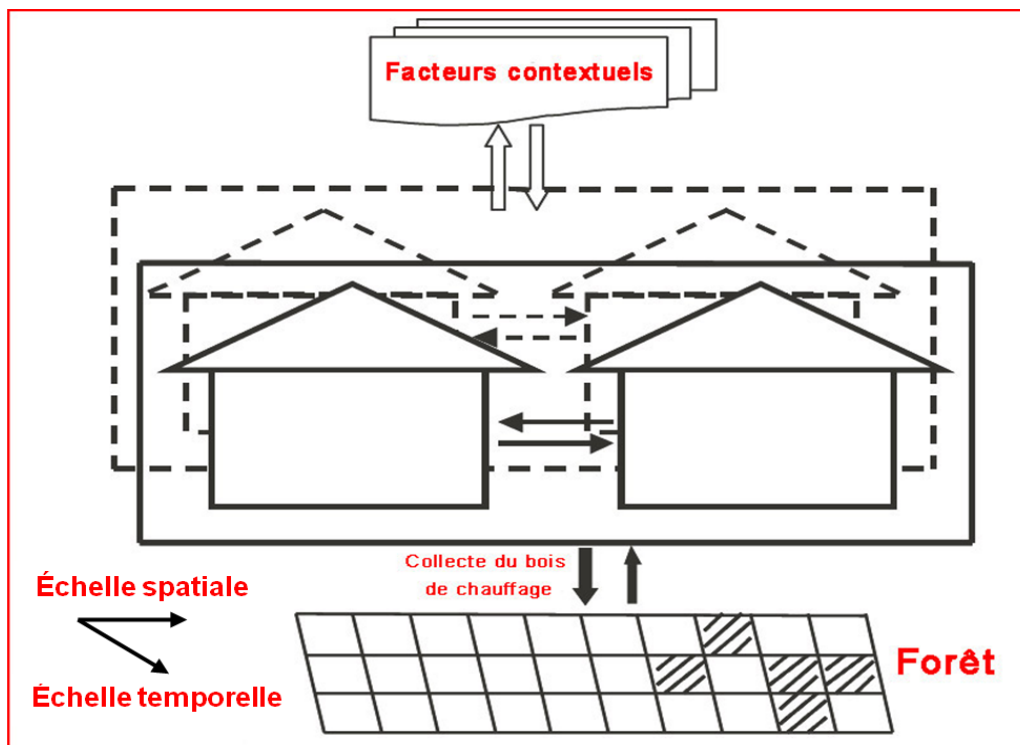


Figure 18: Le modèle conceptuel *IMSHED* [An *et al.*, 2005].

Ainsi, la Figure 18 montre les différents éléments de ce modèle. En effet, les maisons avec des lignes en pointillés représentent les foyers de la réserve pendant une période ou un temps déterminé. Par contre, les maisons avec des lignes en continu représentent les foyers dans un même espace, mais à une période de temps ultérieure. Pendant cet intervalle de temps, les caractéristiques des foyers peuvent changer (natalité, décès, délocalisation, etc.). De plus et à chaque unité de temps, la demande en bois de chauffage est modélisée en fonction par exemple de la taille des foyers. D'un autre côté, l'interaction entre les paysans et la forêt est mise en évidence par la collecte du bois de chauffage. En effet, un collecteur appartenant à un foyer possède un certain besoin en bois de chauffage. Il doit se déplacer à l'endroit le plus proche (zones hachurées dans une grille qui est modélisée par un automate cellulaire) afin de couper du bois dans la forêt. L'augmentation de la distance parcourue peut faire diminuer l'offre en bois étant donné qu'il serait de plus en plus difficile de le ramasser. De ce fait, les foyers peuvent songer à l'électricité comme substitut énergétique. Des facteurs physiques et sociaux, nommés des facteurs contextuels, incluant justement la distance, l'altitude et les décisions politiques peuvent exercer un impact sur plusieurs processus comme ceux concernant la demande en bois de chauffage et en électricité [An *et al.*, 2005].

Le deuxième exemple touche directement nos propres travaux de recherche et il s'agit du projet *MUSCAMAGS* (*Multi-Scale Multi-Agent Geo-Simulation*) [Site Web: *MUSCAMAGS*]. En effet, l'objectif principal de ce projet est de développer des méthodes et des plateformes logicielles permettant de développer des geosimulations multi-agent et multi-niveau pour supporter l'aide à la décision dans des systèmes comportant des situations spatiales dynamiques faisant intervenir une multitude d'acteurs à plusieurs niveaux de granularité (*MADSS* : *Multi-Actor Dynamic Spatial Situations*). Les *MADSSs* impliquent un grand nombre d'acteurs de différents types (humains, animaux, etc), agissant dans des espaces géographiques de diverses étendues [Moulin, 2012]. Certaines *MADSSs* se produisent sur une base régulière (par exemple le trafic routier dans une zone urbaine), tandis que d'autres peuvent évoluer rapidement comme une conséquence de certains événements et/ou un changement de comportements individuels (souvent dans des situations de crise). Pour chaque *MADSS* particulière, les décideurs ont besoin d'obtenir une compréhension globale de la situation, de suivre son évolution, d'élaborer des stratégies et

des tactiques appropriées pour l'intervention, de développer et de comparer des scénarios alternatifs d'intervention et d'anticiper les conséquences de l'application de tels scénarios. Certaines *MADSSs* peuvent se produire dans le cadre d'une autre *MADSS* à un niveau de granularité plus élevé. Par exemple, une manifestation de milliers de personnes dans le centre d'une ville aura une influence sur une *MADSS* plus globale représentant la circulation automobile d'une région faisant partie de la même ville. En raison de cette intégration de *MADSSs*, les décideurs ont souvent besoin d'examiner les diverses situations simultanément à différents niveaux de granularité (spatiaux, temporels, catégories des acteurs). Cette question est importante puisque les phénomènes observés peuvent être différents d'un niveau de détail à l'autre. De plus, ces phénomènes peuvent interférer s'ils appartiennent à des *MADSSs* inter-reliés.

En outre et dans le cadre de ce projet, Chaker et Moulin [Chaker et Moulin, 2008] ont travaillé sur la modélisation multi-échelle d'environnements urbains virtuels. Ils visent à simuler des déplacements de personnes à différents niveaux de détail. En effet, la mobilité de chaque individu ou d'un groupe d'individus dépend de plusieurs contraintes dont certaines sont reliées aux déplacements d'autres individus. Le trafic routier, par exemple, est la conséquence de la mobilité de plusieurs personnes. Un excès du trafic peut entraîner des modifications dans les habitudes de déplacements d'autres personnes, soit en évitant les heures de pointe, soit en empruntant d'autres chemins, ou encore en changeant de mode de transport. De plus, dans une activité de déplacement on tient compte de la personne qui se déplace (composante population), des lieux d'origine et de destination du déplacement (composante lieux d'activités) et du moyen de déplacement (composante réseaux de transport). Afin de simuler ces phénomènes, les auteurs ont commencé par modéliser l'infrastructure de transport (dimensions lieux et réseaux) par la création d'un environnement multi-échelle basé sur une structuration sémantique et topologique. Ils ont également créé une population synthétique afin de l'affecter à l'environnement multi-échelle (dimension population). Cette population permet d'estimer la composition des ménages en individus, leurs caractéristiques et leurs répartitions dans l'espace [Chaker et Moulin, 2008].

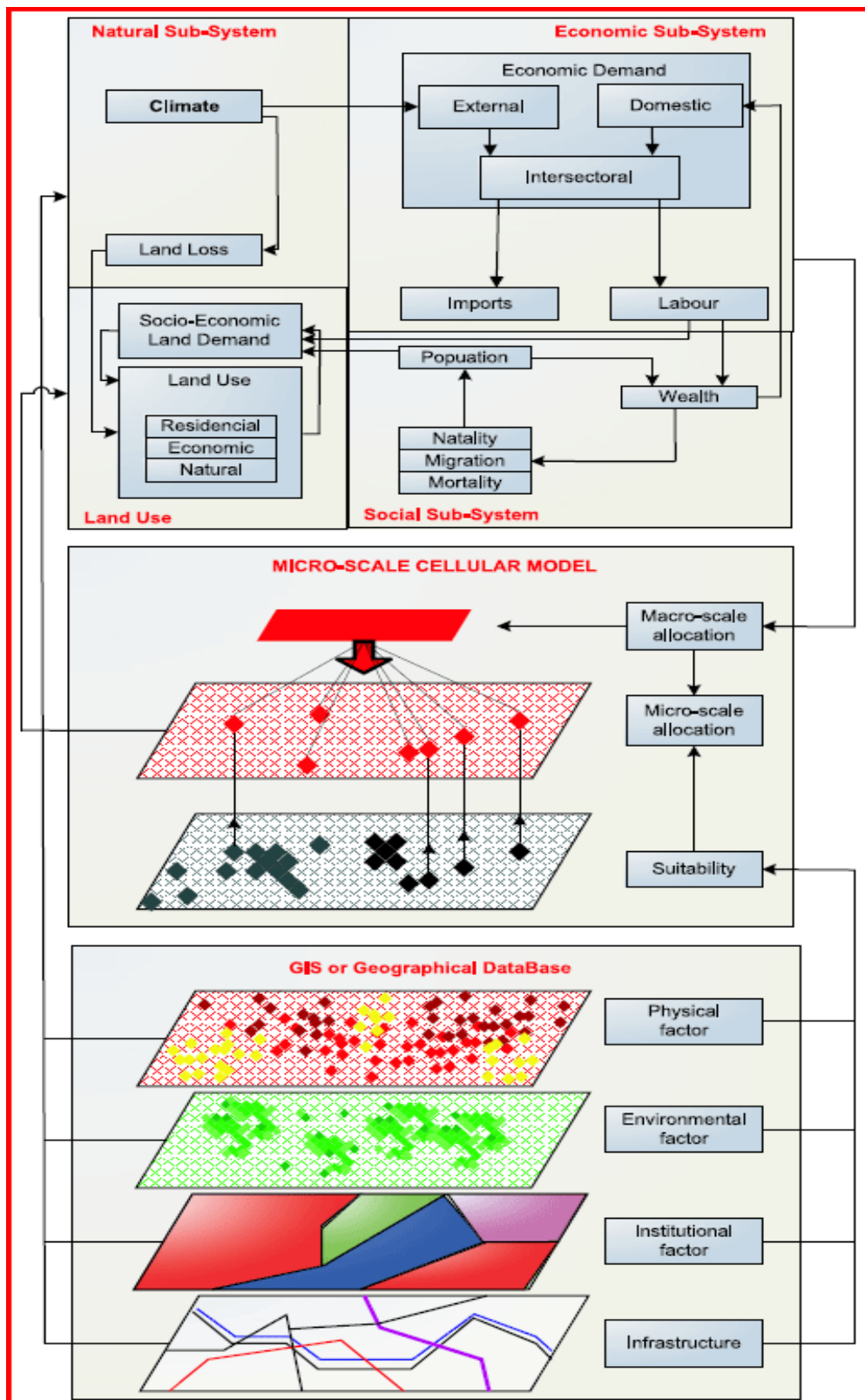


Figure 19: Approche de simulation multi-niveau [White *et al.*, 2000].

D'un autre côté, White et ses collaborateurs [White *et al.*, 2000] ont élaboré un modèle multi-niveau pour simuler l'impact des changements climatiques sur l'île St Lucia. La Figure 19 montre les différents niveaux utilisés par ces auteurs. En effet, ils ont utilisé un niveau Macro pour modéliser les aspects socio-économiques et naturels de toute l'île. À ce niveau, ils ont fait le lien entre des scénarios de changements climatiques et les changements de certaines variables économiques en rapport entre autres avec le tourisme et l'agriculture, ainsi que les changements au niveau démographique. Les auteurs ont utilisé également un niveau Micro pour la représentation de la couverture du terrain en faisant appel aux automates cellulaires. Finalement, les auteurs ont utilisé un troisième niveau représentant la base de données géographique. Ce niveau comprend toutes les caractéristiques géographiques et alimente le mécanisme de simulation par des informations précises issues de divers niveaux physique, environnemental, institutionnel et infrastructurel [White *et al.*, 2000].

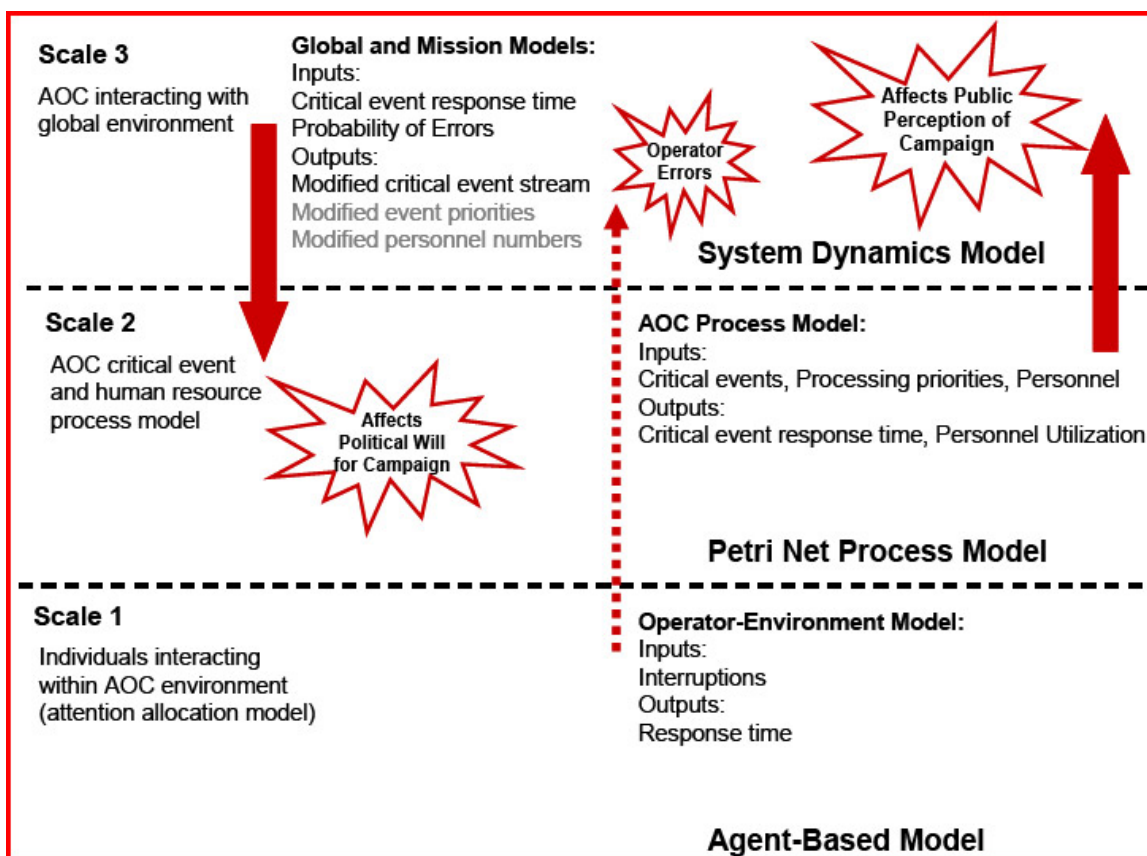


Figure 20: Modélisation à 3 niveaux d'abstraction [Mathieu *et al.*, 2007].

Nous citons finalement les travaux de Mathieu et ses collaborateurs [Mathieu *et al.*, 2007] qui ont modélisé les interactions qui peuvent se produire dans les processus du centre des opérations de l'air et de l'espace (*Air and Space Operations Center: AOC*). Les auteurs utilisent ce qu'ils appellent des "échelles" afin de modéliser ces interactions à trois niveaux d'abstraction. En effet, l'interaction entre l'opérateur avec l'interface de l'ordinateur a été modélisée en utilisant un modèle à base d'agent. L'analyse des cibles durant un combat a été modélisée en utilisant un réseau de Petri. Finalement, les objectifs, stratégies et processus de la mission incluant les interventions possibles de l'ennemi ont été modélisés en utilisant la dynamique des systèmes. La Figure 20 montre les échanges qui peuvent se produire entre les différents modèles. Les auteurs mentionnent que cette approche multi-modèle à plusieurs niveaux d'abstraction est très efficace pour simuler un système aussi complexe faisant intervenir des acteurs et des entités de différents types [Mathieu *et al.*, 2007].

3.7 Systèmes holoniques

L'approche holonique introduite par Koestler en 1967 [Koestler, 1967] est une théorie prometteuse permettant de spécifier un système complexe. En effet, on peut l'appliquer aux systèmes multi-agents, car elle se caractérise par la prise en compte des notions de hiérarchies et de structures ordonnées. Le terme "Holon" provient du mot grec "Holos" qui signifie « ensemble » et le suffixe «-on» qui désigne «une partie». Un holon peut être vu comme une structure fractale qui est stable, cohérente et qui se compose de plusieurs holons en tant que sous-structures d'elle-même et faisant partie d'un autre ensemble plus grand. Un exemple de holon est l'être humain, qui comprend des organes et qui sont à leur tour composés de cellules qui peuvent être décomposées en organites et ainsi de suite. En outre, l'être humain fait partie d'une famille et d'une société. Aucun de ces composants ne peut être entièrement compris sans ses sous-composantes ou sans l'élément se trouvant à une hiérarchie supérieure [Shillo et Fischer, 2003].

Par ailleurs, la plupart des SMA actuels considèrent les agents comme des entités que l'on voit à un niveau micro. Cependant, un groupe, une organisation ou une population d'agents, à un certain niveau d'abstraction, peut se comporter comme une entité à part entière. La

façon de représenter ce genre d'entités demeure ouverte dans le domaine des systèmes multi-agents. De nombreuses approches ont tenté de modéliser cette idée d'agents composés d'agents par des agents collectifs ou encore des méta-agents. Cependant, le fait de considérer ce type d'agents comme un holon est une approche qui n'a pas cessé d'évoluer, surtout pour modéliser les systèmes complexes. Elle a été utilisée, dans les années 90, dans le domaine des systèmes manufacturiers comme un nouveau paradigme de modélisation. Il a depuis fait l'objet de beaucoup d'attention et de nombreux modèles sont apparus. Cependant, c'est Gerber et ses collaborateurs [Gerber *et al.*, 1999] qui ont introduit la notion de holon dans les systèmes multi-agents où ils ont proposé une vision des holons correspondant davantage à la notion d'agent récursif ou composé. Plus récemment, d'autres chercheurs comme Rodriguez et ses collaborateurs [Rodriguez *et al.*, 2006] ont développé des environnements de simulation basés sur les systèmes multi-agents holoniques. En effet, chaque holon modélise une composante de l'environnement qui elle-même peut être décomposée en sous-entités. De ce fait, l'environnement définit une holarchie¹¹ et les agents appartenant à un holon dans cette holarchie peuvent satisfaire plus facilement leurs besoins.

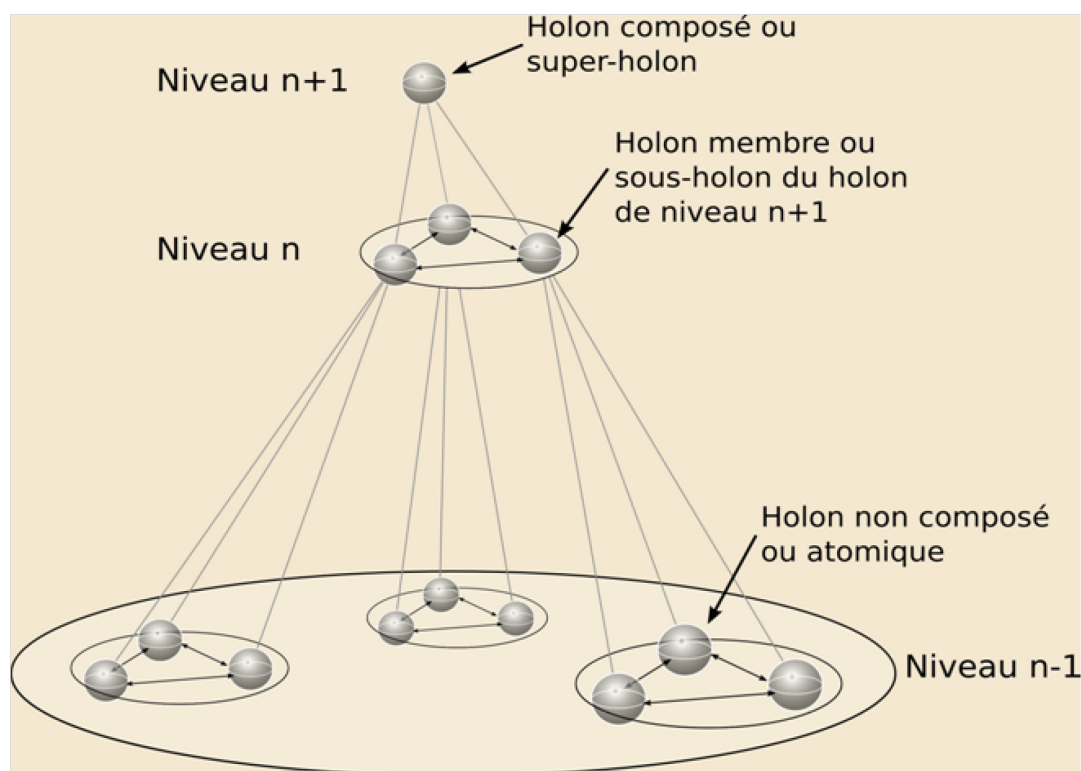


Figure 21: La notion de Holon dans un système multi-agent [Cossentino *et al.*, 2007].

¹¹ La structure hiérarchique composée de holons est appelée holarchie [Rodriguez *et al.*, 2006]

La Figure 21 résume la définition de la notion de holon telle qu'elle est considérée dans les SMA. Les holons de niveau n sont regroupés en organisations qui au niveau $n+1$ peuvent être considérées comme des entités individuelles. Inversement, des holons considérés comme des entités individuelles de niveau $n + 1$ peuvent être considérés comme des organisations au niveau n . Ce processus de composition hiérarchique peut être réitéré sur un nombre arbitraire de niveaux [Cossentino *et al.*, 2007].

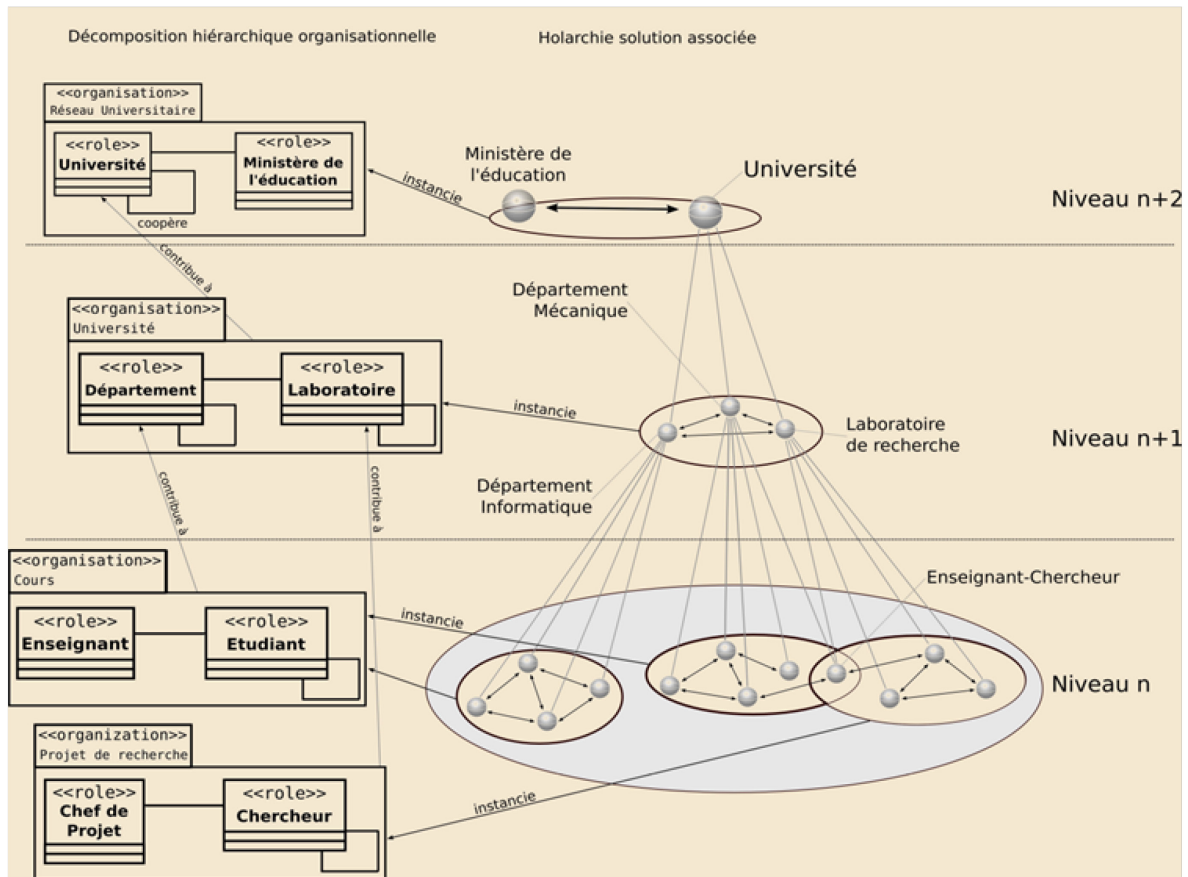


Figure 22: Exemple de la structure holonique d'une université [Cossentino *et al.*, 2007].

Un holon peut être vu, en fonction du niveau d'observation, tantôt comme une entité atomique, tantôt comme un groupe de holons en interaction. À un niveau d'observation donné, le holon composé est qualifié de super-holon. Les holons qui composent un super-holon sont appelés sous-holons ou holons membres. Si l'on considère une université d'un point de vue holonique, on peut la décomposer en différents départements et laboratoires. De la même manière, un département peut être décomposé en un ensemble d'organisations (conseil de département, cours, etc.) peuplées d'enseignants et d'étudiants. Cette holarchie

de l'université est décrite dans la Figure 22. Au niveau le plus haut (niveau $n + 2$) se trouve le holon université, qui dans cet exemple est composé de trois sous-holons: le département informatique, le département mécanique et un laboratoire de recherche (niveau $n + 1$). Au niveau le plus bas se trouvent les enseignants, les étudiants et les chercheurs (niveau n).

4 Conclusion

Dans ce chapitre, nous avons présenté les systèmes complexes d'une façon générale afin de comprendre leur nature et les conséquences de leur complexité. Nous avons également donné un aperçu sur différentes approches de modélisation et de simulation de ces systèmes complexes. Nous nous sommes surtout attardés sur celles qui nous semblent importantes afin d'atteindre nos objectifs de recherche. D'ailleurs, nous aimerions conclure ce chapitre par une constatation issue de notre recherche bibliographique. En effet, il semble que l'approche de géosimulation combinant à la fois tous ses avantages (voir Section 3.5) avec celles d'un système multi-niveau (voir Section 3.6) permettrait à remédier aux lacunes des autres approches actuellement qui sont utilisées pour simuler certains phénomènes complexes.

Afin d'illustrer nos idées de recherche, nous avons sélectionné un domaine d'application où l'utilisation des méthodes et technologies abordées dans ce chapitre est possible et surtout pertinente. Dans le chapitre suivant, nous présentons les maladies infectieuses ainsi que les différents aspects qui nous intéressent et qui sont en rapport avec la modélisation et la simulation de la propagation d'une zoonose.

Chapitre III: Domaine d'application: Propagation des zoonoses

1 Introduction

Dans ce chapitre, nous commençons par donner un aperçu général sur les maladies infectieuses en essayant de les définir et de les caractériser. Nous présentons d'ailleurs deux types de maladies qui vont nous servir d'exemples d'application. Nous présentons également les principales approches et modèles qui sont utilisés actuellement pour simuler la propagation de ces maladies.

2 Présentation des maladies infectieuses

Les maladies infectieuses sont la première cause de mortalité sur la planète surtout après leur prolifération suite au réchauffement climatique. En effet, des millions de personnes décèdent chaque année dans le monde des suites d'une infection. La liste est longue de ce type de maladies qui jalonnent l'histoire de la santé des hommes. Certaines récemment ressurgies avec la multiplication des échanges internationaux ou l'augmentation de la résistance aux antibiotiques, d'autres nouvellement apparues avec l'émergence d'agents infectieux jusque-là inconnus. Les grandes maladies comme le VIH/Sida, la tuberculose, le paludisme et la rougeole continuent de peser lourdement sur les économies et les sociétés du monde, notamment dans les pays en développement. Pour un grand nombre de ces maladies, il n'existe toujours pas de médicaments, de vaccins ou d'autres traitements efficaces [Site Web: *OMS*].

2.1 Définition et caractéristiques

Comme nous l'avons déjà mentionné brièvement dans le premier chapitre, les maladies infectieuses sont le résultat de la transmission d'un micro-organisme représentant l'agent pathogène (bactérie, virus, champignon ou parasite) d'un individu infecté (hôte: homme ou animal) à un individu sain. On parle d'une épizootie quand la transmission se fait entre animaux. Si l'épizootie touche un continent ou le monde, on parlera de panzootie, alors que

si elle frappe une région d'une façon constante, on parlera d'enzootie. Par ailleurs, l'épizootie peut se transformer en zoonose si elle se transmet à l'humain. C'est le cas par exemple de l'Encéphalopathie spongiforme bovine qui a frappé la Grande-Bretagne et s'est transmise à l'homme sous le nom de maladie de Creutzfeldt-Jakob. On distingue deux types de zoonoses: (1) une anthroponose qui est une maladie transmise de l'homme à l'animal et (2) une zooanthroponose qui est une maladie transmise de l'animal à l'homme. D'un autre côté, l'épizootie peut également évoluer en épidémie si elle se transmet de l'homme à l'homme. C'est le cas par exemple de la grippe aviaire (une épizootie) qui pourrait devenir contagieuse pour l'homme (une zoonose) et devenir très contagieuse entre les hommes eux-mêmes (une épidémie). On parle généralement d'épidémie quand les cas de transmission se multiplient dans un lieu et une période limitée. Si la diffusion est beaucoup plus généralisée, on parle alors de pandémie. Par contre, l'endémie désigne la présence habituelle d'une maladie dans une région déterminée [Site Web: OMS], [Krauss *et al.*, 2003].

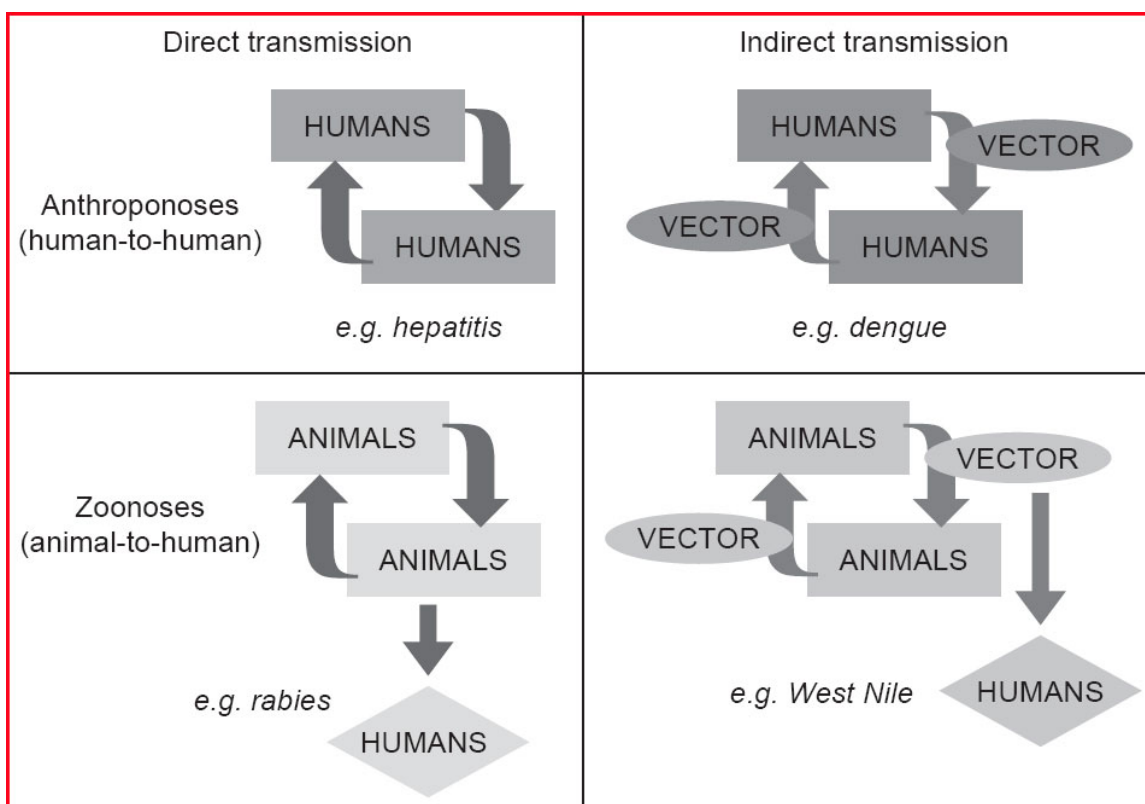


Figure 23: Principaux types de maladies infectieuses [McMichael et Woodruff, 2008].

En outre, l'infection peut se faire d'une façon directe quand l'agent pathogène se transmet de l'humain à l'humain ou de l'animal à l'animal. Elle peut se faire également d'une façon indirecte quand la transmission se fait par l'intermédiaire d'un vecteur (Figure 23). Ce vecteur est justement responsable de la propagation de la maladie et il est le plus souvent un arthropode (insecte, tique, etc.). Cependant, la circulation des humains et des animaux (d'élevages, domestiques) joue probablement un rôle majeur dans la diffusion et l'extension mondiale de nombreuses maladies infectieuses. D'autre part, ces maladies sont particulièrement sensibles aux changements écologiques tels que le réchauffement climatique susceptibles de modifier l'aire de répartition ou de diffusion de certains pathogènes et/ou vecteurs et de ce fait, favoriser la propagation de la maladie [Tran *et al.*, 2005].

Par ailleurs, le mode de transmission indirect des maladies fait intervenir de nombreux acteurs. Parmi ces acteurs, il y a bien sûr les hôtes et les vecteurs, mais également la présence d'un réservoir assurant le maintien de l'agent pathogène dans la nature. L'espèce-réservoir permet à l'agent pathogène (p.ex. le virus de la rage utilise des chauves-souris comme espèces-réservoir) de proliférer. Ces espèces jouent souvent le rôle de support de développement pour le pathogène, surtout si ce dernier cherche à déjouer l'immunité de son hôte. Ces espèces peuvent représenter également le moyen de transport et de large diffusion de cet agent pathogène, parfois sur de longues distances quand l'espèce-réservoir est un oiseau ou un mammifère très mobile. De plus, les excréments, les cadavres ou d'autres contaminants possibles peuvent faire disperser le pathogène vers d'autres espèces ou vers d'autres individus de la même espèce [McMichael et Woodruff, 2008].

Nous sommes particulièrement intéressés par les zoonoses à transmission indirecte. Ce type de maladie fait intervenir une multitude d'acteurs qui engendrent des interactions responsables de la propagation de l'agent pathogène. Nous présentons dans ce qui suit deux exemples de zoonoses qui nous semblent intéressantes à étudier surtout qu'elles représentent un risque important pour la santé humaine en Amérique du Nord [Hannon et Ruth, 2009].

2.2 Exemples de zoonoses

Il existe bien sûr plusieurs zoonoses (voir la liste des principales zoonoses ici [Site Web: *INMA*]), mais nous présentons dans cette section deux zoonoses qui nous intéressent étant donné qu'elles vont nous servir pour leur appliquer l'approche proposée. Une des raisons de ce choix est que nous disposons des données de leurs programmes de surveillance pour certaines régions du Canada (p.ex. le sud du Québec et la région métropolitaine d'Ottawa (voir Chapitre VII).

2.2.1 Virus du Nil Occidental (VNO)

Le virus du Nil occidental (VNO) est un flavivirus qui a été isolé pour la première fois en 1937. Son nom vient du district de West Nile en Ouganda. Il a été détecté chez les hommes, les oiseaux et les moustiques en Égypte au début des années 50, et a depuis été retrouvé dans divers pays. Ce n'est toutefois qu'avec l'écllosion importante de 1996 survenue à Bucarest, en Roumanie, que le VNO est devenu une préoccupation dans le domaine de la santé publique. D'ailleurs, on ne connaît pas de traitements spécifiques de la maladie et aucun vaccin n'est encore disponible pour l'être humain. Le VNO a été détecté sur le continent américain en 1999 et plus spécifiquement dans la ville de New York. Au Canada, le VNO a atteint le sud de l'Ontario en 2001. Cependant, les premiers cas humains ont été détectés au mois d'août 2002 [Tardif *et al.*, 2003]. En 2003, plus de 1200 cas humains d'infection au VNO ont été signalés, dont dix décès, au Canada et 9858 cas humains d'infection, dont 262 décès, aux États-Unis [Vincent *et al.*, 2004].

Le VNO a fait son apparition au Québec en juillet 2002. Le virus s'est alors propagé, comme partout ailleurs, par l'intermédiaire des moustiques (vecteurs) et des corvidés (réservoirs). L'expansion de cette zoonose a amené le gouvernement québécois à adopter un plan d'intervention. Ce plan prévoyait notamment la mise en place d'un système de surveillance. En effet, le Système intégré des données de vigie sanitaire (SIDVS-VNO) a été développé en 2003 par l'Institut national de santé publique du Québec (INSPQ). Cet outil vise notamment à effectuer la représentation cartographique en ligne, via un réseau Extranet, des principales données recueillies sur le terrain. Ces données renferment entre autres les signalements et analyses de corvidés morts ainsi que les analyses de pools de moustiques [Gosselin *et al.*, 2005]. Après 2003, les responsables de la santé publique au

Québec ont remarqué qu'il n'y a pas eu une augmentation significative de cas humains d'infection. C'est pour cette raison qu'ils ont décidé en 2006 d'arrêter leur système de surveillance animale dans la province. Cependant, le VNO continue sa progression surtout dans les autres provinces (p.ex. Saskatchewan, Manitoba, Alberta, etc.). D'ailleurs, on avait signalé en 2007, 2215 cas humains d'infection, dont 12 personnes décédées au Canada. Pendant la même année, il y a eu 3623 cas humains d'infection, dont 124 décès aux États-Unis. Plus récemment, on avait signalé en 2012, 433 cas humains d'infection, dont 6 personnes décédées au Canada. Pendant la même année, il y a eu 4725 cas humains d'infection, dont 219 décès aux États-Unis. Le nombre de cas cliniques au Québec a augmenté de 37 en 2011 à 126 (parmi les 433 au Canada) en 2012 [Site Web : *ASPC*].

2.2.1.1 Cycle de transmission du VNO

Le VNO a la caractéristique, comme tous les arbovirus, de se transmettre essentiellement par piqûre d'arthropodes hématophages. En effet, les moustiques propagent le virus après avoir contracté l'infection en se nourrissant du sang d'un hôte infecté porteur de ce virus. Les oiseaux sont considérés comme étant les réservoirs principaux et naturels du VNO, mais d'autres vertébrés, tels que les humains et les animaux, peuvent aussi être infectés, les chevaux y étant particulièrement sensibles (Figure 24).

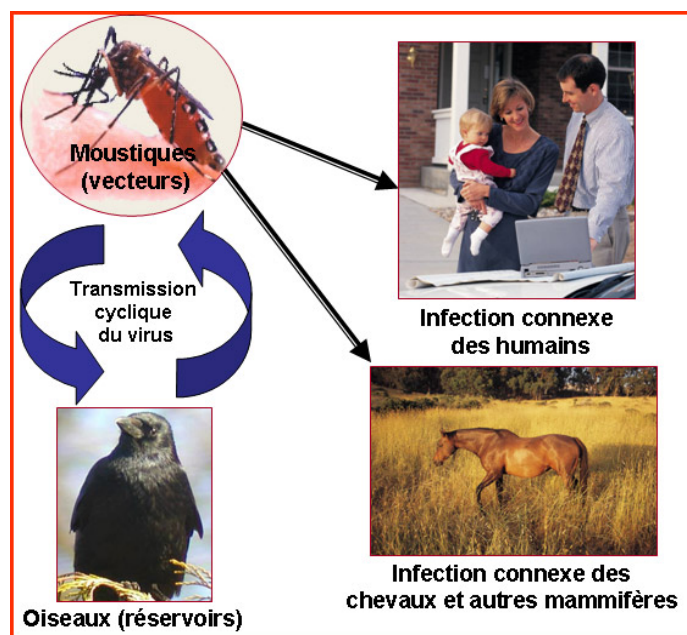


Figure 24 : Cycle de transmission du virus du Nil occidental [Site Web : *ASPC*].

Par ailleurs, l'infection des humains ainsi que des autres mammifères semble être moins fréquente (en dehors de la transmission cyclique du virus : *moustiques-oiseaux-moustiques*) étant donné que les moustiques impliqués dans la transmission du VNO sont majoritairement ornithophiles (c.-à-d. attirés par les oiseaux.). Cependant, des travaux de recherche plus récents ont démontré que certaines espèces de moustiques de l'Amérique du Nord sont hybrides (par rapport à leurs cousins de l'Europe) et peuvent chercher à piquer les mammifères d'une façon beaucoup plus fréquente [Fonseca *et al.*, 2004].

2.2.1.2 Populations de moustiques et leur rôle vectoriel

Il existe plus de 3000 espèces de moustiques partout dans le monde. Cependant, il y a 74 de ces espèces qui ont été répertoriées au Canada. Parmi elles, il y en a 52 qui sont présentes presque partout au Québec du printemps jusqu'à l'automne [Koné *et al.*, 2003].

Le cycle biologique des moustiques comprend, comme chez tous les insectes diptères, trois périodes (Figure 25). En effet, les larves se développent progressivement (augmentation de la taille) après l'éclosion des œufs. Elles sont aquatiques et se nourrissent de petites algues et d'animaux microscopiques. Elles utilisent de l'oxygène atmosphérique grâce à un siphon. Les nymphes de moustiques ont une forme globuleuse et elles sont moins mobiles. Après métamorphose, les moustiques adultes (période de reproduction sans croissance) émergent de l'eau. Les mâles sucent des sucres végétaux, mais ne piquent pas alors que les femelles ont besoin du sang pour la maturation des œufs [Site Web : *ASPC*].

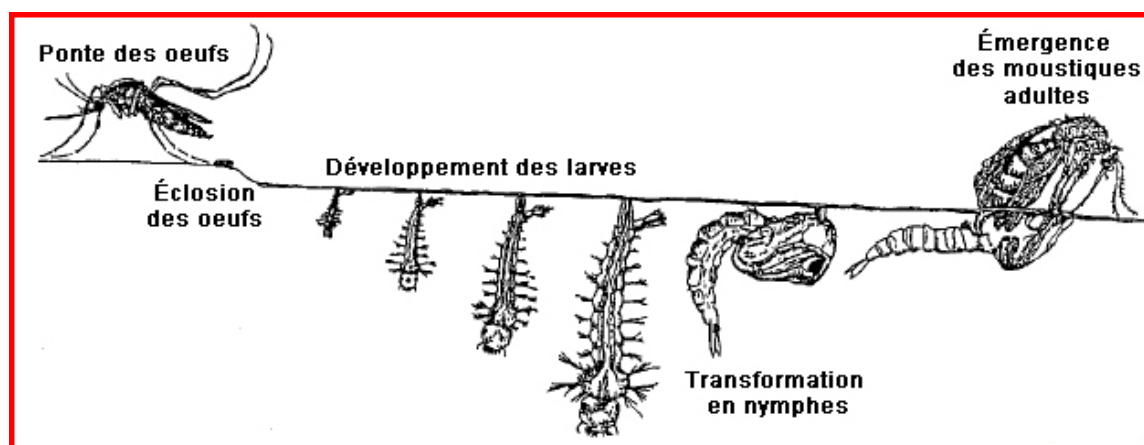


Figure 25 : Cycle biologique des moustiques [Site Web : *ASPC*].

Il y a essentiellement deux types de moustiques jouant un rôle important dans la transmission du VNO. Il s'agit des "vecteurs primaires" qui sont capables de maintenir le

virus dans le cycle naturel de transmission moustique-oiseaux-moustique (Figure 24) et des "vecteurs passerelles" qui permettent au virus de passer du cycle naturel aux mammifères. Les vecteurs primaires les plus importants au Québec sont *Culex pipiens* et *Culex restuans*. Ils sont présents de la fin avril à la fin septembre. Les vecteurs passerelles les plus prépondérants sont: *Anopheles sp.*, *Aedes vexans*, *Coquillettidia perturbans* (surtout en début d'été) et *Ochlerotatus canadensis* [GDG, 2002].

La femelle de moustique est infectée par le repas de sang qu'elle fait afin de pondre ses œufs. Ce sang contient un certain nombre de particules virales vivantes. Le processus d'infection est enclenché quand le virus traverse la membrane de l'intestin du moustique pour s'échapper dans son sang. Le virus peut alors se multiplier dans tout le corps du moustique, y compris les glandes salivaires et être éventuellement libéré dans la salive. De cette façon, le moustique nouvellement infecté peut transmettre le VNO lorsqu'il pique un hôte susceptible d'attraper la maladie. La période d'incubation du virus après son introduction dans le moustique (période d'incubation extrinsèque) est de 5 à 15 jours. Il est à noter que le moustique ne meurt pas suite à une infection au VNO [Koné *et al.*, 2003].

Le VNO peut se transmettre d'une génération de moustique à l'autre par l'intermédiaire des œufs (transmission verticale). Ces derniers peuvent entrer dans une période de repos pour éclore au printemps ou à l'été suivant. De plus, les femelles de certaines espèces de moustiques comme les *Culex pipiens* et *Culex restuans* peuvent survivre à l'hiver à l'état adulte. Elles se réfugient sous des abris naturels ou artificiels où elles demeurent immobiles dans un état de repos relatif, jusqu'au retour des conditions favorables à la reprise de leurs activités.

On estime que le pourcentage de femelles infectées d'une espèce vectrice du VNO dépasse rarement 1%, ce qui fait que la probabilité de piqûre infectée est presque toujours faible. Bien qu'il soit possible d'être piqué par un moustique infecté partout où le virus est actif, la plupart des infections au Canada se sont produites en juillet et en août, à un moment où les taux d'infection chez les moustiques étaient les plus élevés [Site Web: *ASPC*].

Les moustiques se reproduisent dans des milieux aquatiques ou sur des sols humides. Les pneus, les objets en plastique, les carcasses d'automobiles et d'appareils ménagers fournissent, lorsque remplis d'eau, un habitat idéal aux larves de moustiques. D'ailleurs, les larves de *Culex* se développent généralement dans des gîtes artificiels tels que les puisards urbains. Les mâles et les femelles peuvent s'éloigner du milieu strictement aquatique dans un rayon variant, selon les espèces, de quelques mètres à plusieurs kilomètres. Pour le cas du *Culex* par exemple, il peut parcourir une cinquantaine de mètres en moyenne pendant une semaine. Cependant, certains peuvent franchir une distance variant entre 500 et 800 mètres pendant la même période. Par ailleurs, différents facteurs, comme l'humidité et la température, influencent la durée de vie des moustiques qui peut varier d'une semaine à plusieurs mois. D'ailleurs et dans le contexte des changements climatiques, si un scénario de l'accroissement des températures moyennes et de l'humidité se concrétisait, il pourrait se produire une augmentation de la distribution et de l'abondance des moustiques vecteurs [Hubalek et Halouzka, 1999].

2.2.1.3 Populations d'oiseaux et leur rôle amplificateur

Le VNO a infecté plus de 150 espèces d'oiseaux en Amérique du Nord. Cependant, certaines de ces espèces sont particulièrement sensibles aux effets de la maladie. Les oiseaux de l'ordre des passériformes (corvidés, moineau domestique, mésange, etc.) sont généralement des réservoirs importants pour le maintien non seulement du VNO mais aussi d'autres flavivirus tels que le virus de l'encéphalite de Saint-Louis et le virus de l'encéphalite japonaise [Komar *et al.*, 2003]. Au Québec, les corvidés comptent quatre représentants. Nous retrouvons le Grand Corbeau (*Corvus corax*), la Corneille d'Amérique (*Corvus brachyrhynchos*), le Geai du Canada (*Perisoreus canadensis*) et le Geai bleu (*Cyanocitta cristata*). Ces passereaux regroupent des oiseaux de moyenne à grande taille (20 à 71 cm de longueur). Ce sont les corneilles et les corbeaux qui sont les plus grands. La Corneille d'Amérique profite des milieux habités par l'homme, alors que le Grand Corbeau, plus forestier et plus nordique, tend à s'en éloigner [Gauthier et Aubry, 1995].

La pathogénie du VNO chez les corvidés engendre une mortalité relativement rapide. Par ailleurs, la surveillance des oiseaux morts a montré qu'il y a un petit pic de mortalité au mois de mai suivi d'une accalmie (dont on pense que près de 100 % sont non reliées au VNO), et ensuite un pic important en août (dont on pense que 50 % des décès sont dus au

VNO) [Koné *et al.*, 2003]. Par conséquent, il est peu probable qu'au mois de mai, les corvidés puissent mourir suite à des piqûres de moustiques du Québec surtout que ces derniers ne sont pas très actifs au printemps. Cela semble renforcer l'hypothèse que des corvidés infectés aux États-Unis remontent vers le Québec au printemps dans leur cycle habituel de migration et qu'ils infectent les moustiques du Québec. Ces corvidés meurent alors au Québec. De ce fait, le cycle de transmission du début de la saison est soit *moustique-oiseaux-moustique*, soit *oiseaux-moustique-oiseaux*.

Après la saison de reproduction (mi-mars à début juillet pour le Québec), les groupes familiaux des corneilles, par exemple, se déplacent à la recherche de nouvelles sources de nourriture et forment, au fil des mois, des regroupements de plus en plus importants. Ces grands rassemblements constituent probablement la manifestation la plus spectaculaire de comportement social de l'espèce. En effet, à partir de la mi-juillet, les corneilles se réunissent dans des dortoirs pour la nuit, en groupes très variables de quelques dizaines d'individus à quelques dizaines de milliers. Aux premières lueurs du jour, elles se dispersent en suivant des couloirs aériens bien précis, sur des distances pouvant atteindre 80 km, et se mettent en quête de nourriture. Elles reviennent à leur dortoir par les mêmes couloirs à la fin de la journée [Gauthier et Aubry, 1995]. Il semble alors que les dortoirs soient une des cibles de choix pour les *Culex* qui sont, rappelons-le, ornithophiles et également nocturnes.

2.2.1.4 Interactions entre les moustiques et les corvidés

Dans la Figure 26, nous proposons une vue synthétique des principaux phénomènes qui nous intéressent. En effet, nous avons essayé de représenter d'une façon simple les interactions entre les moustiques et les corvidés [Bouden *et al.*, 2008a]. De plus, nous avons ajouté les comportements, qui nous semblent importants à considérer lors de la simulation, pour chacune des populations impliquées dans la transmission du VNO.

Les moustiques pondent des œufs dans des milieux contenant de l'eau comme les puisards ou les autres gîtes artificiels (p.ex. sites de pneus usagés). Ces œufs vont éclore pour donner vie à des larves qui se développent progressivement pour se transformer par la suite en nymphes. Ces dernières vont se métamorphoser en moustiques adultes. La période

d'émergence est propre à chaque espèce et dépend énormément des conditions climatiques comme la température et les précipitations. Par ailleurs, les œufs peuvent être infectés par la transmission verticale, mais ce mode d'infection demeure négligeable surtout chez les espèces de *Culex* [Nasci *et al.*, 2001]. Il est à noter également que l'intervention humaine, afin de diminuer la propagation du VNO, se fait généralement par des traitements de larvicides en utilisant essentiellement le Bti¹² ou encore le Méthoprène¹³.

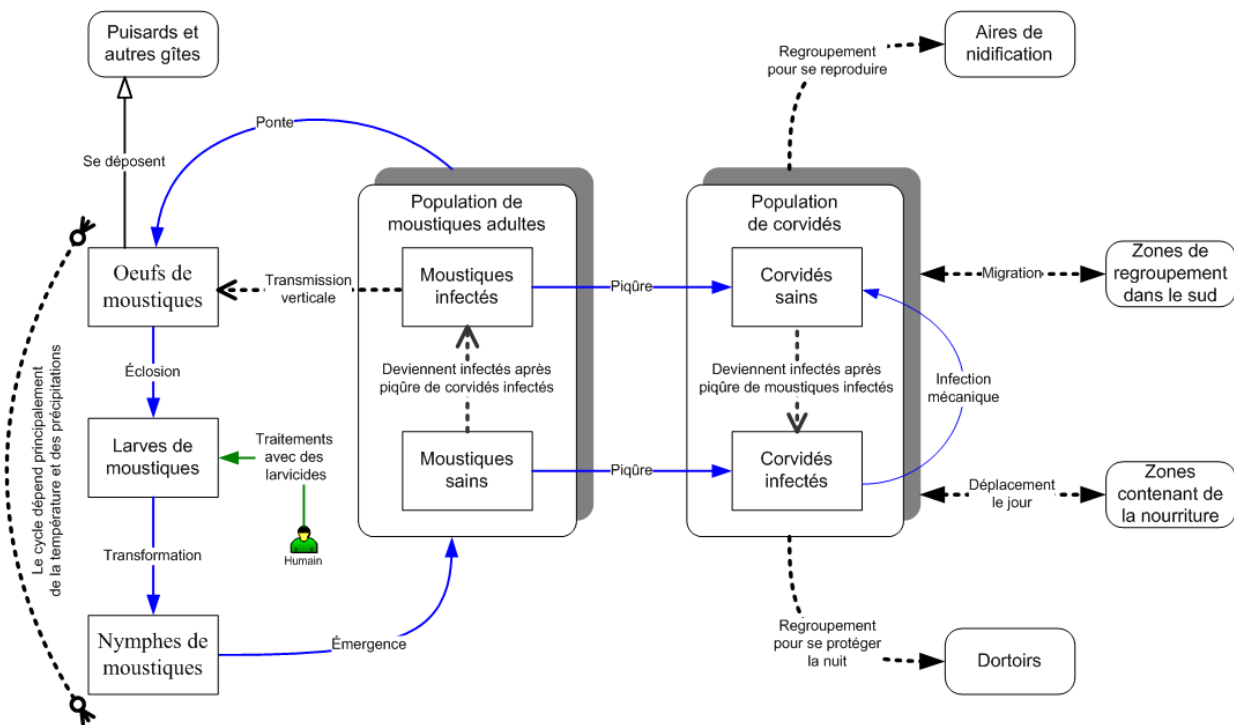


Figure 26: Diagramme de la vue d'ensemble des phénomènes qui nous intéressent [Bouden *et al.*, 2008a].

Les femelles de moustiques adultes sains qui vont piquer les corvidés infectés peuvent devenir des vecteurs du VNO. Ces vecteurs peuvent à leur tour infecter les corvidés sains en leur transmettant le virus après piqûre. Par ailleurs, cette transmission du VNO d'une population à l'autre peut être estimée en utilisant entre autres les probabilités d'infection.

¹² Bti (*Bacillus thuringiensis israelensis*): Larvicide biologique fabriqué à partir de bactéries du sol qui produisent des protéines cristallisées détruisant la paroi intestinale des larves. Cet effet ne peut survenir que dans un système digestif très basique. Celui des humains et autres animaux est si acide que le Bti ne représente aucun danger. Il est vaporisé dans les eaux stagnantes à l'aide d'épandeurs portatifs individuels (http://www.pmra-arla.gc.ca/francais/pdf/fact/fs_bti-f.pdf).

¹³ Méthoprène: Larvicide qui agit en perturbant l'action d'une hormone de régulation de la croissance chez les insectes. Sans risque notable pour les mammifères et les humains, mais légèrement à modérément toxique pour certains poissons et oiseaux, et très dommageable pour certains invertébrés d'eau douce. Nécessaire entre autres pour les puisards où l'effet du Bti est moins durable en raison des matières organiques en suspension (http://www.pmra-arla.gc.ca/francais/pdf/fact/fs_methoprene-f.pdf).

En effet, Wonham et ses collaborateurs [Wonham *et al.*, 2004] ont utilisé des paramètres pour leur modèle de propagation du VNO. Nous pouvons citer par exemple qu'en moyenne, 9% des individus d'une population de *Culex* vont piquer des corneilles en une journée et qu'il y a 88% des chances qu'un *Culex* puisse transmettre le VNO à une corneille après l'avoir piqué (voir Tableau E3 de l'Annexe E).

Par ailleurs, la population de corvidés est caractérisée par plusieurs types de déplacements dont la destination change en fonction de la saison. En effet, au début du printemps, les corvidés se regroupent dans des aires de nidification afin de se reproduire. Après cette période, les groupes familiaux se déplacent à la recherche de nouvelles sources de nourriture et se rassemblent peu à peu en bandes plus importantes. On assiste alors à un comportement grégaire qui se manifeste surtout par un regroupement dans des dortoirs. D'ailleurs, les corvidés parcourent des distances variables le jour afin de se nourrir et retournent par la suite à leurs dortoirs pour passer la nuit. Certains groupes de corvidés passent l'hiver au Québec alors que d'autres migrent plus ou moins loin vers le sud à partir du début de l'automne et reviennent au début du printemps. En outre, il existe un autre type de transmission du VNO qui ne fait pas intervenir les moustiques. En effet, le regroupement dans les dortoirs pourrait favoriser l'infection des corvidés essentiellement par voie orale, d'oiseau à oiseau. Cette transmission mécanique peut être enclenchée par un comportement agressif, un toilettage ou par simple proximité [Koné *et al.*, 2003].

2.2.2 Maladie de Lyme

La maladie de Lyme est une borréliose¹⁴ causée par une bactérie qui est véhiculée et transmise à l'homme par les tiques. La première description de cette maladie a été faite aux États-Unis en 1977 suite à l'observation d'un agrégat de cas d'arthrite rhumatoïde chez des enfants de la ville de Lyme au Connecticut. Ce n'est qu'en 1982 que William Burgdorfer a identifié l'agent responsable, *Borrelia burgdorferi*, isolé chez la tique *Ixodes scapularis*. La

¹⁴ La borréliose doit son nom à Amédée Borre, un célèbre bactériologiste français. Les borrelies sont un genre (*Borrelia*) de bactéries spiralées du groupe des spirochètes, découvert il y a plus de 100 ans. Le genre *Borrelia* regroupe aujourd'hui quelques dizaines d'espèces (36 fin 2008, mais d'autres pourraient ne pas encore avoir été découvertes) <http://fr.wikipedia.org/wiki/Borrelia>

maladie de Lyme se caractérise par l'apparition, au site de piqûre de la tique, d'une plaque rouge sur la peau, en forme de cible, dont la surface augmente progressivement durant les semaines suivant le contact. D'autres symptômes peuvent survenir, tels que des douleurs musculaires ou articulaires, des maux de tête, de la fièvre ou de la fatigue. Dans les semaines ou les mois suivants, des complications au niveau du cœur, du système nerveux ou des articulations peuvent apparaître [Site Web : *ASPC*].

Aux États-Unis, on fait état de quelque 20 000 infections annuellement [Site Web: *CDCP*]. Le nombre de cas de maladie de Lyme rapportés au Canada est relativement faible (de 20 à 60 nouveaux cas par an). Les populations de tiques infectées sont établies dans certaines régions du sud de l'Ontario, l'extrémité sud-est du Manitoba et certaines régions de la Rive-Sud de la Nouvelle-Écosse ainsi qu'en Colombie-Britannique [Ogden *et al.*, 2005]. Au Québec, seulement quelques cas sont recensés chaque année et les médecins estiment généralement que les patients ont été infectés lors de séjours aux États-Unis. Cependant, la situation pourrait changer puisqu'en 2008, des tiques infectées ont été identifiées entre autres en Montérégie et sur la Rive-Sud du Saint-Laurent entre Longueuil et Châteauguay [Ogden *et al.*, 2008a].

Les tiques vivent généralement dans les régions boisées ou les hautes herbes. Elles ont un cycle de développement qui dure environ deux ans et demi (voir Figure F2 de l'Annexe F et Figure 27a), pendant lequel on peut les retrouver sous trois stades (Figure 27b): larve, nymphe et adulte (mâle et femelle). Elles doivent trouver un hôte, animal ou humain, pour effectuer un "repas sanguin" à chacun des stades d'évolution afin de passer au stade suivant ou, pour les femelles adultes, afin de pouvoir pondre leurs œufs. À la fin de leur repas sanguin, les tiques gorgées augmentent de volume de façon significative, ce qui les rend plus facilement décelables [Ogden *et al.*, 2007].

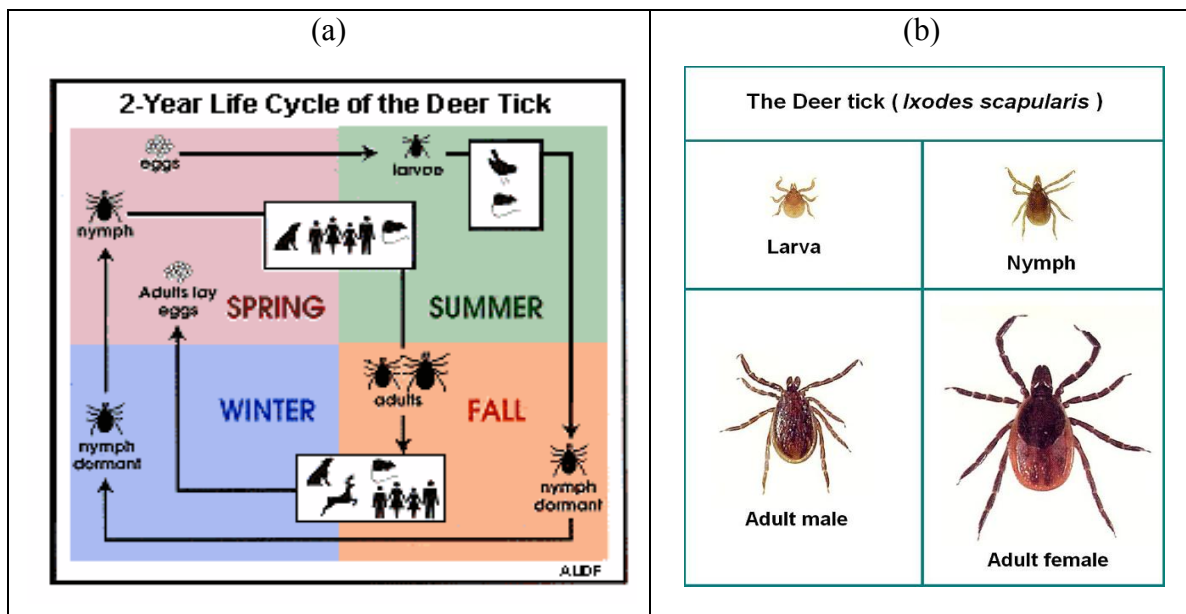


Figure 27: Stades (a) et cycle de développement des tiques (b) [Site Web: CDCP].

Les petits rongeurs ainsi que certains types d'oiseaux (surtout migrateurs) sont considérés comme les réservoirs naturels de la bactérie. En Amérique du Nord, les petites souris à pattes blanches (*Peromyscus leucopus*) sont les hôtes les plus fréquents pour les stades immatures (larves et nymphes). Pour le stade adulte, l'hôte le plus fréquent est le cerf de Virginie (*Odocoileus virginianus*), appelé également chevreuil, sur lequel peut avoir lieu la phase de reproduction. On peut également retrouver ces différents stades sur les animaux domestiques (p.ex. les chiens) et, occasionnellement, chez les humains. Les tiques doivent demeurer sur leur hôte de deux à cinq jours, selon le stade, pour compléter leur repas sanguin. Pour ce faire, elles ont tout intérêt à ne pas causer de désagrément à leur hôte au moment de la piqûre, pour éviter de se faire repérer. Les nymphes, de plus petite taille, sont probablement plus aptes que les formes adultes à transmettre l'agent pathogène puisqu'elles passent plus facilement inaperçues. De plus, elles sont présentes dans l'environnement durant l'été, où les activités extérieures sont plus fréquentes [Ogden *et al.*, 2008]

Les tiques ne se déplacent pas beaucoup et ne sont généralement présentes que dans des habitats précis (p.ex. forêt). Cependant, leur transport notamment par les oiseaux migrateurs explique l'identification de cas de maladie de Lyme chez les humains dans certaines régions du Canada où les tiques n'ont pas de présence connue. En effet, les

oiseaux (p.ex. passereaux) migrant au printemps font une escale pour se nourrir au moment qui coïncide avec l'activité saisonnière des tiques immatures. Ces tiques se nourrissent sans arrêt du sang des oiseaux durant cinq jours (ce qui équivaut à un parcours migratoire d'environ 500 km) et tombent par la suite dans leurs habitats [Ogden *et al.*, 2007].

D'un autre côté, l'augmentation de la température ambiante en raison des changements climatiques pourrait également mener à une expansion de la maladie. Cela pourrait avoir une incidence sur l'étendue future des tiques ainsi que les zones à risque pour l'infection humaine au Canada et plus spécifiquement au Québec. Il faut signaler également que pour l'instant et contrairement au VNO, les décideurs de la santé publique ne peuvent que prévenir la population humaine par des messages de sensibilisation afin de limiter la propagation de la maladie [Ogden *et al.*, 2006].

3 Modélisation et simulation de la propagation des maladies infectieuses

Nous avons déjà présenté dans le chapitre II les notions de base des principales approches qui sont utilisées pour la modélisation et la simulation des systèmes complexes d'une façon générale. Nous présentons dans cette section les approches qui sont utilisées actuellement pour modéliser et dans certains cas simuler la propagation des maladies infectieuses, et ce, à travers des exemples de travaux qui nous semblent intéressants. Nous commençons par la présentation brève de quelques types de modèles mathématiques étant donné que ces derniers nous seront très utiles pour développer notre outil générique permettant de calculer l'évolution et les interactions des acteurs responsables de la propagation des maladies infectieuses. Nous introduisons par la suite certains projets qui ont utilisé d'autres approches de simulation. Nous terminons en présentant des outils informatiques qui sont utilisés actuellement pour simuler la propagation des maladies infectieuses.

3.1 Modèles mathématiques

Nous avons mentionné auparavant (voir section 3.1.1 du chapitre II) que les modèles mathématiques présentent certaines lacunes et ne peuvent pas à eux seuls satisfaire les attentes des décideurs dans le domaine de la santé publique. Cependant, ces modèles sont

indispensables, étant donné qu'ils permettent de calculer la dynamique des populations impliquées dans la propagation des zoonoses.

D'un autre côté, il faut mentionner que nous collaborons avec plusieurs chercheurs affiliés au projet *CODIGEOSIM* et qui travaillent fréquemment avec des modèles mathématiques. Il s'agit notamment de l'équipe du Dr. Ogden N., directeur de GREZOSP (Groupe de Recherche en Épidémiologie des Zoonoses et Santé Publique) [Site Web: GREZOSP] et professeur à la Faculté de médecine vétérinaire de Saint-Hyacinthe. Il s'agit également de l'équipe du Dr. Wu J., directeur du *CSM (Centre for Disease Modelling)* [Site Web: *CDM*] et professeur au département de mathématiques et de statistiques à l'Université *York* de Toronto. De ce fait, nous présentons dans ce qui suit, les principaux modèles que nous pouvons éventuellement utiliser pour nos propres besoins.

3.1.1 Modèles compartimentaux

Il existe différents types de modèles compartimentaux (*Compartment models*), mais ils sont à la base de la modélisation mathématique en épidémiologie. Nous citons par exemple les modèles *SI* (susceptible-infecté) qui ne considèrent que les individus infectés et les individus qui sont susceptibles de le devenir. C'est donc des modèles à deux compartiments qui prennent comme hypothèse que la transmission de la maladie se passe dans un temps suffisamment court pour que les naissances et morts "naturelles" aient un impact négligeable sur la dynamique de la maladie. On note "*S*" le compartiment contenant les individus susceptibles. En un instant *t*, le compartiment *S* comporte *S(t)* individus. De la même façon, on note "*I*" le compartiment contenant les individus infectés et qui renferme *I(t)* individus au temps *t*. On considère maintenant que la probabilité qu'un individu susceptible devienne infecté est proportionnelle au nombre d'individus actuellement infectés en utilisant un coefficient de proportionnalité *i*. Sous forme d'équations différentielles, ceci devient :

$$\frac{dS}{dt} = -iIS$$
$$\frac{dI}{dt} = iIS$$

C'est le modèle le plus simple, mais il existe d'autres modèles plus complexes faisant intervenir plusieurs paramètres comme par exemple les modèles *SIS*, *SEI*, *SEIS*, *SEIR* et *SEIRS*. Par exemple, si on ajoute le facteur guérison, les individus infectés peuvent redevenir, à une probabilité g , de nouveau susceptibles. Les équations deviennent:

$$\frac{dS}{dt} = -iIS + gI$$

$$\frac{dI}{dt} = iIS - gI$$

De ce fait, on peut introduire d'autres compartiments dans le modèle comme le compartiment "X" représentant les individus décédés suite à une infection, Il y a également le compartiment "E" représentant les individus exposés à la maladie, mais qui ne sont pas encore contagieux puisque l'agent pathogène a besoin d'une période d'incubation. Il y a aussi le compartiment "R" représentant les individus ayant guéri d'une infection et qui développent dans certains cas une certaine immunité à l'infection [Noël, 2007]. Nous présentons d'ailleurs dans la Figure 28 un exemple de modèle mathématique comportant 8 équations différentielles pouvant calculer la dynamique des populations de moustiques (y compris les larves) et des populations d'oiseaux impliquées dans la propagation du VNO [Wonham *et al.*, 2004].

<p>(1) $\frac{dS_B}{dt} = -abI_M \frac{S_B}{N_B}$</p> <p>(2) $\frac{dI_B}{dt} = abI_M \frac{S_B}{N_B} - \mu_V I_B - gI_B$</p> <p>(3) $\frac{dR_B}{dt} = gI_B$</p> <p>(4) $\frac{dX_B}{dt} = \mu_V I_B$</p> <p>(5) $\frac{dL_M}{dt} = \beta_M(S_M + E_M + I_M) - mL_M - \mu_L L_M$</p> <p>(6) $\frac{dS_M}{dt} = -acS_M \frac{I_B}{N_B} - \mu_A S_M + mL_M$</p> <p>(7) $\frac{dE_M}{dt} = acS_M \frac{I_B}{N_B} - kE_M - \mu_A E_M$</p> <p>(8) $\frac{dI_M}{dt} = kE_M - \mu_A I_M$</p>	<p>S_B: Susceptible birds I_B: Infected birds R_B: Recovered birds X_B: Dead birds L_M: larvae of mosquitoes S_M: Susceptible mosquitoes E_M: Exposed mosquitoes I_M: Intected mosquitoes</p>
--	--

Figure 28: Modèle mathématique de VNO à 8 équations différentielles [Wonham *et al.*, 2004].

3.1.2 Modèles à *patches* (*patchy models*)

Les modèles que nous venons de voir se basent sur la "réaction-diffusion". En effet, on distingue un processus local, dans lequel les différentes populations, impliquées dans la propagation de la maladie, évoluent. On distingue également un processus de diffusion qui provoque une répartition de ces populations d'une façon continue dans l'espace. Cependant les modèles à *patches* (c.-à-d. *patchy models*) essaient de simuler l'apparition d'une façon discontinue de ce que Liu et ses collaborateurs [Liu *et al.*, 2006a] appellent des "motifs de dispersion spatiale". Ces motifs reflètent la présence de la maladie dans des régions pas nécessairement contigües dans l'espace. Les *patches* correspondent dans la plupart des cas à des régions administratives, dont le nombre et le découpage sont choisis en fonction de la disponibilité des données permettant d'alimenter les modèles.

$$\begin{aligned}
 \frac{dB_{Si}}{dt} &= b_i - d_{bi}B_{Si} + \sum_{\substack{j=-m+i \\ j \neq i}}^{m+i} D_{bji}B_{Sj} - \sum_{\substack{j=-m+i \\ j \neq i}}^{m+i} D_{bij}B_{Si} - \frac{C_{mbi}M_{Ii}B_{Si}}{N_{Bi}}, \\
 \frac{dB_{Ii}}{dt} &= -d_{b2i}B_{Ii} + \frac{C_{mbi}M_{Ii}B_{Si}}{N_{Bi}} + \sum_{\substack{j=-m+i \\ j \neq i}}^{m+i} D_{bji}B_{Ij} - \sum_{\substack{j=-m+i \\ j \neq i}}^{m+i} D_{bij}B_{Ii}, \\
 \frac{dM_{Si}}{dt} &= m_i - d_{mi}M_{Si} + \sum_{|k-i|=1} D_{mki}M_{Sk} - \sum_{|k-i|=1} D_{mik}M_{Si} - \frac{C_{bmi}M_{Si}B_{Ii}}{N_{Bi}}, \\
 \frac{dM_{Ii}}{dt} &= -d_{mi}M_{Ii} + \sum_{|k-i|=1} D_{mki}M_{Ik} - \sum_{|k-i|=1} D_{mik}M_{Ii} + \frac{C_{bmi}M_{Si}B_{Ii}}{N_{Bi}},
 \end{aligned}$$

Figure 29: Équations faisant évoluer les populations de moustiques et d'oiseaux pour chacun des *patches* subdivisant la région étudiée [Gourley *et al.*, 2008].

Par ailleurs, Gourley et ses collaborateurs [Gourley *et al.*, 2008] ont modélisé l'évolution et les interactions des moustiques et des oiseaux qui sont responsables de la propagation du VNO dans une région subdivisée en *patches* qui ne peuvent pas normalement se chevaucher. La distance moyenne que peuvent parcourir les moustiques durant leur vie est utilisée comme unité de mesure pour la partition de la région étudiée en *patches*. En effet, les moustiques appartenant au $i^{\text{ème}}$ *patch* ne peuvent se déplacer qu'aux *patches* voisins $i+1$ et $i-1$. Cependant, les oiseaux appartenant à un *patch* spécifique peuvent se déplacer au $m^{\text{ième}}$ *patches* voisins $i-m, \dots, i+i, \dots, i+m$, avec $m \geq 1$. La Figure 29 montre les équations

différentielles que les auteurs utilisent pour modéliser l'évolution des populations appartenant à chacun des *patches*. La variable i est un identifiant du *patch* correspondant. Cette variable est donc entre 1 et N , où N est le nombre total des *patches*. Dans chaque *patch* i , le nombre total des oiseaux est $N_{Bi} = B_{Si} + B_{Ii}$ et le nombre total des moustiques est $N_{Mi} = M_{Si} + M_{Ii}$ (sachant que S : susceptible et I : infecté) [Gourley *et al.*, 2008].

3.1.3 Modèles de métapopulations

La dynamique des métapopulations est relativement récente dans le domaine des mathématiques et de l'épidémiologie, mais elle a été utilisée depuis longtemps en écologie [Fulford *et al.*, 2002]. En effet, une métapopulation est un groupe de populations de même espèce séparées spatialement. Ce terme a été utilisé pour la première fois par Levins [Levins et Lewontin, 1985] pour décrire un modèle de dynamique de population d'insectes nuisibles à l'agriculture. L'idée a été reprise par la suite afin de modéliser des populations soumises par ce que Hanski [Hanski, 1999] appelle la "fragmentation géographique". Les habitats favorables et défavorables alternent, et les animaux risquent leur vie s'ils essaient de passer d'un endroit à l'autre. La survie d'une métapopulation dépend d'une recolonisation périodique par dispersion. Quand le taux de recolonisation est supérieur au taux d'extinction, la métapopulation survit [Hanski, 1999].

Les modèles mathématiques utilisant les métapopulations et s'intéressant à la propagation des maladies infectieuses (*Metapopulation Models*), ont été utilisés entre autres par le Dr. Arino, professeur au département de mathématiques de l'Université de Manitoba à Winnipeg et qui est membre également du projet *CODIGEOSIM*. Ces modèles représentent des graphes avec des systèmes d'équations différentielles dans chacun de leurs sommets. Les sommets sont appelés aussi des "*patches*". En effet, un *patch* est une unité au sein de laquelle la population est considérée homogène. Ce *patch* peut également représenter une localisation géographique (p.ex. quartier, ville, région, pays, etc.). Les *patches* peuvent ou ne peuvent pas se chevaucher. Ils peuvent être contigus ou séparés dans l'espace. Ils sont normalement reliés par le mouvement des espèces entre eux en utilisant des arcs reliant les sommets des graphes. Par conséquent, chaque sommet ou plutôt chaque *patch* contient un certain nombre de sous-populations d'espèces [Arino, 2009].

3.2 Autres approches de simulation

Nous présentons dans cette section quelques travaux utilisant d'autres approches permettant de simuler la propagation des maladies infectieuses. Nous commençons par les projets qui ont utilisé la dynamique des systèmes. Nous citons par exemple les travaux de Hannon et Ruth [Hannon et Ruth, 2009] qui ont développé plusieurs modèles épidémiques en utilisant le logiciel Stella [Site Web: *Stella*]. La Figure 30 montre un exemple simple de ces modèles. En effet, le réservoir "Sick" représente le nombre d'individus malades. La variable "Getting Sick" permet de contrôler ce nombre en utilisant une variable de transformation "Contagion Rate" pouvant estimer le nombre d'individus qui deviennent malades par unité de temps ($Getting\ Sick = Contagion\ Rate * Sick$). La variable "Awareness level" est une fonction qui affecte le nombre de malades en modifiant le "Contagion Rate". De plus et après une période de temps "Sick Time", les individus malades peuvent guérir en utilisant une variable de contrôle "Getting Well" [Hannon et Ruth, 2009].

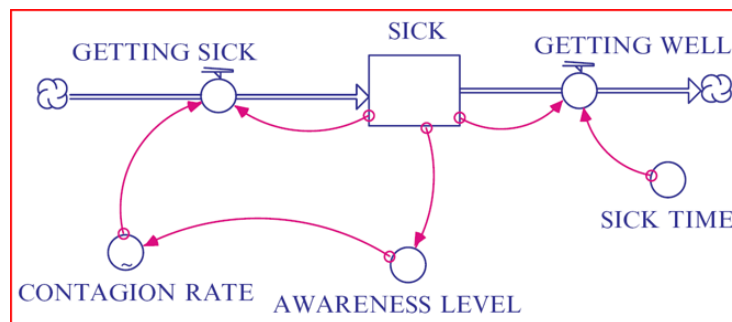


Figure 30: Modèle épidémique [Hannon et Ruth, 2009].

Ces mêmes auteurs [Hannon et Ruth, 2009] ont simulé la propagation de la maladie de Lyme (voir Section 2.2.2 du chapitre III) dans un écosystème représentant un hectare de forêt de chênes. Ils ont utilisé des modèles de la dynamique des systèmes faisant intervenir plusieurs acteurs notamment les tiques, les petites souris à pattes blanches et les chevreuils. La Figure 31 montre un de ces modèles représentant la dynamique de la population des tiques à travers leur cycle biologique [Hannon et Ruth, 2009].

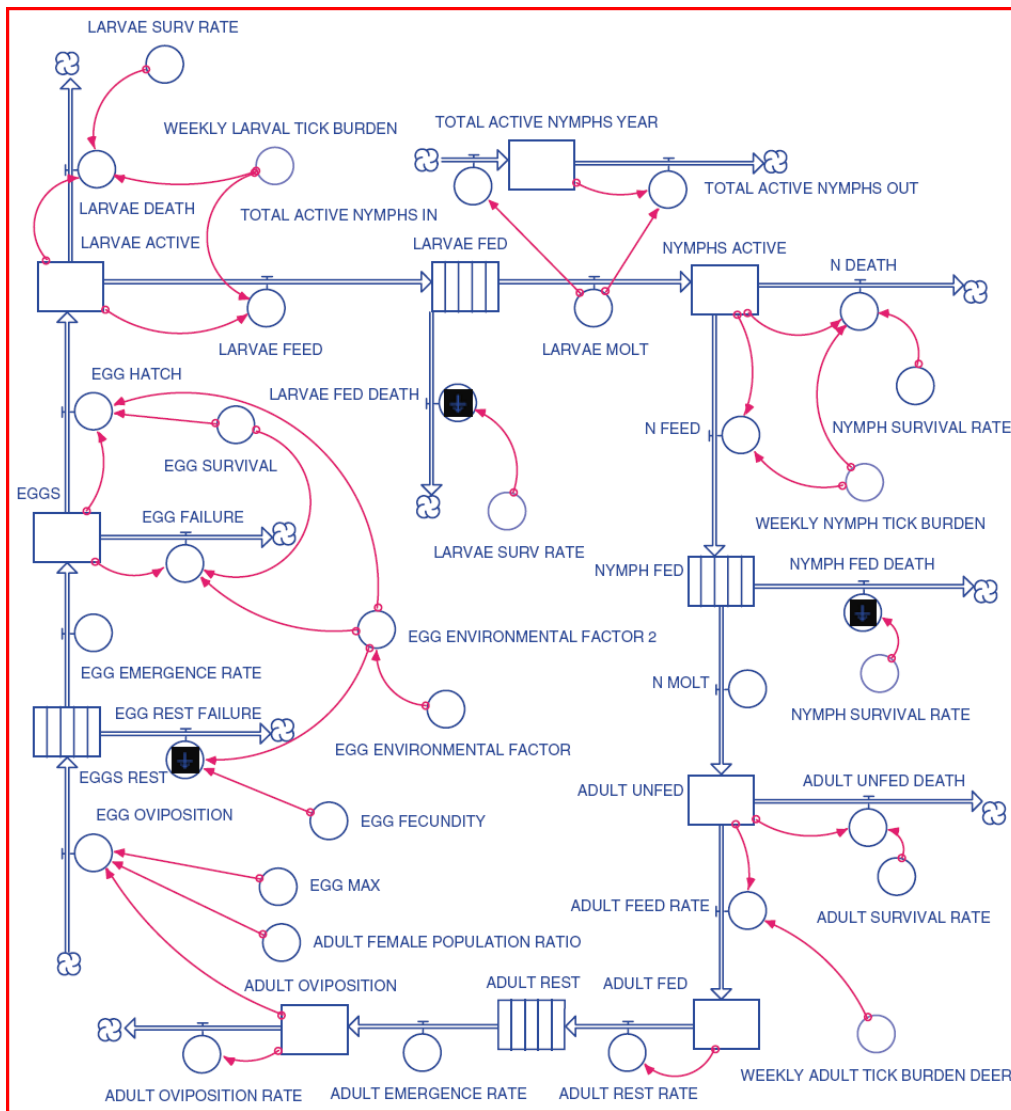


Figure 31: Modélisation du cycle biologique des tiques en utilisant la dynamique des systèmes [Hannon et Ruth, 2009].

Par ailleurs, Ogden et ses collaborateurs [Ogden *et al.*, 2005] ont également utilisé la dynamique des systèmes et le logiciel *Stella* pour modéliser l'impact de la température sur l'évolution des populations des tiques responsables de la propagation de la maladie de Lyme. La Figure 32 montre justement les différents états représentant chacun un stade spécifique de la vie des tiques. Ce modèle est conçu pour être capable de simuler l'évolution des populations de tiques dans des régions géographiques bien déterminées en utilisant les données de température de ces régions [Ogden *et al.*, 2005].

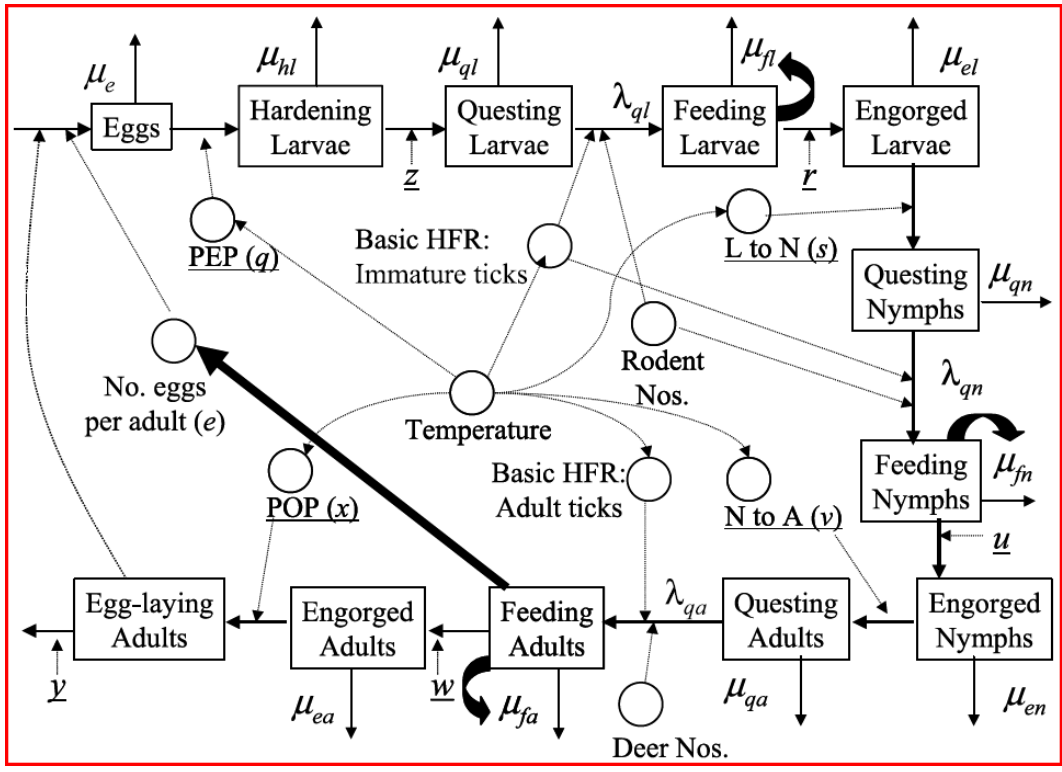


Figure 32: Dynamique des populations de tiques [Ogden *et al.*, 2005].

Par ailleurs, d'autres travaux ont utilisé les automates cellulaires [Fu et Milne, 2003], [Beauchemin *et al.*, 2005], [Liu *et al.*, 2006b]. Nous citons, par exemple, les travaux de White et ses collaborateurs [White *et al.*, 2009]. Ces derniers ont utilisé un automate cellulaire à deux dimensions afin de simuler d'une façon générique la propagation d'une maladie infectieuse. Chaque cellule de l'automate représentant une subdivision régulière de l'environnement et contient une même population homogène d'individus qui sont impliqués dans la propagation de la maladie. L'état de chaque cellule à chaque pas de temps est obtenu à partir de la fraction du nombre d'individus infectés. Cet état dépend également de l'état des 8 cellules avoisinantes et aussi de l'état de la cellule elle-même au pas de temps précédent. Les auteurs utilisent un code de couleur pour représenter les différents états des cellules. Ces auteurs mentionnent à la fin de leur article que les simulations obtenues semblent être réalistes alors qu'ils n'ont fait aucune calibration de leur modèle [White *et al.*, 2009].

Finalement, nous aimerions mentionner que d'autres travaux ont simulé la propagation des maladies infectieuses en utilisant les systèmes multi-agents [Emrich *et al.*, 2007], [Deng *et*

al., 2008], [Bauer *et al.*, 2009]. Nous citons, par exemple, les travaux de Liu et ses collaborateurs [Liu *et al.*, 2008] qui ont proposé un modèle multi-agent pour simuler la transmission spatio-temporelle d'une épidémie. Les auteurs utilisent 4 groupes d'agents: (1) les agents susceptibles, (2) les agents exposés, (3) les agents infectés et (4) les agents guéris. Leur modèle est formé de 4 composantes: (1) Les attributs associés aux agents comme par exemple le jour d'infection ou encore la localisation, (2) les règles de mobilité comme par exemple le fait qu'un agent ne peut pas se déplacer dans une cellule que d'autres agents occupent, (3) les règles de voisinage comme par exemple le fait qu'un agent saint peut devenir infecté s'il se rapproche suffisamment d'un agent infecté et (4) les règles des états de transition permettant de faire changer l'appartenance d'un agent à un des 4 groupes mentionnés auparavant [Liu *et al.*, 2008].

3.3 Outils informatiques

Les chercheurs qui ont développé des modèles mathématiques en utilisant simplement les équations différentielles ou encore en utilisant la DS pour simuler la propagation des maladies infectieuses utilisent généralement des outils comme *Stella* qui a été déjà mentionné, mais également *Powersim* [Site Web: *Powersim*] ou bien *Vensim* [Site Web: *Vensim*]. Cependant, les chercheurs qui ont utilisé les autres approches comme les systèmes multi-agents développent pour la plupart des cas de nouveaux outils en utilisant des langages comme C⁺⁺ ou encore Java.

Par ailleurs, nous aimerions citer l'exemple de *AnyLogic* [Site Web: *AnyLogic*] qui est un logiciel multi-paradigme pouvant simuler des modèles en utilisant plusieurs approches notamment la DS et les SMA. Cet outil permet donc de simuler un système hybride qui est une combinaison entre un système à événements discrets et un système continu. Il faut également savoir que ce logiciel utilise les diagrammes d'états (*StateCharts*) qui font partie des standards UML [Site Web : *OMG*] afin de modéliser les comportements d'agents dans des SMA. Le système hybride consiste à ajouter entre autres des équations différentielles au niveau des états des *statecharts*. Par exemple, Rahmandad et Sterman [Rahmandad et Sterman, 2006] ont utilisé *AnyLogic* pour simuler la propagation d'une maladie infectieuse en utilisant à la fois la DS et les SMA. La Figure 33a montre le modèle de DS utilisant le

diagramme des réservoirs et des valves et la Figure 33b montre le modèle des SMA utilisant le diagramme de *statechart* [Rahmandad et Sterman, 2006].

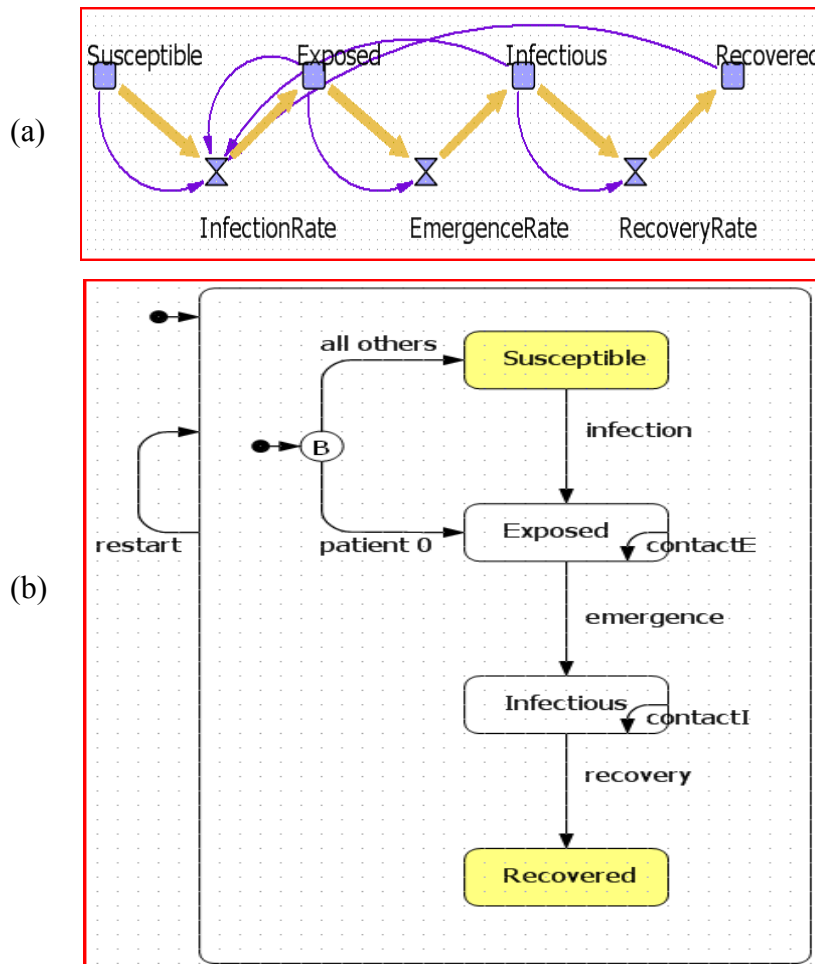


Figure 33: Modèle DS (a) et modèle SMA (b) en utilisant *Anylogic* [Rahmandad et Sterman, 2006].

Les outils que nous venons de présenter sont quand même assez limités et ne peuvent pas répondre à nos attentes surtout que nous avons besoin d'un outil générique capable de produire des géosimulations multi-niveaux. Par contre, nous aimerions citer l'outil *STEM* (*Spatiotemporal Epidemiological Modeler*) [Site Web: *STEM*] qui a été développé initialement par *IBM* et qui est rendu maintenant "*open source*". Il est basé surtout sur des modèles compartimentaux (SI, SIR, ou SEIR). Il présente certains avantages puisqu'il permet par exemple d'intégrer les données SIG, mais il présente par contre l'inconvénient d'être un outil très technique puisqu'il utilise par exemple des "Triggers" pour simuler des événements et également il est très lent à l'exécution puisqu'il est implanté en Java.

Néanmoins, cet outil nous a fourni des pistes lors de la phase d'implémentation de notre propre outil générique.

4 Conclusion

Nous venons de voir dans ce chapitre deux types de maladies infectieuses qui ont des similarités ainsi que des différences. Les deux maladies font intervenir plusieurs acteurs permettant de propager l'agent pathogène à travers le territoire. C'est en fait l'interaction entre ces acteurs qui permet au virus du VNO ainsi qu'à la bactérie de la maladie de Lyme de pouvoir se transmettre, et de ce fait, infecter leurs hôtes. Nous avons pu survoler également les principales approches utilisées pour modéliser et simuler une zoonose. Nous constatons d'ailleurs que des approches comme les systèmes multi-niveaux ou les systèmes holoniques n'ont pas encore été utilisées dans le domaine qui nous intéresse et semblent être une bonne alternative pour trouver des éléments de réponse à notre problématique. De plus, nous remarquons que les rares outils informatiques utilisés actuellement pour la propagation d'une zoonose ne répondent pas aux attentes des décideurs dans le domaine de la santé publique. L'approche que nous proposons et que nous présentons dans le chapitre suivant nous permettra de produire notre propre outil qui aura l'avantage d'être générique afin de simuler, entre autres, le VNO et la maladie de Lyme.

Chapitre IV: Approche de géosimulation multi-niveau

1 Introduction

Dans ce chapitre, nous présentons notre approche de géosimulation multi-niveau [Bouden et Moulin, 2010]. En effet, nous commençons par la présentation d'une vue d'ensemble de notre approche en listant entre autres les différents aspects qui nous intéressent. Par la suite, nous présentons la modélisation des différents niveaux de granularité faisant partie de notre approche. Finalement, nous discutons dans une conclusion la contribution que nous apportons par cette approche.

2 Vue d'ensemble de notre approche

Nous présentons dans cette section une vue d'ensemble de notre approche en commençant par expliquer l'utilité d'un outil d'aide à la décision. Par la suite, nous présentons une vue globale des différents aspects qui devront normalement être traités lors du développement d'un outil de géosimulation de la propagation des maladies infectieuses. Finalement, nous présentons une esquisse des différents aspects qui nous intéressent réellement dans le cadre de cette thèse et qui nous permettront éventuellement de satisfaire les objectifs de notre recherche.

2.1 Utilité d'un outil d'aide à la décision

Nous aimerions montrer dans cette section l'utilité du développement d'un outil d'aide à la décision. En effet, nous proposons, dans la Figure 34, une vue très synthétique montrant comment les décideurs peuvent intervenir quand ils sont confrontés à un phénomène réel qui a des conséquences sur la société et/ou l'environnement. En effet, ces décideurs vont tenter d'évaluer les conséquences du phénomène étudié. Cette évaluation leur permettra de prendre la décision d'intervenir pour tenter de modifier ces conséquences. Les décideurs peuvent également décider d'ignorer ces conséquences s'ils jugent par exemple que le phénomène étudié n'a pas un effet significatif sur la société et/ou l'environnement permettant de justifier une intervention sur le terrain. Cette intervention, rappelons-le, demande souvent beaucoup de ressources à la fois matérielles et en personnel. De ce fait,

quand les décideurs ignorent parfois les conséquences d'un phénomène ou du moins quand ils limitent les actions d'intervention, ce n'est pas toujours par choix, mais souvent par nécessité. Nous avons d'ailleurs déjà donné un tel exemple qui est lié à la décision des responsables de la santé publique au Québec d'arrêter le programme de surveillance animale du VNO en 2006.

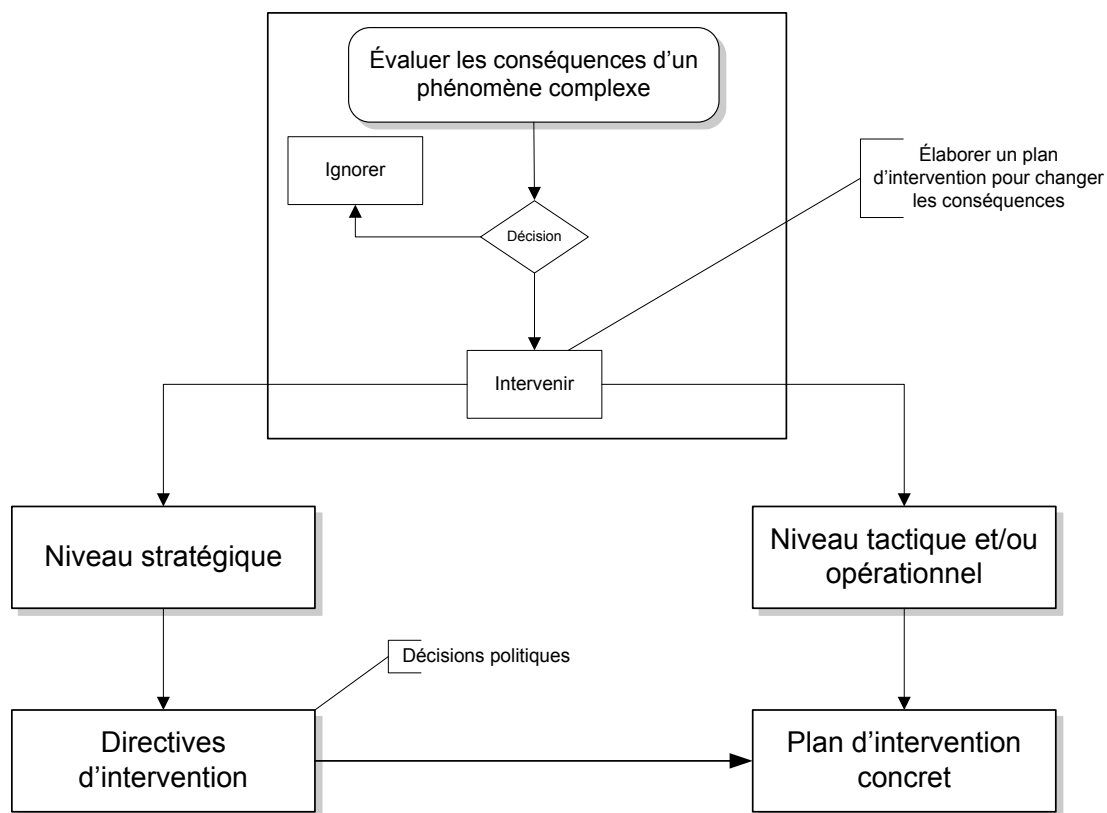


Figure 34: Système de prise de décision et d'intervention.

Par ailleurs et selon l'importance des conséquences d'un phénomène réel, l'élaboration d'un plan d'intervention pourrait être incontournable si par exemple les effets de ce phénomène ont des répercussions jugées néfastes sur la société et/ou l'environnement. Ce plan d'intervention pourrait se situer à plusieurs niveaux. Le niveau stratégique a un objectif global et à plus long terme alors que les niveaux tactique et/ou opérationnel ont un objectif plus local et limité dans le temps. En effet, le plan d'intervention au niveau stratégique est préparé généralement au plus haut niveau de l'État, dans un dialogue entre responsables politiques, diplomatiques et/ou militaires. Il contient généralement des directives d'intervention comme par exemple celles de la déclaration du G8 sur la lutte contre les maladies infectieuses à Saint-Petersbourg en 2006 [Site Web : *DG8LCMI*]. Cependant, le plan d'intervention au niveau tactique et/ou opérationnel est préparé généralement par un

responsable d'un théâtre d'opérations ou le responsable d'une unité engagée dans une action particulière. C'est donc un plan d'intervention concret comme celui préparé par la direction des opérations de la société de protection des forêts contre les insectes et maladies (SOPFIM) permettant de faire des applications de larvicides pour réduire la propagation du VNO [Site Web: SOPFIM].

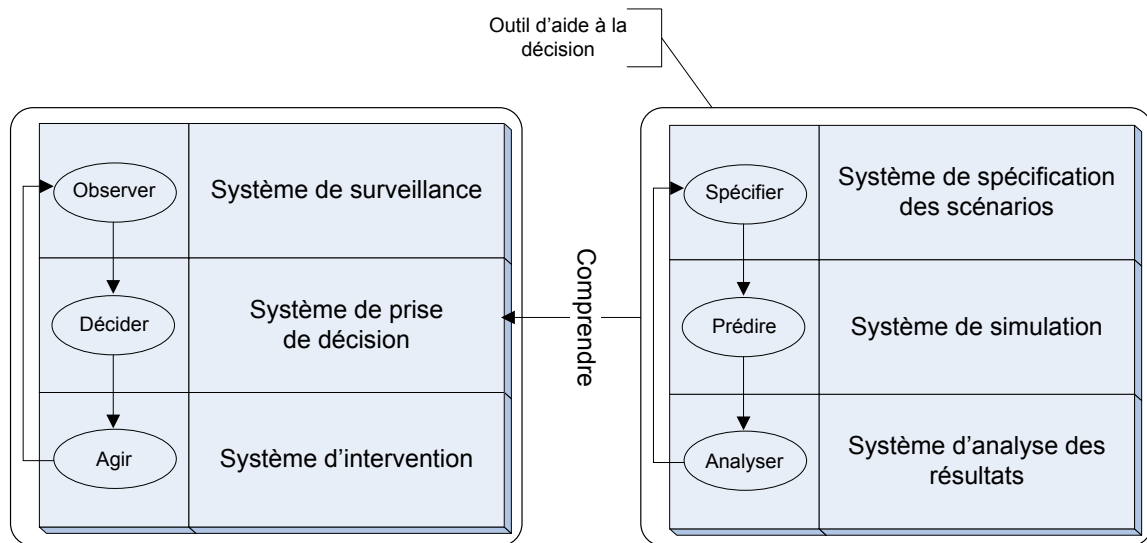


Figure 35: L'utilisation d'un outil d'aide à la décision par les décideurs afin de prendre des décisions informées.

En outre, nous avons mentionné dans le premier chapitre (voir Figure 3) que les décideurs observent normalement le phénomène étudié avant de pouvoir décider comment, quand et où agir pour tenter de modifier ses conséquences. Toutefois, il n'est pas facile de prendre des décisions éclairées permettant d'établir un plan d'intervention (stratégique, tactique ou opérationnel) si les décideurs se basent uniquement sur l'observation du phénomène. De ce fait, ils ont besoin d'outils d'aide à la décision permettant de simuler ce phénomène selon divers scénarios alternatifs d'intervention. En utilisant ces outils, les décideurs pourraient spécifier différents scénarios, prédire l'évolution du phénomène selon ces scénarios et finalement comprendre ce phénomène en analysant les résultats de la simulation (Figure 35).

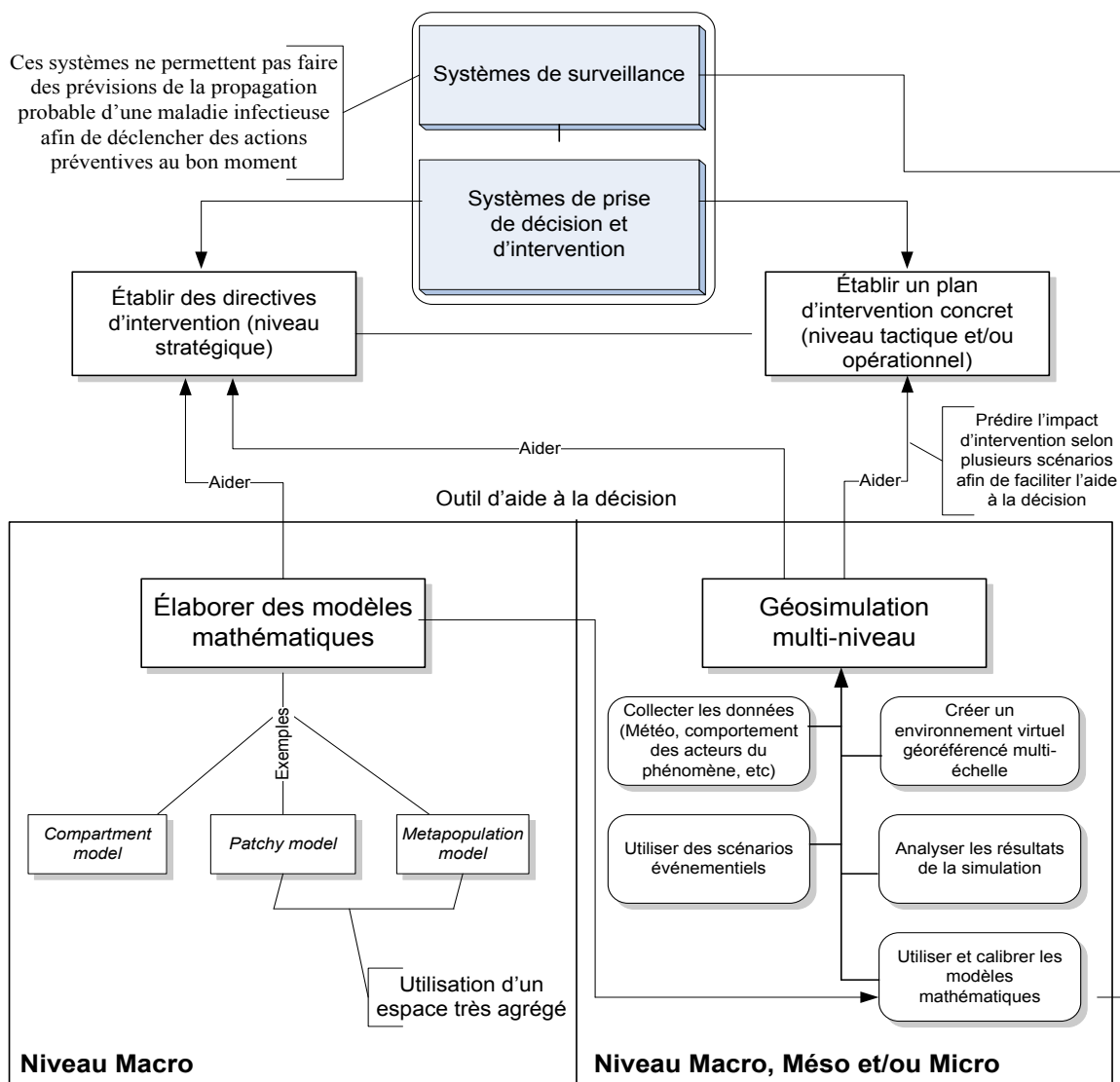


Figure 36: Les différents niveaux de granularité (Macro, Mésos et/ou Micro) qui sont utilisés pour établir les différents plans d'intervention (stratégique, tactique et/ou opérationnel).

Nous présentons dans la Figure 36, les approches de simulation qui nous intéressent et qui sont utilisées pour développer un outil d'aide à la décision permettant aux décideurs d'établir des plans d'intervention à différents niveaux d'abstraction (stratégique, tactique et/ou opérationnelle). En effet, la simulation à base de modèles mathématiques ne peut donner que des résultats à un niveau Macro. Elle pourra aider à établir des directives d'intervention à un niveau stratégique, suite par exemple à des décisions politiques. Parmi les exemples de ces modèles mathématiques, nous avons déjà présenté les *compartment models*, les *patchy models* ou encore les *metapopulation models* (voir section 3.1 du Chapitre III). Ces deux derniers types de modèles utilisent un espace très agrégé qui ne se base pas sur les données *GIS* reflétant la réalité sur le terrain (Figure 36).

Par contre, la géosimulation multi-niveau que nous proposons utilise un environnement géographique virtuel géoréférencé. De plus, elle est caractérisée par plusieurs aspects (p.ex. utiliser des scénarios événementiels, analyser les résultats de la simulation, utiliser des modèles mathématiques ainsi que les données qui les alimentent, etc.). Par ailleurs et étant donné que cette approche devrait produire des simulations à plusieurs niveaux de granularité (p.ex. Macro, Méso et/ou Micro), elle pourra aider non seulement les décideurs politiques pour établir les directives d'intervention à un niveau stratégique, mais également les décideurs tactiques ou opérationnels pour établir des plans d'intervention concrets. D'autre part, la Figure 36 illustre le fait que les systèmes de surveillance ne permettent pas de faire des prévisions de la propagation probable d'une maladie infectieuse afin de déclencher des actions préventives au bon moment. Cependant, ces systèmes sont indispensables à la géosimulation multi-niveau puisqu'ils permettent de calibrer les données alimentant les modèles mathématiques (Figure 36).

2.2 Vue globale des différents aspects qui nous semblent importants

Nous sommes conscients que l'approche de géosimulation multi-niveau de la propagation des maladies infectieuses est un domaine très vaste et complexe, et évidemment, il ne sera pas facile de résoudre tous les problèmes qui s'y rattachent dans le cadre de cette thèse. Nous présentons d'ailleurs les différents aspects qui nous semblent importants et qui représentent une vue d'ensemble de ce qu'il faut faire normalement pour trouver des éléments de réponses à une bonne partie des problèmes mentionnés auparavant. Ceci nous permettra premièrement d'avoir une vision globale d'une solution plus ou moins complète à ces problèmes et deuxièmement de nous situer par rapport à ce que nous pouvons traiter. Par conséquent, nous allons dégager, à partir de cette vue d'ensemble, les aspects qui nous intéressent le plus. De ce fait, certains de ces aspects pourront être des contributions concernant l'avancement des connaissances. D'ailleurs, pour certains de ces aspects, il existe des solutions que nous pouvons déjà appliquer, étant donné qu'elles semblent être satisfaisantes. Pour d'autres, il faut soit leur appliquer des solutions améliorées ou nouvellement conçues.

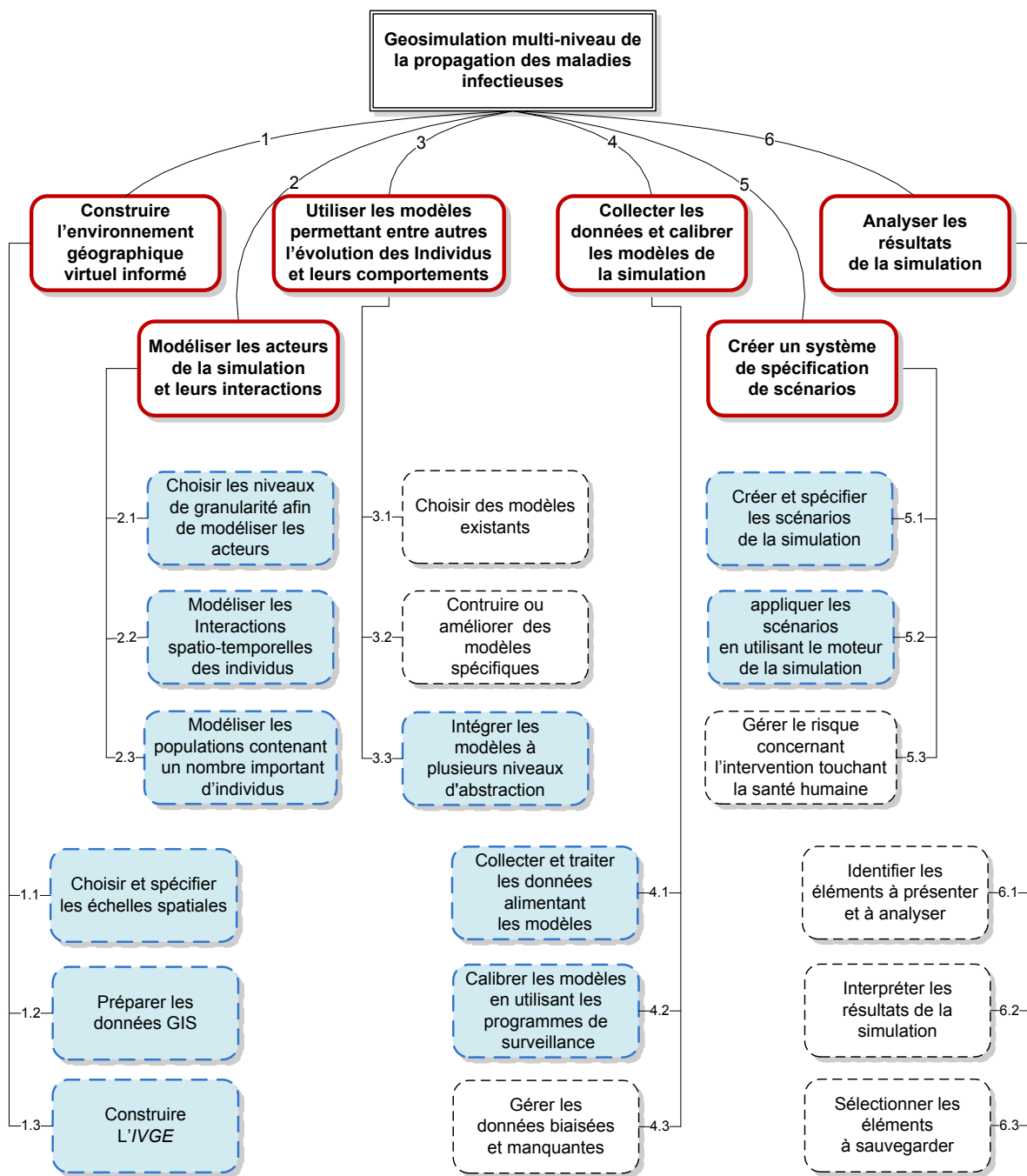


Figure 37: Les différents aspects qui nous semblent importants.

Ainsi, la Figure 38 montre une liste non exhaustive des différents aspects qui, selon nous, doivent être pris en considération afin de résoudre les problèmes liés à la géosimulation multi-niveau de la propagation des maladies infectieuses. Chacun de ces aspects (rectangles avec des lignes continues et rouges et numérotés de 1 à 6) est associé à une liste de solutions (rectangles en pointillés et numérotés de x.1 à x.3, sachant de x est le numéro de l'aspect de plus haut niveau). Ceux qui sont les plus importants et que nous aimerions

traiter sont présentés en bleu et ils sont pleins. Nous détaillons maintenant chacun de ces aspects.

1. **Construire l'environnement géographique virtuel informé (c.-à-d. l'IVGE):** Nous pensons qu'un aspect important pour notre cas est de bien choisir et spécifier les échelles spatiales de l'environnement de la simulation (aspect 1.1 de la Figure 37). Cet environnement devrait utiliser des données SIG (aspect 1.2) qui sont générées pour représenter les connaissances spatiales dont nous aurons besoin (p.ex. limites administratives de la zone géographique concernée, les routes, la couverture du sol, etc.). De ce fait, nous pouvons construire l'environnement géographique virtuel informé (aspect 1.3) dans lequel la simulation peut avoir lieu, et ce, selon différentes échelles spatiales (p.ex. Macro, Méso, Micro). Nous pensons nous inspirer des travaux effectués dans notre laboratoire (GRIC) sur les environnements géoréférencés [Chaker et Moulin, 2008], [Paris *et al.*, 2009], [Mekni et Moulin, 2011]. D'ailleurs et pour que l'outil envisagé soit générique, il faut que l'IVGE soit apte à prendre en considération les besoins spécifiques de la majorité des zoonoses. Par exemple, dans le cas du VNO, nous avons besoin des routes pour calculer une population initiale de moustiques étant donné que ces derniers émergent notamment des puisards. Par contre, pour le cas de la maladie de Lyme, nous avons besoin, entre autres, des forêts pour représenter les habitats des tiques puisque ces dernières vivent généralement dans les régions boisées. D'un autre côté, un environnement en 2D semble être largement suffisant étant donné que les zoonoses se propagent généralement sur un large territoire, et de ce fait une simulation en 3D ne serait pas pertinente, sauf bien sûr dans des cas exceptionnels (p.ex. l'utilisateur aimerait simuler le comportement individuel de certains individus dont le nombre n'est pas très important).
2. **Modéliser les acteurs de la simulation et leurs interactions:** Plusieurs acteurs jouent un rôle important dans la propagation d'une zoonose. Il est donc essentiel de s'intéresser de près à leur modélisation afin de bien les simuler. Premièrement, il s'agit de déterminer quels sont les différents niveaux de granularité (aspect 2.1) qu'il faut utiliser pour modéliser les catégories d'acteurs (individus, groupes ou populations d'humains et/ou d'animaux) qui sont impliqués dans la propagation de la maladie. Deuxièmement, il faut modéliser les interactions spatio-temporelles entre ces individus en tenant compte de l'IVGE (aspect 2.2). Troisièmement, nous avons besoin de modéliser des populations contenant un nombre important d'individus (aspect 2.3). En effet, et selon le niveau de granularité recherché, il faut trouver des alternatives

facilitant la modélisation de ce type de populations, étant donné que les approches actuelles ne permettent pas de le faire, surtout à un niveau de granularité fin.

3. **Utiliser les modèles permettant entre autres l'évolution des individus et leurs comportements:** Nous ne pouvons pas simuler les acteurs impliqués dans la propagation des maladies infectieuses sans concevoir, ou avoir déjà en main, des modèles qui vont prédire le changement dynamique des individus dans l'espace et dans le temps (aspects 3.1 et 3.2). En fait, ce changement est normalement de deux natures. D'une part, nous avons besoin de faire évoluer certains acteurs qui représentent généralement des populations contenant un nombre important d'individus (p.ex. moustiques, tiques, etc.) et ce, selon leurs cycles biologiques. D'autre part, nous avons besoin de simuler les comportements d'autres individus qui peuvent être représentés surtout au niveau individuel ou des groupes d'individus (p.ex. oiseaux, chevreuils, etc.). D'un autre côté, l'intégration de plusieurs modèles (aspect 3.3) semble être une technique prometteuse, surtout si elle permet une combinaison efficace des avantages de chacun d'eux. D'ailleurs, la présence de plusieurs modèles dans une simulation multi-niveau nous laisse penser à plusieurs façons de les utiliser. Nous pouvons par exemple coupler ces modèles pour qu'ils puissent interagir tout en fonctionnant de façon indépendante. Nous pouvons également les intégrer pour qu'ils puissent fonctionner ensemble afin de simuler le même phénomène. Dans ce dernier cas, nous pensons surtout à l'exemple qui a été donné auparavant (voir Section 3.4 du Chapitre II) et qui permet d'intégrer la DS et les SMA dans un même modèle.

4. **Collecter les données et calibrer les modèles de la simulation:** Nous distinguons deux types de données qu'il faut collecter (aspect 4.1). Premièrement, il y a les données permettant d'alimenter les modèles de la simulation. Parmi ces données nous pouvons citer par exemple les paramètres des modèles mathématiques utilisés pour faire évoluer les populations impliquées dans la propagation de la maladie. Il y a également les données relatives aux comportements de certains acteurs comme le regroupement des corneilles dans des dortoirs, la migration des passereaux ou encore le déplacement des chevreuils dans la forêt. De plus, il y a aussi les données relatives aux facteurs influençant les acteurs de la simulation. Les données climatiques comme la température et les précipitations sont un exemple de ce type de données. Deuxièmement, il y a les données permettant de calibrer les modèles de la simulation. Ces données sont habituellement obtenues à partir des systèmes de surveillance (aspect 4.2). D'un autre côté, il faut gérer les données manquantes ainsi que le biais présent dans les données disponibles (aspect 4.3) et qui est engendré par la variabilité spatio-temporelle dans la

collecte effectuée par les systèmes de surveillance. Cette gestion pourrait se faire, entre autres, par des traitements statistiques.

5. **Créer un système de spécification de scénarios:** Les décideurs ont besoin de créer et de spécifier des scénarios (aspect 5.1) qui permettront d'étudier diverses stratégies d'intervention et envisager des alternatives d'évolution de la situation. Les scénarios peuvent être constitués par un ensemble d'événements typiques influençant la propagation de la maladie (p.ex. des fortes pluies, des applications de larvicides, des programmes de sensibilisation ou de mise en quarantaine pour les humains, etc.). Cependant, les utilisateurs pourraient spécifier d'autres paramètres et caractéristiques à incorporer dans les scénarios. Par exemple, les décideurs peuvent choisir le ou les modèles qu'il faut utiliser pour simuler les acteurs de la simulation et leurs comportements. Ils peuvent également modifier les paramètres de ces modèles. Ils peuvent même spécifier certaines caractéristiques de l'environnement géographique virtuel, ou encore choisir le type de données qu'il faut appliquer pour alimenter certains modèles. Par ailleurs, les décideurs peuvent appliquer (aspect 5.2) plusieurs scénarios alternatifs afin de tester, par exemple, plusieurs combinaisons possibles d'événements et voir leurs effets sur le déroulement de la simulation. D'un autre côté, étant donné que la décision d'intervenir sur le terrain suite à l'application de ces scénarios peut toucher la santé humaine, il faut évaluer le risque de prendre une telle décision (aspect 5.3). Du point de vue du degré de risque attaché à la prise de décision, on parle habituellement de décision "certaines", de décisions "aléatoires", et de décisions incertaines. Ce sont surtout les décisions incertaines qui ont un plus grand degré de risque [Vansnick, 1990].
6. **Analyse des résultats de la simulation:** Chaque scénario donnera lieu à une ou plusieurs simulations. Cependant, avant de déclencher la simulation, il faut identifier premièrement les éléments à recueillir pour des fins d'analyse (aspect 6.1). Une analyse des besoins est donc nécessaire à cette étape pour cerner les attentes des décideurs vis-à-vis des résultats de la simulation. Deuxièmement, il faut interpréter les résultats (aspect 6.2). L'interprétation ou l'analyse de ces résultats pourra se faire de plusieurs manières, notamment en utilisant l'approche *SOLAP* [Site Web: *SOLAP*], [Ali *et al.*, 2007] ou encore l'approche multi-critère [Roy, 1985]. Finalement, il faut bien sélectionner les éléments à sauvegarder (aspect 6.3). En effet, les simulations pourront engendrer des quantités importantes de données qui ne devraient pas être toutes nécessairement enregistrées. Dans ce cas, un choix s'impose pour ne pas se retrouver avec une énorme base de données difficilement exploitable.

2.3 Esquisse des différents aspects qui nous intéressent

Comme nous l'avons mentionné, nous ne pourrions pas traiter tous les aspects qui sont mentionnés ci-dessus. Nous pouvons par contre nous attarder à ceux qui nous permettront d'apporter une contribution à la fois théorique et applicative dans notre domaine. Les 4 objectifs spécifiques de notre travail qui ont été présentés précédemment (voir Section 4 du chapitre I) peuvent être atteints si nous trouvons des solutions aux aspects qui nous intéressent. Nous présentons d'ailleurs dans la Figure 38 les liens entre nos objectifs et ces aspects. En effet, nous pensons que les aspects 1.1, 2.1 et 3.3 sont en rapport avec le premier objectif. Les aspects 2.2 et 2.3 sont en rapport avec le deuxième objectif. Les aspects 1.2 et 1.3 sont en rapport avec le troisième objectif. Par contre, les aspects 4.1, 4.2, 5.2 et 5.2 sont en rapport avec le quatrième objectif.

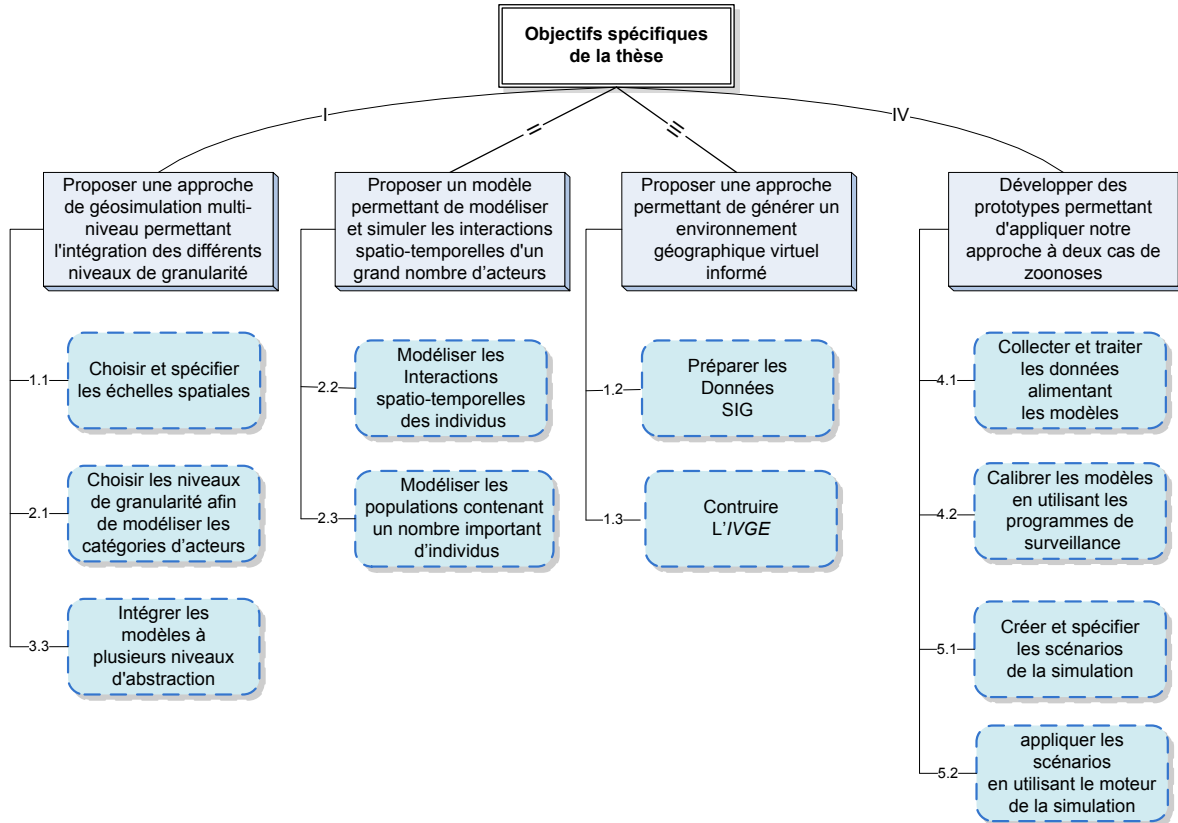


Figure 38: Liens entre les objectifs spécifiques de la thèse et la solution proposée.

3 Modélisation des différents niveaux de granularité

Nous avons déjà abordé la notion de multi-niveau dans notre état de l'art (voir Section 3.6 du chapitre II) et nous avons justement souligné l'importance de développer une approche à

plusieurs niveaux de granularité. D'ailleurs, nous avons besoin de déterminer quelles sont les niveaux de granularité nécessaires pour modéliser un phénomène complexe afin d'en arriver à produire des simulations plausibles. En effet, nous pensons que le choix de ces niveaux dépend de trois principaux facteurs (Figure 39). Il s'agit: (1) des besoins de l'utilisateur vis-à-vis de sa compréhension du phénomène, (2) de la disponibilité des modèles représentant les acteurs et leurs comportements et (3) de la disponibilité des données alimentant les modèles de la simulation.

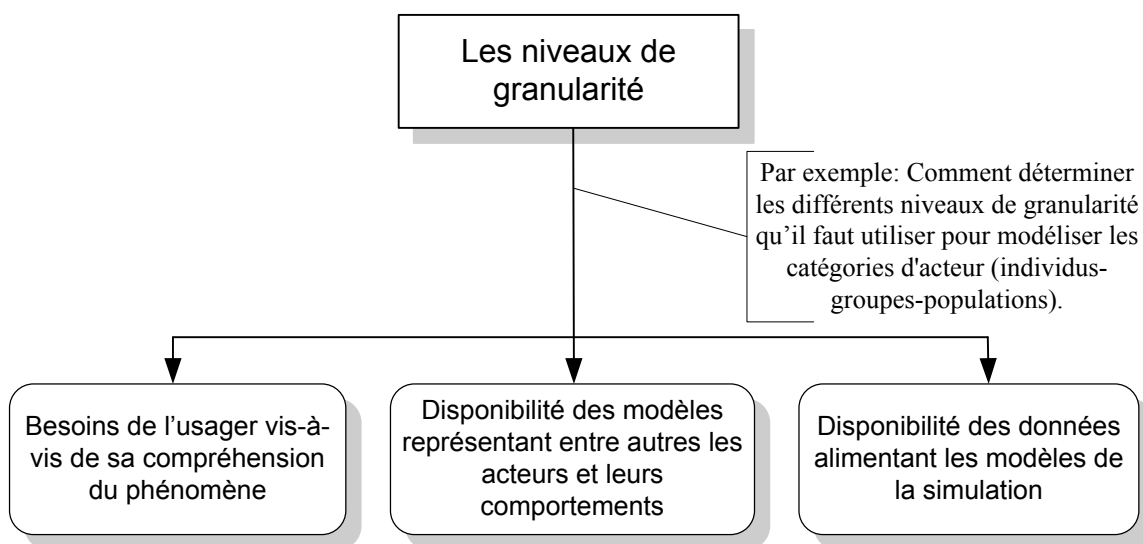


Figure 39: Principaux facteurs influençant le choix des niveaux de granularité.

En outre, nous distinguons plusieurs types de niveaux de granularité. Nous présentons d'ailleurs dans la Figure 40, ces différents types qui sont déterminés à partir des propriétés du phénomène qu'on veut simuler. En effet, un phénomène se produit généralement dans une zone géographique bien déterminée. La variation entre les niveaux est fonction de l'échelle spatiale (ES). Par exemple, la propagation d'une zoonose peut être simulée à l'échelle du monde, d'un continent et/ou d'un pays. De plus, un phénomène est généralement caractérisé par une évolution de plusieurs situations temporelles¹⁵. Dans ce

¹⁵ Une situation temporelle consiste à associer une situation à un intervalle de temps permettant de la localiser sur un axe temporel [Moulin, 1997]. On peut distinguer trois types de situations temporelles: les états, les événements et les processus. Un état est une situation statique exprimant la stabilité et l'absence de changement. Un événement est représenté par un intervalle temporel dont la durée est exprimée par la distance entre les deux bornes qui séparent l'état d'avant l'événement et l'état d'après l'événement. Par contre, un processus exprime un changement avec un événement initial qui indique le début du processus et pourrait avoir un événement final marquant la fin du processus et qui produit un état résultant [Haddad et Moulin, 2008].

cas, la variation entre les niveaux est fonction de l'échelle temporelle (ET). Par exemple, la propagation d'une zoonose peut être simulée en utilisant un pas de simulation d'une journée, d'une semaine et/ou d'une saison. Par ailleurs, un phénomène fait intervenir plusieurs catégories d'acteurs. Dans ce cas, la variation entre les niveaux est fonction de ces catégories (NGCA: Niveau de Granularité des Catégories d'Acteurs). Par exemple, la propagation d'une zoonose peut être simulée au niveau des individus, des groupes et/ou des populations. Finalement, un phénomène peut être simulé en utilisant des modèles à différents niveaux de détail (NGM : Niveau de Granularité des Modèles). Par exemple, un modèle peut avoir une très grande précision afin de représenter des détails importants comme le comportement de chaque individu participant à la simulation. Par contre, un autre modèle peut représenter un phénomène à un niveau de détails macro comme les flux de déplacement des individus (Figure 40).

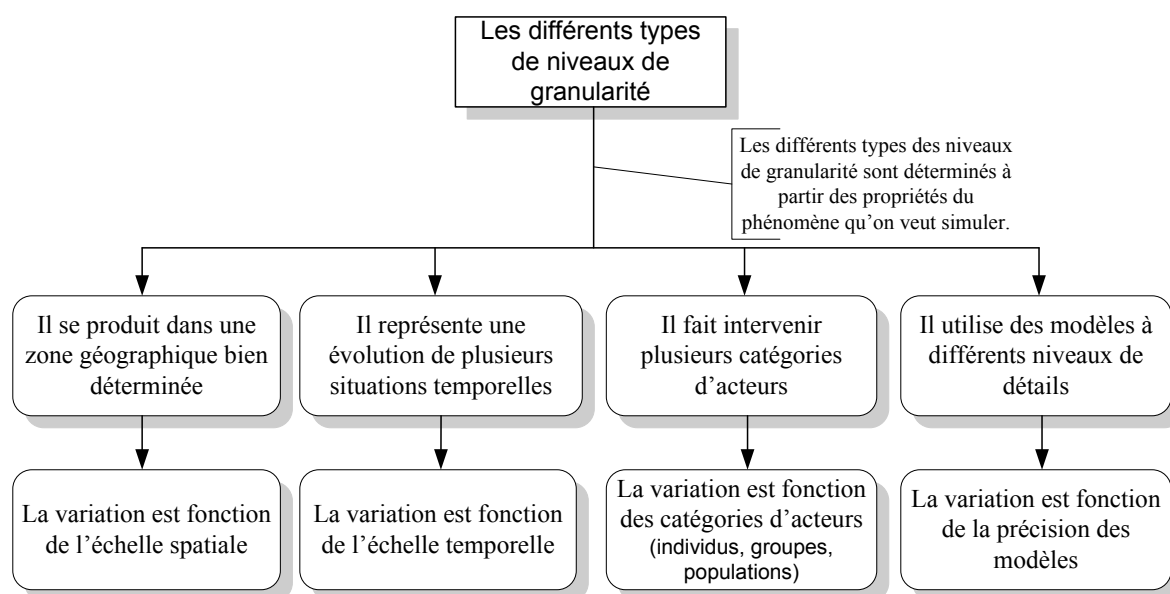


Figure 40: Les différents types de niveaux de granularité.

Ainsi, une simulation d'un phénomène complexe peut être caractérisée par un ou plusieurs ES, ET, NGCA et NGM. Nous présentons d'ailleurs dans la Figure 41, les différentes combinaisons des niveaux de granularité en utilisant trois axes représentant les trois premiers types de ces niveaux (ES, ET et NGCA). Nous savons que les différents niveaux (Macro, Méso et Micro) appartenant à un même type sont déterminés selon les trois principaux facteurs qui ont été déjà mentionnés (besoins de l'utilisateur, disponibilité des modèles et disponibilité des données). De plus, étant donnée une simulation faite à un

niveau bien déterminé d'un type particulier (p.ex. Macro de l'ES), nous pouvons utiliser différents niveaux d'un autre type (p.ex. Macro, Méso et Macro de l'ET). Par exemple, la propagation d'une maladie peut être simulée à l'échelle de la région (p.ex. la province de Québec) en utilisant plusieurs échelles temporelles représentant des pas de simulation différents. D'autre part, nous représentons également dans la Figure 41 une quatrième dimension représentant le choix de niveau de détail des modèles qui sont utilisés dans la simulation. En effet, les modèles peuvent offrir des degrés d'abstraction différents. Par exemple, la propagation d'une zoonose peut être simulée en utilisant un modèle faisant évoluer les acteurs de la simulation et ne montrant que les flux de déplacement de l'infection. Cette propagation peut être également simulée en utilisant un modèle offrant plus de détails et de ce fait, pouvant générer des simulations à un niveau d'abstraction plus fin (p.ex. simulation des différents comportements des acteurs de la simulation comme le regroupement des corneilles en dortoirs). Cet aspect multi-modèle à différents niveaux d'abstraction va être d'ailleurs présenté dans la section 3.3.

En outre, le choix d'une combinaison des 4 niveaux de granularité mentionnés auparavant doit normalement être relié à certaines contraintes afin que le choix d'un niveau de granularité reste compatible avec le choix d'un autre niveau de granularité. Par exemple, le choix d'un niveau macro pour l'échelle spatiale n'est pas nécessairement compatible avec le choix d'un niveau micro pour l'échelle temporelle surtout s'il faut prendre en considération les limites computationnelles. En effet, essayer par exemple d'obtenir les résultats de la simulation d'un phénomène se produisant dans une grande région à toutes les heures pourrait être impossible à réaliser, étant donné le nombre gigantesque de calculs à faire.

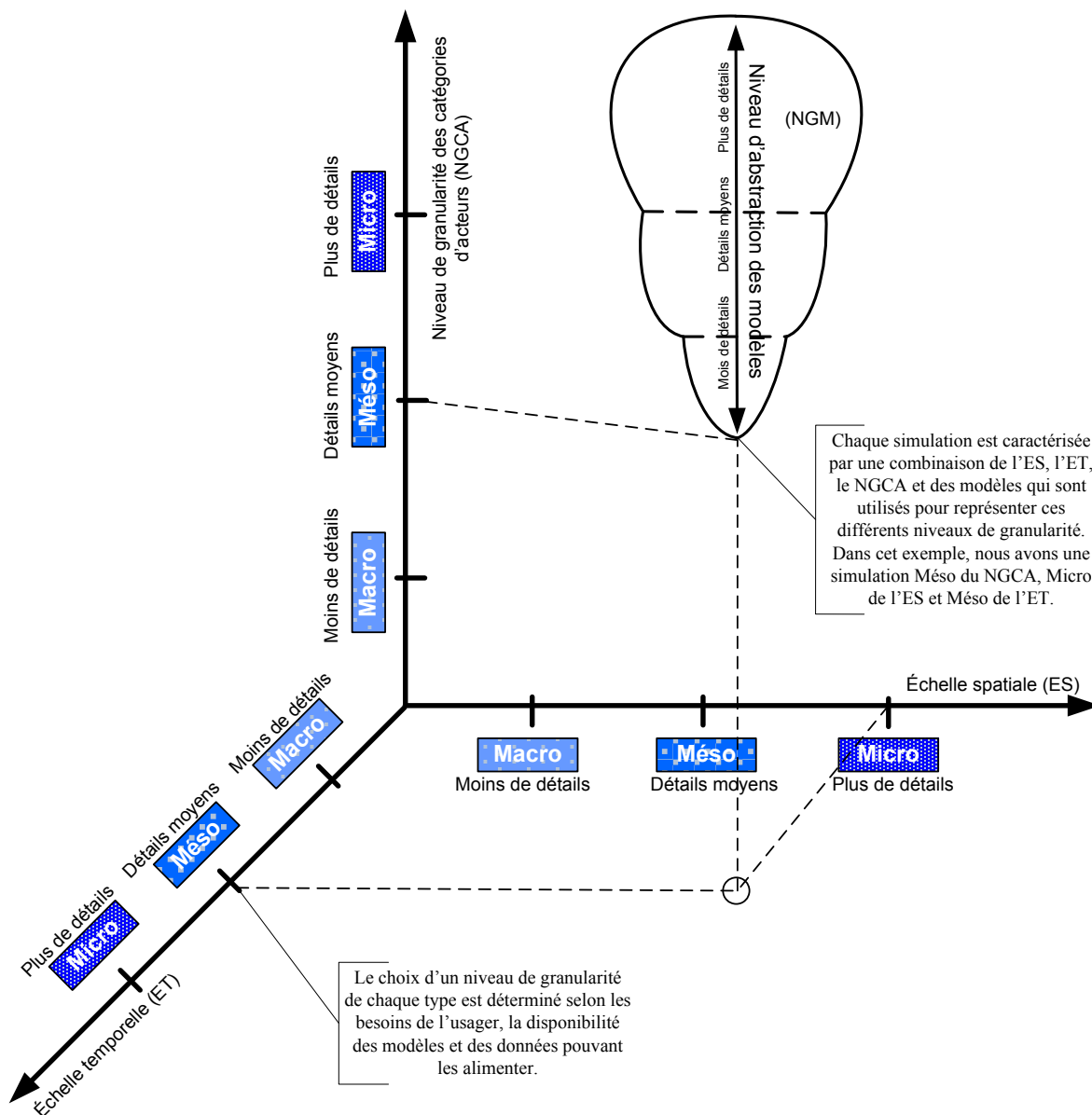


Figure 41: Les différentes combinaisons des niveaux de granularité.

3.1 Modélisation des échelles spatio-temporelles

En s'inspirant des travaux qui ont été effectués en écologie, nous pensons que la propagation d'une maladie infectieuse doit être simulée en utilisant plusieurs échelles spatio-temporelles. En effet, les chercheurs en écologie modélisent la variabilité spatio-temporelle des écosystèmes selon différentes échelles. Non seulement les processus physiques et biologiques, ainsi que la réponse des individus à l'hétérogénéité environnementale changent avec l'échelle, mais les phénomènes corrélés positivement à une échelle peuvent être corrélés négativement à une autre échelle. D'ailleurs, les

écosystèmes sont hétérogènes dans l'espace et dans le temps. Cette hétérogénéité écologique est fondamentale dans la structure et le fonctionnement des écosystèmes. De plus, de nouvelles propriétés écologiques peuvent apparaître à différentes échelles d'observation [Avois-Jacquet, 2002]. De ce fait, les résultats de la simulation d'une maladie infectieuse peuvent être également différents en fonction de l'échelle choisie. Par ailleurs, les échelles spatiales les plus souvent utilisées sont : Macro, Méso et/ou Micro. Par conséquent, nous proposons l'utilisation de ces trois échelles (Figure 42):

- **Une échelle Macro:** La simulation macroscopique représente le phénomène avec une abstraction la plus élevée et elle est utilisée pour simuler un large territoire (p.ex. la province de Québec).
- **Une échelle Méso:** La simulation méso est utilisée pour simuler, avec plus de détail, une région bien déterminée (p.ex. la région de Québec).
- **Une échelle Micro:** La simulation micro est utilisée pour simuler avec encore plus de détails une ville (p.ex. la ville de Québec).

D'un autre côté, les échelles temporelles représentent les pas de simulation choisis pour suivre l'évolution de la maladie. Un pas de simulation peut être une heure, une journée, une semaine, un mois ou même une année selon le niveau de granularité recherché. Nous aimerions d'ailleurs donner un exemple typique de l'influence de l'échelle temporelle sur les résultats de la simulation. En effet, pour le cas du VNO, dérouler la simulation en utilisant un pas de simulation d'une journée, ou à la rigueur d'une semaine, semble satisfaisant pour les besoins des décideurs, surtout que les moustiques qui sont impliqués dans la propagation de la maladie ont un cycle biologique relativement rapide (quelques semaines). Cependant, le cas de la maladie de Lyme est différent surtout que le cycle des tiques responsables de la propagation de cette maladie est d'une durée d'environ 2.5 années (cette durée est donc beaucoup plus longue que celle des moustiques). De ce fait, une simulation en utilisant par exemple un pas de simulation d'une semaine ou d'un mois pourrait être éventuellement prise en considération pour que les décideurs puissent se rendre compte rapidement de l'évolution des populations de tiques. Cela dit, un pas de simulation d'une journée est toujours faisable surtout si l'utilisateur a besoin de plus de précision, mais le temps de calcul pourrait par contre s'accroître rapidement.

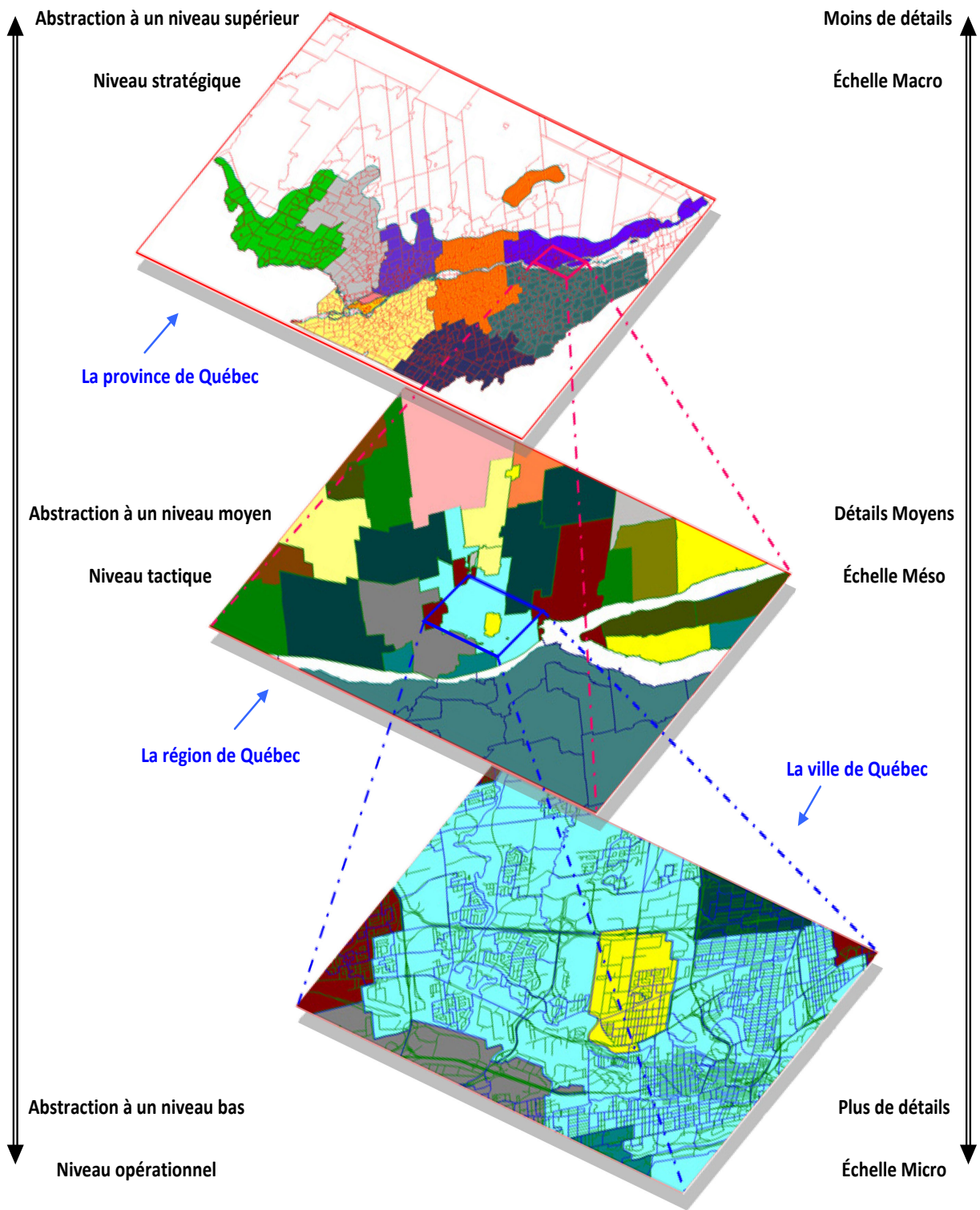


Figure 42: Les différentes échelles spatiales pour représenter l'environnement géographique virtuel.

3.2 Modélisation des différentes catégories d'acteurs participant dans la propagation des zoonoses

Nous définissons un acteur comme étant une entité dont l'action a un effet notable sur le phénomène étudié. Les acteurs sont donc normalement responsables en grande partie de l'évolution de ce phénomène.

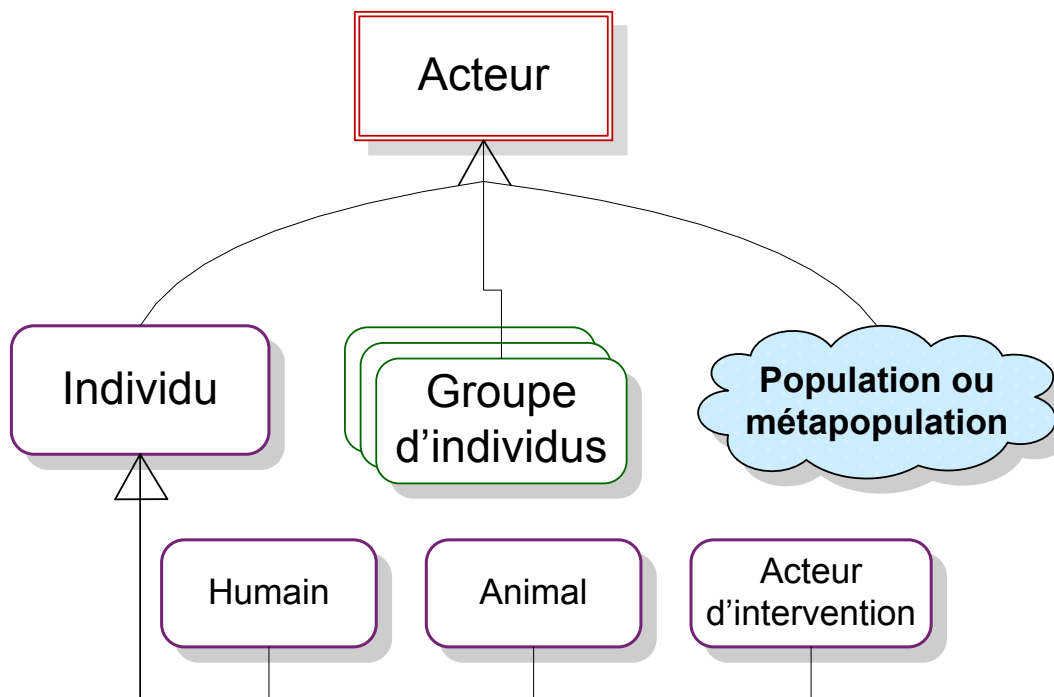


Figure 43: Spécification des acteurs impliqués dans la propagation d'une maladie infectieuse.

Nous envisageons trois catégories d'acteurs (Figure 43) : (1) une population ou un ensemble de populations (métapopulation¹⁶) qui représente un ensemble d'individus qui ont les mêmes caractéristiques et qui vivent généralement dans le même endroit, (2) un groupe d'individus qui représente quelques individus qui ont également les mêmes caractéristiques et (3) un individu qui possède des caractéristiques individuelles et représente dans notre cas un humain, un animal ou un acteur d'intervention. De ce fait, nous proposons trois niveaux de granularité (Figure 44):

¹⁶ Une métapopulation est un groupe de populations d'individus d'une même espèce, séparées spatialement (ou temporellement) et qui réagissent réciproquement à un niveau quelconque [Hanski, 1999].

- La simulation à un niveau de faible granularité afin de représenter les individus avec une abstraction la plus élevée. Elle est employée pour évaluer les dynamiques globales des agrégats d'individus situés sur un vaste territoire. Il s'agit des populations contenant un nombre important d'individus comme c'est le cas pour les moustiques et les tiques. Nous pouvons à ce niveau utiliser les modèles mathématiques pour faire simuler l'évolution de ces populations. Nous pouvons éventuellement utiliser les automates cellulaires pour pouvoir les représenter dans l'environnement géographique. Cependant, nous pouvons également attacher ces populations à des cellules irrégulières si notre environnement géographique virtuel utilise des données vectorielles.

- La simulation à un niveau de granularité moyenne est utilisée pour simuler des groupes d'individus par l'intermédiaire de comportements simples. C'est le cas par exemple des oiseaux ou des chevreuils. Nous pouvons utiliser des agents logiciels ou des particules pour simuler le comportement de ces groupes.

- La simulation à un niveau de granularité fine est utilisée pour simuler des individus caractérisés par des comportements relativement complexes et qui ont une capacité d'interaction et de perception de l'environnement. Nous pouvons également utiliser les agents ou les particules pour simuler ces individus.

Par ailleurs, nous pensons qu'une approche holonique (voir Section 3.7 du chapitre II) serait envisageable pour gérer les différents types d'individus à différents niveaux de granularité. En effet, nous pouvons utiliser une holarchie permettant une agrégation progressive des caractéristiques des acteurs de la simulation et une diminution de la complexité de leurs comportements. Le niveau le plus bas de cette holarchie correspond au niveau Micro où chaque individu est associé à un holon. Les individus sont ensuite regroupés, et chaque groupe est associé à un super-holon dont les membres sont des individus (niveau Méso). Ce processus de composition se poursuit jusqu'à l'obtention au sommet de l'holarchie d'un autre super-holon. Ce dernier représente la population des individus dans sa globalité (niveau Macro). À chaque niveau de cette holarchie, les holons disposent de rôles différents régissant, entre autres, leurs comportements (Figure 45).

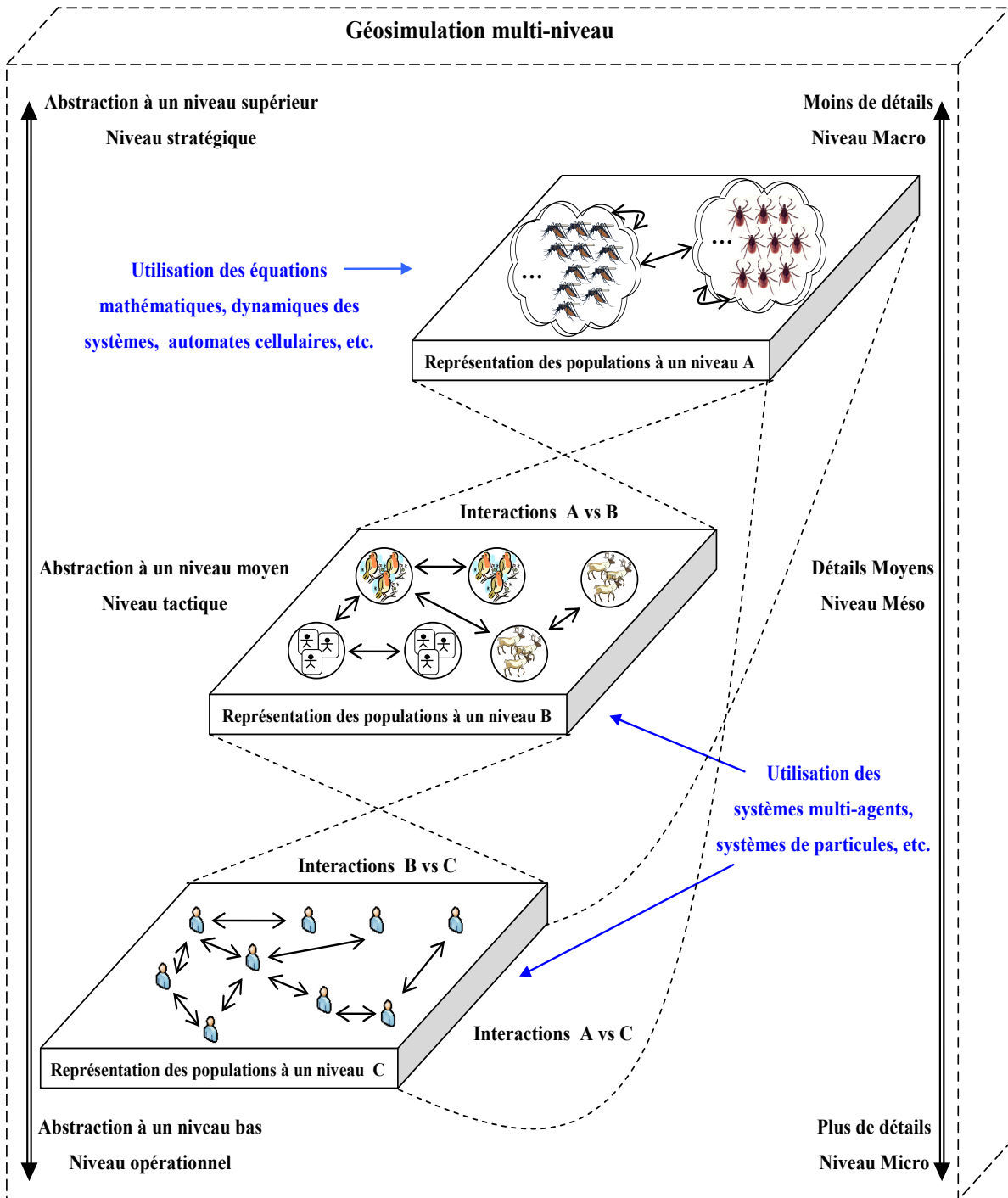


Figure 44: La représentation multi-niveau des différentes catégories d'acteurs.

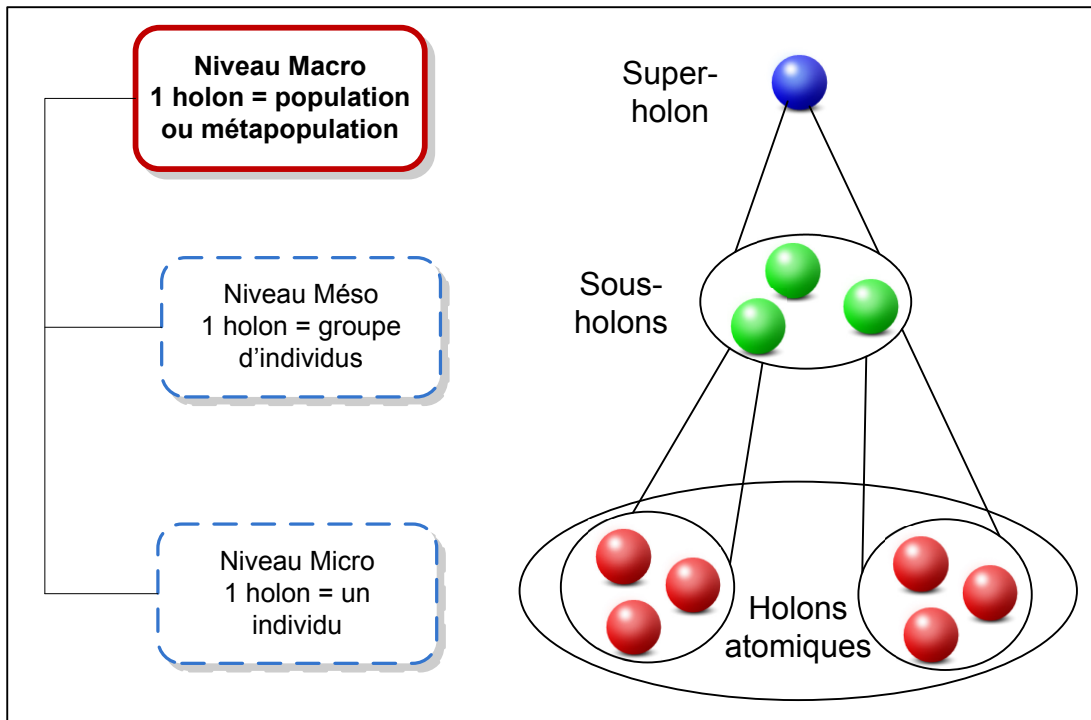


Figure 45: Une holarchie représentant les différentes catégories d'acteurs.

3.2.1 Modélisation des individus et des groupes d'individus

Comme nous venons de le mentionner, la modélisation des individus ou des groupes d'individus peut se faire en utilisant soit un système multi-agent soit un système de particules à condition que le nombre d'individus ne soit pas trop élevé en fonction du niveau de granularité recherché. Les agents ou les particules peuvent être dotés de plusieurs caractéristiques, ainsi que de règles de comportements leur donnant une certaine autonomie. Nous pouvons d'ailleurs bénéficier des modules de *MAGS* (*Agents Module* et *Particules Module*) pour simuler ce type d'acteurs [Bouden, 2004].

Par ailleurs, nous savons que les agents peuvent être réactifs ou cognitifs. Dans le cas où nous avons besoin de représenter les humains, ces derniers seront représentés par des agents cognitifs. Cependant, un agent cognitif peut passer à un état réactif s'il subit par exemple une situation de crise ou un changement important dans son état de santé (infection par une maladie) lui incitant à penser essentiellement à sa survie (instinct animal). Par contre, les animaux ou les groupes d'animaux sont représentés normalement par des agents réactifs. Nous pouvons également les représenter par des particules surtout si le nombre d'individus est élevé. D'autre part, nous pouvons combiner les deux approches. Par exemple, les

dortoirs de corneilles peuvent être modélisés par des agents et chacun de ces agents peut être modélisé avec un système de particules représentant l'ensemble de corneilles. Dans ce cas, une particule peut représenter une corneille ou un groupe de corneilles selon le nombre de corneilles présentes dans le dortoir (voir Section 2 du Chapitre VII).

3.2.2 Modélisation des populations

Nous utilisons une classification qualitative afin de distinguer les différents types de populations selon leurs caractéristiques (quantité et mobilité d'individus). En effet, nous distinguons les populations suivantes:

- **Les populations contenant un nombre important d'individus:** Nous avons mentionné que dans le cas des populations dont le nombre d'individus est considérable, il n'est pas facile de les modéliser en utilisant les approches classiques comme les systèmes de particules ou les systèmes multi-agents. De plus et selon le niveau de granularité recherché, il est souvent inutile, et généralement impossible, de représenter ce genre de population au niveau des individus ou même des groupes d'individus. Ainsi, nous proposons de modéliser ce type de population en leur associant ce que nous appelons "un espace d'occupation". En effet, la population sera liée à une densité d'individus dans cet espace. Par exemple, les moustiques ne se déplacent pas beaucoup et ils sont pratiquement présents partout sur le territoire. Ces moustiques peuvent être considérés comme une caractéristique de l'environnement de la simulation au même titre qu'une caractéristique biophysique comme la végétation. D'ailleurs, *MAGS* est capable justement de faire réagir les acteurs (agents et/ou particules), présents dans la simulation, en fonction de l'environnement dans lequel ils évoluent. D'un autre côté, nous distinguons deux types de populations parmi celles contenant un nombre important d'individus. Il y a des populations peu mobiles comme les moustiques et des populations très mobiles comme les criquets. En effet, ces derniers peuvent parcourir de très grandes distances et se déplacent en très grand nombre. De ce fait, nous proposons de modéliser une population peu mobile en lui associant un espace d'occupation statique (Figure 46a) et une population très mobile en lui associant un espace d'occupation dynamique (Figure 46b).

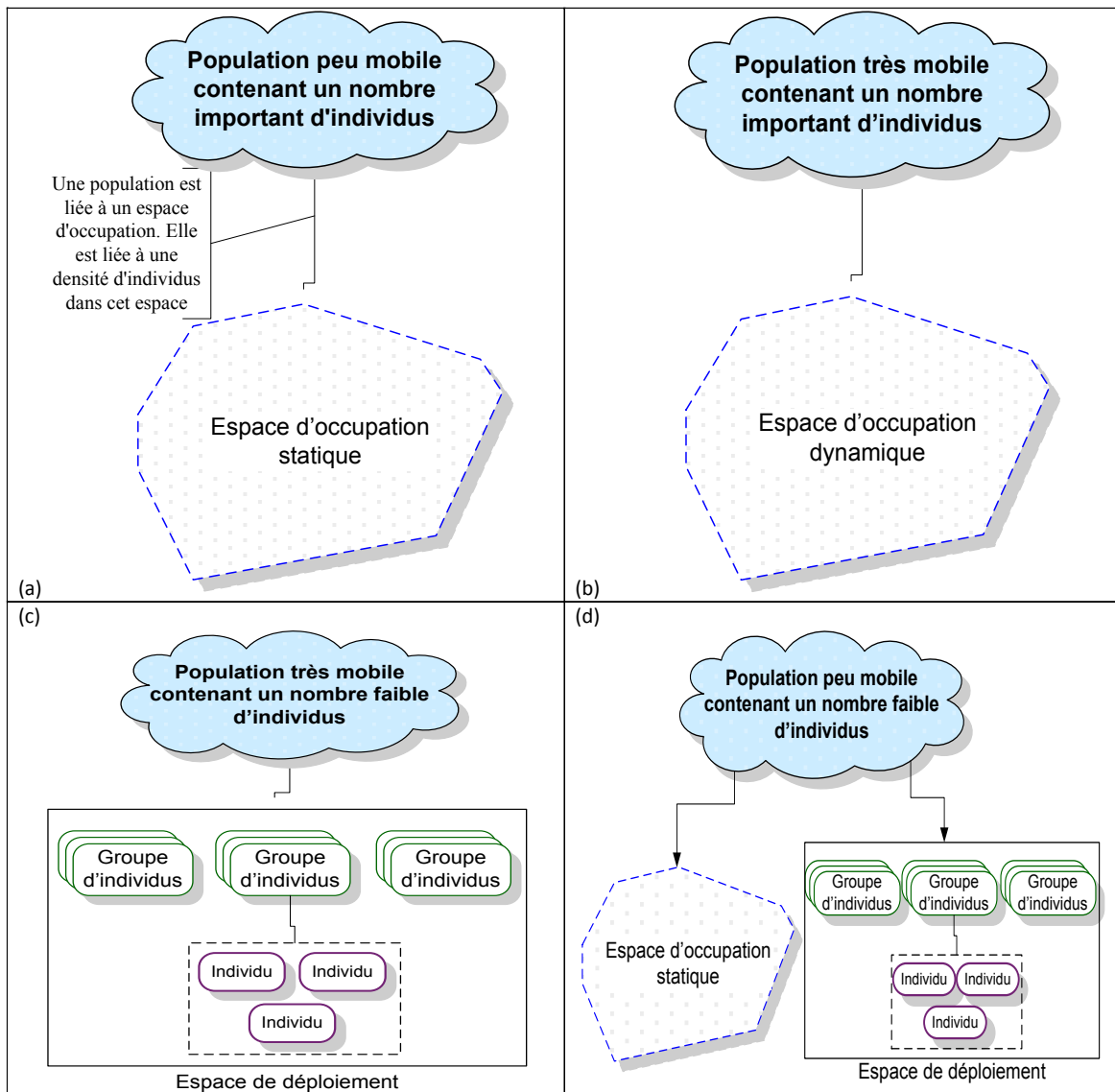


Figure 46: Modélisation différente selon la catégorie de population.

- **Les populations contenant un nombre faible¹⁷ d'individus:** Selon le niveau de granularité recherché, ces populations peuvent être modélisées en les décomposant en groupes d'individus ou même en individus. Dans ce cas, nous proposons d'associer à chacune de ces populations ce que nous appelons "un espace de déploiement" afin de représenter leur mobilité (Figure 46c). Nous pouvons donc utiliser un système multi-agent ou un système de particules pour modéliser ce type de populations. Quand la population est peu mobile, nous proposons de lui associer soit un espace de déploiement

¹⁷ Le terme "faible" utilisé ici veut dire que la population ne contient pas un nombre important d'individus, mais plutôt un nombre relativement petit. Dans ce cas, la population peut être modélisée au niveau des individus et/ou des groupes d'individus.

soit un espace d'occupation statique (Figure 46d) et ce, selon le niveau de granularité recherché.

En outre, nous rappelons que la notion d'agent peut être utilisée pour représenter différents types d'entités dans la simulation comme un objet mobile (p.ex. véhicule, avion, etc) ou encore un être vivant (p.ex. humain, animal). Elle peut également servir à représenter un objet statique (p.ex. bâtiment), une entité physique (p.ex. fumée, gaz, feu, etc) ou encore une entité géographique (p.ex. parcelle, municipalité, etc.). De ce fait, nous pouvons éventuellement modéliser les populations par des agents qui seront considérés comme des entités globales indivisibles.

3.2.3 Modélisation des interactions entre les différents acteurs

Le terme "interaction" tel qu'on peut le trouver dans un dictionnaire, signifie une action réciproque de deux ou plusieurs corps. Il est important donc de définir ce que veut dire une action. En effet, l'action est une opération qui produit un effet sur quelqu'un ou sur quelque chose. Dans le domaine de la simulation multi-agent, il s'agit d'étudier la manière dont les agents peuvent agir, et donc interagir, étant donné une situation et un objectif bien déterminé. L'importance de l'étude de ces interactions découle du fait qu'elles participent étroitement à la dynamique globale du système étudié. Nous distinguons d'ailleurs deux types d'interactions: (1) les interactions entre les acteurs eux-mêmes et (2) les interactions entre les acteurs et l'environnement de la simulation. Un système tel que *MAGS* permet de gérer ce dernier type d'interactions.

En outre, nous proposons dans la Figure 47, un diagramme représentant les multiples interactions entre les différents types d'acteurs (individus, groupes et populations). Nous donnons un sigle aux différentes catégories de populations proposées pour faciliter leur identification dans ce diagramme (Figure 47):

- ✓ Une population peu mobile contenant un nombre important d'individus (*Slow Moving Population with a Large Number of Individuals*): **SMP-LNI**

- ✓ Population peu mobile et contenant un nombre faible d'individus (*Slow Moving Population with a Small Number of Individuals*): **SMP-SNI**
- ✓ Une population très mobile et contenant un nombre important d'individus (*Fast Moving Population with a Large Number of Individuals*): **FMP-LNI**
- ✓ Population très mobile et contenant un nombre faible d'individus (*Fast Moving Population with a Small Number of Individuals*): **FMP-SNI**

La *SMP-LNI* ainsi que dans certains cas (selon les besoins de l'utilisateur) la *SMP-SNI* sont représentées en utilisant un espace d'occupation statique. La *FMP-SNI*, ainsi que dans d'autres cas (selon les besoins de l'utilisateur) la *SMP-SNI*, occupent un espace de déploiement et peuvent être modélisées (selon le niveau de granularité recherché) au niveau des individus et/ou des groupes d'individus appartenant à ces populations. Finalement, la *FMP-LNI* est représentée par un espace d'occupation dynamique. Les interactions entre ces populations vont se traduire par des chevauchements d'espaces. Par exemple, des corneilles à la recherche de leur nourriture, vont se déplacer sur une longue distance (espace de déploiement), vont rencontrer des moustiques se trouvant dans une municipalité (espace d'occupation statique) et vont éventuellement être infectées par les moustiques porteurs du virus ou transmettre la maladie à ces derniers.

D'un autre côté, les interactions entre un individu et une population peuvent se faire quand l'individu traverse (en suivant une certaine trajectoire) un espace d'occupation et/ou un espace de déploiement. Le calcul de la portion de la trajectoire de l'individu incluse dans l'espace occupé par la population permettra de déterminer la période de temps que l'individu est resté en contact avec cette population. De ce fait, nous pouvons refléter l'effet de la présence de cet individu sur le modèle d'évolution ou sur le modèle comportemental de la population selon par exemple une probabilité d'infection. Par ailleurs, les interactions entre un individu et un groupe d'individus peuvent se faire quand l'individu traverse l'espace de déploiement de ce groupe. De plus, les interactions entre une population et un groupe d'individus vont se traduire par des chevauchements d'espaces comme c'est le cas pour les interactions entre populations (Figure 47).

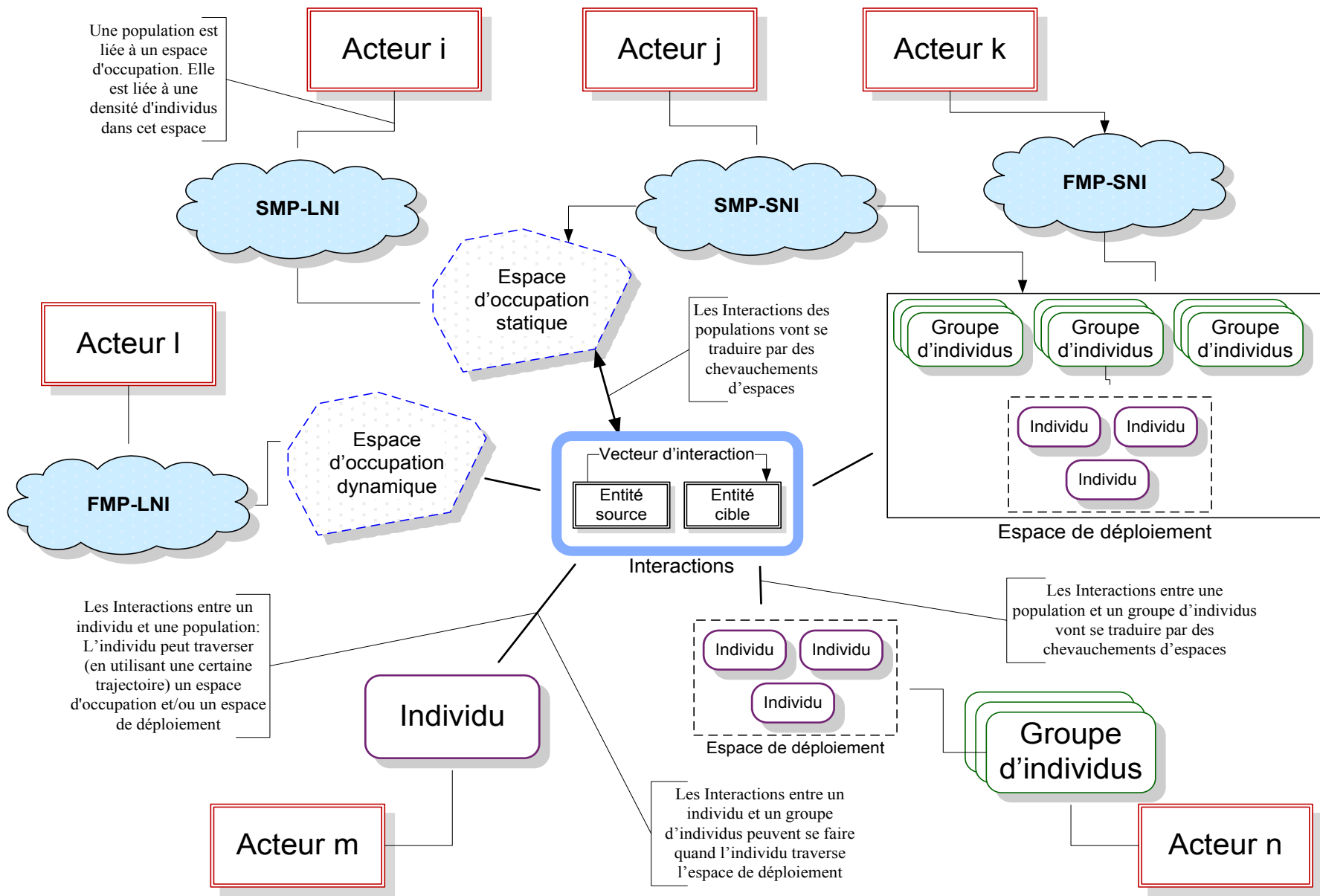


Figure 47: Interactions des différents acteurs.

Ainsi, nous mettons en évidence qu'un problème théorique fondamental abordé dans cette thèse est celui de la modélisation et de la simulation efficace de tous ces types d'interactions spatio-temporelles. En effet, nous pensons que les interactions peuvent être modélisées en combinant des interactions élémentaires. D'ailleurs et suivant le niveau de granularité recherché, les différentes catégories d'acteurs (individus groupes, populations) peuvent être représentées par des agents. Nous aurons alors n agents présents dans la simulation. Chacun de ces agents peut être soit une source, soit une cible d'une interaction élémentaire. De ce fait, une interaction se produit entre une entité source et une entité cible. Une entité peut être par exemple un agent ou un groupe d'agents. Elle peut être aussi une particule ou un groupe de particules. De plus, l'interaction élémentaire modifie normalement la source, la cible ou les deux à la fois. L'action représentant l'interaction entre la cible et la source peut se faire par l'intermédiaire de ce que nous appelons "un vecteur d'interaction" (Figure 47). Ce dernier va transmettre l'effet de l'action de l'entité source à l'entité cible. Cette transmission peut se faire d'une façon discrète (p.ex. un agent pathogène) ou continue (p.ex. flux d'énergie). Nous pouvons d'ailleurs modéliser l'aspect temporel de l'interaction en nous basant sur cette transmission de l'effet d'action. Nous pouvons par exemple utiliser un processus dynamique pour modifier les états initiaux des entités cibles et des entités sources en des états finaux. Cette modification se produira pendant le pas de temps choisi pour la simulation. Par exemple, ce pas peut être une heure, une journée, une semaine, un mois ou même une année, toujours selon le niveau de granularité recherché.

3.3 Niveau de détail des modèles utilisés pour la propagation des zoonoses.

Nous savons que la précision qu'offre une simulation est fonction des modèles qu'elle intègre. Cette précision implique par contre, un coût au niveau de l'implémentation. De ce fait, plus les modèles sont détaillés, plus ils sont difficiles à implémenter et plus ils nécessitent un temps de simulation important. Cependant, l'intégration de plusieurs modèles à différents niveaux d'abstraction permet de créer une synergie entre eux, mais surtout de combiner leurs forces afin d'en arriver à produire un meilleur résultat lors de la

simulation. Ainsi, nous pensons utiliser différents modèles afin de simuler la propagation des maladies infectieuses, et ce, à différents niveaux de granularité.

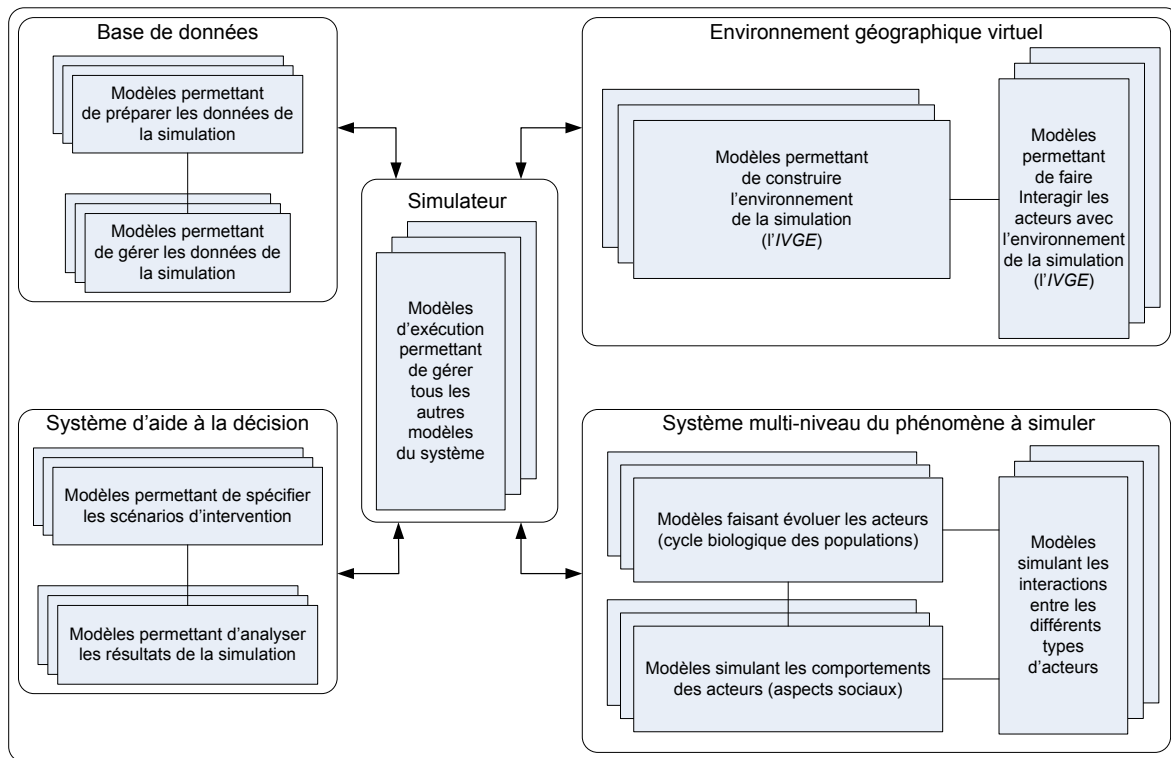


Figure 48: Système multi-modèle à différents niveaux de granularité.

Nous présentons d'ailleurs dans la Figure 48 un système multi-modèle à différents niveaux de granularité. En effet, nous pensons utiliser des modèles d'exécution¹⁸ permettant d'exécuter et éventuellement gérer tous les autres modèles de notre système à différents niveaux de granularité. En fait, c'est le moteur de la simulation (simulateur) qui utilise ces modèles d'exécution qui doivent être adaptés afin d'incorporer les outils nécessaires à la synchronisation et à la transition entre les différents niveaux de granularité. Par ailleurs,

¹⁸ Un modèle d'exécution est un système logiciel permettant l'exécution, sur un ordinateur, d'un programme en utilisant un langage de programmation donné. Il peut y avoir plusieurs modèles d'exécution dans un environnement de programmation (p.ex. contrôleur, ordonnanceur, lanceur, etc.) pouvant chacun exécuter ou gérer l'exécution d'un ou plusieurs autres types de modèles [Gaud *et al.*, 2008].

nous distinguons principalement quatre autres catégories de modèles. Premièrement, il y a les modèles permettant de préparer et gérer les données de la simulation. La précision de ces modèles est toujours en relation avec le niveau de granularité choisi et en fonction surtout de la disponibilité de ces données. Deuxièmement, il y a les modèles permettant de construire l'environnement géographique virtuel, mais aussi de faire interagir les acteurs de la simulation avec cet environnement. D'ailleurs, la simulation exige que le système soit clairement distingué de son environnement afin d'autoriser une gestion indépendante et spécifique des différentes échelles spatio-temporelles de cet environnement. Troisièmement, nous savons qu'une simulation nécessite la création ou l'utilisation de modèles spécifiques pour le phénomène étudié (modèles mathématiques, modèles de comportements, modèle d'interactions, etc.). Ces modèles du phénomène étudié doivent être enrichis afin d'intégrer différents niveaux de granularité. Finalement, nous avons besoin de modèles permettant de spécifier les scénarios d'intervention et d'analyser les résultats de la simulation. Ces modèles sont étroitement liés au système d'aide à la décision. De plus, il faut mentionner que tous les modèles sont normalement inter-reliés puisqu'un modèle à un niveau de granularité peut influencer les résultats d'un autre modèle à un niveau de granularité différent (Figure 48).

4 Discussion et conclusion

Nous avons présenté dans les sections précédentes les différents niveaux de granularité qui sont utilisés dans notre approche. Nous venons par la même occasion de franchir un pas important dans la réalisation des objectifs de notre thèse étant donné que notre premier objectif était de proposer une approche permettant l'intégration des différents niveaux de granularité auxquels une zoonose peut être observée et comprise par les décideurs de la santé publique et selon la disponibilité des données qui alimentent les modèles de simulation. Nous pensons d'ailleurs que l'approche de géosimulation multi-niveau proposée est innovatrice puisque la plupart des approches actuelles (p.ex. les modèles mathématiques, les automates cellulaires, les systèmes multi-agents classiques et la géosimulation multi-agent) n'intègrent pas ces différents niveaux de granularité. De plus, certains systèmes hybrides (voir Section 3.4 du Chapitre II) disposent d'un nombre limité

de niveaux de granularité. Deux niveaux sont très souvent considérés (p.ex. Micro et Marco) comme par exemple pour les travaux de Schieritz et Größler [Schieritz et Größler, 2003] qui ont combiné la dynamique des systèmes et les systèmes multi-agents afin de réduire la complexité dans la gestion de la chaîne logistique. Ces systèmes sont généralement dépendants du domaine d'application étudié, ce qui a pour effet de limiter leur généralité. Cependant, notre approche peut être appliquée non seulement au phénomène de la propagation des zoonoses, mais également à d'autres phénomènes faisant intervenir des interactions spatio-temporelles entre différents acteurs de plusieurs types.

Dans la suite de cette thèse, nous proposons de concrétiser notre approche en commençant par formaliser tous les aspects qui nous intéressent, ce qui nous permettra de bien expliquer ces aspects afin par la suite de les implanter correctement. En fait, le formalisme que nous proposons représente la réalisation du deuxième objectif et notre thèse qui est de proposer un modèle théorique permettant de modéliser et simuler les interactions spatio-temporelles d'un grand nombre d'acteurs de différents types (individus, groupes ou populations d'humains et d'animaux), y compris ceux qui représentent des populations contenant un nombre important d'individus. Ce modèle théorique va être justement présenté dans le chapitre suivant (voir chapitre V). Par la suite, nous proposons dans le chapitre VI une approche pour générer un environnement géographique virtuel informé en tenant compte des spécificités et des contraintes du phénomène étudié qui est la propagation des zoonoses. Ce chapitre permettra donc de répondre au troisième objectif de notre thèse. Finalement et afin d'implanter notre solution, nous proposons dans le chapitre VII deux prototypes illustrant l'application de notre approche à deux cas de zoonoses, afin de montrer sa généralité. Ces prototypes sont des outils d'aide à la décision permettant aux décideurs en santé publique d'évaluer les résultats de plusieurs scénarios alternatifs d'intervention dans le but de comprendre et estimer l'ampleur de l'évolution d'une zoonose. Ces décideurs pourraient éventuellement intervenir sur le terrain après avoir pris des décisions informées et espérons-le, ralentir la progression de ces zoonoses. Ce chapitre permettra donc de répondre au quatrième objectif de notre thèse.

Chapitre V: Modèle théorique *MASTIM*

1 Introduction

Nous proposons dans ce chapitre un nouveau modèle théorique permettant de modéliser et de simuler les interactions spatio-temporelles des acteurs de différents types, y compris ceux qui représentent des populations contenant un nombre important d'individus. Nous avons nommé ce modèle *MASTIM* (*Multi-Actor Spatio-Temporal Interaction Model*) [Bouden et Moulin, 2012c]. Nous commençons par la présentation de ce modèle en spécifiant son utilité par rapport à notre approche de géosimulation multi-niveau qui a été présentée auparavant. Nous présentons par la suite le formalisme de *MATSIM* qui permet de spécifier les concepts importants dont nous avons besoin pour simuler les populations impliquées dans la propagation d'une zoonose. Nous présentons finalement des applications de ce formalisme en utilisant le cas du VNO et celui de la maladie de Lyme.

2 Présentation de *MASTIM*

L'approche de géosimulation multi-niveau présentée dans le chapitre précédent nous a permis d'avoir entre autres une vue d'ensemble des différents aspects qui nous intéressent, dont l'objectif ultime est de développer un outil générique permettant de simuler à différents niveaux de granularité la propagation d'une zoonose. Pour ce faire, nous avons besoin de spécifier l'évolution des populations impliquées dans la propagation de cette zoonose. Nous avons également besoin de spécifier les interactions qui peuvent avoir lieu entre ces populations elles-mêmes, mais également entre ces populations et l'espace géographique représentant l'environnement de la simulation. Cependant, le nombre important d'individus faisant partie de ces populations présente une problématique de modélisation. En effet, les approches de modélisation existantes comme les systèmes multi-agent classiques [Emrich *et al.*, 2007] ne sont pas capables de simuler ces populations étant donné qu'il est difficile de représenter chaque individu par un agent. Les limites des ressources computationnelles des ordinateurs ainsi que le manque des données font en sorte

que nous ne pouvons pas simuler des populations contenant des millions ou même des milliards d'individus en utilisant ce type d'approche. Ceci est d'ailleurs le cas des populations de moustiques impliquées dans la propagation du VNO et des populations de tiques impliquées dans la propagation de la maladie de Lyme. D'un autre côté et dans la littérature, les épidémiologistes et les mathématiciens créent des modèles à base de compartiments (voir Section 3.1.1 du Chapitre III) pour représenter l'évolution des populations impliquées dans la propagation des maladies infectieuses en utilisant un ensemble d'équations différentielles [Wonham *et al.*, 2004] [Ogden *et al.*, 2005]. Par contre, la plupart de ces modèles n'intègrent pas les caractéristiques spatiales du phénomène étudié et ne peuvent pas tenir compte de l'influence des caractéristiques environnementales et géographiques sur les interactions des populations, et de ce fait sur la propagation de la maladie d'une façon générale. De plus, ces modèles impliquent généralement un ensemble de paramètres et des équations que les décideurs de la santé publique ont de la difficulté à bien interpréter pour éventuellement suggérer de les améliorer. Pour surmonter ces difficultés, nous proposons le modèle *MASTIM* permettant de simuler des populations contenant un nombre d'individus très élevé en s'inspirant des modèles compartimentaux pour spécifier l'évolution de ces populations. Toutefois, nous avons ajouté les aspects relatifs aux interactions spatio-temporelles de ces populations, incluant leurs déplacements dans l'environnement de simulation. Nous avons également spécifié dans notre modèle d'autres aspects qui sont indispensables et qui seront présentés dans la section suivante.

3 Formalisme de *MASTIM*

Nous proposons dans cette section le formalisme du modèle théorique *MASTIM* en présentant les notations formelles pour chacun des aspects importants qui nous intéressent. En effet, ce formalisme vise à modéliser entre autres les populations impliquées dans la propagation d'une zoonose à partir d'un point de vue global («le système écologique») en mettant l'accent sur la dimension spatiale des interactions de ces populations. Nous commençons d'ailleurs par la présentation du système écologique et nous présentons par la

suite tous les autres concepts fondamentaux qui nous permettent de développer plus facilement, mais efficacement notre outil d'aide à la décision (voir Chapitre VII).

3.1 Le système écologique

Un système écologique (noté Σ_k) est une partie ou tout un écosystème¹⁹ qui se compose d'organismes vivants qui évoluent et interagissent dans un espace et dans une période bien définis. De ce fait, tous les phénomènes que notre formalisme peut spécifier auront lieu à l'intérieur de ce système écologique. Bien qu'il soit possible pour de nombreux systèmes de coexister, notre formalisme ne traite que des phénomènes qui se passent à l'intérieur d'un système donné. Étant donné que les territoires étudiés sont énormes et que nous devons tenir compte des caractéristiques géographiques qui influencent les cycles biologiques des organismes vivants concernés et leurs comportements, nous proposons de subdiviser l'environnement géographique dans lequel se trouve le système écologique en cellules spatiales élémentaires. Une telle subdivision peut être réalisée à l'aide de données matricielles (c.-à-d. *raster*) ou vectorielles. Dans un système matriciel, l'environnement est divisé en cellules de taille uniforme. Ce genre de subdivision (carrée ou hexagonale) est utilisé généralement par un automate cellulaire [White *et al.*, 2009]. Cependant, la grille utilisée est habituellement considérée comme artificielle et simplificatrice puisqu'elle ne permet pas d'exploiter les caractéristiques spatiales naturelles du phénomène étudié. Nous proposons donc de représenter les cellules du système écologique en utilisant des primitives géométriques telles que des polygones (régions de l'espace de forme irrégulière qui sont bien définies à l'aide des données SIG). Ces cellules pourront être agrégées à différents niveaux de détail dans d'autres cellules à des niveaux hiérarchiques supérieurs. Cette hiérarchie reflète le système multi-niveau que nous avons proposé dans le chapitre précédent. Par ailleurs et dans notre système écologique, le temps est également divisé en étapes discrètes d'une durée sélectionnée et qui représentent les pas de la simulation. De plus, les organismes vivants de notre système écologique sont classés en des espèces

¹⁹ Un écosystème est défini en écologie comme l'ensemble des organismes vivants (plantes, animaux et microorganismes) qui interagissent entre eux et avec le milieu (sol, climat, eau, lumière...) dans lequel ils vivent. C'est donc une subdivision de la biosphère constituée d'un ensemble d'espèces (biocénose) et du milieu (biotope) où il se déploie.

représentant la composante majeure du modèle que nous présentons dans ce chapitre. Chaque espèce a sa propre dynamique d'évolution représentée par un modèle similaire à un modèle enrichi à base de compartiments.

3.2 Les cellules

Une cellule est une région donnée de l'espace (généralement de forme irrégulière) représentant un polygone qui est bien défini à l'aide des données SIG et qui peut être prétraitée afin de refléter les caractéristiques spatiales qui sont importantes pour l'observation et l'analyse du phénomène naturel étudié. Comme nous venons de le mentionner, les cellules de base peuvent être regroupées en cellules de plus haut niveau dans ce qu'on appelle un niveau hiérarchique. Un niveau hiérarchique H_i contient n_i cellules qui sont des agrégats de plusieurs cellules définies au niveau hiérarchique inférieur H_{i-1} . L'ensemble des cellules d'un système \sum_k avec m niveaux hiérarchiques est noté de la façon suivante :

$$C^{\sum_k} = \bigcup_{i=1}^m \{C_1^{H_i}, C_2^{H_i}, C_3^{H_i}, \dots, C_{n_i}^{H_i}\}$$

Une cellule $C_j^{H_i}$ est donc représentée en utilisant un indice j pour bien l'identifier dans un niveau hiérarchique H_i . De ce fait, $C^{\sum_k}(H_i) = \{C_1^{H_i}, C_2^{H_i}, \dots, C_{n_i}^{H_i}\}$ représentent l'ensemble des n_i cellules se trouvant dans le niveau hiérarchique H_i .

Par ailleurs, nous définissons les voisins d'une cellule $C_j^{H_i}$ comme étant un ensemble de cellules se trouvant dans le même niveau hiérarchique H_i et qui ne sont pas nécessairement toutes adjacentes à cette cellule selon les données SIG. Cet ensemble est noté de la façon suivante :

$$V^{C_j^{H_i}} = (\{C_k^{H_i}\}, \text{Constraints})$$

$\{C_k^{H_i}\}$ est donc l'ensemble des cellules voisines qui sont calculées éventuellement en fonction de certaines contraintes reliées par exemple aux caractéristiques de l'habitat

favorable pour les espèces impliquées dans la propagation des zoonoses. Nous pouvons par exemple ne considérer que les cellules voisines qui sont aptes à accueillir une espèce donnée selon la couverture du sol ou les conditions climatiques de ces cellules. Un autre exemple de contraintes est qu'il est possible d'obtenir les voisins d'une cellule selon un angle permettant de ne choisir que des cellules selon une certaine orientation. Nous détaillons ces notions dans le chapitre suivant concernant la génération de l'environnement géographique virtuel informé (voir Chapitre VII).

3.3 Les régions

Une région est une cellule qui est définie à un niveau hiérarchique i :

$$G_j^{H_i} = (Id_{G_j^{H_i}}, Type_{G_j^{H_i}}, C_p^{H_i})$$

$Id_{G_j^{H_i}}$ est l'identificateur ou le nom de la région (exemple : Montérégie). $Type_{G_j^{H_i}}$ est le type de la région (exemple : Région socio-sanitaire : RSS). $C_p^{H_i}$ est la cellule utilisée pour définir la région au niveau hiérarchique i . À un niveau hiérarchique k inférieur à i , nous représentons la région $G_j^{H_i}$ par sa projection dans le niveau k . Cette projection est un conteneur représentant la décomposition de la région $G_j^{H_i}$ (région parent) en un ensemble de régions (région filles) qui sont définies dans le niveau k . Nous notons ce conteneur de la façon suivante :

$$cG_{k < i}^{H_k} = (G_j^{H_i}, \{G_1^{H_k}, G_2^{H_k}, \dots, G_m^{H_k}\})$$

Les régions (parents et filles) sont surtout utilisées pour la décomposition administrative (région socio-sanitaire, municipalité, secteur de recensement, aire de diffusion, etc.). Il est à noter qu'une région peut ne pas être décomposée en d'autres régions (exemple : une municipalité représentant une région rurale qui ne peut pas être décomposée en secteurs de recensement par manque de données), mais qu'elle peut être décomposée en d'autres cellules (exemple : une municipalité décomposée en cellules représentant la couverture du sol). Dans ce cas, le conteneur de la région $G_j^{H_i}$ dans le niveau k est noté de la façon suivante :

$$cG_{k<i}^{H_k} = (G_j^{H_i}, \{C_1^{H_k}, C_2^{H_k}, \dots, C_m^{H_k}\})$$

3.4 Les zones sélectionnées

Une zone sélectionnée est une zone rectangulaire faisant partie de l'environnement de la simulation et que l'utilisateur a sélectionné afin de lui appliquer un certain traitement (exemple : application d'un larvicide, présence de pluies torrentielles, etc.). Nous notons cette zone de la façon suivante :

$$Z_j^{H_i} = (Id_{z_j^{H_i}}, P1_{z_j^{H_i}}, P2_{z_j^{H_i}}, Constraints_{z_j^{H_i}})$$

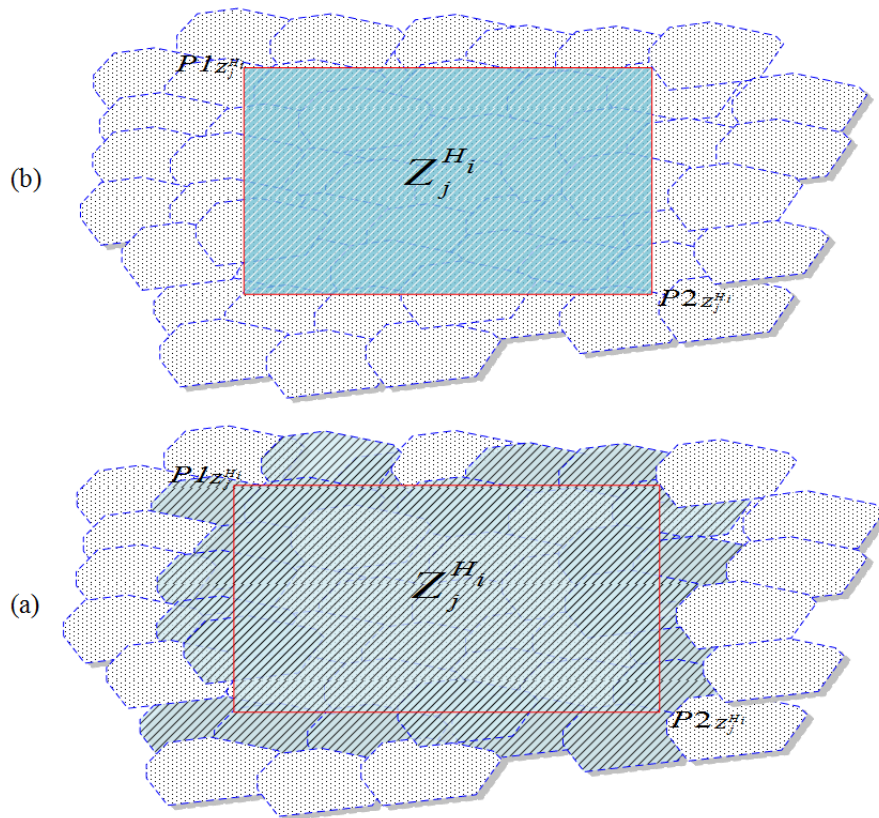


Figure 49: Exemples de zones sélectionnées incluant (a) ou non (b) la totalité des cellules qui touchent les frontières.

$Id_{z_j^{H_i}}$ est l'identificateur ou le nom de la zone sélectionnée. $P1_{z_j^{H_i}} = (X1_{z_j^{H_i}}, Y1_{z_j^{H_i}})$ est le point en haut à gauche de la zone sélectionnée. $X1_{z_j^{H_i}}, Y1_{z_j^{H_i}}$ représente la latitude et la

longitude du point $P1_{Z_j^{H_i}}$. $P2_{Z_j^{H_i}} = (X2_{Z_j^{H_i}}, Y2_{Z_j^{H_i}})$ est le point en bas à droite de la zone sélectionnée. $X2_{Z_j^{H_i}}, Y2_{Z_j^{H_i}}$ représente la latitude et la longitude du point $P2_{Z_j^{H_i}}$. $Constraints_{Z_j^{H_i}}$ représentent les contraintes que l'utilisateur veut appliquer sur la zone sélectionnée. Une contrainte est en fait une expression simple retournant Vrai ou Faux (p.ex. variable booléenne) permettant par exemple d'inclure une partie ou la totalité des cellules qui touchent les frontières de la zone sélectionnée (Figure 49).

3.5 Les espèces

Chaque cellule d'un niveau hiérarchique donné peut accueillir un nombre variable d'individus d'espèces différentes. Les espèces de notre système écologique ne sont pas limitées à un type spécifique et pourraient représenter des organismes comme les mammifères, les oiseaux, les insectes, les virus et les bactéries. La dynamique de l'espèce et son évolution dans une cellule sont modélisées en utilisant une forme étendue d'un modèle à base de compartiments. Ce modèle est unique pour l'espèce, mais il est instancié pour chaque cellule où des individus de l'espèce peuvent vivre et croître. L'ensemble des espèces définies pour un système écologique \sum_k s'écrit de la façon suivante:

$$S^{\sum_k} = \{S_1, S_2, \dots, S_n\}$$

L'indice i de S_i est un identificateur unique pour l'espèce dans le système \sum_k . Par ailleurs, nous définissons l'habitat favorable pour une espèce donnée comme étant un endroit où des conditions appropriées pour l'alimentation, la croissance et la reproduction de cette espèce sont remplies. Dans notre modèle, un habitat favorable est surtout représenté par les contraintes sur la couverture du sol ainsi que sur le climat. Éventuellement, la présence de la nourriture pourrait s'ajouter à ces contraintes. L'habitat favorable s'écrit de la façon suivante :

$$B^{S_k} = \left\{ \begin{array}{l} \{(LType_1, \dots, LType_n, MinBSD), \dots, (LType_1, \dots, LType_m, MaxBSD)\} \\ MinTemp, MaxTemp \\ MinPrecipitations, MaxPrecipitations \end{array} \right\}$$

B^{S_k} représente l'habitat favorable de l'espèce S_k . Les contraintes sur la couverture du sol sont un ensemble de tuples contenant plusieurs valeurs. Les premières valeurs d'un tuple ($LCtype_1, \dots, LCtype_n$) représentent des types de couverture du sol selon les données SIG. En fait, chaque type est représenté par un numéro unique (p.ex. 50 pour les arbustes, 200 pour les forêts, etc.) qui provient par exemple de la base de données géospatiales canadiennes Géobase [Site Web: *Geobase*]. Par contre, la dernière valeur ($MinBSD$: *lower value of the basic suitability degree*) est le degré de favorabilité de base représentant un attribut de l'espèce dans une cellule (voir Section 3.7) et qui est un pourcentage représentant la favorabilité de ces différents types de couverture par rapport à l'espèce. Il est déterminé soit selon la littérature soit selon les avis des experts. Nous détaillons cette notion de favorabilité dans le chapitre suivant concernant la génération de l'IVGE (voir Chapitre VII). Sinon, $MinTemp$, $MaxTemp$ représentent les deux températures extrêmes que peut supporter l'espèce. De la même façon, $MinPrecipitations$ et $MaxPrecipitations$ représentent respectivement le minimum et le maximum de précipitations (en mm) que devront tomber pour que le milieu soit favorable à l'espèce.

3.6 Les compartiments

Nous utilisons dans notre formalisme une forme étendue d'un modèle compartimental. En effet, les différentes étapes de la vie d'une espèce donnée sont représentées par un ensemble de compartiments. Ces étapes représentent en réalité l'évolution biologique de l'espèce comme le passage des stades juvéniles aux stades adultes. Ils représentent également le changement du statut de l'espèce par rapport à son état d'infection (c.-à-d. susceptible, infecté, rétabli) si bien sûr nous ajoutons à notre modèle compartimental l'effet de la maladie infectieuse. L'ensemble des compartiments types²⁰ d'une espèce S_k donnée est noté:

$$O(S_k) = \{O_1^{S_k}, O_2^{S_k}, \dots, O_n^{S_k}\}$$

²⁰ Un compartiment type peut être instancié afin de le cloner dans chacune des cellules de l'environnement de simulation. Par exemple, le compartiment des œufs des moustiques est unique, mais les instances de ce compartiment sont bien sûr multiples.

L'indice i de $O_i^{S_k}$ est un identificateur unique pour le compartiment de l'espèce S_k . En outre, nous pouvons avoir un regroupement de compartiments pour la même espèce qu'on note: $O_{1..m}^{S_k}$. Nous aurons besoin de cette notion de regroupement de compartiments lors de la représentation graphique de nos modèles (voir Section 4). De plus, nous introduisons une notation permettant de regrouper les compartiments de plusieurs espèces et qui est définie comme suit: $O_{1..m}^{S_{1..n-\{p,..,k\}}}$. Dans ce cas, $S_{1..n-\{p,..,k\}}$ veut dire que les espèces concernées sont autres que S_p à S_k . Par ailleurs, les individus se trouvant dans un compartiment peuvent avoir des origines ou des comportements différents. De ce fait, nous introduisons la notion de réservoirs²¹ (c.-à-d. *stocks*) qui peuvent être utilisés pour spécifier des subdivisions de la population dans un compartiment représentant des cohortes²². Par exemple, si l'on veut garder la trace des individus dans un compartiment qui ont des origines différentes (p.ex. les oiseaux migrateurs par rapport aux oiseaux résidents) ou un degré de maturation différent (p.ex. faire évoluer progressivement la maturation des larves pour se transformer par la suite en nymphes), un réservoir peut être créé pour chaque catégorie. L'ensemble des réservoirs d'un compartiment $O_p^{S_k}$ est noté comme suit: $R(O_p^{S_k}) = \{R_1^{O_p^{S_k}}, R_2^{O_p^{S_k}}, \dots, R_m^{O_p^{S_k}}\}$. L'indice i de $R_i^{O_p^{S_k}}$ est un identificateur unique pour le réservoir du compartiment $O_p^{S_k}$.

3.7 Les attributs.

Nous proposons de stocker les données de la simulation dans des attributs qui nous permettront de suivre l'évolution des espèces se trouvant dans les cellules, mais aussi de calculer les interactions spatio-temporelles de ces espèces. Nous distinguons des attributs globaux et des attributs locaux. En effet, un attribut global est une propriété typique du

²¹ C'est en fait une notion qui se rapproche de celle utilisée en dynamique des systèmes où les réservoirs sont des récipients dont le niveau de liquide peut changer en fonction de l'évolution dynamique du modèle. À ne pas confondre donc avec l'espèce-réservoir du pathogène en épidémiologie.

²² Une cohorte désigne un ensemble d'individus ayant vécu un même événement au cours d'une même période. Pour la naissance, on utilise plus le mot de génération. Par exemple, la génération 2013 est la cohorte des individus nés en 2013.

système écologique comme c'est le cas par exemple du pas de simulation. Cet attribut système est noté de la façon suivante :

$$A_k^{\Sigma_k} = (Id_{A_k^{\Sigma_k}}, V_{A_k^{\Sigma_k}})$$

$Id_{A_k^{\Sigma_k}}$ est l'identificateur de l'attribut pour le système écologique Σ_k et $V_{A_k^{\Sigma_k}}$ est la valeur de cet attribut. Par ailleurs, nous distinguons plusieurs types d'attributs locaux. Il y a les attributs qui sont en rapport avec la cellule elle-même (une propriété typique d'une cellule) et qui n'ont rien avoir avec les espèces qui se trouvent dans cette cellule, comme par exemple la superficie et les caractéristiques géo-climatiques de la cellule (p.ex. couverture du sol, température, précipitations, etc.). Nous proposons donc d'attacher à chacune des cellules du système écologique un ensemble d'attributs représentant les informations mentionnées auparavant. Un attribut d'une cellule est noté de la façon suivante :

$$A_k^{C_j^{H_i}} = (Id_{A_k^{C_j^{H_i}}}, V_{A_k^{C_j^{H_i}}})$$

$Id_{A_k^{C_j^{H_i}}}$ est l'identificateur de l'attribut pour la cellule $C_j^{H_i}$ et $V_{A_k^{C_j^{H_i}}}$ est la valeur de cet attribut. L'autre type d'attributs locaux est celui en rapport avec l'espèce se trouvant dans une cellule (p.ex. nom de l'espèce, nombre des individus, etc.). Cet attribut peut être défini selon la notation suivante :

$$A_k^{C_j^{H_i} S_m} = (Id_{A_k^{C_j^{H_i} S_m}}, V_{A_k^{C_j^{H_i} S_m}})$$

$Id_{A_k^{C_j^{H_i} S_m}}$ est l'identificateur de l'attribut pour l'espèce S_j se trouvant dans la cellule $C_j^{H_i}$ et $V_{A_k^{C_j^{H_i} S_m}}$ est la valeur de cet attribut. Nous introduisons également des attributs par rapport aux compartiments se trouvant dans une cellule. Ce type d'attributs peut être défini selon la notation suivante :

$$A_k^{C_j^{H_i} O_n^{S_m}} = (Id_{A_k^{C_j^{H_i} O_n^{S_m}}}, V_{A_k^{C_j^{H_i} O_n^{S_m}}})$$

$Id_{A_k^{C_j^{H_i} O_n^{S_m}}}$ est l'identificateur de l'attribut pour le compartiment $O_n^{S_m}$ de l'espèce S_m se trouvant dans la cellule $C_j^{H_i}$ et $V_{A_k^{C_j^{H_i} O_n^{S_m}}}$ est la valeur de cet attribut. Par ailleurs. Les attributs d'une région $G_j^{H_i}$ au niveau hiérarchique i (noté $A_k^{G_j^{H_i}}$) sont les mêmes que les

attributs de la cellule $C_p^{H_i}$ utilisée pour définir cette région. $G_j^{H_i}$ peut avoir d'autres attributs à des niveaux hiérarchiques inférieurs à i . Ces attributs sont en fait des agrégations d'attributs des régions filles. De plus, une zone sélectionnée peut avoir également des attributs (notés $A_k^{Z_j^{H_i}}$). Certains de ces attributs peuvent être des agrégations d'attributs des cellules appartenant à cette zone.

En outre, nous pouvons citer deux autres exemples d'attributs locaux qui dépendent de l'espèce se trouvant dans une cellule et qui nous semblent importants. Il s'agit premièrement de la capacité d'accueil d'une cellule pour une espèce donnée et qui représente le nombre maximum d'individus de cette espèce que la cellule peut accueillir. Sa valeur est calculée en tenant compte de l'habitat favorable de l'espèce ainsi que de la couverture du sol de la cellule. Cette valeur peut-être ajustée en fonction de la densité représentant la présence habituelle de l'espèce dans la cellule ou dans une région qu'englobe la cellule en question. Le deuxième exemple est le coefficient d'attraction qui est un paramètre déterminant le pourcentage d'individus d'une espèce pouvant être attirés par une cellule²³. La valeur du coefficient d'attraction est calculée en utilisant la capacité d'accueil de la cellule. Donc, plus la capacité d'accueil est grande plus le coefficient est élevé. Ce coefficient est l'équivalent d'un taux de distribution permettant de bien répartir les individus en fonction de la disponibilité des places dans les cellules. Nous discuterons de ces deux exemples d'attributs dans le chapitre suivant concernant la distribution des individus en utilisant une trajectoire.

3.8 Les transitions

Une transition lie généralement deux compartiments (d'une même espèce pouvant se trouver dans la même cellule ou dans deux cellules différentes) et indique qu'un groupe d'individus dans un compartiment source peuvent entrer dans un compartiment destination si un ou plusieurs événements déclencheurs se produisent suite à la vérification des

²³ Une cellule contenant des conditions favorables à la survie et à la reproduction d'une espèce peut attirer plus d'individus de cette espèce qu'une autre cellule contenant des conditions non favorables (p.ex. peu de nourriture, climat ou couverture du sol non propice à l'évolution de l'espèce, etc.).

certaines conditions (voir Section 3.9). Nous distinguons trois types de transitions : (1) les transitions d'évolution, (2) les transitions de mortalité et (3) les transitions de transfert.

3.8.1 Les transitions d'évolution

Une transition d'évolution (*ET* : *Evolutionary Transition*) permet de faire passer un groupe d'individus d'une espèce d'un compartiment à un autre se trouvant dans la même cellule. Une espèce S_k possède un ensemble de transitions d'évolution types que nous notons: $ET(S_k) = \{ET_{i,j}^{S_k}, ET_{p,l}^{S_k}, \dots, ET_{g,c}^{S_k}\}$. En fait, une transition d'évolution type $ET_{s,d}^{S_k}$ permet de faire le passage d'un compartiment type $O_s^{S_k}$ à un autre compartiment type $O_d^{S_k}$ sachant que s et d sont compris entre 1 et n (n étant le nombre de compartiments types possibles et pertinents pour l'espèce S_k). D'ailleurs $ET(S_k)$ est un sous-ensemble de la matrice de transitions suivante :

Compartiment Type Destination Compartiment Type Source	$O_1^{S_k}$	$O_2^{S_k}$..	$O_n^{S_k}$
$O_1^{S_k}$	$ET_{1,1}^{S_k}$	$ET_{1,2}^{S_k}$..	$ET_{1,n}^{S_k}$
$O_2^{S_k}$	$ET_{2,1}^{S_k}$	$ET_{2,2}^{S_k}$..	$ET_{2,n}^{S_k}$
⋮	⋮	⋮	⋮	⋮
$O_n^{S_k}$	$ET_{n,1}^{S_k}$	$ET_{n,2}^{S_k}$..	$ET_{n,n}^{S_k}$

La pertinence d'un compartiment ou d'une transition dépend de l'espèce étudiée, mais aussi du choix du concepteur du modèle selon par exemple les hypothèses qu'il a considérées (voir Section 4 pour des exemples de sous-ensemble de matrice de transitions). En outre, nous considérons qu'une transition peut être déclenchée par un ou plusieurs événements et/ou par une ou plusieurs conditions (voir Section 3.9). Les conséquences de ce déclenchement se traduisent par une ou plusieurs actions (ou activités) qui vont faire passer les individus d'un compartiment à un autre. Techniquement, un événement ou une condition est spécifié par une expression qui retourne *Vrai* ou *Faux*. Les conditions sont en réalité associées à des événements et sont utilisées pour spécifier les combinaisons de

facteurs (valeurs des attributs) qui doivent être vérifiées pour déclencher une action, avant ou après que l'événement se soit produit. Par ailleurs, la transition d'évolution est notée comme suit :

$$ET_{s,d}^{Sp} = (O_s^{Sk}, O_d^{Sk}, \{E_i\}, \{N_j\}, FR_{s,d}^{ET_{s,d}^{Sp}})$$

O_s^{Sk} et O_d^{Sk} représentent respectivement le compartiment source et le compartiment destination dans la même cellule. $\{E_i\}$ et $\{N_j\}$ représentent respectivement un ensemble d'événements (qui peut contenir zéro ou plusieurs événements.) et un ensemble de conditions. Cet ensemble peut être vide si par exemple une transition peut se déclencher quand un événement s'est produit et qu'il n'y a aucune condition à évaluer. Par contre, $FR_{s,d}^{ET_{s,d}^{Sp}}$ représente des débits (c.-à-d. *flow rates*) d'absorption et de production entre les réservoirs appartenant aux deux compartiments qui interviennent dans la transition d'évolution. En effet, l'exécution d'une transition affecte ses réservoirs sources (dans son compartiment source) et ses réservoirs de destination (dans son compartiment de destination). La transition absorbe un nombre d'individus dans les réservoirs sources selon le calcul de ce que nous appelons un processus d'absorption. Il se produit alors un certain nombre d'individus qui seront ajoutés aux réservoirs de destination tel que calculé par un processus de production qui prend comme paramètres le nombre d'individus qui ont été absorbés dans les réservoirs sources. D'ailleurs $FR_{s,d}^{ET_{s,d}^{Sp}}$ est un sous-ensemble des débits suivants :

Réservoirs du compartiment source O_s^{Sk}	$R_1^{O_s^{Sk}}$	$R_2^{O_s^{Sk}}$..	$R_n^{O_s^{Sk}}$
Réservoirs du compartiment destination O_d^{Sk}	$R_1^{O_d^{Sk}}$	$R_2^{O_d^{Sk}}$		$R_n^{O_d^{Sk}}$
$FR_{s,d}^{ET_{s,d}^{Sp}}$	$FR_{R_1^{O_s^{Sk}} R_1^{O_d^{Sk}}}$	$FR_{R_2^{O_s^{Sk}} R_2^{O_d^{Sk}}}$..	$FR_{R_n^{O_s^{Sk}} R_n^{O_d^{Sk}}}$

Sachant que $FR_{R_1^{O_s^{Sk}} R_1^{O_d^{Sk}}}$ définit l'absorption des individus du réservoir source $R_1^{O_s^{Sk}}$ et la production des individus dans le réservoir destination $R_1^{O_d^{Sk}}$ et il est noté comme suit :

$$FR^{R_1^{O_s^k} R_1^{O_d^k}} = \{Absorption(R_1^{O_s^k}, \alpha_1, \beta_1, \dots, \omega_1), Production(R_1^{O_d^k}, \alpha_2, \beta_2, \dots, \omega_2)\}.$$

Les paramètres $\alpha_x, \beta_x, \dots, \omega_x$ représentent des taux d'absorption et de production. L'utilisation de plusieurs taux d'absorption et de production peut être pratique pour définir par exemple un taux différent pour les interactions qui ont comme conséquence la même transition d'évolution (voir Section 3.10).

3.8.2 Les transitions de mortalité

La mortalité est une sorte de transition qui permet de réduire le nombre d'individus se trouvant dans un compartiment. Cette mortalité est représentée graphiquement en utilisant une flèche partant d'un compartiment et n'allant nulle part. L'extrémité de cette flèche se termine par une barre représentant le fait que les individus sont morts et sont donc résorbés du système (voir Section 4). Voici comment cette transition de mortalité est spécifiée:

$$MT^{O_m^{S_p}} = (O_k^{S_p}, \{E_i\}, \{N_j\}, FR^{MT^{O_m^{S_p}}})$$

Ce type de transition est défini de la même façon qu'une transition d'évolution. La différence est l'absence d'un compartiment de destination. Par ailleurs, plusieurs facteurs peuvent provoquer la mortalité des individus. Nous pensons à l'exemple des chevreuils qui se frottent contre des arbres pour se débarrasser des tiques accrochées à eux (c.-à-d. comportement de « *grooming* »). Un autre exemple est l'application de larvicides ou le lessivage des puisards provoquant une grande mortalité chez les larves de moustiques.

3.8.3 Les transitions de transfert

Une transition de transfert (*CT: Transition between Cells*) permet de faire passer des individus d'un compartiment d'une espèce S_p se trouvant dans une cellule source $C_s^{H_i}$ à un autre compartiment de même type de la même espèce, mais se trouvant dans une autre cellule $C_d^{H_i}$ qui est la cellule de destination. Nous notons cette transition comme suit :

$$CT_{s,d}^{O_m^{S_p}} = (O_k^{S_p C_s^{H_i}}, O_k^{S_p C_d^{H_i}}, \{E_i\}, \{N_j\}, FR^{CT_{s,d}^{O_m^{S_p}}})$$

Ce type de transition est défini de la même façon qu’une transition d’évolution. La seule différence est dans les caractéristiques des compartiments impliqués dans la transition. En effet, ils sont de même type, mais appartenant à deux cellules différentes (voir Figure F1 de l’Annexe F).

3.9 Les événements et les conditions

Un événement est généré par la satisfaction (c.-à-d. passage de Faux à Vrai) d’une expression booléenne représentant une condition sur des valeurs d’attributs. Il s’agit donc d’attendre qu’une ou plusieurs conditions soient satisfaites afin de déclencher un événement. De ce fait, l’événement peut se produire pendant la simulation suite au changement de l’état du système par une modification de certains de ses attributs. Nous distinguons d’ailleurs deux types d’événements. (1) un événement biologique ou (2) un événement temporel. Un événement n’a pas de durée, mais plutôt un moment où il se produit. Il peut déclencher une ou plusieurs transitions, mais aussi un ou plusieurs processus (voir Section 3.11) qui vont modifier l’état du système écologique. Nous n’avons pas besoin de spécifier les processus ou les transitions dans la notation de l’événement étant donné que ces informations sont déjà présentes dans chacun des processus et des transitions. Par contre, un événement peut être répétitif (survenant plusieurs fois pendant une certaine durée). Nous notons d’ailleurs un événement de la façon suivante :

$$E_i = (Id_{E_i}, D_i, \{N_k\}, [D_f, Periodicity])$$

Id_{E_i} est le nom et/ou l’identificateur de l’événement. D_i est la date de début de l’événement. $\{N_k\}$ est l’ensemble des conditions qui vont faire en sorte que l’événement se déclenche. Il ne faut pas confondre ces conditions avec les conditions d’une transition. En effet, ces dernières vont être évaluées après le déclenchement de l’événement et non pas avant. D_f est optionnel et représente la date de fin des événements répétitifs alors que $Periodicity$ représente le nombre de jours qui séparent deux instances de l’événement répétitif.

En outre, une condition N_i est une expression booléenne qui est constituée d'un ensemble d'expressions reliées entre elles par des opérateurs logiques (ET, OU). Chacune des expressions est composée d'opérandes et d'opérateurs de comparaison (<, >, <=, >=, !=). Chacun des opérandes est composé de paramètres et d'opérateurs arithmétiques. Les paramètres peuvent être soit des valeurs soit des attributs. Nous avons utilisé une grammaire *BNF* pour spécifier différentes sortes de conditions [Sedrati, 2013]. Voici d'ailleurs des exemples de conditions :

Condition	Commentaire
$value(A_k^{\Sigma_k}) \geq 5 \text{ min}$	Cette condition peut être utilisée par exemple pour déclencher un événement après 5 min du début de la simulation
$value(A_k^{C_j^{H_i}}) > 120$	Cette condition peut être utilisée par exemple pour déclencher la maturation des larves étant donné que l'attribut de la cellule concernant les degrés-jours cumulés a dépassé 120. Ce chiffre représente le nombre de jours pour lesquels on a mesuré une température au dessus d'un certain seuil qui serait favorable à l'activité biologique des moustiques.
$value(A_1^{C_j^{H_i} O_n^{S_m}}) \leq 1000$ and $value(A_2^{C_j^{H_i} O_n^{S_m}}) > 500$	Cette condition peut être utilisée par exemple comme une des conditions d'une transition qu'il faut évaluer suite au déclenchement d'un événement.
$value(A_1^{C_j^{H_i} O_n^{S_m}}) \leq value(A_2^{C_j^{H_i} O_n^{S_m}})$	On peut également avoir une condition faisant intervenir une comparaison entre attributs de même type.

3.10 Les interactions

Les interactions spatio-temporelles entre les espèces sont très fréquentes dans la nature. Dans notre formalisme, nous considérons qu'une espèce S_i située dans la même cellule qu'une autre espèce S_j peut déclencher une ou plusieurs transitions de S_j , en tenant compte d'une probabilité d'interaction. Les interactions entre les espèces sont un aspect important

pour les transitions étant donné que très souvent la transition d'un compartiment à un autre dépend de la probabilité que l'espèce concernée puisse trouver des hôtes. Cependant, ce n'est pas toutes les transitions qui nécessitent une interaction. Dans notre cas, nous considérons les interactions entre deux espèces qui dépendent l'une de l'autre. En effet, nous nous intéressons à une espèce hôte qui permet à une autre espèce d'évoluer en prenant par exemple un repas de sang (comme dans le cas des moustiques et des tiques). Par ailleurs, nous distinguons deux types d'interactions, compte tenu de leurs conséquences possibles: (1) les interactions instantanées et (2) les interactions durables. En effet, une interaction instantanée peut déclencher des transitions d'évolution. Par exemple, les individus d'une espèce peuvent pondre des œufs après un repas de sang (p.ex. lorsque les moustiques piquent des oiseaux) et donc peuvent devenir infectés si l'espèce hôte est un réservoir de l'agent pathogène. En revanche, les interactions durables peuvent éventuellement déclencher des transitions de transfert. Ces interactions peuvent par exemple conduire à l'accrochage de certains individus à une espèce hôte (p.ex. lorsque les tiques s'accrochent aux chevreuils), permettant le transfert de ces individus d'une cellule à une autre en raison de la mobilité de l'espèce hôte. La liste des interactions pertinentes entre les différents compartiments types de deux espèces est notée comme suit:

$$Interaction_{S_k S_p} = \{I_{O_i^{S_k} O_j^{S_p}}, I_{O_p^{S_k} O_l^{S_p}}, \dots, I_{O_g^{S_k} O_c^{S_p}}\}$$

Rappelons que les compartiments types des deux espèces sont représentés de la façon suivante :

$$O(S_k) = \{O_1^{S_k}, O_2^{S_k}, \dots, O_n^{S_k}\}$$

$$O(S_p) = \{O_1^{S_p}, O_2^{S_p}, \dots, O_m^{S_p}\}$$

$Interaction_{S_k S_p}$ est un sous-ensemble d'une matrice à deux dimensions. La première dimension représente le nombre de compartiments appartenant à l'espèce S_k et la deuxième dimension représente le nombre de compartiments appartenant à l'espèce S_p :

Compartiment Type S_p				
Compartiment Type S_k	$O_1^{S_p}$	$O_2^{S_p}$..	$O_m^{S_p}$
$O_1^{S_k}$	$I_{O_1^{S_k} O_1^{S_p}}$	$I_{O_1^{S_k} O_2^{S_p}}$..	$I_{O_1^{S_k} O_m^{S_p}}$
$O_2^{S_k}$	$I_{O_2^{S_k} O_1^{S_p}}$	$I_{O_2^{S_k} O_2^{S_p}}$..	$I_{O_2^{S_k} O_m^{S_p}}$
:	:	:	:	:
$O_n^{S_k}$	$I_{O_n^{S_k} O_1^{S_p}}$	$I_{O_n^{S_k} O_2^{S_p}}$..	$I_{O_n^{S_k} O_m^{S_p}}$

Voici ce que représente par exemple l'interaction entre les deux premiers compartiments des deux espèces S_k et S_p :

$$I_{O_1^{S_k} O_1^{S_p}} = (Id_{I_{O_1^{S_k} O_1^{S_p}}}, ProbI_{O_1^{S_k} O_1^{S_p}}, [\{ET_{1,j}^{S_k}, \dots, ET_{1,c}^{S_k}\}, \{ET_{1,v}^{S_p}, \dots, ET_{1,z}^{S_p}\}], [Drop_{O_1^{S_k} O_1^{S_p}}, \{CT_{s,d}^{O_1^{S_k}}, CT_{s,d}^{O_1^{S_p}}\}])$$

$Id_{I_{O_1^{S_k} O_1^{S_p}}}$ représente le nom et/ou l'identificateur de l'interaction entre les deux compartiments. $ProbI_{O_1^{S_k} O_1^{S_p}}$ représente la probabilité d'interaction qui peut se produire entre les individus du premier compartiment type de l'espèce S_k avec les individus du premier compartiment type de l'espèce S_p . Cette probabilité peut être nulle s'il n'y a pas d'interaction entre les individus de ces deux compartiments. Les deux autres termes (entre crochets) sont optionnels et représentent les conséquences de l'interaction. Le premier terme représente deux possibilités de conséquences d'une interaction instantanée. La première possibilité (notée $\{ET_{1,j}^{S_k}, \dots, ET_{1,c}^{S_k}\}$) peut se traduire par l'activation ou non des transitions d'évolution qui partent du compartiment $O_1^{S_k}$ vers les autres compartiments de l'espèce S_k . Sachant que $\{ET_{1,j}^{S_k}, \dots, ET_{1,c}^{S_k}\}$ est un sous-ensemble de la matrice suivante :

Compartiment Type Destination				
Compartiment Type Source	$O_1^{S_k}$	$O_2^{S_k}$..	$O_n^{S_k}$
$O_1^{S_k}$	$ET_{1,1}^{S_k}$	$ET_{1,2}^{S_k}$..	$ET_{1,n}^{S_k}$

La deuxième possibilité de conséquences (notée $\{ET_{1,v}^{S_p}, \dots, ET_{1,z}^{S_p}\}$) peut se traduire par l'activation ou non des transitions d'évolution qui partent du compartiment $O_1^{S_p}$ vers les autres compartiments de l'espèce S_p . Sachant que $\{ET_{1,v}^{S_p}, \dots, ET_{1,z}^{S_p}\}$ est un sous-ensemble de la matrice suivante :

Compartiment Type Destination Compartiment Type Source	$O_1^{S_p}$	$O_2^{S_p}$..	$O_m^{S_p}$
$O_1^{S_p}$	$ET_{1,1}^{S_p}$	$ET_{1,2}^{S_p}$..	$ET_{1,m}^{S_p}$

Par ailleurs, le second terme optionnel représente les conséquences d'une interaction durable. $Drop_{O_1^{S_k} O_1^{S_p}}$ représente une fonction permettant de calculer entre autres la durée moyenne d'accrochage avant que l'espèce transportée (c.-à-d. *Hitchhiker*) se laisse tomber de l'espèce transportrice (c.-à-d. *Carrier*). $\{CT_{s,d}^{O_1^{S_k}}, CT_{s,d}^{O_1^{S_p}}\}$ permet le transfert des individus d'une cellule à une autre. Ce transfert peut entraîner l'activation de deux transitions de transfert. Le premier permet le passage des individus dans le compartiment de l'espèce transportée à un compartiment du même type, mais situé dans une cellule différente. Le second permet le passage des individus à partir du compartiment de l'espèce transportrice à un compartiment du même type, mais situé également dans une cellule différente.

3.11 Les processus

Un processus est un ensemble d'activités corrélées ou interactives qui utilisent des ressources (personnel, équipement, matériels, informations, etc.) pour transformer des éléments entrants en éléments de sortie dont le résultat final attendu est un produit ou un service. Nous pouvons également le définir comme étant une opération complexe exécutable par un ordinateur et définie par un ensemble d'instruction à exécuter (un programme) et un espace d'adressage en mémoire pour lire et/ou stocker les données de

travail. Dans notre cas, nous nous intéressons essentiellement à trois grands types de processus. Un processus naturel, un processus représentant une intervention humaine et un processus système qui est ajouté pour les besoins de la simulation (Figure 50). Il est à noter que les interventions humaines peuvent modifier le processus naturel et vice-versa, le processus naturel peut éventuellement inciter les humains à modifier leurs interventions.

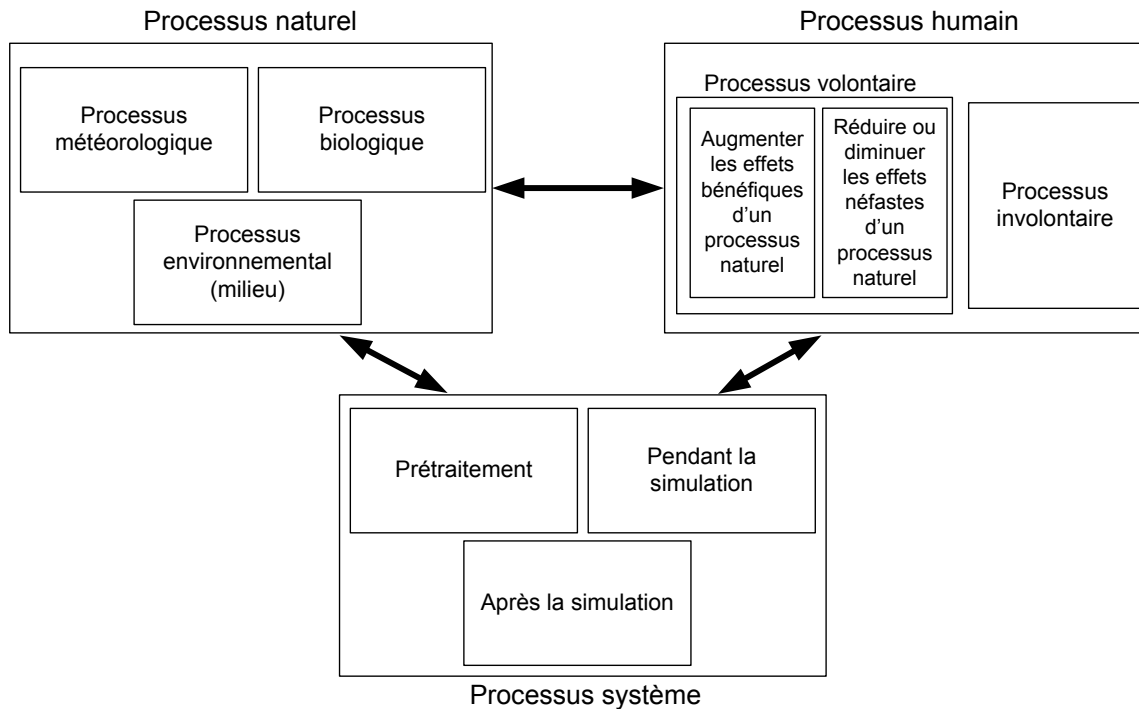


Figure 50: Les différents types de processus qui nous intéressent.

Dans le cadre de notre projet, nous pouvons distinguer trois types de processus naturels. Un processus qui est en rapport avec les phénomènes météorologiques (exemple: pluie torrentielle, vague de chaleur, etc.), un processus qui est en rapport avec les phénomènes biologiques (exemple: migration des oiseaux, etc.) et un processus qui est en rapport avec le milieu (exemple : modification de la couverture du sol, etc.). Nous pouvons appeler ce dernier processus, un processus environnemental.

Concernant le processus représentant une intervention humaine, nous distinguons deux types de processus. En effet, l'humain peut intervenir d'une façon intentionnelle ou volontaire afin de modifier l'effet d'un processus naturel (exemple : application d'un larvicide pour tuer les larves causant la propagation d'une maladie, etc.). Pour ce type de processus volontaire, nous pouvons éventuellement distinguer un processus permettant d'augmenter les effets bénéfiques d'un processus naturel et un processus permettant de réduire ou diminuer les effets néfastes d'un processus naturel. Par contre, l'humain peut intervenir aussi dans un but qui n'a pas de rapport direct avec le processus naturel, mais qui a un effet sur ce processus (exemple : émissions de gaz à effet de serre, etc.). Nous l'appelons un processus involontaire. Concernant le processus système, nous distinguons trois types : (1) Un prétraitement permettant entre autres de préparer les données de la simulation, (2) un processus en cours d'exécution de la simulation comme par exemple la gestion du temps de la simulation et finalement, (3) un processus après la simulation permettant par exemple d'analyser les résultats de la simulation ou encore de sauvegarder ces résultats.

Un processus peut se produire pendant la simulation affectant les cellules se trouvant dans une ou plusieurs régions ou encore dans une ou plusieurs zones sélectionnées. Une région peut être une partie ou la totalité de l'environnement de la simulation. Le processus change l'état du système en modifiant les attributs de la région considérée. Nous notons un processus générique de la façon suivante :

$$P_m = (Id_{P_m}, Type_{P_m}, \{G_j^{H_i}\}, \{Z_s^{H_i}\}, \{E_d\}, \Delta_{P_m}, f_{P_m}, \{g_k(A_j)\})$$

Id_{P_m} est le nom et/ou l'identificateur du processus. $Type_{P_m}$ représente le type du processus. $\{G_j^{H_i}\}$ est l'ensemble des régions où le processus se produit. Si cet ensemble est vide, il faut alors avoir au moins une zone sélectionnée où le processus se produit. $\{Z_s^{H_i}\}$ est un ensemble de zones sélectionnées où le processus se produit. Si cet ensemble est vide, il faut alors avoir au moins une région où le processus se produit. $\{E_d\}$ est l'ensemble des

évènements qui font déclencher le processus. Δ_{P_m} est la durée du processus estimée en nombre de pas de simulation (exemple : heures, jours, semaines, mois, etc.). Cette durée n'est pas nécessairement connue à l'avance, mais qui pourrait être mise à jour pendant ou après l'exécution du processus. f_{P_m} est la fréquence à laquelle le processus se répète en fonction du pas de simulation. $\{g_k(A_j)\}$ est l'ensemble de fonctions permettant de modifier certains attributs A_j à la fréquence f_{P_m} . Le choix de l'attribut A_j parmi $\{A_j^{\Sigma_k}, A_j^{G_m^{H_i}}, A_j^{Z_m^{H_i}}, A_j^{C_m^{H_i}}, A_j^{C_m^{H_i}S_p}, A_j^{C_m^{H_i}O_n^{S_p}}, \dots\}$ est fonction du type de processus.

4 Applications de *MASTIM*

Nous présentons dans cette section l'application de notre modèle *MASTIM* à deux cas différents de zoonoses afin d'illustrer sa généralité.

4.1 Un modèle pour le VNO

Le VNO a déjà été présenté dans le Chapitre III (voir Section 2.2.1). Dans ce cas, le système écologique permettant de simuler la propagation du VNO est noté Σ_1 . Il est constitué d'un ensemble de cellules se trouvant dans plusieurs niveaux hiérarchiques :

$$C^{\Sigma_1} = \bigcup_{i=1}^m \{C_1^{H_i}, C_2^{H_i}, C_3^{H_i}, \dots, C_{n_i}^{H_i}\}$$

Le système contient deux espèces $S^{\Sigma_1} = \{S_1, S_2\}$ qui interagissent entre elles. S_1 représente l'espèce de *Culex* et S_2 représente l'espèce de Corneilles. De ce fait, une cellule $C_j^{H_i}$ d'un niveau hiérarchique i peut contenir ces deux espèces. Pour les *Culex*, nous utilisons 6 compartiments types: $O_1^{S_1}$: Œuf *Culex*, $O_2^{S_1}$: Larve *Culex*, $O_3^{S_1}$: Adulte *Culex* Sain, $O_4^{S_1}$: Adulte *Culex* exposé, $O_5^{S_1}$: Adulte *Culex* infecté, et $O_6^{S_1}$: Adulte *Culex* en diapause.

L'ensemble des compartiments types des *Culex* est donc noté:

$$O(S_1) = \{O_1^{S_1}, O_2^{S_1}, O_3^{S_1}, O_4^{S_1}, O_5^{S_1}, O_6^{S_1}\}$$

Nous avons par contre omis le compartiment type nymphe afin de simplifier le cycle biologique du *Culex* sans pour autant affecter sa plausibilité. La matrice de transition pour les *Culex* est la suivante :

Compartiment Type Destination Compartiment Type Source	$O_1^{S_1}$ (œuf <i>Culex</i>)	$O_2^{S_1}$ (larve <i>Culex</i>)	$O_3^{S_1}$ (adulte sain)	$O_4^{S_1}$ (adulte exposé)	$O_5^{S_1}$ (adulte infecté)	$O_6^{S_1}$ (adulte diapause)
$O_1^{S_1}$ (œuf <i>Culex</i>)	$ET_{1,1}^{S_1}$	$ET_{1,2}^{S_1}$	$ET_{1,3}^{S_1}$	$ET_{1,4}^{S_1}$	$ET_{1,5}^{S_1}$	$ET_{1,6}^{S_1}$
$O_2^{S_1}$ (larve <i>Culex</i>)	$ET_{2,1}^{S_1}$	$ET_{2,2}^{S_1}$	$ET_{2,3}^{S_1}$	$ET_{2,4}^{S_1}$	$ET_{2,5}^{S_1}$	$ET_{2,6}^{S_1}$
$O_3^{S_1}$ (adulte sain)	$ET_{3,1}^{S_1}$	$ET_{3,2}^{S_1}$	$ET_{3,3}^{S_1}$	$ET_{3,4}^{S_1}$	$ET_{3,5}^{S_1}$	$ET_{3,6}^{S_1}$
$O_4^{S_1}$ (adulte exposé)	$ET_{4,1}^{S_1}$	$ET_{4,2}^{S_1}$	$ET_{4,3}^{S_1}$	$ET_{4,4}^{S_1}$	$ET_{4,5}^{S_1}$	$ET_{4,6}^{S_1}$
$O_5^{S_1}$ (adulte infecté)	$ET_{5,1}^{S_1}$	$ET_{5,2}^{S_1}$	$ET_{5,3}^{S_1}$	$ET_{5,4}^{S_1}$	$ET_{5,5}^{S_1}$	$ET_{5,6}^{S_1}$
$O_6^{S_1}$ (adulte diapause)	$ET_{6,1}^{S_1}$	$ET_{6,2}^{S_1}$	$ET_{6,3}^{S_1}$	$ET_{6,4}^{S_1}$	$ET_{6,5}^{S_1}$	$ET_{6,6}^{S_1}$

Les transitions en gris sont pertinentes pour notre modèle, tandis que les transitions barrées ne sont pas pertinentes. Par exemple, la transition $ET_{1,2}^{S_1}$ qui représente le passage d'œufs à larves est pertinente, mais la transition $ET_{2,1}^{S_1}$ qui représente le passage de larves à œufs n'a pas de sens. Nous utilisons dans cet exemple d'application de notre modèle un seul réservoir par compartiment. Par exemple, le compartiment Adulte *Culex* Sain est composé d'un seul réservoir type:

$$R(O_3^{S_1}) = \{R_1^{O_3^{S_1}}\}.$$

En supposant que la transmission verticale est négligeable, l'ensemble des transitions types pertinentes pour notre cas est le suivant (Figure 52):

$$ET(S_1) = \{ET_{1,2}^{S_1}, ET_{2,3}^{S_1}, ET_{3,1}^{S_1}, ET_{3,4}^{S_1}, ET_{3,6}^{S_1}, ET_{4,5}^{S_1}, ET_{5,1}^{S_1}, ET_{5,6}^{S_1}, ET_{6,1}^{S_1}\}$$

- $ET_{1,2}^{S_1}$ = Passage de œuf à larve (éclosion des œufs)

- $ET_{2,3}^{S_1}$ = Passage de *larve* à *adulte sain* (émergence des *Culex*)
- $ET_{3,1}^{S_1}$ = Passage de *adulte sain* à *œuf* (ponte des œufs)
- $ET_{3,4}^{S_1}$ = Passage de *adulte sain* à *adulte exposé* (exposition des *Culex*)
- $ET_{3,6}^{S_1}$ = Passage de *adulte sain* à *adulte en diapause* (dormance des *Culex*)
- $ET_{4,5}^{S_1}$ = Passage de *adulte exposé* à *adulte infecté* (infection des *Culex*)
- $ET_{5,1}^{S_1}$ = Passage de *adulte infecté* à *œuf* (ponte des œufs)
- $ET_{5,6}^{S_1}$ = Passage de *adulte infecté* à *adulte en diapause* (dormance des *Culex*)
- $ET_{6,1}^{S_1}$ = Passage de *adulte en diapause* à *œuf* (ponte des œufs)

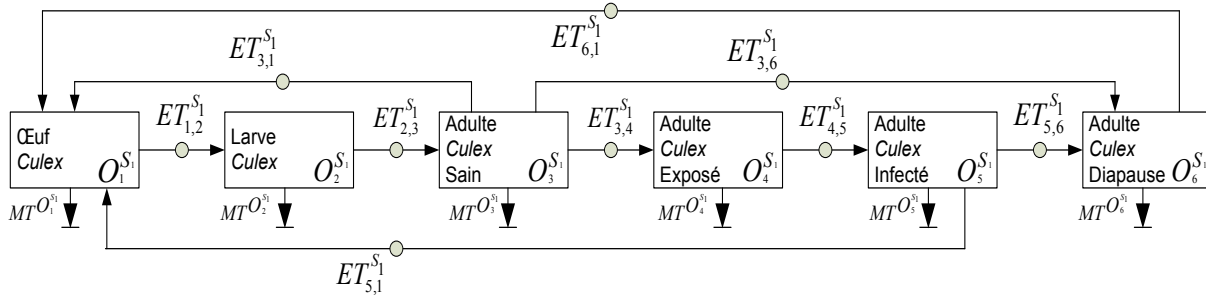


Figure 51: Cycle biologique des *Culex* en tenant compte de l'infection.

Pour les corneilles, nous utilisons 4 compartiments types ($O_1^{S_2}$: Corneille saine, $O_2^{S_2}$: Corneille infectée, $O_3^{S_2}$: Corneille morte, $O_4^{S_2}$: Corneille guérie). L'ensemble des compartiments types des corneilles est donc noté :

$$O(S_2) = \{O_1^{S_2}, O_2^{S_2}, O_3^{S_2}, O_4^{S_2}\}$$

La matrice de transition pour les corneilles est la suivante :

Compartiment Type Destination \ Compartiment Type Source	$O_1^{S_2}$ (Corneille saine)	$O_2^{S_2}$ (Corneille infectée)	$O_3^{S_2}$ (Corneille morte)	$O_4^{S_2}$ (Corneille guérie)
$O_1^{S_2}$ (Corneille saine)	$ET_{1,1}^{S_2}$	$ET_{1,2}^{S_2}$	$ET_{1,3}^{S_2}$	$ET_{1,4}^{S_2}$
$O_2^{S_2}$ (Corneille infectée)	$ET_{2,1}^{S_2}$	$ET_{2,2}^{S_2}$	$ET_{2,3}^{S_2}$	$ET_{2,4}^{S_2}$
$O_3^{S_2}$ (Corneille morte)	$ET_{3,1}^{S_2}$	$ET_{3,2}^{S_2}$	$ET_{3,3}^{S_2}$	$ET_{3,4}^{S_2}$
$O_4^{S_2}$ (Corneille guérie)	$ET_{4,1}^{S_2}$	$ET_{4,2}^{S_2}$	$ET_{4,3}^{S_2}$	$ET_{4,4}^{S_2}$

De ce fait, l'ensemble des transitions types pertinentes pour notre cas est le suivant (Figure 52):

$$ET(S_2) = \{ET_{1,2}^{S_2}, ET_{2,3}^{S_2}, ET_{2,4}^{S_2}\}$$

- $ET_{1,2}^{S_2}$ = Passage de *corneille saine* à *corneille infectée* (infection des corneilles)
- $ET_{2,3}^{S_2}$ = Passage de *corneille infectée* à *corneille morte* (mort des corneilles)
- $ET_{2,4}^{S_2}$ = Passage de *corneille infectée* à *corneille guérie* (guérison des corneilles)

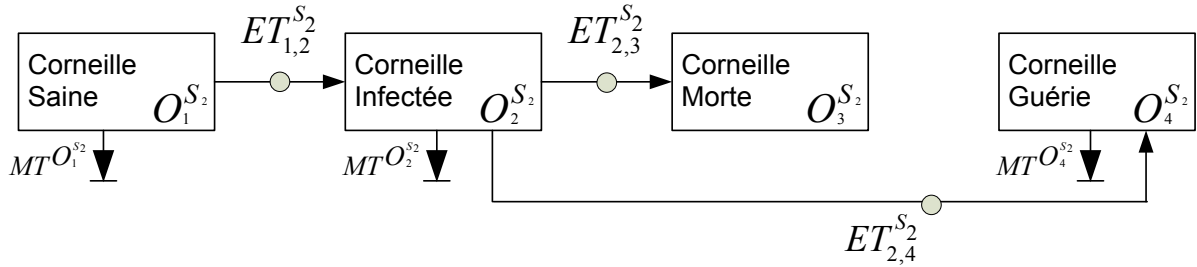


Figure 52: Cycle biologique des corneilles en tenant compte de l'infection.

Voici maintenant la matrice montrant toutes les interactions possibles entre les différents compartiments de ces deux espèces :

Compartiment Type Corneilles \ Compartiment Type Culex	$O_1^{S_2}$ (<i>Corneille saine</i>)	$O_2^{S_2}$ (<i>Corneille infectée</i>)	$O_3^{S_2}$ (<i>Corneille morte</i>)	$O_4^{S_2}$ (<i>Corneille guérie</i>)
$O_1^{S_1}$ (<i>œuf Culex</i>)	$I_{O_1^{S_1} O_1^{S_2}}$	$I_{O_1^{S_1} O_2^{S_2}}$	$I_{O_1^{S_1} O_3^{S_2}}$	$I_{O_1^{S_1} O_4^{S_2}}$
$O_2^{S_1}$ (<i>larve Culex</i>)	$I_{O_2^{S_1} O_1^{S_2}}$	$I_{O_2^{S_1} O_2^{S_2}}$	$I_{O_2^{S_1} O_3^{S_2}}$	$I_{O_2^{S_1} O_4^{S_2}}$
$O_3^{S_1}$ (<i>adulte sain</i>)	$I_{O_3^{S_1} O_1^{S_2}}$	$I_{O_3^{S_1} O_2^{S_2}}$	$I_{O_3^{S_1} O_3^{S_2}}$	$I_{O_3^{S_1} O_4^{S_2}}$
$O_4^{S_1}$ (<i>adulte exposé</i>)	$I_{O_4^{S_1} O_1^{S_2}}$	$I_{O_4^{S_1} O_2^{S_2}}$	$I_{O_4^{S_1} O_3^{S_2}}$	$I_{O_4^{S_1} O_4^{S_2}}$
$O_5^{S_1}$ (<i>adulte infecté</i>)	$I_{O_5^{S_1} O_1^{S_2}}$	$I_{O_5^{S_1} O_2^{S_2}}$	$I_{O_5^{S_1} O_3^{S_2}}$	$I_{O_5^{S_1} O_4^{S_2}}$
$O_6^{S_1}$ (<i>adulte diapause</i>)	$I_{O_6^{S_1} O_1^{S_2}}$	$I_{O_6^{S_1} O_2^{S_2}}$	$I_{O_6^{S_1} O_3^{S_2}}$	$I_{O_6^{S_1} O_4^{S_2}}$

Donc, la liste des interactions pertinentes entre les différents compartiments types de ces deux espèces est :

$$Interaction_{S_1 S_2} = \{I_{O_3^{S_1} O_1^{S_2}}, I_{O_3^{S_1} O_2^{S_2}}, I_{O_3^{S_1} O_4^{S_2}}, I_{O_5^{S_1} O_1^{S_2}}, I_{O_5^{S_1} O_2^{S_2}}, I_{O_5^{S_1} O_4^{S_2}}\}$$

Nous supposons dans ce cas que nous sommes en train de faire une simulation pour une seule saison. De ce fait, la reprise d'activités du *Culex* adulte après sa période de diapause n'est pas prise en compte. C'est pour cette raison que le *Culex* adulte en diapause n'entrera pas en interaction avec les corneilles puisqu'il n'a pas besoin de prendre un repas de sang pour se reproduire pour la saison en cours. De plus, l'adulte exposé a déjà pris un repas de sang et il n'y aura donc pas d'interaction avec les corneilles.

Dans la Figure 53, nous présentons une représentation graphique de l'évolution et des interactions (conformément aux matrices de transitions et d'interactions présentées auparavant) des *Culex* et des corneilles impliquées dans la propagation du VNO. Le cycle biologique de la population de *Culex* incluant les effets de la maladie est représenté par 6 compartiments et 9 transitions d'évolution. En outre, 4 compartiments et 3 transitions d'évolution sont utilisés pour représenter le cycle biologique de la population de corneilles. Chaque rectangle représente un compartiment alors qu'une flèche contenant un petit cercle représente une transition entre les compartiments. Ce cercle agit comme un conteneur qui peut recevoir d'autres flèches (représentées avec des petits points) représentant les conséquences des interactions. De plus, nous utilisons d'autres flèches (représentées en traits interrompus) et d'autres petits cercles (ils sont isolés et donc ne sont pas comme ceux qui sont attachés aux transitions) pour justement représenter les interactions entre les différents compartiments des *Culex* et des corneilles. Par exemple, $I_{O_3^{S_1} O_1^{S_2}}$ (voir en bas de la figure 54) est une interaction qui représente la piqûre d'un groupe de corneilles susceptibles par un groupe de *Culex* adultes infectés. À la suite de cette interaction, les corneilles susceptibles peuvent devenir infectées (c.-à-d. transition $ET_{1,2}^{S_2}$) et les *Culex* adultes infectés peuvent pondre des œufs (c.-à-d. transition $ET_{5,1}^{S_1}$). Par ailleurs et comme nous l'avons déjà mentionné, la transition de mortalité est illustrée à l'aide d'une flèche qui commence à partir d'un compartiment et se termine par une barre qui représente le fait que les individus sont

morts et sont résorbés par le système. De plus, nous utilisons un rectangle avec des coins arrondis afin de représenter globalement des espèces autres que les corneilles (noté $S_{l..n(1,2)}$) et qui peuvent être piquées par les *Culex*. Nous pensons que cet ajout est primordial afin d'éviter d'avoir un modèle trop simpliste dans lequel les corneilles sont les seules espèces piquées par les moustiques (Figure 53).

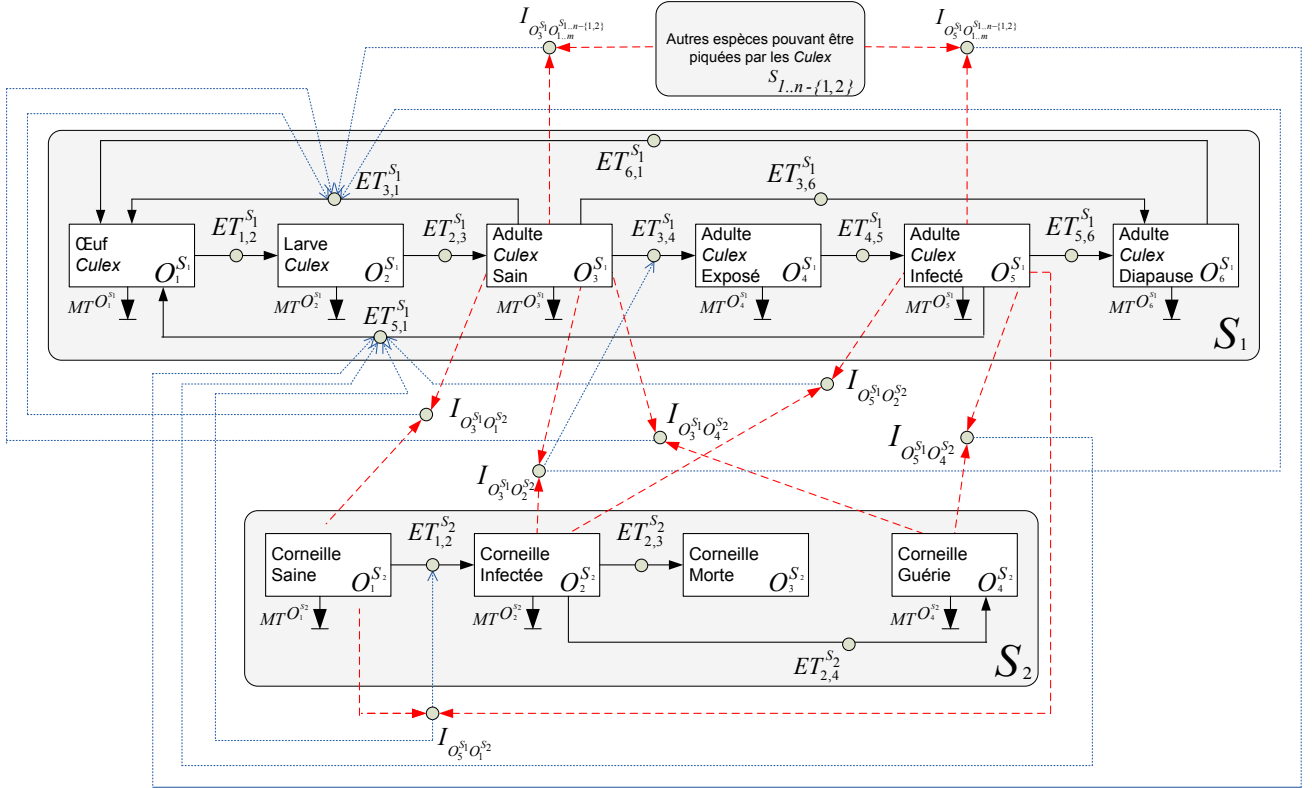


Figure 53: Les différentes interactions entre les *Culex* et les corneilles.

Voici maintenant le détail de chacune des interactions en nous basons sur le formalisme présenté auparavant :

$$\begin{aligned}
 I_{O_3^{S_1}O_1^{S_2}} &= (Id_{I_{O_3^{S_1}O_1^{S_2}}}, \text{Probc}_{O_3^{S_1}O_1^{S_2}}, \{ET_{3,1}^{S_1}\}), & I_{O_3^{S_1}O_2^{S_2}} &= (Id_{I_{O_3^{S_1}O_2^{S_2}}}, \text{Probc}_{O_3^{S_1}O_2^{S_2}}, \{ET_{3,1}^{S_1}, ET_{3,4}^{S_1}\}), \\
 I_{O_3^{S_1}O_4^{S_2}} &= (Id_{I_{O_3^{S_1}O_4^{S_2}}}, \text{Probc}_{O_3^{S_1}O_4^{S_2}}, \{ET_{3,1}^{S_1}\}), & I_{O_5^{S_1}O_1^{S_2}} &= (Id_{I_{O_5^{S_1}O_1^{S_2}}}, \text{Probc}_{O_5^{S_1}O_1^{S_2}}, \{ET_{5,1}^{S_1}\}, \{ET_{1,2}^{S_2}\}), \\
 I_{O_5^{S_1}O_2^{S_2}} &= (Id_{I_{O_5^{S_1}O_2^{S_2}}}, \text{Probc}_{O_5^{S_1}O_2^{S_2}}, \{ET_{5,1}^{S_1}\}), & I_{O_5^{S_1}O_4^{S_2}} &= (Id_{I_{O_5^{S_1}O_4^{S_2}}}, \text{Probc}_{O_5^{S_1}O_4^{S_2}}, \{ET_{5,1}^{S_1}\})
 \end{aligned}$$

Sachant que :

$Id_{I_{o_3^s o_1^s}} = \ll \text{susceptible } Culex \text{ biting susceptible crows} \gg$, $Id_{I_{o_3^s o_2^s}} = \ll \text{susceptible } Culex \text{ biting infectious crows} \gg$, $Id_{I_{o_3^s o_4^s}} = \ll \text{susceptible } Culex \text{ biting recovered crows} \gg$, $Id_{I_{o_5^s o_1^s}} = \ll \text{infectious } Culex \text{ biting susceptible crows} \gg$, $Id_{I_{o_5^s o_2^s}} = \ll \text{infectious } Culex \text{ biting infectious crows} \gg$, $Id_{I_{o_5^s o_4^s}} = \ll \text{infectious } Culex \text{ biting recovered crows} \gg$.

L'interaction $I_{o_3^s o_1^s}$ représente le fait qu'un *Culex* adulte sain a piqué une corneille saine afin de prendre un repas de sang. Suite à cette interaction, la transition $ET_{3,1}^{S_1}$ peut être déclenchée. Le *Culex* adulte sain peut donc pondre des œufs.

L'interaction $I_{o_3^s o_2^s}$ représente le fait qu'un *Culex* adulte sain a piqué une corneille infectée afin de prendre un repas de sang. Suite à cette interaction, les deux transitions $ET_{3,1}^{S_1}, ET_{3,4}^{S_1}$ peuvent être déclenchées. Le *Culex* adulte sain peut donc pondre des œufs et il peut devenir infecté.

L'interaction $I_{o_3^s o_4^s}$ représente le fait qu'un *Culex* adulte sain a piqué une corneille guérie afin de prendre un repas de sang. Suite à cette interaction, la transition $ET_{3,1}^{S_1}$ peut être déclenchée. Le *Culex* adulte sain peut donc pondre des œufs.

L'interaction $I_{o_5^s o_1^s}$ représente le fait qu'un *Culex* adulte infecté a piqué une corneille saine afin de prendre un repas de sang. Suite à cette interaction, les deux transitions $ET_{5,1}^{S_1}, ET_{1,2}^{S_2}$ peuvent être déclenchées. Le *Culex* adulte infecté peut donc pondre des œufs et la corneille saine peut devenir infectée.

L'interaction $I_{o_5^s o_2^s}$ représente le fait qu'un *Culex* adulte infecté a piqué une corneille infectée afin de prendre un repas de sang. Suite à cette interaction, la transition $ET_{5,1}^{S_1}$ peut être déclenchée. Le *Culex* adulte infecté peut donc pondre des œufs.

L'interaction $I_{O_3^{S_1} O_4^{S_2}}$ représente le fait qu'un *Culex* adulte infecté a piqué une corneille guérie afin de prendre un repas de sang. Suite à cette interaction, la transition $ET_{3,1}^{S_1}$ peut être déclenchée. Le *Culex* adulte infecté peut donc pondre des œufs.

L'interaction $I_{O_3^{S_1} O_{1..m}^{S_1..n-(1,2)}}$ représente le fait qu'un *Culex* adulte sain a piqué une autre espèce qu'une corneille afin de prendre un repas de sang. Suite à cette interaction, la transition $ET_{3,1}^{S_1}$ peut être déclenchée. Le *Culex* adulte sain peut donc pondre des œufs.

L'interaction $I_{O_3^{S_1} O_{1..m}^{S_1..n-(1,2)}}$ représente le fait qu'un *Culex* adulte infecté a piqué une autre espèce qu'une corneille afin de prendre un repas de sang. Suite à cette interaction, la transition $ET_{5,1}^{S_1}$ peut être déclenchée. Le *Culex* adulte infecté peut donc pondre des œufs.

4.2 Un modèle pour la maladie de Lyme

La maladie de Lyme a déjà été présentée dans le Chapitre III (voir section 2.2.2). Nous avons utilisé le modèle d'Ogden [Ogden *et al.*, 2005] afin de représenter les différents compartiments impliqués dans le cycle biologique des tiques. Nous ne présentons dans cette section que les représentations graphiques de notre modèle puisque le principe de l'utilisation des matrices a été bien élucidé dans l'exemple précédent (VNO). Nous commençons par présenter dans la Figure 54 le modèle biologique des tiques sans maladie.

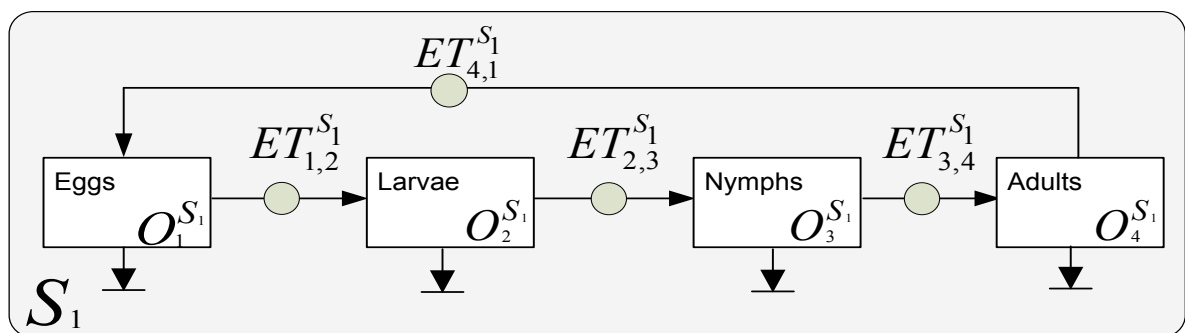


Figure 54: Le modèle biologique des tiques sans maladie.

Nous détaillons maintenant dans la Figure 55 les trois compartiments *larvae*, *nymphs* et *adults* en ajoutant les stades intermédiaires qui sont utilisés dans le modèle d'Ogden (*hardening*, *questing*, *feeding*, *engorged*). Cependant, nous avons modifié le nom du compartiment *questing* par le nom *maturing/questing* étant donné que les nymphes ainsi que les adultes devraient normalement passer par une période de maturation avant de chercher un hôte. Il est à noter que la flèche représentant la mortalité du compartiment $O_2^{S_1}$ est la somme des mortalités des sous-compartiments $O_{2.1}^{S_1}$, $O_{2.2}^{S_1}$, $O_{2.3}^{S_1}$ et $O_{2.4}^{S_1}$. Un nouvel aspect qui mérite d'être mentionné est l'utilisation de transitions entre les compartiments et sous-compartiments. Par exemple, la transition $ET_{1,2.1}^{S_1}$ représente le passage des individus entre le compartiment *Eggs* $O_1^{S_1}$ vers le sous-compartiment *Hardening* $O_{2.1}^{S_1}$.

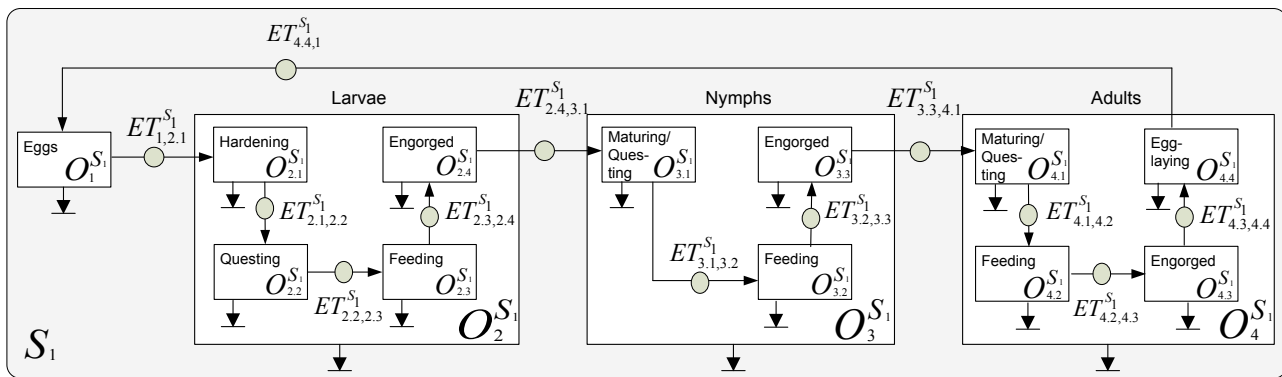


Figure 55: Le modèle biologique des tiques sans maladie.

Nous présentons maintenant dans la Figure 56 le modèle biologique des tiques en ajoutant la maladie (*susceptible*, *infected*). Nous ne considérons pas dans ce cas le compartiment *recovered* puisque les larves et les nymphes vont redevenir susceptibles. De plus et pour simplifier le modèle, nous considérons que les adultes ne vont piquer que des chevreuils. De ce fait, ces adultes vont nécessairement devenir non infectés suite à la mortalité de la bactérie [Ogden *et al.*, 2005]. D'ailleurs, ces adultes ne sont plus susceptibles d'attraper la maladie puisqu'ils vont mourir après avoir pris un repas de sang pour pondre. Dans notre

cas, les nymphes infectées peuvent redevenir susceptibles si elles ont piqué un hôte (c.-à-d. chevreuil) qui les a stérilisés. De plus et étant donné que nous ne tenons pas compte de la transmission verticale, les larves ne peuvent pas être stérilisées puisqu'elles ne peuvent pas être déjà infectées avant de prendre leur premier repas de sang.

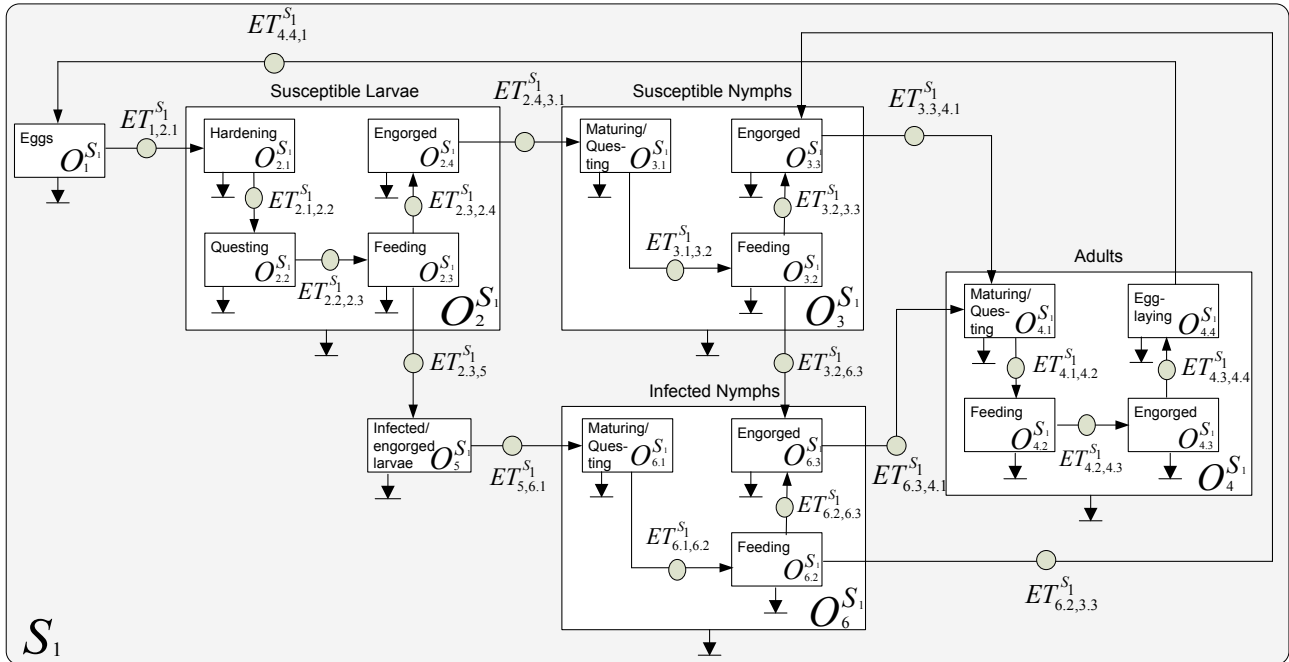


Figure 56: Le modèle biologique des tiques avec maladie.

Étant donné que le modèle est trop complexe sachant qu'il y a plusieurs acteurs qui sont impliqués dans la propagation de la maladie de Lyme, nous ne pouvons pas présenter les interactions en une seule étape comme c'était le cas le VNO. Nous présentons donc dans les sous-sections suivantes les interactions entre les tiques et chacun de ses hôtes (c.-à-d. rongeurs, oiseaux et chevreuils). De plus, nous présentons dans l'annexe B (Figures B1 à B5) une représentation graphique plus complète pour la maladie de Lyme en intégrant les différentes équations différentielles utilisées par Ogden [Ogden *et al.*, 2005].

4.2.1 Interaction entre les tiques et les rongeurs

La Figure 58 montre les différentes interactions entre les tiques et les rongeurs. Pour la population de rongeurs, nous n'utilisons que deux compartiments (*susceptible rodents* et *infected rodents*). En fait, nous avons fait une simplification en ne tenant pas compte des jeunes rongeurs (comme dans le modèle d'Ogden) puisque nous n'avons pas besoin d'aller à ce niveau de détail dans le modèle actuel. En outre, la figure 58 montre six interactions entre les tiques et les rongeurs. Par exemple, $I_{O_{6,2}^{S_1}O_1^{S_2}}$ est une interaction qui représente un accrochage d'une nymphe infectée en train de prendre son repas de sang ($O_{6,2}^{S_1}$) sur un rongeur susceptible ($O_1^{S_2}$). Suite à cette interaction, ce rongeur susceptible peut devenir infecté ($ET_{1,2}^{S_2}$) et cette nymphe peut devenir gorgée de sang ($ET_{6,2,6,3}^{S_1}$).

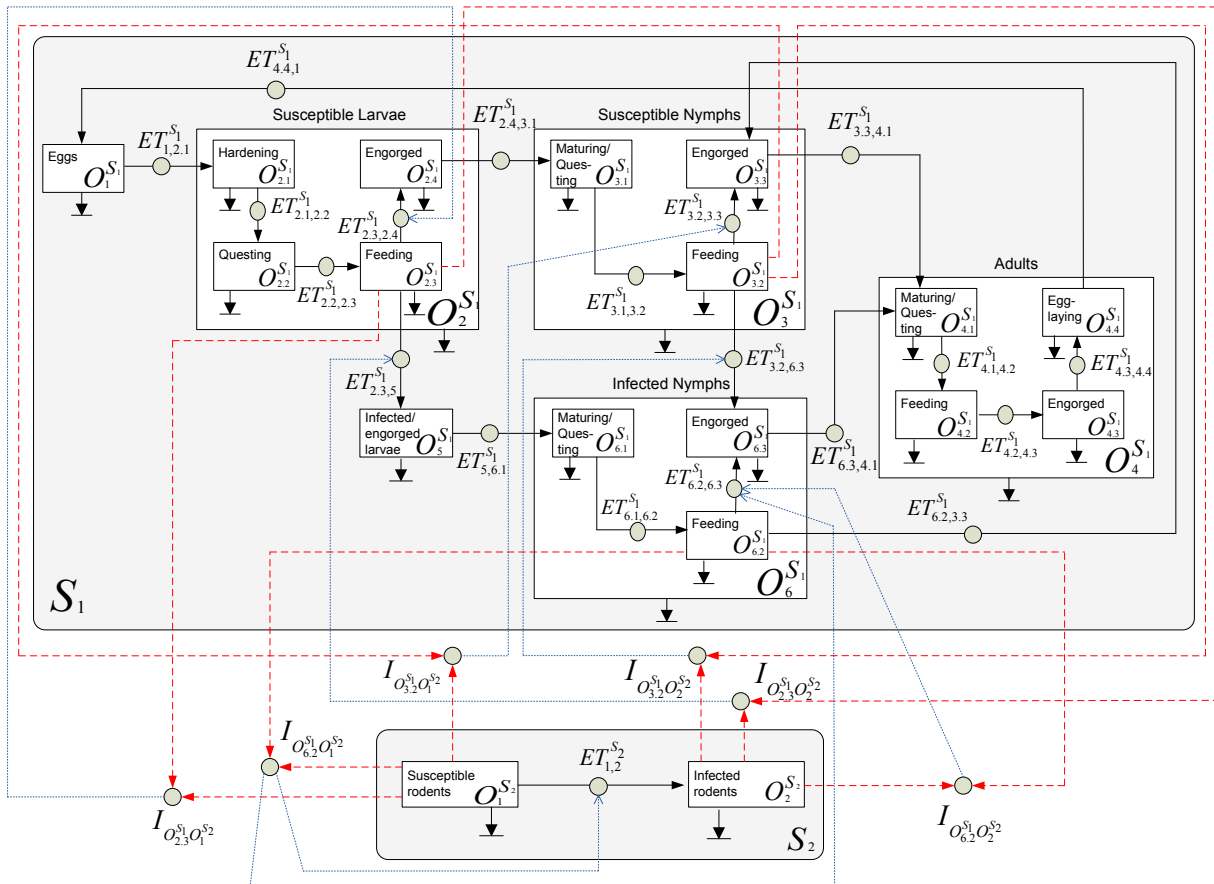


Figure 57: Les interactions entre les tiques et les rongeurs.

4.2.2 Interaction entre les tiques et les oiseaux

La Figure 59 montre les différentes interactions entre les tiques et les oiseaux. Ce sont quasiment les mêmes interactions que pour les tiques et les rongeurs, étant donné que ce sont également les stades juvéniles des tiques qui piquent généralement les oiseaux.

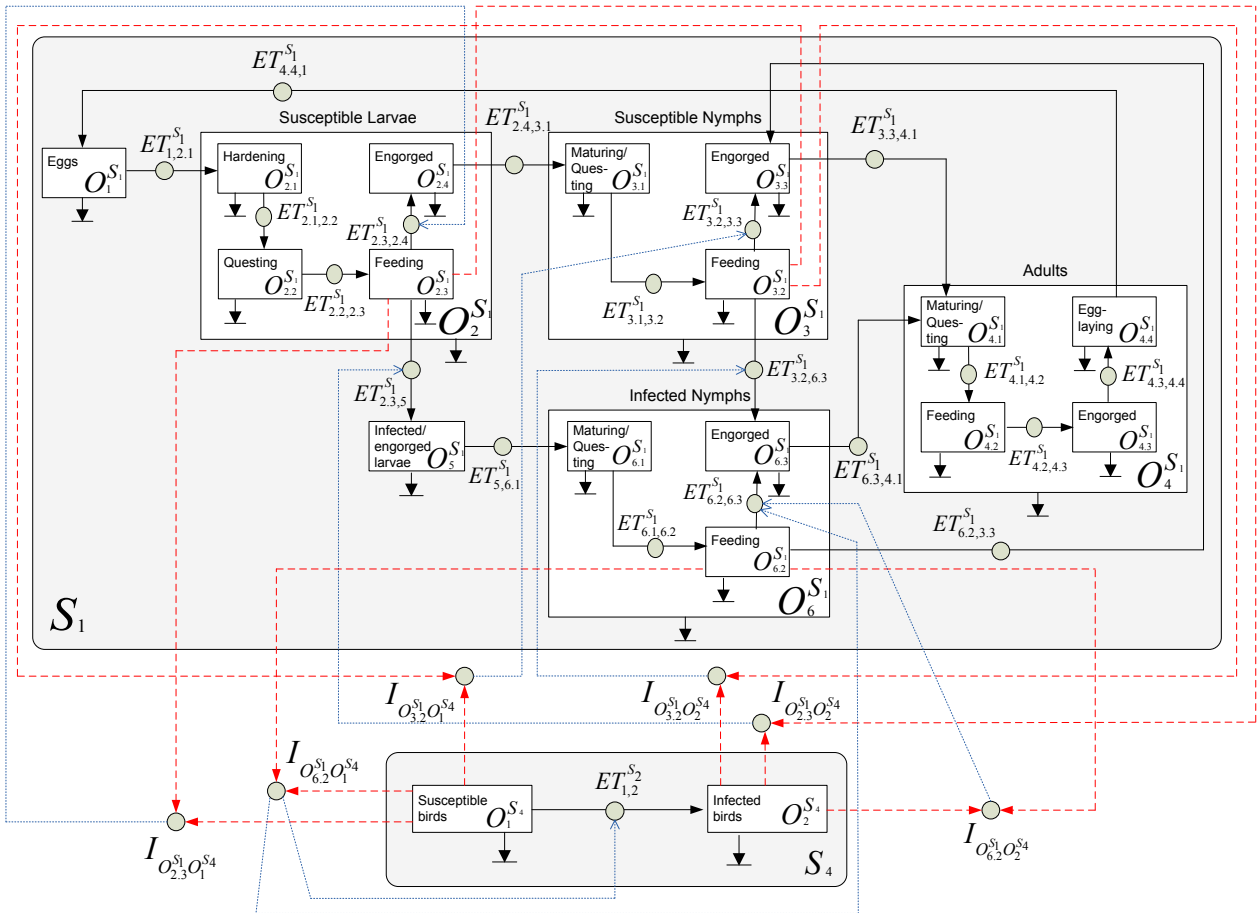


Figure 58: Les interactions entre les tiques et les oiseaux.

4.2.3 Interaction entre les tiques et les chevreuils

La Figure 59 montre les différentes interactions entre les tiques et les chevreuils. Aux fins de simplification, nous n'utilisons qu'un seul compartiment (*Adult deer*) pour les chevreuils. Nous constatons six interactions possibles entre les tiques et les chevreuils. Par

exemple, $I_{O_{4,2}^{S_1} O_1^{S_3}}$ est une interaction qui représente un accrochage d'une tique adulte en train de prendre son repas de sang ($O_{4,2}^{S_1}$) sur un chevreuil adulte ($O_1^{S_3}$). Suite à cette interaction, cette tique peut devenir gorgée de sang ($ET_{4,2,4,3}^{S_1}$).

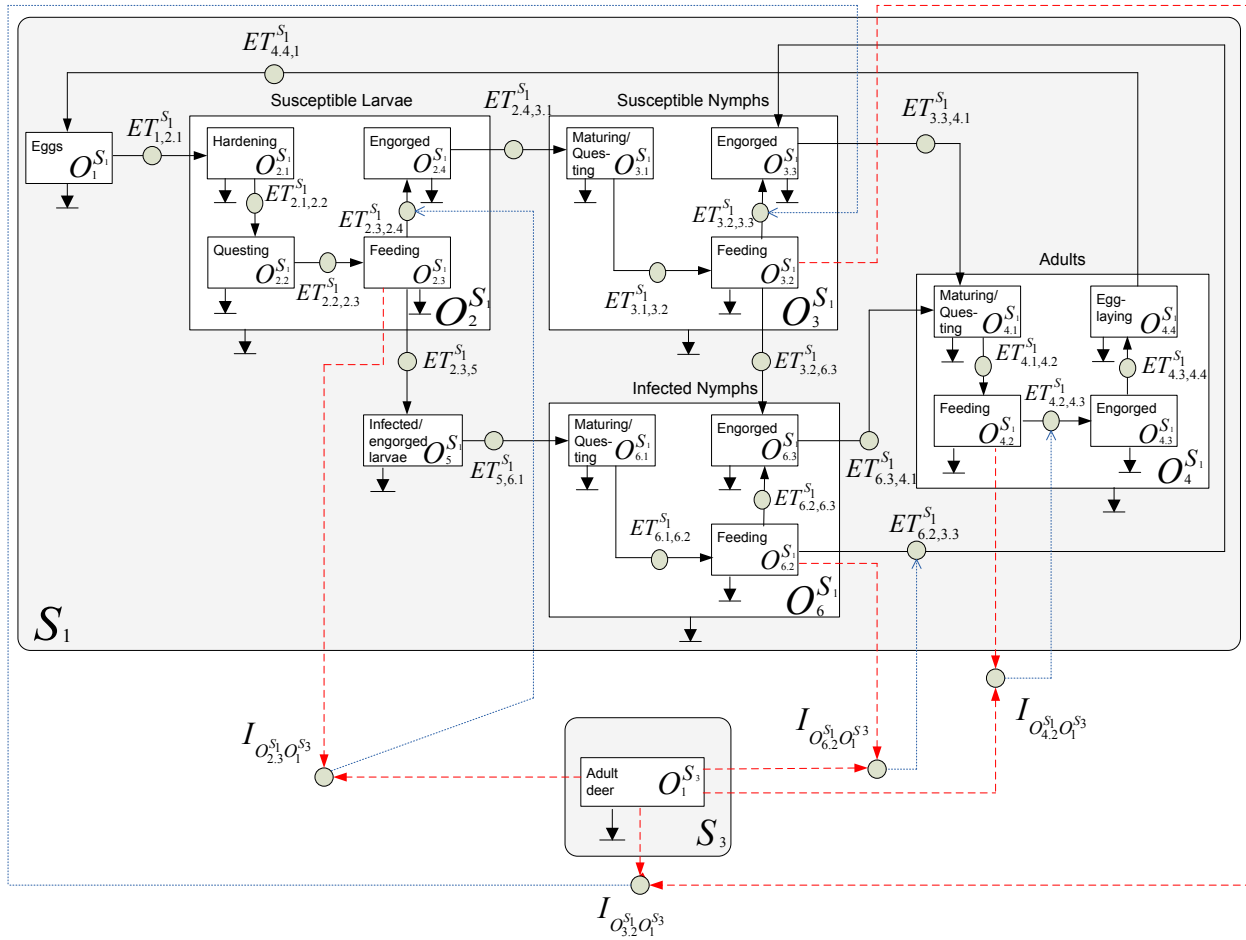


Figure 59: Les interactions entre les tiques et les chevreuil.

5 Discussion et conclusion

Dans ce chapitre, nous avons présenté un nouveau modèle théorique utilisé pour formaliser les différents concepts qui nous intéressent pour simuler la propagation d'une zoonose, en

tenant compte des caractéristiques spatio-temporelles de ce phénomène. Nous avons appliqué ce modèle à deux cas différents de zoonoses: le VNO et la maladie de Lyme. Notre modèle fournit beaucoup d'avantages si on le compare aux modèles classiques à base de compartiments qui ont été utilisés pour simuler la propagation d'épidémies jusqu'à présent [Ogden *et al.*, 2005], [Wonham *et al.*, 2004]. En effet, les modèles compartimentaux classiques ne tiennent pas compte de l'espace géographique dans lequel les populations évoluent et interagissent. Ainsi, ces modèles calculent l'évolution et une partie des interactions des espèces impliquées dans la propagation des zoonoses dans un espace géographique homogène, le plus souvent dans une seule cellule. Cette approche permet entre autres d'avoir une idée sur les différents facteurs qui influencent la propagation de l'agent pathogène, mais elle ne pourra pas être utilisée pour simuler ce qui se passe réellement sur le terrain, surtout que ce dernier est caractérisé pour un espace étendu et très hétérogène. De plus, les modèles compartimentaux classiques ne peuvent donner des résultats qu'à un niveau macro et de ce fait, ils ne peuvent aider qu'à établir des directives d'intervention à un niveau stratégique (p.ex. soutenir des décisions politiques). En revanche, le modèle *MASTIM* que nous proposons utilise un environnement de simulation obtenu à partir des données SIG. Cet environnement est subdivisé en un ensemble de cellules qui peuvent être agrégées à différents niveaux de détail dans d'autres cellules à des niveaux hiérarchiques supérieurs. Dans chacune de ces cellules, nous clonons un modèle enrichi à base de compartiments qui permet non seulement de calculer l'évolution des espèces impliquées dans la propagation du phénomène étudié, mais permet également de calculer les différentes interactions spatio-temporelles se produisant entre ces espèces qui sont de différents types et représentant généralement des populations contenant un nombre très important d'individus. Par ailleurs, *MASTIM* matérialise les différentes échelles spatiales (p.ex. Macro, Méso et/ou Micro) de notre approche de géosimulation présentée auparavant (voir Chapitre IV) en utilisant les différents niveaux hiérarchiques des cellules. De ce fait, notre modèle pourrait aider non seulement les décideurs politiques pour établir les directives d'intervention à un niveau stratégique, mais également les décideurs tactiques ou opérationnels pour établir des plans d'intervention concrets.

En outre, le modèle *MASTIM* tient compte de l'influence des caractéristiques de l'environnement (p.ex. les différents habitats favorables se trouvant dans la zone géographique représentant la région d'études) dans le calcul des interactions spatio-temporelles entre les différentes espèces impliquées dans la propagation d'une zoonose. Cependant, la génération de l'environnement de la simulation intégrant ces caractéristiques représente un grand défi de modélisation. Nous essayons d'ailleurs de relever ce défi en développant une nouvelle approche permettant de générer automatiquement un environnement géographique virtuel informé permettant entre autres d'améliorer l'efficacité et la plausibilité de la simulation du phénomène qui nous intéresse. Cette approche sera présentée dans le chapitre suivant, ce qui nous permettra de répondre au troisième objectif de notre thèse. D'ailleurs et dans le chapitre suivant, nous complétons le formalisme présenté dans ce chapitre en ajoutant d'autres concepts qui nous semblent importants et qui permettent entre autres de faciliter la présentation de la nouvelle approche proposée.

Chapitre VI: Approche de génération de l'*IVGE*

1 Introduction

Nous proposons dans ce chapitre une nouvelle approche permettant de générer d'une façon automatique un environnement géographique virtuel informé en bénéficiant des concepts proposés auparavant lors de la formalisation du modèle *MASTIM*, mais en introduisant de nouveaux concepts permettant de répondre à nos besoins et par conséquent enrichir ce modèle. Nous avons nommé cet environnement de simulation un *IVGE (Informed Virtual Geographic Environment)* [Bouden et Moulin, 2012a]. Nous commençons par la présentation de l'*IVGE* en spécifiant entre autres les raisons qui nous ont amené à développer une nouvelle approche pour sa création et à lui donner le caractère « informé ». Nous présentons par la suite les différentes étapes permettant de créer l'*IVGE* en donnant une vue d'ensemble de l'approche proposée. Nous présentons finalement une application de l'*IVGE* concernant la distribution spatiale des flots d'individus, en montrant l'utilité de l'*IVGE* par rapport à notre approche de géosimulation multi-niveau.

2 Présentation de l'*IVGE*

L'*IVGE* est en réalité une représentation virtuelle de la zone géographique qui nous intéresse (p.ex. le sud de Québec pour le VNO) et qui fait partie du système écologique que nous voulons simuler. Ainsi, il joue le rôle d'un support permettant de géosimuler l'évolution des espèces impliquées dans la propagation d'une zoonose ainsi que leurs interactions spatio-temporelles. Cependant, une des problématiques qui nous semble cruciale est de savoir comment représenter efficacement l'*IVGE* dans lequel la géosimulation sera effectuée, en tenant compte du fait que les territoires étudiés sont énormes et que nous devons tenir compte des caractéristiques géographiques qui influencent les cycles biologiques des espèces concernées (p.ex. zones géographiques qui sont favorables ou non favorables à la survie et à la prolifération des vecteurs faisant propager la maladie) et leurs comportements (p.ex. alimentation, déplacements, migrations,

etc.). Nous avons mentionné dans le chapitre précédent que l'environnement géographique dans lequel se trouve le système écologique est subdivisé en cellules spatiales élémentaires en utilisant des données vectorielles. Chaque cellule est représentée par un polygone défini à l'aide des données SIG. Plusieurs types de polygones peuvent être utilisés pour faire la subdivision spatiale de l'IVGE. Le choix des types de polygones dépend des caractéristiques spatiales qui sont importantes pour l'observation et l'analyse de la propagation des zoonoses.

Dans le cas du VNO (voir Section 2 du Chapitre VII), nous avons utilisé des polygones représentant des municipalités ou des secteurs de recensement²⁴, selon la région d'intérêt (les municipalités sont utilisées pour couvrir de grandes surfaces telles que le sud de la province de Québec, alors que les secteurs de recensement sont utilisés pour caractériser des petites zones telles que la région métropolitaine d'Ottawa). Cette subdivision administrative correspond aux données de surveillance qui étaient disponibles pour calibrer la simulation de la propagation du VNO. Par exemple, le choix de la zone d'application des larvicides permettant de limiter cette propagation se faisait en se basant sur les municipalités et/ou les secteurs de recensement qui sont à risque. De ce fait, le choix de la subdivision administrative pour le cas du VNO était assez raisonnable bien que le phénomène étudié ne soit pas limité aux régions urbaines. Par contre, dans le cas de la simulation de l'implantation des colonies de tiques et la propagation de la maladie de Lyme, il ne serait pas raisonnable d'utiliser une telle subdivision administrative. En effet, les cellules représentant des municipalités ou des secteurs de recensement sont trop grandes et n'ont aucune signification biologique, étant donné que les populations de tiques ne peuvent survivre que dans des zones herbeuses au bord (ou dans) des forêts pas trop denses [Ogden *et al.*, 2008a]. Ainsi, nous avons pensé à utiliser une subdivision spatiale basée sur les caractéristiques de la couverture du sol (c.-à-d. *Land Cover*). Les données de cette

²⁴ Au Canada, les secteurs de recensement (SR) sont de petites régions géographiques relativement stables qui comptent habituellement entre 2 500 et 8 000 habitants. Ils sont créés au sein de régions métropolitaines de recensement et d'agglomérations de recensement dont le noyau urbain compte 50 000 habitants ou plus [Site Web: *StatCan*].

couverture du sol peuvent être fournies en format vectoriel (c.-à-d. *shape file*) dans lesquelles les cellules polygonales peuvent couvrir de larges territoires (en utilisant par exemple la base de données GéoBase [Site Web: *Geobase*]).

Néanmoins, nous avons été confrontés à un autre problème majeur concernant l'utilisation de ce genre de subdivision (c.-à-d. en utilisant la couverture du sol) afin de simuler la propagation des zoonoses. Il s'agit du nombre élevé de cellules à traiter, sachant que nous devons cloner notre modèle compartimental enrichi dans chacune de ces cellules. De ce fait, nous avons eu l'idée de fusionner les cellules de couverture du sol ayant des caractéristiques similaires relativement au phénomène à simuler. Ainsi, nous pouvons avoir les plus grands polygones possibles permettant de minimiser les besoins computationnels sans pour autant affaiblir la plausibilité de la simulation. La méthode de fusion que nous proposons est basée sur des seuils sélectionnés par l'utilisateur qui modélise le phénomène. Ces seuils sont liés à la favorabilité de l'habitat qui peut affecter la survie, l'alimentation et/ou la reproduction des espèces impliquées dans la propagation de la zoonose.

Dans ce chapitre, nous proposons donc une approche pour créer un *IVGE* qui porte le nom d'informé, car il est renforcé (p.ex. biologiquement, topologiquement, etc.) avec des données qui sont attachées aux cellules. Nous détaillerons d'ailleurs (voir Section 3.5) ces données ainsi que leur utilité (voir Section 4) afin de surtout améliorer l'efficacité et la plausibilité de la géosimulation de la propagation d'une zoonose.

3 Création de l'*IVGE*

Nous commençons cette section par la présentation d'une vue d'ensemble de l'approche proposée permettant de créer automatiquement l'*IVGE*. Ensuite, nous présentons le critère utilisé pour fusionner les cellules. Après cela, nous discutons des différentes étapes de notre approche, y compris le prétraitement des données SIG et la suppression des trous. Nous présentons finalement la fusion progressive des cellules et l'enrichissement de l'*IVGE* par l'ajout des données dans les cellules obtenues après fusion.

3.1 Vue d'ensemble de l'approche proposée

La Figure 60 montre une vue d'ensemble de notre approche permettant de générer automatiquement l'*IVGE* présenté précédemment. La partie gauche de cette figure (p.ex. rectangles en gris transparent) montre comment notre système peut générer l'*IVGE* en utilisant une subdivision administrative. Comme mentionnée auparavant, cette subdivision utilise par exemple des cellules représentant des municipalités ou des secteurs de recensement selon le choix de l'utilisateur et en fonction de l'échelle spatiale sélectionnée. Les cellules sont par la suite attachées avec des données de populations et/ou des données climatiques en utilisant une base de données. Dans cet exemple, l'*IVGE* obtenu est utilisé pour géosimuler la propagation du VNO (voir Section 2 du Chapitre VII).

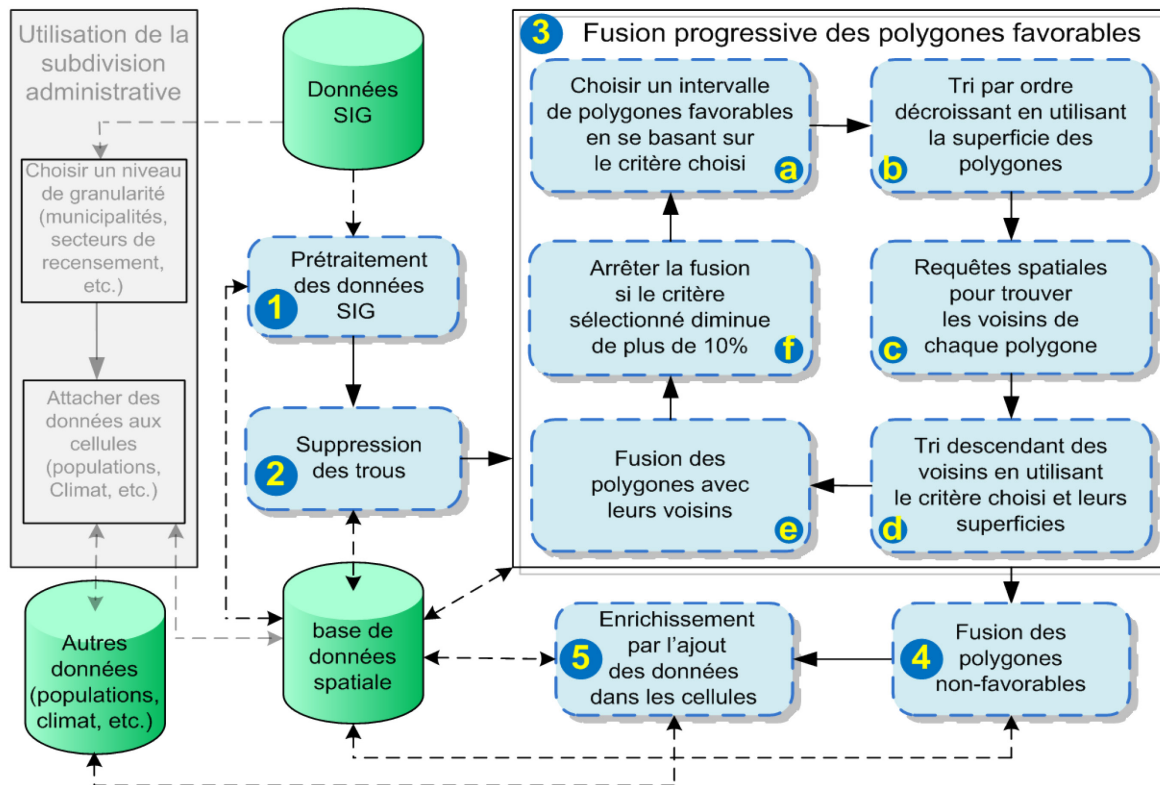


Figure 60: Vue d'ensemble de l'approche permettant de générer l'*IVGE*.

La partie droite de la Figure 60 (rectangles bleus contenant soit des numéros soit des lettres) montre comment notre système peut générer l'*IVGE* en utilisant une subdivision initiale basée sur un critère de favorabilité selon par exemple la couverture du sol. Cette subdivision est par la suite raffinée en faisant la fusion de cellules représentant les polygones fournis par les données SIG. En raison du grand nombre de cellules à traiter et parce qu'il est difficile de prévoir à quel moment arrêter le processus de fusion pour obtenir les plus grands polygones qui peuvent satisfaire le critère choisi, nous proposons une approche qui se déroule en plusieurs étapes. Dans un premier temps, notre système fait un prétraitement des données SIG en sélectionnant la région qui nous intéresse et il fait un calcul de la favorabilité de tous les polygones par rapport au critère retenu (voir Sections 0). Ensuite, nous créons des polygones pleins en retirant leurs trous afin de réduire le nombre de cellules élémentaires (voir Section 3.3). Le résultat est sauvegardé dans une base de données spatiale (nous avons utilisé pour nos premiers tests une base spatiale *Access*, mais nous avons migré par la suite vers une base de type *PostGIS* [Site Web: *PostGIS*] pour améliorer l'efficacité des traitements) qui est mise à jour au fur et à mesure que notre système exécute les différentes étapes. Par ailleurs, l'étape la plus importante de l'approche proposée est la fusion progressive des polygones favorables. Notre système commence ce processus en sélectionnant les polygones ayant la meilleure favorabilité (100%) selon le critère choisi. Ensuite, il trie ces polygones selon la superficie en ordre décroissant (le plus grand est le premier qui va être traité). Il fait par la suite des requêtes spatiales afin d'obtenir les voisins qui sont adjacents au polygone en cours de traitement. Ces voisins sont également triés (le plus grand et le meilleur selon le critère de favorabilité choisi est le premier à être traité) afin d'essayer de les fusionner au polygone d'origine.

Notre système arrête l'agrégation des voisins si la fusion d'un voisin supplémentaire diminue la favorabilité, du polygone obtenu après cette étape, de plus de 10%. De cette manière, nous préservons la favorabilité des polygones résultants et ainsi nous permettons un processus progressif de fusion. Notre système itère ensuite le processus de fusion sur un nouvel ensemble de polygones avec une valeur de favorabilité se trouvant dans l'intervalle [90%, 100%] que nous appelons *ATI* «*Absorption Threshold Interval*». Le système

continue donc d'appliquer ce processus plusieurs fois en augmentant progressivement l'intervalle *ATI* pour les polygones traités et en réduisant simultanément le seuil utilisé pour arrêter la fusion. Enfin, le système fait le contraire en sélectionnant les polygones ayant la plus mauvaise favorabilité (0%) en fonction du critère choisi et en essayant de les fusionner avec d'autres polygones non favorables (voir Section 3.4). Comme c'est le cas pour les polygones favorables, ce processus permet de générer des zones non favorables homogènes (agrégation des polygones non favorables) avec une taille maximale. De plus, ce processus permet de réduire encore le nombre de polygones utilisés pour générer l'*IVGE*. Toutes les étapes mentionnées ci-dessus sont effectuées jusqu'à atteindre un résultat satisfaisant selon l'appréciation de l'utilisateur. Par ailleurs, la dernière étape de notre approche est l'enrichissement de l'*IVGE* dans lequel la propagation d'une zoonose peut être simulée avec surtout un peu plus d'efficacité computationnelle (voir Section 3.5).

En outre, notre approche de fusion progressive est basée sur des heuristiques²⁵ sous la forme de règles d'agrégation en utilisant des critères et des seuils permettant d'obtenir une solution approchée de la solution optimale. En effet, les premiers polygones qui sont traités sont ceux qui ont la meilleure favorabilité et qui ont la plus grande superficie (notre système fait un tri descendant de ces polygones en utilisant la favorabilité et la superficie comme critères pour l'ordre de comparaison). De ce fait, nous nous assurons que les polygones sélectionnés sont les meilleurs candidats pour la suite du processus. De plus, ces polygones sont fusionnés avec leurs meilleurs voisins, étant donné que le système fait également un tri descendant sur ces voisins selon leur favorabilité et leur superficie. La fusion de chaque polygone à ses propres voisins s'arrête si la favorabilité descend en dessous d'un certain seuil défini par l'utilisateur (nous avons utilisé dans notre cas 10%). Par conséquent, les polygones résultants après fusion progressive auront un raisonnable rapport favorabilité/superficie afin de les utiliser comme cellules élémentaires de l'environnement de la simulation. Ainsi, notre système applique un algorithme vorace (voir

²⁵ Les heuristiques sont des règles empiriques permettant de faire un calcul fournissant une solution approchée, pas nécessairement optimale, pour un problème d'optimisation difficile.

Annexe C) permettant de construire la solution d'une manière incrémentale où chaque étape prend la direction la plus prometteuse (c.-à-d. le choix des polygones à fusionner) suivant les règles mentionnées auparavant.

3.2 Utilisation des habitats favorables comme critère de fusion

Le phénomène biologique que nous étudions est étroitement lié aux caractéristiques du paysage. En effet, la propagation des zoonoses dépend beaucoup de la survie des populations impliquées dans la transmission des agents pathogènes (c.-à-d. virus, bactéries). Par exemple, la capacité des populations de tiques à survivre dans des habitats appropriés (comme pour 'les forêts de feuillus clairsemées' selon la terminologie de Géobase [Site Web: *Geobase*]) est un facteur important qui influence la propagation de la maladie de Lyme. En fait, une espèce peut s'installer dans une zone seulement si cette zone est favorable à sa survie. Par conséquent et pour chaque polygone des données SIG (voir Section 3.3), nous proposons de déterminer ce que nous appelons un degré de favorabilité de base (*BSD : Basic Suitability Degree*) qui a été introduit dans le chapitre précédent (voir Section 3.5 du Chapitre V) afin d'estimer la qualité des habitats appropriés pour une espèce donnée.

À cette fin, nous avons ajouté un nouvel attribut à la table de la base de données spatiale (*Access* pour nos premiers tests et *PostGIS* par la suite). Cet attribut qui représente la favorabilité du polygone pour une espèce donnée est un réel qui prend une valeur entre 0 et 1. La valeur 0 correspond à une zone qui n'est pas favorable pour les individus de l'espèce concernée (0% de favorabilité comme par exemple une zone urbaine pour les tiques) et la valeur 1 correspond à une zone très favorable (100% de favorabilité comme par exemple une zone chaude et humide pour les moustiques) (voir Tableau 2). Ce degré de favorabilité de base est un attribut de la cellule sachant que sa valeur est déterminée soit selon la littérature soit selon les avis des experts. Il peut bien sûr être modifié selon le choix de l'utilisateur afin de tester des résultats différents d'agrégation de cellules.

Tableau 2: Degrés de favorabilité de base (BSD) des habitats pour les tiques en fonction des types de couverture du sol de Géobase.

Code	Type	BSD	Code	Type	BSD
0	Pas de données	0	100	Plantes herbacées	0.2
10	Non classifié	0	101	Toundra de graminoides parsemée de buttes de gazon	0.2
11	Nuage	0	102	Carex humide	0.2
12	Ombre	0	103	Toundra de graminoides et d'arbustes nains, humide à sèche sans butte de gazon	0.2
20	Eau	0	104	Toundra d'arbustes nains prostrés et de graminoides	0.2
30	Stérile/non végétalisé	0	110	Prairies, herbes indigènes	0.2
31	Neige/glace	0	120	Terres agricoles cultivées	0.2
32	Roche/blocaille	0	121	Cultures annuelles	0.2
33	Terrain découvert	0	122	Cultures pérennes et pâturages	0.2
34	Zones développées	0	200	Forêts – classes d'arbres	0.8
35	Substrat rocheux à végétation épars	0	210	Forêt de conifères	0.8
36	Till-colluvion à végétation épars	0	211	Coniférien - dense	0.6
37	Sol à nu avec croûte de cryptogames - éminences dues au gel	0	212	Coniférien - ouvert	0.9
40	Bryophytes et lichens	0	213	Coniférien - clairsemé	0.9
50	Arbustes	0,8	220	Forêt de feuillus	0.85
51	Grands arbustes	0.75	221	Feuillu - dense	0.7
52	Petits arbustes	0.75	222	Feuillu - ouvert	1
53	Arbustes nains prostrés	0.2	223	Feuillu - clairsemé	1
80	Terres humides	0.2	230	Forêt mixte	0.8
81	Zone humide - boisée	0.2	231	Mixte - dense	0.65
82	Zone humide - arbustive	0.2	232	Mixte - ouvert	0.95
83	Zone humide - herbacée	0.2	233	Mixte - clairsemé	0.95

Nous présentons d'ailleurs un exemple de la fonction que nous appliquons (pour le cas des tiques) afin d'attribuer un degré de favorabilité de base à chacune des cellules avant la suppression des trous et avant le processus de fusion progressive:

BSD = *Suitability-BSD-Function* (Code : Type de couverture du sol){
 SI (Code=222 Ou Code =223) ALORS BSD = 1
 SINON SI (Code =232 Ou Code =233) ALORS BSD =0.95
 SINON SI (Code =212 Ou Code =213) ALORS BSD = 0.9
 SINON SI (Code =220) ALORS BSD = 0.85
 SINON SI (Code =50 Ou Code =200 Ou Code =210 Ou Code =230) ALORS BSD = 0.8


```

    SINON SI (Code=51 Ou Code=52) ALORS BSD=0,75
    SINON SI (Code =221) ALORS BSD = 0.7
    SINON SI (Code =231) ALORS BSD = 0,65
    SINON SI (Code =211) ALORS BSD = 0.6
    SINON SI (Code =53 Ou (Code >=80 Et Code <=122) ALORS BSD = 0.2
    SINON BSD = 0
}

```

De ce fait et pour illustrer le formalisme adopté concernant l’habitat favorable d’une espèce (voir Section 3.5 du Chapitre V), nous pouvons donner l’exemple suivant :

$$\{(LCtype_1, \dots, LCtype_n, MinBSD), \dots, (LCtype_1, \dots, LCtype_m, MaxBSD)\} =$$

$$\{(53,80,81,82,83,100,101,102,103,104,110,120,121,122,0.2),$$

$$(211,0.6),(231,0.65),(221,0.7),(51,52,0.75),(50,200,210,0.8),$$

$$(220,0.85),(212,213,0.9),(232,233,0.95),(222,223,1)\}$$

Dans ce cas, $MinBSD = 0.2$ et $MaxBSD = 1$. Nous n’avons pas besoin de mentionner dans cet exemple les types de couverture qui sont non favorables ($BSD = 0$) puisqu’ils peuvent être déterminés par élimination.

3.3 Présentation et prétraitement des données SIG

Nous avons utilisé les fichiers de couverture du sol (c.-à-d. *shapefiles*) fournis par la base de données spatiale GéoBase [Site Web: *Geobase*]. Ces données de la couverture terrestre sont le résultat de la vectorisation de données matricielles thématiques issues de la classification d'ortho-images des satellites Landsat5 et Landsat7 des zones forestières et agricoles du Canada et des territoires du Nord. Ces données couvrent l'ensemble des territoires du Canada qui sont divisés en différentes régions en utilisant les cartes d'index du Système National de Référence Cartographique du Canada (SNRC). Chaque région est identifiée par un numéro unique (p.ex. 22M, 30K, etc.).

Étant donné que pour le cas de la maladie de Lyme, nous nous intéressons au sud de la province de Québec et plus particulièrement aux régions de la Montérégie et de l'Estrie, nous considérons six régions (c.-à-d. 21E, 21L, 31G, 31H, 31I et 31J) qui sont représentées dans la Figure 61. Par exemple, la région 31H contient 133 780 polygones distincts et sa

dimension est presque de 156 x 110 km. Chaque polygone a des attributs descriptifs tels que le type de la couverture sol qui est basé sur la classification de l'OTDD (Observation de la Terre pour le Développement Durable) du Canada. En effet, l'attribut correspondant au type de la couverture du sol prend sa valeur entre 0 et 233 (voir Tableau 2 et Tableau 3) et représente les différentes catégories de couverture terrestre (p.ex. 20 pour l'eau, 34 pour une zone urbaine, 50 pour les arbustes, 220 pour une forêt de feuillus, etc.). Il est également important de mentionner que certains polygones peuvent avoir des formes complexes comme le montre la Figure 62. Nous devons d'ailleurs tenir compte de cette complexité lors de la fusion des cellules, mais aussi lors de l'enrichissement de l'IVGE (voir Section 3.5).

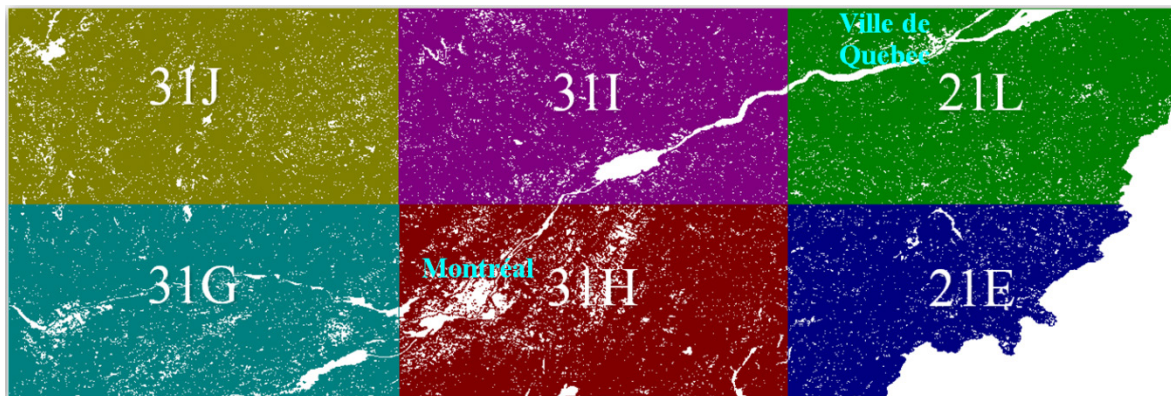


Figure 61: Les 6 régions SNRC qui nous intéressent.

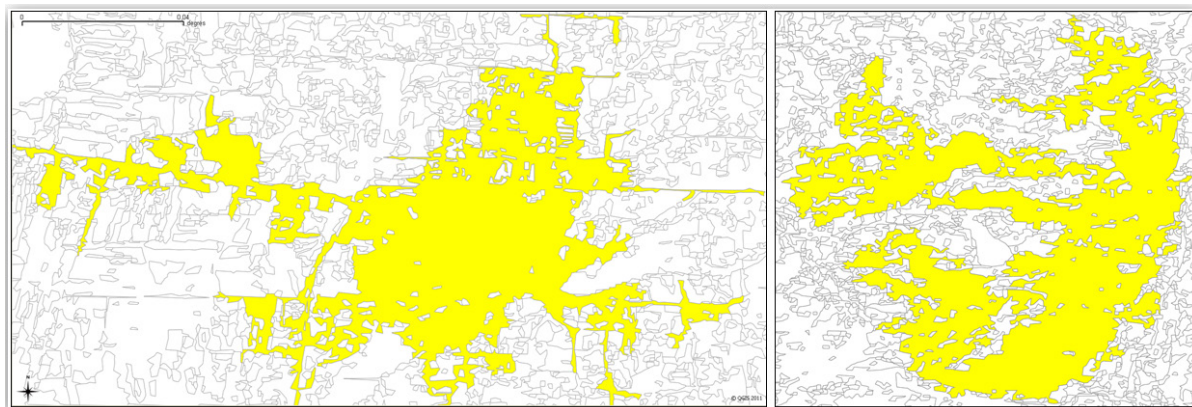


Figure 62: Deux polygones complexes représentant une zone urbaine (à gauche) et une forêt (à droite).

Tableau 3: Description des principaux types de couverture du sol présents dans Géobase.

Code	Type	Description
50	Arbustes	Principalement les étendues occupées par une végétation ligneuse relativement basse (généralement ± 2 mètres). Commentaires : Elles peuvent inclure des étendues herbeuses ou des prairies humides portant une végétation ligneuse et les forêts en voie de régénération.
51	Grands arbustes	La couverture végétale est d'au moins 20 % et se compose d'au moins un tiers d'arbustes et d'arbrisseaux dont la hauteur moyenne est égale ou supérieure à 2 m. Dans le Nord, grands arbustes dressés en milieux ripariens à secs de plus de 40 cm constituant plus de 25 % de la couverture végétale composée surtout de bouleaux glanduleux (<i>Betula</i>), de saules (<i>Salix</i>) et/ou d'aulnes (<i>Alnus</i>). La couverture restante est composée de graminoides, de lichen, peut contenir moins de 10 % d'arbustes nains prostrés et de sol nu.
52	Petits arbustes	La couverture végétale est d'au moins 20 % et se compose d'au moins un tiers d'arbustes et d'arbrisseaux dont la hauteur moyenne est inférieure à 2 m. Dans le Nord, petits arbustes dressés ripariens de plus de 40 cm constituant plus de 25 % de la couverture végétale composée surtout de bouleaux glanduleux (<i>Betula</i>) et/ou de saules (<i>Salix</i>). La couverture restante est composée de graminoides et de lichen et peut contenir des arbustes nains prostrés et du sol nu.
200	Forêts-classes d'arbres	Principalement des étendues boisées ou avec végétation arborescente. Commentaires : Cette classe est représentée uniquement s'il est impossible de distinguer les sous-classes de couvert forestier (210, 220, 230).
210	Forêt de conifères	Forêt principalement composée de conifères ou végétation arborescente à feuillage persistant. Peut inclure des étendues de forêt mixte et d'arbustes.
211	Coniférien - dense	La fermeture du couvert est supérieure à 60 %; au moins 75 % de la surface terrière totale est occupée par des arbres du groupe des conifères.
212	Coniférien - ouvert	La fermeture du couvert est de 26 à 60 %; au moins 75 % de la surface terrière totale est occupée par des arbres appartenant au groupe des conifères.
213	Coniférien -clairsemé	La fermeture du couvert est de 10 à 25 %; au moins 75 % de la surface terrière totale est occupée par des arbres appartenant au groupe des conifères.
220	Forêt de feuillus	Forêts et étendues boisées principalement composées de feuillus. Peut inclure des étendues de forêt mixte et d'arbustes.
221	Feuillu - dense	La fermeture du couvert est supérieure à 60 %; au moins 75 % de la surface terrière totale est occupée par des arbres feuillus.
222	Feuillu - ouvert	La fermeture du couvert est de 26 à 60 %; au moins 75 % de la surface terrière totale est occupée par des arbres feuillus.
223	Feuillu - clairsemé	La fermeture du couvert est de 10 à 25 %; au moins 75 % de la surface terrière totale est occupée par des arbres feuillus.
230	Forêt mixte	Forêts et étendues boisées composées de conifères et de feuillus.
231	Mixte-dense	La fermeture du couvert est supérieure à 60 %; aucun des deux grands types d'arbres (conifères et feuillus) n'occupe 75 % ou plus de la surface terrière totale.
232	Mixte-ouvert	La fermeture du couvert est de 26 à 60 %; aucun des deux grands types d'arbres (conifères et feuillus) n'occupe 75 % ou plus de la surface terrière totale.
233	Mixte-clairsemé	La fermeture du couvert est de 10 à 25 %; aucun des deux grands types d'arbres (conifères et feuillus) n'occupe 75 % ou plus de la surface terrière totale.

Nous avons testé notre approche en utilisant le fichier *shapefile* représentant la région 31H (voir Figure 63). Ce fichier contient 133 780 polygones représentant différents types de couverture du sol. Nous avons utilisé Geomedia [Site Web: *Geomedia*] afin de convertir ce fichier en une base de données spatiale Access (en mode lecture et écriture) afin d'avoir la possibilité de le modifier. Pour cette conversion, nous avons utilisé le système de coordonnées de référence du Canada (c.-à-d. *NAD83-CSRS*). Ensuite, nous avons ajouté et calculé l'attribut représentant le degré de favorabilité de base (c.-à-d. *BSD*) pour chaque

polygone tel que mentionné dans la section précédente. Nous avons également développé une nouvelle application en utilisant le langage Visual Basic qui est capable d'interroger (via des requêtes spatiales) la base de données spatiale Access et modifier ses attributs.

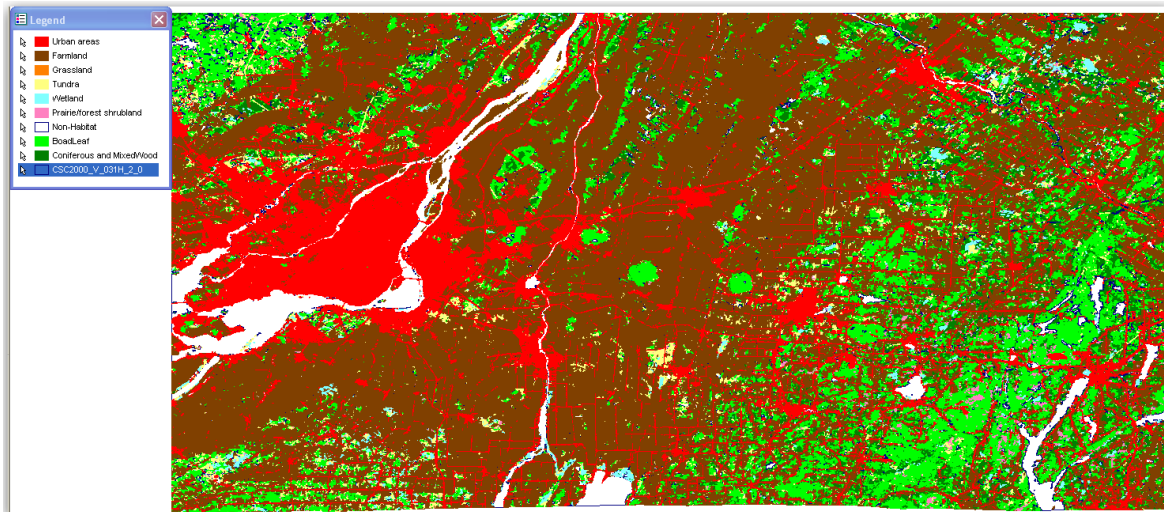


Figure 63: Région 31H avec une représentation de certains types de couverture du sol.

Nous avons remarqué qu'un grand nombre de polygones ont des trous. Ces trous sont considérés comme des polygones dans Géobase avec leurs propres attributs de couverture du sol. Pour réduire le nombre de polygones, nous avons décidé de retirer ces trous et de créer des polygones pleins. Nous avons supposé que le remplissage des trous ne réduira pas considérablement la favorabilité du polygone, puisque la superficie de la plupart des trous représente une petite partie de la taille de ce polygone.

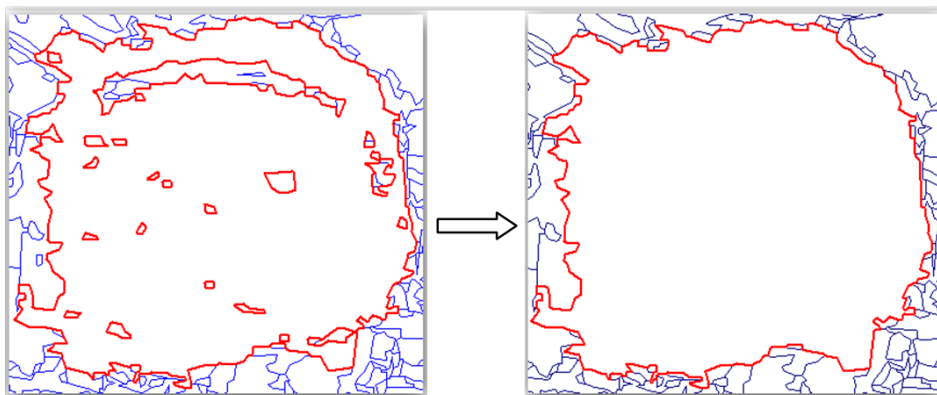


Figure 64: Création d'un polygone plein à partir d'un polygone contenant 20 trous et 15 sous-trous.

Pour expliquer notre démarche, prenons l'exemple du polygone se trouvant à gauche de la Figure 64. En effet, ce polygone représente une forêt de feuillus dense ($BSD = 0.7$) et contient 20 trous. Certains trous contiennent eux-mêmes d'autres trous, ce qui porte le nombre total de trous à 35. Le Tableau 4 montre comment nous calculons le nouveau degré de favorabilité après la suppression de ces trous. Nous nommons cet attribut un degré de favorabilité calculé ($CSD : Computed Suitability Degree$). Notre système va remplacer le degré de favorabilité de base attaché à chacune des cellules de la base de données spatiale. Le calcul est effectué en prenant compte la superficie de chaque trou par rapport à la superficie totale du polygone plein après suppression des trous. Par exemple, il y a 16 trous représentant une forêt mixte et dense ($BSD = 0.65$). La somme des superficies de ces trous est égale à 890025.2 m^2 . Nous calculons une proportion de la superficie ($AP : Area Proportion$) qui est égale au rapport entre la somme des superficies des trous et la surface totale du polygone sans trous ($890025.2/21368279.7 = 0.041651701$). Ensuite, nous calculons une proportion du degré de favorabilité ($SDP : Suitability Degree Proportion$) qui est égale à la multiplication entre AP et BSD ($0.041651701 \times 0.65 = 0.027073606$).

Tableau 4: Calcul du nouveau degré de favorabilité.

Type de couverture du sol	Nb	Code	Superficie (m^2)	BSD	AP	SDP
Eau	1	20	96894.1	0	0.004534483	0
Grands arbustes	3	51	35420.8	0.75	0.001657635	0.001243226
Plantes herbacées	1	100	11284	0.2	0.000528072	0.000105614
Cultures annuelles	1	121	60536.5	0.2	0.002833008	0.000566602
Cultures pérennes et pâturages	1	122	21789.3	0.2	0.001019703	0.000203941
Coniférien - dense	1	211	73049.7	0.6	0.003418605	0.002051163
Feuille - dense	2	221	19914816.3	0.7	0.931980327	0.652386229
Feuille - ouvert	1	222	13574.2	1	0.00063525	0.00063525
Mixte - dense	16	231	890025.2	0.65	0.041651701	0.027073606
Mixte - ouvert	9	232	250889.6	0.95	0.011741217	0.011154156
Total	36		21368279.7		1	0,695419786

De ce fait, la somme de tous les SDP représente en réalité le CSD qui est la nouvelle valeur du degré de favorabilité que nous attachons à la cellule représentant le polygone plein après suppression des trous. Il convient de mentionner que le nombre total de polygones énumérés dans le tableau 5 est de 36 étant donné que nous ajoutons aux 35 trous, le

polygone contenant des trous et qui a bien sûr la plus grosse superficie. En outre et afin de supprimer les trous, notre application (Figure 65) lance une requête spatiale permettant de trouver tous les polygones situés à l'intérieur de celui qui contient les trous. Ensuite, elle fusionne tous ces polygones pour obtenir un polygone plein. Dans notre exemple, le polygone en question a maintenant un degré de favorabilité calculé égal à 0.69542 (au lieu de 0.7). Nous constatons que l'hypothèse mentionnée auparavant est vérifiée étant donné que la favorabilité n'a pas beaucoup diminué bien que nous ayons enlevé quand même un bon nombre de trous.

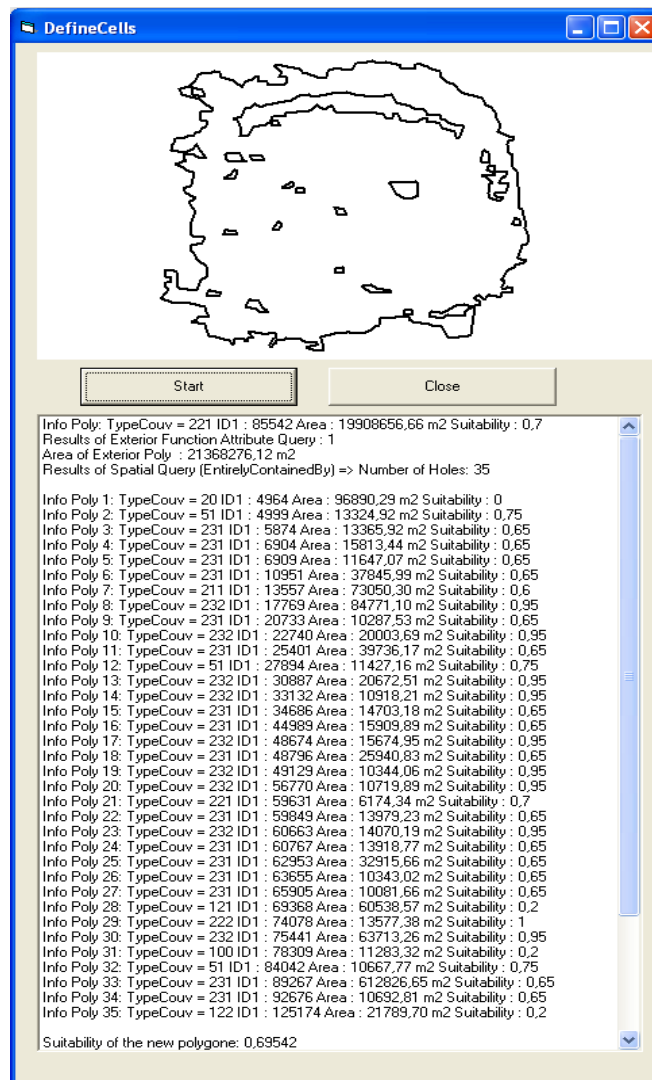


Figure 65: Une fenêtre de l'application *Visual Basic* permettant de faire les requêtes spatiales.

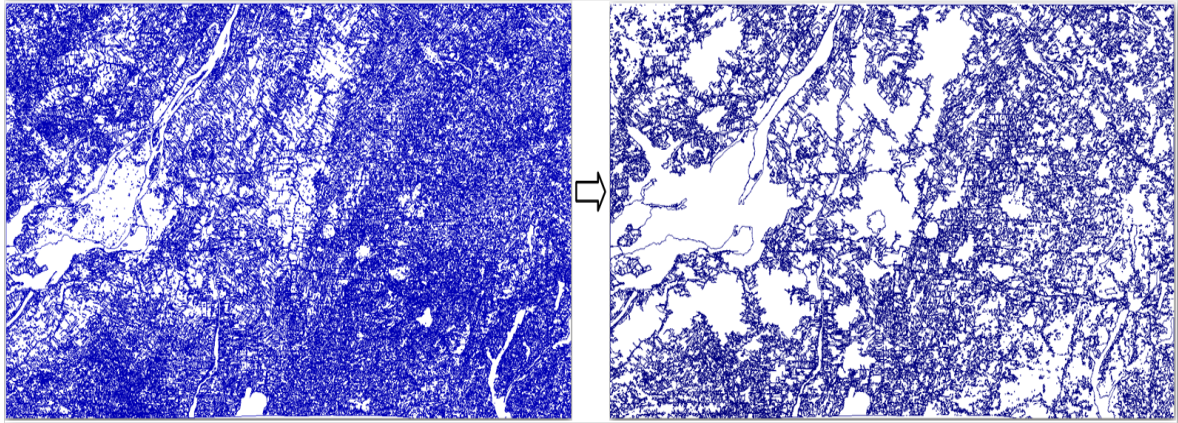


Figure 66: Le résultat de la suppression des trous des polygones de la région 31H.

Par ailleurs, nous avons ajouté une contrainte à notre application afin d'éviter la suppression de trous qui dépassent une certaine superficie et un certain seuil de favorabilité. Cela permet un paramétrage de l'approche afin de l'adapter aux besoins spécifiques de l'utilisateur. Après avoir supprimé les trous des polygones appartenant à la région de 31H (Figure 66), nous avons réussi à éliminer 34 108 trous (25.5% du nombre initial de polygones). Le nombre restant de polygones est donc de 99 672 (74.5%). Il ne reste qu'un seul trou (Figure 67) en tenant compte des seuils que nous avons fixés pour la superficie (10 km^2) et la favorabilité (0,6).

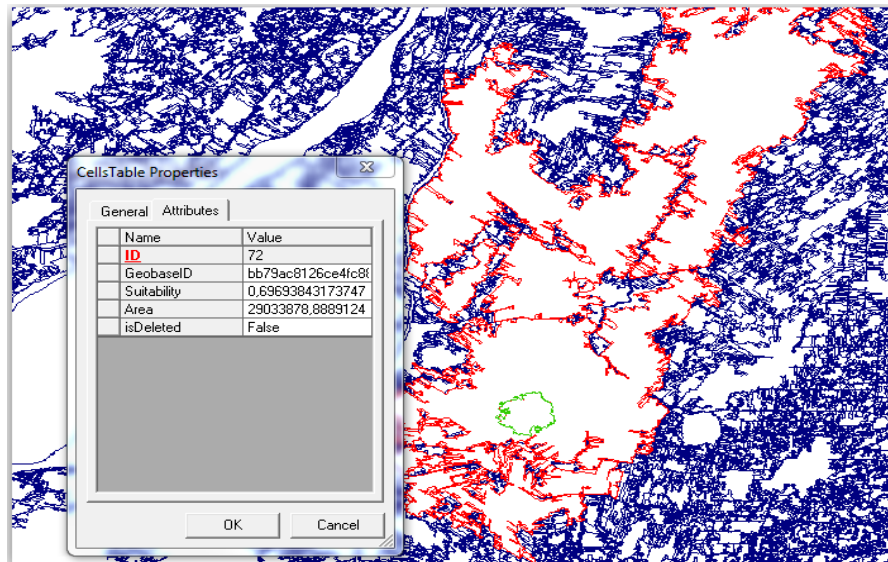


Figure 67: Le polygone (en rouge) contenant le seul trou non supprimé (en vert).

3.4 Fusion progressive des cellules

Après la suppression des trous pour la région 31H, le système stocke le résultat dans la base de données spatiale. Puis, le processus de fusion des cellules s'exécute (en mettant à jour cette base de données) en différentes étapes en utilisant différents intervalles de favorabilité (c.-à-d. *ATI*). Dans une première étape, le système interroge la base de données pour avoir la liste des polygones ayant un degré de favorabilité égal à 1 (100% de favorabilité). Ensuite, le système applique aux 6 261 polygones obtenus un tri décroissant selon leurs superficies. Nous avons expliqué auparavant que ce tri est nécessaire pour commencer le traitement par les meilleurs candidats à la fusion, et de ce fait, obtenir une solution proche de l'optimale. Ainsi, le premier polygone traité sera celui ayant la meilleure favorabilité par rapport à l'habitat de l'espèce étudiée et la plus grande superficie. Pour chacun de ces polygones, le système applique une requête spatiale afin de trouver ses voisins immédiats. Ces voisins sont triés dans l'ordre décroissant en fonction de leurs degrés de favorabilité ainsi que de leurs superficies. Par conséquent, le premier voisin traité sera celui ayant également la meilleure favorabilité et la plus grande superficie. Ce voisin représente donc le meilleur premier candidat permettant de le fusionner au polygone en cours de traitement. Le système essaye donc de fusionner à ce polygone ses voisins l'un après l'autre, jusqu'à obtenir un polygone dont le degré de favorabilité ne descend pas en dessous d'un seuil de favorabilité choisi de 0.9 (ce seuil permet ainsi d'arrêter cette étape de fusion).

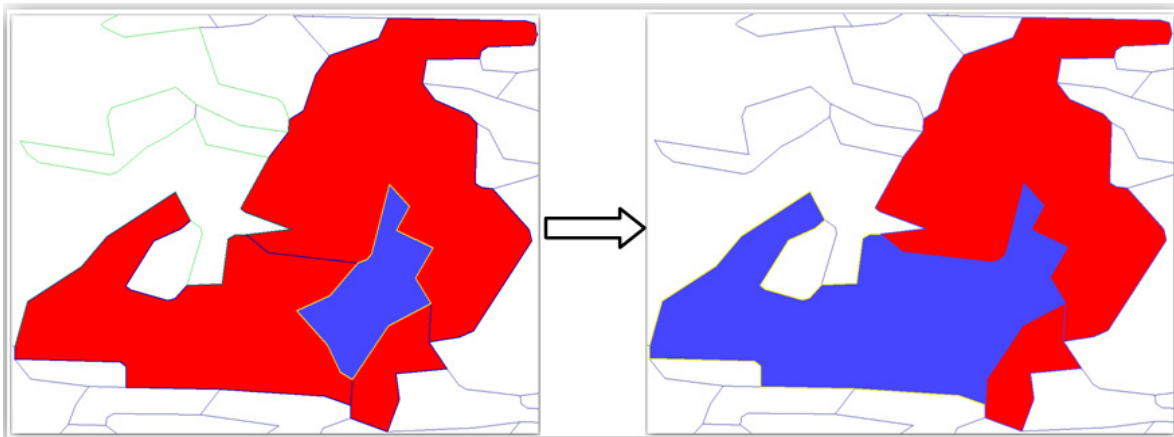


Figure 68: Fusion avec les cellules voisines en utilisant un seuil de favorabilité.

Après ce premier passage de fusion, 7 137 polygones ont été supprimés de la base de données. Ainsi, jusqu'à présent, nous avons réussi à éliminer 41 245 (30.8%) des polygones et le nombre restant de polygones est de 92 535 (69.2% des polygones GéoBase initiales concernant la région 31H). La Figure 68 montre (à gauche) un polygone (en bleu) avec trois voisins (en rouge). Ce polygone est fusionné à l'un de ses voisins (à droite) et le polygone obtenu conserve un degré de favorabilité supérieur ou égal au seuil choisi. La Figure 69 montre le résultat de la première itération après ce premier passage de fusion dans une partie de la région 31H. Les polygones rouges sont les polygones initiaux dont la favorabilité est à 100% alors que les polygones verts sont ceux qui ont fusionné.

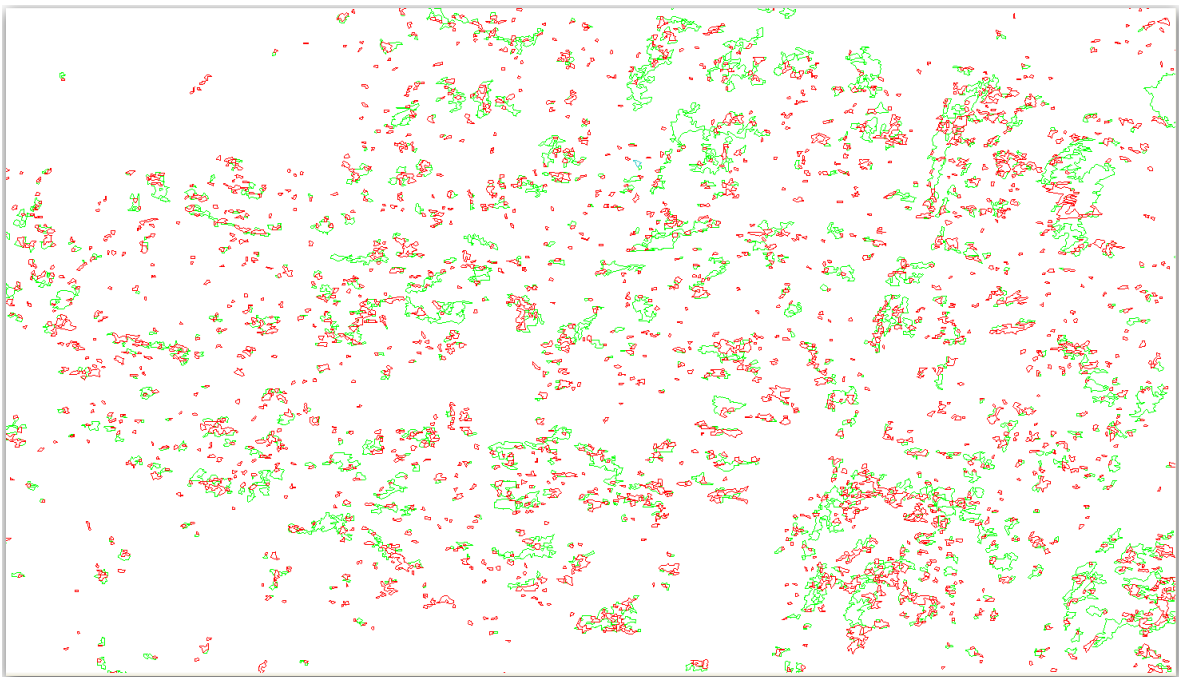


Figure 69: Première itération du processus de fusion progressive.

Après ce premier passage, le système applique le processus de fusion par étapes successives, en augmentant l'intervalle (c.-à-d. *ATI*) de favorabilité (utilisée pour sélectionner les polygones initiaux qui devraient être traités) et en diminuant le seuil de favorabilité (utilisé pour arrêter la fusion). En fait, dans le processus suivant, les polygones utilisés sont ceux qui ont un degré de favorabilité appartenant à l'intervalle $[1, 0.9]$. De plus, le seuil de favorabilité utilisé est celui permettant l'arrêt de la fusion si le degré de favorabilité descend en bas de 0.8. Pour les prochaines itérations de fusion, nous avons

utilisé respectivement les intervalles et les seuils suivants: $([1, 0.8], 0.7)$, $([1, 0.7], 0.6)$, $([1, 0.6]; 0.5)$. Ce choix d'intervalles et de seuils est basé sur le fait que nous cherchons à obtenir des polygones ayant la plus grande taille possible sans pour autant réduire considérablement leur favorabilité. Étant donné que chaque polygone est fusionné avec ses meilleurs voisins, la favorabilité du polygone résultant ne baissera pas énormément. D'ailleurs, les seuils choisis (0.9, 0.8, 0.7, etc.) permettent de réduire la favorabilité des polygones d'une façon progressive. Ces seuils font en sorte que la fusion s'arrête à chaque fois que le système prévoit l'agrégation d'un voisin ayant une favorabilité relativement faible. D'un autre côté, ces intervalles et seuils sont paramétrables par l'utilisateur afin de modifier le résultat obtenu selon ses besoins.

Après l'achèvement de la fusion progressive des polygones favorables, le système applique un autre processus dont le but est de fusionner les polygones non-favorables tel que mentionné dans la Section 3.1. En effet, nous cherchons à obtenir également avec ce processus, les polygones non favorables ayant la plus grande taille possible. De ce fait, le nombre de cellules obtenues après toutes les fusions sera encore réduit. Pour ce faire, le système sélectionne, en utilisant une requête spatiale, les polygones ayant un degré de favorabilité égal à zéro (0% de favorabilité). Ces polygones seront par la suite triés en ordre décroissant par rapport à leurs superficies. Par conséquent, le premier polygone non favorable traité sera celui ayant la plus grande superficie. Le système essaye ensuite de fusionner ce polygone avec ses voisins qui ont un degré de favorabilité inférieur à 0.5. Nous utilisons cette valeur, car elle représente le dernier seuil de favorabilité (permettant d'arrêter la fusion) qui est utilisé par le processus de fusion de polygones favorables.

Les pourcentages des cellules enlevées après chacun des traitements de fusion appliqués sont présentés dans la Figure 70 (le premier processus est la suppression des trous alors que le dernier processus est la fusion de polygones non favorables). Nous constatons qu'à l'issue de la fusion des polygones favorables, le nombre des cellules est diminué de 69% et

après la fusion des polygones non favorables, le pourcentage de réduction du nombre des cellules passe à 83%.

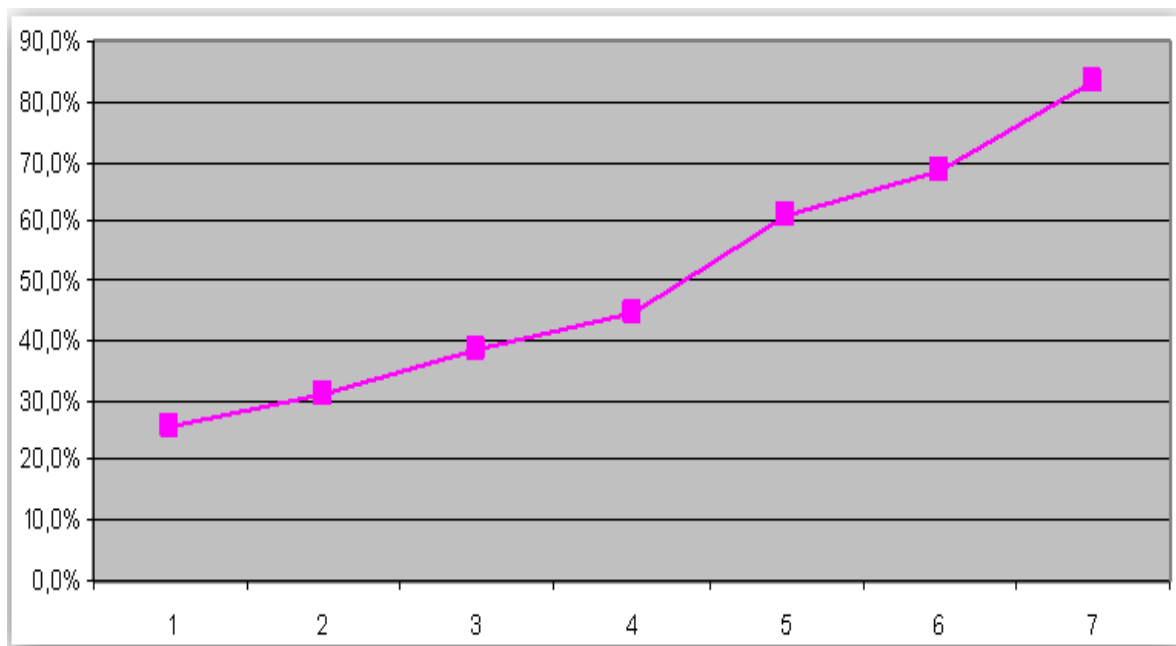


Figure 70: Pourcentages des cellules supprimées après les différentes itérations de fusion.

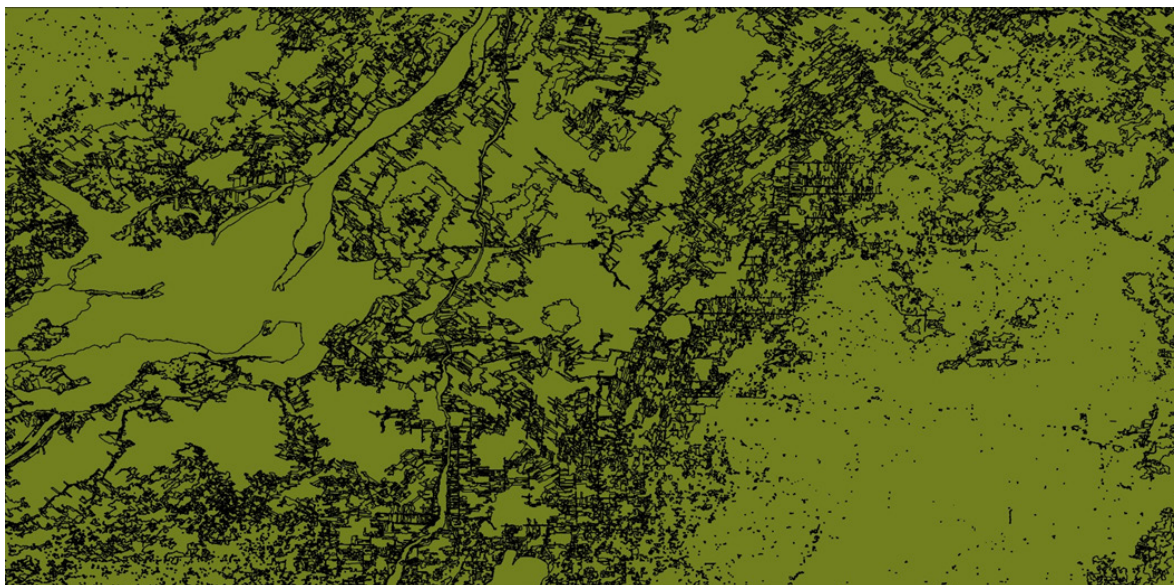


Figure 71: Résultat final de la fusion progressive des polygones de la région 31H.

Finalement et après la fusion des cellules de GéoBase en utilisant les différents processus mentionnés auparavant, il reste seulement 17% (22 698) du nombre des polygones initiaux appartenant à la région 31H (Figure 71). Ce résultat nous semble satisfaisant étant donné qu'il nous permettra de limiter le nombre de modèles compartimentaux à cloner et de ce fait, utiliser beaucoup moins de ressources computationnelles lors de la simulation de la propagation d'une zoonose utilisant l'*IVGE*.

3.5 Enrichissement par l'ajout des données dans les cellules

Dans le but d'optimiser la géosimulation de la propagation des zoonoses, nous proposons d'enrichir l'*IVGE* en ajoutant des informations supplémentaires dans chaque polygone. Nous présentons chaque type de ces informations dans les sous-sections suivantes.

3.5.1 Données biologiques et climatiques

Nous pensons qu'il serait intéressant d'attacher aux cellules les données de populations des espèces qui nous intéressent (p.ex. le nombre des individus). Ces données sont en réalité des attributs de cellules comme nous l'avons déjà présenté dans le modèle *MASTIM* (voir Section 3.7 du Chapitre V). De plus, l'ajout de données climatiques (p.ex. température moyenne, degré jour cumulé, etc.) dans ces cellules pourrait améliorer la plausibilité de la simulation surtout que le climat influence énormément l'évolution, ainsi que l'infection des espèces impliquées dans la propagation des zoonoses (p.ex. la température influence la durée de vie des *Culex*, responsables de la propagation du VNO, qui peut varier d'une semaine à plusieurs mois [Madder *et al.*, 1983]).

3.5.2 Identification des voisins et leurs orientations

Nous proposons d'attacher aux cellules certaines informations topologiques. En effet, nous avons constaté que certains processus biologiques sont très importants lors de la propagation d'une zoonose. Par exemple, la migration printanière des oiseaux (qui ont des tiques juvéniles accrochées à eux) peut avoir une influence significative sur la propagation de la maladie de Lyme. De ce fait, nous avons pensé à trouver une solution efficace afin de

simuler la distribution des oiseaux migrateurs sur le territoire et par la même occasion de simuler l'apport des tiques du Sud des États-Unis vers le Nord (c.-à-d. les régions de Québec qui nous intéressent comme la Montérégie). Afin de réaliser cette distribution (voir Section 4), nous pensons qu'il est important d'attacher aux cellules des informations pertinentes comme par exemple les identificateurs ainsi que les orientations de leurs voisins. Ainsi, lors du traitement d'une cellule, notre système pourra utiliser ces informations pour distribuer les individus se trouvant dans cette cellule vers ses voisins. Par exemple, des individus se trouvant dans une cellule peuvent être distribués vers une cellule voisine si l'orientation de cette dernière est proche de la direction du corridor suivi par ces individus.

Afin d'attacher ces informations topologiques aux cellules, notre système ajoute à chaque polygone après fusion un nouvel attribut (nommé *Neighbors*) qui contient entre autres la liste des identifiants de ses voisins. Pour ce faire et après toutes les étapes de fusion présentées auparavant (chaque cellule a maintenant un identifiant unique et final), le système affecte à l'attribut *Neighbors* de chaque polygone la liste des identifiants de ses voisins immédiats en utilisant des requêtes spatiales (les mêmes requêtes effectuées pour le processus de fusion progressif). De plus, le système ajoute à cet attribut l'orientation de chaque voisin par rapport au polygone en cours de traitement. En fait, nous proposons de calculer deux types de données d'orientation : (1) une information quantitative représentant une valeur exacte en radian (c.-à-d. un angle comme par exemple 4.3121) et (2) une information qualitative représentant une valeur floue stockée en utilisant une chaîne de caractères (c.-à-d. une direction présélectionnée obtenue selon un cercle trigonométrique comme par exemple "SW"). L'utilisation de l'une ou de l'autre (c.-à-d. orientation quantitative ou qualitative) lors de la distribution des individus dépend de la précision recherchée pour les résultats de la simulation en fonction de la qualité des données disponibles. Pour ce faire, le système trace une ligne virtuelle qui passe par le centroïde d'un polygone donné et le centroïde de son voisin. Ainsi, le système calcule un angle en radian en utilisant une fonction spatiale qui retourne l'azimut (c.-à-d. *FA: Forward Azimuth*) de la ligne tracée par rapport à l'axe du Nord (voir Figure 72a). La valeur

retournée, représentant cette information quantitative, est normalisée pour être comprise entre zéro et 2π en utilisant le sens des aiguilles d'une montre. Par la suite, le système utilise cette valeur en radian pour calculer l'information qualitative qui représente l'une des quatre directions cardinales (c.-à-d. Nord, Est, Sud, Ouest) ou l'une des quatre directions ordinales (c.-à-d. Nord-Est, Sud-Est, Sud-Ouest, Nord-Ouest) ou l'une des huit autres subdivisions représentées dans la Figure 72b.

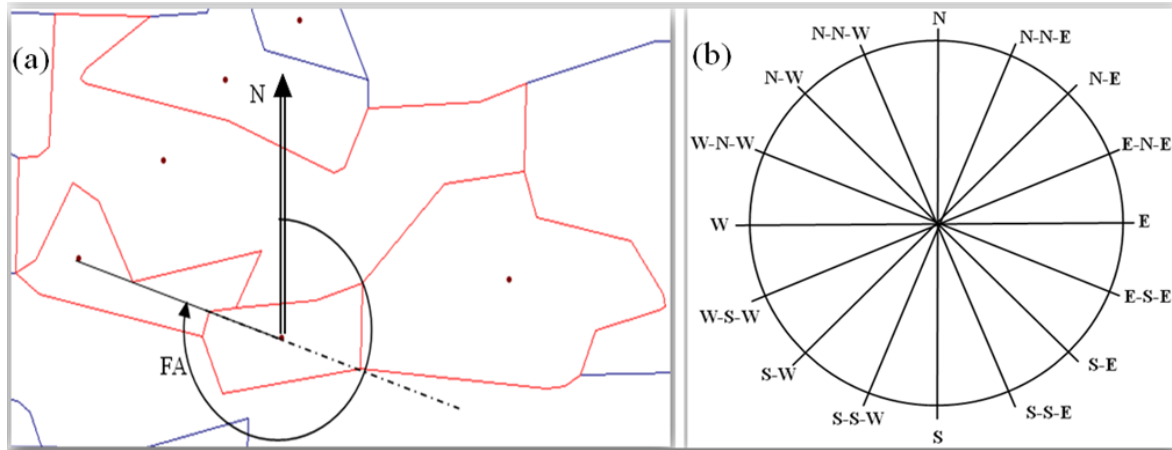


Figure 72: Information quantitative (a) et qualitative (b) utilisée pour stocker l'orientation des voisins.

La fonction que notre système utilise pour faire ce calcul est la suivante (les valeurs sont affichées en degré pour faciliter leur lecture, mais ils sont convertis en radian):

```
Orientation = Qualitative-Information-Function (Angle : Information quantitative){
  SI (Angle >= 348.75 ET Angle < 11.25) ALORS QAngle = "N"
  SINON SI (Angle >= R(11.25) ET Angle < R(33.75)) ALORS Orientation = "NNE"
  SINON SI (Angle >= R(33.75) ET Angle < R(56.25)) ALORS Orientation = "NE"
  SINON SI (Angle >= R(56.25) ET Angle < R(78.75)) ALORS Orientation = "ENE"
  SINON SI (Angle >= R(78.75) ET Angle < R(101.25)) ALORS Orientation = "E"
  SINON SI (Angle >= R(101.25) ET Angle < R(123.75)) ALORS Orientation = "ESE"
  SINON SI (Angle >= R(123.75) ET Angle < R(146.25)) ALORS Orientation = "SE"
  SINON SI (Angle >= R(146.25) ET Angle < R(168.75)) ALORS Orientation = "SSE"
  SINON SI (Angle >= R(168.75) ET Angle < R(191.25)) ALORS Orientation = "S"
  SINON SI (Angle >= R(191.25) ET Angle < R(213.75)) ALORS Orientation = "SSW"
  SINON SI (Angle >= R(213.75) ET Angle < R(236.25)) ALORS Orientation = "SW"
  SINON SI (Angle >= R(236.25) ET Angle < R(258.75)) ALORS Orientation = "WSW"
  SINON SI (Angle >= R(258.75) ET Angle < R(281.25)) ALORS Orientation = "W"
  SINON SI (Angle >= R(281.25) ET Angle < R(303.75)) ALORS Orientation = "WNW"
  SINON SI (Angle >= R(303.75) ET Angle < R(326.25)) ALORS Orientation = "NW"
  SINON SI (Angle >= R(326.25) ET Angle < R(348.75)) ALORS Orientation = "NNW"
}
```

Par conséquent, l'attribut *Neighbors* aura le format suivant :

$(idV_1, angleV_1, orientationV_1), \dots, (idV_n, angleV_n, orientationV_n)$

Sachant que idV_1 est l'identifiant du premier voisin V_1 , $angleV_1$ est l'information quantitative représentant l'angle en radian (l'azimut entre V_1 et le polygone traité), $orientationV_1$ est l'information qualitative représentant l'une des directions (une chaîne de caractères) mentionnées auparavant et n est le nombre de voisins. La Figure 73 montre des exemples d'enregistrement de l'attribut *Neighbors*.

ID	Suitability	Area	Neighbors	PropLandCover
1	0,202887559	15547483,4	(2, 4.3121, WSW),(9, 1.6004, E),(13, 0.9804, NE),(22, 4.8994, W),(27, 3.8128, SW),(32, (34,0.5642),(100,0.2133),(121,1.6850),(122,96.8380),(221,0.3801),(222,0.1	
2	0,0371411764	16789453,4	(1, 1.1705, ENE),(6, 0.0312, N),(14, 3.2824, S),(15, 2.0224, ESE),(17, 1.3076, ENE),(25, 4	(20,0.0610),(33,1.0245),(34,87.0748),(100,0.6728),(121,2.5878),(122,5.860
3	0,6535072137	10783558,9	(11, 0.5279, NNE),(16, 1.3791, E),(28, 1.7427, E),(112, 5.0248, WNW),(117, 0.9778, NE)	(221,2.5965),(231,96.6672),(232,0.7363)
4	0,7	8046575,6	(39, 2.5608, SSE),(41, 1.1873, ENE),(151, 4.6339, W),(233, 4.4710, WSW),(434, 4.6794,	(221,100)
5	0,1987261221	7588276,29	(47, 3.8463, SW),(55, 6.0842, NNW),(58, 0.7419, NE),(91, 5.2537, WNW),(99, 5.1841, V	(34,0.9692),(100,0.1438),(121,0.9423),(122,97.7969),(231,0.1477)
6	0,2191310335	7964106,38	(2, 3.1728, S),(17, 2.5212, SE),(22, 1.8521, ESE),(83, 4.5980, W),(84, 4.9829, WNW),(86	(34,1.0867),(51,0.4788),(100,0.1873),(121,0.8702),(122,93.5367),(221,2.78
7	0,6484620397	7500127,67	(23, 0.2516, NNE),(40, 1.5814, E),(144, 2.2037, SE),(240, 1.1461, ENE),(241, 2.2948, SE)	(100,0.2923),(122,0.1843),(221,1.2134),(231,98.3100)
8	0,6575324798	6248883,74	(23, 4.3660, WSW),(47, 0.2878, NNE),(56, 1.3973, E),(73, 2.5608, SSE),(87, 1.0091, ENE	(221,2.2326),(222,0.1876),(231,95.6599),(232,1.9199)
9	0,6959059393	6052971,48	(1, 4.7420, W),(21, 1.1596, ENE),(63, 4.9957, WNW),(96, 0.1486, N),(106, 4.3204, WSW	(20,0.2725),(34,0.1779),(51,0.4868),(100,0.2133),(221,97.4184),(222,0.170
10	0,1996583638	5372499,25	(12, 1.0352, ENE),(20, 2.8647, SSE),(93, 5.2336, NNW),(114, 0.4683, NNE),(125, 4.0082	(34,0.1708),(122,99.8292)
11	0,6976353373	4682024,43	(3, 3.6695, SSW),(24, 0.9755, NE),(42, 4.4474, WSW),(43, 0.1210, N),(112, 4.0015, SW)	(33,0.2588),(51,0.2787),(221,98.0772),(231,1.3853)
12	0,65940077454	4936103,37	(10, 4.1768, WSW),(15, 4.5852, W),(61, 5.0437, NNW),(114, 4.4040, WSW),(124, 4.772	(221,0.9452),(231,96.0764),(232,2.9784)
13	0,7	4299941,35	(1, 4.1220, SW),(210, 2.9541, S),(264, 4.1188, SW),(494, 4.1889, WSW),(509, 4.3332, W	(221,100)
14	0,6613168700	4446469,99	(2, 0.1408, N),(45, 2.1755, SE),(49, 4.7391, W),(50, 3.8674, SW),(53, 4.3930, WSW),(65,	(122,0.3781),(221,2.2702),(222,1.3427),(231,93.6144),(232,2.3946)
15	0,2149877864	3764748,26	(2, 5.1640, WNW),(12, 1.4436, E),(26, 4.7575, W),(30, 5.1934, WNW),(67, 5.0717, WNW	(34,1.0782),(100,1.3399),(121,2.0597),(122,92.7910),(231,1.1135),(232,1.6
16	0,1993309954	3411193,15	(3, 4.5207, W),(18, 1.8334, ESE),(28, 4.1543, WSW),(34, 0.3810, NNE),(97, 1.0293, ENE	(34,0.3345),(121,3.1420),(122,96.5235)
17	0,6529151409	3310199,58	(2, 4.4492, WSW),(6, 5.6628, NW),(86, 5.8424, NNW),(88, 4.6848, W),(98, 0.7010, NE),	(221,1.7279),(222,0.3039),(231,97.6390),(232,0.3292)
18	0,6532963487	3002228,34	(16, 4.9750, WNW),(33, 2.1245, ESE),(97, 0.4809, NNE),(199, 1.4817, E),(203, 1.0949, E	(221,0.7938),(231,98.2397),(232,0.9665)
19	0,65	2930541,82	(29, 4.0578, SW),(81, 2.4226, SE),(111, 4.5668, W),(217, 2.0444, ESE),(472, 4.3215, WSV	(231,100)
20	0,6490568834	2947154,32	(10, 6.0063, NNW),(33, 4.1879, WSW),(60, 3.5686, SSW),(169, 2.6544, SSE),(204, 4.090	(122,2.0761),(221,1.2266),(222,1.1429),(231,94.2925),(232,1.2619)
21	0,2	2625370,8	(9, 4.3012, WSW),(41, 2.2994, SE),(96, 5.0623, WNW),(210, 5.2089, WNW),(417, 5.032	(122,100)
22	0,6470042387	2816371,68	(1, 1.7578, E),(6, 4.9937, WNW),(130, 3.5768, SSW),(170, 4.4537, WSW),(235, 4.3675, V	(122,3.7825),(221,1.9194),(222,0.3772),(231,90.0057),(232,3.9152)
23	0,2	2337475,41	(7, 3.3932, SSW),(8, 1.2244, ENE),(133, 0.1424, N),(142, 6.1227, N),(249, 6.2802, N),(2	(122,100)
24	0,6585145726	2288534,67	(11, 4.1171, SW),(26, 1.2747, ENE),(52, 5.1293, WNW),(57, 1.8181, ESE),(117, 2.7178, S	(221,4.8985),(222,1.7330),(231,93.3686)
25	0,1933659057	2243511,42	(2, 1.8222, ESE),(48, 2.6909, SSE),(58, 6.0197, NNW),(94, 0.4201, NNE),(95, 2.2095, SE)	(34,3.3170),(121,2.4126),(122,94.2703)
26	0,6994322638	2102064,01	(15, 1.6159, E),(24, 4.4163, WSW),(57, 3.1050, S),(67, 6.2109, N),(82, 2.3101, SE),(140,	(221,98.8645),(231,1.1355)
27	0,6623898399	2164503,25	(1, 0.6712, NE),(2, 4.7167, W),(66, 1.9153, ESE),(158, 1.4750, E),(160, 4.0024, SW),(162	(221,2.5834),(231,93.7175),(232,3.6991)
28	0,7	1908867,86	(3, 4.8843, W),(16, 1.0127, ENE),(166, 0.1697, N),(488, 1.0942, ENE),(568, 0.9682, NE),	(221,100)
29	0,2074838800	1916370,4	(19, 0.9163, NE),(58, 1.8595, ESE),(69, 2.1330, ESE),(91, 4.1549, WSW),(111, 6.0922, N)	(52,0.4389),(122,98.8851),(232,0.6760)
30	0	1796370,66	(2, 5.1394, WNW),(15, 2.0518, ESE),(100, 1.4549, E),(181, 1.7392, E),(274, 1.6126, E),(3	(20,100)
31	0,6617470515	2016121,47	(40, 3.8848, SW),(46, 1.9232, ESE),(104, 2.6289, SSE),(146, 5.5916, NW),(214, 5.1172, V	(221,11.8343),(222,0.5756),(231,86.3184),(232,1.2718)
32	0,6558480538	1849495,68	(1, 5.7690, NNW),(2, 4.8571, W),(36, 1.4037, E),(100, 4.0416, SW),(221, 4.4983, WSW),	(81,0.9113),(221,2.5410),(222,0.5669),(231,93.7494),(232,2.2314)
33	0,2	1849495,68	(18, 5.2661, WNW),(20, 1.0463, ENE),(199, 0.1411, N),(204, 1.1796, ENE),(342, 5.5350,	(122,100)
34	0,6742190914	1813987,03	(16, 3.5225, SSW),(57, 0.2304, NNE),(117, 5.3692, NW),(134, 1.5388, E),(174, 4.0363, S	(221,0.7268),(222,0.7980),(231,91.4543),(232,7.0209)
35	0,2092768347	1732137,02	(55, 4.3670, WSW),(58, 2.3056, SE),(155, 5.0577, WNW),(201, 3.3203, S),(206, 3.0862,	(121,2.7814),(122,95.3632),(221,1.8554)

Figure 73: Enrichissement de l'IVGE par l'ajout des données dans les cellules.

3.5.3 Informations de couverture du sol

Nous proposons de garder une trace de toutes les fusions effectuées en ajoutant à chaque cellule un attribut (nommé *PropLandCover*) contenant la liste des polygones fusionnés en incluant leur proportion de superficie et leur type de couverture du sol. Cet attribut est formé de n couples de la façon suivante :

$(TypeLandCover_1, Proportion_1), \dots, (TypeLandCover_n, Proportion_n)$

Sachant que $TypeLandCover_1$ est le type de couverture du sol selon Géobase et $Proportion_1$ représente le pourcentage de ce type de couverture du sol par rapport à la superficie totale de la cellule. Il faut bien sûr que la somme de ces pourcentages soit égale à 100% (voir le champ $PropLandCover$ dans la Figure 73). La sauvegarde de ces informations dans la base de données spatiale permet à l'utilisateur d'avoir une idée sur la signification du degré de favorabilité calculé (c.-à-d. BSD) en consultant les pourcentages des différents types de couverture du sol appartenant aux polygones qui ont participé à la fusion.

3.5.4 Informations sur la forme des cellules

Nous savons que les cellules représentant des polygones de couverture du sol ont une forme très complexe (voir Figure 62). De ce fait, la distance nécessaire pour que des individus traversant une cellule afin d'atteindre un des voisins de cette cellule peut changer significativement en fonction de la direction choisie. Pour cette raison, nous proposons un autre enrichissement de l' $IVGE$ en ajoutant à chaque cellule une information reflétant sa forme.

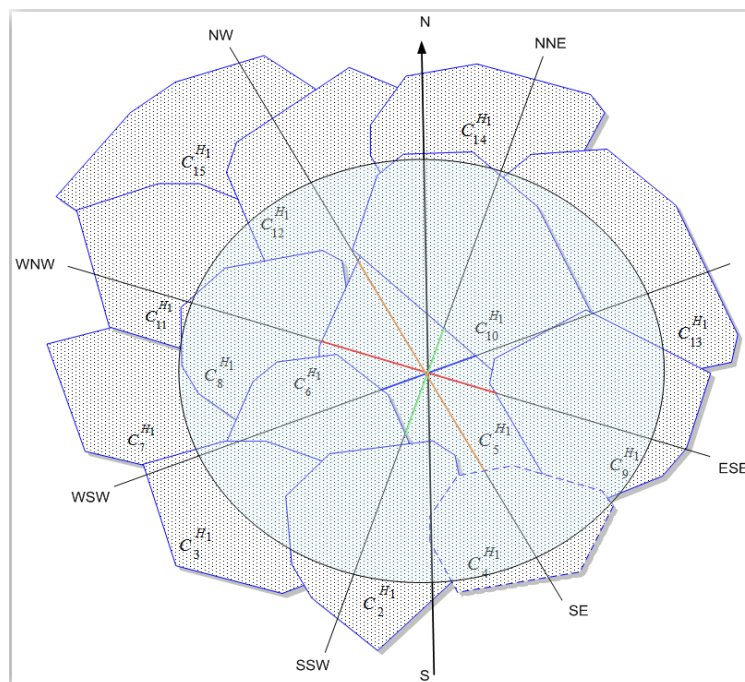


Figure 74: Calcul des distances pour traverser une cellule dans les différentes directions de ses voisins.

En effet, le système utilise les lignes virtuelles qu'il a déjà tracées pour calculer l'orientation des voisins, mais cette fois-ci pour calculer des portions en mètre, dont les deux extrémités représentent les intersections entre ces lignes et le contour de la cellule en cours de traitement (Figure 74). La sauvegarde de ces informations dans la base de données spatiale nous permettra entre autres de faire une estimation du temps nécessaire pour qu'un individu puisse traverser un polygone dans la direction d'un de ses voisins.

3.5.5 Informations sur les trajectoires

Nous proposons d'ajouter dans l'*IVGE* des informations sur les trajectoires de certaines espèces impliquées dans la propagation d'une zoonose afin de faciliter la distribution des flots d'individus (p.ex. la migration des oiseaux qui sont considérés comme un des plus importants réservoirs de plusieurs zoonoses). Nous détaillons cet aspect dans la section 4.1.2.

4 Application de l'*IVGE*: Distribution spatiale des flots d'individus

Nous présentons dans cette section une application de l'*IVGE* qui nous semble importante. Il s'agit de la distribution spatiale des flots d'individus. En effet, l'enrichissement de l'*IVGE* avec entre autres des données biologiques et topologiques peut faciliter la modélisation d'une variété de processus qui ont un effet significatif sur la propagation des zoonoses. Nous avons présenté dans la section 3.11 du Chapitre V les différents types de processus qui nous intéressent. La migration des oiseaux est un des plus importants processus naturels qui influence la propagation de certaines zoonoses comme la maladie de Lyme. En effet, les oiseaux importent des tiques juvéniles (certaines d'entre elles étant infectées) au Canada à partir des États-Unis. Ainsi, la modélisation de la mobilité de ces oiseaux est nécessaire pour simuler la propagation de la maladie de Lyme. D'ailleurs, nous pensons qu'il serait intéressant de modéliser les migrations printanières comme des vagues qui distribuent à travers les cellules, des oiseaux de diverses espèces portant des tiques. En fait, les tiques juvéniles s'accrochent à certaines espèces d'oiseaux comme les Merles

d'Amérique provenant des États-Unis. Une bonne partie de ces tiques se laissent tomber après que les oiseaux arrivent dans le sud du Québec. Si ces tiques tombent dans une zone favorable à leur survie, elles vont pouvoir s'établir dans cette zone. De ce fait, les tiques juvéniles vont évoluer en des tiques adultes. De plus, si ces tiques sont infectées par l'agent pathogène (c.-à-d. bactérie), elles peuvent propager la maladie de Lyme [Ogden *et al.*, 2008b]. Ainsi, la migration des oiseaux joue un rôle prépondérant dans la propagation de la maladie et c'est pour cette raison que nous sommes intéressés par ce phénomène en modélisant la distribution des flots d'individus à travers les cellules. Afin de faire cette modélisation en bénéficiant des avantages de l'enrichissement de l'*IVGE* (voir Section 3.5), nous proposons d'ajouter à *MASTIM* (voir Section 3 du Chapitre V) d'autres concepts pour formaliser le transfert spatial des individus d'une cellule à une autre. Nous présentons ces concepts dans la section suivante. Par la suite, nous présentons une implémentation de la distribution spatiale des flots d'individus afin de tester notre approche.

4.1 Formalisme du transfert spatial des individus

Nous présentons dans cette section les notations formelles pour chacun des aspects importants concernant le transfert spatial des individus. En effet, nous formalisons dans les sous-sections suivantes le patron de mobilité qui va servir à instancier des trajectoires. Ces dernières forment un réseau de mobilité permettant à un flot d'individus d'être distribué à travers les cellules de l'*IVGE*. Ces nouvelles notations complètent le formalisme du modèle *MASTIM* dont la liste complète se retrouve dans l'annexe A.

4.1.1 Le patron de mobilité

Un patron de mobilité est une forme récurrente permettant de caractériser d'une façon générique la mobilité d'un individu, d'un groupe d'individus ou d'une population. Généralement, un patron de mobilité découle d'une étude statistique sur les données dont on dispose afin de mettre en évidence la récurrence de ce patron (comportement moyen de mobilité). Le terme de mobilité quant à lui, concerne le mouvement dans l'espace géographique d'individus, de biens ou d'informations [Chardonnel *et al.*, 2004]. Le patron

est représenté par une séquence de symboles [Wilson, 2008]. Chaque symbole représentant un état de l'individu au cours de ses déplacements ou de ses activités. Par exemple, un patron de mobilité pour les humains «Ménage-Déplacement-Travail» caractérise des individus ayant une séquence d'activités similaire. Nous pouvons définir des patrons simples comme par exemple: Arrêt (pas de mobilité), Départ-Déplacement-Arrivée (mobilité sans faire un arrêt), Départ-Déplacement-Arrêt-Déplacement-Arrivée (mobilité en faisant un arrêt). Pour le cas du phénomène qui nous intéresse, le départ, l'arrêt ou l'arrivée doivent normalement se faire dans une zone favorable. Par exemple, les chevreuils qui transportent les tiques adultes se déplacent de la zone d'hivernage à une zone de pâturage en suivant leur biocorridor²⁶. Ils cherchent notamment à se déplacer de, et vers, les forêts de feuillus ou mixtes (feuillus et conifères). Ainsi, nous définissons un patron de mobilité d'une espèce de la façon suivante :

$$MobilityPattern_m^{S_p} = (Id_{MP_m^{S_p}}, \{StartingAreaType - StartingPeriodType - ActivityType - TranscientAreaType - DestinationAreaType - DestinationPeriodType \}, Parameters_{MP_m^{S_p}})$$

$Id_{MP_m^{S_p}}$ représente le nom et/ou l'identificateur du patron de mobilité. $StartingAreaType$ représente un type de zone de départ (p.ex. Sud des États-Unis), $StartingPeriodType$ représente un type de la période de départ (p.ex. début du printemps), $ActivityType$ représente une activité type qui est en fait l'activité principale que font les individus de l'espèce pendant leur déplacement de la zone de départ vers la zone d'arrivée (p.ex. vol en groupe). $TranscientAreaType$ représente un type de zone de transit (p.ex. une zone de repos). $DestinationAreaType$ représente un type d'une zone d'arrivée (p.ex. Sud du Québec). $DestinationPeriodType$ représente un type de la période d'arrivée (p.ex. fin du printemps). $Parameters_{MP_m^{S_p}}$ représentent par exemple les contraintes sur la zone de départ et la zone d'arrivée ou les autres zones autorisées pour les activités. Ces contraintes sont généralement établies par rapport au degré de favorabilité de ces zones. Nous pensons également à des paramètres de déplacement comme par exemple la distance maximale que

²⁶ L'expression « corridor biologique » (ou biocorridor) désigne un ou des milieux reliant fonctionnellement entre eux différents habitats vitaux pour une espèce, une population, une métapopulation ou un groupe d'espèces ou métacommunauté.

les individus de l'espèce peuvent parcourir en une journée. Pour des fins d'illustration, prenons l'exemple de deux patrons de mobilité pour les corneilles qui migrent du sud au printemps et se regroupent en dortoirs en été (mi-juillet) et qui vont se déplacer pendant la journée pour chercher de la nourriture et revenir la fin de la journée à ces dortoirs:

$$MobilityPattern_1^{S_{Crows}} = (Id_{MP_1^{S_{Crows}}}, \{SouthUSA - EarlySpring - Migration - RestArea - SouthQuebec - LateSpring\}, Parameters_{MP_1^{S_{Crows}}})$$

$$MobilityPattern_2^{S_{Crows}} = (Id_{MP_2^{S_{Crows}}}, \{Summer - Roost - Foraging - Roost\}, Parameters_{MP_1^{S_{Crows}}})$$

4.1.2 La trajectoire

Une trajectoire est l'instanciation d'un patron de mobilité en spécifiant la liste des cellules visitées ou survolées par les individus de l'espèce concernée. Elle est notée :

$$Trajectory_m^{S_p} = (Id_{Trajectory_m^{S_p}}, \{C_{src}^{H_i}, \dots, C_{dest}^{H_i}\}, MobilityPattern_n^{S_p}, Extent_{Trajectory_m^{S_p}})$$

$Id_{Trajectory_j^{S_p}}$ représente le nom et/ou l'identificateur de la trajectoire. $\{C_{src}^{H_i}, \dots, C_{dest}^{H_i}\}$ représente l'ensemble des cellules effectivement traversées par la trajectoire à partir de la cellule de départ ($C_{src}^{H_i}$) jusqu'à la cellule d'arrivée ($C_{dest}^{H_i}$). $MobilityPattern_n^{S_p}$ est le patron de mobilité utilisé pour instancier la trajectoire des individus de l'espèce S_p .

De plus, nous introduisons la notion d'extension de la trajectoire ($Extent_{Trajectory_m^{S_p}}$) qui représente une zone tampon de l'axe (traversant les cellules visitées) de cette trajectoire (Figure 75). En effet et afin de représenter la trajectoire dans l'IVGE, nous proposons d'étendre l'axe de la trajectoire (c.-à-d. la ligne représentant la direction du biocorridor de l'espèce étudiée) en utilisant un rayon d'extension défini par l'utilisateur du système. Ce rayon fait en sorte que les extrémités de l'extension sont équidistantes par rapport aux points les plus proches appartenant à l'axe de la trajectoire. Étant donné que cette extension s'étale sur les cellules en quelque sorte comme une ombre, il est possible de calculer la

proportion de chevauchement de cette ombre avec chacune des cellules. Par exemple, la Figure 75 montre que la cellule C_1 est presque recouverte entièrement par l'extension alors que cette extension ne touche qu'une petite partie de la cellule C_2 . Ces proportions seront utiles entre autres pour que notre système de distribution de flots d'individus soit en mesure de déterminer si une cellule est traversée ou non par une trajectoire. D'ailleurs, plus le rayon de l'extension est grand, plus le nombre de cellules traversées par la trajectoire est élevé.

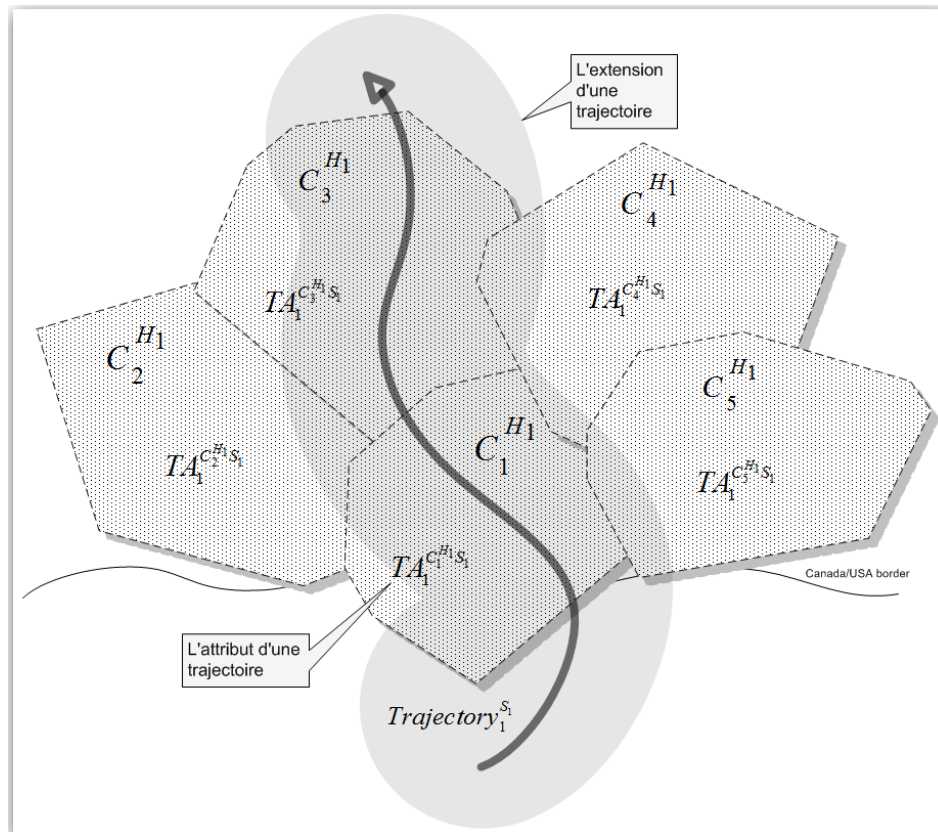


Figure 75: Un exemple de trajectoire avec son extension.

Par ailleurs, nous proposons d'ajouter dans les cellules traversées par la trajectoire (incluant celles qui sont sélectionnées en utilisant l'extension de la trajectoire) un attribut ($TA_k^{C_j^{H_i} S_p}$: *Trajectory Attribute*) afin de matérialiser cette trajectoire dans l'IVGE. En effet, nous

définissons l'attribut d'une trajectoire pour une espèce S_p et pour une cellule $C_j^{H_i}$ de la façon suivante :

$$TA_k^{C_j^{H_i} S_p} = (Id_{TA_k^{C_j^{H_i} S_p}}, \{(Id_{Trajectory_m^{S_p}}, TOrientation_m^{C_j^{H_i}}, TProportion_m^{C_j^{H_i}})\})$$

Cet attribut contient en plus de son identifiant ($Id_{TA_k^{C_j^{H_i} S_p}}$), un ensemble de triplets représentant les informations sur la ou les trajectoires de l'espèce S_p qui passent à travers la cellule $C_j^{H_i}$. Chaque triplet contient premièrement l'identifiant ($Id_{Trajectory_m^{S_p}}$) de la trajectoire. Il contient également une orientation ($TOrientation_m^{C_j^{H_i}}$) qui peut être soit une information quantitative soit une information qualitative comme mentionné auparavant (voir Section 3.5). Elle est calculée en utilisant les tangentes des deux points d'intersection entre l'axe de la trajectoire et le contour de la cellule en question. En effet, puisque la trajectoire représente le corridor de l'espèce étudiée, elle devrait normalement avoir une forme plus ou moins linéaire à l'intérieur de la cellule traitée surtout que nous travaillons sur un large territoire. De ce fait, nous pensons que la moyenne de ces tangentes est une bonne approximation de l'orientation de la trajectoire. Cette orientation est celle que devront prendre les individus afin de suivre la trajectoire lors de leurs déplacements. D'ailleurs et puisque nous connaissons l'orientation des voisins d'une cellule, il serait facile de déterminer le ou les voisins que les individus devront choisir comme destination lors de leurs déplacements afin de suivre la trajectoire préétablie dans l'IVGE. Finalement, le triplet contient un pourcentage ($TProportion_m^{C_j^{H_i}}$) représentant, comme mentionné auparavant, le rapport entre la superficie de la partie de l'extension qui touche la cellule $C_j^{H_i}$ et la superficie totale de la même cellule. Ce pourcentage pourra être utile afin de répartir proportionnellement le nombre des individus lors de leur distribution vers les cellules voisines.

4.1.3 Le réseau de mobilité

Le réseau de mobilité d'une espèce représente l'ensemble des trajectoires spécifiées dans l'IVGE par l'intermédiaire des attributs dans les cellules. Nous représentons un réseau de mobilité de la façon suivante :

$$MobilityNetwork_m^{S_p} = (Id_{Network_m^{S_p}}, \{Trajectory_1^{S_p}, Trajectory_2^{S_p}, \dots\})$$

$Id_{Network_i^{S_p}}$ représente le nom et/ou l'identificateur du réseau. $\{Trajectory_1^{S_p}, Trajectory_2^{S_p}, \dots\}$ est l'ensemble des trajectoires appartenant à ce réseau. De ce fait et étant donné que les trajectoires peuvent se chevaucher (Figure 76), les individus peuvent suivre une trajectoire ou l'autre lors de leur déplacement. Nous pensons que dans ce cas, le choix de la trajectoire à suivre peut se faire en fonction des coefficients d'attraction (voir Section 4.2).

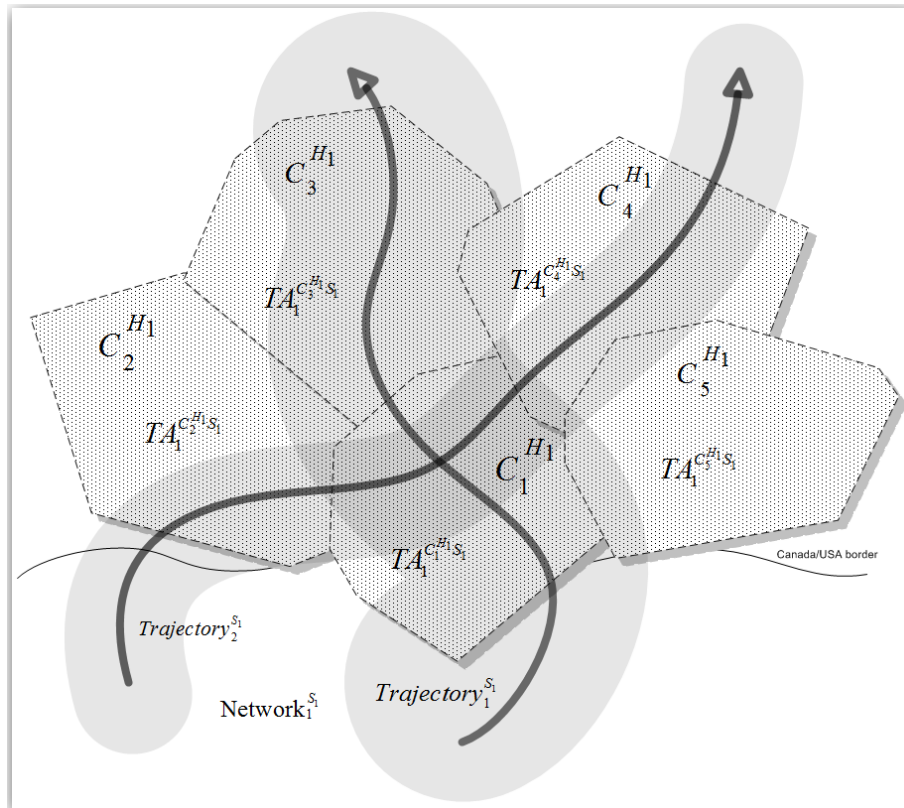


Figure 76: Un réseau de mobilité montrant deux trajectoires.

4.1.4 Le flot d'individus

Avant de définir le flot d'individus, nous définissons un groupe d'individus qui est un certain nombre d'individus qui se sont rassemblés dans une ou plusieurs cellules et à une certaine période pour commencer ensemble un ou plusieurs déplacements. Nous le notons :

$$\text{Group}_m^{S_j} = (\text{Id}_{\text{Group}_m^{S_j}}, \text{Nb}_{\text{Group}_m^{S_j}}, \{C_1^{H_i}, \dots, C_n^{H_i}\}, \text{StartingPeriod}_{\text{Group}_m^{S_j}})$$

$\text{Id}_{\text{Group}_m^{S_j}}$ représente le nom et/ou l'identificateur du groupe. $\text{Nb}_{\text{Group}_m^{S_j}}$ représente le nombre des individus de ce groupe. $\{C_1^{H_i}, \dots, C_n^{H_i}\}$ est l'ensemble des cellules représentant la zone de départ de ce groupe. Cet ensemble doit contenir au moins une cellule. $\text{StartingPeriod}_{\text{Group}_m^{S_j}}$ représente la période à laquelle les individus de ce groupe commencent à faire leurs déplacements.

Nous définissons également le déplacement comme étant le mouvement d'un groupe d'individus d'une cellule source à une cellule destination. Cette cellule de destination n'est pas nécessairement une cellule de destination finale. Elle peut être donc une cellule intermédiaire. Nous notons ce déplacement de la façon suivante :

$$\text{Displacement}_t^{S_p} = (\text{Id}_{\text{Displacement}_t^{S_p}}, \text{Group}_u^{S_p}, C_s^{H_i}, C_d^{H_i})$$

$\text{Id}_{\text{Displacement}_t^{S_p}}$ représente le nom et/ou l'identificateur du déplacement. $\text{Group}_u^{S_p}$ est le groupe qui fait ce déplacement. $C_s^{H_i}$ est la cellule de départ pour ce déplacement alors que $C_d^{H_i}$ est la cellule d'arrivée.

À la lumière de ces deux définitions (groupe et déplacement), nous définissons un flot d'individus comme étant un ensemble de déplacements d'un groupe d'individus qui partent d'un endroit (une ou plusieurs cellules) approximativement en même temps ou à la même période (l'intervalle de temps peut être défini par l'utilisateur) pour emprunter une ou plusieurs trajectoires appartenant à un réseau de mobilité. Nous notons le flot d'individus comme suit :

$$\text{Flow}_u^{S_p} = (\text{Id}_{\text{Flow}_u^{S_p}}, \text{Group}_m^{S_p}, \{\text{Displacement}_t^{S_p}\}, \text{MobilityPattern}_n^{S_p}, \text{MobilityNetwork}_r^{S_p}, \{P_r\})$$

Le flot est donc défini par rapport à un phénomène de déplacement collectif (comportement de groupes) des individus de l'espèce. Pour cela, nous utilisons le patron de mobilité d'une espèce pour définir le flot d'individus. $\{P_r\}$ est l'ensemble des processus que le flot lance afin de distribuer effectivement les individus appartenant au groupe $\text{Group}_m^{S_p}$.

4.2 Implémentation de la distribution d'un flot d'individus

Dans cette section, nous présentons une application que nous avons développée afin de tester un algorithme de distribution d'un flot d'individus en nous basant sur le formalisme présenté dans la section précédente (les différents concepts du transfert spatial des individus). Nous commençons en présentant l'approche que nous proposons et en détaillant les données et les paramètres qui sont utilisés avant l'exécution de l'application. Nous présentons par la suite les différentes étapes d'exécution de notre application permettant justement de faire la distribution spatiale d'un flot d'individus représentant la migration d'oiseaux.

4.2.1 Présentation de l'approche de distribution

L'approche proposée est basée sur le fait d'instancier un groupe d'individus dans un endroit bien précis (p.ex. au sud de la frontière Canada/États-Unis) et de suivre toutes les étapes de distribution (c.-à-d. les différents déplacements) de ces individus cellule par cellule jusqu'à ce qu'il ne reste plus d'individus à distribuer ou jusqu'à ce que le système atteigne les limites de la zone de distribution choisie. Les individus se déplacent à partir du point de départ en empruntant une trajectoire appartenant à un réseau de mobilité. Chaque cellule qui reçoit des individus va jouer le rôle d'une cellule génératrice afin de distribuer une partie de ces individus (étant donné que certains individus peuvent devenir résidents) à ses cellules voisines. L'ordre des cellules voisines lors de la distribution des individus est

important. Nous proposons d'ailleurs de faire un tri croissant en comparant les différences positives entre l'orientation de chaque cellule voisine et l'orientation de la trajectoire. De ce fait, l'application commence la distribution des individus par les cellules qui ont une orientation la plus proche par rapport à celle de la trajectoire. Cependant, le calcul des proportions des individus qui sont réellement distribués se fait en utilisant le coefficient d'attraction de chaque cellule voisine. Ce mécanisme de distribution est détaillé dans la section 4.2.2, mais nous présentons dans la suite de cette section l'environnement (incluant les données et les paramètres initiaux) utilisé par notre application pour faire les tests de l'approche proposée.

Dans cet exemple, on considère une espèce S_1 représentant les Merles d'Amérique. Cette espèce est l'une des plus importantes espèces Nord Américaine qui importe les tiques juvéniles du Sud des États-Unis [Ogden *et al.*, 2008b].

Le flot que notre application utilisera est défini comme suit:

$$\text{Flow}_1^{S_1} = (\text{Id}_{\text{Flow}_1^{S_1}}, \text{Group}_1^{S_1}, \{\text{Displacement}_i^{S_1}\}, \text{MobilityPattern}_1^{S_1}, \text{MobilityNetwork}_1^{S_1}, \{P_1\})$$

Le processus que peut lancer ce flot d'individus en nous basant sur le formalisme de *MATSIM* (voir Section 3.11 du Chapitre V) est défini de la façon suivante:

$$P_1 = ('MigrationOfBirds', 'Biological', \{Z_1^{H_1}\}, \{E_1\}, \Delta_{P_1}, f_{P_1}, \{g_1(A_j^{C_{H_1}^{H_1}})\})$$

L'événement utilisé pour déclencher ce processus est le suivant:

$$E_1 = ('ArrivalOfBirds', D_i, \{N_k\})$$

$\{Z_j^{H_i}\}$ contient une seule zone sélectionnée où le processus se produit. Pour faire nos tests, nous proposons de choisir 14 cellules représentant cette zone. Δ_{P_1} (la durée du processus) ne sera pas initialisée étant donné que cette information n'est pas connue à l'avance. Par contre, f_{P_1} (la fréquence à laquelle le processus se répète) est initialisé à 1 puisque dans notre exemple, le processus va se faire une seule fois. $\{g_1(A_j^{C_{H_1}^{H_1}})\}$ est l'ensemble contenant

la fonction g_I que nous implantons (l'annexe D présente l'algorithme de cette fonction qui est implanté en java) pour modifier les attributs des cellules $A_j^{C^{H_i}}$ au fur et à mesure que nous distribuons les individus à travers ces cellules.

Par ailleurs, nous présentons dans la Figure 77 une capture d'écran de la fenêtre principale de l'application afin de tester la distribution d'un flot d'individus sur les 14 cellules appartenant à l'*IVGE*. Cette figure montre en bas le nombre des individus (initialisé à 1500) représentant le groupe d'individus se trouvant au sud de la frontière Canada/États-Unis que l'application va distribuer. L'utilisateur peut augmenter ou diminuer ce nombre à sa guise en utilisant la barre de défilement ou en entrant directement le nombre souhaité. D'ailleurs, ce nombre peut être initialisé en utilisant la base de données géoréférencées des oiseaux traversant la frontière [Francis et Hussell, 1998].

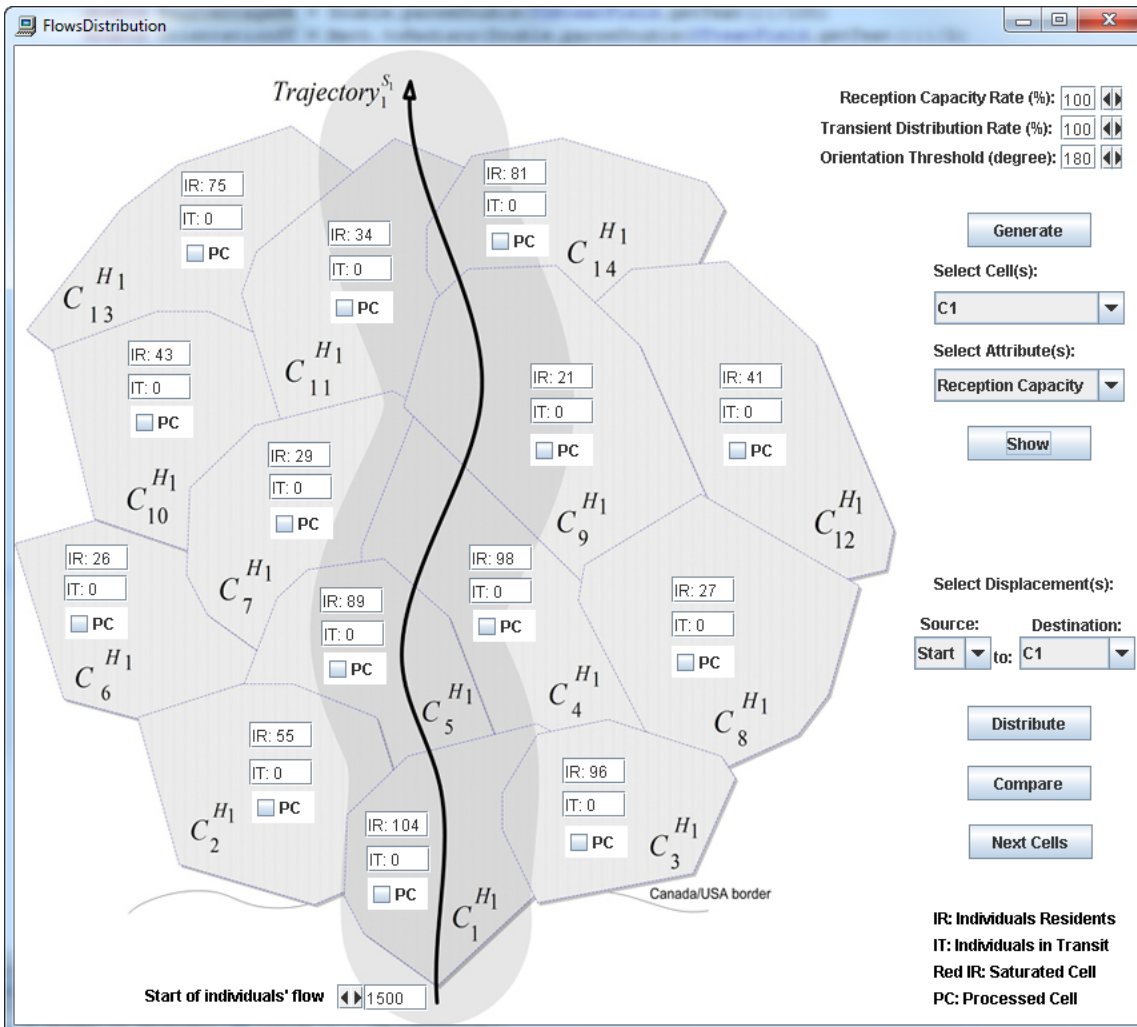


Figure 77: Application de distribution des individus à travers les cellules.

De plus, la Figure 77 montre une trajectoire (c.-à-d. $Trajectory_1^{S_1}$) ayant une direction vers le Nord (les individus vont essayer de suivre la même direction que cette trajectoire lors de leurs déplacements). Chaque cellule contient deux cases représentant le nombre d'individus résidents (c.-à-d. IR: *Individuals Residents*) et le nombre d'individus non résidents (c.-à-d. IT: *Individuals in Transit*). Ces informations représentent des attributs attachés aux cellules. Cependant, il y a d'autres attributs que nous présentons dans le Tableau 5, mais dont nous expliquons l'utilité dans la section 4.2.2.

Tableau 5: Les principaux attributs attachés aux cellules.

Nom de l'attribut	Type	Description
<i>idAttribute</i>	Entier	L'identificateur de la cellule.
<i>ReceptionCapacity</i>	Entier	La capacité d'accueil de la cellule telle que définie dans la Section 3.7 du Chapitre V.
<i>NbResidents</i>	Entier	Le nombre des individus résidents.
<i>NbInTransition</i>	Entier	Le nombre des individus non résidents et qui sont donc en transition.
<i>AttractionCoefficient</i>	Réel	Le coefficient d'attraction de la cellule tel que défini dans la Section 3.7 du Chapitre V.
<i>AttractionProportion</i>	Réel	Une proportion d'attraction permettant de distribuer les individus vers les voisins d'une cellule en comparant les coefficients d'attraction de ces voisins.
<i>OrientationDifference</i>	Réel	La différence entre l'orientation d'une cellule voisine et l'orientation de la trajectoire de sa cellule génératrice. Cette information permet de choisir la cellule voisine qui devrait être traitée en premier.
<i>isProcessed</i>	Booléen	Un Booléen qui retourne Vrai si une cellule génératrice a été déjà traitée, Faux sinon.
<i>infoNeighbors</i>	Vecteur	Un vecteur contenant les informations sur les cellules voisines : l'identificateur du voisin qui est un entier, l'information quantitative (angle en radian) et l'information qualitative (orientation qui une chaîne de caractères) représentant l'orientation d'un voisin par rapport à une cellule génératrice.
<i>TrajectoryAttributes</i>	Vecteur	Un vecteur contenant les informations sur les trajectoires qui traversent la cellule : l'identificateur de la trajectoire qui un entier, l'information quantitative (angle en radian) et l'information qualitative (orientation qui une chaîne de caractères) représentant l'orientation que les individus devront prendre pour suivre la trajectoire et finalement la proportion de l'extension de la trajectoire dans la cellule.

En cliquant sur le bouton «*Generate*» de l'application (voir Figure 77), l'application génère des valeurs pour tous les attributs mentionnés dans le Tableau 5 (certains attributs sont attribués d'une façon aléatoire comme par exemple le nombre des individus résidents alors que d'autres attributs sont attribués selon la configuration des cellules choisies comme par exemple l'identification et l'orientation des voisins). D'ailleurs, en sélectionnant une cellule et en cliquant sur le bouton «*Show*», l'application affiche une fenêtre montrant les valeurs attribuées à certains attributs. La Figure 78a (voir attribut *Neighbors*) montre par exemple que la cellule C_1 possède 4 voisins qui sont C_2 , C_3 , C_4 et C_5 (nous omettons ici la hiérarchie dans le nom des cellules pour des fins de simplification, étant donné que nous faisons nos tests à un seul niveau de granularité). De plus, la cellule C_2 se trouve à un angle de 5.0471 radians et à une orientation Ouest-Nord-Ouest (c.-à-d. *WNW*) par rapport à la

cellule C_1 . En outre et puisque nous utilisons une seule trajectoire pour faire nos tests (cette trajectoire passe par la cellule C_1), la Figure 78a (voir *Trajectory attribute*) montre que la trajectoire a 1 comme identificateur, 0.1079 comme angle en radians, Nord (c.-à-d. N) comme orientation (utilisés pour la direction que devront prendre les individus pour suivre la trajectoire) et 0.97 comme proportion de l'extension de la trajectoire (l'application calcule cette valeur, comme mentionnée auparavant, en faisant l'intersection de l'extension avec la cellule en question). Cette proportion est proche de 100% car l'extension de la trajectoire englobe quasiment la totalité de la cellule C_1 (voir Figure 77).

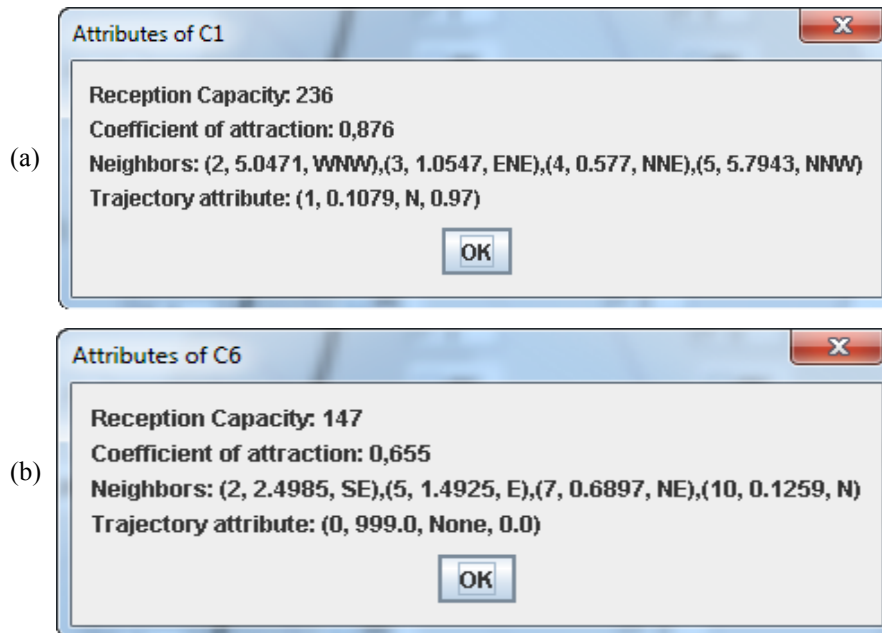


Figure 78: Quelques attributs des cellules C_1 (a) et C_6 (b)

Dans le cas où la trajectoire ne passe pas par une cellule, l'application initialise l'attribut de la trajectoire dans cette cellule avec une valeur de l'angle égale à 999.0 et une valeur de l'orientation égale à la chaîne de caractères « None ». La figure 78b montre justement un exemple avec la cellule C_6 (Figure 78b). En fait et puisque les cellules voisines seront triées en ordre croissant selon la proximité de leur orientation par rapport à l'orientation de la trajectoire, celles qui ne sont pas traversées par la trajectoire (c.-à-d. elles ont une valeur élevée de l'angle égal à 999.0) seront les dernières après ce tri. De ce fait, elles seront les dernières à être traitées. Nous détaillons l'utilité de ce mécanisme dans la section 4.2.2.

En outre et en sélectionnant « *All* » et « *Neighbors* » et en cliquant sur le bouton « *Show* » encore une fois, l'application affiche une fenêtre montrant l'orientation de toutes les cellules les unes par rapport aux autres (voir Figure 79).

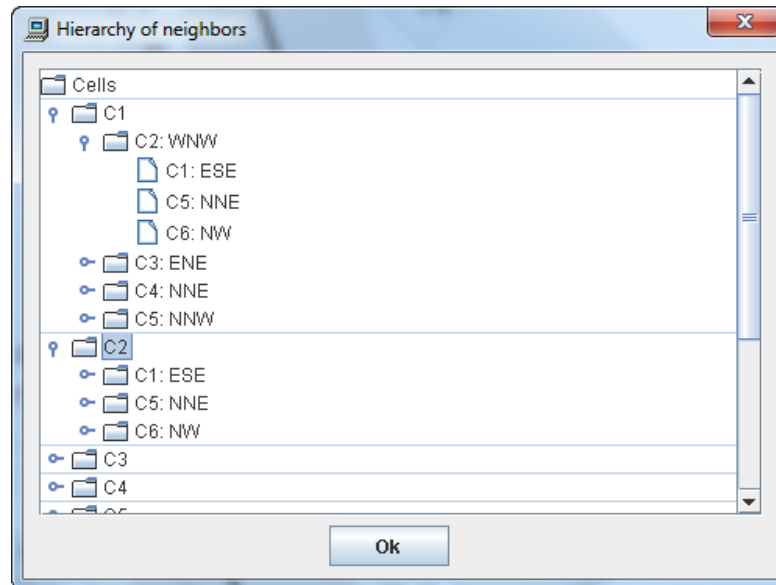


Figure 79: Présentation de l'orientation des différentes cellules.

4.2.2 Étapes d'exécution de l'application de distribution

Nous présentons dans cette section les différentes étapes que le système suit pour distribuer le groupe d'individus à travers les 14 cellules contenant chacune les attributs initialisés auparavant. En effet, l'utilisateur peut choisir le ou les déplacements à faire (voir Figure 77), mais le premier qui est sélectionné par défaut représente le déplacement des 1500 individus vers la cellule C_I . D'ailleurs et en cliquant sur le bouton « *Distribute* », nous remarquons que le nombre des individus de départ devient égal à zéro puisque tous ces individus sont transférés à la cellule C_I . De ce fait, le nombre des individus résidents dans la cellule C_I augmente ainsi que celui des individus non résidents (voir Figure 80). En effet et puisque la capacité d'accueil de la cellule C_I est de 236 individus (voir Figure 78a), le nombre des individus résidents de cette cellule a changé de 104 à 236. La capacité d'accueil est donc très utile afin de limiter le nombre des individus résidents pour ne pas faire épuiser les ressources disponibles dans cette cellule.

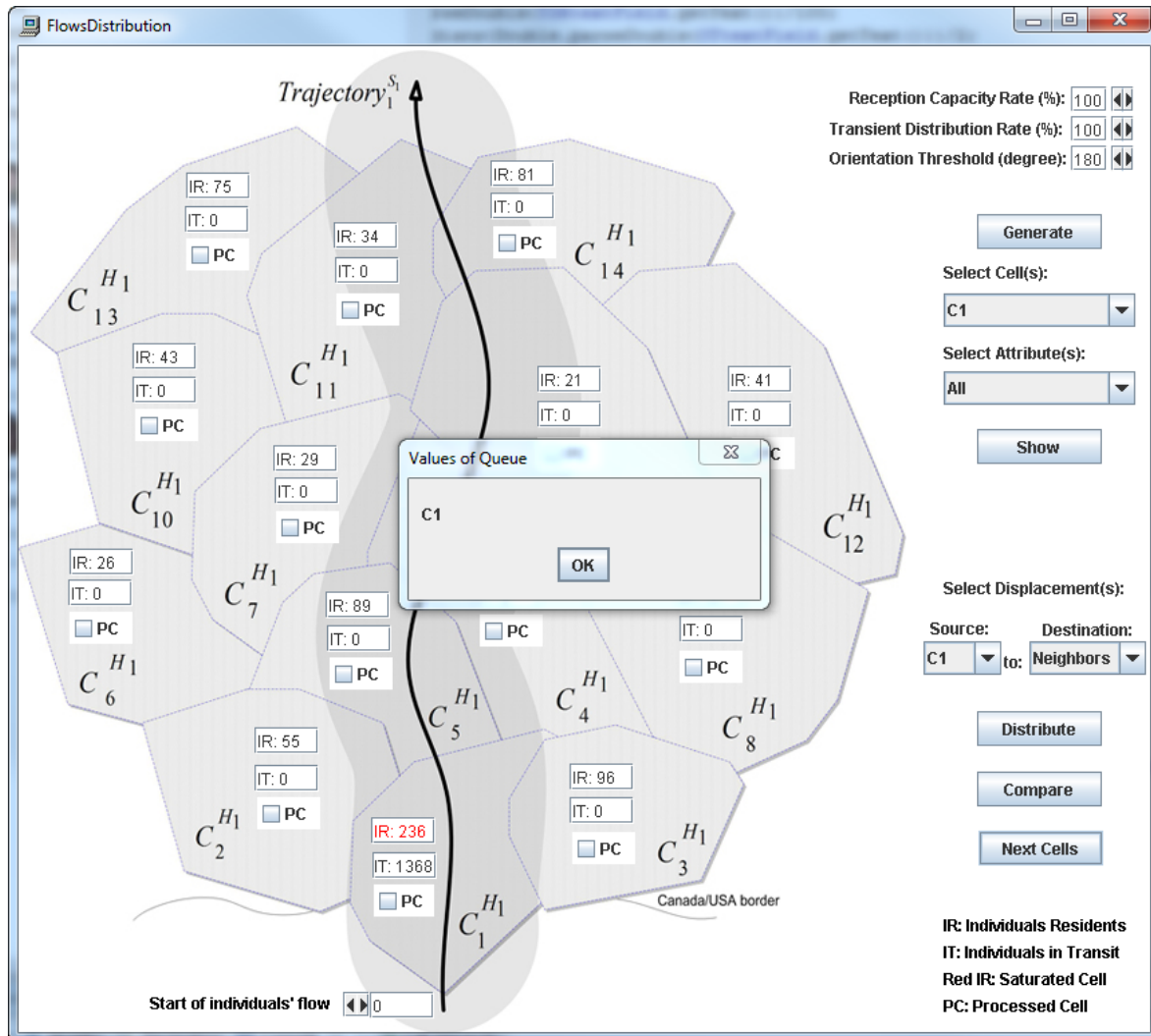


Figure 80: Premier déplacement des individus et prochaine cellule à traiter.

Par ailleurs, l'application utilise un paramètre modifiable (voir la Figure 80 en haut à droite) par l'utilisateur nommé le pourcentage de la capacité d'accueil (c.-à-d. *Reception Capacity Rate*). Ce paramètre est égal par défaut à 100%. De ce fait, notre algorithme essaye d'augmenter le nombre des individus résidents provenant d'autres cellules jusqu'à saturer la cellule traitée avec 100% de sa capacité d'accueil. Cependant, si nous baissons ce paramètre à 80%, l'application calcule une capacité d'accueil réelle en multipliant la capacité d'accueil par son pourcentage défini par l'utilisateur. Par exemple, la capacité d'accueil réelle de la cellule C_1 changera pour $236 * 80 / 100 = 189$ individus. Ainsi, le

nombre de ses individus résidents serait également égal à 189. L'utilisation de ce paramètre permettrait donc à l'utilisateur de ne pas saturer les cellules à 100% de leurs capacités d'accueil afin de tester par exemple un scénario permettant de favoriser le transfert d'un plus grand nombre d'individus étant donné qu'il y aura moins d'individus résidents dans les cellules. D'autre part, le nouveau nombre des individus non résidents est calculé en additionnant l'ancien nombre (NbNonResidents) avec le nombre initial du groupe à distribuer (NbGroupe). L'algorithme enlève au résultat la différence entre la capacité réelle (CAReelle) et le nombre des individus résidents (NbResident):

$$\text{NbNonResidents} = \text{NbNonResidents} + \text{NbGroupe} - (\text{CAReelle} - \text{NbResidents})$$

Dans notre exemple, ce nombre est égal à $0 + 1500 - (236 * 100 / 100 - 104) = 1368$ (Figure 80). De plus et quand une cellule est saturée (elle ne peut plus accueillir des individus résidents), le système change en rouge la couleur du texte se trouvant dans la case montrant le nombre des individus résidents. En outre et en cliquant sur le bouton «*Next Cells*», une fenêtre apparaît et montre les cellules qui sont en attente d'être traitées. L'algorithme utilise une file (c.-à-d. *Queue*) pour simuler cette liste d'attente (voir fenêtre "*Values of Queue*" de la Figure 80). Dans l'exemple, la prochaine cellule qui va être traitée est la cellule C_1 . Cette cellule va jouer le rôle d'une cellule génératrice. De ce fait, les individus non résidents se trouvant dans la cellule C_1 (c.-à-d. 1368 individus) vont être distribués aux cellules voisines de C_1 (c.-à-d. C_2 , C_3 , C_4 et C_5). Il faut par contre mentionner que l'application utilise un paramètre modifiable par l'utilisateur (voir en haut à droite de la Figure 80) nommé le pourcentage de la distribution des individus non résidents (c.-à-d. *Transient Distribution Rate*). Ce paramètre est égal par défaut à 100%. Ainsi, tous les individus non résidents seront distribués. Il ne restera donc pas d'individus non résidents dans la cellule génératrice. Cependant, si nous baissons ce paramètre à 80%, l'application ne distribuera que $1368 * 80 / 100 = 1094$ individus. L'utilisation de ce paramètre permettrait donc à l'utilisateur de ne pas distribuer tous les individus non résidents afin de tester par exemple un scénario permettant de prolonger la période de transit de ces individus pour les distribuer éventuellement un peu plus tard en utilisant un autre processus.

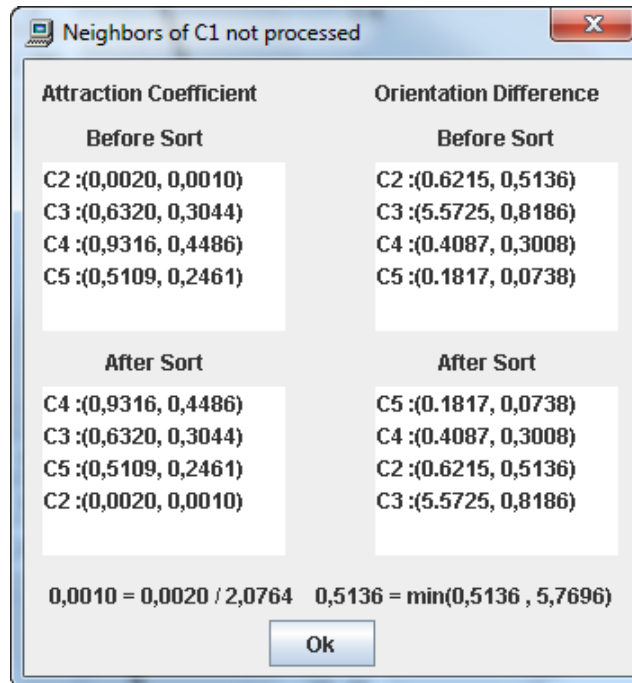


Figure 81: Tris des cellules voisines de C_1 et qui ne sont pas encore traitées.

Sinon et pour effectuer la distribution des individus non résidents de la cellule génératrice à ses cellules voisines, l'algorithme fait un tri des cellules voisines de deux façons différentes. La première consiste à trier les cellules selon leurs coefficients d'attraction. Pour ce faire, l'algorithme calcule pour chaque cellule le rapport entre son coefficient d'attraction et la somme de tous les coefficients d'attraction de toutes les cellules voisines prises en compte. Par exemple, la Figure 81 montre à gauche la liste des cellules voisines avant et après le tri en ordre décroissant. De ce fait, c'est la cellule C_4 qui a le coefficient le plus élevé et c'est la cellule C_2 qui a le coefficient le moins élevé. Nous utilisons ces rapports (0.4486 pour C_4 , 0.3044 pour C_3 , 0.2461 pour C_5 et 0.0010 pour C_2) dont la somme est égale à 1 afin de calculer le nombre d'individus qui doivent être répartis proportionnellement dans les cellules voisines. Par exemple, la cellule C_4 va recevoir $1368 \times 0.4486 = 613$ individus alors que la cellule C_2 va recevoir $1368 \times 0.0010 = 1$ seul individu. Cependant, on doit déterminer par quelle cellule commencer pour faire cette distribution. Nous proposons d'utiliser le deuxième tri présenté dans la partie droite de la Figure 81. En effet, les cellules voisines sont triées cette fois-ci en ordre croissant et en

utilisant la valeur absolue de la différence entre l'orientation d'une cellule et l'orientation de la trajectoire de sa cellule génératrice (nous utilisons l'attribut *OrientationDifference* qui a été présenté dans le Tableau 5). Ainsi, la première cellule voisine traitée sera celle qui a l'orientation la plus proche de l'orientation de la trajectoire de la cellule génératrice. En fait, la valeur calculée par notre algorithme est le minimum entre la différence des deux orientations et la valeur absolue de 2π moins cette différence : $\text{Minimum}(\text{Différence}, \text{ValeurAbsolue}(2*\text{PI}-\text{Différence}))$. De ce fait, ce tri permet de choisir la première cellule à essayer de saturer en utilisant le nombre des individus calculé en fonction des coefficients d'attraction. Dans l'exemple, l'algorithme commence par C_5 , étant donné que cette cellule se trouve dans l'axe de la trajectoire (voir Figure 80). La dernière cellule à saturer est la cellule C_3 comme le montre la Figure 81.

Par ailleurs et pour ne pas distribuer des individus dans le sens contraire d'une trajectoire, l'application utilise un autre paramètre modifiable par l'utilisateur nommé le seuil d'orientation (c.-à-d. *Orientation Threshold*). Ce paramètre est par défaut égal à 180 degrés. Ainsi, l'algorithme répartit les individus provenant d'une cellule génératrice qu'à seulement ses cellules voisines ayant une orientation se trouvant dans l'intervalle de 90 degrés à gauche de l'orientation de la trajectoire et de 90 degrés à droite de l'orientation de la trajectoire. La diminution de ce seuil permet donc de ne pas trop éloigner les individus de la trajectoire alors que son augmentation permet éventuellement de faire le contraire si par exemple l'utilisateur l'initialise à sa valeur maximale qui est de 360 degrés.

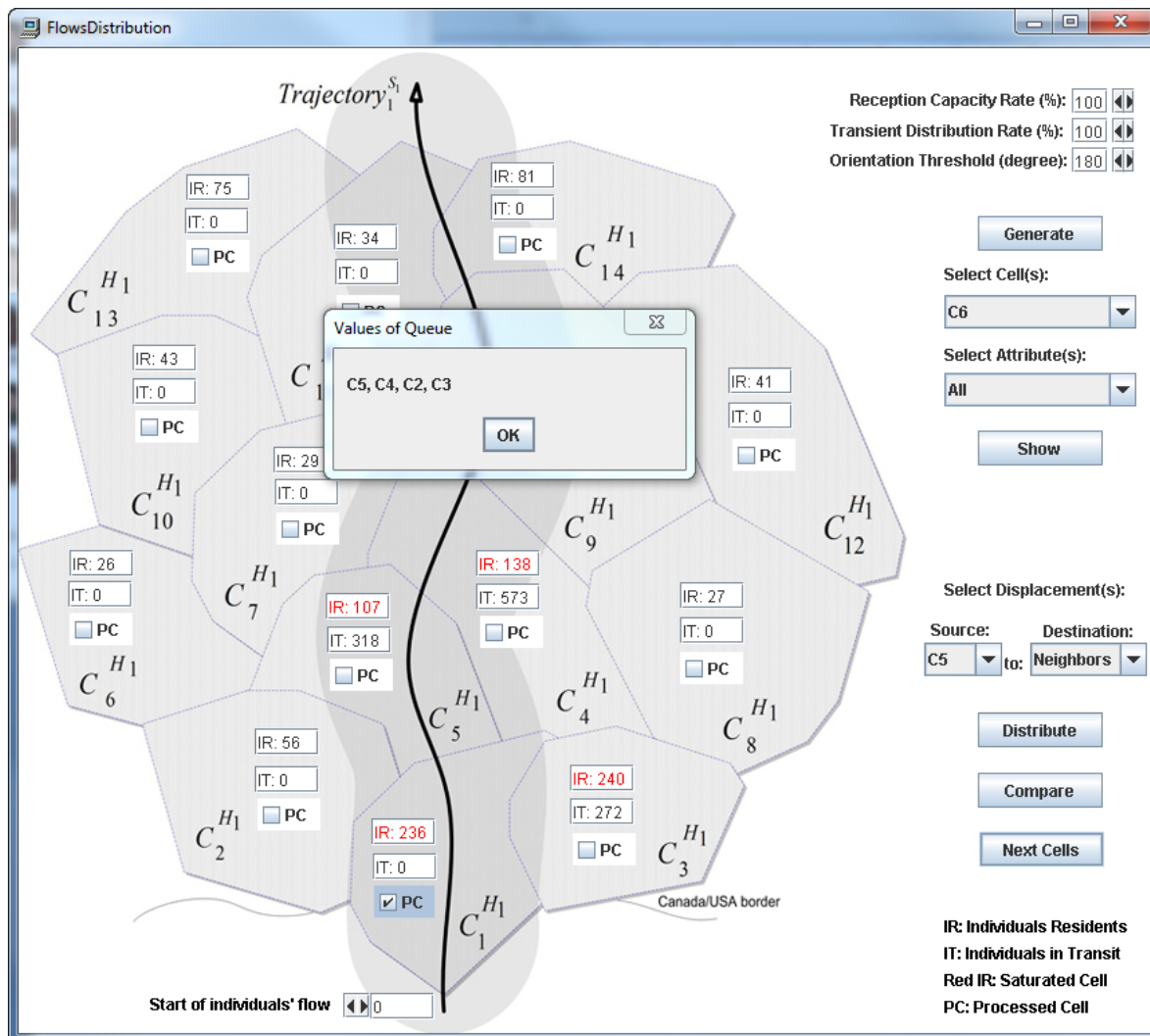


Figure 82: Deuxième distribution des individus et prochaines cellules à traiter.

En cliquant encore une fois sur le bouton « *Distribute* », les individus qui se trouvaient dans la cellule génératrice C_1 sont distribués à ses cellules voisines. La Figure 82 montre d'ailleurs le résultat de cette distribution. Nous remarquons que les cellules C_3 , C_4 et C_5 sont saturées alors que la cellule C_2 ne l'est pas, car elle n'a reçu qu'un seul individu comme nous l'avons précisé auparavant. Étant donné que la cellule C_1 vient d'être traitée, le système coche la case *PC* (c.-à-d. *Processed Cell*) en modifiant sa couleur en bleu. Cette cellule ne sera plus considérée dans les deux tris présentés auparavant et de ce fait, elle ne recevra plus d'individus d'autres cellules. Nous remarquons également que les cellules C_5 ,

C_4 , C_2 et C_3 ont été ajoutées dans la file d'attente (fenêtre "Values of Queue" de la Figure 82). C'est donc la cellule C_5 qui est la prochaine à être traitée, étant donné qu'elle est la première selon le tri croissant qui utilise les différences des orientations. D'ailleurs, l'application sélectionne automatiquement la cellule C_5 comme étant la cellule source du prochain déplacement (Figure 82).

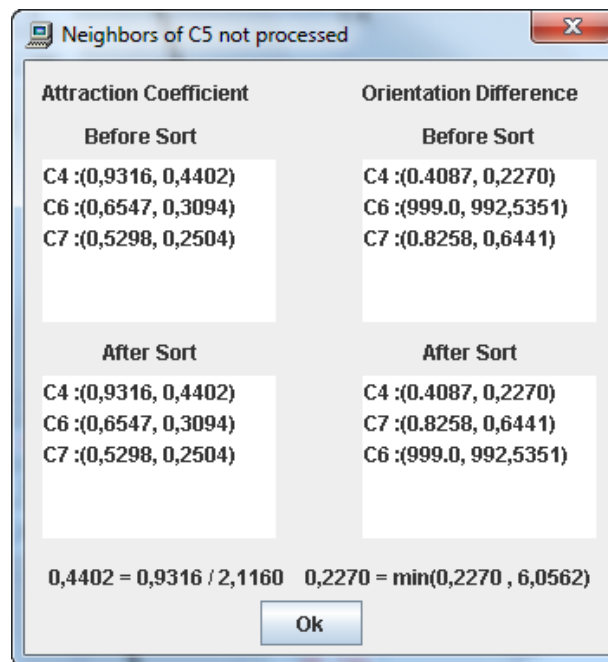


Figure 83: Tris des cellules voisines de C_5 et qui ne sont pas encore traitées.

La prochaine étape est maintenant de distribuer les individus non résidents se trouvant dans la cellule C_5 à ses cellules voisines. Pour ce faire, l'algorithme trie encore une fois ces cellules voisines selon les deux techniques mentionnées auparavant. Cependant, la cellule C_1 n'est pas prise en compte, étant donné qu'elle a été déjà traitée. De plus, la cellule C_2 n'est pas également prise en compte dans les tris puisqu'elle ne respecte pas le seuil d'orientation (c.-à-d. 180 degrés) présenté auparavant. La Figure 83 présente les résultats de ces deux tris. La cellule C_5 va donc distribuer ses 318 individus non résidents en commençant par la cellule C_4 , ensuite C_7 et finalement C_6 . Par contre, C_6 va recevoir plus d'individus (car elle possède un meilleur coefficient d'attraction) que C_7 bien que l'orientation de cette dernière soit plus proche de l'orientation de la trajectoire. En fait, nous

favorisons une cellule plus attractive (qui contient un milieu plus favorable) par rapport à une cellule plus proche de la trajectoire, car nous pensons que dans la nature les individus peuvent se dévier un peu de leur trajectoire surtout s'ils sont attirés par un milieu favorable. Ils peuvent éventuellement retrouver le chemin de la trajectoire dans des cellules subséquentes.

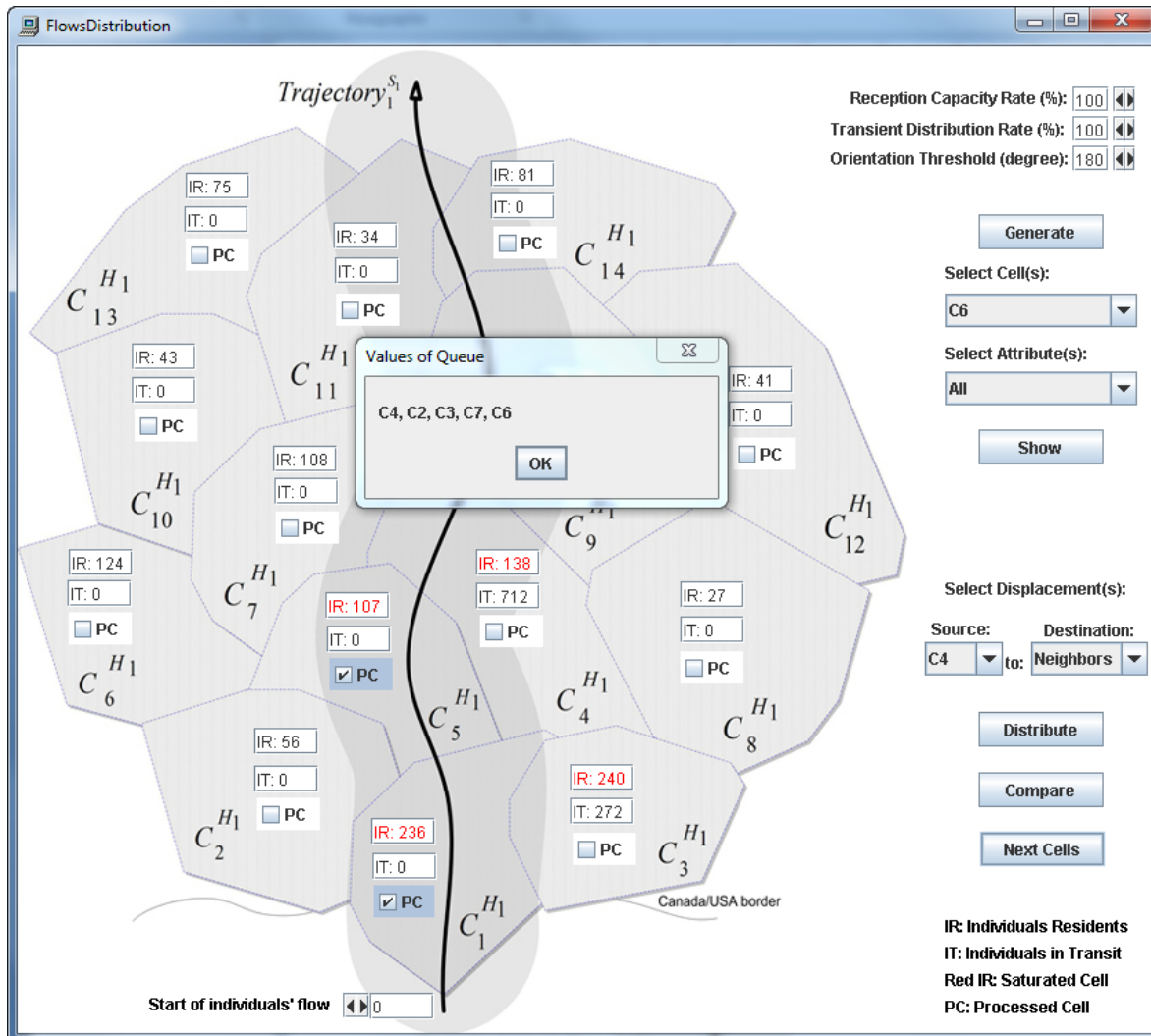


Figure 84: Troisième distribution des individus et prochaines cellules à traiter.

La Figure 84 montre le résultat de la distribution de la cellule génératrice C₅ qui ne sera désormais plus traitée. Ainsi, le nombre d'individus résidents de C₆ et C₇ a augmenté (de 26 186

à 124 pour C_6 et de 29 à 108 pour C_7), mais ces cellules n'ont pas été saturées. De plus, le nombre des individus non résidents de C_4 a également augmenté (de 573 à 712). À la suite de cette troisième distribution, les cellules C_7 et C_6 sont ajoutées dans la file d'attente pour être traitées prochainement après les cellules C_4 , C_2 et C_3 . (voir fenêtre "Values of Queue" de la Figure 84).

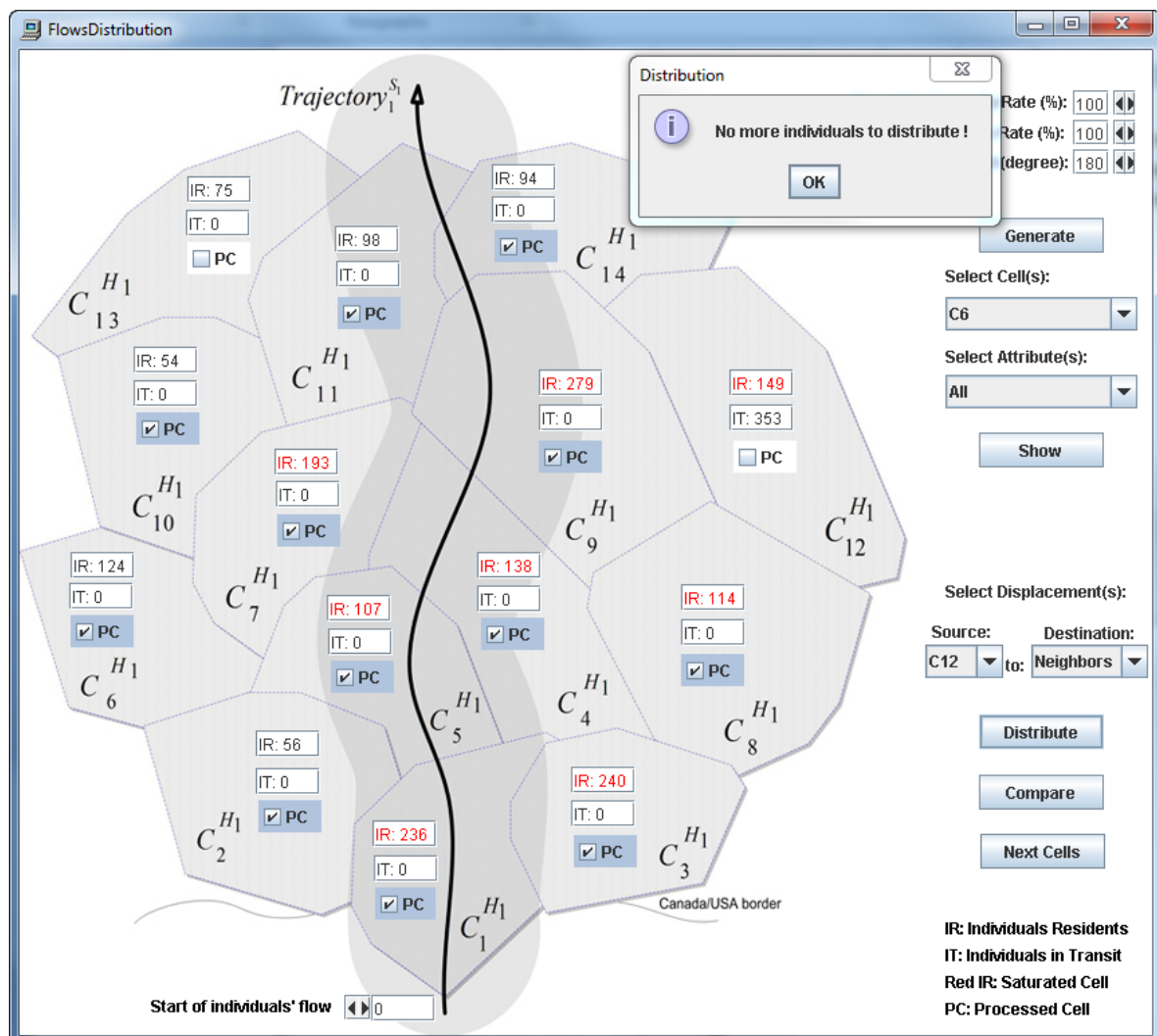


Figure 85: Fin du processus de distribution.

Ce processus de distribution continue jusqu'à ce qu'il n'y ait plus d'individus à distribuer ou jusqu'à ce que toutes les cellules deviennent saturées. La Figure 85 montre d'ailleurs qu'il n'y a plus d'individus à distribuer après la saturation de 8 cellules et le traitement de

11 cellules. Les 353 individus non résidents de la cellule C_{12} n'ont pas pu être distribués, étant donné que notre zone géographique est limitée à 14 cellules. Ces individus sont donc bloqués, car C_{12} n'a pas d'autres voisins à part C_8 , C_9 et C_{14} qui ont été déjà traitées. De plus, C_8 n'est pas considérée puisqu'elle se trouve en dehors de l'angle permis selon le seuil d'orientation (Figure 85).

5 Discussion et conclusion

Dans ce chapitre, nous avons présenté une nouvelle approche pour générer un environnement virtuel géographique informé (c.-à-d. l'*IVGE*) qui est utilisé pour géosimuler la propagation d'une zoonose, en tenant compte surtout des caractéristiques spatiales de ce phénomène. Nous avons exploité entre autres les données vectorielles de la couverture du sol afin de progressivement fusionner les polygones en fonction de leur degré de favorabilité. Les polygones obtenus sont utilisés comme cellules de base qui sont les éléments fondamentaux de l'*IVGE*.

Nous pensons que cette approche innovatrice offre plusieurs avantages si nous la comparons aux approches qui ont été utilisées jusqu'à présent. Par exemple, l'environnement de simulation utilisé par les automates cellulaires [White *et al.*, 2009] ne tient pas compte de certains facteurs importants liés à la propagation de la maladie telle que la survie des populations dans un espace géographique spécifique. Un autre exemple est l'algorithme *GDBSCAN* [Sander *et al.*, 1998] (c.-à-d. *Generalizing Density-Based Clustering Algorithm*) qui peut être utilisé pour regrouper des polygones en un ensemble d'agrégats (c.-à-d. *clusters*), mais repose sur la sélection des polygones voisins (en utilisant un critère de densité) se trouvant dans une région circulaire autour du polygone en cours de traitement. Par ailleurs, la géosimulation de la propagation des zoonoses doit traiter des populations énormes de diverses espèces qui se trouvent dans de vastes territoires. De ce fait, l'utilisation des cellules initiales de couverture du sol (comme c'est le cas pour les cellules administratives qui sont utilisées sans prétraitement) pour générer l'*IVGE*

représentant ces vastes territoires est presque impossible, surtout parce que le nombre de ces cellules peut atteindre (pour les six régions représentées à la Figure 61) près d'un million. Heureusement, notre approche permet de réduire considérablement le nombre de ces cellules tout en respectant les caractéristiques du phénomène étudié (p.ex. la favorabilité du milieu), mais aussi selon les besoins de l'utilisateur (choix des critères et des seuils). De plus, la réduction du nombre de cellules dans lesquelles les populations qui nous intéressent doivent évoluer et interagir permettra d'améliorer l'efficacité du moteur de simulation en réduisant les ressources computationnelles nécessaires pour géosimuler la propagation des zoonoses.

En outre, l'enrichissement de l'*IVGE* en ajoutant des informations supplémentaires (p.ex. biologiques, topographiques, climatiques, etc.) dans chaque polygone permet entre autres de rendre disponible des données pertinentes alimentant le modèle mathématique enrichi. Ce modèle qui est cloné dans chacune des cellules permet de calculer l'évolution et les interactions spatio-temporelles des espèces impliquées dans la propagation de la zoonose étudiée. Ce calcul est possible grâce entre autres à l'attachement des données aux cellules comme par exemple le nombre initial des individus ou encore certaines conditions climatiques (p.ex. la température moyenne). De plus, l'enrichissement de l'*IVGE* permet de géosimuler certains processus biologiques d'une façon plus efficace. Par exemple, la migration des oiseaux qui importent les tiques juvéniles au Québec depuis les États-Unis est modélisée en utilisant un algorithme de distribution utilisant entre autres l'identification de chaque voisin de chaque cellule. Nous rappelons que ces informations sont disponibles dans l'*IVGE* étant donné que nous avons stocké dans chaque cellule l'identification de ses voisins, mais également leurs orientations par rapport à cette cellule. Nous utilisons également des attributs comme le coefficient d'attraction, la capacité d'accueil et l'attribut de la trajectoire (matérialisant un corridor de migration) afin de bien distribuer les individus à travers les cellules.

Par ailleurs et afin d'améliorer notre approche dans un travail futur (voir Section 3 du Chapitre VIII), nous pensons rendre le processus de fusion progressive un peu plus

efficace. En effet, l'algorithme traite actuellement les voisins immédiats de chaque polygone. Il passe par la suite au polygone suivant sans traiter les voisins des voisins du polygone courant. Une recherche en largeur (c.-à-d. *Breadth-first search*) ou une recherche en profondeur (c.-à-d. *depth-first search*) pourrait éventuellement améliorer le résultat, car chaque polygone aura la possibilité de choisir parmi un nombre plus important de candidats pour la fusion. Cette recherche de meilleurs voisins pourrait continuer jusqu'à atteindre un résultat satisfaisant selon l'appréciation de l'utilisateur en se basant sur les critères et les seuils retenus. Finalement, nous croyons que notre approche permettant de générer un *IVGE* d'une façon automatique peut être utilisée non seulement pour simuler la propagation des zoonoses, mais aussi qu'elle peut être adaptée à d'autres phénomènes qui ne sont pas nécessairement liés à la propagation des maladies infectieuses comme par exemple la modélisation des zones de pêche ou des zones de chasse.

Chapitre VII: Développement d'outils d'aide à la décision

1 Introduction

Nous proposons dans ce chapitre de présenter les outils que nous avons développés ou ceux pour lesquels nous avons participé au développement afin d'appliquer certaines notions théoriques que nous avons présentées auparavant et de ce fait, répondre au quatrième objectif de notre thèse (voir Section 4 du Chapitre I). En effet, nous avons acquis une première expérience avec le développement d'un outil permettant la géosimulation de la propagation du VNO (voir Section 2.2.1 du Chapitre III). Pour cela, nous avons côtoyé les experts du domaine afin de mieux comprendre le phénomène à simuler. Nous avons alors essayé de trouver des éléments de réponses permettant d'atteindre nos objectifs. Notre solution pour le VNO a donné de bons résultats, surtout après la calibration du modèle faisant évoluer les populations impliquées dans la propagation de la maladie, et ce à deux niveaux de granularité (deux échelles spatiales). Cette première expérience nous a amenés à la conclusion que notre solution peut être appliquée à d'autres types de maladies. Nous avons alors choisi d'explorer des pistes afin de rendre notre solution générique.

Nous commençons par la présentation de l'essentiel de l'expérience que nous avons eue avec le projet *WNV-MAGS* (*West Nile Virus Multi-Agent Geosimulation*) permettant de développer un outil de simulation de la propagation du VNO. Nous présentons par la suite un outil générique permettant de simuler la propagation d'une zoonose comme la maladie de Lyme (voir Section 2.2.2 du Chapitre III) dans le cadre du projet *Zoonosis-MAGS*.

2 WNV-MAGS

Nous présentons dans cette section l'essentiel de l'expérience que nous avons eue avec la simulation du VNO [Bouden *et al.*, 2008a], [Bouden *et al.*, 2008b], [Bouden *et al.*, 2010]. En effet, nous avons développé un système dont l'objectif est de simuler les comportements

des populations de moustiques et des oiseaux qui sont impliquées dans la propagation et la transmission du VNO. Cette simulation est faite dans un environnement géographique virtuel représentant un large territoire (p.ex. Québec) et selon divers scénarios climatiques et de traitements par larvicides. Nous avons déjà présenté dans le troisième chapitre une vue d'ensemble des phénomènes qui sont reliés à la propagation du VNO (voir Section 2.2.1.4 du Chapitre III, Figure 26). Cependant, nous ne pouvons pas simuler tous ces phénomènes qui sont réellement présents dans la nature. La complexité de représenter tous ces phénomènes ainsi que la non-disponibilité de toutes les données nécessaires afin de les modéliser sont parmi les raisons qui nous obligent à opter pour des simplifications. Nous avons alors formulé un certain nombre d'hypothèses qui nous ont conduits à identifier les paramètres qui semblent déterminants pour la simulation et qui concernent notamment les espèces de moustiques et de corvidés, les facteurs influençant l'évolution des populations, la zone géographique considérée, la période de simulation et les échelles spatio-temporelles qu'il faut utiliser (voir Tableau 6).

Tableau 6 : Spécifications des paramètres de la simulation.

Paramètres	Spécifications
Espèces considérées	Espèce représentative des populations de corvidés : <i>Corvus brachyrhynchos</i> (la Corneille d'Amérique). D'autres espèces sont aussi considérées. Espèces représentatives des populations de moustiques : <i>Culex pipiens</i> et <i>Culex restuans</i> .
Principaux facteurs influençant les comportements des <i>Culex</i>	Conditions climatiques (essentiellement la température et les précipitations).
Principaux facteurs influençant les comportements des corneilles	Zones et périodicité de déplacements et de rassemblements.
Zones géographiques	Deux zones géographiques : (1) le sud du Québec représentant la zone écumène des régions administratives suivantes: Québec, Chaudière-Appalaches, Mauricie-Centre du Québec, Montérégie, Estrie, Montréal-centre, Laval, Laurentides, Lanaudière et Outaouais et (2) la région métropolitaine d'Ottawa.
Période de simulation	Nous avons décidé de simuler la propagation du VNO du premier juillet jusqu'au premier octobre étant donné que c'est la période d'apparition des cas d'infection (animaux et humain).
Échelle temporelle	Un intervalle d'un pas quotidien avec un bilan hebdomadaire.
Échelles spatiales	Deux échelles spatiales : (1) une échelle macro représentant le sud du Québec qui donnera une idée sur l'ampleur de la propagation du VNO sur un large territoire et (2) une échelle méso représentant la région métropolitaine d'Ottawa.

Le Tableau 6 montre par exemple que nous avons sélectionné la corneille d'Amérique comme espèce représentative des populations de corvidés. Ce choix n'est pas arbitraire puisque c'est une espèce qui est utilisée dans les programmes de surveillance comme étant une espèce indicatrice de la propagation du VNO (les corneilles mortes sont signalées et ramassées pour analyse au laboratoire). De plus, nous avons sélectionné les *Culex* comme espèce représentative des populations de moustiques. C'est une espèce qui se trouve souvent dans les régions urbaines et de ce fait, elle peut transmettre la maladie aux humains. Ces hypothèses nous ont conduits à produire un modèle conceptuel pour la propagation du VNO (voir Figure 86).

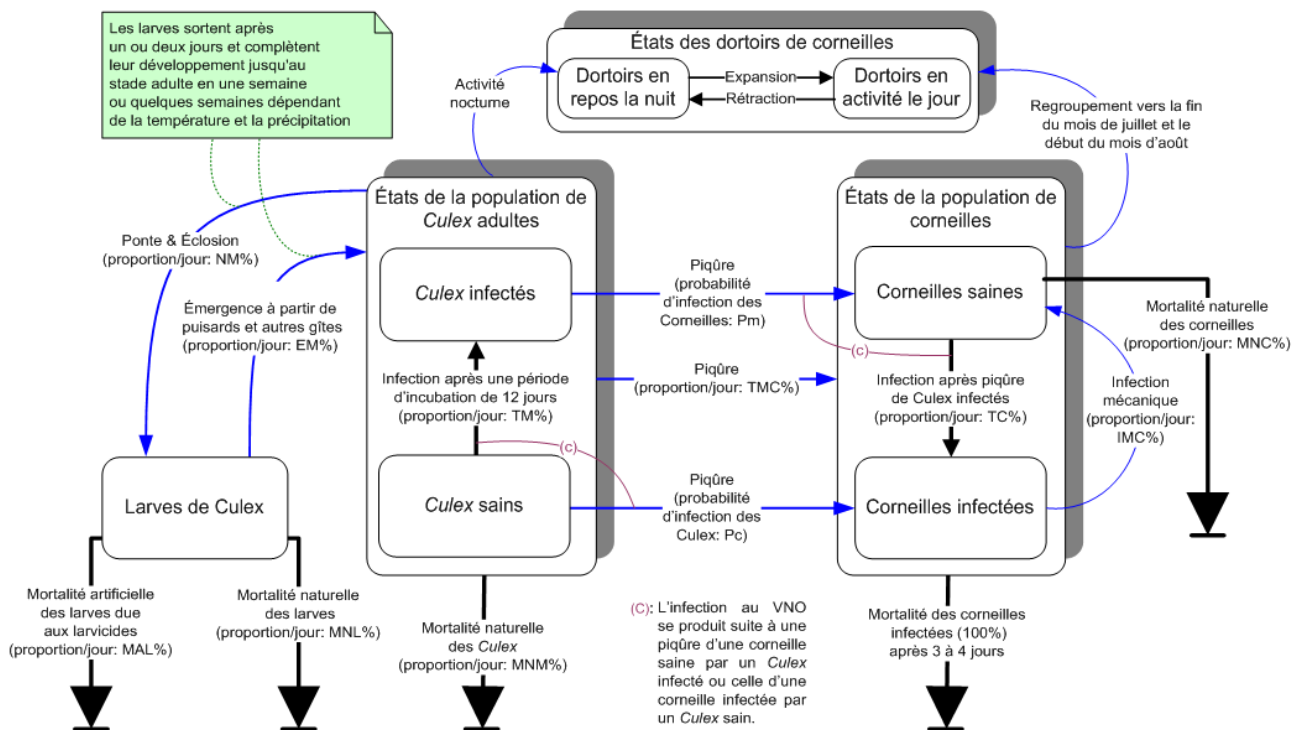


Figure 86: Modèle conceptuel représentant une vue dynamique du système.

L'objectif de ce modèle conceptuel est de présenter une vue synthétique de la dynamique des populations de *Culex* et de corneilles ainsi que les interactions entre ces populations. La Figure 86 montre par exemple que les corneilles commencent à se regrouper, vers la fin du

mois de juillet et le début du mois d'août, dans des dortoirs afin de passer la nuit. Pour simplifier la simulation de ce phénomène de rassemblement des corneilles en dortoirs et compte tenu du grand nombre d'individus de corneilles à considérer, nous avons choisi d'introduire des dortoirs dont les caractéristiques spatiales changent en fonction de l'heure de la journée. En effet, un dortoir peut être considéré comme l'extension spatiale d'un agrégat de corneilles (une sous-population de corneilles qui prend l'habitude de se rassembler dans ce dortoir pendant la période de l'année qui nous intéresse). Le jour, les corneilles se dispersent à la périphérie du dortoir en quête de leur nourriture. Mais, elles reviennent pour passer la nuit sur le lieu du dortoir. Le phénomène spatial de rassemblement et de dispersion de cette sous-population de corneilles peut être représenté de façon synthétique sous la forme d'une expansion et d'une contraction de la région occupée par cette sous-population. Dans le modèle conceptuel, nous représentons deux états principaux pour les dortoirs, celui de leur rétraction maximale («dortoir en repos la nuit») et celui de leur expansion maximale («dortoir en activité le jour»). Cette expansion formera ce que nous appelons «l'Aire d'influence d'un dortoir». Il est à noter que les dortoirs en repos représentent une cible de choix pour les *Culex* étant donné que ces derniers sont nocturnes et que les corneilles sont plus faciles à piquer puisqu'elles ne sont pas en déplacement.

En se basant sur ce modèle conceptuel, nous avons conçu l'architecture du système *WNV-MAGS* (voir Figure E1 de l'annexe E). Ce système est le résultat de l'ajout d'un nouveau module (implanté en utilisant C++) à la plateforme *MAGS* (voir Section 3.5.2 du Chapitre II) permettant de géosimuler la propagation du VNO selon divers scénarios alternatifs d'intervention. Étant donné que nous avons utilisé deux niveaux de granularité (c.-à-d. deux échelles spatiales: voir Tableau 6) afin d'appliquer notre approche de géosimulation multi-niveau, nous présentons dans les sections suivantes le travail effectué ainsi que les résultats que nous avons obtenus pour ces deux niveaux.

2.1 Échelle spatiale Macro (Sud de Québec)

Nous présentons dans cette section les travaux effectués afin de géosimuler la propagation du VNO sur un large territoire représentant le sud de Québec. Le choix de cette région est basé essentiellement sur le fait que nous disposons des données de surveillance (en provenant du système *SIDVS*) nous permettant de calibrer notre système. Nous commençons en présentant la préparation des données ainsi que la création de l'*IVGE* en utilisant une subdivision administrative. Nous présentons par la suite brièvement le modèle mathématique à base de compartiments que nous avons utilisé pour entre autres faire évoluer les populations impliquées dans la propagation du VNO. Ensuite, nous présentons les résultats de la géosimulation. Nous terminons cette section en présentant les différents scénarios utilisés, ainsi que la calibration du système.

2.1.1 Préparation des données et création de l'*IVGE*

Nous avons commencé par générer la subdivision administrative de l'*IVGE* représentant les municipalités de la zone géographique qui nous intéresse (c.-à-d. le sud de Québec). En effet, nous avons utilisé le logiciel Geomedia [Site Web: *Geomedia*] afin de manipuler les données SIG du *DMTI*²⁷ et de l'INSPQ. Nous avons exporté 945 municipalités se trouvant dans la zone géographique qui nous intéresse sur un total de 1476 (province du Québec). La Figure 87 montre justement les polygones représentant ces municipalités et appartenant à la zone écoulène²⁸ des régions sociosanitaires (10 RSS) de Québec (Figure 87).

Par la suite, nous avons préparé les données dont nous avons besoin et qui sont mentionnées dans l'architecture conceptuelle de notre système (voir Tableaux E1 à E4 de l'Annexe E). Ces données vont être attachées aux municipalités représentant les cellules de l'environnement de simulation (c.-à-d. l'*IVGE*) afin de rendre ce dernier informé. En effet,

²⁷ *DMTI Spatial Inc* est un fournisseur canadien de cartes très détaillées qui intègrent par exemple les routes, la végétation ou encore les lacs d'une région particulière (<http://www.dmtispatial.com/>).

²⁸ Le terme « écoulène » est utilisé par les géographes pour désigner la surface habitée. Il s'applique généralement aux régions où des habitants ont établi leur résidence permanente, ainsi qu'à toutes les zones de travail occupées ou utilisées à des fins agricoles ou pour d'autres activités économiques.

nous avons estimé les populations initiales de *Culex* et de corneilles au début de la simulation. Nous avons utilisé les logiciels SAS²⁹ et MapInfo^{©30} afin de calculer une densité de corneilles par région sociosanitaire (nombre d'individus par km²) en faisant une moyenne des mentions de la base ÉPOQ³¹ qui sont situées à l'intérieur de l'écoumène (voir Tableau E1 de l'Annexe E). Nous avons également traité les données relatives aux dortoirs (reçues par courriel à la suite d'une enquête que nous avons menée auprès d'ornithologues ou contenues dans la base ÉPOQ). Ces données comprenaient pour la majorité des cas, les coordonnées spatiales (latitude/longitude), ainsi qu'une approximation du nombre d'individus pour chacun des dortoirs mentionnés (voir Tableau E4 de l'Annexe E). Afin que la représentation soit la plus significative possible, nous avons calculé une moyenne du nombre d'individus pour le cas de plusieurs signalements d'un même dortoir (même latitude et même longitude).

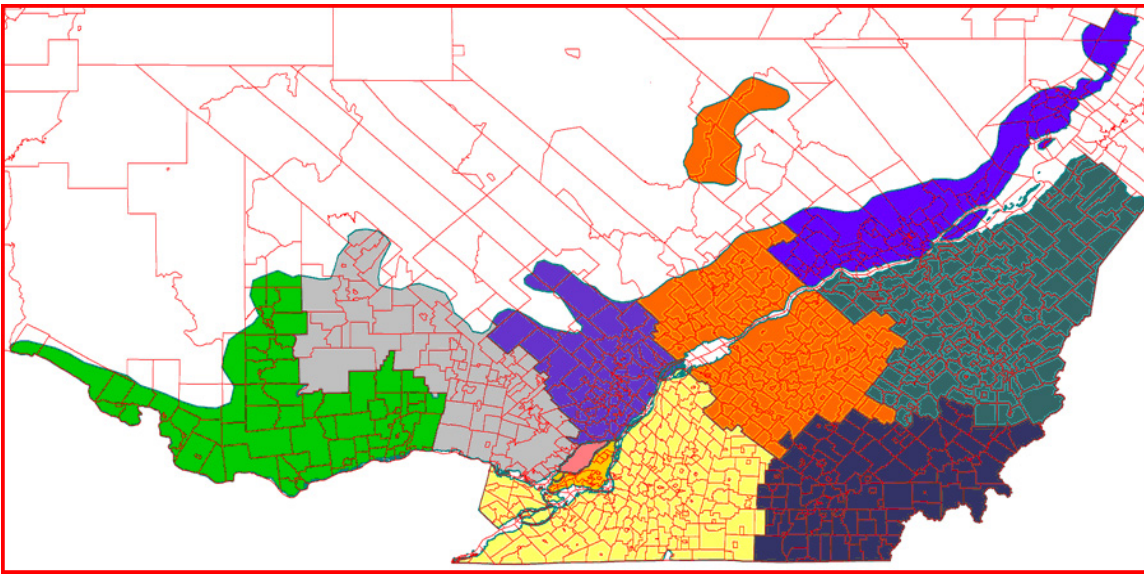


Figure 87: La zone géographique qui nous intéresse.

²⁹ SAS est logiciel de statistique intégrant de nombreuses fonctionnalités (<http://www.sas.com/>).

³⁰ MapInfo est un puissant outil d'analyse cartographique (<http://www.mapinfo.com/>).

³¹ ÉPOQ : Base de données consacrée à l'étude des populations d'oiseaux du Québec de 1997 à 2005. Cette base contient plus de cinq millions d'observations faites par les ornithologues au cours de leurs excursions de tous les jours. Les observations sont recueillies à l'aide d'un feuillet d'observations quotidiennes des oiseaux. On y retrouve le nombre total d'individus de chaque espèce vue ou entendue.

Concernant les *Culex*, nous avons calculé le nombre initial des populations en estimant le nombre des individus qui émergent des puisards (Figure 88). En effet, nous avons supposé que ces puisards, collectant l'eau de pluie tout au long des routes, représentent un endroit propice pour la ponte des œufs de *Culex* dans les régions urbaines. À cette fin, nous avons utilisé le même type d'application *Visual Basic* mentionnée auparavant (voir Section 3.3 du Chapitre VI: Figure 65) afin d'interroger la base géoréférencée de Geomedia pour calculer la longueur des routes de chaque municipalité se trouvant dans chacune des dix régions écumène (voir Tableau E2 de l'Annexe E). Nous avons donc calculé le nombre de puisards se trouvant dans chaque municipalité en nous basant sur la longueur des routes. Nous avons supposé qu'il y a 1 puisard à tous les 30 m de route. Cette distance moyenne de 30 m est conforme aux normes utilisées par le ministère des Transports lors de l'installation des puisards. Cependant, nous avons laissé à l'utilisateur du système la liberté de modifier cette valeur, étant donné qu'elle peut varier en fonction des régions. Nous avons également supposé qu'il y a seulement 20% des puisards contenant des larves. Cette valeur par défaut peut également être modifiée par l'utilisateur, ainsi que le nombre moyen de *Culex* adultes qui émergent d'un puisard au début de la simulation. Ces chiffres ont été obtenus en consultant les données de la SOPFIM (Société de Protection des Forêts contre les Insectes et Maladies [Site Web: SOPFIM]) concernant le nombre de *Culex* adultes capturés par semaine à proximité d'un puisard pendant la période estivale.

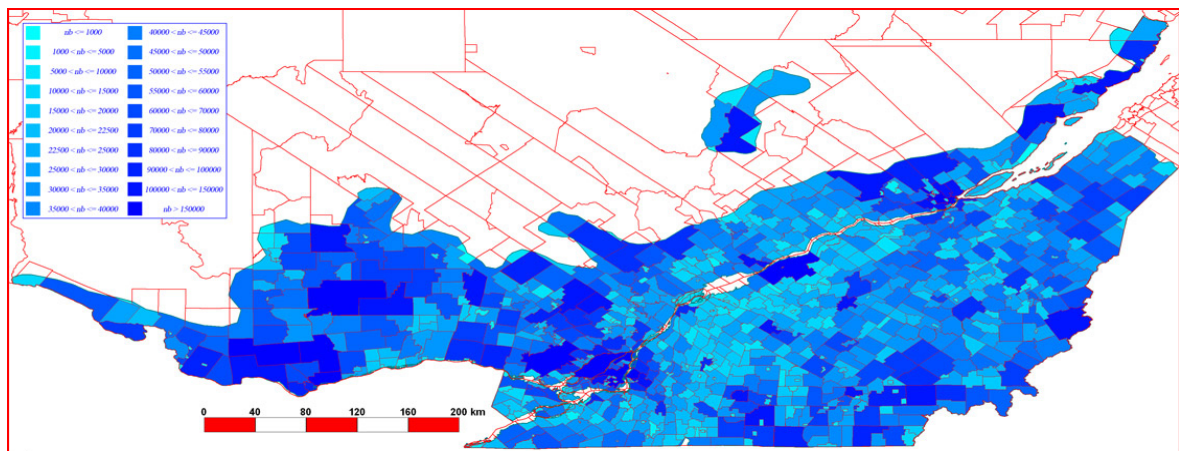


Figure 88: Répartition des populations initiales de *Culex* (les municipalités plus foncées contiennent plus de moustiques).

En outre, nous avons utilisé une *DLL* (c.-à-d. *Dynamic Link Library*) qui permet d'intégrer les données climatiques (p.ex. température, précipitations, degrés jours cumulés, etc.) dans l'*IVGE* (voir Figure 89 et Figure E2 de l'Annexe E). Cette DLL représente une des fonctionnalités du logiciel BIOSIM [Regniere et Saint-Amant, 2008] que nous avons utilisée pour interpoler les valeurs de température et de précipitations à des endroits précis sur le territoire, en prenant en compte des données des quatre stations météorologiques les plus proches et des données d'élévation du terrain. Ce calcul peut être fait en utilisant soit les données réelles ou les normales climatiques du Canada qui sont produites sur plusieurs années [Site Web: *Environment Canada*].

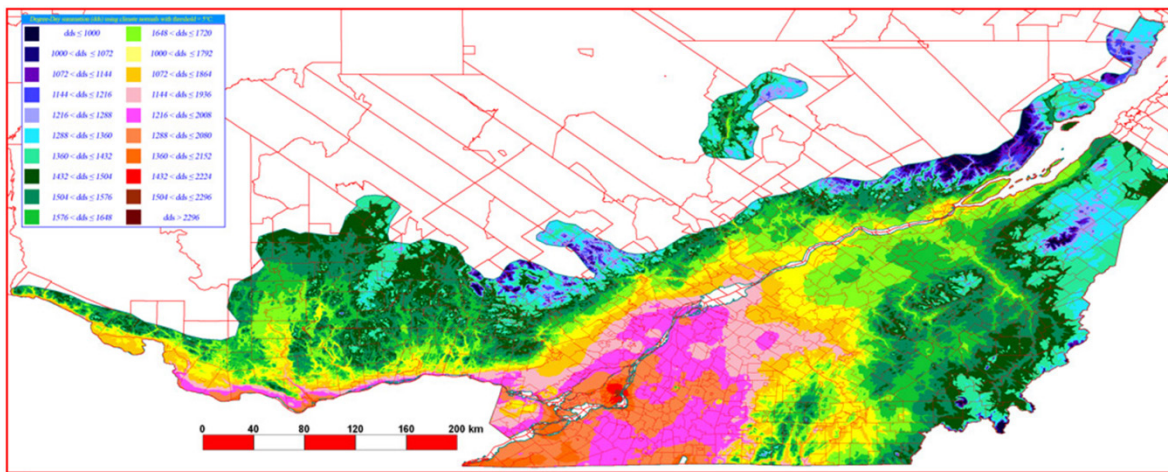


Figure 89: Répartition des degrés jours cumulés (p.ex. les régions en rouge sont plus chaudes que les régions en bleue).

2.1.2 Description du modèle mathématique à base de compartiments

Nous avons besoin de calculer l'évolution des populations de *Culex* et des populations de corneilles, ainsi que leurs interactions. Pour ce faire, nous avons choisi le modèle proposé par Wonham et ses collaborateurs [Wonham *et al.*, 2004] pour calculer la dynamique des deux populations. Ce modèle est basé sur 8 équations différentielles (voir Figure 90a) qui spécifient l'évolution temporelle des individus appartenant aux différents compartiments de chaque espèce (les corneilles susceptibles, infectées, guéries et mortes ainsi que les larves

de *Culex* et les *Culex* adultes susceptibles, exposés et infectés). Les paramètres de ce modèle sont présentés dans le Tableau E3 de l'Annexe E.

(a)

$$(1) \frac{dS_B}{dt} = -abI_M \frac{S_B}{N_B}$$

$$(2) \frac{dI_B}{dt} = abI_M \frac{S_B}{N_B} - \mu_V I_B - gI_B$$

$$(3) \frac{dR_B}{dt} = gI_B$$

$$(4) \frac{dX_B}{dt} = \mu_V I_B$$

$$(5) \frac{dL_M}{dt} = \beta_M(S_M + E_M + I_M) - m_L M - \mu_L L_M$$

$$(6) \frac{dS_M}{dt} = -acS_M \frac{I_B}{N_B} - \mu_A S_M + m_L M$$

$$(7) \frac{dE_M}{dt} = acS_M \frac{I_B}{N_B} - kE_M - \mu_A E_M$$

$$(8) \frac{dI_M}{dt} = kE_M - \mu_A I_M$$

S_B : Susceptible birds
 I_B : Infected birds
 R_B : Recovered birds
 X_B : Dead birds
 L_M : larvae of mosquitoes
 S_M : Susceptible mosquitoes
 E_M : Exposed mosquitoes
 I_M : Infected mosquitoes

(b)

$$\frac{dS_{Bj}}{dt} = -\frac{i_{Bj}\omega_{Bj}\psi_{MB}S_{Bj}}{\tilde{N}_O + \tilde{N}_B}$$

$$\frac{dI_{Bj}}{dt} = \frac{i_{Bj}\omega_{Bj}\psi_{MB}S_{Bj}}{\tilde{N}_O + \tilde{N}_B} - r_j I_{Bj} - \mu_{Bj} I_{Bj}$$

$$\frac{dR_{Bj}}{dt} = r_j I_{Bj}$$

$$\frac{dX_{Bj}}{dt} = \mu_{Bj} I_{Bj}$$

$$\frac{\partial L_M}{\partial t} = D_{x_L} \frac{\partial^2 L_M}{\partial x_L^2} - v_{x_L} \frac{\partial L_M}{\partial x_L} - \mu_L L_M$$

$$\frac{\partial S_M}{\partial t} = D_{x_A} \frac{\partial^2 S_M}{\partial x_A^2} - v_{x_A} \frac{\partial S_M}{\partial x_A} - \mu_A S_M$$

$$\frac{\partial E_M}{\partial t} = D_{x_A} \frac{\partial^2 E_M}{\partial x_A^2} + D_{y_A} \frac{\partial^2 E_M}{\partial y_A^2} - v_{x_A} \frac{\partial E_M}{\partial x_A} - v_{y_A} \frac{\partial E_M}{\partial y_A} - \mu_A E_M$$

$$\frac{\partial I_M}{\partial t} = D_{x_A} \frac{\partial^2 I_M}{\partial x_A^2} - v_{x_A} \frac{\partial I_M}{\partial x_A} - \mu_A I_M$$

$$\frac{dR_M}{dt} = \phi_{RM} - \mu_A R_M$$

R_M : Diapausing mosquitoes

Figure 90: (a) modèle mathématique de Wonham *et al.* (2004) vs (b) modèle amélioré qui a été utilisé pour le système *WNV-MAGS*.

Cependant, nous avons proposé quelques modifications afin de corriger certains problèmes que nous avons trouvés dans le modèle. Nous avons d'ailleurs inclus les effets du climat dans le modèle en utilisant le travail de Madder et ses collaborateurs [Madder *et al.*, 1983]. Ce fut une tâche difficile parce que le modèle n'était plus dans son état d'équilibre et cela a nécessité plusieurs modifications surtout dans les équations différentielles [Noël, 2007]. La Figure 90b présente un aperçu des nouvelles équations du modèle proposé incluant ces modifications. Par exemple, les équations des oiseaux (dS_{Bj}/dt , dI_{Bj}/dt , dR_{Bj}/dt , dX_{Bj}/dt) ont

un indice j qui représente une espèce d'oiseaux qui peut être différente des corneilles et que nous pouvons inclure dans la simulation (voir Section 2.1.3). Un autre exemple est l'utilisation d'un taux de maturation des larves (c.-à-d. v_{xL} dans l'équation dL_M/dt) qui dépend de la température afin d'inclure l'effet du climat dans le modèle comme mentionné auparavant. Le modèle ajusté a donné des résultats satisfaisants sur le plan qualitatif (p.ex. distribution des générations des moustiques). En effet, l'allure des courbes établies reflète les comportements biologiques des espèces étudiées selon l'avis des experts du domaine. Cependant, les résultats quantitatifs (p.ex. le nombre des œufs, des larves, des *Culex* adultes qui ont émergé, des corneilles mortes, etc.) n'ont pas été concluants. Nous avons corrigé ce problème grâce à la calibration du système comme présenté dans la section 2.1.4.

2.1.3 Geosimulation des populations

Selon notre modèle conceptuel (voir Figure 86), la progression du VNO implique un grand nombre d'individus appartenant principalement aux populations de deux espèces. De ce fait, nous avons modélisé ces populations en bénéficiant des avantages de l'approche de géosimulation que nous avons proposée, et en particulier de l'*IVGE*. Ainsi, puisque les *Culex* représentent une population peu mobile et contenant un nombre important d'individus (c.-à-d. *SMP-LNI*, voir Section 3.2.3 du Chapitre IV), nous avons utilisé ce que nous avons appelé une "carte de densité intelligente" (Figure 91) pouvant associer les sous-populations de *Culex* à des espaces d'occupation statiques représentant des municipalités dans l'*IVGE*. Notre idée est donc d'attacher à chaque zone de référence (c.-à-d. municipalité) une liste de variables (c.-à-d. attributs : voir Section 3.7 du Chapitre V) correspondant au nombre des individus se trouvant dans les différents compartiments (larves, adultes sains et infectés) situés à cet endroit.

Par ailleurs, nous avons utilisé les dortoirs de corneilles afin de représenter les populations de cette espèce (*FMP-SNI*, voir Section 3.2.3 du Chapitre IV). Le calcul d'un rayon d'action de chacun des dortoirs pendant la journée nous a permis d'évaluer les distances parcourues par les corneilles à la recherche de nourriture. Nous avons modélisé chaque

dortoir par un agent qui est doté de caractéristiques initiales telles que le nombre d'individus ou encore la position sur la carte. Les dortoirs représentés par des agents prennent de l'expansion au début de la journée pour se rétracter au début de la nuit. Chaque agent-dortoir est modélisé par un système de particules. Chacune de ces particules représente une ou plusieurs corneilles dépendamment du nombre d'individus rattaché au dortoir correspondant. Dans la Figure 92, nous pouvons voir une capture d'écran de *WNV-MAGS* dans laquelle les dortoirs sont affichés comme des "nuages de particules bleues". Chaque particule a des caractéristiques différentes (p.ex. vitesse, direction de déplacement, etc.) qui lui permettent de se déplacer à une certaine distance à partir de l'emplacement d'origine du dortoir. Par conséquent, l'ensemble des particules associé à un dortoir donné couvre une zone circulaire avec un rayon maximal défini par le système. Ce rayon est calculé en fonction du nombre des individus associé au dortoir et de la densité des corneilles se trouvant dans la zone couverte par l'expansion du dortoir (Figure 92).

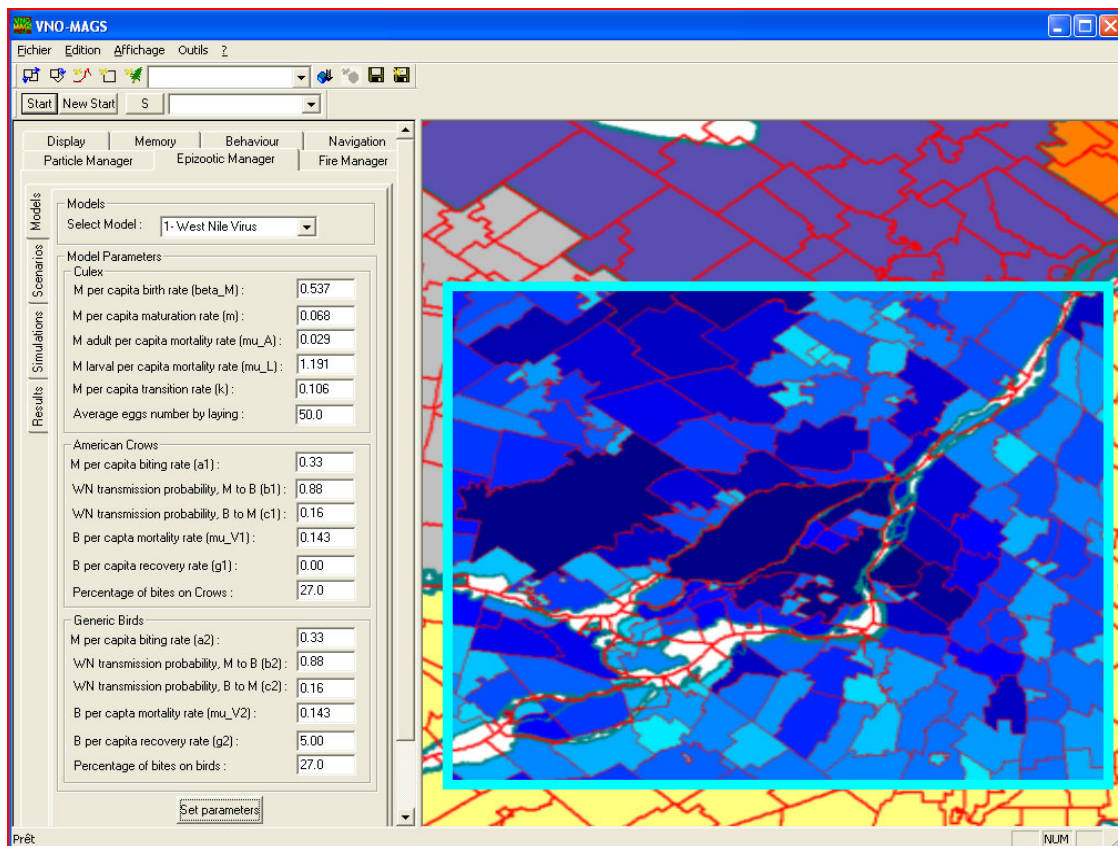


Figure 91: Utilisation d'une carte de densité intelligente pour représenter les *Culex*.

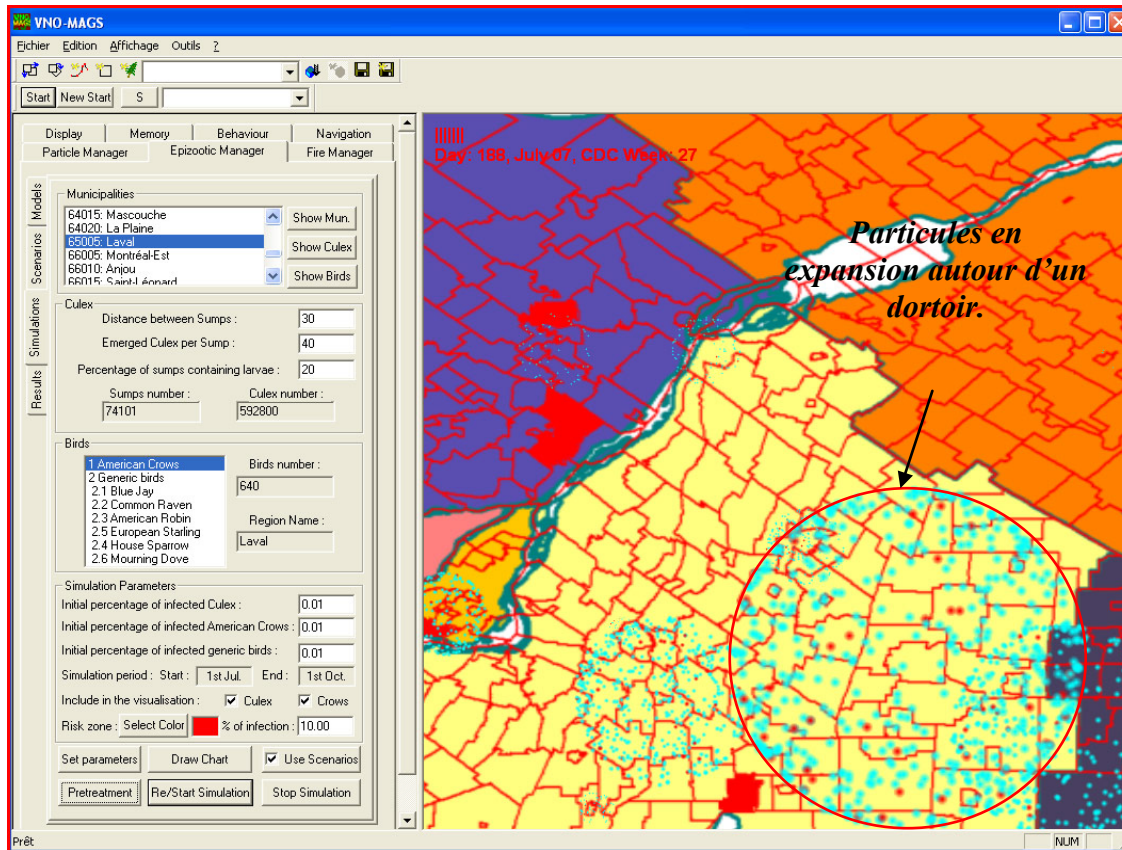


Figure 92: Utilisation des dortoirs (systèmes de particules) pour représenter les corneilles.

Nous avons également modélisé les interactions entre les deux types de populations en utilisant entre autres les équations du modèle mathématique décrit à la section 2.1.2, mais aussi en utilisant les informations stockées dans l'IVGE. Ces interactions sont très importantes pour notre système puisqu'elles permettent de propager le VNO à travers le territoire. En effet, un système de particules représentant un dortoir peut, à un certain moment de la simulation, chevaucher une zone dont la densité de *Culex* est assez importante. Par conséquent, le système *WNV-MAGS* estime le nombre d'individus infectés en se basant sur la probabilité que les corneilles puissent contracter le VNO en présence des *Culex*. Par ailleurs, l'utilisateur peut visualiser l'ampleur de la propagation du VNO sur la carte de deux façons. Le système peut soit modifier la couleur des particules représentant

les corneilles infectées (une particule bleue devient rouge) ou la couleur du polygone représentant une municipalité contenant une grande densité de *Culex* infectés (Figure 92).

En outre, nous avons ajouté d'autres espèces d'oiseaux dans la simulation afin d'augmenter les opportunités de piqûre pour les *Culex*, de sorte que la population de corneilles ne disparaît pas à la fin de la période de simulation. En effet, nous avons introduit dans le système *WNV-MAGS* une autre population 'globale' d'oiseaux que nous avons appelée "les oiseaux génériques" (Corbeau, Geai Bleu, Merle d'Amérique, Moineau Domestique, Étourneau Sansonnet et Tourterelle Triste) qui sont résidents dans les municipalités et qui sont connues pour être potentiellement porteurs du VNO [Koenig *et al.*, 2007]. Dans la simulation, la distinction que nous avons établie entre les corneilles et les "oiseaux génériques" s'appuie sur la supposition que ces oiseaux ne se déplacent pas à l'extérieur des municipalités (pas comme des corneilles qui peuvent voler loin de leur dortoir). Par conséquent, les oiseaux génériques restent en contact avec la même population de *Culex* pendant la simulation. D'ailleurs, nous avons déjà représenté ces oiseaux en utilisant la notation $O_{1..m}^{S_{1..n-\{p,..,k\}}}$ dans le formalisme de *MATSIM* (voir Section 3.6 du Chapitre V).

2.1.4 Application de différents scénarios et calibration du système

Le système *WNV-MAGS* est un outil d'aide à la décision. Il a été implanté en utilisant une approche de géosimulation qui est quelque peu différente des simulations traditionnelles qui sont plus utilisées à des fins de prédiction [Site Web: *Geosimulation*]. En effet, le système *WNV-MAGS* simule la propagation du VNO en proposant à l'utilisateur de spécifier des scénarios afin d'explorer diverses situations, y compris des changements climatiques et différentes stratégies d'intervention. Le choix de ces scénarios a été le résultat de concertations avec les experts du domaine. L'utilisateur peut d'ailleurs choisir parmi quatre types de scénarios différents qui influencent surtout la dynamique des populations de *Culex* (Figure 93).

Le premier scénario est le scénario par défaut qui utilise les conditions moyennes de température et de précipitations (p.ex. les normales canadiennes [Site Web: *Environment*

Canada]). Dans un second type de scénario, l'utilisateur peut choisir une date au cours de laquelle des pluies abondantes peuvent lessiver les puisards dans certaines municipalités. Ce scénario a pour conséquence de tuer les larves présentes dans ces puisards. Cependant et étant donné que les *Culex* adultes continueront à pondre des œufs, ils vont faire en sorte de régénérer les populations larvaires (voir Figure E3 de l'Annexe E). De la même manière, le troisième scénario est utilisé pour simuler l'application de larvicide dans une certaine zone (c.-à-d. municipalité). Le dernier scénario est une combinaison du deuxième et du troisième scénario. Par conséquent, il est possible de choisir plusieurs dates pour le lessivage des puisards et pour l'application des larvicides. Pour tous ces scénarios, l'utilisateur peut changer les conditions climatiques en chargeant les données de température et/ou de précipitations.

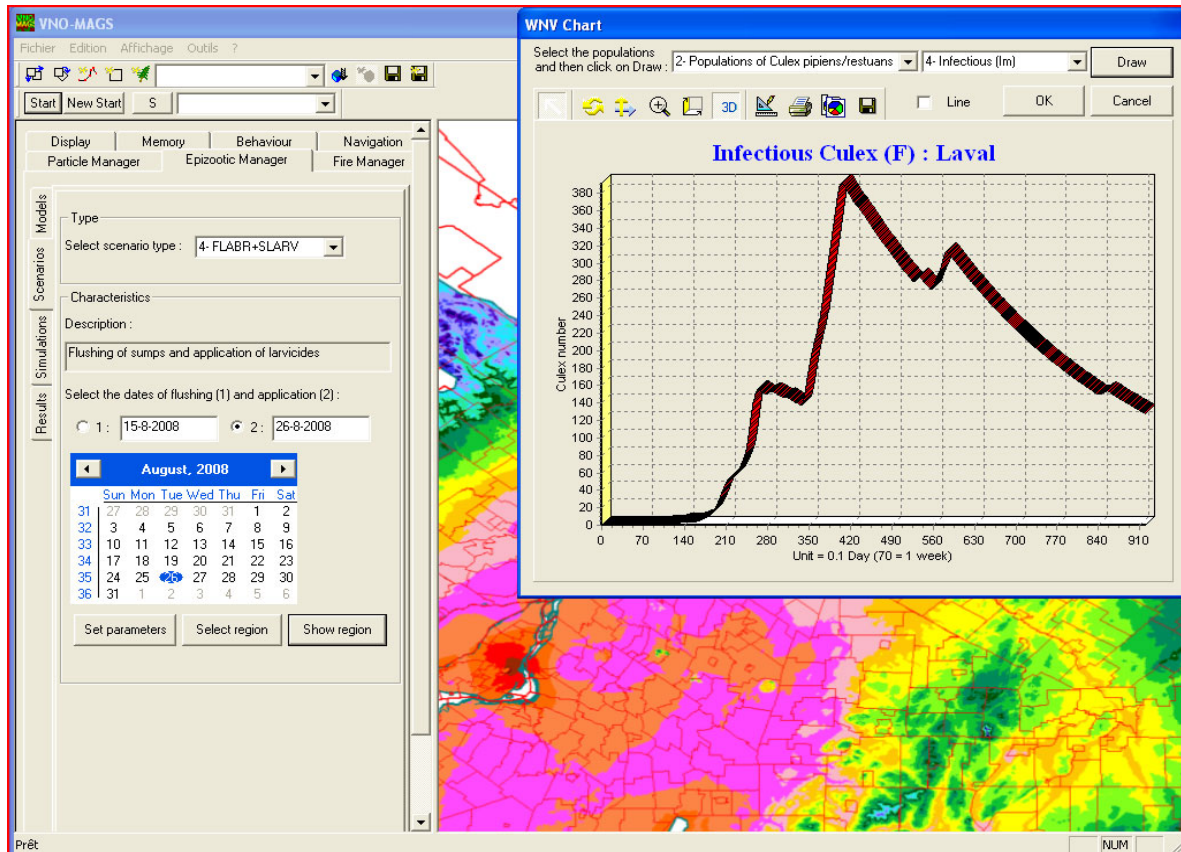


Figure 93: Utilisation des scénarios alternatifs d'intervention.

En outre, le système *WNV-MAGS* offre plusieurs autres fonctionnalités à l'utilisateur comme par exemple: (1) la modification des paramètres du modèle mathématique, (2) la génération de graphiques montrant l'évolution des individus appartenant aux différents compartiments des *Culex* et des corneilles et (3) la production des résultats représentant différents types de cartes comme la répartition des populations sur le territoire, le modèle numérique d'élévation du terrain, la variation des degrés jours cumulés ou encore des précipitations sur le territoire (voir Figure E2 de l'Annexe E).

Par ailleurs et comme nous l'avons déjà mentionné, nous avons calibré le système *WNV-MAGS* en utilisant les données du système de surveillance SIDVS de l'INSPQ (voir Section 1.2 du chapitre I). Par exemple, nous avons estimé les populations réelles de *Culex* en utilisant les données de captures en nous basant sur les travaux de Reisen et ses collaborateurs [Reisen *et al.*, 1991], [Reisen *et al.*, 1992] (un moustique capturé représente approximativement une population de 300 *Culex* par km²). Nous avons, par contre, dû faire plusieurs traitements statistiques afin de gérer les données manquantes ainsi que le biais présent dans les données (p.ex. certaines données ne sont pas répétées dans la même zone et/ou à la même période). De plus et étant donné que nous n'avons pas de données pour toutes les régions, nous avons calibré les résultats de la simulation pour certaines municipalités clés dans lesquelles des infections humaines ont été signalées. Il est apparu par la suite qu'il y a une différence significative entre les données générées par le système *WNV-MAGS* et celles obtenues sur le terrain. De ce fait, nous avons apporté des modifications à certains paramètres du système (p.ex. modifier le nombre de *Culex* qui émergent des puisards permettant de calculer une population initiale, ajuster le nombre initial de *Culex* infectés au début de la saison, etc.) ainsi qu'à certains paramètres du modèle mathématique (p.ex. proportion de moustiques qui vont piquer des corneilles par jour, probabilité de transmission du VNO des moustiques vers les corneilles ou des corneilles vers les moustiques, etc.). Néanmoins, les paramètres biologiques n'ont pas été modifiés (p.ex. nombre de pontes par *Culex* adulte) pour ne pas réduire la plausibilité des résultats. Ces changements nous ont aidés à calibrer quantitativement les résultats de la simulation pour les municipalités clés choisies.

D'un autre côté et afin d'évaluer notre outil avant et après calibration, nous avons rencontré à plusieurs reprises les experts du domaine (p.ex. les spécialistes de l'INSPQ et du Ministère de la Santé et des Services sociaux, les entomologistes de la SOPFIM, les biologistes de l'Université du Québec à Trois Rivières, etc.) pour bénéficier de leurs rétroactions par rapport aux résultats obtenus. Ainsi, ces experts nous ont aidés à mieux calibrer notre système. En effet, nous avons constaté qu'il y avait initialement une différence significative entre les résultats du modèle mathématique et les données sur le terrain. Après consultation avec ces experts, nous nous sommes rendu compte que nous avons oublié d'introduire dans le scénario de certaines municipalités les applications de larvicides qui avaient été faites sur le terrain pendant la période couverte par la simulation. En introduisant effectivement l'application des larvicides dans le scénario aux dates fournies par la SOPFIM, les résultats de la simulation se sont significativement rapprochés des données recueillies sur le terrain (voir Figure E4 de l'Annexe E). Cet exemple montre bien l'importance d'avoir une rétroaction des experts du domaine lors du développement d'un outil d'aide à la décision comme *WNV-MAGS*.

2.2 Échelle spatiale Méso (Ottawa)

Nous venons de voir dans la section précédente que l'utilisation des données du sud de Québec nous a permis de calibrer le système *WNV-MGAS* et surtout les paramètres du modèle mathématique utilisé. La validation de ce modèle nous semble une étape importante et nous pensons qu'il faut le faire en l'appliquant sur une autre région. La région métropolitaine d'Ottawa a été choisie pour cette validation en raison de la disponibilité des données du VNO et des similarités géographiques avec les régions du sud de Québec. Nous commençons cette section en expliquant le changement effectué au niveau de l'échelle spatiale. Nous présentons par la suite la préparation des données et la création de l'*IVGE*. Nous terminons cette section en présentant le résultat de la simulation et la validation du modèle mathématique.

2.2.1 Changement de l'échelle spatiale

La simulation du VNO en utilisant les données de la région métropolitaine d'Ottawa implique un changement d'échelle spatiale. En effet et jusqu'à maintenant, nous avons travaillé à l'échelle de la province de Québec. Nous avons donc subdivisé notre zone géographique en municipalités. Cependant et afin de valider notre modèle avec les données de la région métropolitaine d'Ottawa, il faut choisir à quel niveau de détail nous avons besoin de représenter les données sur la carte surtout qu'on travaille maintenant au niveau d'une plus petite région (voir Section 3.1 du Chapitre IV, Figure 42).

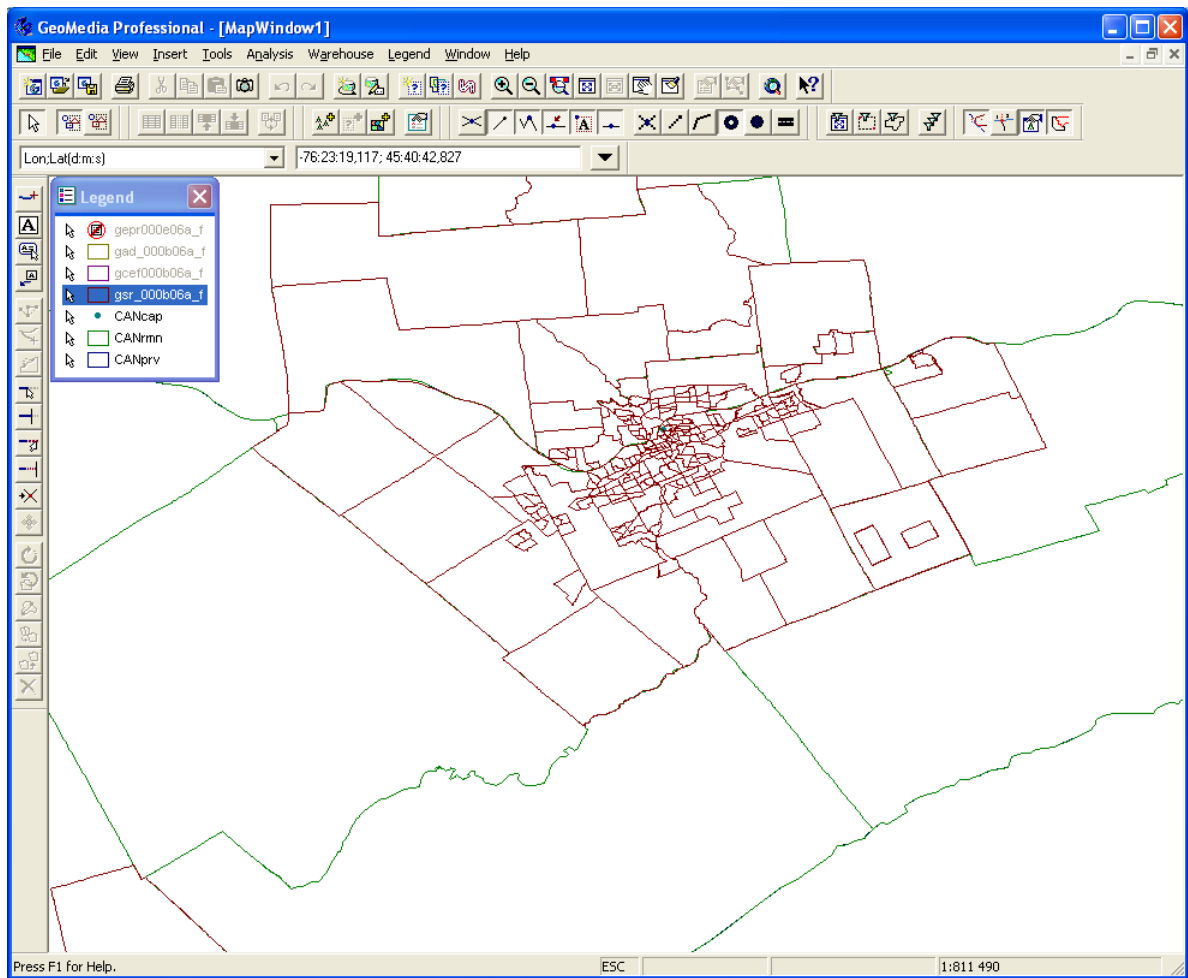


Figure 94: Subdivision de la région d'Ottawa selon les secteurs de recensement (SR).

Nous avons donc besoin de choisir une subdivision significative pour cette région par rapport aux données de capture de moustiques et de collectes d'oiseaux. Nous avons

identifié la possibilité d'utiliser trois subdivisions : (1) les circonscriptions électorales fédérales (CEF)³², (2) les secteurs de recensement³³ (SR) ou encore (3) les aires de diffusion³⁴ (AD). Nous avons finalement opté pour l'utilisation des secteurs de recensement étant donné qu'ils sont utilisés par les spécialistes du domaine de la santé publique (voir Figure 94 et Figures E5 et E6 de l'Annexe E).

2.2.2 Préparation des données et création de l'IVGE

Pour adapter notre système à une région comme celle d'Ottawa, nous avons commencé par récupérer les données SIG représentant cette région. Nous avons utilisé les données du *DMTI*, mais aussi les données géospatiales de la bibliothèque de l'Université Laval [Site Web: *GéoStat*] pour avoir les polygones représentant les secteurs de recensement (Figure 95). Ces polygones représentent donc les cellules élémentaires de l'IVGE et qui sont identifiées d'une façon unique par les mêmes valeurs numériques (un SR est identifié par un numéro unique, exp: 9330005.00) utilisées par Statistique Canada [Site Web: *StatCan*].

³² Une CEF est une région représentée par un député à la Chambre des communes [Site Web: *StatCan*].

³³ Un SR est une petite région géographique relativement stable qui compte habituellement entre 2 500 et 8 000 habitants [Site Web: *StatCan*].

³⁴ Une AD est une petite unité géographique relativement stable formée de un ou de plusieurs îlots de diffusion (territoire équivalant à un pâté de maisons dont les côtés sont délimités par des rues formant des intersections) avoisinants [Site Web: *StatCan*].

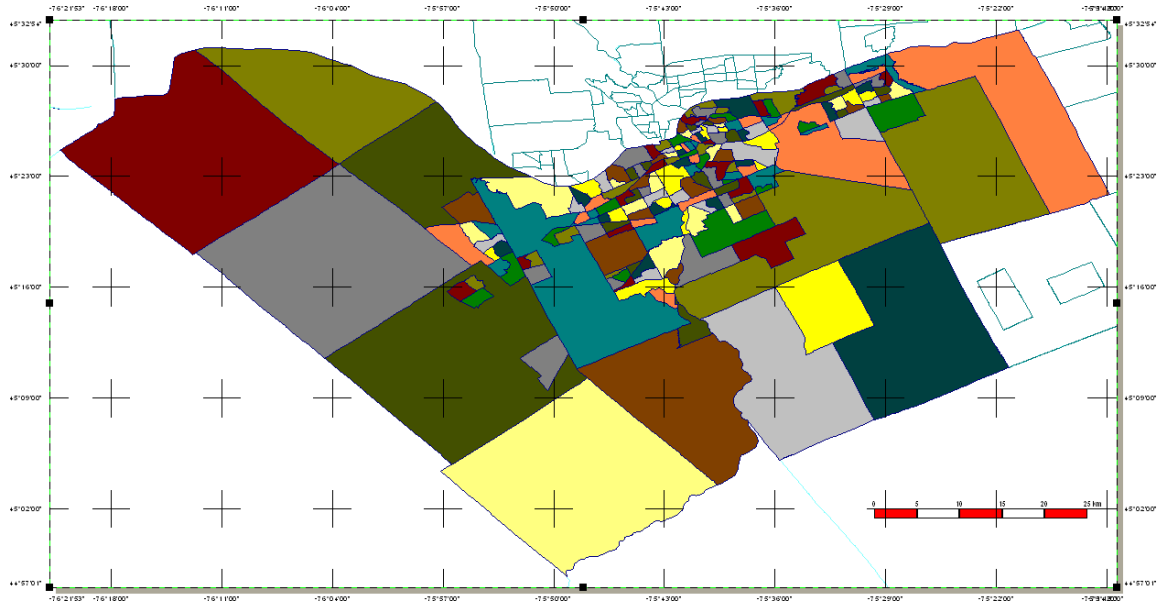


Figure 95: Zone géographique qui nous intéresse pour la région d'Ottawa.

D'un autre côté et comme nous l'avons fait pour le sud de Québec, nous avons attaché aux cellules les données de populations de *Culex* et de corneilles en ajoutant des attributs à ces cellules. Par exemple, nous avons utilisé l'application *Visual Basic* mentionnée auparavant (il a fallu faire des ajustements pour l'adapter aux données ontariennes) afin de calculer la longueur des routes dans chaque secteur de recensement. Ce calcul a permis d'estimer le nombre des puisards et de ce fait, le nombre des populations initiales de *Culex* (Figure 96).

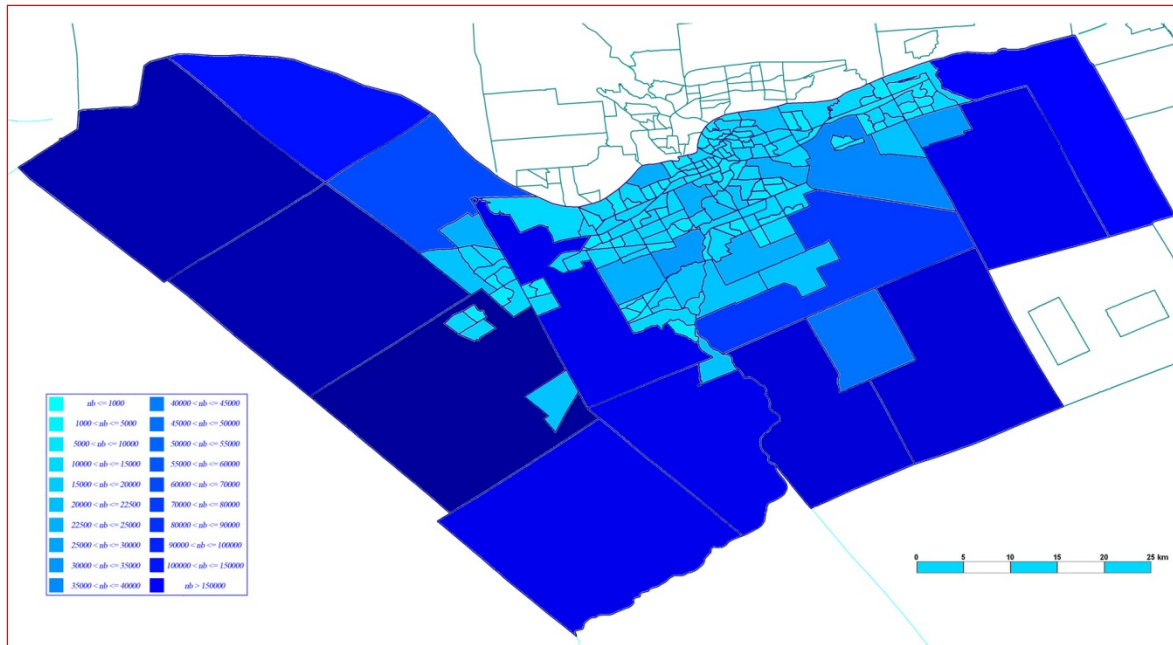


Figure 96: Population initiale de *Culex* selon les secteurs de recensement.

En outre, Il a fallu trouver une base de données qui permet de calculer la densité de populations d'oiseaux réservoirs (corneilles et/ou autres) dans le but d'alimenter les données initiales du modèle. Pour le Québec, nous avons utilisé la base ÉPOQ. Cependant, il ne semble pas exister une base semblable à ÉPOQ pour la région d'Ottawa. Nous avons alors demandé l'avis d'experts et ils nous ont conseillé d'utiliser la base *eBirds* [Site Web: *eBird*]. De ce fait, nous avons calculé une densité moyenne des oiseaux impliqués dans la propagation du VNO et nous avons trouvé quelques signalements de dortoirs de corneilles dans la zone géographique qui nous intéresse. Par ailleurs, nous avons pensé au début calculer les données climatiques de la même façon que nous avons procédé pour la province de Québec. Cependant et après avoir testé les données de plusieurs secteurs, nous avons constaté que les variations de température ne sont pas significatives pour la région qui nous intéresse, compte tenu de sa faible étendue et que le relief est passablement plat. Ainsi, nous n'avons pas eu besoin de traiter les données climatiques pour chaque secteur de recensement. Ceci démontre que le choix de l'échelle spatiale peut influencer la nature des données alimentant la simulation.

2.2.3 Résultat de la simulation et validation du modèle

Nous avons simulé la propagation du VNO pour la région métropolitaine d'Ottawa en utilisant les mêmes paramètres qui ont été ajustés après calibration des résultats des municipalités clés du sud de Québec. La Figure 97 montre une capture d'écran du système *WNV-MAGS* avant le début de la simulation. L'utilisateur peut d'ailleurs sélectionner un secteur de recensement (p.ex. polygone en bleu) et afficher les populations initiales de moustiques (c.-à-d. *Culex*) et d'oiseaux (c.-à-d. corneilles et oiseaux génériques).

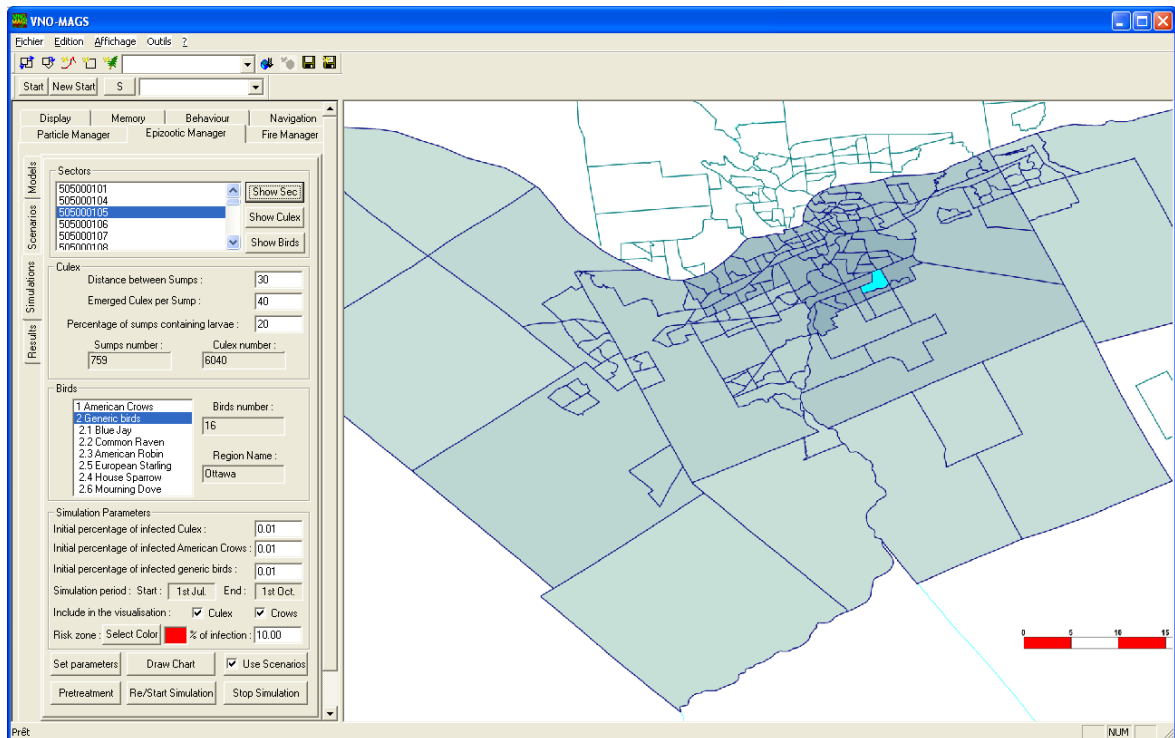


Figure 97: Capture d'écran de *WNV-MAGS* avant la simulation pour la région d'Ottawa.

Après avoir sélectionné le type de scénario et modifié éventuellement les paramètres de la simulation, l'utilisateur peut lancer la simulation en cliquant sur le bouton "*Re/Start Simulation*" (Figure 97). De ce fait, la simulation démarre en affichant jour après jour (du 1^{er} juillet jusqu'au 1^{er} octobre) l'évolution des populations de corneilles (c.-à-d. mouvement des dortoirs) et les changements des couleurs des particules ainsi que des secteurs de recensements indiquant la propagation de l'infection (Figure 98). *WNV-MAGS* affiche en haut de la fenêtre (en rouge) le jour et la semaine *CDC* (*Centers for Disease Control and Prevention*) pendant le déroulement de la simulation. En fait, *CDC* numérote les semaines

de l'année civile de 1 à 52 afin de permettre la comparabilité des données d'une année à l'autre. De ce fait et étant donné que les décideurs de la santé publique utilisent également les semaines *CDC* (programme de surveillance, application des larvicides, etc.), notre système *WNV-MAGS* leur permet de mieux comprendre les résultats de la simulation (Figure 98).

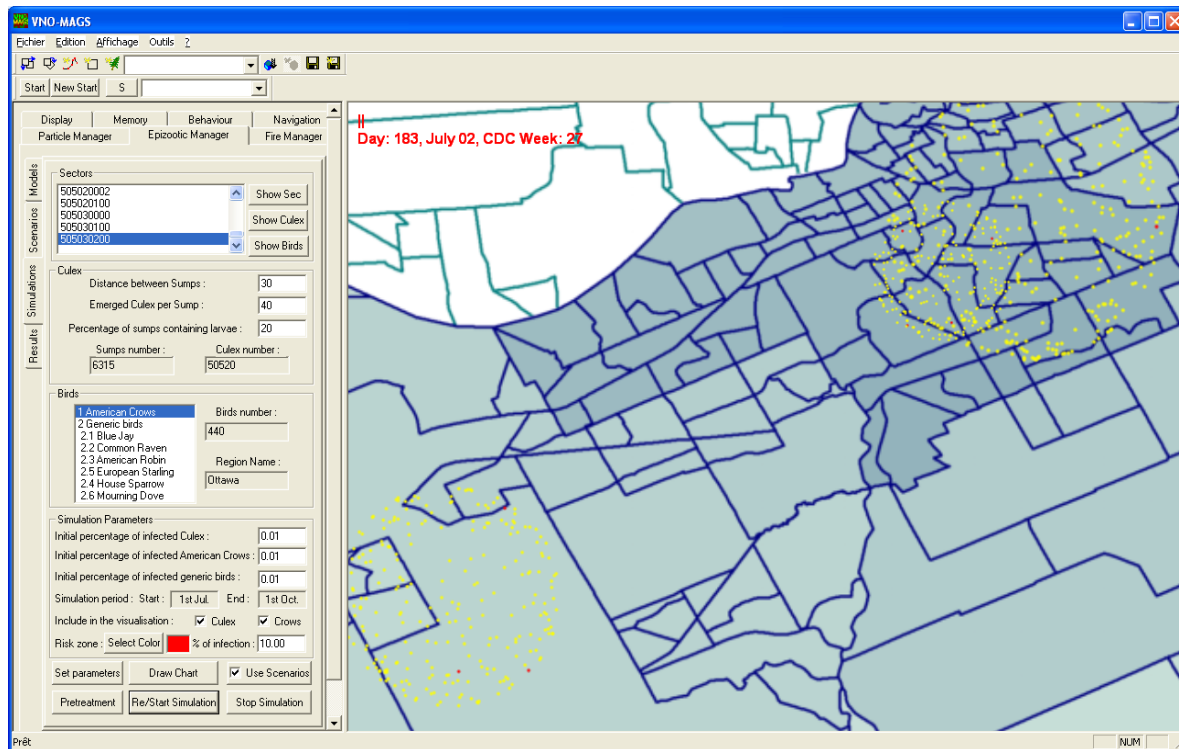


Figure 98: Simulation du VNO pour la région d'Ottawa.

En outre et afin de montrer la validité du modèle mathématique, nous avons sélectionné un secteur de recensement clé pour voir si les données de la simulation sont calibrées par rapport aux données recueillies sur le terrain (c.-à-d. données de capture de *Culex* provenant de la SOPFIM pour la région d'Ottawa). Il s'agit du secteur 5050011.04 contenant le plus gros dortoir de corneilles (situé sur le *General Hospital Campus*) et qui a le plus de stations de capture de *Culex* (Figure 99). La Figure 100 montre justement l'évolution de la population des *Culex* (données du modèle par rapport aux données sur le terrain) après traitement des données manquantes pour le secteur de recensement 212

5050011.04. Cette figure montre que les données du modèle mathématique se rapprochent de celles sur le terrain, surtout que nous avons utilisé un scénario qui faisait l'application des larvicides aux mêmes périodes où elle a été effectuée sur le terrain (Figure 100).

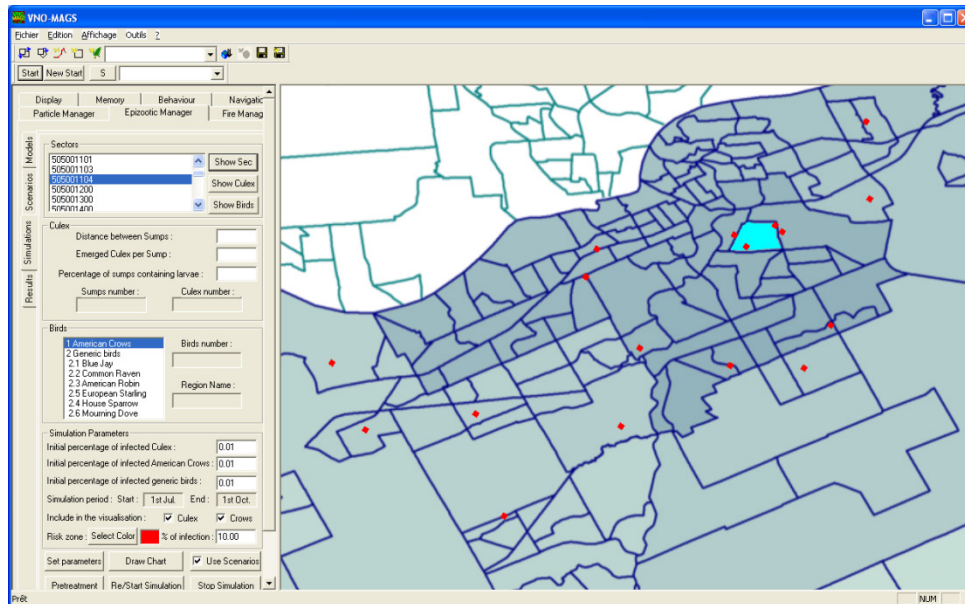


Figure 99: Stations de capture de *Culex* pour la région d'Ottawa (le secteur 5050011.04 contient 4 stations).

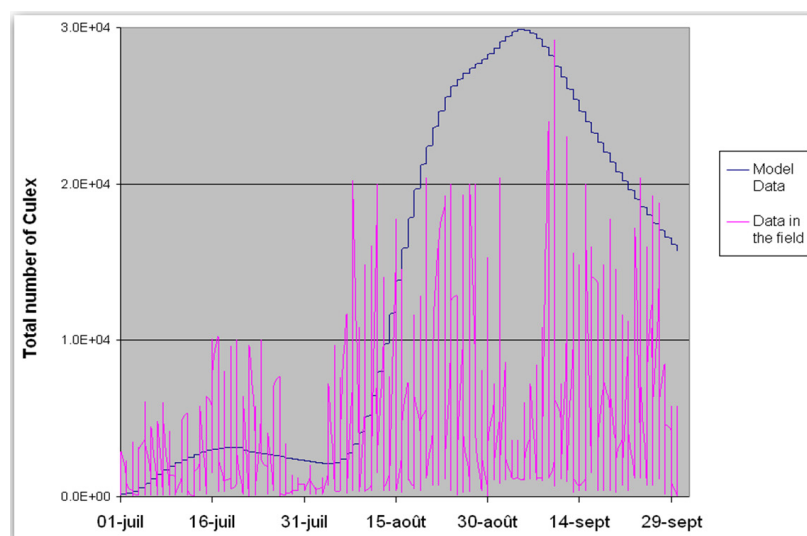


Figure 100: Évolution de la population de *Culex* (données du modèle vs données de terrain) après application de larvicides pour le secteur de recensement 5050011.04.

2.3 Discussion

L'outil de géosimulation *WNV-MAGS* que nous venons de présenter offre beaucoup d'avantages surtout si nous le comparons à d'autres outils comme par exemple *STEM* qui a été déjà présenté (voir Section 3.3 du Chapitre III) ou encore *GLEaMviz* (*Global Epidemic and Mobility Model Visualization* [Site Web: GLEaMviz]). Ce dernier outil (c.-à-d. *GLEaMviz*) permet de simuler une épidémie à une seule et unique échelle Macro (c.-à-d. à l'échelle mondiale). Par contre, le système *WNV-MAGS* permet comme nous venons de le voir de simuler le VNO en utilisant plusieurs échelles spatiales. Les avantages du changement d'une échelle spatiale, ou d'une façon plus générale les avantages du changement d'un niveau de granularité, offrent à l'utilisateur la possibilité de mieux comprendre le phénomène étudié, mais aussi d'établir des plans d'intervention en fonction du niveau choisi. Nous avons d'ailleurs parlé dans notre approche de géosimulation multi-niveau (voir Figure 36) qu'une simulation à un niveau Macro pourrait aider les décideurs à établir des directives d'intervention à un niveau stratégique, suite par exemple à des décisions politiques. Cependant, une simulation Méso et/ou Micro pourrait aider à établir des directives d'intervention à un niveau tactique ou opérationnel (c.-à-d. établir des plans d'intervention concrets). De plus, nous rappelons que le choix du niveau de granularité dépend des besoins de l'utilisateur, de la disponibilité des modèles et des données alimentant ces modèles (voir Figure 39). C'est pour ces raisons d'ailleurs que nous avons choisi le sud de Québec ainsi que la région métropolitaine d'Ottawa, afin de simuler la propagation du VNO puisque les données étaient entre autres disponibles pour ces régions. Évidemment, si les modèles et les données sont disponibles pour une autre région, il serait possible d'adapter *WNV-MAGS* à cette région surtout que le système est bien calibré.

En outre, *WNV-MAGS* présente certaines limites surtout par rapport à sa flexibilité, étant donné qu'il est conçu spécialement pour le VNO et il serait difficile de l'adapter à une autre

zoonose. Par exemple, le modèle mathématique est intégré dans le système et nous ne pouvons pas le modifier sans changer le code C⁺⁺ du module l'implantant. C'est d'ailleurs la raison principale qui nous a amenés à penser à développer un outil générique.

3 Zoonosis-MAGS

Nous présentons dans cette section les travaux effectués, ainsi que les résultats obtenus jusqu'à maintenant, afin de généraliser notre solution pour geosimuler la propagation des zoonoses comme le VNO ou la maladie de Lyme [Bouden et Moulin, 2010], [Bouden et Moulin, 2012b]. Nous commençons par la présentation de l'architecture conceptuelle de *Zoonosis-MAGS*. Nous présentons par la suite les concepts que nous proposons d'ajouter à *MASTIM* (voir Section 3 du Chapitre V) pour formaliser la simulation, étant donné que notre objectif est de produire un outil générique d'aide à la décision. Finalement, nous présentons la géosimulation de la maladie de Lyme incluant les résultats obtenus.

3.1 Architecture conceptuelle

Nous avons choisi d'utiliser la méthode EPAS (Entrée, Processus, Accumulation, Sortie) [Moulin, 1985] afin de représenter l'architecture conceptuelle du système *Zoonosis-MAGS* en nous inspirant bien sûr de l'approche de géosimulation décrite auparavant. En effet, cette représentation est basée sur un ensemble d'éléments en interaction dynamique. Les rectangles dont la couleur est verte et le contour régulier représentent les processus et les rectangles dont la couleur est bleue et le contour ovale représentent les accumulations de données. Tous les processus sont numérotés en P_i et toutes les accumulations sont numérotées en A_{j_k} . En fait, le processus global du système *Zoonosis-MAGS* porte le numéro P_0 . Ce processus peut contenir des sous-processus qui sont numérotés de P_1 à P_n et des accumulations qui sont numérotées de A_{0_1} à A_{0_m} . Le processus P_1 , qui est un sous-processus de P_0 , peut contenir des sous-processus qui sont numérotés de $P_{1.1}$ à $P_{1.p}$ et des accumulations qui sont numérotées de A_{1_1} à A_{1_s} . Une accumulation A_{0_1} peut contenir des sous-accumulations qui sont numérotées de $A_{0_{1.1}}$ à $A_{0_{1.t}}$.

La Figure 101 présente le processus P0 qui est le processus global du système *Zoonosis-MAGS*. Nous retrouvons à l'extérieur de ce processus, les bases de données qui vont alimenter le système, mais aussi les autres systèmes comme le système de surveillance (p.ex. *SIDVS-VNO*) et le système d'information géographique (p.ex. *Geomedia*, *PostGIS*, etc.). Nous retrouvons également les acteurs qui vont manipuler ce système : l'utilisateur (par exemple un responsable de la santé publique) et le développeur (qui joue le rôle d'analyste et de concepteur). Par ailleurs et en allant du bas vers le haut de la figure, le processus P7 permet de préparer les données nécessaires pour la simulation comme par exemple les données des populations (p.ex. la densité des oiseaux en utilisant les bases ÉPOQ et/ou eBird, le nombre initial des *Culex* en utilisant le nombre de puisards, le nombre moyen de tiques juvéniles qui arrivent du Sud en s'accrochant aux oiseaux migrateurs, etc.). Ce processus permet aussi de préparer les données climatiques qui peuvent être des données réelles (en utilisant par exemple les serveurs d'Environnement Canada) ou les normales canadiennes. Le processus P7 permet également de préparer d'autres données qui sont reliées à la nature du phénomène simulé, comme par exemple les paramètres du modèle mathématique utilisé pour propager la maladie étudiée. En outre, le processus P8 permet de préparer les données SIG qui vont servir de données de base pour la génération de l'*IVGE*. Les processus P7 et P8 vont produire les accumulations A0₆, A0₇ et A0₈ qui sont stockées dans la base de données de l'application *Zoonosis-MAGS* (nous utilisons une base de données relationnelle de type *PostgreSQL* contenant une base spatiale *PostGIS* [Site Web: *PostGIS*]).

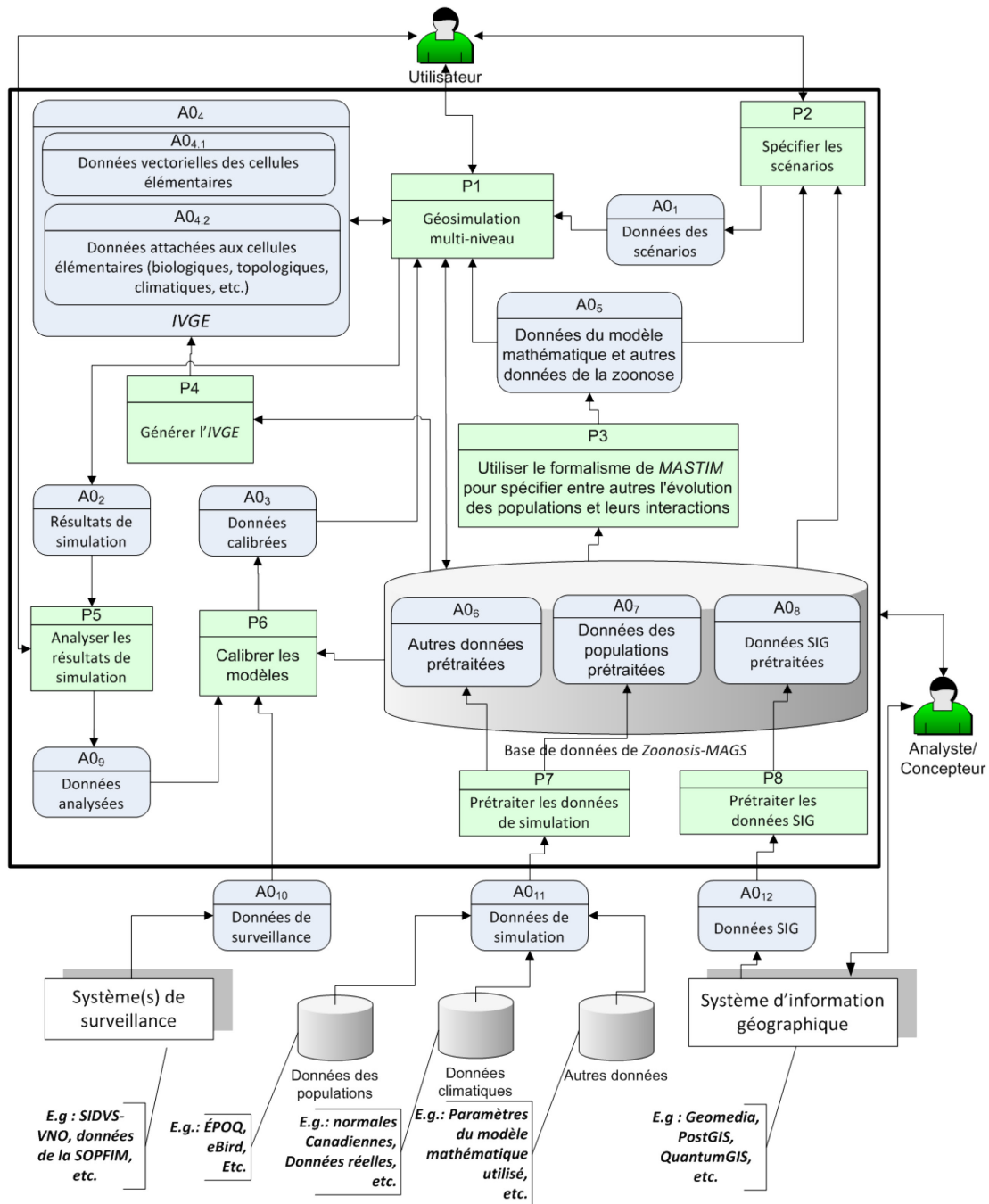


Figure 101: Architecture conceptuelle de Zoonosis-MAGS (basée sur la méthode EPAS).

Par la suite, le processus P4 permet de générer l'IVGE en utilisant les données prétraitées se trouvant dans la base de données et en produisant l'accumulation A04. Cette accumulation contient une sous-accumulation A041 représentant les données vectorielles des cellules

élémentaires de l'*IVGE*. Elle contient également une autre sous-accumulation A0₄₂ représentant les données (biologiques, topologiques, climatiques, etc.) attachées aux cellules élémentaires. C'est cette dernière sous-accumulation qui donne le caractère informé à l'environnement de simulation (c.-à-d. l'*IVGE*) comme nous l'avons expliqué dans le chapitre précédent (voir Section 3.5 du Chapitre VI). Par ailleurs, le processus P3 utilise les données contenues dans la base de données de *Zoonosis-MAGS*, ainsi que le formalisme de *MASTIM*, pour spécifier entre autres l'évolution et les interactions des populations impliquées dans le phénomène simulé (c.-à-d. propagation d'une zoonose). Ce processus peut spécifier par exemple les espèces, les compartiments types, les transitions d'évolution types, les interactions instantanées et durables, etc. Par conséquent, le processus P3 produit l'accumulation A0₅ contenant les données du modèle mathématique et les autres données de la zoonose. Pour sa part, le processus P2 permet de spécifier les scénarios de la simulation en utilisant également les données de la base de données et produisant l'accumulation A0₁ contenant les données de ces scénarios (voir Section 3.2.2).

Toutes les accumulations mentionnées auparavant, y compris celles de l'*IVGE*, vont alimenter le processus P1 qui a l'importante tâche de faire la géosimulation multi-niveau de la propagation de la zoonose. Ce processus contient entre autres le moteur de la simulation permettant de calculer par exemple les interactions spatio-temporelles des populations impliquées dans la propagation de la maladie. D'un autre côté, le processus P1 va bien sûr produire des résultats (accumulation A0₂) qui sont analysés par le processus P5. Les données analysées (accumulation A0₉) pourront être utilisées par le processus P6 pour calibrer le ou les modèles mathématiques utilisés. Ce dernier processus va utiliser également les données du système de surveillance (accumulation A0₁₀) ainsi que le regroupement d'accumulations se trouvant dans la base de données (A0₆ à A0₈). Pour sa part, l'utilisateur du système peut intervenir dans la spécification des scénarios, la géosimulation multi-niveau (au niveau de l'interface de l'application pour spécifier par exemple les paramètres de la simulation) et l'analyse des résultats. De son côté,

l'analyste/concepteur peut intervenir bien sûr dans tous les processus, étant donné que c'est lui qui développe toute l'application.

Par ailleurs, la méthode EPAS se base sur une approche de raffinement qui consiste à présenter premièrement le processus global du système comme nous venons de le faire. Ensuite, il s'agit de détailler chacun des sous-processus appartenant à ce processus global. Ce principe hiérarchique permettant de détailler au fur et à mesure les processus d'un système jusqu'à en arriver à la précision recherchée avec la possibilité d'un raffinement par un retour en arrière, nous semble très efficace surtout pour le développement d'un système multi-niveau. En effet, nous pensons qu'il y a un lien étroit entre la conception d'un système et le système lui-même. Ce système même s'il est livré à ses utilisateurs, reste toujours inachevé. Ceci découle de l'ingénierie des systèmes complexes qui s'adresse aux problèmes et solutions qui changent constamment [Kuras, 2007]. D'ailleurs, la conception d'un système évolue au cours de son cycle de vie en passant des besoins à la solution par un processus itératif. L'aspect itératif est dû à la complexité du problème global.

3.2 Formalisme de la simulation

Nous présentons dans cette section les notations formelles des plus importants aspects qui nous intéressent et qui concernent la simulation. Ces notations compléteront le formalisme du modèle *MASTIM* comme nous l'avons fait pour le transfert des individus. Nous rappelons que toutes ces notations se retrouvent dans l'Annexe A.

3.2.1 L'état du système

Nous commençons par définir l'état du système comme étant sa situation à un moment précis de la simulation. En fait, nous distinguons un état différent pour chaque pas de simulation. Nous rappelons qu'un pas de simulation reflète l'échelle temporelle choisie par l'utilisateur (p.ex. un jour, une semaine, etc.). De ce fait, le changement du pas de simulation implique normalement un changement de l'état du système, à moins qu'il n'y ait eu aucune modification du système. D'ailleurs, la modification d'un système se fait par le changement des valeurs de ses attributs (voir Section 3.7 du Chapitre V). Par conséquent,

un état du système représente, pour un certain pas de simulation, les valeurs que peuvent prendre les attributs du système. Voici comment nous notons un état du système :

$$State_j^{\Sigma_k} = \{A_j^{\Sigma_k}, A_j^{G_m^{H_i}}, A_j^{Z_m^{H_i}}, A_j^{C_m^{H_i}}, A_j^{C_m^{H_i}S_p}, A_j^{C_m^{H_i}O_n^{S_p}}, \dots\}$$

Un état représente donc les valeurs des attributs appartenant aux différentes catégories déjà définies comme par exemple $A_j^{\Sigma_k}$ représentant un attribut global du système (p.ex. temps de la simulation), $A_j^{G_m^{H_i}}$ représentant un attribut d'une région, $A_j^{Z_m^{H_i}}$ représentant un attribut d'une zone sélectionnée, $A_j^{C_m^{H_i}}$ représentant un attribut d'une cellule (p.ex. donnée attachée à la cellule comme la couverture du sol), $A_j^{C_m^{H_i}S_p}$ représentant un attribut d'une espèce se trouvant dans une cellule particulière (p.ex. nombre des individus dans la cellule) et $A_j^{C_m^{H_i}O_n^{S_p}}$ représentant un attribut d'un compartiment d'une espèce se trouvant dans une cellule. Il peut y avoir bien sûr d'autres attributs comme par exemple l'attribut de la trajectoire (voir Section 4.1.2 du Chapitre VI).

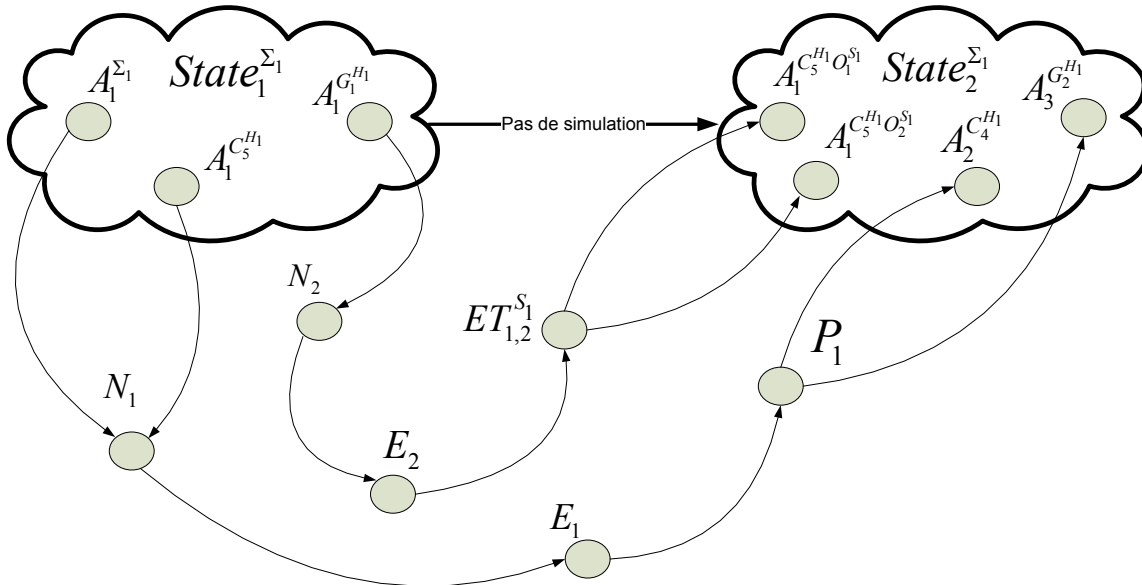


Figure 102: Passage du système *Zoonosis-MAGS* d'un état à un autre.

Afin d'illustrer le changement d'un état à un autre, nous présentons dans la Figure 102 un exemple où le système effectue un pas de simulation en modifiant certains attributs. Cette

figure montre un changement d'un premier état du système $State_1^{\Sigma_1}$ à un deuxième état du système $State_2^{\Sigma_1}$. Dans cet exemple, le système Σ_1 possède 7 attributs (c.-à-d. $A_1^{\Sigma_1}, A_1^{G_1^{H_1}}, A_1^{C_5^{H_1}}, A_1^{C_5^{H_1}O_1^{S_1}}, A_1^{C_5^{H_1}O_2^{S_1}}, A_2^{C_4^{H_1}}, A_3^{G_2^{H_1}}$). Une modification des attributs $A_1^{\Sigma_1}, A_1^{C_5^{H_1}}$ entraîne la satisfaction de la condition N_1 , ce qui va déclencher l'événement E_1 qui à son tour va déclencher le processus P_1 . Ce dernier va modifier les attributs $A_2^{C_4^{H_1}}, A_3^{G_2^{H_1}}$. D'un autre côté, une modification de l'attribut $A_1^{G_1^{H_1}}$ entraîne la satisfaction de la condition N_2 , ce qui va déclencher l'événement E_2 qui à son tour va déclencher la transition d'évolution $ET_{1,2}^{S_1}$. Cette dernière transition va modifier les attributs $A_1^{C_5^{H_1}O_1^{S_1}}, A_1^{C_5^{H_1}O_2^{S_1}}$ qui représentent dans ce cas le nombre des individus dans les compartiments $O_1^{S_1}$ et $O_2^{S_1}$.

3.2.2 Le scénario de la simulation

Le terme scénario est inspiré de la cinématographie. En effet, avant de produire un film le scénariste doit choisir l'endroit et l'époque où va se passer la scène, les acteurs qui vont participer à la scène et l'histoire sur laquelle va tourner la scène. Le décideur de la santé publique représentant l'utilisateur du système *Zoonosis-MAGS* doit essentiellement suivre des étapes similaires pour spécifier un scénario de simulation. Il va préciser par exemple le lieu et la période où va se dérouler la simulation, les acteurs (p.ex. individus, groupes, populations, etc.) qui vont participer à la simulation, le ou les modèles qui vont faire évoluer ces acteurs et calculer leurs interactions, les événements pouvant déclencher des processus au courant de la simulation, etc. Avant de définir un scénario, définissons ce que nous appelons un patron de scénario :

$$ScenarioPattern_m^{Zoonosis_p} = (Id_{SP_m}^{Zoonosis_p}, \Sigma_k, (SimulationModel_k, ModelParameters_k), \dots)$$

Un patron de scénario (identifié par $Id_{SP_m}^{Zoonosis_p}$) donne essentiellement la spécification et les paramètres (p.ex. nombre de pontes par *Culex* adulte) du modèle de simulation

($SimulationModel_k, ModelParameters_k$) qui est utilisé pour simuler un système écologique Σ_k en faisant entre autres évoluer les populations impliquées dans la propagation d'une zoonose et en calculant les interactions de ces populations avec l'*IVGE*. Ce modèle se base bien sûr sur le formalisme de *MASTIM* pour spécifier par exemple les espèces choisies, les compartiments types, les transitions types, les interactions, etc. Le patron de scénario est partiellement spécifié (en ajoutant les deux petits points à la fin de la formule) afin de permettre au concepteur de la simulation d'ajouter d'autres caractéristiques à ce patron comme certains comportements des espèces impliquées dans la propagation de la zoonose. Cette façon de partiellement spécifier une formule est inspirée du type d'un objet ouvert dans le paradigme orienté objet. En fait, un objet ouvert a un type (ensemble de ses méthodes) polymorphe qui peut être utilisé pour instancier plusieurs objets différents (utilisant d'autres méthodes). Cet aspect du type d'objet ouvert permet au langage d'être plus générique [Site Web: OCaml]. Nous définissons maintenant un scénario de la façon suivante :

$$Scenario_j = (Id_{Scenario_j}, ScenarioPattern_m^{Zoonosis_p}, \Delta_{Scenario_j}, \{RefData_n^{Zoonosis_p}\}, \{E_i\}, \dots)$$

$Id_{Scenario_j}$ est l'identificateur du scénario. $ScenarioPattern_m^{Zoonosis_p}$ est le patron de ce scénario. En fait, un scénario est une instance d'un patron de scénario. Par exemple, un patron de scénario spécifie un compartiment type d'une espèce alors qu'un scénario crée des instances de ce compartiment dans toutes les cellules de l'*IVGE* en initialisant le nombre des individus de ces instances. $\Delta_{Scenario_j}$ est la durée de la simulation (p. ex. du 1^{er} juillet au 1^{er} octobre pour le VNO). $\{RefData_n^{Zoonosis_p}\}$ est un ensemble contenant des références à toutes les données qui sont utilisées pour justement instancier le scénario à partir de son patron de scénario (p.ex. les données initiales des populations, les données climatiques, les données vectorielles, etc.). $\{E_i\}$ est l'ensemble des événements qui sont spécifiés dans le scénario et qui pourront déclencher éventuellement des processus. La liste

de ces événements est normalement établie avec des experts du domaine (p.ex. un événement qui va déclencher une application de larvicides, des fortes pluies, des programmes de sensibilisation ou de mise en quarantaine, etc.). Le scénario est également spécifié d'une façon partielle comme c'est le cas pour le patron de scénario.

3.2.3 La simulation d'une zoonose

Une simulation est un procédé selon lequel on exécute un programme informatique sur un ordinateur en vue de simuler un phénomène complexe en se basant sur un ou plusieurs modèles théoriques. Pour exécuter une simulation, il faut se référer à ce qui se trouve dans le scénario de la simulation comme la durée de la simulation. En fait, une simulation s'exécute pendant une certaine durée en modifiant éventuellement l'état du système à chaque pas de simulation. De ce fait, nous définissons une simulation de la façon suivante:

$$Simulation_m^{Zoonosis_p} = (Id\ Sim_m^{Zoonosis_p}, Scenario_j, SimStep_m, floop(State_n^{\Sigma_k}), SimParam_m)$$

$Id\ Sim_m^{Zoonosis_p}$ est l'identificateur de la simulation. $Scenario_j$ est le scénario utilisé par la simulation et qui se base sur un patron de scénario. $SimStep_m$ est le pas de simulation qu'il faut utiliser pour changer les états du système. $floop(State_n^{\Sigma_k})$ est une fonction prenant comme paramètre d'entrée l'état initial du système et utilisant une boucle pour faire changer l'état précédent du système à chaque pas de simulation. $SimParam_m$ représentent les paramètres de la simulation comme par exemple les paramètres de visualisation (p.ex. ne pas afficher les systèmes de particules représentant les dortoirs des corneilles pour ne pas ralentir le système) ou encore les paramètres permettant de sauvegarder les résultats de la simulation.

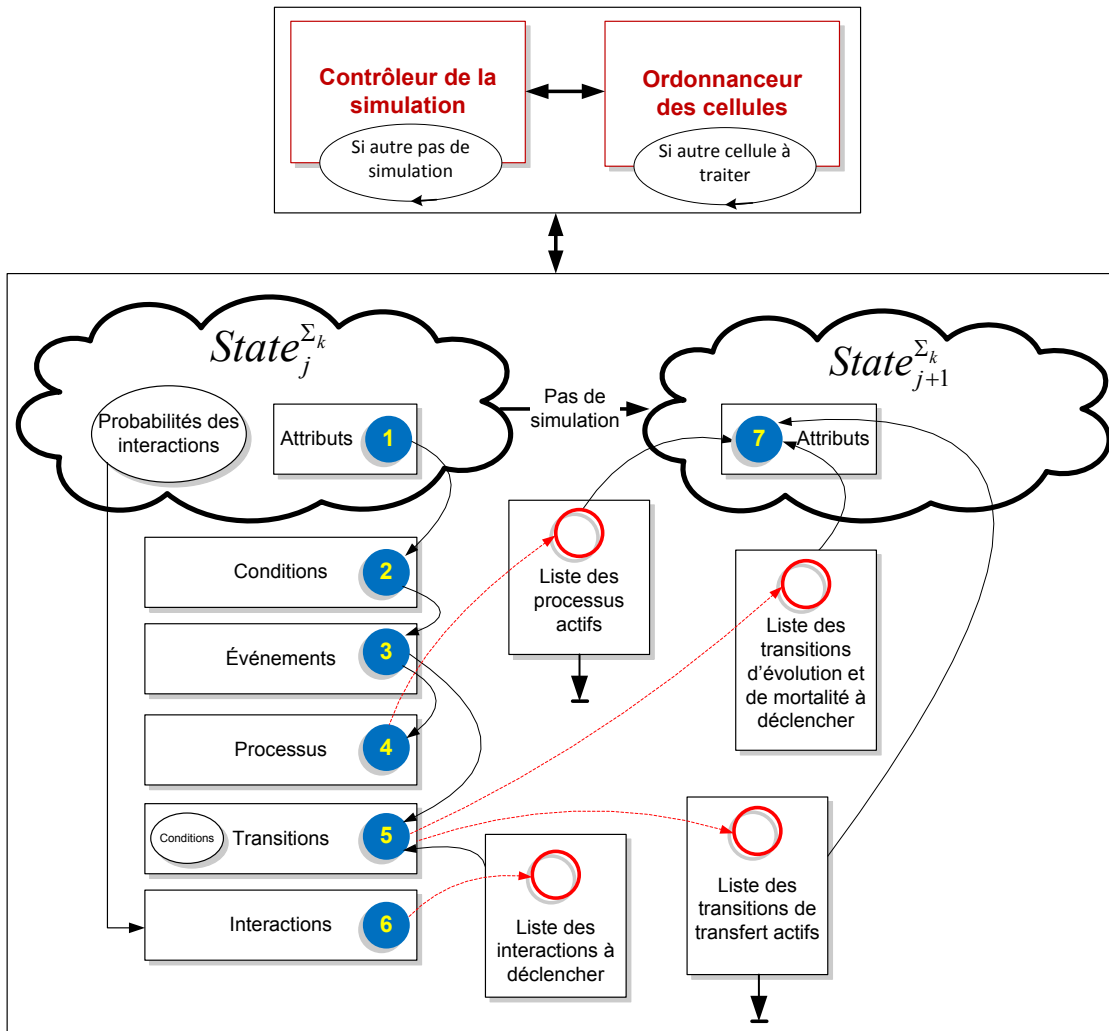


Figure 103: Ordre des étapes au cours d'un pas de simulation.

En outre, la fonction $floop(State_n^{\Sigma_k})$ représente en réalité une façon simple de formaliser le moteur de simulation. En fait, ce moteur de simulation est beaucoup plus complexe puisqu'il fait tous les calculs nécessaires pour mettre à jour par exemple les données se trouvant dans l'*IVGE* (données attachées aux cellules représentées par des attributs dans le système). L'ordre d'exécution de toutes les tâches de ce moteur de simulation est très important étant donné qu'il peut influencer le résultat du calcul. D'ailleurs, nous présentons dans la Figure 103 une vue sommaire (c.-à-d. non exhaustive) des principales étapes qui

sont effectuées par la simulation ainsi que leur ordre d'exécution en s'inspirant du formalisme de *MASTIM*. En effet, le moteur de simulation de *Zoonosis-MAGS* possède un contrôleur permettant de gérer entre autres l'initialisation et l'exécution de chaque pas de simulation (il permet de vérifier s'il y a un nouveau pas de simulation à exécuter) ainsi qu'un ordonnanceur des cellules permettant de vérifier s'il y a une nouvelle cellule à traiter. En fait, ce sont les valeurs des attributs de chaque pas de simulation qui font en sorte que les conditions deviennent vraies. La satisfaction de ces conditions va déclencher éventuellement des événements. Ces événements vont déclencher éventuellement des processus et/ou des transitions. Ces transitions peuvent avoir besoin de certaines conditions pour être déclenchées comme nous l'avons déjà mentionné dans le formalisme de *MASTIM* (voir Section 3.8 du Chapitre V). Par ailleurs, nous distinguons deux types de listes permettant de gérer les processus, les transitions et les interactions: (1) une liste qui déclenche les transitions d'évolution et de mortalité ainsi que les interactions à la fin du pas de simulation et (2) une liste qui active les processus et les transitions de transfert étant donné que ces derniers peuvent durer plus longtemps qu'un seul pas de simulation. En effet, les processus deviennent actifs et s'ajoutent à la liste des processus actifs. Ces processus vont changer les attributs des états futurs si ces processus durent plus longtemps. Certains processus vont devenir inactifs si leur durée est terminée. La flèche qui rappelle un peu la transition de mortalité représente le fait que des processus sont désactivés. Les probabilités des interactions vont déclencher les interactions (instantanées et/ou durables). Ces interactions vont s'ajouter à la liste des interactions à déclencher. Ces interactions auront comme conséquence le déclenchement des transitions. Ces transitions s'ajoutent à la liste des transitions d'évolution et de mortalité à déclencher qui vont changer les attributs de l'état suivant du système. Par contre, les transitions de transfert vont s'ajouter à une liste de transitions de transfert actives. Certaines de ces transitions vont devenir inactives si leur durée est terminée et elles peuvent bien sûr changer les attributs à la fin du pas de simulation.

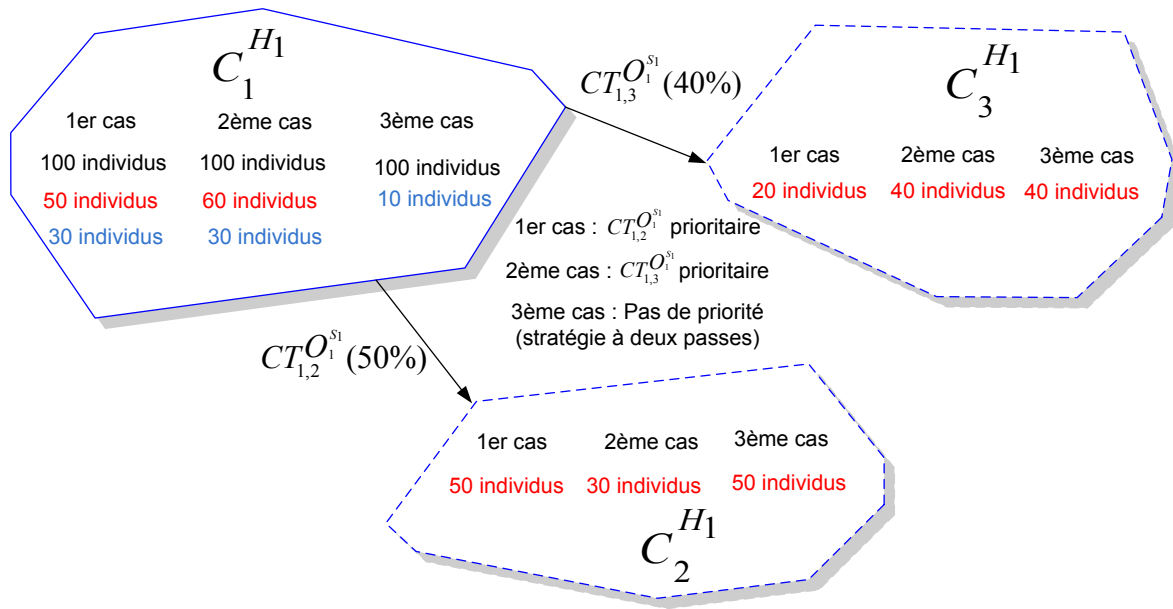


Figure 104: Exemples pour illustrer la stratégie à deux passes.

Par ailleurs, nous proposons une stratégie en deux passes pour régler les problèmes de conflits concernant la modification des valeurs des attributs, étant donné que ces dernières peuvent être partagées (c.-à-d. les processus et les transitions peuvent changer au même temps les mêmes attributs). La stratégie proposée consiste à calculer dans une première passe les nouvelles valeurs des attributs et les conserver en mémoire sans les appliquer réellement jusqu'à ce que tous les processus et toutes les transitions aient fini leur exécution. Dans une deuxième passe et à la fin du pas de simulation, le système applique les modifications aux attributs. Ainsi, l'ordre d'exécution des processus et des transitions n'a plus d'importance et tout se passe comme s'ils s'exécutaient en parallèle. Afin d'illustrer cette stratégie à deux passes, nous présentons dans la Figure 104 un exemple où il y a 3 cas de transfert d'individus entre cellules. Le premier cas consiste à donner une priorité à la transition de transfert entre la cellule C_1 et la cellule C_2 (c.-à-d. $CT_{1,2}^{O_1^{S_1}}$). Nous supposons qu'il y a initialement 100 individus dans cellule C_1 et que 50% des individus vont changer de la cellule C_1 vers la cellule C_2 alors que seulement 40% des individus vont changer de la cellule C_1 vers la cellule C_3 . De ce fait, 50 individus vont se transférer à la

cellule C_2 et 50 individus vont rester dans la cellule C_1 . Par la suite, 40% de ces 50 individus qui restent dans la cellule C_1 vont se transférer dans la cellule C_2 (c.-à-d. 20). Le deuxième cas consiste à donner une priorité à la transition de transfert entre la cellule C_1 et la cellule C_3 (c.-à-d. $CT_{1,3}^{O_1^{s_1}}$). Dans ce cas, nous obtenons un résultat différent du premier cas. 30 individus vont se transférer dans la cellule C_2 et 40 individus vont se transférer dans la cellule C_3 alors que 30 individus vont rester dans la cellule C_1 . Le dernier cas est celui où nous appliquons la stratégie en deux passes. Nous ne considérons donc pas de priorité dans ce cas. Le calcul des pourcentages se fait, mais ne sera effectif qu'à la fin du pas de simulation. Par conséquent, 50 individus vont se transférer dans la cellule C_2 et 40 individus vont se transférer dans la cellule C_3 alors que 10 individus vont rester dans la cellule C_1 .

3.3 Géosimulation de la maladie de Lyme

Nous présentons dans cette section l'essentiel de l'expérience que nous avons eue avec la simulation de la maladie de Lyme dans le cadre de notre travail avec toute l'équipe de notre laboratoire de recherche (c.-à-d. GRIC) [Moulin *et al.*, 2013a], [Moulin *et al.*, 2013b]. Nous commençons par la présentation d'un aperçu rapide de l'architecture technique de l'application *Zoonosis-MAGS*. Nous présentons par la suite la spécification du modèle de simulation ainsi que du scénario. Nous terminons en présentant les résultats de la simulation.

3.3.1 Architecture technique

La Figure 105 présente l'architecture technique du système *Zoonosis-MAGS* qui est développé en C^{++} en utilisant Qt³⁵. En effet, ce système est conçu en utilisant un ensemble de modules qui sont indépendants les uns des autres, mais tels que chacun peut ajouter de nouvelles fonctionnalités au système en spécifiant et/ou en exécutant une partie de la simulation. Certains modules sont des modules de base comme par exemple le moteur de

³⁵ Qt est une bibliothèque logicielle permettant de gérer entre autres des interfaces graphiques (<http://qt.digia.com/>).

simulation (c.-à-d. *Simulation Engine*). D'autres modules sont des modules d'extension (c.-à-d. *plugins*) qui peuvent se greffer au système. Ce principe permet au système d'être très flexible étant donné que le développeur peut ajouter de nouveaux modules afin d'adapter le système à une nouvelle zoonose (ce qui rend le système générique). Les flèches en pointillé de la Figure 105 représentent des dépendances inspirées de la notation *UML* (c.-à-d. *Unified Modeling Language* [Site Web: *OMG*]). Le sens de la flèche veut dire qu'un module dépend de l'autre. Nous ne présentons que les plus importantes dépendances pour ne pas alourdir la figure, surtout que l'architecture technique est très complexe et sa représentation nécessite des dizaines de pages (Figure 105).

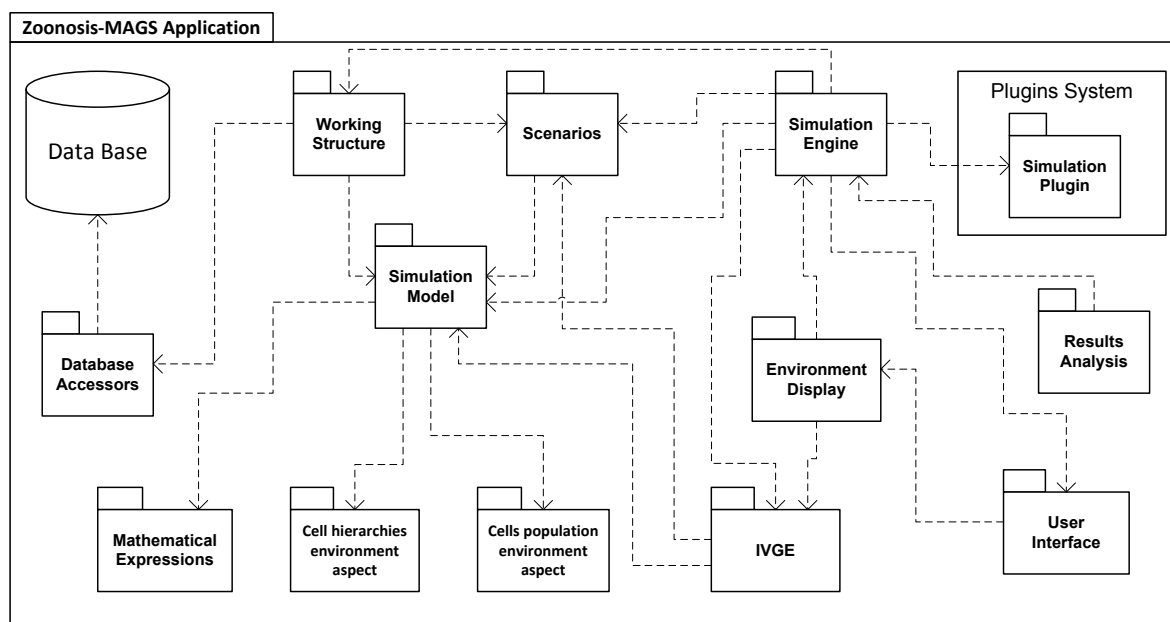


Figure 105: Les plus importants modules de l'architecture technique de *Zoonosis-MAGS*.

Par ailleurs, le module du modèle de simulation (c.-à-d. *Simulation Model* utilisé dans la formule du *ScenarioPattern* présenté auparavant), qui est un des plus importants modules de *Zoonosis-MAGS*, permet de spécifier la majorité des concepts que nous avons mentionnés dans le formalisme de *MASTIM*. Ainsi, ce module représente en quelque sorte une partie importante du squelette de la simulation. Il contient un système de spécification permettant à l'utilisateur de configurer les cellules, les espèces, les compartiments, les transitions, les interactions, etc. En effet, le modèle de simulation utilise trois principaux

modules d'extension: (1) le *plugin* des hiérarchies des cellules (c.-à-d. *Cell hierarchies environment aspect*) permettant de spécifier les cellules ainsi que leurs niveaux hiérarchiques qui seront utilisés dans l'*IVGE*, (2) le *plugin* des populations des cellules (c.-à-d. *Cells population environment aspect*) permettant de spécifier le modèle enrichi à base de compartiments faisant évoluer et interagir les espèces impliquées dans la zoonose et (3) le *plugin* des expressions mathématiques (c.-à-d. *Mathematical Expressions*) permettant de spécifier les équations du modèle mathématique.

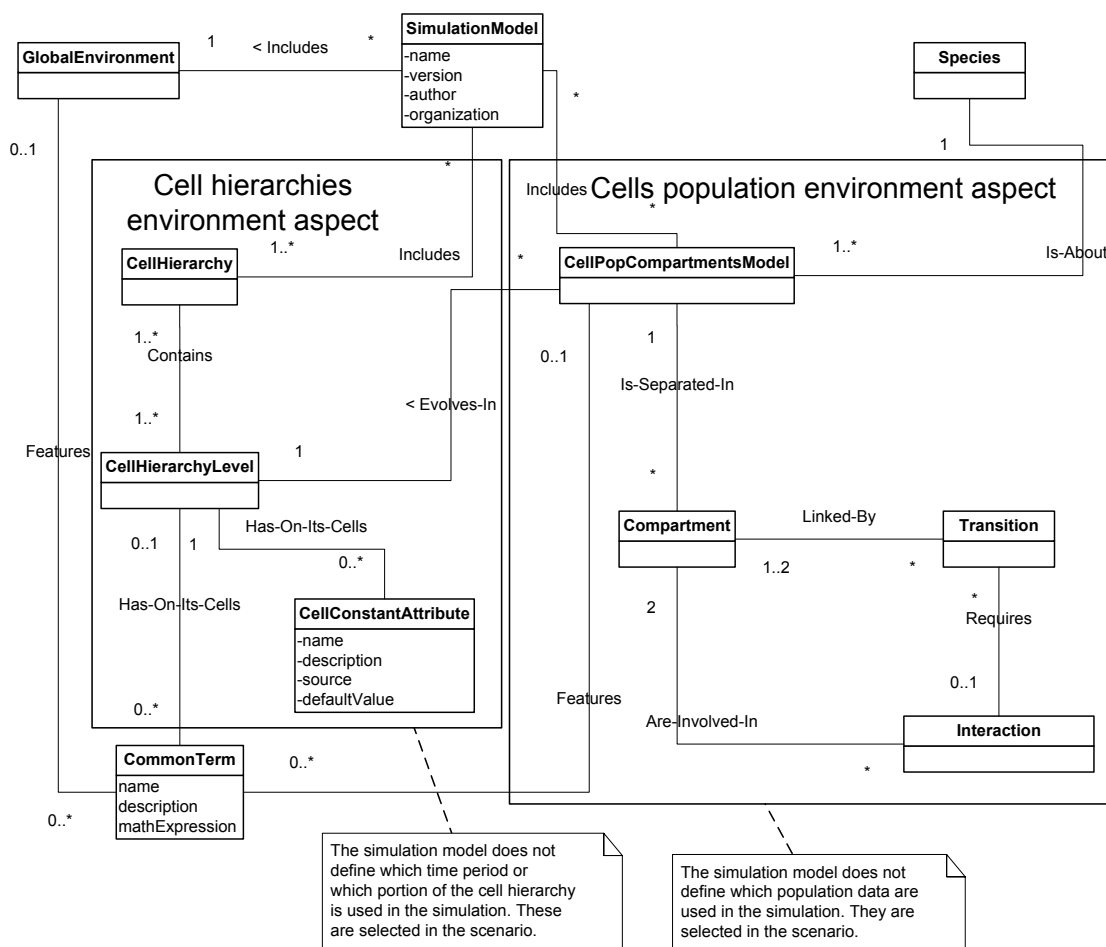


Figure 106: Les plugins du modèle de simulation.

Nous présentons d'ailleurs dans la Figure 106 les principales classes utilisées par deux de ces trois *plugins*. À gauche de cette figure, nous trouvons les principales classes spécifiant les niveaux hiérarchiques des cellules, alors qu'à droite nous trouvons les classes spécifiant les compartiments, les transitions, les interactions, etc. De plus, la Figure 106 explique la

différence entre le modèle de simulation et le scénario. Ce dernier est une instantiation du modèle de simulation en spécifiant par exemple les données vectorielles ainsi que les données de populations qui seront utilisées lors de la simulation.

3.3.2 Spécification du modèle de simulation et du scénario

Nous présentons dans cette section les principales étapes que l'utilisateur de *Zoonosis-MAGS* doit faire afin de bien spécifier le modèle de simulation ainsi que le scénario qu'il veut appliquer lors de la simulation. En effet, la première fenêtre de *Zoonosis-MAGS* (Figure 107a) permet à l'utilisateur de saisir les paramètres pour se connecter à la base de données de l'application (c.-à-d. *PostgreSQL* contenant *PostGIS*). La deuxième fenêtre permet à l'utilisateur de sélectionner les *plugins* qu'il veut utiliser pour spécifier le modèle de simulation. Dans notre cas, nous avons sélectionné les 3 *plugins* mentionnés auparavant. Il s'agit du *plugin* des hiérarchies des cellules, du *plugin* des populations des cellules et du *plugin* des expressions mathématiques (Figure 107b).

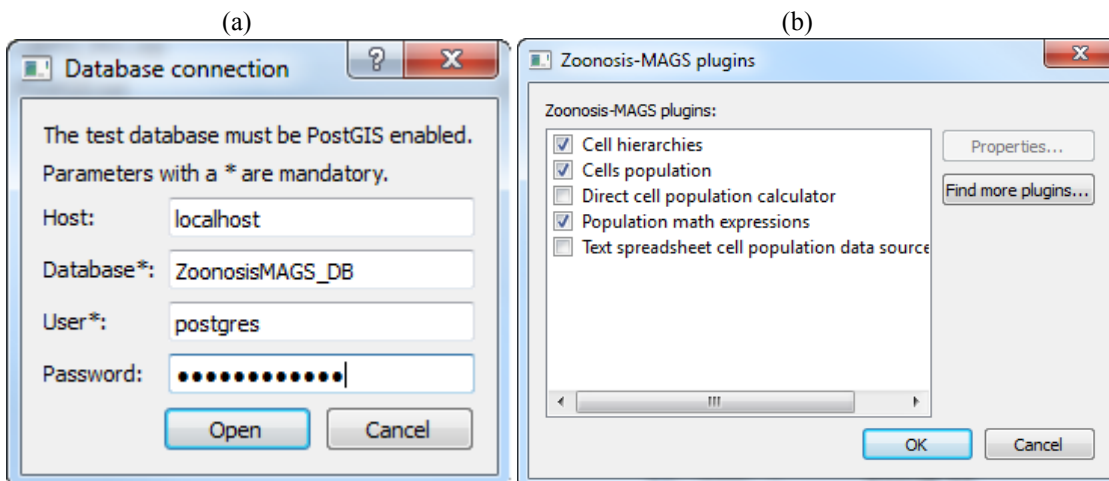


Figure 107: (a) Connexion à la base de données, (b) Sélection des plugings.

Les prochaines fenêtres permettent à l'utilisateur de spécifier les niveaux hiérarchiques des cellules qu'il faut utiliser en important dans la base *PostGIS* les données vectorielles représentant dans notre cas l'*IVGE* qui est prétraité, rappelons le, par une autre application (voir Chapitre VI). Pour des fins de test, nous avons sélectionné un échantillon de la zone

31H représentant une bonne partie de la Montérégie (une région au sud et à l'est de Montréal). L'utilisateur peut également sélectionner les données des populations en utilisant l'onglet « *Cells population* » (Figure 108).

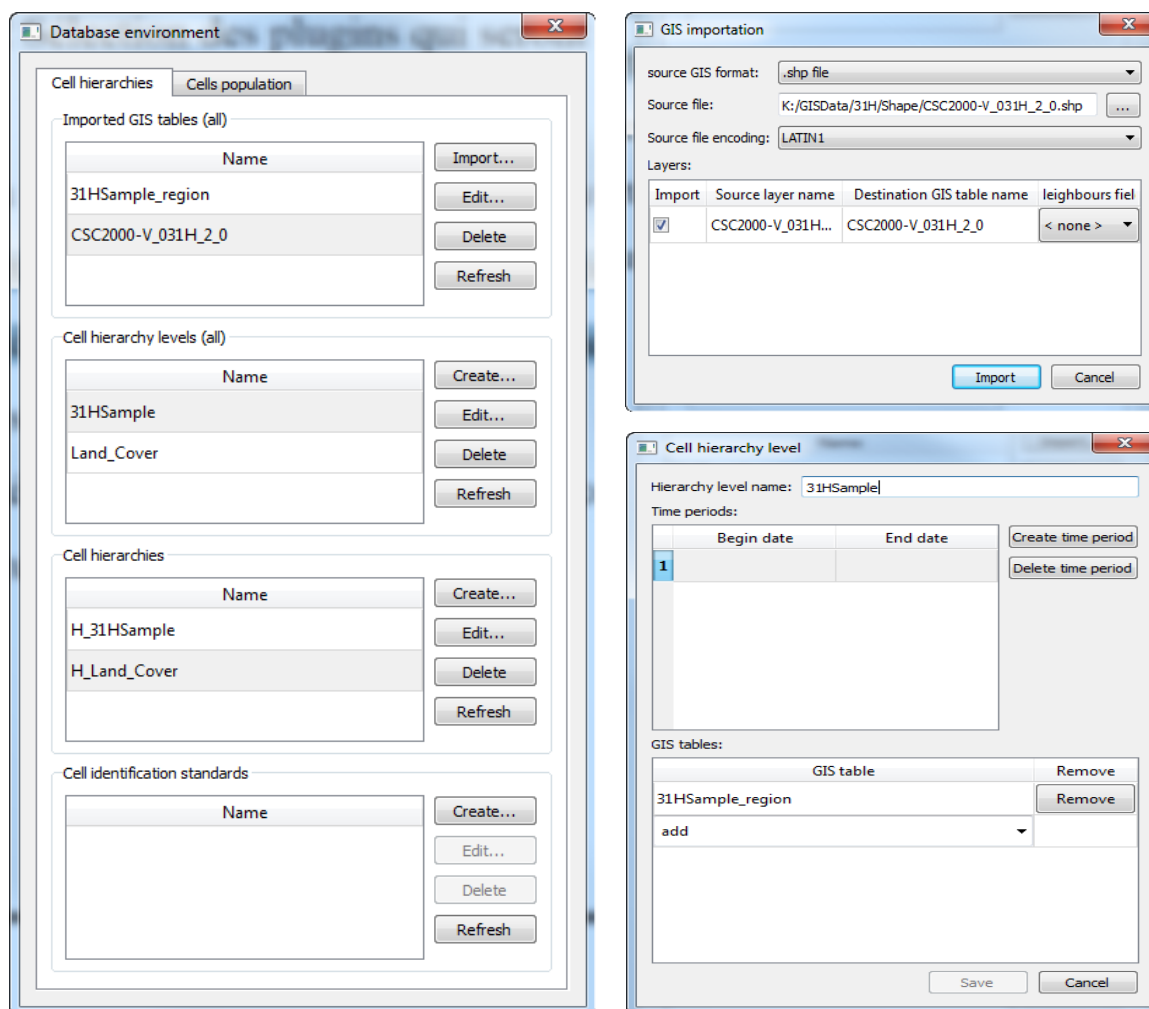


Figure 108: Spécification des hiérarchies des cellules et des données de populations.

Le système permet par la suite à l'utilisateur de visualiser l'IVGE dans une fenêtre graphique. L'utilisateur peut également sélectionner une cellule (c.-à-d. un polygone) et afficher les attributs associés à cette cellule. La Figure 109 montre justement les propriétés d'une cellule sélectionnée en affichant le degré de favorabilité de la cellule, mais également les informations par rapport aux voisins de cette cellule (identificateur et orientation quantitative et qualitative de chaque voisin) ainsi que la trace des fusions effectuées (c.-à-d. l'attribut *PropLandCover* qui contient la liste des types de couverture du sol avec les

proportions des superficies des polygones qui ont participé dans la fusion). La Figure 109 visualise les cellules avec un gradient de couleur. Les cellules qui ont un degré de favorabilité plus faibles sont plus claires et les cellules qui ont un degré de favorabilité plus élevé sont plus foncées (Figure 109).

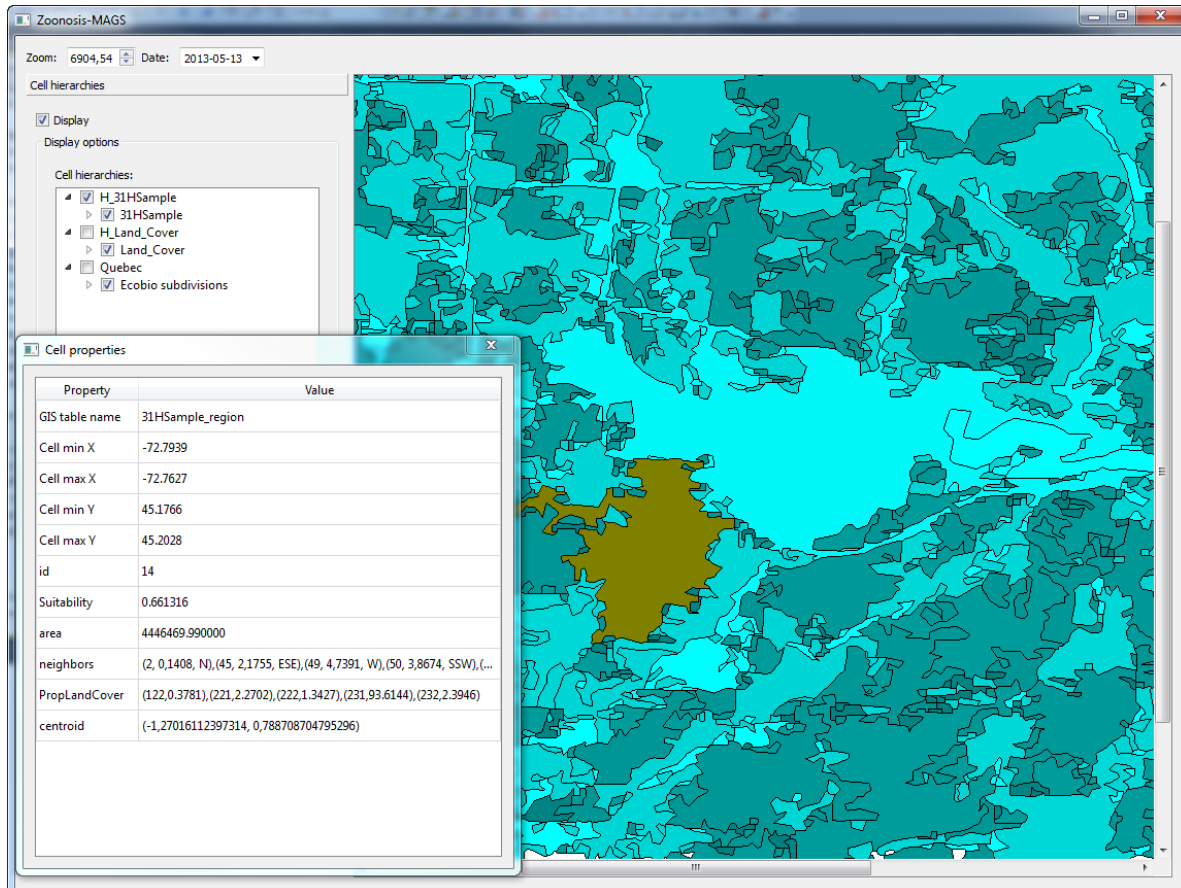


Figure 109: Utilisation de l'IVGE pour la simulation de la maladie de Lyme.

En outre, la Figure 110 montre que le système permet à l'utilisateur d'éditer le modèle de simulation ou de charger un modèle existant. C'est une étape importante avant de pouvoir éditer ou charger un scénario afin de faire partir par la suite la simulation (Figure 110). En cliquant sur le bouton « *Edit simulation model* », l'utilisateur peut spécifier tous les détails du modèle de simulation en ajoutant les espèces impliquées dans la propagation de la maladie de Lyme. Dans notre cas, il s'agit des tiques, des chevreuils, des oiseaux et des

rongeurs. La Figure 111 montre la liste des compartiments et des transitions qui ont été ajoutés pour les tiques.

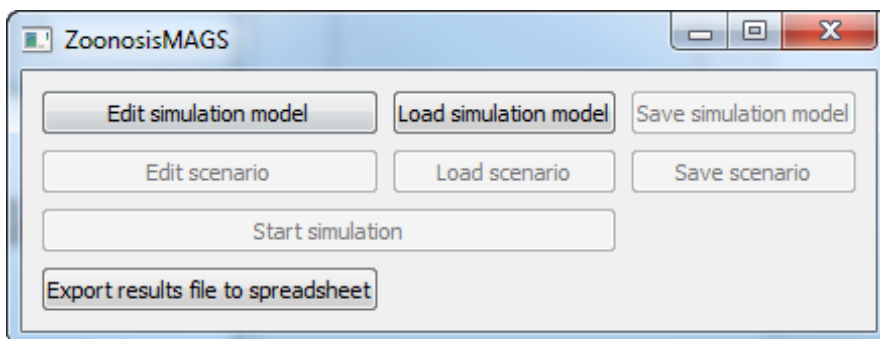


Figure 110: Fenêtre de commandes de la simulation.

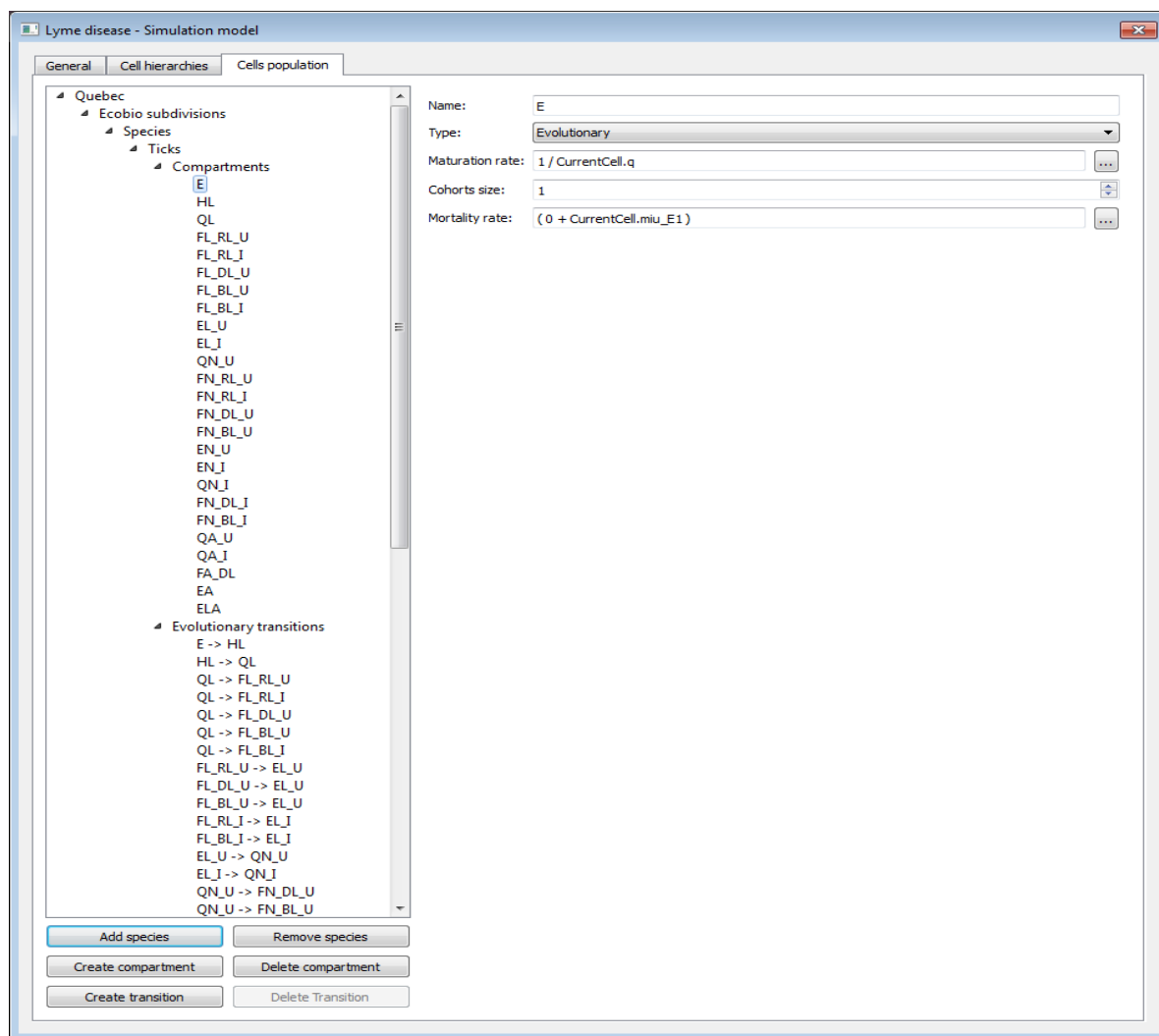


Figure 111: Spécification du modèle à base de compartiments pour la maladie de Lyme.

Le compartiment E (c.-à-d. oeufs) est sélectionné afin de voir à droite de la Figure 111 les propriétés de ce compartiment. L'utilisateur peut par exemple entrer ou modifier le taux de maturation ou le taux de mortalité des individus appartenant à ce compartiment. Ces taux sont spécifiés en utilisant des équations mathématiques qui peuvent être simples, mais dans la majorité des cas elles sont complexes, surtout pour les équations qui sont utilisées pour les transitions. La Figure 112 montre d'ailleurs comment l'utilisateur peut spécifier une équation mathématique. Ensuite, il peut sélectionner un patron parmi une liste de patrons d'expressions mathématiques, et de saisir les paramètres de ces expressions (Figure 112).

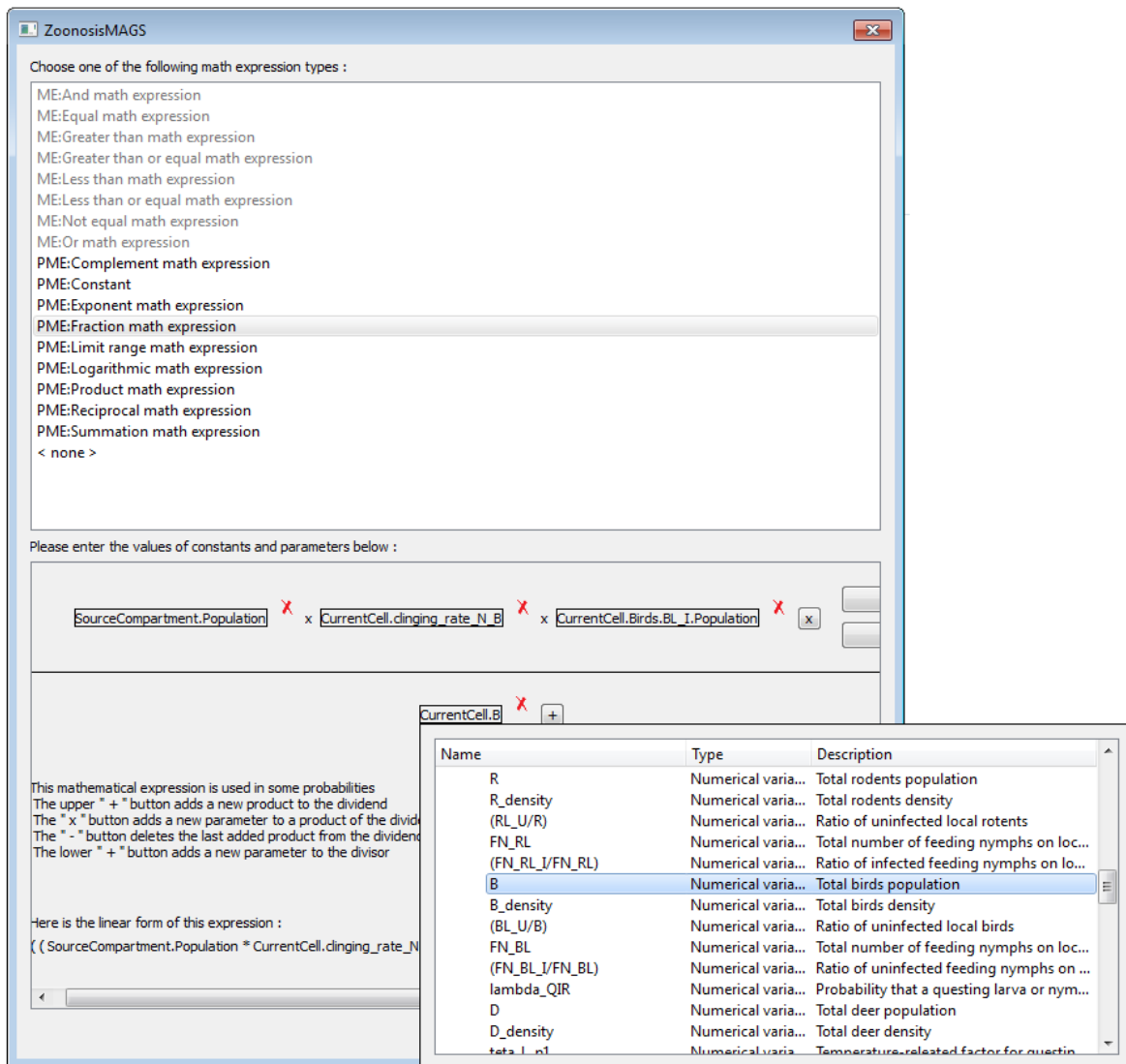


Figure 112: Spécification des équations mathématiques.

D'un autre côté, l'utilisateur peut choisir de spécifier le modèle de simulation en utilisant une interface graphique plus conviviale et qui a été développé en Visio, mais totalement intégré à *Zoonosis-MGAS* [Sedrati, 2013]. En effet, la Figure 113 montre un exemple de spécification d'une transition entre le compartiment *ELA* (c.-à-d. tiques adultes en train de pondre des oeufs) et le compartiment *E* (c.-à-d. oeufs).

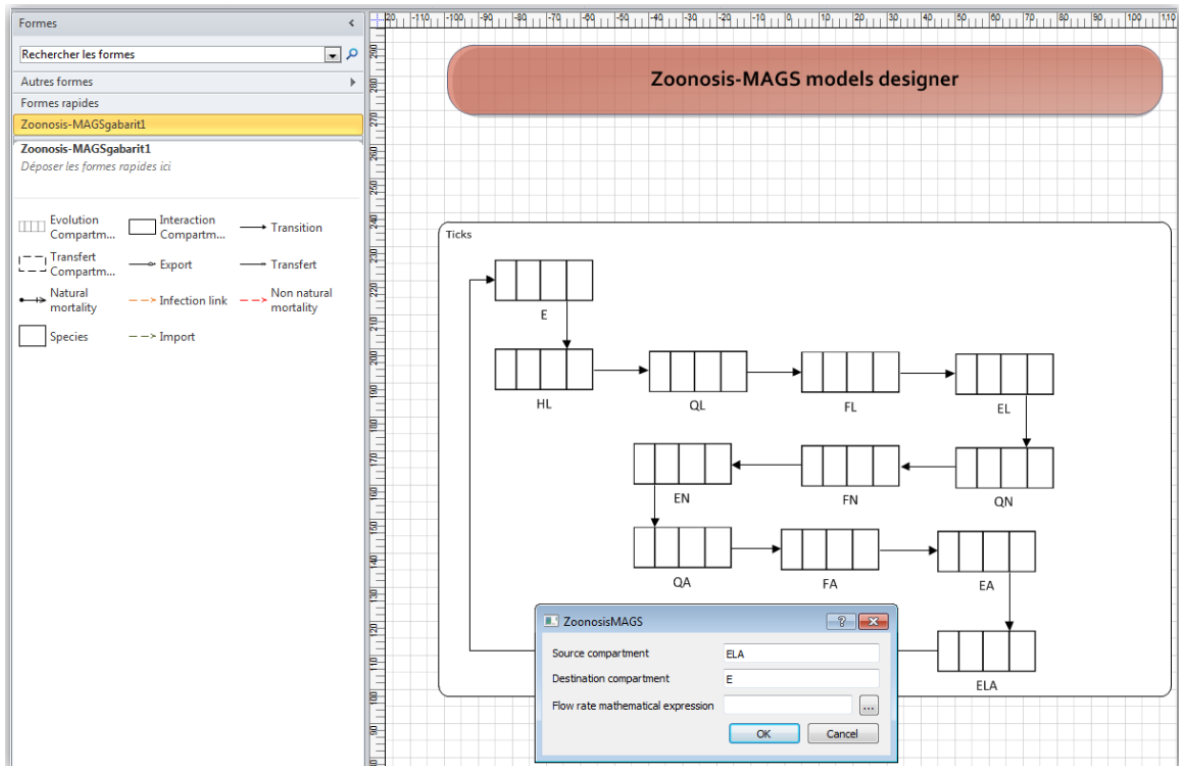


Figure 113: Spécification du modèle de simulation en utilisant Visio.

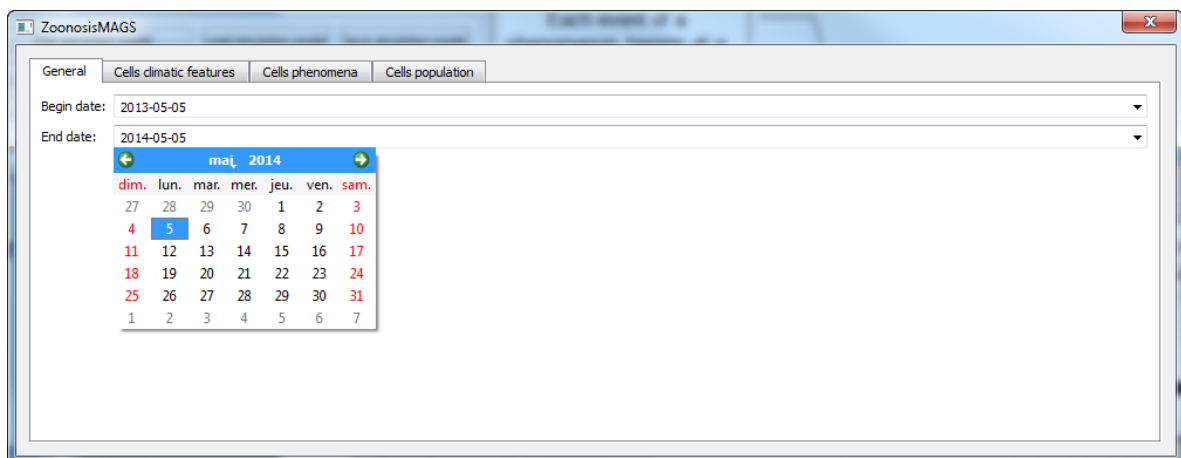


Figure 114: Spécification d'un scénario.

Après la spécification du modèle de simulation, l'utilisateur peut maintenant spécifier un ou plusieurs scénarios en cliquant sur le bouton « *Edit scenario* » de la Figure 110. De ce fait, la Figure 114 s'ouvre permettant à l'utilisateur de choisir par exemple la période de la simulation en sélectionnant la date de début et la date de fin de la simulation (Figure 114).

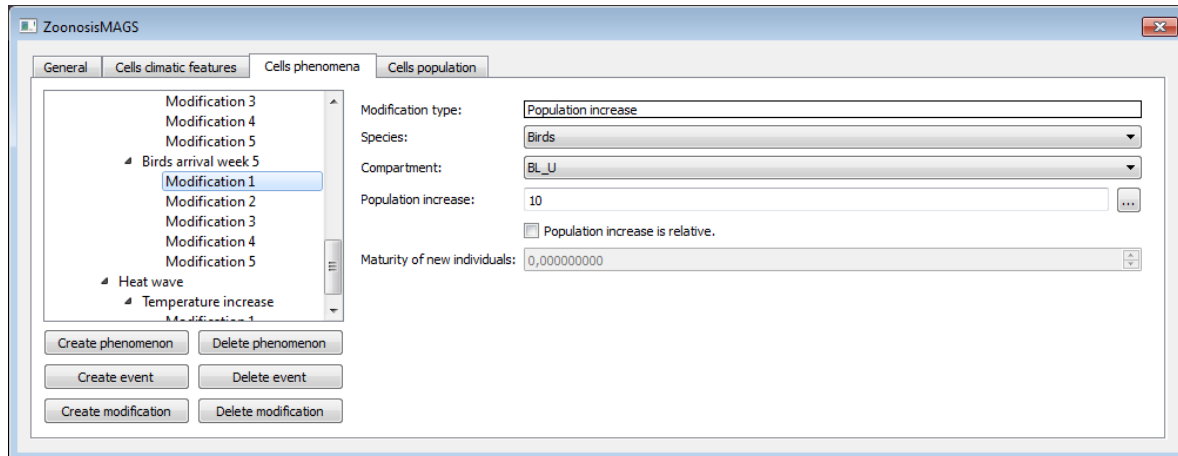


Figure 115: Spécification des événements dans un scénario.

L'utilisateur peut également spécifier des événements dans le scénario comme par exemple l'arrivée des oiseaux migrateurs ou encore une vague de chaleur afin de produire des simulations différentes (Figure 115).

3.3.3 Résultats de la simulation

Comme mentionné auparavant, nous nous sommes (c.-à-d. toute l'équipe du GRIC) inspirés du modèle d'Ogden [Ogden *et al.*, 2005] afin de simuler la propagation de la maladie de Lyme (voir Section 4.2 du Chapitre V). Cependant et avant de pouvoir le faire dans *Zoonosis-MAGS*, nous avons dû passer par un système de prototypage en utilisant *MATLAB*³⁶ [Navarro, 2013]. Nous avons d'ailleurs fait la même chose pour la simulation de la propagation du VNO étant donné que le modèle mathématique a été conçu dans *MATLAB*, mais converti par la suite en C⁺⁺. Cette façon de faire permet de tester rapidement plusieurs scénarios dans *MATLAB*, mais aussi de comparer les résultats afin de

³⁶ *MATLAB* est un environnement interactif pour le calcul numérique, la visualisation et la programmation. <http://www.mathworks.fr/products/matlab/index.html>

les valider avec éventuellement les experts du domaine. Actuellement, le modèle est fonctionnel dans les deux plateformes et les résultats concordent. *Zoonosis-MAGS* permet de montrer dans des fichiers les résultats de la simulation, mais il n'est pas encore prêt pour montrer ces résultats dans l'*IVGE* comme c'est le cas pour *WVN-MAGS*. En cliquant sur le bouton « *Start Simulation* » de la Figure 110, l'utilisateur peut configurer certains paramètres de la simulation en sélectionnant par exemple les options de sauvegarde pour chaque espèce impliquée dans la propagation de la maladie de Lyme (Figure 116). Par la suite, *Zoonosis-MGAS* affiche une fenêtre montrant le nombre des itérations (c.-à-d. les pas de simulation) ainsi que le temps d'exécution de la simulation (Figure 117).

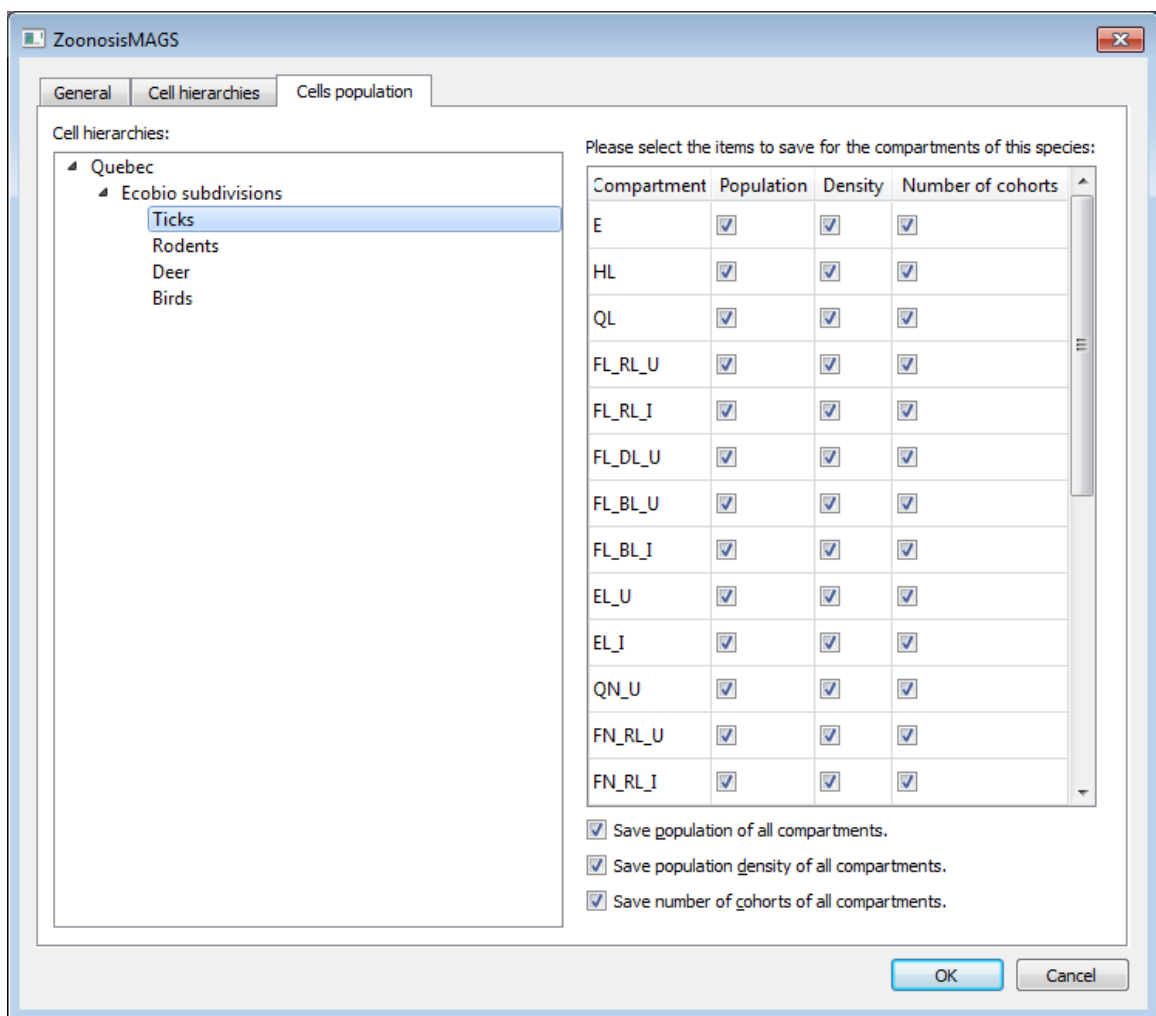


Figure 116: Les options de sauvegarde avant de démarrer la simulation.

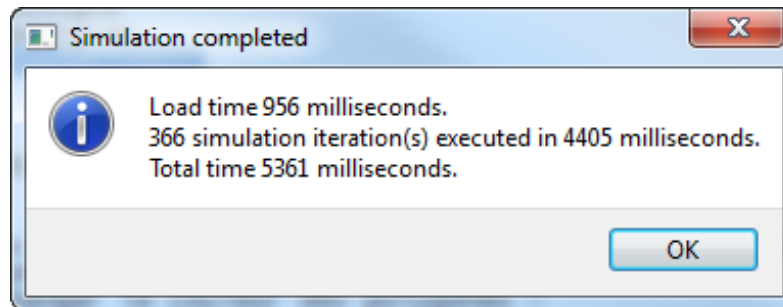


Figure 117: Fin de la simulation en affichant le temps d'exécution.

	A	B	C	D	E	F	AE	AF	AG	AH	AI	AJ
4	Iter	Date	Temp	Temp	Temp	Diff	EL_U			EL_I		
5			ML	C++	abs. dif		ML	C++	abs. dif	ML	C++	abs. dif
6			Temp				EL_U			EL_I		
7				temp				EL_U			EL_I	
8	0				0	6,1964E-12		0	1,0809E-12		0	1,0408E-15
9	1	2001-01-01	0	0	0	0	0	0	0	0	0	0
10	2	2001-01-02	0	0	0	0	0	0	0	0	0	0
11	3	2001-01-03	0	0	0	0	0	0	0	0	0	0
325	317	2001-11-13	4,2	4,2	0	1,934E-14	6,65856138	6,65856138	1,6875E-14	0,03364612	0,03364612	0
326	318	2001-11-14	4,2	4,2	0	1,3128E-14	6,66664744	6,66664744	1,0658E-14	0,03368577	0,03368577	0
327	319	2001-11-15	4,2	4,2	0	2,4594E-15	6,67440162	6,67440162	0	0,03372377	0,03372377	0
328	320	2001-11-16	4,2	4,2	0	1,3997E-14	6,68182827	6,68182827	1,1546E-14	0,03376015	0,03376015	0
329	321	2001-11-17	4,2	4,2	0	2,4473E-15	6,68893169	6,68893169	0	0,03379493	0,03379493	0
330	322	2001-11-18	4,2	4,2	0	2,4408E-15	6,69571616	6,69571616	0	0,03382812	0,03382812	0
331	323	2001-11-19	4,2	4,2	0	3,6633E-15	6,70218587	6,70218587	0	0,03385975	0,03385975	0
332	324	2001-11-20	4,2	4,2	0	1,4865E-14	6,70834499	6,70834499	1,2434E-14	0,03388984	0,03388984	0
333	325	2001-11-21	4,2	4,2	0	1,3085E-14	6,71419763	6,71419763	1,0658E-14	0,03391841	0,03391841	0
334	326	2001-11-22	4,2	4,2	0	2,4238E-15	6,71974787	6,71974787	0	0,03394548	0,03394548	0
335	327	2001-11-23	4,2	4,2	0	1,5754E-14	6,72499972	6,72499972	1,3323E-14	0,03397106	0,03397106	0
336	328	2001-11-24	4,2	4,2	0	2,413E-15	6,72995716	6,72995716	0	0,03399519	0,03399519	0
337	329	2001-11-25	4,2	4,2	0	2,4269E-15	6,73462412	6,73462412	0	0,03401788	0,03401788	0
338	330	2001-11-26	4,2	4,2	0	2,403E-15	6,73900448	6,73900448	0	0,03403914	0,03403914	0
339	331	2001-11-27	4,2	4,2	0	2,4026E-15	6,74310209	6,74310209	0	0,034059	0,034059	0
340	332	2001-11-28	4,2	4,2	0	2,4026E-15	6,74692074	6,74692074	0	0,03407747	0,03407747	0
341	333	2001-11-29	4,2	4,2	0	1,5716E-14	6,75046418	6,75046418	1,3323E-14	0,03409458	0,03409458	0
342	334	2001-11-30	4,2	4,2	0	1,5705E-14	6,75373613	6,75373613	1,3323E-14	0,03411034	0,03411034	0
343	335	2001-12-01	0	0	0	1,3042E-14	6,75674025	6,75674025	1,0658E-14	0,03412477	0,03412477	0
344	336	2001-12-02	0	0	0	2,3714E-15	6,75948017	6,75948017	0	0,03413789	0,03413789	0
345	337	2001-12-03	0	0	0	2,3705E-15	6,76194159	6,76194159	0	0,03414971	0,03414971	0
346	338	2001-12-04	0	0	0	1,3916E-14	6,79367329	6,79367329	1,1546E-14	0,03430884	0,03430884	0
347	339	2001-12-05	0	0	0	1,4798E-14	6,82434054	6,82434054	1,2434E-14	0,03446292	0,03446292	0
348	340	2001-12-06	0	0	0	1,7452E-14	6,85400526	6,85400526	1,5099E-14	0,03461204	0,03461204	0
349	341	2001-12-07	0	0	0	1,789E-14	6,88270037	6,88270037	1,4211E-14	0,03475632	0,03475632	0
350	342	2001-12-08	0	0	0	3,6772E-15	6,91044878	6,91044878	0	0,03489584	0,03489584	0
351	343	2001-12-09	0	0	0	2,244E-14	6,93726827	6,93726827	1,8652E-14	0,03503069	0,03503069	0
352	344	2001-12-10	0	0	0	2,3358E-15	6,96317629	6,96317629	0	0,03516097	0,03516097	0
353	345	2001-12-11	0	0	0	1,8875E-14	6,98819005	6,98819005	1,5099E-14	0,03528676	0,03528676	0
354	346	2001-12-12	0	0	0	1,4782E-14	7,01232641	7,01232641	1,2434E-14	0,03540814	0,03540814	0
355	347	2001-12-13	0	0	0	1,9203E-14	7,03560196	7,03560196	1,6875E-14	0,0355252	0,0355252	0

Figure 118: Comparaison entre les résultats de *MATLAB* et ceux de *Zoonosis-MAGS*.

En fait, les résultats de la simulation sont sauvegardés dans un fichier binaire, mais qui peut être converti par exemple en un fichier Excel. La Figure 118 montre d'ailleurs les résultats

obtenus dans un fichier Excel. Ce fichier contient l'évolution des individus de tous les compartiments. Les colonnes E, AG et AJ montrent la différence entre les données de *MATLAB* et les données de *Zoonosis-MAGS*. Cette différence est soit nulle soit très minime. La variation minime est due essentiellement à la précision du calcul qui peut changer en fonction du microprocesseur utilisé. Par ailleurs et étant donné que le modèle mathématique à base de compartiments est cloné dans chaque cellule de l'*IVGE*, les résultats obtenus sont évidemment différents d'une cellule à une autre. Cette différence est attribuée entre autres à la densité des espèces impliquées dans la propagation de la maladie puisque la superficie n'est pas la même pour toutes les cellules (Figure 118).

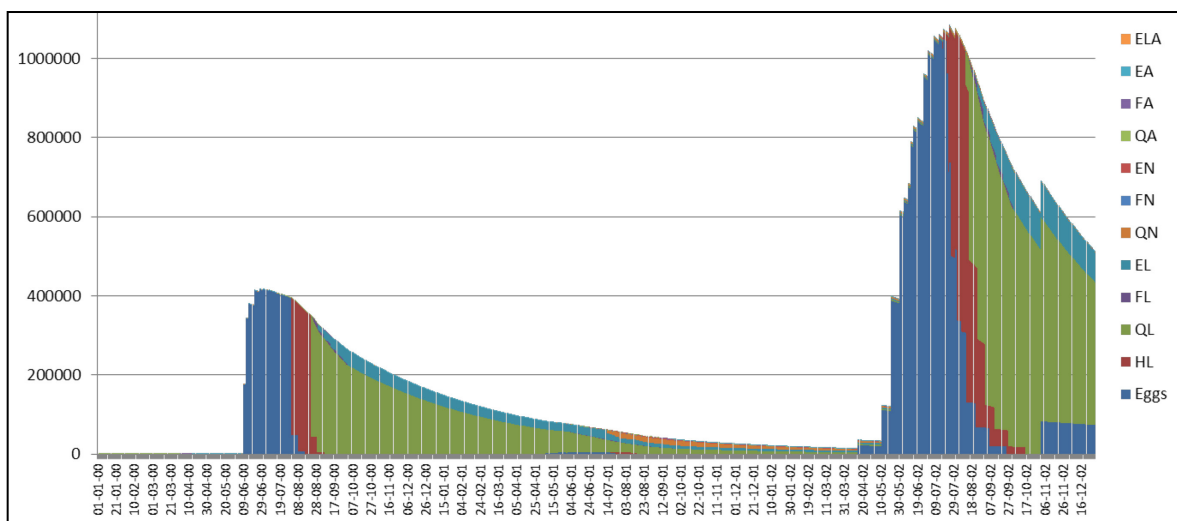


Figure 119: Évolution des différents compartiments des tiques sur deux ans.

En outre et à partir des fichiers contenant les résultats obtenus, il est bien sûr possible de produire des graphiques montrant entre autres l'évolution des populations. D'ailleurs et puisque le cycle biologique des tiques dure plus longtemps que celui des moustiques, la Figure 119 montre l'évolution des différents compartiments des tiques sur deux ans. Par exemple, les œufs (*Eggs*: courbe en bleue) qui sont pondus au mois de juin de la première année vont donner des adultes (*QA*: courbe en vert) au mois d'août de la deuxième année. (Figure 119). Par ailleurs, un des scénarios que nous avons testés est la possibilité d'établissement des colonies de tiques au Canada en faisant migrer des oiseaux à partir des États-Unis. Les tiques juvéniles qui sont accrochées à ces oiseaux vont se laisser tomber

pour éventuellement s'établir dans une région du Nord dont les conditions de l'habitat sont favorables (voir Figure F2 de l'Annexe F). Dans ce scénario, nous avons considéré une cellule au Sud (États-Unis) et une autre cellule au Nord (Long Point, Ontario : ce choix est basé sur les travaux d'Ogden [Ogden *et al.*, 2007]). En effet, la Figure 120 montre les résultats de ce scénario représentant l'évolution de la population totale des tiques sur 40 ans dans les deux cellules en utilisant les données des températures moyennes qui sont ajustées en fonction des changements climatiques (augmentation de la température de 1970 à 2010 [Hansen *et al.*, 2006]). Il est à noter que nous avons supposé que les températures dans la cellule au Sud sont 5% plus élevées que celles dans la cellule au Nord.

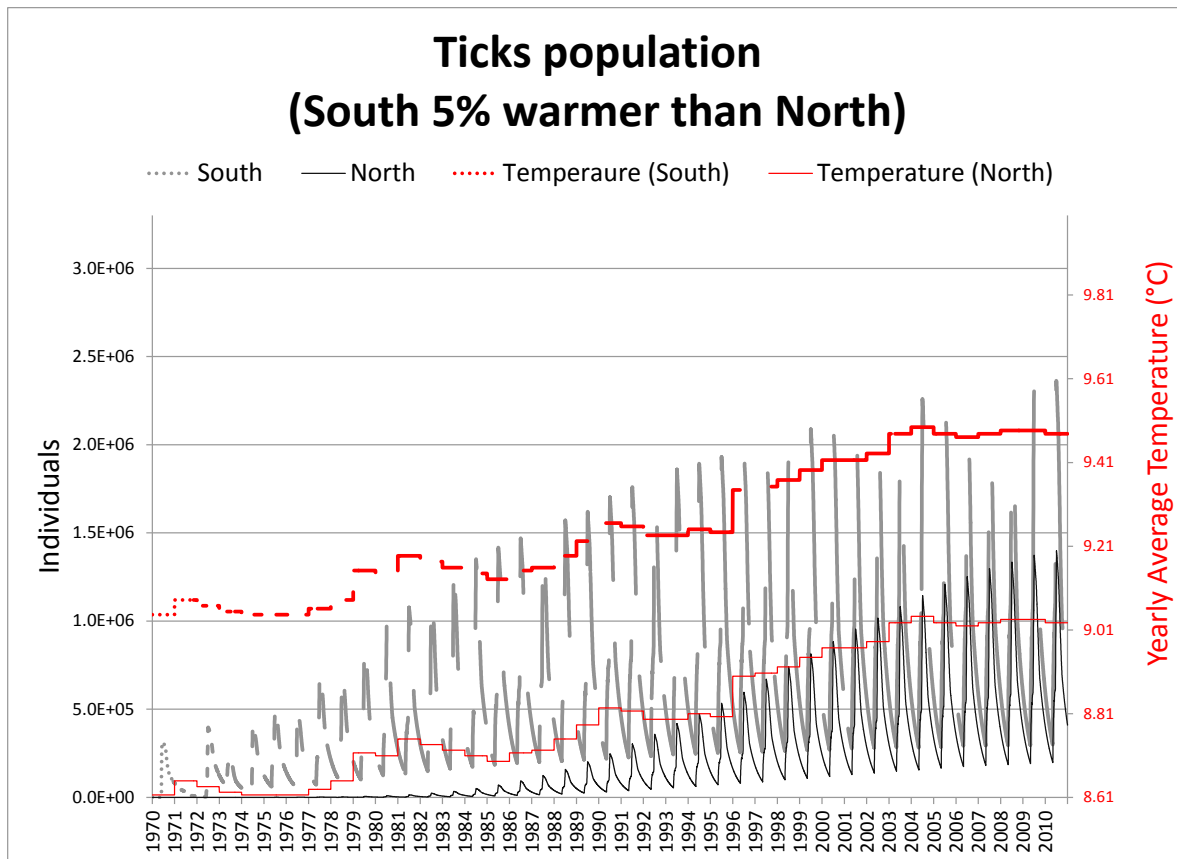


Figure 120: Évolution de la population totale des tiques sur 40 ans en utilisant une cellule au Sud dont la température est supérieure de 5% qu'une cellule au Nord.

D'autre part, nous avons introduit dans ce scénario un patron de mobilité (voir Section 4.1.1 du Chapitre VI) simple des oiseaux. Ces derniers vont migrer du Sud vers le Nord au printemps et revenir en automne chaque année. Nous avons supposé que la moitié des oiseaux va rester dans la cellule du Sud alors que l'autre moitié va migrer vers la cellule du Nord. Nous avons également supposé que les tiques sont présentes au début de la simulation (c.-à-d. en 1970) dans la cellule du Sud, mais qu'il n'y a pas de tiques dans la cellule du Nord. La Figure 120 montre que l'évolution de la population de tiques dans la cellule du Sud augmente de façon significative pendant les 40 ans de simulation. De plus et après 15 ans, nous constatons que les populations de tiques commencent à s'implanter dans la cellule du Nord suite à l'apport des tiques juvéniles (c.-à-d. larves et nymphes) par les oiseaux migrateurs (p.ex. Merle d'Amérique, Bruant à gorge blanche, Paruline masquée, Carouge à épauettes, Moqueur Chat) qui peuvent transporter plus de 10 tiques juvéniles par oiseau [Ogden *et al.* 2008b]. Ces populations de la cellule du Nord augmentent par la suite de la même façon que celles de la cellule du Sud pour finalement se stabiliser à un certain niveau. Par conséquent, ce scénario montre qu'une population de tiques peut s'établir lentement dans une nouvelle zone. Cependant, les responsables de la santé publique peuvent ne pas se rendre compte de cet établissement étant donné qu'il se fait sur plusieurs années et que la surveillance sur le terrain utilise un système d'échantillonnage qui peut considérer que cet établissement est non significatif. Mais, après une longue période d'établissement, la population de tiques peut littéralement exploser. Ceci aura pour effet des réels risques de propagation de la maladie. Ce scénario montre l'importance de notre outil d'aide à la décision puisqu'il permet aux décideurs de comprendre certains phénomènes qui ne sont pas nécessairement visibles dans la nature à cause de leur progression très lente.

Nous terminons cette section en présentant dans la Figure 121 le résultat des tests que nous avons effectués afin d'évaluer le temps d'exécution (en milliers de millisecondes : 300 mille ms représentent 5 minutes) de la simulation en fonction du nombre de cellules se trouvant dans l'*IVGE*. La courbe bleue présente l'évolution du temps de chargement des données à partir de la base de données de *Zoonosis-MAGS*. La courbe rouge présente

l'évolution du temps de simulation alors que la courbe verte présente l'évolution du temps total (p.ex. pour 20000 cellules, le temps total est de 24 min). Ce temps augmente linéairement et nous pensons qu'il est assez raisonnable étant donné le nombre important de calculs que le système fait pour chacune des cellules. Heureusement d'ailleurs que nous avons pu réduire le nombre des cellules à traiter en utilisant les processus de fusions successives permettant de créer un *IVGE* permettant d'améliorer l'efficacité de la simulation. Sans cette approche de génération de l'*IVGE*, il serait quasiment impossible de produire des simulations dans un temps raisonnable même avec un ordinateur très puissant.

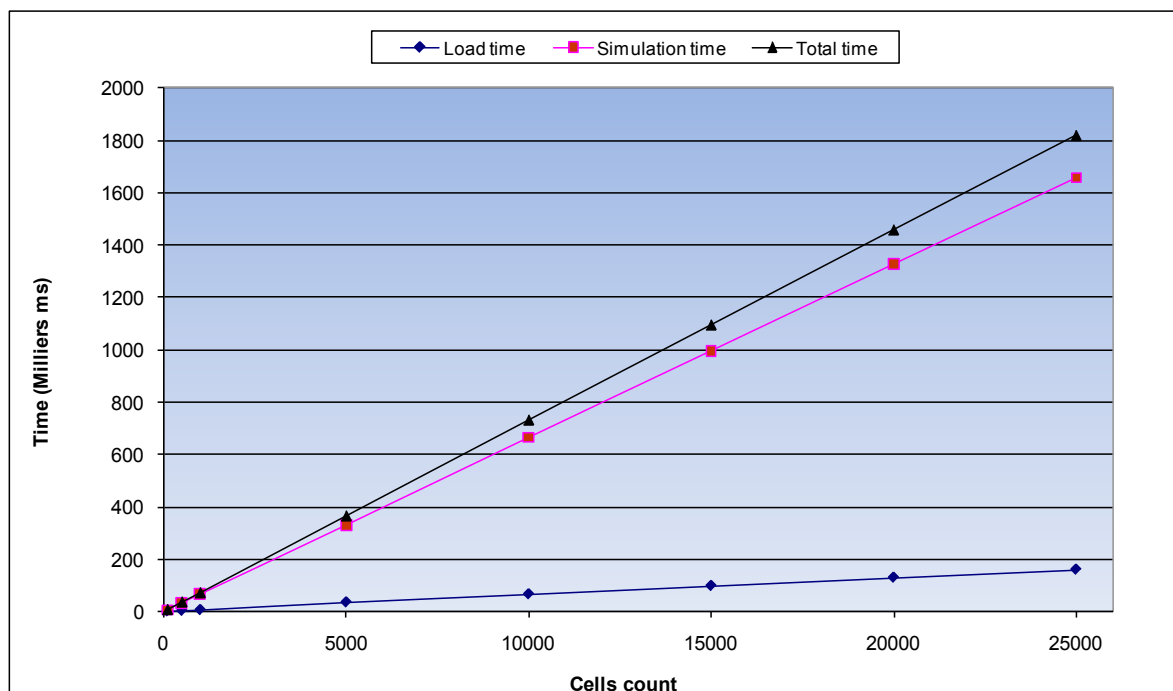


Figure 121: Augmentation du temps d'exécution de la simulation en fonction du nombre de cellules dans l'*IVGE*.

3.4 Discussion

L'outil de géosimulation *Zoonosis-MAGS* est conçu afin de simuler plusieurs zoonoses. En fait, nous le qualifions de système générique (ce n'est pas le cas du système *WNV-MAGS*)

étant donné qu'il utilise par exemple des modules d'extension (c.-à-d. *plug-ins* ou plugiciels) qui le rendent assez flexible. De plus, l'utilisation d'un système de spécification du modèle de simulation permet à l'utilisateur d'ajouter, en fonction de la zoonose qu'il veut étudier, les composantes fondamentales de la simulation qui sont par exemple les espèces impliquées dans la propagation de cette zoonose, les compartiments, les transitions, les interactions, etc. L'utilisateur peut également utiliser un système de spécification des équations mathématiques qui se basent sur des patrons d'expressions assez génériques pour les zoonoses. L'utilisateur peut donc concevoir ou modifier un modèle mathématique, faisant évoluer et interagir les espèces étudiées beaucoup plus facilement par exemple qu'avec un outil comme *Stella* [Site Web: *Stella*].

Par ailleurs, *Zoonosis-MAGS* permet d'utiliser les données prétraitées de l'*IVGE* (données vectorielles et données des populations attachées aux cellules) afin de créer selon les besoins de l'utilisateur différents niveaux hiérarchiques comme cela était spécifié dans le formalisme de *MASTIM*. Par exemple, l'utilisateur peut ajouter un niveau utilisant une subdivision selon les habitats favorables d'une espèce ou de plusieurs espèces. Il peut également ajouter un niveau supérieur utilisant une subdivision administrative comme les municipalités ou les secteurs de recensement. De ce fait, la simulation pourra se faire par exemple dans le niveau inférieur (c.-à-d. le modèle est cloné dans les cellules fusionnées en se basant sur le critère de favorabilité selon la couverture du sol). Les résultats pourront par la suite être agrégés au niveau supérieur (c.-à-d. subdivision administrative). De ce fait, les décideurs de la santé publique auront des résultats qu'ils peuvent probablement interpréter plus facilement, étant donné que ces décideurs sont plus habitués de travailler avec des cellules représentant des municipalités ou des secteurs de recensement.

En outre, le simulateur *Zoonosis-MAGS* est fonctionnel étant donné qu'il permet de produire des résultats dans des fichiers binaires selon différents scénarios. Cependant, il lui manque pour l'instant certaines fonctionnalités comme l'affichage des résultats dans l'environnement de la simulation. En fait et étant donné que le projet est limité par le temps, notre équipe a choisi de prioriser l'obtention de résultats en utilisant entre autres un

système de prototypage. Nous pensons que le système peut être amélioré rapidement si nous obtenons une extension du projet.

4 Conclusion

Nous avons présenté dans ce chapitre deux outils d'aide à la décision : *WNV-MGAS* et *Zoonosis-MAGS*. Le premier permet de simuler la propagation du VNO alors que le deuxième permet de simuler la propagation d'une zoonose d'une façon générique. Le développement de *WNV-MGAS* a été une étape primordiale pour la suite de nos recherches, étant donné qu'elle nous a permis d'évaluer entre autres la faisabilité du développement d'un outil d'aide à la décision dans le domaine de la santé publique. De plus, le développement de *Zoonosis-MAGS* nous a permis premièrement de nous rendre compte de la complexité du phénomène que nous voulons simuler (la maladie de Lyme est beaucoup plus complexe que le VNO puisqu'elle fait intervenir plusieurs espèces) et deuxièmement d'appliquer les concepts théoriques que nous avons conçus comme le modèle *MASTIM* ou la génération de l'*IVGE*. D'un autre côté, la simulation de deux zoonoses différentes nous a permis non seulement de les comparer, mais aussi de généraliser certains concepts fondamentaux comme par exemple les interactions qui sont instantanées pour le VNO, mais plus ou moins durables pour la maladie de Lyme.

Chapitre VIII : Conclusion générale

Nous présentons dans ce chapitre la conclusion générale de nos travaux de recherche. Nous commençons par la présentation d'une synthèse des approches et modèles que nous avons proposés en mentionnant surtout leurs principales caractéristiques. Par la suite, nous présentons nos contributions en nous situant par rapport à ce qui existe dans littérature, mais aussi en faisons un retour sur nos objectifs de recherche. Nous terminons ce chapitre et donc cette thèse en présentant les limites de notre travail, ainsi que les travaux futurs.

1 Synthèse

Dans cette thèse, nous avons proposé une nouvelle approche de géosimulation multi-niveau permettant de remédier aux lacunes des méthodes actuelles (p.ex. les modèles mathématiques, les automates cellulaires et les systèmes multi-agents classiques) quand elles sont utilisées seules et en bénéficiant de leurs avantages quand elles sont utilisées ensemble, et ce, à différents niveaux de granularité. Ces niveaux sont déterminés en fonction des besoins de l'utilisateur vis-à-vis de sa compréhension du phénomène, de la disponibilité des modèles représentant les acteurs et leurs comportements et de la disponibilité des données alimentant ces modèles. De plus, ces niveaux peuvent varier en fonction des échelles spatiales de l'environnement géographique virtuel géoréférencé, des échelles temporelles caractérisant les pas de la simulation, des différentes catégories d'acteurs (individus, groupes, populations ou métapopulations) et de la précision (niveau de détail) des modèles qui sont utilisés dans la simulation. Par ailleurs, nous étions intéressés à appliquer notre approche de géosimulation multi-niveau dans le domaine de la santé publique, afin d'aider les décideurs à mieux comprendre un phénomène aussi complexe que la propagation d'une zoonose et par conséquent, pouvoir prendre des décisions informées après avoir évalué différentes stratégies d'intervention dans le but de limiter la propagation de cette zoonose. De ce fait, nous avons présenté une vue d'ensemble des différents aspects qui devront normalement être traités lors du développement d'un outil de géosimulation de la propagation des zoonoses. Étant donné qu'il ne sera pas facile de traiter tous ces aspects dans le cadre de cette thèse surtout que le phénomène étudié est très complexe, nous avons

présenté dans une revue synthétique les différents aspects qui nous intéressent particulièrement et qui nous permettront éventuellement de satisfaire les objectifs de notre recherche.

En outre et afin justement de satisfaire ces objectifs, nous avons proposé un modèle théorique original que nous avons nommé *MASTIM (Multi-Actor Spatio-Temporal Interaction Model)*. Ce modèle permet de simuler des populations contenant un nombre d'individus très élevé en utilisant un modèle compartimental enrichi pouvant spécifier non seulement l'évolution de ces populations, mais aussi les aspects relatifs aux interactions spatio-temporelles de ces populations, incluant leurs déplacements dans l'environnement de simulation. Nous avons également proposé une nouvelle approche permettant de générer un environnement géographique virtuel informé (*IVGE*) dans lequel se déroule la géosimulation. Cet environnement est composé d'un ensemble de cellules élémentaires (correspondant à des polygones qui sont des régions de l'espace de forme irrégulière et qui sont bien définis à l'aide des données SIG) dans lesquelles les transitions des différents stades biologiques des populations concernées, ainsi que leurs interactions peuvent être simulés. Nous avons utilisé entre autres une subdivision spatiale basée sur la couverture du sol afin de refléter dans l'environnement de simulation les caractéristiques de l'habitat favorable pour les espèces impliquées dans la propagation d'une zoonose. Nous avons fusionné les cellules de couverture du sol ayant des caractéristiques similaires relativement au phénomène à simuler. Ainsi, nous avons obtenu les plus grands polygones possible permettant de minimiser les besoins computationnels sans pour autant diminuer la plausibilité de la simulation. La méthode de fusion proposée est basée sur des seuils sélectionnés par l'utilisateur qui modélise le phénomène. Ces seuils sont liés à la favorabilité de l'habitat qui peut affecter la survie, l'alimentation et/ou la reproduction des espèces. Après ce processus de fusion, nous avons ajouté à chacune des cellules des informations pertinentes comme par exemple la liste des cellules voisines, ainsi que leurs orientations quantitatives (angles par rapport à l'axe du Nord) et qualitatives (e.g *N*, *S*, *SW*, *ESE*, etc.). Nous avons également gardé une trace de toutes les fusions effectuées en

ajoutant à chaque cellule la liste des polygones fusionnés en incluant leur proportion de superficie et leur type de couverture du sol. De plus, nous avons ajouté dans l'*IVGE* des informations sur les trajectoires de certaines espèces impliquées dans la propagation de la maladie ce qui permettra de faciliter la distribution des flots d'individus comme par exemple la migration des oiseaux qui sont considérés comme un des plus importants réservoirs de plusieurs zoonoses.

D'autre part, nous avons appliqué nos travaux de recherche au développement d'outils d'aide à la décision. Nous avons acquis une première expérience avec le développement d'un outil (c.-à-d. *WNV-MAGS*) dont l'objectif principal est de simuler les comportements des populations de moustiques (*Culex*) et des oiseaux (corneilles) qui sont liés à la propagation et à la transmission du VNO. Cette simulation est faite dans un environnement géographique virtuel représentant un large territoire et selon divers scénarios climatiques et de traitements par larvicides. Notre solution pour le VNO a donné de bons résultats, surtout après la calibration du modèle compartimental faisant évoluer les populations impliquées dans la propagation de la maladie avec les données du sud de la province de Québec et la région métropolitaine d'Ottawa (Ontario). Nous avons par la suite participé au développement d'un outil générique (c.-à-d. *Zoonosis-MAGS*) permettant non seulement de simuler la propagation du VNO, mais aussi la propagation d'autres zoonoses comme la maladie de Lyme. Nous pensons que cet outil permettra aux décideurs dans le domaine de la santé publique d'appliquer et d'analyser les résultats de plusieurs scénarios alternatifs d'intervention dans le but de comprendre et estimer l'ampleur de l'évolution d'une zoonose et d'intervenir sur le terrain en connaissance de cause, après avoir pris des décisions informées.

2 Contributions

Nous commençons en soulignant le fait que notre travail est multi-disciplinaire. Il touche plusieurs domaines, dont notamment l'intelligence artificielle distribuée, les mathématiques, la géomatique, la recherche opérationnelle (aide à décision), les sciences naturelles (p. ex. biologie, écologie, épidémiologie, entomologie, ornithologie, etc.) ainsi

que les sciences humaines et sociales (p. ex. sciences médicales). En effet et comme nous l'avons déjà signalé à plusieurs reprises dans notre thèse, nous nous intéressons à la simulation d'un phénomène complexe faisant intervenir un grand nombre d'acteurs de différents types, interagissant dans un environnement géographique virtuel géoréférencé. Nous avons essayé de contribuer à l'élaboration d'approches, techniques et modèles informatiques permettant la géosimulation multi-niveau de la propagation des maladies infectieuses et ce, selon divers scénarios alternatifs d'intervention dans un objectif d'aide à la décision.

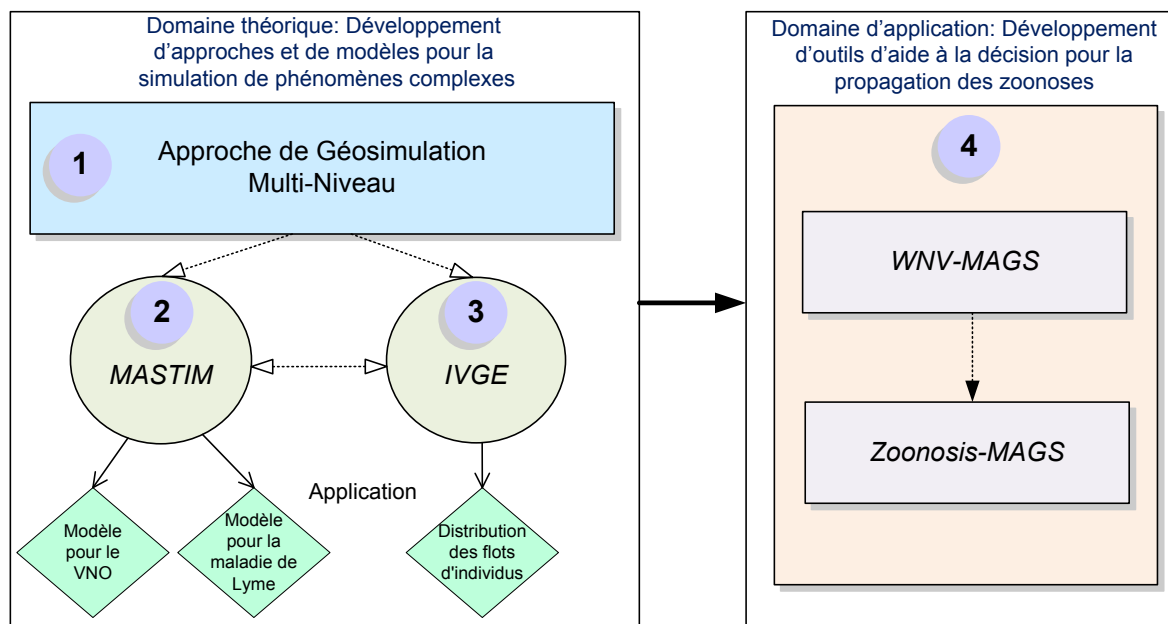


Figure 122: Nos principales contributions de recherche.

La Figure 122 présente nos 4 principales contributions de recherche qui sont effectivement: (1) l'approche de géosimulation multi-niveau, (2) le modèle *MASTIM*, (3) l'*IVGE* et (4) le développement d'outils d'aide à la décision. Le modèle *MASTIM* et l'*IVGE* représentent une suite logique de l'approche de géosimulation multi-niveau étant donné qu'ils permettent de résoudre des problèmes spécifiques que nous avons rencontrés dans nos travaux de recherche. Nous avons appliqué le modèle *MASTIM* à deux zoonoses en

spécifiant un modèle pour le VNO et un modèle pour la maladie de Lyme. De plus, nous avons appliqué l'*IVGE* à la distribution des flots d'individus. Par ailleurs et en utilisant ces approches et modèles, nous avons pu développer des outils d'aide à la décision (c.-à-d. *WNV-MAGS* et *Zoonosis-MAGS*) afin d'appliquer les concepts théoriques dans le domaine de la santé publique et plus spécifiquement pour simuler la propagation des zoonoses (Figure 122). Nous présentons d'ailleurs dans les paragraphes suivants les caractéristiques de chacune de nos contributions en nous situant par rapport à ce qui existe dans littérature:

1. Approche de géosimulation multi-niveau: Nous pensons que notre approche de géosimulation multi-niveau a beaucoup d'avantages si nous la comparons aux approches existantes. En effet et comme nous l'avons déjà mentionné, la simulation à base de modèles mathématiques qui utilise généralement des équations différentielles [Bowman *et al.*, 2005] ne permet pas de prendre en considération l'espace géographique dans lequel les populations évoluent. De plus, elle ne peut donner des résultats qu'à un niveau de granularité très élevé [Van Dyke Parunak *et al.*, 1998]. Par contre, la simulation à base d'automates cellulaires [Li *et al.*, 2012] modélise les caractéristiques spatiales des zones géographiques contenant des populations infectées par la maladie, mais elle ne permet pas de représenter les individus ainsi que leur mobilité. De plus, le découpage de l'espace géographique en un ensemble de cellules de forme régulière est considéré comme simpliste, ce qui ne reflète pas la réalité sur le terrain [Badariotti et Weber, 2002]. Quant à la simulation à base d'un système multi-agent classique [Bauer *et al.*, 2009], elle prend en compte explicitement la mobilité des individus en modélisant les comportements des agents ou des groupes d'agents. Cependant, cette approche ne profite pas généralement des données fournies par un SIG afin de bien repérer ces agents dans l'environnement de simulation. En outre, la géosimulation multi-agent [Moulin *et al.*, 2003] se base entre autres sur des données SIG qui fournissent le contexte spatial virtuel de la simulation. De plus, la géosimulation s'écarte de l'objectif par défaut de la simulation traditionnelle qui se limite à la prédiction. Elle tend, à travers les projets réalisés à ce jour, vers des simulations permettant d'inclure et de modifier des scénarios événementiels. Toutefois, cette approche présente elle aussi certaines limites puisqu'elle n'intègre pas les différents niveaux de

granularité auxquels le phénomène étudié peut être observé par les décideurs et ainsi, dans beaucoup de situations, elle a pour résultat une grande perte surtout dans sa puissance explicative. D'ailleurs, de nouvelles propriétés écologiques peuvent apparaître à différents niveaux d'observation. Ainsi, les résultats de la simulation d'une zoonose peuvent être différents en fonction du niveau de granularité choisi, étant donné que des phénomènes corrélés positivement à un niveau peuvent être corrélés négativement à un autre niveau. De ce fait, notre approche de géosimulation multi-niveau permet de remédier aux lacunes des méthodes actuelles en modélisant un système complexe à plusieurs niveaux de granularité.

2. Modèle théorique *MASTIM*: La propagation d'une zoonose est un phénomène très complexe puisque les mécanismes de transmission indirecte de la maladie font intervenir de nombreux acteurs ("hôtes", "vecteurs", "réservoirs") dont le nombre est souvent très élevé [Alexander *et al.*, 2012]. Ce nombre considérable d'individus présente justement une problématique de modélisation, surtout qu'il faut prendre en considération les multiples interactions spatio-temporelles des populations impliquées dans la propagation de la zoonose (p.ex. moustiques pour le VNO et tiques pour la maladie de Lyme). Les approches existantes comme les systèmes multi-agent classiques ne sont pas capables de simuler ces populations, étant donné qu'il est difficile de représenter chaque individu par un agent [Auchincloss et Diez Roux, 2008]. Les limites des ressources computationnelles des ordinateurs ainsi que le manque des données font en sorte que nous ne pouvons pas simuler des populations contenant des millions ou même des milliards d'individus en utilisant ce type d'approche. D'un autre côté, les épidémiologistes et les mathématiciens créent des modèles mathématiques à base de compartiments pour représenter l'évolution des populations impliquées dans la propagation des zoonoses en utilisant un ensemble d'équations différentielles [Bowman *et al.*, 2005]. Cependant, ces modèles n'intègrent pas les caractéristiques spatiales du phénomène étudié et ne peuvent pas tenir compte de l'influence des caractéristiques environnementales et géographiques sur les interactions des populations et de ce fait sur la propagation de la maladie d'une façon générale. De plus, ces modèles impliquent généralement un ensemble de paramètres et d'équations que les

décideurs de la santé publique ont de la difficulté à bien interpréter pour éventuellement suggérer de les améliorer. Nous pensons que notre modèle *MASTIM* propose une solution originale aux problèmes de modélisation mentionnés auparavant, surtout qu'il formalise entre autres les aspects spatiaux en les intégrant aux aspects des compartiments permettant de faire évoluer et interagir ces populations entre elles (contenant un nombre élevé d'individus de différentes espèces), mais aussi avec l'environnement de simulation. Par ailleurs, *MASTIM* propose, en plus de son formalisme, une représentation graphique de l'évolution et des interactions (ainsi que leurs conséquences) des espèces étudiées (voir Section 4 du Chapitre V). Nous pensons que cette représentation devrait permettre aux décideurs de la santé publique de mieux comprendre le phénomène simulé.

3. Approche de génération d'un *IVGE*: Les approches actuelles comme les automates cellulaires [White *et al.*, 2009] ne permettent pas d'intégrer les caractéristiques biologiques (p.ex. habitat favorable) et comportementales (p.ex. distribution à travers des trajectoires) des espèces impliquées dans la propagation d'une zoonose lors de la génération de l'environnement de simulation représentant des territoires étendus. De plus, l'utilisation des données vectorielles représentant par exemple les cellules de couverture du sol sans prétraitement n'est pas faisable. Le nombre très élevé de ces cellules et le fait que la simulation de la zoonose nécessite de cloner le modèle compartimental enrichi dans chacune de ces cellules ne permettent pas à la simulation d'obtenir des résultats dans un temps raisonnable. La fusion des cellules selon des critères et des seuils semble être une étape très importante pour la génération d'un *IVGE* permettant d'améliorer l'efficacité de la simulation. Certaines approches de fusion sont disponibles dans la littérature, mais nous pensons qu'elles ne peuvent pas être appliquées à notre domaine d'application. Nous avons déjà cité l'algorithme *GDBSCAN* [Sander *et al.*, 1998], mais il existe par exemple l'outil *AZTool* [Cockings *et al.*, 2011] qui permet de suggérer des agrégations de polygones (*Groups of Building Blocks*) pour des fins de zonage urbain (reconfigurer le territoire municipal en zones selon par exemple le nombre d'individus habitant dans chaque zone). Cet outil utilise uniquement une subdivision administrative (dans notre cas, les polygones de couverture de sol sont beaucoup plus complexes) et il ne modifie pas directement les

données vectorielles. Son résultat consiste entre autres à identifier les polygones qui peuvent être fusionnés formant ainsi des agrégats (appelé *Tracts*). Nous pensons que les résultats obtenus par notre *IVGE* sont beaucoup plus adaptés à notre cas. En plus du fait que nos algorithmes générant l'*IVGE* minimisent le nombre de cellules à traiter en appliquant un processus de fusion répétitif basé par exemple sur des seuils de favorabilité, ils ajoutent des informations supplémentaires (p.ex. biologiques, topographiques, climatiques, etc.) dans chaque polygone. Cet ajout donne le caractère informé à l'*IVGE* et permet donc à la simulation de bénéficier de tous les avantages des données qui sont stockées dans chacune des cellules. Par ailleurs, notre algorithme de distribution des flots d'individus montre justement ces avantages étant donné que cette distribution se base sur les attributs qui sont attachés aux cellules comme par exemple le coefficient d'attraction (basé sur le degré de favorabilité), les informations sur les voisins ainsi que le coefficient de trajectoire.

4. Développement d'outils d'aide à la décision: Les systèmes de surveillance des zoonoses fournissent principalement des statistiques et des outils de cartographie permettant d'informer les agents de la santé publique des différents indicateurs liés aux maladies comme par exemple la présence d'animaux infectés [Gosselin *et al.*, 2005], [Chen *et al.*, 2007]. Ces systèmes sont évidemment utiles, mais ils ne fournissent pas les moyens de bien comprendre et surtout de prédire la dynamique de la propagation d'une zoonose sur un vaste territoire, en tenant compte de l'influence du paysage et du climat sur les interactions spatio-temporelles des espèces étudiées. Par ailleurs, les outils de simulation qui existent actuellement (p.ex. *STEM* et *GLEaMviz*) ne peuvent pas être utilisés pour simuler la propagation des zoonoses. En effet et à part leurs inconvénients (p.ex. les modèles de maladies existants dans *STEM* sont prédéfinis et limités, ce qui restreint le nombre de maladies qu'il peut simuler), ces outils ont été plus conçus pour simuler des épidémies, mais surtout celles qui peuvent transmettre la maladie entre humains comme l'influenza. Nous pensons que nos approches et nos modèles nous a permis de développer des outils innovateurs permettant de simuler la propagation des zoonoses. D'ailleurs, l'outil *WNV-MAGS* est le seul qui existe actuellement permettant de simuler la propagation du

VNO en utilisant deux niveaux de granularité (c.-à-d. deux échelles spatiales). Ce simulateur est un outil d'aide à la décision offrant à l'utilisateur (c.-à-d. le décideur de la santé publique) la possibilité de choisir plusieurs types de scénarios (p.ex. applications de larvicides, fortes pluies, etc.) et de visualiser la propagation de la maladie sur un large territoire. La force de *WNV-MAGS* réside également dans la fiabilité des résultats obtenus puisque le modèle mathématique utilisé a été calibré pour le sud de Québec et validé pour la région métropolitaine d'Ottawa. D'autre part, l'outil *Zoonosis-MAGS* permettant de généraliser la solution obtenue avec *WNV-MAGS* est un simulateur pouvant simuler la propagation d'une zoonose d'une façon générique. C'est un outil très flexible puisque l'utilisateur peut lui-même créer ou modifier le modèle permettant de faire évoluer et interagir les espèces impliquées dans la propagation de la maladie. La spécification du modèle de simulation (incluant la spécification de la hiérarchie des cellules, des espèces, des compartiments, des transitions, des interactions, etc.) en utilisant des modules d'extension (p.ex. *plugin* des expressions mathématiques) rend le système *Zoonosis-MAGS* unique dans son genre. De plus, la spécification d'un scénario qui est une instanciation du modèle de simulation permet à l'utilisateur d'ajouter des événements déclenchant des processus (p.ex. migration des oiseaux, vague de chaleur, etc.). La spécification de plusieurs scénarios permet à l'utilisateur de comparer les résultats obtenus et de ce fait, mieux comprendre un phénomène aussi complexe que la propagation d'une zoonose.

En tenant compte de ces contributions et du travail présenté dans les chapitres de cette thèse, nous pensons que nos objectifs de recherche ont été atteints. En effet, notre premier objectif était de proposer **une approche de géosimulation multi-niveau permettant l'intégration des différents niveaux de granularité** auxquels une zoonose peut être observée et comprise par les décideurs de la santé publique et selon la disponibilité des données qui alimentent les modèles de simulation. Le chapitre IV présente cette approche représentant l'objet de notre première contribution et qui a permis la réalisation des autres contributions mentionnées ci-dessus. Notre deuxième objectif était de proposer un **modèle théorique permettant de modéliser et simuler les interactions spatio-temporelles d'un grand nombre d'acteurs** de différents types (individus, groupes ou populations d'humains

et d'animaux), y compris ceux qui représentent des populations contenant un nombre considérable d'individus. Le chapitre V présente ce modèle (c.-à-d. *MASTIM*) représentant l'objet de notre deuxième contribution. Notre troisième objectif était de proposer **une approche permettant de générer un environnement géographique virtuel informé** dans lequel se déroule la géosimulation en utilisant des données géoréférencées. Le chapitre VI présente cette approche (c.-à-d. *l'IVGE*) représentant l'objet de notre troisième contribution. Finalement, notre dernier objectif était de développer des **prototypes permettant d'appliquer notre approche à deux cas de zoonoses** afin de montrer sa généralité. Le chapitre VII présente ces outils (c.-à-d. *WNV-MAGS* et *Zoonosis-MAGS*) représentant l'objet de notre quatrième contribution.

3 Limites et travaux futurs

Bien que nous soyons convaincus de la pertinence et de l'originalité de notre travail, nous sommes tout à fait conscients qu'il présente des limitations auxquelles nous n'avons pas pu remédier faute de temps. Nous résumons les plus importantes limites dans les principaux points suivants en suggérant des travaux futurs permettant d'y répondre:

1. Le changement du niveau de granularité lors de la simulation: Nous avons distingué, dans notre approche de géosimulation multi-niveau, plusieurs types de niveaux de granularité (les échelles spatio-temporelles, les catégories d'acteurs et les niveaux de détail des modèles utilisés). De plus, les différents niveaux (c.-à-d. Macro, Méso et Micro) appartenant à un même type sont déterminés selon les besoins de l'utilisateur, la disponibilité des modèles et la disponibilité des données alimentant ces modèles. De ce fait, l'utilisateur peut choisir une combinaison de niveaux compatibles afin de simuler la propagation d'une zoonose. Par exemple, cette simulation peut se faire au niveau des populations sur un large territoire en utilisant un pas de simulation d'une semaine et un modèle dont le niveau de détail représente les flux de déplacements des individus. Nous pensons que le fait de fixer à l'avance un niveau pour chacun des types et ne pas les changer d'une façon dynamique ne posera généralement pas de problème pour la

simulation. Cependant, si au moins deux niveaux de granularité (p.ex. Macro et Méso) sont considérés au sein d'un même type (p.ex. échelle spatiale) pour une même simulation, la question de changement d'un niveau à un autre lors de la simulation mérite une réflexion. En effet, quand l'utilisateur fait un zoom (en avant ou en arrière) de l'environnement de simulation, les résultats de la simulation des niveaux de granularité utilisés doivent être cohérents. Ainsi, il faut s'assurer que les modèles utilisés par deux niveaux de granularité soient compatibles pour pouvoir faire le passage entre eux (p.ex. passer d'un niveau Macro à un niveau Méso en faisant un zoom avant). Il faut par contre mentionner que le passage d'un niveau de granularité détaillé à un haut niveau de granularité moins détaillé peut se faire en agrégeant des données (p.ex. passer d'un pas de simulation d'un jour à un pas de simulation d'une semaine en agrégeant les données pour avoir des résultats à toutes les semaines). D'ailleurs, nos outils de simulation supportent cette agrégation de données. Nous avons déjà mentionné que *Zoonosis-MAGS* peut agréger les données d'un niveau inférieur (p.ex. subdivision selon la couverture du sol) à un niveau supérieur (p.ex. subdivision administrative). Cependant, le passage entre certains types de niveaux de granularité (surtout d'un niveau supérieur à un niveau inférieur) peut poser un problème de compatibilité entre les modèles utilisés par ces niveaux. Par exemple et dans une simulation Macro, la population impliquée dans le phénomène étudié est considérée comme une structure caractérisée par un ensemble de variables (p.ex. utilisation de la dynamique des systèmes pour spécifier ces variables). Cependant et dans une simulation Micro, en supposant qu'elle est pertinente pour la propagation d'une zoonose (p.ex. l'utilisateur aimerait simuler le déplacement individuel des chevreuils pour le cas de la maladie de Lyme), le système modélise les comportements des individus composant une population (p.ex. utilisation d'un système multi-agent pour spécifier ces comportements). De ce fait, les deux modèles utilisés (c.-à-d. un modèle de dynamique des systèmes et un modèle de système multi-agent) doivent être compatibles afin que les interactions des individus au niveau Micro fassent émerger un comportement global de la population paramétrisée par des variables au niveau Macro. Pour remédier à cette limite, nous envisageons d'explorer des pistes afin d'adapter par exemple les travaux de Gaud et ses collaborateurs [Gaud *et al.*, 2008] à notre domaine d'application. En effet, ces derniers proposent une gestion efficace

de la transition entre les niveaux d'abstraction pour la simulation des piétons en utilisant une approche holonique. Les holons atomiques appartenant au niveau le plus bas, correspondant au niveau microscopique, sont regroupés et associés à un super-holon. Le mécanisme de regroupement successif des holons est ainsi reproduit jusqu'à obtenir la holarchie d'exécution du système dont le sommet correspond au niveau macroscopique. Chaque super-holon joue à la fois un rôle applicatif (exécuter une approximation du comportement des rôles de ses membres) et un rôle d'ordonnancement (exécuter effectivement chaque comportement de ses membres). Le fait de simuler le comportement d'un groupe d'entités à un niveau d'abstraction donné signifie que le super-holon du niveau considéré n'exécutera pas le comportement de ses membres, mais son rôle applicatif. Le système est simulé au niveau microscopique si tous les holons, composant le plus bas niveau de la hiérarchie, sont exécutés. Les auteurs proposent qu'en fonction des contraintes de la simulation, un super-holon déterminera s'il doit exécuter son rôle applicatif ou celui de ses membres. En fonction de cette décision, le niveau de précision de la simulation pourra changer. Si tous les super-holons de la simulation exécutent leurs membres respectifs, la simulation est alors globalement à son niveau le plus précis. Par contre, si seul le rôle applicatif associé au holon situé au sommet de la holarchie est exécuté, la simulation est à son niveau de précision le plus faible. La holarchie d'exécution permet ainsi de moduler dynamiquement le niveau de précision d'une simulation [Gaud *et al.*, 2008].

2. Raffinement de *MASTIM* et application à d'autres zoonoses: Le modèle *MASTIM* que nous avons proposé formalise tous les concepts qui nous intéressent comme les cellules, les espèces, les compartiments, les transitions, les interactions, les conditions, le patron de mobilité, la trajectoire, le scénario, la simulation, etc. Cependant, nos formules ne sont pas totalement détaillées. Il leur manque par exemple la spécification des paramètres utilisés. Par exemple, *SimParam_m* représentant les paramètres de la simulation (voir Section 3.2.3 du Chapitre VII) a été mentionné sans être formalisé. Nous pouvons éventuellement nous inspirer de certains travaux (p.ex. [Ferber et Müller, 1996], [Zeigler *et al.*, 2000]) afin de raffiner notre modèle théorique. Par ailleurs, il serait également intéressant d'appliquer

notre modèle à d'autres zoonoses afin de prouver encore plus sa généralité. En fait, chaque zoonose a ses propres caractéristiques et de ce fait, il est probable que le modèle *MASTIM* aurait besoin d'être étendu pour formaliser des nouveaux concepts ou enrichir des concepts existants. Il existe bien sûr plusieurs zoonoses que nous puissions étudier comme par exemple la Rage [Wiraningsih *et al.*, 2010], la Dengue [Esteva *et al.*, 1998], l'Anthrax (maladie du charbon) ou encore l'Ébola (fièvre hémorragique à virus Ébola) [Alexander *et al.*, 2012].

3. Génération de l'IVGE et son application à la maladie de Lyme: Nous avons déjà mentionné (voir Section 5 du Chapitre VII) que nous pensons rendre le processus de fusion progressive des cellules un peu plus efficace lors de la génération de l'IVGE en faisant une recherche en largeur ou une recherche en profondeur jusqu'à atteindre un résultat satisfaisant selon l'appréciation de l'utilisateur en se basant sur les critères et les seuils retenus. De plus, nous pensons créer une interface conviviale permettant à l'utilisateur d'entrer ces critères et ces seuils afin de générer automatiquement l'IVGE en se connectant à la base de données spatiale *PostGIS*. D'un autre côté, l'outil *Zoonosis-MAGS* peut charger l'IVGE, mais il ne l'utilise pas pour afficher les résultats de la simulation étant donné que le moteur d'affichage implanté actuellement n'est pas apte à le faire. Pour remédier à cette limite, nous envisageons l'utilisation d'une librairie permettant d'optimiser l'affichage des cellules ainsi que les données qui sont attachées à ces cellules. Nous pensons par exemple à utiliser Mapnik [Site Web: Mapnik] ou encore Saga [Site Web: Saga].

4. Manque et hétérogénéité des données de la simulation: Nous avons déjà mentionné que nous avons éprouvé beaucoup de problèmes avec la collecte des données. En effet, le manque des données pour alimenter les modèles de simulation, ainsi que le degré de validité des données qui sont disponibles pour calibrer ces modèles, sont parmi les principaux problèmes de la simulation d'un phénomène tel que la propagation d'une zoonose. D'ailleurs, il est parfois impossible de trouver ces données dans la littérature ou même auprès des experts du domaine. De ce fait, nous avons été obligés à poser certaines hypothèses afin de traiter le problème des données manquantes. Cependant, certaines de ces

hypothèses peuvent réduire le degré de réalisme de la simulation. D'un autre côté, les données permettant de calibrer les modèles de simulation et issues généralement des programmes de surveillance présentent un certain biais. Par exemple, la collecte ou l'analyse des animaux infectés ne se fait pas sur une base régulière et la plupart du temps dans un sous ensemble restreint de zones géographiques. L'échantillonnage est parfois déterminé d'une façon aléatoire dans le temps et dans l'espace. Par conséquent, la très forte variabilité de la collecte des données par les programmes de surveillance entraîne très souvent un manque de données pour alimenter les simulations. Le problème réside dans le fait que les concepteurs des programmes de surveillance ne pensent pas toujours à l'utilité des données ramassées afin de pouvoir bien les interpréter. D'ailleurs, l'utilisation des données collectées par les programmes de surveillance est nécessaire pour le développement d'un outil d'aide à la décision utilisant la géosimulation tel que *WNV-MAGS* ou *Zoonosis-MAGS*. Nous pensons donc qu'il est important de faire certaines recommandations afin d'enrichir les programmes de surveillance de zoonoses en vue de la géosimulation. En effet, nous croyons qu'il est plus intéressant de se concentrer sur quelques régions bien déterminées (p.ex. quelques municipalités) que de vouloir tout surveiller sur un large territoire. De ce fait, on ne se retrouvera pas avec des ensembles de données gigantesques qui sont dans la plupart des cas inutilisables. Au contraire, on aura des données obtenues en utilisant un échantillonnage bien réparti dans le temps et dans l'espace. De plus, il faut procéder à certaines répétitions dans la collecte des données d'année en année afin de pouvoir valider la fiabilité de ces données, et en suivre l'évolution.

5. Analyse des résultats: Les outils que nous avons développés ou pour lesquels nous avons participé au développement permettent aux décideurs de la santé publique d'appliquer plusieurs scénarios alternatifs afin de tester par exemple plusieurs combinaisons possibles d'événements et voir leurs effets sur le déroulement de la simulation. L'application de ces scénarios va engendrer des résultats qu'il faut analyser pour aider les décideurs à les comprendre. Nous n'avons malheureusement pas eu le temps

d'aborder ce sujet dans notre thèse. Pour remédier à cette limite, nous envisageons d'explorer certaines approches permettant l'analyse des résultats de la simulation. En effet, certaines techniques et/ou outils statistiques pourront éventuellement répondre aux besoins des décideurs de la santé publique comme par exemple ce que proposent Nakayama [Nakayama, 2008] ou Grier [Grier, 1992]. Cependant, ces techniques d'analyse classique qui sont essentiellement basées sur des tableaux et/ou des graphiques sont un peu limitées pour l'analyse spatiale d'un phénomène aussi complexe que la propagation d'une zoonose. De ce fait, nous pensons que l'utilisation par exemple de l'approche *SOLAP* (*Spatial Online Analytical Processing*) [Site Web: SOLAP] pourrait améliorer l'interprétation des résultats de simulation de ce type de phénomène. D'ailleurs, cette approche permet l'exploration et l'analyse spatio-temporelle des données selon une approche multidimensionnelle à plusieurs niveaux d'agrégation via un affichage cartographique, tabulaire ou en diagramme statistique [Bédard, 2004]. Nous pouvons donc bénéficier de l'aspect multidimensionnel de cette approche afin d'analyser les résultats des différents niveaux de granularité que nous proposons pour géosimuler la propagation d'une zoonose. Cependant, il faut choisir avec soin les données qu'il faut analyser et également bien les préparer pour être directement utilisables par SOLAP. Par ailleurs, une analyse multi-critère [Roy, 1985] basée sur une comparaison entre les résultats des scénarios pourrait également aider les décideurs à avoir une vision plus précise des options qui s'offrent à eux [Mourits et Oude Lansink, 2006]. Cette comparaison permettrait aux décideurs d'opter pour la meilleure stratégie d'intervention en fonction de certains critères choisis. Cependant, il faut bien spécifier ces critères comme par exemple la date et la zone d'application de larvicides pour le cas de la simulation de la propagation du VNO.

Enfin, nous aimerions conclure cette thèse par le fait que le phénomène qui nous intéresse (c.-à-d. la propagation d'une maladie infectieuse) est une réalité que nous vivons tous. Par exemple, nous avons entendu parler de l'éclosion d'une souche du virus de la grippe A (H1N1)³⁷ au Mexique et juste après, nous avons constaté l'apparition de cas confirmés au

³⁷ La grippe A (H1N1) (initialement appelée grippe porcine) est une maladie infectieuse très contagieuse d'origine virale (http://fr.wikipedia.org/wiki/Influenzavirus_A_sous-type_H1N1).

Canada et dans plusieurs pays du monde, ce qui a déclenché un programme massif de vaccination. Cet exemple de propagation d'une maladie infectieuse qui ne peut que nous inquiéter tous, souligne l'importance de notre travail.

Bibliographie

1. [Abbott, 2007] Abbott R. (2007). Putting complex systems to work. *Complexity*, vol. 13, issue 2, pp. 30–49. <http://onlinelibrary.wiley.com/doi/10.1002/cplx.20200/pdf>
2. [Alexander *et al.*, 2012] Alexander K.A., Leis B.L., Marathe M., Eubank S., Blackburn J. K. (2012). Modeling of wildlife-associated zoonoses: Applications and caveats. *Vector-Borne and Zoonotic Diseases*, 12(12), pp. 1005-1018.
3. [Ali *et al.*, 2007] Ali, W., Moulin, B., Bedard, Y., Proulx, M.J., Rivest, S. (2007). Coupling multiagent geosimulation and spatial OLAP for better geosimulation data analysis. *URISA Journal*, Washington DC, vol. 19 (2), pp. 23-32.
4. [Ali et Moulin, 2007] Ali A., Moulin B. (2007). Multiagent geosimulation of human behaviors in micro-scale geographic environments: The case of the shopping behavior in a mall. GeoComputation International Conference, the National University of Ireland, Maynooth, Ireland). <http://ncg.nuim.ie/geocomputation/sessions/8B/8B2.pdf>
5. [An *et al.*, 2005] An L., Linderman M., Qi J., Shortridge A., Liu J. (2005). Exploring complexity in a human-environment system: An agent-based spatial model for multi-disciplinary and multiscale integration. *Annals of the Association of American Geographers*, 95(1), pp. 54–79.
6. [Arino, 2009] Arino J. (2009). Diseases in metapopulations. *Modeling and Dynamics of Infectious Diseases*. Series in Contemporary Applied Mathematics, vol. 11, pp. 65-123. http://www.math.umanitoba.ca/~jarino/papers/Arino_metapopulations.pdf
7. [Arthur, 1993] Arthur W.B. (1993). Why do things become more complex? *Scientific American*, May, pp. 92.
8. [Auchincloss et Diez Roux, 2008] Auchincloss A.H., Diez Roux A.V. (2008). A new tool for epidemiology: The usefulness of dynamic-agent models in understanding place effects on health. *American Journal of Epidemiology*, 168(1), pp. 1-8.
9. [Avois-Jacquet, 2002] Avois-Jacquet C. (2002). “Variabilité spatiale multi-échelle du zooplancton dans un lagon récifal côtier”. Thèse de doctorat, Université de Montréal. 302 pages.
10. [Axelrod et Cohen, 2000] Axelrod, R., Cohen M. (2000). Harnessing complexity. *Free Press*, 2000. <http://slashdot.org/books/00/10/06/1444228.shtml>
11. [Badariotti et Weber, 2002] Badariotti D., Weber C. (2002). La mobilité résidentielle en ville. Modélisation par automates cellulaires et système multi-agents à Bogota. *Espace géographique*, tome 31, pp. 97-108.
12. [Bally, 2006] Bally N. (2006). “L’utilisation des concepts de la théorie de la complexité de Kolmogorov en épistémologie”. Mémoire de Master Recherche, philosophie mention histoire des sciences et épistémologie. 41 pages.
13. [Bauer *et al.*, 2009] Bauer A.L., Beauchemin C., Perelson A.S. (2009). Agent-based modeling of host-pathogen systems: The successes and challenges. *Inform. Sciences*, 179(10), pp.1379-1389.
14. [Beauchemin *et al.*, 2005] Beauchemin C., Samuel J., Tuszynski J. (2005). A simple cellular automaton model for influenza: A viral infections. *Journal of Theoretical Biology*, 232(2), pp. 223-234. <http://arxiv.org/abs/q-bio.CB/0402012>.
15. [Bédard, 2004] Bédard Y. (2004). Amélioration des capacités décisionnelles des SIG par l'ajout d'un module SOLAP. Université de Provence, Centre de Mathématiques et Informatique, Laboratoire des Sciences de l'Information et des Systèmes (LSIS), Marseille, 8 avril.

16. [Benenson et Torrens, 2004] Benenson I., Torrens P.M. (2004). "Geosimulation: Automata-Based Modeling of Urban Phenomena". John Wiley & Sons. 312 pages.
<http://onlinelibrary.wiley.com/book/10.1002/0470020997>
17. [Benkirane, 2006] Benkirane R. (2006). "La complexité, vertiges et promesses". Édition le Pommier, 411 pages.
18. [Borshchev et Filippov, 2004] Borshchev A., Filippov A. (2004). From system dynamics and discrete event to practical agent based modelling: Reasons, techniques, tools. The 22nd International Conference of the System Dynamics Society, July 25 - 29, 2004, Oxford, England.
19. [Bouden et Moulin, 2012a] Bouden M., Moulin B. (2012). Generating an informed virtual geographic environment through cell merging in order to geosimulate the propagation of zoonoses. In Proceedings of International Conference on Modeling, Simulation and Visualization Methods. World Congress in Computer Science, Computer Engineering, and Applied Computing. Las Vegas, Nevada, USA, July 16-19, 2012. <http://elrond.informatik.tu-freiberg.de/papers/WorldComp2012/MSV4726.pdf>
20. [Bouden et Moulin, 2012b] Bouden M., Moulin B. (2012). Zoonosis-MAGS: A generic multi-level geosimulation tool for zoonosis propagation. In Proceedings of Global Geospatial Conference 2012. Spatially Enabling Government, Industry and Citizens. Quebec City, Canada, 14-17 May 2012. <http://www.gsdi.org/gsdiconf/gsdi13/papers/208.pdf>
21. [Bouden et Moulin, 2012c] Bouden M., Moulin B. (2012). A spatio-temporal interaction model used to geosimulate the zoonosis propagation. In Proceedings of Symposium on Theory of Modeling and Simulation (TMS'12). Spring Simulation Multi-Conference. The Society for Modeling & Simulation International. March 26-29 2012, Orlando, FL, USA.
<http://dl.acm.org/citation.cfm?id=2346619>
22. [Bouden *et al.*, 2010] Bouden M., Moulin B., Gosselin P. (2010). A geosimulation tool to assess intervention scenarios in relation to the spread of infectious disease and environmental factors. International Environmental Modelling and Software Society (iEMSs). 2010 International Congress on Environmental Modelling and Software. Modelling for Environment's Sake, Fifth Biennial Meeting, July 5-8 2010, Ottawa, Canada.
<http://www.iemss.org/iemss2010/proceedings.html> (See Section S27).
23. [Bouden et Moulin, 2010] Bouden M., Moulin B. (2010). Multi-level geosimulation of zoonosis propagation: A multi-agent and climate sensitive tool for risk management in public health. Chapter in the book "Advances in Risk Management", (Ed.), ISBN: 978-953-307-138-1, Sciyo.
<http://sciyo.com/articles/show/title/multi-level-geosimulation-of-zoonosis-propagation-a-multi-agent-and-climate-sensitive-tool-for-risk->
24. [Bouden *et al.*, 2008a] Bouden M., Moulin B., Gosselin P. (2008). The Geosimulation of west Nile virus propagation: A Tool for Risk Management in Public Health. *International Journal of Health Geographics*, 7:35, 2008. <http://www.ij-healthgeographics.com/content/7/1/35>
25. [Bouden *et al.*, 2008b] Bouden M., Moulin B., Gosselin P. (2008). Epidemic Propagation of West Nile Virus Using a Multi-Agent Geo-Simulation under various short-term climate scenarios. Agent-directed Simulation (ADS 2008). Ottawa, Canada, April 14-17, 2008.
<http://dl.acm.org/citation.cfm?id=1400562>
26. [Bouden *et al.*, 2007] Bouden M., L'heureux G., Mekni M., Moulin B., Sahli N. (2007). Literature review on the modeling and simulation of crowd behaviours, monitoring and non-

- lethal weapons. Intermediate deliverables of the first phase of the Crowd-MAGS project. CR 2007-047, January 2007, 259 pages.
27. [Bouden, 2004] Bouden M. (2004). "Utilisation de systèmes de particules pour la propagation de la fumée ou d'un gaz dans un monde virtuel URBAIN 3D". Mémoire de maîtrise, Département d'informatique et génie logiciel, Université Laval, avril 2004, 148 pages.
http://books.google.ca/books/about/Utilisation_des_syst%C3%A8mes_de_particules.html?id=vHUJOAAACAAJ&redir_esc=y
 28. [Boulding, 1956] Boulding K. (1956). General Systems Theory, the skeleton of science. *Management Science*, 2, 3 (Apr. 1956) pp.197-208.
 29. [Bowman *et al.*, 2005] Bowman C., Gumel A.B., Van Den Driessched P., Wue J., Zhue H. (2005). A mathematical model for assessing control strategies against West Nile virus. *Bulletin of Mathematical Biology*, 67 (2005) pp. 1107–1133.
 30. [Branicky et Mattson, 1997] Branicky M. S., Mattson S.E. (1997). Simulation of hybrid systems. In Proceedings of Hybrid Systems, IV Workshop, Lecture Notes in Computer Science, 1273, (Éd. P. Antsaklis, W. Kohn, A. Nerode, S. Sastry), pp. 34-56, Springer Computer Science.
 31. [Brodu, 2007] Brodu N (2007). "Practical investigations of complex systems". Thèse de doctorat de l'Université Concordia, Montréal. Juin 2007. 225 pages.
 32. [Burger et Todd, 2006] Burger O., Todd L. C. (2006). Grain, extent, and intensity: The components of scale in archaeological survey. *Confronting Scale in Archaeology, Issues in Theory and Practice*, pp. 235-256. <http://www.springerlink.com/index/q5h6884x87384366.pdf>
 33. [Carvalho *et al.*, 2006] Carvalho J.P., Carola M., Tomé J.A.B. (2006). Using rule-based fuzzy cognitive maps to model dynamic cell behaviour in voronoi based cellular automata. IEEE International Conference on Fuzzy Systems, Vancouver, BC, Canada, July 16-21, 2006.
<http://www.inesc-id.pt/pt/indicadores/Ficheiros/3290.pdf>
 34. [Chaker et Moulin, 2008] Chaker W., Moulin, B. (2008). Forecasting travel supply and demand by modeling multiscale urban environments. *Transportation Research Record, Journal of the Transportation Research Board*, vol. 2064, pp. 65-72.
 35. [Chardonnel *et al.*, 2004] Chardonnel S., du Mouza C., Fauvet M.-C., Josselin D., Rigaux P. (2004). Patrons de mobilité: Proposition de définition, de méthode de représentation et d'interrogation. Conférence Nationale avec comité de lecture : 7ème Journées CASSINI, Janvier 2004, pp.19-25.
 36. [Chen *et al.*, 2007] Chen D., Cunningham, J. Moore K., Tian J. (2011). An overview of syndromic surveillance: Applications of GIS to the Epidemiology of infectious disease control. *Annals of GIS*, vol. 17(4), pp. 211-220.
 37. [Cockings *et al.*, 2011] Cockings S., Harfoot D., Martin D., Hornby D. (2011). Maintaining existing zoning systems using automated zone design techniques: methods for creating the 2011 Census output geographies for England and Wales. *Environment and Planning A*, 43(10), pp. 2399-2418.
 38. [Cossentino *et al.*, 2007] Cossentino M., Gaud N., Hilaire V., Galland S., Koukam A. (2007). A metamodel and implementation platform for holonic multi-agent systems. Fifth European Workshop on Multi-Agent Systems (EUMAS'07). Hammamet, Tunisia - December 13th to 14th, 2007.
 39. [Delorme, 1998] Delorme M. (1998). An introduction to cellular automata. In "Cellular Automata: A parallel model". M. Delorme and J. Mazoyer Eds, Mathematics and Its Application, Kluwer.

40. [Deng *et al.*, 2008] Deng C., Tao H., Ye Z. (2008). Agent-based modeling to simulate the dengue spread. *Geoinformatics 2008 and Joint Conference on GIS and Built Environment: Geo-Simulation and Virtual GIS Environments*, In Proceedings of SPIE, vol. 7143 71431O-1.
41. [Dong-Hwan et Jae-Ho, 1997] Dong-Hwan K., Jae-Ho J. (1997). System dynamics as a modelling platform for multi-agent systems. In Proceedings of the 15th International Conference of the System Dynamics Society, Istanbul.
42. [Drogoul et Ferber, 1994] Drogoul A., Ferber J. (1994). Multi-agent simulation as a tool for modeling societies: Application to social differentiation in ant colonies. *Artificial Social Systems*, 830(8), pp. 3–23.
43. [Emrich *et al.*, 2007] Emrich S., Suslov S., Judex F. (2007). Fully agent based. modellings of epidemic spread using Anylogic. In Proceedings EUROSIM, 9-13 Sept. 2007, Ljubljana, Slovenia.
44. [Esteva *et al.*, 1998] Esteva L., Vargas C. (1998). Analysis of a dengue disease transmission model. *Mathematical Biosciences*, 150(2), pp. 131-151.
45. [Farenc *et al.*, 1999] Farenc N., Boulic R., Thalmann D. (1999). An informed environment dedicated to the simulation of virtual humans in urban context. *Computer Graphics Forum (Eurographics '99)*, Brunet P. and Scopigno R. Eds, vol. 18(3), pp. 309-318.
46. [Farge, 1988] Farge M. (1988). L'approche numérique : Simulation ou simulacre des phénomènes ?. Jean Petitot Ed., “Logos et Theorie des Catastrophes ”, Édition Patino (1988), pp. 119-139.
47. [Ferber et Müller, 1996] Ferber J, Müller J. (1996). Influences and reaction: A model of situated multi-agent systems. In Proceedings of the 2nd International Conference on Multi-agent Systems (ICMAS-96). pp. 72–79.
48. [Ferber, 1995] Ferber J. (1995). Les systèmes multi-agents. Vers une intelligence collective. InterEditions, Paris, 522 pages. http://www.lirmm.fr/~ferber/publications/LesSMA_Ferber.pdf
49. [Fishwick, 1995] Fishwick P.A. (1995). “Simulation model design and execution: Building digital worlds”. Prentice Hall, 450 pages.
50. [Fonseca *et al.*, 2004] Fonseca D.M., Keyghobadi N., Malcolm C.A. (2004). Emerging vectors in the *Culex pipiens*. *Complex. Science*, vol. 303(5663), pp. 1535-1538.
51. [Forrester, 1958] Forrester J. W. (1958). Industrial Dynamics: A major breakthrough for decision makers. *Harvard Business Review*, vol. 36, no. 4, pp. 37–66.
52. [Francis et Hussell, 1998] Francis C.M., Hussell D.J.T. (1998). Changes in numbers of land birds counted in migration at Long Point Bird Observatory, 1961–1997. *Bird Population*, vol. 4, pp. 37-66.
53. [Fu et Milne, 2003] Fu S.C. and Milne G. (2003). Epidemic modelling using cellular automata. In Proceedings of the Australian Conference on Artificial Life, 2003.
54. [Fulford *et al.*, 2002] Fulford G.R., Roberts M.G., Heesterbeek J.A.P. (2002). The metapopulation dynamics of an infectious disease: tuberculosis in possums. *Theoretical Population Biology*, vol. 61, pp. 15–29.
55. [Gareau, 1997] Gareau A. (1997). “Utilisation des systèmes de particules pour la simulation de phénomènes naturels”. Thèse de doctorat présenté à l'Université Claude Bernard-Lyon 1, 90 pages.
56. [Gaud *et al.*, 2008] Gaud N., Galland S., Gechter F., Hilaire V., Koukam A. (2008). Holonic multilevel simulation of complex systems: Application to real-time pedestrians simulation in

- virtual urban environment. *Simulation Modelling Practice and Theory*, vol. 16, issue 10, November 2008, pp. 1659-1676.
57. [Gauthier et Aubry, 1995] Gauthier J., Aubry Y. (1995). L'Atlas des oiseaux nicheurs du Québec méridional. Édité par l'association québécoise des groupes d'ornithologues. Publié en collaboration avec la Société québécoise de protection des oiseaux et le Service canadien de la faune, Environnement Canada, Région du Québec.
 58. [GDG, 2002] Gestion et contrôle biologique des insectes piqueurs, 2002. Données biologiques sur *Culex pipiens* Linnaeus et *Culex restuans* Theobald et implications sur la transmission du VNO. GDG Environnement Ltée. 17 pages.
 59. [Gerber *et al.*, 1999] Gerber C., Siekmann J.H., Vierke G. (1999). Holonic multi-agent systems. Technical Report, DFKIRR-99-03, Deutsches Forschungszentrum für Künstliche Intelligenz - GmbH, Postfach 20 80, 67608 Kaiserslautern, FRG, May 1999.
 60. [Gosselin *et al.*, 2005] Gosselin P, Lebel G, Rivest S, Douville-Fradet M. (2005). The Integrated System for Public Health Monitoring of west Nile virus (ISPHM-WNV): a real-time GIS for surveillance and decision-making. *International Journal of Health Geographics*, 4:21, 2005. <http://www.ij-healthgeographics.com/content/4/1/21>
 61. [Gourley *et al.*, 2008] Gourley S.A., Liu R., Wu J. (2008). Spatiotemporal patterns of disease spread: interaction of physiological structure, spatial movements, disease progression and human intervention. *Structured population models in biology and epidemiology*. Lecture Notes in Mathematics, vol. 1936, pp. 165-208.
 62. [Grier, 1992] Grier D.A. (1992). Graphical techniques for output analysis. In Proceedings of the 1992 Winter Simulation Conference, pp. 314-319.
 63. [Haddad et Moulin, 2008] Haddad, H., Moulin, B. (2008). An agent-based geosimulation multidisciplinary approach to support scenarios evaluation in dynamic virtual geographic environments. In Proceedings of the Agent-Directed Simulation Symposium ADS 2008, pp. 53-60.
 64. [Hannon et Ruth, 2009] Hannon B., Ruth M. (2009). "Dynamic modeling of diseases and pests". Springer Science, Business Media LLC 2009. <http://www.springerlink.com/content/n307w2/>
 65. [Hansen *et al.*, 2006] Hansen J., Sato M., Ruedy R., Lo K., Lea D.W. Medina-Elizade M. (2006). Global temperature change. In Proceedings of the National Academy of Sciences of the United States of America. September 26, 2006, vol. 103, no. 39, pp. 14288–14293.
 66. [Hanski, 1999] Hanski I. (1999). "Metapopulation Ecology". Oxford University Press. 1999. ISBN 0-19-854065-5. <http://www.amazon.com/Metapopulation-Ecology-Oxford-Evolution/dp/0198540655>.
 67. [Hsieh *et al.*, 2006] Hsieh J.L., Sun C.T., Kao G.Y.M., Huang C.Y. (2006). Teaching through Simulation: Epidemic dynamics and public health policies. *SIMULATION*, 2006, vol. 82, no. 11, pp. 731-759.
 68. [Hu *et al.*, 2008] Hu B., Gong J., Sun J., Cao W., Fang L. (2008). GIS-based epidemical disease information visualization. Geoinformatics 2008 and Joint Conference on GIS and Built Environment. In Proceedings of SPIE, vol. 7143, 71431J-1.
 69. [Hubalek et Halouzka, 1999] Hubalek Z., Halouzka J. (1999). West Nile fever-A reemerging mosquito-borne viral disease in Europe. *Emerg Infect Dis*, 5(5), pp. 643-650.
 70. [Jakovljevic et Basch, 2004] Jakovljevic G., Basch D. (2004). Implementing multiscale traffic simulators using agents. Information Technology Interfaces, 2004. 26th International

Conference on Publication Date: 7-10 June 2004.

http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1372469

71. [Karsky, 2004] Karsky M. (2004). Dynamique des systèmes complexes. Concepts et méthodologies. Édition TI - Sciences et Techniques. 12 pages.
72. [Kirkwood, 1998] Kirkwood C. W. (1998). System dynamics methods: A quick introduction. Department of Management. Arizona State University, Tempe, AZ 85287-4006, 125 pages. <http://www.public.asu.edu/~kirkwood/sysdyn/SDIntro/SDIntro.htm>
73. [Koenig *et al.*, 2007] Koenig W.D., Marcus L., Scott T.W., Dickinson J.L. (2007). West Nile virus and California breeding bird declines. *EcoHealth*, no. 1/mars 2007, vol. 4, pp. 18-24.
74. [Koestler, 1967] Koestler A. (1967). "The Ghost in the Machine". Hutchinson. (1990 reprint edition), Penguin Group. ISBN 0-14-019192-5. http://en.wikipedia.org/wiki/The_Ghost_in_the_Machine
75. [Komar *et al.*, 2003] Komar N., Langevin S., Hinten S., Nemeth N., Edwards E., Hettler D., Davis B., Bowen R., Bunning M. (2003). Experimental infection of North American birds with the New York 1999 strain of West Nile Virus. *Emerg. Infect. Dis.*, 9(3), pp. 311-322.
76. [Koné *et al.*, 2003] Koné P., Lambert L., Milord F., Gariépy C. (2003). Épidémiologie et effets de l'infection par le virus du Nil occidental sur la santé humaine. Direction de risques biologiques, environnementaux et occupationnels. Institut de santé publique du Québec. 74 pages. http://www.inspq.qc.ca/pdf/publications/265-Epidemiologie_Effets_VNO_Mise-a-jour.pdf
77. [Krauss *et al.*, 2003] Krauss H, Weber A., Appel M., Enders B., Isenberg H.D., Schiefer H.G., Slenczka W., von Graevenitz A., Zahner H. (2003). "Zoonoses: Infectious diseases transmissible from animals to humans". ASM Press, 3rd Edition. 474 pages. <http://estore.asm.org/viewItemDetails.asp?ItemID=318>
78. [Kuras, 2007] Kuras M. L. (2007). Complex-system engineering. Symposium on Complex Systems Engineering, 11-12 January 2007. <http://cs.calstatela.edu/wiki/images/c/c5/Kuras.pdf>
79. [Levins et Lewontin, 1985] Levins R., Lewontin R.C. (1985). "The Dialectical Biologist". Harvard University Press, 336 pages.
80. [Li *et al.*, 2012] Li S., Hartemink N., Speybroeck N., Vanwambeke S. (2012). Consequences of landscape fragmentation on Lyme disease risk: A cellular automata approach. *PLoS ONE*, 7(6) e39612. <http://www.ncbi.nlm.nih.gov/pubmed/22761842>
81. [Liu *et al.*, 2006a] Liu R., Shuai J., Zhu H., Wu J. (2006). Modeling spatial spread of West Nile virus and impact of directional dispersal of birds. *Math. Biosci. Eng.*, vol. 3, pp. 145-160.
82. [Liu *et al.*, 2006b] Liu Q-X, Jin Z., Liu M-X. (2006). Spatial organization and evolution period of the epidemic model using cellular automata. *Phys Rev E Stat Nonlin Soft Matter Phys.*, 74(3 Pt 1):031110. Epub 2006 Sep 13.
83. [Liu *et al.*, 2008] Liu T., Li X., Fu J., Ai B., Zhang X. (2008). A multi-agent based approach for simulating epidemic transmission. Geoinformatics 2008 and Joint Conference on GIS and Built Environment: Geo-Simulation and Virtual GIS Environments. In Proceedings of SPIE, vol. 7143 71431V-1.
84. [Mabrouk *et al.*, 2007] Mabrouk N., Deffuant G., Lobry C. (2007). Confronting macro, meso and micro scale modelling of bacteria dynamics. Third International Model-to-Model Workshop Marseille. <http://m2m2007.macaulay.ac.uk/M2M2007-Mabrouk.pdf>

85. [Madder *et al.*, 1983] Madder D.J., Surgeoner G.A., Helson B.B. (1983): Number of generations, egg production, and developmental time of *Culex pipiens* and *Culex restuans* in southern Ontario. *J Med Entomol*, 1983, vol. 20, pp. 275-287.
86. [Mandiau *et al.*, 2002] Mandiau R., Crislin-Le Strugeon E. (2002). "Systèmes multiagents". Édition TI - Sciences et Techniques, 16 pages.
87. [Mathieu *et al.*, 2007] Mathieu J., Melhuish J., James J., Mahoney P., Boiney L., White B. (2007). Multi-scale modeling of the air and space operations center. Symposium on Complex Systems Engineering, 11-12 January 2007. <http://cs.calstatela.edu/wiki/images/d/db/White.pdf>
88. [McMichael et Woodruff, 2008] McMichael A.J., Woodruff R.E. (2008). Climate change and infectious diseases. *The Social Ecology of Infectious Diseases*, pp. 378-407.
89. [Mekni *et al.*, 2008] Mekni M., Nabil S., Bernard M. (2008): A geosimulation approach involving spatially-aware agents a case study on the identification of risky areas for trains. In Proceedings of the 2008 Spring Simulation Multiconference 2008, pp. 37-44. <http://portal.acm.org/citation.cfm?id=1400555>.
90. [Mekni et Moulin, 2011] Mekni M., Moulin B. (2011). Informed virtual geographic environments: A geometrically precise and semantically enriched model for multi-agent geosimulations. Winter Simulation Conference, 2011, pp. 381-392.
91. [Minar *et al.*, 1996] Minar N., Burkhart R., Langton C., Askenazi M. (1996): The swarm simulation system: A toolkit for building multi-agent simulations. Swarm Development Group. http://en.wikipedia.org/wiki/Swarm_Development_Group
92. [Moulin *et al.*, 2013a] Moulin B., Bouden M., Navarro D. (2013): The Zoonosis-MAGS project (part 1): Population-based geosimulation of zoonoses in an informed virtual geographic environment. Chapter in the forthcoming book "Spatial and Temporal Dynamics of Infectious Diseases", D. Chen, B. Moulin, J. Wu (eds.), to be published by Wiley, 39 pages.
93. [Moulin *et al.*, 2013b] Moulin B., Navarro D., Marcotte D., Sedrati S., Bouden M. (2013). The Zoonosis-MAGS project (part 2): Complementarity of a rapid-prototyping tool and of a full-scale geosimulator for population-based geosimulation of zoonoses. Chapter in the forthcoming book "Spatial and Temporal Dynamics of Infectious Diseases", D. Chen, B. Moulin, J. Wu (eds.), to be published by Wiley, 27 pages.
94. [Moulin, 2012] Moulin B. (2012). Twelve years of geoid-sponsored research and development on multi-agent and population-based geo-simulation for decision support. Chapter in the book "The Added Value of Scientific Networking: Perspectives from the Geoid Network Members 1998-2012", Chrisman N. et Wachowics M. Eds. <http://www.gsdi.org/gsdiconf/gsdi13/abstracts.html#abMoulin>
95. [Moulin *et al.*, 2004] Moulin B., Chaker W., Gancet J. (2004). PADI-Simul: An agent-based geosimulation software supporting the design of geographic spaces. *Computers, Environment and Urban Systems*, 28(4), pp. 387-420. <http://www.sciencedirect.com/science/article/pii/S0198971503000632>
96. [Moulin *et al.*, 2003] Moulin B., Chaker W., Perron J., Pelletier P., Hogan J., Gbei E (2003). MAGS Project: Multi-agent geosimulation and crowd simulation. In Proceedings of the COSIT'03 Conference, Ittingen (Switzerland), Kuhn, Worboys and Timpf (eds.), Spatial Information Theory, Springer Verlag LNCS 2825, pp. 151-168.
97. [Moulin, 1985] Moulin B. (1985). La méthode EPAS, plan de systèmes et modélisation conceptuelle des données. Rapport de recherche DIUL-RR-8508, Université Laval, Département d'informatique, septembre 1985, 120 pages.

98. [Moulin, 1997] Moulin B. (1997). Temporal contexts for discourse representation: An extension of the conceptual graph approach. *Applied Intelligence*, vol. 7, 1997, pp. 227-225. <http://link.springer.com/article/10.1023%2FA%3A1008224616031>
99. [Mourits et Oude Lansink, 2006] Mourits M.C.M., Oude Lansink A.G.J.M. (2006). Multi-criteria decision making to evaluate quarantine disease control strategies. Chapter 9 in “New Approaches to the Economics of Plant Health”. Springer, pp.131-144.
100. [Murray, 1995] Murray G-M. (1995). What is complexity? *Complexity*, 1, vol. 1, no. 1, pp. 16-19. <http://complexity.martinsewell.com/Gell95.pdf>
101. [Nakayama, 2008] Nakayama M. K. (2008). Statistical analysis of simulation output. In Proceedings of the 2008 Winter Simulation Conference, pp. 62-72.
102. [Nasci *et al.*, 2001] Nasci R. S., Savage H. M., White D. J., Miller J.R., Cropp B.C., Godsey M. S., Kerst A.J., Bennett P., Gottfried K., Lanciotti R.S. (2001). West Nile virus in overwintering *Culex* mosquitoes, New York City, 2000. *Emerging Infect. Dis.*, vol. 7, pp. 742-744. <http://www.ncbi.nlm.nih.gov/pubmed/11585542>
103. [Navarro, 2013] Navarro V.D. (2013). “Geosimulation of the dynamics of ecological communities”. Mémoire de maîtrise, Université Laval, Québec, Canada.
104. [Noël, 2007] Noël P.A. (2007). “Dynamique de modèles épidémiologiques: Applications au cas du virus du Nil occidental”. Mémoire de maîtrise, Université Laval, Québec, Canada. 150 pages.
105. [Ogden *et al.*, 2008a] Ogden N.H., St-Onge L, Barker I.K., Brazeau S., Bigras-Poulin M., Charron D.F., Francis C.M., Heagy A., Lindsay L.R., Maarouf A., Michel P., Milord F., O’Callaghan C.J., Trudel L., Thompson R.A. (2008). Risk maps for range expansion of the Lyme disease vector, *Ixodes scapularis*, in Canada now and with climate change. *International Journal of Health Geographics*, 22, 7:24.
106. [Ogden *et al.*, 2008b] Ogden N.H., Lindsay L.R., Hanincová K., Barker I.K., Bigras-Poulin M., Charron D.F., Heagy A., Francis C.M., O’Callaghan C.J., Schwartz I., Thompson R.A. (2008). Role of migratory birds in introduction and range expansion of *Ixodes scapularis* ticks and of *Borrelia burgdorferi* and *Anaplasma phagocyte-philum* in Canada. *Appl Environ Microbiol*. 2008 Mar;74(6):1780-90. <http://www.ncbi.nlm.nih.gov/pubmed/18498647>
107. [Ogden *et al.*, 2007] Ogden N.H., Bigras-Poulin M., O’Callaghan C.J., Barker I.K., Kurtenbach K., Lindsay L.R., Charron D.F. (2007). Vector seasonality, host infection dynamics and fitness of pathogens transmitted by the tick *Ixodes scapularis*. *Parasitology*, vol. 134, pp. 209-227.
108. [Ogden *et al.*, 2006] Ogden N.H., Maarouf A., Barker I.K., Bigras-Poulin M., Lindsay L.R., Morshed M.G., O’callaghan C.J., Ramay F., Waltner-Toews D., Charron D.F. (2006). Climate change and the potential for range expansion of the Lyme disease vector *Ixodes scapularis* in Canada. *Int J Parasitol*, 2006, 36(1), pp. 63-70. <http://www.ncbi.nlm.nih.gov/pubmed/16229849>
109. [Ogden *et al.*, 2005] Ogden N.H., Bigras-Poulin M., O’Callaghan C.J., Barker I.K., Lindsay L.R., Maarouf A., Smoyer-Tomic K.E., Waltner-Toews D., Charron D. (2005). A dynamic population model to investigate effects of climate on geographic range and seasonality of the tick *Ixodes scapularis*. *Int. J. Parasitol*, 35(4), pp. 375-389. <http://www.ncbi.nlm.nih.gov/pubmed/15777914>
110. [O’Sullivan et Torrens, 2000] O’Sullivan D., Torrens P., (2000). Cellular models of urban systems. In Bandini, S & Worsch, T (eds) 2000, Theoretical and Practical Issues on Cellular Automata. In Proceedings of the Fourth International Conference on Cellular Automata for Research and Industry (ACRI 2000), Karlsruhe, Germany, pp. 108-116.

- 111.[Paris *et al.*, 2009] Paris S., Mekni M., Moulin B. (2009) Informed virtual geographic environments: an accurate topological approach. The International Conference on Advanced Geographic Information Systems & Web Services (GEOWS).pp.1-6, IEEE Computer Society Press, 2009.
- 112.[Pereira, 2002] Pereira G. M. (2002). A typology of spatial and temporal scale relations. *Geographical Analysis*, vol. 34, no. 1 (January 2002), The Ohio State University. http://muse.jhu.edu/demo/geographical_analysis/v034/34.1pereira.pdf
- 113.[Parrott, 2002] Parrott L. (2002). Complexity and the Limits of Ecological Engineering. *American Society of Agricultural Engineers*, vol. 45(5). <http://complexity.ok.ubc.ca/files/2013/04/be34final.pdf>
- 114.[Pouvreau, 2013] Pouvreau D. (2013). “Une histoire de la "systémologie générale" de Ludwig von Bertalanffy - Généalogie, genèse, actualisation et postérité d'un projet herméneutique”. Thèse de doctorat en Sciences Sociales. <http://tel.archives-ouvertes.fr/tel-00804157>
115. [Rahmandad et Sterman, 2006] Rahmandad H., Sterman J. (2006). Heterogeneity and network structure in the dynamics of diffusion: Comparing agent-based and differential equation models. MIT Sloan School of Management, Cambridge MA 02142.
- 116.[Reeves, 1983] Reeves W. T. (1983). Particle Systems – a technique for modeling a class of fuzzy objects. *ACM Trans. Graphics*, vol. 2, pp. 91–108.
- 117.[Regniere et Saint-Amant, 2008] Regniere J, Saint-Amant R. (2008). BioSIM 9: A computer-based decision support tool for seasonal planning of pest management activities. User's Manual. Information Report LAU-X-134. 68 pages.
- 118.[Reisen *et al.*, 1991] Reisen W.K., Milby M.M., Meyer R.P., Pfuntner A.R., Spoehel J., Hazelrigg , Webb J.P. Jr (1991). Mark-release-recapture studies with *Culex* Mosquitoes (Diptera: Culicidae) in Southern California. *J Med Entomol*, 1991, 28(3), pp. 357-371.
- 119.[Reisen *et al.*, 1992] Reisen W.K., Marilyn M.M., Meyer R. (1992). Population dynamics of adult *Culex* Mosquitoes (Diptera: Culicidaw) along the Ken river, Kern County, California, in 1990. *J Med Entomol*, 992, 29(3), pp. 531-543.
- 120.[Reynolds, 1987] Reynolds C. W. (1987). Flocks, herds, and schools: A distributed behavioral model. In Proceedings of SIGGRAPH '87, the 14th annual conference on Computer graphics and interactive techniques, pp. 25-34.
- 121.[Rodriguez *et al.*, 2006] Rodriguez S., Hilaire V., Koukam A. (2006). Holonic modeling of environments for situated multi-agent systems. *Environments for Multi-Agent Systems II*, Springer Verlag, Lecture Notes in Computer Science, vol. 3830, 2006, pp. 18-31. <http://www.springerlink.com/content/vv68px430n2k5431/>
- 122.[Rosnay, 1975] Rosnay J.D. (1975). Le Macroscopie : Vers une vision globale. Editions du Seuil, 1975, ISBN 2-02004-567-2. Prix de l'Académie des sciences morales et politiques 1975.
- 123.[Roy, 1985] Roy B. (1985). Méthodologie multicritère d'aide à la décision. *Economica*, Collection Gestion. Série Production et techniques quantitatives appliquées à la gestion, 423 pages.
- 124.[Sahli et Moulin, 2007] Sahli N., Moulin B. La géo-simulation multi-agent : Un support pour la planification dans un environnement réel. *Revue d'Intelligence Artificielle*, 21(2), pp. 151-182. <http://ria.revuesonline.com/article.jsp?articleId=9768>
- 125.[Sander *et al.*, 1998] Sander J., Ester M., Kriegel H.P., Xu X. (1998). Density-based clustering in spatial databases: The algorithm GDBSCAN and its applications. *Data Mining Knowledge Discovery*. 2(2), pp. 169-194.

126. [Sawyer, 2005] Sawyer R.K. (2005). "Social emergence: Societies as complex systems". Cambridge University Press, 383 pages.
127. [Schieritz et Größler, 2003] Schieritz N., Größler A. (2003). Emergent structures in supply chains: A study integrating agent-based and system dynamics modelling. In Proceedings of the 36th Annual Hawaii International Conference on System Sciences, Washington. <http://csdl.computer.org/comp/proceedings/hicss/2003/1874/03/187430094a.pdf>
128. [Schieritz et Milling, 2003] Schieritz N., Milling P.M. (2003). Modelling the forest or modelling the trees - a comparison of system dynamics and agent-based simulation. In Proceedings of the 21st International Conference of the System Dynamics Society. Hrsg., Robert L. Eberlein *et al.*. New York, 2003, S. <http://is.bwl.uni-mannheim.de/Forschung/pr/sd03/p-na.pdf>
129. [Scholl, 2001] Scholl H. J. (2001). Agent-based and System Dynamics modelling: A call for cross study and joint research. In Proceedings of the 34th Annual Hawaii.
130. [Schüle *et al.*, 2004] Schüle, M., Herrler, R., Klugl, F. (2004). Coupling GIS and multi-agent simulation – towards infrastructure for realistic simulation. Lindemann-v. Trzebiatowski, G.; Denzinger, J.; Timm, I.J.; Unland, R. (Eds.). Chapter in "Multiagent System Technologies", Springer Verlag, Lecture Notes in Computer Science, vol. 3187, pp. 228-242.
131. [Sedrati, 2013] Sedrati S. (2013). "Un outil de spécification et d'évaluation efficace des expressions mathématiques des modèles épidémiologiques pour la simulation de zoonoses". Mémoire de maîtrise de l'Université Laval, Québec, Canada, 216 pages.
132. [Shi et Pang, 2000] Shi W., Pang, M.Y.C. (2000). Development of Voronoi-based cellular automata—an integrated dynamic model for Geographical Information Systems. *International Journal of Geographical Information Science*, vol. 14 (5), pp. 455-474.
133. [Shillo et Fischer, 2003] Schillo M., Fischer K. (2003) Holonic multiagent systems. A foundation for the organisation of multiagent systems. In Proceedings of the First International Conference on Applications of Holonic and Multiagent Systems (HoloMAS'03). <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.14.1097>
134. [Sichman *et al.*, 1998] Sichman J.S., Conte R., Gilbert N. (1998). Multi-agent systems and agent-based simulation (MABS). First International Workshop, MABS '98, Paris, France.
135. [Stemate et Taylor, 2007] Stemate L., Taylor I. (2007). A comparison between system dynamics and agent based modeling and opportunities for cross-fertilization. Submitted to the 2007 Winter Simulation Conference. December 9-12, 2007. Washington, D.C.
136. [Stermann, 2000] Stermann J. (2000). "Business Dynamics: Systems thinking and modelling for a complex world". Irwin/McGraw-Hill, ISBN 0-07-231135-5. <http://www.amazon.fr/Business-Dynamics-Systems-Thinking-Modeling/dp/007238915X>
137. [Tardif *et al.*, 2003] Tardif I., Bolduc D.G., St-Laurent L., Samuel O., Pinsonneault L., Chevalier P. (2003). Pertinence et faisabilité, en 2003, d'un programme préventif de réduction du risque de transmission du virus du Nil occidental avec des larvicides. Direction de risques biologiques, environnementaux et occupationnels. Institut de santé publique du Québec. 55 pages. http://www.inspq.qc.ca/pdf/publications/263-VNO_Larvicides2003.pdf
138. [Torrens, 2006] Torrens P.M. (2006). Geosimulation and its application to urban growth modeling. *Complex Artificial Environments*, J. Portugali (Eds.), London: Springer-Verlag, pp. 119-134.
139. [Tran *et al.*, 2005] Tran A., Biteau-Coroller F., Guis H., Roger F. (2005). Modélisation des maladies vectorielles. *Epidémiol. et santé anim.*, 2005, vol. 47, pp. 35-51.

140. [Van Dyke Parunak *et al.*, 1998] Van Dyke Parunak H., Savit R., Riolo R.L. (1998). Agent-based modeling vs. equation-based modeling: A case study and users' guide. In Proceedings of Multi-Agent Systems and Agent-Based Simulation (MABS'98), 10-25, Springer, LNAI 1534.
141. [Vansnick, 1990] Vansnick J.C. (1990). Measurement theory and decision aid. Bana e Costa (ed.), Readings in Multiple Criteria Decision Aid, Springer-Verlog, Berlin, pp. 81-100.
142. [Vincent *et al.*, 2004] Vincent C., Desjardins F., Tanguay M. (2004). Virus du Nil occidental : État de situation 2004. *Bulletin zoosanitaire*. Réseau d'alerte et d'information zoosanitaire. Institut national de santé animale, no. 42, 21 mai 2004, 4 pages.
143. [Volle, 2003] Volle M. (2003). Complexité et complication. *La revue Génie Logiciel*, no. 64, mars 2003, pp. 49-58. <http://cat.inist.fr/?aModele=afficheN&cpsidt=14790698>
144. [Weisbuch, 1989] Weisbuch G. (1989). Dynamique des systèmes complexes. éditions du CNRS/InterEditions, 1989. <http://www.lps.ens.fr/~weisbuch/livre/livesd.html>
145. [White *et al.*, 2000] White R., Engelen G., Uljee I. (2000). Modelling land use change with linked cellular automata and socio-economic models: A tool for exploring the impact of climate change on the island of St Lucia. *Spatial Information for Land Use Management*. p. 189-204.
146. [White *et al.*, 2009] White S.H., Martin del Rey A., Rodriguez Sanchez G. (2009). Using cellular automata to simulate epidemic diseases. *Applied Mathematical Sciences*, vol. 3, 2009, no. 20, pp. 959-968.
147. [Wilson, 2008] Wilson C. (2008). Activity patterns in space and time: Calculating representative Hagerstrand trajectories. *Transportation*, vol. 35, issue 4, pp. 485-499. <http://link.springer.com/article/10.1007%2Fs11116-008-9162-z>
148. [Wiraningsih *et al.*, 2010] Wiraningsih E. D., Widodo, Aryati L., Toaha S., Lenhart S. (2010). Optimal control for SEIR Rabies model between dogs and human with vaccination effect in dogs. In Proceedings of the 6th IMT-GT Conference on Mathematics, Statistics and its Applications (ICMSA2010), pp. 1161-1175.
149. [Wolfram, 2002] Wolfram S. (2002). Continuous versus discrete systems, Chapter 4 in "A New Kind of science", Wolfram S. Ed., Publisher Wolfram Media Inc. 1125 pages.
150. [Wolfram, 1984] Wolfram S. (1984). Universality and complexity in cellular automata. *Physica D*, no.10 (Janvier 1984), pp. 1-35. <http://www.stephenwolfram.com/publications/articles/ca/84-universality/>
151. [Wonham *et al.*, 2004] Wonham M.J., De-Camino-Beck T., Lewis M.A. (2004). An epidemiological model for west Nile virus: Invasion analysis and control applications. *Proc R Soc Lond B Biol Sci*, 2004 Mar., 7, 271(1538), pp. 501-507. <http://www.math.ualberta.ca/~mathbio/publications/Wonham.etal.2004.pdf>
152. [Wooldridge, 2002] Wooldridge M. (2002). "An introduction to multiagent systems". Published by John Wiley & Sons (Chichester, England), 340 pages. <http://www.cs.ox.ac.uk/people/michael.wooldridge/pubs/imas/IMAS2e.html>
153. [Zeigler *et al.*, 2000] Zeigler B.P., Praehofer H., Kim T.G. (2000). "Theory of modeling and simulation: Integrating discrete event and continuous complex dynamic systems". Academic Press. 517 pages.

Webliographie

1. [Site Web : *ASPC*] Agence de la santé publique du Canada (ASPC). <http://www.phac-aspc.gc.ca/wnv-vwn/index.html>
2. [Site Web: *AnyLogic*] AnyLogic™ simulation tool. <http://www.xjtek.com/anylogic/>
3. [Site Web: *CDCP*] Centers for Disease Control and Prevention. <http://www.cdc.gov/>
4. [Site Web: *CDM*] Centre for Disease Modelling. <http://www.cdm.yorku.ca/>
5. [Site Web: *CODIGEOSIM*] Geosimulation tools for simulating spatial-temporal spread patterns and evaluating health outcomes of communicable diseases. <http://gis.geog.queensu.ca/CODIGEOSIM/>
6. [Site Web: *Complex Systems*] Complex Systems. <http://www.complex-systems.com/>
7. [Site Web: *CSCS*] The Center for the Study of Complex Systems. <http://www.lsa.umich.edu/cscs/>
8. [Site Web: *DG8LCMI*] Déclaration du G8 sur la lutte contre les maladies infectieuses. <http://www.g8.utoronto.ca/francais/2006stpetersbourg/maladies.html>
9. [Site Web: *eBird*] Uniting birders around the world. <http://ebird.org>
10. [Site Web: *ECSS*] European Complex Systems Society. <http://www.complexssociety.eu/>
11. [Site Web: *Environment Canada*] Canadian climate normals or averages 1971–2000. http://climate.weatheroffice.gc.ca/climate_normals/index_e.html
12. [Site Web: *Geobase*] Spatial GEOBASE. <http://www.geobase.ca/>
13. [Site Web: *Geomedia*] GeoMedia: GIS management platform. <http://geospatial.intergraph.com/products/GeoMedia/Details.aspx>
14. [Site Web: *Geosimulation*] Geosimulation Web Site. <http://www.geosimulation.org/geosim/>
15. [Site Web: *GéoStat*] Centre GéoStat de la bibliothèque de l'Université Laval. <http://www.bibl.ulaval.ca/services/centregeostat/donnees-geospaciales>
16. [Site Web: *GLEaMviz*] Global Epidemic and Mobility Model Visualization. <http://www.gleamviz.org/>
17. [Site Web: *GREZOSP*] Groupe de Recherche en Épidémiologie des Zoonoses et Santé Publique. <http://www.medvet.umontreal.ca/grezosp>
18. [Site Web: *INMA*] Institut National de Médecine Agricole : Les principales zoonoses. <http://inma.fr/fr-162-les-principales-zoonoses.html>
19. [Site Web: *LANL*] Los Alamos National Laboratory. <http://www.lanl.gov/>
20. [Site Web: *Mapnik*] Mapnik is a Free Toolkit for developing mapping applications. <http://mapnik.org/>
21. [Site Web: *MAS*] Multiagent Systems and Agent-Based Modeling. <http://www.multiagent.com/>
22. [Site Web: *MIT*] MIT Sloan School of Management. <http://mitsloan.mit.edu/>

23. [Site Web: **MUSCAMAGS**] The MUSCAMAGS Project.
<http://www.ift.ulaval.ca/~muscamags/>
24. [Site Web: **OCaml**] Langage orienté objet qui se base sur le paradigme fonctionnel.
<http://caml.inria.fr/ocaml/index.fr.html>
25. [Site Web: **OMG**] Object Management Group (OMG), UML Resource Page.
<http://www.uml.org/>
26. [Site Web: **OMS**] Organisation Mondiale de la Santé <http://www.who.int/fr/>
27. [Site Web: **PostGIS**] PostGIS: Spatial database extender for PostgreSQL object-relational database. <http://postgis.net/>
28. [Site Web: **Powersim**] Powersim Software: Business intelligence and strategic business simulations. <http://www.powersim.no/>
29. [Site Web: **Saga**] SAGA - System for Automated Geoscientific Analyses.
<http://sourceforge.net/projects/saga-gis/>
30. [Site Web: **Science**] Science Magazine. <http://www.sciencemag.org/>
31. [Site Web: **SDS**] System dynamics society. <http://www.systemdynamics.org/>
32. [Site Web: **SFI**] The Santa Fe Institute is the world center for the study of complexity.
<http://www.santafe.edu/index.html>
33. [Site Web: **SOLAP**] Spatial OLAP. <http://www.spatialbi.com/>
34. [Site Web: **SOPFIM**] La société de protection des forêts contre les insectes et maladies. <http://www.sopfim.qc.ca/>
35. [Site Web: **StatCan**] Statistique Canada. <http://www.statcan.gc.ca/>
36. [Site Web: **Stella**] Stella: Systems Thinking for Education and Research.
<http://www.iseesystems.com/software/Education/StellaSoftware.aspx>
37. [Site Web: **STEM**] The Spatiotemporal Epidemiological Modeler.
<http://www.eclipse.org/stem/>
38. [Site Web: **Wikipedia**] Wikipedia, l'encyclopédie libre. <http://www.wikipedia.org/>
39. [Web Site: **Vensim**] Vensim Simulation Software: developing, analyzing, and packaging high quality dynamic feedback models.
<http://www.vensim.com/software.html>

ANNEXES

Annexe A

Formalisme de MASTIM

Nous présentons dans le tableau suivant les plus importantes notations du formalisme employé afin de retrouver plus rapidement une notation particulière.

Notation	Explication
Σ_k	Un système écologique
$C^{\Sigma_k} = \bigcup_{i=1}^m \{C_1^{H_i}, C_2^{H_i}, C_3^{H_i}, \dots, C_{n_i}^{H_i}\}$	Un ensemble de cellules se trouvant dans plusieurs niveaux hiérarchiques
$C^{\Sigma_k}(H_i) = \{C_1^{H_i}, C_2^{H_i}, \dots, C_{n_i}^{H_i}\}$	Un ensemble de cellules se trouvant dans le niveau hiérarchique i
$C_j^{H_i}$	Une cellule j d'un niveau hiérarchique i
$V^{C_j^{H_i}} = (\{C_k^{H_i}\}, Constraints)$	Les voisins d'une cellule $C_j^{H_i}$
$G_j^{H_i} = (Id_{G_j^{H_i}}, Type_{G_j^{H_i}}, C_p^{H_i})$	Une région représentant une cellule qui est définie à une certaine hiérarchie i
$cG_{k<i}^{H_k} = (G_j^{H_i}, \{G_1^{H_k}, G_2^{H_k}, \dots, G_m^{H_k}\})$	Une projection de la région $G_j^{H_i}$ qui est un conteneur représentant la décomposition de la région mère en un ensemble de régions filles.
$cG_{k<i}^{H_k} = (G_j^{H_i}, \{C_1^{H_k}, C_2^{H_k}, \dots, C_m^{H_k}\})$	Une projection de la région $G_j^{H_i}$ qui est un conteneur représentant la décomposition de la région mère en un ensemble de cellules
$Z_j^{H_i} = (Id_{Z_j^{H_i}}, P1_{Z_j^{H_i}}, P2_{Z_j^{H_i}}, Constraints_{Z_j^{H_i}})$	Une zone sélectionnée qui est une zone rectangulaire faisant partie de l'environnement de la simulation et que l'utilisateur a sélectionnée afin de lui appliquer un certain traitement
$S^{\Sigma_k} = \{S_1, S_2, \dots, S_n\}$	Un ensemble d'espèces qui interagissent entre elles
$B^{S_k} = \left\{ \begin{array}{l} \{(LType_1, \dots, LType_n, MinBSD), \dots, \\ (LType_1, \dots, LType_m, MaxBSD)\} \\ MinTemp, MaxTemp \\ MinPrecipitations, \\ MaxPrecipitations \end{array} \right\}$	L'habitat favorable de l'espèce S_k
$O(S_k) = \{O_1^{S_k}, O_2^{S_k}, \dots, O_n^{S_k}\}$	Un ensemble de compartiments types d'une espèce S_k

$R(O_p^{S_k}) = \{R_1^{O_p^{S_k}}, R_2^{O_p^{S_k}}, \dots, R_m^{O_p^{S_k}}\}$	Un ensemble de réservoirs types (Stocks) d'un compartiment type $O_p^{S_k}$
$A_k^{\Sigma_k} = (Id_{A_k^{\Sigma_k}}, V_{A_k^{\Sigma_k}})$	Un attribut de système
$A_k^{C_j^{H_i}} = (Id_{A_k^{C_j^{H_i}}}, V_{A_k^{C_j^{H_i}}})$	Un attribut de cellule
$A_k^{C_j^{H_i} S_m} = (Id_{A_k^{C_j^{H_i} S_m}}, V_{A_k^{C_j^{H_i} S_m}})$	Un attribut d'espèce
$A_k^{C_j^{H_i} O_n^{S_m}} = (Id_{A_k^{C_j^{H_i} O_n^{S_m}}}, V_{A_k^{C_j^{H_i} O_n^{S_m}}})$	Un attribut de compartiment
$ET(S_k) = \{ET_{i,j}^{S_k}, ET_{p,l}^{S_k}, \dots, ET_{g,c}^{S_k}\}$	Un ensemble de transitions types d'une espèce S_k
$ET_{s,d}^{S_k} = (O_s^{S_k}, O_d^{S_k}, \{E_i\}, \{N_j\}, FR^{ET_{s,d}^{S_k}})$	Une transition type de $O_s^{S_k}$ au $O_d^{S_k}$ se trouvant dans une même cellule (<i>Evolutionary Transition</i>)
$FR^{R_1^{O_s^{S_k}} R_1^{O_d^{S_k}}} = \{Absorption(R_1^{O_s^{S_k}}, \alpha_1, \beta_1, \dots, \omega_1), Production(R_1^{O_d^{S_k}}, \alpha_2, \beta_2, \dots, \omega_2)\}$	Un débit (flow rate) permettant l'absorption des individus du réservoir source et production des individus dans le réservoir destination
$MT^{O_m^{S_p}} = (O_k^{S_p}, \{E_i\}, \{N_j\}, FR^{MT^{O_m^{S_p}}})$	Une transition de mortalité (<i>Mortality Transition</i>)
$CT_{s,d}^{O_m^{S_p}} = (O_k^{S_p C_s^{H_i}}, O_k^{S_p C_d^{H_i}}, \{E_i\}, \{N_j\}, FR^{CT_{s,d}^{O_m^{S_p}}})$	Une transition entre compartiments de même type, mais de cellules différentes (<i>Cells' Transition</i>)
$E_i = (Id_{E_i}, D_i, \{N_k\}, [D_f, Periodicity])$	Un Événement
N_i	Une condition p.ex. $value(A_k^{\Sigma_k}) \geq 5 \text{ min}$
$Interaction_{S_k S_p} = \{I_{O_i^{S_k} O_j^{S_p}}, I_{O_p^{S_k} O_l^{S_p}}, \dots, I_{O_g^{S_k} O_c^{S_p}}\}$	Un ensemble d'interactions pertinentes entre les différents compartiments types des deux espèces
$I_{O_i^{S_k} O_l^{S_p}} = (Id_{I_{O_i^{S_k} O_l^{S_p}}}, Probi_{O_i^{S_k} O_l^{S_p}}, [\{ET_{1,j}^{S_k}, \dots, ET_{1,c}^{S_k}\}, \{ET_{1,v}^{S_p}, \dots, ET_{1,z}^{S_p}\}], [Drop_{O_i^{S_k} O_l^{S_p}}, \{CT_{s,d}^{O_i^{S_k}}, CT_{s,d}^{O_l^{S_p}}\}])$	L'interaction entre les deux premiers compartiments des deux espèces S_k et S_p
$P_m = (Id_{P_m}, Type_{P_m}, \{G_j^{H_i}\}, \{Z_s^{H_i}\}, \{E_d\}, \Delta_{P_m}, f_{P_m}, \{g_k(A_j)\})$	Un processus
$MobilityPattern_m^{S_p} = (Id_{MP_m^{S_p}}, \{StartingAreaType - StartingPeriodType - ActivityType - TranscientAreaType - DestinationAreaType - DestinationPeriodType\}, Parameters_{MP_m^{S_p}})$	Un patron de mobilité

$Trajectory_m^{S_p} = (Id_{Trajectory_m^{S_p}}, \{C_{src}^{H_i}, \dots, C_{dest}^{H_i}\},$ $MobilityPattern_n^{S_p}, Extent_{Trajectory_m^{S_p}})$	Une trajectoire
$TA_k^{C_j^{H_i} S_p} = (Id_{TA_k^{C_j^{H_i} S_p}}, \{(Id_{Trajectory_m^{S_p}},$ $TOrientation_m^{C_j^{H_i}}, TProportion_m^{C_j^{H_i}})\})$	Un attribut de la trajectoire
$MobilityNetwork_m^{S_p} = (Id_{Network_m^{S_p}},$ $\{Trajectory_1^{S_p}, Trajectory_2^{S_p}, \dots\})$	Un réseau de mobilité
$Group_m^{S_j} = (Id_{Group_m^{S_j}}, Nb_{Group_m^{S_j}},$ $\{C_1^{H_i}, \dots, C_n^{H_i}\}, StartingPeriod_{Group_m^{S_j}})$	Un groupe d'individus
$Displacement_t^{S_p} = (Id_{Displacement_t^{S_p}},$ $Group_u^{S_p}, C_s^{H_i}, C_d^{H_i})$	Un déplacement
$Flow_u^{S_p} = (Id_{Flow_u^{S_p}}, Group_m^{S_p}, \{Displacement_t^{S_p}\},$ $MobilityPattern_n^{S_p}, MobilityNetwork_r^{S_p}, \{P_r\})$	Un flot d'individus
$State_j^{\Sigma_k} = \{A_j^{\Sigma_k}, A_j^{G_m^{H_i}}, A_j^{Z_m^{H_i}}, A_j^{C_m^{H_i}}, A_j^{C_m^{H_i} S_p}, A_j^{C_m^{H_i} O_n^{S_p}}, \dots\}$	Un état du système
$ScenarioPattern_m^{Zoonosis_p} = (Id_{SP_m^{Zoonosis_p}}, \Sigma_k,$ $(SimulationModel_k, ModelParameters_k), \dots)$	Un patron de scénario
$Scenario_j = (Id_{Scenario_j}, ScenarioPattern_m^{Zoonosis_p},$ $\Delta_{Scenario_j}, \{RefData_n^{Zoonosis_p}\}, \{E_i\}, \dots)$	Un scénario de la simulation
$Simulation_m^{Zoonosis_p} = (Id_{Sim_m^{Zoonosis_p}}, Scenario_j,$ $SimStep_m, floop(State_n^{\Sigma_k}), SimParam_m)$	Une simulation

Annexe B

Représentation graphique plus complète pour la maladie de Lyme en intégrant les différentes équations différentielles utilisées par Ogden [Ogden *et al.*, 2005].

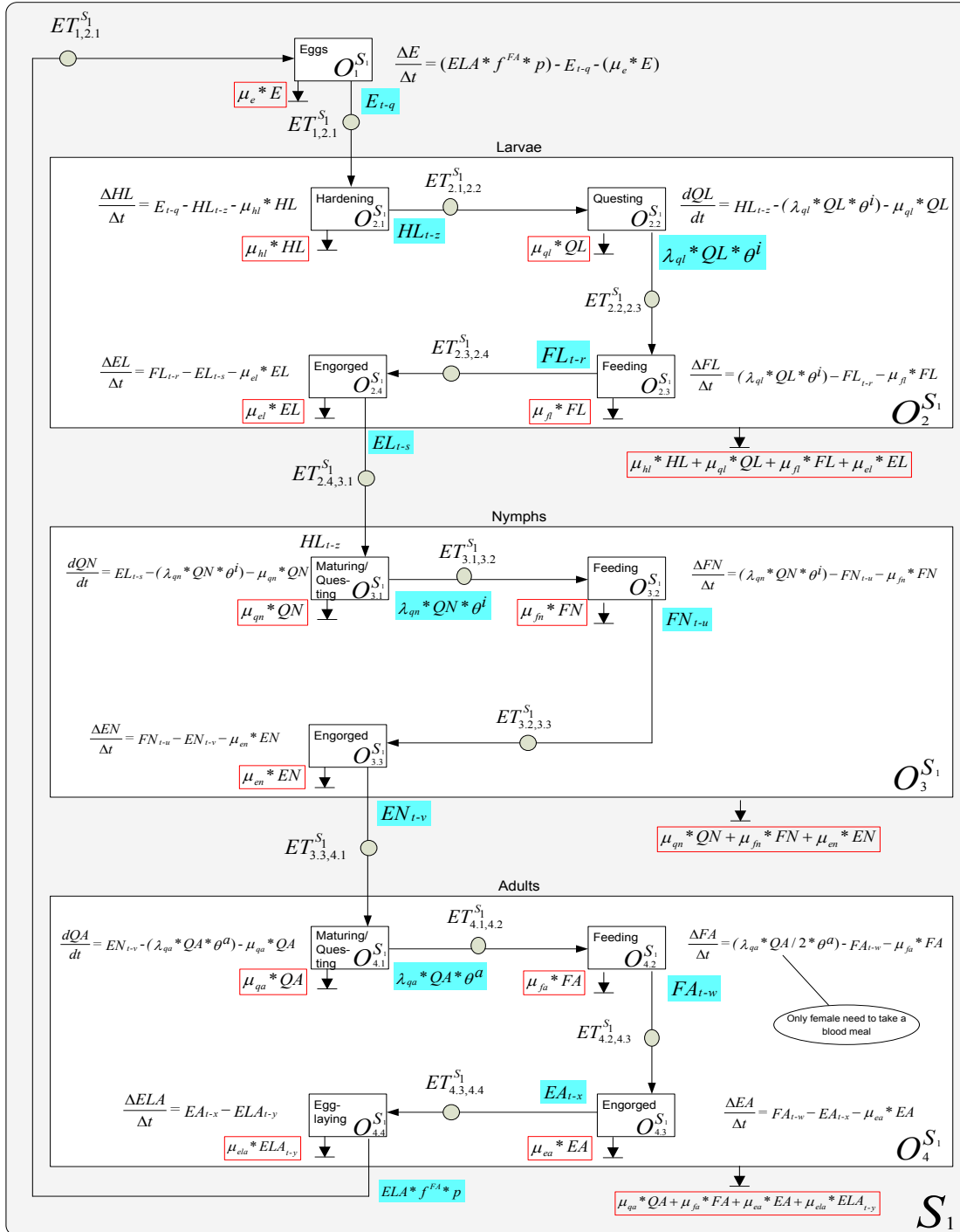


Figure B1 : Modèle biologique des tiques sans maladie.

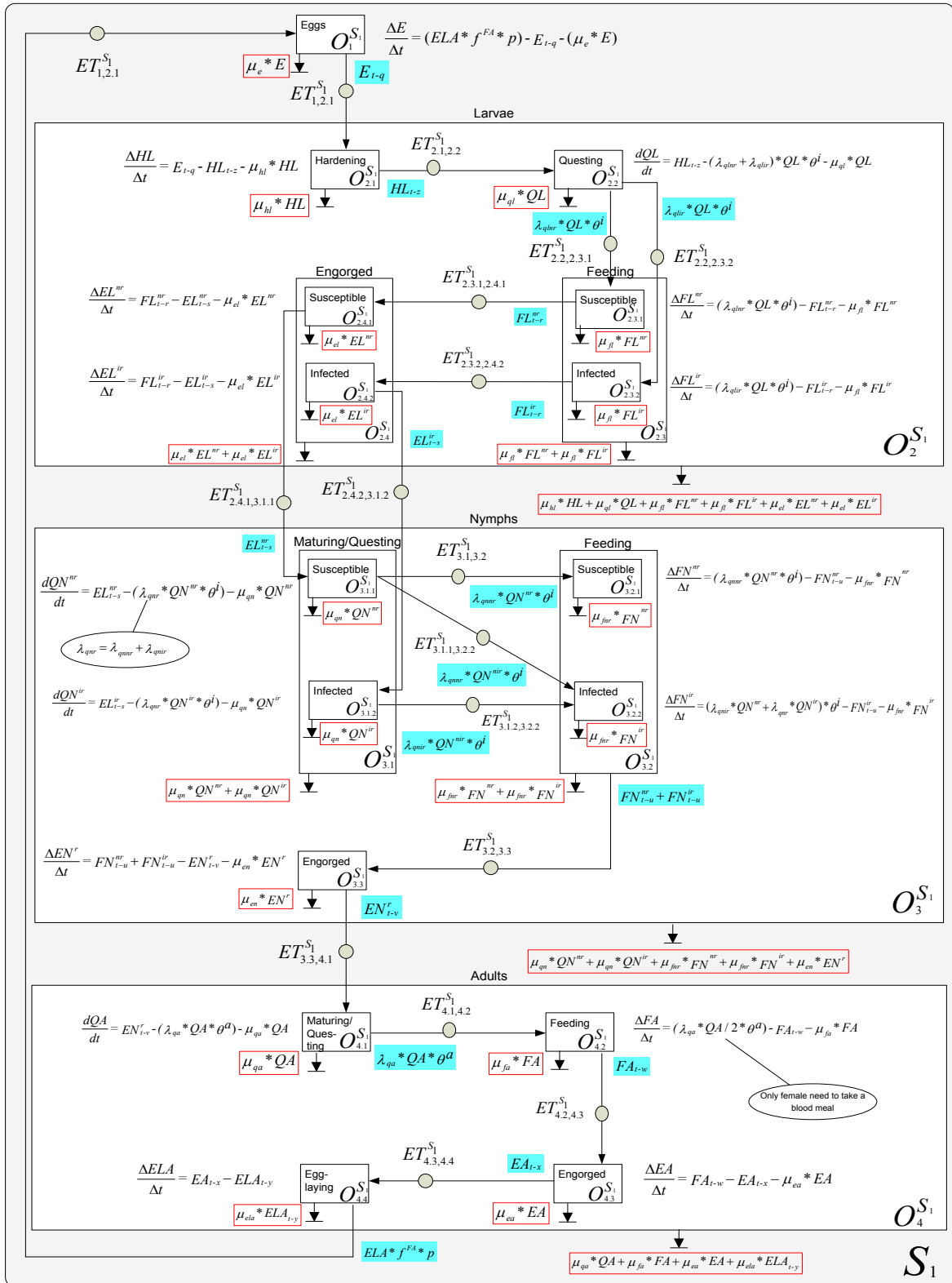


Figure B2 : Interactions des tiques avec les rongeurs.

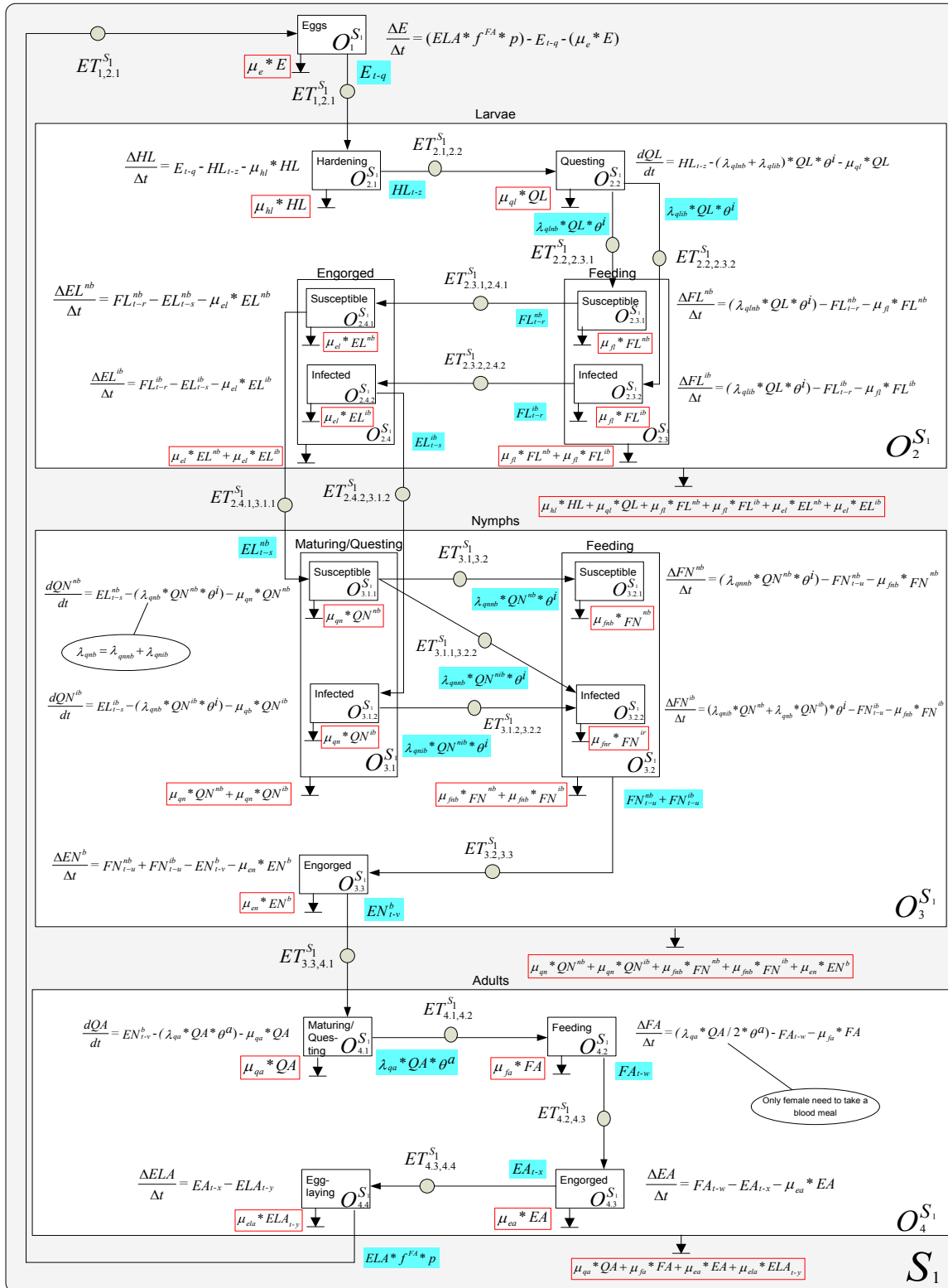


Figure B3 : Interactions des tiques avec les oiseaux.

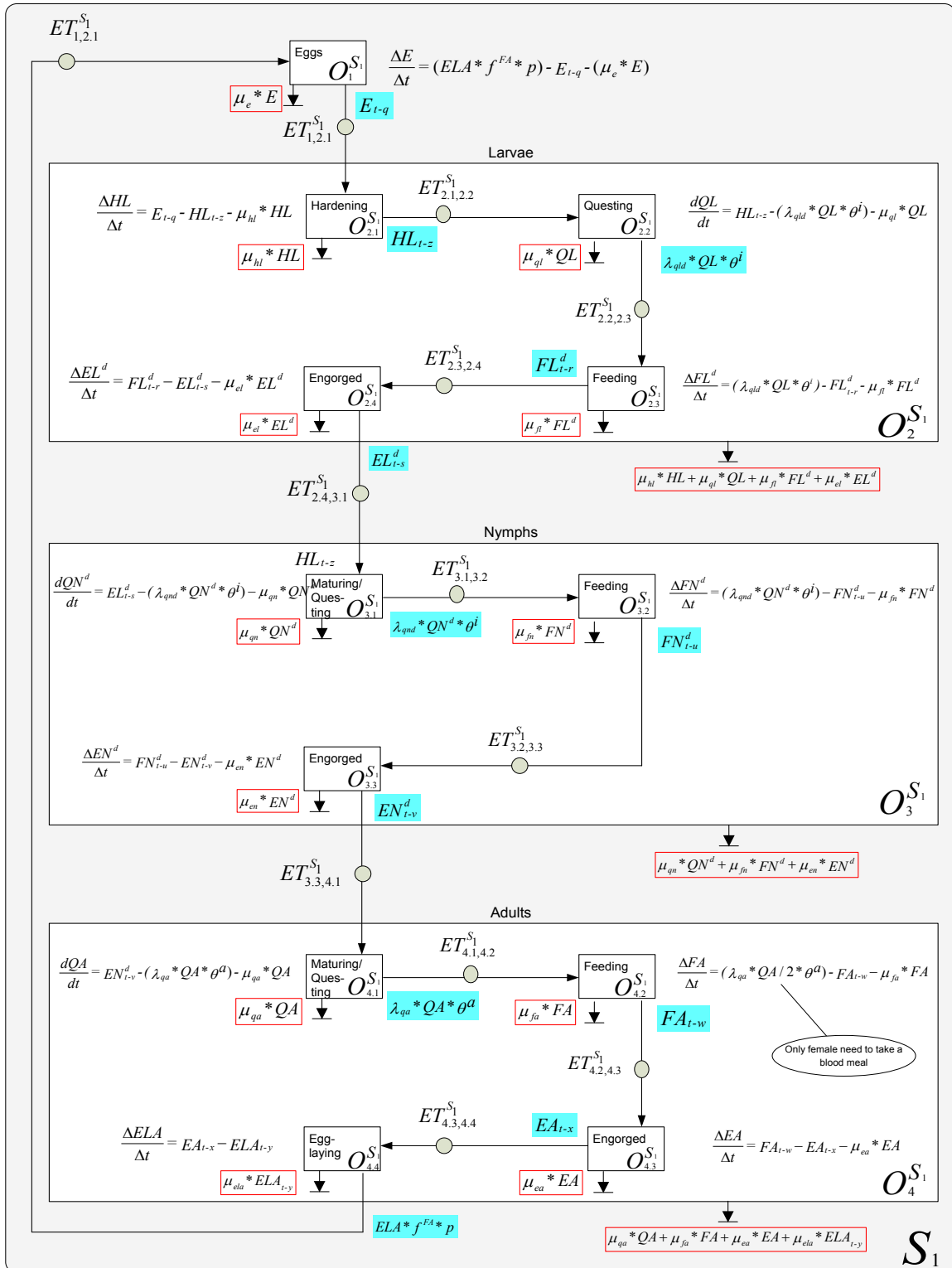


Figure B4 : Interactions des tiques avec les chevreuils.

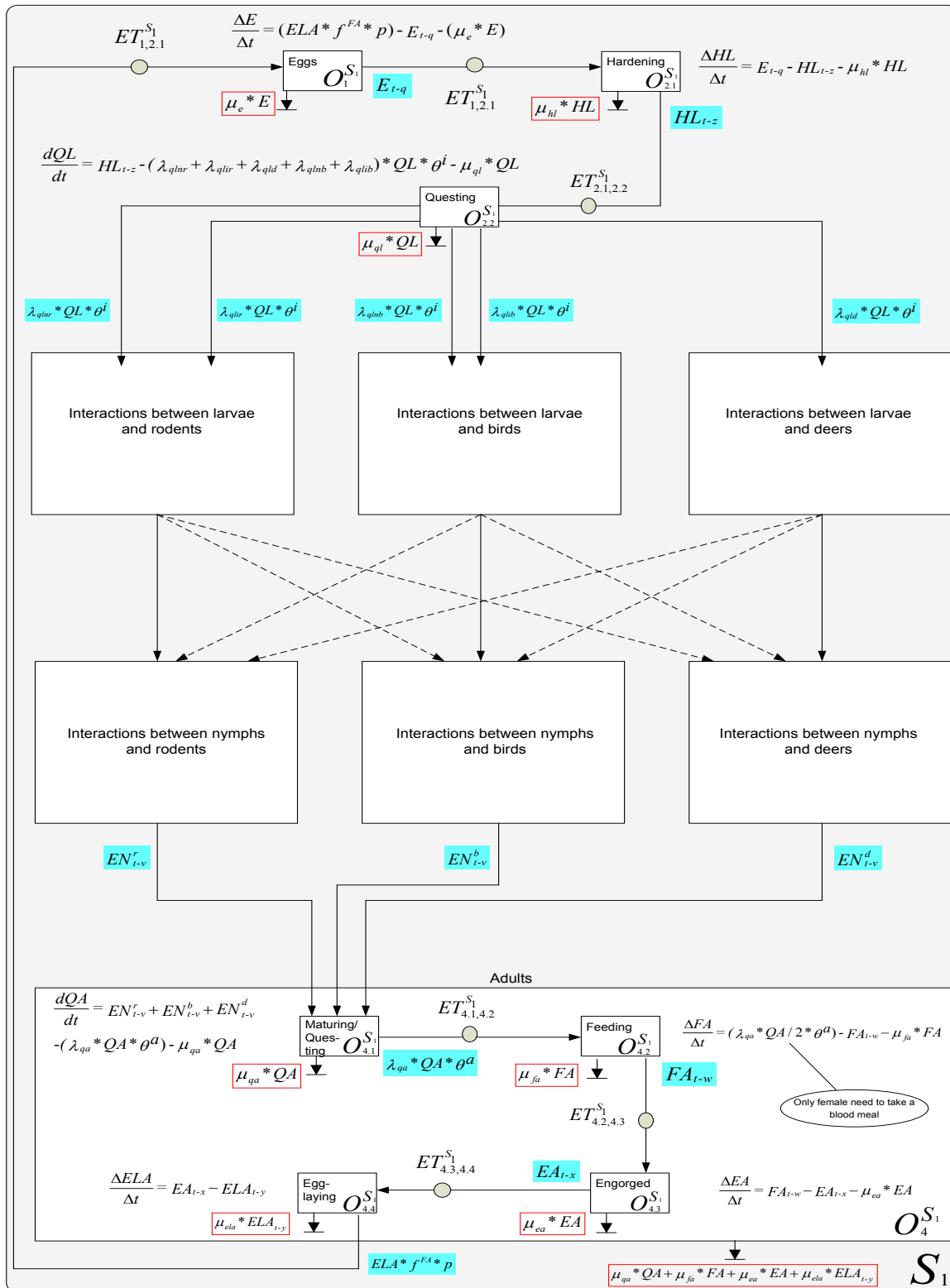


Figure B5 : Interactions des tiques avec tous les hôtes.

Annexe C

Algorithme de la principale fonction permettant la fusion des cellules (après l'enlèvement des trous) pour la création de l'IVGE.

Base de données SIG

Les données d'entrée qui sont utilisées par cet algorithme représentent le résultat d'un prétraitement afin d'obtenir les polygones de couverture de sol ne contenant pas de trous. Chacun de ces polygones possède entre autres un identifiant unique, un attribut représentant la superficie ainsi qu'un attribut représentant le facteur de favorabilité (Suitability) qui est un chiffre entre 0 et 1. Il y a également deux autres attributs « isDeleted » et « isMerged » permettant de savoir si la cellule en cours doit être supprimée ou fusionnée respectivement.

Algorithme

Fonction : Fusion des cellules

/*

Objectif : Fusion des cellules selon un seuil et des heuristiques.

Entrées : La base de données SIG mentionnée auparavant.

Sorties : La base de données spatiale modifiée.

*/

Début

CELL ListeCellules[] // Un tableau ou un recordset contenant une liste de cellules

NUM NbCellules ← 0 // Nombre des cellules

CELL ListeVoisines[] // La liste de cellules voisines

NUM NbVoisines ← 0 // Nombre des cellules voisines

CELL ListeFusion[] // La liste des cellules à fusionner

NUM NbFusion ← 0 // Nombre des cellules à fusionner

CELL MergedCell // Une cellule contenant le résultat de la fusion

CELL ListeSupprim[] // La liste des cellules à supprimer après la fusion

NUM NbSupprim ← 0 // Nombre des cellules à supprimer après la fusion

NUM IdPoly ← 0 // L'identifiant unique d'une cellule

NUM PolySuit ← 0 // La favorabilité d'une cellule

NUM PolyArea ← 0 // La superficie d'une cellule

NUM IdNeighbor ← 0 // L'identifiant unique d'une cellule voisine

NUM NeighborSuit ← 0 // La favorabilité d'une cellule voisine

NUM NeighborArea ← 0 // La superficie d'une cellule voisine

NUM OldSuit ← 0 // L'ancienne valeur de favorabilité

NUM OldArea ← 0 // L'ancienne valeur de superficie

NUM AllArea ← 0 // La superficie totale des polygones fusionnés

NUM NewSuit ← 0 // La nouvelle valeur de favorabilité

NUM Const Threshold ← 0.9 // Une constante représentant un seuil
// qui est utilisé pour arrêter la fusion des cellules

BOOL isMergeNeighbors // Booléen pour savoir s'il y a des cellules à fusionner

// Faire une requête pour retourner les cellules dont la favorabilité est égale à 1

NbCellules, ListeCellules[] ← Lancer_requête "Select*from CellsTable where SUITABILITY=1"

```

// Notre heuristique est de commencer par la cellule ayant la superficie la plus élevée.
// On doit donc trier les cellules dans le tableau ou le recordset en ordre décroissant.
TrierCellules(ListeCellules[], NbCellules) // Criterion = Area

// Traiter maintenant les cellules une par une
REPETER NbCellules FOIS (i DE 0 A NbCellules - 1)
  Début
    //Cette fonction retourne la valeur de l'attribut de la cellule mentionnée dans l'input
    IdPoly ← Retourner_Valeur_Attribut (ListeCellules[i], "ID")
    PolySuit ← Retourner_Valeur_Attribut (ListeCellules[i], " SUIABILITY ")
    PolyArea ← Retourner_Valeur_Attribut (ListeCellules[i], " AREA ")

    // Trouver la liste des cellules voisines de la cellule en cours
    // en faisant une requête spatiale «Touch»
    NbVoisines, ListeVoisines[] ← TrouverCellulesVoisines(ListeCellules[i])

    // Notre heuristique est de commencer par la cellule voisine ayant
    // la favorabilité et la superficie les plus élevées.
    // On doit donc trier les cellules voisines en ordre décroissant.
    TrierCellules(ListeVoisines[], NbVoisines) // Criteria = Suitability + Area

    OldSuit ← PolySuit
    OldArea ← PolyArea
    AllArea ← 0
    NewSuit ← 0
    isMergeNeighbors ← FAUX

    // Traiter maintenant les cellules voisines une par une
    REPETER NbVoisines FOIS (j DE 0 A NbVoisines - 1)
      Début
        //Cette fonction retourne la valeur de l'attribut de la
        //cellule mentionnée dans l'input
        isDeleted ← Retourner_Valeur_Attribut (ListeVoisines[j], " isDeleted ")

        //Si la cellule traitée n'est pas une cellule qu'il faut supprimer
        SI (NON(isDeleted)) ALORS
          Début

            IdNeighbor ← Retourner_Valeur_Attribut (ListeVoisines[j], "ID")
            NeighborArea ← Retourner_Valeur_Attribut (ListeVoisines[j], "AREA")
            NeighborSuit ← Retourner_Valeur_Attribut (ListeVoisines[j],
                                                       " SUIABILITY ")

            //Calculer la nouvelle favorabilité en supposant que la cellule traitée
            //va être fusionnée avec la cellule en cours.
            AllArea ← OldArea + NeighborArea
            NewSuit ← (OldSuit * (OldArea / AllArea))
                      + (NeighborSuit * (NeighborArea / AllArea))

            SI (NewSuit >= Threshold) ALORS
              Début
                isMergeNeighbors ← VRAI
                OldSuit ← NewSuit
                OldArea ← AllArea
                Valeur_Attribut (ListeVoisines[j], " isDeleted ") ← VRAI

```



```

        Valeur_Attribut (ListeVoisines[j], " isMerged ") ← VRAI
    Fin
    SINON
    Début
        NewSuit ← OldSuit
        AllArea ← OldArea
    Fin
    Fin
    Fin
    SI (isMergeNeighbors) ALORS
    Début
        // Faire une requête pour retourner les cellules dont la favorabilité est égale à 1
        NbFusion, ListeFusion[] ← Lancer_requête
        "Select * from CellsTable where ID = " & IdPoly & " OR isMerged = True"

        // C'est ici la fonction la plus importante permettant de fusionner les cellules
        // Elle retourne une cellule après fusion de toutes les autres
        MergedCell ← FussionerCellules (ListeFusion[],NbFusion)

        // Donner les nouvelles valeurs à la cellule en cours
        // en modifiant aussi sa géométrie
        Valeur_Attribut (ListeCellules[i], " SUITABILITY") ← NewSuit
        Valeur_Attribut (ListeCellules[i], " Area") ← AllArea
        Valeur_Attribut (ListeCellules[i], " Geometry") ←
            Retourner_Valeur_Attribut (MergedCell, " Geometry ")

        REPETER NbVoisines FOIS (j DE 0 A NbVoisines - 1)
        Début
            Valeur_Attribut (ListeVoisines[j], " isMerged ") ← FAUX
        Fin
    Fin
    Fin
    // Faire une requête pour retourner les cellules qu'il faut supprimer
    NbSupprim, ListeSupprim[] ← Lancer_requête
        "Select * from CellsTable where isDeleted = True"

    // Traiter maintenant les cellules qu'il faut supprimer une par une
    REPETER NbSupprim FOIS (i DE 0 A NbSupprim - 1)
    Début
        SupprimerCellule(ListeSupprim[i])
    Fin
    Fin

```

Annexe D

Algorithme de la principale fonction permettant le transfert d'un flot d'individus à travers les cellules de l'IVGE.

Fonction : Distribution d'un flot d'individus

/*

Objectif : Distribuer un flot d'individus à travers les cellules de l'IVGE

Entrées :

*- ListeCellules[] : les cellules de l'IVGE. Chaque cellule contient ses attributs comme le nombre des individus résidents, le nombre des individus non résidents, le coefficient d'attraction, l'attribut de la trajectoire, la liste des cellules voisines et leurs orientations, etc.

*- NbCellules : nombre des cellules.

*- NbFlot : le nombre des individus à distribuer.

*- ProportionCA: le pourcentage de la capacité d'accueil (*Reception Capacity Rate*).

*- PourcentageNR: la distribution des individus non résidents (*Transient Distribution Rate*).

*- OrientationST: le seuil d'orientation (*Orientation Threshold*).

*- PremiereCelluleATraiter : Booléen pour commencer par traiter la première cellule.

*- CellulesVoisinesATraiter : Booléen pour traiter les autres cellules.

Sorties : Les cellules mises à jour avec les nouvelles valeurs d'individus distribués.

*/

Début

CELL ListeVoisines[] // La liste des cellules voisines

NUM NbVoisines ← 0 // Nombre des cellules voisines

NUM NbResidents ← 0 // Nombre des individus résidents

NUM NbNonResidents ← 0 // Nombre des individus non résidents

NUM CAREelle ← 0 // Capacité réelle à utiliser (Capacité d'accueil * ProportionCA)

NUM NRAdistribuer ← 0 // Nombre des individus non résidents à distribuer

NUM OriCG ← 0 // Orientation de la trajectoire de la cellule génératrice

NUM OriV ← 0 // Orientation du voisin

NUM DiffOri ← 0 // Différence des orientations

NUM SC ← 0 // Numéro de la cellule génératrice

NUM sumCoefficients ← 0 // La somme des coefficients d'attraction

BOOL CelluleTraitee ← FAUX // Booléen pour savoir si la cellule est déjà traitée

NUM i ← 0 // Variable d'incrémement

// Commencer par distribuer les individus pour la première cellule

SI (PremiereCelluleATraiter = VRAI) ALORS

Début

// Remarque: on a besoin de ne rien faire si le nombre d'individus qui arrivent est nul

SI (NbFlot > 0) ALORS

Début

NbResidents ← Retourner_Valeur_Attribut (ListeCellules[0], "NbResidents")

NbNonResidents ← Retourner_Valeur_Attribut (ListeCellules[0], "NbNonResidents")

// Calculer la capacité d'accueil de C1 qui va être réellement utilisée

CAREelle = Retourner_Valeur_Attribut (ListeCellules[0], "CapaciteAccueil") * ProportionCA

SI (NbResidents < CAREelle) ALORS

Début

SI (nbFlot > (CAREelle - NbResidents)) ALORS

Début

NbNonResidents ← NbNonResidents + nbFlot - (CAREelle - NbResidents)

NbResidents ← CAREelle

```

        Fin
    SINON
        Début
            NbResidents ← NbResidents + nbFlot
        Fin
    Fin
SINON SI (NbResidents = CAReelle) ALORS
    Début
        NbNonResidents ← NbNonResidents + nbFlot
    Fin
SINON
    Début
        Afficher "Message Erreur : NbResidents ne peut pas être > CAReelle"
    Fin
    //Affecter les nouvelles valeurs de la cellule C1
    Valeur_Attribut (ListeCellules[0], "NbResidents") ← NbResidents
    Valeur_Attribut (ListeCellules[0], "NbNonResidents") ← NbNonResidents
    //Tous les individus du flot ont été envoyés à la première cellule
    nbFlot = 0
    Fin
Fin
// Distribuer maintenant les individus d'une cellule génératrice à ses cellules voisines
SI (CellulesVoisinesATraiter = VRAI) ALORS
    Début
        /* Traiter maintenant les cellules voisines. Notre heuristique est de commencer
        premièrement par la cellule voisine dont l'orientation est la plus proche de celle
        de la trajectoire. Par la suite, nous considérons, le coefficient d'attraction et la capacité
        d'accueil les plus élevés. On doit donc trier les cellules */

        // Trouver le numéro de la cellule génératrice
        SC = Retourner_Numero_Cellule ("CelluleGénératrice")

        NbNonResidents = Retourner_Valeur_Attribut (ListeCellules[SC], "NbNonResidents")
        // Remarque: on a besoin de ne rien faire si le nombre d'individus qui arrivent est nul
        SI (NbNonResidents > 0) ALORS
            Début
                //Trouver l'orientation de la trajectoire dans la cellule génératrice
                OriCG = Retourner_Valeur_Attribut (ListeCellules[SC], "AttributAngleTrajectoire")
                //Trouver les cellules voisines de la cellule génératrice
                NbVoisines, ListeVoisines[] ← TrouverCellulesVoisines(ListeCellules[SC])

                REPETER NbVoisines FOIS (i DE 0 A NbVoisines - 1)
                    Début
                        CelluleTraitee ← Retourner_Valeur_Attribut (ListeVoisines[i], " CelluleTraitee")
                        SI (CelluleTraitee = VRAI) ALORS
                            Début
                                //Cette fonction enlève la cellule du tableau pour ne pas la trier
                                Enlever_Cellule (ListeVoisines[i])
                            Fin
                            // Calculer la somme des coefficients d'attraction
                            sumCoefficients ← sumCoefficients +
                                Retourner_Valeur_Attribut (ListeCellules[i],
                                    "AttributCoefficientAttraction")
                        Fin
                    Fin
            Fin
        // Modifier les valeurs des proportions et des différences dans les cellules voisines pour les trier

```

```

REPETER NbVoisines FOIS (i DE 0 A NbVoisines - 1)
  Début
    Valeur_Attribut (ListeVoisines[i], "AttractionProportion") ←
      Retourner_Valeur_Attribut (ListeVoisines [i],
        "AttributCoefficientAttraction") / sumCoefficients
    SI (OriCG != 999.0) ALORS
      Début
        //Trouver l'orientation de la cellule voisine par rapport à la cellule génératrice
        OriV = Retourner_Valeur_Attribut (ListeVoisines[i], "AttributAngleVoisin")
        DiffOri = Math.abs(OriCG - OriV)
        DiffOri = Math.min(DiffOri, Math.abs(2*Math.PI- DiffOri))
        //Changer la valeur de la différence d'orientation qui va être utilisé pour le tri
        Valeur_Attribut (ListeVoisines[i], "OrientationDifference") ← DiffOri
      Fin
    Fin

/* Trier les cellules selon les orientations les plus proches de la trajectoire
ou selon les proportions d'attraction */
SI (OriCG != 999.0) ALORS
  Début
    TrierCellules (ListeVoisines[], NbVoisines) // Criteria = AttractionProportion
  Fin
SINON
  Début
    TrierCellules (ListeVoisines[], NbVoisines) // Criteria = OrientationDifference
  Fin

// Distribuer les individus cellule par cellule
REPETER NbVoisines FOIS (i DE 0 A NbVoisines - 1)
  Début
    // Ajouter la cellule voisine dans la file d'attente si elle n'existe pas déjà
    SI (ChercherCelluleDansFile (ListeVoisines[i]) = FAUX) ALORS
      Début
        Ajouter_Cellule (ListeVoisines[i], File_Attente)
      Fin
    // Calculer la capacité d'accueil de la cellule voisine qui va être réellement utilisée
    CAREelle = Retourner_Valeur_Attribut (ListeVoisines[i], "CapaciteAcceuil")
      * ProportionCA
    // Chercher les valeurs des individus résidents et non résidents
    NbResidents ← Retourner_Valeur_Attribut (ListeVoisines[i], "NbResidents")
    NbNonResidents ← Retourner_Valeur_Attribut (ListeVoisines[i],
      " NbNonResidents ")
    // Trouver le nombre des individus à distribuer selon la proportion d'attraction
    NRAdistribuer ← Retourner_Valeur_Attribut (ListeCellules[SC],
      "NbNonResidents") * Retourner_Valeur_Attribut (ListeVoisines[i],
        "AttractionProportion")
    SI (NbResidents < CAREelle) ALORS
      Début
        SI (NRAdistribuer > (CAREelle – NbResidents)) ALORS
          Début
            NbNonResidents ← NbNonResidents + NRAdistribuer –
              (CAREelle – NbResidents)
            NbResidents ← CAREelle
          Fin
        Fin
      Fin
    SINON

```

```

        Début
            NbResidents ← NbResidents + NRAdistribuer
        Fin
    Fin
SINON SI (NbResidents = CAREelle) ALORS
    Début
        NbNonResidents ← NbNonResidents + NRAdistribuer

    Fin
SINON
    Début
        Afficher "Message Erreur : NbResidents ne peut pas être > CAREelle"
    Fin

    //Affecter les nouvelles valeurs de la cellule
    Valeur_Attribut (ListeVoisines[i], "NbResidents") ← NbResidents
    Valeur_Attribut (ListeVoisines[i], "NbNonResidents") ← NbNonResidents
    Fin
    Fin
// Si la file est vide, on arrête le processus sinon on continue avec la prochaine cellule génératrice
SI (File_Est_Vide = FAUX) ALORS
    Début
        Valeur_Attribut (ListeVoisines[i], " CelluleTraitee") ← VRAI
    Fin
SINON
    Début
        Afficher "Il ne reste plus d'individus à distribuer !"
    Fin
    Fin
Fin

```

Annexe E

Architecture, données et résultats de calibration de *WNV-MAGS*

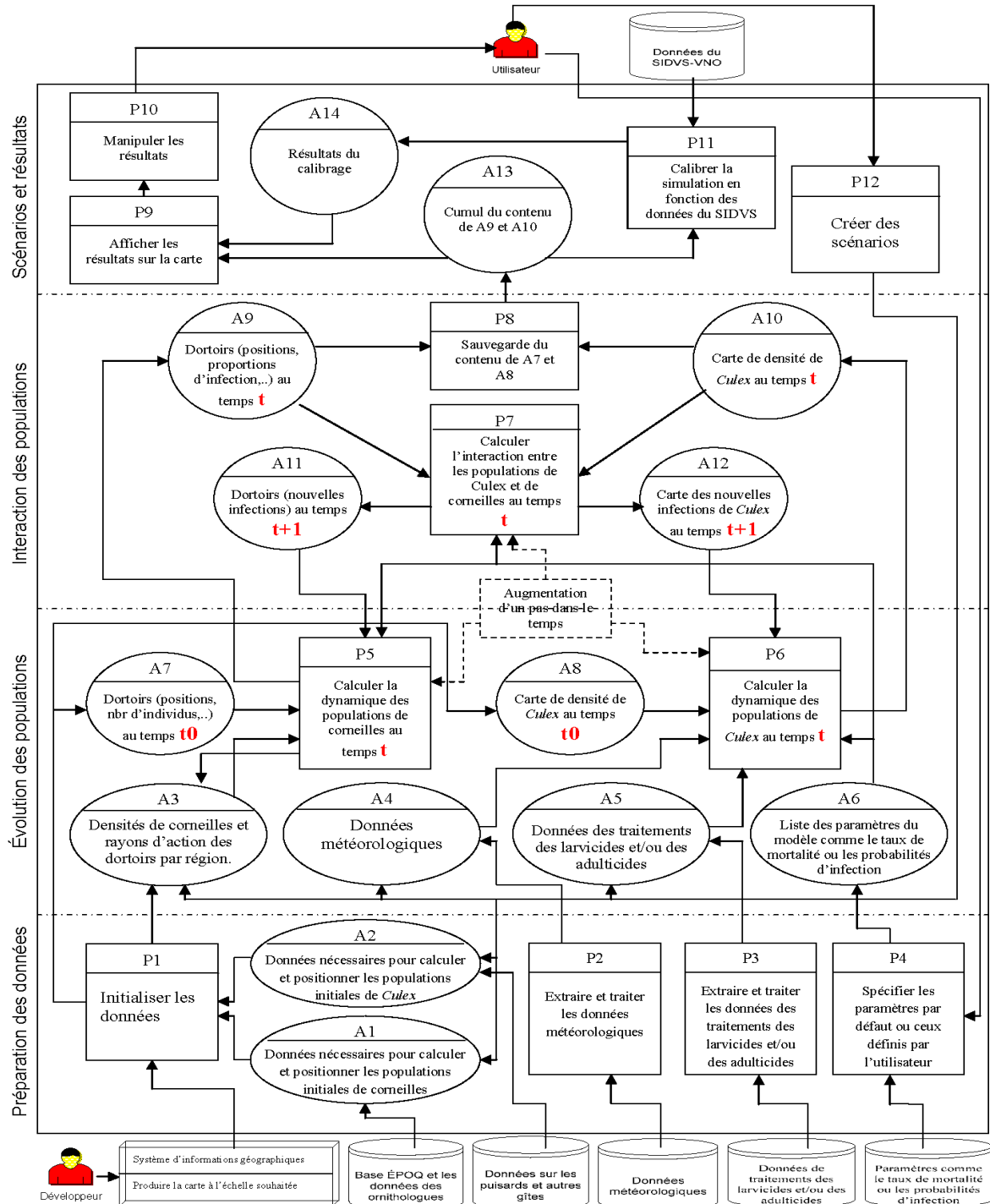


Figure E1 : Architecture conceptuelle du système *WNV-MAGS*.

Tableau E1 : Densité moyenne de corneilles par région.

Code RSS	Nom_RSS	Sup écoumène	Nb année	Moy.	Min.	Max.	Somme RSS	Densité (Nb/km ²)
4	Mauricie-Centre-du-Québec	11760	8	415	63	814	3320	0,035289116
15	Laurentides	10990	8	2367,125	23	4801	18937	0,21538899
5	Estrie	10430	8	2478,375	573	4718	19827	0,237619847
7	Outaouais	11950	7	5205,142857	364	7727	36436	0,435576808
12	Chaudière-Appalaches	15160	8	6616	2	23655	52928	0,436411609
14	Lanaudière	7060	8	3503,875	324	6913	28031	0,496299575
3	Québec	6896	8	9496,75	136	12794	75974	1,377138921
16	Montérégie	11270	8	27858,875	11323	40836	222871	2,471949867
13	Laval	246,6	8	652,75	193	1149	5222	2,646999189
6	Montréal-Centre	500,9	8	3088,375	1411	5645	24707	6,165651827

Tableau E2 : Exemple des résultats obtenus afin de calculer la longueur des routes des municipalités.

Ville de Sainte-Foy

Fichier	Description	Nombre des morceaux de routes	Longueur totale des routes (m)
QChwy	Highways	100	20406,61
QChrd	Major Roads and Highways	435	58876,30
QCrd	Roads	2921	441635,34

Total :	520918,25
Nb puisards s'il y a 1 puisard/30 m	17364

Ville de Laval

Fichier	Description	Nombre des morceaux de routes	Longueur totale des routes (m)
QChwy	Highways	655	92781,06
QChrd	Major Roads and Highways	2728	391500,05
QCrd	Roads	12481	1738758,69

Total :	2223039,80
Nb puisards s'il y a 1 puisard/30 m	74101

Tableau E3 : Paramètres du modèle mathématique du VNO [Wonham *et al.*, 2004].

Paramètre	Moyenne (Intervalle)	Nom de la variable
Proportion de moustiques (par rapport à une population) qui vont piquer des corneilles par jour.	0.09 (0.03-0.16)	a
Probabilité de transmission du VNO des moustiques vers les corneilles par piqûre	0.88 (0.8-1.00)	b
Probabilité de transmission du VNO des corneilles vers les moustiques par piqûre	0.16 (0.02-0.24)	c
Proportion de moustiques (par rapport à une population) qui vont naître par jour	0.537 (0.036-42.5)	β_M
Proportion de larves (par rapport à une population) qui vont se transformer en adultes par jour	0.068 (0.051-0.093)	m
Proportion de moustiques adultes (par rapport à une population) qui vont mourir par jour	0.029 (0.016-0.070)	μ_A
Proportion de larves (par rapport à une population) qui vont mourir par jour	1.191 (0.213-16.9)	μ_L
Proportion de moustiques (par rapport à une population) qui vont être infectés par le VNO par jour après une période d'incubation du virus (12 jours).	0.106 (0.087-0.125)	k
Proportion de corneilles (par rapport à une population) qui vont mourir par jour après une infection au VNO	0.143 (0.125-0.200)	μ_V
Proportion de corneilles (par rapport à une population) qui vont se rétablir par jour après une infection au VNO	0	g

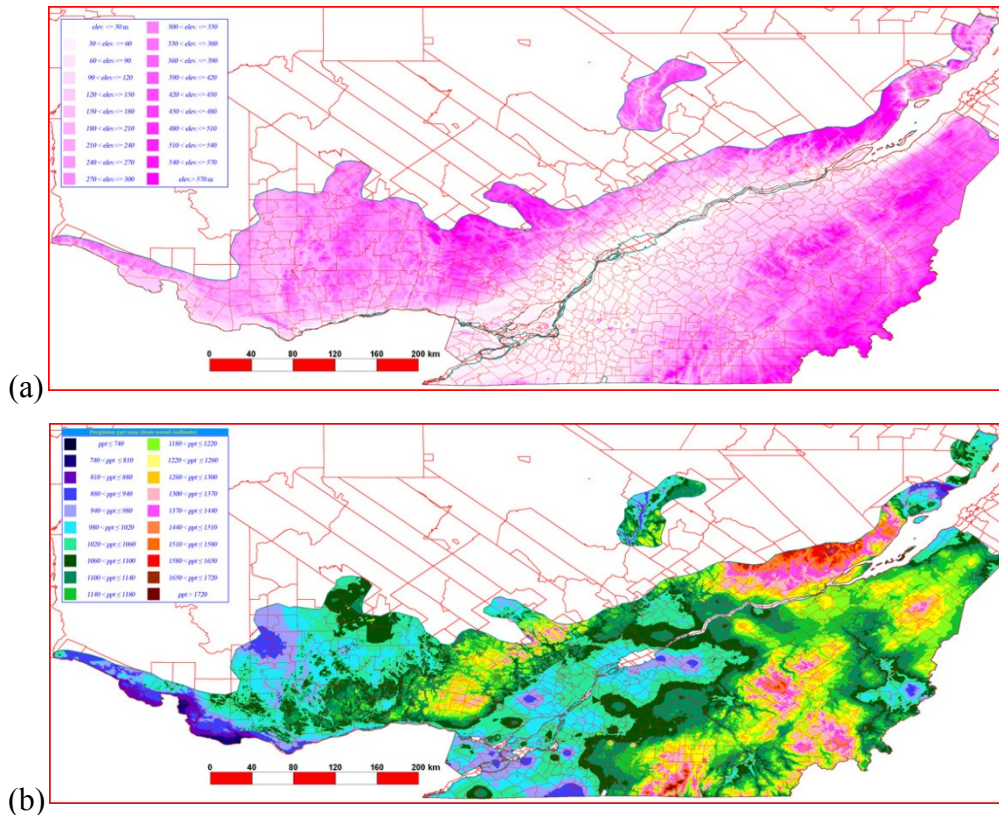


Figure E2 : Cartes produites par le système *WNV-MAGS*: (a) modèle numérique d'élévation du terrain et (b) les précipitations.

Tableau E4 : Mentions de dortoirs signalées par les ornithologues ou se trouvant dans la base ÉPOQ.

Nb	Lat	Long	Lieu du dortoir	Nb_mentions	Nb_individus_moy
1	48,43	-71,07	Bonsai (Chicoutimi)	1	40
2	47,07	-70,80	Cap Tourmente	1	500
3	45,45	-73,32	Carignan	1	100
4	46,72	-71,27	Charny	1	450
5	45,58	-71,62	Dudswell	1	70
6	45,00	-74,05	Dundee	1	200
7	46,73	-75,47	Ferme-Neuve	1	450
8	45,24	-72,44	Granby_1	1	8000
9	45,39	-72,75	Granby_2	2	20450
10	46,02	-73,45	Joliette	1	200
11	48,03	-70,87	La Baie	1	1500
12	48,53	-71,38	Lac Duclos	1	150
13	45,08	-73,37	Lacolle	1	200
14	46,80	-71,18	Lévis	2	161
15	46,54	-75,45	Mont-Laurier	1	300
16	46,55	-75,05	Mont-Laurier, ruisseau Villemaire	3	222
17	45,48	-73,67	Montréal	1	3000
18	47,58	-68,80	Notre-Dame-du-Lac	6	780
19	46,75	-71,13	Pintendre	1	185
20	48,43	-68,55	Rimouski	4	71
21	48,48	-68,43	Saint-Anaclet	1	563
22	45,30	-73,23	Saint-Athanase, rivière Richelieu	2	5000
23	48,05	-68,25	Saint-Donat (Rimouski)	2	575
24	45,62	-72,95	Saint-Hyacinthe_1	1	1000
25	45,63	-72,98	Saint-Hyacinthe_2	1	300
26	45,32	-73,27	Saint-Jean-sur-Richelieu	5	3540
27	48,47	-78,05	Saint-Marc-de-Figuery	1	1300
28	45,25	-71,54	Sherbrooke	1	200
29	45,52	-71,81	Stoke	1	150
30	46,02	-73,15	Tracy	21	825
31	46,05	-71,91	Victoriaville	1	500
32	47,55	-68,65	Ville-Dégelis	1	400

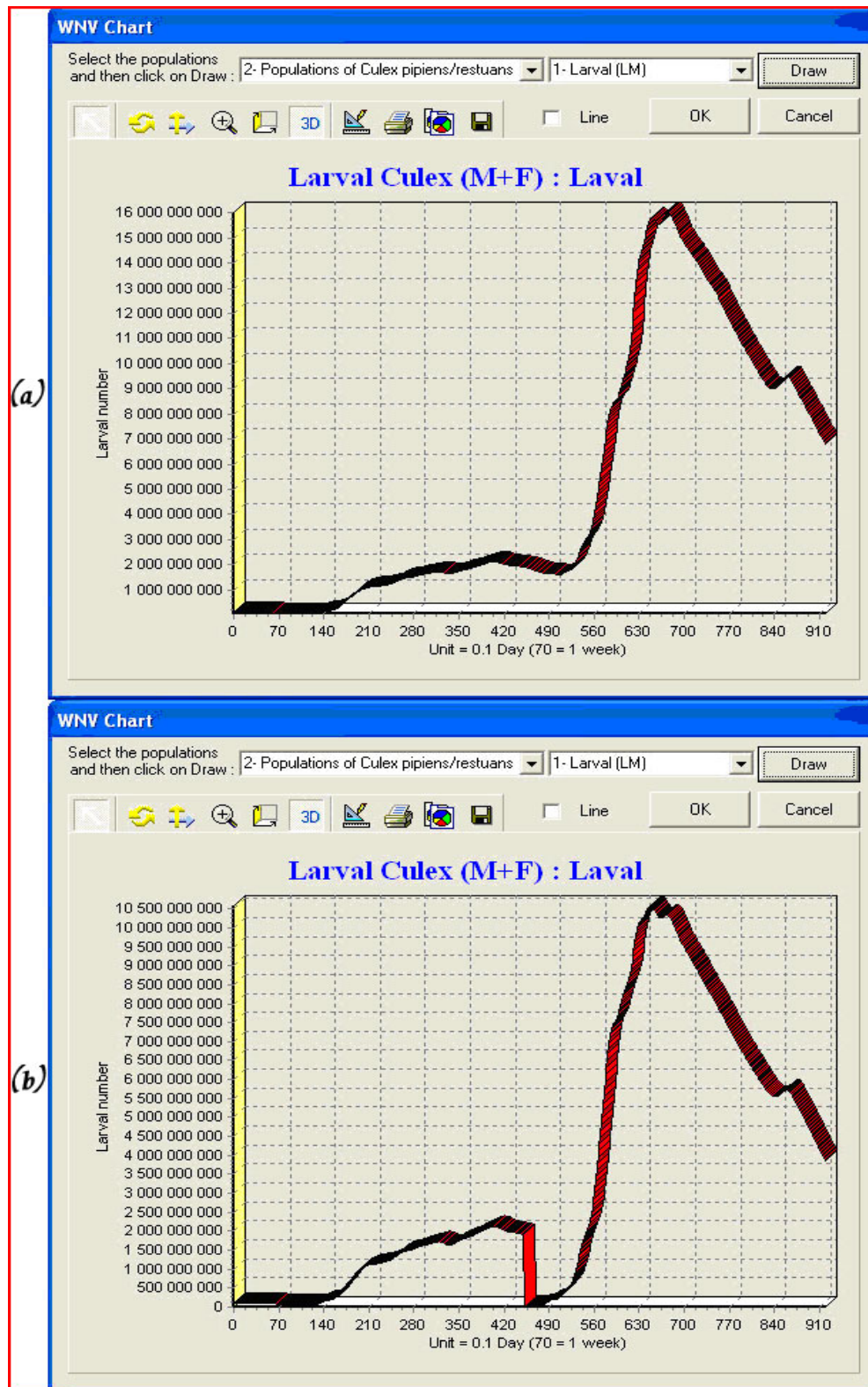


Figure E3 : Dynamique des populations larvaires avant (a) et après (b) le lessivage des puisards dans la municipalité de Laval le 15 août (scénario hypothétique défini par l'utilisateur).

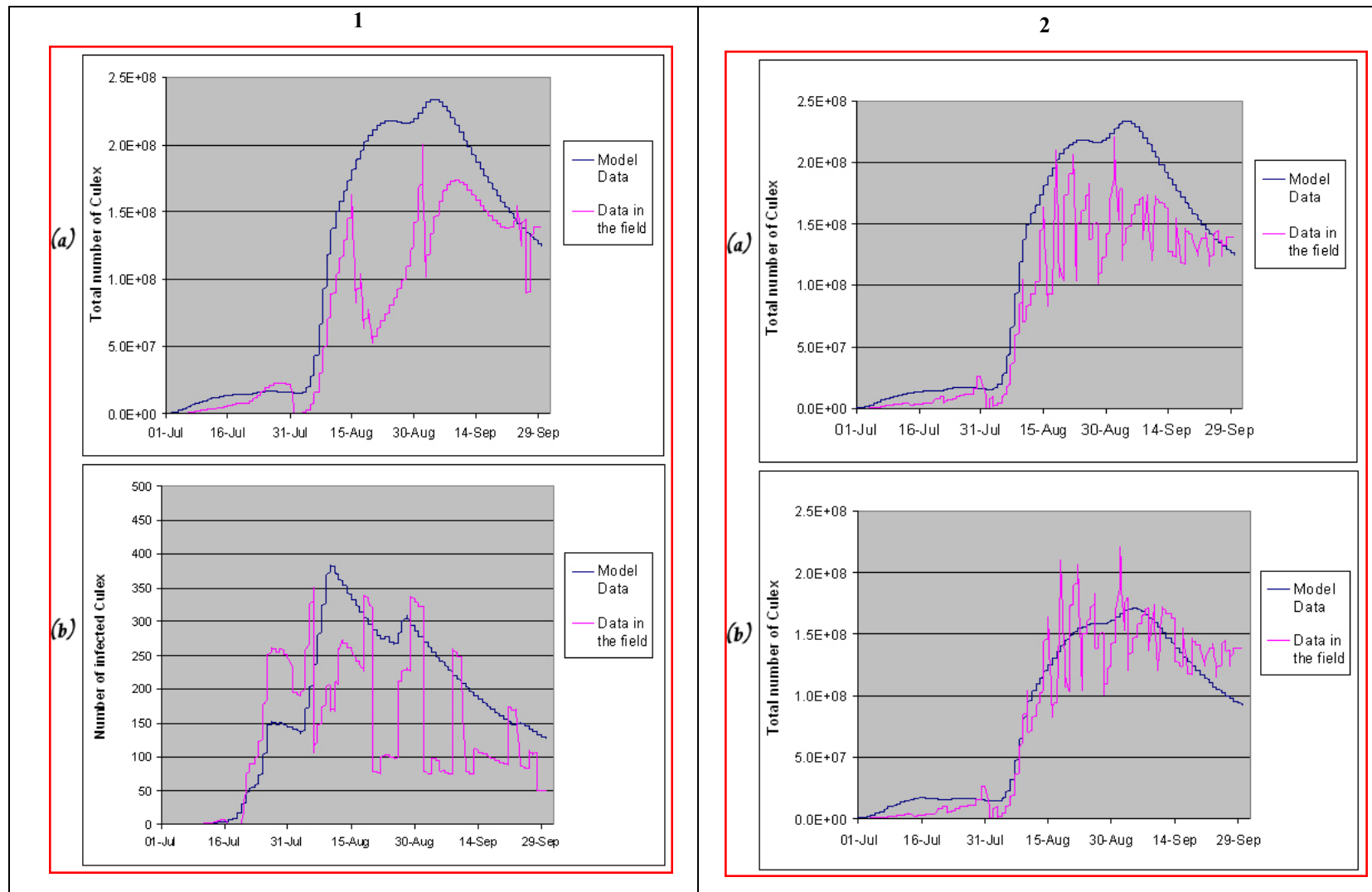


Figure E4 : Calibrage du modèle de *WNV-MAGS* qui est utilisé pour faire évoluer les populations de moustiques (a: totales, b: infectés): (1) sans applications de larvicides en utilisant les moyennes de 2003 à 2006; (2) avec applications de larvicides (18 juin, 17 juillet et 13 août) en utilisant les données de 2003 uniquement pour la municipalité de Laval.

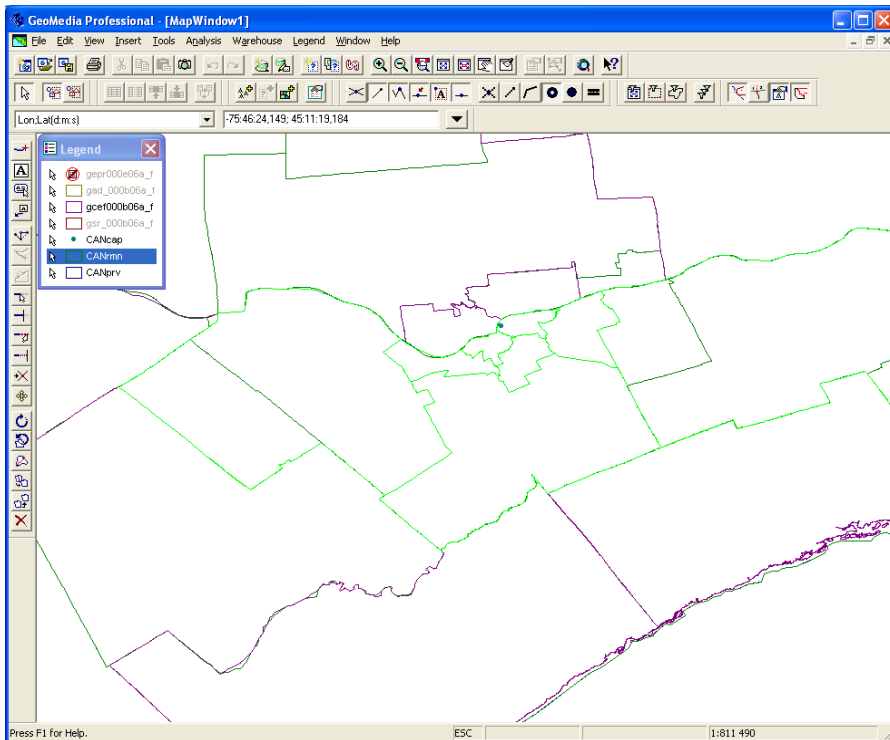


Figure E5 : Subdivision de la région d’Ottawa selon les circonscriptions électorales fédérales (CEF).

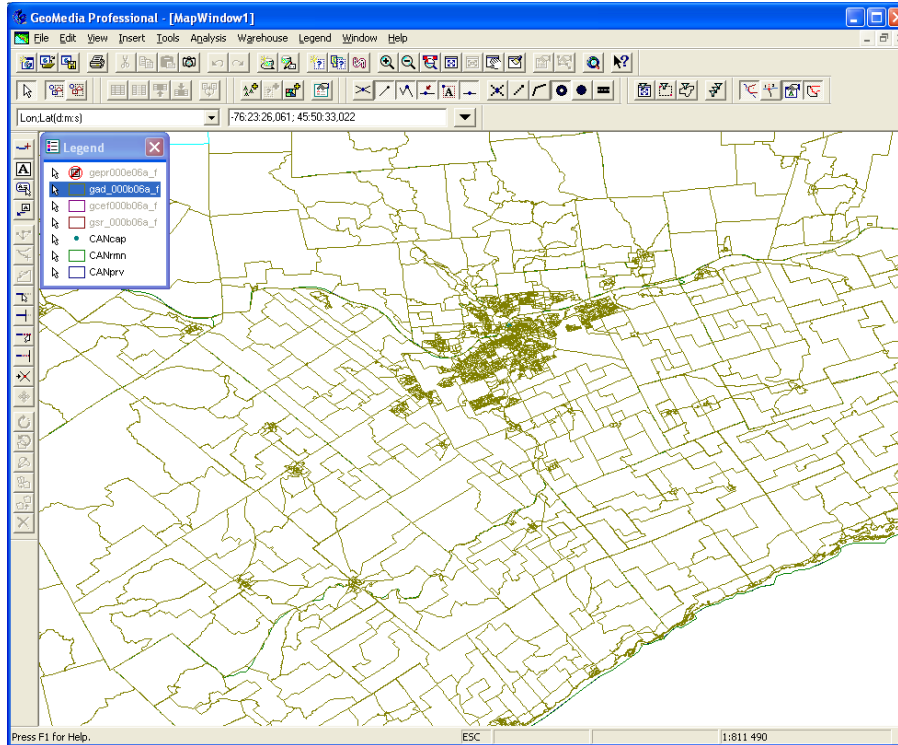


Figure E6 : Subdivision de la région d’Ottawa selon les aires de diffusion (AD).

Annexe F

Transfert et cycle biologique des tiques en considérant le déplacement des hôtes.

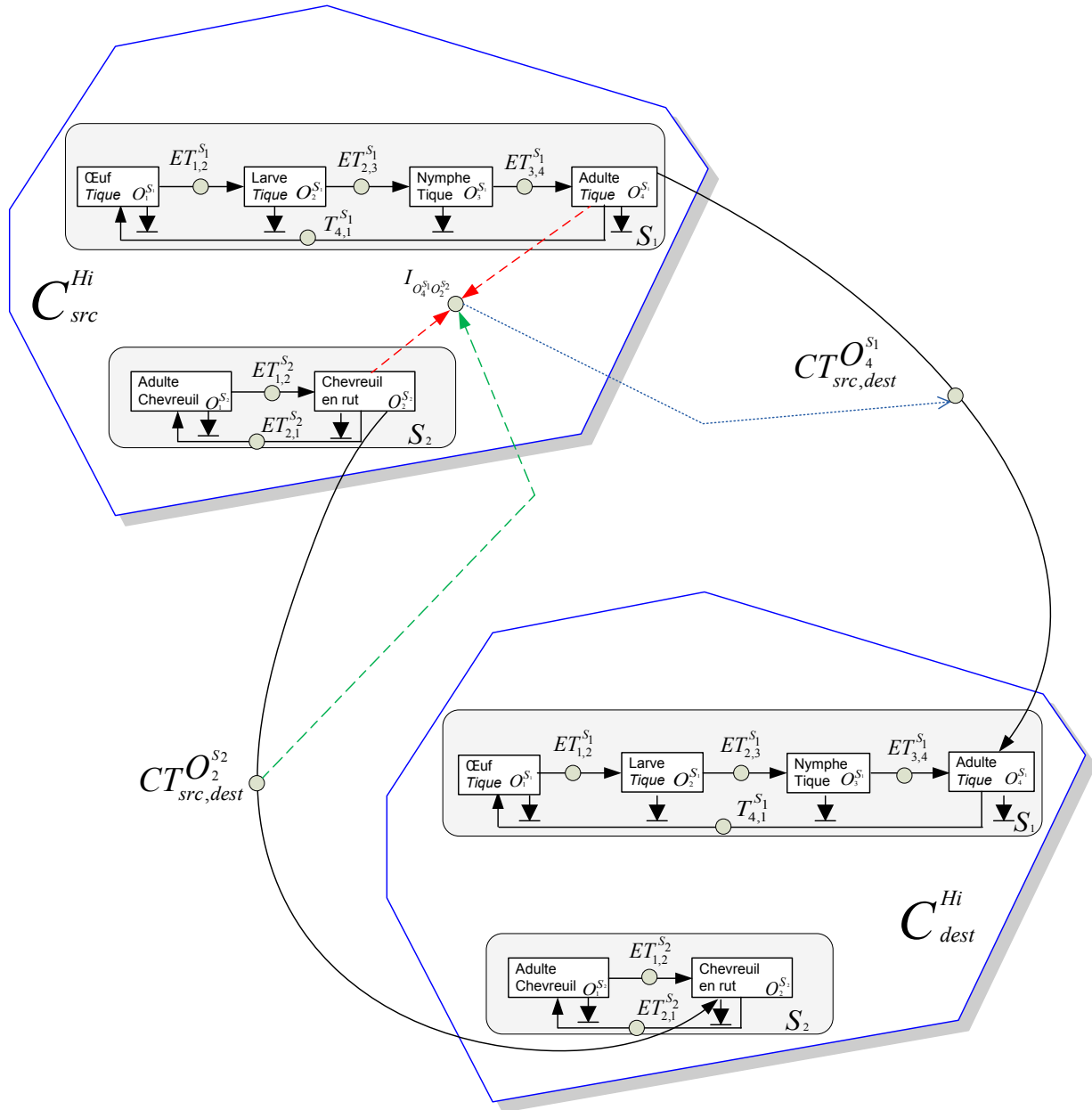


Figure F1 : Transfert des tiques accrochées sur les chevreuils lors du déplacement de ces derniers (utilisation de deux transitions de transfert).

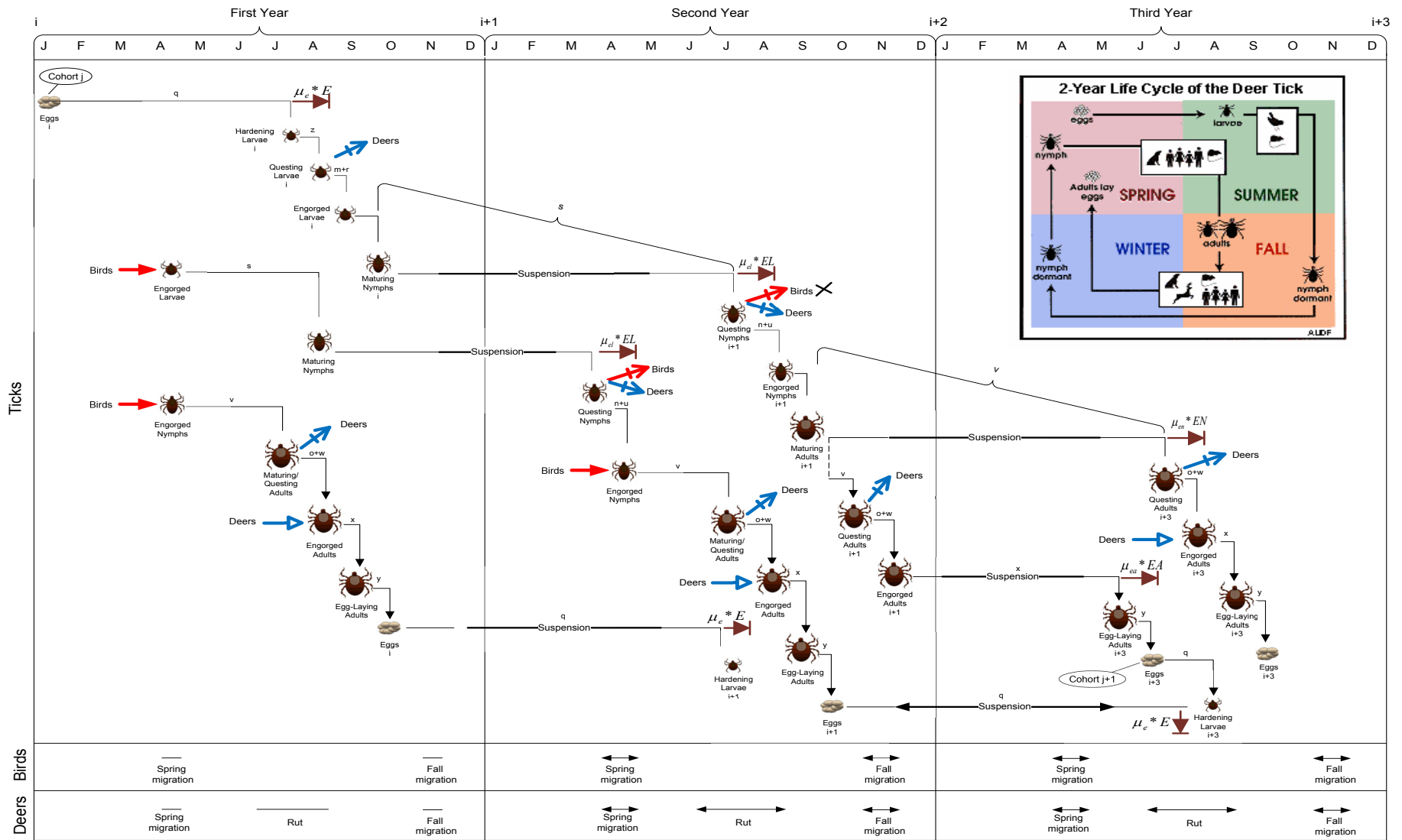


Figure F2 : Évolution des tiques sur 3 ans en relation avec le déplacement des hôtes (oiseaux et chevreuils).