BULLETIN

DE

L'AFIA

ASSOCIATION FRANCAISE POUR
L'INTELLIGENCE ARTIFICIELLE

JANVIER 2012

N° 75

Association Française pour l'Intelligence Artificielle

Présentation du bulletin

Le **Bulletin de l'Association Française pour l'Intelligence Artificielle** vise à fournir un cadre de discussions et d'échanges au sein de la communauté universitaire et industrielle. Ainsi, toutes les contributions, pour peu qu'elles aient un intérêt général pour l'ensemble des lecteurs, sont les bienvenues. En particulier, les annonces, les comptes rendus de conférences, les notes de lecture et les articles de débat sont très recherchés. Le Bulletin de l'AFIA publie également des dossiers plus substantiels sur différents thèmes liés à l'IA. Le comité de rédaction se réserve le droit de ne pas publier des contributions qu'il jugerait contraire à l'esprit du bulletin ou à sa politique éditoriale. En outre, les articles signés, de même que les contributions aux débats, reflètent le point de vue de leurs auteurs et n'engagent qu'eux-mêmes.

Pour contacter l'AFIA

Président
Yves DEMAZEAU
L.I.G./C.N.R.S., Maison Jean Kuntzmann, 110, avenue de la Chimie, B.P. 53, 38041 Grenoble cedex 9
Tel : +33 (0)4 76 51 46 43
Fax : +33 (0)4 76 51 49 85
Yves.Demazeau@imag.fr
http://membres-lig.imag.fr/demazeau

Serveur WEB
http://www.afia.asso.fr

Adhésions, liens avec les adhérents
Thomas GUYET
Laboratoire Informatique d'Agrocampus-Ouest
65, rue de Saint-Brieuc
35042 Rennes cedex
Mél. : tresorier@afia.asso.fr

Membres d'honneur

Marie-Odile Cordier (1999), Jean-Paul Haton (1999), Jacques Pitrat (1999), Jean-Marc David (2000), Daniel Kayser (2000), Claude Vogel (2000), Henri Farreny (2001), Alain Colmerauer (2002), Jean-Louis Laurière (2002), Gérard Sabah (2003), Jean-Claude Latombe (2004), Yves Kodratoff (2004), Malik Ghallab (2005), Marie-Christine Rousset (2005), Christian Bessière (2006), Luis Farinas del Cerro (2006), Pierre Marquis (2009), Jérôme Lang (2009), Michèle Sebag (2010), Jean-Gabriel Ganascia (2010), Serge Dupuy (2011).

Personnes morales adhérentes à l'AFIA

ENSMSE, Université Paris Dauphine, LORIA, LIRIS, LIMSI, IRIT/SMAC, EDF/STEP, LIPADE, IFFSTAR, LIRMM, TAO, LIFL, GREYC, LIG, ONERA, IRSTEA-TETIS, INRA, LITIS

Conseil d'Administration de l'AFIA

Yves DEMAZEAU, président
Jean-Denis MULLER, vice-président
Olivier BOISSIER, secrétaire
Catherien TESSIER, secrétaire adjoint
Thomas GUYET, trésorier
Patrick REIGNIER, webmestre
Amélie CORDIER, webmestre adjoint
Membres :
Patrick ALBERT, Christine BOURJOT, Serge GARLATTI, Vincent LEMAIRE, Nicolas MAUDET, Philippe MORIGNOT, Gérald PETITJEAN, Jean ROHMER, Dominique SCIAMMA, Laurent VERCOUTER, Pierre ZWEIGENBAUM.

Comité de Rédaction

Laurence Capus
Rubrique « I.A. au Québec »
Dept. d'Informatique et de Génie Logiciel
Université Laval, Québec, Canada
laurence.capus@ift.ulaval.ca

Nicolas Maudet
Rubrique « I.A. sur les blogs »
LIP6, Université Pierre et Marie Curie
4, place Jussieu, 75005 Paris
maudetn@lip6.fr

Philippe Morignot
Rédacteur en chef
EPITA
14, rue Voltaire, 94270 Le Kremlin-Bicètre
pmorignot@yahoo.fr

Patrick Reignier
Rubrique « Résumés de thèse et HDR »
PRIMA, INRIA Rhône-Alpes
655, avenue de l'Europe, 38334 Saint-Ismier cedex
Patrick.Reignier@inrialpes.fr

Laurent Vercouter
Rédacteur en chef adjoint
LITIS, INSA de Rouen
avenue de l'université, BP8
76801 St-Étienne-du-Rouvray
laurent.vercouter@insa-rouen.fr

EDITORIAL

I.A. et Agronomie

Ce numéro de printemps est consacré à une revue des travaux menés dans plus de dix laboratoires français sur le thème de l'I.A. et de l'Agronomie. Les contributions de l'Intelligence Artificielle au domaine de l'agronomie y sont plurielles, répondant à des besoins en aide à la décision, ingénierie des connaissances, extraction d'information et en apprentissage. Remercions Marie-Odile Cordier (Université de Rennes1, IRISA) et Thomas Guyet (AgroCampus Ouest, IRISA) pour avoir préparé un panorama aussi complet.

Précédant ce dossier spécial, nos lecteurs trouveront la rubrique I.A. au Québec préparée par Laurence Capus (Université Laval). Elle est consacrée dans le présent numéro à une présentation des formations universitaires au Québec relatives à l'I.A.

Lors des dernières journées francophones en Systèmes Multi-Agents (JFSMA), le prix du meilleur papier, sponsorisé par l'AFIA, a été décerné à Patricia Everaere, Maxime Morge (LIFL) et Gauthier Picard (EMSE) pour une proposition d'algorithme distribué de résolution du problème d'appariemment stable. Nos lecteurs trouveront un résumé de la contribution primé dans ce bulletin. Félicitatons aux primés.

Rappellons la nouvelle adresse (depuis 2012) du site web de notre association, relooké et simplifié par rapport à celui que nous vous avons proposé pendant des années :

http://www.afia.asso.fr

Vous y trouverez des brèves nationales et internationales, des supports de cours, des petites annonces, et beaucoup d'autres choses encore.

Notons enfin que les nouvelles conditions d'adhésion à l'AFIA (appliquées depuis le début de l'année 2012) sont résumées dans la fiche d'adhésion présente à la fin de ce bulletin.

Philippe Morignot & Laurent Vercouter
Rédacteurs en chef

I.A. au Québec

Nouvelles du Québec
Laurence Capus (Université Laval)

Quel est la place de l'intelligence artificielle dans la formation universitaire au Québec ? Voici quelques données issues des universités francophones pour avoir un aperçu de cette place.

La formation universitaire au Québec se compose d'un premier diplôme de 1er cycle appelé baccalauréat, puis d'un diplôme de 2ème cycle, la maîtrise, et enfin d'un diplôme de 3e cycle, le programme de doctorat (Ph.D.). D'autres diplômes plus spécifiques sont proposés pour compléter une formation passée, mettre à jour ses connaissances, aller chercher d'autres compétences, prenant la forme d'un microprogramme appelé souvent certificat.

Dans le domaine de l'informatique, plusieurs diplômes sont disponibles offrant une formation générale ou plus spécifique comme l'informatique de gestion, le génie logiciel, la bio-informatique, le génie informatique, etc. Une spécificité peut également se traduire par une concentration dans le programme général (sécurité informatique, systèmes intelligents, génie logiciel, etc.).

Quel que soit le diplôme de baccalauréat, au moins un cours d'intelligence artificielle est proposée reprenant les concepts de recherche heuristique, représentation des connaissances, systèmes à base de connaissances ou encore traitement automatique des langues naturelles. Bien sûr, un tel cours est plus ou moins obligatoire selon la spécificité de la formation. Lorsque l'université propose une concentration plus axée sur l'intelligence artificielle ou la recherche de manière générale, d'autres cours plus avancés sont offerts. Les domaines traités sont l'apprentissage automatique, la planification, l'informatique cognitive, la gestion des connaissances, les systèmes multi-agents.

Ce sont dans les programmes de 2e et 3e cycles qu'on retrouve le plus de formations en intelligence artificielle et ceci en fonction des expertises des professeur(e)s de l'unité de formation. En effet, à ce niveau d'études, les professeur(e)s proposent des cours dans leur domaine de recherche et les étudiant(e)s se spécialisent en fonction de leur projet de recherche.

La plupart des universités ont des professeur(e)s spécialisé(e)s dans le domaine de l'intelligence artificielle. L'Université de Sherbrooke (`www.usherbrooke.ca`) et l'Université de Montréal (`www.umontreal.ca`) semblent les mieux pourvues, elles offrent beaucoup de cours dans ce domaine quel que soit le diplôme.

Dossier IA et Agronomie

L'intérêt sociétal pour la préservation de l'environnement et la qualité des produits agricoles conduit les décideurs à s'intéresser de manière prioritaire au développement d'une agriculture durable. Ces réflexions sont nourries par les travaux menés dans des laboratoires de recherche en agronomie. Comme toutes les sciences du vivant, l'agronomie est face à des problèmes qui impliquent des données complexes, imprécises, incomplètes et hétérogènes. Si on ajoute que ces problèmes prennent généralement en compte des acteurs humains, le terrain ainsi décrit est propice à la mise en place de méthodes d'intelligence artificielle pour apporter des solutions originales et pertinentes à ces problèmes.

Dans ce dossier, nous recensons les activités de recherche française où l'intelligence artificielle est utilisée pour aider à résoudre des problèmes agronomiques en vraie grandeur. Nous nous sommes ainsi focalisés sur les équipes de recherche dont l'activité repose sur une collaboration approfondie entre les deux disciplines informatique et agronomie. Le terme "agronomie" doit ici être compris comme la science des agro-systèmes, qui cherche à étudier et mieux gérer un territoire, en intégrant dans un même système complexe le comportement des différents acteurs et les processus biophysiques.

Plus d'une dizaine de laboratoires ont participé à ce dossier, provenant de structures de recherche très diversifiées : CIRAD [1], écoles supérieures d'agronomie, INRA [2], INRIA, IRSTEA [3], universités. La localisation géographique de ces équipes est également très diversifiée. On notera en particulier la présence de deux laboratoires d'outre-mer (Nouvelle-Calédonie et Réunion). Vous trouverez pour chacun des laboratoires les coordonnées de la personne à contacter, la problématique de recherche, la description des projets et travaux en cours, ainsi que les principales publications.

Dans ce dossier, le lecteur pourra constater que deux grandes problématiques sont traitées par la plupart des équipes de recherche : l'aide à la décision à partir de données complexes et l'information spatio-temporelle.

Dans une première partie du bulletin, nous avons regroupé les équipes qui s'intéressent aux **outils d'aide à la décision par la modélisation et la simulation d'agro-systèmes** (équipes BIA, DREAM, ITAP, RECORD, Recyclage et Risque) dans un but d'évaluation des stratégies de gestion d'un territoire. Parmi les thématiques communes à ces travaux, la représentation des décisions et le raisonnement sur celles-ci sont particulièrement développés.

Dans une seconde partie, nous avons réuni les équipes dont les travaux cherchent à aider à la décision par l'exploitation de l'information disponible sur les agro-systèmes. Leurs thématiques principales sont l'**ingénierie des connaissances et des systèmes** (équipes Copain, FruitQual, Met@risk, MISTIC, SISO). Les données acquises dans le cadre des activités agronomiques sont de nature complexe. Elles sont généralement imprécises, hétérogènes, de nature spatio-temporelle avec plusieurs échelles. Les méthodes développées dans ces équipes sont la représentation floue (MISTIC, Met@risk), les ontologies (FruitQual, Met@risk), les entrepôts de données (SISO, Copain).

Nous avons enfin rassemblé les équipes et projets dont les activités sont principalement l'**extraction d'information et l'apprentissage automatique** (équipes AAA, PPME et projets Foster, Fresqueau et Payote2). Les travaux menés dans ces équipes ont pour point commun de prendre en compte la nature spatio-temporelle des données traitées ainsi que de l'information acquise.

Le lien entre l'intelligence artificielle et les problématiques agronomiques reste à renforcer. Nous espérons que ce bulletin sera une première occasion de le faire, et peut-être le premier pas vers la création d'une communauté sur ce thème, permettant d'augmenter les collaborations et les échanges entre les différentes équipes.

Nous tenons à remercier les contributeurs à ce dossier qui ont répondu positivement à notre sollicitation et ont réalisé un travail spécifique pour l'AFIA.

Marie-Odile Cordier (Université Rennes1/IRISA-Dream) et Thomas Guyet (Agrocampus Ouest/IRISA-Dream)

1. CIRAD : Centre de coopération internationale en recherche agronomique pour le développement.
2. INRA : Institut National de la recherche Agronomique.
3. IRSTEA : Institut national de Recherche en sciences et technologies pour l'environnement et l'agriculture (anciennement CEMAGREF).

UMR ITAP (Irstea-SupAgro Montpellier), équipe « Modélisation et Décision Agro-environnementale »

Ce groupe mène des recherche sur l'aide à la décision. Il appartient à l'UMR ITAP (Information, Technologies, Analyse environnementale, Procédés agricoles) des organismes Irstea et Supagro.

- CONTACTS : Irstea : prenom.nom@irstea.fr, Supagro : nom@supagro.inra.fr
- ADRESSE : 361, rue J.F. Breton - BP 5095 - 34196 Montpellier Cedex 5, France
- WEB : `https://itap.cemagref.fr` → `http://itap.irstea.fr/`

Membres permanents de l'équipe

- Serge Guillaume, Ingénieur de recherche Irstea, HdR
- Jean-Luc Lablée, Ingénieur d'étude Irstea
- Olivier Naud, Ingénieur de recherche Irstea
- Bernard Palagos, Ingénieur de recherche Irstea
- Bruno Tisseyre, Maître de conférences SupAgro, HdR

Thème général de l'équipe

Les systèmes complexes souffrent d'un manque de modèles de simulation et de prévision qui prennent en compte la multitude des paramètres d'influence et leur variabilité. La gestion des systèmes agro-environnementaux complexes peut néanmoins s'appuyer sur une connaissance opérationnelle qui est disponible du fait de l'action permanente de l'homme. Les outils d'aide à la décision ont vocation à intégrer l'ensemble de cette connaissance opérationnelle, qu'elle soit issue de l'élicitation de l'expertise ou extraite à partir des données. La conception de ces systèmes d'aide à la décision est l'objet de notre contribution scientifique.

A cet égard, nous pouvons énoncer quelques spécificités s'appliquant aux domaines de l'agronomie et de l'environnement. Les phénomènes observés sur lesquels l'homme agit sont spatio-temporels, par exemple *la zone de haut rendement pour l'année 2011*. Les systèmes de décision doivent travailler à plusieurs échelles : parcelle, exploitation, bassin versant, appellation contrôlée. Les données disponibles sont acquises avec des résolutions spatio-temporelles différentes, très élevées pour les capteurs embarqués ou les images aériennes, très faibles pour les mesures manuelles coûteuses comme les analyses de sol ou la mesure du potentiel hydrique des plantes. La variabilité qui caractérise le monde biologique induit une incertitude qu'il convient de prendre en compte. Enfin, les systèmes de décision doivent valoriser l'expertise humaine disponible.

Travaux et projets

Notre champ d'application privilégié est celui de la viticulture. Les dimensions spatio-temporelles sont appréhendées au plan opérationnel par les notions de zone et d'étape. Nous avons choisi de présenter trois exemples de travaux. Le premier concerne ce que nous appelons un processus opérationnel de décision, appliqué à la protection phytosanitaire des cultures, dans le but de réduire le nombre des applications tout en assurant une production de qualité. Il illustre la prise en compte de la dimension temporelle de la décision. Le second propose une modélisation de la structure spatiale du stress hydrique de la vigne afin d'extrapoler une mesure locale, coûteuse, sur un domaine spatial plus large. Enfin, le dernier présente des logiciels libres au service de la coopération entre l'expertise et les données, basés sur la logique floue, ainsi que des exemples de leur utilisation en agronomie et environnement.

Processus Opérationnels de Décision (POD) pour la protection des cultures

Nos recherches sur la modélisation temporelle de la décision agronomique par systèmes à événements discrets (SED) ont débuté en 2005 sur des questions d'optimisation du nombre de traitements fongicides sur vigne. Elles sont menées sur vigne avec l'INRA, et avec d'autres partenaires académiques et économiques, par exemple avec Arvalis pour la protection fongicide du blé tendre.

La première question qui se posait à nous était la pertinence des SED pour représenter un système décisionnel à base d'expertise. C'est ainsi qu'une élicitation de connaissances, dans le langage graphique et formel des Statecharts, a été menée auprès d'experts pathologistes ayant conçu une manière de décider les traitements contre l'oïdium et le mildiou de la vigne [4]. Les Statecharts sont un langage permettant la composition d'automates, au sens de la hiérarchie et au sens des processus concurrents. Les Statecharts font partie de la norme UML2.

L'élicitation a été faite directement dans le langage ciblé pour la représentation et la simulation du système décisionnel, pour favoriser, a priori et au plan déclaratif, la fidélité du modèle à l'expertise. Pour mieux nous assurer de cette fidélité, le modèle a été simulé sur la base de scénarios bioclimatiques et comparé avec les décisions en situation des experts [5]. Cette comparaison a conclu à la bonne qualité du modèle, qui est devenu référence pour la rédaction

de protocoles de tests agronomiques de ce qui a été appelé « POD Mildium ».

Cependant, cette comparaison a également permis de constater que les experts optimisaient leur procédé de décision, et notamment le positionnement temporel des traitements, grâce à l'anticipation. Or, la simulation du modèle Statechart correspond à la simulation d'un système réactif, l'action consécutive à la décision étant, dans nos hypothèses d'alors, supposée réalisée « le plus vite possible », sans tenir compte des conditions de mise en œuvre. En effet, alors que le modèle décisionnel simule la décision parcelle par parcelle pour une décision précise optimisée, la mise en œuvre d'un traitement suppose l'utilisation d'un pulvérisateur permettant de traiter à la suite plusieurs hectares, c'est-à-dire plusieurs parcelles. Ceci soulevait entre autres la nécessité d'une modélisation temporelle plus précise, permettant de représenter les dates au plus tôt et au plus tard de l'exécution de chaque décision, ainsi que la nécessité d'un mécanisme de décision pour regrouper les traitements décidés d'une façon optimale pour l'exploitant. Enfin, il nous est apparu que la terminologie et les outils de « workflow » étaient de bons candidats pour mettre à disposition de conseillers agricoles et d' agronomes expérimentateurs, des modèles comme le prototype « POD Mildium ». La terminologie « Processus Opérationnels de Décision » tire d'ailleurs son origine de la traduction de *decision workflow*.

Les automates temporisés, ainsi que notamment les réseaux de Petri temporels, permettent de dépasser le cadre du modèle réactif et de représenter les contraintes permettant d'établir, in fine, par exemple grâce à un model-checking adapté [4] les dates au plus tôt et au plus tard d'un événement tel que la mise en œuvre d'une décision de traitement. Dans le cadre d'une hypothèse de modélisation par automates temporisés, nous avons établi que l'optimisation pouvait se décrire comme un problème MTVRPTW (Multi-Trip Vehicle Routing Problem with Time Windows), et nous avons fourni, en collaboration avec le LIRMM et le CMP de l'Ecole des Mines de Saint-Etienne, le premier algorithme exact permettant de résoudre ce problème [3], dont les applications urbaines potentielles sont au moins aussi nombreuses que nos questions d'agriculture durable.

Un simulateur de POD pour la protection fongicide du blé, conçu avec Arvalis, et grâce à son expertise, a été réalisé dans le formalisme des réseaux de Petri colorés, avec l'outil CPN tools.

Nous avons donc identifié un ensemble de formalismes de modélisation à événements pertinents pour modéliser des procédures décisionnelles et ainsi faciliter leur validation et leur diffusion. Les problèmes décisionnels agronomiques les plus tactiques, notamment la protection des cultures, soulèvent des questions de représentation de logique temporelle, de planification dans l'incertain, d'optimisation. Les formalismes mobilisés relèvent du croisement des disciplines de l'Informatique, de l'Automatique et de la Recherche Opérationnelle. Parce qu'expertise, actions humaines et organisation d'entreprise sont impliquées, nos développements méthodologiques en gestion de la connaissance et calcul décisionnel participent des applications de l'Intelligence Artificielle au domaine de l'agronomie.

Modélisation spatiale : application à l'état hydrique de la vigne

L'état hydrique de la vigne est un paramètre qui varie de manière importante dans le temps en fonction du climat de l'année mais aussi dans l'espace en fonction des hétérogénéités de facteurs environnementaux. La connaissance de ses variations spatiales et temporelles représente un enjeu important pour la filière viticole. En effet, l'état hydrique, et son évolution aux différents stades de la culture, détermine le potentiel quantitatif et qualitatif de la récolte. Cette variable intégratrice et très informative permet de déclencher des actions correctives. Elle est ainsi à la base d'une gestion optimale de la qualité du raisin et de la ressource en eau dans le cas de l'irrigation. Au-delà des enjeux de cette filière, savoir spatialiser les évolutions d'états hydriques de la vigne sur les espaces cultivés est un pré-requis incontournable pour diagnostiquer l'état actuel de la ressource en eau.

L'objectif de cette recherche est de proposer un modèle d'estimation de l'état hydrique de la vigne dans le temps et dans l'espace. Afin de produire une aide à la décision opérationnelle pour le pilotage de la culture, les échelles visées sont celle de l'exploitation et celle de l'aire de production (cave coopérative par exemple). Ce modèle d'estimation met en synergie une nouvelle configuration de données, dont la disponibilité est désormais envisageable à ces échelles grâce aux avancées technologiques récentes : i) un réseau de mesures locales géo-référencées de suivi de l'état hydrique de la vigne (ii) des données à haute résolution relatives à la plante (images aériennes proche infra-rouge) ou au sol (conductivité électrique apparente des sols).

Une démarche originale a été proposée [1]. L'approche fait collaborer un modèle de la structure spatiale, préalablement déterminé par une base de données de mesures ou

4. Pour des exemples de model-checking adapté à la recherche de dates pour une situation donnée, on pourra par exemple se reporter aux travaux en modélisation qualitative suivants : C. Largouët, M-O. Cordier, Y-M. Bozec, Y. Zhao, and G. Fontenelle. Use of timed automata and model-checking to explore scenarios on ecosystem models. *Environmental Modelling and Software*, 30 :123–138, 2012.

par des mesures à haute résolution spatiale disponibles, et une mesure de référence qui caractérise l'état du système à l'instant t. L'approche proposée est résumée par l'équation 1. Elle consiste, sur un domaine D, à effectuer une estimation $\hat{z}(s_i, t_j)$ de l'état hydrique des plantes en un site s_i et à une date t_j à partir d'une mesure de référence $z(s_{re}, t_j)$ réalisée sur un site de référence s_{re} à la même date t_j. L'estimation est basée sur une fonction $f(s_i)$ dont les variables sont des mesures stables dans le temps qui caractérisent le site s_i.

$$\hat{z}(s_i, t_j) = f(s_i) \times z_{re}(s_{re}, t_j); s_{re} \in D, \forall s_i \in D \qquad (1)$$

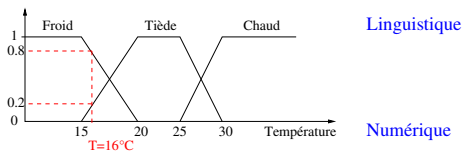
Cette approche a été appliquée avec succès sur deux parcelles, sur plusieurs années avec des cépages différents. Elle a également été mise en œuvre avec succès à l'échelle d'un bassin de production. La figure 1 présente l'extrapolation d'une mesure de référence sur l'aire d'apport d'une coopérative, avec une qualité d'estimation sur les points de mesures non utilisés pour l'apprentissage de 0.10 Mpa [6]. La méthode proposée s'est avérée peu sensible au choix du site de référence et permet d'expliquer une part de la variabilité spatiale d'autant plus importante que le niveau de contrainte hydrique est élevé.

Des logiciels pour la coopération entre expertise et données

Cette partie est conduite en étroite collaboration avec Brigitte Charnomordic, INRA, UMR MISTEA.

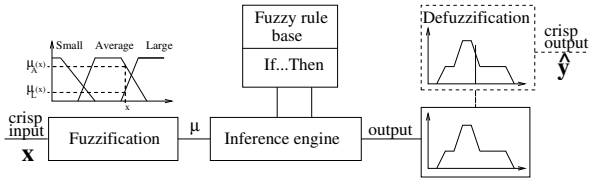
En sciences du vivant, l'incertitude provient à la fois de la variabilité biologique et de l'imprécision des concepts utilisés, comme celui de *zone inondable*.

La logique floue a été choisie comme cadre de représentation pour sa capacité à modéliser l'incertitude. Elle constitue une interface commode entre les mondes symbolique, celui du raisonnement humain, et numérique, celui du calcul et de l'informatique. La notion de variable linguistique est au cœur des systèmes d'inférence floue.



Les termes linguistiques utilisés par le raisonnement, tels que *Faible dose* ou *Rendement élevé*, sont modélisés par des ensembles flous. L'interface avec le numérique se fait par le degré d'appartenance d'une valeur, 16°C sur la figure, à chacun des ensembles de la partition.

L'ensemble des règles de raisonnement forme le moteur d'inférence du système. Leur caractère approché permet de qualifier des situations ou de prendre des décisions d'action à partir d'informations imprécises.



Les systèmes d'inférence floue présentent un double avantage : ils permettent de représenter simplement la connaissance experte, et ce sont également des approximateurs universels, capables de représenter des relations entre des valeurs d'entrée et de sortie avec une grande précision.

La recherche de la meilleure précision ne doit pas conduire à sacrifier ce qui fait la spécificité de ces systèmes : leur interprétabilité, qui est une condition nécessaire pour la coopération entre l'apprentissage automatique et l'expertise.

Le logiciel libre *FisPro* est le fruit de dix années de recherche dans ce domaine : il implémente des méthodes de conception et d'optimisation de ces systèmes dont l'originalité est de garantir la sémantique [2].

Son interface est disponible en quatre langues, ses modules pédagogiques sont appréciés. Il possède de puissantes fonctionnalités au service de l'interaction homme-machine : possibilité de désactiver des variables, des règles ou des exemples de données, répercussion dynamique de toute modification de paramètre, indices d'évaluation quantitative de la performance du système, outils d'appréciation de son comportement (surface de réponse, liens entre règles et exemples). Proposé sous la forme d'une bibliothèque de fonctions autonomes, il peut être intégré dans d'autres programmes.

Parmi ses utilisations en agronomie et environnement, citons notamment : l'étude de l'impact de la combinaison spatio-temporelle des pratiques viticoles sur la qualité de l'eau à l'échelle du bassin versant (Colin, 2011) ; la modélisation des interactions entre composantes de la durabilité d'un agro-système (Rajaram, 2010) , la conception d'un indicateur de la vigueur de la vigne (Coulon, 2010) ou encore la caractérisation des zones de fonctionnement agronomique à partir de données viticoles (Grelier, 2007). La base de règles peut intégrer des connaissances issues de trois types d'information complémentaires : une série temporelle d'images satellite, un modèle de croissance de la plante, et la connaissance experte, pour suivre en continu

5. Les références mentionnées sont disponibles sur le site de FisPro http://www.inra.fr/mia/M/fispro

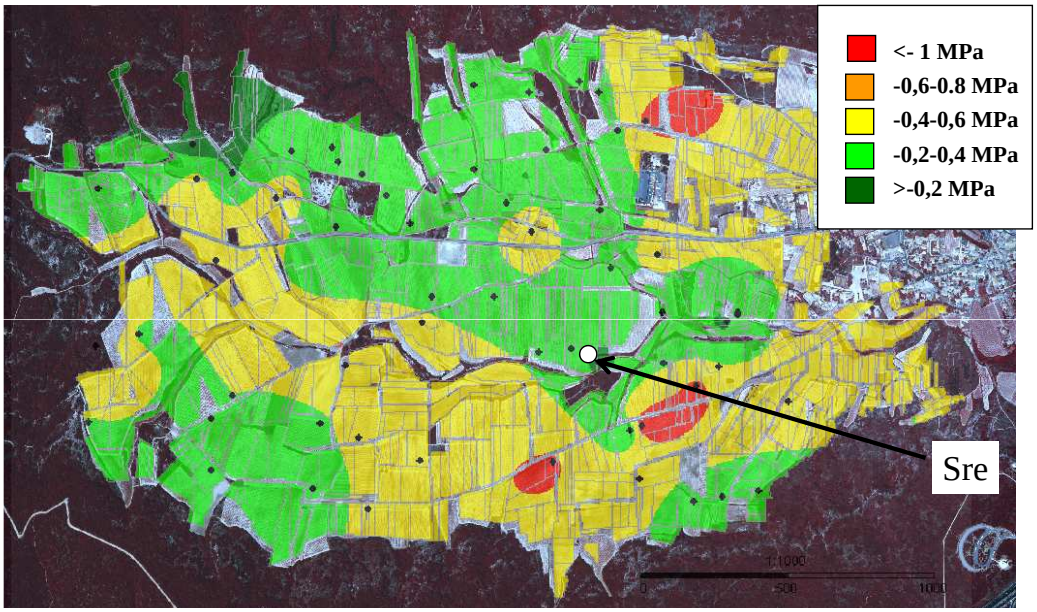


FIGURE 1 – Extrapolation de l'état hydrique de la vigne (potentiel hydrique de base) sur l'ensemble d'un bassin de production (Tavel, Gard) à partir d'une mesure positionnée sur un site de référence (Sre). (Collaboration IFV)

la récolte de la canne à sucre sur l'île de la Réunion (El Hajj, 2009) [5].

Le raisonnement à partir de données spatialisées devient un enjeu majeur dans les domaines de l'agriculture et de l'environnement. Cela s'explique par le vaste déploiement d'applications et de systèmes de mesures géoréférencés comme les images aériennes ou les réseaux de capteurs. Dans ce contexte, la zone apparaît comme l'entité de base de la gestion des données spatialisées, avec la nécessité de méthodes adaptées de représentation et de raisonnement. Cela suppose le développement de concepts, tels que la zone floue ou bien le raisonnement spatialisé approché, mais aussi celui d'outils logiciels spécifiques, pour l'interaction homme machine et la représentation de l'incertitude. Le logiciel en cours de développement, GeoFIS, bénéficiera des acquis de FisPro pour offrir une boîte à outils ouverte et efficace dans ce nouveau domaine.

Plus d'information : https://itap.cemagref.fr/pages-personnelles/serge-guillaume

Références

[1] C. Acevedo-Opazo, B. Tisseyre, J. A. Taylor, H. Ojeda, and S. Guillaume. A model for the spatial prediction of water status in vines (Vitis vinifera L.) using high resolution ancillary information. *Precision Agriculture*, 11(4) :358–378, 2010.

[2] S. Guillaume and B. Charnomordic. Learning interpretable fuzzy inference systems with FisPro. *International Journal of Information Sciences*, 181 :4409–4427, 2011.

[3] F. Hernandez. *Méthodes de résolution exactes pour le problème de routage de véhicules avec fenêtres de temps et routes multiples.* thèse de doctorat, Université Montpellier II, 2010.

[4] B. Léger and O. Naud. Experimenting statecharts for multiple experts knowledge elicitation in agriculture. *Expert Systems with Applications*, 36(8) :11296 – 11303, 2009.

[5] B. Léger, O. Naud, V. Bellon-Maurel, M. Clerjeau, L. Delière, P. Cartolaro, and L. Delbac. GrapeMilDeWS : a formally designed integrated pest management decision process against grapevine powdery and downy mildews. In *Decision Support Systems in Agriculture, Food and the Environment :Trends, Applications and Advances*, pages 246–269. IGI Global, 2010.

[6] J. A. Taylor, C. Acevedo-Opazo, S. Guillaume, H. Ojeda, and B. Tisseyre. A comment on inter-field spatial extrapolation of vine (Vitis vinifera L.) water status. *Journal International Des Sciences De La Vigne Et Du Vin*, 45(2) :121–124, 2011.

UR INRA BIA, équipe MAD « Modélisation des Agro-écosystèmes et Décision »

- CONTACT : Régis Sabbadin regis.sabbadin@toulouse.inra.fr,,
- ADRESSE : INRA. Chemin de Borderouge, BP 52627, Castanet-Tolosan, 31326 Cedex, France.
- WEB : http://carlit.toulouse.inra.fr/wikiz/index.php/Accueil
- TEL : (+33)5.61.28.54.76

Membres permanents de l'équipe

- Stéphane Couture (CR INRA)
- Robert Faivre (DR INRA)
- Sylvain Jasson (IR INRA)
- Roger Martin-Clouaire (DR INRA)
- Nathalie Peyrard (CR INRA)
- Gauthier Quesnel (CR INRA)
- Jean-Pierre Rellier (IR INRA)
- Régis Sabbadin (DR INRA)

À ces membres permanents s'ajoutent en moyenne une demi-douzaine de membres non-permanents de longue durée (doctorants, post-doctorants ou CDD).

Thème général de l'équipe

La gestion des agro-écosystèmes devient un enjeu majeur de développement durable. Elle est appelée à évoluer considérablement pour répondre à des problèmes environnementaux cruciaux et pour faire face à d'importants changements de contexte concernant, par exemple, la politique agricole, la demande des consommateurs, l'aspiration des agriculteurs à une meilleure qualité de vie, ou encore le changement climatique. La réponse aux demandes du développement durable passe par un accroissement des travaux entre sciences de la nature et sciences humaines et sociales et, de plus en plus, les sciences de la formalisation et de l'expérimentation virtuelle. Un besoin croissant se fait sentir d'outils mathématiques et informatiques permettant de :

- mieux comprendre les interactions entre les composantes physiques, biologiques et anthropiques des agro-écosystèmes ;
- faciliter la conception, la mise en oeuvre et l'évaluation des interventions pratiques qui favorisent la durabilité des systèmes agricoles dans des contextes particuliers ;
- améliorer les liens entre la recherche fondamentale et les démarches d'innovation.

Ces outils contribuent ainsi à élaborer de nouvelles manières de produire, d'organiser la production et de gérer l'espace agricole, à différentes échelles, pour concilier compétitivité, qualité des produits, qualité de vie de ceux travaillant sur ces systèmes de production, et respect de l'environnement.

Forte de ses compétences pluri-disciplinaires en intelligence artificielle, recherche opérationnelle, statistiques et économie l'équipe mène des travaux méthodologiques et/ou appliqués dans les domaines suivants :

- Modélisation ;
- Simulation ;
- Optimisation.

Description des travaux

Modélisation

Les systèmes étudiés sont définis par l'interaction entre activités agricoles, processus biophysiques et processus de décision individuelle ou collective, aux échelles de la parcelle agricole, de l'exploitation dans sa globalité (voire d'un territoire plus vaste), pour des horizons temporels allant de la journée à plusieurs décennies. On est ainsi amené à coupler des modèles de dynamique du système piloté (modèles de culture par exemple), de comportements décisionnels (stratégies de conduite par exemple) et, éventuellement, d'exécution des actions décidées.

L'équipe développe des approches de modélisation des comportements décisionnels de pilotage de systèmes de production mettant l'accent sur la prise en compte de l'incertitude, la caractérisation de stratégies robustes et flexibles, et sur la coordination des activités de gestion de production, intégrant des contraintes structurelles, agronomiques, environnementales et économiques [6]. Les recherches portent sur la caractérisation de ces processus décisionnels et sur le développement de formalismes génériques de représentation permettant leur simulation [8] et/ou leur utilisation dans des démarches de conception participative [4] ou d'optimisation. Elles s'appuient et contribuent à des approches issues de l'intelligence artificielle, des sciences économiques, et des sciences des organisations et de la gestion de production.

Simulation

La simulation à événements discrets est largement utilisée dans l'équipe pour l'analyse des agro-écosystèmes et des stratégies de gestion associées. Elle est également un

axe de recherche propre de l'équipe, via le développement des environnements de simulation DIESE [6] et VLE [7].

DIESE [7] repose sur un formalisme original et générique de représentation des systèmes de production et, en particulier, de l'agent décideur en charge du pilotage. L'ontologie développée constitue un méta-modèle à partir duquel peuvent être construits des modèles de système par particularisation des concepts de haut niveau fournis (e.g. opération, activité, plan, ressource). DIESE permet par ces modèles de systèmes de production d'effectuer des expérimentations informatiques de modes de conduite flexibles et sous contraintes de ressources.

Les développements réalisés par l'équipe autour de VLE [10] s'appuient sur le formalisme à événements discrets DEVS (*Discrete Event System Specification*). Ce formalisme, issu des mathématiques discrètes, possède un grand nombre de propriétés mathématiques permettant de modéliser, coupler et simuler des systèmes dynamiques modulaires et hiérarchiques complexes. Les propriétés mathématiques de DEVS permettent également de représenter d'autres formalismes de systèmes dynamiques et de les coupler entre eux. Ces propriétés et ces formalismes de haut niveau se retrouvent dans l'implémentation informatique de DEVS que nous avons réalisée avec VLE conjointement depuis 2003 avec l'université ULCO et le Cirad. Nous proposons un certain nombre de formalismes ainsi que la possibilité de les coupler entre eux. Les formalismes proposés sont soit des formalismes connus comme les réseaux de Petri, les *statecharts* d'UML et les équations différentielles mais aussi des extensions regroupant un certain nombre de formalismes ou fonctions comme l'extension de décision que nous proposons dans VLE et qui regroupe une partie des travaux réalisés dans l'équipe MAD avec un planificateur de tâches, un gestionnaire de ressources et un résolveur de contraintes pondérées pour représenter les agents décisionnels. VLE est un incubateur des travaux méthodologiques réalisés dans l'équipe dans le cadre de la théorie de la modélisation et de la simulation et plus particulièrement sur les systèmes à événements discrets, sur le couplage de systèmes hybrides continus et discrets, sur les travaux de traduction de formalismes et plus récemment sur la parallélisation de la simulation. L'environnement VLE sert, de plus, de support à la plate-forme RECORD [8] (voir page 13).

Optimisation/Conception

L'équipe développe également des recherches à la frontière entre mathématiques (probabilités) et informatique (intelligence artificielle, optimisation combinatoire), sur la formalisation et la résolution de problèmes de décision séquentielle dans l'incertain. Outre le cadre des Processus Décisionnels de Markov sur lequel nous possédons une longue expérience [1], nous développons des outils dédiés à la résolution de problèmes de décision spatialisée, exprimés dans le cadre des modèles graphiques, stochastiques [12] ou déterministes [2]. Dans le cadre de collaboration avec nos partenaires agronomes, écologues ou épidémiologistes, nous les appliquons à la conception de stratégies d'échantillonnage [9] ou de gestion dans les agro-écosystèmes [11, 5, 3, 2].

Projets en cours

Nous sommes (ou avons été) impliqués dans plusieurs projets, en cours ou en construction, dans lesquels des méthodes de l'intelligence artificielle sont mises en œuvre ou développées :

Gestion spatialisée des agro-écosystèmes :
- Projet ANR-ADD « Comment Exploiter Durablement les REsistances (CEDRE) » (2005-2008). Développement d'un cadre basé sur les PDMs spatialisés pour la gestion des résistances spécifiques des cultures aux pathogènes.
- Projet européen « Pesticide Use Resistance in Europe (PURE) » (2011-2014). Travaux dans la suite du projet ANR CEDRE.
- Projet ANR-SYSTERRA « Organismes et Organisations Localement Adaptés » (O2LA) (2010-2013).

Développement de méthodes d'échantillonnage pour la cartographie d'espèces invasives :
- Projet Australian Research Council « Applying search theory for eradicating invasive species » (2007-2009). Développement de méthodes d'optimisation de l'échantillonnage de processus spatiaux appliquées à la reconstruction de cartes d'invasion de l'espèce « fire ants », dans le Queensland.
- Projet ANR « SPatial SAmpling of Pests (SPASAP) » (soumis à l'ANR en 2012). Travaux dans la suite du projet précédent, appliqués à la reconstruction de cartes d'adventices ou de ravageurs des cultures.

Références

[1] *Processus Décisionnels de Markov en Intelligence Artificielle.* Hermes / Lavoisier, 2008.

6. http://carlit.toulouse.inra.fr/diese/
7. http://www.vle-project.org/
8. http://www4.inra.fr/record

[2] M. Akplogan, J. Dury, S. de Givry, G. Quesnel, A. Joannon, A. Reynaud, J.-E. Bergez, and F. Garcia. A weighted CSP approach for solving spatio-temporal planning problem in farming systems. In *Proc. of the Soft'2011 workshop*, pages 1–15, 2011.

[3] M. Akplogan, G. Quesnel, F. Garcia, A. Joannon, and R. Martin-Clouaire. Towards a deliberative agent system based on DEVS formalism for application in agriculture. In *Proccedings of the Summer Computer Simulation Conference (SCSC'10)*, July 2010.

[4] M. Duru and R. Martin-Clouaire. Cognitive tools to support learning about farming system management : a case study in grazing systems.

[5] N. Forsell, P. Wikström, F. Garcia, R. Sabbadin, K. Blennow, and L. Eriksson. Management of the risk of wind damage in forestry : a graph-based Markov decision process approach. *Annals of Operations Research*, 190(1) :57–74, 2009.

[6] G. Martin, R. Martin-Clouaire, J.-P. Rellier, and M. Duru. A simulation framework for the design of grassland-based beef-cattle farms. *Environmental Modelling and Software*, 26 :371–385, 2011.

[7] R. Martin-Clouaire and J.-P. Rellier. Modelling and simulating work practices in agriculture. *Int. J. of Metadata, Semantics and Ontologies*, 4(1) :42–53, 2009.

[8] R. Martin-Clouaire and J.-P. Rellier. Dynamic resource allocation in a farm management simulation. In *Proc. of MODSIM11, International Congress on Modelling and Simulation*. Modelling and Simulation Society of Australia and New Zealand, 2011.

[9] N. Peyrard, R. Sabbadin, D. Spring, B. Brook, and R. Mac Nally. Model-based adaptive spatial sampling for occurrence map construction. *Statistics and Computing*, pages 1–14, 2011.

[10] G. Quesnel, R. Duboz, and E. Ramat. The Virtual Laboratory Environment - An Operational Framework for Multi-Modelling, Simulation and Analysis of Complex Systems. *Simulation Modelling Practice and Theory*, 17 :641–653, 2009.

[11] C.-E. Rabier, R. Sabbadin, and D. Spring. Dynamic reserve site selection under contagion risk of deforestation. *Ecological Modelling*, 201 :75–81, 2007.

[12] R. Sabbadin, N. Peyrard, and N. Forsell. A framework and a mean-field algorithm for the local control of spatial processes. *International Journal of Approximate Reasoning*, 53(1) :66 – 86, 2012.

UR INRA BIA, équipe RECORD

Plate-forme pour la modélisation et la simulation des agro-écosystèmes

- CONTACT : Raynal Hélène Helene.Raynal@toulouse.inra.fr,
- ADRESSE : INRA - Unité BIA - Chemin de Borde Rouge - BP 52627 - 31326 Castanet Tolosan cédex, France
- WEB : http://www.inra.fr/record

Membres de l'équipe

- Patrick Chabrier (IE INRA)
- Hélène Raynal (IR INRA)
- Nathalie Rousse (IR INRA)
- Ronan Trépos (IR INRA)

L'équipe est renforcée par des ingénieurs mis à disposition de l'équipe dans le cadre de projets de recherche financés ou dans le cadre de collaborations INRA.

Thème général de l'équipe

Après quatre ans de conception et de développement, la plate-forme RECORD [3] (« REnovation et COoRDination de la modélisation de cultures pour la gestion des agro-écosystèmes »), dédiée à la modélisation et à la simulation informatique des agro-écosystèmes a été lancée fin 2010. Elle est basée sur le logiciel VLE [6], et dispose d'une bibliothèque de modules et de méthodes. C'est l'équipe Plate-forme RECORD de l'Unité de Biométrie et Intelligence Artificielle qui assure sa maintenance, et son évolution en particulier par l'intégration de développements méthodologiques concernant les formalismes de modélisation et les algorithmes d'optimisation. Parmi ces développements, plusieurs relèvent du domaine de l'intelligence artificielle, et sont des outils importants pour la conception de systèmes de culture innovants, permettant de répondre aux défis actuels adressés à l'agriculture comme l'adaptation au changement climatique ou la diminution de l'impact agricole sur l'environnement.

Description des travaux

Modélisation de la décision

Les processus de décision sont très présents dans la gestion des agro-écosystèmes. Ainsi, l'itinéraire technique d'une culture, qui correspond à l'ensemble des opérations culturales à effectuer et implique des processus de décision, doit être formalisable et modélisable pour la simulation de ces systèmes. Ces processus de décision sont à l'échelle de la parcelle mais il est souvent nécessaire de considérer d'autres échelles de décision comme l'exploitation agricole ou le petit territoire. En collaboration avec l'équipe MAD, l'implantation d'un formalisme dédié à la modélisation de la décision a été réalisé. Ainsi, les processus de décision sont modélisés en s'appuyant sur la spécification d'activités du décideur. Des contraintes d'activation d'activité, de précédence entre ces activités sont exprimées dans ce formalisme. La simulation produit dynamiquement une instanciation du plan d'activités, de manière à être conforme aux contraintes.

Couplage de modèles

Modéliser et simuler les agro-écosystèmes pose le problème de considérer conjointement la simulation de différents types de processus tels que décisionnels et agronomiques, ce qui amène à considérer le couplage de différents formalismes de modélisation. Par ailleurs, les agro-écosystèmes modélisés dans le contexte de RECORD sont majoritairement des systèmes hybrides faisant apparaître des processus à la fois discrets et continus, qu'il est nécessaire de coupler dynamiquement. DEVS (Discrete Event System Specification) [7] est le formalisme sous jacent à VLE qui permet de répondre à ces problèmes de couplage. Par ailleurs, pour faciliter l'implantation des modèles sur RECORD, la traduction en DEVS des formalismes couramment utilisés pour la modélisation des agro-écosystèmes reste un des axes de travail privilégiés de l'équipe.

Exploration de modèles

L'équipe RECORD a également pour objectif de proposer la mise en oeuvre des méthodes d'exploration de modèles. Ce point est traité en collaboration avec l'équipe MAD et le réseau MEXICO (http://www.reseau-mexico.fr/welcome) et concerne l'analyse de sensibilité de modèles, l'optimisation par simulation [1] et l'apprentissage par renforcement [4]. Dans le cadre de RECORD, ces méthodes contribuent de manière significative à l'aide à la prise de décision car elles permettent de déterminer les valeurs optimales des variables de décision.

Des outils et des modèles

En tant que plate-forme ouverte sur une communauté, le rôle de RECORD est également de proposer des outils facilitant la construction, la simulation et l'analyse de modèles. Pour cela, des modules génériques, réutilisables, sont développés et mis à disposition dans une base de données,

accessible via une application web. L'exploration de modèles, telle que l'analyse de sensibilité, requiert des capacités de calcul importantes ainsi que des méthodes d'analyse statistiques spécifiques. C'est pourquoi des outils autour de la parallélisation des simulations et un couplage au logiciel R sont développés et maintenus en collaboration avec l'équipe MAD. Nous avons également le souci de l'ergonomie des outils et proposons des interfaces graphiques dédiées, facilitant le travail du modélisateur.

Projets en cours

La plate-forme RECORD est utilisée dans plusieurs projets de recherche dans lesquels des méthodes d'intelligence artificielle sont mises en oeuvre. Ci-dessous un ensemble de projets représentatifs :

- Participation au projet INRA-UMT Eau « Outils et méthodes pour la gestion quantitative de l'eau du bloc d'irrigation au collectif d'irrigants ». Plusieurs travaux ont été menés dans ce contexte. Un d'entre eux a consisté à réaliser le couplage sous VLE d'un modèle de culture (STICS), d'une representation spatiale du territoire d'une exploitation agricole (posGRES), d'un algorithme d'optimisation par satisfaction de contraintes (Toolbar [5]) pour choisir les assolements des cultures. [2]
- ANR POPSY (2008-2012) : « Systèmes de Production en Grandes Cultures, Environnement Politiques Publiques ». Dans le cadre de cette ANR, nous nous intéressons à construire automatiquement, à l'aide de la simulation, des règles de conduite d'une culture qui permettent l'optimisation d'une sortie du modèle. Pour cela, nous utilisons des algorithmes d'apprentissage par renforcement tels que XITEK [4]. Ce travail est appliqué à un modèle de culture du colza et a permis d'induire par simulation un ensemble de règles de gestion de la culture.
- ANR MicMac (2010-2014) : « Modelling for Integrated Crop Management in low input farming, Assessment and Cropping System Design ». Ce projet a pour originalité d'utiliser en complémentarité de l'expérimentation aux champs, l'expérimentation virtuelle pour la conception de systèmes de culture innovants à bas niveau d'intrants à l'échelle de la parcelle agricole. Un travail particulier de représentation et simulation des pratiques agricoles pour la détermination des itinéraires techniques optimaux a été réalisé.

Références

[1] O. Crespo. *Conception par simulation pour la conduite de culture.* PhD thesis, Université de Toulouse III, 2008.

[2] J. Dury. *Aide au choix d'assolement et de conduite de culture en environnement incertain par modèles dynamiques de simulation.* PhD thesis, Université de Toulouse III, 2011.

[3] Bergez et al. RECORD : an open platform to build, evaluate and simulate integrated models of farming and agro-ecosystems (in revisions). *Environmental Modelling and Software*, 2012.

[4] F. Garcia. Use of reinforcement learning and simulation to optimise wheat crop technical management. In *Proceedings of the International Congress on Modelling and Simulation (MODSIM 99)*, pages 801–806, 1999.

[5] J. Larrosa and T. Schiex. In the quest of the best form of local consistency for Weighted CSP. In *Proceedings of IJCAI*, 2003.

[6] G. Quesnel, R. Duboz, and É. Ramat. The Virtual Laboratory Environment – An operational framework for multi-modelling, simulation and analysis of complex dynamical systems. *Simulation Modelling Practice and Theory*, 17 :641–653, 2009.

[7] B. P. Zeigler. *Theory Of Modeling and Simulation.* Wiley Interscience, 1976.

UMR IRISA (CNRS, UR1, ENS Bretagne, INSA), équipe DREAM « Diagnostic, Recommandation d'Action, Modélisation »

Diagnostic, Recommandation d'Action, Modélisation

- CONTACT : Marie-Odile Cordier cordier@irisa.fr,
- ADRESSE : Campus de Beaulieu, F - 35 042 Rennes Cedex
- WEB : http://www.irisa.fr/dream/site/

Membres de l'équipe

Membres permanents :

- Louis Bonneau de Beaufort, Ingénieur de recherche, Agrocampus-Ouest
- Marie-Odile Cordier, Professeur, Université Rennes 1
- Thomas Guyet, Maître de conférences, Agrocampus-Ouest
- Véronique Masson, Maître de conférences, Université Rennes 1
- Yves Moinard, Chargé de recherche, INRIA
- Florimond Ployette, Ingénieur de recherche, INRIA
- René Quiniou, Chargé de recherche, INRIA
- Sophie Robin, Maître de conférences, Université Rennes 1
- Laurence Rozé, Maître de conférences, INSA Rennes
- Karima Sedki, Maître de conférences, Agrocampus-Ouest

Collaborateurs extérieurs :

- Philippe Besnard, Directeur de recherche, IRIT-CNRS

Doctorants :

- Tassadit Bouadi, Doctorante, Université Rennes-1
- Yulong Zhao, Doctorant, Université Rennes-1

Thèmes généraux de l'équipe

Les travaux de l'équipe-projet DREAM concernent la surveillance des systèmes complexes. Ils s'intéressent en particulier

- à la modélisation qualitative de ces systèmes complexes, ainsi qu'à l'acquisition automatique des connaissances permettant de construire ces modèles,
- aux techniques de diagnostic à base de modèles, en s'appuyant sur des formalismes des systèmes à événements discrets (principalement décrits par des automates) ou à celui des chroniques (une chronique est un ensemble d'événements contraints temporellement),
- à l'analyse des résultats de simulation fournis par ces modèles, et plus généralement à l'analyse de grande masse de données, par des méthodes d'apprentissage symbolique et de fouille de données ; un accent particulier est mis sur l'acquisition de motifs temporels et spatiaux ainsi que l'apprentissage incrémental à partir de flux de données,
- à la prise en compte de l'utilisateur, en lui fournissant des outils d'aide à la décision et en tenant compte de son interaction dans l'exploitation des modèles.

Les applications cibles sont de deux sortes : la surveillance de système à base de composants, comme les réseaux de télécommunications et les systèmes logiciels comme des services Web et la protection de l'environnement avec le développement de systèmes d'aide à la décision pour aider à la gestion des parcelles agricoles et à la qualité des eaux menacées par des pollutions.

Description des travaux appliqués à la protection de l'environnement

Aide à la gestion de la qualité des eaux

La nécessité de systèmes d'aide à la décision dans le domaine de l'environnement est maintenant bien reconnue. C'est particulièrement vrai dans le domaine de la qualité de l'eau. Par exemple le programme nommé Bretagne Eau Pure a été lancé il y a quelques années afin d'aider les directeurs régionaux à protéger cette ressource importante en Bretagne. Le défi est de préserver la qualité de l'eau contre les polluants comme les nitrates et les herbicides. La difficulté est alors de trouver des solutions qui répondent à des intérêts contradictoires.

Dans ce contexte, nous coopérons avec l'INRA (Institut National de Recherche Agronomique). L'approche que nous préconisons est de s'appuyer sur une modélisation qualitative, afin d'obtenir des modèles biophysiques à une échelle pertinente vis-à-vis des objectifs [2, 7] et de les coupler à des modèles décisionnels représentant le comportement des acteurs (exploitants agricoles, gestionnaires du bassin versant). Le modèle Sacadeau couple ainsi un modèle biophysique et un modèle décisionnel, pour représenter le transfert des pesticides à l'échelle d'un bassin versant.

Des techniques d'apprentissage et d'extraction de données ont été développées pour découvrir, à partir des résultats de simulation du modèle (cf. Figure 2), les variables discriminantes et acquérir automatiquement des règles concernant ces variables. Une caractéristique importante est la nature spatio-temporelle des données. Les règles apprises sont ensuite analysées afin de recommander des actions visant à améliorer la situation actuelle [9, 10].

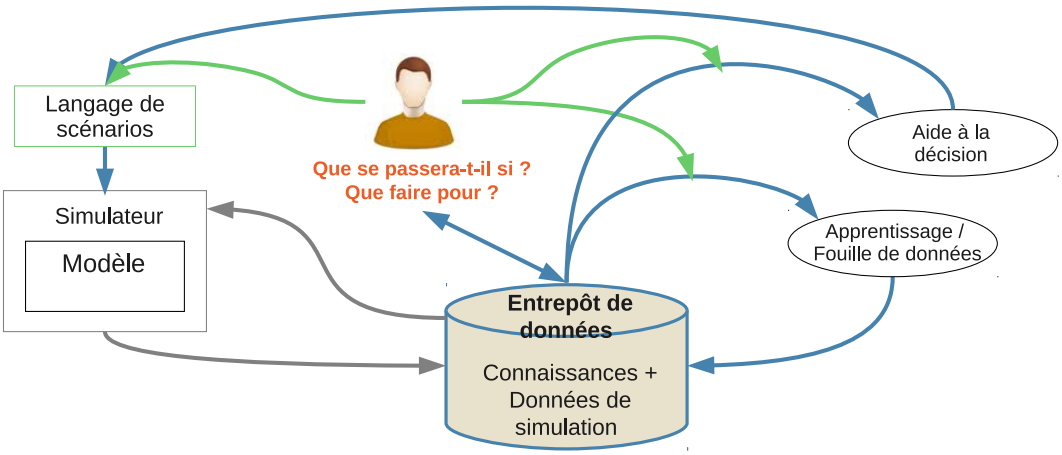


FIGURE 2 – Schéma regroupant les différentes activités de recherche de l'équipe DREAM autour de l'analyse de résultats issues de simulateurs.

Ce travail a été effectué dans le cadre du projet SACADEAU, financé par une AIP INRA, et de l'ANR-ADD APPEAU. Il a donné lieu au logiciel SACADEAU [6].

Entrepôt de données de simulations

Un des objectifs du projet Accassya (Accompagner l'évolution Agro-écologique deS SYstèmes d'élevAge dans des bassins versants côtiers, 2009-2012), financé par l'ANR, est de développer des outils d'aide à la décision pour une meilleure gestion des écosystèmes, en particulier la gestion des systèmes d'élevage dans les bassins versants côtiers. Dans ce contexte, nous avons comme objectif de fournir aux experts, gestionnaires et agriculteurs un outil pour évaluer l'impact des pratiques agricoles sur la qualité de l'eau en s'appuyant sur les résultats du modèle de simulation TNT2, modèle de transfert des nitrates à l'échelle des bassins versants.

Les résultats de simulation consistent en des quantités de données spatio-temporelles très importantes. Nous avons proposé de stocker ces données de simulation dans un entrepôt de données, en l'occurrence utilisant le framework Pentaho et de les interroger grâce à des requêtes OLAP. Nos travaux actuels portent sur le traitement efficace de réponses à des requêtes *skyline*. Celles-ci permettent de retrouver les données de l'entrepôt selon une analyse multicritères, satisfaisant au mieux les préférences des utilisateurs, ou offrant le meilleur compromis. Ceci est étudié dans le contexte d'une analyse interactive du cube de données [1].

EcoMata : modélisation qualitative des écosystèmes

Dans le domaine de l'écologie, il est important de mieux comprendre les interactions souvent complexes entre différents acteurs, l'écosystème lui-même, l'environnement (climat par exemple), les humains en charge de l'exploitation de l'écosystème.

La modélisation qualitative de type automates temporisés, permet de représenter de manière modulaire les différents modèles ainsi que de proposer des patrons de scénarios facilitant l'expression, par un utilisateur non spécialiste, de requêtes dites prédictives de type « Peut-on atteindre telle situation ? », « Quand atteindra-t-on telle situation ? » (atteignabilité), « Est-il possible (oui ou non) d'éviter telle situation ? » (sûreté) [5]. L'idée est de tirer parti des méthodes et outils de *model-checking* pour fournir une réponse à ces requêtes traduites dans la logique temporelle TCTL.

Nos travaux en cours [11] s'intéressent à répondre à des requêtes dites proactives de type « Que faire pour ? », et en particulier les requêtes dites de sûreté de type « Que faire pour éviter telle situation ? » en utilisant des méthodes de synthèse d'automate toujours à partir d'une modélisation par des automates temporisés.

Ce travail a été effectué sur un écosystème marin décrivant l'évolution d'espèces de poisson soumis à des pressions de pêche et à des perturbations de l'environnement et a donné lieu à une expérimentation en grandeur nature pour explorer un écosystème corallien en Nouvelle-Calédonie. Nous l'utilisons actuellement pour la gestion des troupeaux d'élevage dans un bassin versant breton.

Apprentissage et fouille de données temporelles et spatiales

Les travaux menés dans le cadre d'applications agro-écologiques mettent en avant le besoin d'adapter les outils d'acquisition automatique de modèles pour des données spatio-temporelles. La découverte de motifs spatiaux ou de relations temporelles dans une séquence d'évènements nécessite deux étapes : 1) le choix de la représentation des données et des motifs, 2) le choix de la méthode d'apprentissage. Nous sommes principalement intéressés par des méthodes d'apprentissage symbolique, supervisées ou non-supervisées. La spécificité de notre approche est l'apprentissage de modèles expressifs qui prennent explicitement en compte les dimensions temporelles et spatiales.

L'apprentissage inductif (ILP – Inductive Logic Programming) offre la possibilité d'extraire des motifs spatiaux prenant en compte les relations spatiales symboliques entre les éléments à partir d'une représentation logique des données. Ces méthodes ont été utilisées pour l'apprentissage de la structure des arbres de ruissellement discriminant le comportement de ces arbres par rapport à leur apport en pesticide dans la rivière [10].

Pour des données temporelles, nous explorons particulièrement les méthodes de fouilles de données. Nous nous intéressons à l'extraction de motifs séquentiels associant des intervalles de temps aux événements d'une séquence [4]. Ces méthodes sont utilisées pour caractériser la dynamique d'un système à partir de traces de fonctionnement. Dans le cadre du projet MOZAE, de telles approches ont été utilisées pour détecter les chaleurs à partir de signaux de température d'une vache [8].

Dans la dimension spatiale, nous nous intéressons à l'extraction de structures locales fréquentes dans des paysages agricoles (voir la présentation projet Payote, page 46) [3].

Projets

Les projets auxquels nous participons sont :

- Projets portant sur l'amélioration de la qualité des eaux :
  - SACADEAU/APPEAU : « Système d'Acquisition de Connaissances pour l'Aide à la Décision pour la qualité de l'EAU »,
  - Accassya (2009-2013) : « Accompagner l'évolution Agro-écologique deS SYstèmes d'élevAge dans des bassins versants côtiers »,
  - Projet CLIMASTER (2009-2011) : « Changement climatique, systèmes agricoles, ressources naturelles et développement territorial », projet PSDR GO (programmes de recherche « Pour et Sur le Développement Régional Grand Ouest »),
- Projet INRA/INRIA Payote-II (2010-2011) : « Paysage ou territoire » caractérisation des paysages agricoles par des méthodes de fouilles de données,
- Projet PayTal (2011-2014) : « Paysage et étalement urbain », financé par le minitère de l'écologie.

Références

[1] T. Bouadi, M.-O. Cordier, R. Quiniou, and C. Gascuel-Odoux. Analyse multidimensionnelle interactive de résultats de simulation d'un agro-système. In *Actes de la conférence INFORSID*, 2011.

[2] C. Gascuel-Odoux, P. Aurousseau, M.-O. Cordier, P. Durand, F. Garcia, V. Masson, J. Salmon-Monviola, F. Tortrat, and R. Trépos. A decision-oriented model to evaluate the effect of land use and agricultural management on herbicide contamination in stream water. *Environmental Modelling and Software*, 24 :1433–1446, 2009.

[3] T. Guyet. Fouille de données spatiales pour la caractérisation de paysages en lien avec des fonctionnalités agro-écologiques. In *Actes de la conférence SAGEO*, page 3, 2010.

[4] T. Guyet and R. Quiniou. Extracting temporal patterns from interval-based sequences. In *Proceedings of International Joint Conference on Artificial Intelligence (IJCAI)*, pages 1306–1311, 2011.

[5] C. Largouët and M.-O. Cordier. Patrons de scenarios pour l'exploration qualitative d'un écosystème. In *Actes de la conférence RFIA*, 2010.

[6] V. Masson and F. Ployette. Logiciel SACADEAU - Outil d'aide à la gestion de la qualité des eaux d'un bassin versant : simulation, apprentissage de règles de caractérisation et recommandation d'actions. In *Actes de la conférence RFIA*, 2012.

[7] J. Salmon-Monviola, C. Gascuel-Odoux, F. Garcia, F. Tortrat, M.-O. Cordier, V. Masson, and R. Trépos. Simulating the effect of techniques and environmental constraints on the spatio-temporal distribution of herbicide applications and stream losses. *Agriculture, Ecosystems and Environment*, 140 :382–394, 2011.

[8] E. Timsit, S. Assié, R. Quiniou, H. Seegers, and N. Bareille. Early detection of bovine respiratory disease in young bulls using reticulo-rumen temperature boluses. *The Veterinary Journal*, 190(1) :136–142, 2011.

[9] R. Trépos, A. Salleb, M.-O. Cordier, V. Masson, and C. Gascuel. A Distance Based Approach for Action Recommendation. In *Proceedings of European Conference on Machine Learning (ECML)*, pages 425–436, 2005.

[10] R. Trépos, A. Salleb-Aouissi, M.-O. Cordier, V. Masson, and C. Gascuel-Odoux. Building Actions From Classification Rules. *Knowledge and Information Systems (KAIS)*, 2012.

[11] Y. Zhao, M.-O. Cordier, and C. Largouët. Répondre aux questions "Que faire pour" par synthèse de contrôleur sur des automates temporisés - Application à la gestion de la pêche. In *Actes de la conférence RFIA*, 2012.

UPR Cirad « Recyclage et Risque »

Modélisation intégrée de systèmes de production agricoles – Application à la gestion territoriale de produits résiduaires organiques

- CONTACT : François Guerrin
  francois.guerrin@cirad.fr
- ADRESSE : Station de La Bretagne - BP 20 – 97408 Saint-Denis Messag. Cedex 9 - La Réunion
- WEB : http://ur-recyclage-risque.cirad.fr
- TEL : +33 (0) 262 52 81 14 (secrétariat : +33 (0) 262 52 80 20)

Membres de l'équipe

- François Guerrin, Ingénieur de recherches Inra (mis à disposition) HDR
- Jean-Marie Paillat, chercheur Cirad HDR
- Tom Wassenaar, chercheur Cirad
- Daniel David, chercheur post-doctorant Cirad
- Zoubida Afoutni, étudiante en thèse Cirad et Université de la Réunion
- François Dumoulin, étudiant en thèse Ademe/Cirad et Université de la Réunion

L'équipe collabore depuis 1998 avec Rémy Courdier, professeur au « Laboratoire d'Informatique et de Mathématiques » (LIM) de l'Université de La Réunion sur la modélisation multi-agents.

Thème général de l'équipe

L'unité de recherche « Recyclage et Risque » du Cirad étudie les voies permettant le recyclage agricole de produits résiduaires organiques avec un impact environnemental contrôlé. Au sein de l'unité, notre équipe se consacre à la construction de modèles permettant de simuler des scénarios de gestion de ces produits aux échelles de l'exploitation et du territoire et d'évaluer leurs performances agronomiques et environnementales.

Description des travaux

Contexte

Améliorer la durabilité de l'agriculture est devenu indispensable pour faire face aux défis tels que satisfaire les besoins alimentaires mondiaux en forte croissance tout en préservant l'environnement. Le recyclage agricole de déchets et sous-produits organiques, afin de limiter l'usage d'engrais chimiques pour fertiliser les cultures, contribue à atteindre ce but. Concevoir des systèmes de recyclage agricole intégrés nécessite l'usage de connaissances nombreuses et hétérogènes, à la fois sur les composants biophysiques (produits résiduaires organiques, sols, cultures) et les composants humains (les pratiques de gestion des agriculteurs) des systèmes de production. Nous développons ainsi des modèles de simulation permettant d'intégrer ces connaissances pour représenter ces systèmes comme des ensembles d'unités de production et de consommation réparties au sein de territoires. Nous avons pour cela privilégié jusqu'à présent deux approches de modélisation : les systèmes dynamiques hybrides, constitués de variables continues et discrètes, et les systèmes multi-agents.

Modélisation et analyse de flux de matières à l'échelle de territoires

Le fil conducteur de nos recherches a été formalisé dans une démarche comportant plusieurs étapes : (i) enquêtes, rassemblant les données sur les systèmes de production au sein des territoires étudiés ; (ii) typologies, caractérisant les pratiques de gestion des agriculteurs ; (iii) modèles conceptuels, synthétisant les connaissances acquises par les enquêtes ; (iv) modèles informatiques, permettant de simuler l'interaction entre les flux de matières et les pratiques ; (v) scénarios de gestion, évalués sur des critères agronomiques et environnementaux ; (vi) outils et méthodes, pour aider les acteurs agricoles. En suivant cette démarche nous avons développé :

- la représentation conceptuelle des pratiques de gestion des effluents d'élevage à la Réunion [2] ;
- des modèles dynamiques de flux pour raisonner : la gestion d'effluents d'élevage au sein d'exploitations [4, 8] ou de groupes d'exploitations [3] ; l'approvisionnement d'unités collectives de traitement d'effluents [5] ; l'interaction entre pratiques de gestion et flux biophysiques dans des exploitations laitières [12] ; l'organisation individuelle et collective de l'épandage de lisier de porc au sein d'un territoire [10] ;
- un cadre de modélisation de l'activité humaine au niveau opérationnel [7].

Cadre de modélisation conceptuel

En généralisant les représentations mises en œuvre dans nos modèles, nous avons conçu une ontologie 'Action-Flux-Stock' [6] suivant laquelle les systèmes de production agricoles sont formalisés comme des ensembles de stocks connectés par des flux. Deux types de flux sont distingués : les flux « agissables », directement générés par les activités humaines, et les flux « biophysiques », qui ont lieu même en l'absence d'activités humaines. Ces flux interagissent par

l'intermédiaire des activités humaines, qui visent à guider les flux biophysiques par l'intermédiaire des flux agissables qu'elles génèrent. La gestion d'un système de production est donc vu comme le contrôle d'un ensemble de stocks par des activités.

Action « planifiée » vs *« située »*

L'une des questions que nous avons eues à traiter est donc de représenter l'action humaine en contexte agricole. La confrontation entre notre ontologie de modélisation (cf. section ci-dessus) et celles conçues par d'autres auteurs, nous a conduit à remettre en question le paradigme de l'action « planifiée » au profit de la théorie de l'action « située » [11]. Alors que l'approche classique considère que toute action est nécessairement basée sur un plan préétabli (e.g. voir [9]), de nombreuses études ont montré, qu'en fait, une très large part de l'activité humaine ne nécessite ni plan ni représentation mais, plutôt, qu'elle est issue des interactions continues et quasi-spontanées entre les agents et leur environnement. Notre travail se base donc sur cette théorie de l'action située [7, 1]. C'est donc le niveau opérationnel des systèmes de production que nous nous efforçons de représenter de façon réaliste, en nous focalisant plus sur l'action (vue comme le résultat de l'interaction « naturelle » d'un agent avec son environnement) que sur la décision (vue classiquement comme un processus délibératif rationnel).

Evaluation de systèmes de production agricoles

Nous distinguons deux aspects à évaluer : les performances agronomiques et les impacts environnementaux. Dans le premier cas, nous avons besoin de modéliser les flux biophysiques pour simuler leurs interactions avec les flux agissables. Pour cela, l'essentiel de la connaissance disponible est synthétisée par des expressions liant le plus simplement possible les causes avec leurs effets principaux, sans représenter le détail des processus sous-jacents. Dans le second cas, nous devons comparer des stratégies de gestion entre elles. La problématique du développement durable nous conduit à évaluer les impacts de ces systèmes sur des échelles temporelles et spatiales souvent plus larges que celles auxquelles nous avons l'habitude de les considérer. De là, notre intérêt pour les approches d'évaluation globales, telles que l'analyse de cycle de vie (ACV), permettant cette comparaison pour différentes catégories d'impacts et que nous essayons d'articuler avec notre démarche de simulation [10].

Aide à la gestion

Chercher comment aider les acteurs agricoles à évaluer et à concevoir des stratégies de gestion nous a amené à considérer des questions relatives aux protocoles et aux outils de nature à permettre l'expérimentation virtuelle des utilisateurs en utilisant nos modèles de simulation. Quoique nous ayons déjà établi un protocole de simulation, basé sur une logique expérimentale, et que nos modèles sont utilisés par différents collègues agronomes, ceux-ci n'ont cependant jamais vraiment servi en situation « réelle » avec des acteurs agricoles. Cet échec, auquel nous tentons de remédier, a principalement été attribué à notre difficulté à saisir correctement les jeux sociaux des acteurs dans les processus organisationnels ou politiques qui sous-tendent les problématiques de gestion, notamment au niveau collectif du territoire.

Projets en cours

Nous sommes en train d'étendre notre cadre de représentation de l'action en y introduisant les concepts d'agent et d'espace physique [1] nécessaires, notamment, à la représentation de processus de coordination complexes. Ce travail est réalisé dans la thèse de Z. Afoutni au sein du projet « Isard » [9].

Notre mode d'évaluation des stratégies de gestion des systèmes de production doit également être repensé à la lumière d'objectifs mieux définis : que doit-on évaluer, dans quels buts, avec quels acteurs, sur quelles échelles de temps et d'espace ? Une partie de ce questionnement sera traité dans la thèse de F. Dumoulin, qui vise à établir une méthodologie d'évaluation des impacts et des services environnementaux liés au recyclage de déchets organiques sur un territoire. Ce travail est réalisé en lien avec le projet « Girovar » [10], fondé sur une démarche participative impliquant tous les acteurs concernés par la gestion globale des déchets organiques (agricoles et urbains) sur le territoire de la côte Ouest de la Réunion. L'une des ambitions de ce projet est aussi d'apprécier comment des modèles de simulation pourraient s'insérer dans un tel processus collectif de gestion. Ceci permettra peut-être de pallier les limitations signalées ci-dessus.

Comme nous sommes aussi confrontés à la nécessité de considérer des systèmes de production de plus en plus larges, le choix des « échelles de représentation » pertinentes devient un aspect de modélisation à traiter, en explorant les méthodes mathématiques pour modifier le rap-

9. *« Intensification écologique des systèmes de production agricoles par le recyclage des déchets »* (financement ANR, 2008). http://isard.cirad.fr/
10. *« Gestion intégrée des résidus organiques par leur valorisation agronomique à la Réunion »* (financement Cas Dar, ministère de l'agriculture). http://girovar.com/

port entre la granularité et l'étendue de représentation de nos modèles (*model scaling*).

Références

[1] Z. Afoutni, R. Courdier, and F. Guerrin. A model to represent human activities in farming systems based on reactive situated agents. In *Modsim 2011, Int. Congress on Modelling and Simulation*, pages 2901–2907, 2011.

[2] C. Aubry, J.-M. Paillat, and F. Guerrin. A conceptual model of animal wastes management at the farm scale. The case of the Reunion Island. *Agricultural Systems*, 88 :294–315, 2006.

[3] R. Courdier, F. Guerrin, F.-H. Andriamasinoro, and J.-M. Paillat. Agent-based simulation of complex systems : Application to collective management of animal wastes. *Journal of Artificial Societies and Social Simulation*, 5(3), 2002. //jasss.soc.surrey.ac.uk/5/3/4.html.

[4] F. Guerrin. Magma : A model to help manage animal wastes at the farm level. *Computers and Electronics in Agriculture*, 33(1) :35–54, 2001.

[5] F. Guerrin. Simulation of stock control policies in a two-stage production system. Application to pig slurry management involving multiple farms. *Computers and Electronics in Agriculture*, 45(1-3) :27–50, 2004.

[6] F. Guerrin. Modelling agricultural production systems using an action-flow-stock ontology. In *MAS 2008, Int. Workshop on Modelling and Applied Simulation*, pages 7–12, 2008.

[7] F. Guerrin. Dynamic simulation of action at operations level. *Journal of Autonomous Agents and Multi-Agent Systems*, 18(1) :156–185, 2009.

[8] A. Hélias, F. Guerrin, and J.-P. Steyer. Using timed automata and model-checking to simulate material flow in agricultural production systems–Application to animal waste management. *Computers and Electronics in Agriculture*, 63(2) :183–192, 2008.

[9] R. Martin-Clouaire and J.-P. Rellier. Modelling and simulating work practices in agriculture. *International Journal on Metadata, Semantics and Ontologies*, 4(1-2) :42–53, 2006.

[10] J.-M. Paillat and F. Guerrin. Combining individual and collective management of animal manure to reduce environmental impacts on a territory scale. In *Modsim 2011, Int. Congress on Modelling and Simulation*, pages 843–849, 2011.

[11] L.A. Suchman. *Plans and Situated Actions : The Problem of Human-Machine Communication*. Cambridge University Press, Cambridge, MA, USA, 1987.

[12] J. Vayssières, F. Guerrin, J.-M. Paillat, and P. Lecomte. GAMEDE : A global activity model for evaluating the sustainability of dairy enterprises. Part I-Whole-farm dynamic model. *Agricultural Systems*, 101 :128–138, 2009.

TR Irstea Motive, équipe COPAIN, « Systèmes d'information communicants et agri-environnementaux »

Les activités de l'équipe COPAIN portent sur les systèmes d'information environnementaux. Ces systèmes d'information ont pour principales caractéristiques de mobiliser des données spatio-temporelles à des fins d'aide à la décision. Ils permettent d'appréhender les impacts des activités humaines sur l'environnement et de proposer aux décideurs des outils permettant d'assurer la traçabilité de ces activités et de gérer au mieux les milieux.

- CONTACT : Jean-Pierre Chanet
  jean-pierre.chanet@irstea.fr
- ADRESSE : Irstea-Cemagref, 24 avenue des Landais, BP 50085, 63172 Aubière
- WEB : `http://motive.cemagref.fr`

Membres de l'équipe

L'équipe est composée de 9 permanents et de 9 doctorants, post-doctorants et ingénieurs CDD.

Membres permanents :

- André Géraldine, Assistant Ingénieur
- Bernard Stephan, Ingénieur d'Étude
- Bimonte Sandro, Chargé de Recherche
- Boffety Daniel, Ingénieur d'Étude
- Chanet Jean-Pierre, Ingénieur de Recherche
- De Sousa Gil, Ingénieur de Recherche
- Pinet François, Chargé de Recherche HDR
- Roussey Catherine, Chargé de Recherche
- Soulignac Vincent, Ingénieur Divisionnaire de l'Agriculture et l'Environnement

Membres non permanents :

- Barbier Guillaume, Doctorant
- Bendadouche Rimel, Doctorante
- Boulil Kamal, Doctorant
- Chen Yibo, Doctorant
- Edoh-Alove Elodie, Doctorante
- Kone Cheick Tidjane, Post-doctorant
- Lambert Eva, Ingénieur d'Étude
- Wermeille Anais, Ingénieur d'Étude
- Zaamoune Mehdi, Doctorant

Thème général de l'équipe

Les activités de l'équipe COPAIN portent sur la conception et le développement de systèmes d'information pour aider à gérer la qualité environnementale et la production agricole. Ces systèmes d'information mobilisent des données spatio-temporelles, ce qui pose des problèmes de méthodes spécifiques. L'équipe propose aux décideurs des outils permettant d'assurer la traçabilité de certaines activités et de visualiser des indicateurs de différents types, par exemple de qualité ou d'impacts. L'équipe Systèmes d'Information se consacre principalement à deux thématiques :

- développer des méthodes et des outils pour gérer des réseaux de capteurs sans fil et optimiser leur fonctionnement, notamment sur le plan énergétique ;
- développer des méthodes et des outils de conception de systèmes d'information sur l'environnement, prenant en compte des contraintes liées à la qualité des données, aux besoins utilisateurs, etc.

Une liste de quelques publications illustrant nos travaux est proposée à la fin de cette présentation.

Description des travaux

Les réseaux de capteurs sans fil

Nous nous intéressons à la problématique de l'acquisition de données par des réseaux de capteurs sans fil (RCSF) utilisés dans des applications de gestion de l'environnement. Nous avons choisi d'articuler ces recherches autour de plate-formes technologiques mettant en œuvre des capteurs en conditions réelles (surveillance de parcelles agricoles, de bassins versants, etc.).

Les résultats de ces recherches concernent la gestion des données dans ces réseaux en optimisant leur consommation énergétique (qui est la contrainte majeure des RCSF). Ce sont :

- des méthodes pour stocker et gérer les données au sein d'un capteur sans fil sous la forme d'un micro-système de fichiers optimisant l'utilisation des différents types de mémoires (RAM, Flash) ;
- un outil de supervision mettant en œuvre des estimateurs permettant d'administrer à distance un RCSF et de récupérer (extraire) les données collectées par le capteur ainsi que leurs métadonnées associées, tout en minimisant le volume de données échangées.

Les contraintes d'intégrité spatiale

Les Contraintes d'Intégrité Spatiale (CIS) sont des règles de contrôle de la qualité des données des SI spatiaux. Nos recherches sur les CIS sont motivées par le manque de moyens de représentation conceptuelle et de contrôle des CIS dans les SI spatiaux en général et donc en particulier dans les SI sur l'environnement. Nous avons proposé une extension d'OCL (Object Constraint Language) pour spéci-

fier des CIS au sein de représentations orientées-objet UML de données spatiales.

De manière cohérente, nous avons utilisé les techniques orientées objet basées sur UML proposées pour représenter les données et OCL pour représenter les CIS sur ces données. Les objets représentés dans des SI sur l'environnement possèdent souvent des limites géographiques imprécises. Nous avons donc aussi traité le problème de la spécification des CIS des données incertaines (vagues). Nous avons proposé :

1. une représentation des données spatiales vagues facilement utilisable au sein des BD,
2. un modèle de relations spatiales que nous avons intégré à OCL

Les travaux sur les CIS se poursuivent maintenant dans le contexte des entrepôts de données.

Les entrepôts de données spatiales

C'est en 2007-2008 que nous avons pris la décision de porter une part de notre activité sur les Entrepôts de Données (ED) et les outils de type « Online Analytical Processing » (OLAP), car il nous est apparu que ces solutions surtout utilisées dans le secteur tertiaire (grande distribution, finances, gestion, etc.) présentaient un grand potentiel pour stocker, analyser et recouper des données environnementales géo-référencées. Les outils OLAP spatial (SOLAP) et les entrepôts de données spatiales (EDS) permettent l'analyse multidimensionnelle de grands volumes de données géo-référencées. Dans ce contexte, les recherches de l'équipe ont porté sur la représentation, le requêtage et la restitution visuelle de l'information spatio-multidimensionnelle.

Nous avons proposé des solutions originales pour la représentation et le requêtage dans les outils SOLAP des données spatiales complexes comme les trajectoires, les données spatiales continues, les données à différentes échelles et les données provenant de différentes sources, qui caractérisent les applications liées à la gestion de l'environnement. Ces travaux ont été validés par la mise en œuvre de plusieurs prototypes SOLAP. De plus, nous avons défini un cadre conceptuel et d'implémentation pour l'analyse SOLAP des données produites par les modèles de simulation numériques car l'analyse des données issues de simulation est un enjeu important dans la compréhension des dynamiques écologiques et sociales. Nous avons également proposé des techniques nouvelles pour prendre en compte des CIS dans la représentation conceptuelle des ED.

Nous avons également exploré les nouvelles bases de données pour le Cloud Computing en proposant un nouveau format de stockage et des opérateurs en Pig pour l'analyse multidimensionnelle. Ces premières idées ont montré toute la pertinence des solutions de type Cloud dans les problématiques de gestion de données environnementales.

Les ontologies et la gestion de connaissances

Sur le thème des ontologies pour l'agriculture, l'équipe possède une bonne expertise en terme d'ingénierie (participation à différents projets de construction de référentiel en agriculture, GIEA / GIEEA) et une bonne connaissance des acteurs socio-économiques. De plus, dans le cadre du projet ANR Forum, l'équipe a participé à la mise en place d'un système de médiation visant l'interopérabilité et l'intégration de sources de données hétérogènes ; l'équipe a développé un schéma médiateur dans le domaine agricole et fourni des sources de données.

Nous travaillons actuellement sur le développement d'un système de gestion des connaissances sur le web à base d'ontologies. Pour apporter des solutions concrètes au partage et au manque d'information en agriculture biologique, l'équipe COPAIN a investi dans le développement d'un système de gestion des connaissances sur le web intitulé KOFIS « Knowledge for Organic Farming and Innovative System ». Ce système se compose de deux applications web collaboratives à base de réseaux sociaux :

- ***KOFIS-Innovation*** est une application web sociale fondée sur la plate-forme AgriDrupal. Cette application permet à tous les acteurs du monde agricole de demander ou proposer de l'aide pour résoudre un problème rencontré dans une culture de type biologique. Un problème constitue un thème représenté par un billet de blog. Les utilisateurs commentent ce billet pour essayer d'aider à trouver une solution. Ainsi à chaque billet est associé un fil de discussion : un ensemble de commentaires. Chaque billet est annoté à l'aide de mots clés.
- ***KOFIS-Knowledge*** est une application Web Sémantique fondée sur la plate forme Semantic Media Wiki (SMW). Cette application permet à des experts de formaliser des connaissances en agriculture biologique sous forme de pages wiki annotées avec une ontologie en agriculture.

Des ontologies, qu'il convient d'aligner, permettent de maintenir la cohérence des connaissances entre les deux espaces.

11. http://db-ictagri.eu

Quelques exemples de projets

Projet européen FP7 Eranet ICT-Agri

The overall goal of ICT-AGRI [11] « *Coordination of European Research within ICT and Robotics in Agriculture and Related Environmental Issues* » is to strengthen the European Research Area and develop a common European research agenda concerning ICT and robotics in agriculture. ICT-AGRI is funded by the European Commission's ERA-NET scheme under the 7th Framework Programme for Research. The project will finish in 2013.

Projet FUI DISP'EAU

L'outil d'aide à la décision qui sera développé dans le projet FUI DISP'eau, « *Développement d'un outil logiciel innovant d'aide à la décision pour la viticulture* », va permettre de quantifier le stress hydrique au niveau de la plante en continu et de faire le lien avec la qualité potentielle du produit final grâce à la modélisation mécaniste du continuum Sol-Vigne-Atmosphère. Ce modèle sera ensuite couplé à une grille de diagnostic permettant de faire le lien entre le stress hydrique de la vigne et la qualité du raisin ainsi qu'à un système expert. Au final, l'outil permettra de piloter les itinéraires hydriques ou le mode de conduite de la vigne, dans les régions où l'irrigation n'est pas possible, en fonction d'un objectif de qualité des raisins. L'optimisation des trajectoires hydriques grâce à l'outil permettra une meilleure gestion de la ressource en eau, de substantielles économies d'eau dans les zones déjà irriguées, et de dimensionner au mieux les aménagements hydrauliques d'irrigation. Pour l'acquisition des données, un réseau sans fil connecté à un SI sera utilisé. Les partenaires du projet sont : Association Climatologique de l'Hérault, Irstea-Cemagref, Chambre d'Agriculture de l'Hérault, CIRAD, Geocarta, INRA, Les Vignobles Foncalieu, Netafim France, Société ITK (coordinateur), Vignerons du Sieur d'Arques.

Projet MIRIPHYQUE

L'objectif général du projet MIRIPHYQUE, « *MIse au point de descripteurs du RIsque de contaminations des eaux de surface par les PHYtosanitaires à l'échelle du bassin versant* » est de construire une méthode d'évaluation du potentiel de contamination des eaux de surface par les produits phytosanitaires à l'échelle du petit Bassin Versant (BV), qui inclut les dimensions spatiales et temporelles, sans avoir la complexité d'un modèle complet de devenir des pesticides à l'échelle du BV. L'idée centrale est d'agréger à l'échelle du BV les résultats issus de la mise en œuvre de modèles de transfert de phytosanitaires à l'échelle de la parcelle, et d'intégrer l'influence des éléments du paysage sur les flux transférés. Les partenaires du projet sont : ARVALIS-Institut du végétal, Irstea-Cemagref, CNRS, INRA (Coordinateur), UIPP.

Références

[1] L. Bejaoui, F. Pinet, M. Schneider, and Y. Bédard. OCL for formal modelling of topological constraints involving regions with broad boundaries. *Geoinformatica*, 14 :353–378, 2010.

[2] S. Bimonte, M. Schneider, F. Pinet, and J.P. Chanet. Overlay-drill across : a new drill across operator for spatial data warehouses. *Journal of Decision Systems*, 19(3) :261–290, 2010.

[3] S. Bimonte, M. Villanova-Oliver, and J. Gensel. *A multidimensional model for correct aggregation of geographic measures.* IGI Global, 2010.

[4] O. Corcho, C. Roussey, O. Svab-Zamazal, and F. Scharffe. SPARQL-based detection of antipatterns in owl ontologies. In *EKAW2010 workshop proceedings, vol 674*, page 2, 2010.

[5] L. D'Orazio and S. Bimonte. Multidimensional arrays for warehousing data on clouds. In *Data Management in Grid and Peer-to-Peer Systems*, 2010.

[6] A. Jacquot, J.P. Chanet, K.M. Hou, G. De Sousa, and A. Monier. A new management method for wireless sensor networks. In *9th IEEE IFIP Annual Mediterranean Ad Hoc Networking Worshop*, page 9, 2010.

[7] F. Pinet, M. Duboisset, and M. Schneider. Modelling spatial integrity constraints with OCL. *Revue Internationale de Géomatique*, 21(numéro spécial) :95–123, 2011.

[8] F. Pinet and M. Schneider. Precise design of environmental data warehouses. *Operational Research*, 10(3) :349–369, 2010.

[9] C. Roussey, J.P. Chanet, V. Soulignac, and S. Bernard. Les ontologies en agriculture. *Ingénierie des Systèmes d'Information*, 16(3) :55–84, 2011.

[10] C. Roussey and F. Pinet. Détection automatique d'incohérences spatiales à partir de logiques de description. In *SAGEO'10 (Spatial Analysis and GEOmatics 2010)*, pages 306–320, 2010.

[11] C. Roussey, F. Pinet, A.K. Kang, and O. Corcho. Ontologies for interoperability. *Advanced Information and Knowledge Processing*, 1(2011) :39–54, 2011.

[12] C. Roussey, F. Pinet, M.A. Kang, and O. Corcho. An introduction to ontologies and ontology engineering. *Advanced Information and Knowledge Processing*, 1(2011) :9–38, 2011.

[13] C. Roussey, F. Scharffe, O. Corcho, and O. Zamazal. Une méthode de débogage d'ontologies OWL basées sur la détection d'anti-patrons. In *21èmes Journées Francophones d'Ingénierie des Connaissances*, 2010.

[14] V. Soulignac, J.L. Ermine, J.L. Paris, O. Devise, and J.P. Chanet. Gestion informatisée des connaissances pour une agriculture durable. In *3ème Conférence Francophone Gestion des Connaissances, Société et Organisations (GECSO 2010)*, page 23, 2010.

[15] V. Soulignac, J.L. Ermine, J.L. Paris, O. Devise, and J.P. Chanet. A knowledge server for sustainable agriculture. In *8th International Conference on Intellectual Capital, Knowledge Management and Organisational Learning*, page 14, 2011.

[16] M. Visoli, S. Bimonte, S. Ternès, F. Pinet, and J.P. Chanet. *Towards a spatial decision support system for animal traceability*. IGI Global, 2011.

EA CReSTIC, équipe SIC-MISTI « Management de l'Information SpatioTemporelle Imparfaite »

Le groupe « Signal, Image et Connaissance » (SIC) rassemble des enseignants-chercheurs ayant des activités de recherche autour du signal, du traitement et de l'analyse d'images, de l'informatique graphique et de l'intelligence artificielle au service de la modélisation des connaissances. Au sein du SIC, l'équipe projet « Management de l'Information SpatioTemporelle Imparfaite » (MISTI) travaille sur la représentation, la gestion, l'analyse et la visualisation de connaissances spatiotemporelles sujettes à de l'imperfection.

- CONTACT : Eric Desjardin eric.desjardin@univ-reims.fr,
- ADRESSE : IUT de Reims Châlons Charleville, Rue des Crayères, 51687 Reims Cedex 2, France
- WEB : http://www.univ-reims.fr/crestic
- TEL : +33 (0) 3.26.91.84.58

Membres de l'équipe

Membres permanents :
- Eric Desjardin,
- Michel Herbin,
- Cyril de Runz,
- Frédéric Blanchard,
- Herman Akdag

Membres non-permanents :
- Karima Zayrit,
- Asma Zoghlami

Thème général de l'équipe

Ce groupe aborde sous des angles multiples et complémentaires les divers aspects d'un axe scientifique intitulé « des Données à la Décision Assistée » qui vise :

– la maîtrise et le développement des technologies de manipulation de données complexes et/ou volumineuses, structurées ou non (acquisition, extraction, traitement, fouille, classification, analyse, interprétation, visualisation et interaction) comme de leur enchaînement,
– pour les utiliser dans le cadre de la modélisation, la simulation et l'expérimentation virtuelle,
– en vue de produire des systèmes d'aide à la compréhension et à la décision.

Prenant son essence dans la convergence des préoccupations du *Traitement de l'incertitude* et des *Systèmes d'Information Géographique*, les travaux menés au sein de l'équipe projet MISTI associent la gestion de bases de données spatiales et/ou temporelles, la reconstruction virtuelle, la fouille de données et la visualisation. Nous cherchons dans nos approches à prendre en considération la nature de l'information exploitée (imperfection, qualité interne) et son usage (qualité externe/fonction des usages, profils d'utilisateurs).

Description des travaux

Les travaux en lien avec le thème du dossier sont effectués dans l'axe « Systèmes d'information pour le patrimoine et le développement durable » au sein de l'équipe-projet MISTI. Depuis 2007, des travaux autour de la mise en place, du développement et de l'exploitation des Systèmes d'Information Géographique pour le développement durable et particulièrement l'agronomie et la viticulture ont émergé au CReSTIC.

Dans ce cadre, le projet OBSERVOX est effectué en collaboration avec l'équipe de Marc Benoit de l'INRA Mirecourt. Des collaborations sont aussi en cours avec le monde viticole sur la fouille de données.

Projet Observox

Contexte L'enjeu fondamental du contrat d'objectif AQUAL dans son ensemble est la préservation de la qualité des eaux souterraines et superficielles vis-à-vis des pollutions issues des activités agri-viticoles sur le bassin de la Vesle. L'objectif d'AQUAL ne peut être réalisé qu'avec une bonne connaissance des pratiques contribuant à la dissémination des produits phytosanitaires.

Il est indispensable de mieux connaître l'origine des pesticides retrouvés dans les eaux, en organisant l'information sur leur utilisation passée et présente, leur nature (matières actives et adjuvants), et leurs liens avec l'évolution des systèmes de culture et les stratégies sociales, économiques et politiques (types de cultures, types d'usages, types d'exploitations), sur leurs localisations d'application (spatialisation des processus) tout en maitrisant la qualité de ces connaissances. Le dispositif d'observatoire OBSERVOX, co-construit par l'ensemble des acteurs, organise la restitution des résultats à des fins d'analyse, de modélisation, de préconisation.

Dans ce but, de multiples sources d'informations (statistiques, enquêtes, géographiques) sont mobilisées. Or la circulation d'informations ainsi que la difficulté pour les utilisateurs d'apprécier leur qualité peuvent déboucher sur des utilisations ou des interprétations erronées et avoir un

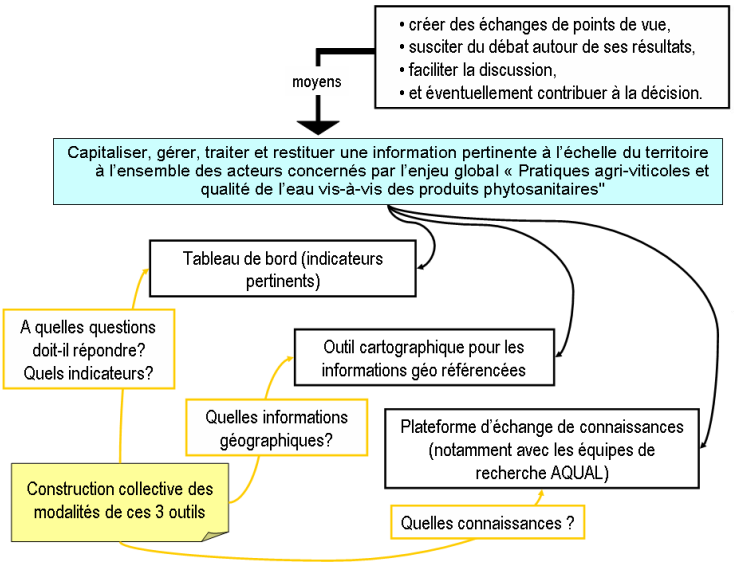


FIGURE 3 – Un outil d'échange et d'aide à la discussion

impact négatif sur les résultats ou sur la crédibilité de l'observatoire.

Ces problèmes viennent de la nature même des données que nous pouvons collecter. En effet, lorsque l'utilisation de chaque molécule sur chacune des parcelles d'un territoire agricole ne donne pas lieu d'une manière accessible à une déclaration par les agriculteurs, cette information doit être déduite de différentes sources relatives aux pratiques phytosanitaires et aux localisations des surfaces cultivées. Ces données sont alors définies avec des échelles spatiales et temporelles différentes entre elles (certaines ne sont d'ailleurs connues que d'une manière globale). En dernier lieu, les pratiques agricoles sont évolutives et la connaissance que nous en avons est partielle.

Si tant est que l'ensemble des processus sous-jacents aux pratiques agricoles (par exemple tel produit est utilisé pour telle culture avec des quantités normalisées) puisse être modélisé, l'utilisation de ces règles est dépendante de la connaissance fine et dynamique du parcellaire et des cultures mises en place sur ces parcelles.

Modélisation des données agro-viticoles Ainsi, l'un des objectifs d'Observox est de traiter l'imperfection des données spatiales et quantitatives et de toujours associer une évaluation de la qualité aux données acquises ou inférées (*c.f.* [2]). En effet, la composante spatiale des informations étudiées dans un SIG est souvent aussi imprécise (et/ou incertaine) que leurs composantes attributaires (quantitatives ou non).

Selon la littérature, le cadre théorique des sous-ensembles flous est bien adapté à la représentation et le traitement des données imprécises. Afin de répondre aux objectifs de l'observatoire, à l'aide de cette théorie, nous sommes en mesure de construire des entités imprécises aussi bien sur leurs composantes spatiales que quantitatives. Nous appelons les entités ainsi formées « entités agronomiques floues » dont la définition a été introduite dans [4].

Pour en permettre la gestion, nous avons travaillé à la prise en compte des imperfections des données dès la modélisation du système (diagramme de classe) en augmentant le langage Perceptory de nouveaux pictogrammes [5], à la structuration de la base de données afin de permettre la gestion de données spatiotemporelles floues (discrétisées en données multivaluées) et à la définition du modèle [6].

Propagation de l'imprécision Cette modélisation permet de représenter l'imprécision des données mais, de fait, rend aussi possible le chevauchement des composantes spatiales de deux (ou plus) entités. La question qui se pose concerne alors l'évaluation de la valeur floue des attributs quantitatifs imprécis dans les zones de chevauchement.

Afin d'y répondre, nous propageons l'imprécision du spatial vers le quantitatif à l'aide d'un opérateur [3] de caractère additif qui prend en compte à la fois l'information spatiale et quantitative, et qui fournit une information quantitative locale et floue. Cet opérateur a été reformulé afin d'obtenir une représentation floue des quantités

de produits phytosanitaires disséminés à chaque endroit du territoire étudié [5].

Classification et fusion des données Pour permettre le suivi spatiotemporel de la diffusion des pesticides, il nous faut connaître et combiner la nature des couverts végétaux sur le territoire et leur évolution avec la connaissance des pratiques phytosanitaires qui leur sont associées.

L'objectif est de mettre en place des algorithmes de traitement et de classification d'images de télédétection, issues de sources multiples, pour identifier les cultures présentes sur le territoire, avec intégration de données du Registre Parcellaire Graphique. Afin de gérer l'imperfection des sources, ainsi que la fusion des connaissances, ces algorithmes de classification pourront fournir et/ou utiliser les différentes théories de l'incertain et de l'imprécis (fonctions de croyances, possibilités/nécessités, ensembles flous, etc.).

Fouille de données viticoles

Le champagne est exploité par 16555 vignerons et couvre, à l'heure actuelle, 34 045 hectares ce qui représente environ 3.4 % de la superficie du vignoble français. Cette aire de production délimitée en 1927 comprend 319 crus ou communes regroupés en 4 grandes régions : La Montagne de Reims, La Vallée de la Marne, la Côte des Blancs et la Côte des Bar. Héritage de trois siècles d'histoire, le Champagne est un vin mondialement connu, souvent synonyme de luxe et d'événements festifs.

Dans ce contexte, l'un des objectifs du territoire est d'aller vers la viticulture raisonnée et de conseiller au mieux les viticulteurs en termes de pratiques agronomiques en étudiant toutes les variables qui peuvent influencer la récolte et l'impact environnemental.

La compréhension du fonctionnement de la vigne dans son environnement est importante pour expliquer la variabilité en quantité et qualité (au sens descriptif du terme) du raisin à la vendange. En effet, de la tige à la récolte du raisin, beaucoup de paramètres entrent en jeu : les caractéristiques génétiques comme le cépage, le porte greffe, des choix de pratiques culturales (système de conduite, fertilisation, traitements phytosanitaires), et enfin des caractéristiques du territoire (sol, sous-sol, et la climatologie). Connaître et prendre en compte cette variabilité est le principe de la viticulture de précision.

Les évolutions technologiques permettent l'acquisition de masses de données de plus en plus importantes, ce qui donne une connaissance plus approfondie du terroir mais aussi complexifie les traitements. Pour obtenir ces informations, de nombreuses mesures ont été acquises sur une parcelle expérimentale, gérée par le CIVC, à l'aide de capteurs fixes et embarqués. L'objectif est de faire émerger différentes zones de « potentiel » à l'aide de techniques de fouilles de données complexes.

Ainsi, en fonction du matériel végétal, c'est-à-dire les types de cépages (Chardonnay, Meunier, et Pinot Noir), nous avons des données qui quantifient la géologie et le climat. Par le biais de ses différentes données mesurées, nous avons établi une problématique dont le but est de comprendre le fonctionnement de la parcelle et de définir des zones de potentiel.

Des analyses bidimensionnelles confirment l'existence de certaines corrélations (par exemple entre pente et quantité de récolte). Le but de la collaboration est de mettre en lumière des relations plus complexes et interprétables entre variables au regard des configurations spatiales et temporelles. Un premier travail a consisté à proposer une « classification d'une parcelle viticole en zones de potentiel homogène » [1].

Références

[1] S. Camara. Classification d'une parcelle viticole en zones de potentiel homogène. Master's thesis, CIVC, CReSTIC, Université de Reims, France, 2011.

[2] C. de Runz and É. Desjardin. Spatialisation qualifiée des pratiques et données agricoles pour l'étude des risques de contamination de ressources en eau du bassin de la vesle. revue des problématiques. In *SAGEO*, 2009.

[3] C. de Runz, É. Desjardin, and H. Akdag. Study of spatial fusion of geographical entities and quantitative information in accordance with their imprecision : application to agricultural information in observox. In *Spatial Accuracy*, pages 33–36, 2010.

[4] K. Zayrit, É. Desjardin, C. de Runz, and H. Akdag. Propagation of spatial imprecision in imprecise quantitative data in agronomy. In *ISSDQ*, pages 145–150. INESC Coimbra, 2011.

[5] A. Zoghlami, C. de Runz, H. Akdag, M. Zaghdoud, and H. Ben Ghezala. Handling imperfect spatiotemporal information from the conceptual modeling to database structures. In *ISSDQ*, pages 165–170. INESC Coimbra, 2011.

[6] A. Zoghlami, K. Zayrit, C. de Runz, É. Desjardin, and H. Akdag. Vers la construction d'un observatoire des pratiques agricoles : gestion et propagation de l'imprécision des données agronomiques. In *EGC*, 2012.

UMR Mét@risk (INRA - AgroParisTech), équipe « Ingénierie des Connaissances »

- CONTACT : Juliette Dibie-Barthélemy dibie@agroparistech.fr,
- ADRESSE : Mét@risk & AgroParisTech, 16 rue Claude Bernard, 75231 Paris cedex 05
- WEB : http://www.paris.inra.fr/metarisk/research_unit/knowledge_engineering
- TEL : +33 (0) 1 44 08 18 48

Membres de l'équipe

- Damien Berry, Assistant Ingénieur
- Stéphane Dervaux, Ingénieur d'étude
- Juliette Dibie-Barthélemy, Maître de conférences d'AgroParisTech, HDR
- Liliana Ibanescu, Maître de conférences d'AgroParisTech
- Lydie Soler, Ingénieur d'étude
- Rim Touhami, Doctorante

Thème général de l'équipe

La société actuelle doit faire face à des évolutions rapides et à des enjeux aux multiples facettes (environnementaux, sociétaux, économiques). C'est particulièrement vrai pour le monde des sciences du vivant, de l'agronomie et de l'agro-alimentaire, soumis en plus aux aléas naturels. Pour répondre à ces enjeux et évolutions, de nombreux outils informatiques ont été développés. Ces outils se sont concentrés soit sur des approches spécifiques à une partie du domaine d'application étudié (e.g. population à risque comme les jeunes enfants), soit sur des enjeux bien déterminés (e.g. problèmes sanitaires). Il en a résulté une dispersion des ensembles de données et de la connaissance, ce qui s'est traduit par une grande difficulté à faire cohabiter ces connaissances et ces données, et, à les exploiter conjointement.

La nécessité d'acquérir, d'intégrer et d'utiliser des sources de données hétérogènes alimente de nombreuses thématiques de recherche en Intelligence Artificielle. Le travail de recherche de l'équipe Ingénierie des Connaissances s'inscrit dans ce courant. Il porte plus précisément sur l'intégration de données hétérogènes provenant de sources multiples afin de pouvoir les utiliser et les exploiter dans des outils d'aide à décision dédiés à l'évaluation des risques sanitaires microbiologiques et chimiques des aliments.

Description des travaux

Intégration des données

Le travail de recherche de l'équipe Ingénierie des Connaissances porte sur l'intégration de données hétérogènes provenant de sources multiples, en particulier le Web, cette intégration étant guidée par une ontologie. Les quatre principaux axes de recherche de l'équipe sont :

- l'enrichissement automatique d'un entrepôt par des données extraites de documents du Web et sémantiquement annotées à l'aide d'une ontologie ;
- l'interrogation flexible d'un entrepôt de données ouvert sur le Web ;
- la construction et l'évolution d'ontologie ;
- l'alignement d'ontologies.

Ce travail de recherche est mené dans le cadre de la construction d'un système d'intégration de données thématiques ouvert sur le Web, appelé ONDINE (Ontology based Data INtEgration) [8, 4], qui repose sur les standards du Web sémantique (XML, RDF, OWL, SPARQL). Le système ONDINE (cf. Figure 4) propose un processus complet d'acquisition de données du Web, plus précisément des tableaux, qui contiennent en général une synthèse des données publiées dans les articles scientifiques, pour enrichir des bases de données locales. Il repose sur une Ressource Termino-Ontologique (RTO), qui donne une représentation, formalisée et structurée, du vocabulaire spécifique à un domaine d'application donné. La RTO est composée d'une partie générique contenant les concepts dédiés à la tâche d'intégration de données et d'une partie spécifique ainsi que d'une terminologie, dédiés au domaine d'application étudié. Le système ONDINE est composé de deux sous-systèmes : (1) le sous-système @Web [14, 11, 12, 13] qui permet d'alimenter un entrepôt de données avec des données extraites de tableaux du Web qui ont été sémantiquement annotés avec des concepts de la RTO ; (2) le sous-système MIEL++ [2, 3, 6, 9, 5] qui propose un mécanisme d'interrogation unifié et flexible des bases locales et de l'entrepôt de données issu de l'annotation sémantique des données du Web. La RTO joue un rôle central dans le système d'intégration ONDINE : (i) elle permet une extraction pertinente des données du Web et une annotation sémantique fine de ces données ; (ii) elle permet une interrogation unifiée des données locales et des données du Web ; (iii) elle est le pivot entre l'annotation et l'interrogation des données du Web et assure leur intégration avec les données locales ; (iv) elle permet de détecter des redondances entre les données du Web sémantiquement annotées avant leur intégration dans l'entrepôt [15].

L'équipe Ingénierie des Connaissances étudie en particulier la représentation de connaissances imprécises et

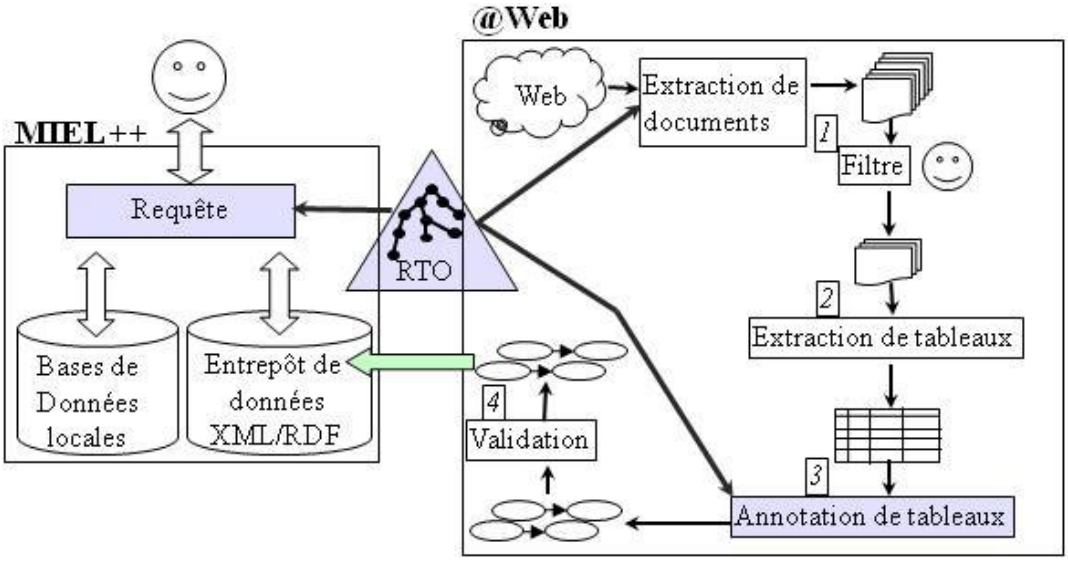


FIGURE 4 – Le système ONDINE.

leur interrogation flexible [5]. Dans le cadre des applications en risque alimentaire, l'imprécision des données est de deux natures différentes : (i) elle est propre aux données qui sont des données expérimentales et donc par nature imprécises (variabilité liée à la complexité intrinsèque des processus biologiques, imprécision des capteurs, imprécision dans l'expression des résultats publiés dans les sources de données) ; (ii) elle résulte du processus d'annotation sémantique des données du Web et correspond à une mesure de similarité entre ces données du Web et des valeurs de la RTO. Le système ONDINE propose de traiter cette imprécision des données à différents niveaux à l'aide de sous-ensembles flous : représentation de données imprécises dans les bases locales et l'entrepôt de données XML/RDF, expression de similarité entre des valeurs du Web et des valeurs de la RTO, expression de préférences dans l'interrogation unifiée des données locales et des données du Web.

De l'Intégration des données à l'Aide à la Décision

Le travail méthodologique de recherche, décrit ci-dessus, mené par l'équipe Ingénierie des Connaissances est une étape préalable indispensable de préparation des données, qui se situe en amont du raisonnement et de la prise de décision. Une fois intégrée ensemble, les données sont utilisées pour alimenter des outils d'aide à la décision dans le domaine de l'évaluation du risque alimentaire, permettant de mettre en évidence des populations à risque, de modéliser des comportements des consommateurs, d'aider des gestionnaires de risque dans leur prise de décisions,...

L'équipe Ingénierie des Connaissances développe notamment le logiciel CARAT afin d'étudier les comportements alimentaires et leurs conséquences sur la santé d'une population cible. Le logiciel CARAT est un outil d'évaluation du niveau d'exposition d'une population cible à un contaminant chimique ou à un nutriment sur la base de sa consommation et de la contamination ou composition des aliments permettant ainsi d'effectuer des analyses risque/bénéfice. Afin de pouvoir effectuer une évaluation quantitative de l'exposition d'une population cible à un risque chimique, il faut être capable d'exploiter conjointement des données de consommation qui peuvent provenir de sources diverses (INCA1, INCA2,KANTAR) et des données de contamination qui peuvent également provenir de sources diverses (DGAL, TDS, ONIVIN). Cela implique d'être capable de mettre en correspondance les systèmes de codification des aliments utilisés dans les différentes sources de données : entre chaque source de données de consommation, entre chaque source de données de contamination et entre les données de consommation et les données de contamination. Afin de faciliter cette tâche, l'équipe Ingénierie des Connaissances s'intéresse aux méthodes de mise en correspondance semi-automatique d'ontologies [7, 10].

Projets

L'équipe Ingénierie des Connaissances participe notamment aux projets suivants :

– le projet ANR Holyrisk (2009-2013), qui porte sur « Analyse de l'incertitude dans l'évaluation des

risques alimentaires et ses conséquences sur la décision publique ». Dans ce projet, l'équipe Ingénierie des Connaissances étudie comment aider un gestionnaire de risque à prendre sa décision sur un risque donné en fonction de la chaîne chronologique de documents disponibles sur l'évaluation de ce risque et éventuellement de documents préalables de gestion de ce risque [1].

– le projet ANR/ALIA MAP'OPT (2011-2015), qui porte sur « composition, dynamique des gaz et optimisation de la protection des denrées dans les emballages sous atmosphère modifiée ». Dans ce projet, l'équipe Ingénierie des Connaissances travaille sur la problématique de construction et d'évolution d'ontologie [16] dans une optique d'intégration de données.
– le projet GEMS/Food de l'OMS (Organisation Mondial de la Santé), l'équipe étant centre collaborateur de l'OMS depuis 2007, qui vise à construire un Système d'Information au niveau mondial permettant de collecter des données de consommation et de contamination, et d'évaluer l'exposition d'une population cible.
– le projet AgroParisTech ExtraEx (2012-2013) qui vise à construire un système d'Extraction de paramètrEs cruciaux pour une aide à la synthèse et à la prise de décision à partir de grandes bases de données bibliographiques et est appliqué à l'estimation des émissions de gaz à effets de serre. Dans ce projet, l'équipe Ingénierie des Connaissances travaillera sur la problématique d'annotation de documents textuels.

L'équipe participe également au GIS Sym'Previus (réseau national d'experts sur la prévention du risque alimentaire) qui a été mis en place en décembre 1999 et est financé par les Ministères de l'Agriculture et de la Recherche. Elle est membre du PAP, Pôle Alimentation Parisien, qui a pour principal objectif de collecter et produire des données et des connaissances sur la consommation, la composition et la contamination des aliments, en vue de contribuer à l'élaboration et l'évaluation de dispositifs de régulation des problèmes sanitaires et à la prévention des problèmes nutritionnels.

Références

[1] S. Blanchemanche, P. Buche, J. Dibie-Barthélemy, E. Feinblatt Mélèze, L. Ibanescu, and Akos Rona-Tas. Ontology building : an application in food risk analysis. In *8th International Conference on Terminology and Artificial Intelligence*, 2009.

[2] P. Buche, O. Couvert, J. Dibie-Barthélemy, G. Hignette, Eric Mettler, and L. Soler. Flexible querying of web data to simulate bacterial growth in food. *Food Microbiology*, 28(4) :685–693, 2011.

[3] P. Buche, J. Dibie-Barthélemy, and H. Chebil. Flexible sparql querying of web data tables driven by an ontology. In *FQAS*, volume 5822 of *Lecture Notes in Computer Science*, pages 345–357, 2009.

[4] P. Buche, J. Dibie-Barthélemy, O. Haemmerlé, and G. Hignette. Fuzzy semantic tagging and flexible querying of XML documents extracted from the web. *Journal of Intelligent Information Systems*, 26(1) :25–40, 2006.

[5] P. Buche, J. Dibie-Barthélemy, O. Haemmerlé, and R. Thomopoulos. The MIEL++ architecture : when RDB, CGs and XML meet for the sake of risk assessment in food products. In *14th International Conference on Conceptual Structures*, volume 4068 of *LNCS*, pages 158–171, 2006.

[6] P. Buche, J. Dibie-Barthélemy, and G. Hignette. Flexible querying of fuzzy rdf annotations using fuzzy conceptual graphs. In *16th International Conference on Conceptual Structures*, volume 5113 of *LNCS*, pages 133–146, 2008.

[7] P. Buche, J. Dibie-Barthélemy, and L. Ibanescu. Towards large-size ontology alignment by mapping groups of concepts. In *OTM Workshops*, pages 43–44, 2010.

[8] P. Buche, J. Dibie-Barthélemy, L. Ibanescu, and L. Soler. Fuzzy web data tables integration guided by an ontological and terminological resource. *IEEE Transactions on Knowledge and Data Engineering*, 99 :Preprints, 2011.

[9] P. Buche, J. Dibie-Barthélemy, and F. Wattez. Approximate querying of XML fuzzy data. In *FQAS*, volume 4027 of *Lecture Notes in Computer Science*, pages 26–38, 2006.

[10] D. Doussot, P. Buche, J. Dibie-Barthélemy, and O. Haemmerlé. Using fuzzy conceptual graphs to map ontologies. In *OTM Conferences (1)*, pages 891–900, 2006.

[11] G. Hignette, P. Buche, O. Couvert, J. Dibie-Barthélemy, D. Doussot, O. Haemmerlé, Eric Mettler, and L. Soler. Semantic annotation of web data applied to risk in food. *Food Microbiology*, 128(1) :174–180, 2008.

[12] G. Hignette, P. Buche, J. Dibie-Barthélemy, and O. Haemmerlé. An ontology-driven annotation of data tables. In *WISE Workshops. Web Data Integration and Management for Life Sciences.*, volume 4832 of *LNCS*, pages 29–40, 2007.

[13] G. Hignette, P. Buche, J. Dibie-Barthélemy, and O. Haemmerlé. Semantic annotation of data tables using a domain ontology. In *10th International Conference on Discovery Science*, volume 4755 of *LNCS*, pages 253–258, 2007.

[14] G. Hignette, P. Buche, J. Dibie-Barthélemy, and O. Haemmerlé. Fuzzy annotation of web data tables driven by a domain ontology. In *ESWC*, volume 5554 of *Lecture Notes in Computer Science*, pages 638–653, 2009.

[15] R. Khefifi, P. Buche, J. Dibie-Barthélemy, and Fatiha Saïs. Détection de redondances dans les tableaux guidée par une ontologie. In *EGC*, pages 563–568, 2011.

[16] R. Touhami, P. Buche, J. Dibie-Barthélemy, and L. Ibanescu. An ontological and terminological resource for n-ary relation annotation in web data tables. In *OTM Conferences (2)*, pages 662–679, 2011.

UMR IRHS (INRA-Agrocampus-Ouest), équipe FruitQual

Facteurs génétiques, transcriptomiques et écophysiologiques influant sur la qualité des pommes et en particulier les caractéristiques de texture

- CONTACT : Julie Bourbeilon julie.bourbeillon@agrocampus-ouest.fr,
- ADRESSE : IRHS [12] Bâtiment B, 42 rue Georges Morel, 49070 BEAUCOUZE CEDEX
- WEB : http://www.angers-nantes.inra.fr/irhs

Chercheurs permanents de l'équipe

- Catherine Bernard, Maître de conférences à l'Université d'Angers
- Julie Bourbeillon, Maître de conférences à AGROCAMPUS OUEST
- Gerhard Buck-Sorlin, Professeur à AGROCAMPUS OUEST
- Mickaël Delaire, Maître de conférences à AGROCAMPUS OUEST
- Pascale Guillermin, Maître de conférences à AGROCAMPUS OUEST
- François Laurens, IR INRA
- Mathilde Orsel Baldwin, CR INRA
- Jean-Pierre Renou, DR INRA

Thème général de l'équipe

Après l'attractivité visuelle, les caractéristiques sensorielles de texture sont les facteurs les plus importants dans la décision d'achat de pommes des consommateurs. Ces caractéristiques de texture sont des éléments déterminants du succès de toute nouvelle variété. Or, la texture est un trait composite, qui évolue dans le temps, et qui est très influencé par des facteurs génétiques, physiologiques et environnementaux. Afin de mieux comprendre les facteurs impliqués dans cet élément de la qualité des fruits, l'équipe FruitQual mène des recherches intégratives aux niveaux génomique, génomique fonctionnelle et écophysiologique. Ces études visent principalement des caractéristiques de texture, mais aussi, à moindre échelle, des aspects nutritionnels tels que les fibres ou les polyphénols.

12. IRHS : Institut de Recherche en Horticulture et Semences

Description des travaux

Deux volets des activités de l'équipe sont adossées à des activités en Intelligence Artificielle, principalement en ingénierie des connaissances et aide à la décision.

Intégration de données pluri-annuelles et multi-échelles

Au sein de l'équipe FruitQual, trois approches complémentaires, à des échelles allant de la molécule aux vergers, sont appliquées sur du matériel biologique partagé, avec un objet de recherche commun, la texture des pommes :

- Bases génétiques d'attributs majeurs de la qualité des fruits telles que celles présentées dans [5],
- Génomique fonctionnelle, comme exposé dans [3],
- Études écophysiologiques telles que [6],

Le but est de mettre en place un modèle à la fois génétique et écophysiologique de l'élaboration de la qualité des fruits, tel que celui proposé chez la rose fleur coupée [2] ou un modèle de type « Fruit Virtuel » [4]. Cet objectif de modélisation génère des problématiques d'intégration de données pluri-annuelles et multi-échelles.

En effet, le contexte est celui de cultures pérennes. Les mesures sont donc réalisées d'une part à divers stades de développement de l'arbre (plantule, jeune arbre, arbre adulte) et d'autre part sur des fruits ou lots de fruits récoltés des années différentes. De plus, les données acquises concernent soit un arbre particulier entier (données génomiques ou architecture de la cime par exemple), soit des ensembles d'organes particuliers (telles que des expressions différentielles de gènes entre organes), soit des lots de fruits (caractéristiques biochimiques de jus), des fruits entiers ou des morceaux de fruits (mesures de pénétrométrie selon la face du fruit, orientée vers l'extérieur ou l'intérieur de l'arbre, par exemple). Le traitement conjoint de l'ensemble de ces données est difficile et implique :

1. d'identifier de manière précise les échantillons et en particulier où ils se situent sur les différents axes considérés (stades de développement, années de récolte, échelle),
2. de définir et appliquer des méthodes permettant le traitement de données où les échantillons correspondants sont à des niveaux différents sur ces axes : par exemple, comment traiter conjointement une mesure réalisée sur 3 fruits et une mesure réalisée sur un lot de fruits du même arbre ? Comment intégrer au niveau du lot ? Comment extrapoler au niveau du fruit ?
3. d'aider au choix et à l'application des méthodes du point 2.,

Ces éléments impliquent le recours à des outils informatiques basés sur des méthodes d'Intelligence Artificielle : représentations de connaissances, sous forme d'ontologie de l'arbre en tant que source d'échantillons situés à différents niveaux d'échelle, etc., et associés à des méthodes d'intégration / extrapolation particulières pour changer de niveau ; aide à la décision pour faciliter le choix des méthodes d'intégration / extrapolation à utiliser pour un traitement de données particulier. L'objectif à terme est la mise en place d'une plateforme informatique intégrée de type LIMS [13], destinée à un public non informaticien, où l'interaction avec le système est située à un niveau d'abstraction proche de la question biologique, dans l'esprit de ce qui a été développé dans [1].

Représentation et exploitation de données de pratiques culturales

Les facteurs environnementaux tels que le climat, la pédologie, mais aussi la conduite des vergers par les arboriculteurs, ont une influence forte sur les caractéristiques de texture des pommes. Dans un contexte de politique de réduction d'intrants (Ecophyto2018 par exemple), l'évaluation de systèmes de cultures innovants, permettant de prendre en compte les enjeux environnementaux, la sécurité du consommateur, la qualité des produits finis, tout en assurant une durabilité économique des exploitations, devient cruciale. Cette évaluation peut être faite par la conduite d'expérimentations en conditions réelles (pluri-annuelles, chez les arboriculteurs), dont l'objectif est la comparaison entre pratiques culturales classiques et innovantes. Dans ce cadre, plusieurs problématiques émergent.

La conduite d'expérimentations sur le long terme induit l'acquisition d'un volume important de données, données qui sont aussi originales. En effet, il faut représenter des pratiques culturales, éventuellement variables d'une année à l'autre, qui peuvent soit être assimilées à des règles de type *Si* `<Conditions>` *Alors* `<Actions>`, soit consistent en propriétés particulières d'une parcelle. De plus, l'application des pratiques culturales en conditions réelles n'est pas nécessairement parfaite, pour des raisons organisationnelles ou météorologiques, etc. Or, la différence entre les pratiques culturales recommandées et leur mise en pratique réelle doit être prise en compte dans l'évaluation des pratiques culturales, soit dans le calcul d'indicateurs techniques, environnementaux, économiques, etc. Il s'agit alors d'une part d'être en mesure de capturer comment les pratiques culturales sont appliquées et d'autre part de formaliser et quantifier les déviations. L'approche adoptée pour résoudre ces questions passe par la mise en place d'une base de données associée à des représentations de connaissances pour la capture des informations.

À plus long terme, par delà la gestion et l'analyse des données d'expérimentation, l'objectif est aussi de proposer un système de recommandation de pratiques culturales en fonction des caractéristiques des parcelles, prenant en compte les habitudes des exploitants, c'est-à-dire la façon dont ils appliquent les pratiques culturales recommandées. Il s'agirait alors d'un système d'aide à la décision à base de règles adaptatif, permettant la proposition d'alternatives spécifiques aux utilisateurs.

Projets

Les deux thèmes évoqués sont traités dans le cadre de projets de recherche en collaboration avec d'autres organismes. Ainsi l'intégration de données pluri-annuelles et multi-échelles est un des volets du projet « AI-Fruit », financé par la Région Pays de Loire, qui vise à développer des approches intégratives du déterminisme structural, génétique et écophysiologique de la qualité des fruits. Cet axe du projet est mené en commun avec le LASQUO (Université d'Angers) et l'USC (ONIRIS). La représentation et l'exploitation de données de pratiques culturales sert de support au projet « Verger Cidricole de Demain », qui vise à expérimenter en conditions réelles (chez des arboriculteurs, producteurs de pommes à cidre) la faisabilité de pratiques culturales envisagées pour réduire l'usage des produits phytosanitaires et évaluer l'efficacité environnementale et l'incidence technico-économique de ces approches. Ce projet est mené en collaboration avec des instituts techniques agricoles, chambres d'agriculture, entreprises et associations professionnelles régionales.

13. Laboratory Information Management System

Références

[1] J. Bourbeillon, C. Garbay., and F. Giroud. Mass data exploration in oncology: An information synthesis approach. *Journal of Biomedical Informatics*, 42(4):612–623, 2009.

[2] G. Buck-Sorlin, P. H. B. de Visserand M. Henke, V. Sarlikioti, G. W. A. M. van der Heijden, L. F. M. Marcelis, and J. Vos. Towards a functional-structural plant model of cut-rose: simulation of light environment, light absorption, photosynthesis and interference with the plant structure. *Annals of Botany*, 108(6):1121–1134, 2011.

[3] J-M. Celton, S. Martinez, M-J. Jammes, A. Bechti, S. Salvi, J-M. Legave, and E. Costes. Deciphering the genetic determinism of bud phenology in apple progenies: a new insight into chilling and heat requirement effects on flowering dates and positional candidate genes. *New Phytologist*, 192(2):378–392, 2011.

[4] M. Génard, N. Bertin, H. Gautier, F. Lescourret, and B. Quilot. Virtual profiling: a new way to analyse phenotypes. *The Plant Journal*, 62(2):344–355, Apr 2010.

[5] A. Kouassi, C.-E. Durel, F. Costa, S. Tartarini, E. van de Weg, K. Evans, F. Fernandez-Fernandez, C. Govan, A. Boudichevskaja, F. Dunemann, A. Antofie, M. Lateur, M. Stankiewicz-Kosyl, A. Soska, K. Tomala, M. Lewandowski, K. Rutkovski, E. Zurawicz, W. Guerra, and F. Laurens. Estimation of genetic parameters and prediction of breeding values for apple fruit-quality traits using pedigreed plant material in Europe. *Tree Genetics & Genomes*, 5:659–672, 2009.

[6] C. Renard, N. Dupont, and P. Guillermin. Concentrations and characteristics of procyanidins and other phenolics in apples during fruit growth. *Phytochemistry*, 68(8):1128–1138, 2007.

UMR TETIS (AgroParisTech - Cirad - Irstea), équipe SISO « Système d'Information Spatialisée »

Système d'Information Spatialisée : modélisation, extraction et diffusion des dOnnées et connaissances

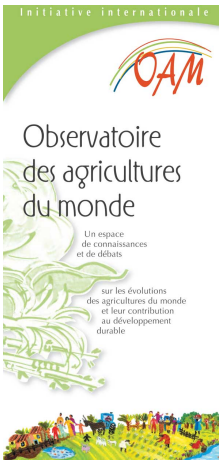
- CONTACT : Teisseire Maguelonne (Directrice de recherche)
  maguelonne.teisseire@irstea.fr
- ADRESSE : UMR TETIS (Territoires, Environnement, Télédétection et Information Spatiale) - Maison de la Télédetection 500, rue J.F.Breton 34093 MONTPELLIER Cedex 5
- WEB : http://www.tetis.teledetection.fr
- TEL : +33 (0) 4 67 54 87 62

Membres de l'équipe

- Vincent Bonnal, Ingénieur-chercheur, Cirad
- Sandra Bringay, MCF UM3, (associée)
- Vincent Douzal, Ingénieur, Irstea
- Dino Ienco, CR, Irstea
- Philippe Lemoisson, Ingénieur-chercheur, Cirad
- Philippe Morant, Ingénieur-chercheur, Cirad
- André Miralles, Ingénieur de recherche, Irstea
- Michel Passouant, Ingénieur-chercheur, Cirad
- Pascal Poncelet, Pr UM2, (associé)

Thème général de l'équipe

L'équipe SISO s'intéresse aux problèmes d'analyse, de conception et d'implémentation des Systèmes d'Information Spatialisée en y incluant les défis associés à l'intégration, au stockage des données, à la modélisation et à l'extraction des connaissances dans un objectif décisionnel. Ces travaux sont effectués dans une perspective de fiabilité, de qualité et de traçabilité afin d'assurer la réelle utilisation, la pérennité et la remobilisation de l'information et des connaissances en vue d'une exploitation ou d'une diffusion ultérieure. Les domaines d'application privilégiés de l'unité étant les territoires et l'environnement, les informations associées sont souvent multi-sources, hétérogènes et possèdent fréquemment une composante spatiale, temporelle voire spatio-temporelle. Partant de l'hypothèse que l'information est une ressource clé pour les acteurs du développement durable, l'équipe SISO traite donc des conditions nécessaires à l'appropriation et à l'usage des systèmes d'informations au service des processus de gouvernance territoriale multi-acteurs et multi-niveaux. L'équipe SISO est structurée selon trois directions qui orientent les travaux de recherche et de transfert : 1) la conception - de la formalisation des besoins à la modélisation du système d'information, 2) la gestion et l'exploitation des données - de l'intégration des sources diverses et l'interopérabilité entre systèmes à la construction de connaissances à partir des grands jeux de données, 3) la restitution et la diffusion des informations - de la cartographie dynamique au webmapping. Enfin, les chercheurs jouent un rôle majeur dans des projets fédérateurs tels que l'Equipex Geosud et OAM, l'Observatoire des Agricultures du Monde.



Description des travaux

L'utilisation de données complexes associées à l'agriculture (nature et cycle des cultures, nature du bétail utilisant les parcelles ...), à l'environnement, au territoire (découpage parcellaire ...) et aux politiques locales ou régionales (communiqués de presses, discours, orientation politique agricole ...) offre des vues à différents niveaux sur les données et sur les évènements observés dans le temps et dans l'espace. L'exploitation et la restitution de ces données en apparence très hétérogènes reste un challenge scientifique. Les travaux de recherche menés visent à lever les verrous qui concernent essentiellement les données dans leur hétérogénéité sémantique et syntaxique mais également les méthodes de modélisation, d'exploration et d'extraction de connaissances ainsi que les techniques de validation à mettre en œuvre.

Développement itératif des Systèmes d'Information

Pour rendre possible le processus de développement itératif et interactif, et satisfaire l'objectif de capitalisation, nous avons conçu et implémenté un environnement, appelé Software Development Process Model, qui permet de

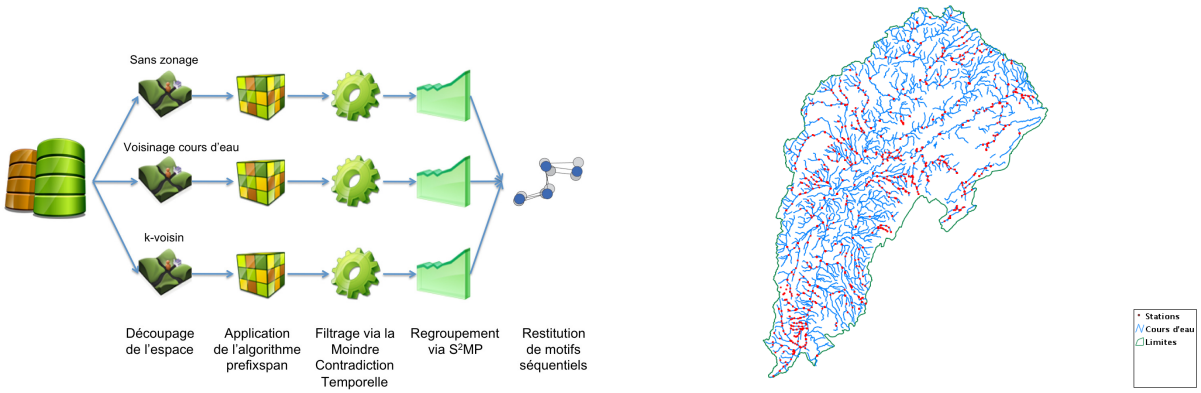


FIGURE 5 – À gauche : des données hydrobiologiques à l'évaluation des connaissances, à droite : le bassin versant de la Saône

manipuler simultanément plusieurs modèles, chacun représentant une phase de la méthode de développement (analyse, conception et implémentation). Les propriétés spatiales, temporelles et spatio-temporelles des objets géographiques sont annotées en analyse via un langage pictogrammique [4]. Une des transformations convertit les pictogrammes en concepts UML exploitables par les générateurs de code. Cette approche a été validée sur le projet européen Operational Management and Geodecisional Prototype to Track and Trace Agricultural Production (OTAG) sur la traçabilité de la viande bovine au Brésil et sur le projet Système d'Information Environnementale pour les Pesticides (SIE Pesticides). Ces travaux se poursuivent dans le cadre du projet Miriphyque en partenariat avec l'INRA, l'UIPP, Arvalis et AgroParisTech. D'autres travaux, menés en collaboration avec le LIRMM, portent sur des méthodes de factorisation favorisant l'émergence de nouveaux concepts thématiques [6].

Extraction de connaissances à partir de données environnementales

Pour l'extraction de connaissances, l'un des principaux challenges consiste à mettre en évidence des corrélations entre les informations, éventuellement historisées, qui ne pouvaient l'être au préalable. Dans ce contexte, la découverte de relations spatio-temporelles est une solution pertinente et adaptée. En effet, parmi des événements observés sur différents lieux, ces relations permettront de mieux appréhender la manière dont les éléments d'un système interagissent les uns avec les autres (voir figure 5). En se basant sur les définitions de motifs multidimensionnels et spatio-séquentiels [2], il s'agit de mettre en œuvre des mécanismes d'extraction de connaissances typiques et atypiques adaptés à la richesse et complexité des données issues des domaines de la santé, de l'environnement et des territoires et pour lesquelles aucune solution n'est encore satisfaisante. Ces travaux couvrent tout le processus d'extractions de connaissance et permettent par exemple la prise en compte des caractéristiques spatio-temporelles des données hydrobiologiques [1]. Ils sont réalisés en collaboration avec l'Université de Nouvelle Calédonie et en partenariat avec l'Institut de Veille Sanitaire pour le suivi des épidémies de dengue. Ils se poursuivront dans le projet Fresqueau - (2011-2014), financement ANR Modèle Numérique - Masse de données (voir page 44).

Classification, Cartographie et Propagations basées sur les motifs fréquents

L'utilisation des motifs fréquents comme base d'un mécanisme de classification est assez récent. Nous avons participé à ces contributions en définissant des algorithmes de classification à partir d'itemsets fréquents et de mesures issues de la théorie de l'information dans le contexte des réseaux sociaux [3] ou à partir de motifs séquentiels pour construire une cartographie automatique des types de cultures à partir d'images satellites moyenne résolution Modis au Mali [7] (voir figure 6). Pour ce dernier point, nous avons proposé une approche originale de sélection des dimensions statiques de description permettant de se focaliser dans un deuxième temps sur les motifs discriminants résumant les séries de données temporelles [8]. Ces travaux se poursuivent afin d'identifier les informations pertinentes permettant de décrire et modéliser les propagations spatiales et temporelles au sein de gros volumes de données. Nous avons ainsi défini des règles graduelles et spatio-temporelles [5] permettant de décrire les évolutions successives. Une application sur des données d'occupation des sols est en cours.

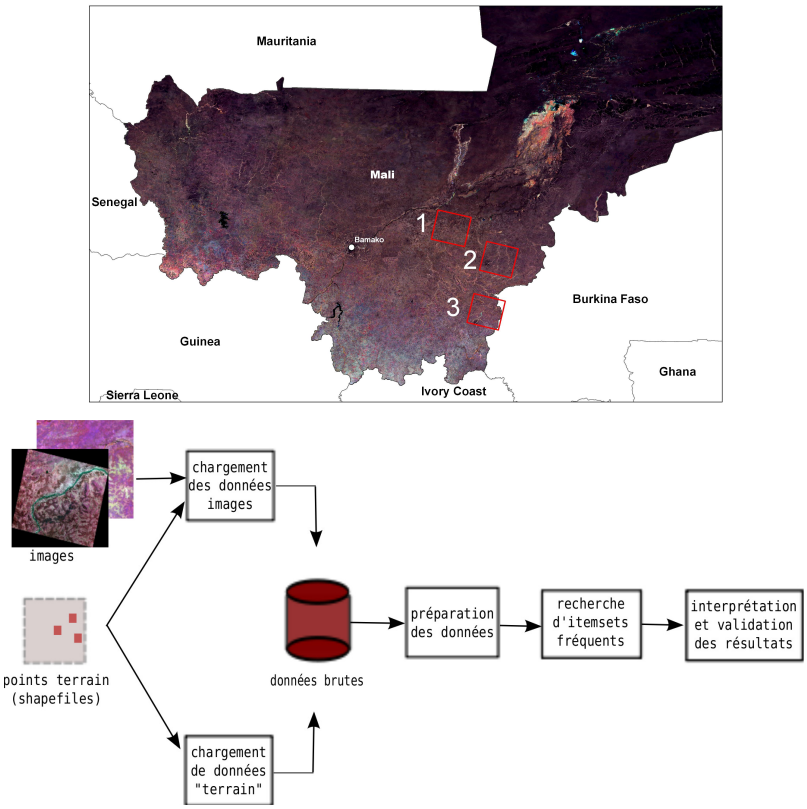


FIGURE 6 – En haut : Carte du Mali sur fond de série temporelle MODIS 2007 et emprise géographique des trois images SPOT de validation (Cinzana, Koutiala, Sikasso), en bas : Processus d'extraction de connaissance pour la cartographie du Mali à l'aide de motifs

Références

[1] H. Alatrista Salas, S. Azéand J. Bringay, F. Cernesson, F. Flouvat, N. Selmaoui-Folcher, and M. Teisseire. Recherche de séquences spatio-temporelles peu contredites dans des données hydrologiques. *Revue des nouvelles technologies de l'information (RNTI)*, page 24, 2011.

[2] H. Alatrista Salas, S. Bringay, F. Flouvat, N. Selmaoui-Folcher, and M. Teisseire. The pattern next door: Towards spatio-sequential pattern discovery. *16th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD)*, page 12, 2012.

[3] F. Bonchi, C. Castillo, and D. Ienco. Meme ranking to maximize posts virality in microblogging platforms. *Journal of Intelligent Information Systems*, page 31, 2011.

[4] A. Miralles, F. Pinet, and Y. Bédard. Describing spatio-temporal phenomena for environmental system development: An overview of today's needs and solutions. *IJAEIS*, 1(2):68–84, 2010.

[5] P. Nhat Hai, P. Poncelet, and M. Teisseire. Moving objects: Combining gradual rules and spatio-temporal patterns. In *ICSDM, IEEE International Conference on Spatial Data Mining and Geographical Knowledge Services*. IEEE, 2011.

[6] A. Osman Guedi, A. Miralles, M. Huchard, and C. Nebut. Analyse de l'évolution d'un modèle : vers une méthode basée sur l'analyse formelle de concepts. In *Inforsid*, page 8, 2011.

[7] Y. Pitarch, E Vintrou, F. Badra, A Begue, and M Teisseire. Mining sequential patterns from modis time series for cultivated area mapping. *Advancing Geoinformation Science for a Changing World*, pages 45–62, 2011.

[8] E. Vintrou, Y. Pitarch, A Begue, and M Teisseire. Des relevés terrain aux données satellitaires : application au mali. *Revue Internationale de Géomatique*, page 20, 2011.

EA UNC-PPME, équipe « Informatique »

Le thème développé par l'équipe Informatique du PPME[14] est « Extraction des connaissances et modélisation spatio-temporelle des phénomènes dynamiques en environnement et en santé ».

- CONTACT : Nazha Selmaoui-Folcher
  nazha.selmaoui@univ-nc.nc,
  http://pages.univ-nc.nc/~selmaoui
- ADRESSE : Université de la Nouvelle Calédonie (UNC) - PPME
  BP R4 98851 Nouméa Cedex
  Nouvelle Calédonie
- WEB : http://ppme.univ-nc.nc
- TEL : +687 29 03 14

Membres de l'équipe

- Hugo Alatrista Salas, Doctorant
- Frédéric Flouvat, Maître de conférences de l'UNC
- Claude Pasquier, CR CNRS (mis à disposition)
- Isabelle Rouet, Ingénieur Contractuel de l'UNC
- Jérémy Sanhes, Doctorant de l'UNC
- Nazha Selmaoui-Folcher, Maître de conférences de l'UNC

Thème général de l'équipe

L'équipe développe ses compétences dans l'analyse et la gestion des données spatio-temporelles. Elle a pour but de modéliser des phénomènes dynamiques en environnement et santé. L'objectif scientifique est de construire, à partir des masses de données spatio-temporelles disponibles, des modèles dynamiques permettant d'observer et d'étudier l'évolution de l'environnement sur un questionnement précis (par exemple l'érosion, propagation de la Dengue, suivi de l'état de santé des récifs coralliens, etc.).

A l'heure actuelle, les scientifiques qui travaillent sur les problématiques environnementales se heurtent à des difficultés grandissantes dans le traitement et l'analyse des données qu'ils recueillent. En effet, les bases de données sont hétérogènes, multi-échelles, incomplètes, imprécises, et contiennent des objets complexes (images multisources satellitaires, aériennes ou terrestres ; modèles numériques de terrain, mesures physiques au sol, observations qualitatives). L'exploitation de ces masses de données spatio-temporelles générées pose donc un grand nombre de problèmes. Dans le contexte calédonien et plus largement, dans le sud du pacifique, des initiatives se sont multipliées pour accompagner des programmes de recherche dans l'environnement et la santé (GOBS, STRATOM, Projet CNRT-PBV, CNRTHyperspectral, MOM-Dengue, etc.). Dans ce contexte, et dans le cadre de ces programmes pluridisciplinaires, le PPME collecte et exploite depuis longtemps des données spatio-temporelles. Il met en oeuvre des approches multiscalaires permettant de mieux comprendre les phénomènes environnementaux et sociaux qui se dérobent aux analyses conventionnelles. Dans cette optique, l'équipe Informatique a pour objectif de concevoir, développer et mettre en oeuvre des nouvelles méthodes de modélisation, supervision et surveillance des phénomènes spatio-temporels adaptées aux masses de données (MNT, images satellites, données météorologiques, ...), dans le but d'une gestion améliorée de la santé et de l'environnement.

Nous nous intéressons donc à l'extraction et la gestion des connaissances intégrant l'information spatiale et temporelle. D'un point de vue valorisation, l'équipe informatique met aussi l'accent sur le transfert des résultats vers la société civile.

Projets

Les projets auxquels nous participons sont :

- Projet ANR COSINUS 2010 intitulé « Fouille de données spatio-temporelles : application à la compréhension et à au suivi de l'érosion » (voir page 42).
- Projet CNRT Nickel & Environnement intitulé « Fonctionnement des petits bassins versant ».
- Projet Labex CORAIL intitulé « Les récifs coralliens face aux changements globaux de la planète ».

14. Pôle Pluridisciplinaire de la Matière et de l'Environnement EA-3325

UMR MIA (AgroParisTech - INRA), Groupe AAA « Apprentissage Artificiel AgroParisTech »

- CONTACT : Antoine Cornuéjols (Pr) antoine.cornuejols@agroparistech.fr
- ADRESSE : AgroParisTech – Département MMIP, 16 rue Claude Bernard, F-75231 Paris Cedex 05
- WEB : http://www.agroparistech.fr/mia/
- TEL : +33 (0) 1 44 08 16 16

Membres du groupe

Permanents :
- Antoine Cornuéjols
- Christine Martin
- Laurent Orseau

Doctorant :
- Ghazal Jaber

Thème général de recherche du groupe

Notre groupe est spécialisé sur l'apprentissage artificiel et la fouille de données. Sur le plan du développement des concepts et des méthodes, notre groupe s'intéresse spécialement à l'apprentissage à partir de flux de données et à l'apprentissage à partir d'exemples positifs seuls.

L'apprentissage à partir de flux de données prend une importance croissante, spécialement dans les sciences du vivant et de l'environnement avec le suivi d'organismes, de cultures ou de tendances environnementales. D'un point de vue méthodologique, il introduit des problèmes d'apprentissage à partir de données qui ne sont plus indépendantes et identiquement distribuées comme c'est supposé être le cas dans la théorie statistique de l'apprentissage artificiel. Il faut alors savoir identifier des ruptures dans les données temporelles, adapter les prédictions réalisées aux dérives de concept, transférer éventuellement ce qui a été appris d'une tâche à une autre. Cela introduit des questions sur la gestion de la mémoire du passé, le niveau d'abstraction optimal et la meilleure manière de transférer des connaissances d'un instant à un autre ou d'un domaine à un autre. Notre groupe étudie ces questions sur un plan théorique, par exemple l'apprentissage en-ligne par renforcement, et par le développement de nouvelles méthodes d'apprentissage : apprentissage par comités d'experts adaptatifs.

Le second grand axe de nos recherches concerne l'apprentissage à partir d'exemples positifs seuls et s'étend à des questions plus générales d'apprentissage non supervisé. Par exemple, dans le cas de l'amarrage de protéines, les bases de données ne répertorient que des exemples connus d'amarrages, pas de contre-exemples. Plus généralement, les données sur l'environnement ne sont souvent pas contrôlées, elles ne contiennent donc pas d'exemples négatifs, comme on pourrait cherche à en avoir en médecine. Dans ces conditions, la mesure de performance et la définition d'un critère à optimiser deviennent des questions à part entière et des algorithmes d'apprentissage spécifiques sont à développer. Les contributions de notre groupe concernent en particulier le développement de méthodes de comités d'experts pour l'apprentissage non supervisé et pour des problèmes de sélection d'attributs.

Les champs d'application étudiés dans notre groupe concernent principalement la bioinformatique et les données environnementales. Nous nous focalisons sur ces dernières dans ce dossier. Plusieurs types de questions sont à l'ordre du jour sur l'environnement. Celles-ci couvrent un spectre depuis des questions locales jusqu'à des questions globales. Par exemple, quelle est la production de N2O dans les champs en fonction des cultures, de la quantité et de la composition des engrais utilisés, du climat, des saisons, etc. ? Quelle est la production de méthane par les animaux d'élevage, en particulier les bovins et les ovins ? Comment se déplacent les polluants dans la basse atmosphère à l'échelle d'un territoire, d'un pays, d'un continent ? Les grands courants océaniques, tel le Gulf-Stream, présentent-ils des variations significatives de leur régime et peut-on prédire leur comportement dans les décennies à venir ?

La résolution de ces questions passe par l'exploitation de grandes masses de données hétérogènes dans leur qualité, dans l'objectif ayant présidé à leur recueil, et souvent glanées dans des lieux et à des dates très variés. De plus, à côté des données brutes, pas toujours faciles à obtenir, il existe de nombreux articles sur ces questions publiés au cours du temps, là aussi de caractéristiques hétérogènes. Il est essentiel de pouvoir tirer parti de ces publications en les mettant en regard, et en réalisant ce que l'on appelle une méta-analyse : combiner des analyses locales pour obtenir une réponse générale à certaines questions.

L'étude de l'environnement demande donc que soient mises en œuvre des méthodes d'apprentissage à partir de données dépendantes spatialement et temporellement, des méthodes d'analyse de publications, ainsi que de fusion et réconciliation de connaissances, c'est-à-dire de méta-analyse.

Projets en cours

Projet ANR « Holyrisk »

Ce projet porte sur l'étude de l'analyse de risques sanitaires. Des experts, américains et français, analysent des rapports sur des recommandations sanitaires et étiquettent des parties de texte en fonction des types de risques mentionnés, des niveaux de certitude exprimés et des catégories d'évidences servant à l'appui des recommandations. Deux questions principales motivent cette étude. D'une part, ces analyses d'experts sont coûteuses et longues, d'autre part, elles sont sujettes à variations, aussi bien entre experts qu'au cours du temps. Une analyse automatisée permettrait à la fois de réduire le coût de l'analyse, mais aussi d'étudier de manière objective les différences d'appréciation entre les experts des deux côtés de l'Atlantique. D'un point de vue recherche, ce travail demande le développement de méthodes d'étiquetage de textes, ces étiquettes appartenant à des ontologies. Il faut également tenir compte d'interdépendances textuelles à plus ou moins longue distance. Par ailleurs, une capacité d'apprentissage incrémental est souhaitée pour que les experts puissent utiliser l'outil pour développer et préciser leur propre expertise. Finalement, il est également envisagé de chercher à détecter des éventuelles dérives d'appréciation par les experts au cours du temps, reflétant par exemple de nouvelles connaissances ou de nouvelles prises de conscience sociétales.

Projet « ExtraEx » (Extraction de paramètreEs cruciaux pour une aide à la synthèse et à la prise de décision à partir de grandes bases de données bibliographiques. Application à l'estimation des émissions de gaz à effets de serre)

Le projet ExtraEx associe quatre équipes de AgroParisTech et de l'INRA. Son objectif est de fournir des outils et des méthodes d'estimation de chiffres à partir de bases de données bibliographiques et d'y associer une mesure d'incertitude. Très spécifiquement, dans le cadre de ce projet, l'objectif est de ré-examiner l'estimation de deux nombres. L'un est la quantité de N2O produite par hectare de culture en moyenne dans le monde. L'autre est la quantité de méthane produite par les animaux d'élevage. La valeur actuellement admise pour ces deux nombres, et utilisée par le GIEC dans ses prévisions sur le réchauffement climatique, résulte d'analyses datant d'une dizaine d'années à partir de corpus de quelques centaines d'articles. En outre, l'incertitude sur la valeur de ces deux nombres est importante, et, plus encore, il manque un intervalle de confiance fiable autour de ces valeurs. L'impact de ces incertitudes sur les prévisions du GIEC est potentiellement significatif. Dans ce contexte, le projet vise à reprendre et étendre ces travaux en recourant à des méthodes d'apprentissage artificiel, de méta-analyse et de statistiques avancées développées à cet effet et en incluant des données récemment publiées sur ces questions environnementales. Les méthodes développées pour ce projet trouveront un champ d'application beaucoup plus large dans des études portant sur des questions environnementales globales.

Références

[1] A.F Bouwman. Direct emission of nitrous oxide from agricultural soils. *Nutrient Cycling in Agroecosystems*, 1996.

[2] A. Cornuéjols and C. Martin. Unsupervised object ranking using not even weak experts. In *proceedings of 18th International Conference on Neural Information Processing (ICONIP)*, 2011.

[3] G. Jaber, A. Cornuéjols, and P. Tarroux. Predicting concept changes using a committee of experts. In *proceedings of 18th International Conference on Neural Information Processing (ICONIP)*, 2011.

[4] C. Martin and A Cornuéjols. Détection non supervisée d'une sous-population par méthode d'ensemble et changement de représentation itératif. In *Actes de la conférence Extraction et Gestion des Connaissances (EGC-2012)*, 2012.

[5] L. Orseau. Optimality issues of universal greedy agents with static priors. In *Proc. 21st Int. Conf. on Algorithmic Learning Theory (ALT'10)*, 2010.

[6] K. Smith, L. Bowman, and Braatz. N20: direct emissions from agricultural soils, 2000.

Projet FOSTER

- CONTACT : Nazha Selmaoui-Folcher
  nazha.selmaoui@univ-nc.nc,
  http://pages.univ-nc.nc/~selmaoui
- ADRESSE : Université de la Nouvelle Calédonie (UNC) - PPME
  BP R4 98851 Nouméa Cedex
  Nouvelle Calédonie
- WEB : http://foster.univ-nc.nc/
- TEL : +687 29 03 14

Membres du projet

- PPME, EA 3325 Université de Nouvelle Calédonie (N. Selmaoui-Folcher)
- LIRIS, CNRS UMR 5205, Lyon (C. Rigotti)
- LSTII, CNRS UMR 7005, Strasbourg, (P. Gançarski)
- LISTIC, EA 3703, Annecy, (N. Méger)
- Bluecham SAS (D. Lille)

Le projet FOSTER "FOuille de données Spatio-Temporelles : application à la compréhension et à la surveillance de l'ERosion" est un projet ANR-10-COSINUS (Conception et Simulation) coordonné par le PPME et a débuté en janvier 2011 pour une durée de 39 mois. Il regroupe un consortium de 4 laboratoires de recherche (LSIIT de l'Université de Strasbourg, LIRIS de l'INSA de Lyon, LISTIC de l'Université de Savoie) et une entreprise privée calédonienne (Bleucham SAS)

- Coordinateur du projet et dans l'équipe du PPME : Nazha Selmaoui-Folcher (nazha.selmaoui@univ-nc.nc)
- Site web du projet : http://foster.univ-nc.nc

Contexte

La compréhension de phénomènes liés à l'environnement se heurte à deux principaux verrous : la complexité des données et la complexité des phénomènes étudiés. De plus, les objectifs en terme de modélisation peuvent être variés (par exemple, détection de changements abrupts ou suivi de phénomènes évolutifs). L'étude de systèmes dynamiques évoluant dans l'espace et dans le temps soulève donc aussi des questions sur la nature de la relation spatiale et le type d'évolution temporelle à étudier.

L'érosion est un des problèmes environnementaux prioritaire et important en Nouvelle Calédonie. Elle a un fort impact sur les écosystèmes terrestres et côtiers (bassin versant, la mangrove, les récifs coraliens etc.). Les feux de brousse, la déforestation et/ou les activités humaines telles que l'exploitation minière à ciel ouvert accélèrent le phénomène d'érosion dans les zones montagneuses en particulier sur les sols latéritiques. Les précipitations fréquentes et intenses (dépressions tropicales, cyclones) favorisent le transport sédimentaire provenant des sols dégradés et sensibles vers les plaines côtières et à la mer le long des lignes de drainage principales avec un effet immédiat sur les activités humaines telles que l'agriculture et la pêche. Il est donc nécessaire d'identifier les éléments clés des processus d'érosion pour la gestion durable de l'environnement et de planification.

Objectif scientifique du projet

Le projet FOSTER a pour objectif de construire, à partir des masses de données disponibles, des modèles permettant d'observer et d'éventuellement prédire l'évolution de l'environnement sur un questionnement précis. Ce projet va s'intéresser aux principales phases de la découverte de connaissances à partir de données : la préparation des données i.e. prétraitement et détection des objets d'intérêt (par exemple segmentation des images), puis calculs de motifs et de modèles suivis de leurs interprétations. Un critère de succès sera la mise en oeuvre d'un processus complet d'exploitation des données environnementales disponibles pour l'étude de l'érosion sur le territoire de la Nouvelle- Calédonie.

L'originalité du projet FOSTER tient à la fois dans le questionnement environnemental retenu (l'érosion) et dans les méthodes originales de découverte de connaissances qui devraient émerger du fait de la complémentarité des laboratoires partenaires spécialistes en fouille de données et en imagerie.

Ainsi, nous proposons de fournir aux spécialistes de l'érosion et de la gestion du risque, de nouveaux outils et méthodes pour exploiter les données hétérogènes qu'ils collectent ou peuvent collecter. Notre objectif scientifique principal est donc d'étudier la découverte de connaissances à partir de données environnementales spatio-temporelles pour la compréhension et l'évaluation de l'aléa érosion. Le processus mis au point sera mis au service de la compréhension de l'érosion hydrique, courante en Nouvelle Calédonie. Il sera également appliqué et testé dans le cadre de l'érosion par glissement de terrain (Bassin de Barcelonnette, Sud-Est des Alpes françaises). Ce double contexte d'application permettra de valider la généricité des méthodes et outils proposés.

Dans cette optique nous voulons plus particulièrement proposer de nouvelles méthodes de fouille de données spatio-temporelles avec des avancées dans les domaines suivants :

– L'extension des techniques d'extraction de motifs, notamment les méthodes complètes d'extraction sous contraintes, à un contexte spatio-temporel.
– Le problème très important de la classification automatique sous contraintes pourra être étudié dans ce contexte stimulant des données spatio-temporelles.

Nous proposons d'étudier une variété de méthodes de fouilles de données qui vont de la recherche de motifs locaux (épisodes, motifs séquentiels) jusqu'à la construction de modèles globaux (clustering, classifieurs) à partir de ces motifs en prenant en compte l'aspect spatial et temporel.

Travaux en lien avec la problématique de l'érosion

Plusieurs publications sur les travaux réalisés par l'équipe Informatique du PPME en lien avec la problématique de l'érosion sont référencées dans [3, 5, 6, 2, 7, 4, 1]. Les publications sur les travaux en cours dans le cadre du projet FOSTER sont disponibles sur le site WEB http://foster.univ-nc.nc.

Références

[1] E. Desmier, F. Flouvat, D. Gay, and N. Selmaoui-Folcher. A clustering-based visualization of colocation patterns. In *IDEAS*, pages 70–78, 2011.

[2] F. Flouvat, N. Selmaoui-Folcher, D. Gay, I. Rouet, and C. Grison. Constrained colocation mining : application to soil erosion characterization. In *SAC*, 2010.

[3] D. Gay, I. Rouet, M. Mangeas, P. Dumas, and N. Selmaoui. Assessment of classification methods for soils erosion risks. In *International Congress on Modelling and Simulation 2007 (MODSIM'07)*, pages 2659–2665, 2007.

[4] L. Mabit, N. Selmaoui-Folcher, and F. Flouvat. Modélisation de la dynamique de phénomènes spatio-temporels par des séquences de motifs. In *EGC*, pages 455–466, 2011.

[5] I. Rouet, D. Gay, M. Allenbach, N. Selmaoui, J. Maura, and D. Lille. Semi-automatic detection and labeling of erosion areas using domain knowledge. an appraisal from southeast new caledonia (sw pacific). In *ISPIE Asia-Pacific Remote Sensing*, volume 7150, 2007.

[6] I. Rouet, D. Gay, Michel Allenbach, N. Selmaoui, A.-G. Ausseil, M. Mangeas, J. Maura, P. Dumas, and D. Lille. Tools for soil erosion mapping and hazard assessment: application to new caledonia, sw pacific. In *International Congress on Modelling and Simulation (MODSIM'09)*, pages 1986–1992, 2009.

[7] N. Selmaoui-Folcher, F. Flouvat, D. Gay, and I. Rouet. Spatial pattern mining for soil erosion characterization. *IJAEIS*, 2(2):73–92, 2011.

Projet FRESQUEAU

Fouille de données pour l'évaluation et le suivi de la qualité hydrobiologique des cours d'eau

- CONTACT :
  - Florence Le Ber
    florence.leber@engees.unistra.fr,
  - Maguelonne Teisseire
    maguelonne.teisseire@teledetection.fr
- WEB : http://engees-fresqueau.unistra.fr/

Laboratoires et entreprises participant au projet

- UMR LHYGES, http://lhyges.u-strasbg.fr/
- UMR TETIS, http://tetis.teledetection.fr/
- UMR LIRMM, http://www.lirmm.fr/
- UMR LSIIT, http://lsiit.u-strasbg.fr/
- AQUASCOP, http://www.aquascop.fr/
- AQUABIO, http://www.aquabio-conseil.fr/

Thème général

Le projet FRESQUEAU a été selectionné dans le cadre de l'appel d'offre modèles numériques 2011 de l'ANR. C'est un projet de trois ans, qui s'inscrit dans la thématique "masse de données" et s'applique à l'étude de la qualité des écosystèmes aquatiques. Pour cela, le consortium regroupe un ensemble d'experts en structuration et extraction de connaissances à partir de données et un ensemble d'experts en hydroécologie.

Objectifs

L'objectif de préserver ou restaurer le bon état des masses d'eau, imposé par la Directive Cadre Européenne sur l'eau (2000), met en exergue la nécessité de disposer d'outils opérationnels pour aider à l'interprétation des informations complexes concernant les cours d'eau et leur fonctionnement, ainsi que pour évaluer l'efficacité des programmes d'actions engagés. Dans cette perspective, le projet FRESQUEAU a pour but de développer de nouvelles méthodes pour étudier, comparer et exploiter l'ensemble des paramètres disponibles concernant l'état des cours d'eau ainsi que les informations décrivant les usages et les mesures prises. Les outils développés seront intégrés dans une plateforme d'aide à l'interprétation du fonctionnement des cours d'eau. L'originalité de l'approche proposée est de mettre en relation des données structurelles et des données fonctionnelles *via* la mobilisation d'un ensemble de méthodes novatrices, et de mettre ainsi en place un véritable processus d'aide à la découverte de connaissances.

Enjeux

L'enjeu est à la fois applicatif et théorique. Il s'agit (1) de développer un outil permettant d'évaluer le fonctionnement global d'un cours d'eau sur la base des différents compartiments de l'écosystème ; (2) de progresser dans les méthodes d'extraction de connaissances à partir de gros volumes de données hétérogènes, temporelles et spatiales.

Plus précisément, d'un point de vue applicatif, le projet contribuera à répondre à deux enjeux spécifiques : (1) approfondir la connaissance du fonctionnement des cours d'eau par l'analyse des taxons à la base des indices biologiques et (2) relier les sources de pressions sur le milieu à la qualité physico-chimique et biologique des cours d'eau. Le projet s'appuie pour cela sur les données physico-chimiques et biologiques produites par les Agences de l'Eau et par l'Office National de l'Eau et des Milieux Aquatiques (ONEMA), et complétées par les mesures fines opérées par le LHYGES [4]. Sont également disponibles des données décrivant le réseau hydrographique, l'occupation du sol, les stations d'épuration, complétées localement par des enquêtes sur les activités agricoles et sur les actions de restauration, et par des cartographies fines dans les espaces rivulaires, produites par TETIS. Toutes ces données se caractérisent par une grande hétérogénéité et une grande complexité, tant par leurs formes propres (données incomplètes, entachées d'incertitudes, provenant de différentes sources, etc.) que par leurs structures et échelles spatiales et temporelles.

Démarche

Pour exploiter ces données, nous adopterons une démarche d'extraction de connaissances. Travaillant d'abord à la structuration et à la mise en forme des données, nous proposons d'explorer et faire collaborer différentes approches de fouille de données dans un souci constant d'évaluation par les experts. Nous travaillerons à la définition de méthodes de fouille adéquates et efficaces, en particulier pour la découverte de relations spatio-temporelles. Nous étudierons la mise en œuvre de méthodes qualita-

tives sur des gros tableaux de données, posant des problèmes de changement d'échelle. La participation de deux bureaux d'études nous garantit les moyens d'une validation *sur le terrain*. La plateforme finale inclura un entrepôt de données, une typologie de stations, et un ensemble de méthodes d'analyse et de fouille de données. La typologie des stations permettra de guider l'analyse et l'interprétation des relevés en station, en lien avec le fonctionnement hydrobiologique du cours d'eau et les pressions constatées, dans le but d'évaluer son état global.

Le programme scientifique se décompose en cinq grandes tâches. La première est une tâche thématique qui s'attache aux données (recensement, évaluation, structuration dans des outils adaptés [2], pré-traitement) et à l'évaluation des résultats des analyses effectuées dans les tâches suivantes. Les deux tâches suivantes sont des tâches méthodologiques concernant l'évaluation puis la combinaison de méthodes de fouille ou d'analyse de données spatio-temporelles (recherche de motifs séquentiels fréquents [7, 6], analyse de concepts formels et treillis relationnels [5, 1], propositionalisation de données relationnelles [3], approches statistiques). La tâche 4 porte sur le développement d'un outil à destination des thématiciens, outil qui sera testé par les bureaux d'études partenaires du projet. Finalement la tâche 5 est une tâche d'évaluation et de valorisation de l'ensemble de la démarche.

Références

[1] A. Bertaux, F. Le Ber, A. Braud, and Trémolières M. Identifying ecological traits: a concrete fca-based approach. In S. Ferré and S. Rudolph, editors, *7th International Conference on Formal Concept Analysis, ICFCA 2009, Darmstadt*, pages 224–236, 2009.

[2] S. Bimonte. A web-based tool for spatio-multidimensional analysis of geographic and complex data. *International Journal of Agricultural and Environmental Information Systems*, 1(2):42–67, 2010.

[3] A. Braud and N. Lachiche. Propositionaliser des attributs numériques sans les discrétiser, ni les agréger. In *Extraction et Gestion des Connaissances (EGC 2011)*, pages 1–6, 2011.

[4] C. Grac, A. Braud, F. Le Ber, and M. Trémolières. Un système d'information pour le suivi et l'évaluation de la qualité des cours d'eau – Application à l'hydro-écorégion de la plaine d'Alsace. *RSTI - Ingénierie des Systèmes d'Information*, 16:9–30, 2011.

[5] M. Huchard, M. Rouane Hacene, C. Roume, and P. Valtchev. Relational concept discovery in structured datasets. *Annals of Mathematics and Artificial Intelligence*, 49(1-4):39–76, 2007.

[6] M. Plantevit, A. Laurent, D. Laurent, M. Teisseire, and Y. W. Choong. Mining multidimensional and multilevel sequential patterns. *Transactions on Knowledge Discovery from Data (TKDD)*, 4(1):1–37, 2010.

[7] C. Raïssi, P. Poncelet, and M. Teisseire. Towards a new approach for mining frequent itemsets on data stream. *Journal of Intelligent Information Systems*, 28(1):23–36, 2007.

Projet PAYOTE 2

Caractérisation et simulation de paysages agricoles réalistes

Membres du projet

- Orpailleur, INRIA NGE LORIA, Nancy (F. Le Ber, S. Da Silva)
- DREAM, INRIA Centre Bretagne Atlantique, Rennes (T. Guyet, M.-O. Cordier)
- INRA-SAD Paysage, Rennes (H. Boussard)
- INRA MIAJ, Jouy-en-Josas, (H. Monod, K. Adamczyk)
- INRA UBIA, Toulouse (P. Chabrier)

Thème général du projet

Le projet PayOTe 2 (Paysage Ou Territoire) est un projet INRA-INRIA de deux ans (2010-2011) qui a hérité d'un premier projet (2009-10) et s'est poursuivi dans un réseau (PayOTe 3) porté par l'INRA. Il regroupe des équipes de recherche françaises, INRA et INRIA, qui s'intéressent à la caractérisation et à la génération de paysages ou de territoires agricoles. "Paysage" s'entend ici comme une mosaïque parcellaire 2D, incluant la configuration (géométrie des parcelles) et la composition (couverture du sol).

L'objectif du projet est de renforcer les collaborations établies et de privilégier des thématiques communes et porteuses définies précédemment en concertation avec des équipes d'agronomes et d'écologues : d'une part la fouille et la caractérisation des paysages, d'autre part la génération de paysages, à base de pavages ou de bibliothèques de formes (ou motifs locaux). La représentation spatio-temporelle des paysages (librairie APILand, INRA Rennes) a été également étudiée. Ces travaux sont menés en lien avec la Zone Atelier (ZA) Armorique [15] qui a fourni les données nécessaires aux réalisations des expérimentations.

Description des travaux

Contexte

En recherche agronomique, les processus agro-écologiques sont étudiés en tenant compte des interactions avec le paysage. Un paysage est vu ici comme une mosaïque (spatialement) bidimensionnelle de quelques km^2 à quelques dizaines de km^2. Cette mosaïque est composée d'éléments surfaciques (parcelles agricoles, bâtis, routes, ...) ou linéaires (haies, routes, ...), voire ponctuels (arbres). Elle évolue dans le temps (succession de cultures, modification du parcellaire, etc.).

De plus en plus de simulateurs de processus agro-écologiques, développés par les thématiciens, écologues, agronomes ou géographes, intègrent l'influence de l'organisation spatio-temporelle du paysage. Le paysage devient une donnée d'entrée qu'il faut caractériser et faire varier pour en comprendre le rôle. En pratique, les thématiciens n'utilisent souvent que des paysages réels particuliers pour faire leurs expérimentations. Il se fait alors sentir le besoin non seulement d'outils pour analyser les paysages réels disponibles, mais également d'outils pour générer, de façon plus ou moins contrôlée, des paysages artificiels – complémentaires des données réelles disponibles – afin d'examiner différents scénarios ou de tester plus largement certains modèles pour en évaluer les résultats.

Génération de paysages agricole réalistes

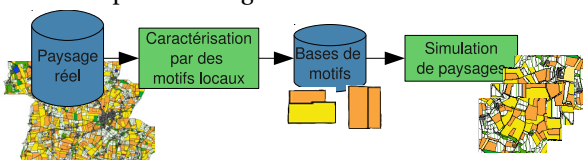
Le réalisme d'un paysage agricole vu "à plat" porte sur deux aspects :

1. la géométrie (i.e. le parcellaire) : les limites en 2D des parcelles agricoles, les regroupements de parcelles (séparés par les routes ou les haies).
2. l'assolement : l'attribution des occupations du sol à chaque parcelle.

Dans certains travaux (e.g. [1]), la dimension temporelle est également prise en compte (principalement les rotations de cultures).

L'approche de génération de paysage développée dans la projet Payote est issue des travaux menés pour l'étude de flux de transgènes [4]. Cette approche peut être résumée par le schéma ci-dessous :

1. un algorithme de fouille de données extrait de manière automatique une caractérisation d'un paysage à partir de paysages réels (numérisés),
2. cette caractérisation du paysage sert à générer un nouveau paysage qui reproduira les caractéristiques des paysages d'entrées,
3. les paysages générés servent de support à la simulation de processus agro-environnementaux.



Dans les travaux menés principalement par l'équipe Orpailleur (Nancy) et MIAJ (Jouy-en-Josas), l'objectif est de

15. Zone Atelier Armorique : http://osur.univ-rennes1.fr/zoneatelier-armorique/.

réaliser un pavage réaliste de l'espace vis-à-vis du processus étudié. Dans le cadre de l'étude du transfert des gènes, les caractères importants du paysage sont la taille et la distribution spatiale des parcelles. La solution développée depuis plusieurs années utilise des modèles statistiques de la distribution des centres de parcelles. Cette distribution peut être apprise à partir d'un paysage réel et est utilisée pour générer de nouveaux parcellaires sous la forme d'un ensemble de points dont la distribution spatiale correspond à celle d'un parcellaire réel.

Les limites des parcelles sont ensuite définies à l'aide de méthodes de construction de pavages à partir d'une distribution de points (e.g. la tessélation de Voronoï). Dans des travaux antérieurs [4], il a été montré que les parcelles obtenues par une tessélation de Voronoï reproduisent correctement la distribution des tailles et des formes observées sur le paysage réel (paysage de grandes cultures), vis-à-vis du processus étudié (flux de transgènes). Pour traiter d'autres types de paysages et d'autres processus, des travaux sont menés sur de nouvelles méthodes de pavage [5].

Dans les travaux menés principalement par les équipes DREAM et INRA-SAD Paysage, un paysage est représenté par un graphe de parcelle [3]. Chaque parcelle est un nœud du graphe associé à des attributs catégoriels (en particulier, l'occupation du sol), tandis que les arcs entre nœuds représentent des relations entre ces parcelles. Comme dans le langage Ocelet [16], les relations entre parcelles peuvent être des relations spatiales (e.g. proximité ou connexité), mais également des relations fonctionnelles (e.g. appartenance à une même exploitation).

Des algorithmes de fouille de graphe sont alors utilisés pour caractériser ce graphe de parcelles par l'ensemble des sous-graphes fréquents qu'il comporte. Dans la phase de génération, ces sous-graphes fréquents sont réutilisés pour associer des parcelles agricoles vierges à des occupations du sol. En procédant ainsi, les parcellaires générés reproduisent les structures locales fréquentes du paysage réel. Les résultats montrent que les surfaces globales des cultures sont également reproduites.

En mixant les deux approches, nous avons pu mettre en place un outil capable de générer à la fois des parcellaires agricoles s'appuyant sur des réseaux des routes existants et de générer les structures locales d'occupations du sol à partir de sous-graphes fréquents. La figure 7 montre à gauche le paysage réel et à droite le paysage généré selon cette méthode (la méthode de pavage utilisée est la tessélation de Voronoï).

16. Projet ANR STAMP (2007).

Caractérisation des linéaires dans les paysages

Le travail engagé dans le cadre de la thèse de Sébastien da Silva, en collaboration entre l'équipe Orpailleur (INRIA NGE – LORIA) et l'unité INRA PSH (Avignon) a pour objectif de caractériser puis de générer les structures linéaires de haies rencontrées dans les paysages de vergers et de bocages. Deux paysages types ont été considérés, l'un en basse vallée du Rhône (données fournies par l'INRA Avignon), l'autre en Bretagne (données fournies par la ZA Armorique). Dans un premier temps, les relations entre haies et autres linéaires (routes, canaux) sont caractérisées à l'aide de descripteurs simples, en lien avec leurs fonctionnalités écologiques (longueur, angle, distance au voisin ...). La figure 8 (à droite) montre par exemple la sur-représentation de haies parallèles entre elles à courte et moyenne distance, sur une zone typique des alignements de haies brise-vents présents dans les vergers de la basse vallée du Rhône (à gauche).

Dans un deuxième temps nous tenterons de reproduire la distribution des haies, en combinant différentes approches, tirées de travaux en intelligence artficielle et en statistiques (modèles de Markov, chemins de Hilbert Peano, processus ponctuels marqués) [2].

Références

[1] M.S. Castellazzi, J. Matthews, F. Angevin, C. Sausse, G.A. Wood, P.J. Burgess, I. Brown, K.F. Conrad, and J.N. Perry. Simulation scenarios of spatio-temporal arrangement of crops at the landscape scale. *Environmental Modelling & Software*, 25(12):1881 – 1889, 2010.

[2] S. Da Silva, C. Lavigne, and F. Le Ber. Analyse de la structure des haies dans les vergers pour la définition de paysages mieux adaptés contre les bioagresseurs. In *Actes du colloque SAGEO*, pages 1–4, 2011.

[3] T. Guyet. Fouille de données spatiales pour la caractérisation de paysages en lien avec des fonctionnalités agro-écologiques. In *Actes du colloque SAGEO*, 2010.

[4] F. Le Ber, C. Lavigne, K. Adamczyk, F. Angevin, N. Colbach, J.-F. Mari, and H. Monod. Neutral modelling of agricultural landscapes by tessellation methods – application for gene flow simulation. *Ecological Modelling*, 220:3536–3545, 2009.

[5] F. Le Ber, C. Lavigne, K. Adamczyk, J.-F. Mari, F. Angevin, H. Monod, and N. Colbach. Neutral modelling of agricultural landscapes by tessellation methods: the

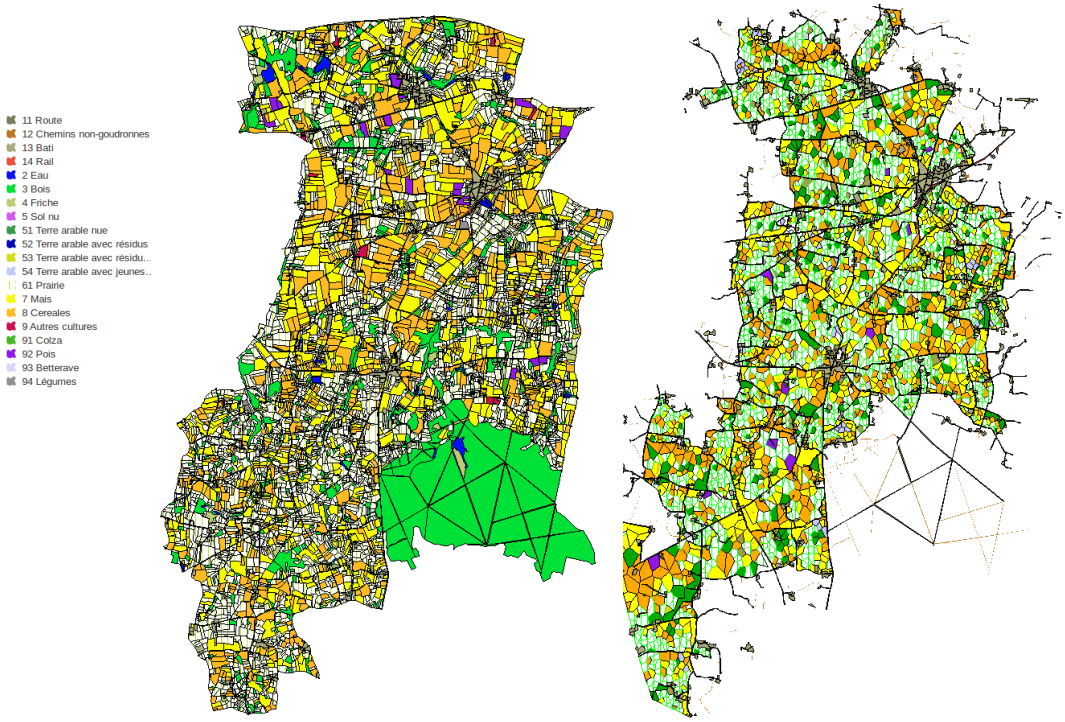


FIGURE 7 – À gauche : parcellaire réel, à droite : mosaïque reconstituée par pavage aléatoire et affection des occupations du sol à partir de sous-graphes fréquents

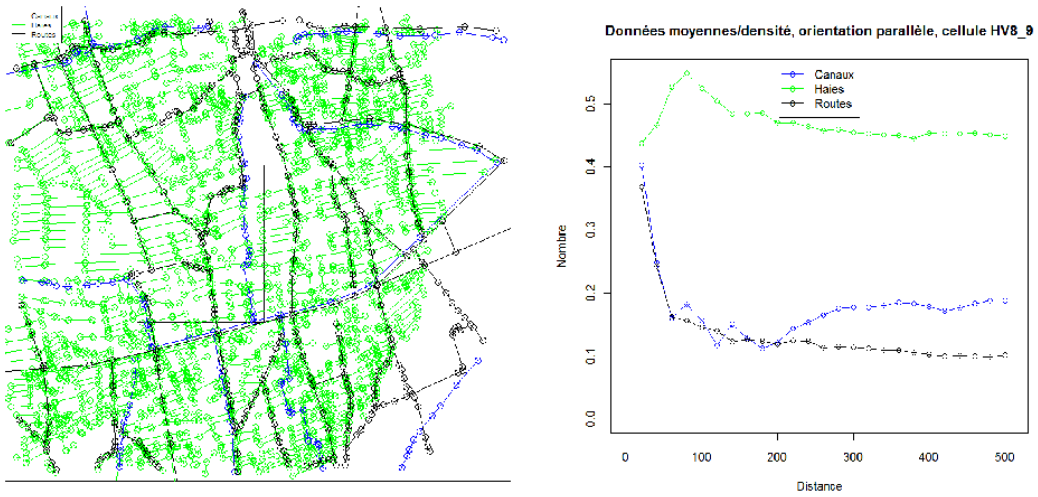


FIGURE 8 – Localisation (à gauche) et caractérisation des voisinages (à droite) des haies brise-vent dans la basse vallée du Rhone.

GENEXP-LANDSITES software – Application to the simulation of gene flow. In *International Conference on Integrative Landscape Modelling – LandMod*, pages 1–9, 2010.

Prix AFIA du meilleur papier aux JFSMA'12

Casanova : un comportement d'agent pour l'équité des mariages préservant la privacité

Patricia Everaere (Laboratoire d'Informatique Fondamentale de Lille, Université Lille 1)
Maxime Morge (Laboratoire d'Informatique Fondamentale de Lille, Université Lille 1)
Gauthier Picard (ISCOD / Institut Henri Fayol, École Nationale Supérieure des Mines de Saint-Étienne)

Introduction

Dans cet article, nous proposons un cadre multi-agent pour résoudre de manière distribuée des problèmes d'appariements stables. Dans notre approche, chaque individu est représenté par un agent négociant selon ses préférences. L'objectif d'une telle procédure est de trouver une affectation qui est *optimale*. À cette fin, nous pouvons considérer différentes notions de *bien-être social*. Dans ce travail, nous proposons *Casanova*, un algorithme distribué pour résoudre des *problèmes de mariages stables* (ou SMP). Cette stratégie de négociation présente des propriétés intéressantes : *(i)* elle assure la *privacité* puisque les agents ne révèlent pas leur liste de préférences et *(ii)* elle améliore l'optimalité de la solution et l'*équité* des partenaires (entre hommes et femmes, par exemple).

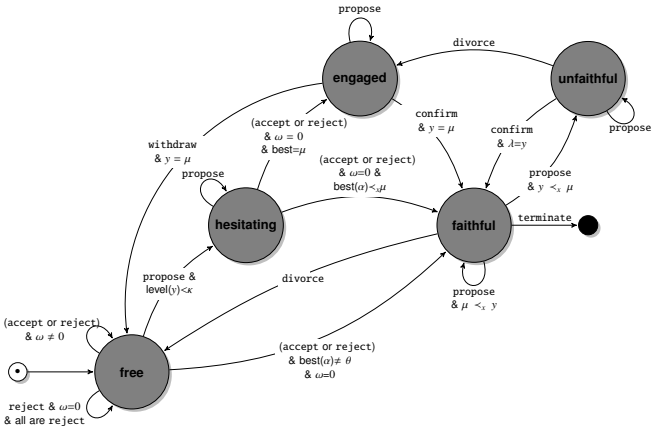
Algorithme Gale-Shapley

Gale et Shapley ont décrit dans [4] un algorithme centralisé (GS) qui trouve à coup sûr un appariement stable pour toute instance de SMP. Ils ont également remarqué que l'algorithme produit une affectation qui est optimale pour les hommes.

Une version distribuée et étendue de GS (DisEGS) est proposée par [2]. Dans cette version, contrairement au GS classique, chaque agent encapsule ses propres préférences, ce qui représente un pas intéressant vers plus de privacité des données, mais pas pour l'équité, car les hommes proposent et les femmes disposent.

Algorithme Casanova

Dans cette étude, nous considérons les appariements comme un phénomène émergent résultant de négociations locales. L'algorithme Casanova est une stratégie de négociation pour résoudre un SMP. Contrairement à la version distribuée "directe" DisEGS, nous ne distinguons pas hommes et femmes : ils envoient tous des propositions et répondent par des acceptations ou des rejets, ce qui représente la principale difficulté de ce travail. Nous cherchons à fournir des comportements d'agents menant des négociations jusqu'à des appariements socialement optimaux. Casanova consiste notamment à appliquer la stratégie de concession minimale. Selon cette stratégie, les agents débutent la négociation avec leurs partenaires préférés. Lors de la négociation, un agent concède minimalement dès que son partenaire optimal refuse une proposition.



Comportement des agents représenté comme un automate fini déterministe

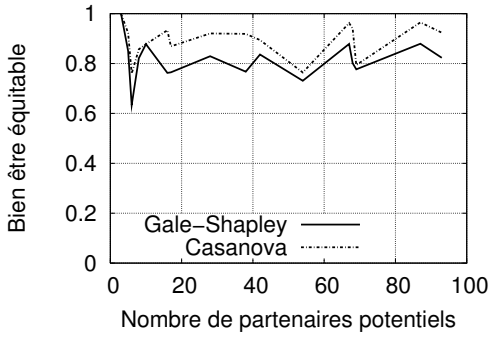
Évaluation

Casanova a été implémenté avec Jason [1], qui est un interpréteur d'une version étendue du langage AgentSpeak. Il met en oeuvre la sémantique opérationnelle du langage et fournit une plateforme de développement multi-agent. Jason est distribué sous license GNU LGPL.

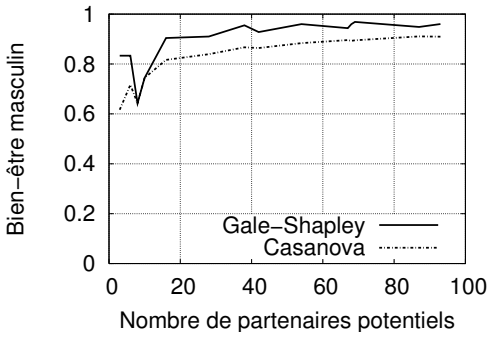
Conclusion

Casanova présente des propriétés intéressantes, alors que les agents n'agissent pas différemment en fonction de leur communauté (sexe). Comme présenté dans les résultats expérimentaux, Casanova trouve des appariements qui *améliore l'optimalité et l'équité entre communautés*.

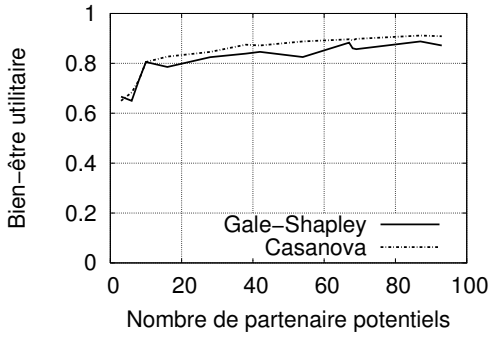
Le problème du mariage stable est un problème abstrait qui peut être instancié par une large gamme d'applications pratiques comme l'affectation des projets à des étudiants. On peut penser également au co-voiturage qui est un problème d'appariement entre deux communautés : des conducteurs et des passagers. D'une manière plus générale,



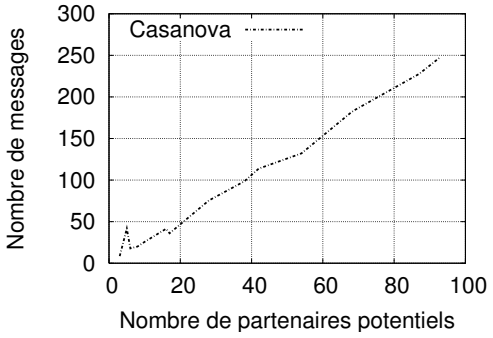
Bien-être inéquitable en fonction de la taille des communautés



Valeur normalisée du bien-être utilitaire en fonction de la taille des communautés



Valeur normalisée du bien-être des hommes en fonction de la taille des communautés



Nombre de messages en fonction de la taille des communautés

le problème SM est un cas particulier du problème d'appariement stable (en anglais, *Stable Matching Problem*). Parmi les extensions possibles, on compte : (i) le problème de colocation (en anglais, *roomate problem*) qui consiste à partitionner un ensemble fini d'individus en couples ou singletons ; (ii) la formation de coalitions où les individus sont partitionnés en fonction de leurs préférences sur les coalitions ; (iii) les marché sociaux où des liens peuvent être établis et les individus ont des préférences sur ces liens (voire sur la structure entière du réseau). Afin de résoudre ce genre de problèmes, l'approche centrée individu permet d'envisager des algorithmes adaptatifs, *anytime* et préservant la privacité des préférences.

Références

[1] Rafael H. Bordini, Jomi Fred Hubner, and Michael Wooldridge. Programming Multi-Agent Systems in AgentSpeak Using Jason. John Wiley & Sons Ltd, 2007.

[2] I. Brito and P. Meseguer. Distributed Stable Marriage Problem. *6th Workshop on Distributed Constraint Reasoning at IJCAI 05*, pp. 135-147, 2005.

[3] I. Brito and P. Meseguer. Distributed Stable Matching Problems with Ties and Incomplete Lists. *Proc. of CP-06*, pp. 675-680, Springer Verlag, LNCS 6057, 2006.

[4] D. Gale and L. S. Shapley. College Admissions and the Stability of Marriage. *American Mathematical Monthly*, vol. 69, pp. 9-14, 1962.

[5] Maxime Morge and Gauthier Picard. Privacy-Preserving Strategy for Negotiating Stable, Equitable and Optimal Matchings. *Proc. of PAAMS*, pp. 97-102, 2011.

Adhésion individuelle et abonnement

❏ **Demande** ❏ **Renouvellement**

Nom : Prénom :
Affiliation :
Adresse postale :
N° de téléphone : N° de télécopie :
Adresse électronique :
Activité (à titre professionnel / à titre privé (*rayer la mention inutile*)) :

Type d'adhésion	Consultation du bulletin sur WEB (pour une personne)
❏ Simple :	30 €
❏ Étudiant (sur justificatif) :	15 €
❏ Soutient :	Sans objet

❏ Adhésion au collège SMA : ajouter 7,5 € pour les étudiants, 15 € pour les autres
❏ Adhésion au collège IC : ajouter 7,5 € pour les étudiants, 15 € pour les autres
❏ Adhésion au collège *FERA (Apprentissage)* : ajouter 7,5 € pour les étudiants, 15 € pour les autres

Adhésion personne morale

❏ **Demande** ❏ **Renouvellement**

Organisme :
Adresse postale commune aux bénéficiaires couverts par cette adhésion :
Nom et prénom du représentant : Fonction :
Mél : Tél : Fax :
Adresse postale :

Le tarif d'adhésion comprend une partie fixe et une partie par bénéficiaire.
Coordonnées des bénéficiaires (10 maximum) :

Nom, prénom	Mél.	Tél.	Fax

	Tarif de base fixe :	Tarif par bénéficiaire :
❏ Laboratoire universitaires/PME	150 €	Gratuit pour 5 personnes (30 € par bénéficiaire supplémentaire)
❏ Personnes morales non universitaires	450 €	Gratuit
❏ Adhésion de soutien	600 €	Sans objet

❏ j'accepte que les renseignements ci-dessus apparaissent dans l'annuaire de l'AFIA
❏ j'accepte que les renseignements ci-dessus soient transmis à l'ECCAI pour constituer un ficher européen

Veuillez trouver un règlement (à l'ordre de l'AFIA) de ..... Euros

Trésorier AFIA : Thomas GUYET, Lab. Inf. d'Agrocampus-Ouest, 65, rue de Saint-Brieuc, 35042 Rennes cedex.
Mode d'adhésion :
De préférence, en ligne via le site Internet de l'AFIA : `http://www.afia.asso.fr`
A défaut, cette page doit être envoyée au trésorier.
Modes de paiement :

1. par chèque, à l'ordre de l'AFIA, envoyé au trésorier ;
2. par bon de commande administratif, à l'ordre de l'AFIA, envoyé au trésorier ;
3. par virement bancaire sur le compte de l'AFIA : Société Générale, 1 place du Maréchal Foch, 35000 Rennes, France. Code banque 30003, code guichet 01902, numéro de compte 00037283856 clef RIB 39.

Sommaire du Bulletin n° 75

Editorial 3

I.A. au Québec 4

Dossier IA et Agronomie 5

UMR ITAP (Irstea-SupAgro Montpellier), équipe « Modélisation et Décision Agro-environnementale » 6
UR INRA BIA, équipe MAD « Modélisation des Agro-écosystèmes et Décision » 10
UR INRA BIA, équipe RECORD 13
UMR IRISA (CNRS, UR1, ENS Bretagne, INSA), équipe DREAM « Diagnostic, Recommandation d'Action, Modélisation » 15
UPR Cirad « Recyclage et Risque » 19
TR Irstea Motive, équipe COPAIN, « Systèmes d'information communicants et agri-environnementaux » 22
EA CReSTIC, équipe SIC-MISTI « Management de l'Information SpatioTemporelle Imparfaite » 26
UMR Mét@risk (INRA - AgroParisTech), équipe « Ingénierie des Connaissances » 29
UMR IRHS (INRA-Agrocampus-Ouest), équipe FruitQual 33
UMR TETIS (AgroParisTech - Cirad - Irstea), équipe SISO « Système d'Information Spatialisée » 36
EA UNC-PPME, équipe « Informatique » 39
UMR MIA (AgroParisTech - INRA), Groupe AAA « Apprentissage Artificiel AgroParisTech » 40
Projet FOSTER 42
Projet FRESQUEAU 44
Projet PAYOTE 2 46

Prix AFIA du meilleur papier aux JFSMA'12 49

Calendrier de parution du bulletin de l'AFIA

Hiver	*Été*
Réception des contributions:	Réception des contributions:
15 décembre	**15 juin**
Sortie le **31 janvier**	Sortie le **31 juillet**
Printemps	*Automne*
Réception des contributions:	Réception des contributions:
15 mars	**15 septembre**
Sortie le **30 avril**	Sortie le **31 octobre**