

Simulation informatique au service des Sciences de la Vie

David R. C. HILL





Presses Universitaires Blaise Pascal ©

*Maison des Sciences de l'Homme
4, rue Ledru – 63057 Clermont-Ferrand Cedex 1
Tel. 04 73 34 68 09 – Fax 04 73 34 68 12
Publi.Lettres@univ-bpclermont.fr
www pubp fr*

*Maquette et Illustration de couverture :
Montage, Diazo I*

Simulation informatique

au service des Sciences de la Vie

David R. C. HILL

Pour mon épouse Anne (Eph. 5, 25)

TABLE DES MATIÈRES

CHAPITRE 1 Introduction Générale

1	Systémique et Complexité	9
2	Les écosystèmes : archétype des systèmes complexes	10
3	La simulation	11
4	La modélisation d'écosystèmes	13
5	Contexte de nos travaux de recherche	15
6	Organisation de l'ouvrage	16

CHAPITRE 2 Réflexions sur la Modélisation et la Simulation

1	Introduction	19
2	Qu'est-ce qu'un système ?	20
3	Complexité des systèmes et intérêt de la modélisation	21
4	Qu'est-ce qu'un modèle ?	24
	<i>4.1. La notion de modèle</i>	24
	<i>4.2. Modèles discrets, continus, stochastiques, déterministes</i>	25
	<i>4.3. Les qualités d'un modèle</i>	27

5	À quoi sert un modèle ?	27
	<i>5.1. Questions et problèmes</i>	27
	<i>5.2. Les modèles sont des instruments scientifiques</i>	29
6	La simulation : le modèle plongé dans le temps	30
	<i>6.1. Introduction</i>	30
	<i>6.2. Simulation par objets, acteurs et agents</i>	31
	<i>6.3. Langages, méthodes et outils visuels de simulation</i>	33
7	Choix méthodologiques	35
	<i>7.1. Niveau d'abstraction</i>	35
	<i>7.2. Niveau de détail</i>	36
	<i>7.3. Granularité du temps</i>	38
8	Conclusion : exploitation d'un modèle	39

CHAPITRE 3 Modélisation par Objets de Systèmes à Flux discrets

6

1	Introduction	41
2	Proposition d'un processus et d'une méthode de modélisation par objets	41
	<i>2.1. La synthèse des approches existantes</i>	44
	<i>2.2. La proposition d'un processus</i>	45
	<i>2.3. La proposition d'une méthode de modélisation</i>	48
	<i>2.4. Conclusion</i>	52
3	L'analyse et la conception d'outils d'animation de résultats de simulation	53
	<i>3.1. Le contexte historique</i>	53
	<i>3.2. Les techniques proposées</i>	55
	<i>3.3. Conclusion</i>	57
4	La génération automatique de code pour l'animation et la simulation	58
5	Conclusion	60

CHAPITRE 4 Une approche de la Modélisation d'Écosystèmes

1	Introduction	61
2	Le nécessaire couplage	
	Système d'Information Géographique / simulation	63
	2.1. <i>Introduction</i>	63
	2.2. <i>Intérêt du couplage SIG - SAED</i>	64
	2.3. <i>Conclusion sur les relations entre les données d'une SAED et les informations fournies par un SIG</i>	66
3	Les simulations en foresterie	68
	3.1. <i>Introduction</i>	68
	3.2. <i>Application à la croissance de forêts avec prise en compte de l'effet spatial</i>	71
	3.3. <i>Conclusion</i>	74
4	Les simulations appliquées à l'Océanographie	76
	4.1. <i>Introduction</i>	76
	4.2. <i>Le modèle d'expansion de l'algue Caulerpa raxifolia</i>	78
	4.3. <i>Résultats de la simulation de la croissance de Caulerpa taxifolia</i>	81
	4.4. <i>Simulation de la croissance de l'herbier de Posidonie</i>	85
	4.5. <i>La lutte par un agent de contrôle biologique</i>	88
	4.6. <i>Métamodélisation par un réseau de neurones</i>	92
	4.7. <i>Conclusion</i>	97
5	Les simulations multi-agents pour l'éthologie	99
	5.1. <i>Introduction</i>	99
	5.2. <i>Étude de la mémoire des moutons</i>	101
	5.3. <i>L'entretien des paysages par des herbivores</i>	108
	5.4. <i>Conclusion</i>	118
6	Intégration des techniques du Web	119
	6.1. <i>Introduction</i>	119
	6.2. <i>Quelques applications</i>	122
	6.3. <i>Conclusion sur les simulations de type « Web-based »</i>	125
7	Perspectives	126

CHAPITRE 5	Les Problèmes de Validation et de Vérification des Modèles	
1	Introduction	131
2	Les cadres expérimentaux	134
3	La vérification des programmes de simulation	134
4	La validation des modèles et des résultats	138
	<i>4.1. Introduction</i>	138
	<i>4.2. Analyse et validation des données</i>	141
	<i>4.3. La validation du modèle conceptuel</i>	142
5	Utilité de l'animation pour la validation de résultats de simulation	143
6	L'analyse spectrale : une technique d'aide à la validation de modèles stochastiques spatialisés	146
	<i>6.1. Interprétation de résultats de couplage SIG - SAED</i>	146
	<i>6.2. Analyse spatiale et analyse statistique</i>	147
	<i>6.3. Un exemple appliqué en océanographie</i>	148
7	Une approche logicielle pour la conception de plans d'expériences	154
8	Une application à grande échelle des plans d'expériences	158
9	Conclusion	163
8		
CHAPITRE 6	Conclusion	165
	Réflexions	165
	Quelques apports	168
	Derniers travaux et perspectives	170
ANNEXE	Références bibliographiques	175
	TABLE DES FIGURES	203

CHAPITRE 1

INTRODUCTION GÉNÉRALE

Peu de gens parlent de l'humilité humblement... La vanité est si ancrée au cœur de l'homme... que ceux qui écrivent contre veulent avoir la gloire d'avoir bien écrit ; et ceux qui le lisent veulent avoir la gloire de l'avoir lu, et moi qui écrit ceci, ai peut-être cette envie ; et peut-être que ceux qui liront...

Blaise Pascal (1632-1662)

1. SYSTÉMIQUE ET COMPLEXITÉ

Cet ouvrage synthétise des travaux menés en informatique pour les Sciences de la Vie depuis 1994 au sein du Laboratoire d'Informatique de Modélisation et d'Optimisation des Systèmes de l'Université Blaise-Pascal (LIMOS). Les dernières décennies sont caractérisées par un développement impressionnant des technologies de l'information. Cette puissance de traitement de l'information a cependant engendré une augmentation non négligeable de la complexité des systèmes à la base de l'activité économique des entreprises et des administrations qui, dans les pays « développés », reposent maintenant majoritairement sur du matériel et du logiciel informatique visant à améliorer la productivité et les communications. Par ailleurs, l'aug-

mentation de la puissance de calcul et de mémorisation des informations a permis aux scientifiques d'augmenter leurs connaissances des phénomènes physiques et de la nature en général. C'est précisément la confrontation avec des situations complexes de type naturel (physique, chimie, biologie, écologie, médecine, géographie,...) qui amenèrent les premiers travaux sur la notion de système et sur la systémique en général [Von Bertalanffy 1968] [Le Moigne 1977] [Simon 1991]. Ces réflexions furent ensuite adaptées aux problèmes technologiques, d'ingénierie, d'architecture, d'économie, d'organisation... En effet, pour toutes les situations où apparaissent des difficultés de compréhension, de prise de décision individuelle ou collective, il convient d'essayer d'étudier le système qualifié de « complexe ». Un système est un ensemble d'éléments en interaction ; par opposition à un système qualifié de « compliqué » qui peut être compris en le décomposant en éléments simples analysables séparément (approche réductionniste classique), un système qualifié de « complexe » ne peut pas se réduire à la somme de ses parties [Atlan 1979]. Cette vision globale du système porte le nom d'holisme. En effet, lorsque le système global présente des propriétés qui ne sont pas directement déductibles de celles des éléments qui le composent, et que les informations apportées par les éléments pris dans le tout sont plus riches que celles issues d'éléments pris isolément, le fonctionnement global du système doit être étudié par simulation.

La théorie retenue est donc l'approche systémique. Les objets étudiés dans cet ouvrage sont des systèmes complexes et les outils utilisés sont les techniques informatiques. Mon objectif est de montrer comment des solutions informatiques permettent d'approfondir la compréhension de systèmes complexes, principalement en écologie. Toutes les réalisations effectuées reposent donc sur l'intégration d'outils logiciels pour la simulation de systèmes complexes.

2. LES ÉCOSYSTÈMES : ARCHÉTYPE DES SYSTÈMES COMPLEXES

On peut définir l'écologie comme une science qui étudie les conditions d'existence d'un être vivant ainsi que les rapports entre ce dernier et son environnement. La notion d'écosystème découle de cette définition : il s'agit d'un ensemble constitué par un milieu (sol, eau, etc.) et des êtres vivants, et entre lesquels existent des relations énergétiques trophiques (qui se rappor-

tent à la nutrition des tissus) [Gotelli 1998]. Un ensemble d'écosystèmes forme un complexe d'écosystèmes caractérisés par une origine commune ou des processus dynamiques communs [Coquillard et Hill 1997]. Dans son ouvrage sur l'apprentissage de la complexité, Gérard Clergue [Clergue 1997] terminait son chapitre sur l'apprentissage par simulation par le paragraphe suivant : « *En résumé, la simulation est la seule façon d'aborder de plain pied la connaissance des systèmes complexes dont l'archéotype pourrait être fourni par les écosystèmes.* » Le lecteur intéressé par une approche théorique de l'écologie pourra se reporter utilement aux ouvrages suivants : [May 1973] [Frontier 1977] [Roughgarden 1989] [Yodziz 1989] [Bulmer 1994] [Ågren et Bosatta 1996]. Nous présentons dans [Hill et Coquillard 2007] une vue d'ensemble réactualisée des travaux internationaux en matière de modélisation et de simulation des écosystèmes.

3. LA SIMULATION

Dans cet ouvrage, on désignera par « simulation informatique » ou plus simplement « simulation », les programmes informatiques permettant de faire évoluer un modèle discret en fonction du temps (un temps virtuel, appelé « temps de simulation »). En conséquence, les techniques décrites dans cet ouvrage ne s'apparentent pas aux programmes de résolution numérique de modèles continus basés sur des systèmes d'équations différentielles (eux aussi appelés « simulations »), même si elles peuvent, dans le cadre de simulations hybrides ou de multi-modèles par exemple, utiliser ou collaborer avec de tels programmes.

11

Certains scientifiques considèrent que la simulation n'est pas un domaine de recherche. Il est vrai que la simulation est par essence un domaine appliquatif plus qu'un domaine de recherche fondamentale au sens propre avec ses innovations ; l'utilisation de cette technique permet cependant de faire avancer la recherche fondamentale dans de nombreux domaines. En effet, la simulation informatique est une activité essentiellement appliquée et elle force les collaborations pluridisciplinaires qui sont souvent le cadre d'échanges fructueux. Enfin, pour un chercheur dans ce domaine, il reste rassurant de savoir que le CNRS attache, dans ses communiqués sur les procédures d'évaluation, autant d'importance au développement d'applications logicielles complexes qu'à la recherche fondamentale.

L'intérêt grandissant pour les techniques de modélisation dans de nombreux domaines scientifiques est essentiellement motivé par les possibilités de prédictions qui leur sont associées. Toutefois, cet espoir est souvent déçu. L'intérêt majeur que l'on peut retirer d'un modèle reste, à notre avis, l'approfondissement de la connaissance des scientifiques qui ont collaboré à sa construction, ou bien exploré de nombreux scénarios à l'aide de ce modèle. Cependant, n'omettons pas trop vite l'aspect prévisionnel : les modèles prédictifs sont nécessaires, la simulation possède intrinsèquement une réelle capacité d'aide à la décision et on peut élaborer des modèles à des coûts raisonnables. De ce fait, les besoins en simulation connaissent actuellement un essor, d'autant plus que l'infographie, la réalité virtuelle et la possibilité d'exécuter des modèles sur le Web permettent de communiquer avec le grand public ou avec des décideurs politiques. Cependant les scientifiques restent attachés aux résultats statistiques, plus exploitables pour affiner ou remettre en question leurs connaissances.

12

Les systèmes étudiés par simulation, même très complexes, voire chaotiques, ont une organisation interne cohérente, mais encore faut-il être capable de la spécifier. Les aspects statiques sont souvent assez faciles à décrire, mais la description du comportement dynamique d'un système est beaucoup plus délicate à réaliser. Les comportements dynamiques sont étudiés en exécutant les modèles dans un cadre expérimental réaliste. Les descriptions mathématiques ou algorithmiques des modèles permettent de représenter les changements d'état d'un système au cours du temps, ces changements pouvant être déterministes ou stochastiques dès lors que l'on introduit une part de hasard à l'aide de nombres pseudo-aléatoires. Lorsque l'on simule un système, on cherche à prévoir à partir de l'état de ce système à l'instant ' t ', ce que pourrait être son état à l'instant ' $t+dt$ '. Pendant une phase (souvent longue) de mise au point, de calibration et de remise en question d'un modèle, les résultats obtenus sont confrontés à la réalité. S'ils sont en accord avec la réalité, on suppose que le modèle explique certains mécanismes du système réel. Dans les meilleurs cas, il est parfois possible d'utiliser le modèle pour tenter des « prédictions », mais dans la majorité des cas c'est uniquement le rôle exploratoire des modèles qui sera d'un grand secours au scientifique. Même lorsque le modèle donne des résultats prédictifs médiocres, sa construction aura néanmoins contribué à parfaire la connaissance des experts qui pourront peut-être identifier une mauvaise hypothèse sur le fonctionnement du système réel.

4. LA MODÉLISATION D'ÉCOSYSTÈMES

Les modèles mathématiques d'écosystèmes ont déjà largement été étudiés, mais ils ont rapidement montré leurs limites (le lecteur intéressé peut se reporter aux ouvrages suivants : [Maynard Smith 1974] [Pielou 1977] [Okubo 1980] [Vandermeer 1981] [Levin 1989] [Hallam et Levin 1986] [Jeffries 1989] [Pavé 1994]. Ces modèles sortent du cadre de notre domaine d'étude, tout comme les travaux théoriques de modélisation pour l'écologie des populations qui sont par ailleurs bien décrits dans la littérature : [Pielou 1974] [Nisbet et Gurney 1982] [Caswell 1989] [Renshaw 1993].

Mon expérience m'a conduit à me spécialiser dans la simulation stochastique à événements discrets. Une des meilleures références en français sur le sujet reste [Leroudier 1980]. L'introduction d'éléments stochastiques dans les modèles est bien une conséquence des limites de notre connaissance. Dans [Stewart 1992] et [Clergue 1997] nous trouvons : « *Spinoza écrivait : "Il n'y a rien d'aléatoire dans la nature... une chose paraît aléatoire seulement à travers l'insuffisance de nos connaissances". Ainsi conçu le déterminisme implique une prédictibilité absolue pour peu que toutes les causes soient connues. C'est la vision classique de la science de Newton à Einstein, pour qui Dieu ne joue pas aux dés* ». La simulation à événements discrets qui était encore peu utilisée en écologie dans le début des années 1990 [Keen et Spain 1992], commence maintenant à être très prisée par les écologues [Grimm 1999]. Cette technique de simulation permet non seulement une représentation structurelle du système étudié, mais surtout elle autorise la modélisation d'interactions stochastiques discrètes dans le temps et dans l'espace entre les différentes entités du système à modéliser [Schneider 1994]. Les phénomènes écologiques que nous modélisons présentent des discontinuités spatiales et temporelles difficiles à reproduire avec des outils purement mathématiques.

13

La modélisation d'écosystèmes est une entreprise complexe, de longue haleine, parsemée de pièges et de difficultés. L'établissement d'objectifs clairs dès la décision d'élaborer un modèle (qu'attend-on du modèle ?) évitera bien des modifications ultérieures hasardeuses. Pour un modèle d'écosystème, les choix du niveau d'abstraction et du degré de détail sont cruciaux. L'expression de Frontier : « ni réductionisme, ni holisme », pourrait être la devise des concepteurs de modèles [Frontier 1977]. Un des points les plus délicats dans la modélisation constitue à définir une fermeture de l'écosystème, alors que celui-ci par essence se trouve être ouvert. Cette fermeture constitue l'hy-

pothèse simplificatrice la plus forte. Quels que soient les choix, la simplicité restera une vertu en matière de modélisation.

En ce qui concerne le niveau d'abstraction, notre choix a été celui des modèles individus centrés [Huston *et al.* 1988] [De Angelis et Gross 1992] [Breckling et Müller 1994]. Une étude récente de Volker Grimm publiée dans la revue *Ecological Modelling* présente une synthèse de l'utilisation des modèles individus centrés lors de ces dix dernières années [Grimm 1999]. Notre thématique de recherche aborde précisément ce type de modèle, encore appelés « Individual Based Models » (IBM) dans la terminologie anglo-saxonne. Si par contre, on a retenu le niveau d'une population biologique, il y a fort à parier que dans un grand nombre de cas, on soit néanmoins amené à recueillir des données relatives aux individus. En effet, le comportement de la population (résultat attendu) résulte des interactions individuelles. Dans d'autres cas, on pourra effectuer une agrégation, c'est-à-dire simuler le comportement global de la population au moyen d'une fonction que l'on estime suffisamment représentative de celui-ci. Les individus eux-mêmes ne sont pas exempts d'influences provenant des niveaux supérieurs de la hiérarchie (autres populations, facteurs de l'environnement). Doit-on prendre en compte ou non ces facteurs ? Lesquels d'entre eux peuvent-être considérés comme négligeables ? Il n'y a pas de règle absolue dans ce domaine, ce qui peut faire dire que la modélisation s'apparente quelque peu à un art.

14

La qualité des données écologiques constitue un deuxième impératif : leur adaptation aux objectifs fixés, bien sûr, mais aussi la qualité de leur échantillonnage. L'écologie de terrain est saisonnière. Quoi de plus désappointant, en fin de saison, que de constater un nombre insuffisant de relevés ou d'expériences sur le terrain, des lacunes, des données manquantes, trop d'incertitudes dans les mesures, toutes choses qui ont pour effet de retarder encore la mise au point du modèle ?

Le choix de la technique à utiliser découle directement des objectifs de la modélisation. Méthodes analytiques et simulations discrètes n'offrent pas les mêmes possibilités. Seul l'expert du domaine peut influer sur ce choix par les objectifs qu'il fixe et les éléments qu'il souhaite prendre en compte.

Nous avons fait le choix d'utiliser pour ces modélisations des méthodes mêlant déterminisme et stochasticité. Nous pensons que ce sont précisément ces méthodes qui permettent d'approcher la complexité des écosystèmes en améliorant sensiblement le réalisme des modèles. De même, nous avons retenu les techniques à objets pour le développement de tous mes modèles

et, c'est également le choix retenu par de nombreux autres collègues pour des applications variées (ex : [Lhotka 1991] [Breckling et Müller 1994] [Baveco et Smeulders 1994]). En effet, en biologie et en écologie, l'approche orientée-objet, par ses concepts si proches des raisonnements en matière de taxonomie, est une aide précieuse ; elle s'adapte remarquablement bien aux modèles nécessitant la prise en compte de l'espace et des comportements individuels. Cette approche facilite largement la réutilisabilité et la maintenance, elle oblige l'expert (l'écologue) à un recensement exhaustif et hiérarchisé de tous les facteurs qu'il souhaite prendre en compte, et elle facilite grandement le dialogue entre le biologiste et l'informaticien qui trouvent ici un langage commun. Ce dernier point est un élément de garantie pour limiter les ouboris et ne négliger aucun des aspects importants d'un écosystème.

Malgré tout, l'apprentissage par un biologiste ou un écologue d'une méthode d'analyse et de conception par objets, ainsi que d'un langage à objets reste un sérieux investissement en temps et en efforts d'abstraction ; mais les bénéfices compenseront largement les sacrifices. Par ailleurs, l'obtention, l'expression et l'analyse des résultats doivent être adaptées à la méthode. L'expression des résultats doit faire l'objet d'un soin tout particulier. Ainsi, dans le domaine des méthodes de simulations discrètes stochastiques, l'usage d'intervalles de confiance est impératif. On ne peut que regretter le peu d'attention porté aux phases de validation en matière de modélisation d'écosystèmes [Hill 1995b]. À la décharge des modélisateurs, la validation de modèles d'écosystèmes ne peut pas être poussée aussi loin que dans le domaine industriel où l'ensemble des paramètres peuvent être spécifiés par l'homme. Une attention toute particulière doit être cependant portée à la validation. En effet, la validation par confrontation avec la réalité prend ici toute son importance.

5. CONTEXTE DE NOS TRAVAUX DE RECHERCHE

La thématique de recherche en modélisation et en simulation pour les Sciences de la Vie s'inscrit dans le cadre de la participation du LIMOS aux actions de recherche interdisciplinaires. Après avoir consacré les années 1990 à 1993 à la simulation visuelle par objets, de systèmes de production, de transports, de systèmes informatiques ou de systèmes complexes en général, nous avons

estimé que ces techniques pouvaient s'appliquer à la modélisation de certains écosystèmes où les aspects spatiaux limitent le champ d'application des outils mathématiques usuels. Ainsi depuis 1993, nous avons recentré notre thématique sur la modélisation des écosystèmes et du vivant en tant que systèmes complexes naturels.

La résolution des problèmes posés par la modélisation d'écosystèmes nécessite de nombreuses collaborations. Les principaux laboratoires avec lesquels nous avons pu collaborer sont : le Laboratoire d'Ecologie Végétale et Cellulaire de l'Université d'Auvergne, le Laboratoire d'Environnement Marin Littoral de l'Université de Nice Sophia-Antipolis, plusieurs unités de l'INRA et du CEMAGREF.

6. ORGANISATION DE L'OUVRAGE

Le chapitre qui suit présente les concepts généraux que nous manipulons. Les notions de système, de modèle et de complexité sont abordées. Le rôle d'un modèle et celui de la simulation sont également traités avant de discuter les choix méthodologiques.

16

Le chapitre 3 présente nos activités de recherche lors de mon doctorat. Celui-ci était consacré au développement d'outils logiciels pour la modélisation et la simulation de systèmes complexes. Au début des années 1990, nous avons conçu et réalisé des outils de simulation visuelle par objets pour des systèmes à flux discrets (de production, de transports, de systèmes informatiques et des systèmes administratifs). Nous avons également proposé un processus et une méthode de modélisation par objets et un environnement de programmation visuelle de modèles de systèmes à flux discrets. Cet environnement était capable, à partir d'une saisie graphique et interactive, de générer automatiquement du code pour différents outils et langages de simulation (Siman IV, Qnap 2, Simula 67) en se basant sur le formalisme des réseaux de files d'attentes. J'avais également développé un outil d'animation de résultats de simulation reposant sur le même type de programmation visuelle. Les outils développés avaient été validés sur de nombreux systèmes de production du groupe Valeo et commercialisés par la société Simulog. La capacité à générer automatiquement des codes de simulation à partir de spécifications graphiques et grâce à des méta-modèles, était préconisée et mise en œuvre depuis le début des années 1990. En 2000, l'Object Manage-

ment Group a introduit des concepts similaires au sein de son architecture dirigée par les modèles (MDA : Model Driven Engineering) ; les techniques de transformation mise en œuvre de façon plus générale dans l'Ingénierie de Modèles au début du nouveau millénaire sont fortement inspirées de ces travaux.

Le chapitre 4 présente synthétiquement les principaux projets de modélisation que nous avons réalisés et encadrés en écologie terrestre et en océanographie. Les détails techniques sont omis afin de ne pas alourdir cet ouvrage. Pour chaque projet l'apport principal réside dans le développement et l'intégration concrète de techniques logicielles pour la simulation d'écosystèmes. Nous avons été amené à associer le Web, les Systèmes d'Information Géographique, l'infographie, les outils d'analyse statistique et la conception de plans d'expérience. Étant donnée la grande complexité des écosystèmes étudiés, nous préconisons dans un grand nombre de situations, l'utilisation de modèles utilisant la simulation à événements discrets couplés à des Systèmes d'Information Géographique. Cette approche présente le double avantage de pouvoir prendre en compte à la fois les phénomènes spatiaux discontinus des écosystèmes et les aspects individuels (intégrant aussi bien la diversité génétique des individus que les aspects sociaux des groupes d'individus).

Le chapitre 5 est consacré exclusivement aux problèmes de validation des modèles. La complexité des modèles d'écosystèmes nécessite un intérêt particulier pour les techniques de validation et de vérification applicables au cours du cycle de développement d'un modèle de simulation stochastique, ainsi que lors de l'analyse statistique des résultats (analyse spectrale, détermination d'intervalles de confiance, etc.). Pour tout modèle d'écosystème réel, il est difficile d'obtenir des résultats de qualité. Les résultats des simulations stochastiques sous contraintes spatiales produisent néanmoins des cartes utiles et une technique d'analyse spectrale de séries de cartes est présentée. Les outils de visualisation et d'animation des résultats sont abordés car nous avons travaillé sur ces aspects qui sont des atouts supplémentaires pour la validation des modèles et des résultats.

Le dernier chapitre conclura en présentant une réflexion sur les travaux de simulation réalisés et abordera également un ensemble de perspectives de recherche.

CHAPITRE 2

RÉFLEXIONS SUR LA MODÉLISATION ET LA SIMULATION

*Là où il n'y a pas d'amour,
semez de l'amour et vous recolterez de l'amour.*

Saint Jean de la Croix

1. INTRODUCTION

L'étude de systèmes qui semblent complexes à nos pauvres esprits constitue une dimension essentielle de l'approche scientifique. La réalisation d'une représentation simplifiée qui aide à comprendre le fonctionnement du système est alors l'essence de l'activité de modélisation. Cette activité est fréquemment pluridisciplinaire car, en effet, elle nécessite le rapprochement d'outils, de techniques, de méthodes et d'experts de différents domaines. C'est précisément dans l'intégration de techniques, d'outils et de méthodes pour divers domaines d'application que je situe mes travaux de recherche. Avant de présenter les applications que j'ai pu réaliser, je tiens à définir plusieurs termes génériques, notamment : système, modèle, simulation, complexité dont la signification peut varier entre deux informaticiens qui n'abordent pas le même thème de recherche.

2. QU'EST-CE QU'UN SYSTÈME ?

La notion de système embrasse une grande variété d'objets : système d'équations, système naturel, système mécanique... Aussi une définition unique du terme système est-elle malaisée à trouver. La définition suivante pourrait à la rigueur convenir à l'ensemble des entités que nous reconnaissons comme des systèmes : un système est une collection d'objets en interactions. Même si les systèmes statiques existent (minéraux, roches, métaux...), les plus intéressants sont naturellement les systèmes dans lesquels les interactions provoquent fréquemment des changements d'état.

Il est trivial de constater qu'en matière de systèmes « naturels », il n'existe pas de systèmes juxtaposés, indépendants – isolés au sens thermodynamique –, c'est-à-dire n'entretenant avec le milieu ambiant aucun échange d'énergie ni de matière ; les systèmes naturels sont dits ouverts. Par commodité, nous distinguons des sous-systèmes, c'est-à-dire des entités qui, au sein du système naturel, fonctionnent de façon apparemment autonome mais en entretenant des relations avec le reste du système naturel. Ceci conduit naturellement à considérer un système et son environnement. Ainsi, les écosystèmes dont nous parlerons sont-ils des sous-systèmes de notre biosphère.

20

Un système est donc constitué d'un ensemble d'objets en interaction, constituant autant d'entités du système caractérisées par un (des) attribut(s) et une (des) activité(s). Nous appelons activité tout processus susceptible de changer l'état du système. Pour prendre un exemple simple, un végétal constitue une entité caractérisée par les attributs taille, type de feuillage, etc. Ses activités seront essentiellement d'absorber des substances nutritionnelles, de croître et de se reproduire. Chacune de ces activités est susceptible de modifier l'état d'un système « collection de végétaux », comme une forêt par exemple. On appelle état d'un système la description de l'ensemble des entités, attributs et activités qui le composent à un instant 't' donné.

Un système, nous l'avons mentionné, entretient des relations avec son environnement. Les activités de l'environnement qui peuvent affecter l'état du système sont dites exogènes par opposition aux activités endogènes, *i.e.* internes au système.

Par ailleurs, l'activité d'un système semble pouvoir être déterminée complètement par les entrées du système. Ainsi, il semble que l'on peut décrire le développement d'un arbre en connaissant l'intensité lumineuse, les ressour-

ces en nutriments et en eau comme paramètres d'une loi de croissance. Un tel système est dit déterministe.

À y regarder de plus près, cette hypothèse doit être levée car bien des entrées du système varient aléatoirement, notamment les paramètres climatiques, les attaques parasites et les activités humaines qui vont perturber la croissance de l'arbre. De même, un système de production automatisé sera sujet à des pannes. L'activité résultante comporte donc une part d'aléatoire et le système est dit stochastique. En réalité un système n'est jamais totalement déterministe ou stochastique, mais comporte une part des deux types d'activités. On dira qu'un système est stochastique si son activité comporte au moins un processus stochastique, et il sera dit déterministe dans le cas inverse.

Enfin, une dernière distinction doit être faite entre les systèmes continus et discrets. Typiquement, la croissance d'un arbre au cours de la saison de végétation constitue une activité continue – sans à-coups. Par opposition, certaines activités peuvent modifier le système de manière discontinue à certaines dates du déroulement de l'activité. Par exemple, l'assemblage de pièces au sein d'un système de production est une activité discrète, puisqu'elle intervient à date fixe et fournit un nombre entier de pièces. On devra cependant porter attention à la nature réelle de cette activité. Si l'on s'intéresse par exemple à la reproduction au sein d'une population de bactéries, l'activité « reproduction » pourra à la rigueur être considérée comme continue en raison du temps très court de génération, du nombre considérable d'individus en jeu et de la désynchronisation des divisions bactériennes. Mais il s'agit bien, en dernière analyse, d'une activité discrète. À l'inverse, au sein d'une population d'animaux supérieurs, cette activité, considérée comme continue chez les bactéries, devra préférentiellement être considérée comme discrète.

3. COMPLEXITÉ DES SYSTÈMES ET INTÉRÊT DE LA MODÉLISATION

La réalité est souvent beaucoup plus complexe que ce que nous imaginons. L'illusion de pouvoir un jour tout expliquer scientifiquement est tenace, et nos tentatives de compréhension, de classification sont parfois dérisoires. Il faut définitivement raisonner dans des espaces multivariés, c'est-à-dire possédant de nombreuses variables avec toutes leurs interactions ; il

faut alors pouvoir explorer ces espaces notamment en essayant de trouver les variables sensibles, en agissant sur elles afin d'améliorer notre compréhension des phénomènes observés. L'analyse statistique de données permet d'analyser des situations complexes, telles qu'elles sont dans la réalité (ou presque), et d'étudier simultanément plusieurs variables et plusieurs facteurs de variations. Il faut cependant être conscient que si l'on ne dispose pas d'un minimum de compréhension d'un système toutes les statistiques que l'on pourrait obtenir n'auront aucun contenu sémantique. Par contre, dans de nombreux cas, il n'est pas possible ou il est relativement difficile d'obtenir des données sur un système réel complexe, la construction d'un modèle sera alors toujours très difficile. Toutefois elle s'impose comme technique exploratoire d'hypothèses.

Dans tous les domaines d'application des techniques de simulation que nous avons pu aborder lors de ces dix dernières années, la mesure de la complexité reste difficile à quantifier [Cellier 1991]. En effet, comment pouvons nous déterminer qu'un système est plus complexe qu'un autre ? La taille du système, la quantité d'informations qu'il manipule, le comportement « imprévisible » ou « chaotique » sont souvent des caractéristiques mises en avant pour justifier du caractère complexe d'un système. En réalité la complexité peut commencer avec des systèmes apparemment « très simples ». Ce qui signifie qu'il faut différencier la complexité du système, que l'on peut associer grossièrement au nombre de relations internes (l'aspect « compliqué » [Atlan 1979]) et la complexité du comportement, qui est en fait le résultat du fonctionnement du système. Des équations simples telles que celle présentée en 1848 par Verlhust pour modéliser la croissance d'une population a fait émerger un comportement complexe (bifurcation). Ainsi il apparaît clairement que la complexité ne saurait véritablement se mesurer comme une grandeur du nombre d'attributs ou de relations affectant les entités d'un système. Il est désormais admis par l'ensemble de la communauté scientifique que les systèmes complexes (ne pouvant être analysés avec une approche réductionniste), se caractérisent par des phénomènes de réitération au cours desquels se produisent un auto-contrôle, une régulation du résultat. Une des particularités des systèmes écologiques est qu'ils reposent sur de tels types de rétrocontrôles agissant dans un environnement plus ou moins stable mais fluctuant sur de courtes périodes (journalières, saisonnières, annuelles). Les rétrocontrôles des effectifs de population s'effectuent par exemple par le biais des naissances et des disponibilités des ressources, par les relations proie-prédateur, etc. L'étude des phénomènes de régulation se rapporte à la la dynamique des systèmes [Forrester 1961].

Nous pensons, tout comme [Legay 1996], que l'organisation interne d'un système est à même de justifier la complexité. En effet, le concept de complexité est lié pour un nombre croissant de scientifiques au concept d'organisation. Les scientifiques sont de plus en plus persuadés de l'organisation du monde vivant et des systèmes naturels. Einstein ne se demandait-il pas : « *Pourquoi le monde est-il compréhensible, au lieu de nous apparaître comme un chaos sans ordre lisible ?* »

L'organisation des systèmes se présente souvent sous la forme de réseaux de relations et de structures s'articulant entre elles, régies par un ensemble de contraintes. Il n'est pas étonnant de voir le succès des méthodes d'analyse et de conception par objets qui reposent toutes sur des schémas permettant de mettre en évidence l'organisation des systèmes sous forme de classes d'objets avec leurs relations. Les avancées scientifiques dans le domaine des systèmes naturels, les plus complexes que nous connaissons, font découvrir de nouveaux mécanismes, de nouvelles contraintes plus précises et plus fines, de nouvelles relations qui n'avaient pas été envisagées. Il est alors possible de donner une nouvelle représentation du système présentant une organisation plus fine, statique et dynamique, quitte à remettre en cause les recherches antérieures ce qui est le propre de la recherche scientifique. Nos connaissances humaines dans tous les domaines sont limitées ; le « flou artistique » qui règne lorsque des « informaticiens » posent aux experts d'un domaine des questions précises pour construire leur modèle reflète plus l'image de notre ignorance collective que celle de la non-organisation. Comme nous l'avons dit, un des principaux gains que nous obtenons en réalisant des modèles est précisément l'augmentation de la connaissance, qui conduit à une meilleure spécification de l'organisation du système. Même si cette amélioration est infime, elle devrait être le but premier de toute recherche qui utilise des modèles.

23

Les modèles informatiques sont utilisés non seulement pour la conception d'outils ou de systèmes, mais leur intérêt est unique lorsqu'on essaye de « percer » toutes les grandes énigmes scientifiques, que ce soit en astrophysique, en biochimie, ou plus généralement pour les Sciences de la Terre et de la Vie. Dans certaines situations et dans certains domaines, des états d'équilibre peuvent apparaître. On a également appris que les variations pouvaient être importantes sans entraîner de catastrophes, tant que le retour à ces états d'équilibre était assuré. Par contre dans de nombreuses autres situations l'irrégularité des phénomènes renvoie les chercheurs à leurs modèles, et ceci quel que soit le domaine abordé. La constante reste que, dans tous les cas,

la modélisation informatique est un outil précieux pour étudier et découvrir les détails des processus qui régissent des systèmes dits complexes.

« La variabilité du monde vivant n'est pas le chaos ; on la sent liée au milieu physique, au climat, à sa propre histoire ; » ... « Les variations les plus fines suivent elles-mêmes des lois complexes, et c'est leur exploration qui constitue le champ habituel de la science. Mais ces variations ne sont pas isolées, indépendantes ; elles sont celles des éléments d'un système et c'est leur approche multiple et simultanée qui rend maintenant nécessaire l'usage des moyens informatiques comme outils de la recherche, tant au niveau technique que méthodologique. » [Legay 1996]

4. QU'EST-CE QU'UN MODÈLE ?

4.1. *La notion de modèle*

Avant d'aborder la notion de modèle proprement dit, il convient d'attirer l'attention sur un point particulier. Il ne faudra pas confondre la nature d'un système et celle de son modèle. Ainsi, un système peut être continu et déterministe, mais le modèle discret et stochastique. L'introduction d'éléments stochastiques dans un modèle reflète, dans bien des cas, notre incapacité à modéliser l'ensemble d'une activité dont il faudrait une connaissance très approfondie pour la reproduire fidèlement. La reproduction des événements climatiques en est un bon exemple.

La notion de modèle n'est pas récente en science. Ainsi le modèle astronomique héliocentrique de Copernic révolutionna-t-il la conception géocentrique d'inspiration aristotélicienne et judéo-chrétienne. Le modèle copernicien fut lui-même par la suite enrichi pour être remplacé par le modèle planétaire actuel de notre système solaire. On le voit aisément, les modèles ne sont pas destinés à survivre indéfiniment. C'est qu'ils représentent, à un instant donné, la somme des connaissances accessibles afférentes à un domaine particulier. Que l'expérimentation ou l'observation viennent à prendre en défaut le modèle, et tout est à reconsidérer... Il s'agit là du processus même de l'avancement de la science. Poser une hypothèse – ou une série d'hypothèses – relative à un phénomène observable et mesurable fonde l'acte de la modélisation. L'ensemble des observations ultérieures du système réel en fonctionnement permettra la validation ou l'invalidation du modèle.

La notion de modèle est si intimement liée à la pensée scientifique que nous pouvons nous surprendre très souvent à confondre modèle – l'idée que nous nous faisons de... – et réalité, au point que cette prise de conscience nous demande parfois quelques efforts. En un mot, un modèle est une abstraction qui simplifie le système réel étudié en ignorant de nombreuses caractéristiques de celui-ci, pour se focaliser sur les aspects qui intéressent le modélisateur et qui définissent la problématique du modèle.

En 1965, Marvin Minsky proposait la définition suivante :

« *To an observer B, an object A* is a model of an object A to the extent that B can use A* to answer questions that interest him about A* »
[Minsky 1965]

C'est le moins que l'on puisse demander à un modèle, direz-vous ! Pourtant, cette phrase est bien plus lourdement chargée de sens qu'il n'y paraît. Si le modèle nous permet d'apprendre quelque chose d'utile sur le fonctionnement du système, l'observateur peut se considérer comme satisfait. Voilà qui peut lever bien des réticences et des préjugés à propos de la modélisation. Que l'on se reporte au modèle logistique : les bifurcations, les régimes cycliques et le comportement chaotique furent découverts au travers de ce modèle bien avant d'être reconnus dans les systèmes réels [Gleick 1989]. Pourtant, bien peu de systèmes réels peuvent être modélisés très exactement par l'équation logistique. C'est bien sous l'impulsion de la modélisation que ces phénomènes ont été découverts dans la nature. Deux autres exemples : la double hélice de l'ADN (acide désoxyribonucléique) de Watson et Crick ainsi que les α -hélices des chaînes polypeptidiques proposées par Pauling ont été des modèles avant que de trouver confirmation dans l'expérimentation. Ici, la théorie peut précéder l'observation. Si le modèle reproduit fidèlement en termes quantitatifs le système réel, tout est pour le mieux. S'il ne le peut, contentons-nous tout d'abord d'explorer ses divers comportements et interrogeons-nous sur la validité de ceux-ci dans le réel. En tout cas, aurons-nous appris quelque chose, à savoir que nos connaissances du système à modéliser sont encore insuffisantes.

25

4.2. Modèles discrets, continus, stochastiques, déterministes

Les critères de classification évoqués précédemment pour les systèmes s'appliquent aux modèles : suivant la technique de représentation des changements d'états au sein d'un modèle celui-ci va être considéré comme étant discret ou continu. Si les changements d'états du modèle s'opèrent de manière

continue au cours du temps on parle de modèle et de simulation continue. Une simulation de barrage (remplissage, vidange,...) sera effectuée par un modèle continu reposant sur des équations différentielles. Lorsque les changements d'états interviennent seulement à des dates précises, appelées dates d'occurrence des événements, on parle de modèle discret et de simulation à événements discrets. Il est également possible de coupler des modèles discrets à des modèles continus, on parle alors de modèles combinés. Les derniers travaux de Zeigler en sont un bon exemple [Zeigler *et al.* 2000].

Les techniques de simulation discrète ne sont pas limitées par une formalisation mathématique du problème. Cependant elles sont aussi les plus complexes à mettre en œuvre et les plus gourmandes en temps de calcul.

La distinction que nous avions faite pour les systèmes vaut pour les modèles. Lorsqu'un modèle fait apparaître explicitement le hasard (en utilisant par exemple des tirages de nombres pseudo-aléatoires distribués suivant des lois de probabilités), il est classé parmi les modèles stochastiques. Par opposition, un modèle déterministe ne reproduit pas de comportements aléatoires. Nous avions vu que les systèmes réels qui ne présentent pas d'incertitudes de fonctionnement sont rares (qu'elles soient dues à la complexité interne du système ou à son environnement) ; c'est pourquoi les modèles stochastiques (ou encore probabilistes) sont les plus intéressants. De plus, ces modèles permettent de compenser notre manque de connaissance en introduisant volontairement des fluctuations aléatoires guidées par des lois de probabilités. Il est par exemple souvent possible de trouver la loi de distribution des graines autour du tronc d'un arbre, ou encore la loi de distribution des arrivées d'appels téléphoniques sur un central. On peut ainsi, grâce à la génération de nombres pseudo-aléatoires reproduire des phénomènes complexes pour lesquels une mise en équation totale est encore impossible. Il se peut cependant qu'en fonction des objectifs fixés on ait décidé d'ignorer les phénomènes aléatoires s'ils ont un impact négligeable et, dans ce cas, le choix d'un modèle déterministe est tout à fait justifié. Un modèle déterministe ne contient plus que des relations certaines. Un modèle de croissance de la population mondiale basé sur une série mathématique peut être un modèle simple et déterministe, mais si les objectifs sont d'avoir un modèle prédictif, il n'est pas certain que ce soit le meilleur modèle.

4.3. *Les qualités d'un modèle*

La dernière caractéristique fondamentale que nous voulons mettre en évidence pour un modèle réside dans le fait qu'il est construit en fonction d'un ensemble d'objectifs. Ce sont ces objectifs qui déterminent les hypothèses de modélisation et le degré de simplification du modèle. Finalement, qu'est-ce donc qu'un modèle ? Sa principale caractéristique est d'être une simplification de la réalité, une approximation qui, bon an mal an, reproduit à peu près la réalité. [Popper 1973] relève trois caractéristiques communes à tous les modèles :

1. Un modèle doit avoir un caractère de ressemblance avec le système réel,
2. Un modèle doit constituer une simplification du système réel,
3. Un modèle est une idéalisation du système réel.

auxquelles nous ajouterons : un modèle est dépendant des objectifs fixés par la problématique d'une étude précise et c'est dans ce cadre qu'il doit reproduire le mieux possible le comportement du système réel (*cf.* premier point de Popper).

27

5. À QUOI SERT UN MODÈLE ?

5.1. *Questions et problèmes*

Nous pourrions esquiver la question en affirmant qu'un modèle doit satisfaire les objectifs qui lui sont associés. Détaillons cependant le sujet en factorisant les objectifs communs à toute modélisation. Un modèle sert principalement à répondre à des questions et dans le meilleur des cas à résoudre un certain nombre de problèmes posés par un système (qualifié de complexe du fait de sa taille ou de son organisation), ainsi que, comme nous l'avons déjà souligné, à améliorer notre connaissance sur ce système. En effet, la complexité engendre des interrogations et des problèmes, aussi bien lors de l'étude et de la conception que lors de l'exploitation d'un système. Citons quelques exemples en modélisation d'écosystèmes :

- la détermination d'un dimensionnement adapté : par exemple la taille des zones à laisser en friche, le nombre d'animaux à placer sur une estive (notion de chargement), le nombre de limaces nécessaires pour une lutte biologique contre une autre espèce ; si l'on considère les

systèmes de production, de transport, les systèmes informatiques et administratifs... ; les problèmes à résoudre concernent : le nombre de machines, de camions, de processeurs ou de fonctionnaires qu'il faut prévoir, la taille des zones des stocks-tampons, de la mémoire cache, des aires de parking ou des salles d'attentes,...

- la compréhension de leur fonctionnement : comment les ovins organisent-ils leur mémoire, comment s'en servent-ils ? L'algue tropicale *Caulerpa taxifolia* peut-elle se déplacer avec des courants alors que sa flottabilité est négative ? Combien peut-on ajouter de chevaux sur une estive pâturée par des bovins sans induire trop de perturbations... ?
- l'amélioration de leur productivité : qu'il s'agisse d'estives pâturées par des vaches, des agnelles ou des chevaux, ou d'une cellule d'assemblage flexible dans une usine, comment peut-on s'assurer que les ressources sont utilisées au mieux ? N'y a-t-il pas des ressources actives qui passent plus de temps à attendre qu'à produire, des zones non exploitées par les vaches qui pourraient l'être par des chevaux ?
- les problèmes de maintenance : doit-on arrêter la traite des vaches pendant certaines périodes pour éviter l'apparition de mammites ? Doit-on « forcer » une personne à prendre des vacances avant qu'elle ne tombe malade, et avec quelle fréquence ?
- les problèmes d'aléas : qu'il s'agisse de propagation de boutures au gré des vents ou des marées, de pannes dans un matériel de production automatisé, il est intéressant de connaître le comportement stochastique d'un système. Dans un système naturel, la prise en compte des aspects stochastiques est souvent nécessaire pour calibrer un modèle avec des données réelles.
- les problèmes d'ordonnancement des flux dans les systèmes : existe-t-il un ou plusieurs ordonnancements qui améliorent le rendement d'un système ? Que ce soit au niveau de l'alimentation en fourrage d'un troupeau de vaches à lait ou du déplacement des troupeaux sur plusieurs estives, il convient d'essayer d'optimiser la production.
- ...

28

Pour appréhender le comportement d'un système et donc pour essayer de résoudre les problèmes évoqués précédemment, on essaye de recueillir de l'information, de prendre des mesures, tout en sachant que toute prise de mesure perturbe le système étudié. L'analyse des valeurs mesurées conduit à une meilleure compréhension du système étudié. Dans certains cas, la prise

de mesure est impossible, trop complexe ou trop coûteuse à mettre en œuvre et on ne peut comprendre le fonctionnement du système qu'à partir d'un modèle. Si le système réel n'existe pas, et que l'on cherche à le concevoir, on parlera de modélisation « *a priori* », ce qui n'empêche pas de chercher à évaluer et à améliorer les « performances du système futur ». Évaluer signifie « déterminer une quantité par le calcul sans recourir à la mesure directe ». L'évaluation est donc toujours effectuée à l'aide d'un modèle. Les modèles de simulation sont souvent utilisés en tant qu'aide pour la prise de décisions sur des projets coûteux. Les utilisateurs de ces modèles doivent donc être en mesure de faire confiance aux modèles. Pour que cette confiance soit justifiée, il est toujours nécessaire de vérifier et de valider ces modèles en fonction des objectifs que l'on s'est fixés. Nous reviendrons plus tard sur les notions de vérification et de validation.

En matière d'écologie, les connaissances acquises ces dernières décennies ont connu une croissance spectaculaire corrélativement à l'acquisition de nouvelles techniques d'échantillonnage (télédétection et imagerie spatiale ; suivi radiogoniométrique des animaux ; automatisation de l'acquisition de données physico-chimiques de l'air et de l'eau, etc.), de techniques d'analyse numérique (analyse statistique de données multidimensionnelles, analyse de séries), ainsi que de l'outil informatique (matériel et logiciel). Dans le même temps, ce flot de données fit prendre conscience à un nombre croissant de décideurs qu'une organisation de la gestion des activités humaines était nécessaire afin d'améliorer la gestion de notre environnement.

Face à l'insoudable complexité du fonctionnement des écosystèmes, seules les techniques de modélisation, associées aux performances croissantes de l'outil informatique, permettent dans un certain nombre de cas de proposer des aides fiables à la gestion raisonnée de notre environnement : modèles de gestion des quotas de pêche, de croissance forestière, de fonctionnements lacustres, de dispersion d'émissions polluantes, de prévision d'extension des incendies, etc. Les besoins sont ressentis aujourd'hui comme considérables.

5.2. *Les modèles sont des instruments scientifiques*

Les modèles écologiques ne diffèrent pas fondamentalement – pas même par leur complexité – des modèles développés dans d'autres disciplines. L'in-croyable complexité que présente le fonctionnement d'un écosystème ne peut être saisie par la simple acquisition de l'ensemble des paramètres qui le caractérise. Or, pendant longtemps l'écologie s'est réduite à cette seule

démarche descriptive. Cette démarche fut une étape nécessaire, mais les multiples interactions et rétroactions (effets feed-back) au sein de tels systèmes font apparaître des comportements que l'on ne saurait saisir par la simple juxtaposition des données recueillies sur l'ensemble du système, furent-elles exhaustives. En d'autres termes, le comportement du système n'est pas équivalent à la somme des comportements des parties. Seule une modélisation mettant en interaction les différentes parties du système peut faire apparaître les comportements émergeants. Il n'est donc pas surprenant que les modèles en matière d'écologie soient de plus en plus utilisés pour une meilleure compréhension des écosystèmes [Pavé 1994]. Jorgensen résume en quatre points les avantages de la modélisation [Jorgensen 1994] :

- Les modèles ont leur utilité dans la surveillance de systèmes complexes.
- Les modèles peuvent être utilisés pour révéler les propriétés des systèmes écologiques.
- Les modèles peuvent montrer des carences dans nos connaissances et être utilisés pour définir des priorités dans la recherche.
- Les modèles sont utiles pour tester des hypothèses scientifiques, dans la mesure où le modèle peut simuler les réactions de l'écosystème, lesquelles peuvent être comparées aux observations.

30

6. LA SIMULATION : LE MODÈLE PLONGÉ DANS LE TEMPS

6.1. *Introduction*

Beaucoup d'auteurs associent le terme de simulation à une technique de résolution de problème. Ainsi, la génération d'une variable aléatoire X de distribution exponentielle négative par la méthode de l'anamorphose est appelée simulation de la variable X. Nous considérons d'une façon plus générale que la simulation consiste à faire évoluer le modèle d'un système au cours du temps et que ce n'est pas de la méthode de génération des événements qui fait évoluer ce modèle, cette vision purement technique constituant à notre avis une restriction. La simulation associe étroitement modèle et temps. Nous retenons la définition suivante de la simulation :

« La simulation consiste à faire évoluer une abstraction d'un système au cours du temps afin d'aider à comprendre le fonctionnement et le compor-

tement de ce système, et à appréhender certaines de ses caractéristiques dynamiques dans l'objectif d'évaluer différentes décisions. » [Hill 1993b].

6.2. Simulation par objets, acteurs et agents

Rappelons qu'en introduction nous précisions qu'un système est considéré comme un « ensemble d'objets en interaction ». Après avoir abordé les concepts clés de la modélisation en esquissant des définitions de la notion de système et de modèle sans se préoccuper des techniques informatiques, nous allons utiliser la notion d'objet logiciel puis esquisser une introduction aux techniques de simulation par objets.

Le modèle objet, issu du monde de la simulation dès 1957 lors du projet du missile Minuteman [Tend Dyke et Kunz 1989], a été concrétisé à la fin des années soixante par Simula [Dahl *et al.* 1966]. Depuis, il a conquis la communauté du génie logiciel en se montrant adapté à la production industrielle de logiciels de qualité. Par rapport aux approches procédurales ou fonctionnelles, un logiciel utilisant les possibilités de la technologie à objets devient, selon Bertrand Meyer, un modèle opérationnel où les objets logiciels reflètent la réalité. À ce propos, nous tenons à citer un extrait de son ouvrage de référence : « *Tout cela est particulièrement frappant dans le domaine de la simulation. Ce n'est pas par accident que depuis Simula 67, la simulation est un domaine d'application privilégié des techniques à objets. Pour modéliser le monde réel en vue de le simuler, quoi de mieux que de décrire les objets à simuler.* » [Meyer 1990, p. 78]. La technologie à objets s'applique aussi bien aux simulations continues [Cellier 1991] qu'aux simulations discrètes [Hill 1993a].

31

La simulation d'un système complexe fait intervenir de nombreux processus concurrents et en interaction. Des objets actifs, autonomes et concurrents (implémentés par un parallélisme physique ou logique) sont communément appelés acteurs et sont aptes à représenter les processus d'un modèle de simulation. Le mode de communication de référence pour les acteurs reste l'envoi de messages asynchrones non bloquants et il existe bien sûr des extensions pour les communications bloquantes. Les acteurs tels qu'ils sont présentés ci-dessus et dans [Briot 1989] sont à différencier du modèle de calcul des acteurs introduit par Agha [Agha 1990]. Les modèles d'acteurs sont des modèles à objets particuliers, aptes à appréhender le parallélisme des systèmes réels, et susceptibles d'être implantés sur des architectures parallèles (multiprocesseurs ou autres). Plusieurs modes de concurrence peuvent

être isolés pour permettre une collaboration des acteurs fonctionnant en parallèle [Yonezawa *et al.* 1987]. L'utilisation de la programmation concurrente permet d'obtenir des simulations qui peuvent profiter des techniques à objets distribués et des réseaux à hauts débits. L'utilisation et la normalisation IEEE du DIS (Distributed Interactive Simulation) par le Département de la Défense américaine en est un exemple concret.

Une évolution des acteurs a conduit à la notion d'agent. Le but initial de l'intelligence artificielle était la modélisation des connaissances et du raisonnement humain. Un autre objectif ambitieux est connu sous le nom d'Intelligence Artificielle Distribuée et se base sur la collaboration d'une multitude d'agents simples et autonomes, *organisés en société pour résoudre collectivement un problème* qui peut être complexe [Drogoul 1993]. Les systèmes multi-agents prennent donc comme référence les interactions sociales élémentaires, et ceci pour favoriser l'émergence d'organisations complexes. Grâce au concept d'agent, il est possible de simuler des systèmes économiques, biologiques,... où chaque individu peut présenter un comportement différent. Cette approche de la simulation est utile dans des domaines très divers même si l'éthologie, où l'on souhaite étudier le comportement des animaux, reste un domaine de prédilection [Drogoul 1993] [Ferber 1995, 1999]. Le lecteur intéressé par l'utilisation de ces techniques pour la robotique se reportera utilement à [Drogoul 2000].

32

Les systèmes multi-agents se séparent en deux catégories principales : les systèmes d'agents cognitifs et les systèmes d'agents réactifs. Les systèmes d'agents réactifs présentent de nombreux agents simples, sans mémoire et avec une vision locale de leur environnement. Ces agents réagissent à des stimuli élémentaires leur permettant d'exprimer leur comportement, au besoin de coopérer, de s'organiser, de se reproduire... On parle d'écorréolution lorsque l'ensemble des interactions non déterministes d'agents réactifs cherchant à se satisfaire permet l'apparition d'états stationnaires sur des systèmes initialement instables [Ferber 1990]. D'une manière opposée, les systèmes d'agents cognitifs ne présentent que peu d'agents : par contre, ces agents possèdent une mémoire du passé, connaissent leur environnement, ainsi que les autres agents avec lesquels ils s'organisent afin d'arriver à leurs objectifs [Wayner 1995].

Les agents communiquent entre eux soit en partageant de l'information (mécanisme de tableau noir), soit par échanges de message. Il est possible de tenter une classification des agents basée sur les protocoles de communication (asynchrone, synchrone, synchrone différé,...). Le lecteur intéressé par

une synthèse présentant plus de 20 catégories d'agents pourra se référer à [Wolridge 1997]. En ce qui concerne la conception et la réalisation de système multi-agents, les travaux de Guessoum sont également une référence utile [Guessoum 1996].

6.3. *Langages, méthodes et outils visuels de simulation*

Si l'on parle de langages de simulation par objets, Simula reste incontournable : les objets actifs et passifs sont issus de la pratique de la simulation et de la modélisation, et si l'on compare les possibilités offertes par Simula à celles de récentes bibliothèques dédiées à la simulation, on est obligé d'admettre que ce langage est loin d'être obsolète. Smalltalk, également, propose les mécanismes de base pour écrire des modèles de simulation [Bézivin 1987]. Des langages tels que ROSS, dérivé de Simula et des langages d'acteurs, ont été développés aux États-Unis pour les simulations de champs de bataille... Il existe également des dizaines d'outils commerciaux pour la simulation. La « Society for Computer Simulation International » en recense plus de 300, dont au moins 50 sont dits « orientés objets ». Parmi ceux-ci on distingue les environnements dédiés à certaines classes d'applications (réseaux, systèmes de production,...) des langages de simulation plus puissants mais plus difficiles à maîtriser. En outre, des bibliothèques de simulation dérivées de Java, d'Eiffel [Howard 1995] ou de C++ sont disponibles sur Internet (Simpack et Sim++, C++Sim, SimEX,...) [Fishwick 1995] [Hill 1996]. En ce qui concerne les langages de simulation par objets, il convient de citer Modsim III (dérivé de Modula 2), Simple++ (dérivé de C++) et Qnap2 (qui repose sur la théorie des réseaux de files d'attente et qui présente des caractéristiques basées sur les objets). De même, des langages de spécification comme VHDL ou SDL sont sortis de leurs cadres d'application respectifs (modélisation de circuits et télécommunications) et sont utilisables pour la conception et la simulation de nombreux types de systèmes communicants. SDL (Specification and Description Language) est standardisé par l'« International Telecommunication Union » depuis 1980 et propose des extensions orientées objets depuis 1992. De même, OOVHDL propose maintenant une version orientée objet de VHDL.

33

Il n'est plus concevable d'utiliser directement des langages sans méthode. La communauté de la simulation propose des cadres méthodologiques où les objets sont sous-jacents, comme CS (Condition Specification) d'Overstreet [Overstreet 1982], CM (Conical Methodology) [Nance 1987], ou encore

STA/DEVS (System Theoretic Approach / Discrete Event Specification) de Zeigler [Zeigler 1990] [Zeigler *et al.* 2000] qui tire avantage des techniques de spécifications formelles avec le formalisme DEVS (Discrete Event System Specification). Toutes ces méthodologies se basent sur une analogie forte entre les objets réels du système physique à modéliser et les objets logiciels, des cycles de vie associés à la modélisation ont été proposés par la communauté du génie logiciel, tels que ceux dérivés du cycle en spirale de Boehm [Boehm 1988], ou du modèle de la fontaine d'Edwards et Henderson-Sellers [Edwards et Henderson-Sellers 1990]. Les méthodes traditionnelles du génie logiciel sont souvent peu adaptées à la représentation de problèmes de simulation où de nombreux objets se mettent à jour d'eux-mêmes au cours du temps. Un rapprochement entre les communautés du génie logiciel et de la simulation a été tenté avec un cycle de vie pour la modélisation orientée objets associé à une méthode d'analyse de conception par objets [Hill 1993a]. Par la suite, des cycles de vie comme OSM (Object Select and Merge) [Barbier et Bézivin 1993], reposant sur l'ingénierie des besoins [Jacobson *et al.* 1993] [Bézivin 1995], ont permis d'intégrer plus simplement les applications de simulation. D'après Rumbaugh, les simulations sont parmi les applications les plus simples à concevoir avec OMT et une approche orientée objets [Rumbaugh 1991] alors que, dans la classification des logiciels de Lehman, les modèles de simulation se retrouvent dans la catégorie de difficulté de réalisation maximale [Lehman 1980]. La méthode unifiée d'analyse et de conception qui était promise en 1996, s'est concrétisée par la proposition d'un langage uniifié : UML (Unified Modeling Language). UML se présente actuellement comme un langage autorisant la modélisation par objets de toutes catégories de systèmes et qui facilite la spécification de modèles de simulation complexes grâce à ses aspects temps réels [Booch 1996] [Lai 1997] [Fowler et Scott 1997] [Muller 1997] [Booch *et al.* 2000]. Un processus uniifié de développement le USDG (Unified Software Development Process) est également en train de voir le jour [Jacobson et Bylund 2000].

Que ce soit pour des ateliers de génie logiciel ou pour des environnements de simulation, les produits commerciaux de qualité ne peuvent se passer d'interfaces de programmation visuelle et de sorties graphiques simples ou animées. On parle de simulation visuelle interactive lorsque, comme pour les jeux électroniques, il est possible d'interagir avec une représentation graphique du modèle en cours de simulation. Cette technique est adaptée à une classe de problèmes, comme la formation des pilotes par exemple, mais elle limite considérablement la production de statistiques prévisionnelles.

Les techniques graphiques et de réalité virtuelle reposent maintenant sur des technologies à objets qui peuvent s'interfacer avec des outils de simulation par objets.

7. CHOIX MÉTHODOLOGIQUES

7.1. Niveau d'abstraction

Dans la pratique, l'élaboration d'un modèle s'appuie sur les deux contraintes suivantes qu'il convient de ne jamais perdre de vue au cours du travail :

- les objectifs à atteindre (quels types de résultats attendons-nous ?, quelles sont les questions auxquelles nous souhaitons répondre ? ou quels sont les problèmes que nous souhaitons résoudre ?),
- l'état des connaissances concernant le système et les données à disposition (ou du moins qu'il est raisonnable d'espérer acquérir dans le temps imparti).

Ces deux contraintes vont définir le niveau d'abstraction du modèle. En écologie, les experts parlent plus volontiers de l'échelle d'étude. Étant donné que nos applications concernent principalement la biologie et l'écologie, nous allons raisonner sur ce domaine. Le choix de l'échelle est en effet en relation directe avec ce qui vient d'être exposé. Nous intéresserons-nous à des individus biologiques, à des parties d'individus, à des ensembles d'individus, à des populations tout entières voire à des ensembles de populations ? Il doit y avoir une cohérence étroite entre les objectifs et les données. Il faut également bien avoir à l'esprit que le niveau d'abstraction va directement influer, d'une part sur la complexité du modèle, et d'autre part, sur les techniques de modélisation à mettre en œuvre.

35

Par ailleurs, l'identification du niveau d'abstraction impose un examen approfondi des données : les objectifs étant fixés et donc le niveau d'abstraction déterminé, encore faut-il que les données relatives à ce niveau d'abstraction soient disponibles.

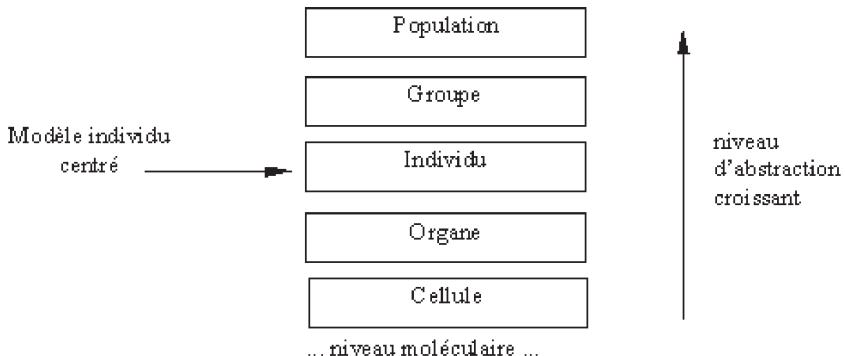


Figure 1. Hiérarchie des niveaux d'abstraction en modélisation d'écosystèmes

7.2. Niveau de détail

Le niveau d'abstraction étant choisi en fonction des objectifs, il convient de préciser ce que l'on entend par niveau de détail de la modélisation. Ainsi, pour un modèle de croissance forestière (avec un niveau d'abstraction correspondant à une population), devra-t-on ou non, tenir compte des paramètres suivants ? :

- variation saisonnière de l'intensité lumineuse,
- pression partielle en CO₂ de l'atmosphère,
- compétition pour les ressources en eau et les nutriments du sol,
- compétition pour l'espace,
- densité du peuplement,
- compétition pour la lumière,
- compétition avec les espèces arbustives et herbacées,
- ...

On voit rapidement qu'objectifs et données vont interagir et que les choix vont être fonction du cadre de l'expérimentation. Ainsi, si l'on s'assigne la modélisation de la croissance des arbres sur les cinq années à venir, il n'y a certainement pas lieu de tenir compte d'éventuelles et infinitésimales variations de la teneur en CO₂ sur cet intervalle de temps. Enfin, la compétition pour les ressources avec les espèces arbustives/herbacées devra être écartée dans la mesure où les arbres y échappent largement et où l'on peut supposer que l'écosystème forestier est relativement homogène de ce point de

vue. Une constante de prélèvement des ressources par les espèces herbacées donnera sans doute un résultat acceptable. Remarquons qu'il en irait tout autrement si nous devions modéliser un peuplement se répartissant sur deux versants, un ubac et un adret. Il y aurait alors lieu de tenir compte de la variation lumineuse. L'effet de la répartition spatiale des individus se ferait alors sentir.

Il n'est donc pas exact que l'accroissement systématique de la complexité du modèle par ajout de variables d'état et de forçage donnera de substantiels gains en termes de validité du modèle et de connaissances sur le système étudié. Au-delà d'un certain point, l'ajout de variables supplémentaires ne fait qu'augmenter la complexité du modèle et accumuler les incertitudes. Il se peut même que ces accroissements d'incertitudes fassent alors radicalement diverger le modèle par rapport au système réel. Il convient souvent de préférer des variables holistiques¹ à plusieurs variables élémentaires dont les incertitudes ne peuvent qu'oblitérer la qualité du modèle ; on parle aussi dans ce cas d'agrégation. Il va de soi, par ailleurs, qu'accroître la complexité du modèle ne peut qu'ajouter aux difficultés de mise en œuvre informatique : très rapidement des problèmes de capacité de mémoire, de facilité de vérification de la cohérence interne du logiciel, de vitesse de calcul peuvent apparaître. La simplification – dans certaines limites – est une vertu en modélisation.

37

À l'inverse, une simulation sur le long terme imposerait certainement un changement radical de stratégie, et aux paramètres ci-dessus il conviendrait d'ajouter la régénération des arbres, leur espérance de vie, les chablis, etc. De la même manière, le choix d'intégrer un certain nombre de sous-modèles repose sur l'arbitraire. Seuls les objectifs et les données disponibles peuvent guider empiriquement l'élaboration du modèle.

Le choix du niveau de détail constitue donc une phase majeure du processus de modélisation, aux répercussions importantes en termes de coût économique, de temps et d'effort de travail, d'acquisition de données, et, bien entendu, de temps de calcul des machines. Une discussion dans [Coquillard et Hill 1997] montre que la majorité des chercheurs ont développé des modèles de faible articulation (peu de variables) mais de forte précision, à savoir qu'ils ont cherché à restreindre le champ de modélisation tout en se focalisant sur des processus de plus en plus fins, abandonnant la complexité.

1. C'est-à-dire de variables globales dont on estime qu'elles sont la résultante d'actions d'un ensemble de variables agissant de manière conjointe et simultanée.

« *Ces modèles nous informent beaucoup à propos de très peu* » écrit Jorgensen [Jorgensen 1994], qui fait remarquer à juste titre qu’« *il est impossible de décrire le comportement de toutes les espèces d’un système en fonction des combinaisons possibles des variables de forçage* [c'est-à-dire les entrées du système] *en utilisant des réponses précises à des questions étroites* ». Il suggère de s’ache-miner vers des modèles « *informant peu à propos de beaucoup* », et donc d’augmenter la complexité et le pouvoir explicatif des modèles au détriment de la précision.

Ces dernières remarques nous renvoient à la fois à la question de la qualité des données (nombre de mesures) et du nombre de variables (lié au niveau de détail). Comment procéder alors pour effectuer ces choix ? Il n’existe pas à notre connaissance de méthode véritablement efficace. Le lecteur intéressé peut cependant trouver quelques indications, suggérées par Jorgensen [Jorgensen 1994]. Il ne s’agit cependant que d’approches empiriques, ne pouvant s’appliquer qu’aux modèles à formalisme mathématique.

7.3. Granularité du temps

38

Quel pas de temps choisir ? Voilà sans aucun doute l’une des questions les plus importantes en matière de simulation. La cohérence est de rigueur : si l’on s’intéresse à des populations dont la génération est d’une année, il vient naturellement à l’esprit de retenir le pas de temps d’une année. Mais attention ! Les pas de temps ne sont pas forcément fixes et il faut pouvoir prendre en compte tous les événements élémentaires déterminants pour le comportement de la population. Si on choisit de retenir un pas fixe, une bonne règle pourrait être la suivante : le pas de temps à retenir est celui des événements pris en compte dans la modélisation et intervenant le plus fréquemment dans le système réel. Ici encore, les objectifs de la modélisation induisent les événements à prendre en compte et sont déterminants dans le choix. Notons aussi que si la modélisation sur plusieurs pas de temps très différents est possible, elle présente cependant peu d’intérêt. En effet, lorsque les échelles de temps déterminant les fréquences d’événements sont très différentes, le rapport de la fréquence la plus rapide par rapport à la fréquence la plus lente devrait être inférieur à 10^6 . En effet, s’il fallait générer environ 10^6 événements à fréquence rapide pour ne générer qu’un seul événement à fréquence lente, l’obtention de résultats statistiques cohérents nécessiterait la génération d’un nombre suffisant d’événement à fréquence lente, et cela imposerait la génération d’un nombre gigantesque d’événements à

fréquence rapide et généreraient des problèmes de précision arithmétique lors du calcul des statistiques.

8. CONCLUSION : L'EXPLOITATION D'UN MODÈLE

Un processus de modélisation consiste en un ensemble d'opérations à effectuer en séquence, avec, dans de nombreux cas, des retours en arrière. Le processus de modélisation que nous retenons a été présenté dans [Hill 1993b]. Il va de l'analyse fonctionnelle de l'écosystème et du recueil des données à la mise à disposition d'un logiciel apte à l'exploitation. Dès lors que la phase d'analyse est terminée et que le modèle est conçu, on procède à l'implémentation. Vient ensuite un ensemble de phases cruciales que nous résumons ci-dessous :

- Vérification. Elle consiste en la vérification de la cohérence interne du logiciel et l'élimination des erreurs de programmation.
- Calibration. C'est l'initialisation des paramètres du modèle permettant l'ajustement du comportement de celui-ci au comportement du système étudié, dans un cas connu.
- Robustesse et fiabilité logicielle. C'est l'examen du comportement du logiciel dans des conditions extrêmes.
- Validation. Test de la cohérence des résultats du modèle sur plusieurs séries de données.
- Exploitation, exploration. C'est l'utilisation pour la découverte de comportements insoupçonnés ou pour la prédiction.

39

Il est utile de dissocier la notion de validation de celle de vérification, la distinction entre ces deux notions étant souvent confuse selon [Balci et Sargent 1981]. La validation d'un modèle cherche à constater qu'il possède une marge de précision de ses résultats suffisante et si ces résultats sont cohérents avec ceux attendus de ce modèle dans son cadre d'application. La vérification consiste à s'assurer que le programme correspondant au modèle est fonctionnellement correct, et qu'il traduit bien les choix et les hypothèses effectués. Sachant qu'une étude de simulation doit être menée dans un but précis et pour une utilisation bien définie, la validation et la vérification d'un modèle doivent donc être faites dans le cadre de cet objectif.

CHAPITRE 3

MODÉLISATION PAR OBJETS DE SYSTÈMES À FLUX DISCRETS

*Que reste-t-il à l'homme de toute la peine et de tous les calculs
pour lesquels il se fatigue sous le soleil ?*

L'Écclésiaste, 2,21

1. INTRODUCTION

Les logiciels de simulation se retrouvent dans la catégorie de difficulté de réalisation maximale dans la classification des logiciels de Lehman [Lehman 1980]. Ma thèse de doctorat portait sur la conception et la réalisation d'outils logiciels pour la modélisation de systèmes complexes [Hill 1993a,b]. Un des objectifs majeurs de mes travaux à cette époque consistait précisément à rapprocher le génie logiciel et la modélisation. J'ai cherché à recenser les techniques avancées du génie logiciel qui étaient exploitables pour la simulation, puis j'ai proposé des outils du génie logiciel pour l'analyse, la conception, la vérification et la validation de modèles de simulation. Ce recensement a été effectué sur la base du modèle objet exploité par la communauté du génie logiciel et celle de la simulation. L'un des nombreux intérêts des techniques à objets est qu'elles facilitent les liens

entre, d'une part, l'analyse et la spécification d'un système réel et, d'autre part, la conception et l'implémentation du modèle par un simulateur. Des techniques de génération automatique de code à partir de spécification graphique étaient déjà présentes. On s'appuyait bien sûr sur des métamodèles même si la terminologie liée maintenant à l'ingénierie des modèles n'était pas encore présente.

La conception et l'exploitation de systèmes à flux discrets, tels que les systèmes de production, les systèmes informatiques, les systèmes de transport, systèmes administratifs,... a toujours été problématique. Parmi les différents problèmes posés aux concepteurs et aux ingénieurs, on retrouve bien sûr des soucis de compréhension du fonctionnement de ces systèmes (que ce soit en fonctionnement nominal ou en mode dégradé) ; le dimensionnement de tels systèmes est également un problème ; de plus les objectifs économiques imposent systématiquement l'amélioration de leur productivité et donc l'évaluation de leurs performances.

Pour résoudre ces problèmes, la simulation aléatoire à événements discrets est quasiment indispensable notamment pour l'étude de phénomènes transitoires. La maîtrise de cette technique n'est pas évidente malgré la disponibilité de logiciels de simulation puissants qu'ils soient dédiés ou généraux. Les outils dédiés ne savent traiter qu'une partie restreinte des problèmes, et les logiciels généraux reposent sur un formalisme difficile à acquérir rapidement. Ils autorisent cependant la modélisation de systèmes très complexes.

42

Partant de ce constat à la fin des années 1980, il était intéressant de concevoir, à partir de logiciels généraux de simulation, des environnements de modélisation conviviaux qui permettent, dans un domaine d'application donné de construire facilement un modèle de connaissance d'un système, et de déduire automatiquement un programme de simulation valide. Dans les débuts des années 1990, plusieurs chercheurs du LIMOS avec lesquels nous travaillions, ont défini un environnement de modélisation, comme un ensemble comprenant : un logiciel d'évaluation des performances (constituant le noyau de l'environnement), des outils graphiques, des outils de statistiques et de recherche opérationnelle, un système d'aide à la décision, un système de gestion de bases de données, une méthode d'analyse, des outils de spécification, une méthodologie de modélisation et les interfaces logicielles permettant de relier ces outils [Breugnot *et al.* 1990].

Au sein d'un environnement, tel que nous le définissions, la méthodologie de modélisation reposait sur un ensemble de méthodes, certaines orientées objet, et s'appuyait sur un processus de modélisation [Kellert 1992]. Les principaux objectifs que nous fixions au LIMOS pour un environnement de modélisation de systèmes à flux discrets étaient les suivants :

- fournir à l'utilisateur des outils de spécification de son système lui permettant de construire le modèle de connaissance de ce système,
- construire une méthodologie évitant d'élaborer pour chaque système un simulateur dédié,
- traduire automatiquement ce modèle de connaissance en un modèle d'action,
- aider à vérifier et à valider le modèle d'action.

Entre 1990 et 1993, j'ai principalement abordé les deux derniers points, tout en ayant contribué aux deux premiers. Mes objectifs étaient :

- de trouver une méthode d'analyse et de conception par objets adaptée aux problèmes de la modélisation des systèmes discrets afin de construire des programmes de simulation de qualité,
- de concevoir des outils graphiques de vérification et de validation en utilisant des techniques d'animation par objets,
- de générer automatiquement du code de simulation à partir d'outils de programmation visuelle. On parlerai aujourd'hui de transformation de modèle.

43

Les principales applications que j'avais traitées à l'époque concernaient les systèmes d'assemblage flexibles à flux discrets. Cependant dans un souci de généricté, nous avions avec d'autres doctorants également abordé un système administratif (une préfecture), un service hospitalier (urgences de l'hôpital de Bastia), un système informatique parallèle à base de transputers et également la spécification de systèmes parallèles à flux discrets par des réseaux de Petri.

2. PROPOSITION D’UN PROCESSUS ET D’UNE MÉTHODE DE MODÉLISATION PAR OBJETS

2.1. *La synthèse des approches existantes*

Pour proposer une méthode de modélisation par objets dans le début des années 1990, j’ai tout d’abord cherché à recenser les travaux invariants des différentes méthodes d’analyse et de conception par objets [Hill 1992]. Les résultats de cette synthèse m’ont permis de proposer un processus et une méthode de modélisation (M2PO) adaptée à la simulation à événements discrets [Hill 1993a]. Il ne s’agissait pas d’imposer une nouvelle méthode. Dans l’attente du langage de modélisation unifié (UML Unified Modeling Language), le nombre de méthodes d’analyse et de conception par objet allait croissant. Je préférerais à l’époque ne pas introduire de nouvelles notations graphiques en m’attachant aux notations les plus couramment utilisées. Le but n’était cependant pas de rassembler au sein d’une méthode un maximum d’éléments composites mais de proposer un modèle d’une complexité raisonnable, intégrant une partie des préoccupations de la communauté de la simulation.

44

Les méthodes d’analyse et de conception par objets, existant à l’époque, proposaient toutes un certain nombre d’étapes accompagnées d’une démarche propre à chaque auteur. Comme je l’annonçais, ma démarche a tenté une classification des travaux invariants présents dans les méthodes que j’avais étudiées [Hill 1992]. J’avais effectué ce travail de synthèse car si la majorité des méthodes avaient en commun un grand nombre de travaux élémentaires, chacune possédait des particularités intéressantes qu’il est souhaitable d’intégrer dans une méthode d’analyse et de conception par objets. J’ai donc pu présenter en 1993 une synthèse des travaux élémentaires qu’un concepteur est susceptible d’effectuer [Hill 1993a]. Ces travaux n’étaient pas destinés à être réalisés séquentiellement car le processus de conception par objets restait incrémental et itératif. Pour chaque travail à réaliser, il convient de vérifier s’il est possible de réutiliser des résultats d’analyse ou de conception existants. Les travaux élémentaires que j’avais sélectionnés à l’époque parmi les méthodes étudiées, sont donnés ci-après :

- Identification des classes relatives au système logiciel que l’on conçoit,
- Classification au moyen de l’héritage et de l’agrégation,

- Établissement des relations de communication existantes entre les différentes classes,
- Définition des structures permettant aux objets de collaborer pour atteindre les objectifs fixés,
- Choix de l'interface des classes et de leur représentation interne,
- Évaluation de la qualité des abstractions obtenues,
- Amélioration des résultats obtenus par chaque travail élémentaire jusqu'à obtenir le niveau de détail souhaité. (Test d'arrêt du processus itératif).

Le résultat de ces travaux conduisait à l'obtention d'un modèle objet qui comportait une partie statique et une partie dynamique. La partie statique décrit la structure des classes en termes de classification par héritage et composition, de relations entre classes mais également en termes de méthodes et d'attributs (que ce soit pour les instances ou pour les classes). La partie dynamique élémentaire permet de spécifier, d'une part, les interactions existantes entre les différents objets, et d'autre part, le comportement individuel des objets.

45

2.2. *La proposition d'un processus*

Lors de l'étude des méthodes d'analyse et de conception par objets, j'avais effectué une synthèse comparative des principaux cycles de développement par objets. Le processus de modélisation par objets que nous proposions était dérivé de ces cycles tout en prenant en compte le point de vue de la simulation. Ceci permettait de se placer dans un cadre de modélisation plus général que celui de la création de logiciels.

La vue générale du cycle de modélisation que j'avais proposé prend en compte successivement, une analyse de domaine, une analyse par objets, une conception par objets et enfin une implémentation. Par analogie avec les logiciels, la maintenance constitue la vie des modèles en matière d'évolution et de correction d'erreurs. Le cycle que je proposais était incrémental et itératif, chaque étape étant susceptible d'alimenter les autres. Ce cycle est arrêté dès que le niveau d'information atteint est jugé suffisant par l'expert en modélisation. La Figure 2 schématisé le processus de modélisation proposé à l'époque.

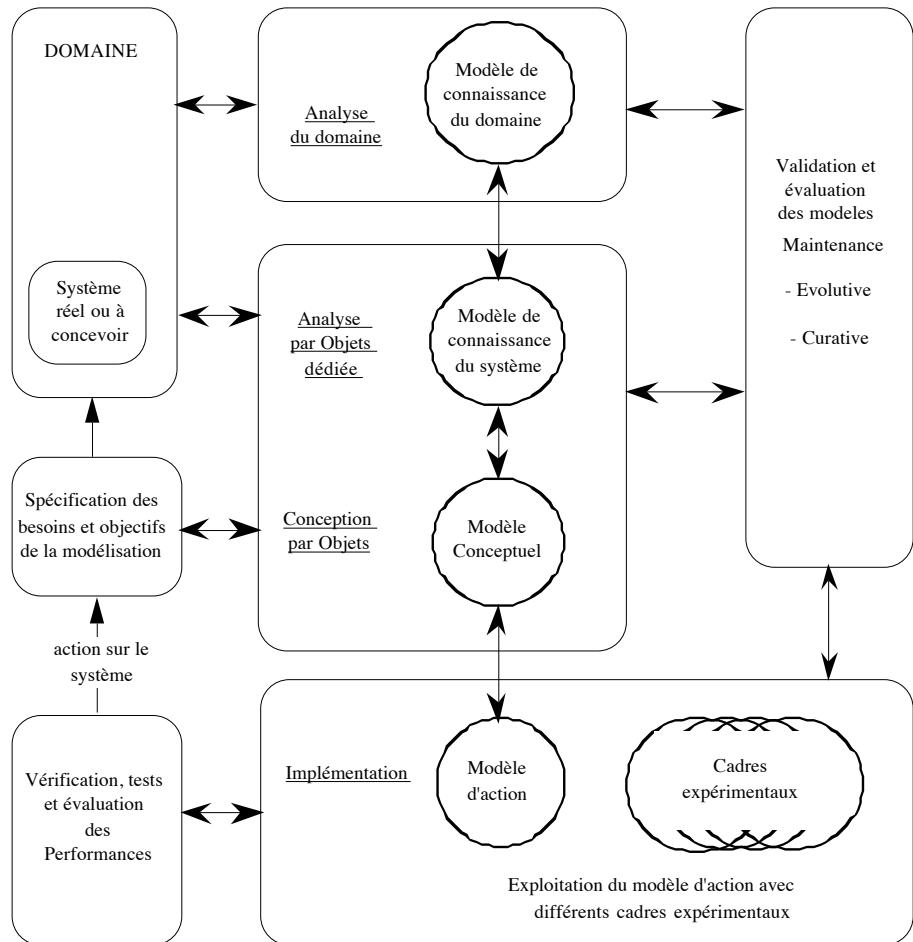


Figure 2. Vue générale du processus de modélisation par objets

Je tiens à préciser que je ne proposais pas de technique particulière pour les phases de spécification des besoins, de vérification et de validation des modèles et des logiciels. Les tests, la conception de plans d'expériences, les techniques de validation et de vérification des modèles mériteraient à elles seules un ouvrage entier [Kleijnen 1987].

L'identification des classes (au sens « objet ») communes à toutes les applications d'un domaine particulier, reste, encore aujourd'hui, une des tâches les plus ardues de l'analyse. La qualité des abstractions obtenues

reste encore fortement liée à l'expérience des spécialistes en modélisation. Les objets, les relations et les comportements découverts et retenus doivent être ceux qui sont perçus comme étant importants par les experts du domaine. Le résultat de ce travail constitue un modèle de connaissance du domaine, qui comporte aussi un glossaire donnant le vocabulaire de ces experts. Avec la démarche préconisée, il est possible de réutiliser les résultats d'une analyse de domaine, afin d'initialiser des analyses par objets pour d'autres systèmes du même domaine. Puis pour chaque analyse on peut « concevoir » différentes solutions, elles mêmes implémentables de différentes manières. Les relations hiérarchiques entre les différents modèles sont données par la **Figure 3**. Les problèmes de terminologie ne doivent pas être négligés, une étude avait été menée par nos collègues dans le domaine des systèmes de production [Breugnot *et al.* 1991c], la constitution de glossaires et leur intégration au sein du modèle de connaissance du domaine s'est révélée être indispensable.

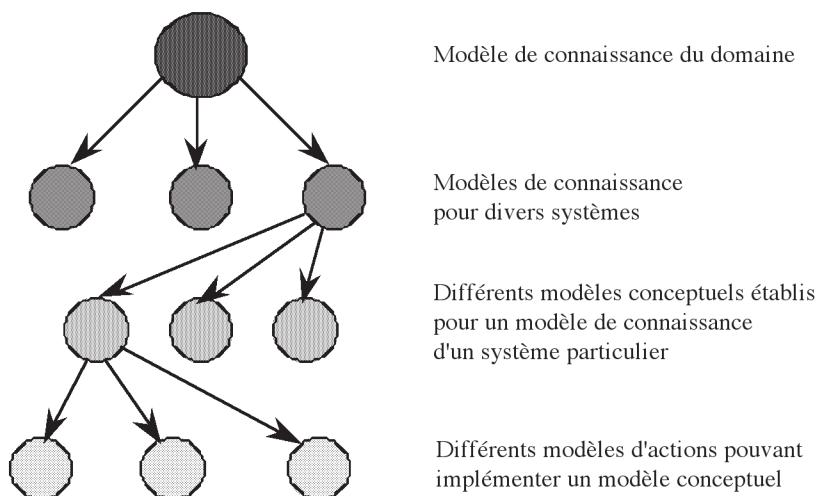


Figure 3. Relations hiérarchiques entre les différents modèles

La phase d'implémentation comprend le codage d'un simulateur sur une plateforme cible, suivi des différentes étapes de tests puis de validation du modèle d'action lors d'utilisations réelles. Pour un même modèle concep-

tuel, il peut y avoir plusieurs modèles d'action (machines cibles différentes, outils d'évaluation des performances ou langages de simulation différents,...). La séparation du modèle et de ses cadres expérimentaux, est basée sur les travaux de Zeigler [Zeigler 1976]. Le modèle d'action d'un système implémente les caractéristiques statiques et dynamiques du système. Le cadre d'expérimentation définit les conditions sous lesquelles un modèle d'action est exécuté. Il est ainsi possible d'effectuer de nombreuses exploitations du modèle d'action en modifiant seulement la valeur de certains paramètres. La phase de maintenance, qu'elle soit curative ou évolutive, peut entraîner des remises en question et des modifications dans les résultats obtenus avec les autres phases. Aucun travail ne peut être considéré comme « parfait ou complet », le refus d'éventuels changements est comparé à la stratégie de la ligne Maginot par Cox [Cox 1986].

2.3. La proposition d'une méthode de modélisation

La méthode M2PO (Méthode de Modélisation Par Objets) que je proposais suit le processus de modélisation présenté précédemment. M2PO peut être considérée comme une méthode hybride car elle était basée, d'une part, sur la synthèse des différents travaux élémentaires des principales méthodes d'analyse et de conception par objets et, d'autre part, sur un ensemble de critères jugés essentiels à la réalisation de modèles objets pour la simulation et l'animation de systèmes à flux discrets. Le point de vue d'un spécialiste en simulation permettait d'aborder la modélisation de systèmes complexes en général sans se limiter aux systèmes logiciels.

48

La méthode M2PO comportait les phases suivantes, qui découlaient du processus de modélisation par objets proposé précédemment (**Figure 4**) :

- Phase 1 :** une analyse de domaine,
- Phase 2 :** une analyse par objets pour le système que l'on souhaite modéliser,
- Phase 3 :** une conception par objets apportant une solution au problème de modélisation posé.

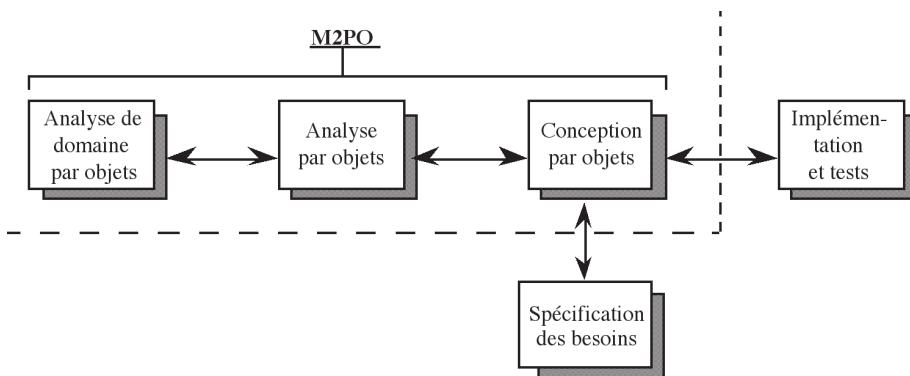


Figure 4. Les phases de la méthode M2PO

Les phases de spécification des besoins, d'implémentation et de tests sortaient du cadre de M2PO et étaient abordées par d'autres méthodes. L'implémentation par objets fournissant un simulateur n'était pas détaillée, ni celle qui consiste à élaborer des tests, les diverses méthodes ou solutions existantes suffisaient sans que nous ayons à rajouter de nouvelles techniques personnelles [Sommerville 1993]. De même, la spécification des besoins était abordée avec succès [Jacobson *et al.* 1993] et il nous semblait inutile de proposer d'autres techniques alors que celles qui existaient alors rencontraient un succès croissant (spécification à base de cas d'utilisation).

49

Précisons maintenant le détail des objectifs qui étaient fixés pour M2PO :

1. Utiliser tous les concepts majeurs de l'approche orientée objets, en préconisant l'utilisation de classes abstraites partout où cela est possible. Garder cette approche pour toutes les phases du processus de modélisation, de l'analyse de domaine à l'implémentation, afin de disposer de modèles cohérents.
2. Dissocier les aspects statiques et dynamiques des systèmes à modéliser.
3. Utiliser une notation textuelle et graphique simple et unifiée pour toutes les phases du processus de modélisation. Cette notation doit être indépendante de tout langage de programmation par objets, et doit pouvoir supporter la majorité des concepts du modèle objet. Ces caractéristiques doivent rendre la notation facilement adaptable à d'autres méthodes d'analyse et de conception par objets.

4. Favoriser la réutilisabilité, non seulement au niveau des codes des objets produits, mais surtout au niveau de l'analyse et de la conception. Il est indéniable que l'efficacité de la réutilisabilité au niveau de l'analyse passe par la constitution de modèles de connaissance pour les domaines.

D'autres objectifs visaient à inclure dans la méthode M2PO les aspects spécifiques à la modélisation des systèmes à flux discrets, en vue de leur simulation :

- Prise en compte de la décomposition en trois sous-systèmes (physique, logique et décisionnel) introduite dans [Breugnot *et al.* 1990].
- Détail des transactions dans les systèmes modélisés avec introduction des notions de chemins et de phases [Gourgand 1984].
- Mise en valeur des aspects dynamiques des systèmes modélisés ainsi que de leur potentiel de réaction piloté par un sous-système décisionnel.
- Prise en compte de la séparation du modèle et de ses cadres expérimentaux suivant le processus de modélisation de M2PO.
- Validation et vérification des différents modèles en vue d'effectuer des simulations et des animations cohérentes.
- Constitution de glossaires techniques propres à la terminologie des domaines concernés. Ces glossaires constituent la généralisation des dictionnaires de données.

50

Les résultats de M2PO se répartissaient sur deux dimensions. La première dimension concerne les aspects statiques (descriptifs) du système, et la seconde concerne les aspects dynamiques du système. Ce découpage se retrouve dans les différents modèles (modèle de connaissance du domaine, modèle de connaissance d'un système et modèle conceptuel). Les catégories de résultats que je proposais étaient, pour la partie statique :

- Des diagrammes de classification avec la structure des classes.
- La description des classes sous forme de fiches.
- La description des instances sous forme de fiches.
- Des diagrammes d'instance.
- Un glossaire pour chaque domaine.

pour la partie dynamique :

- Les interactions entre objets et entre sous-systèmes avec des diagrammes de flots de messages.

- Les comportements internes des objets spécifiés avec des réseaux de Petri étendus.
- Les séquences d'opérations associées à des diagrammes de circulation des flux.
- Des scénarios ou traces d'événements.
- Des relations de classification dynamique.

Des scénarios ou traces d'événements pouvaient être associés aux diagrammes de flux de messages, aux spécifications des comportements internes ainsi qu'aux transactions réalisées par les flux du système.

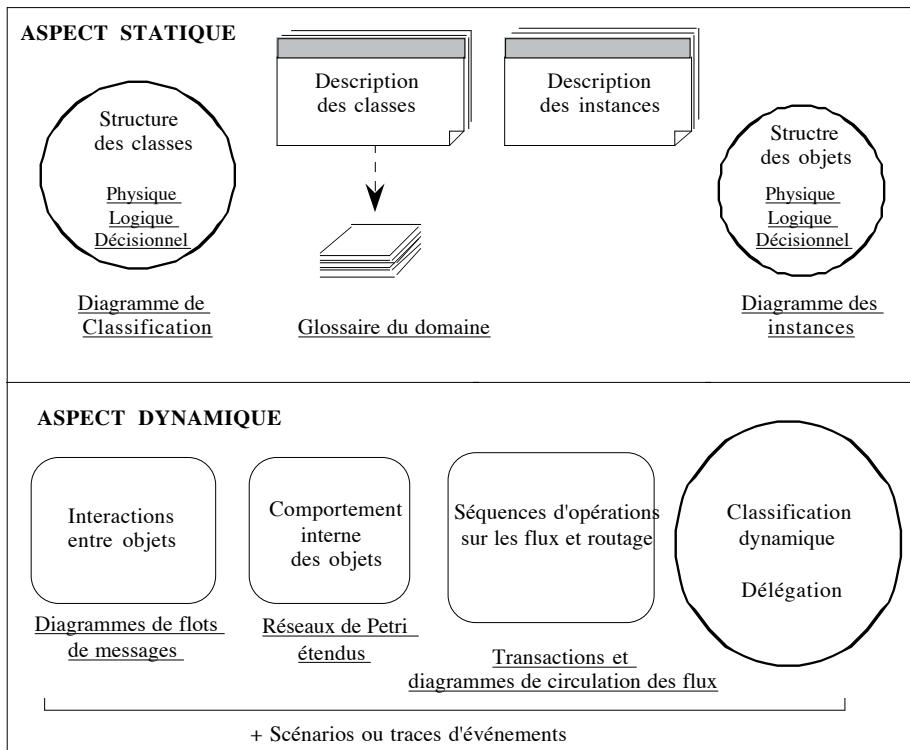


Figure 5. Les différentes catégories de résultats

2.4. Conclusion

La méthode M2PO a été utilisée pour analyser, concevoir et implémenter des logiciels de simulation. Son développement s'est appuyé sur l'expérience de notre laboratoire concernant le développement d'environnements de modélisation [Breugnot *et al.* 1990], sur divers projets de simulation de systèmes industriels [Kellert 1992], et également sur mes développements d'outils orientés objets pour la génération de code et l'animation de résultats de simulation [Hill et Junqua 1990] [Hill 1993a] [Hill 1996]. Quelques thèses et articles y font référence [Tanguy 1993] [Combes *et al.* 1994] [Ruch 1994] [Kellert *et al.* 1997] [Laizé 1998]. Dès sa conception, M2PO était dédiée aux projets de simulation, si elle n'a pas connu de réelle diffusion ni de développements commerciaux, il est intéressant de remarquer un ensemble de similitudes avec les notations proposées par UML, cinq ans avant la standardisation de ce langage. Reprenons maintenant ces principales caractéristiques, M2PO :

- 52
- permettait de couvrir tout le cycle de modélisation par objets présenté en restant indépendante des différents langages à objets (de simulation ou non) susceptibles d'implémenter les modèles d'action.
 - reposait sur les principaux concepts du modèle objet et dissociait les aspects statiques et les aspects dynamiques d'un système réel à modéliser.
 - pouvait être utilisée à un haut niveau d'abstraction en prenant en compte l'analyse de domaine (et la réutilisabilité qui en découle), mais également en mettant en avant les classes abstraites, aussi bien pour les aspects statiques que pour les aspects dynamiques.
 - utilisait une notation graphique simple reportant sur la notation textuelle les détails qui surchargent habituellement les schémas, aussi bien pour la partie statique que pour la partie dynamique. Cette simplicité est également due à la limitation volontaire du nombre de symboles.
 - prenait en compte les principaux aspects dynamiques du modèle objet. Ces aspects dynamiques sont utilisés pour représenter au mieux le fonctionnement des systèmes réels (allant de la classification dynamique à la prise en compte de la concurrence au sein du comportement interne d'un objet). Les interactions entre objets sont étudiées avec des diagrammes de flots de messages permettant d'examiner le couplage entre les objets et entre les sous-systèmes d'objets. Le com-

portement interne des objets est spécifié avec les réseaux de Petri interprétés et communicants.

- utilisait la décomposition en trois sous-systèmes (physique, logique et décisionnel) introduite lors du projet SIGMA [Breugnot *et al.* 1990]. Cette décomposition était intéressante pour notre approche de la modélisation des systèmes à flux discrets.
- établissait pour les éléments de flux des chemins et des phases permettant d'appréhender le fonctionnement des systèmes en vue d'effectuer des simulations par objets utilisant une approche transaction.
- améliorait la traçabilité des objets à travers les différents modèles en permettant de savoir aisément si un objet appartenait au domaine du problème ou au domaine de la solution.

3. L'ANALYSE ET LA CONCEPTION D'OUTILS D'ANIMATION DE RÉSULTATS DE SIMULATION

3.1. *Le contexte historique*

53

La disponibilité de micro-ordinateurs et de stations de travail, possédant de bonnes capacités graphiques à bas prix, a largement contribué à la démocratisation des techniques graphiques qui, avant les années 1980, étaient uniquement accessibles sur des calculateurs de taille imposante. Nous nous intéressons plus particulièrement à l'impact du graphisme pour la modélisation de systèmes complexes, utilisant la simulation à événements discrets comme technique d'évaluation des performances. Les premières utilisations du graphisme, en tant qu'outil d'aide à la résolution de problèmes de simulation, remontent à la fin des années 1960 [Bell 1969]. En 1976 une communication de Bazjanac, au neuvième symposium de simulation, présente une étude visuelle interactive de l'évacuation en urgence des personnes présentes dans un gratte-ciel, au moyen d'ascenseurs [Bazjnac 1976]. Cette situation illustre parfaitement l'apport de l'animation qui permet d'analyser le comportement transitoire d'un système et qui vient compléter efficacement les résultats statistiques globaux. À notre connaissance, une des premières tentatives d'animation par objets remonte à Palme [Palme 1977] qui expliquait les avantages de disposer d'une animation générée par des modèles écrits avec Simula. Les travaux de Hurzion, présentés en 1976 dans

sa thèse, introduisent les concepts de la simulation visuelle interactive [Hurzion et Secker 1978]. La dernière décennie vit l'explosion de l'utilisation des techniques d'animation graphiques de résultats de simulations dans de nombreux domaines, et tout particulièrement celui des systèmes automatisés de production. Une revue des différents logiciels d'animation graphique de résultats de simulation, ainsi que des logiciels de simulation visuelle interactive est présentée dans [Bell et O'Keefe 1987]. Green et Sun présentaient à la fin des années 1980 les différentes méthodes utilisées pour implémenter une simulation visuelle interactive [Green et Sun 1988].

L'animation des modèles constitue maintenant une partie non négligeable d'un projet de simulation, voire essentielle, pour mener à bien certaines études comportant de nombreux phénomènes transitoires. Les avantages de l'animation de résultats de simulation sont indéniables, Shannon et Gipps présentent les différents atouts de l'animation en matière de validation de modèles [Shannon 1986] [Gipps 1986]. Smith donne une partie des apports de l'animation en matière de modélisation et de simulation des systèmes d'assemblage [Smith et Platt 1987]. Je développe les différents apports des techniques graphiques et de l'animation dans le paragraphe suivant.

54

Deux approches différentes se sont développées en matière d'utilisation de techniques graphiques pour la simulation, l'une étant principalement soutenue par les chercheurs travaillant aux États-Unis, et l'autre par les chercheurs Britanniques. La technique préconisée aux États-Unis consiste à animer les résultats de simulation, sans laisser l'utilisateur interagir avec le modèle en cours d'exécution. La technique utilisée en Grande-Bretagne autorise des interactions entre l'utilisateur et le modèle animé ; il s'agit des techniques de simulation visuelle interactive introduites par Hurzion, elles n'autorisent plus de validation statistique des résultats. Je tiens à préciser ici que je n'avais pas abordé à l'époque les simulations visuelles interactives combinant des techniques de réalité virtuelle et de synthèse d'images. Le matériel dont nous disposions ne nous le permettait pas. Ces aspects ont cependant pu être abordés dans la thèse d'André Campos que j'ai encadrée, notamment en réalisant sur des stations Silicon Graphix des simulations sur le Web avec les langages Java et VRML (Virtual Reality Markup Language). Cependant, si un réalisme minimum est nécessaire, pour beaucoup de projets de simulation, un graphisme excessif peut être anti-productif. Il reste en effet à démontrer que les techniques avancées de CAO et de synthèse d'images soient réellement nécessaires pour la majorité des applications communes de simulation.

3.2. Les techniques proposées

Durant les années 1990 à 1992, j'ai proposé une boîte à outils d'aide à la construction d'environnements d'animation et de simulation nommée GIGA. Après avoir dressé un état de l'art des techniques d'animation de résultats de simulation, j'ai pu préciser les apports de ces techniques dans un projet de simulation. Puis, j'ai porté une attention particulière aux techniques de simulation visuelle interactive pour, par la suite, présenter les concepts et les objectifs de cette boîte à outils GIGA.

L'animation graphique de résultats de simulation avait souvent entraîné un important travail de programmation. Dans le début des années 1990, il existait plusieurs logiciels de simulation associés à un animateur (Siman/ Cinema, Simscript ou Modsim II avec Simgraphics, GPSS avec divers outils, Proof Animation,...). Il existait également des environnements intégrant l'animateur et le simulateur Arena basé sur Siman et Cinema, Slam-system basé sur Slam II, ou d'autres environnements Witness, Cadence, Hocus, INSIGHT,... D'autre part, on trouvait également des animateurs indépendants du logiciel de simulation. Nous avons, avec des collègues du LIMOS, mené une étude des divers outils précités et d'un certain nombre d'autres ; cette étude a donné lieu à un rapport pour un projet Européen Tempus [Hill *et al.* 1992].

55

Je me suis intéressé à l'époque plus particulièrement à la catégorie des animateurs susceptibles de s'adapter à différents outils de simulation à événements discrets. C'est en effet cette catégorie d'outils qui s'avère la plus apte à s'intégrer dans ce que nous nommions au LIMOS des environnements de modélisation composés d'outils hétérogènes. Les objectifs fixés m'ont permis de proposer des solutions pour augmenter le potentiel d'aide, que peut apporter une animation orientée objets, à la validation et à la vérification de modèles de simulation. J'avais notamment développé des points tels que l'intégration de techniques de mise au point des programmes de simulation au sein des objets graphiques utilisés pour l'animation.

La boîte à outils nommée GIGA comportait un ensemble de composants logiciels réutilisables et indépendants de tout logiciel de simulation (se présentant sous la forme d'un cadriel : « framework »). Ces composants concernaient essentiellement :

- l'édition de modèles constitués d'objets graphiques,
- l'animation d'objets graphiques suivant leurs comportements propres.

Le but principal de GIGA était d'aider les spécialistes en modélisation dans la réalisation de logiciels permettant de construire des animations valides. Chaque logiciel réalisé avec GIGA était dédié à un domaine particulier et autorisait la construction d'animations cohérentes et réalistes. Contrairement aux outils existants, qui intégraient très peu de potentiel de validation / vérification dans leurs animations (simples déplacements d'icônes à l'écran), nous avions choisi d'implémenter, dans le code des objets d'un domaine, des éléments de contrôle de leur comportement interne. Ainsi, toute incohérence avec les comportements prédefinis dans ce domaine pouvait être immédiatement signalée lors d'une animation. Cette solution permettait d'isoler, d'une part, les composants réutilisables indépendants des domaines de problèmes et, d'autre part, les objets propres aux domaines que l'on souhaitait traiter. Ceci autorisait la réalisation d'outils hybrides possédant des traits communs avec les logiciels généraux d'édition / animation de modèles, tout en disposant de la puissance des outils dédiés. Les outils d'animation produits avec GIGA intégraient donc des techniques de validation et de vérification. Cependant, ils n'en restaient pas moins des outils puissants de communication et d'aide à la décision. Les principaux objectifs que je me fixais pour GIGA étaient les suivants :

56

- Les outils créés avec GIGA devaient rester indépendants du logiciel de simulation. Cette indépendance permettait de garantir une composabilité importante avec tous les logiciels de simulation ouverts.
- Les communications entre le logiciel de simulation et le logiciel d'animation reposaient sur l'envoi de messages.
- Les outils développés avec GIGA devaient permettre d'augmenter sensiblement le potentiel d'analyse, de validation et de vérification des modèles de simulation en comparaison avec les outils d'animation existants. Une animation par objets permet de se placer à un niveau d'abstraction élevé pour la vérification et la validation de modèles de simulation. Par exemple, la mise au point doit pouvoir être réalisée au niveau d'abstraction des objets graphiques du système. Ceci permet d'appréhender des phénomènes parallèles et transitoires impossibles à aborder au niveau du code source d'un modèle de simulation.
- La structure de GIGA et sa conception devaient reposer sur un cadre méthodologique autorisant la réutilisabilité non seulement du code des composants standards de GIGA, mais également de l'analyse et de la conception pour la création d'autres outils d'édition et d'animation.

- Les composants standards de GIGA devaient être indépendants du domaine traité.
- La création de bibliothèques d'objets pour aborder différents domaines devaient suivre le cadre méthodologique (cadriel) proposé par GIGA.
- Le graphisme associé aux objets spécifiques à un domaine devait posséder un degré de réalisme suffisant, tout en restant suffisamment symbolique pour ne pas détourner l'attention de l'utilisateur par un effet « cinéma ».
- Le comportement visuel des objets d'un domaine doit correspondre aux états significatifs des objets réels.

Les composants logiciels de GIGA étaient écrits en C++ sur micro-ordinateur compatible PC et sur station de travail HP/APOLLO, une partie des composants était également portée en Objective-C sur les anciennes stations de travail NeXT.

3.3. Conclusion

Pour réaliser un outil d'animation hybride, possédant les traits des logiciels généraux, tout en disposant de la puissance des outils dédiés, un spécialiste en modélisation pouvait se concentrer sur le domaine de son choix et réutiliser les composants d'édition et d'animation d'objets fournis par GIGA. Avec les choix effectués à l'époque, la qualité de l'élaboration des objets propres au domaine déterminait le potentiel de l'outil final. On parlerait aujourd'hui de DSL (Domain Specific Language) dans le contexte de l'ingénierie des modèles. Lorsque l'on disposait d'une bibliothèque d'objets propres à un domaine, la réalisation d'un outil avec la version C++ de GIGA passait par un ensemble de manipulations pouvant prendre une journée. Avec un langage à typage dynamique et des techniques d'héritage dynamique il aurait été possible de fournir un outil automatique, donnant la possibilité à un utilisateur de choisir et de charger une bibliothèque d'objets propres à un domaine, puis d'utiliser l'outil qu'il venait de construire dynamiquement.

GIGA a prouvé son utilité et son potentiel de réutilisabilité sur différents domaines d'application. La boîte à outils GIGA nous a par exemple permis de réaliser en une semaine un outil de simulation de réseaux de Petri (PSA, 11 000 lignes C++) [Caux *et al.* 1991]. Les études menées dans d'autres domaines que celui des systèmes de production (systèmes administratifs,

systèmes informatiques, systèmes hospitaliers, systèmes de transports), ont montré l'intérêt porté à cet outil, mais aussi le caractère générique des logiciels construits avec GIGA. Ces logiciels donnaient la possibilité d'intégrer dans le comportement des objets spécifiques au domaine, les tests de vérification et de validation que l'on souhaitait.

4. LA GÉNÉRATION AUTOMATIQUE DE CODE POUR L'ANIMATION ET LA SIMULATION

Parmi les diverses applications de la boîte à outils GIGA j'avais abordé la génération automatique de code de simulation à partir d'une spécification graphique de modèle. Le terme consacré aujourd'hui serait « transformation » de modèles. Grâce aux classes développées, le code de simulation généré permettait, ensuite, d'animer le modèle graphique saisi. La conception de ces logiciels était réalisée avec le processus et le langage graphique (notation) de M2PO.

58

Pour tous mes développements, j'avais réutilisé, d'une part, les résultats de l'étude de GIGA (de l'analyse de domaine à l'implémentation) et, d'autre part, les résultats de l'analyse de domaine effectuée pour le projet SIGMA concernant un environnement de modélisation pour les systèmes d'assemblage à flux discrets [Breugnot *et al.* 1990]. Cette dernière analyse a permis d'identifier les classes d'objets du domaine des systèmes d'assemblage à flux discrets, ainsi que les relations existantes entre ces classes. Après cette étape de réutilisation, il fallait concevoir un ensemble de classes d'objets graphiques adaptés pour l'animation et spécifier leurs comportements internes de manière à augmenter le potentiel de validation et de vérification du logiciel d'animation final. Dans un premier temps, nous utilisions donc GIGA pour réaliser un animateur de modèles de simulation de systèmes de production pour le groupe Valeo¹. Le développement a été effectué en relation directe avec les experts des différents sites Valeo existants en France. Par la suite cet animateur, nommé Game, a été commercialisé sous le nom de Viewmod par la société SIMULOG, grâce à un transfert de technologie entre cette société et notre Université. La description détaillée de ce logiciel se trouve dans [Hill et Junqua 1990] et dans d'autres articles [Breugnot *et al.* 1991a,b,c]

1. Valeo est un équipementier automobile.

[Hill 1993b]. Ce logiciel communiquait avec le simulateur suivant les protocoles établis pour GIGA, et il était capable de détecter et de signaler les incohérences d'un programme de simulation (erreurs de programmation ou de modélisation). L'analyse propre à la réalisation de Game, ainsi que la spécification des comportements des objets d'animation, utilisait la notation graphique de M2PO.

Pour être en mesure de générer automatiquement du code de simulation orienté objet à partir de la saisie graphique d'un modèle de système de production, j'utilisais le découpage en sous-systèmes physique, logique et décisionnel tels qu'il était préconisé dans le projet SIGMA et dans M2PO. La technique proposée pour la génération de code utilisait une combinaison de l'approche transaction et de l'approche station (propre à la simulation des réseaux de files d'attente). La génération de code de simulation permettait d'éviter la majorité des erreurs de programmation d'un modèle de simulation. L'animation facilitait alors la mise en évidence des erreurs de modélisation, aidant ainsi à la validation du modèle de connaissance d'un système. Les langages de simulation cibles pour la génération de code étaient principalement Qnap2, Siman IV, Simula et Occam 2. Les logiciels utilisant GIGA ont pu être testés sur plus d'une dizaine de cas industriels, mais aussi sur une simulation du fonctionnement administratif d'une préfecture, sur des systèmes de transports [Ruch 1994] et sur la simulation du système des urgences de l'hôpital de Bastia [Combes *et al.* 1994].

59

La génération de code de simulation orientée objet nous avait également permis d'identifier des lacunes au niveau des fonctionnalités du langage à objets de QNAP2, notamment concernant la classification dynamique et le polymorphisme. J'avais donc présenté une solution permettant de combler ces lacunes, en utilisant un compromis entre les techniques d'implémentation de la délégation et de l'héritage dynamique de méthodes [Hill 1993c].

Parmi les autres utilisations de la boîte à outil GIGA citons les réseaux de Petri. La partie édition et animation des réseaux que nous avons détaillée dans [Caux et Hill 1991], et les techniques utilisées pour modéliser le fonctionnement d'un ensemble de réseaux de Petri interprétés et communicants sur un réseau de transputers avaient été étudiées dans la fin des années 1980, elles sont présentées dans [Gourgand et Hill 1990].

5. CONCLUSION

Durant les années 1990 à 1993, nous avons réalisé des outils de simulation visuelle par objets pour des systèmes à flux discrets (de production, de transports, de systèmes informatiques et des systèmes administratifs). J'ai pu proposer un processus de modélisation, une méthode et un environnement de programmation visuelle de modèles de systèmes à flux discrets. Un des aspects intéressants des outils construits concernait la génération automatique de code de simulation pour différents outils et langages de simulation (Siman IV, Qnap 2, Simula 67) en se basant sur le formalisme des réseaux de files d'attentes. Cette génération automatique de code a pu être validée sur des cas industriels à partir d'une saisie graphique de modèles. L'approche utilisée effectuait une combinaison entre une approche station et une approche transaction en ce qui concerne la circulation des flux au sein du réseau de files d'attentes sous-jacent. Le logiciel Qnap2 a pu être adapté en proposant des techniques pour implémenter la classification dynamique et le polymorphisme. Le fait que des codes de simulation aient pu être générés pour plusieurs langages, et également la simulation de réseau de Pétri, a permis de tester la réutilisabilité des outils construits.

60

Un autre point relatif aux travaux de cette époque concerne le développement d'un outil d'animation de résultats de simulation reposant sur le même type de saisie graphique, que l'on nommerait aujourd'hui programmation visuelle. Cet outil a pu être utilisé et commercialisé par la société Simulog dans le cadre d'un accord avec le Laboratoire d'Informatique de l'Université Blaise-Pascal. Tous les outils développés durant cette période ont pu être validés sur de nombreux systèmes de production du groupe Valeo [Hill 1993b, 1996]. En 1995, dans un cadre liant les systèmes administratifs et les systèmes de production de logiciels, nous avons travaillé avec Éric Vigor (pendant l'encadrement de sa thèse) sur la proposition d'un métamodèle UML des systèmes de production de logiciels. Ce métamodèle avait été conçu pour l'aide à la gestion des grands projets informatiques de la Société Générale [Vigor 1998] [Hill et Vigor 1998a, 1998b]. Nous trouvons dans cette approche tous les éléments qui ont fondé l'architecture MDA (Model Driven Architecture) proposée en 2000 par l'OMG et tous les éléments théoriques qui fondent aujourd'hui l'ingénierie des modèles. Indépendamment de ces travaux, j'ai eu, dès 1993, grâce à Patrick Coquillard, l'opportunité d'appliquer à la modélisation d'écosystèmes les connaissances acquises en génie logiciel à objets et en modélisation.

CHAPITRE 4

UNE APPROCHE DE LA MODÉLISATION D'ÉCOSYSTÈMES

*S'ils ont poussé la science jusqu'à pouvoir prédire
le cours naturel des choses, comment n'ont-ils pas découvert
auparavant le Maître de celles-ci ?*

Livre de la Sagesse, 13,9

1. INTRODUCTION

Dans ce chapitre, je présenterai quelques activités de modélisation des écosystèmes au LIMOS. Nous avons eu à traiter des projets concrets de modélisation d'écosystèmes principalement en collaboration avec le Laboratoire d'Ecologie Végétale et Cellulaire de l'Université d'Auvergne (LEVC), l'INRA, le CEMAGREF, le Laboratoire d'Environnement Marin Littoral (LEM) de l'Université de Nice et d'autres partenaires internationaux. Pour tous ces projets j'avais principalement un rôle d'intégrateur. L'intégration de techniques logicielles pour la simulation d'écosystèmes nous a amené à associer le Web, les Systèmes d'Information Géographique (SIG), l'infographie ainsi que des outils d'analyse statistique. Les simulations sur le Web initiées en 1996 ont débouché en 1998 sur la première conférence de « Web-Based Simulation » à San Diego que nous avions organisée avec Paul Fishwick.

Dans un premier temps j’aborderai le couplage entre la simulation à événements discrets et les Systèmes d’Information Géographique car toutes nos simulations prennent en compte de fortes contraintes spatiales. Puis je présenterai l’étude et l’optimisation de modèles de simulation de forêts dans le Massif central, avec la présentation de la conception d’un modèle de couverture spatiale. Dans ce contexte j’ai pu avec des collègues biologistes simuler la génétique lors de la reproduction sexuée des bruyères en compétition avec les pins sylvestres. Cette partie ne sera pas présentée pour ne pas alourdir la présentation du manuscrit, le lecteur intéressé peut se reporter à l’article [Coquillard *et al.* 1997]. Cet article expose nos travaux dans ce domaine.

Une étude importante que j’ai choisi de présenter concerne la modélisation et la simulation de la propagation de l’algue *Caulerpa taxifolia*. Ce thème de recherche se situe au sein d’une affaire médiatique et controversée. J’ai tenté de développer avec le LEVC et le LEML un modèle sous contrainte spatiale qui permette de reproduire l’évolution de l’expansion de cette algue tropicale en Nord Méditerranée. En 1996, un contrat européen Life m’a donné la responsabilité scientifique de la partie modélisation, des travaux de calibrage ont permis une validation du modèle sur plusieurs sites à différentes échelles spatiales. Nous présenterons aussi les simulations de compétition entre la Caulerpe et *Posidonia oceanica*, une espèce Méditerranéenne protégée par l’U.E.. Un contrat du ministère de l’Environnement nous a permis en 1999 de continuer les travaux, en abordant à la fois des techniques de métamodélisation et la simulation du contrôle biologique par des limaces Ascoglosses qui sont des prédateurs naturels pour la Caulerpe.

62

Par la suite je présenterai des simulations multi-agents appliquées à l'éthologie. Une de ces applications avait été réalisée au sein d'un Groupement d'Intérêt Scientifique pour la modélisation de la gestion de la végétation et de l'entretien des milieux en moyenne montagne par des herbivores. Cette activité de modélisation des interactions entre vaches et chevaux entraînait à l'époque des retombées médiatiques en raison du suivi des animaux par satellites GPS.

Enfin j’aborderai les simulations sur le Web incluant la visualisation et l’animation de résultats de simulation. Tous les modèles que j’ai pu développer ont une interface graphique autorisant une visualisation et une animation au cours des simulations. Certaines simulations étant réalisées pour explorer les théories de nos collègues biologistes, nous avons développé avec André Campos des simulations avec des possibilités d’interaction à distance sur le Web (dans le cadre de la thèse d’André que j’avais eu la chance d’encadrer).

2. LE NÉCESSAIRE COUPLAGE SYSTÈME D'INFORMATION GÉOGRAPHIQUE / SIMULATION

2.1. *Introduction*

Les premiers modèles d'écosystèmes que j'ai eu à réaliser étaient des modèles spatialisés visant à comprendre le comportement d'écosystèmes réels. Nous avons donc, parallèlement aux développements des modèles, mené avec nos collègues du LEVC une réflexion sur les possibilités de couplage entre, d'une part, des outils de simulation et, d'autre part, des Systèmes d'Information Géographique (SIG) [Coquillard *et al.* 1995]. Par la suite nous avons présenté une nouvelle catégorie de simulateurs intégrant des contraintes spatiales [Coquillard *et al.* 1996], puis nous nous sommes également attachés à l'interprétation des résultats de couplage de SIG et d'outils de simulation discrète d'écosystèmes [Mazel *et al.* 1996]. Nous présentons dans ce qui suit les raisons qui ont motivé ces travaux, ainsi que l'approche retenue.

De nombreuses modélisations d'écosystèmes reposent sur un ensemble de simplifications réductrices, mais majeures, de l'écosystème, dont la plus importante est, sans aucun doute, la non-prise en compte des effets spatiaux existant au sein des écosystèmes, effets qui constituent une grande part explicative de leur comportement : vitesses d'évolution, structuration de l'espace, compétition, etc. Négliger les effets spatiaux revient à supposer :

1. l'homogénéité des facteurs écologiques stationnels,
2. l'homogénéité de la végétation.

Ces deux hypothèses, on en conviendra aisément, ne sont que très rarement réalisées dans la réalité. Il convient toutefois de préciser que ces simplifications peuvent être considérées comme raisonnables à l'extrême condition de s'adresser à de petites surfaces. Le type de résultats fournis par les modèles classiques sont ceux des mathématiques : courbes et résultats statistiques. En aucun cas ils ne sont aptes à fournir une idée approchée de la répartition et des densités locales des individus, de la structure du couvert végétal, etc., à partir d'une surface hétérogène.

La gestion des problèmes de l'environnement implique donc de plus en plus fréquemment des modélisations faisant interagir l'espace et le temps, notamment lorsque l'on s'intéresse à la prédiction des états des écosystèmes.

La prise en compte de l'espace à différentes échelles est du domaine des SIG. D'autre part, les techniques de simulation permettent de prendre en compte, en fonction du temps, le comportement dynamique de systèmes complexes tels que les écosystèmes. Parmi les techniques de simulation, la Simulation Aléatoire à Événements Discrets (SAED) permet de prendre en compte les aspects stochastiques de ces systèmes. Le couplage SIG - SAED offre donc des possibilités très intéressantes pour étudier les relations très complexes reliant des systèmes dynamiques stochastiques avec des structures spatiales.

L'utilisation de la SAED favorise le couplage avec les SIG, ainsi que le recensement et la description des interactions locales interindividuelles qui décrivent la dynamique d'un écosystème. La SAED permet aussi une discréttisation du temps et de l'espace propice à des descriptions plus réalistes du fonctionnement des écosystèmes. Les principaux avantages des modèles discrets sont [De Angelis et Gross 1992] [Coquillard et Hill 1997] :

- 64
- de différencier chaque individu biologique avec l'intégration de tous les paramètres que l'on souhaite,
 - de prendre en compte les effets spatiaux pour les interactions discrètes entre individus (interactions locales et à distance),
 - d'éviter un recensement des données sur de très longues périodes (nécessaires par exemple pour l'analyse markovienne en écologie terrestre),
 - d'éviter une modélisation mathématique ardue et trop réductionniste,
 - de minimiser les problèmes d'instabilités numériques.

Nous présentons dans le paragraphe suivant l'intérêt du couplage SIG - SAED, puis nous exposerons les relations entre les données d'une SAED et les informations fournies par un SIG.

2.2. *Intérêt du couplage SIG - SAED*

L'outil de simulation permet la reproduction de l'évolution d'un système (sous forme de changements d'états de celui-ci) dans le temps. Il est donc essentiellement basé sur le temps, qui est géré automatiquement par le moteur de simulation (gestion d'une horloge, d'un échéancier,...). Par ailleurs l'utilisation d'un modèle orienté objets permet la prise en compte d'entités biologiques élémentaires. En ce qui concerne le problème de la prise en compte de l'espace dans une simulation, celui-ci n'est qu'un para-

mètre parmi d'autres, devant être modélisé et géré par le concepteur de la simulation lui-même. Suivant l'importance accordée à ce paramètre, l'espace sera représenté par un simple attribut individuel ou bien par une structure de données plus ou moins sophistiquée, comme par exemple une matrice de listes chaînées des objets à simuler (**Figures 6 et 7**).

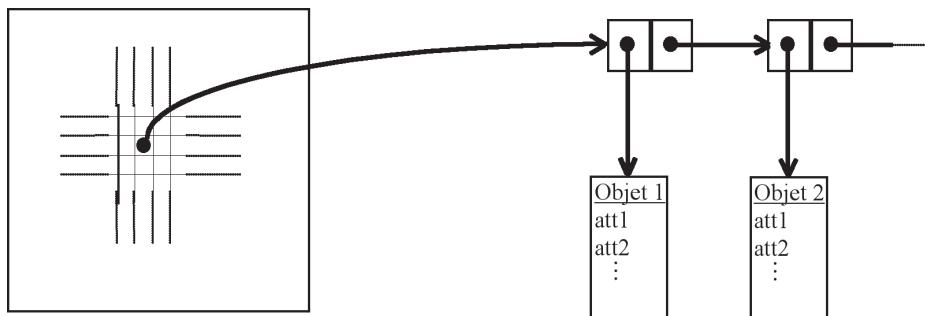


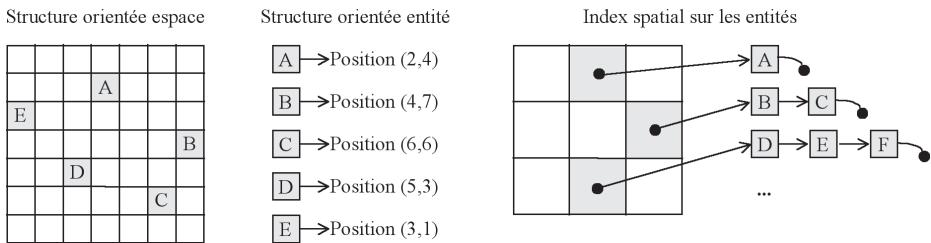
Figure 6. Une matrice de listes chaînées d'objets à simuler

De manière duale, la vocation d'un SIG est la gestion et la représentation de données géographiques, donc liées à l'espace. La prise en compte de l'évolution de ces données dans le temps est alors un problème délicat, car les SIG « *sont statiques et ne comportent pas de procédures de gestion du temps* » [Vasconcelos *et al.* 1994]. Même si les avancées dans ce domaine sont significatives et intéressent les chercheurs [Parker 1996], il n'existe pas encore de SIG commercial intégrant pleinement la notion de temps [Peterson 1995]. En fonction de la nature plus ou moins dynamique des entités représentées, les solutions apportées par les SIG sont plus ou moins complètes : animations à partir de cartes obtenues à différentes dates [Batty et Howes 1996], ou par extrapolations de cartes connues, fonctions de calcul de temps de diffusion [Aronoff 1989].

65

La complémentarité des potentialités des SAED et des SIG, concernant respectivement la gestion du temps et de l'espace, met en évidence l'intérêt de couplages SIG - SAED pour la modélisation spatiale de systèmes dynamiques [Zeigler 1984] [Fishwick 1995] [Coquillard *et al.* 1995] [Mazel *et al.* 1997] [Hill 1997c]. Toutefois la mise en œuvre effective de couplages étroits entre SAED et SIG, c'est-à-dire de couplages qui permettraient des modifications et des consultations totalement interactives du simulateur sur le

SIG, représente une très grande difficulté, aussi bien technique que conceptuelle (problème de la représentation de phénomènes dynamiques) [Sinton 1978]. Parmi les travaux menés dans cette direction, mentionnons ceux de [Vasconcelos *et al.* 1993] qui présentent une méthodologie de couplage entre un SIG et DEVS-Scheme (qui est un environnement de modélisation orienté-objets intégrant un formalisme de SAED avec des schémas de représentation de la connaissance issus des techniques de l’intelligence artificielle) [Zeigler 1990]. Vasconcelos présente un exemple d’application de cette méthodologie à un modèle de simulation de propagation d’incendie [Vasconcelos *et al.* 1994].



66

Figure 7. Exemple de différentes structures de données possibles pour gérer les données spatiales au sein d’une simulation à événement discrets [Campos et Hill 1998a]

Il est probable que l’utilisation croissante de la SAED pour la réalisation de modèles d’écosystèmes de grande taille qui nécessitent l’utilisation de nombreuses données géoréférencées, ainsi qu’une prise en compte des aspects dynamiques de ces systèmes, fera apparaître des besoins en outils permettant des couplages interactifs entre SAED et SIG.

Lorsqu’il est couplé, de manière interactive ou non, avec une SAED, un SIG conserve ses fonctions fondamentales : être une source de données, bien sûr, mais aussi permettre l’archivage des résultats [Peterson 1995]. Les deux points suivants présentent respectivement pour chacune de ces deux fonctions les contraintes liées au couplage :

1. l’utilisation d’un SIG comme source de données permet, pour le moins, de disposer de données réalistes pour initialiser une simulation. Cependant les données fournies par le SIG ne sont pas forcément utilisables directement, et il peut être éventuellement nécessaire

- de leur faire subir un traitement préalable permettant de mettre en adéquation les modèles spatiaux du SIG et de la simulation,
2. utiliser un SIG pour récupérer des résultats de simulation nous semble être une idée très intéressante. Ainsi, les résultats de simulation peuvent sortir du simple cadre des courbes et des résultats statistiques pour aboutir à un ensemble de données réellement manipulables et visualisables par les experts du domaine. Toutefois comme nous le verrons, l'interprétation de tels résultats n'est pas sans poser quelques problèmes ; nous y reviendrons dans le chapitre dédié à la validation.

2.3. Conclusion sur les relations entre les données d'une SAED et les informations fournies par un SIG

Quel que soit le modèle spatial adopté par le SIG, celui-ci peut fournir à la SAED, de manière directe ou après traitement des données, un modèle géographique représenté par un maillage qui fournit une partition de l'espace en entités géographiques élémentaires, désignées par 'EG' dans la suite de ce paragraphe. Pour chaque EG, le SIG fournit un ensemble d'attributs géographiques, correspondant à des caractéristiques du terrain. L'avantage de cette approche est de pouvoir stocker le maillage en mémoire afin d'atteindre les performances nécessaires à la SAED.

67

Par ailleurs, l'analyse orientée objets du système aboutit à la définition d'objets dotés de méthodes qui leur permettent d'agir sur les paramètres du terrain, de réagir aux modifications de ces paramètres, d'interagir directement entre eux,... Nous appellerons ces objets des entités biologiques élémentaires actives, désignées par 'EB' dans la suite de cette section, et qui ne sont pas forcément des individus au sens biologique du terme. Le principe du couplage SIG - SAED est de faire évoluer ces EB sur le modèle géographique qui représente leur environnement. Ces EB possèdent un certain nombre d'attributs, dont certains permettent de les caractériser (attributs biologiques) et d'autres de les localiser (attributs géographiques). Par exemple pour une simulation de croissance de forêt, les EB sont des arbres dont les attributs biologiques sont leur espèce, leur taille, leur âge,... et les attributs géographiques leurs coordonnées, leur numéro de parcelle,... [Mazel *et al.* 1996].

Il est important que la représentation de l'espace utilisée dans la SAED soit cohérente avec la nature des EG (et donc avec le modèle spatial utilisé par

le SIG) et qu’en particulier les attributs géographiques fassent directement référence à l’EG d’appartenance. Si les données fournies par le SIG sont discrètes, comme par exemple dans le cas de mesures effectuées en différents points d’un terrain, et que chaque EG correspond à un point de mesure, il est bien sûr préférable que la SAED gère également l’espace de manière discrète.

De même que les EB doivent disposer d’attributs géographiques permettant de les localiser, les EG, en particulier dans le cas où elles peuvent être occupées par un nombre important d’EB, peuvent disposer d’attributs biologiques fournissant des informations sur leur population. Pour l’exemple d’une simulation de croissance de forêt, l’information nécessaire à l’initialisation de la simulation peut être générée au moyen d’un modèle théorique [Pukkala et Kolstrom 1991], mais elle peut également être fournie, au moins partiellement par le SIG : si par exemple chaque EG dispose de la répartition des effectifs des différentes espèces sur son territoire, le modèle théorique permettant l’initialisation peut être appliqué à chaque EG en respectant les proportions d’espèces, et en permettant ainsi d’initialiser le modèle de simulation avec des données théoriques, mais probablement plus proches de la réalité [Brzeziecki *et al.* 1994].

3. LES SIMULATIONS EN FORESTERIE

3.1. *Introduction*

La simulation de croissance de forêt a été une des premières applications que j’ai abordée avec le LEVC [Coquillard *et al.* 1995] [Hill *et al.* 1995b]. Nous avions retenu une approche individu centrée malgré le fait que cette technique soit parfois lourde lorsque l’on s’attache à modéliser de vastes territoires. La prise en compte des contraintes spatiales nous a permis d’optimiser sensiblement ce type de simulation [Mazel *et al.* 1996]. Nous avons pu également participer à l’encadrement de la thèse de Bernard Prévosto, au CEMAGREF, et plus précisément à la réalisation d’un modèle permettant de définir la dynamique d’évolution des boisements spontanés, en nous limitant aux peuplements monospécifiques de pin sylvestre et de bouleau [Prévosto *et al.* 1999a,b,c] [Prévosto 2000].

Il existe différents types de modèles qui sont couramment utilisés pour la simulation de la dynamique forestière. Les modèles markoviens [Horn 1975a] peuvent converger vers des forêts stables. Des modèles markoviens non homogènes, où les probabilités de transition évoluent au cours du temps, ont également été proposés [Horn 1975b] [Acvedo 1981]. Ces modèles sont intéressants mais incapables de simuler les processus spatiaux tels que la dispersion des graines, la reproduction végétative ou la propagation de trouées. De plus, les modèles markoviens de forêts se concentrent sur la composition globale de la forêt et supposent qu'il y a un état stationnaire stable (« climax »).

Avec nos collègues, nous souhaitons montrer l'émergence d'états stables en modélisant les interactions dynamiques locales. Une parcelle de petite taille n'est pas vraiment stable en raison du couvert végétal et des nombreuses perturbations provoquées lorsqu'un arbre dominant tombe et que la compétition entre ses successeurs est relancée. Ces constats ont entraîné un concept de « climax » différent : « the shifting-mosaic steady state » présenté par [Bormann et Likens 1979]. La forêt peut être visualisée comme un tableau de parcelles ayant des dynamiques locales fortes sans empêcher le fait que l'ensemble de la forêt soit perçue comme stable. Cette dynamique des parcelles peut être simulée à l'aide de ce que l'on appelle les « gap models ».

69

Ce nom de « gap model » provient du constat que les premiers modèles en écologie animale se concentraient sur la dynamique d'une ou de deux population(s) en faisant abstraction de leur environnement, alors que les modèles en écologie végétale se concentraient sur la relation environnement - végétation [Botkin 1993]. C'est en faisant ce constat que Botkin et ses collègues ont proposé, dès 1972, un modèle qui comblerait le fossé « gap » entre ces deux approches pour l'étude des dynamiques forestières. C'est ainsi que le modèle JABOWA a été conçu [Botkin *et al.* 1972]. JABOWA est un modèle hybride basé principalement sur une dynamique de population, mais traitant différentes espèces, incluant des processus stochastiques, gérant la croissance et la régénération et utilisant des variables d'environnement. JABOWA est à la source d'un nombre très important de « gap models » qui peuvent varier suivant la forêt étudiée mais qui partagent un grand nombre d'hypothèses et une logique de fonctionnement identiques. Quelques « gap models » sont décrits dans [Botkin *et al.* 1972] pour JABOWA I, dans [Botkin 1993] pour le modèle JABOWA II, dans [Shugart et West 1977] pour le modèle FORET et dans [Leemans et Prentice, 1987]

pour le modèle FORSKA. [Bugmann *et al.* 1996] fournit une étude comparative et qualitative de différents « gap models ».

La fin des années quatre-vingts vit le développement de choix de modélisation reposant sur des techniques de recherche opérationnelle telles que la programmation linéaire (en nombre entier, mixte,...) : FORPLAN (FORest PLANning) est un outil de planification et d’optimisation des forêts qui utilise ces techniques [Johnson *et al.* 1986]. Ce type de modèle permet à l’utilisateur de fixer des contraintes, principalement des critères économiques et financiers. L’utilisateur peut alors minimiser ou maximiser certains coûts en combinant différentes activités de maintenance forestière. FORPLAN peut travailler au niveau d’une forêt entière ou au niveau de grandes parcelles sur des échelles de temps allant de 1 à 100 ans. FORPLAN est principalement un outil de planification des activités qui cherche à optimiser l’utilisation et l’allocation des ressources. Il est très bien diffusé et s’utilise couramment pour déterminer les moments de coupe [Kent *et al.* 1991].

Les modèles individus centrés tels que ceux proposés par [Urban et Shugart 1992] dérivent des travaux initiaux de [Huston *et al.* 1988] et [De Angelis et Gross 1992]. Avec cette approche, les arbres sont représentés individuellement, avec leur espèce, leur taille (hauteur, diamètre à hauteur de poitrine et même leur vigueur en prenant en compte leur histoire récente). L’influence des arbres les uns sur les autres est prise en compte. Le plus bel exemple de simulation de croissance de forêt avec une approche individu centrée a été développé par le CIRAD avec le logiciel AMAP issu de l’équipe de Philippe de Reffye [Saito *et al.* 1993], le niveau de détail et la qualité visuelle des résultats de simulation dérivent notamment des travaux de Frédérique Blaise. Ce type de modèle, très minutieux, qui s’applique à de petites parcelles, prend en compte de manière détaillée la croissance végétale, et offre une visualisation réaliste de très grande qualité. Les modèles auxquels je me suis attachés s’appliquent quant à eux à de grandes surfaces (plusieurs centaines d’hectares), ils n’offrent par contre pas le même niveau de détail (la croissance végétale, par exemple est modélisée avec une fonction globale). La principale difficulté réside alors dans la technique d’optimisation qui va permettre d’implémenter un modèle individu centré. C’est cette technique que je présente ci-après, elle est applicable à d’autres domaines que la foresterie et nous utilisons une technique similaire pour des modèles en océanographie. De plus, la manipulation d’un grand nombre d’individus biologiques par un programme de simulation suppose la re-programmation

de la gestion de la mémoire dynamique, c'est ce que nous avons fait à chaque fois que cela était nécessaire.

3.2. Application à la croissance de forêts avec prise en compte de l'effet spatial

L'application qui est présentée dans cette partie permet d'étudier l'impact des effets spatiaux sur la croissance d'une forêt dont la surface peut atteindre quelques milliers d'hectares. L'espèce étudiée est le pin sylvestre, mais les concepts qui sont exposés sont applicables à d'autres espèces arborescentes. Posons tout d'abord quelques définitions. La *couverture foliaire* d'une forêt est déterminée par l'ensemble des *couvertures foliaires* des pins qui composent la forêt. Si l'on considère une vue d'avion à la verticale, un arbre recouvre une surface délimitée par sa circonférence foliaire. La circonférence foliaire représente, comme le montre la **Figure 8**, la limite d'encombrement des feuilles et des branches. Nous nommons *couverture foliaire* d'un arbre la surface qui s'étend du tronc à la circonférence foliaire. Cette surface définit une zone hostile, où le peu de graines qui y germent n'ont aucune chance de survivre à cause du manque de lumière (compétition photique), par influence chimique (pluvio-lessivats)... Sur la **Figure 8**, la *couverture foliaire* est représentée en gris. La *zone d'influence* d'un pin correspond à la surface où il peut disperser des graines viables.

71

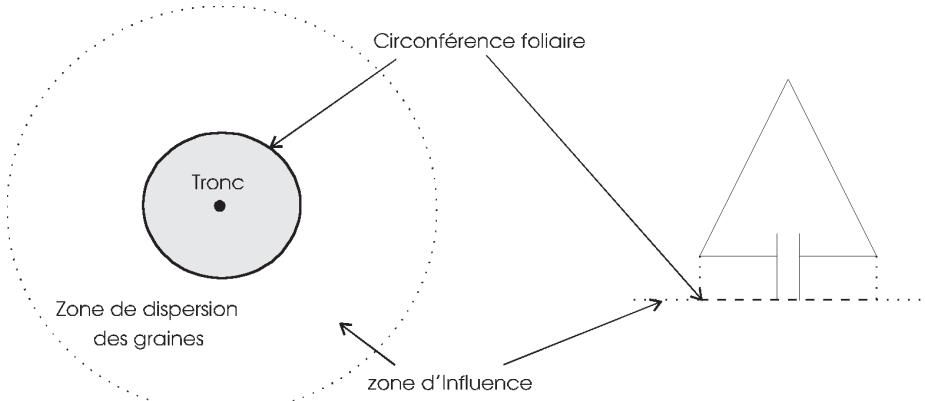


Figure 8. Circonference foliaire, zone de dispersion des graines et zone d'influence

La *couverture foliaire* correspond à la zone où les graines qui germent donnent naissance à de jeunes arbres qui ne peuvent survivre pour les raisons exposées précédemment. La *zone d'influence* d'une forêt est donnée par l'union de toutes les *zones d'influences* des pins. La croissance de la forêt se fait donc par conquête vers l'extérieur de la *couverture foliaire* et dans la limite de la *zone d'influence*. Cette conquête peut bien sûr s'effectuer vers l'intérieur de la forêt dans le cas de clairières.

La croissance spatiale d'une forêt est donc uniquement déterminée par l'évolution de la *couverture foliaire* et de sa *zone d'influence*. Ces deux surfaces permettent de déterminer quels sont les individus qui se trouvent « en limite extérieure » de la forêt et qui participent directement à son expansion. Les simulations de croissance de forêts réalisées dans le passé ont rarement pris en compte ce facteur spatial. Cette constatation se retrouve dans la synthèse proposée par Stout sur ce domaine [Stout 1991].

72

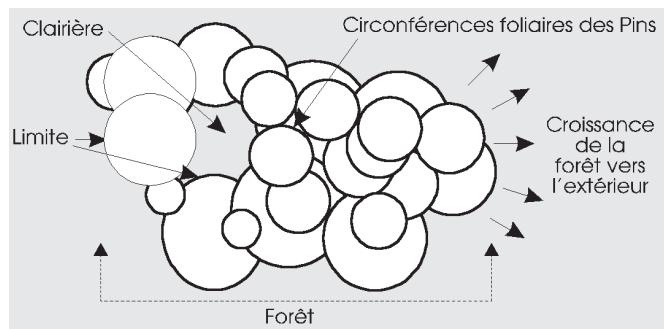


Figure 9. Couverture foliaire d'une forêt

Pour simuler la croissance d'une forêt de pins sylvestres, il est nécessaire d'étudier des événements significatifs tels que ceux liés à la reproduction, à la mort, à la croissance annuelle... Chaque année, les pins en âge de reproduction (20 ans et plus) produisent environ 54 000 graines qui se répartissent autour du tronc suivant une distribution gaussienne, comme le montre la **Figure 10**. Les campagnes de mesure réalisées ont montré, dans la zone géographique que nous étudions (Massif central) qu'en moyenne seulement 36 graines par pins sont viables et qu'en moyenne 0,6 jeunes pins survivaient par semencier et par an.

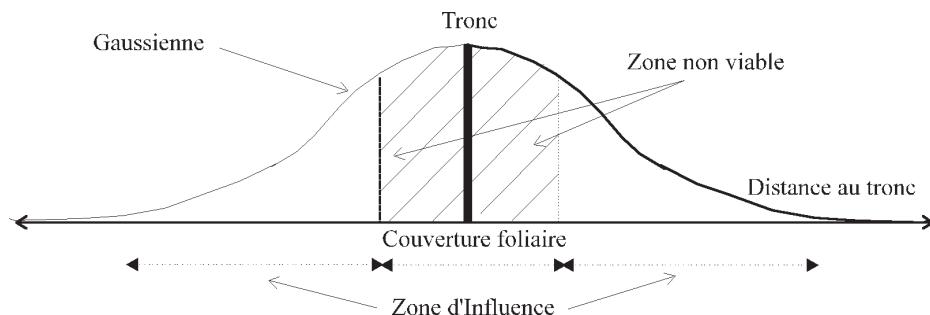


Figure 10. Répartition des graines autour du tronc

La *zone d'influence* est facile à déterminer pour un arbre isolé, mais elle devient complexe pour une forêt. En effet, il ne s'agit pas uniquement de la recherche d'une couverture convexe, car la forme d'une forêt ne correspond pas à un convexe. Pour modéliser l'expansion d'une forêt entière, il serait maladroit pour des raisons d'efficacité (temps de calcul), de prendre en compte les comportements individuels de la totalité des pins. En effet, seuls les pins qui se situent en bordure extérieure ou en bordure de clairières jouent un rôle pour la colonisation de l'espace par germination.

73

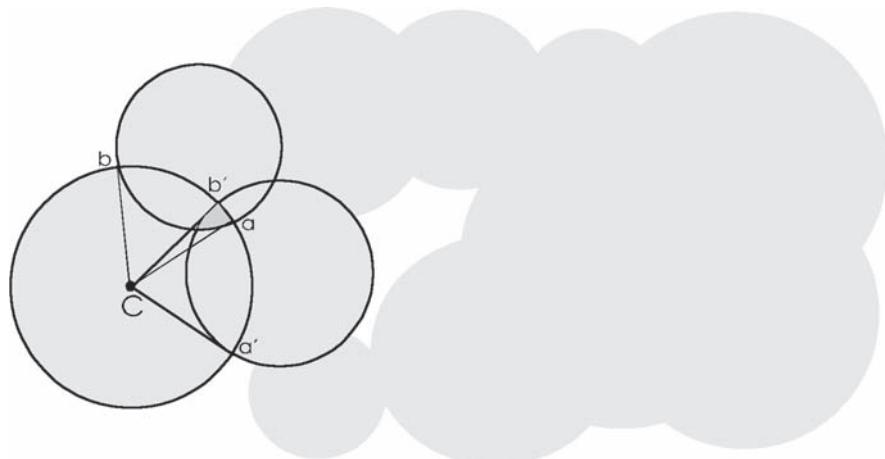
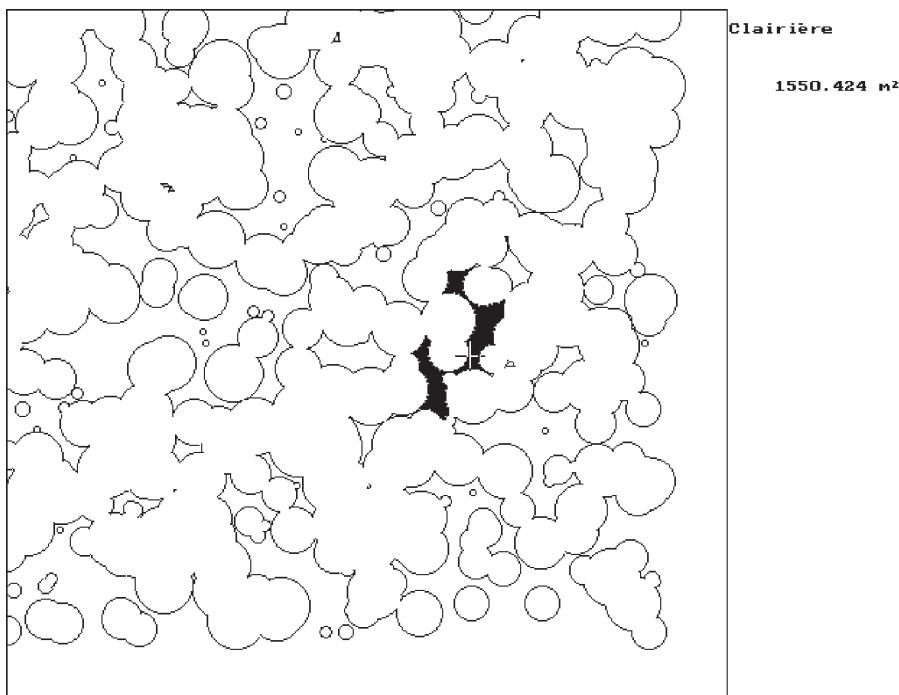


Figure 11. Modèle de forêt

Étudions sur des schémas quelques éléments supplémentaires utilisables par une simulation de croissance forestière. La surface grisée de la **Figure 11** représente la zone d'influence d'une forêt. Les rayons des *zones d'influence* ont des valeurs différentes afin de prendre en compte l'âge (et donc la hauteur), l'état de santé ainsi que l'espèce (type de graine et aptitude à la dispersion de ces graines). La **Figure 11** représente des *zones d'influence* circulaires, mais d'autres formes convexes (ellipse ou autre) permettraient de prendre en compte le vent ou la topographie. La surface en gris représente une zone qui est influencée par trois arbres et qui peut donc recevoir les graines qu'ils disséminent.

3.3. Conclusion

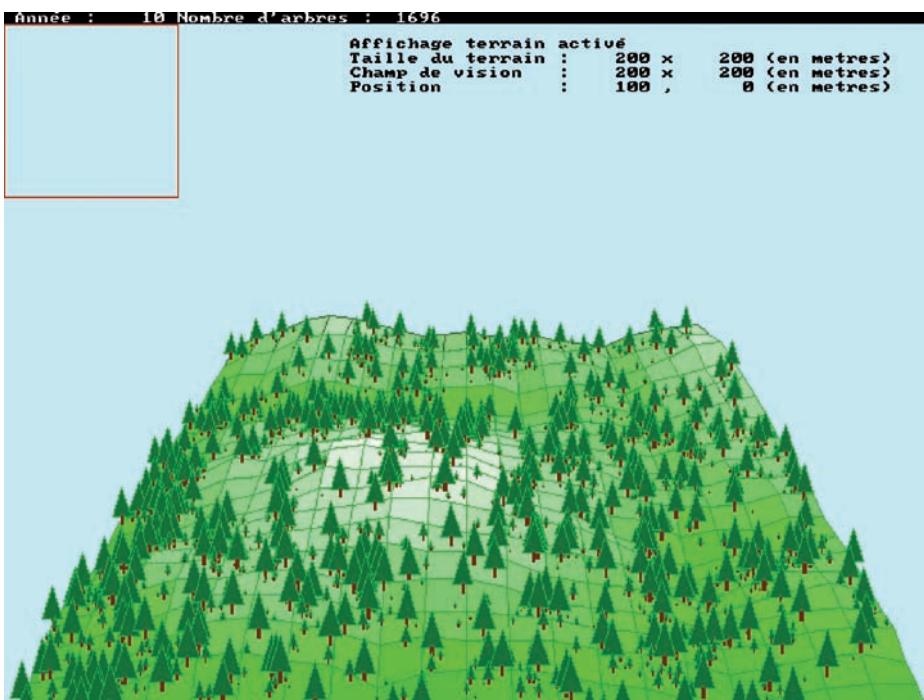
Actuellement, il n'existe aucune méthode analytique permettant de déterminer simplement cette *zone d'influence* modélisée par un ensemble de cercles. En attendant une solution mathématique efficace, nous avons présenté un algorithme de calcul d'une couverture d'influence généralisable pour un ensemble de convexes [Mazel *et al.* 1996]. Le lecteur intéressé par plus de détails se reportera à [Coquillard et Hill 1997] qui présentent cette méthode en tant qu'outil permettant de manipuler plusieurs visions d'un même ensemble d'objets. La **Figure 12** présente le résultat visuel d'un calcul de couverture foliaire. Enfin je tiens à ajouter une remarque technique sur l'implémentation de ce type de modèle qui manipule un grand nombre d'individus biologiques. Rappelons que même si la technique d'optimisation présentée ci-dessus autorise la simulation de ce type de modèle sur de grandes surfaces, il faut y ajouter l'implémentation d'une gestion de la mémoire dynamique dédiée. En effet, l'allocation dynamique telle qu'elle est implémentée dans les langages à objets tels que C++, que ce soit avec un système d'exploitation Unix ou avec Windows n'est pas adaptée. Le très grand nombre de petits objets ainsi que les limites en taille des segments de pile n'autorisent pas le codage de telles applications. Il convient d'utiliser de la mémoire statique et de re-programmer le codage de l'allocation dynamique sur cette mémoire statique en utilisant un algorithme dédié aux objets manipulés. De plus, le choix d'utiliser de la mémoire statique permet d'adresser la totalité de l'espace mémoire disponible sur une machine.



75

Figure 12. Écrans de couvertures foliaires calculées avec sélection d'une clairière et calcul de sa surface

Un logiciel implémentant l'algorithme de calcul qui a été présenté ainsi que les techniques de gestion de la mémoire dynamique a été développé à l'origine sur micro-ordinateurs compatibles PC. Le codage de l'algorithme de calcul de couverture spatiale a été réalisé à la fois en Simula et en C++, uniquement à des fins de comparaison académique. Un rendu réaliste en trois dimensions des simulations peut être obtenu avec des outils simples comme Vistapro. La simulation visuelle interactive qui a été développée utilise un rendu plus simple pour des raisons d'efficacité (*cf. Figure 13*). L'aspect visuel est principalement lié à la présentation et à la communication des résultats de simulation. La version actuelle du logiciel est implémentée en C++ sous Unix.



76

Figure 13. Écran de simulation visuelle interactive pour la croissance d'une forêt

4. LES SIMULATIONS APPLIQUÉES À L’OCÉANOGRAPHIE

4.1. Introduction

En 1994, j’ai été contacté par le Laboratoire d’Écologie Végétale et Cellulaire (LEVC) de l’Université d’Auvergne et le Laboratoire d’Environnement Marin Littoral (LEM) de l’Université de Nice Sophia-Antipolis, pour simuler l’expansion de l’algue *Caulerpa taxifolia* (Caulerpe) qui commençait à se propager dangereusement dans le nord de la Méditerranée autour de Monaco [Meinesz et Hesse 1991] [Meinesz et Belscher 1993]. La simulation de l’expansion de l’algue *Caulerpa taxifolia* a donc démarré en 1994 à la demande du LEM. C’est le travail considérable de ce laboratoire en matière de suivi cartographique de l’expansion de la Caulerpe [Vaugelas 1996] qui a permis la réalisation d’un modèle informatique (Figure 14). Dès 1994, nous avons pu réaliser un prototype de simulateur nommé SIMCT afin de

présenter des résultats préliminaires fin 1994 [Hill *et al.* 1994b]. Ce premier prototype a été affiné en 1995 [Hill *et al.* 1995b] et j'ai pu présenter les premiers résultats validés sur un site concret au Centre de Recherche Océanographique de Tokyo lors d'une conférence invitée. Ensuite, nous avons commencé à étudier l'impact sur la biodiversité [Vaugelas *et al.* 1996] et plus particulièrement les éventuelles compétitions avec une espèce protégée, l'herbier *Posidonia oceanica* [Hill *et al.* 1997a]. La validation sur plusieurs sites a pu être menée avec succès en modifiant sensiblement le modèle. Les évolutions du prototype sont présentées dans [Hill *et al.* 1997b] [Vaugelas *et al.* 1997]. Par ailleurs nous avons pu développer des techniques d'analyses spectrales pour étudier les aspects stochastiques des cartes produites par le simulateur SIMCT [Hill 1997b] et [Hill *et al.* 1998b]. Les temps de calcul sur de longues périodes (une dizaine d'années) et sur de vastes sites étant rapidement prohibitifs, notamment du fait des réplifications, nous avons envisagé le développement d'une approche de métamodélisation, où un réseau de neurones entraîné par les simulations discrètes a été capable de reproduire, avec un gain de temps très significatif, les résultats obtenus uniquement en utilisant les simulations [Aussem et Hill 1999]. Nous étudions actuellement la possibilité de réaliser l'apprentissage du réseau de neurones à partir de plusieurs sites réels, mais également à partir de sites virtuels créés avec une approche fractale et des caractéristiques statistiques identiques à celles rencontrées dans la nature [Aussem et Hill 2000].

Des simulations d'éradication ont pu récemment être réalisées et le lecteur intéressé trouvera une description des techniques envisagées dans [Escoubet *et al.* 1998] [Vaugelas *et al.* 1999]. Dans ce contexte, nous avons également pu simuler une lutte biologique avec une limace tropicale *Elysia subornata* et l'algue *Caulerpa taxifolia*. Ce contrôle biologique fut suggéré à l'Académie des Sciences dès 1997 par le Professeur Alexandre Meinesz [Meinesz 1997]. Les premiers résultats de simulation concernant cette technique de lutte ont été publiés en 2000 [Hill *et al.* 2000a] et [Coquillard *et al.* 2000]. Dans ce contexte je peux co-encadrer, sur les aspects modélisation informatique, la thèse de Thierry Thibaut, doctorant au LEML. L'introduction d'une autre espèce tropicale n'étant pas envisagée pour l'instant, d'autres espèces de limaces autochtones (*Oxynoe olivacea* et *Lobiger serradifalcii*) sont en cours d'étude [Thibaut *et al.* 2000] et [Thibaut et Meinesz 2000].

L'ampleur de ce projet lié à l'expansion de *Caulerpa taxifolia* en Nord-Méditerranée nous a permis de participer à un projet Européen Life (LIFE n° 95/A31/EPT/782) en tant que responsable des aspects modélisation. Ce

sujet reste d’actualité, la Caulerpe faisant son apparition au sud de la Californie et en Australie.

4.2. Le modèle d’expansion de l’algue *Caulerpa taxifolia*

Depuis son introduction en Méditerranée il y a seize ans, l’algue d’origine aquariologique [Meinesz et Hesse 1991] [Jousson *et al.* 1998] *Caulerpa taxifolia* ne cesse de s’étendre [Meinesz *et al.* 1996] et présente toutes les caractéristiques d’une invasion biologique majeure. Actuellement plus de 6 500 hectares sont concernés par cette algue et de nombreux pays sont affectés (France, Italie, Croatie, Espagne et Monaco...). Les zones atteintes ne cessent d’augmenter et sa découverte récente dans le Pacifique au sud de San Diego sensibilise nos collègues d’outre-atlantique.



78

Figure 14. Carte des zones atteintes par la Caulerpe en 1997

La réalisation de ce modèle est le résultat d’un réel travail d’équipe pluridisciplinaire. Les principaux objectifs assignés à la modélisation de l’expansion de *Caulerpa taxifolia* en Nord-Méditerranée se résument en six points :

- Simuler sur ordinateur l’expansion de *Caulerpa taxifolia* sur des sites précis, à partir d’un modèle de connaissance issu des travaux du LEML. Cette simulation est particulièrement utile dans les tentati-

ves de prévision de l'expansion sur les zones largement envahies et notamment sur les colonisations profondes (au-delà de 40 m). Cette simulation peut couvrir plusieurs années. En estimant la probabilité d'extension de l'algue dans un milieu difficile à cartographier comme l'herbier de Posidonie, il doit être possible de répertorier des colonisations non encore signalées.

- Explorer le comportement intrinsèque du modèle en le comparant en permanence avec la réalité. Seule la comparaison avec ce qui se passe réellement « sous l'eau » permet de caler les hypothèses de départ et d'affiner progressivement l'adéquation entre le modèle et la réalité du terrain. Ces tests permettent par ailleurs de se poser un certain nombre de questions fondamentales sur la stratégie de croissance de cette algue prolifique, sur les causes et les modalités de sa dissémination, ainsi que sur les réactions des espèces avec lesquelles elle est susceptible d'entrer en compétition. Nous avons déjà évalué les principaux paramètres de dispersion des fragments de *Caulerpa taxifolia* en fonction de divers milieux.
- Évaluer l'importance de divers paramètres de l'environnement (bathymétrie, biocénoses, substrats, température, courants,...) du (des) site(s) étudié(s) sur la dynamique de l'espèce [Boudouresque *et al.* 1994] [Ribera *et al.* 96] [Komatsu *et al.* 1997].
- Permettre diverses estimations de paramètres biologiques tels que biomasse et productivité, mais également des statistiques mensuelles liées à la colonisation : surfaces colonisées, évolution des densités, biomasses résiduelles d'espèces en compétition (la phanérogame marine *Posidonia oceanica* notamment).
- Explorer divers scénarios liés à la dispersion anthropique, au développement ou à l'introduction de prédateurs spécifiques (lutte biologique), etc. De même, l'efficacité et le coût d'une opération d'éradication peuvent être évalués. En simulant une éradication incomplète (un certain nombre de boutures « échappent » à la vigilance des plongeurs), il sera possible de calculer approximativement l'effort à déployer les années suivantes pour parvenir à garder le contrôle de la zone envahie.
- Évaluer rétrospectivement, pour un site donné, les différents stades d'invasion de manière à déduire l'année et le point le plus probable d'introduction.

Figure 15. Photo de l’algue *Caulerpa taxifolia* avec ses frondes (grandes feuilles) et les rhizoïdes (racines rampantes qui présentent pour cette algue une croissance très rapide)



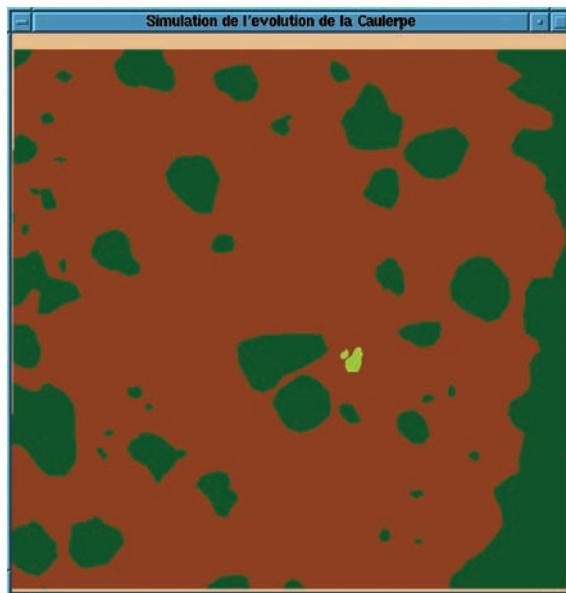
La prise en compte de l'espace dans le modèle d'expansion de *Caulerpa taxifolia* (et notamment la progression discrète des stolons et la dispersion discrète de boutures de manière naturelle ou anthropique) a impliqué l'utilisation des techniques de simulation aléatoire à événements discrets. Le simulateur que nous avons conçu pouvait, dès 1994, produire de l'information exploitable par le S.I.G Mapgraphix qui est utilisé au LEML de Nice [Hill *et al.* 1994b]. Il est important que les résultats de simulation puissent être présentés sous forme de cartes directement exploitables par les experts du domaine. On peut également obtenir ainsi des images de la répartition dans l'espace des entités et de leurs attributs. Ces images successives de l'écosystème en évolution sont trompeuses, car elles ne définissent que l'une des multiples trajectoires possibles, en raison de la nature intrinsèquement stochastique du modèle (et du système réel). Pour de nombreux modèles de simulation sous contrainte spatiale, il peut s'avérer intéressant d'essayer de construire avec des analyses spectrales des images cartographiques « minimale, moyenne, maximale » obtenues après de nombreuses réplications des simulations stochastiques, et ceci afin de prendre en compte les multiples évolutions possibles de l'environnement. Nous traiterons plus particulièrement ces techniques dites d'analyse spectrale dans un chapitre dédié à la validation des modèles. Avec cette technique on peut obtenir en chaque point un intervalle de confiance sur la probabilité de présence des entités biologiques.

La simulation utilise en entrée un fichier d'expérience qui permet de spécifier la connaissance des experts biologistes en réglant un ensemble de paramètres. On trouve également en entrée du modèle, pour chaque site, les cartes précisant les substrats et la bathymétrie. Le modèle peut également

être affiné en précisant sur les cartes les zones portuaires et les courants lorsqu'ils sont connus.

4.3. Résultats de la simulation de la croissance de *Caulerpa taxifolia*

Un des sites de référence pour la modélisation se situe à Passable dans un trou de bombe de la Seconde Guerre mondiale. Sur la **Figure 16**, en foncé on retrouve les zones d'herbier de Posidonie et en clair les taches de Caulerpe ($3,5 \text{ m}^2$) (la largeur de la zone est ici de 60 m). Une cartographie très détaillée au 1:1000^{ème} et le repérage de chaque bouture avec un balisage a permis de calibrer le modèle, et notamment les taux de bouturage ainsi que les paramètres de dispersion. Le modèle stochastique utilise une loi de probabilité discrète pour la dispersion des boutures car le manque de données ne permet pas de faire un ajustement à une fonction de densité.



81

Figure 16. Situation initiale dans le trou de bombe sur une zone de 60 x 60 m. La tache claire au milieu de l'image correspond à $3,5 \text{ m}^2$ couverts par *Caulerpa taxifolia*, le gris foncé correspond à des parties d'un herbier de posidonie et le gris clair à de la matte morte de l'herbier

La Figure 17 est le résultat fourni par une simulation stochastique sur une période de un an. Pour cette simulation, la surface colonisée par l’algue durant cette période est de 58 m^2 . Tous les résultats fournis sur une carte, comme celle de la Figure 17, peuvent être considérés comme des réalisations de variables aléatoires, c’est le cas des résultats synthétiques comme la surface colonisée, par exemple. Il convient donc d’étudier statistiquement les résultats en établissant les bornes supérieures et inférieures, les moyennes et les intervalles de confiance. Après une validation des surfaces colonisées lors des premières années, les résultats de simulation indiquaient systématiquement la présence de boutures à l’ouest du trou de bombe, comme l’indique la Figure 18 (qui utilise une échelle plus petite car elle représente une surface d’environ $500 \times 600 \text{ m}$).

Les plongeurs sont allés vérifier la présence de l’algue et ils ont pu confirmer cette hypothèse en découvrant des taches nouvelles au sein de l’herbier dans une zone où personne n’était allé auparavant. Ce cas précis montre l’utilité de ce type de modèle dans les milieux difficiles d’accès : herbiers ou eaux profondes. Un modèle calibré peut, malgré la complexité du problème, se révéler prédictif dans un cadre expérimental précis

82

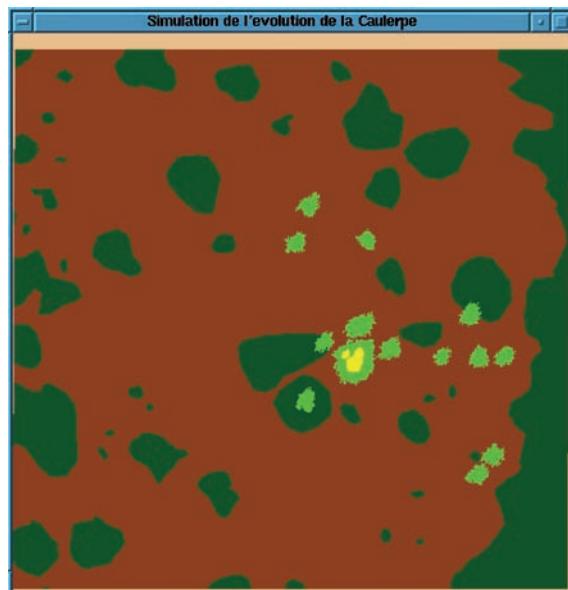


Figure 17. Après un an de simulation, 58.4 m^2 sont colonisés



Figure 18. La même zone est étudiée sur une échelle de 500 x 600 m. Cette simulation donne un résultat de 5 240 m² colonisés après 5 ans

83

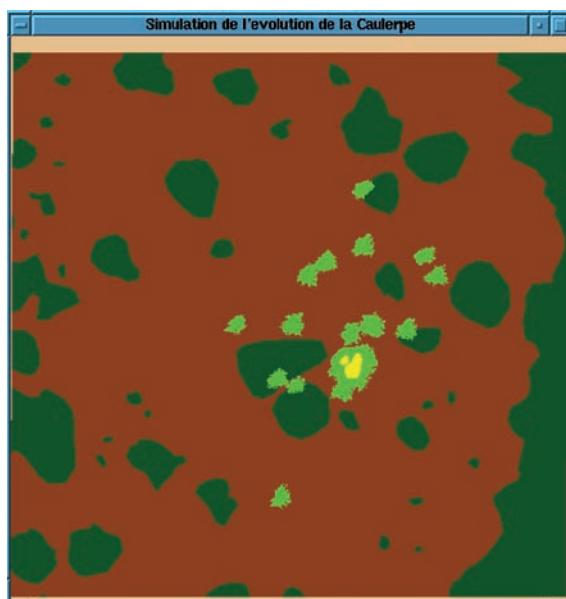


Figure 19. Deuxième réplique en rapport avec la Figure 17, 62.1 m² colonisés

Comme nous l’avions indiqué, les résultats fournis sur une carte comme celle de la **Figure 17** peuvent être considérés comme des réalisations de variables aléatoires. Voici sur la **Figure 19** une autre simulation sur une période d’un an au sein du trou de bombe, sur la même surface de 60×60 m. on obtient bien un résultat différent de 62.1 m^2 colonisés (par rapport aux 58.4 m^2 de la **Figure 17**), tout comme si l’on avait éradiqué l’algue mais que sa réintroduction eut mené à un résultat différent mais statistiquement équivalent. Pour toute simulation stochastique, il faut effectuer des réplications afin de fournir des résultats cohérents. Des techniques statistiques existent, par exemple pour calculer le nombre de réplications nécessaires pour obtenir une précision donnée (rayon de l’intervalle de confiance).

Les histogrammes présentés sur les **Figures 20** et **21** montrent les résultats obtenus avec respectivement mille et dix mille réplications de la simulation sur le site de Passable. La **Figure 22** présente un des écrans de contrôle du simulateur SIMCT avec notamment la saisie du nombre de réplications.

84

Figure 20. 1 000 réplications (distribution des surfaces sur un an de simulation)

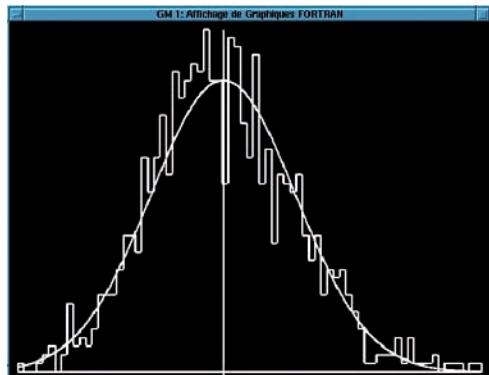
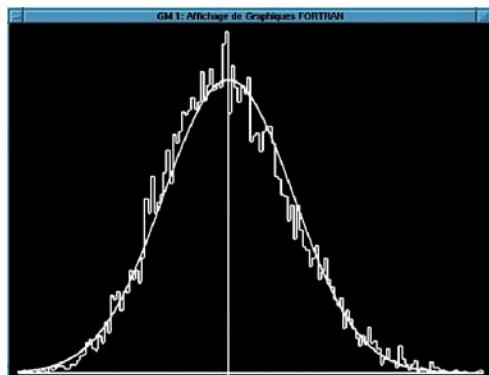


Figure 21. 10 000 réplications dans les mêmes conditions



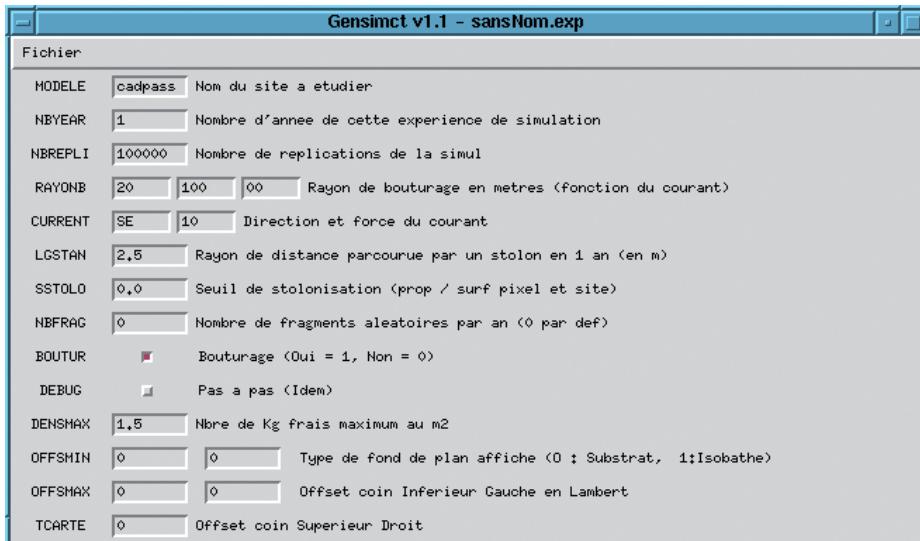


Figure 22. Fenêtre de saisie de quelques paramètres du modèle SIMCT

85

4.4. Simulation de la croissance de l'herbier de Posidonie

La Posidonie est un végétal marin, mais ce n'est pas une algue. Il s'agit en fait d'une plante à fleurs descendant d'un ancêtre terrestre qui devait ressembler aux joncs. L'espèce *Posidonia oceanica* ne se rencontre qu'en Méditerranée (**Figure 23**). Les herbiers ceinturent presque complètement la Méditerranée, ne s'interrompant qu'à l'embouchure des grands fleuves, et constituent un véritable oasis de vie. Au total, 400 espèces d'algues et plusieurs milliers d'espèces animales vivent dans l'herbier (qui peut descendre jusqu'à 42 m dans les eaux claires de Corse). L'herbier de Posidonie représente l'écosystème-pivot des espaces littoraux méditerranéens. Il constitue en effet un lieu de frayère et de nurserie pour de nombreuses espèces qui y trouvent nourriture et protection. Mais l'herbier protège également les plages contre l'érosion, en atténuant la puissance des vagues et des courants. La Posidonie est une espèce protégée par la loi depuis 1988.



Figure 23. Un herbier de Posidonie souvent comparé à un champ de poireaux avec de longues feuilles de plusieurs dizaines de centimètres

L'herbier est malheureusement menacé par les activités humaines et régresse un peu partout sur le littoral. À sa place, de vastes étendues désertiques de mattes mortes (zones où l'on trouve des « souches » d'herbiers morts) occupent les fonds. Les principales causes de sa régression sont les aménagements littoraux et la pollution (rejets urbains et industriels). Les ports et les plages artificielles ensevelissent l'herbier de manière irréversible. C'est ainsi qu'entre Marseille et Nice 10 % des petits fonds entre la surface et 20 m de profondeur sont définitivement détruits, recouverts de béton. Depuis une décennie c'est l'algue *Caulerpa taxifolia* qui colonise les Herbiers [Villèle et Verlaque 1995] [Bellan-Santini *et al.* 1996] (Figure 24).

86

Avec le logiciel SIMCT, j'ai également abordé la compétition avec l'herbier de Posidonie qui ralentit l'expansion de l'algue tropicale [Hill *et al.* 1997a]. La croissance de l'herbier est très lente de l'ordre de 3 centimètres par an, par

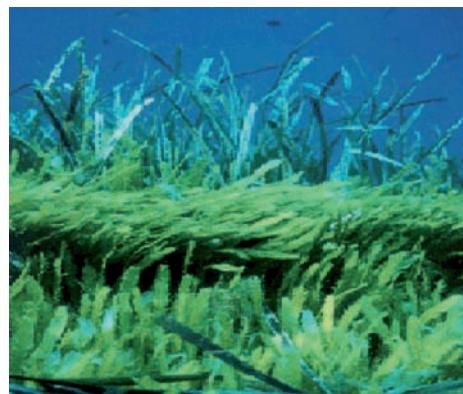


Figure 24. Un herbier de posidonie envahi par deux strates superposées d'algue *Caulerpa taxifolia*

contre celle de l'algue tropicale peut aller jusqu'à 3 centimètres par jour. Tout comme pour les simulations de croissance de forêt j'ai pris en compte les aspects spatiaux pour optimiser la simulation. L'évolution de l'herbier est très complexe, très lente et peut être fortement influencée par les courants locaux. J'ai simplifié l'étude en prenant plusieurs hypothèses de croissance annuelle. Le modèle de colonisation de l'espace par l'herbier repose tout d'abord sur l'identification des cellules en bordure de celui-ci, car ce sont celles qui seront actives pour la croissance. Nos simulations ont considéré une progression annuelle réaliste de 3 cm. Pour les besoins de l'étude j'ai également simulé des progressions de 6 cm., et même de 10 cm. par an, puis j'ai essayé une loi uniforme entre 3 et 10 cm. La **Figure 25** présente à gauche la carte initiale d'une zone de 500 x 600 m. dans la rade de Passable. L'herbier de Posidonie y figure en gris foncé. On note, comme précisé précédemment, la présence d'un trou de bombe de la Deuxième Guerre mondiale au milieu de l'herbier. La partie droite de la figure présente une hypothétique recolonisation du trou de bombe par l'herbier après 100 ans avec une croissance de 3 cm par an. Malheureusement comme il n'y a pas de réelle compétition avec la Caulerpe, qui s'introduit même dans les herbiers, cette recolonisation ne pourra pas avoir lieu étant donnée la vigueur de l'algue tropicale. La **Figure 26** présente la situation après cinq ans de simulation de croissance de la Caulerpe sur la même carte.



Figure 25. À gauche la situation initiale dans la rade de Passable et à droite une hypothétique recolonisation par l'herbier avec une croissance de 3 cm par an sur 100 ans

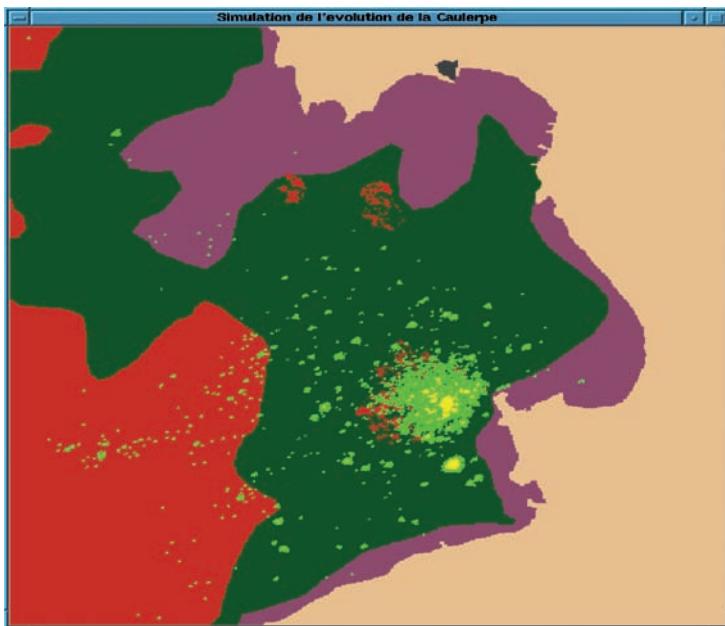


Figure 26. Simulation sur 7 ans avec prise en compte de l'introduction de l'algue tropicale

88

4.5. La lutte par un agent de contrôle biologique

Depuis 1992, différentes techniques d'éradication chimique et physique ont été testées pour éradiquer *Caulerpa taxifolia*. Outre leur coût de revient important, elles ne constituent qu'un élément de contrôle limité car elles ne sont réellement efficaces que sur de faibles surfaces, à faible profondeur et sur un substrat faiblement accidenté [Jaffrenou et Odonne 1994] [Cottalorda *et al.* 1996] [Gavach *et al.* 1998].

Il est couramment admis que lorsqu'aucune technique ne permet d'endiguer une invasion biologique, il convient de rechercher des prédateurs spécifiques de l'agent envahissant, en vue de leur utilisation potentielle comme agent de contrôle biologique [Lafferty et Kuris 1996]. Cette technique, couramment utilisée en milieu terrestre, n'a encore jamais été tentée en milieu marin. Actuellement plusieurs projets de ce type sont à l'étude dans le monde : contre le crabe européen *Carcinus maenas* qui prolifère en Tasmanie et en Californie, contre l'étoile de Mer *Asterias amurensis* qui envahit les côtes

australiennes et contre le cténophore *Mnemiopsis leidyi* qui envahit la mer Noire [Meinesz 1999]. Sous l'égide du CIEM (Conseil International pour l'Exploration de la Mer) un protocole d'études pour la mise en œuvre d'une lutte biologique en mer a été défini en avril 1997 (ICES/CIEM 1997).

Depuis 1994, le Laboratoire d'Environnement Marin Littoral a initié des recherches sur les possibilités d'utilisation de prédateurs spécifiques de *Caulerpa taxifolia*. Quatre espèces de mollusques gastéropodes (ordre des Ascoglosses) ont été étudiées afin d'évaluer leur potentiel d'utilisation comme agent de lutte biologique contre *Caulerpa taxifolia*. Deux sont d'origine méditerranéennes, *Oxynoe olivacea* et *Lobiger serradifalci*, et deux sont d'origine tropicales, *Elysia subornata* et *Oxynoe azuropunctata* [Meinesz *et al.* 1996] [Meinesz 1997] [Thibaut *et al.* 1998] [Thibaut et Meinesz 2000].

Ces espèces sont des consommateurs exclusifs d'algue du genre *Caulerpa* et présentent un fort degré de spécialisation pour leur nourriture [Jensen 1997]. En effet, leur morphologie buccale (radula unisériée) font qu'elles ne peuvent consommer que des algues coenocytiques comme les *Caulerpa* [Jensen 1993]. De plus, ces ascoglosses utilisent les toxines contenues dans leur nourriture afin d'assurer leur défense, notamment les métabolites secondaires des *Caulerpa* tels que la caulerpine et la caulerpicine [Paul et Van Alstyne, 1988]. Enfin pour *Elysia*, le stockage et l'utilisation des chloroplastes des *Caulerpa* (phénomène de kleptoplastie) est un besoin vital qui lui permet de survivre [Hinde et Smith 1974]. Elles fonctionnent alors, grâce à ces chloroplastes, comme des plantes en utilisant les produits de la photosynthèse [Clark 1992]. Les ascoglosses à coquilles comme *Oxynoe* et *Lobiger* sont exclusivement inféodés au genre *Caulerpa* [Jensen 1990]. Toutes ces caractéristiques physiologiques font de ces mollusques des agents potentiels de contrôle biologique de *Caulerpa taxifolia*.

Des résultats préliminaires acquis sur ces mollusques ont permis d'apprécier les taux de consommation de *Caulerpa taxifolia* ainsi que la capacité reproductive de cette espèce [Meinesz *et al.* 1996] [Thibaut *et al.* 1998]. Les ascoglosses tropicaux consomment plus de *Caulerpa taxifolia* que les espèces indigènes [Thibaut *et al.* 1998]. Par contre, il a été démontré que ces espèces tropicales étaient sensibles aux températures de la Méditerranée : *Elysia subornata* meurt à 15°C et ne se reproduit plus en dessous de 21°C [Meinesz *et al.* 1996]. Les cycles de vie d'*Elysia subornata* et d'*Oxynoe azuropunctata* sont les mieux connus et leurs élevages sont très bien maîtrisés. Ces espèces présentent un développement larvaire direct

ce qui, d'une part, facilite leur élevage et, d'autre part, permet d'espérer, dans l'objectif d'un contrôle de *Caulerpa taxifolia*, un rapide développement de leur population.



Figure 27. Photo d'un Ascoglosse *Elysia Subornata* présentant son meilleur profil (A. Meinesz)

Les risques de changement de nourriture ne peuvent être avancés à cause de la forte sténophagie de ces espèces tropicales, mais celles-ci seraient capables de se nourrir des autres *Caulerpa* indigènes de Méditerranée : *Caulerpa prolifera*, *Caulerpa mexicana*, *Caulerpa racemosa*, *Caulerpa ollivieri*, *Caulerpa scalpelliformis*.

90

En ce qui concerne la modélisation, nous avons pu développer lors des deux dernières années un modèle de contrôle de *Caulerpa taxifolia* par *Elysia subornata*, en collaboration avec le LEML de l'Université de Nice et le Laboratoire d'Écologie Végétale et Cellulaire [Hill *et al.* 2000a] [Coquillard *et al.* 2000]. Une taxonomie des techniques de modélisation réalisée dans le milieu des années quatre-vingt a introduit le terme de multimodèle [Ören 1984]. Ce terme a conduit à une technique dite de multi-modélisation qui est principalement dérivée de ce que Bernie Zeigler appelait les approches multiformalismes [Zeigler 1979]. Nous considérons un multi-modèle comme étant composé de différents sous-modèles homogènes ou hétérogènes qui ont pu être réalisés à différents niveaux d'abstraction (Figure 28). Cette approche facilite la construction de modèles hiérarchiques de systèmes réels qui ne peuvent pas être simulés aisément par des modèles simples [Fishwick 1993, 1995]. L'application des concepts orientés objets au domaine de la multi-modélisation a été discutée dans la présentation du langage « Omola language » [Matteson et Anderson 1993].

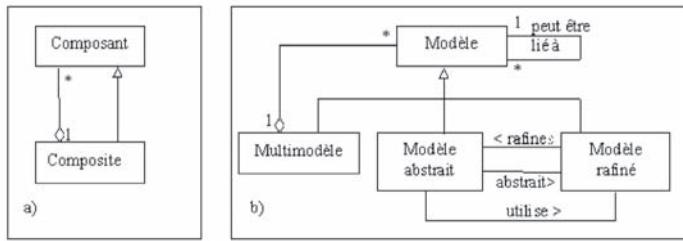


Figure 28. Un métamodèle représentant les multi-modèles avec un diagramme de classe UML

Le développement de multi-modèles peut conduire à une utilisation systématique de patrons d'analyse et de conception par objet [Hill *et al.* 2000a]. Celui qui a été développé pour simuler la lutte biologique fait usage de deux patrons connus : le patron « adapter » et le patron « composite » [Gamma *et al.* 1995]. Le modèle final comporte des équations déterministes, une matrice de Leslie pour la prise en compte du processus de vie et de mort, un automate cellulaire et un simulateur stochastique à événements discrets. La Figure 29 donne un extrait du diagramme de classe qui décrit les choix de conception statique. La Figure 30 donne un aperçu des écrans de contrôle de la simulation.

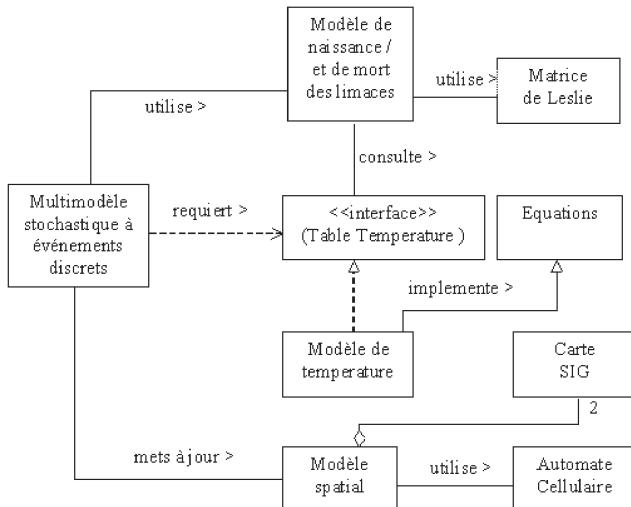


Figure 29. Extrait du diagramme de classe UML pour le modèle de lutte biologique

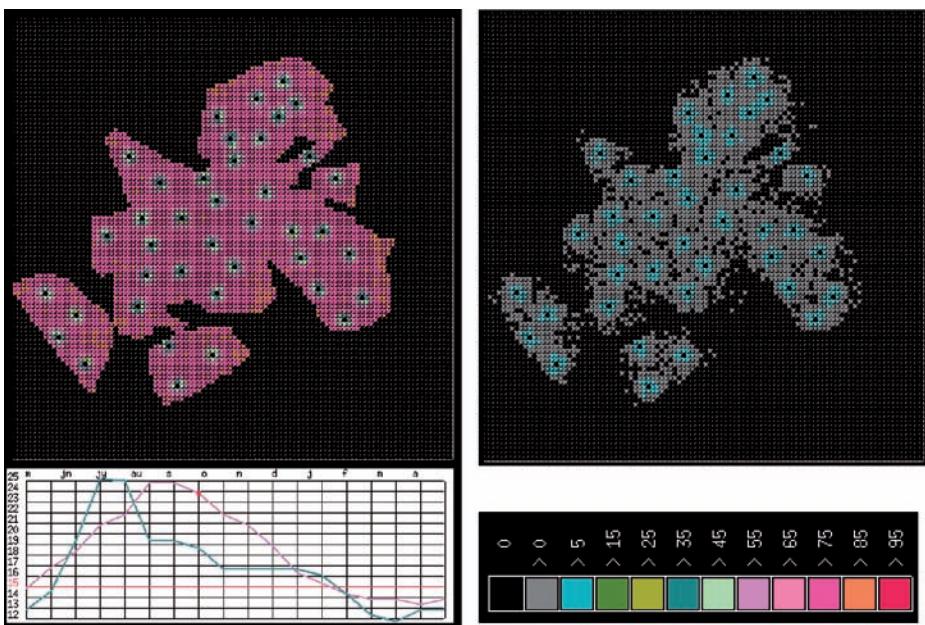


Figure 30. Écran de simulation sur le site de Le Brusc. La dispersion maximum des limaces est atteinte (à droite) et l'impact est visible sur la colonie entière de *Caulerpa taxifolia* (à gauche)

92

Les résultats publiés dans [Coquillard *et al.* 2000] et [Hill *et al.* 2000a] sont suffisamment significatifs pour débuter des expériences à grande échelle. Cependant, une réduction significative de la biomasse de la Caulerpe dépend fortement de la stratégie à retenir pour le lâcher d'individus sur une colonie de *Caulerpa taxifolia* (répartition et nombre des individus, période d'introduction, classe d'âge des limaces, quantité, etc). La température de l'eau est notamment un critère très sensible. Ces paramètres sont étudiés avec précaution afin d'obtenir des valeurs pertinentes.

4.6. Métamodélisation par un réseau de neurones

Une des solutions possibles pour mener une exploration en profondeur des paramètres influençant un système est de modéliser le simulateur afin d'approcher son comportement. Cette opération est appelée métamodélisation et le modèle obtenu est un métamodèle. Nous avons souhaité appliquer cette technique au simulateur de l'expansion de la Caulerpe.

Il existe plusieurs types d'approche pour la construction de métamodèle. Les premières approches étaient basées sur des modèles de régression [Kleijen 1979] [Huber et Szczerbicka 1994] [Pierreval 1992]. D'autres approches utilisent les réseaux de neurones et leurs différentes variantes [Berthold *et al.* 1993]. Ceux-ci sont en effet particulièrement efficaces pour approcher des fonctions arbitraires [Pierreval 1996]. Une fois entraînés à l'aide de données artificiellement générées par les simulations, les réseaux de neurones sont reconnus comme étant des approximations fidèles de l'écosystème d'origine, fournissant des résultats satisfaisants pour un temps de calcul nettement inférieur à celui du simulateur [Aussem et Hill 1999]. L'inconvénient majeur des réseaux de neurones que nous ne devons pas cacher restera qu'ils ne permettent pas une prédiction de la répartition spatiale de l'algue.

Un réseau de neurones multi-couches est un modèle à entrées-sorties, largement connu [Rumelhart *et al.* 1986], utilisant des fonctions de transfert non-linéaires et permettant d'approcher les fonctions universelles [Cybenko 1989]. Il est important de noter qu'un réseau de neurones est un modèle déterministe. Ceci signifie qu'à un ensemble de données en entrée, il fera correspondre une sortie unique (contrairement au simulateur stochastique). Cette approche possède plusieurs avantages. Tout d'abord, un modèle à base de réseaux de neurones peut manipuler une combinaison de variables continues et discrètes. De plus, l'entraînement d'un seul modèle neuronal permet de modéliser la totalité de l'espace des scénarios de simulation, contrairement à la métamodélisation par régression polynomiale, où la surface de régression est ajustée à une partie seulement des trajectoires possibles. Finalement, l'architecture parallèle des couches de neurones fournit une robustesse aux ensembles de données incomplets ou erronés et offre une tolérance aux erreurs, et ce en temps réel.

Par contre, un réseau de neurones de taille fixe ne peut approcher la fonction optimale qu'en introduisant un biais. Un réseau plus large (travaillant à partir de plus de variables) introduira un biais plus faible, mais nécessitera une quantité plus importante de données d'entraînement et aura une variance plus grande. Ce phénomène est typique de la théorie de la régression [Geman *et al.* 1992]. Les réseaux de neurones sont donc des outils très efficaces, mais aussi forts complexes, et nous invitons le lecteur à se reporter à la littérature pour de plus amples informations [Cellier 1991].

L'entraînement d'un réseau de neurones supervisé est effectué à partir d'un ensemble de données d'entrée et de sortie, appelé base d'apprentissage. L'apprentissage du réseau de neurones se fait en lui présentant les uns après les

autres les jeux de données constituées de variables d’entrée et de sortie leur correspondant. Le réseau compare ensuite son estimation des variables de sortie à leurs valeurs effectives. Il utilise ensuite cette comparaison pour modifier ses paramètres internes de façon à réduire l’écart avec la fonction modélisée.

Dans le cas de la modélisation d’un simulateur stochastique, la valeur obtenue à l’issue de la simulation varie d’une réPLICATION à l’autre. Pour obtenir une valeur unique, on effectue plusieurs réPLICATIONS avec le même jeu de paramètres et on calcule une moyenne. Une fois l’apprentissage effectué, on teste son efficacité sur une base de test. Cette base est constituée de données du même type que celle de la base d’apprentissage, mais ces données n’ont pas été présentées au réseau de neurones lors de son apprentissage.

Pour « métamodéliser » le simulateur de l’expansion de *Caulerpa taxifolia*, nous avons utilisé avec Alexandre Aussem un réseau de neurones traditionnel à plusieurs couches, avec un algorithme de rétro-propagation standard. Sur ce modèle basique ont été greffées deux techniques (très simples) utiles pour un bon fonctionnement des réSEAUX de neurones [Aussem et Hill 1999]. Étant donné le temps de calcul nécessaire à la simulation stochastique discrète l’accent a été mis sur des prédictions à long terme (12 ans) où le gain de temps serait le plus important. Le choix de la variable à prédire s’est porté sur la surface contaminée qui est considérée comme une bonne caractéristique quantitative de l’expansion [Meinesz *et al.* 1996]. L’objectif de la métamodélisation a donc été de prédire la surface contaminée 12 ans à l’avance sur un site entre Menton et Villefranche-sur-Mer. La Figure 31 présente un résultat de simulation après 8 ans de propagation de l’algue. Cette carte est en adéquation avec les relevés effectués en 1997.

L’observation du modèle de simulation fut nécessaire pour imaginer ce que la métamodélisation par un réseau de neurones pouvait apporter. Cette étude a montré que la courbe contenant le plus d’information était celle représentant la surface contaminée en fonction du temps. L’observation de cette courbe avec une échelle logarithmique fait apparaître deux choses :

- l’expansion de *Caulerpa taxifolia* suit des cycles annuels : forte croissance durant les mois d’été (chaud et lumineux) et période de stagnation durant l’hiver.

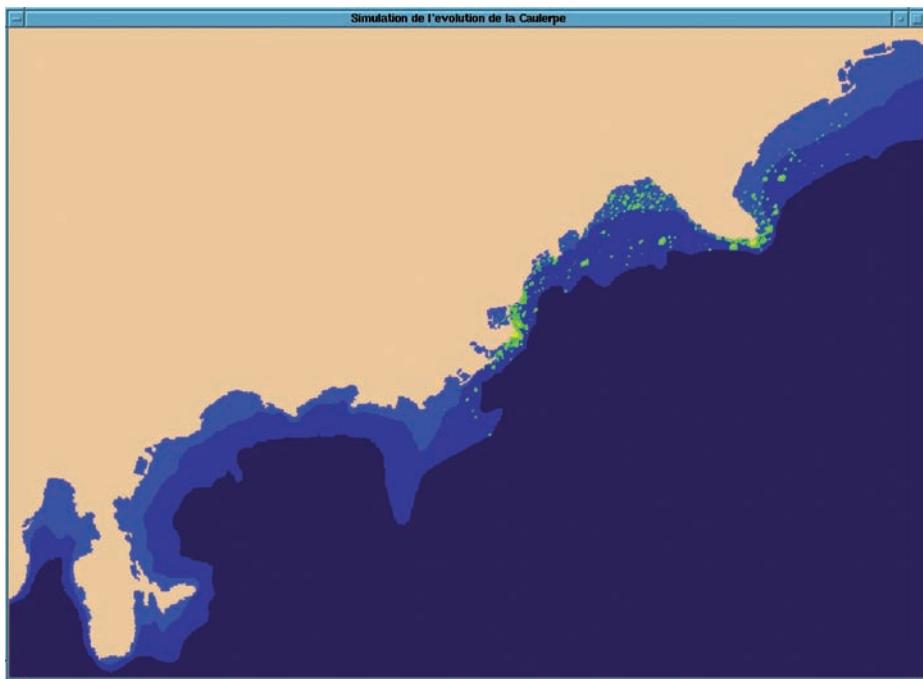


Figure 31. Résultat de 8 années de simulation entre Menton et Villefranche-sur-Mer

95

- la colonisation de *Caulerpa taxifolia* n'est pas régulière, mais s'effectue en trois temps :
 - i. une phase de croissance régulière,
 - ii. un temps de stagnation dû à la rencontre de caps et aux limites bathymétriques,
 - iii. une seconde phase d'expansion, moins forte que la première, où la Caulerpe suit la côte et se développe en eau profonde.

Une phase d'exploration a été entreprise pendant laquelle plusieurs paramètres ont été changés. Il en est ressorti que l'allure générale de la courbe ne changeait que très peu. En conséquence, il fut donc décidé que, plutôt que d'essayer de reconstruire la courbe à partir de la date initiale, le réseau de neurones prédirait la surface contaminée à la date finale.

Une fois l'objectif fixé, il a fallu sélectionner les variables pertinentes. En effet, une représentation pertinente de l'information à prédire peut produire des informations explicites, tout en écartant les données inutiles et encombrantes. Des représentations peuvent être équivalentes en termes de

résultats de sortie mais très différentes dans leur efficacité à les fournir. De plus, comme nous l'avons vu précédemment, l'augmentation du nombre de paramètres augmente le risque d'introduire un biais important. Après concertation avec les écologues, quatre des paramètres d'entrée du simulateur ont été retenus.

Après ces choix, l'étape suivante a été la construction d'une base d'apprentissage. Cette construction s'est heurtée à quelques problèmes. Les simulations d'une période aussi longue prenant beaucoup de temps, il est extrêmement long d'effectuer autant de réplications que souhaité. Or il est difficile d'obtenir une moyenne significative à partir de peu de termes. Finalement, l'apprentissage a pu être effectué à partir de cinq bases d'apprentissage, pour un total de 600 jeux de données. Les résultats obtenus ont été très prometteurs : la fonction de prédiction de la surface contaminée a été parfaitement apprise par un réseau de neurones de taille relativement petite (*cf. Figure 32*).

96

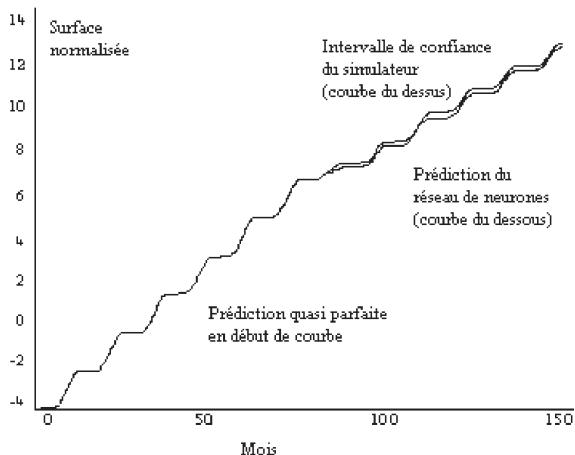


Figure 32. Courbe d'expansion de *Caulerpa taxifolia* en fonction du temps (exprimé en mois) ; la courbe issue du simulateur est la courbe du dessus, et la courbe issue du réseau de neurones est celle du dessous [Aussem et Hill 2000]

La reconstruction de la courbe représentant la surface contaminée en fonction du temps a été plus délicate. L'ajustement de la courbe prototype n'a pas donné les résultats escomptés. Mais en entraînant de nouveaux réseaux de neurones afin de prédire les coordonnées des points de cassure, des résultats encourageants ont été obtenus.

Cependant, cette approche possède quelques points faibles :

- Peu de réplications ont été effectuées pour chaque jeu de paramètres. On ne peut donc pas être certain que l'intervalle de confiance de la surface couverte prédict par le réseau de neurones est fiable.
- Cette expérience n'a pris en compte qu'un seul site. Les prédictions du réseau ne sont donc valables que pour ce site.
- Il n'est pas possible de faire des prévisions spatiales.

4.7. Conclusion

Tous nos travaux de recherche en océanographie ont été effectués jusqu'à ce jour dans le contexte méditerranéen et ont porté sur l'expansion de *Caulerpa taxifolia* et sur les problèmes qui en découlent. Depuis deux années, une autre Caulerpe envahissante est arrivée en France : *Caulerpa racemosa* [Verlaque *et al.* 2000]. Grâce au réseau d'observation de *Caulerpa taxifolia* et de nombreux relais d'observateurs mis en place, le LEML a recueilli et vérifié la présence de cette espèce à Marseille, Porquerolles, Menton et Villefranche-sur-Mer.

Le développement des différents modèles présentés est toujours en cours. Il devrait permettre d'apporter des réponses à des questions plus complexes portant notamment sur le fonctionnement simultané de plusieurs foyers de dispersion. Actuellement, le modèle est calibré pour travailler sur un seul site à la fois, et il ne permet pas de faire démarrer les divers foyers connus dans l'ordre chronologique de leur apparition. De même, la dispersion anthropique à longue distance n'est pas encore implémentée ; or ce processus est certainement la clé de voûte de l'expansion de *Caulerpa taxifolia* en Méditerranée. En effet, le modèle actuel confirme que la dispersion de proche en proche n'est pas suffisante pour expliquer la rapidité de dispersion de cette algue, et de fait, l'étendue de la zone envahie. La flottabilité négative de l'algue et sa présence dans de nombreux ports réfute l'hypothèse d'une propagation exclusive par les courants marins. Par la suite, afin de prendre en compte les interventions humaines, nous envisageons d'indiquer, grâce au

Système d’Information Géographique mis en œuvre par le LEML, les zones de plaisance et de pêche, ainsi que les zones et les dates d’éradiation. Le recours à des techniques de simulation distribuée [Hill 1997c] pour répartir le calcul et les cartes sur plusieurs calculateurs puissants n’est pas simple à mettre en œuvre. Un des verrous technologiques reste la distribution de nombres pseudo-aléatoires, cependant cela a permis de mettre en évidence les mécanismes de la dispersion anthropique.

Malgré les limites précédemment évoquées et les difficultés liées à la validation de tels modèles, le caractère « prédictif » du modèle a été vérifié sur des sites précis à plusieurs reprises, attestant ainsi de son réalisme. Nous avons également présenté des techniques de validation de ces simulations stochastiques en utilisant des analyses spectrales, qui se révèlent longues et coûteuses en temps de calcul. Les résultats préliminaires obtenus démontrent cependant l’intérêt de cette technique de modélisation par simulation aléatoire à événements discrets sous contraintes spatiales.

Des espèces indigènes de limaces, *Oxynoe olivacea* et *Lobiger serradifalci*, habituellement associées à *Caulerpa prolifera* [Murillo *et al.* 1986], se rencontrent maintenant dans les prairies de *Caulerpa taxifolia*. Elles présentent un développement larvaire planctonique (larve planctonophage) qui en font des animaux rares aussi bien dans les prairies de *Caulerpa taxifolia* que dans les prairies de *Caulerpa prolifera* [Thibaut et Meinesz 2000]. Le LEML a pu étudier (entre avril et juillet 1999) la dynamique de population d’*Oxynoe olivacea* et *Lobiger serradifalci* dans le port de la Darse (Villefranche-sur-Mer) depuis l’apparition de milliers d’individus jusqu’à leur extinction. Le LEML a profité de cette abondance de mollusques pour effectuer des lâchers de populations d’*Oxynoe olivacea* et *Lobiger serradifalci* sur des colonies isolées de *Caulerpa taxifolia*. Les résultats de ces expériences préliminaires montrent que les mollusques sont capables de diminuer significativement l’abondance de *Caulerpa taxifolia* (données non publiées). Des essais similaires ont eu lieu en Croatie (Zuljevic) avec des individus maintenus dans des cages. Là encore, la présence d’ascoglosses a entraîné une diminution significative de la densité de *Caulerpa taxifolia* sous les enclos. Compte tenu des premiers résultats concernant *Oxynoe olivacea* et *Lobiger serradifalci*, le moyen le plus efficace d’utiliser ces mollusques indigènes consisterait à les éléver artificiellement afin d’augmenter leur population et à en lâcher un grand nombre sur les prairies de *Caulerpa taxifolia*. Actuellement, l’élevage n’est pas maîtrisé en partie car la durée et le mode de nutrition des larves sont encore inconnus. De même l’impact des incisions de limaces sur le

thalle des Caulerpes doit être étudié avec plus de précision (existe-t-il des parties épargnées, des risques de fragmentation accélérée des frondes,... ?). Il conviendrait également de tester et d'évaluer les meilleures espèces d'ascoglosses méditerranéennes et tropicales, ainsi que les poissons *Siganidae* pour lutter contre *Caulerpa racemosa*.

5. LES SIMULATIONS MULTI-AGENTS POUR L'ÉTHOLOGIE

5.1. Introduction

Nous abordons maintenant des problèmes en éthologie. Dans ce contexte il nous semble important d'intégrer aux simulations des techniques d'Intelligence Artificielle Distribuée (IAD). Un système multi-agents (SMA) est système perçu comme un ensemble d'agents en interaction [Jennings *et al.* 1998]. Les types d'interactions incluent : la coopération, la coordination, la négociation et la résolution de conflits entre les agents. Avec un système multi-agents dérivé des modèles individus-centrés [Huston *et al.* 1988] [De Angelis et Gross 1992] [Breckling et Müller 1994], il est possible de représenter des phénomènes environnementaux comme la conséquence d'interactions entre des agents agissant en parallèle, chaque agent étant un objet actif, autonome et intégrant un comportement social.

Il existe de nombreuses variétés d'agents, nous en présentons quelques-unes en reprenant essentiellement les travaux et les définitions de [Ferber 1995]. Un agent dit *cognitif* dispose d'une représentation symbolique et explicite du monde où il agit et, parfois, de ses actions passées. De ce fait, cet agent possède la capacité de planifier les actions pour l'accomplissement de ses tâches. Lorsqu'il ne possède pas un but explicite, c'est-à-dire, lorsqu'il agit seulement quand d'autres agents lui font une demande, on parlera d'*Agent Module*. Un agent dit *réactif* est un agent dont les actions sont réalisées à travers des mécanismes de réaction à des stimuli (internes ou externes). Il ne dispose donc pas d'une capacité de planification des actions. Lorsqu'il dispose d'un mécanisme de motivation le poussant à accomplir une tâche, on parlera d'*Agent Pulsionnel*. Lorsqu'il dispose d'une mémoire pour accomplir son but, on l'appellera *Agent Hystérétique*, contrairement aux *Agents Tropiques* dont les actions sont le fruit exclusif de leur perception de l'environnement. Des agents qui échangent directement entre eux des messages

sont des agents communiquants. Lorsque l'environnement n'est pas utilisé comme support de la communication et que les agents ne font que communiquer par messages, on parle des *Systèmes Multi-Agents Purement Communicants*. Un agent qui se trouve positionné sur un environnement (élément du SMA) et qui en a une vision locale est dit être un *agent situé* [Drogoul 2000]. On parlera de *Système Multi-Agent Purement Situé* pour désigner le SMA dans lequel les agents ne communiquent pas par envois de messages, mais seulement par propagation de signaux sur l'environnement [Ferber 1995]. Les agents émotionnels produisent une impression de vie propre à travers la présence de comportements « émotionnels » (surprise, joie, etc.) [Bates 1994]. Ils sont utilisés surtout dans les animations, les jeux et dans les systèmes de réalité virtuelle.

Les *agents mobiles* sont des objets qui possèdent leur propre code et qui peuvent être transmis dans la communication des participants d'un système distribué [Knabe 1996]. Ils sont largement utilisés dans les applications sur Internet. Ils gardent des concepts d'autonomie, mais n'ont pas forcément les mêmes caractéristiques que les agents originaires de l'IAD. En ce qui concerne nos applications à l'éthologie, nous pensons que c'est la prise en compte des comportements sociaux qui justifie l'utilisation d'agents et il ne s'agit pas d'adhérer à un effet de mode.

100

Pour Guessoum, un *agent hybride* est un agent qui possède des capacités de raisonnement individuelles, comme les agents cognitifs, mais qui intègrent également des capacités réactives pour pouvoir réagir « au plus vite » dans certains cas [Guessoum 1996]. Une distinction est proposée par Nwana qui considère un agent hybride comme un agent dont la constitution est une combinaison de deux ou plus philosophies d'agents différentes [Nwana 1996]. Le lecteur intéressé par ces techniques pourra se reporter aux travaux et ouvrages suivants : [Gasser et Briot 1992] [Drogoul 1993] [Ferber 1995] [Bousquet *et al.* 1996] [Doram 1997] [Uhrmacher 1997] [Campos et Hill 1998b].

Toutes les études décrites par la suite sont pluridisciplinaires, elles ont été réalisées en collaboration avec des laboratoires d'éthologie, d'environnement et même de psychologie. Les agents que nous avons retenus sont hybrides (selon la terminologie préconisée par Guessoum) et nous permettent d'étudier les collaborations de groupes d'animaux qui s'organisent collectivement pour assouvir certains besoins. Le fait que les SMA prennent comme référence les interactions sociales élémentaires favorise l'émergence d'organisations complexes, telles que celles observées par des éthologues.

Nous avons travaillé principalement dans deux directions : la première concerne des simulations « classiques » en éthologie animale et la seconde traite des SMA sur le Web ; dans cette deuxième voie nous avons utilisé des agents émotionnels. Voici les quelques précisions sur les travaux entrepris dans ces deux directions :

1. Des simulations multi-agents ont été développées dans le cadre d'un Groupement d'Intérêt Scientifique se préoccupant de l'entretien du paysage en moyenne montagne. Nous avons pu aborder la simulation de vaches et de chevaux au pâturage sur un site de plusieurs dizaines d'hectares dans le Massif central [Michelin *et al.* 1995] [Michelin *et al.* 1998]. Ces travaux ont pu aboutir à un simulateur complet, interfacé avec des bases de données géographiques et avec des relevés de positionnement GPS. Cette étude est présentée plus loin et le lecteur intéressé par plus de détails se reportera à [Mechoud *et al.* 1998] et à [Hill *et al.* 2000b]. Une autre application concrète que nous avons réalisée en collaboration avec l'INRA en 1998, étudiait les mammites des vaches laitières [Perrochon *et al.* 1998].
2. Des simulations multi-agents sur le Web nous ont permis d'étudier des agents primitifs avec des prototypes de sociétés virtuelles (notamment de poissons et de fourmis). Les particularités du développement de simulations sur le Web sont détaillées dans [Campos et Hill 1998b], [Campos *et al.* 2000]. Parmi les autres applications de simulation sur le Web qui impliquent des agents, nous avons abordé avec l'INRA la simulation de moutons au pâturage afin d'étudier leur mémoire spatiale [Hill *et al.* 1998a] [Dumont et Hill 2001].

Je reviendrai plus précisément sur les aspects liés aux simulations sur le Web dans un paragraphe dédié à cette technique. Dans les paragraphes qui suivent je n'aborderai que deux applications : la modélisation de moutons et de leur mémoire spatiale, puis la modélisation d'une estive pâturée par des vaches et des chevaux.

5.2. Étude de la mémoire des moutons

La capacité des animaux à apprendre et à retenir l'emplacement de placettes d'herbe préférées au pâturage détermine la fréquence avec laquelle celles-ci seront atteintes. Comprendre, puis prédire, l'utilisation de différents types de végétations par les herbivores nécessite donc de mieux appréhender leur mémoire spatiale. Une expérience a été réalisée à partir de l'été 1996 pour

mettre en évidence la mémoire spatiale de groupes de moutons au pâturage. Parallèlement, j’ai cherché à simuler le comportement des animaux à la recherche de sites attractifs dans une prairie. J’ai pu développer un modèle multi-agents stochastique dans lequel il fallait caractériser les règles de déplacement des animaux et leur mémoire spatiale (nombre d’emplacements qu’un animal peut mémoriser, probabilité de se souvenir d’un emplacement déjà visité). Pour ce travail de recherche réalisé en collaboration avec Bertrand Dumont chercheur à l’INRA, j’ai proposé un noyau de synchronisation dédié pour éviter les problèmes de compétition spatiale entre processus simultanés ; nous avons également développé des outils statistiques d’analyse de données dédiés à la validation [Hill *et al.* 1998a] [Dumont et Hill 2001].

Description de l’expérience

Les observations ont été réalisées sur des groupes de trois agnelles dans des parcelles d’herbe rase où étaient cachés 136 bols contenant chacun cinq grammes de concentré (un aliment très attractif pour les animaux). Deux parcelles ont été utilisées. L’une de 80 x 80 m, et l’autre de 160 x 160 m. Les parcelles sont partagées en 16 zones imaginaires (4 x 4). Quatre zones contenaient 25 bols, quatre autres en contenaient 9, et les autres zones ne contenaient aucun bol.

102

Les agnelles étaient introduites dans la parcelle, pour un test d'une demi-heure chaque jour pendant 12 jours successifs. En tout, 4 groupes de 3 agnelles, observés au cours de 3 périodes successives, ont permis d'obtenir 6 courbes synthétiques des évolutions (12 jours de mesures). Entre chaque période l'emplacement des bols était modifié. L'évolution du comportement des agnelles au cours de l'expérience devait permettre de mettre en évidence l'existence de leur mémoire spatiale.

Deux observateurs étaient nécessaires pour réaliser les mesures : l'un suivait une agnelle pour noter l'enchaînement de bols qu'elle visitait, et notait si ceux-ci étaient pleins ou vides. L'autre notait l'activité de chaque agnelle toutes les 30 secondes. Un animal pouvait soit se déplacer tête haute, soit manger dans un bol, soit pâtrurer, soit être inactif. Les animaux restaient 30 minutes dans la parcelle. Une fois ceux-ci sortis, l'un des deux observateurs faisait le tour de la parcelle pour noter les bols visités par le groupe. Les bols étaient ensuite à nouveau remplis pour le groupe suivant.

Les résultats des expériences

Le nombre de bols que visite le groupe augmente au cours des 4-5 premières séances, puis se stabilise. Les animaux sont alors de plus en plus rapides pour trouver un même nombre de bols. Lors des dernières séances, il leur faut en moyenne 12' dans la petite parcelle au lieu de 15'30" le jour 5, et 10'45" dans la grande parcelle au lieu de 16' le jour 6. Ces résultats démontrent clairement la capacité d'un groupe de moutons à mémoriser l'emplacement de sites de pâturage préférés. Quand le nombre de bols visités est devenu stable, les animaux en visitent moins dans la grande que dans la petite parcelle (en moyenne 83 au lieu de 126). Ils explorent aussi plus complètement les zones où la densité des bols est la plus importante, et cela surtout dans la grande parcelle.

Présentation sommaire du modèle

Le modèle d'évolution des agnelles à la recherche d'aliments est un modèle stochastique implémenté par simulation à événements discrets. Le modèle d'une agnelle est configuré par plus de 40 paramètres (**Figure 33**). Certains ont été évalués par l'observation des agnelles, de façon plus ou moins précise, tandis que d'autres, plus difficiles à estimer, nécessitent de nombreux calibrages.

103

Le modèle de comportement d'une agnelle peut être représenté par un automate d'états finis à 5 états. Le comportement dans les différents états est fonction de divers facteurs comme, par exemple, les intentions de l'agnelle (sa motivation pour s'alimenter ou pour se rapprocher de ses congénères), ou le type de zone dans lequel elle se trouve. Les conditions de transitions entre états peuvent être assez complexes et font intervenir de nombreuses règles. Certains états sont par exemple inaccessibles tant que l'agnelle n'a pas atteint un certain taux de rassasiement. Le pas de temps de la simulation est d'une seconde. Le noyau de synchronisation de la simulation à événements discrets utilise une approche par activités, duale de l'approche par événements. Les agnelles possèdent toutes un automate avec différents états et sous-états possibles. On peut rapidement évoquer les états suivants :

- L'agnelle se déplace la tête haute. Elle est soit en phase de recherche de bols dans une zone (déplacements sinueux), soit en déplacement entre zones [Berg 1993]. Cet état est l'état d'exploration. Dans aucun cas elle n'utilise sa mémoire.
- L'agnelle utilise sa mémoire pour se diriger aux environs d'un bol qu'elle a déjà visité, et dont elle a le souvenir.

- L’agnelle a trouvé un bol. Soit elle mange, soit il est vide et elle est en train de l’inspecter avant de le quitter.
- L’agnelle pâture. Les agnelles ont un comportement fortement grégaire dans cet état, et seront automatiquement attirées par une autre agnelle si elle s’en trouve trop éloignée. Une agnelle qui se dirige vers une autre le fait à grande vitesse tout en restant dans l’état « pâturage » au niveau de l’application.
- L’agnelle est inactive. Elle ne se déplace pas et ne pâture pas. Cela correspond aussi aux périodes d’interactions entre animaux.

104

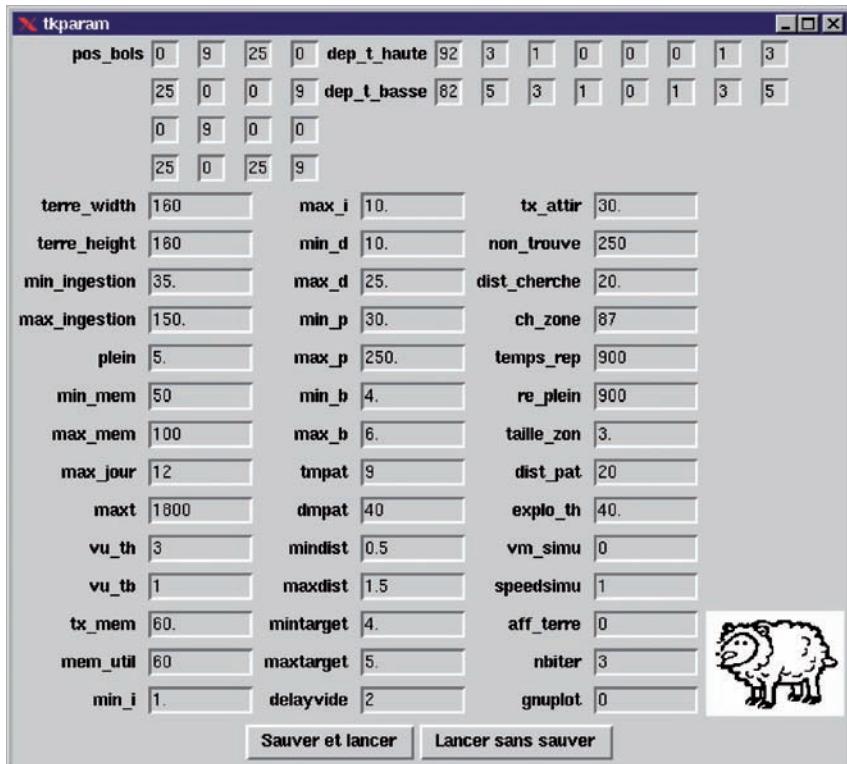
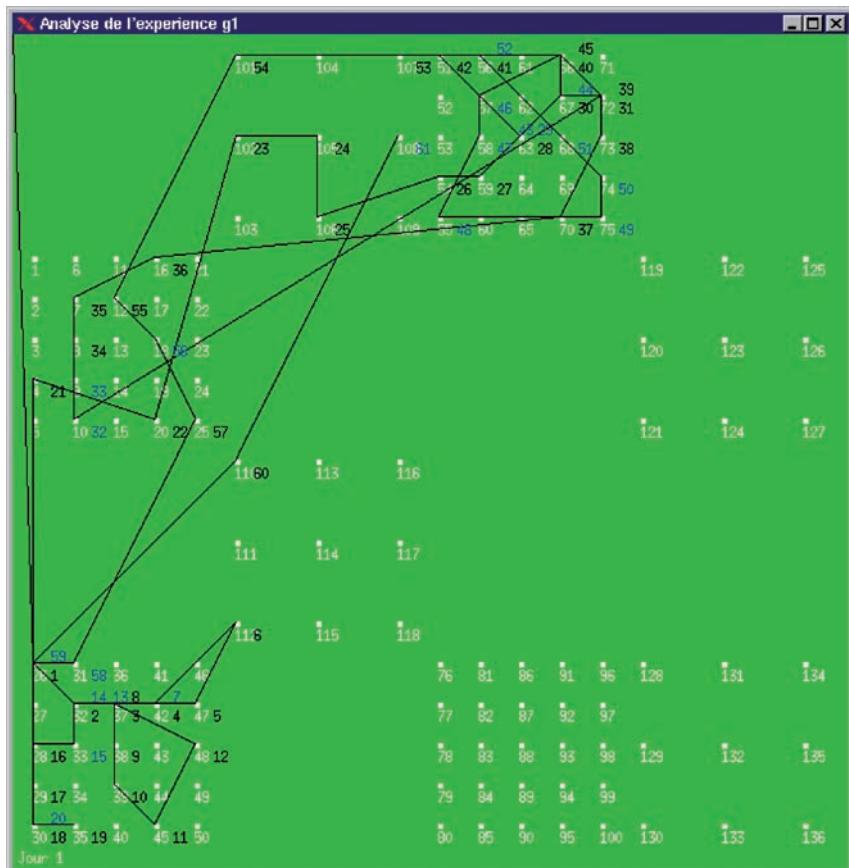


Figure 33. Interface graphique développée avec Tcl/Tk pour paramétrier le logiciel de simulation

En ce qui concerne la validation des résultats de simulation par comparaison avec le système réel, le premier programme réalisé avait pour but de vérifier que la succession des bols visités par simulation était réaliste. Un uti-

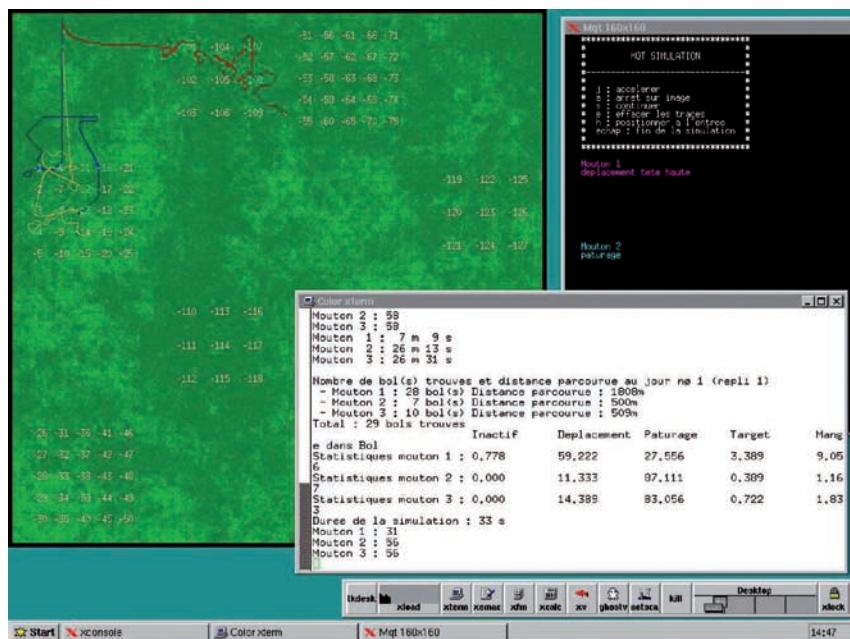
litaire graphique, réalisé en langage C++, dessine le terrain dans une fenêtre X-Windows, puis trace des segments noirs pour schématiser la succession des bols visités. À côté de chaque bol, le programme inscrit un numéro indiquant l'ordre de passage dans le bol au sein du parcours effectué (1 pour le premier bol trouvé, 2, 3,...). Ce numéro est affiché en noir pour les bols trouvés pleins, et en bleu pour les bols trouvés vides. De plus, l'application affiche des données destinées à un traitement statistique. Pour chaque visite dans une zone, le programme affiche le numéro de la zone, suivi du nombre de bols trouvés dans la zone (**Figure 34**).



105

Figure 34. Exemple de tracé fourni par l'application « viewbol ». Ici le tracé est celui de l'agnelle numérotée « 1 » dans une grande parcelle, lors du premier jour de la simulation

Un programme « statzon » a également été réalisé pour la validation de ce modèle. Ce programme calcule pour chaque type de zone (de 9 ou de 25 bols, *cf. Figure 35*) la moyenne et l’écart-type du nombre de bols trouvés au cours du passage dans la zone. Il donne également, pour chaque type de zone (zones de 9 bols et zones de 25 bols) la proportion de bols trouvés par rapport au nombre total de bols dans tous les types de zones (soit 36 bols pour les 4 zones de 9 bols, et 100 bols pour les 4 zones de 25 bols).



106

Figure 35. Une capture de l’environnement de travail. Sur la fenêtre d’animation, on peut observer les déplacements sinueux dans les zones où il y a des bols

Les résultats de ce travail ont permis une meilleure compréhension du fonctionnement de la mémoire des agnelles. Le modèle réalisé par prototypage a entraîné un certain nombre de choix. Pour rendre plus réaliste le système de recherche de bols dans les zones, nous avons ajouté un système de déplacements sinueux dans le but de reproduire le comportement des agnelles. Ceci est un sous-état de l’état « tête haute » qui se manifeste quand une agnelle pense être à proximité de bols. Parmi les autres modifications notons que le comportement gréginaire a été retouché de façon à ce qu’une agnelle essaye

toujours de se rapprocher d'une autre qui a les mêmes motivations. De plus, dès qu'une agnelle se trouve isolée elle est automatiquement attirée par une autre. Ceci permet entre autres de réunir les trois agnelles en fin d'expérience lorsqu'elles pâturent toutes.

Citons un autre problème du premier prototype réalisé. Le premier jour de l'expérience, les agnelles se comportaient comme si elles savaient que la parcelle contenait des bols. Il a donc fallu prendre en compte le cas particulier du premier jour, où l'activité se partage entre exploration tête haute et pâturage tant que les agnelles n'ont pas encore vu de bols. L'ajout dans le modèle du fait qu'une agnelle qui n'a pas trouvé de bols pleins pendant un certain temps puisse commencer à pâtrer même si son seuil en concentré n'est pas atteint, permet également de rendre le modèle plus réaliste. En effet celui-ci traduit le fait que les agnelles se lassent de rechercher des bols, et évite que les agnelles du modèle passent toute la durée de l'expérience à chercher sans jamais pâtrer.

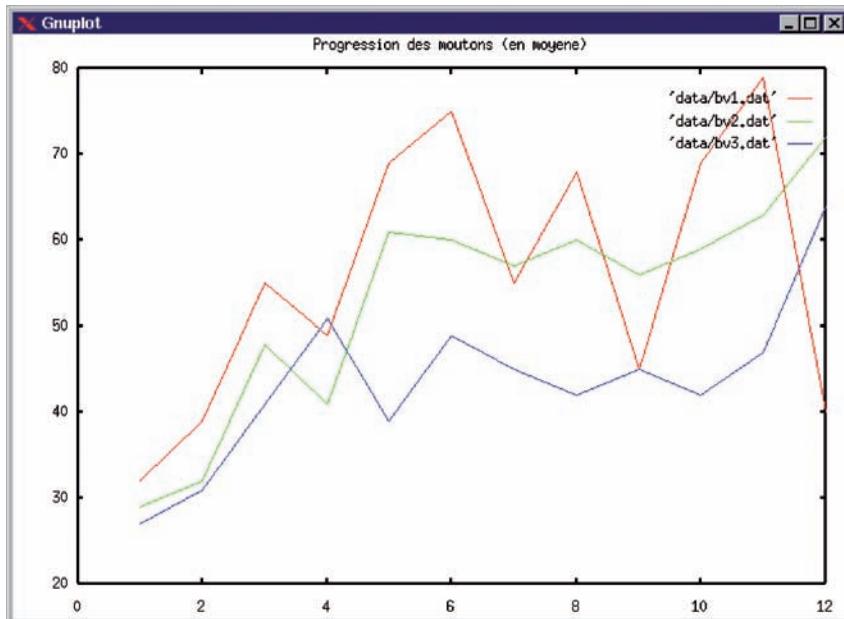


Figure 36. Courbes représentant le nombre de bols trouvés en moyenne par les 3 agnelles (moyennes sur 3 réplications dans la grande parcelle)

Conclusion

Visuellement, les déplacements des agnelles dans la parcelle sont réalistes. Cependant, les données produites par simulation ne concordent pas toutes avec les données réelles. Les résultats de simulation concernant le nombre de bols trouvés sont proches de la réalité. En revanche, l’activité simulée des agnelles est un peu trop dirigée vers les bols. Le fait que les courbes moyennes de progression individuelle soient satisfaisantes, alors que le groupe d’agnelles ne trouve pas suffisamment de bols, pourrait être dû à une attirance excessive entre les agnelles qui trouvent toutes les mêmes bols. Une évolution possible serait justement d'affiner le comportement grégaire des agnelles afin d'améliorer encore les résultats des simulations.

5.3. L’entretien des paysages par des herbivores

Le contexte

Dans le cadre d'un Groupement d'Intérêt Scientifique (GIS), nous avons envisagé l'étude de l'entretien par des herbivores d'espaces sensibles, dans des zones de moyenne altitude à dynamique végétale de reconquête forestière rapide. Ce GIS s'est formé autour du thème « Gestion de la végétation et entretien des milieux par les herbivores domestiques en moyenne montagne : approche expérimentale et modélisation ». Participant à ce GIS des laboratoires de l'INRA, notre Laboratoire inter-universitaire d'Informatique (LIMOS), l'École Nationale d'Ingénieurs en Techniques Agronomiques (ENITA), le Laboratoire d'Écologie Végétale et Cellulaire de l'Université d'Auvergne, et le Parc Régional des Volcans d'Auvergne.

108

Un cas particulier concret a été traité : celui du site de Ternant dans le Puy-de-Dôme. La gestion de ces espaces constitue un enjeu paysager et écologique majeur pour les années à venir. Or, on assiste de plus en plus à un délaissage marqué des pâturages d'altitude de moyenne montagne par l'agriculture et l'élevage. Les éleveurs privilègient aujourd'hui des systèmes d'élevage plus faciles à contrôler s'appuyant sur des terres plus faciles à exploiter au détriment des systèmes fourragers utilisant les prairies d'altitude. Dans ces parcelles, l'abandon de l'utilisation des ressources herbagères pour le pâturage d'herbivores est à l'origine d'un appauvrissement de la diversité écologique. La végétation est progressivement envahie par des espèces indésirables et ligneuses (genêts, callune, ronces, etc.). Ces types de végétation sont, de plus, de faible qualité fourragère et sont souvent plus pauvres sur le plan écologique que les pelouses d'altitude. L'installation pro-

gressive de ces landes hautes rend les parcelles difficiles à pénétrer et très sensibles au feu. L'état final d'enforestation conduit à la fermeture du milieu et à une diminution de la biodiversité. À terme, cette évolution se traduit par le déplacement des activités humaines hors de ces zones et à un déséquilibre au niveau du territoire national.

La gestion de la végétation et des paysages par les herbivores, et leur impact sur l'entretien des milieux ont déjà été étudiés pour certains milieux (pâturages d'altitude non colonisés par les ligneux hauts, landes pures). La complexité de ces systèmes écologiques est telle qu'elle nécessite une approche pluridisciplinaire. Le projet de recherche que nous présentons ici devait permettre :

1. de construire un modèle de connaissance de l'environnement étudié grâce aux données issues d'expérimentations pluridisciplinaires,
2. de mettre en œuvre un programme de simulation qui se déduit du modèle de connaissance précédent, afin que d'éventuelles prévisions puissent servir à l'optimisation de l'utilisation des ressources naturelles,
3. de pouvoir visualiser de manière très réaliste les résultats obtenus afin de pouvoir observer les impacts sur le paysage.

109

Le système étudié

L'expérimentation se déroule sur le site de Ternant, à proximité du puy de Dôme. Par son altitude moyenne (environ 1 000 m), la nature de ses sols et la diversité de sa végétation, elle est représentative de la plupart des situations rencontrées en moyenne montagne humide. Cette ancienne estive ovine domine la ville de Clermont-Ferrand. Elle est gérée par un regroupement d'estives et accueille des génisses durant tout l'été. Pour tester l'hypothèse de complémentarité de l'action animale, l'estive a été divisée en deux (2 x 25 ha) : une moitié accueille un troupeau mono-spécifique constitué uniquement de bovins, et l'autre moitié un troupeau mixte constitué de bovins et de chevaux.

Pour suivre et enregistrer les localisations des animaux sur le terrain, certains d'entre eux sont équipés d'un récepteur-satellite GPS (Global Positioning System, **Figure 37**). Le récepteur enregistre la position de l'animal toutes les 5 secondes. Pour déterminer l'activité de pâturage des animaux, ceux-ci sont également équipés de colliers ETHOSYS [Micol 1997] qui enregistrent toutes les 5 minutes le nombre de secondes où l'animal est actif, et en par-

ticulier son activité de pâturage. Les animaux ainsi équipés sont peu nombreux (3 animaux par parcelle), et changent chaque jour. L'intérêt majeur des techniques présentées consiste à rapprocher les données de localisation et les données d'activité afin d'essayer de déterminer les endroits où les animaux s'alimentent, ainsi que les temps passés aux différentes activités (repos, déplacement, etc.). Les données récoltées sont stockées à l'INRA dans une base de données ACCESS. Un premier travail de filtrage des données est alors nécessaire, les balises GPS pouvant parfois envoyer des localisations « fantaisistes ». Les données de 1996 et 1997 ont un volume de plus de 300 mégaoctets.



110

Figure 37. Garance et Marguerite, équipés de leur balise GPS et de leur collier ETHOSYS

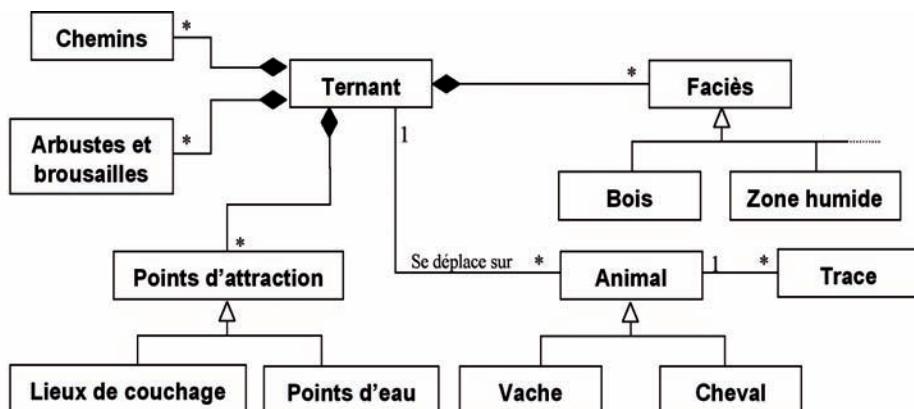
Objectifs de l'étude

L'objectif de la modélisation que nous présentons était de simuler le fonctionnement de l'estive, afin de comprendre l'interaction entre l'évolution de la végétation, et les déplacements et actions des animaux en pâture. Cette compréhension permettra ensuite de connaître les possibilités d'évolution de la végétation sur cette estive. Cela implique les sous-objectifs suivants :

1. comprendre les motivations des animaux (endroits pâturés, trajets, lieux de repos, etc.),
2. comprendre les influences de comportement d'une espèce animale sur une autre,
3. mesurer les variations d'embroussaillement de la parcelle en fonction des taux de chargement,
4. comprendre l'évolution des jeunes repousses d'arbres,
5. mesurer l'offre fourragère (qualité et quantité d'herbe).

Présentation du modèle de simulation

Le modèle que nous proposons a été formalisé avec UML et le logiciel est implémenté en Java. Le schéma de la Figure 38, volontairement épuré, présente les principales classes du système étudié à l'exception des classes d'objets techniques. On distingue déjà le terrain de Ternant, constitué de différents faciès de végétation, de chemins et d'arbustes et broussailles (genêts, ronces, etc.). Ce terrain comporte des points d'attraction que sont les lieux de couchage des animaux et les abreuvoirs [Bailey *et al.* 1998]. Ces points sont fixes, ce qui explique qu'ils ne dépendent pas des animaux. On trouve ensuite des animaux (vaches ou chevaux) qui pâturent sur ce terrain, se servant de ces différents composants (chemins, faciès, etc.) comme de points de repère.



111

Figure 38. Diagramme des classes UML du modèle conceptuel de Ternant

Pour répondre aux différents besoins exprimés, le modèle doit pouvoir fonctionner à différentes échelles :

- 3 pas de temps (instantané, intra-annuel, inter-annuel),
- 3 échelles d'espaces (cellule, faciès, assemblage de faciès),
- 3 niveaux d'individus (animal, groupe d'animaux d'une espèce, troupeau).

Pour respecter ces différentes échelles, le logiciel de simulation fonctionne suivant 2 modes différents. Même si chaque mode utilise un ensemble de données et des paramètres communs, ils se spécialisent en utilisant des paramètres et des fonctionnements spécifiques et en proposant des résultats propres. Ces modes sont les suivants :

1^{er} mode Simulation par trace : étude des trajets GPS et des activités ETHOSYS [Micol 1997] à des fins de compréhension et d’analyse. Ce mode permet de déterminer les lieux de passages des animaux, les lieux de pâture favoris, l’enchaînement et les temps consacrés aux différentes activités. Il permet donc d'affiner les règles comportementales utilisées dans le 2^{ème} mode de fonctionnement.

2^{ème} mode Simuler le comportement d'animaux isolés ou en groupe. Ce mode fournit les données comportementales des troupeaux et les interactions entre les groupes d'animaux.

Analyses des données GPS et ETHOSYS (1^{er} mode)

112

Le but de ce premier module du simulateur est d’analyser et de comprendre les données relevées sur le terrain, afin de pouvoir alimenter les autres modes de fonctionnement avec des données de comportement et de déplacement cohérentes.

Dans ce premier mode, l’utilisateur peut « relire » l’intégralité des trajets d’animaux simulés ou relevés par GPS et les dérouler de façon graphique en les superposant sur des cartes de représentation du terrain d’étude. Les cartes peuvent représenter la végétation du terrain, sa topographie ou seulement les contours et chemins.

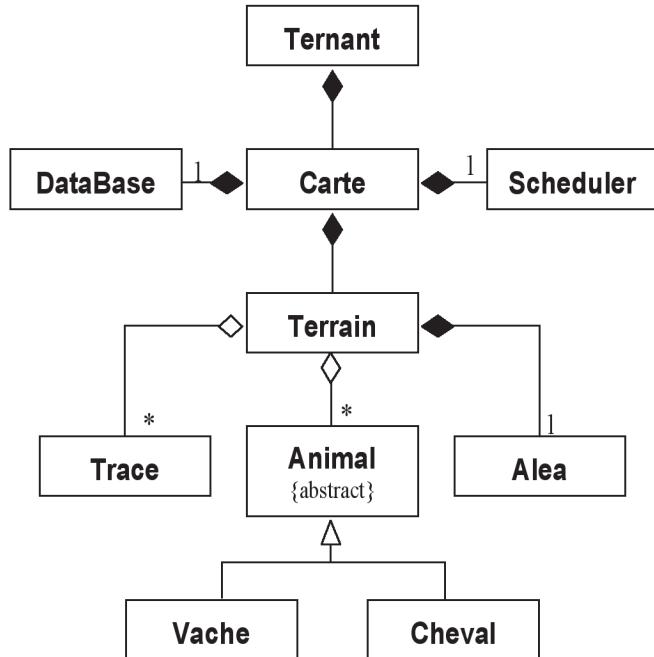
Cette relecture des trajets permet, dans une certaine mesure, de valider le modèle : en effet, elle permet de mieux comprendre le comportement des animaux et de connaître les lieux qu’ils fréquentent. La technique repose alors simplement sur ce que l’on appelle, en modélisation et en physique quantique, la simulation à base de traces. Le simulateur se comporte alors comme un magnétophone qui permet à l’utilisateur de visualiser l’évolution d’un trajet, de l’accélérer, de le ralentir et de le reproduire autant que nécessaire.

L'utilisateur peut également demander la superposition de l'ensemble des traces pour établir une carte de fréquentation du site. Cette superposition peut se faire avec des critères plus restrictifs pour obtenir des cartes plus spécialisées (restriction sur la période, l'espèce étudiée,...). Le programme d'analyse permet également de différencier les traces de simples déplacements de celles de pâturages, afin d'établir une carte de consommation. Ces cartes obtenues permettent en outre de mettre en évidence :

- les trajets effectués et les points de passage fréquents,
- les zones de fort pâturage,
- les zones évitées,
- les zones d'attraction (lieux de couchage, abreuvoirs, etc.),
- les faciès de végétation préférés, en fonction de la période de l'année et de l'espèce étudiée (vache, cheval ou vache « influencée » par la présence de chevaux).

Application de l'approche multi-agents (2^{ème} mode)

Dans ce mode chaque agent simulera le comportement d'une vache ou d'un cheval. Le nombre de ces agents est paramétrable, et, à partir de 2 agents, le programme simule le comportement grégaire des animaux. Chaque agent stimule les autres et interagit avec eux. Les agents agissent également sur le terrain sur lequel ils évoluent, broutant et piétinant la végétation. De ce fait, déplacer des entités virtuelles, changer leur comportement et modifier leur état peut être facilement effectué, toutes les données de simulation et les règles de comportement étant stockées dans des entités séparées. En plus des propriétés classiques des agents, telles que l'autonomie, le comportement social et la réactivité, certaines recherches considèrent qu'un agent doit être implémenté en utilisant une caractéristique propre à un être vivant : l'incertitude de leur comportement et de leur perception [Bates 1994] [Shoham 1993]. Dans notre cas, nous utilisons la logique floue suivant la technique décrite dans [Campos et Hill 1998a].



114

Figure 39. Diagramme UML des classes implémentées dans le modèle

Le diagramme de classes ci-dessus (Figure 39) présente les classes implémentées dans le logiciel pour le 2^{ème} mode. On notera l'apparition de classes « techniques » (DataBase, Scheduler, etc.) car nous sommes ici au niveau de la proposition d'une solution. La classe DataBase permet l'accès aux données Access pour lire et enregistrer des traces d'animaux, ou lire des paramètres.

La Figure 40 montre une simulation multi-agents en cours d'exécution. Les 2 premières icônes en haut de la fenêtre permettent l'accès aux 2 modes de fonctionnement actuellement implémentés (Simulation multi-agents et Suivi de traces GPS). Une portion du terrain peut être agrandie pour suivre le détail d'une trace. On peut également agir sur la vitesse d'exécution, sur les traces laissées par les animaux (trace complète visible, ou seulement animal visible) ainsi que suspendre et reprendre la simulation. On peut également visualiser l'heure, le temps dans le simulateur ne s'écoulant évidemment pas à la même vitesse que pour nous. La simulation en cours porte sur

2 vaches, avec les faciès de végétation en fond d'écran, afin de visualiser les trajets d'un faciès à l'autre entre les repas.

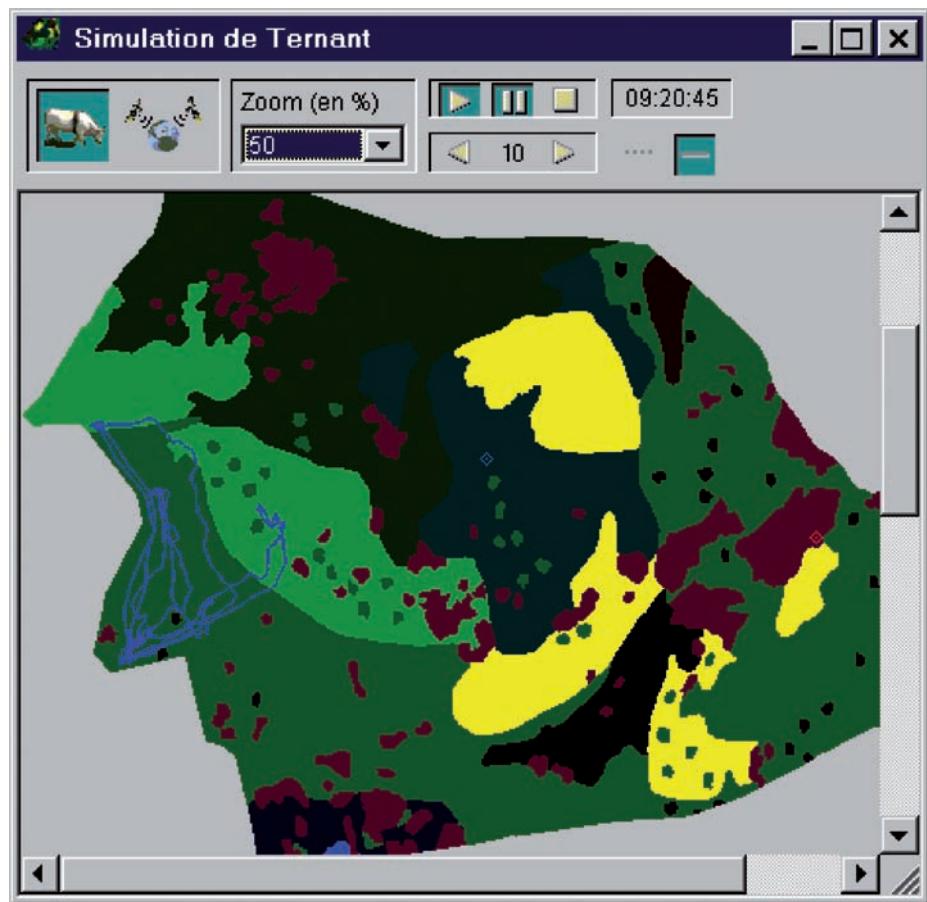


Figure 40. Simulation multi-agents en cours d'exécution

Règles de comportement de la simulation

Les règles de comportement de la simulation multi-agents sont issues de réunions de travail avec des éthologues de l'INRA et des spécialistes de la végétation de l'ENITA, du LEVC, de l'ouvrage suivant : [Arnold et Dudzinski 1978] et de l'analyse des traces GPS et ETHOSYS. Elles peuvent également être le fruit du simple bon sens. Il existe 3 grandes classes de règles :

la dynamique de la végétation (croissance et sénescence de la végétation, etc.), les règles de déplacement des animaux [Bailey *et al.* 1998] et les règles de consommation des animaux [Dumont 1996]. Ces deux dernières classes sont étroitement liées, l’animal se déplaçant principalement pour se nourrir. De même, la croissance de la végétation sera fortement liée à la pression de pâture et au piétinement lié à la présence d’animaux. Dans ce mode de fonctionnement, on ne prendra pas en compte les règles de croissance de la végétation, la durée de simulation étant trop courte pour mesurer une variation significative. L’influence du climat (température, vent, pluie) n’est pas encore utilisée dans le modèle. Il est cependant reconnu qu’il influence fortement le comportement des animaux [Arnold et Dudzinski 1978]. Par exemple, en cas de températures élevées, les animaux auront tendance à se déplacer du côté de l'estive susceptible de recevoir un peu de vent. Cette amélioration est prévue dans une version ultérieure du logiciel de simulation.

Données issues de la simulation

Les informations générées par chaque simulation sont conservées (sous forme de traces identiques aux traces GPS) dans une table « Résultats » de la base de données. Ces données pourront ensuite être traitées de la même manière que les traces relevées par le GPS, et servir, en les cumulant sur une période donnée, à constituer des cartes de fréquentation et de consommation.

Ces cartes calculées permettent de vérifier la cohérence de fonctionnement de ce mode par comparaison avec celles déjà existantes ou obtenues dans le 1^{er} mode du simulateur, notamment cela permet de vérifier si les traces issues du second mode ont été créées avec des paramètres adaptés. On repère ainsi aisément les similitudes (fréquentation des zones, chemins suivis, pression de pâturage, prise en compte du relief, etc.), ou l’absence de similitudes (dans le cas de règles inadaptées ou trop incomplètes) entre les cartes issues des données réelles et les cartes obtenues par simulation. Les résultats obtenus par ce module peuvent donc être utilisés pour subir une validation visuelle.

L’utilisation de simulations interactives et même le rendu réaliste ont été envisagés. Grâce aux collaborations établies, un des objectifs initiaux concernant la visualisation réaliste des résultats a pu être partiellement atteint. La visualisation précise a permis de vérifier les modèles, les échelles et la cohérence des systèmes de coordonnées. Le modèle numérique du terrain de Terrian réalisé par l’ENITA a pu être visualisé avec des résolutions satisfaisantes (nécessitant cependant 800 Mo de mémoire vive pour un modèle 3D avec

un maillage tous les 4 m². Le calcul d'une seule image réaliste sur le terrain de Ternant avec des modèles numériques de vaches et de chevaux nécessitait une heure de calcul sur les machines les plus puissantes de l'ISIMA en 1998 (Silicon Octane et serveur IBM G40 avec 1 Go de RAM) (**Figure 41**). Pour réaliser une animation d'une minute à l'époque, plus d'un millier d'heures de calcul auraient été nécessaires. Les aspects de rendu réaliste concernent essentiellement la communication avec des décideurs non scientifiques tandis que les experts préfèrent souvent des simulations interactives et des résultats statistiques plus exploitables pour une analyse scientifique.

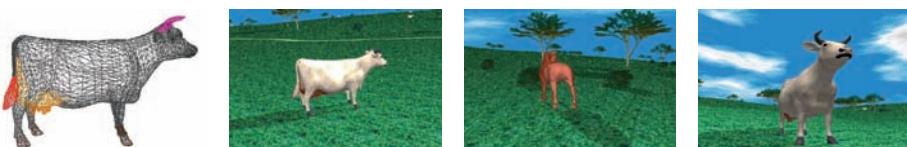


Figure 41. Maillage d'une vache en 3 dimensions et images de synthèse d'animaux sur le modèle du terrain de Ternant

Limites et perspectives

117

Nous devons également préciser les limites de notre démarche. Dans la pratique, peu d'animaux ont été suivis par les systèmes GPS et ETHOSYS. Cette technique demande en effet une logistique très importante pour faire un suivi complet du troupeau. Les données issues d'un tel suivi seraient également bien trop volumineuses pour être traitées, stockées et analysées aisément. Rappelons que les données GPS et ETHOSYS occupent aujourd'hui 300 Mo pour une seule saison de pâturage. Il ressort donc que les données utilisées pour l'étude du comportement animal, même si elles permettent d'avoir une base de renseignement solide, ne représentent pas tous les aspects de la réalité. De même, certaines parties du développement ont subi des limitations liées aux contraintes informatiques. Notamment, le développement a été initialisé en Java, avec l'idée de permettre l'exécution du simulateur *via* le World Wide Web. Nous nous sommes alors heurtés à deux sortes de problèmes. Le premier problème est que les volumes à transmettre (cartes, données GPS, écriture des traces générées, etc.) à travers le réseau sont trop importants pour permettre une exécution à distance. Le deuxième problème porte sur le langage Java lui-même. Java est un langage semi-interprété qui nécessite de s'exécuter sur une JVM (Java Virtual

Machine). Or, nous atteignions à l'époque les limites de la gestion mémoire de la JVM actuelle, car nous étions obligés de stocker en mémoire différentes cartes qui comportent chacune 1 million de points. Sur ce point, un gros travail d'optimisation du volume de ces cartes a été entrepris. Les évolutions futures pourraient concerner la conception et le développement d'un dernier module du simulateur pour la simulation de troupeaux sur un intervalle de temps important (plusieurs années). Ce module incorporerait non seulement le comportement des troupeaux d'animaux, mais également la dynamique de la végétation. Une autre évolution d'envergure concerne-rait la prise en compte du climat dans la détermination du comportement des animaux et des troupeaux. Pour permettre une visualisation réaliste de la simulation et des résultats, il serait également possible d'envisager une visualisation en 3 dimensions à base de VRML.

5.4. Conclusion

Pour ces études nous avons pu réaliser la difficulté de la collecte et du tri des informations servant de base aux règles de comportement de nos agents. Si l'on souhaite obtenir un comportement des agents semblable à celui de leurs modèles réels, cette étape reste fondamentale. Elle implique un travail d'équipe et des réunions de travail fréquentes. Sur ce type de projet ambitieux, la complexité du système écologique en présence a impliqué la participation de beaucoup de personnes de métiers différents (éthologues, agronomes, informaticiens, etc.). Bien évidemment le modèle proposé ne peut pas actuellement répondre à tous les besoins exprimés, notamment parce qu'ils ont été relativement ardu à spécifier. Principalement à cause de la diversité des intervenants, les sous-objectifs du modèle divergent entre les différentes disciplines, entraînant un surcoût d'efforts que ce soit pour les tentatives de validation des résultats préliminaires ou pour la valorisation de ces résultats. Malgré les problèmes liés à cette collaboration d'envergure, nous avons pu présenter des logiciels permettant à la fois de visualiser et de comprendre les données initiales, ainsi que de simuler les comportements individuels et grégaires des animaux grâce à une approche multi-agents [Hill *et al.* 2000b].

6. INTÉGRATIONS DES TECHNIQUES DU WEB

6.1. *Introduction*

La simulation sur le Web a rapidement émergé en tant que nouveau centre d'intérêt pour les chercheurs en simulation et ses utilisateurs. Cette technique correspond à une intégration du Web dans le domaine de la simulation. Cette intégration permet de découvrir différentes pistes d'évolution pour les techniques de simulation. Le Web a rénové une grande partie des technologies logicielles, les experts en simulation se devaient de l'étudier pour appréhender la manière de concevoir des simulations sur le Web. On peut percevoir la simulation sur le Web comme un mariage des technologies du Web avec celles issues de la simulation. Les nouvelles avancées technologiques liées au Web l'ont rendu viable pour développer des mécanismes d'exécution, de diffusion et de distribution de modèles.

Paul Fishwick, expert international en simulation, a été l'un des premiers à s'intéresser au couplage du Web et de la simulation [Fishwick 1996] lors de la *Winter Simulation Conference* qui a eu lieu en décembre 1996 à San Diego. J'avais pu discuter avec Paul en janvier 1996 à Phoenix des possibilités de rendre exécutable des modèles à distance *via* des navigateurs Web. Nous avions à Clermont-Ferrand fait réaliser par un étudiant de maîtrise une version « simulation sur le Web » d'un de nos projets de recherche en éthologie. Nous avions alors convenu avec Paul d'organiser pour janvier 1998, la première conférence internationale de simulation sur le Web [Fishwick *et al.* 1998]. Cette nouvelle conférence s'est maintenue depuis 3 ans [Bruzonne *et al.* 1999] [Signorile et Blais 2000]. Toujours dans une optique de suivre l'évolution de cette approche nouvelle pour la simulation, nous avons édité avec Paul un numéro spécial de la revue *Simulation* [Fishwick et Hill 1999]. Ceci dit, la simulation sur le Web reste un sujet assez vague au sein duquel nous identifions plusieurs pistes d'investigation :

- La simulation en tant qu'hypermédia. Du texte, des images, de l'audio, et des vidéos de simulation. La nature même du Web permet la conception, la production, le stockage et la recherche de « documents » contenant tous les éléments pré-cités. Le fait que des simulations deviennent accessibles *via* n'importe quel poste de travail équipé d'un navigateur est un facteur suffisant pour changer significativement les habitudes et les méthodes d'enseignement, de formation pour toutes les disciplines qui utilisent la simulation, notamment les sciences

de l'ingénieur, la physique, la biologie,... Des concepts pour l'éducation à distance et sur les possibilités d'interaction avec des outils de simulation pour la formation sont en train de voir le jour [Alfonseca *et al.* 1999].

- Les recherches en méthodologie de simulation. Le fait de pouvoir rapidement diffuser des modèles, des résultats, grâce à leur publication sur le Web, ouvre de nouveaux horizons de recherche dans le domaine de la simulation. Les aspects pratiques, notamment concernant la validation des modèles, les aspects économiques et juridiques liés à la publication de modèles, ne doivent cependant pas être négligés.
- L'accès à des programmes de simulation *via* le Web. C'est ce concept qui est souvent associé à la terminologie « Web-based simulation » ou « Simulation sur le Web ». Ce concept comprend d'une part l'exécution à distance d'anciennes simulations à partir de navigateurs proposant des formulaires HTML reliés à des scripts CGI, et d'autre part, le développement de codes mobiles tels que des simulations reposant sur des applets Java et s'exécutant à distance sur des machines clientes.
- La simulation et la modélisation distribuée. Nous n'avons presque pas abordé cet aspect qui comprend les activités rattachées à l'utilisation du Web associé à des technologies dérivées, telles que celle des objets distribués reposant sur CORBA ou sur les Remote Invocation Methods de Java. Ces technologies de distribution d'objets peuvent fournir l'infrastructure nécessaire à l'exécution de simulations distribuées. Les jeux sur Internet font également usage de technologie similaire, très efficace mais qui reste des technologies « propriétaires ». En effet SEGA leader pour les consoles incorporant des capacités de simulation distribuées sur le Net, est en rude compétition avec Nintendo et Sony. La diffusion des technologies sous-jacentes n'est pas à l'ordre du jour. Les applications utilisant les techniques de la réalité virtuelle adaptées au Web telles que celles développées autour du Virtual Reality Modeling Language (VRML) sont, elles, dans le domaine public.
- Le développement d'outils logiciels, d'environnements et/ou de cadriels qui facilitent la conception « collaborative » et distribuée de modèles de simulation sont en train de voir le jour.
- La simulation du Web. Cet aspect concerne la modélisation, l'analyse et l'étude des performances du Web afin de mieux caractériser

ser son fonctionnement, son utilisation et ceci afin d'optimiser ses performances.

Parmi tous les points évoqués ci-dessus, il nous semble que l'utilisation des atouts qu'offre le réseau Internet peut être significativement bénéfique au domaine de la modélisation et de la simulation. L'utilisation de données précises et maintenues à jour sur des bases de données réparties sur le réseau, peut améliorer significativement les processus d'aide à la décision, et ce dans la mesure où les modèles vont être conçus différemment. En effet, il nous semble que les modèles de simulation devraient pouvoir :

1. gérer des données *via* le réseau à partir de bases de données affectées dynamiquement,
2. gérer de nouvelles classes d'objets et/ou de nouvelles données inconnues au début de l'exécution d'un modèle et ce grâce à des métamodèles transmis en cours d'exécution,
3. gérer des composants logiciels distribués et faiblement couplés plutôt que des composants fortement couplés.

Parmi les avancées technologiques permettant de développer des simulations sur le Web, nous souhaitons préciser l'importance du langage Java et des technologies qui se développent autour de ce langage. En effet, la programmation du réseau devient beaucoup plus accessible grâce à des classes capables d'ouvrir des « sockets », de faire des requêtes sur des bases de données distantes,... et surtout de nouvelles classes peuvent être incorporées dynamiquement en cours d'exécution d'un programme grâce à des descriptions en Java IDL (Java Interface Description Language) autorisant ainsi une évolution dynamique des programmes. Les objets présents sur un nœud du réseau peuvent être « sérialisés » (archivés) puis envoyés sur un autre nœud où ils pourraient être incorporés et exécutés dans un modèle existant, s'ils se rattachent à une superclasse commune ou s'ils implémentent le protocole requis. Nous avons déjà évoqué le fait que des objets peuvent en appeler d'autres sur des ordinateurs distants *via* les Remote Method Invocation (RMI).

6.2. Quelques applications

Avant d’aborder les différentes applications que nous avons pu réaliser avec André Campos dans le cadre de sa thèse, je souhaite présenter quelques références montrant l’impact de cette nouvelle approche sur le développement des simulations. Voici une liste non exhaustive de projets qui ont pu aboutir à des publications :

- la mise à disposition de modèles distribués [Praehofer et Schoepl 2000],
- l’implémentation d’applications et d’environnements de simulation basés sur le langage Java [Campos et Hill 1998a],
- le développement coopératif et interactif de modèles à travers le Web [Hirata *et al.* 2000],
- le développement des environnements de simulation multi-utilisateurs [Narayanan *et al.* 1999],
- L’utilisation de ressources multimédia interactives dans les simulations, soit pour l’exécution de la simulation elle-même, soit pour sa documentation [Alfonseca *et al.* 1999],
- la simulation distribuée sur le Web [Page *et al.* 1997] [Rao *et al.* 1999],
- l’implémentation de jeux de simulation sur le Web [Ravid et Rafaeli 2000],
- l’intégration et la standardisation de DIS (Distributed Interactive Simulation) et HLA (High Level Architecture) sur le Web [Page 1998],
- l’intégration du VRML dans les simulations [Schmidt 1999],
- l’utilisation d’objets distribués (OLE/COM, CORBA, HLA, DIS) dans les simulations [Iazeolla et D’Ambrogio 1998] [Karhela *et al.* 2000].
- l’utilisation des scripts CGI dans les modèles distribués [Georgiev et Hoogenboom 1999].

122

En ce qui concerne nos travaux, les premières applications de la simulation sur le Web que nous avons développées reposaient sur l’architecture ECO-SIM [Hill *et al.* 1994a]. En tant que *classe d’environnements de modélisation d’écosystèmes*, cette architecture préconisait l’utilisation de l’Internet pour le développement et la communication des modèles. Un projet de coopération pluridisciplinaire entre l’ISIMA – Institut Supérieur d’Informatique, de Modélisation et de leurs Applications – et l’INRA – Institut National de

Recherche Agronomique – a pu être porté sur le Web en 1996. Les résultats concrets n'ont été publiés qu'en 1998 [Hill *et al.* 1998a]. Il s'agissait de la simulation multi-agents du comportement d'un groupe de moutons au pâturage que nous avons présentée précédemment dans le paragraphe dédié aux applications multi-agents. Le développement était simple car il reposait sur un script CGI qui récupérait les paramètres de la simulation à partir d'un formulaire HTML exporté sur les machines clientes. Les résultats de simulation se présentaient sous forme de fenêtre sur le poste client. Précisons cependant qu'à partir de l'identification des fonctionnalités génériques de cette première application, il a été possible de développer une deuxième application plus abstraite concernant la simulation d'un environnement restreint habité par des proies et un prédateur [Campos et Hill 1998a]. Le but théorique de cette simulation consistait à vérifier et à analyser les situations dans lesquelles le prédateur croit qu'il peut rattraper une proie (ce qui aboutit à une attaque), ainsi que les situations dans lesquelles les proies croient qu'elles peuvent être attrapées (ce qui aboutit à une tentative de fuite). Chaque proie, ainsi que le prédateur, possède son propre comportement, modélisé grâce à la logique floue pour représenter sa perception de la situation (ce qu'il « croit »). Tous apprennent et changent de comportement avec leurs succès et leurs échecs dans les tentatives de fuite ou d'attaque.

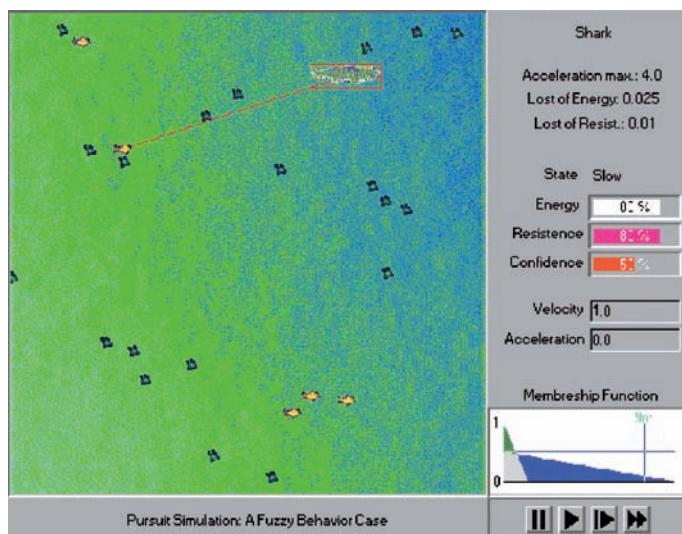


Figure 42. Capture d'écran d'une simulation de proies et prédateurs co-habitant dans un espace restreint [Campos et Hill 1998a]

Cette application a été réalisée en langage Java en utilisant l’AWT comme support pour l’interface utilisateur. Elle prend en charge des modifications visuelles et interactives (manipulation directe sur les entités) au cours de l’exécution de la simulation. La Figure 43 affiche une capture d’écran de l’application en cours de simulation (les proies sont des poissons et le prédateur est un requin).

124

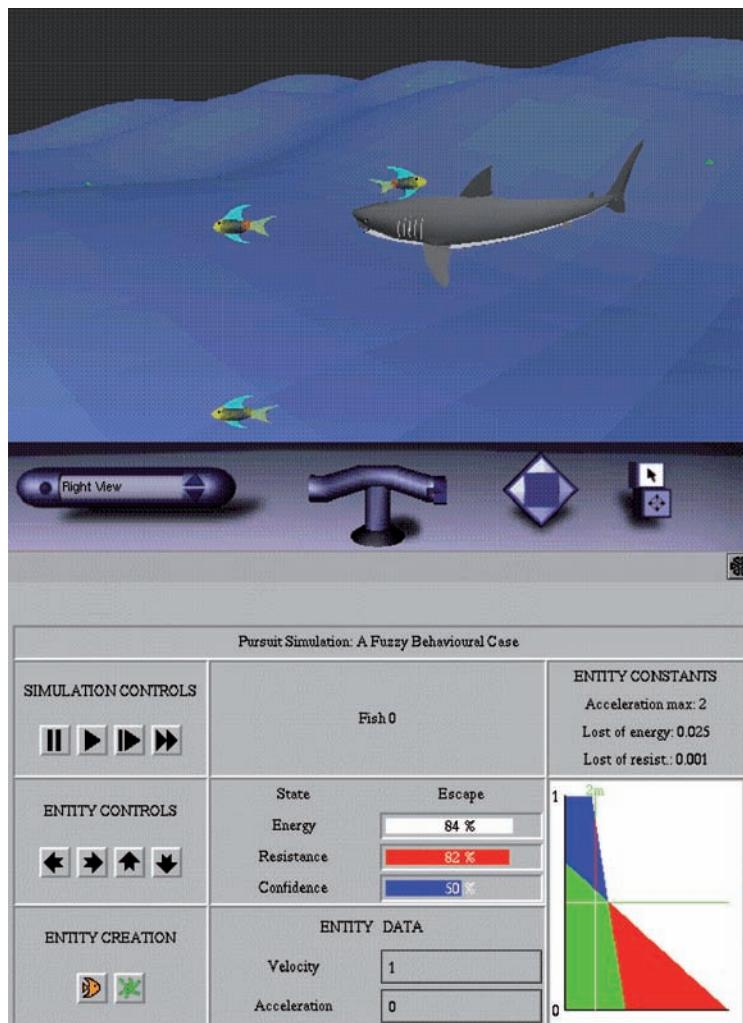


Figure 43. L’application proie-prédateur dans un modèle tri-dimensionnel [Campos et Hill 1998a]

Un autre exemple d'application de simulation sur le Web est la plate-forme Da Vinci [Campos *et al.* 2000]. Le but de cette plate-forme est de parvenir à simuler la construction de volumes creux par des « agents-fourmis » autonomes à partir des ressources disponibles dans leur environnement. Pour cela, il est nécessaire de déterminer les règles élémentaires qui, au niveau individuel, permettent d'une part l'exploitation des ressources de l'environnement et d'autre part la réalisation collective de tels volumes. Ces volumes représentent le nid d'une société de fourmis. La complexité des processus en jeu dans les sociétés d'insectes sociaux, ainsi que celle rencontrée dans la plupart des systèmes naturels, implique la conception de modèles aptes à suivre une modification continue. Cela est dû à notre méconnaissance des paramètres de ce type de systèmes et, par conséquent, les paramètres nécessaires pour les représenter de manière fidèle dans un modèle sont souvent à découvrir. Le prototypage de modèles, en essayant de définir ses règles élémentaires, devient une technique de développement indispensable. Nous avons alors essayé de mettre en place un moyen de prototypage du modèle à travers la définition de nouveaux comportements. La définition de nouveaux comportements est faite par l'utilisateur, en tant que spécialiste du domaine ; elle peut se réaliser à distance à travers le Web. Cette plate-forme est toujours en développement et des extensions sont d'emblée envisageables pour ce qui concerne des nids arboricoles plus complexes (ex : genre Azteca) constitués d'une juxtaposition de volumes de chambres. La Figure 44 (page suivante) montre des captures d'écran de la plate-forme Da Vinci : son applet de simulation ainsi que quelques pages Web.

6.3. Conclusion sur les simulations de type « Web-based »

Nous avons abordé l'état actuel du domaine de la simulation sur le Web et nous avons également survolé deux exemples d'application. Nous pensons que les perspectives d'évolution se situent maintenant au niveau de l'utilisation sur le Web de langages formels de spécification de modèles. En effet, l'évolution vers les langages formels de haut niveau pour la spécification de modèles a été motivée par le souhait de rendre la simulation plus accessible en éliminant les tâches fastidieuses de programmation.

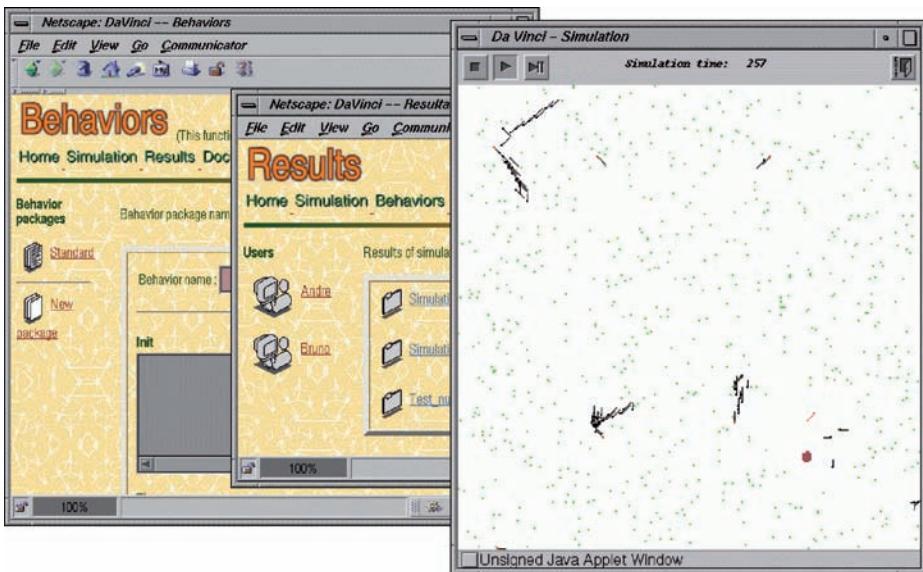


Figure 44. Pages Web pour la configuration de comportements et l'affichage de résultats, et une fenêtre de l'applet de la simulation [Campos *et al.* 2000]

126

7. PERSPECTIVES

Nous avons vu dans ce chapitre quelques applications concrètes que nous avons pu réaliser dans le cadre de divers projets, régionaux, nationaux et européens. Pour chacun de ces projets, différentes techniques informatiques ont été utilisées et intégrées : des bases de données géographiques aux technologies liées au Web. L’expérience acquise lors de ces développements nous a amené à réfléchir sur les méthodes et sur les outils logiciels, sachant que ceux-ci évoluent presque aussi vite que le matériel informatique. Dès 1994, nous évoquions ces problèmes ; le nombre d’écosystèmes susceptibles d’être simulés est énorme, et leurs domaines sont variés. Afin d’implémenter facilement des simulateurs pour différents domaines de l’écologie, sans avoir à reconsidérer à chaque fois tous les éléments théoriques, il est indispensable de considérer le problème à un haut niveau d’abstraction [Hill *et al.* 1994a]. Il est possible de concevoir, pour la simulation d’écosystèmes, un ensemble de concepts intégrant simplicité et réutilisabilité. Plutôt que de tenter une définition d’un hypothétique environnement de modélisation d’eco-

systèmes « universel », il nous semble que la définition d'une architecture d'environnements est plus à même de répondre aux exigences de la simulation d'écosystèmes. Une architecture logicielle de ce type devrait :

- donner aux concepteurs (l'informaticien et les experts du domaine) une méthodologie susceptible de construire rapidement le modèle de connaissance du système,
- permettre une analyse du domaine qui évite à l'informaticien l'analyse et la conception d'un simulateur propre pour chaque écosystème,
- faciliter la construction d'un simulateur et son intégration avec différents outils tels que les Systèmes d'Information Géographique (SIG), les bases de données biologiques et les bases de connaissance,
- autoriser la simulation de sites de taille très importante grâce aux techniques avancées de distribution de l'information et des calculs sur les réseaux à hauts débits,
- proposer une interface conviviale offrant non seulement la possibilité d'effectuer aisément la gestion des expériences et des données en entrées du simulateur, mais aussi la visualisation et l'animation graphique des résultats, y compris à distance *via* le Web,
- aider à la vérification et à la validation du modèle grâce à des outils statistiques et mathématiques, notamment en tenant compte de plans d'expériences.

127

L'expérience acquise lors du développement des modèles qui ont été présentés nous a amené à présenter avec André Campos un cadriciel pour la simulation d'écosystèmes : MAVIS (Multi Agent Visual Interactive Simulation) dédié à la simulation visuelle interactive de modèles individus centrés. Pour tous les prototypes que nous avions pu réaliser, André a analysé la structure des classes, leurs relations et leur mode de fonctionnement,... et ce afin de mettre en place une architecture logicielle réutilisable. Le modèle conceptuel résultant de cette analyse a pu être réutilisé pour le développement de la simulation de l'entretien des paysages par des herbivores en moyenne montagne que nous avons déjà présentée dans le paragraphe dédié aux agents [Hill *et al.* 2000b]. Lors de la conception de MAVIS, plusieurs aspects concernant l'interface utilisateur ont été identifiés. De ce fait, le cadriciel a été fractionné en deux parties : une pour gérer les aspects de mise en œuvre de la simulation et l'autre pour gérer l'affichage des résultats au cours de l'exécution et pour permettre à l'utilisateur d'interagir avec le modèle. La

Figure 45 présente une vue globale de ce cadriel nommé MAVIS (pour Muti-Agent Visual Interactive Simulation).

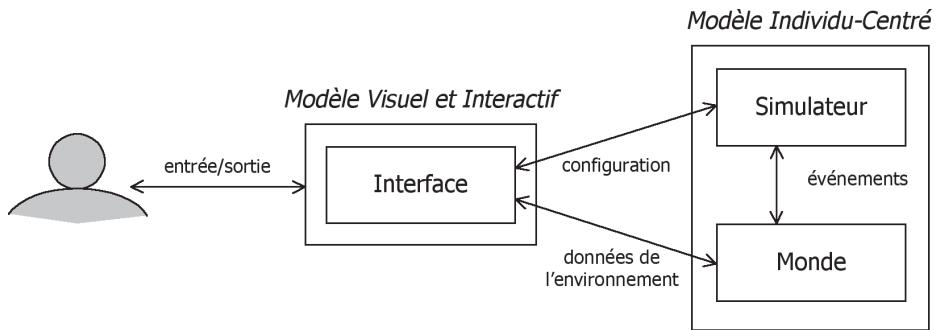


Figure 45. Architecture du cadre MAVIS au plus haut niveau d'abstraction [Campos et Hill 1998b]

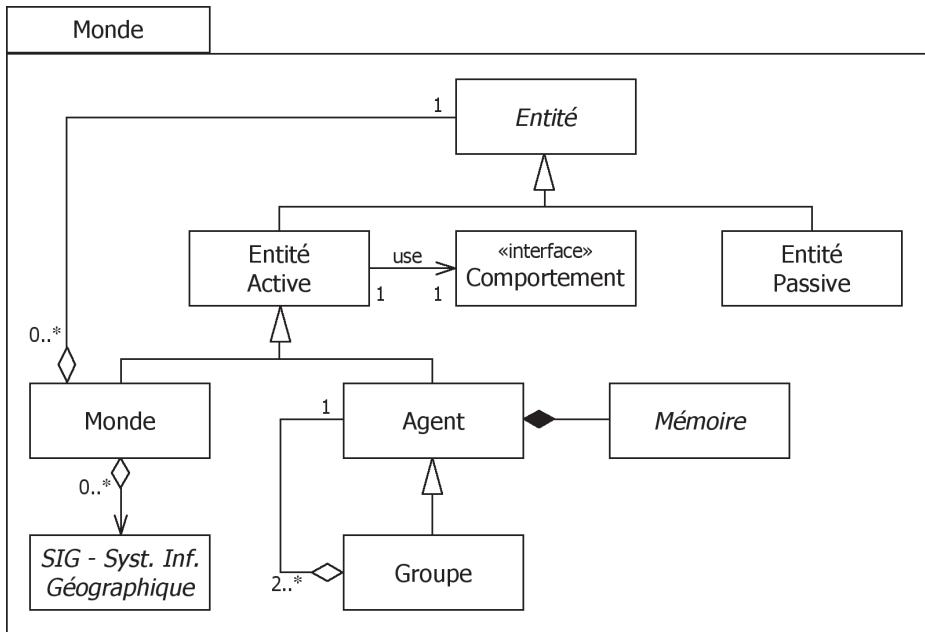


Figure 46. Extrait d'un diagramme de classes UML pour la réalisation d'environnements de modélisation d'écosystèmes [Campos et Hill 1998b]

Les approches visuelles et interactives ont été généralisées dans la deuxième partie du cadiciel. Elles étaient implémentées dans les applications précédentes à partir d'une visualisation bidimensionnelle classique. Il manquait, pour élargir la capacité de réutilisation de MAVIS, une étude des aspects visuels et interactifs avec une visualisation tridimensionnelle. Une application a ainsi été développée en utilisant des paramètres tridimensionnels dans le modèle et en affichant des résultats tridimensionnels [Campos et Hill 1998a]. Cette application reprend l'application test de proies et prédateurs et rajoute la dimension de profondeur dans le modèle. Pour l'affichage, nous avons retenu le langage VRML pour modéliser les entités concrètes du système. La **Figure 43** qui a été présentée dans le paragraphe dédié à la simulation sur le Web montre un affichage de cette application sur une station Silicon Graphique Octane. L'utilisation de classes identiques permettant de créer différentes applications, chacune avec ses particularités et avec des langages variés (C++, Java, Java/VRML) a permis de « valider » les aspects génériques du cadiciel MAVIS (présenté en détail dans [Campos et Hill 1998b]).

La **Figure 46** présente une partie du diagramme de classes de MAVIS, plus précisément les classes qui reflètent les relations de fonctionnement nécessaires à la simulation discrète d'écosystèmes sous contraintes spatiales. Le lecteur intéressé pourra se reporter utilement à la thèse d'André Campos [Campos 2000].

Parmi les perspectives, on peut citer l'utilisation du formalisme DEVS (Discrete Event Specification) qui est actuellement le plus universel pour la simulation. Paru en 1990 [Zeigler 1990], son évolution et ses capacités sont maintenant mondialement reconnues et synthétisées dans l'ouvrage de référence [Zeigler *et al.* 2000]. Nous pensons depuis une décennie qu'il est possible de développer soit un DEVS « ML » (pour markup language), soit une grammaire XML (DTD) autorisant la spécification, la conception et l'exécution de modèles, décrits de manière formelle et interprétable par un navigateur équipé d'un « plugin » adapté pour la simulation. J'ai annoncé cette possibilité dans mon mémoire d'habilitation en 2000, cependant nous avons bien conscience que les environnements de développement à base de spécification formelle ne sont pas encore arrivés à maturité. Ils sont diffusés avec parcimonie et les systèmes spécifiés de manière formelle sont en pratique souvent difficiles à modifier ou à étendre, non seulement parce que peu d'experts en simulation sont formés à ces langages formels, mais aussi parce que la séparation imposée entre la spécification du système et son implé-

mentation conduit souvent à des modèles qui reflètent pauvrement le comportement du système modélisé. De plus le potentiel de réutilisabilité et de diffusion des modèles auprès d’experts non informaticiens ou auprès d’entreprises reste fortement limité tant que des éditeurs interactifs de modèles, générant le code formel interprétable par le «plugin» ne seront pas disponibles pour le Web. La clé se trouve donc dans le camp des développeurs. En 2004, Jean Baptiste Filippi a d’ailleurs défendu sa thèse en montrant qu’une approche de programmation visuelle à base de DEVS était tout fait adaptée à la simulation pour l’environnement [Filippi et Bisgambiglia 2004].

CHAPITRE 5

LES PROBLÈMES DE VALIDATION ET DE VÉRIFICATION DES MODÈLES

*Pesons le gain et la perte, en prenant croix que Dieu est.
Estimons ces deux cas : si vous gagnez, vous gagnez tout ;
si vous perdez, vous ne perdez rien.
Gagez donc qu'IL EST sans hésiter*

Blaise Pascal

1. INTRODUCTION

Tous les modèles que nous avons présentés reposent sur l'implémentation de programmes. Quel que soit le modèle construit, il convient de cerner les limites d'utilisation et les conditions d'interprétation des résultats. Dans tous les cas, il faut *vérifier* que le code du programme fonctionne de manière saine, qu'il correspond à une traduction correcte du modèle et des hypothèses retenues et que le code informatique du modèle est exempt d'erreurs de codage notoires. De même, il est essentiel de *valider le modèle et ses résultats* en les comparant si possibles à ceux du système réel modélisé de manière à déterminer si le modèle est, dans son cadre expérimental, une représentation valable du système réel. La validation d'un modèle doit bien sûr prendre également place au niveau conceptuel, c'est-à-dire avant l'écriture d'un pro-

gramme. En fait, il est primordial que la validation s'effectue tout au long du cycle de développement d'un modèle de simulation et non pas uniquement une fois le modèle implanté sur un ordinateur. Nous avons abordé les problèmes de vérification et de validation des modèles dans [Hill 1993a, 1996] et ce qui suit dans ce chapitre reprend essentiellement les travaux que nous avons présentés avec mes collègues Coquillard et Mazel dans [Hill *et al.* 1996b], [Coquillard et Hill 1997] et dans [Mazel *et al.* 1997] ainsi que des travaux plus récents sur les plans d'expériences et leur exécution en parallèle dans le cadre de programmes de recherche avec l'INRA.

Comme pour toute terminologie, les définitions associées à la notion de vérification et de validation peuvent varier d'un domaine à l'autre. Ainsi, les définitions de la communauté du génie logiciel diffèrent de celles préconisées au sein de la communauté de la simulation. Voici donc ces définitions :

- **Vérification** : « *Substantiation that a computerized model represents a conceptual model within specified limits of accuracy* » [SCS, 1979].
- **Validation** : « *Substantiation that a computerized model within its domain of applicability possesses a satisfactory range of accuracy consistent with the intended application of the model* » [SCS, 1979].

132

Afin de compléter ces définitions, il est essentiel de rappeler qu'une étude de simulation est menée dans un but précis. Le modèle doit donc avoir une utilisation bien définie, et la validation ne pourra être menée qu'en fonction de l'objectif de la modélisation et uniquement dans un cadre d'utilisation donné. En effet, si l'on décide de réaliser à la même échelle des maquettes de voitures en plastique ainsi que des maquettes de trains avec des wagons capables de transporter des voitures, il sera possible de faire des simulations de chargement des wagons avec différents types de voitures. La modélisation effectuée permet de répondre à des questions telles que : combien de voitures de type X peut-on charger sur un wagon ? Par contre, si l'on souhaite avoir une estimation de la pollution en CO₂ émise par les voitures, les modèles en plastique sont strictement inutiles. De même, une maquette en bois peint d'un futur bâtiment en verre, placé au sein d'une ville miniature peut apporter des réponses à des considérations paysagères ou d'encombrement. Par contre, si l'on souhaite savoir quelle sera la température dans ce bâtiment avec une température extérieure de X°, le positionnement d'un thermomètre dans la maquette en bois ne sera d'aucun secours. Les objectifs de modélisation sont donc cruciaux ; ils vont déterminer les choix de conception du modèle et, par conséquent, le contexte de la validation. Les motivations d'une étude de simulation peuvent être de plusieurs ordres. En

ce qui concerne la modélisation d'écosystèmes nous nous limitons aux deux cas suivants :

1. le modèle est construit dans le but de fournir des prédictions les plus précises possibles concernant le comportement du système réel,
2. le modèle est construit à la manière d'une théorie scientifique, afin d'améliorer la connaissance du fonctionnement interne du système réel et donc de faciliter sa compréhension.

Comme nous le précisons plus haut, les techniques de validation vont différer suivant les objectifs retenus. Il est bien sûr possible, mais hasardeux, de faire des prédictions avec un modèle construit dans un objectif d'amélioration de la connaissance du fonctionnement d'un système. D'après [Balci 1994], les risques encourus lors de la réalisation de tests de validation sont du même ordre que ceux des tests statistiques usuels, à savoir : « *le risque de première espèce consiste à refuser un modèle valide, et le risque de deuxième espèce consiste à accepter un modèle non valide* ».

Nous allons aborder dans un premier temps la notion de cadre expérimental, puis la vérification de code avant de nous focaliser principalement sur la validation. Nous avons pu réaliser une étude générale des techniques de validation applicables à la modélisation des écosystèmes. Elle est présentée dans [Hill 1995b]. Nous avons, avec Claude Mazel et Patrick Coquillard, étudié la validation de modèles couplés à des Systèmes d'Information Géographique dans [Mazel *et al.* 1997] ; de même nous avons abordé l'analyse spectrale de résultats produits par de tels couplages dans [Hill *et al.* 1996a et b] et [Coquillard *et al.* 1997]. Nous n'avons que très peu investi sur les techniques formelles de validation, elles ne seront pas détaillées ici. Le lecteur intéressé pourra cependant se reporter aux travaux que nous avions présentés dans [Attoui et Hill 1995]. Les techniques graphiques et d'animation de résultat de simulation ayant été suffisamment traitées dans nos travaux de 3^{ème} cycle [Hill 1993a], nous n'y reviendrons pas. Citons cependant les travaux récents que nous avons par ailleurs déjà présentés dans le chapitre précédent, évoquant l'intérêt des simulations visuelles sur le Web avec des langages tels que VRML et Java pour la vérification et la validation des modèles [Campos et Hill 1998b] [Campos *et al.* 2000].

2. LES CADRES EXPÉRIMENTAUX

Revenons sur l'un des concepts théoriques fondamentaux pour la validation et la vérification : le concept de cadre expérimental introduit par [Zeigler 1976]. La validation et la vérification d'un modèle doivent absolument être faites pour un « *objectif donné et en référençant un cadre expérimental précis* ». En ce qui concerne les études de simulation, un cadre expérimental doit au minimum comprendre les éléments suivants :

- les données observées (prélèvements, mesures, lois de distribution ajustées en fonction des valeurs mesurées, etc.) qui sont en entrée du modèle,
- les initialisations des paramètres d'entrée,
- les contraintes temporelles en entrée du modèle,
- la définition des sorties du modèle (la quantité des résultats, leurs formats et leur mode de représentation, graphiques, tables, etc.),
- la ou les condition(s) d'arrêt de la simulation.

En effet, un modèle peut être considéré valide pour certains cadres expérimentaux et invalide pour beaucoup d'autres. D'une manière générale, les modèles de simulation sont invalidables dans l'absolu [Sheng *et al.* 1993]. Par nature, les modèles d'écosystèmes sont parmi les plus ardu斯 à valider car, d'une part, les systèmes étudiés ne sont pas conçus par l'homme – donc très incomplètement connus – et, d'autre part, l'échantillonnage des variables est difficile et les données préalablement disponibles souvent peu adaptées aux objectifs de l'étude.

134

3. LA VÉRIFICATION DES PROGRAMMES DE SIMULATION

Les logiciels de simulation sont parmi les plus difficiles à coder [Lehman 1980] et, comme pour tout logiciel, les tests ne doivent pas être négligés. La phase de vérification doit mettre en œuvre un ensemble de tests aidant à prouver que le programme est une *traduction correcte du modèle conceptuel* élaboré par la phase de modélisation. Il existe de nombreux tests : *inspection* du code, *test de régression*, *tests alpha et bêta*, *test de performance*... [Jacobson *et al.* 1993]. Nous retenons ici la taxonomie des techniques de vérification

de Balci et Whitner [Balci et Whitner 1989] qui identifient les six catégories suivantes, des plus informelles aux plus formelles :

1. L'analyse informelle est menée par un modélisateur différent de celui qui a réalisé le modèle ; cette personne doit examiner le code avec une capacité de raisonnement d'autant plus objective qu'elle n'est pas directement impliquée dans le développement.
2. L'analyse statique effectue un ensemble de tests automatiques sur la structure du code source (structure, sémantique,...).
3. L'analyse dynamique effectue des tests de comportement du programme lors de son exécution, afin d'établir des profils de comportement, d'explorer le flot de contrôle...
4. L'analyse symbolique examine, si possible, la transformation par le programme d'un ensemble de données d'entrées symboliques en un résultat symbolique également.
5. L'analyse de contraintes se base sur le modèle de la *programmation par contrat*, consistant à vérifier que le *programme est cohérent avec les hypothèses* faites. Elle utilise fréquemment des assertions. L'analyse de contraintes permet également de déterminer le domaine d'applicabilité du modèle.
6. L'analyse formelle se base sur les techniques de preuves de programmes (utilisant les déductions logiques avec transformation et calcul de prédictats, etc.).

135

Pour l'implantation de ces différents tests nous renvoyons le lecteur aux ouvrages de génie logiciel tels que ceux de [Jacobson *et al.* 1993], [Sommerville 1993] ou à des articles et ouvrages de référence en simulation [Zeigler *et al.* 2000], [Law 1991], [Pace 1992], [Sheng *et al.* 1993], [Balci 1994], [Youngblood 1995]. Nous tenons cependant à attirer l'attention du lecteur sur un test d'analyse dynamique, intéressant pour la vérification des programmes implémentant des modèles stochastiques : la technique des réplications utilisée pour calculer des intervalles de confiance, consistant à effectuer une série de nombreuses simulations sans réinitialiser le générateur de nombres pseudo-aléatoires, peut être très utile pour explorer les nombreuses possibilités du modèle, la combinaison des portions de code du modèle dépendant des événements générés. En effet, s'il est impossible de générer de manière déterministe toutes les séquences de tests pour un modèle stochastique, par contre une exploration aléatoire sur de très longues périodes constitue une

méthode de test remarquable qui, d'après notre expérience, peut révéler des erreurs de programmation après plus de 2 000 réplications !

Dans le cycle de développement d'une application de modélisation de taille importante, les tests précédents représentent une partie non négligeable en coût et en temps. Dans ce cas, un modèle de tests peut regrouper l'ensemble du travail portant sur les tests de l'application. Il définit le type de tests nécessaires, ceux à développer, les éléments à tester, les différentes contraintes à respecter, la stratégie de test utilisée, et bien d'autres informations relatives aux opérations de tests [Sommerville 1993]. Tous ces éléments du modèle de test participent à l'élaboration d'un plan de test qui doit permettre de dire si l'application développée suit les objectifs définis par la modélisation.

136

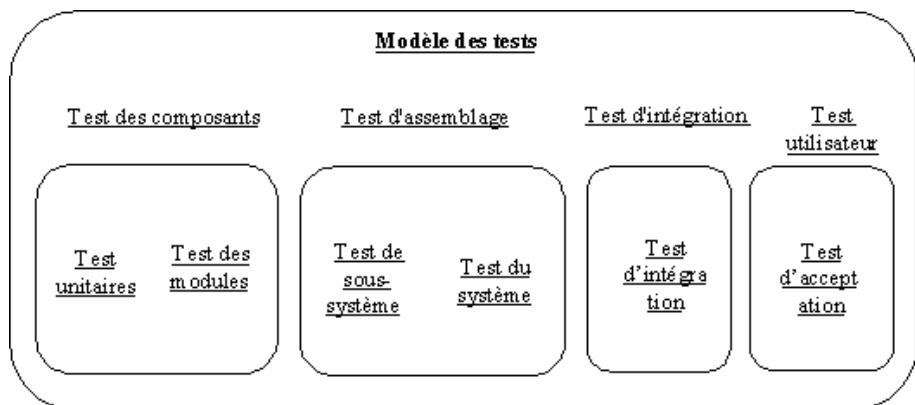


Figure 47. Classification et séquence des tests d'un logiciel de simulation

En utilisant l'architecture interne des programmes développés avec une approche orientée objet ainsi que la séquence des tests au cours du cycle de développement d'un logiciel, nous avons retenu le modèle de test classique de la Figure 47 :

1. Les *tests des composants* permettent de vérifier le fonctionnement correct des objets techniques et des objets du métier dans leur cadre normal d'utilisation. Les méthodes des objets sont testées une à une (tests unitaires), puis les objets en tant qu'entités du programme (test des modules). Les classes d'objets sont celles identifiées lors de la modélisation et de la réalisation du modèle conceptuel.

2. Les *tests d'assemblage* permettent de vérifier la cohérence des différents objets regroupés en sous-systèmes d'objets ayant un objectif commun ; les interactions implémentant les relations entre les différents objets d'un même sous-système doivent être évaluées et vérifiées. Enfin, l'interconnexion des différents sous-systèmes d'objets doit être vérifiée afin de fournir le logiciel final.
3. Les *tests d'intégration* visent à vérifier qu'un logiciel peut être utilisé dans son environnement informatique cible (logiciels, réseau,...) sans le perturber.
4. Les *tests utilisateurs* permettent de vérifier si le logiciel de simulation répond bien aux exigences et aux objectifs définis dans le modèle conceptuel.

Comme le montre la **Figure 47**, les tests des quatre classes peuvent être réalisés des plus particuliers aux plus généraux. Pour chacun des niveaux, les tests permettent de mettre à jour certains types d'erreurs. Avec cette approche hiérarchique, les tests de niveau « N+1 » ne devraient déceler que peu d'erreurs résultant du niveau « N ». Les tests de composants révèlent surtout les erreurs de programmation des classes d'objets (erreurs de codage) et quelquefois des erreurs d'intégration et d'analyse. Les tests d'assemblage mettent plutôt à jour des erreurs de conception, certaines erreurs de programmation (apparues lors de l'interaction des divers objets) et parfois des erreurs d'analyse. Les tests d'intégration doivent révéler les interactions néfastes qui pourraient survenir entre le logiciel de simulation et les autres logiciels de la plate-forme cible, tels que les logiciels systèmes, réseaux ou d'autres applicatifs. Et enfin, les tests de l'utilisateur révèlent tous types d'erreurs (d'intégration, de conception ou de codage) n'étant pas apparues lors des précédentes phases.

137

La vérification du code telle qu'elle est décrite précédemment est maintenant très bien supportée par les ateliers de génie logiciel standard (CASE Tools : « Computer Aided Software Engineering »). En effet, de nombreux outils du marché fournissent non seulement des méthodologies de développement, des supports pour la génération de la documentation, mais aussi des générateurs automatiques de tests. Une vérification rigoureuse du code est essentielle, de manière à pouvoir écarter toutes hypothèses sur ses insuffisances, lors des phases de validation des résultats de simulation.

4. LA VALIDATION DES MODÈLES ET DES RÉSULTATS

4.1. Introduction

Contrairement à la vérification, qui ne peut s'effectuer que lors du codage du programme de simulation, la validation doit se dérouler tout au long du cycle de développement d'un modèle. En fait, pour obtenir une efficacité optimale, c'est l'ensemble des techniques de vérification et de validation qu'il faut appliquer au cours du cycle de vie d'un projet de modélisation. Comme pour le développement d'un logiciel, les tests et la documentation doivent être élaborés en parallèle avec le modèle. En effet, l'application des techniques de vérification et de validation en fin de développement se révèlent très coûteuse et souvent peu efficace.

Phases du cycle de vie d'un modèle de simulation

Phases	Analyse	Analyse et Conception	Implémentation	Interprétation
Validation	validation des données	validation du modèle conceptuel	validation opérationnelle	validation des résultats
Par confrontation	×	×	×	×
Par répétitivité				×
Fonctionnelle		×	×	×
Graphique	×		×	×
Statistique	×	×	×	×

138

Figure 48. Adaptation de la taxonomie des techniques de validation de Mazel

Pour délimiter le cadre de validité d'un modèle, il faut toujours spécifier le cadre expérimental [Sargent 1979]. Pour être considéré valide, un travail de modélisation doit, d'une part, disposer d'un modèle conceptuel valide, et d'autre part, les résultats générés par le modèle de simulation doivent avoir une marge de précision acceptable pour les objectifs et l'ensemble des conditions expérimentales fixées. Comme pour le processus de vérification, il faut effectuer des tests d'évaluation permettant de présumer de la validité d'un modèle. Il faut également rappeler qu'il est impossible de montrer qu'un

modèle de simulation est totalement valide sur l'étendue de son domaine d'application. Le résultat des tests de validation ne constitue donc qu'une présomption de validation.

Parmi les différents tests de validation, aucune procédure et aucun algorithme n'est disponible pour choisir la technique à utiliser suivant le type de modèle et le type d'application. Vous trouverez ici la présentation des techniques de validation élaborée avec mon collègue Claude Mazel. Nous préconisons une taxonomie de ces techniques basée sur les phases du cycle de vie d'un modèle (Figure 48) [Mazel *et al.* 1997]. La phase d'*analyse* d'un modèle comprend l'acquisition et l'analyse des données ; la phase de *conception* doit aboutir au modèle conceptuel ; l'*implémentation* correspond au codage du programme de simulation, et la dernière catégorie cible l'*interprétation* des résultats. Dans un premier temps, nous nous attacherons à présenter les techniques de validation. Nous nous consacrerons ensuite aux différents types de validations pour l'analyse des données et le modèle conceptuel.

L'essentiel des techniques de validation est présenté depuis de nombreuses années par [Zeigler 1976], [Gordon 1978], [Fishman 1978], [Leroudier 1980], [Sargent 1984], [Law 1991], [Hill 1993b], [Sheng *et al.* 1993] et [Youngblood et Pace 1995]. Nous présentons ci-dessous les cinq catégories de validation de la Figure 48 :

139

1. **La validation par confrontation** est basée sur l'expérience des experts du système, acquise par l'observation du système étudié. Elle consiste à demander aux experts du système si *le comportement du modèle* leur paraît cohérent. On utilise donc cette technique pour déterminer de manière *empirique* si la logique de fonctionnement est correcte et si les relations entre les entrées du modèle et les résultats sont acceptables. Cette validation est informelle, l'expert étant supposé capable d'appréhender le bon fonctionnement du modèle. Le **test de Turing** fait partie de cette catégorie. Au cours de ce test, des résultats chiffrés de simulation sont présentés aux écologues. Il leur est alors demandé une appréciation empirique de la qualité de ces résultats (*i.e.* sans comparaison aucune avec des observations réelles du système).
2. **La validité de répétitivité** consiste à comparer le modèle réalisé avec d'autres modèles ou avec la réalité. Il est parfois possible de comparer les résultats de modèles de simulation sur des cas simples avec d'autres modèles valides (modèles analytiques ou modèles markoviens) ou avec la réalité. Des calculs d'intervalles de confiance permettent de

comparer différents échantillons de manière formelle ; des techniques de classement et de sélection permettent d'identifier de manière probabiliste le – ou les – meilleur(s) modèle(s) parmi plusieurs. Pour les tests de **validité prédictive**, le modèle est utilisé pour prédire le comportement du système. Des comparaisons résultats/observations sont faites pour déterminer si le comportement du système et les prédictions sont cohérentes.

- 140
3. **La validité fonctionnelle** vise à exploiter le modèle comme instrument de mesure et d'expérimentation afin de s'assurer de son bon fonctionnement. Plusieurs techniques sont regroupées dans cette catégorie. Nous présentons ici les plus usitées. Avec les **tests de validité des événements**, les événements générés par le modèle sont comparés à ceux du système. On peut aussi tester la **validité structurelle** : un modèle est dit structurellement valide, non seulement s'il fournit des résultats satisfaisants, mais également, si la structure interne du modèle de simulation correspond à la structure de fonctionnement de la réalité. Les comportements de différentes entités du système peuvent être tracés pendant une exécution pour déterminer si la logique du modèle est correcte et si la précision nécessaire est atteinte. L'utilisation de **tests de conditions extrêmes** permet de s'assurer que la structure et les résultats d'un modèle sont plausibles pour toute combinaison extrême, souhaitable ou non, des paramètres du modèle. La **dégénérescence** du comportement du modèle est testée en supprimant des portions du modèle. L'**analyse sensitive** consiste à agir sur les paramètres identifiés comme sensibles pour le modèle, afin de vérifier que son comportement reste toujours cohérent. Des méthodes d'estimation de gradient permettent de tester la sensibilité des paramètres. Dans les cas les plus simples, l'**utilisation de constantes**, fixées pour toutes les variables internes et d'entrée d'un modèle, peut parfois permettre une validation des résultats du modèle par confrontation à des calculs manuels.
 4. **La validation graphique** et **l'animation** permettent l'affichage des résultats du modèle sous forme de courbes, histogrammes et « camemberts », ou d'images donnant l'état de certaines entités du système au cours du temps. La représentation par animation de l'évolution temporelle de certaines entités du modèle permet de représenter *la dynamique transitoire* d'un système. L'animation présente l'énorme avantage d'utiliser la capacité de l'être humain à apprécier les rela-

tions spatiales, parfois essentielles, pour la compréhension du système [Hill 1993b, 1996].

5. **La validation statistique** peut consister d'une part en une comparaison des résultats de simulation à des mesures (ou à des résultats fournis par un autre modèle), et d'autre part, en l'établissement d'intervalles de confiance dans le cas d'un modèle stochastique (plusieurs réplications sont exécutées pour déterminer la variabilité stochastique interne du modèle). Lors de la comparaison de distributions de processus stochastiques, on distingue le cas des processus stationnaires – dont l'évolution converge au cours du temps et pour lesquels les tests d'ajustement classiques conviennent – de celui des processus transitoires dont la convergence n'est pas prévisible, et pour lesquels il faut envisager des traitements statistiques particuliers : analyse spectrale [Fishman et Kiviat 1967] ou ajustement à des modèles paramétriques [Hsu et Hunter 1977].

4.2. Analyse et validation des données

La réalisation d'un modèle passe par une phase d'acquisition des données qui caractérisent le système étudié. Ces données doivent permettre l'élaboration de modèles de données en accord avec les objectifs de la modélisation. La modélisation d'écosystèmes reste un domaine où l'acquisition des données est ardue, coûteuse (écologie terrestre), voire parfois irréalisable. Les données jugées pertinentes par les experts du domaine constituent donc des entrées du modèle et figurent au sein du modèle conceptuel soit sous forme de simples paramètres, soit sous forme de variables aléatoires. Il est toujours intéressant de synthétiser les données recueillies et si possible de leur associer (ajuster) une loi de distribution. Lorsque cela n'est pas possible, on peut se contenter de l'histogramme de ces données. Si besoin est, on peut alimenter le modèle à l'aide des données réelles (on parle alors de *traces*). Ceci ne se justifie que très rarement, par exemple lorsque les résultats sont extrêmement sensibles à la distribution des données, ou, plus sûrement, lorsque l'on souhaite tester la prédictivité du modèle par rapport à un scénario précis dont les résultats exacts sont connus ; il est rare qu'un tirage aléatoire suivant l'histogramme des données réelles ne suffise pas à l'obtention du réalisme souhaité.

141

Dans le cas de la modélisation d'écosystèmes, il est fréquent qu'une partie des données soient indisponibles, et qu'elles ne puissent être échantillonées

dans un délai raisonnable. Dans ce cas, les écologues doivent faire des choix. Les disséminations de graines, de boutures et autres, peuvent être réparties suivant la loi normale ; les vitesses de déplacement d'animaux peuvent être choisies suivant une loi uniforme. Lorsque des dissymétries existent, on peut alors se tourner vers des lois bêta. En cas de méconnaissance totale de la répartition des données, une loi uniforme entre deux valeurs sélectionnées par les experts reste un choix conseillé. De même, si des valeurs sont connues, le choix d'une loi triangulaire ou d'une loi trapézoïdale peut être préconisé [Law et Kelton 1991].

4.3. *La validation du modèle conceptuel*

La validation d'un modèle conceptuel constitue un processus à part entière. Nous rappelons que le modèle conceptuel englobe les théories, les hypothèses, les algorithmes aussi bien que les interactions entre les modules. C'est le modèle conceptuel qui fait le lien entre les objectifs de la modélisation et l'implémentation du modèle. Le but d'une validation du modèle conceptuel est de s'assurer qu'il fournit une représentation raisonnable du système réel en fonction des objectifs.

142

L'obtention d'un modèle conceptuel est le résultat d'un travail d'analyse et de conception qui peut être conduit en suivant une méthode de modélisation dite *orientée objet*. L'avantage d'une méthode *orientée objet* est qu'elle se base sur les entités du domaine étudié, facilitant ainsi le dialogue avec les experts du domaine. Il est donc possible de bâtir un modèle de connaissance du domaine puis, au sein de ce domaine, de réaliser le modèle de connaissance du système que l'on étudie. Enfin, en fonction des objectifs de modélisation et de la connaissance acquise, le modèle conceptuel est élaboré. S'il est bien entendu intéressant de faire valider les modèles de connaissance de manière empirique par les experts du domaine, la validation du modèle conceptuel doit permettre de s'assurer que son élaboration n'a pas dévié sensiblement des objectifs initiaux.

La validation de la connaissance acquise peut être réalisée par confrontation avec les experts du système ou du domaine auquel appartient le système, s'il s'agit de modélisation *a posteriori*. Cette validation de la connaissance comprend, d'une part, le modèle des données dont la validation a été présentée et, d'autre part, la structure du modèle élaboré. La validation du modèle conceptuel entraîne donc une validation structurelle consistant à vérifier que la structure du modèle est conforme à celle du système réel.

L'approche *orientée objet* utilisée lors des phases d'analyse et de conception ne peut que favoriser la constitution d'un modèle qui possède une structure équivalente à celle du système réel. Les tests de sous-systèmes, exposés dans la phase de vérification, ne peuvent qu'en être facilités. Une autre technique de validation qui doit être appliquée au modèle conceptuel est la validation événementielle. Il faut s'assurer que les événements qui seront générés par le modèle sont cohérents avec ceux que produit le système. Les relations de causalité du modèle doivent être identifiées en collaboration avec les experts du système et identiques à celles du système réel [Balci 1994]. Dans le cas de simulation à événements discrets, il convient également de valider les choix de conception concernant la structure du noyau de synchronisation (est-elle en adéquation avec le problème traité ?), la qualité du générateur, les méthodes de génération des lois de distributions... [Law et Kelton 1991] [Kleijnen et Groenendaal 1992]. Dans le cas d'une simulation distribuée, les choix concernant le type de noyau de synchronisation, le mode de communication, le partitionnement des tâches, la distribution des nombres aléatoires, deviennent cruciaux et restent encore aujourd'hui délicats à valider [Palmore 1994]. Des solutions ont été récemment proposées pour le DoD (Department of Defence aux Etats-Unis) [Youngblood et Pace 1995].

143

5. UTILITÉ DE L'ANIMATION POUR LA VALIDATION DE RÉSULTATS DE SIMULATION

Les techniques graphiques sur ordinateur se sont largement répandues depuis maintenant deux décennies. Nous sommes intéressés plus particulièrement à l'apport du graphisme et de l'animation, en tant que technique d'évaluation des performances, pour la modélisation de systèmes complexes utilisant la simulation à événements discrets [Hill 1993b]. Les simulations visuelles en temps réel, qui offrent des possibilités impressionnantes combinant la réalité virtuelle avec des images de synthèse de qualité, sont devenues des réalités dans le domaine de la modélisation d'écosystèmes avec les travaux de Blaise et De Reffye [De Reffye *et al.* 1990] [Saito et De Reffye 1993].

D'une manière générale, l'animation permet d'analyser les comportements transitoires d'un système et vient compléter efficacement les résultats statistiques globaux. Les premières techniques d'animation de résultats de simulations avec une approche *orientée objet* sont dues à [Palme 1977]. Les

concepts de la simulation visuelle interactive grâce auxquels il est possible de modifier des paramètres en cours de simulation, sont dus à [Hurzion et Secker 1978]. Lorsque les paramètres sont changés au cours d'une simulation, on ne peut absolument plus espérer fournir des résultats statistiques globaux. Lorsque l'on anime un modèle, la vitesse à laquelle sont examinés les événements est souvent considérablement ralentie par rapport à la vitesse possible d'une simulation sans animation. Ce ralentissement nous permet de voir et de comprendre le modèle animé (avec une limite de 24 à 30 images par seconde pour les cerveaux rapides !). L'utilisateur qui visualise des animations peut malheureusement être tenté de tirer des conclusions sur des échantillons de taille insignifiante du fait de la faible vitesse de la simulation. Il convient de ne pas oublier que ce que l'on regarde ne constitue qu'un petit échantillon d'un processus stochastique, et il serait injustifié d'en inférer des conclusions. De plus, la possibilité d'interaction directe en cours d'exécution (*simulations interactives*) peut entraîner des modifications du modèle qui conduiront à des situations aberrantes qui ne correspondraient à rien de connu dans le système réel. Lors de l'utilisation de simulations interactives, seul l'aspect visualisation de phénomènes transitoires pour la mise au point d'un modèle ou pour l'enseignement est intéressant. L'animation permet, en effet, de présenter de manière visuelle des cas complexes, mettant simultanément en œuvre de nombreuses entités. Grâce à la simulation visuelle interactive, des analystes ou des modélisateurs débutants peuvent se former en comprenant comment leurs actions affectent les modèles et par conséquent les systèmes réels.

Nous avons présenté différentes techniques d'animation graphique de résultats de simulation, ainsi que les problèmes concernant leur mise en œuvre dans [Hill 1996]. L'animation des modèles constitue une partie essentielle d'un projet de simulation pour mener des études comportant de nombreux phénomènes transitoires. Les avantages de l'animation de résultats de simulation pour la validation de modèles sont présentés par [Shannon 1986] et [Gipps 1986]. Les différentes parties d'un projet de simulation au cours desquelles les apports des techniques graphiques et de l'animation sont significatifs sont résumées ci-dessous :

- la mise au point et la vérification des programmes de simulation,
- la validation des modèles conceptuels et de simulation,
- l'analyse et la conception d'expériences,
- la communication et la présentation des résultats.

Parmi les avantages ressentis par les utilisateurs d'animations graphiques de résultats de simulation, une étude de Hollocks [Hollocks 1984] rapporte trois gains principaux :

- un gain en confiance de l'utilisateur du modèle qui peut le relier à sa connaissance du système réel et du problème à traiter,
- un gain en compréhension du système,
- un gain d'implication des experts du domaine du système, qui se montrent d'autant plus enthousiastes qu'ils ont la possibilité de paramétrier ou d'interagir par eux-mêmes avec les modèles.

En fait, l'animation utilise les capacités de l'esprit humain à reconnaître visuellement des comportements et des relations complexes, permettant ainsi de détecter d'éventuelles déviations par rapport au fonctionnement attendu du modèle. L'animation fournit, de manière conviviale, la séquence des événements générés par un modèle autorisant un suivi simultané de plusieurs entités du système. L'animation est en effet essentielle pour aider à comprendre les différentes interactions entre des entités concurrentes. Cependant, il convient de prendre des précautions concernant les techniques d'animation, si le modèle n'est pas rigoureusement animé, c'est-à-dire si une partie importante des événements que peut fournir une trace ne sont pas animés. Dans ces conditions, une animation correcte ne signifie en aucun cas que le modèle est correct, débogué et encore moins vérifié. Un fort degré de cohérence entre l'animation et le modèle conceptuel est essentiel pour pouvoir considérer l'animation comme un outil de validation et de vérification. La représentation du fonctionnement dynamique d'un système apporte une aide à la validation des hypothèses ainsi qu'à la validation structurelle si l'animation graphique reflète les mécanismes propres à la structure interne de fonctionnement du modèle. Un modèle structurellement valide peut donner des résultats plus précis, ou même des résultats sur des sous-modèles – du fait de la prise en compte de la structure physique du système réel [Leroudier 1980].

145

Aspect à ne pas négliger, l'animation de modèles de simulation présente des atouts considérables au niveau de la présentation des résultats : « *Animation makes lively and accessible what would otherwise be a dry and somewhat obscure presentation of tables and figures* » [Smith et Platt 1987]. Dans la plupart des cas, le spécialiste en modélisation n'est pas l'expert du système. L'animation fournit alors un lien de communication graphique très efficace,

permettant l'amorce de discussions et de critiques constructives (tests de confrontation).

Enfin, il est primordial de pouvoir communiquer les résultats d'une simulation aux décideurs qui ont le pouvoir d'autoriser les modifications des systèmes existants (gestion d'écosystèmes) ou qui valident les décisions de conception des systèmes à réaliser (génie écologique). En effet, l'objectif de tout effort de modélisation est, entre autres, de fournir des informations aux décideurs. C'est pourquoi il est important que les informations essentielles soient crédibles et transmises de la manière la plus souple possible. Il est clair que les résultats statistiques avec leurs intervalles de confiance sont de bons éléments entre les mains des analystes scientifiques, mais l'animation va augmenter considérablement la crédibilité du modèle vis-à-vis des décideurs. Il faut alors veiller à ne pas surexploiter cette démarche ; il est en effet plus facile de développer un dessin animé, plutôt que l'animation d'un programme de simulation garantissant une expression qui ne trahit pas le modèle sous-jacent.

146

6. L'ANALYSE SPECTRALE : UNE TECHNIQUE D'AIDE À LA VALIDATION DE MODÈLES STOCHASTIQUES SPATIALISÉS

6.1. *Interprétation de résultats de couplage SIG - SAED*

Les résultats du couplage d'un Système d'Information Géographique (SIG) et d'une Simulation Aléatoire à Événements Discrets (SAED) peuvent être classés suivant des critères définis, d'une part, en fonction du type de résultats recherchés et de la manière dont ils sont établis et, d'autre part, en fonction de la nature de la population simulée : le type de résultat recherché peut concerner soit le terrain, soit les entités biologiques que la simulation a fait évoluer sur ce terrain [Hill *et al.* 1996b] [Mazel *et al.* 1997]. Dans le premier cas, il est naturel d'effectuer les moyennes pour chaque entité géographique, tandis que, dans le second cas, ces moyennes doivent être calculées pour chaque entité biologique, si cela est possible.

6.2. Analyse spatiale et analyse statistique

La particularité d'un résultat de Simulation Aléatoire à Événements Discrets est liée à l'aspect aléatoire de celle-ci. Il est en effet très important de ne pas commettre l'erreur de considérer le résultat d'une seule simulation comme fiable, car il est en effet lui même aléatoire, et une autre simulation peut donner un résultat sensiblement différent. Lorsque les résultats d'une SAED se présentent comme un ensemble de valeurs numériques indépendantes, il existe des méthodes classiques de calcul d'intervalles de confiance, qui permettent ainsi de quantifier la fiabilité des résultats fournis. Malheureusement, le problème est beaucoup plus compliqué pour des résultats spatiaux, comme un nuage de points dans l'espace, une carte plane,...

Les résultats fournis par la SAED doivent se présenter sous forme de moyennes établies, soit dans le temps (simulation stationnaire), soit sur un ensemble de réplications (simulation de Monte-Carlo). Ces calculs, qui sont nécessaires pour conférer une crédibilité suffisante aux résultats de simulation, présentent cependant deux inconvénients :

- d'une part, rendre leur interprétation plus délicate. Il est en effet très important de comprendre la signification exacte des calculs effectués afin de pouvoir fournir une interprétation correcte,
- d'autre part, écraser éventuellement certaines informations spatiales qui peuvent apparaître au niveau de chaque réplication et ne plus figurer dans le résultat final (la moyenne ne représente en effet pas autant d'information que l'échantillon).

147

Comme nous l'avons mentionné précédemment, un résultat de simulation est obtenu en effectuant une moyenne. Une conséquence négative de ceci est que les propriétés spatiales de ce résultat peuvent être écrasées par ce calcul. Supposons que les entités biologiques (EB) d'un système se répartissent, à tout instant, suivant un motif spatial, par exemple, en amas. Si ces EB se déplacent, tout en maintenant cette structure, il est possible que, de manière globale, toutes les entités géographiques (EG) du terrain aient une densité moyenne de population égale. Le résultat final fourni par une simulation stationnaire ne laissera donc plus apparaître cette structure spatiale. De même, les amas qui seront représentés par les résultats fournis par chaque réplication d'une simulation Monte-Carlo, peuvent être, en fonction des réplications équirépartis sur le terrain (**Figure 49**).

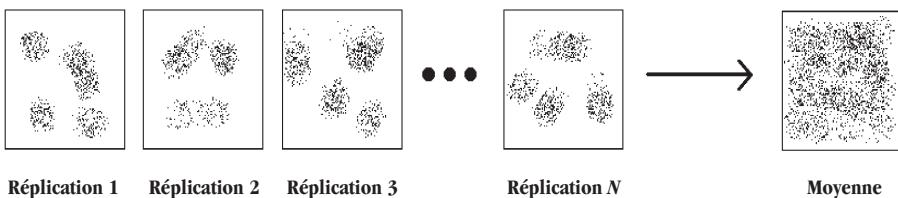


Figure 49. Exemple de structure spatiale qui n'apparaît plus dans le résultat final

Ces considérations simples montrent que l'analyse spatiale de résultats de simulation doit être effectuée de manière prudente : il ne faut pas se contenter d'exécuter une seule simulation, enregistrer les résultats à une date fixée, et réaliser une analyse spatiale sur ces résultats. En effet, cette analyse spatiale, en traitant des données qui ne sont pas statistiquement représentatives, peut apporter une information intéressante, certes, mais elle n'a pas un bon niveau de crédibilité. La nature spatiale du résultat de simulation peut être le résultat d'un hasard ; il est également risqué de vouloir d'abord obtenir des résultats de simulation crédibles, pour pouvoir ensuite appliquer une méthode d'analyse spatiale, puisque comme nous l'avons indiqué, il est possible que des informations intéressantes soient perdues à cause du traitement statistique effectué par la simulation.

148

6.3. Un exemple appliqu   en oc  anographie

Dans le contexte de la simulation de l'expansion de l'algue *Caulerpa taxifolia* d  j   pr  sent  , rappelons que les objectifs de cette mod  lisation t  aient entre autres de pouvoir pr  voir l'expansion de cette algue prolifique sur les zones largement envahies et notamment sur les colonisations profondes (au-del   de 40 m), ainsi que de pouvoir explorer divers sc  narios li  s  la dispersion anthropique, au d  veloppement ou  l'introduction de pr  dateurs sp  cifiques (lutte biologique), etc.

La simulation utilise en entr  e un fichier d'exp  rience qui permet de sp  cifier la connaissance des experts biologistes en r  glant un ensemble de param  tres. On trouve g  n  ralement en entr  e du mod  le, pour chaque site, les cartes précisant les substrats et la bathym  trie. Le mod  le peut g  n  ralement tre affin   en précisant sur les cartes les zones portuaires et les courants lorsqu'ils sont connus. Les entit  s g  ographiques (EG) sont des carr  s dont la surface varie en fonction du site tudi  . Les attributs g  ographiques de

chaque EG sont la profondeur et le type de substrat. L'expansion de cette algue se faisant par croissance de stolons, mais aussi par bouturage, chaque nouvelle bouture correspond à une nouvelle entité biologique (EB). Ces EB sont en nombre variable, et une EG peut être colonisée par plusieurs EB. Cette simulation est de type Monte-Carlo, chaque réplication fournissant une image de la répartition dans l'espace des EB et de leurs attributs. Il est intéressant de construire des images minimale, moyenne, et maximale, obtenues après de nombreuses réplications, afin de prendre en compte les multiples évolutions possibles de l'environnement. On obtient alors une répartition des probabilités de présence des EB et des fréquences des attributs qui leur sont attachés.

Un des sites de référence pour la modélisation se situe à Passable dans un trou de bombe de la seconde guerre mondiale. La Figure 50a représente la situation initiale adoptée par le modèle de simulation : en foncé on retrouve les zones d'herbier de Posidonie, et en clair les tâches de Caulerpe ($3,5 \text{ m}^2$), la largeur de la zone étant ici de 60 m, et la surface des EG de 16 cm^2 . Une cartographie très détaillée (au 1:1000^{ème}), et le repérage de chaque bouture avec un balisage a permis de calibrer le modèle et notamment les taux de bouturage, ainsi que les paramètres de dispersion. Les Figures 50b et 50c présentent deux résultats possibles fournis par deux simulations sur un an.

149

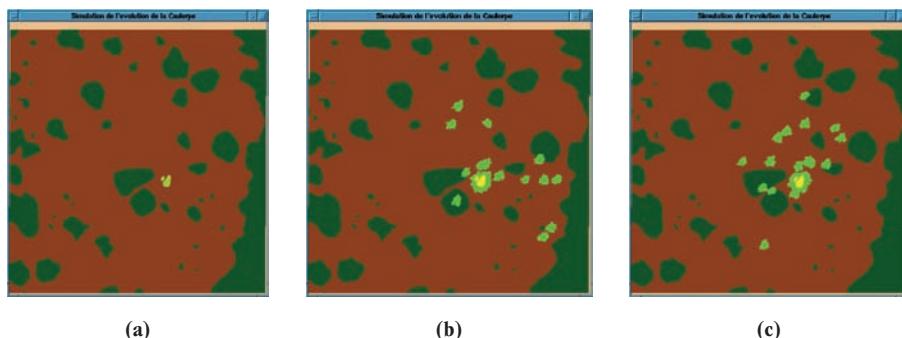


Figure 50. (a) Situation initiale dans le trou de bombe ($3,5 \text{ m}^2$) couvert par *Caulerpa* ; (b) Après un an de simulation, $58,4 \text{ m}^2$ sont couverts ; (c) deuxième réplique, $62,1 \text{ m}^2$

Les **Figures 51, 52 et 53** présentent des résultats d'analyse. Sur la **Figure 51**, nous observons en blanc toutes les cases atteintes après 256 réplications. Ce type de résultat peut aider à déterminer l'enveloppe d'une zone atteignable à partir d'une situation existante. Sur la **Figure 52** avec un dégradé de couleurs nous pouvons identifier les zones qui ont les plus fortes probabilités de colonisation (les plus claires). Il est également possible de présenter des visualisations que nous nommons « à seuil » en précisant le seuil de probabilité de colonisation que l'on souhaite visualiser. Dans les cas les plus simples nous pouvons aussi présenter une visualisation en 3 dimensions du spectre donnant les fréquences de colonisation. Sur la **Figure 53**, on peut même identifier des zones qui sont atteintes à toutes les réplications. Par contre l'obtention de plusieurs pics est souvent délicate car les résultats observés peuvent être induits par plusieurs attracteurs.

150

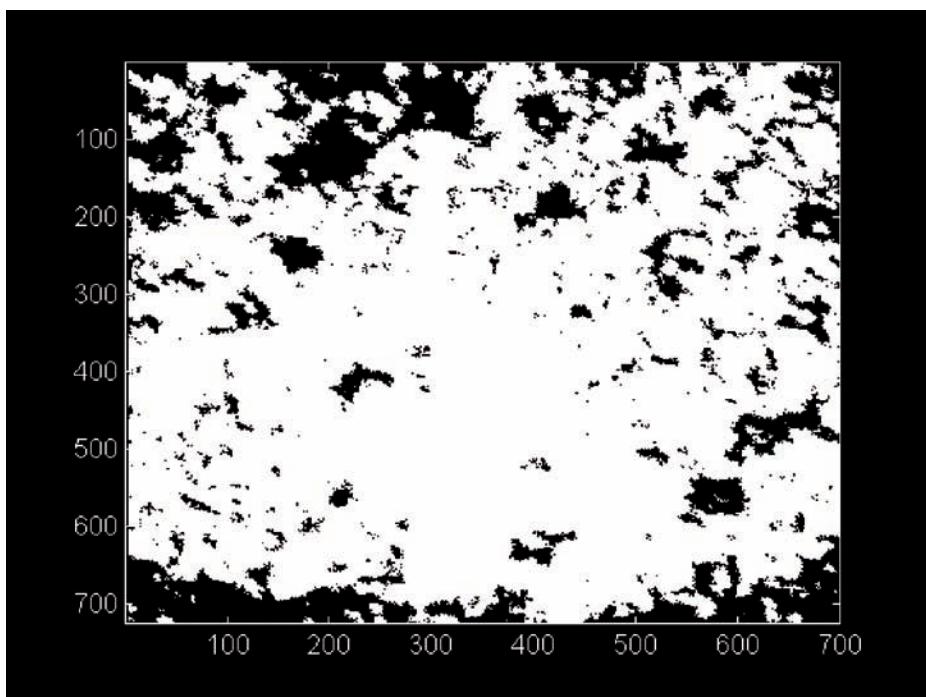
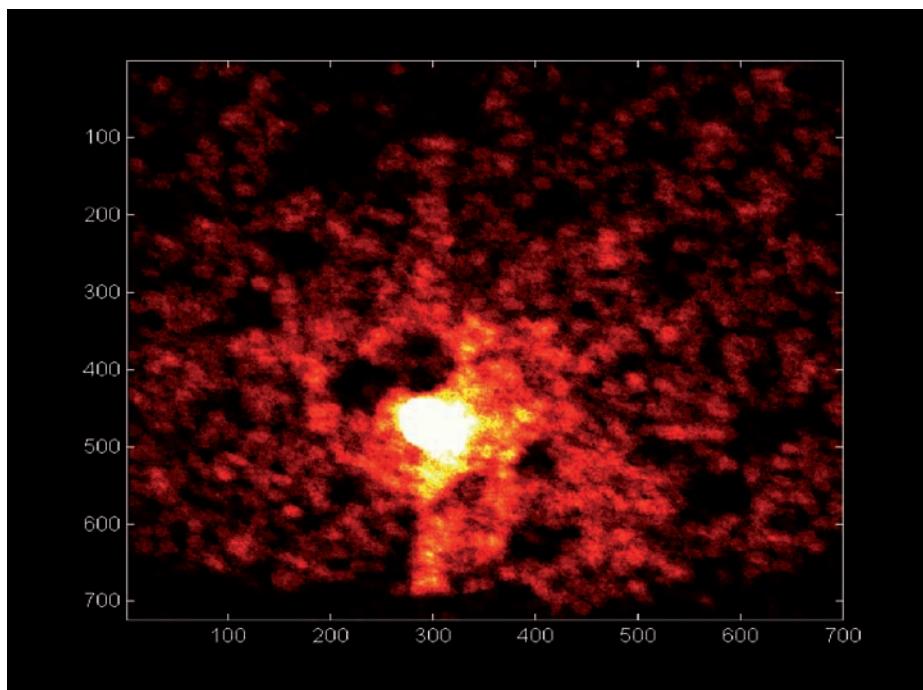


Figure 51. Identification d'une zone atteignable



151

Figure 52. Dégradé montrant le spectre des probabilités de colonisation

La technique d'analyse spectrale que nous utilisons suppose qu'il y ait eu enregistrement des répartitions spatiales pour chaque réPLICATION. Toutefois, ces images devront être confrontées aux résultats numériques correspondants, variance et homogénéité, par exemple, afin de détecter éventuellement plusieurs attracteurs spatiaux dont la « moyenne » ne serait guère informative. La mise en animation des images successives de l'écosystème peut mettre à jour des phases transitoires fugaces et difficilement décelables par d'autres moyens. La mise en évidence des fluctuations du célèbre trou dans la couche d'ozone en est un bon exemple. Parmi les méthodes de validation de modèles, l'analyse spectrale présente un intérêt non négligeable lorsque les résultats des simulations stochastiques montrent de fortes corrélations spatiales.

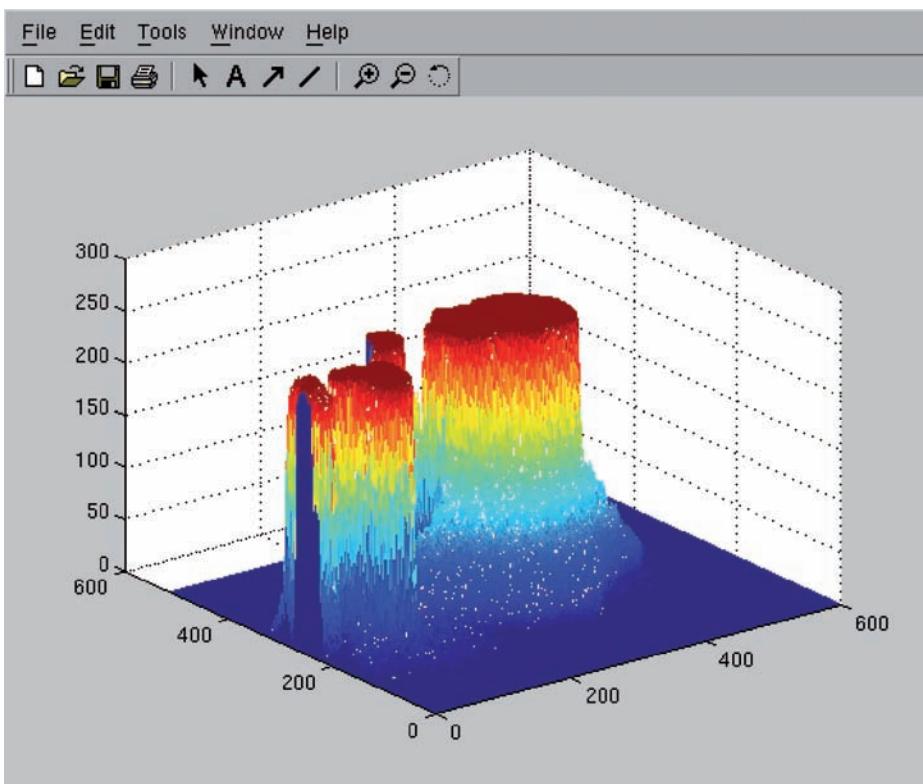


Figure 53. Visualisation d'un spectre en 3 dimensions

Si l'analyse spectrale peut aider à mettre en évidence les zones géographiques qui ont par exemple une grande probabilité d'être colonisées par une espèce ou d'être affectées par un phénomène physique ou chimique, elle entraîne également des difficultés dans l'interprétation des résultats. En effet, puisque le résultat visualisé correspond à une somme dans l'espace des possibilités, l'existence de pics se révèle délicate à analyser et nécessite une étude approfondie pour déterminer les effets possibles d'éventuels attracteurs qui pourraient être exclusifs d'une réPLICATION à l'autre. Supposons qu'à chaque fois qu'une zone A est contaminée, une zone B l'est également avec une probabilité p_1 . Deux pics apparaissent dans les zones A et B avec, par exemple, comme hauteur respectives H et Hp_1 . Supposons maintenant, que chaque fois que la zone A n'est pas contaminée, la zone B l'est, avec une probabilité p_2 , nous observerons à nouveau 2 pics dans les zones A et B, avec comme

hauteur respective H et $(H_{\max} - H)$ p_2 . Dans chaque hypothèse, ces 2 pics peuvent avoir la même hauteur !



Figure 54. Simulation dans le port de la Darse (5 ans)



Figure 55. La tâche en diagonale au sein du port donne la zone de *Caulerpe* cartographiée (5 ans)

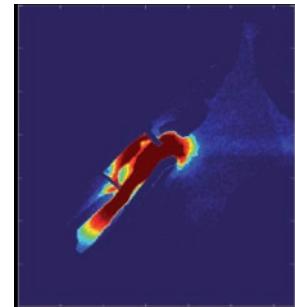


Figure 56. Visualisation du spectre pour 5 ans de simulation (256 répliques)

Étudions précisément un autre site, le port de la Darse à Villefranche-sur-Mer. La Figure 54 présente le résultat d'une simulation à l'intérieur du port, les tâches discrètes, claires et foncées au sein du port (que l'on peut repérer en gris clair avec sa géométrie) matérialisent les zones recouvertes par la Caulerpe. Ce site est beaucoup plus facile à étudier que Passable principalement grâce à une bonne visibilité sous-marine dans le port. La Figure 55 présente la situation qui était cartographiée. On peut remarquer en superposant les cartes, que les résultats de simulations de la Figure 55 s'inscrivent bien dans la carte réalisée par les plongeurs à l'exception de la prédiction par simulation de la présence systématique de boutures dans l'herbier qui entoure le port. En plongée il est presque impossible de repérer les petites boutures qui se développent en sous-strates dans un herbier de Posidonie. Des plongeurs ont alors parcouru la zone située à l'est du port en scrutant avec attention les niveaux inférieurs de l'herbier et ils ont pu localiser de nombreuses petites boutures et touffes qui avaient échappé à l'attention des cartographes. La Figure 56 montre le résultat d'une analyse spectrale avec 256 répliques. Ce type de résultats peut être obtenu avec des temps de calcul très raisonnables avec les machines actuelles. Ces temps étaient encore de quelques heures en 1999 avec des machines qui délivraient 30 SpecFp95 unité de mesure résultant d'une évaluation des

performances de la machine en calcul flottant). L'utilisation de fermes (clusters) et grilles de calcul, avec une approche saine de la parallélisation des flux aléatoires autorise de nos jours un calcul en quelques minutes. En effet, d'un point de vu pratique, nous avons utilisé récemment DistMe un outil conçu par Romain Reuillon dans le cadre de sa thèse (achevée sous ma direction en 2008). Grâce à DistMe nous avons parallélisé l'exploration de paramètres sur la simulation de croissance d'algues. Nous avons généré des spectres en trois dimensions présentant la somme des zones géographiques atteintes. Chacun des spectres est obtenu par le cumul de 1 024 cartes en deux dimensions générées de manière distribuée. Elles correspondent chacune à l'exécution d'une réPLICATION de la simulation. Le calcul de ces spectres représente 308 jours de calcul en séquentiel sur un processeur Intel Xéon à 3 GHz. Celui-ci a été réalisé en 11 jours sur un cluster de 14 bi-Xéons 3 GHz (soient 28 unités d'exécution). La Figure 57 montre deux exemples parmi les résultats obtenus lors de la campagne d'exploration des paramètres réalisée par Romain. Les deux graphiques du haut présentent un scénario optimiste et ceux du bas un scénario pessimiste de la colonisation de l'algue pour la côte entre Menton et Villefranche-sur-Mer (12 années sont simulées). Les cartes en deux dimensions à gauche représentent les zones atteintes par l'algue. Le cumul de 1 024 cartes, issues chacune des exécutions indépendantes du simulateur, permet d'obtenir les spectres en trois dimensions. Ces spectres sont présentés sur la gauche de la figure. Les axes X et Y représentent la zone géographique discrétisée (environ 360 m^2 par pixel) et l'axe Z présente le nombre de réPLICATION indépendantes de la simulation où la colonisation sur un pixel (x, y) a été constatée.

7. UNE APPROCHE LOGICIELLE POUR LA CONCEPTION DE PLANS D'EXPÉRIENCES

Il existe des outils puissants tels que SimLab et les logiciels de statistiques proposent également des aides à la conception de plans d'expériences. Dans un logiciel tel que MatLab la boîte à outil statistique propose déjà de nombreuses techniques. D'une manière générale, les simulations évoquées dans cet ouvrage utilisent des plans d'expériences, et parmi les techniques systé-

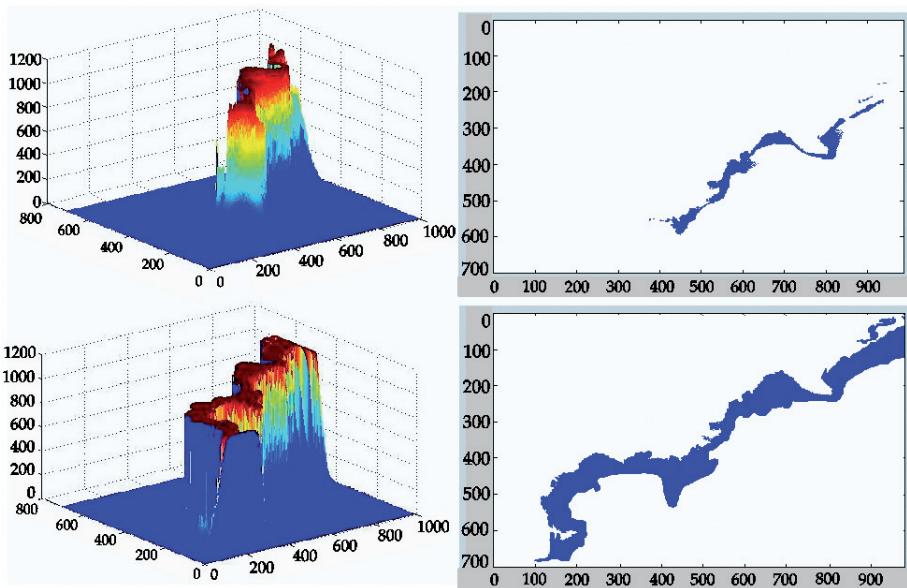


Figure 57. Sur la partie droite, on retrouve les spectres en 3D de simulation de la répartition de l’algue *Caulerpa taxifolia* ; à gauche nous avons la carte des zones colonisées par l’algue en 2D (chaque pixel correspond approximativement à 360 m^2)

155

matiques utilisées pour optimiser et concevoir des plans d’expériences. Les HyperCubes Latins, les plans D-Optimaux, et autres variantes ne sont pas détaillés ici. Voici par contre quelques techniques que nous avons mises en œuvre :

- Les méthodes factorielles permettent de construire des plans d’expérience de manière systématique en effectuant un parcours de l’espace des facteurs. Leur intérêt réside dans la possibilité de déterminer l’influence, sur une réponse, de plusieurs facteurs combinés [Fishman 1978], [Law 1991], [Kleijnen 1987], [Kleijnen et Groenendaal 1992].
- Les méthodes basées sur des méta-modèles de régression [Box et Draaper 1987] ou connexionnistes [Aussem et Hill 1999, 2000] permettent d’exprimer les réponses du modèle sous la forme d’une combinaison linéaire des facteurs obtenue grâce à un système d’équations de régressions multiples

- Les méthodes spectrales se basent sur des plans d’expérience forçant des oscillations des valeurs de chaque facteur avec des fréquences différentes afin d’analyser les fréquences qui peuvent être détectées en sortie du modèle [Sargent et Som 1992] [Hill *et al.* 1996b].

Nous avons avec le CEMAGREF entamé une réflexion sur un outil générique permettant l’exploration de simulation. Les travaux initiaux furent menés en collaboration avec Frédéric Amblard et Jérôme Truffot du Laboratoire d’Ingénierie des Systèmes Complexes dirigé par Guillaume Deffuant. Le but initial était de fournir à un utilisateur final les outils nécessaires à l’exploration de modèles durant les phases de vérification, de calibration, d’étude de sensibilité et de validation (Figure 58).

156

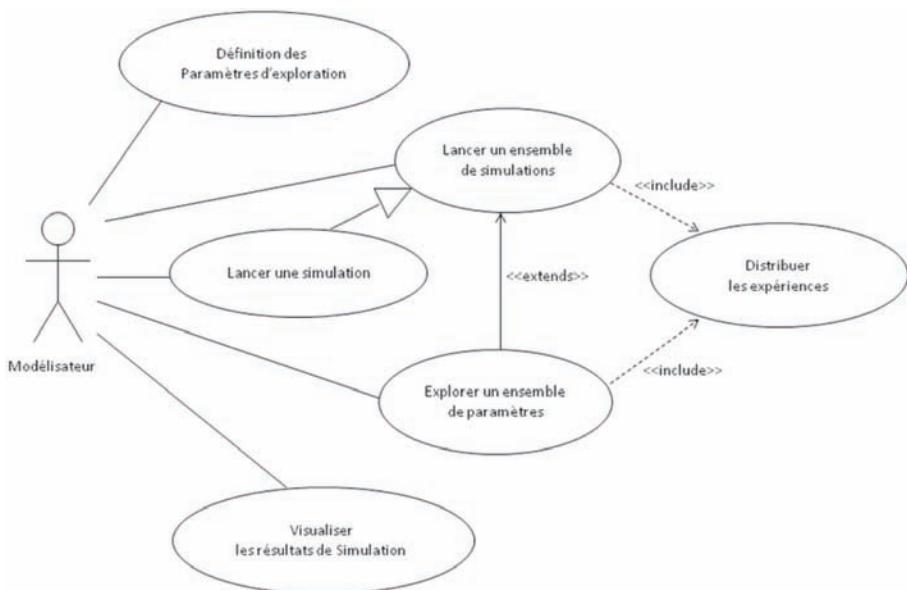


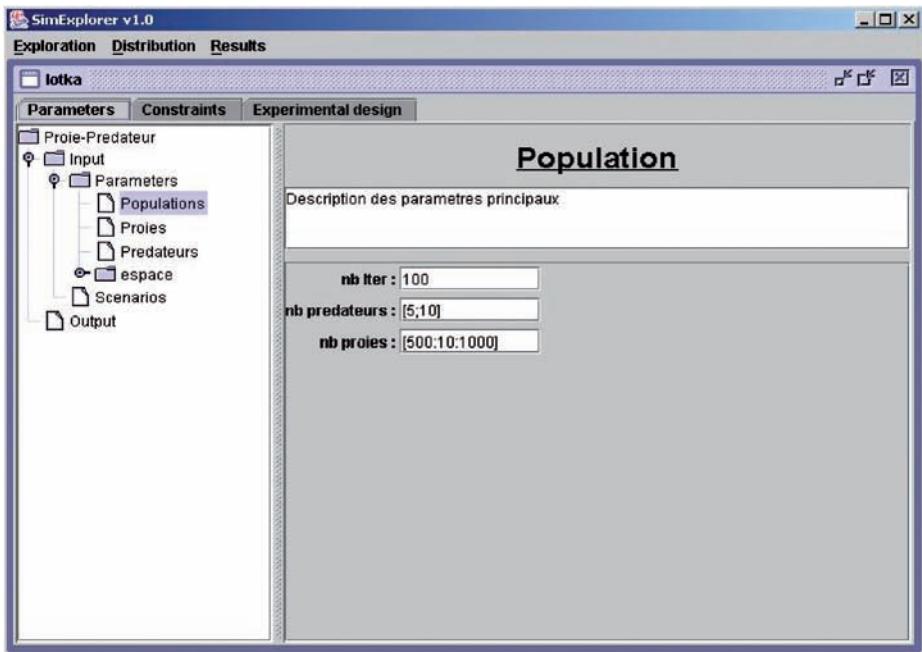
Figure 58. Use-case UML correspondant à ce que nous avions proposé en 2003 [Amblard *et al.* 2003]

Pour plus de souplesse une collection de composants a été proposée indépendamment des modèles de simulation, ces derniers étant considérés comme des entités paramétrables avec un ensemble de variables d'entrée et de sortie. Les fonctionnalités de bases étaient les suivantes :

- Initialisation d'un jeu de paramètre pour une seule simulation,
- Exécution de la simulation correspondante,
- Description d'une zone d'exploration pour les paramètres que l'on considère : en donnant un ensemble de valeurs pour chaque « facteur » selon la terminologie consacrée – l'ensemble des « niveaux » d'un facteur précise toutes les valeurs que pourra prendre un paramètre d'entrée du modèle,
- Choix d'une méthode ou plan d'exploration de l'espace des paramètres : plan complet (tous les facteurs sur tous les niveaux), plan de type Hyper Cube Latin (LHS : Latin Hyper-cube Sampling, plan D-Optimal, utilisation de d'algorithmes de « space filling » à base de nombres quasi-aléatoires. Les travaux de Jean-Pierre Gauchi au département MIA de l'INRA détaillent admirablement ces techniques (Unité de Mathématique et d'Informatique appliquée),
- Exécution de toutes les simulations correspondant à ce plan d'expérience sur une machine locale ou sur une ferme de calcul (cluster),
- Visualisation des résultats de simulation d'une façon synthétique pour pouvoir les interpréter (suivi de plusieurs traces, fouille de données et leur représentation visuelle).

157

L'implémentation initiale a été réalisée en Java et la description des expériences s'effectue grâce à des fichiers au format XML. Un exemple d'exécution pour un modèle proie prédateur est donné ci-après **Figure 59**. Grâce aux développements de Nicolas Dumoulin et Florent Chuffart sous la direction de Thierry Faure et Guillaume Deffuant, l'outil a sensiblement évolué et plus récemment dans le cadre d'un projet LifeGrid et d'un transfert de compétences du LIMOS, Romain Reuillon a pu intégrer à SimExplorer la possibilité de distribuer les expériences importantes sur grille de calcul. Les sources sont disponibles sur www.simexplorer.org sous licence GPL.



158

Figure 59. Exemple d'utilisation de SimExplorer précisant pour 2 facteurs une liste de niveaux

8. UNE APPLICATION À GRANDE ÉCHELLE DES PLANS D'EXPÉRIENCES

Dans la cadre d'un projet porté par l'INRA et financé par l'Agence Nationale pour la Recherche nous avons couplé par la modélisation deux démarches expérimentales disjointes : l'étude des effets de la diversité végétale à même gestion (Jena en Allemagne) et l'étude des effets de la gestion à même diversité initiale). Ce projet a pris en charge des études de diversité biologique, en particulier pour certaines interactions plante-sol (communautés microbiennes liées au cycle de l'azote et traits racinaires). Les expériences réalisées par les collègues biologistes fournissent des données uniques et permettent de paramétriser un modèle nommé GEMINI sur des monocultures grâce aux mesures de traits. Ces mesures permettent de calculer des paramètres qui peuvent être testés.

Grâce au travail de terrain des agronomes, des plans d'expériences complets pour ce modèle ont été mis en œuvre et simulés sur machine parallèle et les résultats ont été confrontés à l'observation. Au sein du LIMOS nous sommes intervenus principalement dans le développement du modèle GEMINI (présenté dans le paragraphe suivant) et de son déploiement sur calculateur parallèle pour des plans d'expériences à grande échelle. Le lancement de simulations en parallèle a induit le développement par le LIMOS de code spécifique pour la parallélisation des simulations (plus de 300 000 simulations !), pour le suivi et la vérification des résultats des simulations parallèles et pour la fouille des données.

Le modèle GEMINI a été développé en C++, sous la direction de Jean-François Soussana, Directeur de Recherche à l'INRA. Il représente environ 25 000 lignes de code dans la plate-forme de modélisation UNIF. Le développement a été principalement réalisé par Bruno Bachelet et des étudiants de l'ISIMA encadré par des chercheurs du LIMOS. Ce modèle propose une arborescence de classes permettant d'instancier une ou plusieurs populations végétales, couplées ou non à un sol et à des herbivores. Le modèle de végétation simule la croissance et la morphogénèse aérienne et souterraine de populations végétales, paramétrées par leurs traits fonctionnels et partageant des ressources (lumière, azote). Le modèle de sol décrit la dynamique de quatre pools de matière organique, sous l'effet de deux types fonctionnels microbiens de décomposeurs. La gestion de la prairie (coupe, fertilisation, pâturage) et le microclimat sont pris en charge par des modules additionnels. Le modèle GEMINI nous offre un cadre assez complet pour assimiler les données expérimentales et pour tenir compte de la plasticité des populations végétales et des interactions avec les organismes du sol. Après une phase de vérification très fine de ce logiciel (débogages finaux) et une finalisation du portage de ce modèle sous CygWin, nous avons réalisé un portage sous UNIX afin de pouvoir lancer à grande échelle des milliers de simulations.

159

Les plans d'expériences mis en œuvre avec le logiciel GEMINI sont principalement dédiés à l'étude de l'effet des traits sur la croissance des plantes. Les deux plans d'expériences complets que nous avons menés ont été conçus avec et principalement par Nicolas Gros également en collaboration avec Vincent Maire et Jean-François Soussana de l'INRA de Crouel (Puy-de-Dôme). Les questions principales auxquelles les biologistes cherchaient des réponses étaient les suivantes :

- Quels sont les impacts des traits morphologiques sur la croissance des plantes ?

- Quels sont les liens entre traits (existence de compromis), e.g. Comment ces liens jouent-ils sur la croissance des plantes ?

Les deux principales hypothèses théoriques à vérifier sont les suivantes :

1. Les valeurs de traits maximisant la croissance des plantes dans GEMINI sont celles observées dans la nature (déjà vérifié sur des analyses trait par trait),
2. Il existe des compromis entre traits, toutes les combinaisons entre traits ne sont pas équivalentes en terme de croissance. Si un compromis entre deux traits existe, alors une seule combinaison de valeur de traits par espèce maximise leur croissance. S'il n'y a pas de compromis, plusieurs valeurs maximisent la croissance des espèces.

Douze espèces paramétrées ont été retenues sur la collection de graminées présentée par Pontes en 2006 et quatre traits ont été sélectionnés en fonction du potentiel adaptatif des espèces : LDMC en relation avec la surface de feuille par gramme de matière sèche (ce trait est relié à la capacité des plantes à acquérir les ressources et à une croissance rapide), Ph0 en relation avec la durée de vie de feuille ainsi que D0 et L0 en relation avec la stature aérienne des plantes (compétition). Le croisement complet entre ces traits est géré par un plan complet avec 11 niveaux pour chaque facteur (les 4 traits dans notre cas). Une pré-étude nous a donné les temps de calculs pour 10 ans de simulation et estimée pour 1 trait et 10 niveaux sur les 12 espèces (**Table 1**). À partir de ces résultats, nous avons exclu l'espèce *Poa-tri* de l'expérimentation car le temps de calcul a été estimé trop long pour un lancement de plus de 10 000 simulations (supérieur à 10 heures pour une simulation – **Table 1**). Dans ce que nous avons pu mener à grande échelle, 10 pas de simulation par espèce ont pu être calculés avec la valeur observée du trait dans la collection (11 simulations par trait et par espèce). Au total, $11 \times 4 \text{ traits} \times 12 \text{ espèces} = 175\,692$ simulations ont été réalisées pour chaque plan complet. Les espèces ont été simulées en condition azote '+' et coupe '-' avec 3 coupes par an et des conditions expérimentales spécifiées par des fichiers. La durée d'une simulation est de 10 ans. Sur ces 10 ans, deux années de simulations se répètent 5 fois. Ces années correspondent aux données climatiques de 2003 et 2004 enregistrées sur la collection. Les simulations sont lancées à partir d'un état d'équilibre. Le fait de changer la valeur d'un paramètre en partant d'un « restart » déstabilise le modèle pendant les deux premières années de simulation. Pour chaque simulation, un fichier .csv rend compte des variables de sortie du modèle. Pour chaque simulation, la taille approximative par fichier est environ de 4 500 Ko pour

une simulation complète. Nous avons pris en compte la taille de stockage des résultats qui est très élevée. En effet, un plan d'expérience complet sur les 12 espèces avec 11 niveaux pour 4 traits produit plus de 800 Giga octets. Il est bien sûr possible de sélectionner les variables de sortie les plus intéressantes et de compresser les résultats. Sur les résultats, nous nous intéresserons particulièrement aux analyses suivantes des données :

- Relation entre traits et variation de traits : identification de valeur de traits maximisant la croissance des plantes,
- Comparaison avec traits observé dans la collection,
- Lien entre combinaison de traits et optimum : identification de trade-offs,
- Analyse de l'utilisation de l'azote RUE et NUE et de la lumière Ea et LAI en fonction des variations de traits.

Les deux campagnes de simulations effectuées au LIMOS ont pu réaliser 1 an de calcul effectif en seulement 30 jours sur un calculateur parallèle. La production de données s'élève à 1 Tera octets de données utilisables stockées sur SAN (Storage Area Network) du LIMOS. Le logiciel GEMINI s'est révélé extrêmement fiable avec plus de 99.99 % de succès sur les 340 000 simulations exécutées. Le transfert de ces données a été effectué grâce à un réseau à haut débit (Gigabits et à une connexion directe de disque externe en e-Sata). À titre d'exemple la simulation d'une seule espèce produit plusieurs dizaines de Giga octets (de 12 à 58 Go). Ces résultats ne peuvent être transférés avec de moyens classiques d'un centre à un autre. La Table 1 ci-après donne les tailles en octet des résultats produits. Il est possible de les compresser significativement (jusqu'à 70 % d'économie) avec des outils de compression capable de passer « à l'échelle », le travail sur micro-ordinateur est actuellement exclu pour ces tailles de fichiers. Dans ces conditions, environ une journée de transfert est nécessaire pour chaque disque de 500 Go. Un transfert sur le disque externe d'un micro-ordinateur à 20 ou 30 Mbits/s nécessiterait une trentaine de jours. Nous avons utilisé un transfert à plus haut débit en e-sata à 80 Mbits/seconde et livré 2 disques de 500 Go de données. L'idéal serait de disposer de disques externes et amovibles avec une technologie et un protocole aussi efficace que le Fiber Channel utilisé actuellement principalement pour les SANs (Storage Area Network).

Table 1. Exemple de taille des résultats, nombres de simulations réussies et en échec

Espèce	Taille des résultats en octets	Nombre de simulations en échec	Nombre de simulations réussies
<i>Alopra</i>	12 850 348 426	0	14 641
<i>Antodo</i>	24 594 380 072	2	14 639
<i>Arrela</i>	36 541 866 244	172	14 469
<i>Clerpin</i>	43 220 325 864	367	14 274
<i>Dacglo</i>	39 478 240 586	0	14 641
<i>Elyrep</i>	30 384 689 821	639	14 002
<i>Fesrub</i>	28 210 084 953	6	14 635
<i>Hollan</i>	46 388 547 776	14	14 627
<i>Lolper</i>	35 027 975 373	242	14 399
<i>Phlpra</i>	58 453 755 637	14	14 627
<i>Poppra</i>	39 532 041 766	17	14 624
<i>Trifla</i>	45 842 701 128	4	14 637

Pour chaque variable de sortie, nous nous intéressons à leur valeur maximum annuelle, à la valeur minimum annuelle ainsi que leur valeur moyenne annuelle. Nous avons développé un logiciel de calcul de la moyenne de ces valeurs pour les dix ans de simulations en excluant les deux premières années de simulations non stables. Il faut également exclure des fichiers de sortie les valeurs aberrantes issus des années non simulées ou non simulées entièrement. Ainsi, une valeur par variable de sortie sera obtenue correspondant à la valeur moyenne sur l'ensemble des dix ans de simulation. Une réflexion sur les variables de sortie intéressantes reste à planifier dans l'étude que nous avons menée. Si toutes les variables de sortie peuvent être enregistrées pour chaque simulation cette réflexion pourra être reportée après l'analyse des résultats. Les résultats préliminaires ont montré que les combinaisons des 4 traits qui sont trouvées idéales pour maximiser la croissance des plantes dans les simulations sont les combinaisons idéales observées sur le terrain pour des conditions identiques. La plasticité des traits observés pour des traitements N+/N- est aussi celle qui maximise la croissance des plantes avec les simulations.

9. CONCLUSION

Sachant que toute simulation informatique repose sur le codage dans un langage de programmation d'un algorithme ou d'un ensemble d'équations, il est important de se préoccuper des problèmes de vérification de code. La traduction du modèle dans un langage informatique n'est jamais parfaite et les techniques de transformation ont aussi leurs limites. Pour minimiser le nombre de bogues, il est important d'appliquer des techniques de *vérification* de code issues du Génie Logiciel. De très bons outils sont maintenant disponibles pour les langages C/C++ et Java. Il conviendra toujours de disposer, les techniques de vérification, qui s'attachent à la qualité du code informatique, des techniques de *validation* qui s'attachent à la qualité du modèle.

Pour nos travaux il était essentiel d'aborder les problèmes de validation. La validation est un processus complexe qui s'effectue pour un cadre d'expérimentation déterminé et avec des objectifs de modélisation précis. La validation utilise des techniques *empiriques, graphiques, statistiques*. Elle s'applique tout au long du processus de modélisation, dans les phases d'*acquisition* et d'*analyse des données*, d'*élaboration du modèle conceptuel* et, bien sûr, lors de l'*analyse des résultats*.

163

Nous nous sommes également attachés aux aspects visuels notamment avec l'animation graphique de résultats de simulation. L'animation se positionne, non seulement comme un outil de *présentation et de communication des résultats*, mais elle se révèle également être un très bon outil de *vérification* (mise au point des programmes) et de *validation* des modèles.

Nous avons aussi présenté l'utilisation d'analyses spatiales de résultats de simulation stochastiques. Cette technique permet d'étudier les corrélations spatiales en 2 et 3 dimensions, qui sont liées à la nature du site étudié [Legendre 1989]. De telles analyses sont possibles en couplant les modèles de simulation avec les Systèmes d'Information Géographique. Elles s'avèrent utiles dès lors que l'on peut supposer des corrélations spatiales dues aux particularités d'un site (substrats, bathymétrie,...). L'outil de simulation couplé au Système d'Information Géographique offre alors un laboratoire virtuel d'étude des aspects spatiaux propres aux sites étudiés. En effet, sur des sites réels, il est impossible d'observer ne serait-ce qu'une partie de l'espace des solutions, étant donné que sur un site unique un seul scénario se produira. Grâce aux simulations, il est possible de répliquer différents scénarios sur

le même site avec soit les mêmes conditions initiales, soit des conditions différentes. Cette technique d'analyse spectrale n'est pas sans poser des problèmes, difficiles et coûteux à résoudre. En effet, une solution peut consister à effectuer de multiples analyses spatiales (soit à différents instants d'une simulation stationnaire, soit pour les différentes réplications d'une simulation Monte Carlo), puis à soumettre les résultats de ces différentes analyses à un traitement statistique avant de pouvoir conclure. Le problème est encore plus complexe lorsque l'état du système converge vers plusieurs attracteurs, car il est alors nécessaire, pour avoir un résultat ayant une signification, de n'effectuer de traitements statistiques que sur des réplications correspondant au même attracteur.

Les derniers points abordés ont montré comment nous prenions en compte les plans d'expérience avec la proposition de logiciels tels que SimExplorer. Dans le cadre d'un projet ANR, nous avons montré comment deux campagnes de calcul on pu être menées pour l'étude de 4 traits de 12 espèces biologiques avec 11 niveaux en utilisant le logiciel Gemini. Plus de 340 000 simulations pour 2 plans d'expérience complets ont nécessité l'équivalent d'une année pleine de calcul pour un processeur actuel. Ces 2 campagnes ont été réalisées en un mois grâce à du calcul intensif et ont produit un Tera octet de données (plus de mille milliards d'octets). Des développements spécifiques ont permis la parallélisation des simulations, les tests de vérification et de validation et la fouille des données pour extraire des résultats pertinents de la masse des données produites.

CHAPITRE 6

CONCLUSION

*La science ne sert guère qu'à nous donner
une idée de l'étendue de notre ignorance.*

Félicité Robert Lamennais (1782-1854)

RÉFLEXIONS

Les modèles sont principalement conçus pour explorer la réalité et c'est à ce titre qu'ils sont des outils indispensables pour la recherche, notamment pour la compréhension de systèmes complexes non réductibles à la somme des parties qui les composent. La majorité des modèles que nous avons développés, que ce soit pour des systèmes à flux discrets ou pour des écosystèmes, étudiaient une multiplicité de facteurs. Ces modèles tentaient d'expliquer une proposition, une hypothèse, certains peuvent être qualifiés de prédictifs, ce sont d'ailleurs ces modèles de prévision et d'aide à la décision qui restent les plus demandés par la société. Ces derniers sont les plus difficiles à réaliser et peuvent ne pas aboutir même si les phases de vérification et de validation ont été réalisées correctement. Rappelons avant tout, que le but d'un modèle est d'être utile pour répondre à des questions qui consti-

tuent l'objectif que l'on s'est fixé, cet objectif ne doit pas être irréaliste et il guide le niveau de simplification que nous retenons. En effet, il faut parfois savoir renoncer au développement de modèles pour lesquels les objectifs apparaissent déraisonnables même si l'amélioration continue des techniques et du matériel informatique ne fait qu'accroître les potentialités des outils de simulation.

Dans le domaine des Sciences de la Vie, le caractère laborieux et forcément insuffisant du travail expérimental constitue souvent une limite au caractère prédictif des modèles. En effet, l'acquisition de données de qualité reste très coûteuse en ce qui concerne l'écologie, l'environnement, la santé, la biologie moléculaire, et la réalisation d'expériences sur le terrain ou « *in vivo* » est essentielle à la validation des modèles [Hill et Coquillard 2007]. D'une manière générale, lorsqu'un modèle est utilisé pour éviter toute expérimentation, il y a danger. Ces expériences sont plus précisément nécessaires pour alimenter et valider les modèles. L'expression anglo-saxonne « *garbage in, garbage out* » illustre bien l'importance de la qualité des données en entrée des modèles si l'on souhaite obtenir des résultats cohérents. En ce qui concerne les possibilités de prédictions de modèles tels que ceux que nous avons réalisés pour le suivi de l'expansion de l'algue *Caulerpa taxifolia* en Méditerranée, l'hypothèse de l'origine de cette algue était déjà fortement suggérée par les résultats du modèle, avant même qu'elle ne soit démontrée biologiquement. Toutefois il convient d'être toujours très prudent avec ce type de prédictions qui ne peuvent en aucun cas constituer des preuves [Hill et al. 2001]. Pour cette application de la modélisation à l'océanographie, plusieurs points de vue ont été retenus avec trois modèles différents qui ont tous entraîné de nouvelles expériences concrètes sur le terrain [Hill et al. 2002]. La confrontation des résultats de ces modèles a progressivement conduit à une meilleure connaissance du système étudié. Si d'une manière générale nous restons prudent sur les capacités de prédition de nos modèles, c'est principalement parce que l'une des particularités de la modélisation d'écosystèmes est qu'elle oblige à les considérer comme des systèmes isolés. Il faut en effet définir une fermeture du système, d'ouvert l'écosystème étudié devient isolé. L'opération de fermeture est sans doute l'opération la plus délicate lors de l'élaboration d'un modèle. Elle constitue dans le processus d'abstraction, l'opération de simplification la plus forte et corollairement la plus dangereuse. Il ne faut donc pas que le modèle soit pris comme une « *seconde nature* » car il déforme forcément la réalité. De même il faut être vigilant sur une utilisation abusive du modèle en dehors de son domaine de validité. La construction de plans d'expérience, l'emploi de techniques

de vérification et de validation a toujours été un de nos soucis. Nous avons essayé, dans la mesure du possible, de bâtir des modèles à partir des relations observables dans les systèmes réels afin de faciliter la validation par confrontation.

Revenons sur l'approche individu centrée qui a été retenue pour toutes nos applications des techniques de modélisation à l'écologie. Son principal avantage est de permettre des hypothèses plus réalistes que celles réalisées avec des modèles réductionnistes plus classiques (sans oublier que le but de la modélisation n'est pas le « réalisme » à proprement parler). Nous avons bien constaté sur nos travaux que le couplage de techniques de simulation discrète avec des Systèmes d'Information Géographique se marie également très bien avec l'approche individu centrée. Cette approche est ascendante elle se base uniquement sur l'étude des parties (individus) d'un système (population) pour chercher à comprendre comment les propriétés du système émergent des interactions entre ces parties. On observe une situation concrète, puis on tente de bâtir un modèle en fonction d'un objectif déterminé, modèle qui permettra peut-être de bâtir une théorie. Cependant, cette approche est plus contraignante, plus longue et plus fastidieuse à mettre en œuvre qu'une approche descendante qui ferait appel explicitement à l'écologie théorique, et qui construirait un modèle en partant de lois très générales avec un petit nombre d'hypothèses.

167

Ajoutons également que le processus de modélisation ne pourra aboutir de façon satisfaisante que dans le cadre d'une étroite collaboration entre les experts du domaine et le ou les réalisateurs techniques du modèle. L'expérience de Jean-Marie Legay nous confirme qu'il n'existe pas de problème en écologie ou en biologie dont une seule discipline soit venue à bout :

Ce qui est essentiel, c'est qu'un chercheur découvre la complexité de l'objet de sa recherche, c'est qu'un spécialiste d'une discipline refuse le ghetto de sa discipline en même temps que la position hiérarchique qui lui a été attribuée par rapport aux autres disciplines. En débordant par nécessité des limites prescrites à son domaine traditionnel, le chercheur se donne l'occasion de voir autrement et de rencontrer d'autres collègues qui ont fait la même démarche que lui à partir de leurs propres disciplines. Des zones d'intersection se forment dont on ne sait pas au départ quelles seront leur importance, leur durée et leur fécondité, mais qu'il faut protéger à tout prix. [Legay 1986]

QUELQUES APPORTS

Parmi les principaux apports des travaux d'intégration de techniques et d'outils pour la modélisation, nous tenons à rappeler le développement de couplages entre Simulateurs Aléatoires à Événements Discrets et Systèmes d'Information Géographique. Ce choix technique permet principalement de proposer des modèles de phénomènes discontinus dans le temps et dans l'espace. Les problèmes de performance ont été abordés et le compromis retenu pour tous nos modèles sous contrainte spatiale consiste à préparer des cartes spécifiques qui peuvent être interrogées et chargées en mémoire pendant l'exécution de la simulation. Une autre originalité de nos travaux réside dans l'étude des possibilités de simulation sur le Web. Dès le début du millénaire, nous avons pensé qu'il était intéressant d'investir dans l'adaptation d'un langage formel tel que DEVS pour décrire et exécuter des modèles sur des navigateurs de modèles.

Concernant les aspects validation, nous avons pu proposer une technique d'analyse spectrale de résultats de simulation stochastiques sous contraintes spatiales [Hill *et al.* 1996b]. Avec cette technique qui a été évoquée dans le chapitre précédent, les résultats ne sont plus seulement des valeurs pour lesquelles on peut donner des intervalles de confiance mais des séries de cartes similaires présentant des aspects stochastiques. L'analyse spectrale de résultats de simulations stochastiques en deux dimensions constitue à notre point de vue une avancée intéressante. Avec nos collègues biologistes, nous avons pu, par exemple, obtenir des cartes de probabilité de colonisation par l'algue *Caulerpa taxifolia* qui ont pu être validées sur des sites méditerranéens à différentes échelles spatiales [Hill 1997b]. Cette technique a également été appliquée en écologie terrestre.

Lorsque les temps d'exécution de modèles de Simulation Aléatoire à Événements Discrets étaient trop longs, par exemple dans le cas de la simulation de l'expansion de l'algue *Caulerpa taxifolia* sur des sites de taille importante et sur des périodes de temps dépassant une dizaine d'année, nous avons proposé une technique de métamodélisation par des réseaux de neurones. Nous avons utilisé un réseau de neurones traditionnel à plusieurs couches, avec un algorithme de rétro-propagation standard. L'inconvénient majeur des réseaux de neurones que nous ne devons pas cacher, restera qu'ils ne permettent pas une prédiction de la répartition spatiale de l'algue.

Concernant les approches individu centrées, elles montrent leurs limites lorsque l'on souhaite traiter de grands nombres d'individus (dans notre cas des dizaines de milliers, voire des centaines de milliers d'individus). Nous avons pu développer une technique d'optimisation du codage des modèles sous contrainte spatiale ; en effet grâce à ces contraintes il est possible pour les processus de colonisation et en fonction des objectifs de la simulation de n'activer qu'un nombre restreint d'individus sur le nombre total d'individus à simuler. Cette approche a été appliquée en écologie terrestre et en océanographie. De plus, confronté aux problèmes de la gestion de zones importantes de mémoire dynamique par les compilateurs C++, nous avons proposé une technique de substitution qui utilise la mémoire statique pour toute allocation d'objet. Sans cette technique, certains de nos modèles satureraient la mémoire virtuelle de nos machines.

Une autre particularité technique de nos travaux réside dans le fait que nous avons proposé et utilisé un noyau de synchronisation adapté à la gestion de processus stochastiques simultanés qui se retrouvent en compétition pour des ressources spatiales. Ce noyau de synchronisation, d'une part, gère les problèmes de parallélisme liés aux compétitions spatiales et, d'autre part, évite le brassage entre le générateur de nombres pseudo-aléatoires du simulateur et celui du système d'exploitation, ce dernier devant parfois faire un choix entre plusieurs processus légers qui se trouvent être potentiellement activables « simultanément ». Ce brassage n'est pas acceptable pour une étude statistique des résultats de simulation. Nous avons également abordé les problèmes d'optimisation et de distribution de nombres pseudo-aléatoires sur plusieurs processeurs dans le cadre de simulations parallèles [Hill 2002]. Même si de nombreuses techniques existent pour essayer d'éviter les auto-corrélations entre les flux aléatoires utilisés sur les différents simulateurs (répartis sur des processeurs différents). Il n'est pas envisageable de fournir des résultats statistiques de qualité pour des simulations stochastiques parallèles sans prêter attention à ces aspects [Hill 1997c] et nous avons proposé des solutions pour mettre en œuvre les meilleures techniques de distribution de flux stochastiques [Reuillon *et al.* 2008].

169

Enfin, tous les développements de nos modèles ont utilisé des techniques du génie logiciel à objets qui sont désormais classiques que ce soit au niveau des techniques d'implémentation ou au niveau des techniques de conception (patrons et cadriel). L'expérience acquise lors du développement de plusieurs applications, dans des domaines passant de l'écologie terrestre à l'océanographie, nous a permis de proposer un cadriel simple pour la simulation

d'écosystèmes avec une approche individu centrée [Campos et Hill 1998a]. Nous avons également utilisé des patrons de conception connus pour bâtir un multimodèle en océanographie [Hill *et al.* 2000a]. Les multimodèles permettent de simuler des systèmes complexes où les différentes parties ne se situent pas au même niveau d'abstraction et pour lesquelles les techniques de modélisation ne sont pas forcément les mêmes.

DERNIERS TRAVAUX ET PERSPECTIVES

Les techniques de simulation aléatoires sous contraintes spatiales s'appliquent dans des domaines variés et nous avons donc une technique assez générique pour une classe de problèmes où les interactions spatiales discrètes et à distance sont importantes. Après avoir travaillé pour la simulation de croissance des plantes [Lafarge *et al.* 2005] [Mazel *et al.* 2005] et également celle des forêts [Prévosto *et al.* 2003], nous avons commencé à aborder les problèmes liés aux changements climatiques [Robert *et al.* 2005]. Dans ce contexte nous envisageons de nombreux travaux avec des plans d'expériences à grande échelle comme ceux que nous avons décris partiellement dans cet ouvrage et qui ont été réalisés dernièrement dans le cadre du projet Discover projet pour l'Agence Nationale pour la Recherche.

170

Depuis quelques années nous travaillons également avec des physiciens sur la simulation de la propagation de feux de forêts [Muzy *et al.* 2005]. Dans ce contexte la prise en compte de phénomènes aléatoires tels que la dispersion de tisons est un facteur important pour suivre la propagation d'un feu (**Figure 60**). Les travaux préliminaires conduisent déjà à des résultats significatifs qui ne sont pas obtenus par des méthodes plus classiques de mécanique des fluides ou à des techniques plus classiques liées à des processus de diffusion. Nous avons aussi développé un outil rapide de visualisation en 3D avec des techniques qui s'apparentent aux jeux vidéo pour un suivi sur une topographie (**Figure 61**) [Muzy *et al.* 2008]. Nous avons pour ambition de paralléliser les simulations pour essayer de suivre de la manière la plus réaliste possible des feux modélisés à une échelle fine par des physiciens, le but étant d'être plus rapide que le temps réel de propagation d'un feu à petite échelle. Nous avons pu obtenir des résultats intéressants dans des travaux récents adaptés aux nouvelles architectures d'ordinateur [Innocenti *et al.* 2009].

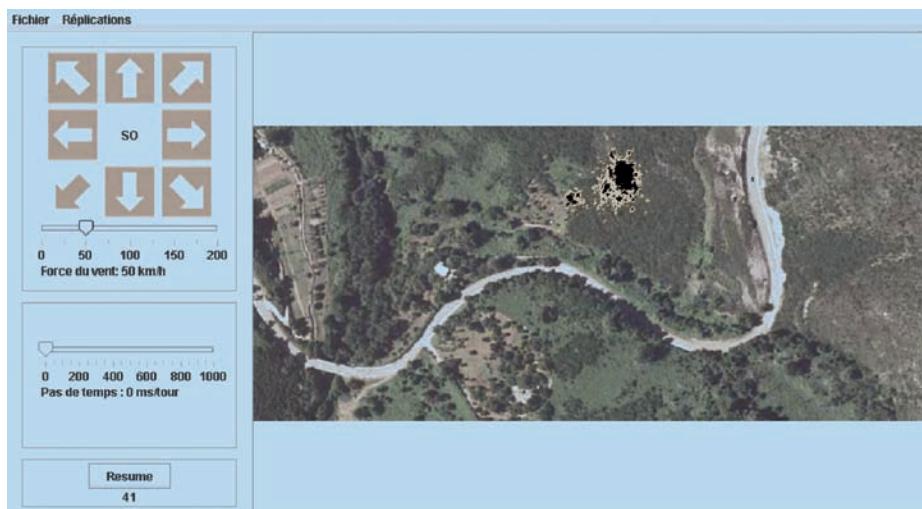


Figure 60. Exemple de suivi et de simulation de propagation de feu avec tison en Corse

Après une formation à la biologie moléculaire, nous avons aussi considéré l'intégration d'outils logiciels dans un domaine proche de l'écologie et de l'environnement : la biologie cellulaire et moléculaire. Nous avons également abordé des problèmes de simulation pour la médecine nucléaire, notamment dans l'amélioration de la reconstruction d'images de scanners [Lazaro *et al.* 2005]; [El Bitar *et al.* 2006]. Nous continuons ces recherches en collaboration avec le Dr. El Bitar de l'IN2P3 de Strasbourg. Toujours dans le domaine de la santé, nous travaillons aussi à la prévision par simulation de la propagation interhumaine du virus de la grippe aviaire en utilisant les données disponibles pour la Corse (Figure 61). La transmission interhumaine du virus H5N1 n'a toujours pas été prouvée dans les récentes études, même si ils sont très rares on observe cependant une multiplication récente (2009) des cas humains. Les scientifiques spécialistes ont démontré qu'une éventuelle propagation interhumaine serait bien sûr facilitée par les transports aériens [Mangili et Gendreau 2005]. De plus, [Ferguson *et al.* 2005], [Doyle *et al.* 2005], [Longini *et al.* 2005] et [Colizza *et al.* 2007], avaient proposé des modèles permettant de prédire l'impact d'une propagation de grippe aviaire. À partir de l'étude de ces différentes publications, nous avions alors pu définir des choix de modélisation pour un simulateur stochastique spatial de grippe aviaire en Corse [Hill *et al.* 2008] (Figure 62).

172

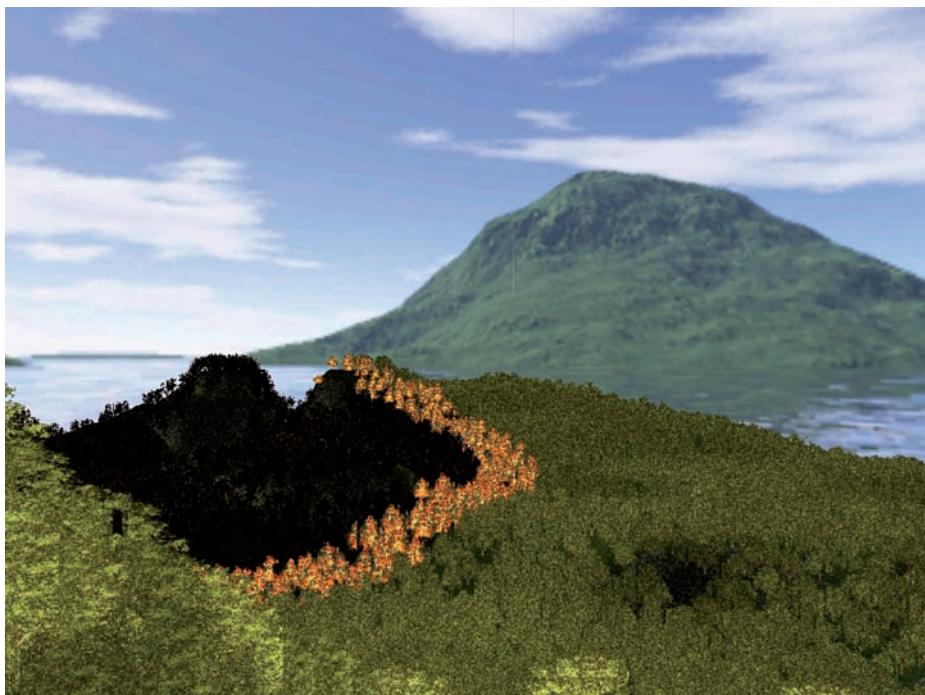


Figure 61. Exemple de rendu presque réaliste avec des techniques graphiques rapides et élémentaires adaptées à une visualisation sur tout type d'ordinateur personnel

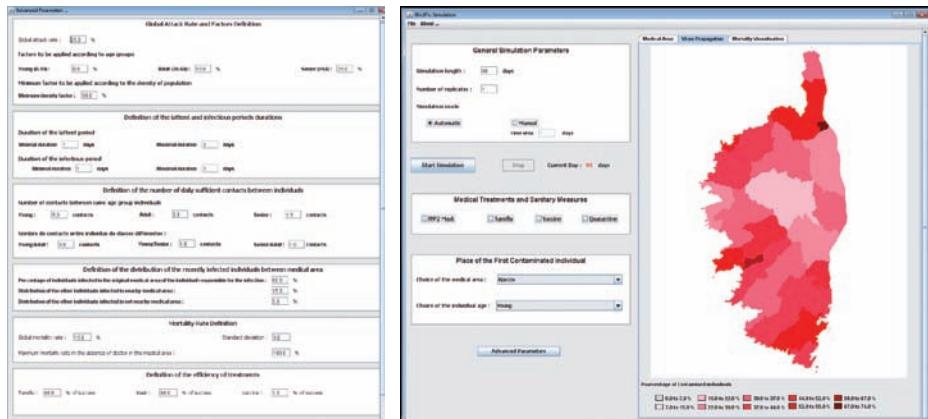


Figure 62. Exemple de visualisation de simulation de la propagation du virus de la grippe aviaire en Corse

La bio-informatique se consacre principalement à l'étude des génomes avec des outils logiciels et la modélisation y joue aussi un rôle important pour l'analyse et l'exploitation de ces données, leur compréhension afin d'en valoriser la signification. Outre les techniques de *data mining* et de classification d'images, nous appliquons des techniques de simulation pour comprendre et tester les mécanismes de la cancérisation, et notamment étudier les mécanismes de résistance aux traitements des cellules tumorales. Nous avons depuis plusieurs années la possibilité de réaliser des biopuces en relation avec le Professeur Peyret. L'approche que nous retenons à chaque fois est pluri- et interdisciplinaire : elle repose sur de fortes collaborations et des doubles compétences. Nous nous sommes spécialisés dans la conception de petites séquences d'ADN appelées oligonucléotides qui permettent d'identifier de manière précise les gènes présents ou plus ou moins exprimés sur des biopuces [Rimour *et al.* 2005], [Militon *et al.* 2007]. Dans ce domaine très ouvert, de nombreux algorithmes sont encore à proposer et nous avons travaillé sur la proposition d'une technique très efficace pour faire de la traduction complète de petits oligopeptides [Hill 2006], [Missaoui *et al.* 2008]. En combinant ce type d'algorithme avec des techniques de biopuces fonctionnelles et phylogénétiques, nous pouvons accélérer la découverte de micro-organismes.

173

Enfin, il est envisageable de faire un pont entre les échelles (de la molécule à l'organe) et d'utiliser la Simulation Aléatoire à Événements Discrets sous contraintes spatiales. Dans le cas des cancers, la multiplication des cellules, l'envahissement de l'organe auquel elles appartiennent, leur migration par la circulation sanguine ou lymphatique pour produire des métastases, se trouvent être analogues à des modèles de colonisation du type de ceux que j'ai réalisés pour la propagation de l'algue *Caulerpa taxifolia*. Or, depuis l'apparition de ce problème en 1984, seules les techniques de simulation sous contraintes spatiales que nous avons proposées ont pu aboutir à des cartes de probabilité de colonisation validées par les océanographes sur de nombreux sites méditerranéens. Les composants logiciels et le modèle que je souhaite développer devraient, d'une part, aider à identifier les gènes qui interviennent dans le dérèglement du fonctionnement cellulaire et, d'autre part, aider à identifier les interactions entre ces gènes. L'expérience que nous avons acquise en éthologie sur des systèmes multi-agents [Dumont et Hill 2001], [Force *et al.* 2002], [Dumont et Hill 2004], nous amène à envisager pour ce type d'application des simulations multi-agents multi-échelles que nous avons mises en œuvre dans un contexte de simulation distribuées pour la SNCF [Feillee *et al.* 2008]. Un nombre croissant d'applications utilisent

le calcul intensif, nous continuerons à utiliser la grille de calcul Européenne EGEE comme nous l'avons fait dans [El Bitar *et al.* 2006] et [Reuillon *et al.* 2008]. Comme dans ces derniers travaux, nous attacherons une importance toute particulière à la distribution des flux stochastiques pour les processus parallèles ou pour les réplications des simulations aléatoires.

Nous espérons que ce tour rapide et vulgarisé, autant que possible, pour des collègues non spécialistes en Simulation Aléatoire à Événements Discrets, aura donné une idée du potentiel de ces méthodes et donné envie de développer des collaborations inter- et pluridisciplinaires.

Pour tout chercheur, par définition spécialisé dans un domaine, il convient souvent de prendre du recul, de réfléchir et d'étudier de manière critique « sa science » en tant que telle. Le produit de nos recherches en simulation, de nos discours sur la modélisation, sur nos domaines de recherche appliquée est provisoire, il faut à juste titre s'interroger sur les conditions dans lesquelles se forment nos connaissances, sur la cohérence des principes qui les gouvernent ; sur l'ajustement des méthodes aux objets que nous étudions, sur le fondement de nos déductions et de nos interprétations des résultats produits par nos simulations.

Quelle que soit l'approche préconisée et les connaissances élaborées, celles-ci gardent toute leur valeur lorsqu'elles restent dans un même contexte, avec des définitions qui cherchent tant bien que mal à être en adéquation avec la réalité. Nos études se fondent sur des concepts qui semblent s'imposer de manière évidente, mais qui en fait ne sont ni simples, ni triviaux à étudier car ils dépassent (metha) l'observable (physique). Dans le domaine de la simulation, avec une attention particulière pour la simulation informatique dédiée aux Sciences de la Vie, un jeune philosophe, historien, ingénieur et maintenant enseignant-chercheur pousse avec bonheur notre réflexion vers plus d'épistémologie [Varenne, 2007].

ANNEXE

RÉFÉRENCES BIBLIOGRAPHIQUES

*Songez que les ouvrages que nous feuilletons le moins,
avec le plus de négligence et de partialité,
ce sont ceux de nos collègues.*

Denis Diderot (1713-1784)

*En vérité, je vous le dis,
aucun prophète n'est bien reçu dans sa patrie.*

Luc, 4,24

A.

[Acvedo 1981] – Acvedo M. F., « On Horn's markovian model of forest dynamics with particular reference to tropical forests », *Theoretical Population Biology*, Vol. 19, pp. 230-250.

[Agha 1990] – Agha G., « Concurrent Object-Oriented Programming », *Communication of the ACM*, Vol. 33, N°9, pp. 125-141.

[Ågren et Bosatta 1996] – Ågren G. I et Bosatta B., *Theoretical Ecosystem Ecology - Understanding Element Cycles*, Cambridge Univ. Press, 233 p.

[Alfonseca et al. 1999] – Alfonseca M., de Lara J. et Pulido E., « Semiautomatic Generation of Web Courses by Means of an Object-Oriented Simulation

- Language », in : Fishwick P. et Hill D. (éds.), *Simulation, Special issue in Web-based Simulation*, 73(1), pp. 5-12.
- [Amblard et al. 2003] – Amblard F., Hill D., Bernard S., Truffot J. et Deffuant G., « MDA compliant Design of SimExplorer, A Software to handle simulation experimental frameworks », in : *Proceedings of SCSC 2003 Summer Simulation Conference*, Montréal, July 20-24, pp. 279-284.
- [Arnold et Dudzinski 1978] – Arnold G. W. et Dudzinski M. L., *Ethology of free-ranging domestic animals*, Elsevier.
- [Aronoff 1989] – Aronoff H. E., « Calculating fire size and perimeter growth », *Fire Management Notes*, Vol. 45, pp. 25-29.
- [Atlan 1979] – Atlan H., *Entre le cristal et la fumée, Essai sur l'organisation du vivant*, Paris, Seuil, « Points-Sciences ».
- [Attoui et Hill 1995] – Attoui A. et Hill D., « A Specification and Validation Method to Improve Concurrent Systems Reliability Based on Object Messaging and Rewriting Logic » (+ article et poster), *International Symposium on Software Reliability Engineering (Toulouse, 24-27 Octobre)*, Toulouse, IEEE / CNRS, pp. 387-392.
- [Aussem et Hill 1999] – Aussem A. et Hill D., « Wedding connectionist and algorithmic modelling: towards forecasting *Caulerpa taxifolia* development in north-western Mediterranean sea », *Ecological modelling*, Vol. 120, pp. 225-236.
- [Aussem et Hill 2000] – Aussem A. et Hill D., « Neural networks metamodelling for the prediction of *Caulerpa taxifolia* development in the Mediterranean sea », *Neurocomputing Letters, Neurocomputing*, Vol. 30, pp. 71-78.
-
- B.**
- [Bailey et al. 1998] – Bailey D. W., Dumont B. et Wallis de Vries M. F., « Utilization of heterogeneous grasslands by domestic herbivores, theory to management », *Ann. Zootech.*, 47, pp. 312-333.
- [Balci 1994] – Balci O., « Validation Verification and Testing Techniques throughout the life cycle of a simulation study », *Annals of Operation Research*.
- [Balci et Sargent 1981] – Balci O. et Sargent R., « A Methodology for Cost-Risk Analysis in the Statistical Validation of Simulation Models », *Communication of the ACM*, Vol. 24, n° 4, pp. 15-29.
- [Balci et Withner 1989] – Balci O. et Withner R. B., « Guidelines for selecting and using simulation model verification techniques », *Winter Simulation Conference*, pp. 559-568.

- [Barazer 1989]** – Barazer P., *Le Calcul Intensif*, Paris, Belin, « Bibliothèque pour la Science », 159 p.
- [Barbier et Bézivin 1993]** – Barbier F. et Bézivin J., « Object-Oriented Design: the OSM scheme », *Proceeding of TOOLS USA 93*, Santa Barbara, Prentice Hall, pp. 57-68.
- [Bates 1994]** – Bates J., « The role of Emotion in Believable Agents », *Communications of the ACM*, vol. 37, n° 7, pp. 122-125.
- [Batty et Howes 1996]** – Batty M. et Howes D., « Exploring Urban Development Dynamics through Visualisation and Animation », publié in : **[Parker 1996]**, pp. 149-161.
- [Baveco et Smeulders 1994]** – Baveco J. M. et Smeulders A. M. W., « Objects for Simulation: Smalltalk and Ecology », *Simulation*, vol. 62, n° 1, pp. 42-57.
- [Bazjnac 1976]** – Bazjnac V., « Interactive Simulation of building evacuation with elevators », in : *Proceedings of the 9th Annual Simulation Symposium (March, Florida, USA)*, pp. 15-29.
- [Bell 1969]** – Bell T. E., *Computer graphics for Simulation-Problem solving*, Santa Monica (California), Rand Corporation.
- [Bell et O'Keefe 1987]** – Bell P. C. et O'Keefe R. M., « Visual Interactive Simulation: History, recent developments and major issues », *Simulation*, vol. 49, n° 3, pp. 109-116.
- [Bellan-Santini et al. 1996]** – Bellan-Santini D., Arnaud P. M., Bellan G. et Verlaque M., « The influence of the introduced tropical alga *Caulerpa taxifolia*, on the biodiversity of the Mediterranean marine biota », *J. Mar. Biol. Ass.*, U.K., 76, pp. 235-237.
- [Berg 1993]** – Berg H. C., *Random walks in biology*, Princeton Univ. Press, 142 p.
- [Berthold et al. 1993]** – Berthold M. R. et Huber Kl.-P., « Neural Network based Construction of Fuzzy Graphs », in : *Proceedings of the 2nd Annual Conference on Information Science*, North Carolina (1995).
- [Bézivin 1987]** – Bézivin J., « Some Experiments in Object-oriented Simulation », *Proceedings of OOPSLA*, 87, pp. 394-404.
- [Bézivin 1995]** – Bézivin J., « Technologie objet et ingénierie des besoins : une réconciliation nécessaire », *L'Objet*, vol. 1, n° 1, pp. 21-26.
- [Bohem 1988]** – Bohem B., « A spiral model of software development and enhancement », *IEEE Computer*, vol. 21, n° 5, pp. 61-72.
- [Booch 1996]** – Booch G., « The Unified Modeling Language », *Unix Review*, December, pp. 41-48.

- [Booch et al. 2000]** – Booch G., Rumbaugh J. et Jacobson I., *Le Guide de l'utilisateur UML*, Eyrolles, 534 p.
- [Bormann et Likens 1979]** – Bormann F. H. et Likens G. E., « Catastrophic disturbance and the steady state in northern hardwood forest », *American Scientist*, vol. 67, pp. 660-669.
- [Botkin 1993]** – Botkin D. B., *Forest Dynamics, An Ecological model*, Oxford, Oxford University Press.
- [Botkin et al. 1972]** – Botkin D. B., Janak J. K. et Wallis J. R., « Some ecological consequences of a computer model of forest growth », *Journal of Ecology*, vol. 60, pp. 849-872.
- [Boudouresque et al. 1994]** – Boudouresque C. F., Meinesz A. et Gravez V. (éds.), *First International Workshop on Caulerpa taxifolia*, GIS Posidonie (éd.), France, 391 p.
- [Bousquet et al. 1996]** – Bousquet F., Dutheoit Y., Proton H. et Weber J., « Tragedy of the Commons, Game Theory and Spatial Simulation of Complex Systems », *Ecological Economics*, May, pp. 1-8.
- [Box et Draper 1987]** – Box G. E. P. et Draper N. R., *Empirical Model Building and Response Surface*, New York, John Wiley.
- [Breckling et Müller 1994]** – Breckling B. et Müller F., « Current trends in ecological modelling and the 8th ISEM conference on the state-of-the-art », *Ecological Modelling*, vol. 75, pp. 667-675.
- [Breugnot et al. 1990]** – Breugnot D., Gourgand M. et Kellert P., « SIGMA: An intelligent and graphical environment for the modelling of flexible assembly-systems », in : *Proceedings of the European Simulation Symposium, (Ghent, Belgium, November 8-11)*, pp. 225-230.
- [Breugnot et al. 1991a]** – Breugnot D., Gourgand M., Hill D. et Kellert P., « Object-Oriented Animation for Flexible Manufacturing Systems Simulation Results », in : *Proceedings of TOOLS 4, Paris, France*, pp. 283-295.
- [Breugnot et al. 1991b]** – Breugnot D., Gourgand M., Hill D. et Kellert P., « GAME: An Object-Oriented Approach to Computer Animation and Flexible Manufacturing System Modelling », in : *Proceedings of the 24th Annual Simulation Symposium*, New Orleans, IEEE / ACM / SCS, pp. 217-227.
- [Breugnot et al. 1991c]** – Breugnot D., Gourgand M. et Kellert P., « Les problèmes de terminologie dans la modélisation des systèmes de production », *Conférence « Terminology work in Subject Fields », TTC 91, 12-14 Nov, Vienne, Autriche*.
- [Briot 1989]** – Briot J. P., « ‘Des Objets aux Acteurs’, 1982-1989 : 7 ans de réflexion », *LITP Research Report*, n° 89-68, Université de Paris VI, 1989.

[Bruzzone et al. 1999] – Bruzzone A., Uhrmacher A. et Page E., *Proceedings of the 2nd International Conference on Web-based Modeling & Simulation (San Francisco, CA, janvier 17-20)*.

[Brzeziecki et al. 1994] – Brzeziecki B., Kienast F. et Wildi O., « Potential impacts of a climate change on the vegetation cover of Switzerland: a simulation using GIS technology », in : Price M. F. et Heywood D. I. (Taylor & Francis), *Mountain Environment & Geographic Information Systems*, pp. 263-279.

[Bugmann et al. 1996] – Bugmann H. K. M., Yan X., Sykes M. T., Martin P., Lindner M., Desanker P. V. et Cumming C. G., « A comparison of forest gap models: model structure and behaviour », *Climatic change*, 34, pp. 289-313.

[Bulmer 1994] – Bulmer M., *Theoretical Evolutionary Ecology*, Sunderland, Sinauer, 352 p.

C.

[Campos 2000] – Campos A., *Une architecture logicielle pour le développement de simulations visuelles et interactives individu-centrés : application à la simulation d'écosystèmes et à la simulation sur le Web*, Thèse de doctorat, Université Blaise-Pascal Clermont-Ferrand II, 8 Septembre 2000.

179

[Campos et Hill 1998a] – Campos A. et Hill D., « An Agent Based Framework for Visual-Interactive Ecosystem Simulations », *Transactions*, SCS, December, vol. 15, n° 4, pp. 139-152.

[Campos et Hill 1998b] – Campos A. et Hill D., « Web-based Simulation of Agent Behaviors », *SCS International Conference on Web-based Modeling and Simulation*, 11-14 Janvier, San Diego (USA), pp. 9-14.

[Campos et al. 2000] – Campos A., Corbara B. et Hill D., « Using web-based facilities in remote simulation for agent behavior prototyping and result analysis », *International Conference on Web-based Modeling and Simulation*, R. Signorile, Curtis Blais, January 23-27, San Diego, USA, pp. 69-74.

[Caswell 1989] – Caswell J., *Matrix Population Models*, Sunderland, Sinauer, 328 p.

[Caux et al. 1991] – Caux C., Gourgand M. et Hill D., « Petri Net Simulation and Animation in a Graphical Object-Oriented Environment », in : *Proceedings of the ISSM Parallel an Distributed Systems Conference (Trany Italy)*, pp. 359-362.

[Cellier 1991] – Cellier F. E., *Continuous System Modeling*, Springer Verlag.

- [Clark 1992] – Clark K. B., « Plant-like animals and animals-like plants: symbiotic coevolution of ascoglossan (= sacoglossan) molluscs, their algal prey, and algal plastids », in : Biopress limited (éd.), *Algae and symbioses: plants, animals, fungi, viruses, interactions explore*, Reisser W. (publ.), pp. 515-530.
- [Clergue 1997] – Clergue G., *L'Apprentissage de la complexité*, Paris, Hermès.
- [Colizza et al. 2007] – Colizza V., Barrat A., Barthelemy M., Valleron A.-J. et Vespignani A., « Modeling the world-wide spread of pandemic influenza: baseline case and containment interventions », *PLoS Medicine*, vol. 4, n° 1 :e13, 23 janvier 2007.
- [Combes et al. 1994] – Combes C., Force C. et Kellert P., « Méthodologie d'élaboration générique d'un modèle de connaissance pour les systèmes hospitaliers », *Revue Informatique et Santé*, Mars, n°16, pp. 44-50.
- [Coquillard et al. 1995] – Coquillard P., Hill D. et Gueugnot J., « Simulation d'Écosystèmes et Systèmes d'Informations Géographiques : une interactivité nécessaire », in : *Actes de la conférence Nationale des Parcs Naturels de France (18 Janvier, École des Mines de St-Etienne)*, pp. 39-46.
- [Coquillard et al. 1996] – Coquillard P., Gueugnot J., Hill D. et Mahy G., « Une nouvelle classe de simulateurs à contrainte spatiale », in : *Actes du Colloque CNRS « Tendances nouvelles en modélisation pour l'environnement »*, Elsevier Éd., pp. 255-265.
- [Coquillard et al. 1997] – Coquillard P., Hill D., Mazel C. et Gueugnot J., « Application de l'analyse spectrale à l'étude de la répartition spatio-temporelle d'événements générés par simulation », in : *Les Temps de l'Environnement, Journées CNRS du PIREVS (5-7 novembre, Toulouse)*, pp. 455-463.
- [Coquillard et al. 1999] – Coquillard P., Hill D., Vaugelas J., de, Meinesz A., « Modelling & simulating Caulerpa taxifolia (Vahl) C. agardh in the north-western mediterranean sea : results and perspectives », United Nations European Program proceedings of the workshop on Invasive Caulerpa species in the mediterranean, Heraklion, Crete (Greece), 18-20 march 1998, MTS; 125, UNEP Athens, pp. 159-175.
- [Coquillard et al. 2000] – Coquillard P., Thibaut T., Hill D., Gueugnot J., Mazel C. et Coquillard Y., « Simulation of the mollusc Ascoglossa Elysia subornata population dynamics: application to the potential biocontrol of Caulerpa taxifolia growth in the Mediterranean Sea », *Ecological modeling*, accepted, in press, Réf : ECOMOD 2668.
- [Coquillard et Hill 1995] – Coquillard P. et Hill D., « Object-oriented Simulation of Scots Pine Growth interactions », in : McLeod J., *Simulation in the Service of Society*, Simulation vol. 65, n° 6, December, p. 411 (sélection d'une page d'un article de la *Summer Simulation Conference [1995, July 24-26, Ottawa, Canada]*), pp. 917-923.

[Coquillard et Hill 1997] — Coquillard P. et Hill D., *Modélisation et Simulation des Écosystèmes*, PARIS, Masson, 273 p.

[Cottalorda et al. 1996] — Cottalorda J. M., Robert P., Charbonnel E., Dimeet J., Menager V., Tillman M., Vaugelas J. de, Volto E., « Éradication de la colonie de *Caulerpa taxifolia* découverte en 1994 dans les eaux du Parc National de Port-Cros (Var, France) », in : Ribera M. A., Ballesteros E., Boudouresque C. F., Gomez A. et Gravez V. (éds), *Second International Workshop on Caulerpa taxifolia*, Barcelone, Publicacions Universitat Barcelona, pp. 149-156.

[Cox 1986] — Cox B. J., *Object-Oriented Programming, An Evolutionary Approach*, Addison Wesley.

[Cybenko 1989] — Cybenko G., « Continuous value neural networks with two hidden layers are sufficient », *Math. Control Signals and Systems*, vol. 2, pp. 303-314.

D.

[Dahl et al. 1966] — Dahl O. J., Myrhaug B. et Nygaard K., « Simula: An Algol Based Simulation Language », *Communication of the ACM*, vol. 9, n° 9, pp. 671-678.

181

[De Angelis et Gross 1992] — De Angelis D. et Gross L. J. (éds.), *Individual-based models and approaches in ecology: populations, communities, and ecosystems*, New York, Chapman and Hall.

[De Reffye et al. 1990] — Reffye Ph. de., Diounard P. et Jaeger M., « Basic concepts of computer plants growth simulation », *Nicograph*, 90, Seminar 9 Tokyo, pp. 219-234.

[Doram 1997] — Doram J., « From Computer Simulation to Artificial Societies », *TRANSACTIONS of SCS*, vol. 14, n° 2, pp. 69-77.

[Doyle et al. 2005] — Doyle A., Bonmarin I., Lévy-Bruhl D., Le Strat Y., Desenclos J. C., « Estimation de l'impact d'une pandémie grippale et analyse de stratégies », Institut de Veille Sanitaire, 17 février.

[Drogoul 1993] — Drogoul A., *De la Simulation Multi-Agents à la Résolution Collective de Problèmes*, Ph.D. Thesis, Université de Paris VI.

[Drogoul 2000] — Drogoul A., *Systèmes mullet-agents situés*, Dossier d'Habilitation à Diriger des Recherches, Université Pierre et Marie Curie, Laboratoire d'Informatique de Paris 6.

[Dumont 1996] — Dumont B., « Préférences et sélection alimentaire au pâturage », *INRA Productions Animales*, 9, pp. 359-366.

- [Dumont et Hill 2001] – Dumont B. et Hill D., « Multi-agent simulation of group foraging in sheep: effects of spatial memory, conspecific attraction and plot size », *Ecological Modelling*, vol. 141, pp. 201-215.
- [Dumont et Hill 2004] – Dumont B. et Hill D., « Spatially explicit models of group foraging by herbivores: What can Agent Based Models offer? », *Animal research (anc. Annales de Zootechnie)*, vol. 53, pp. 419-428.

E.

- [Edwards et Henderson-Sellers 1990] – Edwards M. et Henderson-Sellers B., « Object-oriented Systems Life Cycle », *Communication of the ACM*, September 90, vol. 33, n° 9, pp. 143-159.
- [El Bitar et al. 2006] – El Bitar Z., Lazaro D., Breton V., Hill D. et Buvat I., « Fully 3D Monte Carlo image reconstruction in SPECT using functional regions », *Nucl. Instr. Meth. Phys. Res.*, 569, pp. 399-403.
- [Escoubet et al. 1998] – Escoubet S., Dupeux D. et Escoubet P., « Utilisation du chlorure de sodium comme moyen d'éradication de *Caulerpa taxifolia* (Vahl) C. Agardh. », in : Boudouresque C. F., Gravez V., Meinesz A., Paul-luy F. (éds.), *Third International Workshop on Caulerpa taxifolia, GIS Posidonia Publication, Marseille*, pp. 117-124.

F.

- [Feillee et al. 2008] – Feillee D., Hill D. et Dessagnes G., « Vers une approche multi-agents pour la simulation du système ferroviaire français », *Revue Génie Logiciel*, n° 86, septembre, pp. 22-28.
- [Ferber 1990] – Ferber J., *Eco Problem Solving: How to Solve a Problem by Interactions*, Rapport Technique LAFORIA, n° 5/90.
- [Ferber 1995, 1999] – Ferber J., *Les Systèmes Multi-Agents – vers une intelligence collective*, Paris, InterÉditions, Traduction anglaise chez Addison Wesley.
- [Ferguson et al. 2005] – Ferguson N. M., Cummings D. A. T., Cauchemez S., Fraser C., Riley S., Meeyai A., Iamsirithaworn S., Burke D. S., « Strategies for containing an emerging influenza pandemic in Southeast Asia », *Nature*, Vol. 437, pp. 209-214 (8 Septembre 2005 – doi:10.1038/nature04017).
- [Filippi et Bisgambiglia 2004] – Filippi J. B., Bisgambiglia P., « The JDEVS environmental modeling and simulation environment », *Environmental Modelling & Software*, Elsevier, vol. 19, Issue 3, Mars, pp. 261-274.

- [Fishman 1978]** – Fishman G. S., *Principles of Discrete Event Simulation*, John Wiley & Sons.
- [Fishman et Kiviat 1967]** – Fishman G. S. et Kiviat P. J., « The Analysis of Simulation Generated Time Series », *Management Science*, 13, 7 (July), pp. 525-557.
- [Fishwick 1993]** – Fishwick P. A., « A Simulation Environment for Multimodeling », *Discrete Event Dynamic Systems: Theory and applications*, vol. 3, pp. 151-171.
- [Fishwick 1995]** – Fishwick P. A., *Simulation Model Design and Execution*, Prentice-Hall.
- [Fishwick 1996]** – Fishwick P. A., « Web-based Simulation: Some Personal Observations », *Winter Simulation Conference*, San Diego, CA, December, pp. 772-779.
- [Fishwick et al. 1998]** – Fishwick P. A., Hill D. et Smith R., « International Conference on Web-based Modeling and Simulation », *SCS*, San Diego, Jan. 11-14, p. 203.
- [Fishwick et Hill 1999]** – Fishwick P. A., Hill D. (éds.), *Web-based Simulation*, Special Issue of the *Simulation Journal*, vol. 72, n° 3, 60 p..
- [Force et al. 2002]** – Force C., Perochon P. et Hill D., « Design of a multimodel of a dairy cows herd attacked by mastitis », *Simulation Modelling Practice and Theory*, vol. 10, Issue 8, 31 December, pp. 543-554.
- [Forrester 1961]** – Forester J., *Industrial Dynamics*, Cambridge (MA), MIT Press.
- [Forrester 1969]** – Forester J., *Urban Dynamics*, Cambridge (MA), MIT Press.
- [Forrester 1971]** – Forester J., *World Dynamics*, Cambridge (MA), Wright-Allen Press.
- [Fowler et Scott 1997]** – Fowler M. et Scott K., *UML Distilled*, Addison-Wesley Longman.
- [Frontier 1977]** – Frontier S., « Réflexions pour une théorie des écosystèmes», *Bull. Ecol.*, vol. 8, pp. 445-464.

G.

- [Gamma et al. 1995]** – Gamma E., Helm R., Johnson R. et Vlissides J., *Design Patterns: Elements of Reusable Object-Oriented Software*, Addison-Wesley, reading, MA.
- [Gasser et Briot 1992]** – Gasser L., Briot J. P., « Object-Based Concurrent Programming and DAI », in : Avouris N. M., Gasser L. (éds.), *Distributed*

- Artificial Intelligence: Theory and Praxis*, Dordrecht, Kluwer Academic Press, pp. 81-108.
- [**Gavach et al. 1998**] – Gavach C., Bonnal L., Uchimura M., Sandeaux R., Sandeaux J., Souard R., Lamaze B., Lasserre J-C., Fougairolle C., Combes J-F. et Gravez V., « Destruction de *Caulerpa taxifolia* par la technique de la couverture à ions cuivrées. Développement pré-industriel et premiers essais », in : Boudouresque C. F., Gravez V., Meinesz A. et Paulluy F. (éds), *Third International Workshop on Caulerpa taxifolia*, GIS Posidonie Publication, Marseille, pp. 101-104.
- [**Geman et al. 1992**] – Geman S., Bienenstock E. et Doursat R., « Neural networks and the bias/variance dilemma », *Neural Computation*, vol. 4, n° 1, pp. 1-58.
- [**Georgiev et Hoogenboom 1999**] – Georgiev G. et Hoogenboom G., « Near Real-Time Agricultural Simulations on the Web », in : Fishwick P. et Hill D. (éds.), *Simulation, Special Issue in Web-based Simulation*, 73(1), pp. 22-28.
- [**Gipps 1986**] – Gipps P. G., « The role of computer graphics in validating simulation models », *Mathematics and Computer in Simulation*, vol. 28 pp. 285-289.
- 184**
- [**Gleick 1989**] – Gleick J., *La Théorie du Chaos*, Paris, Flammarion, « Champs ».
- [**Gordon 1978**] – Gordon G., *System simulation*, Englewood Cliffs (N.J.), Prentice-Hall Inc.
- [**Gotelli 1998**] – Gotelli N. J., *A primer of ecology*, Sunderland, Sinauer, 206 p. (2nd Ed.).
- [**Gourgand 1984**] – Gourgand M., *Outils logiciels pour l'évaluation des performances des systèmes informatiques*, Thèse d'État, Université Blaise-Pascal Clermont-Ferrand II, France.
- [**Gourgand et Hill 1990**] – Gourgand M. et Hill D., « Petri Nets modelling on transputers with OCCAM2 ». *SCS European Simulation Symposium*, Ghent (Belgium), pp. 143-148.
- [**Green et Sun 1988**] – Green M. et Sun H., « Interactive animation », *IEEE Computer Graphics and Applications*, vol. 8, n° 6. pp. 52-65.
- [**Grimm 1999**] – Grimm V., « Ten years of individual-based modelling in ecology : what have we learned and what could we learn in the future », *Ecological Modelling*, vol. 115, pp. 129-148.
- [**Guessoum 1996**] – Guessoum Z., *Un Environnement opérationnel de conception et de réalisation de systèmes multi-agents*, Thèse de doctorat de l'Université Pierre et Marie Curie, Paris (France).

H.

- [Hallam et Levin 1986]** – Hallam T. G. et Levin S. A. (éds.), « *Mathematical Ecology: an Introduction, Biomathematics*, vol. 17, Springer, 457 p. (ISBN 3-540-13631-2, 0-387-13631-2).
- [Heudin 1994]** – Heudin J. C., *La Vie artificielle*, Paris, Hermès, 267 p.
- [Hill 1992]** – Hill D., « Étude de quelques concepts pour une analyse et une conception par objets », *Congrès INFORSID (Clermont-Ferrand, 19-22 Mai)*, pp. 307-326.
- [Hill 1993a]** – Hill D., *Outils logiciels pour la modélisation par objets de systèmes complexes*, Doctorat en Informatique, Université Blaise-Pascal Clermont-Ferrand II, Février.
- [Hill 1993b]** – Hill D., *Analyse Orientée-Objets et Modélisation par Simulation*, Addison-Wesley, 362 p.
- [Hill 1993c]** – Hill D., « Enhancing the QNAP2 Object-Oriented Simulation Language for Manufacturing Modelling », in : *Proceedings of the 1993 European Simulation Multi-Conference (June 7-9. Lyon. France)*, pp. 171-175.
- [Hill 1995a]** – Hill D., « Object-oriented Modelling and Simulation », in : *TOOLS 13 tutorial notes (Versailles, France, 7-10 Mars)*, pp. 96-101.
- [Hill 1995b]** – Hill D., « Verification and Validation of Ecosystem Simulation Models », in : *Proceedings of the SCS Summer Simulation Conference (July 24-26, Ottawa, Canada)*, pp. 176-182.
- [Hill 1996]** – Hill D., *Object-Oriented Analysis and Simulation*, Addison-Wesley Longman, 291 p.
- [Hill 1997a]** – Hill D., « Introduction à la Simulation par Objets », *L'Objet*, Paris, Hermès, vol. 3, n° 1, pp. 53-63.
- [Hill 1997b]** – Hill D., « Modélisation des processus d'expansion : application à *Caulerpa taxifolia* », *Conférence Internationale sur la « Dynamique des Espèces invasives » (13-15 mars 1997, Paris, Académie des Sciences)*, Tec & Doc, pp. 219-230.
- [Hill 1997c]** – Hill D., « Object-Oriented Pattern for Distributed Simulation of Large Scale Ecosystems », *SCS Summer Computer Simulation Conference (July 13-17, Arlington, USA)*, pp. 945-950.
- [Hill 2002]** – Hill D., « URNG: A portable optimisation technique for every software application requiring pseudorandom numbers », *Simulation Modelling Practice and Theory*, vol. 11, pp. 643-654.
- [Hill 2006]** – Hill D., « Traduction inverse de Protéines », LIMOS UMR CNRS 6158, *Technical Report*, November, p. 8.

- [Hill *et al.* 1992] – Hill D., Laize E., Ruch S. et Delain T., « Étude d'outils de simulation et d'animation », *Rapport TEMPUS JEP 2605-92/2*.
- [Hill *et al.* 1994a] – Hill D., Pastre J., Coquillard P. et Gueugnot J., « Design of an Ecosystem Modelling Environment: Application to Forest Growth Simulation », *CISS 94, Joint Simulation Societies (Zurich, Switzerland, August 22-26)*, pp. 538-544.
- [Hill *et al.* 1994b] – Hill D., Coquillard P., Vauvelas J. et Meinez A., « Simulation sur ordinateur de l'expansion de l'algue Tropicale *Caulerpa taxifolia* en Méditerranée, Résultats préliminaires », *2nd International Conference on Caulerpa Taxifolia (Dec, 11-15 1994, Barcelone)*, pp. 119-127.
- [Hill *et al.* 1995a] – Hill D., Coquillard P., Gueugnot J., « Object-Oriented Modelling for Forest Growth Simulation Environments », *Object Technology 95, 27-29 March, OXFORD, United-Kingdom*.
- [Hill *et al.* 1995b] – Hill D., Coquillard P., Vauvelas J. et Meinesz A., « A stochastic model with spatial constraints », *EUROSIM Congress (September 11-15, Vienna, Austria)*, pp. 999-1004.
- [Hill *et al.* 1996a] – Hill D., Mazel C., Kellert P. et Coquillard P., « Modèles Dynamiques d'écosystèmes et Simulation à événement discrets », *Colloque CNRS Tendances nouvelles en modélisation pour l'environnement (Paris, Janvier 15-17)*, pp. 117-123.
- [Hill *et al.* 1996b] – Hill D., Mazel C. et Coquillard P., « Integrating V&V in the object-oriented life cycle of ecological modelling simulation projects », *8th SCS European Simulation Symposium (Oct, 24-26, Genova, Italy)*, vol. II pp. 21-25.
- [Hill *et al.* 1997a] – Hill D., Vaugelas J., Campos A. et Meyer U., « Visual Object-Oriented Simulation of Posidonia Oceanica Growth », *9th SCS European Simulation Symposium (October 19-23, Passau, Germany)*, pp. 291-296.
- [Hill *et al.* 1997b] – Hill D., Coquillard P. et De Vaugelas J., « Discrete-Event Simulation of Alga Expansion », *Simulation*, vol. 68, n° 5, pp. 269-277.
- [Hill *et al.* 1998a] – Hill D., Dumont B. et Roux C., « Modelling spatial memory of sheep at pasture with multi-agents », *SCS European Simulation Symposium (Manchester 16-19 Juin)*, pp. 348-352.
- [Hill *et al.* 1998b] – Hill D., Coquillard P., De Vaugelas J. et Meinesz A., « An algorithmic Model for Invasive Species Application to *Caulerpa taxifolia* (Vahl) C, Agardh development in the North-Western Mediterranean Sea », *Ecological Modelling*, vol. 109, pp. 251-265.
- [Hill *et al.* 2000a] – Hill D., Coquillard P., Garcia B., Traore M.K., Mazel C., Campos A. et Thibault T., « Multimodeling and Object-oriented Design

- Patterns Application to Bio-control Simulation », *Proceedings of Artificial Intelligence and Simulation 2000, Arizona*, pp. 219-228.
- [Hill et al. 2000b]** – Hill D., Mechoud S., Campos A., Coquillard P., Gueugnot J., Orth D., Michelin Y., Poix C., L'homme G., Carrere P., Lafarge L., Loiseau P., Micol M., Brun M., Dubuc F., Dumont B., Petit M. et Teuma M., « Modélisation de l'entretien du paysage par des herbivores en moyenne montagne : une approche multi-agents », *Ingénieries*, n° 21, Mars, pp. 63-75.
- [Hill et al. 2001]** – Hill D., Coquillard P., Aussem A., de Vaugelas J., Thibaut T. et Meineisz A., « Modeling the Ultimate Seaweed », *Simulation*, vol. 76, n° 2, pp. 126-134.
- [Hill et al. 2002]** – Hill D., Thibault T. et Coquillard P., « Predicting Invasive species expansion using GIS & Simulation coupling », *Modelling and Simulation*, vol. 1, n° 1, pp. 30-35.
- [Hill et al. 2008]** – Hill D., Muzy A., Barraud R., Crozat B., Madary J. et Leccia F., « Design of a spatial and stochastic simulator for bird flu spreading in Corsica », *2008 International Simulation Multi conference (ISMC'08)*, IEEE / ACM / SCS, Edinburgh, p. 445-452.
- [Hill et Coquillard 2007]** – Hill D. et Coquillard P., « Ecological Modelling and Simulation », in : Fishwick P. (éd.), *Handbook of Dynamic System Modeling*, CRC Press, Chapitre 29, 18 p.
- [Hill et Junqua 1990]** – Hill D. et Junqua A., *GAME : un outil orienté-objet pour l'animation de simulation de systèmes flexibles d'assemblage*, Rapport d'Ingénieur et de DEA, Université Blaise-Pascal, Clermont-Ferrand II.
- [Hill et Vigor 1998a]** – Hill D. et Vigor E., « Dealing with Distributed Dilemmas », *Application Development Advisor* (remplace *Object Expert*), vol. 1, n° 4, pp. 60-63.
- [Hill et Vigor 1998b]** – Hill D. et Vigor E., « Simulation and Software Engineering : bridging the culture gap with UML », *SCS Object-Oriented Simulation Conference (January 11-14, San Diego, USA)*, pp. 81-87.
- [Hinde et Smith 1974]** – Hinde R. et Smith D. C., « “Chloroplast symbiosis” and the extend to which it occurs in Sacoglossa (Gastropoda : Mollusaca) », *Biol. J. Linn. Soc.* 6, pp. 349-356.
- [Hirata et al. 2000]** – Hirata C., Yano E. et Filho W., « A cooperative simulation modeling environment based on the www », *Web-based Modeling & Simulation Conference (San Diego CA, January 2000)*, pp. 28-33.
- [Hollocks 1984]** – Hollocks B. W., « Practical benefits of animated graphics in simulation », in : *Proceedings of the Winter Simulation Conf.*, pp. 323-328.

- [**Horn 1975a**] – Horn H. S., « Forest succession », *Sci. Am.*, vol. 232, pp. 90-98.
- [**Horn 1975b**] – Horn H. S., « Markovian properties of forest succession », *Eco-logy and Evolution of communities*, Cambridge, Harvard University Press, pp. 196-211.
- [**Howard 1995**] – Howard R., « Threaded Simulation », *JOOP*, July-Aug., pp. 59-61.
- [**Hsu et Hunter 1977**] – Hsu D. A., Hunter J. S., « Analysis of Simulation-Generated Response Using Autoregressive Models », *Management Sci.*, 24, pp. 181-190.
- [**Huber et Szczerbicka 1994**] – Huber K. P. et Szczerbicka K., « Sensitivity analysis of simulation models with decision tree algorithms », in : *Proceedings of the European Simulation Symposium ESS '94*, vol. 1, pp. 43-47.
- [**Hurzion et Secker 1978**] – Hurzion R. D. et Secker R. J. R., « Visual interactive simulation, an aid to decision making », *Omega*, vol. 6, n° 5, pp. 419-426.
- [**Huston et al. 1988**] – Huston M., De Angelis D. et Post W., « New Computer Models Unify Ecological Theory, Computer Simulation shows that many ecological patterns can be explained by interactions among individual organisms », *BioScience*, vol. 38, n° 10, pp. 682-691.

I.

- [**Iazeolla et D'Ambrogio 1998**] – Iazeolla G., D'Ambrogio A., « A Web-based environment for the reuse of simulation model », *1st International Conference on Web-based Modeling & Simulation*, San Diego CA, January, 1998.

- [**Innocenti et al. 2009**] – Innocenti E., Silvani X., Muzy A., Hill D., « A software framework for fine grain parallelization of cellular models with OpenMP: Application to fire spread », *Environmental Modelling & Software*, vol. 24, pp. 819–831.

J.

- [**Jacobson et al. 1993**] – Jacobson I., Christerson M., Jhonsson P. et Overgaard G., *Génie Logiciel Orienté Objets*, Addison Wesley.

- [**Jacobson et Bylund 2000**] – Jacobson I. et Bylund S., *The Road to the Unified Software Development Process*, SIGS Ref, 18, 400 p.

- [**Jaffrenou et Odonne 1994**] – Jaffrenou B., Odonne L., « Procédé de maîtrise et de réduction du développement de *Caulerpa taxifolia* en Méditerranée », in :

- Boudouresque C. F., Meinesz A., Gravez V. (éds.), *First International Workshop on Caulerpa taxifolia*, Marseille, GIS Posidonie Publication, pp. 339-340.
- [Jeffries 1989] – Jeffries C., *Mathematical Modeling in Ecology - a Workbook for Students*, Boston, Birkhauser, 193 p.
- [Jennings et al. 1998] – Jennings N., Sycara K., Wooldridge M., « A Roadmap of Agent Research and Development », in : *Autonomous Agents and Multi-Agent Systems*, vol. 1, Boston, Kluwer Academic Publishers, pp. 275-306.
- [Jensen 1990] – Jensen K. R., « Sacoglossa (Mollusca: Opisthobranchia)-Specialists herbivores and partial predators: integrating ecological, physiological and morphological data », in : *Proceedings of the First International Conference on the Marine Biology of Hong Kong and the South China Sea*, Hong Kong, Hong Kong University Press, pp. 437-457.
- [Jensen 1993] – Jensen K. R., « Morphological adaptations and plasticity of radular teeth of the Sacoglossa (=Ascoglossa) (Mollusca: Opisthobranchia) in relation to their food plants », *Biol. J. Lin Soc.*, 48, pp. 135-155.
- [Jensen 1997] – Jensen K. R., *Systematics phylogeny and evolution of the Sacoglossa (Mollusca, Opisthobranchia)*, Kobenhavn, Vestjydsk Forlag, 94 p.
- [Johnson et al. 1986] – Johnson K. N., StuartT. W. et Crim S. A., *FORPLAN Version 2: an overview*, USDA Forest Service, Land.
- 189
- [Jørgensen 1994] – Jørgensen S. E., *Fundamentals of Ecological Modelling*, Elsevier, 632 p. (2nd edition).
- [Jousson et al. 1998] – Jousson O., Pawłowski J., Zaninetti L., Meinesz A. et Boudouresque C. F., *Molecular evidence for the aquarium origin of the green alga Caulerpa taxifolia introduced to the Mediterranean Sea*, Mar Ecol Prog Ser 172, pp. 275-280.
-
- ## K.
- [Karthela et al. 2000] – Karhela T., Mettälä A., Paljakka M., « Component based framework for developing process simulation web user interfaces », *3rd Web-based Modeling & Simulation Conference*, San Diego (CA), January, pp. 61-66.
- [Keen et Spain 1992] – Keen R. E. et Spain J. D., *Computer simulation in biology: a BASIC introduction*», New York (etc.), Wiley-Liss, 516 p., incl. disk.
- [Kellert 1992] – Kellert P., « Définition et mise en œuvre d'une méthodologie orientée objets pour la modélisation des systèmes de production », *Congrès INFORSID (Clermont-Ferrand, 19-22 Mai)*, pp. 415-436.

- [Kellert *et al.* 1997] – Kellert P., Tchernev N., Force C., « Object-oriented methodology for FMS modelling and simulation », *International Journal of Computer Integrated Manufacturing*, vol. 10, n° 6, pp. 405-434.
- [Kent *et al.* 1991] – Kent B., Bare B. B., Field R. C. et Bradley G. A., « Natural Resource Land Management Planning using Large-Scale Linear Programs: the USDA Forest Service experience with FORPLAN », *Operations Research*, 39, pp. 13-27.
- [Kleijnen 1979] – Kleijnen J. P. C., « Regression metamodels for generalizing simulation results », *IEEE Transactions on Systems, Man, and Cybernetics*, 9 (2), pp. 93-96.
- [Kleijnen 1987] – Kleijnen J. P. C., *Statistical Tools for Simulation Practitioners*, New York, Marcel Dekker.
- [Kleijnen et Groenendaal 1992] – Kleijnen J., Van Groenendaal W., *Simulation: A Statistical Perspective*, Chichester, John Wiley.
- [Knaber 1996] – Knabe F., « An Overview of Mobile Agent Programming ». *Proceedings of the 5th LOMAPS Workshop on Analysis and Verification of Multiple-Agent Languages*, Stockholm, Sweden, June.
- [Komatsu *et al.* 1997] – Komatsu T., Meinesz A. et Buckles D., « Temperature and light responses of alga *Caulerpa taxifolia* introduced into the Mediterranean Sea », *Marine Ecology*, vol. 146, pp. 145-153.,

L.

- [Lafarge *et al.* 2005] – Lafarge M., Mazel C. et Hill D., « A modelling of the tillering capable of reproducing the fine-scale horizontal heterogeneity of a pure grass sward and its dynamics », *Ecological Modelling*, vol. 183, pp. 125-141.
- [Lafferty et Kurris 1996] – Lafferty K. D. et Kurris A., « Biocontrol of marine pests », *Ecology*, 77 (7), pp. 1989-2000.
- [Lai 1997] – Lai M., *UML, la notation unifiée de modélisation objet*, InterÉditions.
- [Laizé 1998] – Laizé E., *Modélisation des systèmes manufacturiers complexes : analyse du domaine et spécification de la connaissance*, Thèse de doctorat en informatique, Université Blaise-Pascal Clermont-Ferrand II.
- [Law et Kelton 1991] – Law M., Kelton W. D., *Simulation Modeling and Analysis* McGraw-Hill Inc.

- [Lazaro et al. 2005]** – Lazaro D., El Bitar Z., Breton V., Hill D. et Buvat I., « Fully 3D Monte Carlo reconstruction in SPECT: a feasibility study », *Phys. Med. Biol.*, vol. 50, pp. 3739-3754.
- [Le Moigne 1977]** – Le Moigne J. L., *La Théorie du système général : Théorie de la modélisation, Système Décision*, Paris, PUF.
- [Le Moigne 1990]** – Le Moigne J. L., *La Modélisation des systèmes complexes*, Paris, Dunod.
- [Leemans et Prentice 1987]** – Leemans R. et Prentice I. C., « Description and simulation of tree-layer composition and size distributions in a primaeval Picea-Pinus forest », *Vegetatio*, vol. 69, pp. 147-156.
- [Legay 1986]** – Legay J. M., « Quelques réflexions à propos d'écologie. Défense de l'indisciplinarité », *Acata Oecologica, Oecol. Gener.*, vol. 7, n° 4, pp. 391-398.
- [Legay 1996]** – Legay J. M., *L'Expérience et le modèle : un discours sur la méthode*, Sciences en question, Éd. INRA, 111 p.
- [Legendre 1989]** – Legendre P., Fortin M. J., « Spatial Patterns and ecological Analysis », *Vegetatio*, Vol. 80, pp. 107-138.
- [Lehman 1980]** – Lehman M. M., *Programs, programming and the software life cycle*, Report 80/6, April 15th, Londres, Imperial College of Science and Technology.
- [Leroudier 1980]** – Leroudier J., *La Simulation à événements discrets*, PARIS, Éd. Hommes et Techniques.
- [Levin 1989]** – Levin S. A., Hallam T. G. et Gross L. J., *Applied Mathematical Ecology*, vol. 18 of *Biomathematics*, Springer.
- [Lhotka 1991]** – Lhotka L., « Object-oriented methodology in the field of aquatic ecosystem modeling », in : *TOOLS 4, Technology of O-O Language and Systems*, Prentice Hall, pp. 309-317.
- [Longini et al. 2005]** – Longini M., Nizam A., Xu S., Ungchusak K., Hanshaoworakul W., Cummings D. A. T. et Halloran M., « Containing Pandemic Influenza at the Source », *Science*, 12 juillet.

M.

- [Mangili et Gendreau 2005]** – Mangili A. et Gendreau M. A., 2005, « Transmission of infectious diseases during commercial air travel », *Lancet*, 365 (9463), pp. 989-996.
- [Matteson et Anderson 1993]** – Matteson S. E. et Anderson M., « Omola: An object-oriented modeling language », in : Jamshidi M. et Herget C. J. (éds.),

- Recent Advances in Computer Aided Control Systems Engineering*, vol. 9 of *Studies in Automation and Control*, New York, Elsevier Science Publishers, pp. 291-310.
- [May 1973] – May R. M., *Stability and Complexity in Model Ecosystems*, Monographs in Population Biology, n° 6, Princeton Univ. Press, 247 p. (paperback).
- [Maynard Smith 1974] – Maynard Smith J., *Models in Ecology*, Cambridge University Press, 146 p.
- [Mazel et al. 1996] – Mazel C., Hill D. et Brunner O., « Optimization of Forest Expansion Simulations with an Object-Oriented approach of Spatial Aspects », *22nd ASU Conference in Object-Oriented Modelling & Simulation (July 15-17, Clermont-Ferrand)*, pp. 145-158.
- [Mazel et al. 1997] – Mazel C., Hill D. et Coquillard P., « Interprétation des résultats de couplage de Systèmes d'Information Géographique et d'outils de simulation discrète d'écosystèmes », *Conférence Européenne sur l'Informatique pour l'Environnement*, organisée par l'INRIA, pp. 598-607.
- [Mazel et al. 2005] – Mazel C., Lafarge M. et Hill D., « An individual-based, stochastic and spatial model to simulate the ramification of grass tillers and their distribution in swards ». *Simulation Practice & Theory*, vol. 13, Issue 4, June, pp. 308-334.
- 192**
- [Mechoud et al. 1998] – Mechoud S., Hill D., Campos A., Orth D., Carrere P., Micol D., Poix C., Michelin Y., Coquillard P. et Dumont B., « Simulation Multi-Agents de l'entretien du paysage par des herbivores en moyenne montagne », *Actes de la Conférence SMAGET organisée par le CEMAGREF et l'ENGREF*, pp. 65-78.
- [Meinesz 1997] – Meinesz A., « Utilisation d'Ascoglosses pour la lutte biologique contre *Caulerpa taxifolia* en Méditerranée », in : Académie des Sciences (Paris), *Dynamique d'espèces marines invasives : application à l'expansion de Caulerpa taxifolia en Méditerranée*, Paris, Tec & Doc Publ., pp. 291-300.
- [Meinesz 1998a] – Meinesz A., Cottalorda J.-M., Chiavérini D., Cassar N. et De Vaugelas J., *Suivi de l'invasion de Caulerpa taxifolia en Méditerranée : situation au 31.12.1997*, Éd. LEML-UNSA publ.
- [Meinesz 1998b] – Meinesz A., Cottalorda J.-M. et Chiavérini D., *Suivi de l'invasion de l'algue tropicale Caulerpa taxifolia devant les côtes françaises de la Méditerranée: Situation au 31 décembre 1997*, Université de Nice-Sophia Antipolis, Éd. Laboratoire Environnement Marin Littoral.
- [Meinesz 1999] – Meinesz A., « La lutte biologique contre les espèces introduites en milieu marin », in : INRA (éd.), *Les Dossiers de l'Environnement de l'INRA*, n° 19 : lutte biologique II, pp. 29-34.

- [Meinesz et al. 1993]** – Meinesz A., Vaugelas J., de, Hesse B. et Mari X., « Spread of the introduced tropical green alga *Caulerpa taxifolia* in northern Mediterranean waters », *Journal of applied Phycology*, 5, pp. 141-147.
- [Meinesz et al. 1996]** – Meinesz A., Melnick J., Blachier J. et Charrier S., « Étude préliminaire, en aquarium, de deux ascoglosses tropicaux consommant *Caulerpa taxifolia* : une voie de recherche pour la lutte biologique », in : Ribera M. A., Ballesteros E., Boudouresque C. F., Gomez A., Gravez V. (éds), *Second International Workshop on Caulerpa taxifolia*, Barcelone, Publicacions Universitat Barcelona, pp. 157-161.
- [Meinesz et Belsher 1993]** – Meinesz A. et Belsher T., « Observations en sous-marin de *Caulerpa taxifolia* dans l'étage circalittoral de l'est des Alpes-Maritimes », *Rapport du Laboratoire Environnement Marin Littoral*, Université de Nice-Sophia Antipolis et du Laboratoire d'Ecologie, IFREMER, Centre de Brest-DEL.
- [Meinesz et Hesse 1991]** – Meinesz A. et Hesse B., « Introduction et invasion de l'algue tropicale *Caulerpa taxifolia* en Méditerranée nord-occidentale », *Oceanologica Acta*, vol. 14, n° 4, pp. 415-426.
- [Meyer 1990]** – Meyer B., *Conception et Programmation par objets*, InterÉditions.
- [Michelin et al. 1995]** – Michelin Y., Coquillard P., Hill D., L'homme G., Loiseau P. ET Micol D., « Sustainable herbivores production and ecological land use in mountains areas a multidisciplinary approach », *IV International Symposium on the Nutrition of Herbivores (Clermont-Ferrand, 11-15 Septembre 1995)*. Poster.
- [Michelin et al. 1998]** – Michelin Y., Orth D., L'Homme G., Coquillard P., Gueugnot J., Hill D., Micol D., Loiseau P., Lafarge M., Carrère P. et Teuma M., « Gestion des espaces enrichis par des bovins et des équins : Présentation d'une recherche pluridisciplinaire », *Fourrage*, n° 153, pp. 115-123.
- [Micol 1997]** – Micol D., « Gestion de la végétation et entretien des milieux par les herbivores en moyenne montagne, approche expérimentale et modélisation, présentation du G.I.S. », *Dossiers de l'Environnement*, INRA.
- [Militon et al. 2007]** – Militon C., Rimour S., Missaoui M., Biderre C., Barra V., Hill D., Mone A., Gagne G., Meier H., Peyretaillade E. et Peyret P., *PhyLArray: Phylogenetic Probe Design Algorithm for MicroArray*, *Bioinformatics 2007*, doi: 10.1093/bioinformatics/btm392, vol. 23, pp. 2550-2557.
- [Minsky 1965]** – Minsky M. L., « Matter, Minds and Models », *Proceedings of Inter, Federation of Information Processing Congress*, vol. 1, pp. 45-49.
- [Missaoui et al. 2008]** – Missaoui M., Hill D. et Peyret P. *Comparison of Algorithms for a Complete Backtranslation of Oligopeptides*, 2008, special issue,

- IJCBDD, *International Journal of Computational Biology and Drug Design* 2008, vol. 1, n° 1, pp. 26-38.
- [Muller 1997] – Muller P. A., *Modélisation objet avec UML*, Eyrolles.
- [Murillo *et al.* 1986] – Murillo L, Templado J, Talavera P, « The ascoglossan opisthobranchs of a caulerpan fauna of the Mediterranean Sea », *Shell Sea Life*, 17, pp. 240-243.
- [Muzy *et al.* 2005] – Muzy A., Innocenti E., Aïello A., Santucci J. F., Santoni P. A. et Hill D., « Modelling and simulation of ecological propagation processes: application to fire spread », *Environmental Modelling & Software*, vol. 20, Issue 7, July, pp. 827-842.
- [Muzy *et al.* 2008] – Muzy A., Hill D., Joubert M. et Innocenti E., *A post-processed 3D visualization tool for forest fire simulations, IEEE Simutools 2008 Conference Record, Marseille, March 3-7*, Marseille, CD Proceedings, 6 p.

N.

- [Nance 1987] – Nance R. E., « The Conical Methodology: A Framework for Simulation Model Developement », *Proc. SCS Methodology & Validation*, pp. 38-43.
- 194
- [Narayanan *et al.* 1999] – Narayanan S., Edala N., Geist J., Kumar P., Ruff H., Draper M. et Haas M., « UMAST : A Web-based Architecture for Modeling Future Uninhabited Aerial Vehicles, in Simulation », in : Fishwick P. et Hill D. (éds.), Special issue in *Web-based Simulation*, 73(1), July, pp. 29-39.
- [Nisbet et Gurney 1982] – Nisbet R. M et Gurney W. S. C., *Modelling Fluctuating Population*, Chichester, Wiley, 379 p.
- [Nwana 1996] – Nwana H., « Software Agents: an overview », *The Knowledge Engineering Review*, 11(3).

O.

- [Okubo 1980] – Okubo A., *Diffusion and ecological problems: mathematical models*, vol. 10 of *Biomathematics*, Springer, 254 p..
- [Ören 1984] – Ören T. I., « Model-Based Activities: A Paradigm Shift », in : Ören T. I., Zeigler B. P. et Elsaz M. S. (éds.), *Simulation and Model-Based Methodologies: An Integrative View*, New York, Springer Verlag, pp. 3-40.

[Overstreet 1982] – Overstreet C. M., *Model Specification and Analysis for Discrete Event Simulation*, Ph.D. Dissertation, CS dept, Virginia Tech, Blacksburg, VA.

P.

[Pace 1992] – Pace D. K., « Simulation, the Defense Community, and DMSO », *Simulation*, vol. 58, n° 1, pp. 62-64.

[Page 1998] – Page E., « The Rise of Web-Based Simulation: Implications for the High Level Architecture », *1998 Winter Simulation Conference*, SCS, Washington DD, 13-16 December, pp. 1663-1669.

[Page et al. 1997] – Page E., Moose R. et Griffin S., « Web-based Simulation in SimJava using Remote Method Invocation », *Winter Simulation Conference*, SCS, Atlanta GA, 7-10 December, pp.468-474.

[Palme 1977] – Palme J., « Moving Pictures show simulation to user », *Simulation*, vol. 29, pp. 240-249.

[Palmore 1994] – Palmore J., « A verification and validation framework for analysis in distributed interactive simulation », *Proceedings of the Object-Oriented Simulation Conference (OOS'94)*, pp. 49-54.

195

[Parker 1996] – Parker D., *Innovations in GIS*, Taylor & Francis.

[Paul et Van Alstyne 1988] – Paul V. J. et Van Alstyne K. L., « Use of ingested diterpenoids by Elysia halimeda Macnae (Opisthobranchia: Ascoglossa) as antipredator defenses », *J. Exp. Mar. Bol. Ecol.*, 119, pp. 15-29.

[Pavé 1994] – Pavé A., « Modélisation en Biologie et Écologie », *Aléas*, Lyon.

[Perrochon et al. 1998] – Perrochon L., Force C., Hill D., Coulon J. B. et Gasqui P., « Simulation à l'aide d'un modèle individu centré de l'impact des mammites sur les performances des vaches laitières », *Actes de la Conférence SMAGET organisée par le CEMAGREF et l'ENGREF*, pp. 39-50.

[Peterson 1995] – Peterson P. M., *Interactive and Animated Cartography*, Prentice Hall.

[Pielou 1974] – Pielou E. C., *Population and Community Ecology: Principles and Methods*, New York, Gordon and Breach, 424 p.

[Pielou 1977] – Pielou E. C., *Mathematical Ecology*, New York, Wiley, 385 p.

[Pierreval 1992] – Pierreval H., « Rule-based simulation metamodels », *European Journal of Operational Research*, 61, pp. 6-17.

- [**Pierreval 1996**] – Pierreval H., « A metamodelling approach based on neural networks », *International Journal of Computer Simulation*, vol. 6, n° 2.
- [**Popper 1973**] – Popper J., *La Dynamique des systèmes, principes et applications*, Paris, Éditions d'organisation.
- [**Potier 1977**] – Potier D., *Modèles à files d'attente et gestion des ressources dans un système informatique*, Thèse d'Etat, Grenoble.
- [**Praehofer et Schoepll 2000**] – Praehofer H. et Schoeppl A., « A continuous and Combined Simulation Platform in Java and its Application in Building Paper Mill Training Simulators », *Web-based Modeling & Simulation Conference*, San Diego CA, January, pp. 3-8.
- [**Prévosto et al. 1999a**] – Prévosto B., Curt T., Gueugnot J. et Coquillard P., « Colonization by Scots pine (*Pinus sylvestris L.*) after field abandonment at mountain level on volcanic soils », in : Papanastis V. P., Frame J., Nastis A. S. (éds.), « *Grasslands and woody Plants in Europe* », *Proceedings of the International Occasional Symposium Grassland Federation (Thessaloniki, Greece, May 27-29)*, pp. 105-112.
- [**Prévosto et al. 1999b**] – Prévosto B., Coquillard P., Gueugnot J., « Growth models of silver birch (*Betula pendula*, Roth.) on two volcanic mountains in the French Massif Central », *Plant Ecology* (anciennement *Vegetatio*), vol. 144, n° 2, pp. 231-242.
- [**Prévosto et al. 1999c**] – Prévosto B., Coquillard P., Gueugnot J., Hill D., « Modeling of Silver birch growth at mountain level on volcanic soils », *Second International Workshop on Functional-Structural Tree Models*, Clermont-Ferrand, És. INRA, 95 p.
- [**Prévosto et al. 2003**] – Prévosto B., Hill D. et Coquillard P., « Individual-based modelling of *Pinus sylvestris* invasion after grazing abandonment in French Massif Central », *Plant Ecology*, vol. 168, pp. 121-137.
- [**Pukkala et Kolstrom 1991**] – Pukkala T. et Kolstrom T., « Effect of a spatial pattern of trees on the growth of a Norway spruce stand, A simulation model », *Silva Fennica*, vol. 25, n° 3, pp. 117-131.

R.

- [**Rao et al. 1999**] – Rao D., Radhakrishnan R., Wilsey P., « FWNS: A Framework for Web-Based Network Simulation », *2nd Web-based Modeling & Simulation Conference*, San Francisco.
- [**Ravid et Rafaeli 2000**] – Ravid G. et Rafaeli S., « Multi player, internet and java-based simulation games: learning and research in implementing a com-

- puterized version of the beer-distribution supply chain game », *Web-based Modeling & Simulation Conference*, San Diego CA, January, pp. 15-20.
- [Renshaw 1993]** – Renshaw E., *Modelling biological populations in space and time*, vol. 11 of Cambridge studies in mathematical biology, Cambridge Univ. Press, 403 p.
- [Reuillon et al. 2008]** – Reuillon R., Hill D., El Bitar Z. et Breton V., « Rigorous distribution of stochastic simulations using the DistMe toolkit », *IEEE Transactions On Nuclear Science*, vol. 55, n° 1, February, pp. 595-603.
- [Ribera et al. 1996]** – Ribera M. A., Ballesteros E., Boudouresque C. F., Gomez A. et Gravez V., *Second International Workshop on Caulerpa taxifolia*, 457 p. Université de Barcelone Éd..
- [Rimour et al. 2005]** – Rimour S., Hill D., Militon C. et Peyret P., « GoArrays: highly dynamic and efficient microarray probe design », *Bioinformatics*, vol. 21, Issue 7, April, pp. 1094-1103.
- [Robert et al. 2005]** – Robert A., Prévosto B., Coquillard P. et Hill D. « Modelling the colonization dynamics of Scots pine in French areas. Impact of possible atmospheric carbon changes ». *OICMS 2005, International Conference on Modeling & Simulation (Clermont-Fd, June 12-15)*, pp. 231-239.
- [Roughgarden 1989]** – Roughgarden J., *Perspectives in ecological theory*, Princeton Univ. Press, 394 p.
- [Ruch 1994]** – Ruch S., *OMEGA : un environnement de modélisation multi-domaine des systèmes à flux discrets*, Thèse de doctorat en informatique de l'Université Blaise-Pascal Clermont-Ferrand II.
- [Rumbaugh 1991]** – Rumbaugh J., *Object-Oriented Modeling and Design*, Prentice Hall.
- [Rumelhart et al. 1986]** – Rumelhart D. E., Hinton G. E. et Williams W. J., « Learning internal representations by error propagation », in : Rumelhart D. E. et McClelland J. L. (éds.), *Parallel Distributed Processing Explorations in the Microstructure of Cognition*, Cambridge (MA), MIT Press, Bradford Books, vol, 1, pp. 318-362.
-
- S.**
- [Saito et al. 1993]** – Saito K., Kumagai Y., Honjo T., Ishida Y., de Reffye Ph. et Lecoustre R., « Photo-realistic Forest Landscape Simulation », *Nicograph*, 93, pp. 226-236.
- [Sargent 1979]** – Sargent R. G., « Validation of Simulation Models », in : *Proceedings of the 1979 Winter Simulation Conference*, San Diego, pp. 497-503.

- [**Sargent 1984**] – Sargent R. G. « A tutorial on Verification and Validation of Simulation Models », *Winter Simulation Conference*, pp. 115-121.
- [**Sargent et Som 1992**] – Sargent R. G. et Som T. K., « Current Issues in Frequency Domain Experimentation », *Management Science*, 38(5), pp. 667-687.
- [**Schmidt 1999**] – Schmidt C., « A Remote Laboratory Using Virtual Reality on the Web », in : Fishwick P. et Hill D. (éds.), *Simulation*, Special issue in Web-based Simulation, 73(1), July, pp. 13-21, July.
- [**Schneider 1994**] – Schneider D. C., *Quantitative Ecology: Spatial and Temporal Scaling*, Academic Press, 395 p.
- [**SCS 1979**] – SCS 1979, Technical Committee on Model Credibility 1979, « Terminology for Model Credibility », *Simulation*, 32, 3, pp. 103-104.
- [**Shannon 1986**] – Shannon R. E., « The use of Graphical Models in Model Validation », in : *Proceedings of the 1986 Winter Simulation Conference*, pp. 237-241.
- [**Sheng et al. 1993**] – Sheng G, Elzas M, S, Oren T, I. et Cronhjort B, T., « Model validation; a systemic and systematic approach », *Reliability Engineering and System Safety*, vol. 42, pp. 247-259.
- [**Shoham 1993**] – Shoham Y., « Agent-Oriented Programming », *Artificial Intelligence*, vol. 60, n° 1, pp. 51-92.
- [**Shugart et West 1997**] – Shugart H. H. et West D. C., « Development of an appalachian deciduous forest succession model and its application to assessment of the impact of the chestnut blightg », *J. Environment Management*, vol. 5, pp. 161-179.
- [**Signorile et Blais 2000**] – Signorile R. et Blais C. (éds.), *Proceedings of International Conference on Web-based Modeling and Simulation 2000*, San Diego, 23-27 janvier 2000.
- [**Simon 1991**] – Simon H. A., *Sciences des systèmes, sciences de l'artificiel*, Paris, Dunod, AFCET Systèmes.
- [**Sinton 1978**] – Sinton S, « The inherent structure of information as a constraint to analysis: mapped thematic data as a case of study », in : Dutton E. (éd.), Addison Wesley, Reading, Harvard Papers on Geographic Information Systems, vol. 6.
- [**Smith et Platt 1987**] – Smith R. L et Platt L., « Benefits of animation in the simulation of machining and assembly lines », *Simulation*, vol. 48, n° 1, pp. 28-30.
- [**Sommerville 1993**] – Sommerville I., *Le Génie logiciel*, Addison-Wesley.

[Stewart 1992] — Stewart I., *Dieu joue t-il aux dés ?*, Paris, Flammarion, Trad. fr. en 1992.

[Stout 1991] — Stout B. B., « Forest Growth Simulators - Problems and Prospects », in : *Proceedings of the SCS Western Simulation Multiconferences*, Towards understanding Our Environment, pp. 101-105.

T.

[Tanguy 1993] — Tanguy A., *Modélisation orientée objet des systèmes de production avec une approche transaction : résolution par des méthodes analytiques et par simulation*, Thèse de doctorat en informatique de l'Université Blaise-Pascal Clermont-Ferrand II.

[Ten Dyke et Kunz 1989] — Ten Dyke R. P., Kunz J. C., « Object-Oriented Programming », *IBM System Journal*, vol. 28, n° 3, pp. 465-478.

[Thibault et al. 1998] — Thibaut T., Meinesz A., Burtaire L., Charrier S., Ierardi S., Mangialajo L. et Vidal V., « Biological control of *Caulerpa taxifolia* in the Mediterranean Sea: use of tropical and mediterranean ascoglossans », in : Boudouresque C. F., Gravez V. (éds), *Third International Workshop on Caulerpa taxifolia*, Marseille, GIS Posidone publications, pp. 105-111.

[Thibault et Meinesz 2000] — Thibaut T. et Meinesz A. « Are the Mediterranean ascoglossan molluscs *Oxynoe olivacea* and Lobiger serradifalci suitable agents for a biological control against the invading alga *Caulerpa taxifolia*? », *CR Académie des Sciences, Life Science Paris*, vol. 323, pp. 477-488.

199

U.

[Uhrmacher 1997] — Uhrmacher A. M., « Concepts of Object and Agent-Oriented Simulation », *TRANSACTIONS of SCS*, vol. 14, n° 2, pp. 59-67.

[Urban et Shugart 1992] — Urban D. L. et Shugart H. H., « Individual based models of forest succession », in : Glenn-Lewin D. C. Peet R. K. et Veblen T; T. (éds.), *Plant Succession, Theory and Prediction*, Londres, Chapman & Hall, pp. 249-292.

V.

[Vandermeer 1981] — Vandermeer J. H., *Elementary Mathematical Ecology*, New York, Wiley and Sons, 294 p.

- [**Varenne 2007**] – Varenne F., *Du modèle à la simulation informatique*, Paris, Librairie Philosophique J. Vrin, 249 p.
- [**Vasconcelos et al. 1993**] – Vasconcelos M., Perestrello J., Zeigler B. P. et Graham L. A., « Modelling spatial dynamic ecological processes under the discrete event systems paradigm », *Landscape Ecology*, vol. 8(4), pp. 273-286.
- [**Vasconcelos et al. 1994**] – Vasconcelos M. J., Pereira J. M. C. et Zeigler B. P., « Simulation of fire growth in mountain environments », in : Price M. F. et Heywood D. I. (éds.), Taylor & Francis, *Mountain Environment & Geographic Information Systems*, pp. 167-185.
- [**Vaugelas et al. 1996a**] – De Vaugelas J., Charrier S., Commeau T., Cottalorda J.-M., Delahaye L., Jaffrennou F., Lemée R., Meinesz A., Molenaar H. et Pietkiewicz D., « Cartographie de l'invasion de *Caulerpa taxifolia*, Situation des côtes française de la Méditerranée à la fin de 1994 », *Second international Workshop on Caulerpa taxifolia*, 1996.
- [**Vaugelas et al. 1996b**] – Vaugelas J. De, Meinesz A., Coquillard P. et Hill D., « A partir de quel seuil peut-on modéliser une diminution de la biodiversité ? », *Colloque du Réseau Biodiversité Marine : Biodiversité en milieu dispersif (18-20 Novembre 1996)*, Paris, Muséum d'Histoire Naturelle, 43 p.
- [**Vaugelas et al. 1997**] – Vaugelas J. De, A, Meinesz, P. Coquillard & D, Hill., «A computer simulation to evaluate the impact of *Caulerpa taxifolia* on Mediterranean biodiversity», Vie et Milieu, Volume 47, n°4, pp. 397-400, 1997.
- [**Vaugelas et al. 1998**] – Vaugelas J. De, Meinesz A., Coquillard P. et Hill D., « Modéliser *Caulerpa taxifolia* pour prévoir sa distribution et simuler les opérations d'éradication locale », in : Boudouresque C.-F., Meinesz A. et Gravez V. (éds), *Third International Workshop on Caulerpa taxifolia*, GIS Posidonie Publ., pp. 63-67.
- [**Verlaque et al. 2000**] – Verlaque M., Meinesz A., Boudouresque C. F. et Gravez V., « The *Caulerpa racemosa* complex (Caulerpales, Ulvophyceae) in the Mediterranean Sea », *Bot. Mar.*, 43 (1), pp. 49-68.
- [**Vidal et al. 2000**] – Vidal V., D'Incan C., Santt O., Laplace-Marieze V., Delgado-Viscogliosi P., Baud V., Deval C., Baranova H., Champagnac S., Mazel C., Hill D., Albuison E., Ferrara M., Pradeyrol C. et Bignon Y.-J., « Les biopuces en Auvergne : aspects techniques et bioinformatiques », *Bulletin du cancer*, vol. 87, n° 5, mai, Abstract 63, p. 411.
- [**Vigor 1998**] – Vigor E., *Un métamodèle Objet des Systèmes de Production de Logiciels*, Thèse de doctorat en informatique, Université Blaise-Pascal Clermont-Ferrand II.

[Villèle et Verlaque 1995] – Villèle X. de et Verlaque M., « Changes and degradation in a *Posidonia oceanica* bed invaded by the introduced tropical alga *Caulerpa taxifolia* in the North Western Mediterranean », *Bot, Marina*, (38), pp. 79-87.

[Von Bertalanffy 1987] – Von Bertalanffy L., *Théorie Générale des Systèmes*, Paris, Dunod (Trad, fr, 1973, original en 1968).

W.

[Wayner 1995] – Wayner P., *Agents Unleashed: A Public Domain Look at Agent Technology*, AP Professional.

[Welch 1983] – Welch P. D., « The Statistical Analysis of Simulation Results », in : Lavenberg S. S. (éd.), *The Computer Performance Modeling Handbook*, New York, Academic Press.

[Wooldridge et Jennings 1997] – Wooldridge M. et Jennings N., « Intelligent Agents: Theory and Practice », *Knowledge Engineering Review*.

201

Y.

[Yodziz 1989] – Yodzis P., *Introduction to theoretical ecology*, New York (etc.), Harper & Row, 384 p.

[Yonezawa et al. 1987] – Yonezawa A., Shibayama E., Takada T., et Honda Y., « Modelling and programming in an Object-Oriented Concurrent Language: ABCL/1 », *Object-Oriented Concurrent Programming*, MIT Press, pp. 89-106.

[Youngblood et Pace 1995] – Youngblood S. M. et Pace D. K., « An overview of Model and Simulation Verification, Validation and Accreditation », *John Hopkins APL Technical Digest*, vol. 16, n° 2, pp. 197-205.

Z.

[Zeigler 1976] – Zeigler B. P., *Theory of Modeling and Simulation*, New York, Wiley Interscience.

[Zeigler 1979] – Zeigler B. P. « Multi-level Multiformalism Modelling: An Ecosystem Example », in : Halfton E. (éd.), *Theoretical Systems Ecology*, Academic Press.

- [Zeigler 1984]** – Zeigler B. P., *Multifacetæed modelling and discrete event simulation*, Londres, Academic Press.
- [Zeigler 1990]** – Zeigler B. P., *O-O Simulation with hierarchical modular models: Intelligent agents and endomorphic systems*, Londres, Academic Press.
- [Zeigler et al. 2000]** – Zeigler B. P., Praehofer H., Kim T. G., *Theory of Modeling and Simulation, 2nd Ed. : Integrating Discrete Event and Continuous Complex Dynamic Systems*, Academic Press.

TABLE DES FIGURES

Figure 1	Hiérarchie des niveaux d'abstraction en modélisation d'écosystèmes	36
Figure 2	Vue générale du processus de modélisation par objets	46
Figure 3	Relations hiérarchiques entre les différents modèles	47
Figure 4	Les phases de la méthode M2PO	49
Figure 5	Les différentes catégories de résultats	51
Figure 6	Une matrice de listes chaînées d'objets à simuler	65
Figure 7	Exemple de différentes structures de données possibles pour gérer les données spatiales au sein d'une simulation à événement discrets	66
Figure 8	Circonférence foliaire, zone de dispersion des graines et zone d'influence	71
Figure 9	Couverture foliaire d'une forêt	72
Figure 10	Répartition des graines autour du tronc	73
Figure 11	Modèle de forêt	73
Figure 12	Écrans de couvertures foliaires calculées avec sélection d'une clairière et calcul de sa surface	75
Figure 13	Écran de simulation visuelle interactive pour la croissance d'une forêt	76

Figure 14	Carte des zones atteintes par la Caulerpe en 1997	78
Figure 15	Photo de l'algue <i>Caulerpa taxifolia</i>	80
Figure 16	Situation initiale dans le trou de bombe sur une zone de 60 x 60 m.	81
Figure 17	Après un an de simulation, 58.4 m ² sont colonisés	82
Figure 18	La même zone est étudiée sur une échelle de 500 x 600 m. Cette simulation donne un résultat de 5 240 m ² colonisés après 5 ans	83
Figure 19	Deuxième réPLICATION en rapport avec la Figure 17, 62.1 m ² colonisés	83
Figure 20	1 000 réPLICATIONS (distribution des surfaces sur un an de simulation)	84
Figure 21	10 000 réPLICATIONS dans les mêmes conditions	84
Figure 22	Fenêtre de saisie de quelques paramètres du modèle SIMCT	85
Figure 23	Un herbier de Posidonie souvent comparé à un champ de poireaux avec de longues feuilles de plusieurs dizaines de centimètres	86
Figure 24	Un herbier de posidonie envahi par deux strates superposées d'algue <i>Caulerpa taxifolia</i>	86
Figure 25	À gauche la situation initiale dans la rade de Passable et à droite une hypothétique recolonisation par l'herbier avec une croissance de 3 cm par an sur 100 ans	87
Figure 26	Simulation sur 7 ans avec prise en compte de l'introduction de l'algue tropicale	88
Figure 27	Photo d'un Ascoglosse <i>Elysia Subornata</i> présentant son meilleur profil	90
Figure 28	Un métamodèle représentant les multi-modèles avec un diagramme de classe UML	91
Figure 29	Extrait du diagramme de classe UML pour le modèle de lutte biologique	91
Figure 30	Écran de simulation sur le site de Le Brusc	92
Figure 31	Résultat de 8 années de simulation entre Menton et Villefranche-sur-Mer	95
Figure 32	Courbe d'expansion de <i>Caulerpa taxifolia</i> en fonction du temps (exprimé en mois)	96
Figure 33	Interface graphique développée avec Tcl/Tk pour paramétriser le logiciel de simulation	104

Figure 34	Exemple de tracé fourni par l'application « viewbol »	105
Figure 35	Une capture de l'environnement de travail	106
Figure 36	Courbes représentant le nombre de bols trouvés en moyenne par les 3 agnelles	107
Figure 37	Garance et Marguerite, équipés de leur balise GPS et de leur collier ETHOSYS	109
Figure 38	Diagramme des classes UML du modèle conceptuel de Ternant	111
Figure 39	Diagramme UML des classes implémentées dans le modèle	114
Figure 40	Simulation multi-agents en cours d'exécution	115
Figure 41	Maillage d'une vache en 3 dimensions et images de synthèse d'animaux sur le modèle du terrain de Ternant	117
Figure 42	Capture d'écran d'une simulation de proies et prédateurs co-habitant dans un espace restreint	123
Figure 43	L'application proie-prédateur dans un modèle tridimensionnel	124
Figure 44	Pages Web pour la configuration de comportements et l'affichage de résultats, et une fenêtre de l'applet de la simulation	126
Figure 45	Architecture du cadriel MAVIS au plus haut niveau d'abstraction	128
Figure 46	Extrait d'un diagramme de classes UML pour la réalisation d'environnements de modélisation d'écosystèmes	128
Figure 47	Classification et séquence des tests d'un logiciel de simulation	136
Figure 48	Adaptation de la taxonomie des techniques de validation de Mazel	138
Figure 49	Exemple de structure spatiale qui n'apparaît plus dans le résultat final	148
Figure 50	Situation initiale dans le trou de bombe couvert par <i>Caulerpa</i> ; après un an de simulation ; 2 ^{ème} réPLICATION	149
Figure 51	Identification d'une zone atteignable	150
Figure 52	Degradié montrant le spectre des probabilités de colonisation	151
Figure 53	Visualisation d'un spectre en 3 dimensions	152

TABLE DES FIGURES

Figure 54	Simulation dans le port de la Darse (5 ans)	153
Figure 55	La tâche en diagonale au sein du port donne la zone de <i>Caulerpe</i> cartographiée (5 ans)	153
Figure 56	Visualisation du spectre pour 5 ans de simulation (256 réplications)	153
Figure 57	Spectres en 3D de simulation de la répartition de l'algue <i>Caulerpa taxifolia</i> et carte des zones colonisées par l'algue en 2D	155
Figure 58	Use-case UML correspondant à ce que nous avions proposé en 2003	156
Figure 59	Exemple d'utilisation de SimExplorer précisant pour 2 facteurs une liste de niveaux	158
Figure 60	Exemple de suivi et de simulation de propagation de feu avec tison en Corse	171
Figure 60	Exemple de suivi et de simulation de propagation de feu avec tison en Corse	171
Figure 61	Exemple de rendu presque réaliste avec des techniques graphiques rapides et élémentaires adaptées à une visualisation sur tout type d'ordinateur personnel	172
Figure 62	Exemple de visualisation de simulation de la propagation du virus de la grippe aviaire en Corse	172