



HAL
open science

Modélisation et optimisation de chaînes d’approvisionnement en biomasses pour des bioraffineries

Birome Holo Ba

► **To cite this version:**

Birome Holo Ba. Modélisation et optimisation de chaînes d’approvisionnement en biomasses pour des bioraffineries. Recherche opérationnelle [math.OC]. Université de Technologie de Troyes, 2016. Français. NNT : 2016TROY0001 . tel-03361304

HAL Id: tel-03361304

<https://theses.hal.science/tel-03361304>

Submitted on 1 Oct 2021

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

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

Thèse
de doctorat
de l'UTT

Birome Holo BA

**Modélisation et optimisation
de chaînes d'approvisionnement
en biomasses pour des bioraffineries**

Spécialité :
Optimisation et Sécurité des systèmes

2016TROY0001

Année 2016

THESE

pour l'obtention du grade de

**DOCTEUR de l'UNIVERSITE
DE TECHNOLOGIE DE TROYES
Spécialité : OPTIMISATION ET SURETE DES SYSTEMES**

présentée et soutenue par

Birome Holo BA

le 20 janvier 2016

**Modélisation et optimisation de chaines d'approvisionnement
en biomasses pour des bioraffineries**

JURY

M. O. PÉTON	MAITRE ASSISTANT - HDR	Président
M. J.-C. BILLAUT	PROFESSEUR DES UNIVERSITES	Examineur
M. M. NOUR EL FATH	PROFESSEUR	Rapporteur
M. C. PRINS	PROFESSEUR DES UNIVERSITES	Directeur de thèse
Mme C. PRODHON	MAITRE DE CONFERENCES - HDR	Directrice de thèse
M. A. YASSINE	PROFESSEUR DES UNIVERSITES	Rapporteur

DEDICACES

A la sagesse éternelle

A toute ma famille au sens large

A tous mes amis

Remerciements

Je tiens à remercier à vive voix l'ensemble des personnes qui m'ont accompagné tout au long de mes études et pendant cette thèse. Cette dernière, une expérience inédite, est le résultat d'un long et fastidieux travail, de rencontres fortuites, d'abnégation et d'opportunités. Je remercie particulièrement :

- Mes directeurs de thèse: Christian Prins et Caroline Prodhon. Je vous remercie très sincèrement d'une part de m'avoir fait confiance et donné la chance de faire cette thèse sous votre direction et d'autre part pour tout l'enseignement que vous m'avez transmis durant ces 3 années avec un soutien indéfectible. Merci de votre patience dans les moments difficiles et de m'avoir fait progresser.
- Les membres du jury: Jean-Charles Billaut, Olivier Péton, Adnan Yassine, Mustapha Nourelfath. Je tiens à vous exprimer ma reconnaissance pour avoir pris le temps de juger ce travail et également pour vos commentaires si intéressants et détaillés. Une dédicace particulière à Jean-Charles (mon ancien professeur et tuteur à Tours) pour sa gentillesse, sa disponibilité et son soutien depuis le début.
- Tous mes amis doctorants, post-doctorants et permanents du LOSI et de L'UTT pour votre bonne humeur: Tito, Mustapha, Jorge, Syrine, Sona, Elyn, Oussama, Julien, Yassine, Oumar, Guillermo, Mohamed, Moise, Nassim, Hajar, Yang Li, Xixi Wang, Amadeo, Matthieu, Puca, Komla, Daniela, Joseph, Intissar, Anis, Nassara, Qingxian An, Mourad, D.C.Nguyen, Mohsen, Na Li, Z. Chang, J.Michelet, T.Vidal, Atefeh.
- L'équipe administrative et technique de l'UTT : Véro, Marie-Jo, Bernadette, Patricia, Mme Denis, Mme Leclercq, Mme Kazarian et le Cri particulièrement Cyril et Jean-Philippe Balhadere.
- Monsieur le Ministre de l'enseignement supérieur du Sénégal, Mary Teuw Niane et Monsieur Moise Sarr responsable de la SGEE pour le soutien et les encouragements.
- L'ensemble des étudiants sénégalais et africains que j'ai eu à rencontrer à Troyes. Pas assez de place pour vous citer tous, mais je tiens à vous exprimer toute ma gratitude.
- Mes amis de Lyndiane mon quartier de naissance, du CEM Kabatoki, du MEELJ, du lycée Valdiodio Ndiaye (particulièrement la promo S1 2007), de l'AEEMS, de l'UGB et de Tours.
- Mes anciens camarades à Polytech (Mansour, Imane (ma sœur), Zakaria, Luqman, Wassim, Anas, Jeng, Thibault, Ngor, Kathar, Dieynaba Sall, Asta, Aissatou Niane, Martine, Adama, Nelly et Abdoulaye et bien d'autres.
- Mes nombreux amis personnels, en particulier, mais sans se limiter à Maimouna Niass, Couly Diogoul, Oumar Ndiaye, Mamadou Diallo, Abdoulaye Ndiath, Coumba Kitane, Bamba, Penda, Aissatou, Maodo, Moustapha C.F, Fatou et Cheikh Ibrahim Diop, Karim Diouf, Ousseynou Ndoye/Ndiaye, Makhtar Dia, Awa et Aissatou Lo, Chérif Diouf, Mariama, Roger, Anna, Ousseynou/Assane Badji, Ama et Diez Ndiaye, Mame Astou, Luciao, Charlotte, Fatima, Cindy, Adama Diagne, idrissa camara, Abdoulaye Sarr, Fatma/Khadija Ndiaye, Mustapha Lo, Jean-Gabriel, la famille Haris, la famille Ndoye de Bargny, la famille Faye de Rufisque.
- Mes anciens professeurs qui ont participé à ma formation, je pense particulièrement à Mr Fall, Mr Konaté, Mr Faye, Mr Sidibé, Mr Diouf, Mme Salamata Ba, Mr Saliou Sis, Mr Sagna, Mr Thiaré, Mr Sy et bien d'autres encore.
- Mes soutiens, amis et compagnons de route : Diouldé, Adjil Mané, Aissatou Béye, Sophie, Mada, Rokhaya, Satou, Coumba, Astou, Mbayang Thiam, Bouchra, Ousseynou Diaw, Omar wagué,

Assane Faye, Oulimata et toute la famille Barro, ma sœur Ndéye Adama. C'est d'une manière profondément sincère que je vous dis un grand merci.

- Une dédicace spéciale à toute ma famille : d'abord mes parents (Diariyatou, Dieynaba et Thierno), ensuite une mention très spéciale pour ma famille adoptive (Tonton Mactar Barro, Tante mame Coumba Faye, Tonton Abdoulaye Ngom et tonton Alpha Ba et sa femme Mamy Aw), sans qui, cette thèse n'aurait, peut-être, jamais pu avoir lieu. Merci infiniment du fond du cœur.
- Merci infiniment chère Khady Mbodj pour le soutien et la présence qui m'ont été très bénéfique dans les derniers moments de cette thèse. Un clin d'œil à Absa et Fatou.
- I.T.E PIVERT : Sans argent pas de thèse. Donc ma reconnaissance à PIVERT pour le financement. Merci à toute l'équipe de la SAS PIVERT qui m'a toujours très bien accueilli à Compiègne. Elle a toute ma complaisance. Je n'oublie pas M. Francis Valter du groupe AVRIL pour son soutien et ses conseils.
- Merci également à mon petit frère Paul Darius-Sarmadi pour son soutien et son amitié. Merci à Haroune pour son travail pendant le stage, sa sympathie et sa bonne humeur.
- Merci à Ndéye Adama (Diarama), Nelle et Tonton Adama qui se sont déplacés pour la soutenance.
- Ma future femme et mes futurs enfants 😊 dédicace spéciale à vous !
- Merci enfin à tous ceux dont j'ai pu oublier les noms et non la sympathie.

Birome Holo BA

Résumé

Les travaux de cette thèse concernent la modélisation et l'optimisation de chaînes d'approvisionnement en biomasses pour de futures bio-raffineries. En effet, des chaînes d'approvisionnement efficaces sont essentielles pour fournir aux installations de conversion, de façon régulière, des quantités suffisantes de biomasse de qualité à des prix raisonnables. Le problème est tout d'abord décrit puis modélisé.

Un modèle de réseau et un modèle de données sont ensuite développés pour permettre de décrire la structure de la chaîne d'approvisionnement et ses données, sans affecter le modèle mathématique sous-jacent. Ce dernier est un programme linéaire mixte qui combine pour la première fois divers aspects, soit originaux, soit gérés séparément dans la littérature. A partir des demandes de la raffinerie, une résolution exacte avec le solveur CPLEX précise les activités logistiques dans le réseau (quantités récoltées, mises en balles, transportées, stockées, etc.) et les équipements nécessaires, afin de minimiser le coût total composé des coûts de récoltes, de transport et de stockage. Des études de cas sont décrites pour illustrer ce modèle de planification tactique multi-biomasse et multi-période.

Un modèle plus compact est aussi élaboré pour traiter des instances de très grandes tailles. Il est illustré par une étude de cas réel pour une bio-raffinerie prévue près de Compiègne.

Pour finir, les développements effectués pour la mise en place d'un prototype d'application d'aide à la décision sont présentés et des recommandations en vue de la création d'un futur logiciel commercial sont proposées.

Mots-clefs : Recherche opérationnelle, Optimisation combinatoire, Logistique (organisation), Planification, Théorie des graphes, Programmation linéaire, Biomasse, Énergie de la biomasse, Installations énergétiques, Développement durable.

Abstract

The research works of this thesis address the problem of modeling and optimizing biomass supply chains for biorefineries. Indeed, efficient supply chains are essential to provide conversion facilities with sufficient quantities of quality biomass at reasonable prices. The problem is described and modeled.

A network model and a data model are developed to allow describing the structure of the supply chain and its data, without affecting the underlying mathematical model. The latter is a mixed-integer linear programming that combines for the first time various aspects, either original or tackled separately in the literature. For given refinery needs, its exact resolution by CPLEX specifies the logistic activities in the network (amounts harvested, baled, transported, stored etc.) and the necessary equipment, in order to minimize a total cost including harvesting costs, transport costs and storage costs. Case studies are described to illustrate this multi-biomass and multi-period tactical planning model.

A more compact model is also elaborated to cope with large-scale instances. It is illustrated using a real case study for a bio-refinery planned near Compiègne, France.

Finally, the developments conducted for the implementation of a prototype of decision-support application are presented and recommendations for coming to commercial software are proposed.

Keywords: Operations research, Combinatorial optimization, Business logistics, Planning, Graph theory, Biomass, Biomass energy, Energy facilities, Sustainable development.

Table des matières

Remerciements	3
Résumé.....	5
Abstract	6
Introduction générale.....	13
1. Contexte et défis	13
2. Importance de la logistique.....	15
3. Niveaux de décision dans les chaînes logistiques	15
4. Concepts de durabilité dans les chaînes logistiques de biomasse	16
5. Objectifs et structure de la thèse	17
Présentation du problème	19
1. Introduction.....	19
2. Les chaînes logistiques de biomasse en général.....	20
3. Chaîne d'approvisionnement considérée.....	25
Etat de l'art sur la modélisation des chaînes logistiques de biomasse.....	33
1. Introduction et synthèses déjà publiées	33
2. Modélisation de la durabilité en logistique de biomasse	35
3. Niveaux de décision dans les chaînes logistiques	37
3.1 Niveau de décision stratégique	37
3.2 Niveau de décision tactique	38
3.3 Niveau de décision opérationnel.....	40
4. Modèles d'évaluation de performances	40
4.1 Modèles simples basés sur des calculs de coûts	41
4.2 Modèles à base de SIG	41
4.3 Modèles basés sur la simulation	43
5. Modèles mathématiques pour l'optimisation.....	45
5.1 Modèles déterministes.....	46
5.1.1 Classification selon les caractéristiques des modèles et les méthodes de résolution..	46
5.1.2 Classification selon le nombre de biomasses et de périodes de l'horizon.....	49
5.2 Modèles stochastiques.....	53
5.2.1 Travaux en logistique industrielle a priori pertinents pour la biomasse.....	53
5.2.2 Travaux en logistique de biomasse	54
5.3 Récapitulation sur les modèles d'optimisation	55

6.	Conclusion critique.....	55
	Références :.....	57
	Modélisation des activités logistiques et typologie des données	59
1.	Introduction.....	59
2.	Modèle de réseau (network model).....	60
2.1	Principe des graphes états-tâches.....	60
2.2	Extensions des graphes états-tâches pour la biomasse	61
2.3	Application à notre projet et exemples.....	62
3.	Typologie des données (data model)	67
3.1	Règle de formation des noms et liste des abréviations	68
3.2	Décomposition en tables.....	68
3.3	Table des paramètres généraux (PARAMETERS).....	70
3.4	Table pour l'entité "sites" (LOCATIONS).....	70
3.5	Table pour l'entité "plantes" (CROPS).....	71
3.6	Table pour l'entité "parties de plantes" (CROP_PARTS)	71
3.7	Table pour l'entité "produits" (PRODUCTS)	72
3.8	Table pour l'entité "états" (STATES).....	72
3.9	Table pour l'entité "tâches" (TASKS).....	74
3.10	Table pour la relation "arcs état-tâche" ou "state-task arcs" (ST_ARCS).....	75
3.11	Table pour la relation "arcs tâche-état" ou "task-state arcs" (TS_ARCS).....	76
3.12	Table pour la relation "besoins de la raffinerie" (DEMANDS).....	76
4.	Système de gestion des ressources.....	78
4.1	Principes	78
4.2	Table pour l'entité "ressources" (RESOURCES)	79
4.3	Table pour l'entité "familles de ressources" (FAMILIES).....	80
4.4	Table pour la relation "limitations des ressources" (LOCA_FAM).....	80
4.5	Table pour l'entité "combinaisons de ressources" (COMBINATIONS)	80
4.6	Table pour l'entité "groupes de combinaisons de ressources" (GROUPS).....	81
4.7	Génération automatique de chaînes de récolte (CHAINS).....	82
4.8	Options de gestion des équipements.....	83
5.	Conclusion	83
	Modèle mathématique complet	85
1.	Introduction.....	85
2.	Bases de la génération automatique du modèle	86

2.1	Contraintes liées aux stocks selon Kondili et al. et désagrégation.....	86
2.2	Modéliser les attentes en stock	89
2.3	Modélisation de contraintes "no-wait" pour un produit	90
2.4	Modélisation des dégradations des produits stockés.....	90
2.5	Modélisation de stocks multi-produits.....	90
3.	Description du modèle mathématique complet	91
3.1	Comment définir des données à partir des tables ?	91
3.2	Liste des noms de données et de variables utilisés pour les calculs	92
3.3	Fonction-objectif	96
3.4	Contraintes de stocks et de flux	97
3.5	Contraintes de satisfaction des demandes de la raffinerie.....	99
3.6	Contraintes de gestion des ressources.....	100
3.7	Déclaration des variables	102
3.8	Déduction des consommations d'énergie et rejets de GES	102
4.	Deux études de cas.....	103
4.1	Etude de cas avec une ferme sur 28 jours.....	103
4.2	Etude de cas à 10 fermes sur 364 jours.....	109
5.	Simplification de la gestion de ressources et limites atteintes	112
5.1	Utilisation de coûts horaires "tout compris"	113
5.2	Elimination des variables entières.....	113
5.3	Modifications résultantes du modèle complet	114
5.4	Résultats avec les simplifications	115
5.5	Autres simplifications envisagées.....	115
6.	Conclusion	116
	Références :	117
	Modèle mathématique compact	119
1.	Introduction.....	119
2.	Développement du modèle mathématique compact.....	119
2.1	Principe de la simplification	120
2.2	Modifications dans la base de données	121
2.3	Modification du modèle complet.....	125
2.4	Test sur 10 fermes	129
3.	Etude de cas sur une grande instance.....	130
3.1	Difficultés liées à l'acquisition des données.....	130

3.2	Données sur les fermes de production de biomasse d'un cas réel.....	130
3.3	Données sur les stockages.....	131
3.4	Instance de test d'un cas réel pour le modèle compact	133
4.	Extensions possibles des modèles.....	134
5.	Conclusion	136
	Logiciel-prototype BIOSCO	137
1.	Introduction.....	137
2.	Outils logiciels utilisés.....	137
3.	Architecture fonctionnelle	138
4.	Module DISTANCES	138
5.	Module DATABASE	140
6.	Module OPTIMIZATION	141
7.	Module REPORTING	142
8.	Interface-utilisateur.....	142
9.	Recommandations.....	144
10.	Conclusion	146
	Conclusion générale.....	147
1.	Résumé des contributions et travaux effectués.....	147
2.	Perspectives.....	149
	Bibliographie	153

Introduction générale

1. Contexte et défis

Les deux dernières décennies ont vu un intérêt croissant pour les potentialités des biocarburants comme moyen de réduction de la dépendance aux combustibles fossiles et pour le développement d'une énergie écologique et renouvelable. Ces biocarburants sont élaborés à partir de la biomasse, qui comprend notamment toutes les plantes et matières d'origine végétale. Les sources de biomasse sont en fait plus nombreuses. Elles incluent une grande variété de ressources forestières et agricoles, comme les céréales et oléagineux déjà utilisés pour l'alimentation et l'énergie, mais aussi des sous-produits d'origine animale ou issus de processus industriels, des déchets ménagers etc.

Beaucoup de produits peuvent être obtenus en transformant des biomasses. L'exemple des biocarburants est l'un des plus connus. Les biocarburants issus de la biomasse sont généralement classés en trois générations que nous rappelons ci-dessous. Une présentation détaillée des deux premières générations et de leurs procédés de conversion peut être trouvée dans Naik et al. (2010).

Les plantes sucrières (canne à sucre, betterave, sorgho à sucre), les plantes à amidon (maïs, blé), les plantes oléagineuses (colza) et les graisses animales sont les produits les plus souvent utilisés pour produire des biocarburants de première génération : biodiesel, biogaz, bio-alcool et gaz de synthèse.

Les matières premières pour les biocarburants de première génération peuvent également être utilisées pour l'alimentation animale ou humaine et ont, par conséquent, suscité l'inquiétude du fait qu'elles pourraient conduire à une crise alimentaire mondiale (Rosegrant et al., 2006), d'où la nécessité de trouver une alternative viable et durable.

Les conflits potentiels entre usages énergétiques et alimentaires ont stimulé le développement des biocarburants de deuxième génération. L'éthanol cellulosique, qui est produit à partir de biomasse cellulosique, est un représentant de cette deuxième génération. La biomasse ligno-cellulosique a pour origine les parties non comestibles des cultures vivrières (tiges, feuilles, paille, enveloppes des graines), certaines plantes sans usages alimentaires (comme le panic érigé ou *switchgrass*, le jatropha, les céréales qui portent peu de grains) mais également le bois et des déchets industriels tels que les résidus de scieries (sciure, écorces), les peaux et pulpes provenant du pressage des fruits etc.

Les biocarburants à base d'algues sont étudiés depuis les années 80 (Sheehan et al., 1998) mais sont restés longtemps une curiosité. Ils sont maintenant envisagés pour constituer la troisième génération de biocarburants. Les algues peuvent être cultivées avec un haut rendement à l'aide d'eaux usées, et elles sont biodégradables et relativement inoffensives pour l'environnement en cas de déversement.

Outre la production de biocarburants, les biomasses sont utilisées pour beaucoup d'autres finalités, citons le chauffage urbain, la production d'électricité, la cogénération (chaleur, froid, électricité), la production de bioéthanol, de biomatériaux, de produits intermédiaires pour la chimie etc.

Dans les années 2010-2012, les Etats-Unis, l'Europe, ainsi que certains pays ont pris de grandes décisions de principe favorisant une bioéconomie industrielle. Par exemple, la Commission Européenne a présenté une proposition de directive visant à atteindre d'ici 2020 une part de 20% d'énergies renouvelables, et notamment une utilisation de biocarburants avec un objectif de 10% dans les transports (European Commission, 2008). La valorisation de la biomasse devra jouer un rôle capital pour l'atteinte de cet objectif.

Comme souligné par Schieb et al. (2014) dans leur livre "*Bioraffinerie 2030 – Une question d'avenir*", la bioéconomie industrielle mentionnée précédemment s'intéresse à la conversion de la biomasse en produits d'intérêts pour satisfaire les besoins humains en substituant des matières premières renouvelables à des matières premières d'origine fossile. En pratique, elle se réalise par l'installation de bioraffineries (aussi appelées unités de conversion) en lieu et place de raffineries classiques de type pétrolier. La bioraffinerie, en tant que socle de la bioéconomie industrielle, représente donc des enjeux économiques et environnementaux très importants.

Dans le but de mieux cerner et connaître leurs contours, leurs faiblesses et atouts, leurs conditions de réussite durable ainsi que leurs perspectives d'avenir, les bioraffineries suscitent actuellement de nombreux travaux de recherche sur tous leurs aspects. C'est dans ce contexte que cette thèse a été financée par l'ITE (Institut pour la Transition Energétique) PIVERT (Picardie Innovations Végétales, Enseignements et Recherches Technologiques).

L'ITE PIVERT (www.institut-pivert.com) est spécialisé dans la chimie du végétal, les technologies et l'économie des bioraffineries du futur et les biomasses oléagineuses et ligno-cellulosiques, afin de produire des produits chimiques, des biomatériaux et des biocarburants.

Dans le projet PIVERT, les raffineries sont dites de troisième génération. Ce terme qui concerne habituellement les biocarburants à base d'algues est ici utilisé pour souligner que les raffineries prévues se démarqueront des raffineries de seconde génération par quatre points importants : valorisation de la plante entière, produits finis non limités aux biocarburants, intégration dans les territoires et prise en compte dès la conception des aspects environnementaux à long terme.

Premier ITE labélisé dans le cadre des Investissements d'Avenir, l'ITE PIVERT s'appuie sur des établissements de recherche comme l'Université de Picardie Jules Vernes, l'Université des Sciences et Techniques de Lille (USTL), l'Université de Technologie de Compiègne (UTT), l'Université de Technologie de Troyes (UTT) et l'Institut National de Recherche Agronomique. Les membres incluent aussi des centres techniques (CETIM, CETIOM, IFPEN...), le pôle de compétitivité IAR (Industries et Agro-Ressources) et des industriels (Chimex, Clariant, Maguin, Novance, PCAS, Rhodia, Sofiprotéol (devenu récemment Groupe Avril), Tereos, Veolia, Solvay..). Le budget prévu de l'institut est de 245 millions d'euros sur 10 ans.

Les technologies mises au point par PIVERT seront les outils et les produits de la chimie du végétal permettant à la France d'atteindre ses objectifs, introduits dans le Grenelle de l'Environnement. PIVERT permettra d'optimiser des ressources agricoles et forestières et de produire de façon renouvelable et durable des produits de la vie quotidienne : produits alimentaires, biocarburants, cosmétiques, lubrifiants... en respectant les principes de la chimie verte.

2. Importance de la logistique

En 1960, le dictionnaire Larousse ne donnait qu'une interprétation de la logistique : le soutien aux opérations militaires. Les troupes se concentrent sur leur mission (combattre) tandis que la logistique garantit des approvisionnements en vivres et munitions, assure les transports et évacue les blessés. Dans les années 1960, la définition a été transposée au soutien aux systèmes de production : la logistique garantit des approvisionnements suffisants en matières premières et en composants, gère les stocks et s'occupe des expéditions, le tout pour que les produits commandés par des clients soient réalisés en temps, en heure et à moindre coût.

La logistique a ainsi longtemps été réduite à la gestion des flux de matières. Certains y ajoutent les flux d'information et les flux financiers. Quant à L'ASLOG (Association Française pour la Logistique), elle propose une définition bien plus vaste : "la logistique est l'ensemble des méthodes et techniques pour amener la bonne ressource, au bon endroit, au bon moment, en quantité adéquate et à moindre coût" (www.faq-logistique.com). On y a ajouté ensuite des critères de qualité et d'impact environnemental.

La chaîne logistique (*supply chain* en anglais) est un concept de base pour appréhender et modéliser les relations logistiques entre différents acteurs économiques (Simchi-Levi et al., 2003).

La conception et le pilotage de réseaux logistiques impliquent de nombreuses décisions complexes : où placer des entrepôts, comment gérer les stocks, quelles quantités livrer, combien de camions utiliser, comment mutualiser les moyens, etc.

Ces décisions sont évidemment très présentes dans le contexte d'unités de conversion de biomasse comme les bioraffineries, où l'utilisation de plusieurs biomasses est souvent nécessaire en raison des limites de l'offre, des variations saisonnières et des compositions chimiques différentes de chaque biomasse. La viabilité économique est toujours une question cruciale pour l'implantation, le développement et le fonctionnement d'une bioraffinerie.

Les recherches sur la biomasse se sont d'abord concentrées sur l'amélioration de la productivité des cultures et sur la conception des processus de conversion les plus efficaces pour les bioraffineries. En comparaison, les aspects logistiques ont été longtemps négligés. La logistique est pourtant critique pour approvisionner les bioraffineries de manière régulière, fiable, et avec des quantités suffisantes de biomasse de qualité. De plus, une part significative du coût de la biomasse à l'entrée d'une bioraffinerie réside dans les coûts logistiques. Une chaîne d'approvisionnement efficace est donc un moyen de contribuer à la viabilité économique, sociale et environnementale de l'industrie des bioproduits et biomatériaux issus des biomasses.

3. Niveaux de décision dans les chaînes logistiques

En gestion de production et en logistique industrielle, une approche pour analyser les problèmes est de classer les décisions en trois niveaux selon le degré d'importance et l'horizon temporel concerné : les décisions stratégiques, tactiques et opérationnelles (Mula et al., 2010).

Les *décisions stratégiques* sont des décisions à long terme, impliquant des investissements lourds, qu'on ne peut réviser qu'après plusieurs années, comme l'emplacement d'une nouvelle usine ou la conception d'un produit complexe (avion de ligne par exemple). En général, une décision engageant une entreprise sur au moins un an est considérée comme appartenant au niveau stratégique.

Les *décisions tactiques* sont les décisions à moyen terme, typiquement de quelques jours à quelques mois. Elles prennent en compte les besoins logistiques, la distribution, les flux dans le réseau et la planification des stocks et de la production. Un constructeur automobile va par exemple décider combien de voitures de chaque modèle seront fabriquées par semaine, en se basant sur un carnet de commandes ou des prévisions de vente, sous des contraintes de ressources agrégées (capacité de production, temps machine, main d'œuvre disponible, composants et matières premières).

Enfin, les *décisions opérationnelles* sont des décisions à court terme, typiquement sur une journée. Elles consistent à réaliser de manière détaillée les décisions tactiques. Citons pour la production les décisions d'ordonnancement (définition des heures de début des tâches, des attentes, des mobilisations et libérations de ressources) et, en logistique, les opérations de transbordement dans les plates-formes logistiques et l'élaboration de tournées de véhicules. Un bon critère pour savoir si une décision relève du niveau tactique ou opérationnel est de regarder si on considère l'ordre précis des tâches et leurs heures d'exécution : si oui, on est dans le niveau opérationnel.

Les limites entre niveaux sont assez élastiques selon les secteurs d'activité. Certains disent qu'une décision est stratégique si elle engage l'entreprise sur un an, tandis que pour d'autres ce sera trois ans ou plus. Ces différences ne sont pas critiques : le bon critère pour classer une décision au niveau N est qu'elle peut bloquer des décisions du niveau N+1 tant qu'elle n'est pas prise. Ainsi, on ne peut pas définir les flux mensuels de produits dans un réseau (niveau tactique) si on ne sait pas encore où placer les usines à l'origine de ces flux (niveau stratégique). Et on ne peut pas réaliser des tournées de véhicules (niveau opérationnel) si on n'a pas constitué une flotte de véhicules (niveau tactique).

En résumé, la *conception* d'une chaîne logistique relève essentiellement du niveau stratégique, tandis que son *management* est couvert par les niveaux tactique et opérationnel. Cependant, même les modèles utilisés pour la conception peuvent incorporer des décisions tactiques, pour obtenir des prévisions plus fines et plus réalistes. Comme nous le verrons, la répartition des décisions en trois niveaux est également applicable à la logistique pour la biomasse. Elle permet de fragmenter la difficulté dans l'analyse des systèmes complexes que constituent les chaînes d'approvisionnement en biomasses.

4. Concepts de durabilité dans les chaînes logistiques de biomasse

Le développement durable ou "soutenable" consiste à satisfaire les besoins actuels de l'humanité en combinant l'utilisation raisonnable des ressources naturelles, la préservation de l'environnement, la prospérité économique et la qualité de vie, sans compromettre la capacité des générations futures à répondre à leurs propres besoins. La durabilité doit être analysée en tenant compte des aspects environnementaux, économiques et sociétaux (Van Dam et al., 2005; Amigun et al., 2011; Duku et al., 2011). Elle constitue en particulier un enjeu important dans la logistique de la biomasse : par exemple, le bilan carbone d'une bioraffinerie peut être impacté par une logistique mobilisant trop d'équipements de récolte et de moyens de transport.

La durabilité environnementale vise à empêcher la dégradation de l'environnement. L'impact environnemental d'une activité humaine est caractérisée principalement par trois empreintes écologiques : sur l'atmosphère (gaz à effet de serre, GES), l'eau et les sols (Tan et al., 2012). Les GES sont des gaz atmosphériques qui favorisent l'effet de serre et donc le réchauffement climatique. La mise en œuvre stricte de bonnes pratiques, comme celles recommandées par l'Agence internationale de l'énergie (IEA, 2007), permet de limiter les dégradations induites par les industries dont, en ce qui nous concerne dans cette thèse, celles de la biomasse.

Les enjeux économiques des bioproduits en général, et des biocarburants en particulier incluent la compétition possible avec les productions alimentaires (que l'on peut voir aussi comme un problème sociétal), les considérations de coût, d'efficacité et de combinaison avec les autres énergies (*mix énergétique*) et la grande sensibilité aux politiques incitatives comme les budgets de recherche et une taxation privilégiée (Amigun et al., 2011; Karagiannidis et al., 2009). L'évaluation de la rentabilité nécessite le développement de modèles sophistiqués capables de quantifier et d'optimiser les différents coûts des étapes de la production de bioproduits et biomatériaux: culture, récolte, prétraitement, transport et conversion. Ces coûts se soustraient aux revenus générés, qui sont très sensibles aux aides gouvernementales et aux politiques de taxation des bioproduits. A long terme, la rentabilité dépend aussi des incertitudes sur les prix.

Le mix énergétique choisi par un pays est également un facteur important dans le développement des biocarburants. Outre les questions environnementales, ce mix peut dépendre de facteurs variés, comme les énergies propices au territoire national, l'importance du secteur agricole, la réduction de la dépendance aux importations et la robustesse aux facteurs climatiques.

Enfin, la production de bioproduits comme les biocarburants peut évidemment favoriser le développement économique, par exemple en attirant des investisseurs, en valorisant des sources de biomasse actuellement inutilisées, en augmentant et en diversifiant les revenus des agriculteurs, et en stimulant les industries associées comme celle du machinisme agricole (Demirbas, 2009).

Cependant, le développement de bioproduits soulève de nombreuses questions de société, tels que l'acceptabilité sociale, l'impact sur les populations rurales, l'utilisation des terres et le développement territorial, le potentiel de création d'emplois, et la réduction de la pauvreté. Ainsi, selon Todorov et Marinova (2009), les indicateurs énergétiques pertinents pour une économie en développement ne devraient pas se limiter aux questions environnementales et économiques: ils doivent intégrer des critères sociaux tels que la pauvreté. Quelques auteurs ont présenté des analyses de chaînes d'approvisionnement qui tiennent compte de ces concepts de durabilité, voir par exemple Forsberg (2000) et Cucek et al. (2012).

5. Objectifs et structure de la thèse

L'objectif principal de cette thèse est de modéliser et d'optimiser la logistique d'approvisionnement pour une bioraffinerie dite "de proximité" (rayon d'approvisionnement typique de 50 km), en mettant en place un modèle mathématique réaliste, général et flexible.

Le premier chapitre est consacré à la présentation du problème avec la définition de notions de bases intéressantes pour comprendre le sujet étudié et les contributions scientifiques réalisées durant les travaux de cette thèse. Après des généralités sur les chaînes logistiques, il décrit les différents schémas d'approvisionnement envisagés, le périmètre de la chaîne logistique étudiée, les principes des bioraffineries, les biomasses considérées, les chaînes de récoltes possibles, et enfin les objectifs de l'optimisation.

Le but du deuxième chapitre est de présenter un état de l'art sur la modélisation de chaînes logistiques pour la biomasse. Il présente d'abord les concepts communs à ces chaînes logistiques, puis les différents niveaux de décision avec un accent particulier sur le niveau tactique qui concerne la thèse. Cette distinction en niveaux constitue ainsi un bon guide pour l'analyse et la modélisation. Les différents modèles et méthodes de résolution publiés sont ensuite revus. Une conclusion critique résumant les apports et manques des travaux actuels complète ce chapitre.

Le troisième chapitre détaille un système de modélisation des activités d'une chaîne logistique, basé sur le formalisme des graphes états-tâches. Il liste les données nécessaires au modèle d'activités, les regroupe logiquement sous forme de tables de base de données et détaille les tables décrivant la gestion des ressources, qui est assez complexe dans notre cas. Ce système de modélisation permet à l'utilisateur de décrire la chaîne logistique qu'il souhaite et de préparer la couche de gestion (fichiers de données) des modèles mathématiques des deux chapitres suivants.

Le chapitre 4 décrit un premier modèle mathématique, un programme linéaire mixte, basé sur le modèle de réseau du chapitre 3. À partir d'un calendrier des besoins de la bioraffinerie, il permet de préciser les activités de la chaîne logistique sur un horizon-type d'un an, découpé en jours ou en semaines. Nous décrivons aussi quelques variantes, concernant notamment la façon de gérer les moyens de transport. Ce modèle très fin est testé sur un exemple simple, avec une exploitation agricole et un site de stockage centralisé, puis sur un cas plus complexe à dix exploitations. Il atteint cependant ses limites sur un cas de grande taille avec 100 zones de production. En effet, le nombre de variables et de contraintes explose et le solveur utilisé n'arrive plus à générer le programme linéaire numérique, même sur un PC équipé de 16 giga-octets de mémoire.

Le chapitre 5 propose donc un modèle plus compact pour les cas de grande taille, en agrégeant les activités effectuées dans les différentes exploitations agricoles. Un jeu d'essai en vraie grandeur est décrit puis résolu avec ce modèle. Il concerne une raffinerie prévue près de Compiègne, une centaine de sites de stockage et des exploitations agricoles regroupées en plus de 400 communes.

Un dernier chapitre décrit les principes d'un logiciel-prototype intégrant les modèles précédents. Ce prototype est basé sur des données stockées dans des classeurs Excel, des calculs en langage VBA et le solveur de programmation linéaire CPLEX. En particulier, certains aspects non triviaux comme le calcul du distancier sont détaillés.

Enfin, une conclusion générale clôture ce manuscrit en résumant les différents travaux réalisés ainsi que des perspectives et extensions pour des travaux futurs.

Chapitre 1

Présentation du problème

1. Introduction

Le sujet de la thèse concerne la modélisation et l'optimisation de la chaîne d'approvisionnement en biomasses pour une bioraffinerie de proximité. En effet, la logistique est critique pour approvisionner ce genre d'unité de conversion à moindre coût, de manière régulière et fiable, et avec des quantités suffisantes de biomasses de qualité. La chaîne logistique doit être bien comprise et minutieusement étudiée pour réduire les risques d'erreurs dans l'analyse et la structuration des données, la modélisation et l'optimisation.

Une chaîne d'approvisionnement en biomasse est un réseau complexe avec un nombre considérable de paramètres et de variables de décision. En effet, il faut par exemple déterminer les types de biomasse à mobiliser, la localisation, le nombre, la capacité et la taille des bioraffineries, des sites de stockage et des installations de prétraitement, le transport des matières, etc. Ce genre de chaîne couvre non seulement une grande étendue spatiale mais aussi un horizon temporel de plusieurs mois, pour prendre en compte les fluctuations saisonnières de l'offre et de la demande. A cela s'ajoutent de nombreuses contraintes comme les calendriers de récolte des espèces végétales choisies ou les dégradations dues à des processus biologiques.

La modélisation est très utile pour mettre en place ce genre de chaîne, en évaluant a priori plusieurs structures de chaînes pour quantifier les ressources mises en œuvre, les coûts associés, les consommations d'énergie et les impacts environnementaux. Au-delà d'une simple simulation d'une future réalité, elle permet de rationaliser et d'optimiser les choix de conception. Basée sur des techniques mathématiques et informatiques, la modélisation permet à bas coût, avec simplement des ordinateurs et de la matière grise, d'anticiper l'impact d'investissements considérables et donc de réduire les risques de décisions erronées. Elle soulève cependant de nombreuses difficultés théoriques et pratiques comme la pertinence des modèles et la nécessité de données fiables.

L'objectif de ce chapitre est de brosser un tableau des chaînes logistiques de biomasse en général puis de préciser le type de chaîne envisagée dans PIVERT et le problème à résoudre. Nous présentons d'abord la structure et les activités de la chaîne logistique avant d'indiquer le périmètre de l'étude et les hypothèses qui ont été posées pour éviter toute dispersion. Les types de biomasse à collecter et la notion de chaînes de récolte sont ensuite introduits, avant de donner la fonction-objectif à minimiser et les principales contraintes à respecter. Une conclusion complète le chapitre. Notre but est ici d'introduire pédagogiquement la problématique, sachant que l'analyse des nombreuses données et l'élaboration détaillée des modèles mathématiques font l'objet des chapitres 1 à 5.

2. Les chaînes logistique de biomasse en général

2.1 Vue d'ensemble des chaînes logistiques de biomasse

Un bon moyen d'analyser une chaîne logistique est de la voir comme un système composé de lieux (appelés *nœuds*) organisés en *couches* traversées par des *flux*. Dans le cas de la biomasse, en raison des caractéristiques spéciales des produits, plusieurs questions liées à chaque couche de la chaîne doivent être considérées et traitées pour obtenir un système d'approvisionnement économiquement viable. La figure 1.1 montre un exemple simplifié de chaîne logistique pour des raffineries produisant du bioéthanol à partir d'une plante herbacée sans valeur alimentaire, le panic érigé (*switchgrass*).

La couche d'entrée est formée des zones de production de biomasse. La biomasse doit être cultivée dans des régions à climat approprié. Les exploitations doivent bénéficier de sols acceptables et de ressources suffisantes en eau. Le rendement sur pied dépend de la plante choisie, du système de culture, du temps de croissance, de la période de récolte et de la météo. Quand la biomasse est prête, le rendement et la durée de la récolte dépendent du type et du nombre de machines employées. Ensuite, par rapport aux combustibles fossiles, la biomasse brute a une moindre densité et contient de l'eau, ce qui augmente les coûts de transport. De plus, le stockage est compliqué par la teneur en humidité et nécessite des installations *ad hoc* pour minimiser les dégradations. Par conséquent, les producteurs doivent planifier les opérations de récolte et choisir entre enlèvement immédiat ou stockage local de courte durée, pour assurer une bonne synchronisation entre la maturité des cultures, la météo, la disponibilité des équipements et des ressources humaines.

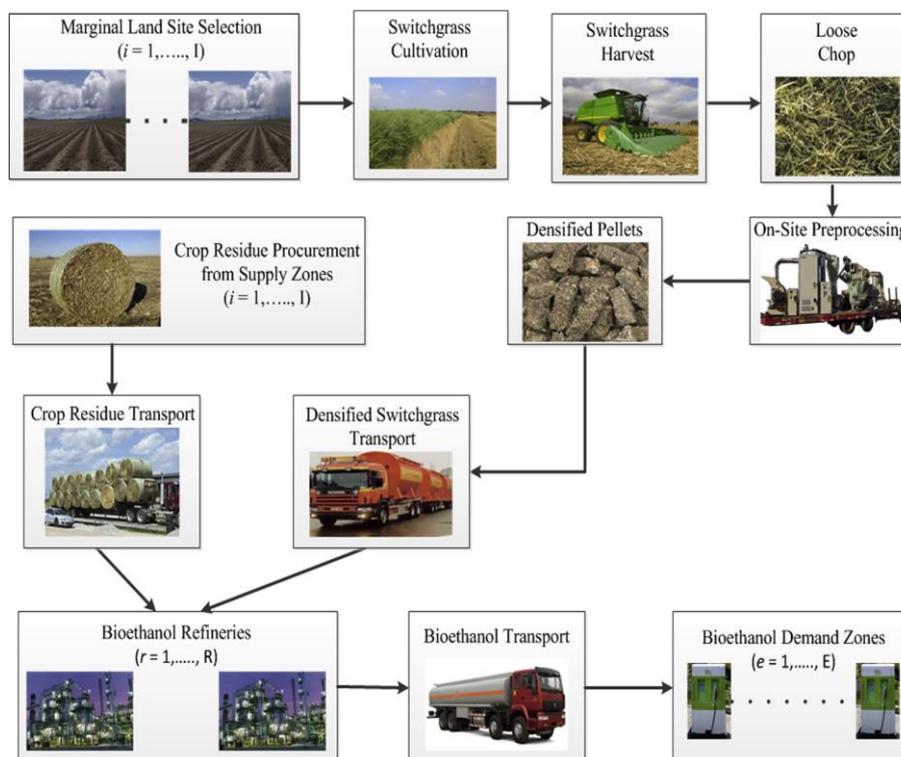


Figure 1.1. Exemple de chaîne pour du bioéthanol (Osmani et Zhang, 2013).

Après la récolte, des installations de stockage et de prétraitements de la biomasse sont souvent nécessaires et constituent une seconde couche de nœuds du système logistique. Les sites de stockage servent de tampons dans les périodes creuses séparant les récoltes de deux espèces végétales. Ils permettent donc d'assurer des approvisionnements continus aux bioraffineries, à condition d'être bien placés et dimensionnés. La biomasse stockée brute continue de se dégrader jusqu'à ce qu'elle soit prétraitée ou convertie. Les prétraitements ont pour but de réduire le taux d'humidité, le poids ou le volume des matières et d'augmenter la densité énergétique, avant d'être livrés à une bioraffinerie pour produire du biocarburant ou d'autres produits. Le niveau d'humidité qui est acceptable dépend de la méthode de conversion utilisée.

La troisième couche de nœuds comporte les bioraffineries. Basées sur différentes technologies de conversion comme la fermentation, la gazéification ou d'autres processus chimiques comme l'estérification, elles peuvent produire non seulement des biocarburants, mais aussi des matières premières pour la chimie et la pharmacie, des lubrifiants, des biomatériaux, etc.

Les deux premières couches forment la chaîne d'approvisionnement (partie amont ou *upstream*), qui conflue vers la couche centrale des bioraffineries (*midstream*). Le système se prolonge en pratique par la chaîne de distribution en aval (*downstream*) avec deux couches de nœuds : des dépôts régionaux, pouvant éventuellement faire des mélanges, et des nœuds de sortie comme des stations-services ou des clients industriels.

Après leur fabrication dans une bioraffinerie, la distribution des biocarburants ressemble donc beaucoup à celle des produits pétroliers. Il est évidemment très difficile de concevoir et de gérer globalement une chaîne logistique complète, avec ses trois parties *upstream*, *midstream* et *downstream*. Sans parler de la masse énorme de données qui serait nécessaire, les zones géographiques concernées, les acteurs économiques, les mécanismes de fixation des prix et les contrats sont très différents en amont et en aval.

2.2 Structure et activités d'une chaîne logistique de biomasse

Une façon de décrire une chaîne d'approvisionnement en biomasse est de recenser ses *processus* ou *activités*, comme la récolte ou collecte, la manutention, le stockage, le transport, le prétraitement et la livraison à l'usine. Il est d'usage d'exclure de la chaîne proprement dite les activités de production aux deux extrémités, c'est-à-dire les systèmes de culture et les processus de la bioraffinerie. En clair, la chaîne commence dans des champs prêts à récolter et s'arrête aux portes de l'usine. Compte tenu de ses caractéristiques, la chaîne d'approvisionnement en biomasse diffère significativement des chaînes observées en logistique industrielle, d'où l'importance d'y planifier les différentes activités dans une perspective globale de la chaîne, plutôt que de planifier chaque activité individuellement.

Une bioraffinerie nécessite un approvisionnement régulier, en biomasses de qualité constante et à coût raisonnable. Selon l'Agence Internationale de l'Energie (IEA, 2007), le transport, le stockage et la manutention de la biomasse nécessitent une évaluation attentive afin de minimiser les risques dans les projets de bioraffineries. Il existe de nombreuses sources de variabilité dans une chaîne d'approvisionnement de biomasse, telles que les incertitudes liées à la météo, la saisonnalité, les

caractéristiques physiques et chimiques comme la densité, la répartition géographique, la localisation des fournisseurs potentiels et leur volonté de cultiver de la biomasse, le transport et les infrastructures locales de distribution, les contrats passés avec les producteurs et les politiques gouvernementales (Fiedler et al., 2007; Iakovou et al., 2010; Gold et Seuring, 2011).

Une description complémentaire est donnée par un *modèle de type réseau*, appelé *graphe* en mathématiques. Un graphe est constitué d'un ensemble de *nœuds*, dans notre cas des sites géographiques, et un ensemble d'*arcs*, représentés graphiquement par des flèches connectant des nœuds.

Dans un graphe de chaîne d'approvisionnement en biomasse, un nœud peut représenter :

- un *nœud d'entrée* ou *nœud-source*, en l'occurrence un site où la biomasse est disponible,
- un *site de transit*, qui représente un carrefour ou une localité importante par où on passe,
- un *stockage*, où reposent les produits en attente de transfert pour une étape ultérieure,
- un *site de transbordement*, là où on change de véhicule ou de mode de transport,
- un *site de prétraitement* comme la granulation ou la densification
- un *nœud de sortie* où on utilise la biomasse, comme une raffinerie.

Les arcs permettant de passer d'un nœud à un autre indiquent des flux potentiels, par exemple :

- une étape de transport routier par camion,
- une étape de transport fluvial ou maritime,
- une étape de transport ferroviaire.

Un modèle de réseau met en évidence l'organisation spatiale de la chaîne et la dynamique des flux. On peut aussi y placer les activités. Ce genre de modèle peut être plus ou moins abstrait. La figure 1.1 déjà citée illustre une chaîne d'approvisionnement pour une bioraffinerie produisant de l'éthanol à partir de switchgrass. La représentation n'est pas encore rigoureusement formalisée sous forme de graphe mais on voit déjà qu'elle est "zoomable" : si nécessaire, un nœud comme une zone de production peut être décomposé en plusieurs nœuds, pour mettre en évidence les activités associées, les stocks temporaires, etc. Les arcs entre ces nœuds correspondent alors à des flux de courte portée, comme des déplacements élémentaires, des chargements ou déchargements etc.

La figure 1.2 montre un réseau mieux formalisé en termes de nœuds et arcs pour un transport de type routier. Il correspond vraiment à un graphe au sens mathématique et met en évidence les sites concernés par la chaîne logistique. Chaque arc représente un trajet routier entre deux sites, avec des distances en kilomètres. En fait, la plupart des modèles n'ont pas besoin des nombreux nœuds par où on se contente de passer (nœuds de transit) pour leurs calculs. Les nœuds conservés sont uniquement les sites importants pour la chaîne, avec ici une symbolique : les cercles correspondent aux nœuds d'entrée, les carrés aux sorties (lieux d'utilisation de la biomasse), les triangles pointe en haut à des sites de stockage et les triangles pointe en bas à des installations de prétraitement.

Chaque arc représente un plus court chemin dans le réseau routier réel, avec la distance correspondante. Ces chemins sont calculables à partir de cartographies routières comme celles de TeleAtlas (repris par TomTom) ou Navteq (repris par Here), par des logiciels comme MapPoint de Microsoft (abandonné fin 2014) ou Google Maps. Il faut noter que ce graphe est très agrégé : nous verrons dans la suite des représentations plus fines permettant notamment de préciser les activités et les ressources.

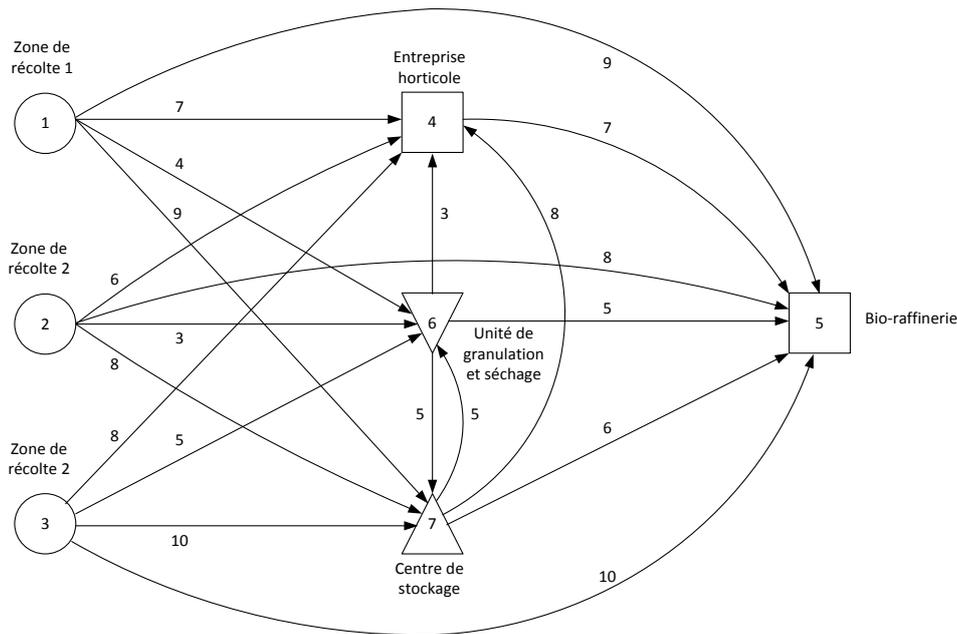


Figure 1.2. Réseau avec plus courts chemins et une symbolique pour les nœuds.

Nous pouvons superposer sur un modèle de type réseau différents types d'activités qui font l'objet des descriptions suivantes.

- **Récolte / collecte** de la biomasse dans des champs ou forêts

Prenons l'exemple des champs pour les cultures oléagineuses. Pour la production de biocarburants, les cultures oléagineuses ont un grand avantage puisqu'elles utilisent la machinerie agricole existante avec de légères modifications. Les méthodes de récolte les plus courantes sont les suivantes :

- La récolte en plusieurs passes (*multi-pass harvesting*) est une procédure en trois ou quatre étapes applicable par exemple au colza. Tout d'abord, une moissonneuse-batteuse récolte les grains et dispose la biomasse résiduelle sur une ligne (andain) derrière l'engin. Les résidus comprennent la paille, formée des tiges, et la menue paille, formée de débris plus petits (enveloppes des graines, poussières, feuilles brisées). Deuxièmement, une presse à balles (*baler*) tirée par un tracteur passe sur les andains pour faire des balles qui sont éjectées quand la capacité de la presse est atteinte. Troisièmement, un tracteur équipé d'une remorque collecte les balles et les transporte au bord d'une route ou vers une installation de stockage.
- La récolte en une seule passe (*single-pass harvesting*) combine par exemple une moissonneuse-batteuse équipée d'un compartiment récupérateur de paille, ce qui permet la récolte simultanée des grains et de la paille. Les produits obtenus sont déchargés en bout de champ et repris ensuite pour des conditionnements ou transports spécifiques.
- La méthode "tout récolté" (*whole-crop harvesting*) consiste à récolter la plante entière pour la mettre sur andain pour un séchage au champ ou la transporter immédiatement à la raffinerie. On l'utilise plutôt pour des plantes herbacées comme le miscanthus.

- **Stockage**

De nombreux types de biomasse ont une disponibilité saisonnière : ils sont récoltables pendant une période précise alors qu'ils peuvent être nécessaires à l'usine pendant presque toute l'année. Il est donc nécessaire de les stocker. Le point de stockage peut être situé en bordure de champ ou de forêt, dans un site de stockage avec une meilleure conservation ou à l'usine (Rentizelas et al., 2009b). On peut stocker du vrac, mais aussi des formes conditionnées (balles) ou prétraitées (granulés).

- **Chargement et déchargement**

Une fois que la biomasse a été déposée en bord de route, il faut la charger dans des véhicules pour le transport vers un centre de stockage, une unité de prétraitement ou un site d'utilisation (centrale électrique, chaufferie, bioraffinerie), où elle devra être déchargée.

- **Transbordement**

Le transbordement consiste à changer de véhicule, ou plus généralement à changer de mode de transport. Il implique un déchargement puis un rechargement dans le nouveau véhicule, avec souvent une attente intermédiaire qui peut être vue comme un stockage temporaire. La synchronisation entre transport routier et un mode à fréquence plus faible (train, bateau) peut même nécessiter des installations spécifiques pour le stockage (silos portuaires par exemple).

- **Transport**

Vu les lieux de récolte (champs et forêts) et les infrastructures de transport disponibles, la route est souvent le seul moyen de transport pour la biomasse (Rentizelas et al., 2009b). Parmi les autres facteurs en faveur du transport routier, on peut noter sa grande souplesse mais aussi les gains kilométriques liés aux trajets directs. Les autres modes comme le transport maritime ou ferroviaire impliquent des transbordements qui ne sont amortis que pour des distances suffisantes.

Outre le mode de transport, il faut choisir des véhicules qui dépendent du type de biomasse, de sa forme, de sa quantité, de la destination et de la distance à parcourir. Le transport du lieu de récolte vers un stockage local peut s'effectuer avec des tracteurs équipés de remorques ou des camions-bennes. Le transport vers la raffinerie peut utiliser des semi-remorques de plus grande capacité, des conteneurs, voire des camions à compartiments pour les granulés. Enfin, le choix pour un producteur de biomasse entre utiliser ses propres véhicules ou ceux d'un prestataire logistique n'est pas évident. Il est discuté par exemple par Huisman et al. (1997) et par Tatsiopoulos et Tolis (2003).

- **Prétraitements**

Le prétraitement de la biomasse sert à améliorer l'efficacité de sa manutention ainsi que la quantité pouvant être transportée ou stockée. Il est également nécessaire lorsque les caractéristiques de la biomasse ne correspondent pas aux exigences fixées par la raffinerie. Il peut se réaliser à n'importe quel stade de la chaîne d'approvisionnement, mais précède souvent les transports routiers et est généralement moins cher quand il est intégré avec la récolte (Allen, 1998).

Des exemples de prétraitements sont le conditionnement en balles (*baling*), la densification plus poussée pour obtenir des granulés ou des briquettes, le broyage, le séchage thermique et la torréfaction. Les deux derniers nécessitent des installations plus lourdes, souvent distantes, mais qui peuvent se justifier si la meilleure efficacité du prétraitement compense le surcoût pour y amener la biomasse brute. Certains auteurs comme Kim et al. (2011a, 2011b) considèrent des sites de pyrolyse qui produisent du bio-pétrole, envoyé ensuite à une bioraffinerie. Mais il vaut mieux parler dans ce cas de pré-conversion puisque la biomasse initiale n'est plus reconnaissable.

Des modèles sont nécessaires pour évaluer le coût et l'efficacité de tous ces processus de transport, de stockage, de prétraitement et de conversion. Comme ces activités consomment de l'énergie et produisent des émissions de gaz à effet de serre (GES), la conception de la chaîne logistique doit être optimisée, non seulement en termes économiques, mais aussi sur le plan environnemental.

Dans cette thèse, nous allons nous limiter à la partie amont ou *upstream*, allant des zones de production de biomasse aux portes d'entrées de l'unité de conversion. Cette limitation vient du fait que l'optimisation globale de chaînes complètes est extrêmement difficile : d'une part, des acteurs et entreprises très différents interviennent en amont et en aval et, d'autre part, l'acquisition de données complètes est quasiment impossible. Par contre, le modèle à développer va gérer finement les flux et les activités de la partie amont, sur un horizon temporel d'un an découpé en jours ou semaines (niveau tactique), tout en étant générique (adaptable à toute structure de chaîne).

3. Chaîne d'approvisionnement considérée

Dans cette section, le périmètre de l'étude et les hypothèses fortes sont identifiés.

3.1 Caractéristiques générales – Bornage du système

Afin de bien délimiter le système abordé, voici la liste des principales caractéristiques et des limitations qui ont été fixées :

- On considère une chaîne d'approvisionnement en biomasse pour une bioraffinerie (BR) oléagineuse de proximité (rayon 50 km) en Picardie. Venette près de Compiègne est pressentie pour installer cette raffinerie, mais le modèle doit être transposable à tout autre emplacement.
- Cette chaîne va faire l'objet d'un modèle mathématique optimisant les activités sur un horizon de planification d'un an, découpé en périodes élémentaires. La "granularité temporelle" (finesse du découpage de l'horizon) doit être ajustable, avec des périodes allant d'un jour à un mois.
- La chaîne va des champs aux stocks de matières premières en entrée de la BR. Les opérations de récolte sont incluses, mais pas le système de culture. Les procédés de conversion et les produits élaborés par la BR ne sont pas précisés et ne concernent pas le système logistique. Les zones de production de biomasse considérées sont au minimum des exploitations agricoles, car le niveau parcelle est inaccessible. La "granularité spatiale" doit aussi être réglable : pour réduire les données nécessaires, on peut considérer des regroupements, par exemple au niveau des communes, des cantons ou des petites régions naturelles, c'est-à-dire des zones homogènes sur le plan pédologique.

- Plusieurs biomasses sont acceptées, principalement oléagineuses (graines de colza, moutarde d'Ethiopie, cameline), mais aussi lignocellulosiques (paille et menue paille de colza, miscanthus).
- Une plante peut avoir plusieurs parties utilisables (par exemple, graines, paille et menue paille).
- Les transports entre sites se font exclusivement par la route : le rayon d'approvisionnement est trop petit pour amortir les transbordements si on utilise d'autres modes de transport.
- La BR fonctionne en "flux tiré" comme dans la plupart des chaînes logistiques industrielles : à partir de contrats ou de prévisions de ventes, elle exprime ses besoins en biomasses, en général semaine par semaine. Le modèle de calcul déterminera ensuite où prélever les biomasses, dans quelles périodes et en quelles quantités, pour satisfaire ces demandes. Il calculera aussi les ressources nécessaires (équipements de récolte et de transport). L'objectif sera de minimiser le coût total du système (coûts en capital et coûts d'utilisation des divers équipements, coûts de stockage, coûts de transport) ainsi que les consommations d'énergie et les rejets de GES.

3.2 La bioraffinerie

Bien que la thèse ne concerne pas les processus de la bioraffinerie, il est important de connaître un peu ses activités pour mieux préciser son impact sur l'approvisionnement en biomasse. Nous allons donner quelques définitions tirées de la littérature puis préciser la structure de bioraffinerie envisagée pour PIVERT. Les biomasses qui seront traitées sont décrites dans la section suivante.

Le Département Américain de l'Énergie (DOE) utilise la définition suivante (Kamm et al., 2006) :

"Une bioraffinerie est un concept générique définissant une usine de transformation où la biomasse est extraite et convertie en un ensemble de produits de valeur. Son fonctionnement est similaire à celui des raffineries pétrochimiques".

Le Laboratoire National Américain de l'Énergie Renouvelable (NREL) utilise pour sa part la définition plus précise suivante (www.nrel.gov/biomass/biorefinery.html) :

"Une bioraffinerie est une installation qui intègre le processus de conversion de la biomasse et le matériel de production des carburants, de l'énergie et des produits chimiques issus de cette biomasse. Le concept est analogue à celui des actuelles raffineries pétrolières qui produisent des carburants et de multiples produits à partir du pétrole. Les bioraffineries industrielles ont été identifiées comme la voie la plus prometteuse pour la création d'une nouvelle industrie nationale domestique".

Quant à l'Agence Internationale de l'Énergie (IEA, 2007), elle présente brièvement la bioraffinerie comme étant *"la transformation soutenable de la biomasse en un ensemble de produits commercialisables et en énergie"*.

Le livre récent de Schieb et al. (2014) introduit la durabilité : *"une bioraffinerie est un site industriel qui transforme la biomasse, de manière soutenable, en produits destinés à l'alimentation humaine et animale, en biomatériaux, en biocarburants, en produits chimiques à forte valeur ajoutée comme par exemple les bases cosmétiques. L'objectif étant de valoriser toutes les composantes des agro-ressources de manière alimentaire et/ou non alimentaire, ainsi, avec les mêmes intrants, les entreprises appartenant à une bioraffinerie auront des sortants beaucoup plus diversifiés"*.

La figure 1.3 ci-dessous, tirée du site de la SAS PIVERT (www.institut-pivert.com), donne une représentation simplifiée du fonctionnement d'une bioraffinerie oléagineuse. L'ensemble des activités logistiques va dépendre étroitement des besoins en biomasses de cette BR.

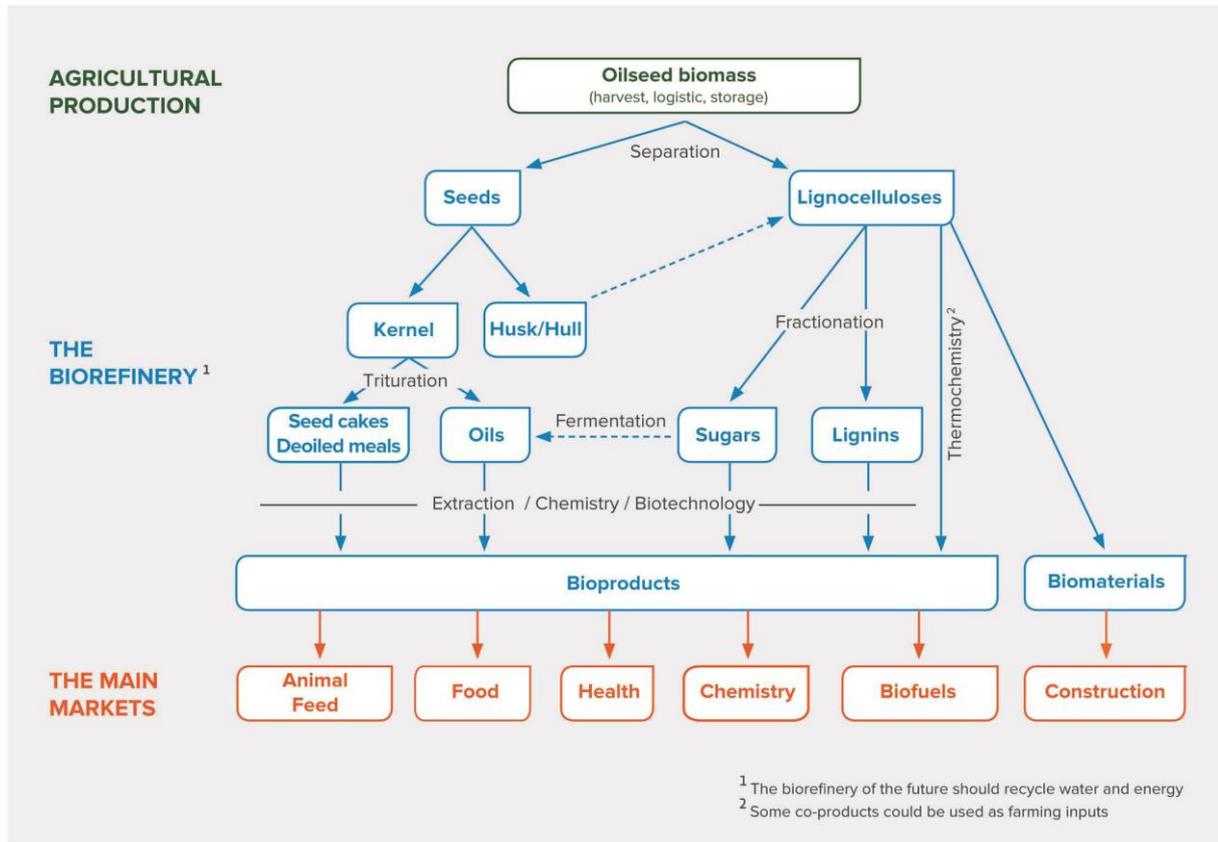


Figure 1.3. Bioraffinerie oléagineuse envisagée.

3.3 Matières premières concernées

Dans cette section, nous présentons les biomasses à livrer à la raffinerie, selon des spécifications établies par Francis Valter (Groupe Avril). Il faut noter que le modèle mathématique qui sera exposé plus loin pourra prendre en compte non seulement ces matières premières mais également d'autres biomasses agricoles qui ne sont pas décrites ici. Autrement dit, le modèle sera suffisamment flexible pour gérer d'autres biomasses, même celles non encore envisagées à l'heure actuelle.

La raffinerie de proximité à étudier est basée sur trois plantes oléagineuses, le colza, la moutarde d'Ethiopie et la cameline, complétées par une sorte de roseau, le miscanthus. Comme le colza, la moutarde d'Ethiopie et la cameline sont des crucifères, mais elles sont de plus petite taille, de récolte plus tardive, et ont des compositions chimiques différentes. La raffinerie nécessitera 80000 tonnes par an. Ces tonnes sont exprimées en matière sèche (MS), c'est-à-dire sans compter l'humidité qui est toujours présente dans les biomasses (même les produits apparemment secs et stables ont toujours une humidité résiduelle). Ce tonnage se décompose en 35000 t MS de graines et 45000 t MS de matière ligno-cellulosique.

a) Colza (*Brassica napus*) : graines, paille et menue paille

- **Graines** : cultures de colza 00 (dit double zéro) et colza érucique. Au final, une distinction de l'approvisionnement est faite entre ces 2 qualités de colza, de composition chimique très différente (colza 00 – *a priori* utilisation alimentaire, colza érucique – utilisation industrielle). Les zones de cultures sont distinctes pour éviter des contaminations croisées mais la période de récolte est identique. Les deux types de colza ne peuvent être conservés ni dans la même cellule de stockage, ni (si possible) sur le même site de stockage. Les graines sont livrées à l'usine en flux tendus sur la période de temps de la récolte et la fraction non directement utilisée est dirigée vers un stockage intermédiaire. Ces graines seront livrées prioritairement jusqu'au 20 octobre.

Graines de colza 00 – 18000 tonnes sèches par an :

Rendement en t/ha de graines : 3,4 t/ha/an à 9% d'humidité (sec : 3,1 t/ha/an); densité moyenne des graines brutes: 645 kg/m³; poids de 1000 graines sèches : 4,5 g; récolte et livraison au stockage ou à la BR, en vrac. Récoltes des graines du 15 au 30 juillet (maturité en Picardie).

Graines de colza érucique – 7000 tonnes sèches par an :

A part la composition chimique, les caractéristiques physiques des graines sont les mêmes que pour le colza 00.

- **Pailles de colzas 00 et érucique (matière lignocellulosique) – 25000 tonnes sèches par an :**

La paille est formée des tiges de colza. Rendement en paille lignocellulosique: 2,3 t/ha/an à 18% d'humidité (sec : 1,9 t/ha/an). Récolte de la paille: entre le 18 juillet et le 2 août (3 jours après les graines à cause du temps de séchage au sol).

- **Menue-paille de colzas 00 et érucique (matière lignocellulosique) – 9500 tonnes sèches par an :**

Rappelons que la menue paille inclut les enveloppes des graines et autres particules. Rendement en menue-paille : 1,25 t/ha/an à 15% d'humidité en cas de récupération immédiate (sec : 1,05 t/ha/an), 1 t/ha/an si déposée sur andain avec la paille. Récolte de la menue-paille du 15 au 30 juillet en même temps que les graines, à la moisson.

- **Pailles et menues-pailles mises en balles parallélépipédiques**

Alors que les graines sont transportées en vrac, la paille et la menue paille sont conditionnées sous forme de balles.

Les balles de paille de 180 kg/m³ (\approx 460 kg) sont stockées en bout de champ ou chargées directement sur un camion (max. 60 balles) par un Maniscopic (chariot élévateur) à fourche (figure 1.4). Coût du chargement-déchargement : 3 €/t. Le camion va vers un site de stockage (coût de stockage : 1 €/t de paille/mois) ou en flux tendu vers l'usine. Pour une balle de menue paille, compter 400 kg. On considère que le stockage de longue durée au champ (plus de 1,5 mois) n'est pas souhaitable en raison des risques (les incendies par exemple) et la "meule" est limitée à 2 000 t (limite appliquée par les assurances).



Figure 1.4. Manutention de paille en bottes.

b) Moutarde d'Ethiopie (*Brassica carinata*) – Graines seulement – 7000 t sèches par an :

Cette plante est cultivée actuellement en France mais à titre expérimental. Rendement en t/ha de graines : 2,2 t/ha/an à 9% d'humidité (2 t/ha/an en matière sèche), récoltes et livraison au stockage ou usine en vrac. Unité prévisionnelle de séchage en intermédiaire. Récolte des graines : 15 au 30 octobre (maturité en Picardie). Restitution au sol de la totalité des pailles, à vérifier si cela est compatible agronomiquement avec la rotation des cultures, en raison de la faible dégradation de la matière organique en hiver et du développement de maladies. Prévoir quand même la possibilité de récupérer cette paille dans l'avenir.

c) Cameline (*Camelina sativa*) – Graines seulement

La cameline servait à faire de l'huile d'éclairage en Picardie au 19^{ème} siècle mais sa culture a quasiment disparu. En attendant les résultats d'études culturales en cours dans PIVERT, on prendra les mêmes paramètres que la moutarde d'Ethiopie.

d) Miscanthus (*Miscanthus giganteus*) :

Rendement en t/ha de paille : 13,5 t/ha/an à 18% d'humidité (11 t/ha/an en matière sèche – référence Picardie); récoltes et livraison au stockage ou usine en vrac sous forme de brins (broyage au champ par ensileuse : 100-130 kg/m³) pour un approvisionnement en flux tendu, ou, alternativement, pressage et stockage en balles parallélépipédiques standards de 170 à 200 kg/m³. Pas de séchage nécessaire sous 20% d'humidité. Récolte des pailles : 10 au 25 mars (maturité en Picardie), extensible du 1^{er} mars au 30 avril.

Le transport de ces matières se fait par camion à partir des stocks en bord de champ, si cela se traduit par au moins 10 h de travail (direction raffinerie en flux tendu, ou vers site de prétraitement). Le même mode de transport est utilisé entre les installations de stockage intermédiaires et la raffinerie. Le transport par les équipements de l'agriculteur (tracteur + remorque) est plus lent et plus cher au km mais est caractérisé par un coût fixe plus faible. Il se fait a priori uniquement vers un site de stockage intermédiaire car l'agriculteur ne livre pas directement la raffinerie.

3.4 Les chaînes de récoltes

Une chaîne de récolte est l'ensemble des activités constituant une pratique particulière de récolte d'une plante cultivée. Elle spécifie les différentes étapes, leur enchaînement temporel, et les équipements associés. Par exemple, le colza peut être moissonné en récupérant les graines, la paille et la menue-paille séparément. Cette pratique caractérise une chaîne de récolte possible. Elle est différente d'une autre manière de moissonner consistant à récupérer les graines de colza d'un côté, et la paille et la menue-paille de l'autre côté. Cette seconde façon de faire définit donc une seconde chaîne de récolte.

Une chaîne de récolte peut être modélisée dans la chaîne logistique par un sous-graphe qui spécifie les activités, les flux et les différents types d'équipements nécessaires à la réalisation correcte de la moisson. Comme il y a très peu de chaînes de récolte possibles pour chaque plante, on retrouvera le même sous-graphe dans toutes les exploitations agricoles ayant choisi la même chaîne de récolte. La section 2 du chapitre 3, consacré à la modélisation des activités de la chaîne logistique, décrit les différentes chaînes de récolte pour les biomasses considérées dans cette étude.

Le modèle élaboré dans cette thèse permet de préciser des chaînes de récolte différentes pour une même plante, selon les exploitations. Cependant, comme il y a très peu de chaînes de récolte possibles pour chaque plante, on retrouvera les mêmes chaînes dans beaucoup d'exploitations agricoles. Pour chaque activité dans une chaîne de récolte, le modèle est aussi capable de choisir dans une base de données un équipement dans la catégorie nécessaire (moissonneuses-batteuses, tracteurs, camions...) et même une association de plusieurs équipements, comme un tracteur plus une remorque. Par exemple, si une activité nécessite un tracteur, on peut choisir un modèle commercial de 200 ou 250 chevaux, avec des coûts et des productivités différents.

3.5 Objectif d'optimisation et contraintes

Le but est de développer un modèle mathématique pour minimiser les coûts des différentes activités de la chaîne d'approvisionnement, depuis les champs (zone de production de la biomasse) jusqu'à l'entrée de la raffinerie, en passant par les unités de prétraitements et les sites de stockages. On déduira des solutions la consommation d'énergie et les émissions de gaz à effet de serre (GES). L'optimisation simultanée des trois objectifs sort du cadre de la thèse (optimisation multicritère).

Les coûts à considérer sont les suivants :

- coûts de récolte,
- coûts de conditionnement (pressage de balles par exemple),
- coûts de manutention,
- coûts de transport,
- coûts de stockage,
- coûts de prétraitement.

Cet objectif doit être optimisé en respectant des contraintes qui peuvent être résumées et regroupées dans les catégories suivantes :

- Respect des quantités récoltables selon les plantes.
- Satisfaction des besoins de la raffinerie, exprimés par exemple en tonnes sèches par semaine pour chaque biomasse consommée.
- Lois de Kirchhoff reliant les flots entrants et sortants pour chaque nœud de la chaîne logistique, certains processus comme le séchage pouvant occasionner des pertes.
- Contraintes liées aux stockages (capacités limitées, stock initial et final requis, dégradations pendant le stockage, débits limités pour l'entrée et la sortie de stock).
- Respect de fenêtres temporelles (fenêtres de récolte, stockages disponibles seulement à certaines période de l'année).
- Respect de certaines attentes (temps de séchage minimal de la paille sur le sol, processus de prétraitement s'étalant sur plusieurs périodes).
- Respect des nombres d'équipements disponibles (récolte, transport, manutention) ainsi que de leurs capacités (temps de travail maximal par jour, débit en tonnes/heures ou hectares/heures pour les équipements de récolte, volume et tonnage limités pour les moyens de transport).

4. Conclusion

La description de la chaîne logistique, un peu fastidieuse, montre déjà une grande complexité et une multitude de détails liés aux plantes, aux chaînes de récolte, aux transports etc. De plus, l'ensemble des plantes à traiter pourra évoluer, ainsi que les chaînes de récolte possibles, les rendements, etc. Le futur modèle ne doit donc pas être prisonnier de la situation actuelle ou d'une structure figée de chaîne logistique, ce qui constitue un défi difficile.

Après l'état de l'art du chapitre 2 qui montrera qu'aucun travail publié n'a traité des chaînes aussi complexes, nous verrons dans le chapitre 3 comment élaborer une représentation générique et flexible de la chaîne logistique, de ses activités et de ses ressources, pour parvenir ensuite dans les chapitres 4 et 5 à des modèles largement découplés de la gestion des données.

Chapitre 2

Etat de l'art sur la modélisation des chaînes logistiques de biomasse

1. Introduction et synthèses déjà publiées

La plupart des recherches sur la biomasse se sont longtemps concentrées actuellement sur l'amélioration des systèmes de culture et la conception des processus de conversion les plus efficaces. En revanche, les aspects logistiques ont été relativement peu étudiés. En effet, les acteurs concernés ont réalisé seulement récemment que le talon d'Achille des systèmes de production de bioénergie pourrait être la logistique. Cette dernière est en effet déterminante pour approvisionner les unités de conversion de manière régulière, fiable, et avec des quantités suffisantes de biomasse de qualité. De plus, une part significative du coût de la biomasse à l'entrée d'une bioraffinerie réside dans les coûts logistiques. Une chaîne d'approvisionnement efficace est donc un moyen de contribuer à la viabilité économique, sociale et environnementale des futures bioraffineries.

Des modèles sont nécessaires pour évaluer en termes de coûts et d'efficacité les processus de récolte, de transport, de stockage, de prétraitement et de conversion de la biomasse. Comme ces activités consomment de l'énergie et produisent des émissions de gaz à effet de serre (GES), la conception de la chaîne logistique doit être optimisée, non seulement en termes économiques, mais aussi sur le plan environnemental.

Plusieurs articles de synthèse sur les chaînes logistiques peuvent être trouvés dans la littérature. Certains concernent la logistique industrielle (Mula et al., 2010) ou les productions agricoles à vocation alimentaire (Ahumada et Villalobos, 2009). Au sujet de la biomasse et des bioénergies, on peut trouver des synthèses générales répertoriant les articles publiés en les classant selon différents critères, comme dans Gold et Seuring (2011). Les revues de la littérature portant sur les articles de modélisation sont encore rares. Nous ne parlerons pas ici de la modélisation des gisements de biomasse et des systèmes de récolte, pour nous concentrer sur les états de l'art relatifs aux systèmes de production de bioénergies et, en particulier, à leurs chaînes logistiques.

Jebaraj et Iniyar (2006) ont examiné les modèles de planification et de prévision utilisés pour la production d'énergies en général, avec une section sur les énergies renouvelables (dont la biomasse) et une autre dédiée aux modèles d'optimisation. Awudu et Zhang (2012) ont réalisé une synthèse focalisée sur les modèles capables de gérer des incertitudes, pour des chaînes logistiques de biocarburants. Ces auteurs considèrent à la fois la biomasse agricole et forestière et incluent les questions de durabilité, d'impact environnemental et d'acceptabilité sociale.

Baños et al. (2011) ont répertorié les publications sur les méthodes d'optimisation appliquées aux énergies renouvelables. Ils distinguent les méthodes de résolution exactes ou heuristiques mais ne consacrent qu'une page aux bioénergies issues de la biomasse. Wang et al. (2009) et Scott et al. (2012) ont examiné les méthodes d'aide à la décision multicritère dans les énergies durables et les systèmes de bioénergie. Il existe deux synthèses spécialisées sur les chaînes logistiques de biomasse forestière, une de Johnson et al. (2012) pour le placement optimal de sites produisant des biocarburants à partir de biomasse forestière et une plus générale de Wolfsmayr et Rauch (2013).

On constate que les synthèses précédentes ne couvrent que certains aspects de la modélisation des chaînes d'approvisionnement en biomasse. Il n'existe à notre connaissance que cinq revues plus larges (voir ci-dessous), dans l'esprit de notre état de l'art.

Iakovou et al. (2010) ont réalisé une synthèse sur la gestion des chaînes logistiques de biomasse en distinguant les décisions stratégiques, tactiques et opérationnelles. Ils citent des travaux de modélisation, mais aussi des articles plus généraux ou relatifs aux technologies de conversion. On constate dans cet article une certaine confusion entre décisions tactiques et opérationnelles. Mafakheri et Nasiri (2014) ont réalisé une petite synthèse, facile à lire car peu technique, qui liste des publications portant sur la modélisation des différentes étapes d'une chaîne logistique de biomasse, mais aussi sur leurs aspects sociaux, environnementaux et financiers.

Sharma et al. (2013a) ont présenté des modèles d'optimisation développés pour les chaînes logistiques de biomasse et ont identifié des défis et des pistes de recherche. Leur article est plutôt une taxonomie des modèles selon la structure de la chaîne, les contraintes prises en compte, les décisions à déterminer et les critères d'optimisation. Il ne couvre ni les approches de type simulation, ni les modèles stochastiques, ni les méthodes de calcul utilisées pour résoudre les modèles. De plus, les articles référencés ne dépassent pas 2011 et il y a eu des nouveautés depuis.

Shabani et al. (2013) passent en revue les études qui ont utilisé des modèles mathématiques déterministes et stochastiques d'optimisation dans les chaînes d'approvisionnement en biomasse forestière pour la production d'électricité, de chaleur ou de biocarburants. Ils soulignent que les futures études devraient prendre en compte des objectifs environnementaux et sociaux, en plus des aspects économiques. L'article cite notamment des publications traitant des incertitudes sur la demande, les rendements des technologies de conversion, l'offre de biomasse, et les prix.

Un bilan plus récent et plus général a été publié par Yue et al. (2014). Il présente une vue d'ensemble de la modélisation logistique pour la biomasse, en incluant l'optimisation des systèmes de culture et de récolte, les procédés de conversion et la distribution de biocarburants. Il se termine par des défis, comme l'intégration avec la logistique des raffineries pétrolières, la compétition avec les autres productions agricoles, et le commerce international des biomasses et des biocarburants.

De Meyer et al. (2014) ont écrit la revue la plus récente et qui nous semble la plus intéressante pour des chercheurs en recherche opérationnelle. Elle se concentre sur les modèles et méthodes d'optimisation déterministes pour des chaînes d'approvisionnement de biomasse. Une description générale des chaînes de biomasse pour la bioénergie est fournie, avec les décisions relatives à leur conception puis à leur gestion. Une classification des articles est également présentée, en fonction du niveau de décision, de l'objectif d'optimisation, et du type de modèle mathématique utilisé.

Par rapport aux synthèses précédentes, notre objectif dans ce chapitre est de fournir un état de l'art actualisé, avec des modèles quantitatifs et des méthodes de calcul dans un domaine en évolution rapide, sous le double éclairage des méthodologies issues de la logistique industrielle et des techniques actuelles de modélisation informatique et mathématique. Nous nous limitons dans la suite à la chaîne d'approvisionnement en biomasse puisque la partie en aval des bioraffineries ne concerne pas le projet dans lequel prend place cette étude.

En plus de l'optimisation déterministe, nous couvrons des modèles d'optimisation stochastique et des méthodes d'évaluation de performances, y compris la simulation et des approches plus simples basées sur des tableurs ou reposant sur des systèmes d'informations géographiques. En fin de chapitre, nous avons souhaité formuler quelques critiques sur les chaînes logistiques étudiées, les types de modèles et techniques de modélisation, et les algorithmes de résolution.

2. Modélisation de la durabilité en logistique de biomasse

Le souci permanent de durabilité a donné naissance à un courant appelé *logistique verte* (*green logistics*). Les aspects environnementaux, économiques et sociétaux que nous avons présentés étant souvent qualitatifs, il est difficile de les prendre en compte dans des modèles numériques, qui sont par essence quantitatifs. Par contre, les aspects quantifiables comme les émissions de GES peuvent donner lieu à des contraintes et à des critères d'optimisation spécifiques dans les modèles utilisés pour concevoir et optimiser les chaînes logistiques. Comme les sections 4 et 5 sur les techniques de modélisation citent des articles avec des critères de durabilité, nous nous concentrons ici sur les principes généraux et quelques articles représentatifs.

Les différents leviers pour améliorer la durabilité dans les chaînes logistiques en général sont discutés par exemple dans Linton et al. (2007). Une synthèse de Nikolopoulou et Ierapetritou (2012) analyse une quarantaine d'articles et donne trois axes d'amélioration : a) l'efficacité énergétique et la gestion des déchets, b) l'impact environnemental et c) la gestion des ressources en eau. Le premier axe inclut les procédés de fabrication économes en énergie, la récupération de l'énergie, la réduction des déchets, mais aussi leur récupération pour recyclage ou *remanufacturing*. Le second vise surtout à minimiser les émissions de GES et de polluants divers. Cette synthèse montre que les techniques de modélisation et de calcul sont très variées : programmation linéaire et non linéaire, métaheuristiques comme les algorithmes génétiques, simulation, logique floue et même théorie des jeux.

La façon la plus simple de traiter un critère de durabilité dans un modèle est de le caractériser par une contrainte supplémentaire, par exemple des émissions de GES respectant un quota alloué. Beaucoup d'articles vont plus loin en minimisant une somme de différents critères, économiques et environnementaux. Paksoy (2010) minimise par exemple les coûts de production, les coûts de transport et les émissions de CO₂ dans une chaîne logistique industrielle à trois étages. En cas de critères conflictuels, ce genre d'approche peut favoriser un critère aux dépens d'un autre dans la solution obtenue. Il faut alors utiliser les méthodologies spéciales de l'optimisation multicritère, exposées par exemple dans le livre de Branke et al. (2008).

Un bon exemple est Zhou et al. (2000), qui ont proposé un modèle linéaire multicritère pour des industries à processus continu. L'objectif est de maximiser les profits et l'utilisation de la capacité industrielle de manière durable, ce qui est modélisé par une minimisation conjointe des

consommations d'énergie et de matières premières. La résolution combine la *programmation par buts* (en anglais *goal programming*) et *l'analyse hiérarchique multicritère* (AHM, en anglais *analytic hierarchy process – AHP*). La programmation par buts consiste à choisir une valeur souhaitée (le but) pour chaque critère d'optimisation et à minimiser l'écart total à ces buts. En général, le but pour chaque critère est la valeur optimale du critère, obtenue en résolvant le problème mono-critère. L'AHM sert à déterminer des pondérations adéquates (priorités) pour les différents buts. Une étude de cas est consacrée à un complexe pétrochimique avec sa logistique amont et aval.

Une approche plus générale pour évaluer un impact environnemental est *l'analyse du cycle de vie* (ACV, LCA pour *Life Cycle Assessment* en anglais). Tout au long de la vie d'un produit, d'un processus ou d'une activité, les consommations de matières et d'énergie ainsi que les divers rejets dans l'environnement (déchets, GES etc.) sont identifiés, quantifiés et évalués en termes de dommages environnementaux. Un bon exemple est l'indicateur Eco-99 (Pré Consultants, 2000), qui agrège 11 catégories de dommages affectant la santé humaine, la qualité des écosystèmes et l'épuisement des ressources naturelles. La valeur de cet indicateur peut même être calculée avec un niveau de confiance donné quand les données sont incertaines, en considérant des lois normales.

Un exemple d'ACV pour une chaîne de biomasse forestière peut être trouvé dans Forsberg (2000). Comme souvent, cet article utilise l'ACV pour quantifier des scénarios, mais il n'y a pas vraiment d'optimisation. La minimisation d'indicateurs ACV dans des modèles mathématiques est en fait très récente et utilise des méthodes multicritères, comme dans les cinq articles suivants.

Guillen-Gosalbez et Grossmann (2009) considèrent la conception de chaînes logistiques dans l'industrie chimique. Leur modèle est un programme non linéaire, stochastique et bi-critère, où l'on cherche à maximiser le profit total de la chaîne, d'une part, et la valeur Ω de l'indicateur Eco-99 avec un niveau de confiance θ (c'est-à-dire $Prob(ECO_{99} \leq \Omega) \geq \theta$), d'autre part. Santibañez-Aguilar (2011) utilisent une méthode très similaire pour une bioraffinerie mexicaine pouvant utiliser des biomasses variées. Ils optimisent les mêmes indicateurs, sauf que leur modèle ne considère pas d'incertitudes dans le calcul de l'indicateur Eco-99.

Elms et El-Halwagi (2010) utilisent également l'ACV, pour la production de biodiesel à partir d'huile de soja et d'huiles de cuisson usagées, mais ils minimisent le profit total en tenant compte des coûts et des subventions en cas de réduction des émissions de GES. Cucek et al. (2012) réalisent aussi une ACV pour optimiser une chaîne de biomasse forestière : leurs critères sont le profit total, la consommation d'eau, la consommation d'énergies non renouvelables, les rejets aquatiques, les émissions dans l'atmosphère et la conversion de cultures alimentaires en cultures énergétiques.

Tan et al. (2012) incorporent dans un modèle d'approvisionnement multirégional en biomasse des contraintes sur les empreintes écologiques sur l'atmosphère, l'eau et les sols. L'originalité de leur travail réside dans l'utilisation de la logique floue, une méthodologie qui convient bien aux données incertaines et aux interactions complexes, sans devoir recourir à des probabilités.

Pour terminer cette section, d'autres méthodologies très générales ont été développées pour évaluer la durabilité des bioénergies. Elghali et al. (2007) proposent ainsi une méthode d'analyse multicritère et des techniques de scoring, dans le but d'élaborer des recommandations pour les politiques concernant le développement des bioénergies au Royaume Uni.

3. Niveaux de décision dans les chaînes logistiques

Cette section expose les principales décisions dans les chaînes logistiques de biomasse, avec des exemples de publications. Un accent particulier est mis sur le niveau de planification tactique qui concerne cette thèse.

3.1 Niveau de décision stratégique

Les décisions stratégiques dans les filières de biocarburants comprennent le choix des technologies de conversion et de production d'énergie, la sélection des techniques de prétraitement de la biomasse, le nombre, la capacité et la localisation des installations (raffineries, entrepôts de stockage, unités de prétraitement), la configuration générale du réseau logistique, les contrats à long terme, les mécanismes d'association (création de coopératives) et les critères de durabilité.

La création d'une bioraffinerie soulève d'abord des questions de faisabilité technique et économique comme l'évaluation des ressources en biomasse et l'estimation des besoins et des débouchés (Gronalt et Rauch, 2007). La conception de l'usine et de sa chaîne d'approvisionnement impliquent ensuite des choix critiques : quels types de biomasse seront utilisés, quels biocarburants ou autres énergies seront produits, et quelles sont les principales infrastructures à construire (Ekşioğlu et al., 2009). D'autres aspects tels que la disponibilité des matières premières, le coût de construction et d'entretien des raffineries, le débat énergie/alimentation, les questions environnementales et la durabilité sont des facteurs importants à considérer.

A cause de leur importance sur la conception des chaînes d'approvisionnement, nous allons détailler, dans la suite de cette section, les décisions stratégiques concernant le choix des types de biomasse, l'emplacement des installations et la conception générale de la chaîne.

Le choix entre utiliser un ou plusieurs types de biomasse est typiquement une décision stratégique. La possibilité de combiner plusieurs types de biomasse est encore peu étudiée, malgré les avantages qu'une telle approche est susceptible d'avoir. L'un des principaux défis techniques de l'approche multi-biomasse est évidemment la capacité de la technologie de conversion à utiliser un mélange composé de plusieurs types de biomasse, avec des caractéristiques variables. Du point de vue de la logistique, seul aspect étudié dans cet état de l'art, le principal avantage est de pouvoir alimenter une unité de conversion sur une plus longue période dans l'année, au lieu de l'utiliser seulement sur une courte période ou d'allonger cette période mais au prix de stocks importants (Hamelinck et al., 2005). L'approche multi-biomasse peut donc permettre de diminuer notablement le coût global du système, d'amortir plus vite les infrastructures, de lisser les stocks, mais aussi de moins dépendre des aléas pesant sur telle ou telle culture.

La création d'une filière de valorisation de biomasse implique la détermination du nombre, des capacités et des emplacements géographiques des unités de conversion et des sites de stockages. La modélisation et l'optimisation de ce genre de décisions relève d'une discipline appelée *facility location* en anglais (Daskin, 1995). Un exemple typique d'articles consacrés au placement des sites en logistique de biomasse est celui d'Ekşioğlu et al. (2009) qui déterminent la taille, le nombre, et la capacité d'unités de conversion et des sites de stockages.

Samsatli et al. (2015) ont développé l'un des modèles stratégiques les plus récents et flexibles appelé BVCM (Biomass Value Chain Model). Ce dernier détermine les surfaces des zones de cultures attribuées à la production de chaque biomasse pour chaque décennie, la quantité de biomasse stockée et transportée, ainsi que le nombre, la capacité et l'emplacement de chaque technologie de conversion, afin de minimiser des coûts, maximiser un profit ou la production totale d'énergie ou d'exergie, minimiser des émissions de GES.

La conception de chaînes complètes incluant la distribution des bio-carburants est peu étudiée, citons Feng et al. (2010) pour les ressources forestières et Ekşioğlu et al. (2009) ainsi que Marvin et al. (2013) pour des bioraffineries acceptant des biomasses variées. En effet, l'optimisation globale de telles chaînes est extrêmement difficile : d'une part, des acteurs très différents interviennent avant et après les bioraffineries et, d'autre part, les données complètes et fiables manquent.

Outre les citations précédentes, des exemples typiques de modèles stratégiques peuvent être trouvés dans différents travaux. Rentizelas et al. (2009a) abordent la production de la chaleur, de l'électricité et du froid à partir de biomasses variées. Rauch et Gronalt (2011) et Frombo et al. (2009) s'intéressent à la biomasse forestière. Parker et al. (2010) évaluent le développement des biocarburants dans l'ouest des USA. Akgul et al. (2011) et Machani et al. (2013) étudient une chaîne logistique de bioéthanol allant jusqu'à la distribution. Awudu et Zhang (2013) proposent un modèle permettant de gérer des incertitudes sur les demandes et les prix. Samsatli et al. (2015) utilisent une représentation générique et flexible inspirée des graphes états-tâches de Kondili et al. (1993).

3.2 Niveau de décision tactique

Les décisions stratégiques précédentes permettent d'établir la structure de la chaîne en plaçant et en dimensionnant ses principales installations. Avant de détailler les opérations dans chaque journée, qui relèvent du niveau opérationnel, il faut planifier les activités sur un horizon de quelques semaines ou mois, pour anticiper les fluctuations, estimer les ressources nécessaires et s'assurer de leur disponibilité le moment venu. C'est le rôle du niveau tactique, qui considère le plus souvent un horizon temporel divisé en créneaux appelés "périodes", qui peuvent être des jours, des semaines ou des mois selon la précision souhaitée. Ce niveau correspond au plan de production dans les systèmes industriels. Il est tentant de prendre des jours, mais une granularité temporelle aussi fine nécessite plus de données, qui peuvent manquer au moment de la planification. Par conséquent, plus l'horizon considéré est long, plus la granularité est grossière en pratique.

Signalons que beaucoup de modèles de la littérature revendiquent des décisions tactiques alors qu'ils ne considèrent qu'une seule période. Par exemple, Akgul et al. (2011) déterminent des quantités récoltées et transportées en tonnes/jour, sur un horizon mono-période d'un an. En fait, ils raisonnent sur des productions moyennes annuelles de biomasse et divisent les tonnages annuels obtenus par le nombre de jours d'activité dans l'année. On a donc l'illusion de travailler sur des créneaux d'une journée alors que le modèle ne fournit pas des résultats variant au jour le jour. Pour nous, ce genre de modèle est à la limite du niveau stratégique.

Par contre, un modèle mono-période peut être considéré comme tactique s'il se limite à un horizon assez court, inférieur à un an. Huisman et al. (1997) ont ainsi élaboré un modèle à une période pour le miscanthus mais ils disent clairement qu'ils ne traitent que la période de récolte, au printemps.

Les principales décisions tactiques sont, pour chaque période, la quantité récoltée dans chaque zone de culture, les quantités entrant ou sortant de chaque stock, la quantité traitée par chaque site de prétraitement, les quantités consommées par l'unité de conversion, et les quantités transportées (les flux) entre tout couple de nœuds du réseau. La détermination de toutes ces décisions constituent la *planification agrégée* ou *planification tactique*. Toutes ces variables doivent bien sûr respecter les contraintes de disponibilité de la biomasse et les capacités de stockage, de production et de transport. Outre les quantités de biomasse récoltées, transportées, prétraitées, stockées ou converties, la planification tactique peut estimer les moyens nécessaires (machines agricoles, véhicules de transport, ressources humaines) et, si nécessaire, les ajuster d'une période à l'autre.

Par exemple, le niveau stratégique va définir un flux annuel de biomasse entre une zone de culture et une raffinerie. Le diviser par le nombre de jours de travail dans l'année et la charge d'un camion donne le nombre de rotations nécessaires. Mais il s'agit d'une moyenne qui peut être dépassée en période de pointe. En prévoyant les quantités période par période, le niveau tactique peut détecter ces insuffisances, ce qui permettra d'acquérir ou de louer à temps des véhicules supplémentaires.

Mapemba et al. (2008) ajustent ainsi le nombre de moissonneuses nécessaires chaque mois pour collecter de la biomasse sur un horizon d'un an. Ils gèrent aussi un type original de décision tactique : la biomasse peut être récoltée avant complète maturité, mais avec un rendement inférieur. Shastri et al. (2011a) sélectionnent à la fois les équipements de récolte et les moyens de transport. Thorsell et al. (2004) comparent les productivités de différentes combinaisons d'équipements de récolte. Pour les cultivateurs, un autre moyen de compléter des moyens limités de récolte et de transport est le recours à des prestataires logistiques (Huisman et al., 1997; Tatsiopoulos et Tolis, 2003).

La décomposition en périodes offre une grande souplesse de modélisation. Par exemple, la flotte de transport peut être de taille constante ou donnée période par période. La variation des besoins permet d'anticiper les réapprovisionnements ou d'ajuster les stocks de sécurité. On peut restreindre la récolte à un nombre donné de périodes consécutives. Cette fenêtre peut en plus dépendre de la date de semis et donc être spécifiée pour chaque zone de culture. Là encore, on voit qu'on peut augmenter à volonté la précision mais en multipliant la quantité de données nécessaires.

Les caractéristiques des exploitations et de leurs équipements permettent aussi d'avoir des productivités réalistes pour les chantiers de récolte. Par exemple, si un champ de 20 ha est prêt à récolter avec 2 tonnes de biomasse à l'hectare et qu'on dispose d'une moissonneuse disponible 12h par jour et pouvant récolter 3 tonnes par heure, on peut en déduire que la récolte devra s'étaler sur au moins deux périodes, avec une quantité inférieure ou égale à 36 tonnes / jour.

Malgré la maille temporelle qui ne permet pas de définir l'ordre exact et les heures des opérations dans chaque période, les modèles tactiques permettent de connaître l'état du système période par période : quantité de biomasse restant à récolter, flux, niveaux des stocks, etc. La détermination de toutes ces variables permet d'estimer assez précisément les coûts et donc les consommations d'énergie et les rejets de GES. Par exemple, le coût d'un transport de 200 tonnes sur 50 km (10 000 t.km) est calculable à partir de barèmes donnant le coût par t.km pour divers types de véhicules. Des auteurs comme Sultana et Kumar (2011) ont décomposé les opérations de récolte, stockage et transport pour établir des coûts précis. En sommant les consommations de ressource et les différents coûts pour chaque période, on obtient des indicateurs de performance globale comme le coût total.

Pour trouver des modèles tactiques représentatifs, le lecteur peut consulter Cundiff et al. (1997) pour une chaîne de biomasse herbacée avec prise en compte de scénarios météorologiques, Akhtari et al. (2013) pour l'approvisionnement en biomasse d'unités de chauffage urbain, Shabani et Sowlati (2013) pour des chaînes de biomasse forestière, Elms et El-Halwagi (2010) pour la prise en compte des émissions de GES, Papapostolou et al. (2011) pour la production de biocarburants, d'électricité et de chauffage, Shastri et al. (2011a) et Lin et al (2014) pour la sélection des moyens de récolte, de transport et des sites de stockages, et Eksioglu et al. (2009) qui combinent des décisions tactiques et des décisions stratégiques de placement d'installations.

3.3 Niveau de décision opérationnel

Quelques modèles d'optimisation ont été développés au niveau opérationnel, mais uniquement pour traiter certains aspects comme l'ordonnancement du transport et du personnel.

Higgins (2002) décrit ainsi un modèle affectant des conducteurs par roulements de 12h à des moissonneuses de canne à sucre. Ravula et al. (2008a) comparent deux politiques de gestion de chargeurs pour remplir des camions et transférer la biomasse vers des centres de stockage : soit les chargeurs passent de champ en champ en effectuant des tournées, soit on pré-affecte les chargeurs à des groupes de champs contigus.

Han et Murphy (2012) étudient l'ordonnancement de camions pour transporter quatre types de biomasse dans l'ouest de l'Oregon. L'objectif est de minimiser le nombre de camions utilisés et leur temps total d'utilisation, sous des contraintes de durée du travail. Recio et al. (2003) déterminent un ordonnancement de coût minimal d'opérations de récolte : chaque tâche peut être réalisée plus ou moins vite, en choisissant parmi des combinaisons de ressources de coûts différents.

Si on renonce à l'optimisation, on peut gérer des aspects opérationnels grâce aux modèles de simulation présentés dans la section 4, qui peuvent décrire des enchaînements de tâches, des priorités, des synchronisations, des phénomènes aléatoires, etc. Un logiciel de simulation déroule ensuite le temps pour reproduire en accéléré le comportement du système réel et produire des statistiques telles que la durée moyenne d'une tâche.

Le plus célèbre modèle de simulation pour la biomasse est IBSAL (Sokhansanj et al., 2006) qui sera détaillé plus loin. En fait, même s'il peut modéliser des décisions opérationnelles, ce genre de modèle est surtout utilisé pour la planification tactique, en laissant le simulateur reproduire plusieurs jours d'activité. Ainsi, Ravula et al. (2008b) simulent une campagne de collecte de balles de tiges de coton sur 75 à 85 jours tout en gérant finement les rotations de camions au niveau opérationnel. Nous reviendrons sur les différences entre modèles d'optimisation et de simulation dans la suite.

4. Modèles d'évaluation de performances

Dans l'évaluation de performances, qui fait l'objet de cette section, l'utilisateur spécifie la structure et les valeurs des paramètres de son système puis détermine des indicateurs de performance, dont le calcul peut être complexe. Cette approche peut modéliser des phénomènes aléatoires mais ne permet pas simplement et rapidement de préconiser des décisions optimales.

Les modèles d'évaluation de performances pour les chaînes de biomasse sont des calculs de coûts, des approches reposant sur des *systèmes d'informations géographiques* (SIG, *GIS* en anglais), ou des modèles de simulation.

D'autres outils d'évaluation de performances ont été appliqués à des chaînes logistiques industrielles, comme les réseaux de Petri (Lindemann, 1998; Chen et al., 2005), les chaînes de Markov (Buffett, 2005) et les réseaux de files d'attente (Kherbache et Mac Gregor Smith, 2004). Mais jusqu'à présent ils ne semblent pas avoir été appliqués à la logistique de biomasse.

4.1 Modèles simples basés sur des calculs de coûts

Dans les modèles les plus simples, l'indicateur de performance est un coût ou profit total qui est déterminé en sommant des coûts partiels correspondant aux différentes étapes de la chaîne. Leur apparente simplicité ne doit pas cacher le fait qu'ils ont besoin de nombreuses données précises, telles que des statistiques de rendements et les coûts des différentes opérations.

Huisman et al. (1997) évaluent ainsi huit options pour la culture du miscanthus sur des sols tourbeux dans le nord-est des Pays Bas. Outre la récolte, le stockage chez les producteurs, et le transport, ces auteurs ont l'originalité d'inclure le système de culture, avec la préparation des sols, la fertilisation, la plantation (à partir de rhizomes ou de plantules obtenues in vitro), le remplacement progressif des plantes mortes et le contrôle des adventices.

Thorsell et al. (2004) se concentrent sur l'étape de récolte pour de la biomasse herbacée telle que du miscanthus. Leur objectif est de chiffrer l'utilisation de différents groupes d'équipements utilisés en synergie (tracteurs, faucheuses, chargeurs, presses à balles...), y compris le personnel nécessaire.

Les tableurs comme Excel sont très utiles pour entrer les données, organiser logiquement les calculs et les automatiser. Hamelinck et al. (2005) proposent un système à base de tableur pour évaluer des chaînes de biomasse comportant des étapes de transport international, et chiffrer leurs coûts, consommations d'énergie et émissions de CO₂.

Un autre modèle reposant sur un tableur est décrit par Delivand et al. (2011) pour évaluer l'approvisionnement en paille de riz en Thaïlande. Une analyse de coûts détaillée est conduite pour trois régions en considérant deux tailles possibles pour les balles de paille. Les grandes balles nécessitent des équipements plus coûteux pour leur constitution mais leur transport nécessite moins de carburant par tonne transportée.

4.2 Modèles à base de SIG

Les SIG sont très appréciés pour leurs capacités graphiques : les représentations sous forme de cartes permettent une visualisation globale du réseau et de l'étendue spatiale de la chaîne logistique. Le principe fondamental des SIG est de pouvoir superposer plusieurs fonds de cartes géo-référencés, par exemple un fond topographique plus un réseau routier plus une carte de parcelles cultivées. On peut même faire coïncider une carte avec une photo satellite infrarouge, ce qui permet d'évaluer la maturité d'une culture, de détecter des plantations illicites ou de diagnostiquer des maladies. La difficulté est évidemment de se procurer les différents fonds de carte, qui sont parfois très chers.

Des fonctions géométriques sophistiquées sont fournies en standard par les SIG. La longueur d'une route est ainsi facilement déterminée, même si elle est formée de courbes ou de plusieurs segments. Le centre de gravité d'un polygone de production de biomasse est aussi aisément calculable, ainsi que la distance moyenne entre ce polygone et une unité de conversion. En cas de problèmes d'accessibilité à une parcelle, il est même possible de trouver la route la plus proche.

Calvert (2011) recense les applications possibles des SIG dans les projets de bioénergies. Comme le fait remarquer l'auteur, les aspects géographiques sont des facteurs importants qui influent sur la faisabilité de ces projets. Les SIG peuvent être très utiles pour analyser les situations géographiques, évaluer des gisements de biomasse, visualiser l'utilisation des terres et les obstacles naturels et estimer les coûts de transport. Cette flexibilité vient du fait que les SIG incluent une base de données qui permet d'associer des données autres que spatiales aux objets géographiques : on peut par exemple associer à une parcelle l'espèce cultivée, la date prévue de maturité et le rendement estimé.

La littérature offre de nombreux articles appliquant les SIG à la logistique d'approvisionnement en biomasse. Graham et al. (2000) ont utilisé un SIG pour évaluer les variations régionales des coûts d'approvisionnement en biomasse et les impacts environnementaux de la conversion de cultures conventionnelles en cultures énergétiques. Voivontas et al. (2001a) placent et dimensionnent des centrales de production d'énergie sur l'île de Crète (Grèce), en s'aidant d'un SIG pour vérifier si une centrale aura assez de biomasse dans un rayon donné. Tatsiopoulos et Tolis (2003) utilisent aussi un SIG pour diviser une région grecque en mailles rectangulaires et évaluer différentes structures de chaînes pour des tiges de coton. Gronalt et Rauch (2007) se basent sur des cartes forestières et de voies navigables pour évaluer l'impact à long terme du transport fluvial de biomasse en Autriche.

Haddad et Anderson (2008) effectuent une analyse géographique pour identifier de nouvelles sources d'approvisionnement potentiel en résidus de maïs dans le Midwest des USA. Brechbill et al. (2011) déterminent des coûts précis de production de biomasse en utilisant les prix récents pour les semences, les engrais, les herbicides, les équipements de récolte, le stockage et le transport des champs à l'usine. Le rôle du SIG utilisé (ArcMap) est dans leur cas de placer les nombreuses données de production de biomasse sur des zones géographiques sélectionnées.

Les SIG sont surtout utilisés pour des décisions stratégiques : il leur manque en effet la dimension temporelle à court et moyen terme qui est nécessaire pour les décisions tactiques et opérationnelles. Ils ne peuvent pas effectuer directement des calculs d'optimisation mais on peut encapsuler dans un SIG un programme de calcul, écrit dans un langage de programmation classique ou dans le langage de script du SIG. Le SIG ArcView permet ainsi le développement d'algorithmes en Visual Basic. Cette possibilité a été exploitée par Muttiah et al. (1996) qui ont programmé une métaheuristique de type recuit simulé pour sélectionner des sites de stockage de déchets dans l'Indiana.

Pour revenir à la biomasse, Frombo et al. (2009) ont conçu un système d'aide à la décision environnementale, appelé EDSS (*GIS-based Environmental Decision Support System*), pour approvisionner en biomasse forestière des installations de production d'énergie. Cet EDSS est organisé en trois modules : un module de visualisation utilisant les fonctions standard du SIG, une base de données métier et un module d'optimisation. L'utilisateur peut afficher les parcelles forestières disponibles, choisir celles à couper, entrer les différents paramètres comme les coûts puis exécuter le module d'optimisation qui va fournir un plan de transport. Les résultats sont stockés dans la base de données et peuvent être ensuite affichés par le SIG.

4.3 Modèles basés sur la simulation

La simulation consiste à reproduire, en le simplifiant, le fonctionnement d'un système réel pour calculer des indicateurs de performance. Il en existe deux grands types. La *simulation à temps continu* utilise souvent des équations différentielles pour modéliser des systèmes qui changent d'état de manière continue, comme des réactions chimiques, des phénomènes hydrodynamiques, des trajectoires de vaisseaux spatiaux, etc. La *simulation à événements discrets* convient aux systèmes où les événements surviennent à des instants discrets, comme dans la plupart des systèmes industriels et logistiques. C'est donc le type de simulation le plus utilisé pour la logistique de biomasse.

Les bases de la simulation figurent dans les ouvrages de Bradley et al. (1987), Ripley (1987) et Fleury et al. (2007). En général, la simulation repose sur un logiciel comme ARENA, EXTEND, SLAM ou SIGMA et procède en deux phases.

Dans la première phase, l'utilisateur modélise son système, soit dans un langage spécial, soit en connectant des blocs graphiques prédéfinis pour définir des liaisons physiques, des événements, etc. On peut aussi simuler des événements aléatoires, avec différentes lois de probabilité. L'utilisateur définit finalement les indicateurs de performance à calculer.

Dans la seconde phase, un composant du logiciel appelé moteur de simulation fait "tourner" le modèle en reproduisant de manière accélérée les activités du système réel et calcule les indicateurs de performance à l'aide de comptages et de statistiques. De rares auteurs programment leur propre moteur mais la plupart des chercheurs se concentrent sur le modèle et font ensuite confiance au moteur du logiciel pour les calculs.

Dans le cas particulier de la logistique de biomasse, les modèles de simulation sont très utilisés pour les décisions tactiques et opérationnelles. Ils sont assez flexibles pour modéliser de nombreuses situations. Ce sont des modèles à événements discrets, mais les logiciels commerciaux de simulation permettent éventuellement d'inclure des parties à temps continu : Sokhansanj et al. (2006) modélisent ainsi des contraintes de dégradation par des équations différentielles. Une autre raison de la popularité de la simulation est qu'un utilisateur peut assez facilement concevoir et modifier un modèle, grâce aux modules de conception graphique qui se sont généralisés dans tous les logiciels.

Les modèles de simulation de chaînes de biomasse se basent sur des analyses de processus et des calculs économiques pour évaluer les coûts, les consommations d'énergies et les émissions de GES. Ils peuvent reproduire toutes les activités de la chaîne logistique : récolte, transport, stockage etc. Ainsi, Gallis (1996) a conçu un modèle de simulation avec le langage de simulation SLAM pour estimer des coûts d'approvisionnement en biomasse forestière, selon plusieurs scénarios.

De Mol et al. (1997) ont proposé un modèle de simulation pour analyser toutes les opérations de la chaîne logistique et calculer les coûts, les consommations d'énergies et les flux de biomasse, de la ferme à l'usine de conversion, au sein d'un réseau d'installations donné. Les pertes en matière sèche et en humidité sont aussi estimées. Le but principal de ce modèle est d'évaluer la faisabilité économique d'une unité de conversion. Réalisé avec le logiciel de simulation PROSIM, il a été encapsulé dans un logiciel d'aide à la décision appelé BIOLOGICS (*BIOmass LOGIstics Computer Simulation*). Comme beaucoup de modèles commercialisables, l'article est opaque : par exemple, il ne donne ni les formules de calcul utilisées, ni d'exemples numériques.

Trois articles de Nilsson (Nilsson, 1999a, 1999b, 2000) décrivent la conception et des applications d'un modèle de simulation appelé SHAM (*Straw HANDling Model*), pour approvisionner en paille des chaufferies. Il permet notamment d'évaluer l'impact de différents prétraitements de la paille.

Sokhansanj et al. (2006) ont proposé plus récemment le modèle IBSAL (*Integrated Biomass Supply Analysis and Logistics*), développé avec le logiciel EXTEND. Contrairement à SHAM qui a été conçu seulement pour l'analyse des systèmes logistiques de paille, IBSAL intègre diverses biomasses agricoles, y compris les résidus de culture et les cultures énergétiques dédiées. IBSAL semble similaire au modèle de De Mol et al. (1997) mais il est bien décrit, en donnant notamment toutes les formules mathématiques utilisées. Ce modèle peut considérer un horizon allant jusqu'à un an, découpé en jours. Il prend en compte la disponibilité de la biomasse, sa teneur en humidité, les facteurs météorologiques qui affectent l'efficacité des opérations sur le terrain, la performance des équipements, la perte de matière sèche et les différents coûts.

IBSAL permet de déterminer le coût total de chaque étape de la chaîne, les quantités finalement livrées et les émissions de carbone associées. Les différentes entrées et sorties sont organisées en modules. Les auteurs ont appliqué leur modèle à la collecte et au transport de résidus de maïs. Des exemples de simulations sont consultables sur Internet (<http://www.biomass.ubc.ca/IBSAL.html>). Ils mettent en évidence l'inconvénient classique des modèles de simulation : il faut leur spécifier tout ce qu'on veut faire, jour par jour : les parcelles à récolter, les équipements à utiliser, etc. Même les conditions météo doivent être fournies. En effet, contrairement à un modèle d'optimisation, le logiciel ne sait pas choisir les champs à traiter ou les matériels à utiliser.

IBSAL a été appliqué ensuite à différentes situations. Kumar et Sokhansanj (2007) puis Sokhansanj et al. (2009) l'ont employé pour étudier la logistique du panic érigé. Il a servi aussi à Stephen et al. (2010) pour évaluer quatre sites possibles pour une bioraffinerie en Alberta (Canada), et à An et Searcy (2012) qui ont complété le logiciel pour gérer des équipements beaucoup plus variés.

La popularité d'IBSAL ne doit pas masquer ses nombreuses limitations. Il fonctionne uniquement en flux poussé, c'est-à-dire qu'il cherche à récolter la biomasse le plus vite possible pour l'évacuer vers l'unité de conversion, même si cette dernière n'est pas si pressée. Cette politique tend à mobiliser tous les équipements de récolte et de transport disponibles, alors qu'on pourrait étaler leur usage. De plus, IBSAL en version de base est déterministe (pas moyen de traiter des incertitudes), une seule biomasse est considérée, toute la région étudiée est assimilée à une zone unique de production, les capacités de stockage sont supposées illimitées et les équipements sont parfaits (pas de pannes).

Ces limitations ont conduit au développement d'IBSAL-MC (IBSAL-Multi Crop) par Ebadian et al. (2011). IBSAL-MC peut gérer plusieurs biomasses, plusieurs exploitations, des stocks de capacité limitée et des incertitudes comme la météo, les pannes de machines et leurs temps de réparation, en leur associant des lois de probabilités. Il continue à récolter le plus vite possible (en flux poussé), mais l'acheminement vers la bioraffinerie se fait ensuite en flux tiré, en fonction des demandes, les stocks servant de tampons. IBSAL-MC a été appliqué à un projet de production d'éthanol cellulosique dans le Saskatchewan (Canada). Les objectifs étaient de déterminer si les besoins en biomasse pouvaient être satisfaits tout au long de l'année par les productions régionales, d'identifier la capacité des sites de stockage, y compris dans les champs et à l'unité de conversion, de déduire le nombre de machines nécessaires pour chaque opération logistique au jour le jour, et enfin de calculer le coût global du système.

Zhang et al. (2012) ont conçu un modèle qui ressemble à IBSAL mais qui nous paraît plus évolué. La chaîne à horizon d'un an découpé en jours est décrite dans le langage graphique de modélisation de processus IDEF, très utilisé dans l'industrie. IDEF décompose le système en activités, identifie leurs entrées-sorties, puis leurs interconnexions (flux de matières, informations, signaux déclenchant des évènements). Le modèle est ensuite traité par le logiciel de simulation ARENA pour déterminer les coûts, les consommations d'énergie et les émissions de GES. Il a été appliqué à la conception d'une chaîne d'approvisionnement de la biomasse pour une usine de biocarburants dans le Michigan.

L'approche de Zhang et al. nous paraît une alternative crédible aux modèles de planification tactique multi-période reposant sur des programmes mathématiques, mais sans les possibilités d'optimisation de ces derniers.

La simulation peut être moins globale que dans les publications précédentes et traiter seulement une étape de la chaîne logistique. Ravula et al. (2008b) ont ainsi étudié le transport de balles de coton, depuis les champs jusqu'à une usine de cardage qui gère elle-même des véhicules en FIFO pour aller chercher les balles. L'étude a montré que d'autres politiques de gestion utilisent moins de véhicules et diminuent les coûts de transport.

Pour terminer, récapitulons les principaux avantages et inconvénients de la simulation.

Avantages :

- grande souplesse de modélisation (flux, évènements, priorités, attentes, synchronisations),
- peut gérer des phénomènes aléatoires et des incertitudes,
- peut traiter de grands systèmes complexes, avec beaucoup de données,
- convient bien au niveau de décision opérationnel et, à la rigueur, au niveau tactique,
- modèles relativement faciles à comprendre et à modifier par un utilisateur éclairé,
- repose sur des logiciels commerciaux de simulation qui se chargent du calcul des indicateurs.

Inconvénients :

- nécessité de tout décrire en détail dans le modèle (processus, flux, évènements...),
- ne convient pas aux décisions stratégiques comme le placement des unités de conversion,
- considère une structure de chaîne figée (ne peut pas prendre des décisions sur la structure),
- inadaptée telle quelle pour l'optimisation (on doit refaire les calculs si on change les paramètres),
- l'extension appelée "optimisation de la simulation" existe mais elle est souvent très chronophage.

5. Modèles mathématiques pour l'optimisation

Ces approches aussi appelées *approches analytiques* diffèrent radicalement de la simulation. L'idée est de représenter toutes les inconnues par des *variables*, de définir les relations entre ces variables par des équations appelées *contraintes*, et de spécifier un critère à optimiser, appelé aussi *fonction-objectif*. L'ensemble constitue un *programme mathématique*. A l'aide d'un logiciel incluant des algorithmes standards d'optimisation, on peut ensuite déterminer les valeurs des variables qui minimisent ou maximisent l'objectif. La difficulté est d'identifier les variables et équations pertinentes, ce qui nécessite des compétences mathématiques et un certain entraînement.

5.1 Modèles déterministes

On observe un nombre croissant de publications sur la modélisation mathématique des chaînes logistiques de biomasse, surtout dans la dernière décennie. Ces travaux ont été rendus possibles par l'avènement de solveurs efficaces, la montée en puissance des ordinateurs et, surtout, l'expérience considérable acquise en logistique industrielle (voir la synthèse de Mula et al., 2010).

Cette section est consacrée aux modèles déterministes, qui sont de très loin majoritaires. Nous nous concentrons sur les publications proposant des programmes mathématiques. Auparavant, nous citons les rares auteurs qui utilisent des approches d'optimisation plus simples.

Nguyen et Prince (1996) déterminent la capacité de production optimale d'une unité de bioéthanol en Australie en calculant le point d'équilibre entre les coûts de transport de la biomasse, qui augmentent avec la taille de l'unité, et les coûts décroissants de production.

Jenkins (1997) considère aussi un problème de détermination de capacité, mais plus générique. Il détermine pour une longue période des fonctions pour le coût d'investissement I , le coût d'exploitation E et le revenu R généré par les biocarburants puis maximise le profit $R - I - E$ par une simple analyse de fonction.

Alfonso et al. (2009) développent une approche sous-optimale pour placer et dimensionner des unités de production d'électricité et/ou de chaleur dans la région de Valencia, en Espagne. Les emplacements optimaux sont d'abord choisis en fonction de la répartition de la biomasse, des quantités disponibles et des coûts de transport. Ces données sont gérées à l'aide d'un SIG et il semble que le module de calcul associé repose sur des méthodes classiques de localisation. Ensuite, la capacité de chaque unité et ses technologies de conversion sont optimisées en fonction des besoins locaux en énergie, selon une méthode de calcul non spécifiée. Le logiciel permet de minimiser au choix deux critères : la durée d'amortissement des unités ou les rejets de CO₂.

5.1.1 Classification selon les caractéristiques des modèles et les méthodes de résolution

Concernant le type de modèle mathématique (programme linéaire – PL, programme linéaire en nombres entiers – PLNE, programme linéaire mixte – PLM, programme non linéaire – PNL), les modèles les plus simples sont des programmes linéaires purs, à variables continues. Ces variables représentent le plus souvent des quantités récoltées, transportées, stockées, prétraitées ou converties (Cundiff et al., 1997; Akhtari et al, 2013). L'objectif est presque toujours le coût total de la chaîne ou le profit total réalisé.

Les contraintes les plus courantes sont les suivantes :

- Des *contraintes de capacité*, qui signifient que la capacité d'une ressource (au sens large) ne doit pas être dépassée. Ainsi, on ne peut pas récolter plus que la quantité de biomasse disponible, une bioraffinerie est limitée par sa capacité de production, la somme des investissements dans une chaîne logistique ne doit pas dépasser le budget alloué, etc. Par exemple, si deux unités de conversion prélèvent des quantités x et y sur une zone de récolte pouvant fournir une quantité p de biomasse, on a la contrainte de capacité : $x + y \leq p$.

- Des *contraintes de conservation de flux*. Si deux flux de biomasse x et y se rejoignent en un nœud du réseau et se divisent ensuite en deux flux différents u et v , on a : $x + y = u + v$. Ces contraintes correspondent à la loi de conservation des flux aux nœuds (loi de Kirchhoff).
- Des *contraintes de conservation des stocks (inventory balance equations)*. Si à la fin de la période $t - 1$ on a un stock s_{t-1} et qu'au cours de la période t une quantité x_t entre en stock et une quantité y_t en sort, le stock s_t en fin de période t sera : $s_t = s_{t-1} + x_t - y_t$. On peut voir ces contraintes comme des contraintes particulières de conservation de flux, dans lesquelles le flux peut être retenu temporairement dans un stock.
- Des *contraintes de transformation* qui expriment comment la biomasse est affectée par une transformation, y compris un conditionnement, un prétraitement ou une dégradation. Par exemple, si on déshydrate et comprime un volume en entrée v_e de biomasse de densité d_e contenant a % d'eau et que le produit en sortie a une densité d_s , le volume en sortie v_s sera égal au poids en sortie divisé par la nouvelle densité, c'est-à-dire : $v_s = v_e d_e (1 - \frac{a}{100}) / d_s$.
- Des *contraintes de mélange ou d'assemblage*, souvent appelées *équations de Leontief* par les auteurs US. Si une unité de conversion prépare chaque mois une quantité x devant contenir 80% d'un type de biomasse et 20% d'un autre type, les quantités à approvisionner y et z des deux types doivent vérifier : $x = 0,8 \times y + 0,2 \times z$.
- Des *contraintes de qualité* qui stipulent des normes minimales et/ou maximales pour un produit. Par exemple, si on mixe x tonnes de biomasse avec a % d'eau et y tonnes d'une autre à b % d'eau pour obtenir un mélange devant avoir entre h_{min} et h_{max} %, il faut respecter les contraintes $h_{min} \leq (xa + yb) / (100(x + y)) \leq h_{max}$. Cette expression est non linéaire si les variables sont x et y mais on peut la linéariser en $100 \cdot h_{min}(x + y) \leq xa + yb$ et $100 \cdot h_{max}(x + y) \geq xa + yb$.

Les modèles les plus répandus actuellement sont en fait les programmes linéaires mixtes (PLM), comme l'ont souligné Dukulis et al. (2010). Des variables binaires servent souvent pour modéliser des décisions oui/non. Tembo et al. (2003) et Eksioglu et al. (2009) les utilisent pour ouvrir ou non des installations sur certains sites. Nagel (2000) y recourt pour décider si des zones urbaines doivent être raccordées à une unité locale de chauffage urbain ou à un réseau de chauffage plus étendu. Shastri et al. (2011a) définissent des variables entières générales (c'est-à-dire non binaires) pour le nombre de machines agricoles et de véhicules de transport.

Les PLM utilisés combinent les contraintes précédentes et des contraintes assurant la cohérence des variables entières. Par exemple, si une variable binaire y doit valoir 1 si un stock de balles de paille de capacité q est créé sur un certain site et que le nombre de balles stockées est x , on doit traduire l'équivalence $x \neq 0 \Leftrightarrow y = 1$. Il faut pour cela deux contraintes $x \leq qy$ et $x \geq y$. La première sert de contrainte de capacité et force en même temps la variable y à 1 si une quantité $x \neq 0$ est stockée, mais elle n'empêche pas d'avoir $x = 0$ et $y = 1$. Ce cas est évité par la seconde contrainte. Si une raffinerie peut être construite sur trois sites et si x , y et z sont les variables binaires d'ouverture pour ces sites, on a $x + y + z = 1$. Enfin, si une quantité x de biomasse est transportée avec des camions de capacité q , la variable entière n spécifiant le nombre de rotations devra vérifier $n \geq x/q$.

Quant aux programmes non linéaires, ils sont peu utilisés en raison de leur difficulté de résolution. Ils ne sont vraiment nécessaires qu'en présence de contraintes qu'on n'arrive pas à linéariser par les techniques classiques. Shabani et Sowlati (2013) considèrent les approvisionnements d'une centrale

électrique en résidus forestiers (écorces, sciure, produits d'élagage) achetés auprès de différents producteurs. La non-linéarité vient dans leur cas de deux contraintes complexes qui lient les types de biomasses consommées, leurs quantités, leur humidité et la quantité d'électricité qu'on peut en tirer.

La grande majorité des modèles considèrent les coûts des différentes étapes de la chaîne et cherchent à minimiser le coût total ou à maximiser le profit ou la valeur actuelle nette (VAN, en anglais *net present value* – NPV). On ne peut pas parler d'optimisation multi-objectif car ces coûts sont cumulés dans une fonction-objectif unique. Les vrais modèles multi-objectifs apparaissent quand on veut optimiser des critères conflictuels comme les coûts et les impacts environnementaux. Une synthèse de Scott et al. (2012) présente les méthodes utilisées pour les résoudre. Par exemple, You et Wang (2011) minimisent le coût total de la chaîne et les rejets de GES sur tout son cycle de vie. Dans Santibañez-Aguilar et al. (2011), le second objectif devient la valeur de l'indicateur Eco-99 (Pré Consultants, 2000). Ces deux articles utilisent la méthode ε -contrainte.

Certains auteurs parviennent à faire du multicritère en évitant l'optimisation au sens de Pareto. Elms et El-Halwagi (2010) remplacent la minimisation des rejets de GES par la maximisation des subventions octroyées pour ces réductions. Comme leur autre critère est de maximiser le profit de la vente des biocarburants, ils peuvent ainsi se ramener à la maximisation d'une somme de deux profits. Alam et al. (2009) considèrent trois critères, le coût d'approvisionnement, la distance moyenne aux gisements de biomasse (pondérée par les quantités), et la qualité (humidité) de la biomasse, mais ils utilisent la programmation par buts (*goal programming*).

Quasiment tous les auteurs intègrent leurs modèles dans des logiciels commerciaux d'optimisation comme CPLEX, dont les temps de calcul augmentent rapidement avec la taille du problème et le pourcentage de variables entières. Pendant longtemps, les études de cas résolues ont donc été assez petites. Par exemple Kim et al. (2011a) rapportent des résultats pour une chaîne complète avec 5 biomasses, 30 zones de récolte, 2 unités de conversion et 10 zones de demande.

Récemment Marvin et al. (2013) ont résolu un cas de grande taille pour placer des bioraffineries dans le Middle West des Etats-Unis, tout en déterminant leurs capacités et les technologies de conversion utilisées. Il s'agit d'une étude stratégique avec 3109 comtés producteurs de biomasse, 8 types de biomasse, 2 modes de transport, 159 unités de conversion existantes, 98 sites candidats, 4 types de biocarburants, 7 technologies de conversion et 3 marchés. Le modèle, un PLM, contient 154 765 variables continues, 5 488 variables binaires et 39 482 contraintes. Le solveur CPLEX obtient une solution à 0,5% de l'optimum en 8 heures. Ce temps de calcul reste cependant tolérable pour une étude stratégique. Le volume de données nécessaire est par contre important et a reposé sur l'existence de bases de données agricoles US très complètes.

Seules quelques publications ne recourent pas uniquement à des solveurs commerciaux et programment leurs propres méthodes d'optimisation. Shastri et al. (2011b) décrivent une méthode itérative alternant entre deux programmes linéaires mixtes, l'un déterminant les productions de biomasse par les exploitants et l'autre optimisant la partie aval (logistique et bioraffinerie). Bai et al. (2011) emploient la relaxation lagrangienne pour résoudre leur modèle, un PNL difficile.

Papadopoulos et Katsigiannis (2002) s'appuient sur la programmation dynamique pour déterminer chaque mois le mélange de biomasse le moins coûteux pour une unité de biocarburants, en fonction des stocks disponibles, des possibilités saisonnières d'approvisionnement et des coûts associés.

De rares auteurs développent des métaheuristiques. Rentizelas et al. (2009a) optimisent un système de tri-génération pour un quartier, en tenant compte des coûts d'approvisionnement en biomasse, de production d'énergie et de sa distribution. Le problème est formulé sous forme d'un PNL difficile à résoudre. Un algorithme génétique fournit une bonne solution initiale qui est ensuite utilisée par une méthode de programmation quadratique séquentielle. La méthode peut traiter des cas avec 5 types de biomasse et 500 clients mais la partie amont (zones de production et chaîne de transport) est ignorée : la biomasse est supposée livrée avec un coût connu.

Vera et al. (2010) comparent trois métaheuristiques pour placer et dimensionner une centrale électrique espagnole fonctionnant avec des produits d'égavage d'oliviers : un algorithme génétique, une méthode à essaim particulière et une approche récente inspirée par la recherche de nourriture par les abeilles (*honey bee foraging* – HBF). Cette dernière donne les meilleurs résultats.

5.1.2 Classification selon le nombre de biomasses et de périodes de l'horizon

On peut aussi classer les modèles logistiques de biomasse selon le nombre de produits (biomasses) considérés (une ou plusieurs) et le nombre de périodes (mono-période ou multi-période). Nous mentionnerons dans chaque catégorie des exemples de modèles multicritères, quand ils existent.

Cas mono-produit mono-période

C'est le cas le plus simple à modéliser, concernant essentiellement des modèles stratégiques raisonnant sur de grandes quantités agrégées. Kanzian et al. (2009) évaluent l'approvisionnement en biomasse forestière pour une région comprenant entre 9 et 16 centrales électriques. Le programme linéaire mixte est résolu avec le solveur XPRESS.

Les quatre articles suivants incluent des variables binaires pour placer des installations sur une liste de sites potentiels. Freppaz et al. (2004) et Frombo et al. (2009) formulent des chaînes de biomasse forestière sous forme de PLM résolus avec le logiciel LINGO. Leurs modèles permettant notamment de déterminer la capacité des unités de conversion et de déterminer si elles vont produire de l'électricité, de la chaleur ou les deux.

Bai et al. (2011) considèrent le placement de bioraffineries en tenant compte de la congestion du réseau routier dans l'estimation des coûts de transport. Leur modèle est un programme linéaire mixte sauf que la fonction-objectif est non linéaire. La résolution utilise une méthode arborescente (*branch and bound*) avec une borne inférieure lagrangienne. L'article de Bowling et al. (2011) cherche à déterminer avec un PLM les emplacements de hubs de prétraitement décentralisés et d'une unité de conversion. Les études de cas avec le logiciel GAMS sont cependant petites : 6 zones de production, 2 emplacements possibles pour les hubs et 2 pour le site central.

Cas multi-produit et mono-période

Considérer plusieurs types de biomasse constitue une extension naturelle des modèles précédents. Les variables sont munies d'un indice supplémentaire pour le type de biomasse.

De Mol et al. (1997) ont proposé un programme linéaire mixte pour choisir des quantités annuelles de biomasses disponibles à prétraiter et à transporter, tout en sélectionnant des sites de prétraitement et des unités de conversion parmi des sites possibles. L'objectif est de minimiser un

coût total qui comprend les coûts de transport, de prétraitement et de manutention, ainsi que les coûts de création des installations. Le cas résolu avec le solveur OMP inclut 5 types de biomasse, 2 prétraitements possibles (broyage et séchage), 3 modes de transport (route, rail, canaux) et 4 sites pour une unité. L'article donne très peu de détails sur les équations et données utilisées.

Le modèle de De Mol et al. a été complété par Diekema et al. (2005) pour constituer un outil d'aide à la décision, BIOLOCO (*BIOmass LOGistics Computer Optimization*). En particulier, BIOLOCO peut optimiser plusieurs critères en utilisant l'approche de programmation par buts (*goal programming*).

Parker et al. (2010) proposent un modèle d'optimisation pour placer dans l'ouest des États-Unis des bioraffineries alimentées en résidus agricoles, déchets urbains, biomasses forestières, cultures classiques (maïs, soja) et énergétiques comme le switchgrass. Le modèle est un PLM dont l'objectif est de déterminer les emplacements, les tailles et les technologies de conversion des bioraffineries pour maximiser le profit total de la chaîne. L'étude à l'horizon 2015 considère 12 types de matières premières et 6 technologies de conversion sur 6 états américains. Les calculs sont réalisés avec le modeleur GAMS et le solveur CPLEX.

Kim et al. (2011a) cherchent à placer et dimensionner un site de pyrolyse acceptant cinq types de biomasse, puis un site de conversion Fischer-Tropsch acceptant trois sous-produits élaborés par le premier. Le premier site a aussi des débouchés locaux pour du gaz et du charbon végétal. Les auteurs comparent une approche centralisée, où les deux unités sont situées au même endroit, avec une approche distribuée, où les deux unités sont disjointes. Les cas traités ont 30 zones de production de biomasse, 29 emplacements possibles pour le site 1, 10 pour le site 2 et incluent la distribution vers 10 zones de consommation. Ici encore les logiciels utilisés sont GAMS et CPLEX.

Quelques auteurs se sont penchés sur des modèles multi-produits et mono-période avec plusieurs critères d'optimisation. Nous donnons deux exemples, qui ne traitent pas de décisions de placement.

Santibañez-Aguilar et al. (2011) cherchent à maximiser le profit total de la chaîne tout en minimisant la valeur de l'indicateur Eco-99. Ils construisent un programme linéaire mixte qui est résolu par la méthode ε -contrainte à l'aide de GAMS et CPLEX. Cucek et al. (2012) modélisent une chaîne complète en considérant plusieurs objectifs, le coût total et divers impacts environnementaux, dont l'impact social lié à la baisse de production de nourriture induit par la conversion de cultures alimentaires en cultures énergétiques. Le modèle proposé est un programme non linéaire. Comme pour l'article précédent, il est résolu via l'approche ε -contrainte et le modeleur GAMS.

Cas mono-produit et multi-période

L'introduction d'un horizon subdivisé en plusieurs périodes permet de traiter soit des problèmes stratégiques avec un horizon de quelques années, découpé en années, soit le plus souvent des problèmes de planification tactique sur des horizons d'un mois à un an, avec des périodes plus ou moins fines allant de la journée au mois. Au niveau tactique, le découpage en périodes améliore la finesse de modélisation et permet de traiter de nombreuses variations saisonnières. Cependant, la taille des modèles augmente car la plupart des variables ont un nouvel indice pour la période.

En descendant au niveau tactique, certains auteurs supposent que les installations sont déjà placées. Cundiff et al. (1997) modélisent l'approvisionnement en switchgrass d'une bioraffinerie de proximité en Virginie (USA), sur un horizon d'un an découpé en mois et un rayon de 50 km. Le modèle résolu

avec CPLEX est un programme linéaire pur. Les cas traités ont 20 producteurs ayant chacun 3 à 10 parcelles et 4 à 7 sites de stockage (à l'air libre ou couvert), dont la capacité peut être choisie.

Un autre exemple de PL sans décisions de placement est celui élaboré par Tatsiopoulos et al. (2003) pour la récolte et le transport de tiges de coton en Grèce. L'horizon de planification d'un an est découpé plus finement, en jours. Le secteur étudié est découpé en mailles rectangulaires à l'aide d'un SIG qui permet aussi de calculer les coûts de transport. L'originalité de l'étude, basée sur un solveur non précisé, est de comparer deux politiques de transport, soit en utilisant les véhicules des producteurs, soit en recourant à des prestataires logistiques.

La planification tactique sur des horizons assez longs permet cependant d'incorporer des décisions de placement. Haque et Eppin (2012) étudient un horizon d'un an divisé en mois, 6 sites possibles pour plusieurs bioraffineries alimentées en switchgrass et un grand territoire, couvrant les 57 comtés de l'Oklahoma. Ils développent un programme linéaire mixte dans lequel des variables binaires spécifient les emplacements et les capacités des bioraffineries et des centres de stockage, tandis que des variables entières permettent de choisir le nombre d'équipements pour la récolte et la manutention. Ce modèle est résolu avec le modèleur GAMS et le solveur CPLEX.

Zhu et al. (2011) traitent avec le modèleur OPL STUDIO (qui appelle le solveur CPLEX) un réseau assez similaire (switchgrass, horizon d'un an découpé en mois), mais comportant deux aspects originaux : il accepte deux modes de transport (route et train) et prend en compte le réemploi des résidus des bioraffineries comme fertilisants. Un cas avec 10 zones de production, 3 entrepôts et 2 raffineries est présenté.

Zhang et Hu (2013) évaluent l'approvisionnement en résidus de maïs d'unités de conversion fabriquant de l'essence par voie thermochimique. L'étude couvre les 99 comtés de l'Iowa et 30 années découpées en mois. On peut ouvrir une unité, avec un stock de biomasse et un stock d'essence, si elle reçoit au moins une certaine quantité de biomasse. Il n'est pas obligatoire de satisfaire la demande en essence d'un comté un mois donné, mais un coût de pénalité est alors appliqué. Le modèle qui comporte 145 000 variables est résolu en 30 minutes avec le solveur CPLEX.

Cas multi-produit et multi-période

Les modèles de ce genre sont les plus riches en termes de contraintes et de décisions gérées mais aussi les plus complexes, vu le nombre de variables nécessaires. Il s'agit essentiellement de programmes linéaires mixtes. Comme les modèles précédents, on peut les partitionner en modèles sans décisions de placement et en modèles incluant de telles décisions. Les variables binaires supplémentaires rendent plus difficiles les modèles du second type.

Un certain nombre de modèles considèrent des unités de conversion dont l'emplacement est déjà fixé, mais ils utilisent tout de même des variables entières. Gunnarson et al. (2004) modélisent une chaîne logistique de biomasse forestière sur 12 mois : des variables binaires permettent de gérer les parcelles à couper dans chaque période et des contrats avec des scieries.

Gomes et al. (2012) étudient également un approvisionnement en biomasse forestière sur des horizons de 7 à 150 jours mais ils cherchent à déterminer pour chaque type d'essence les quantités transportées et stockées, sous forme de troncs ou réduites en copeaux. Leurs variables entières servent à placer des stocks centralisés ou en bord de forêt pour chaque forme.

Le modèle tactique le plus complet avec des bioraffineries déjà placées est celui de Shastri et al. (2011a) pour la production de bioéthanol à partir de switchgrass. La planification est réalisée sur un an avec des périodes d'un jour. Des variables entières interviennent pour choisir et dimensionner les lieux de stockage, les équipements à utiliser, la flotte de véhicules et le nombre de véhicules utilisés chaque jour. Le modèle, appelé BioFeed, est décliné en deux versions : soit en flux tiré (on fixe la capacité de la bioraffinerie et ses besoins par période, et ces besoins doivent être satisfaits), soit en flux poussé (on récolte la biomasse disponible et on détermine en conséquence la capacité).

Passons maintenant aux modèles qui placent les unités de conversion. Ils sont plus complexes que les précédents si on excepte Shastri et al. (2011a), qui détaillent bien le choix des équipements.

Tembo et al. (2003) décrivent un modèle à horizon d'un an et 12 périodes d'un mois pour approvisionner des unités de bioéthanol avec différents types de biomasse ligno-cellulosique. L'originalité de l'article est de considérer des rendements variables selon le niveau de fertilisation et des dégradations fonctions de la durée de stockage. Un cas avec 77 zones de production, 9 types de biomasse, 11 emplacements et 3 tailles possibles d'unités est résolu avec le logiciel XPRESS.

Zhu et Yao (2011) généralisent le modèle de Zhu et al. (2011) (panic érigé, horizon de 12 fois 1 mois) à deux biomasses supplémentaires : des tiges de maïs et de la paille de blé. Le modèle traité avec le modeleur OPL STUDIO permet de placer les sites de stockage (sur les zones de production ou distants) et les bioraffineries, puis de choisir entre transports routiers ou ferroviaires.

Il existe d'autres exemples de modèles tactiques et certains vont même au-delà des bioraffineries, voir par exemple Ekşioğlu et al. (2009, 2010) et Feng et al. (2010).

Enfin, terminons cette section avec les modèles à la fois multi-produits, multi-périodes et multicritères, qui sont évidemment les plus complexes.

You et Wang (2011) conçoivent un modèle minimisant le coût total et les émissions de GES et traitent une étude de cas sur les 99 comtés de l'Iowa, avec un horizon de 12 mois, 3 types de biomasse et 3 types de biocarburants. L'originalité est de placer à la fois des unités de conversion et des unités de pré-conversion (pyrolyse, gazéification...) avec plusieurs technologies. Chaque comté est à la fois zone de récolte, site potentiel pour une unité, et zone de demande. Le modèle est le plus gros jamais traité en logistique de biomasse : 4 294 326 variables continues, 1782 variables binaires, 772 506 contraintes. La résolution avec GAMS et CPLEX utilise la technique ε -contrainte, en testant 20 valeurs pour les émissions de GES utilisées comme ε . Il faut donc résoudre 20 MILP et le temps de calcul total est énorme : 1 060 heures sur un PC à 2.4 GHz, soit une centaine d'heure par MILP.

Bernardi et al. (2013) proposent un modèle tri-objectif pour le nord de l'Italie, avec un horizon de 15 ans découpé en 5 grandes périodes de 3 ans. Un tel horizon montre que les modèles multi-périodes sont aussi utilisables au niveau stratégique. Le modèle va jusqu'aux zones de consommation des biocarburants et distingue 3 modes de transport (camion, train, barges) et 10 technologies de conversion. Trois objectifs sont optimisés : le profit total, les émissions de GES et la consommation d'eau. En fait, l'étude généralise un travail de Giarola et al. (2011), qui se limitait aux deux premiers objectifs. Comme dans You et Wang (2011), le modeleur GAMS avec le solveur CPLEX sont utilisés pour fournir des solutions Pareto-optimales avec la méthode ε -contrainte. Le cas résolu comprend 307 000 variables continues et 9300 variables entières mais le temps de calcul n'est pas mentionné.

5.2 Modèles stochastiques

Les modèles précédents sont déterministes alors que des incertitudes existent à tous les échelons de la chaîne logistique. Afin d'éliminer l'impact de ces incertitudes, il est souhaitable de les intégrer dans le processus décisionnel. En pratique, n'importe quel composant du modèle peut devenir aléatoire : le coefficient d'une variable, la variable elle-même, un second membre, voire la fonction-objectif.

Les modèles stochastiques sont en général très difficiles à résoudre. En effet, en résolvant un programme mathématique déterministe, on récupère par définition une solution qui respecte toutes les contraintes. Le gros problème de la programmation stochastique est qu'on ne peut plus garantir que toutes les contraintes soient satisfaites pour toutes les réalisations des variables aléatoires. Deux grandes approches sont alors envisageables. *Les méthodes à contraintes de chance (chance-constrained programming)* sont les plus simples : elles consistent à rechercher des solutions respectant les contraintes avec une probabilité donnée, par exemple 90%. *Les méthodes à recours* consistent à générer des solutions avant les réalisations des variables aléatoires puis à effectuer une action corrective (le recours) si une contrainte est violée.

La littérature étant peu fournie en ce qui concerne les chaînes d'approvisionnement en biomasse, nous mentionnons d'abord quelques travaux venant de la logistique industrielle.

5.2.1 Travaux en logistique industrielle a priori pertinents pour la biomasse

La recherche sur la conception de chaînes logistiques industrielles en présence d'incertitudes a bien progressé depuis les années 90 et a inspiré la logistique de biomasse. Le cas typique étudié est une chaîne à trois étages : une ou plusieurs usines, des centres de distribution (CD) et des zones de demandes. Il faut décider où ouvrir les usines et les CD (décisions stratégiques), puis définir les flux (décisions tactiques ou opérationnelles) pour minimiser le coût total de la chaîne. L'incertitude concerne en général l'arrivée des commandes-clients et les quantités commandées.

MirHassani et al. (2000) étudient ce cas standard en le modélisant sous forme d'un programme stochastique à deux étages avec recours. L'idée consiste à mettre dans le premier étage les décisions stratégiques (ouverture d'installations), puisque que dans la réalité on devra construire les usines et CD avant d'observer les réalisations des variables aléatoires. Le second étage consiste à évaluer la structure de chaîne préconisée par le premier étage, en présence de différents scénarios d'arrivée des commandes. Le but de ces auteurs est surtout d'accélérer les algorithmes existants (par une décomposition de Benders) en réalisant des implémentations pour des ordinateurs parallèles.

Plusieurs auteurs adoptent cette approche basée sur un programme stochastique avec recours résolu par la décomposition de Benders, comme Santoso et al. (2005). Il existe d'autres approches de modélisation et de résolution. Miranda et Garrido (2004) veulent définir les politiques de gestion des stocks et les niveaux des stocks de sécurité. Comme ils supposent connues la variance des dates d'arrivée des commandes, ils parviennent à éliminer les recours en formulant un modèle non linéaire qui est résolu par une méthode de relaxation lagrangienne. Sodhi et Tang (2009) montrent que certains modèles de gestion de portefeuilles financiers peuvent s'appliquer aux chaînes logistiques.

A part Sodhi et Tang (2009), les modèles précédents ne sont pas multi-périodes : ils cherchent à déterminer le coût moyen du système par unité de temps sur un horizon infini (régime stationnaire)

ou limité. Alonso-Ayuso et al. (2003) et Al-Othman et al. (2008) considèrent un horizon multi-période et ajoutent des incertitudes sur les coûts à celles sur les demandes. Les premiers autorisent même un ajustement de capacité des usines d'une période à l'autre.

Une approche originale a été proposée par Mahnam et al. (2009) pour une chaîne logistique convergeant vers une usine d'assemblage final, avec des problèmes de fiabilité des fournisseurs. La logique floue est utilisée pour évaluer la fiabilité des fournisseurs. Le but est de déterminer les stocks de sécurité dans la chaîne pour minimiser deux objectifs : le coût total, incluant le stockage et les ruptures de stock, et le taux de commandes satisfaites. Les calculs étant complexes, un module de simulation est développé pour déduire les valeurs des objectifs pour des niveaux de stocks donnés en paramètres. Une métaheuristique à essaim de particules trouve ensuite les valeurs des paramètres qui optimisent les critères, en appelant le module de simulation considéré comme une boîte noire. Il s'agit donc d'une approche d'optimisation de la simulation (*simulation-based optimization*).

5.2.2 Travaux en logistique de biomasse

Les modèles stochastiques en logistique de biomasse ont été revus par Awudu et Zhang (2012). Les méthodes les plus simples visent à évaluer les fluctuations de la fonction-objectif en fonction de paramètres tirés au sort (méthode de Monte Carlo). Schmidt et al. (2009) ont par exemple défini un programme linéaire mixte pour placer en Autriche des unités produisant du chauffage et de l'électricité à partir de biomasse forestière. Constatant que la littérature donne des valeurs très variables pour certains paramètres (prix de l'électricité, coût des unités, potentiel de biomasse récoltable), les auteurs utilisent une méthode de Monte Carlo pour évaluer l'impact de variations gaussiennes de ces paramètres autour de leurs valeurs moyennes. Les mêmes auteurs ont réutilisé la même approche en 2010 pour quantifier les émissions de CO₂ d'installations plus diversifiées.

Les modèles suivants utilisent la modélisation sous forme de programme linéaire mixte stochastique, à deux étages et avec recours. Nous avons parlé plus tôt de Cundiff et al. (1997) pour leur modèle linéaire d'une chaîne de switchgrass sur un horizon d'un an découpé en mois. Le même article étend ensuite ce modèle pour gérer des incertitudes météorologiques. Deux périodes (culture et récolte) et deux conditions météorologiques (bonne ou médiocre) sont considérées, ce qui donne 4 scénarios auxquels sont associés des probabilités. Les auteurs pondèrent les données des différents scénarios par leurs probabilités d'occurrence pour minimiser l'espérance mathématique du coût total.

Awudu et Zhang (2013) veulent maximiser le profit escompté pour des bioraffineries déjà placées et produisant des biocarburants sous des demandes et des prix de vente incertains. Les demandes de produits finis suivent des lois normales mais les variations des prix obéissent à un mouvement brownien géométrique. Le premier étage du programme stochastique fixe les quantités de biomasse achetées, les quantités traitées par les bioraffineries et les quantités de biocarburants produites. Le second étage considère les ventes réelles de biocarburants, les commandes non satisfaites et les niveaux de stocks, qui vont dépendre des réalisations des demandes. Dans la décomposition de Benders, les variables du second étage sont approximées à l'aide d'une méthode de Monte Carlo basée sur un ensemble de scénarios. Une étude de cas pour deux unités de bioéthanol dans le Dakota du Nord montre que le modèle stochastique permet d'avoir un profit supérieur (en termes d'espérance mathématique) à celui obtenu par un modèle déterministe à demandes moyennes.

Osmani et le même Zhang que précédemment (2013) étudient un problème similaire mais avec un placement des unités de conversion et des incertitudes plus nombreuses, avec des rendements et des prix incertains pour la biomasse. La technique de résolution est identique sauf que le premier étage du modèle stochastique se contente de placer les unités de conversion. Les tests montrent que ce placement est peu sensible aux incertitudes et que, comme dans l'article précédent, l'approche stochastique donne un meilleur profit qu'un modèle déterministe.

Kim et al. (2011b) modélisent également un réseau avec placements de bioraffineries mais dans leur cas deux types d'unités sont considérés : des unités de pyrolyse produisant du bio-pétrole et des unités Fischer-Tropsch, qui convertissent ce bio-pétrole en bio-diesel par l'intermédiaire d'un gaz de synthèse. De plus, des incertitudes sur les coûts de transport et de conversion s'ajoutent aux incertitudes traitées par Osmani et Zhang (2013). Le premier étage du modèle stochastique fixe les emplacements et les capacités des unités mais, à part ce détail, la même approche de résolution que dans les deux articles précédents (Monte Carlo) est utilisée pour évaluer les recours.

5.3 Récapitulation sur les modèles d'optimisation

Avantages :

- permettent de déterminer les meilleures décisions pour optimiser un ou plusieurs critères,
- grande puissance d'expression,
- idéals pour la planification stratégique ou tactique,
- de puissants solveurs commerciaux et du domaine public sont disponibles.

Inconvénients :

- plus difficiles à comprendre et à modifier pour les utilisateurs (comparé à la simulation),
- conviennent mal à des décisions du niveau opérationnel,
- temps de calcul élevé sur les grands cas, nécessitant le développement d'algorithmes spécifiques.

6. Conclusion critique

Cet état de l'art a permis de mettre en évidence une activité de recherche en forte croissance sur la modélisation des chaînes logistiques de biomasse. Les travaux déjà réalisés montrent des progrès considérables, avec deux grandes familles de modèles : ceux pour la simulation et ceux pour l'optimisation. Les modèles de simulation comme IBSAL (Sokhansanj et al., 2006) permettent une prise en compte fine de toutes les opérations pour une structure de chaîne donnée et peuvent fournir de nombreux indicateurs de performance. Mais ils ne conviennent pas pour l'optimisation. Par exemple, ils sont incapables de prendre une décision sur la structure de la chaîne : l'utilisateur doit faire une simulation pour chaque structure envisagée.

Au début, les modèles d'optimisation étaient des programmes linéaires à variables continues (Cundiff et al., 1997). Actuellement, on voit une majorité de programmes linéaires mixtes (Shastri et al., 2011a) et même quelques programmes non linéaires (Shabani et Sowlati, 2013). Dans les années 90, ces types de modèles étaient hors de portée des ordinateurs disponibles. Les modèles se sont aussi progressivement complexifiés : ils peuvent inclure par exemple plusieurs critères d'optimisation, notamment environnementaux (You et Wang, 2011), ou des aléas (Awudu et Zhang, 2013).

La richesse croissante des modèles a été rendue possible par la montée en puissance des ordinateurs, la mise à disposition de logiciels de plus en plus performants pour la simulation (ARENA, EXTEND...) et l'optimisation (CPLEX, XPRESS, GUROBI...), mais aussi par l'expertise acquise depuis une trentaine d'années sur la modélisation de chaînes logistiques industrielles.

Nous pouvons cependant formuler quelques critiques sur l'état actuel de la recherche et observer certaines lacunes. Nous les avons classées ci-après en critiques portant sur les chaînes logistiques étudiées, les types de modèles et techniques de modélisation, et les algorithmes de résolution.

Critiques sur les chaînes étudiées :

- En nombre de publications, la quasi-totalité des articles se consacrent à la production de bioéthanol, et ensuite à des unités de chauffage et/ou de production d'électricité. Beaucoup reste à faire pour modéliser les spécificités des bioraffineries oléagineuses.
- Après une profusion de modèles stratégiques à données très agrégées, on voit de plus en plus de modèles tactiques multi-périodes mais leur prise en compte des équipements de récolte et de transport est encore limitée, même si on observe quelques progrès récents (Shastri et al., 2011a).
- Quand plusieurs types de biomasse sont pris en compte, il s'agit de sources différentes (cultures, déchets urbains, biomasse forestière), de plantes différentes, ou de formes différentes d'une même biomasse (comme du bois sous forme de sciures ou de chips, ou du switchgrass en vrac ou en balles). Aucun auteur n'aborde la possibilité d'utiliser à des degrés variables des parties différentes d'une même plante, par exemple les graines, la paille et la menue paille.
- La plupart des auteurs de modèles tactiques multi-périodes considèrent des dégradations sous une forme très simple, avec un coefficient de perte pour chaque période passée en stock. Quelques travaux sont allés plus loin et illustrent bien les difficultés de modélisation afférentes (Cundiff et al., 1997; Van Dyken et al., 2010; Shastri et al., 2011a).
- Peu d'articles considèrent des sites de stockage, de prétraitement et/ou de pré-conversion entre les zones de récolte et les bioraffineries. Quand ils sont pris en compte, il s'agit le plus souvent de petits stocks chez le producteur et de prétraitement simples comme constituer des balles.

Critiques sur les modèles utilisés :

- Les modèles manquent de généralité. Ils imposent une granularité temporelle (par exemple, un horizon d'un an découpé en mois), une granularité spatiale (par exemple, une zone géographique découpée en carrés à l'aide d'un SIG) et une structure figée du réseau.
- Plus généralement, beaucoup d'articles mêlent des données pertinentes pour le modèle et des données qui peuvent être pré-calculées par une couche informatique au-dessus. La question des données à préparer et stocker (base de données par exemple) est très rarement abordée.
- Le modèle et ses équations ne peuvent être définis que par le concepteur du modèle. L'utilisateur peut très difficilement les modifier, vu les compétences spéciales nécessaires. Il nous paraît cependant possible que l'utilisateur puisse décrire son réseau, par exemple sous forme graphique, et ensuite que le modèle soit généré automatiquement à partir de cette description.

- Les modèles proposés sont soit stratégiques, soit tactiques, avec des installations déjà placées ou bien à placer. Nous pensons qu'on peut proposer à l'utilisateur le choix entre plusieurs modèles complémentaires, selon la situation à modéliser.

Critiques sur les méthodes de résolution :

- Les modèles d'optimisation sont quasiment tous résolus avec des logiciels standard, qui sont trop généraux pour bien exploiter leur structure mathématique et nécessitent un temps de calcul qui croît très vite en fonction de la taille des cas traités et du nombre de variables entières. Les modèles sont donc encore assez petits, surtout les tactiques avec des contraintes variées.
- Traiter des cas de plus grande taille et/ou avec une modélisation plus fine nécessite de développer des algorithmes *ad hoc*, comme des méthodes de décomposition ou de relaxation et des métaheuristiques. Ce besoin est critique pour les modèles qui impliquent plusieurs appels à l'algorithme de résolution, comme cela est fréquent en optimisation stochastique (évaluation de scénarii) ou multicritère (méthode ε -contrainte).
- Certaines approches de résolution ayant fait leurs preuves en logistique industrielle ne sont pas encore appliquées à la biomasse, citons les algorithmes génétiques multicritères genre NSGA-II (Deb, 2001) et l'optimisation de la simulation (Mahnam et al., 2009).

Pour résumer, la modélisation des chaînes d'approvisionnement en biomasse présente certaines caractéristiques spéciales qu'on ne rencontre pas en logistique industrielle, comme des ressources désynchronisées et disponibles de manière saisonnière ou la possibilité de choisir entre différents prétraitements ou combinaisons d'équipements de récolte et de transport. Ces aspects particuliers et les incertitudes qui pèsent encore sur de nombreuses données comme les coûts et rendements vont nécessiter de nouvelles recherches pour parvenir à des modèles plus pertinents.

En conclusion, notre revue de la littérature montre qu'il n'existe pas de modèle mathématique publié pour la chaîne d'approvisionnement en biomasse que nous avons décrite dans le chapitre 1, avec notamment une structure non figée (laissée au choix du décideur) plusieurs plantes, plusieurs parties de plantes et différentes chaînes de récolte.

Ce chapitre résume un des livrables du projet PIVERT (96 pages, 215 références bibliographiques), que le lecteur peut consulter pour avoir plus de détails. Il a aussi fait l'objet d'une publication dans une conférence internationale et d'un article dans une revue réputée, *Renewable Energy*.

Références :

B.H. Ba, C. Prins, C. Prodhon, 2014. Revue de la littérature sur les modèles logistiques pour la biomasse. Livrable n°1 du WP1-P3, Institut pour la Transition Énergétique (ITE) PIVERT, 05/02/2014, 96 pages et 215 références bibliographiques.

B.H. Ba, C. Prins, C. Prodhon, 2014. An overview of operations research models for biomass supply chains, ICORES 2014 (3rd Int. Conf. on Operations Research and Enterprise Systems), Angers, 06-08/03/2014, 8 pages.

B.H. Ba, C. Prins, C., Prodhon, 2015. Models for optimization and performance evaluation of biomass supply chains: an operations research perspective, *Renewable Energy* (Elsevier), déjà publié en ligne, DOI : 10.1016/j.renene.2015.07.045.

Chapitre 3

Modélisation des activités logistiques et typologie des données

1. Introduction

Les chapitres précédents ont montré que les chaînes d’approvisionnement en biomasse peuvent être complexes et de structures très diverses. Comme elles ont en général une grande taille (par exemple 500 exploitations agricoles et 50 sites de stockages intermédiaires pour une raffinerie de proximité), comment décrire ces chaînes sans ambiguïté et les coder informatiquement, pour qu’elles soient utilisables par le modèle de calcul? De manière plus ambitieuse, comment être assez flexible pour ne pas être prisonnier d’une structure de chaîne figée? Par exemple, comment pouvoir ajouter facilement des sites de prétraitement, de nouvelles plantes, des chaînes de récolte? Enfin, comment recenser et structurer, par exemple sous forme de base de données, les nombreuses informations à chaque étape de la chaîne (équipements nécessaires, tonnages, coûts etc.)?

Ces questions montrent l’importance de la couche de gestion de données dans la réalisation de tout système d’aide à la décision. Bien que le domaine de cette thèse ne soit pas les bases de données, il a été nécessaire de découvrir le monde de l’agriculture, ses processus et sa terminologie, puis d’effectuer un long et fastidieux travail de recensement, d’analyse et de structuration de données avant de réaliser le modèle mathématique.

Ce chapitre résume ce travail préparatoire. Nous avons d’abord réalisé un modèle de réseau (*network model*) permettant de décrire graphiquement les sites géographiques de la chaîne logistique considérée, ses activités et ses flux. Ce modèle basé sur le formalisme des graphes états-tâches fait l’objet de la section 2. Il permet de valider notre compréhension de la chaîne avec les experts du domaine et de préparer l’étape suivante, l’établissement du modèle de données.

Le modèle de réseau étant essentiellement graphique, il n’est pas facilement manipulable par les algorithmes d’optimisation. De plus, il ne représente pas, pour éviter des surcharges, les nombreuses données annexes comme les besoins en biomasse de la raffinerie, les productivités et les coûts des équipements, les coûts de stockage, etc. Le modèle de données (*data model*) a donc pour but de coder informatiquement le modèle de réseau et toutes les autres informations utiles, sous forme de tables de données facilement assimilables par le logiciel d’optimisation. En particulier, le système de gestion des ressources est particulièrement critique. Nous avons choisi une implémentation sous Excel, qui est plus répandu et plus simple que les véritables systèmes de gestion de bases de données comme Access ou Oracle. Le modèle de données obtenu est décrit dans la section 3.

2. Modèle de réseau (network model)

2.1 Principe des graphes états-tâches

Nous avons vu dans le chapitre 1 (figures 1.1 et 1.2) que les chaînes d'approvisionnement en biomasse se prêtent bien à des modélisations sous forme de graphes. Ce genre de modèle a été effectivement employé par divers auteurs comme Eksioglu et al. (2009). Il permet de représenter les sites intéressants de la chaîne sous forme de nœuds et les transports possibles entre ces sites sous forme d'arcs munis de capacités. Les modèles obtenus pour le niveau tactique sont apparentés à des problèmes de flots (ou de multi-flots) de coût minimal, mais sur plusieurs périodes. Cependant, ils ne sont pas assez fins pour prendre en compte, par exemple, des attentes (temps de séchage, séjours en stock), des périodes de récolte ou d'ouverture de certains sites dans l'année et des affectations de ressources. Plus grave, un nœud peut représenter aussi bien un site de stockage, qui ne transforme pas la biomasse, ou une tâche qui la récolte, la transforme ou la transporte.

Pour parvenir à une finesse adéquate, nous avons choisi le formalisme des *graphes états-tâches* (*state-task networks*). Ces graphes ont été introduits par Kondili et al. (1993) et Shah et al. (1993) pour modéliser des processus en pétrochimie, sur un horizon temporel divisé en périodes discrètes. Ils utilisent seulement deux types de nœuds, les états et les tâches.

Un *état* correspond à un produit distinct en termes de composition, de conditionnement, de présentation, etc. Symbolisé par un cercle, il peut s'agir d'un vrai stockage, d'un stock d'en-cours (*work in process – WIP*) ou même d'un produit en transit rapide dans une conduite entre deux unités de fabrication. Un état peut être produit et consommé par une ou plusieurs tâches.

Une *tâche*, ou processus, symbolisée par un rectangle, reçoit un ou plusieurs états d'entrée, dans des proportions fixées, et les transforme pour produire un ou plusieurs états de sortie, également en proportions fixées. Dans chaque période, les flux sont pris simultanément sur les états en entrée, et un temps de fabrication en nombre de périodes peut être spécifié pour chaque produit en sortie. Un arc correspond à un flux d'un seul produit, toujours depuis un état vers une tâche, ou vice-versa. Il faut insister cependant sur une différence importante entre les états et les tâches : les flux entrants sont indépendants pour les premiers tandis qu'ils se font simultanément pour les seconds.

La figure 3.1 adaptée de Kondili et al. (1993), montre un exemple de graphe avec 9 états et 5 tâches pour un processus pétrochimique. Les périodes sont ici des heures. Des capacités de stockage, parfois non limitantes, sont indiquées dans les états. Des fractions et des temps de traitement figurent sur certains arcs. En entrée de tâche, les fractions désignent les pourcentages des différents produits consommés. En sortie, elles donnent les rendements. Ces fractions sont omises pour les tâches à une entrée ou une sortie, car elles valent 100% par défaut.

On peut voir sur cet exemple la grande flexibilité des graphes états-tâches. Le changement d'une seule caractéristique d'un produit en fait un autre produit : ainsi, "Feed A" devient "Hot A" après chauffage. Des circuits sont possibles : le produit intermédiaire Y produit par la réaction 3 subit une séparation donnant l'intermédiaire Z qui revient à la réaction 3. Dans l'article de Kondili et al. (1993), des exemples de ressources sont aussi associés aux tâches : un four de capacité 100 kg, convenant pour le chauffage, un réacteur de 80 kg et un de 50 kg, utilisables par les trois réactions, et une colonne de distillation, utilisable pour la séparation.

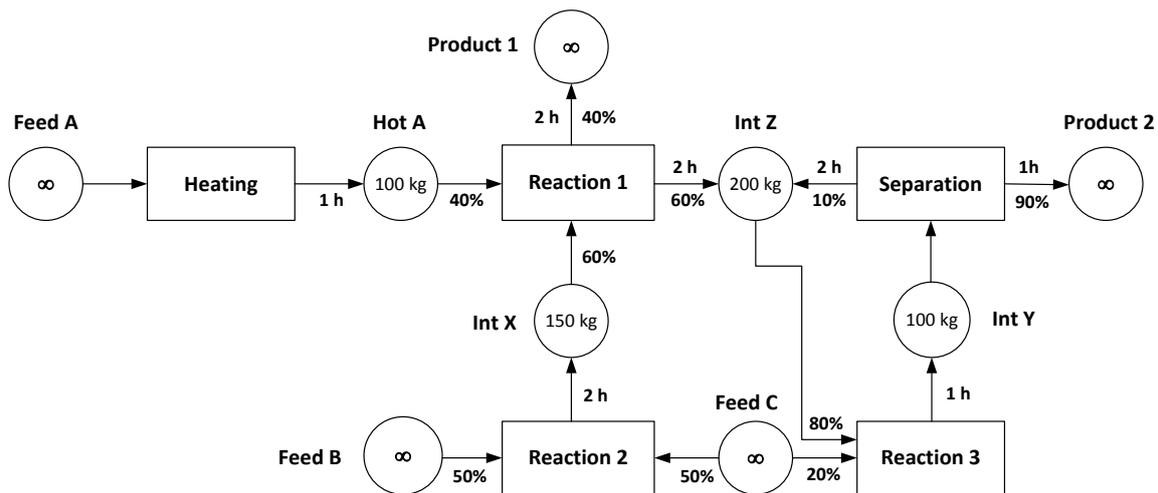


Figure 3.1. Exemple de graphe états-tâches (Kondili et al., 1993).

Les notions d'états et de tâches sont en fait très générales. Un processus chimique peut inclure plusieurs opérations, par exemple une étape de chauffage d'un composant, une étape de réaction et une étape de refroidissement du produit obtenu. Ces étapes peuvent donner lieu à trois tâches mais, si elles sont toujours réalisées dans la même unité de fabrication et sans interruption entre elles, elles peuvent être combinées en une tâche unique. Quant aux états, ils peuvent servir à identifier un produit à existence transitoire, sortant d'un réacteur pour entrer aussitôt dans un autre, ou une accumulation de produit, comme un vrai stock.

Kondili et al. (1993) ont montré que de nombreux problèmes d'optimisation de processus peuvent être modélisés par un programme linéaire mixte, généré à partir d'un graphe états-tâches. En particulier, ce modèle peut gérer des contraintes de capacités (arcs et stocks), de conservation de flux et d'affectation de ressources. Shah et al. (1993) ont proposé diverses techniques pour améliorer la génération de ce genre de modèle. Dans le chapitre 5 nous utilisons ce principe de génération automatique pour découpler au maximum les données et le modèle mathématique. Mais auparavant, il faut adapter les graphes états-tâches au domaine de la biomasse.

2.2 Extensions des graphes états-tâches pour la biomasse

Dunnnett et al. (2007) ont déjà utilisé les graphes états-tâches pour une chaîne logistique fournissant de la biomasse à une installation de chauffage. Une ferme produit du miscanthus qui est récolté en balles ou broyé, puis conservé dans un stockage local avec séchage passif. Les deux formes sont transportées par camion vers une chaufferie. Cette dernière broie les balles, fait un séchage forcé pour atteindre une humidité minimale, puis brûle le produit pour produire de la chaleur. L'horizon de planification couvre une année divisée en mois et les stocks ne sont pas limités. Cependant, les différents équipements sont gérés et l'humidité est suivie tout au long de la chaîne.

Le graphe états-tâches réalisé n'est pas montré et le programme linéaire mixte en annexe de l'article est très peu commenté. Il minimise le coût du système en utilisant des variables continues (quantité traitée par chaque tâche dans chaque période), des variables entières (nombres d'équipements à acheter) et des variables binaires (utilisation d'un équipement dans une période donnée et affectation des équipements aux tâches).

Par rapport à Kondili et al. (pétrochimie), Dunnett et al. (2007) ont réalisé les extensions suivantes :

- des tâches de transport (le produit au départ et à l'arrivée donne donc deux états distincts).
- le produit final immatériel (la chaleur) est aussi considéré comme un état.
- l'humidité est un attribut supplémentaire associé aux états.
- certaines tâches ont des propriétés qui dépendent de la période, par exemple la quantité disponible de miscanthus et son humidité varient selon la période.
- des consommations ("*utilities*") peuvent être associées aux ressources : heures de main d'œuvre, quantités de carburants, etc. Elles permettent par exemple de déduire les rejets de CO₂.

Ainsi, bien que l'étude de Dunnett et al. (2007) soit assez simple (une seule plante, une seule ferme, 12 périodes d'un mois), elle montre que les graphes états-tâches conviennent aux chaînes de biomasse et sont assez facilement extensibles. Pour les chaînes plus complexes étudiées dans cette thèse, nous avons apporté des extensions supplémentaires aux graphes états-tâches :

- planification plus fine, avec des périodes d'un jour et de nombreuses exploitations agricoles,
- introduction de sites géographiques pour représenter un groupe cohérent d'états et de tâches (par exemple un stockage peut regrouper des silos à grains et des plates-formes pour des balles),
- plusieurs plantes oléagineuses et ligno-cellulosiques (colza, cameline, miscanthus...),
- utilisation de plusieurs parties d'une même plante (grains, paille et menue paille pour le colza),
- plusieurs chaînes de récolte possibles pour la même plante,
- choix des ressources des tâches parmi plusieurs modèles (tracteurs de différentes puissances...),
- gestion de combinaisons de ressources, travaillant de manière synchronisée (un tracteur plus une presse à balles, par exemple)
- possibilité de fenêtres de temps sur les états, par exemple des stocks partagés (contenant plusieurs produits) et/ou ayant des périodes d'ouverture limitées dans l'année.

2.3 Application à notre projet et exemples

Nous avons élaboré dans la figure 3.2 un exemple de graphe états-tâches pour une chaîne simple mais complète et conforme aux processus réels, avec une ferme produisant du colza, un site de stockage (coopérative) avec un silo à grains et une plate-forme de stockage de balles, et une raffinerie avec des stocks-tampons. Les périodes sont des jours.

Les grands rectangles gris indiquent les sites géographiques. On voit bien sur la figure que les états S1, S2... correspondent à de vrais stocks (hangar de ferme, silo et plate-forme du site de stockage intermédiaire, stocks en entrée de la raffinerie), mais aussi au champ de colza et aux produits obtenus par la moisson et attendant sur le sol différents traitements (densification, transport). Quant aux tâches T1, T2..., elles incluent la moisson, la densification, le transport et la tâche de conversion de la raffinerie, qui termine la chaîne. Les arcs (flèches) entre les nœuds indiquent les flux possibles.

L'exploitation agricole utilise une chaîne de récolte typique. Une tâche de moisson avec une moissonneuse batteuse récolte le colza (*rape*) et sépare les graines (*seeds*), la paille (*straw*) et la menue paille (*chaff*), dans des proportions connues. Les produits obtenus P1, P2... sont indiqués sous chaque état, avec trois rubriques : la plante d'origine, la partie de plante et la présentation comme le vrac (*bulk*) ou les balles carrées (*SBale*). Par exemple, le produit P4 *Rape/Chaff/SBale* désigne des balles carrées de menue paille provenant du colza.

Les graines sont vidées dans une remorque tirée par un tracteur, qui va ensuite au silo de la coopérative. La menue paille, collectée par la moissonneuse dans un caisson spécial, est déposée en tas en bord de champ. Ce tas est repris par une presse à balles carrées en poste fixe (immobile), mais avec la prise de force d'un tracteur pour actionner la presse. Les balles de menue paille sont acheminées par un tracteur avec une remorque à plateau vers un hangar sur l'exploitation. La paille est disponible après 3 jours de séchage sur andain. Ce temps est simplement exprimé par un délai sur l'arc sortant de la tâche de moisson. Elle est conditionnée en balles carrées par un tracteur équipé d'une presse puis les balles sont emmenées au hangar de la ferme par un tracteur avec une remorque-plateau, ou à la raffinerie, par camion semi-remorque. Des camions transfèrent les balles de paille et de menue paille du hangar de la ferme vers la plate-forme de la coopérative. Le même moyen de transport est utilisé depuis le hangar et la coopérative vers les stocks de la raffinerie.

Nous avons ajouté les capacités des états (à l'intérieur des cercles), les distances (négligeables ou forfaitaires à l'intérieur de l'exploitation) dans les tâches de transport, et les ressources nécessaires au-dessus des tâches. La seule tâche de séparation est ici la moisson, et la seule tâche de mélange est la raffinerie. Cette dernière est particulière car les flux entrants (demandes) vont être donnés période par période : les proportions ne sont donc pas indiquées car elles changent en fonction des demandes. Les proportions sont précisées seulement pour la tâche de moisson (% de graines, paille et menue paille). Sur les autres arcs, elles valent par défaut 100%.

Concernant les capacités, les deux rectangles mauves (S5 avec S13, S9 avec S14) indiquent deux stocks partagés de balles de paille et de menue paille, le hangar de la ferme et la plate-forme de la coopérative. La capacité indiquée dans les deux états de ces stocks indique la capacité totale du stock partagé. La capacité nulle dans l'état qui montre les graines après récolte indique un état sans attente ou "no-wait" : les graines récoltées doivent partir le jour même à la coopérative. Concernant les ressources, on peut évidemment utiliser plusieurs équipements en parallèle s'ils sont disponibles, par exemple deux moissonneuses.

Le seul délai concerne les 3 jours de séchage pour la paille. Les délais pour les autres tâches sont omis car supposés inférieurs à la période d'un jour. Dans ce cas, les opérations peuvent s'enchaîner dans une journée. Par exemple, l'usage est de transporter à un silo, le soir même, les graines de colza récoltées dans la journée. On peut évidemment mettre des délais non nuls dans certains cas. Ainsi, on ne range pas des balles de menue paille dans le hangar de l'exploitation si c'est pour les enlever et les transporter à la coopérative dans la même journée.

En général, pour éviter de multiplier les tâches, les coûts de manutention, de chargement et de déchargement peuvent être comptés forfaitairement, par exemple 1.5 euros/tonne pour le rangement ou le prélèvement dans un silo. Cette méthode est commode car la liste détaillée des petits équipements de manutention est rarement connue, surtout pour les sites de stockage.

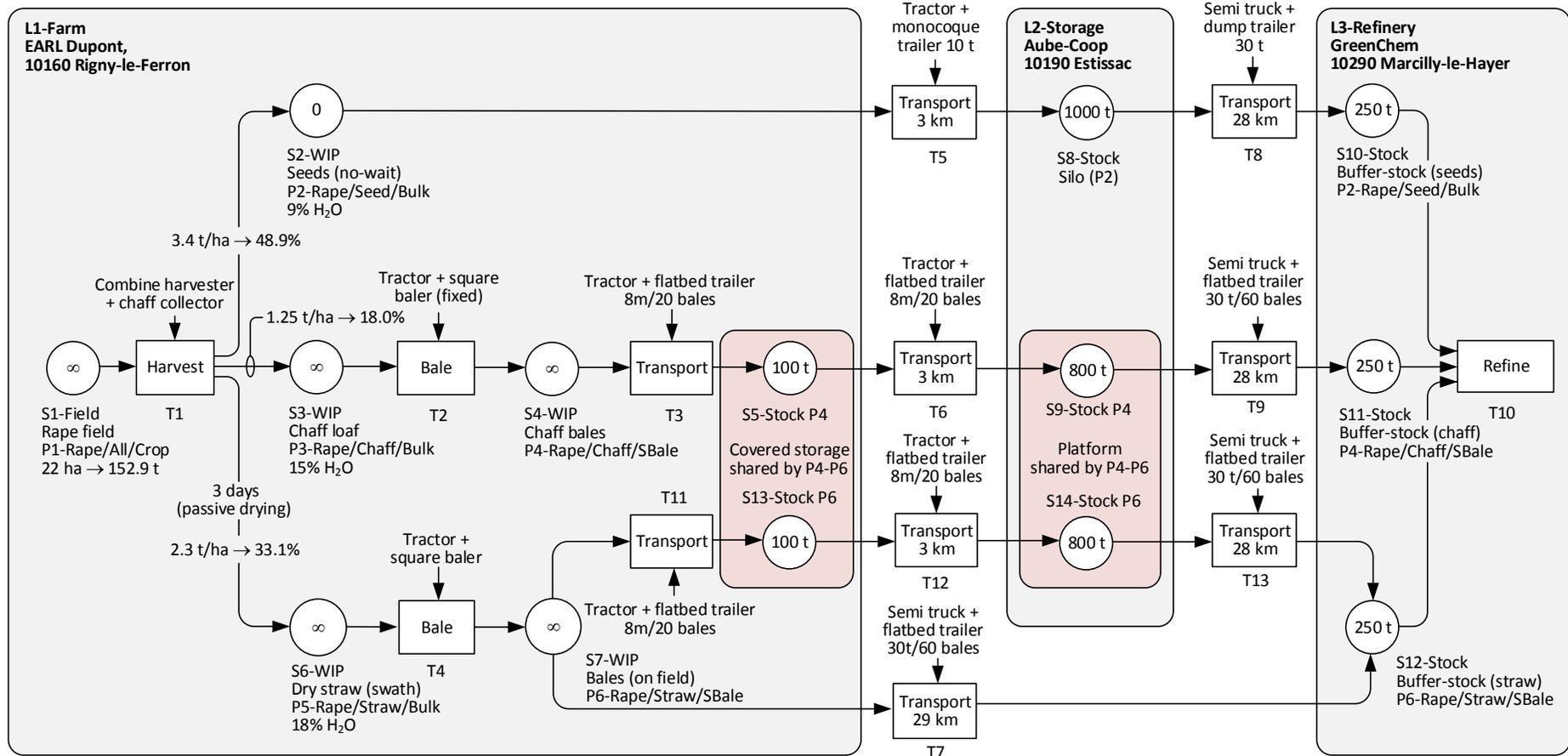


Figure 3.2. Exemple de graphe états-tâches pour une chaîne d'approvisionnement avec deux stocks partagés.

On voit sur cet exemple que le modèle de réseau obtenu est très flexible. Seuls les modèles de simulation peuvent être plus fins mais ils se prêtent mal à l'optimisation. Il convient bien pour des problèmes de planification tactique de chaînes d'approvisionnement (horizon-type d'un an, découpé en créneaux d'une journée à une semaine) et les principaux avantages des graphes états-tâches présentés dans les sections précédentes sont confirmés. Il nécessite évidemment d'autres informations (les divers coûts par exemple), que nous avons omis pour ne pas surcharger la figure.

Dans notre cas, il peut y avoir des centaines d'exploitations agricoles, jusqu'à cinq cultures par exploitation (colza 00, colza érucique, moutarde d'Ethiopie, cameline, miscanthus) et des dizaines de sites de stockage intermédiaire. On peut aussi avoir plusieurs chaînes de récolte par plante. Pour le colza, on peut récupérer ou non la menue paille. Si oui, on peut ajouter un caisson à la moissonneuse pour la récupérer à part (chaîne dite "Colza 1") ou la déposer sur andain pour la collecter avec la paille (chaîne "Colza 2"). Pour le miscanthus, la situation est encore plus flexible. La chaîne "Miscanthus 1" utilise une faucheuse puis un tracteur avec une presse à balle qui passe sur l'andain. La chaîne "Miscanthus 2" emploie un tracteur avec une faucheuse et une ensileuse, cette dernière projetant le miscanthus broyé dans une remorque tirée en parallèle par un autre tracteur. Enfin, le système Kuhn (chaîne "Miscanthus 3") permet de faucher et de faire des balles en une passe.

La chaîne logistique complète pour le projet PIVERT peut être modélisée par un graphe états-tâches, mais ce dernier est trop grand pour être dessiné : il faudra donc le coder et le manipuler sous forme informatique. Pour aider la saisie, nous proposons de prédéfinir les chaînes de récolte possibles, sous forme de sous-graphes déjà saisis avec leurs états et leurs tâches. L'utilisateur pourra alors donner une liste d'exploitations et spécifier quelles plantes sont cultivées et avec quelle chaîne de récolte, comme sur la figure 3.3. Cela évitera de répéter les mêmes états et tâches dans les exploitations qui utilisent une même chaîne de récolte.

Il devra aussi saisir les données pour les sites intermédiaires (une cinquantaine) et la raffinerie, mais cette tâche est mineure par rapport aux 500 exploitations agricoles prévues. Le logiciel de préparation du modèle mathématique pourra ensuite facilement copier les sous-graphes déjà prêts pour les différentes chaînes de récolte.

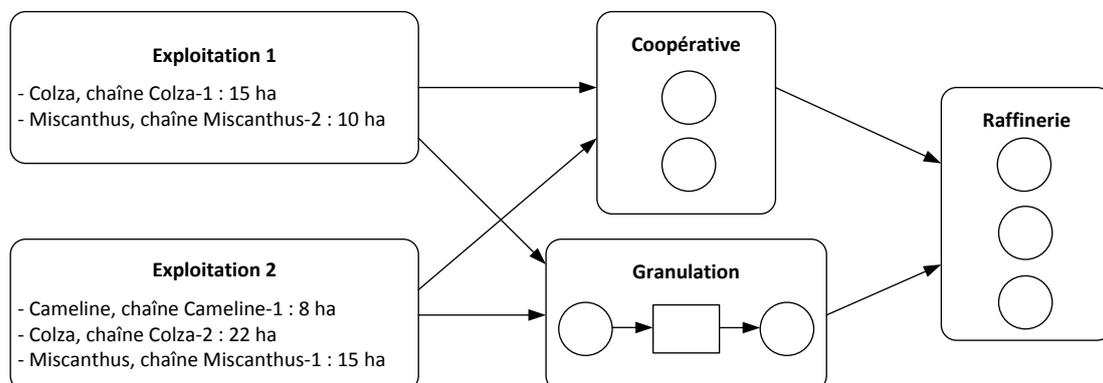


Figure 3.3. Exemple de codage d'une chaîne complète d'approvisionnement.

Les chaînes de récolte prédéfinies ont des codes-chaînes C1, C2... et sont décrites par des états et des tâches numérotés. Quand le logiciel recopie une même chaîne sur plusieurs fermes pour générer le graphe états-tâches complet, il préfixe les nœuds de la chaîne par le code du site géographique des fermes (L1, L2...) pour éviter les ambiguïtés. Le pavé gris obtenu, qui représente tous les états et tâches de la ferme, est ensuite connecté automatiquement aux sites de stockage et à la raffinerie par des tâches de transport.

La figure 3.4 montre un exemple où la chaîne "Colza 1" (C1) est utilisée par la ferme du site L5. Un stock partagé est utilisé pour les balles de paille et celles de menue paille. Dans la figure 3.5, la chaîne "Miscanthus" (C2) est employée pour la ferme du site L6. Enfin, la figure 3.6 illustre une ferme (site L13) cultivant à la fois du colza et du miscanthus, ce qui donne 4 produits puisque 3 parties de plante sont considérées pour le colza. Les balles (paille et menue paille de colza, brins de miscanthus) sont rangées dans un stock partagé. On remarque dans tous les cas que chaque plante dans une exploitation est représentée par un état qui regroupe toutes les parcelles où on cultive la plante. En effet, à cause de la rotation des cultures et de la confidentialité, il est impossible de savoir exactement sur quelles parcelles d'une ferme on cultive du colza, par exemple.

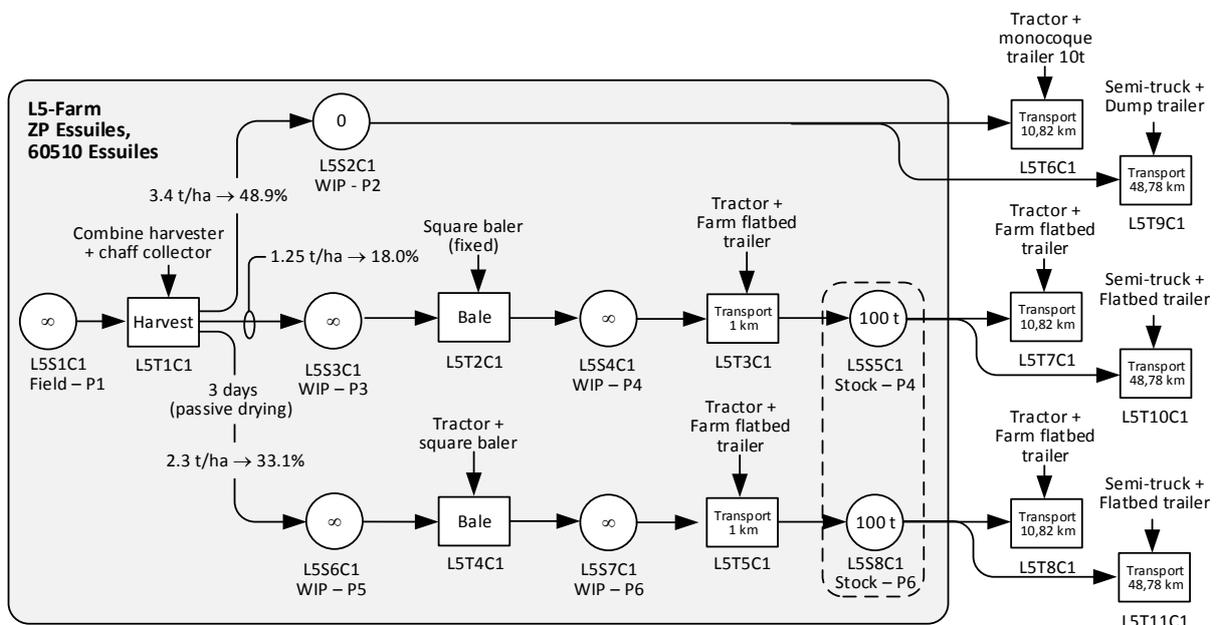


Figure 3.4. Exemple de ferme avec une chaîne de récolte pour le colza (C1).

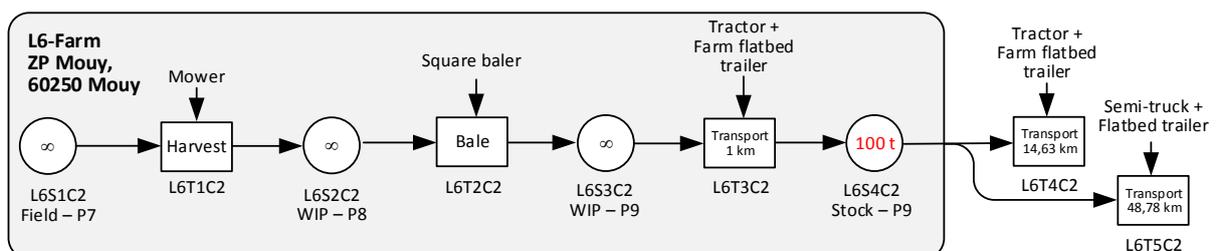


Figure 3.5. Exemple de ferme avec une chaîne de récolte pour le miscanthus (C2).

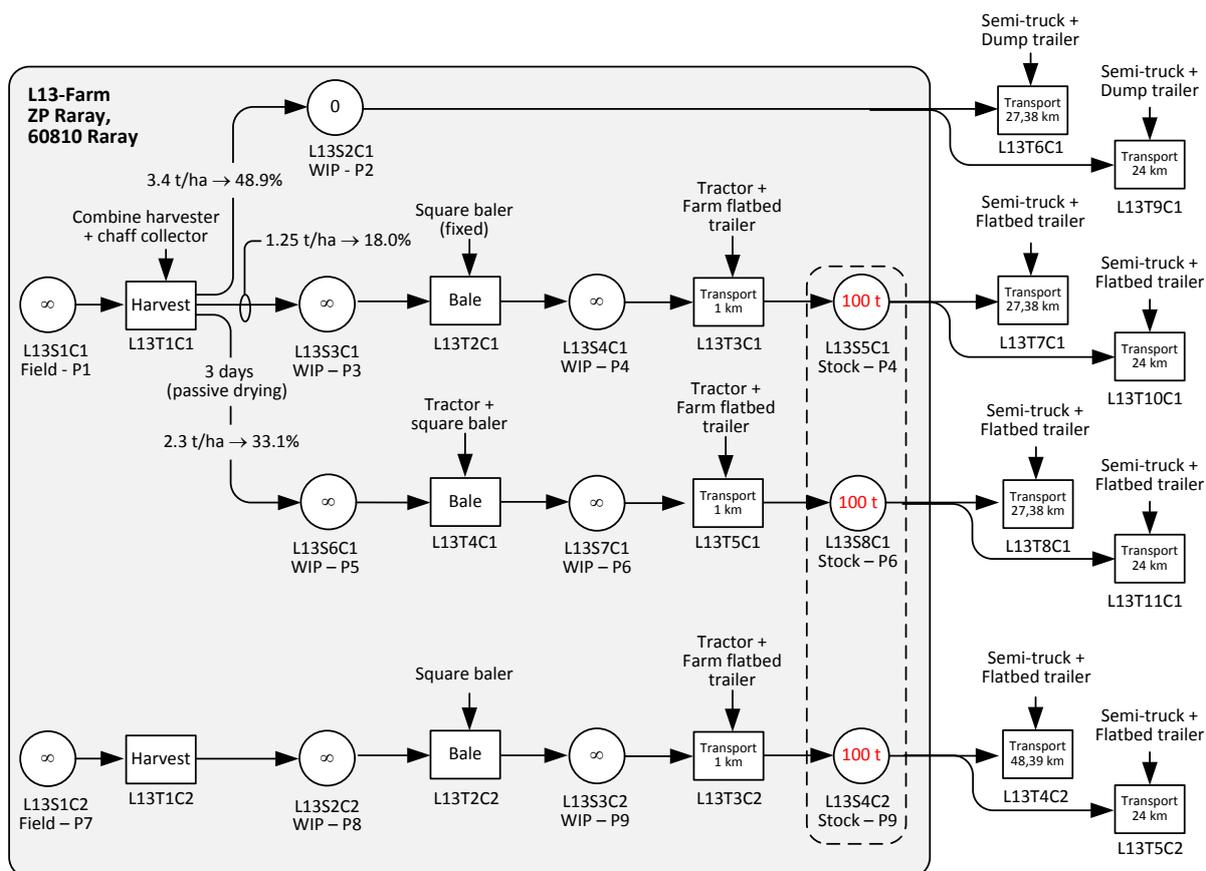


Figure 3.6. Exemple de ferme avec deux plantes ayant chacune leur chaîne de récolte.

3. Typologie des données (data model)

Cette section présente le modèle de données (en anglais *data model*). L'objectif est de recenser et structurer les données utiles pour les calculs, ce qui inclut le codage du modèle de réseau de la section précédente. Nous avons employé le modèle entité-relation qui permet d'aboutir à une décomposition en tables, gérables par un logiciel de bases de données comme Access ou plus simplement Excel. Cette approche définit d'abord les entités (objets réels ou conceptuels), puis leurs attributs (données propres), et enfin les relations qui relient ces entités.

Cette étape est délicate et a demandé beaucoup d'efforts et d'analyses pour aboutir à un modèle conceptuel des données (MCD) cohérent, permettant de gérer plusieurs types de chaînes d'approvisionnement de biomasses. C'est une contribution importante dans la mesure où elle donne lieu à une solide structuration des données, nécessaires aussi bien pour les modèles mathématiques que pour d'autres usages comme réaliser des statistiques, visualiser les résultats dans un système d'information géographique, faire du data mining, etc.

3.1 Règle de formation des noms et liste des abréviations

Les divers noms utilisés pour la typologie des données doivent obéir à des règles de formation pour faciliter la lecture des modèles et la mémorisation. Par convention, tous nos noms sont en anglais.

Les noms d'entités sont en majuscules et au pluriel, par exemple CROPS pour les plantes cultivées. Un nom d'entité sert aussi de nom à la feuille Excel correspondante et à y faire référence quand on charge les données dans des logiciels comme OPL Studio et GUSEK. Il peut être long car il n'est pas utilisé dans les équations du modèle.

Les noms d'attributs servent à nommer les colonnes des tables et les tableaux associés dans le modèle. Ils contiennent la première lettre de l'entité (ou les deux premières en cas d'ambiguïté), suivie de sous-noms débutant chacun par une majuscule pour faciliter la lecture, par exemple CName (*crop name*) pour le nom d'une plante (entité CROPS). Les sous-noms trop longs sont abrégés à 3-4 caractères maximum. Si un sous-nom a déjà une abréviation commune (ID, ZIP), il est en majuscules. Voici la liste des principales abréviations utilisées :

Addr	Address	Inp	Input	N	Number
Beg	Beginning	Inv	Inventory	Out	Output
Dens	Density	Lat	Latitude	Per	Period
Humi	Humidity	Loca	Location	Ref	Refinery
ID	Identifiant	Lon	Longitude	ZIP	Code postal

Quelques données ne peuvent pas être regroupées en entités, on les appelle paramètres généraux. Leurs noms suivent les mêmes règles que pour les attributs (sous-noms) sauf qu'il n'y a pas de nom d'entité au début, par exemple PerDur (*period duration*) pour la durée d'une période en jours.

Certains attributs n'ont que quelques valeurs possibles : un site de la chaîne logistique peut être une ferme, un stockage, un prétraitement ou une raffinerie. Comme ces noms vont être saisis dans une colonne de feuille Excel, les noms de valeurs possibles peuvent être quelconques mais de préférence assez courts. Pour éviter des erreurs, leur saisie peut être faite avec une liste déroulante.

3.2 Décomposition en tables

Le modèle de calcul s'exécutera à partir de données préparées dans des fichiers. Comme nous avons choisi le formalisme des graphes états-tâches, il faut définir un codage informatique du graphe états-tâches qui décrit la chaîne logistique, recenser les données associées, puis les structurer logiquement sous forme de fichiers. Nous proposons d'utiliser Excel, car il facilite la saisie et permet de récupérer des fichiers provenant de différentes origines.

Pour chaque entité (c'est-à-dire un type d'objet comme les états, les tâches, les équipements, etc.), les données nécessaires (appelées aussi attributs ou champs) sont rassemblées dans une table contenue dans une feuille Excel. Chaque colonne correspond à un attribut tandis que chaque ligne correspond à un enregistrement. L'identifiant (ID) ou clé est un attribut important : c'est un code court qui est utilisé comme index par le logiciel, pour retrouver plus rapidement un enregistrement.

Chaque enregistrement doit avoir un ID distinct des autres. Par exemple, les tâches ont pour identifiants T1, T2, T3, etc.

L'ensemble des tables (feuilles Excel) pour un scénario d'optimisation donné sera stocké dans un classeur Excel. L'utilisateur pourra donc évaluer différents scénarios en prenant une copie d'un classeur et en modifiant tout ou une partie des données.

Les tables proposées après analyse sont listées dans la table 3.1. On y trouve évidemment les tables décrivant le graphe états-tâches : une table des états (STATES), une pour les tâches (TASKS), une pour les arcs états-tâches (ST_ARCS) et une pour les arcs tâches-états (TS_ARCS). Les produits sont décrits par trois tables. La table CROPS décrit les plantes et permet d'indiquer des moyennes pour les fenêtres de récolte et les rendements. La table CROP_PARTS décrit les parties de plantes récupérables. Enfin, une partie de plantes ou un mélange de parties peut se présenter vert ou sec, en vrac, en balles etc., ce que nous appelons des "produits" (table PRODUCTS). Les tables DEMANDS à LOCA_FAM décrivent le système de gestion des ressources, décrit plus loin dans la section 4.

Nom de table	Type	Rôle
PARAMETERS	Aucun	Paramètres généraux (nombre de périodes etc.)
LOCATIONS	Entité	Sites géographiques de la chaîne (fermes, stockages...)
CROPS	Entité	Plantes cultivées
CROP_PARTS	Entité	Parties de plantes utilisées (graines, paille, menue paille...)
PRODUCTS	Entité	Parties de plantes sous différentes formes (vrac, balle, granulés)
STATES	Entité	Etats au sens de Kondili et al. (1993) (stocks au sens large)
TASKS	Entité	Tâches au sens de Kondili (récolte, transport...)
ST_ARCS	Relation	Arcs allant d'un état vers une tâche
TS_ARCS	Relation	Arcs allant d'une tâche vers un état
DEMANDS	Relation	Demandes de la raffinerie (par partie de plante et par période)
RESOURCES	Entité	Equipements, exemple une moissonneuse 200 Ch 5m
COMBINATIONS	Entité	Combinaisons de ressources (tracteur 200 Ch + andaineur 5m...)
GROUPS	Entité	Groupe de combinaisons, exemple les (tracteur + remorque)
FAMILIES	Entité	Famille de ressources (les tracteurs, les remorques...)
LOCA_FAM	Relation	Nb max de ressources utilisables par famille sur chaque site
AMOUNTS	Résultat	Quantité traitée par tâche/combinaison de ressources/période.
INVENTORIES	Résultat	Niveau de stock de chaque état par période
ST_FLOWS	Résultat	Flot sur chaque arc état-tâche par période
TS_FLOWS	Résultat	Flot sur chaque arc tâche-état par période
LOCA_RES	Résultat	Nb de ressources de chaque type (fleet size) utilisées par site
TASK_COMB_PER	Résultat	Nb de ressources nécessaires par tâche/combinaison/période

Table 3.1. Liste des tables du modèle de données.

Nous avons aussi ajouté 6 tables (les dernières) pour récupérer les résultats d'une résolution du modèle mathématique. Comme ces tables correspondent simplement aux variables de ce dernier, elles ne sont pas décrites dans ce chapitre.

Cette structuration des données est très flexible, et il est facile d'ajouter des colonnes pour des attributs nouveaux. Les tables sont décrites ci-après. L'identifiant de chaque table est souligné.

3.3 Table des paramètres généraux (PARAMETERS)

Il s'agit d'une petite table à une seule ligne, rassemblant des paramètres généraux :

- PerDur Durée d'une période de base, nombre entier de jours, exemple 1 ou 7.
- NPer Nombre de périodes de l'horizon, exemple 365 si PerDur = 1 ou 52 si PerDur = 7.
- RefPer Durée d'une période-raffinerie, en périodes de base, exemple 7 avec PerDur = 1.
- NA Chaîne spéciale remplaçant les cellules de texte vide ("not available").
- Huge Grande valeur numérique, par exemple pour les stocks de capacité "illimitée".
- KgCO2 Nombre de kg de CO2 rejetés par litre de gazole.

Commentaires :

- Les périodes de l'horizon de planification sont numérotées de 1 à NumPer.
- RefPer a été défini car la raffinerie peut exprimer ses besoins par semaine alors que les opérations de récolte peuvent être détaillées jour par jour. NumPer doit être divisible par RefPer pour avoir un nombre entier de périodes-raffinerie.
- NA et Huge seront utilisés par le logiciel pour initialiser les attributs que l'utilisateur n'a pas saisi dans le fichier Excel. Cette précaution est nécessaire, car des cellule vides, avec un blanc ou avec deux blancs sont considérées comme différentes en informatique.

3.4 Table pour l'entité "sites" (LOCATIONS)

Définition : un site est un lieu géographique contenant un ou plusieurs nœuds du graphe états-tâches de la chaîne, par exemple une exploitation agricole ou un site de stockage avec 4 états (silos). Les sites correspondent aux grands rectangles gris de la figure 3.2. On utilise un nom anglais débutant par "L", puisque "S" est déjà utilisé pour les états (STATES). Les sites servent à pré-calculer une matrice de distances ("distancier"), à faire des regroupements de nœuds, par exemple pour agréger des fermes en communes, et à présenter les résultats par site après résolution du modèle de calcul.

- LID Identifiant numéroté (L1, L2...).
- LName Nom en format libre, exemple "EARL Dupont", "Ferme des Houches".
- LType Type de site, exemple "Farm", "Storage", "Pretreatment", "Refinery".
- LLon Longitude GPS en degrés décimaux, positive à l'est du méridien de Greenwich.
- LLat Latitude GPS en degrés décimaux, positive au nord de l'équateur.
- LAddr Adresse, exemple "17 route de Larcy".
- LZIP Code postal, exemple "10160".
- LCity Commune, exemple "Rigny le Ferron".
- LHours Nombre d'heures d'ouverture du site, par période.

Commentaires :

- Le type Farm doit être interprété au sens large (exploitation agricole, commune, canton...).
- Le type Storage désigne un site de stockage centralisé (hors-ferme), avec états-stocks.

- Le type Pretreatment désigne un site centralisé (hors-ferme) mais faisant des prétraitements.
- Pour limiter les erreurs de distances, utiliser des coordonnées centrales pour les sites étendus.
- Les coordonnées GPS servent à trouver la route la plus proche dans le pré-calcul des distances.
- Pour un site Farm, LHours est le temps de travail sur l'exploitation. Pour un site Storage, c'est le nombre d'heures d'ouverture. LHours s'applique par défaut aux états et tâches à l'intérieur du site. La table STATES présentée plus loin peut spécifier une fenêtre différente pour chaque état.

3.5 Table pour l'entité "plantes" (CROPS)

Définition : il s'agit des plantes cultivées pour produire de la biomasse.

- CID Identifiant numéroté (C1, C2...).
- CName Nom en format libre, exemple "Colza 00", "Miscanthus".
- CBeg Première période de la fenêtre de récolte (valeur moyenne).
- CEnd Dernière période de la fenêtre de récolte (valeur moyenne).

Commentaire :

On peut préciser dans la table STATES (voir plus loin) une fenêtre [SBeg, SEnd], qui correspond à la fenêtre de récolte pour les états de type "Field". La fenêtre [CBeg, CEnd] de CROPS est une moyenne pour la plante, utilisée pour initialiser [SBeg, SEnd] si cette dernière n'est pas renseignée.

3.6 Table pour l'entité "parties de plantes" (CROP_PARTS)

Définition : au moment de la récolte, des plantes comme le colza peuvent être séparées en plusieurs parties de plantes (graines, paille, menue paille). La plante entière peut être considérée comme une partie de plante, avec le type "Whole". L'entité CROP_PARTS doit être introduite car les besoins de la raffinerie sont exprimés en tonnes sèches de parties de plantes par période, peu importe leur forme.

- CPID Identifiant numéroté (CP1, CP2...).
- CPCrop ID de la plante d'origine dans CROP (sert d'index pour la retrouver).
- CPName Nom en format libre, exemple "Colza 00 straw", "Miscanthus".
- CPType Type de partie, exemple "Whole", "Seeds", "Straw", "Chaff".
- CPYield Rendement moyen en t/ha.

Commentaires :

Le rendement correspond à la quantité réellement récupérable, après déductions des pertes à la récolte. Il peut varier selon les fermes (nature du sol, microclimat) et on peut le préciser dans la table des arcs tâche-état vue plus loin (TS_ARCS), pour les arcs en sortie des tâches de moisson. CPYield s'appliquera par défaut aux arcs sans rendement mentionné.

3.7 Table pour l'entité "produits" (PRODUCTS)

Définition : on appelle "produit" une biomasse définie par une plante, une partie de cette plante, une forme (conditionnement), une densité et une humidité. Le changement d'au moins un paramètre en fait un produit différent : ainsi, de la paille de colza fraîche devient de la paille sèche après 3 jours. Dans les graphes états-tâches, rappelons qu'on a un seul produit sur chaque arc et dans chaque état.

- PID Identifiant numéroté (P1, P2...).
- PPart ID de la partie de plante associée dans CROP_PARTS (sert d'index pour la retrouver).
- PName Nom en format libre, exemple "Rape straw in square bales".
- PForm Forme, exemple "Standing plant", "Bulk", "Square bales", "Round bales".
- PDens Densité
- PHumi Humidité en % du poids.

Commentaires :

- "Standing plant" désigne la plante sur pied, non encore récoltée.
- "Bulk" (vrac) concerne les produits après récolte et non conditionnés (graines, paille au sol).
- PDens sert pour la contrainte de volume maximum des véhicules.
- PHumi sert pour respecter les demandes de la raffinerie (exprimées en tonnes sèches).

Dans ce document, nous ne détaillons pas les relations entre les tables car il s'agit d'un aspect délicat de l'implémentation informatique. Ainsi, dans les tables CROPS, CROP_PARTS et PRODUCTS, une plante peut avoir plusieurs parties et chaque partie peut se présenter sous forme de différents produits. En fait, le futur logiciel utilisera les index CPCrop et PPart pour établir ces relations.

3.8 Table pour l'entité "états" (STATES)

Définition : un état au sens de Kondili et al. est une accumulation d'un produit en attente, même très courte. Il est symbolisé par un cercle dans le modèle de réseau, cf. l'exemple de la figure 3.2. Dans cette thèse, les états incluent les vrais stocks (hangars de ferme, silos et plates-formes des sites de stockage centralisés, stocks tampons de la raffinerie), mais aussi les champs de plantes sur pied et les produits en attente sur l'exploitation (paille et menue paille sur andain, meules, balles en attente de chargement en bout de champ). Cette versatilité des états simplifie grandement la modélisation.

- SID Identifiant numéroté S1, S2...
- SName Nom en format libre, exemple "Rape field", "Straw on swath", "Silo of rape seeds".
- SType Type d'état, exemple "Field", "WIP", "Stock".
- SLoca ID du site de rattachement dans LOCATIONS (sert index pour le retrouver).
- SProd ID du produit stocké dans PRODUCTS (sert d'index pour le retrouver).
- SBeg Première période de la fenêtre d'utilisation.
- SEnd Dernière période de la fenêtre d'utilisation.
- SCapa Capacité en tonnes pour le type Stock.
- SArea Surface allouée à l'état, en ha.

- SQBeg Quantité initiale dans le stock, en tonnes (\leq SCapa).
- SQEndMin Quantité finale minimale requise dans le stock, en tonnes (\leq SCapa).
- SQEndMax Quantité finale maximale requise dans le stock, en tonnes (\geq SQEndMin, \leq SCapa).
- SInvCost Coût de stockage (*inventory cost*) en €/t et par période (type Stock).
- SInpCost Coût d'entrée en €/t et par période (type Stock).
- SOutCost Coût de sortie en €/t et par période (type Stock).
- SThru Débit maximal en entrée, en tonnes par heure (type Stock).
- SLoss Perte liée aux dégradations, par période et en % (type Stock).
- SLink Nœud-représentant pour un stock de plusieurs produits (cf. explications plus loin).

Commentaires :

- Pour chaque type d'état on n'aura à saisir que les informations pertinentes. Par exemple, une zone de production de colza est modélisée par un état car elle est vue comme une sorte de stock de plante sur pied, mais on n'a pas besoin de saisir un coût forfaitaire de sortie, car le coût de récolte va être calculé à partir des temps d'utilisation des moissonneuses-batteuses, voir section 4.
- Le type "Field" regroupe en fait tous les champs d'une même plante dans un même site Farm, puisqu'on ne descend pas au niveau parcelle (l'utilisation de telle ou telle parcelle pour cultiver telle ou telle plante une certaine année est une donnée indisponible).
- Le type "Stock" correspond aux vrais stocks (hangar de ferme, silo ou plate-forme de stock centralisé), à ne pas confondre avec le type de site "Storage" dans l'entité LOCATIONS. Une étude effectuée par Coopénergie a permis de recenser les sites disponibles autour de Compiègne.
- Le type "WIP" (*work in progress*) concerne les en-cours de produits dans les champs d'une ferme (paille séchant au sol, tas de menue paille, balles en attente d'expédition en bord de champ).
- La surface est obligatoire pour le type "Field", elle indique la surface cultivée en ha. Avec les rendements des parties de plantes en sortie de la tâche Harvest (voir les tâches plus loin), elle sert à déduire automatiquement le stock de plante sur pied et les fractions des différentes parties.
- Surface et volume peuvent être aussi indiqués pour des stocks de type plate-forme, hangar, silo...
- Pour un état "Field", le stock initial SQBeg est la quantité récoltable. SQBeg et SQEndMin permettent aussi à la BR d'enchaîner les années sans s'arrêter. En effet, si SQBeg = SQEndMin = 0 pour les stocks de la BR, la minimisation des coûts de stockage peut vider les stocks en fin d'année et on doit attendre une récolte pour redémarrer la BR l'année suivante.
- SQEndMax, valant Huge par défaut, permet par exemple de vider un stock. Ainsi, en mettant SQEndMax à zéro, on oblige à récolter un champ complètement ou à vider un stockage à la fin de sa période d'ouverture saisonnière.
- Les derniers attributs concernent surtout les états de type "Stock". SInpCost, SOutCost et SThru s'appliquent au flot total qui entre ou sort de l'état dans une période. Ils permettent de ne pas détailler les tâches de manutention et les équipements associés, grâce à un coût forfaitaire par tonne. SInvCost et SLoss s'appliquent à la quantité constatée en stock à la fin de chaque période.

- SLink permet de gérer des stocks de plusieurs produits, non prévus par Kondili et al. Un stockage de type plate-forme peut mixer des balles de différentes plantes. Chaque produit doit être codé par un état mais on peut indiquer qu'un groupe d'états forme un même stockage : on choisit pour cela un représentant e du groupe et on pose $SLink = e$ pour tous les états du groupe. Pour un état e sans $SLink$ renseigné, le logiciel vérifiera que le stock ne dépasse pas $SCapa_e$. Si on a plusieurs nœuds avec $SLink = e$, il vérifiera que la somme de leurs stocks ne dépasse pas $SCapa_e$. La figure 3.2 montre deux exemples de stocks partagés.
- Les fenêtres [SBeg, SEnd] sont égales par défaut à tout l'horizon : des produits peuvent entrer ou sortir de l'état dans chaque période. Pour un état Field, il s'agit de la fenêtre de récolte. Ainsi, sur un horizon de $52 \times 7 = 364$ jours, la récolte du colza dure 15 jours à partir du 1^{er} juillet, ce qui correspond aux semaines 27-28 et à la fenêtre [183,196]. Les fenêtres des autres états de la ferme ont le même début que la fenêtre de récolte. Comme un stockage de longue durée à la ferme n'est pas souhaitable, leur fenêtre se termine souvent 4 semaines après la fenêtre de récolte. Ceci force le modèle de calcul à évacuer les biomasses vers la raffinerie ou des sites de stockage.

Pré-calculs à effectuer :

- Pour le type Field, récupérer les rendements en t/ha sur les arcs sortant de la tâche Harvest et déduire SQBeg. Comme on peut le voir dans la figure 3.2, pour 22 ha de colza, si la tâche Harvest récupère 3.4 t/ha de graines, 2.3 t/ha de paille humide et 1.25 t de menue paille, le stock initial est $22 \times (3.4 + 2.3 + 1,25) = 152.9$ tonnes.

3.9 Table pour l'entité "tâches" (TASKS)

Définition : une tâche au sens de Kondili et al. est un processus qui transforme des produits d'entrée en proportions données, en produits de sortie également en proportions fixées. Elle est symbolisée par un rectangle dans le modèle de réseau (cf. l'exemple de la figure 3.2). Les tâches utilisent en général des ressources, indiquées au-dessus de la tâche, sauf certaines comme le séchage passif.

- TID Identifiant numéroté T1, T2...
- TName Nom en format libre, exemple "Passive drying", "Transport", "Square baling".
- TType Type, exemple "Harvest", "Dry", "Rake", "Bale", "Pretreat", "Transport", "Refine".
- TLoca ID du site géographique dans LOCATIONS (sert d'index pour le retrouver).
- TDist Distance en km, uniquement pour les tâches de transport.
- TGroup ID du groupe de combinaison de ressources nécessaires, voir section 4.

Commentaires :

- Comme une tâche de transport a un arc entrant et un arc sortant, on préfère associer TDist à la tâche. Cette distance peut être pré-calculée avec un logiciel comme MapPoint (Microsoft) pour les tâches de transport entre deux sites distincts. Ce calcul utilise les coordonnées des sites si elles sont données (méthode la plus précise), sinon l'adresse. Pour le transport dans une ferme (champ vers hangar par exemple), la distance est nulle ou contient une petite valeur, par exemple 1 km.

- Alors qu'un état est dans un site, une tâche peut être interne à un site (tâche "Harvest") ou relier deux sites (tâches de transport). Pour gérer les ressources, il faut savoir si un transport est fait par le fermier ou des prestataires. En général, le fermier effectue en tracteur les transports dans sa ferme et de sa ferme aux stockages externes proches. Les transports ferme-raffinerie et stockage-raffinerie se font par camions. La distinction se fait avec TLoca : les tâches de transport dans les fermes ou entre une ferme et un site de stockage ont pour TLoca le site de la ferme, les autres transports étant rattachés à la raffinerie ou au site d'un transporteur, où sont basés les véhicules.
- Chaque plante dans chaque ferme a une chaîne de récolte imposée, dans laquelle chaque tâche doit utiliser un seul groupe de combinaison de ressources TGroup (par exemple, un tracteur + une presse à balles carrées). Voir la section 4 pour des explications sur la gestion des ressources.
- Il va falloir calculer un distancier entre les sites. Pour une chaîne réelle de 500 fermes et 40 stockages par exemple, on a 2000 distances à calculer, sans compter celles entre les stockages et la raffinerie. Calculer les distances une par une avec un site Internet comme Mappy prend trop de temps. Nous avons donc choisi au début le logiciel Map Point de Microsoft qui inclut une base de données des routes de France et permet de calculer en VBA des distanciers. Hélas, Microsoft a annoncé l'abandon de Map Point fin décembre 2014.

Il existe des solutions de substitution mais la plupart nécessitent d'être connecté à Internet : Maptitude, Map Quest, Google Maps, Maptive, Portatour et Map Business. Seul Map Quest est actuellement 100% gratuit mais il fonctionne uniquement en ligne. Maptitude est le seul qui offre le choix entre un fonctionnement sur PC, après chargement des données cartographiques, et un fonctionnement en ligne. Il est payant mais le prix annuel est raisonnable, autour de 700 dollars. Google Maps est particulier : il est gratuit jusqu'à 2500 requêtes journalières de calcul de distances, ensuite l'accès se bloque et il faut souscrire un abonnement pour faire plus de calculs.

3.10 Table pour la relation "arcs état-tâche" ou "state-task arcs" (ST_ARCS)

Définition. Dans le modèle de réseau, un tel arc est illustré par une flèche allant d'un état à une tâche, voir la figure 3.2. Il modélise un flux de produit au sens large, prélevé sur l'état par la tâche. Dans le cas d'une tâche de transport, on peut le voir comme un chargement d'un produit.

- SID ID de l'état, qui doit être dans STATES.
- TID ID de la tâche, qui doit être dans TASKS.
- STFrac Ratio entre le flot sur l'arc et le flot total consommé par la tâche, en % (fraction du flux consommé par la tâche TID, venant de l'état SID).
- STYield Rendement en tonnes/hectare (arcs sortants des états de type Field).

Commentaires

- Dans le modèle de calcul, un tel arc est identifié par son couple (SID, TID).
- Dans notre réseau, un état a toujours des tâches après lui, car le réseau finit par la tâche *Refine*.
- Le produit sur l'arc est celui qui est stocké dans l'état au début de l'arc.

- Dans la figure 3.2, chaque tâche a un seul arc entrant, sauf la tâche-raffinerie. A part cette dernière, STFrac vaut donc 100%. On pourra spécifier des fractions différentes pour de futures tâches de mélange, par exemple faire des briquettes avec 60% de paille et 40% de menue-paille.

3.11 Table pour la relation "arcs tâche-état" ou "task-state arcs" (TS_ARCS)

Définition. Dans le modèle graphique de réseau, un tel arc est représenté par une flèche allant d'une tâche à un état. Il modélise l'élaboration d'un nouveau produit par la tâche, au sens large. Dans le cas d'une tâche de transport, on peut le voir comme la livraison d'un produit.

- TID ID de la tâche, qui doit être dans TASKS.
- SID ID de l'état, qui doit être dans STATES.
- TSFrac Ratio entre le flux sur l'arc et le flux total en entrée de tâche en % (fraction du flux d'entrée de la tâche TID, allant en sortie vers l'état SID).
- TSYield Rendement en tonnes/hectare (arcs sortants des tâches Harvest).
- TSDura Délai d'obtention du produit à partir des entrées de tâche, en nombre de périodes.

Commentaires

- Dans le modèle de calcul, un tel arc est identifié par son couple (TID, SID).
- Le produit sur l'arc est celui qui est stocké dans l'état de destination de l'arc.
- Dans l'exemple de la page 9, chaque tâche a un seul arc sortant, sauf les tâches Harvest. A part ces dernières, TSFrac vaut moins de 100% seulement quand la tâche induit une perte de masse, comme le séchage. Dans l'exemple, nous aurions pu modéliser le séchage passif de la paille de colza (sur andain) par une tâche. Nous avons préféré dire que la tâche Harvest produit directement de la paille sèche, mais avec un délai de 3 jours. Cette modélisation est équivalente.
- Pour chaque arc sortant d'une tâche Harvest, TSFrac est calculable par le logiciel à partir des rendements et de la surface du champ : il faut sommer les rendements en t/ha sur les arcs sortants de la tâche, multiplier par la surface de l'état Field situé avant la tâche, mettre le résultat dans SQBeg (attribut de l'état Field), puis calculer finalement $TSFrac = TSYield \times SArea / SQBeg$.

3.12 Table pour la relation "besoins de la raffinerie" (DEMANDS)

Définition. La table PARAMETERS décrite au début définit une période de base PerDur (en jours), servant pour planifier la récolte et le transport, et une période RefPer souvent plus longue, avec laquelle la raffinerie exprime ses besoins en biomasse. Par exemple, on peut avoir PerDur = 1 jour et RefPer = 7 jours si les besoins sont donnés par semaine. Dans chaque période-raffinerie, les besoins sont exprimés en tonnes de matière sèche par plante et partie de plante, par exemple 100 t de paille de colza sèche en semaine 2. Pour satisfaire une demande en tonnes sèches pour une plante et une partie de plante, la raffinerie devra donc puiser dans ses stocks d'entrée différentes formes livrées (par exemple des balles et des granulés) et tenir compte de leurs humidités respectives.

- DRefPer Numéro de la période-raffinerie, entre 1 et NumPer/RefPer.
- DPart ID de la partie de plante (sert d'index pour retrouver la partie dans CROP_PARTS).
- DNeed Demande en tonnes sèches pour la partie de plante DPart en période DRefPer.

Commentaires :

- Les demandes sont donc définies par des triplets (période raffinerie, partie de plante, quantité). Ce système prend peu de place car certaines plantes ne sont pas utilisées toute l'année : on peut donc saisir seulement les périodes où la demande est non nulle.
- Vérifier que DRefPer est entre 1 et NumPer/RefPer (inclus).
- Vérifier que DPart existe dans CROP_PARTS.
- Mettre à 0 les demandes non renseignées. Attention : la raffinerie ne produit rien dans les périodes où elle n'a pas de demande, mais on peut amener des produits dans ses stocks!
- La raffinerie doit avoir au moins un stock-tampon de produit pour chaque partie demandée.

Codage de DNeed dans le modèle mathématique :

La raffinerie peut ne pas utiliser certaines plantes toute l'année. Donc, il vaut mieux saisir dans Excel uniquement les couples (DRefPer, DPart) à demandes non nulles, puis les charger dans un sous-ensemble DPairs \subseteq HR \times CPID. HR = 1...NumPer/RefPer désigne l'horizon de la raffinerie.

Synchronisation entre la période de la chaîne logistique et celle de la raffinerie :

La période-raffinerie, par exemple une semaine, correspond souvent à plusieurs périodes de base, par exemple des jours. La question est de savoir quand dans la semaine la raffinerie sort de ses stocks son besoin pour une partie de plante donnée. Les deux extrêmes, en début ou en fin de semaine, sont peu satisfaisants. Une sortie du besoin complet en début de semaine exige d'avoir des stocks plus élevés. Une sortie en fin de semaine ne laisse pas assez de temps pour fabriquer les produits prévus. Une hypothèse moyenne est donc que la raffinerie consomme uniformément son besoin sur chaque période de base qui compose la période-raffinerie.

On peut vérifier sur la figure 3.7 que la période-raffinerie k correspond aux périodes de base $(k - 1) \times RefPer + 1$ à $k \times RefPer$. Pour une partie de plante p , on suppose donc que la demande de la raffinerie dans chacune de ces périodes de base est égale à $DNeed[k, p]/RefPer$. Le flot total sortant des stocks contenant la partie de plante p devra être égal à cette demande (et pas supérieur, sous peine de consommer plus que nécessaire pour éviter les coûts de stockage).

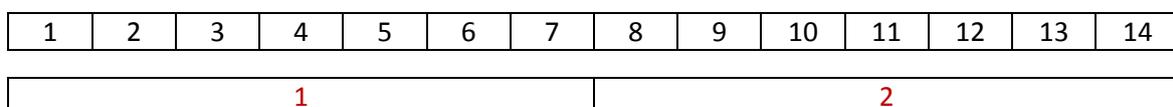


Figure 3.7. Relation entre les périodes de base de l'horizon (en haut) et celles de la raffinerie.

4. Système de gestion des ressources

4.1 Principes

Notre système de gestion des ressources est flexible mais nécessite quelques explications. On appelle ressources ou équipements des engins à moteur ou accessoires utilisés par les tâches dans un graphe états-tâches. Dans la suite, on fait la distinction entre la ressource proprement dite, qui fait référence à un matériel avec des caractéristiques précises (un tracteur 200 ch) et la famille de ressources à laquelle elle appartient (la famille des tracteurs). En général, on ne précisera pas les modèles commerciaux (Claas, New Holland...) car les constructeurs ont des modèles similaires. Les ressources utilisables sont décrites dans la suite dans une table RESOURCES.

Les familles de ressources incluent les faucheuses, les moissonneuses batteuses, les andaineurs, les presses à balles, les camions, les remorques agricoles, les remorques routières, etc. Elles sont stockées dans une table FAMILIES. On a besoin des familles de ressources car notre modèle permet de limiter le nombre d'équipements sur chaque site, par exemple pas plus de 2 tracteurs dans une ferme. Comme le modèle peut panacher les modèles de tracteurs (deux de 200 ch ou un de 200 ch + un de 300 ch, etc.), la limite doit être précisée pour chaque site et chaque famille de ressources.

Les combinaisons de ressources sont des ensembles de ressources qu'on peut associer pour une tâche, par exemple un tracteur de 200 ch + une remorque-plateau de 15 tonnes. Les combinaisons sont nécessaires pour calculer correctement les temps de travail des différentes ressources. En effet, la productivité d'une ressource dépend de la combinaison qui l'utilise. Par exemple, un tracteur seul ou équipé d'un outil porté d'au plus 2,55 mètres de large peut rouler à 40 km/h. Cette vitesse chute à 25 km/h si le tracteur a un outil de plus 2,55 mètres ou s'il tracte une remorque (source : FDSEA 57).

Les combinaisons utilisent les mêmes attributs que les ressources : productivité en ha/h ou en balles/h pour les équipements de récolte, vitesse maximale et charge utile pour les camions, etc. A cause des nombreux cas particuliers, le logiciel ne peut pas déduire facilement la productivité d'une combinaison à partir des productivités des ressources qui la composent. Une personne compétente en machinisme agricole devra donc préciser quelles combinaisons sont possibles à partir des équipements de la table RESOURCES, et lister ces combinaisons dans une table COMBINATIONS.

Enfin, les groupes de combinaisons de ressources ou combinaisons génériques sont aux combinaisons ce que les familles de ressources sont aux ressources : les caractéristiques exactes ne sont pas spécifiées. Par exemple, un tracteur de 200 ch + une remorque-plateau de 15 tonnes constitue une combinaison appartenant au groupe de combinaisons "tracteur + remorque". Ce groupe rassemble plusieurs combinaisons qui précisent les modèles de tracteur et de remorque utilisés. Les groupes sont décrits dans une table GROUPS. Ils sont utiles pour limiter le nombre de variables du modèle. En effet, on doit préciser la chaîne de récolte utilisée pour chaque ferme et chaque plante cultivée. Dans une chaîne donnée, chaque tâche ne peut utiliser qu'un seul groupe de combinaisons. Par exemple une tâche de pressage de paille doit utiliser un tracteur et une presse à balles.

Les sections suivantes décrivent les tables du système de gestion de ressources. Il sera très facile d'ajouter de nouveaux équipements et les combinaisons auxquelles ils peuvent participer.

Hypothèses importantes :

- Pour réduire les familles d'équipements nécessaires, chaque ferme n'utilise qu'une chaîne de récolte pour chaque plante cultivée. Par exemple, la section 2.3 décrit trois chaînes possibles pour le miscanthus : une seule devra être choisie pour chaque ferme produisant cette plante.
- Dans une chaîne de récolte, chaque tâche utilise un seul groupe de combinaisons de ressources (attribut TGroup de la table TASKS), par exemple le groupe "tracteur + presse à balles carrées".
- Pour chaque tâche, le module de calcul peut par contre utiliser une ou plusieurs combinaisons de ressources dans le groupe spécifié, et plusieurs exemplaires de chaque combinaison. Connaissant les ressources dans chaque combinaison, on peut ainsi cumuler les heures de travail de chaque équipement (une même ressource pouvant travailler sur plusieurs tâches et plusieurs combinaisons) et en déduire les coûts et les tailles de flotte.

4.2 Table pour l'entité "ressources" (RESOURCES)

Définition : les ressources ou équipements sont des engins à moteur (camions, tracteurs, moissonneuses) ou des accessoires (presses à balles, andaineurs, remorques), pour traiter les tâches dans le modèle de réseau. Les types d'équipements et leurs attributs sont ceux indiqués dans les barèmes d'entraide, par exemple "tracteur 300 ch".

- **RID** Identifiant numéroté (R1, R2...).
- **RName** Nom en format libre, exemple "Combine harvester 190 ch 5m".
- **RFam** ID de la famille de ressources à laquelle appartient l'équipement.
- **RExam** Exemple de modèle commercial (optionnel), exemple "New Holland CR9080".
- **ROCost** Coût d'utilisation ou d'exploitation (*operating cost*) de la ressource en €/h.
- **RCCost** Coût fixe annuel (*capital cost*) de la ressource en €.
- **RPrice** Prix d'achat ou coût d'acquisition de la ressource en €.
- **RRate** Taux d'amortissement en %.
- **RLife** Durée d'amortissement en années.
- **RThru** Quantité maximum de biomasse traitée par heure (débit), cf. commentaires.
- **RFuel** Quantité de carburant consommée par heure d'utilisation, en l/h.
- **RPower** Puissance en kW pour les équipements motorisés.
- **RWTime** Durée maximum de travail par période, en heures.
- **RSpeed** Vitesse moyenne d'utilisation de la ressource, en km/h.
- **RWeight** Poids maximal de la charge utile, en tonnes.
- **RVolume** Volume maximum de la charge utile, en m³.

Commentaires :

- Le coût fixe peut être pré-calculé à partir des trois attributs RPrice, RRate, RLife.
- L'attribut RThru concerne seulement les équipements de récolte. Il est donné en ha/h pour les ressources traitant les tâches de type Harvest et Rake (moissonneuses et andaineurs) et en balles/h pour les ressources traitant les tâches de type Bale (presses à balles). Le logiciel d'optimisation se ramène à des tonnes/heures pour traiter les flux de manière homogène.

- RSpeed concerne les engins à moteur pouvant faire du transport (camions, tracteurs). RWeight et RVolume s'appliquent aux camions porteurs (= sans remorques) et aux remorques agricoles ou routières. RVolume est défini aussi pour les presses à balles (volume d'une balle).
- RFuel et RPower concernent les engins à moteur. Ils servent à calculer les consommations d'énergie et les émissions de GES.

4.3 Table pour l'entité "familles de ressources" (FAMILIES)

Définition. Une famille de ressources est un groupe d'équipements ayant la même fonction, sans préciser le modèle, par exemple la famille des tracteurs et celle des remorques. La notion de famille sert à limiter le nombre d'équipements sur chaque site.

- RFID Identifiant numéroté (F1, F2...).
- RFName Nom en format libre, par exemple "Balers".

4.4 Table pour la relation "limitations des ressources" (LOCA_FAM)

Définition. Cette table permet de limiter le nombre de ressources de chaque famille sur chaque site de la chaîne logistique, par exemple 3 tracteurs au maximum dans une certaine exploitation agricole. Cette précaution évite certains résultats aberrants, quand par exemple le modèle de calcul mobilise un trop grand nombre d'équipements pour traiter le plus rapidement possible une zone de récolte.

- LID Identifiant du site.
- RFID Identifiant de la famille.
- LFNMax Nombre maximum de ressources utilisables pour cette famille sur ce site.

4.5 Table pour l'entité "combinaisons de ressources" (COMBINATIONS)

Définition. Une combinaison de ressources est un ensemble d'équipements qu'on associe pour effectuer une tâche, par exemple un tracteur avec une presse à balles. Une combinaison peut éventuellement avoir une seule ressource, comme un automoteur. Les attributs d'une combinaison reprennent certains attributs des ressources qui la composent, mais les noms d'attributs commencent par le code RC ("resource combination") au lieu de R.

- RCID Identifiant numéroté (RC1, RC2...).
- RCName Nom en format libre, exemple "Tracteur 240 ch + remorque plateau 12 t".
- RCGroup ID du groupe de combinaisons auquel appartient la combinaison.
- RCThru Quantité max. de biomasse traitée par unité de temps (débit), cf. commentaires.
- RCFuel Quantité de carburant consommée en litres par heure d'utilisation.
- RCPower Puissance en kW pour les équipements motorisés.
- RCWTime Durée maximum du temps de travail par période, en heures.
- RCSpeed Vitesse moyenne de la ressource en km/h.
- RCWeight Poids maximal de la charge utile en t.
- RCVolume Volume de charge maximal de la combinaison en m³.

Commentaires :

- Nous ne détaillons pas dans ce document toutes les relations entre tables. Disons simplement que le modèle connaît les ressources qui composent chaque combinaison et l'ensemble des combinaisons auxquelles une ressource peut participer.
- Les attributs vitesse, poids et volume de la combinaison peuvent être aussi déduits à partir des attributs des ressources, mais il existe des exceptions comme la vitesse des tracteurs, mentionnée dans la section 4.1. C'est pourquoi nous conseillons de saisir manuellement les combinaisons possibles et leurs attributs.
- Le modèle mathématique est introduit dans le chapitre 4 suivant, mais on peut déjà expliquer le calcul des ressources nécessaires avec les attributs décrits précédemment. Le modèle utilise des variables Amp_{ict} (*amount processed*), désignant la quantité traitée par la tâche i à l'aide de la combinaison de ressources c en période t . Si la tâche i est une tâche Harvest sur un état Field e , la combinaison c contient une moissonneuse batteuse et son débit maximum en t/h est $RCThru(c) \times STYield(e,i)$. Pour une tâche de pressage, la productivité des combinaisons "tracteur + presse à balles" est donnée en balles par heure. On se ramène à un débit en tonnes par heure en calculant $RCThru(c) \times RCVolume(c) \times PDens(STProd(e,i))$.

En divisant Amp_{ict} par le débit obtenu, on obtient le nombre d'heures de travail de la combinaison c , qui s'applique aussi à chaque ressource composant la combinaison. En faisant la somme des heures de travail par ressource (une ressource pouvant participer à plusieurs combinaisons) et connaissant le temps de travail maximal journalier de la ressource, on peut déduire le nombre minimal d'exemplaires de chaque ressource à mettre en œuvre chaque jour sur chaque exploitation.

Enfin, en calculant le maximum des nombres d'exemplaires de chaque ressource sur l'ensemble des périodes de l'horizon, on obtient le nombre de ressources qui doivent être disponibles dans chaque exploitation.

4.6 Table pour l'entité "groupes de combinaisons de ressources" (GROUPS)

Définition. Un groupe de combinaisons rassemble les combinaisons qui assurent la même fonction mais sans préciser les caractéristiques exactes des ressources qui la composent. Par exemple, le groupe de combinaisons "tracteur + remorque agricole" peut contenir des combinaisons comme "tracteur 200 ch + remorque-plateau 12 t" ou "tracteur 150 ch + remorque-benne 8 t". Pour chaque tâche d'une chaîne de récolte, on devra préciser le groupe de combinaisons à utiliser (attribut TGroup de la table TASKS). Par contre, le modèle choisira librement une ou plusieurs combinaisons dans ce groupe, en fonction de la quantité à traiter.

- CGID Identifiant numéroté (CG1, CG2...).
- CGName Nom en format libre, exemple "Tracteur + remorque agricole".

4.7 Génération automatique de chaînes de récolte (CHAINS)

Le module de calcul puise ses données dans un classeur Excel incluant toutes les tables précédentes, en particulier les tables STATES et TASKS contiennent tous les états et tâches des différentes fermes. Ce classeur est appelé classeur complet dans la suite. Nous décrivons ici un système assez simple pour éviter la saisie des nombreuses fermes ayant la même chaîne de récolte. Grosso modo, on saisit pour chaque ferme un code de chaîne de récolte au lieu de décrire tous les états et les tâches, ce qui donne un classeur compacté. Un programme VBA génère ensuite automatiquement le classeur complet à partir du classeur compacté. Ce système est déjà implémenté.

Définition. Dans notre modèle de réseau, une chaîne de récolte (harvesting chain) est un sous-graphe (états, tâches et arcs) souvent utilisé par des fermes pour récolter une plante donnée. On considère seulement les tâches dont les états en entrée/sortie sont dans la ferme. De même, les arcs décrits dans le sous-graphe doivent avoir leurs deux extrémités (état et tâche) dans la ferme. Ce sous-graphe a des états sans successeurs car les tâches de transport vers d'autres sites ne sont pas incluses. Des exemples ont été donnés dans les figures 3.4 à 3.6.

Restrictions. Une ferme ne peut utiliser qu'une chaîne de récolte par plante cultivée. En effet, utiliser deux chaînes différentes pour du colza paraît aberrant en termes de ressources nécessaires. De plus, une chaîne de récolte doit utiliser les mêmes catégories d'équipements quelle que soit la ferme. Si deux chaînes de récolte ont le même sous-graphe mais utilisent des catégories d'équipements différents, elles sont considérées comme deux chaînes distinctes.

Implémentation

Chaque chaîne de récolte est stockée dans un classeur Excel CxHy.xlsx, où Cx est l'ID de la plante et Hy le numéro de chaîne pour une plante. Par exemple, le colza (C1) peut avoir des chaînes de récolte C1H1, C1H2 etc. Par rapport à un classeur complet, celui pour une chaîne de récolte contient seulement les feuilles STATES, TASKS, ST_ARCS et TS_ARCS. Ces feuilles ne contiennent que des états, tâches et arcs internes à une ferme qui utilise la chaîne. Les états ont des identifiants S1, S2... tandis que les tâches utilisent T1, T2... L'attribut SLoca de STATES (site géographique) n'est pas utilisé. Les ID des produits doivent être ceux du scénario qui fait référence à la chaîne.

Dans un classeur compacté pour un scénario d'optimisation, les états, les tâches et les arcs internes aux fermes ne sont pas saisis dans les tables STATES, TASKS, ST_ARCS et TS_ARCS. Ceci divise typiquement par 10 la taille du classeur car il y a une seule raffinerie et peu de sites de stockage par rapport au nombre élevé de fermes. Par contre, le classeur a une feuille supplémentaire CHAINS qui donne une liste de tuples (ID d'un site de type Farm, ID de la plante cultivée, surface cultivée, nom du classeur contenant la chaîne à utiliser).

Le programme VBA prend en entrée le classeur compacté et génère un classeur complet. Pour cela, il inspecte chaque tuple de CHAINS. Pour chacun d'entre eux, il ouvre le classeur décrivant la chaîne de récolte et copie les lignes de STATES, TASKS, ST_ARCS et TS_ARCS dans les tables de même nom du classeur complet. Pour éviter les collisions de noms, les nœuds sont préfixés par l'ID de la ferme et l'ID de la plante. Par exemple, l'état S1 de la chaîne C1H1 (type Field) devient L1C1S1 pour une ferme dont l'ID est L1. C'est ce qu'on observe dans les figures 3.4 à 3.6.

Ensuite, les états générés qui n'ont pas de successeurs sont reliés par des tâches de transport aux sites de stockage et/ou à la raffinerie. Les états de destination doivent évidemment avoir le même produit que l'état de départ. Pour éviter trop d'arcs, on peut préciser des conditions, par exemple relier la ferme aux sites de stockage situés à 10 km au maximum ou uniquement au plus proche. L'utilisateur peut évidemment modifier le classeur obtenu, s'il veut introduire quelques différences entre les fermes utilisant une même chaîne de récolte, par exemple des rendements différents.

4.8 Options de gestion des équipements

On a principalement quatre possibilités pour gérer les ressources à utiliser :

- Ressources partagées par tout le monde (appartenant à la raffinerie ou à une coopérative de producteurs) et non encore achetées, donc aucune limite pour le module d'optimisation.
- Ressources partagées mais déjà achetées. Il faut alors préciser la liste des ressources achetées et le nombre d'exemplaires disponibles pour chaque ressource.
- Comme Shastri et al. (2011a), chaque site a ses propres ressources mais elles ne sont pas encore achetées (le modèle va calculer les besoins). Pour des raisons de main d'œuvre par exemple, on limite cependant le nombre de ressources utilisables dans chaque famille, pour chaque site.
- Chaque site a ses propres ressources et elles sont déjà achetées. Il faut donc préciser pour chaque ferme la liste des ressources disponibles et le nombre d'exemplaires disponibles.

Notre modèle de données s'aligne par défaut sur Shastri et al. (2011a) et utilise la table LOCA_FAM pour définir le nombre maximum d'équipements qu'on peut utiliser par site et par famille.

5. Conclusion

Ce chapitre a d'abord présenté un modèle de réseau pour décrire les chaînes logistiques à traiter. Ce modèle, basé sur les graphes états-tâches de Kondili et al. (1993), a été étendu pour traiter certaines subtilités de notre problème, comme les stocks partagés. Nous avons aussi proposé un mécanisme de modélisation des chaînes de récolte par des sous-graphes pouvant être copiés dans toutes les fermes utilisant une même chaîne, ce qui réduit considérablement la saisie.

Les données nécessaires pour traduire informatiquement le graphe états-tâches et toutes les données auxiliaires ont été structurées logiquement en tables Excel. Le MCD obtenu est illustré dans la figure 3.8. Chaque table décrit une entité de la chaîne (tâches, états, ressources...) avec ses attributs (dont un code-identifiant) ou une relation entre 2 ou 3 entités. Certaines relations (en rose) correspondent en fait aux résultats du module d'optimisation : elles n'ont pas été décrites dans ce chapitre car elles correspondent aux variables du modèle mathématique, vu dans le chapitre suivant. Ces tables sont facilement exportables vers un vrai SGBD comme Access.

Le lecteur aura pu constater la quantité de travail à effectuer, rien que pour modéliser les données d'un problème réel d'optimisation. Mais, grâce à ce travail préliminaire, nous disposons maintenant d'un système flexible, permettant de décrire des chaînes d'approvisionnement en biomasse très diverses, et d'une structure de base de données.

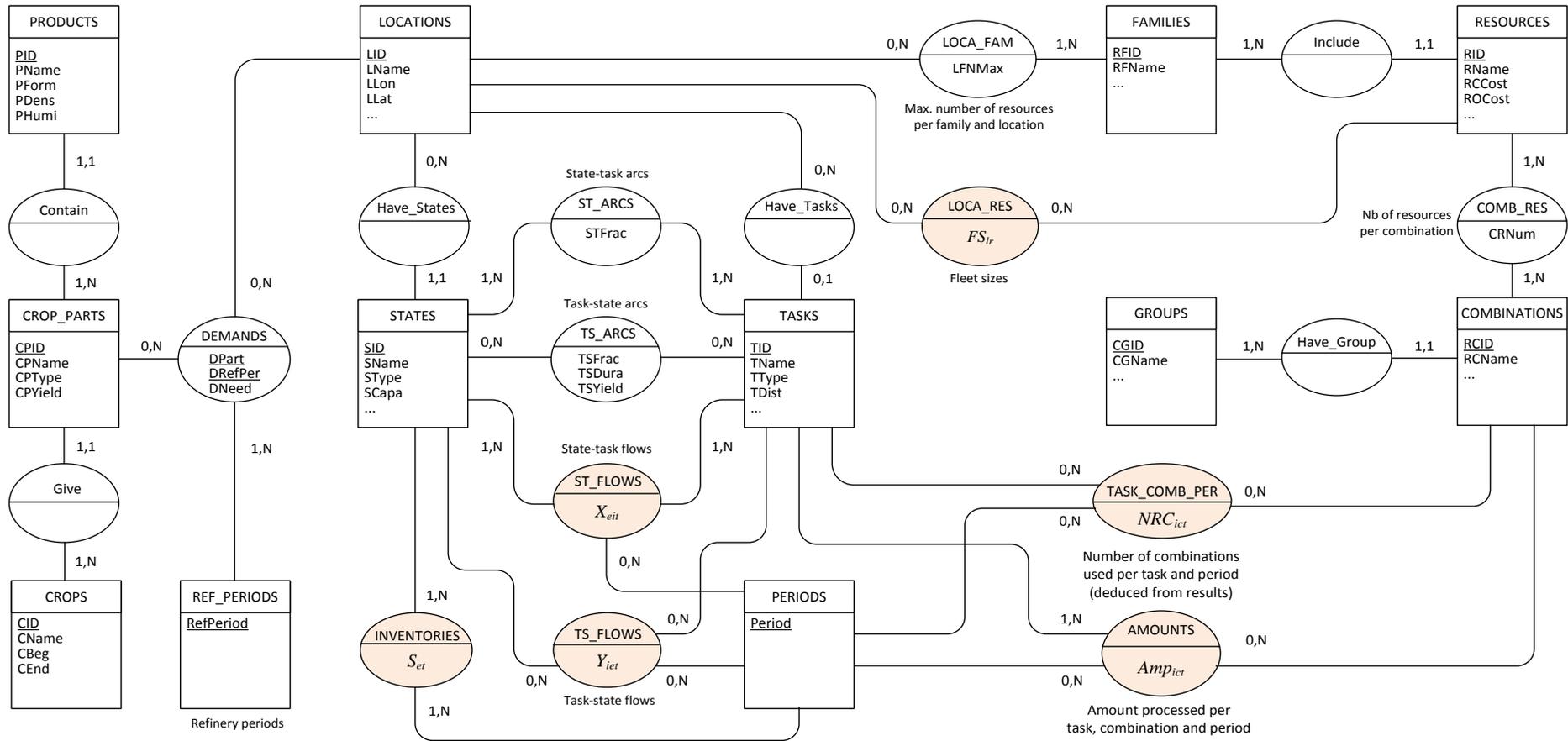


Figure 3.8. Modèle conceptuel de données (MCD) pour une bioraffinerie de proximité.

Chapitre 4

Modèle mathématique complet

1. Introduction

Comme nous l'avons vu dans le chapitre 3, il est essentiel de définir un modèle de réseau et un modèle de données si on veut parvenir à un modèle mathématique relativement indépendant de la structure de la chaîne logistique. De cette façon, l'utilisateur peut décrire sa chaîne en toute liberté, avec l'ensemble des données associées, puis le modèle de calcul peut être généré automatiquement à partir des données. Le module d'optimisation constitue ainsi une boîte noire dans laquelle toute modification est réservée aux concepteurs du modèle, ce qui évite des manipulations intempestives dans les équations. Sans les précautions que nous avons prises, le modèle traduirait une structure de chaîne figée et il faudrait le modifier fréquemment en fonction des desiderata des utilisateurs.

Nous avons présenté dans deux conférences (Ba et al., 2014c ; Ba et al., 2015b) deux modèles préliminaires que nous ne détaillerons pas faute de place. Ils sont définis sur un graphe classique comme dans la figure 1.2 du chapitre 1, avec un seul type de nœuds (sites géographiques). Comme dans Eksioglu et al. (2009), le premier utilise des variables de flot x_{ijt} pour chaque arc (i, j) et chaque période t , mais il modélise plus finement la récolte en distinguant trois étapes (moissonnage, andainage des pailles, mise en balles). Le second modèle étend le premier en remplaçant les x_{ijt} par des variables $x_{ijtt'}$ pour la quantité arrivant au nœud i en période t et expédiée vers le nœud j en période t' , comme dans Shastri et al. (2011a). Ces variables permettent de modéliser des attentes.

Ces modèles se sont révélés trop peu génériques à l'usage, avec une seule chaîne de récolte en trois étapes. De plus, il y avait des ambiguïtés car les nœuds (sites géographiques) ne permettaient pas une distinction entre stocks et tâches. Enfin, les variables de flot à quatre indices conduisaient à un modèle énorme. Ce sont ces difficultés très formatrices qui nous ont orientés vers le modèle de réseau basé sur les graphes états-tâches et le modèle de données de type entité-relation.

En s'appuyant sur le modèle de réseau et le modèle de données du chapitre 3, nous pouvons maintenant présenter un modèle mathématique très général, un programme linéaire à variables mixtes. Il est dit "complet" par rapport au modèle simplifié dit "compact" vu dans le chapitre 5, qui convient mieux aux cas de grande taille. Ce premier modèle étant déjà complexe, la section 2 à vocation pédagogique utilise une petite partie des notations pour montrer comment formuler certaines contraintes courantes à partir d'un graphe états-tâches. La section 3 présente ensuite le modèle mathématique complet, puis deux études de cas pour l'évaluer sont décrites dans la section 4. La section 5 propose des simplifications et teste la taille-limite des problèmes qui peuvent être ainsi résolus. Enfin, la section 6 conclut le chapitre en effectuant un bilan.

2. Bases de la génération automatique du modèle

Avant de présenter le modèle mathématique complet, qui est un peu ardu comme tout modèle comportant des variables et contraintes très diverses, nous passons par une étape pédagogique, pour montrer comment générer des contraintes courantes à partir d'un graphe états-tâches. Nous en profitons aussi pour contourner certaines limitations du graphe états-tâches :

- Les durées sont définies seulement pour les tâches (temps de fabrication). Comment modéliser des durées minimales de stockage, par exemple 2 mois de séchage en stock pour du miscanthus ?
- Certains produits comme les graines sont sensibles à la pluie et doivent donc être transportées vers un site de stockage le jour de leur récolte. Comment modéliser de telles contraintes de type "sans attente" (*no-wait*)?
- Comment modéliser des dégradations pendant le stockage?
- Un état ne peut contenir qu'un produit. Comment modéliser un stock partagé par plusieurs produits, comme une plate-forme stockant des balles de pailles de différentes plantes ?

Nous verrons que la génération automatique de contraintes n'est pas affectée par ces modifications.

2.1 Contraintes liées aux stocks selon Kondili et al. et désagrégation

Pour rendre cette présentation plus digeste, on utilise des notations simplifiées par rapport au modèle complet. En particulier, on ne fait pas de distinction entre les différents types d'états et de tâches et on ne tient pas compte des ressources.

On considère les notations suivantes pour un graphe états-tâches $G = (SID, TID, STA, TSA, H)$:

- SID ensemble des identifiants d'états (*state identifiers*), indicé par e
- TID ensemble des identifiants de tâches (*task identifiers*), indicé par i
- STA ensemble des arcs état-tâche (e, i) (*state-task arcs*)
- TSA ensemble des arcs tâche-état (i, e) (*task-state arcs*)
- H horizon de planification de n périodes (par exemple des jours), indicé par t de 1 à n

Ces entités sont caractérisées par les données suivantes, avec un préfixe S pour les états, ST pour les arcs état-tâche et TS pour les arcs tâche-état :

- $SCapa_e$ capacité maximale de stockage allouée à l'état e
- $STFrac_{ei}$ fraction du flux total d'entrée pris sur l'état e , avec $\sum_{(e,i) \in STA} STFrac_{ei} = 1$
- $TSFrac_{i,e}$ fraction du flux total d'entrée sortant vers l'état e , $\sum_{(i,e) \in TSA} TSFrac_{i,e} < 1$ si pertes
- $TSDur_{i,e}$ temps de traitement en périodes, entre l'entrée de tâche et la sortie vers l'état e

Kondili et al. (1993) proposent d'utiliser seulement une variable par tâche et une variable par nœud, dans chaque période t :

- $F_{i,t}$ la quantité totale traitée en entrée par la tâche i en période t
- $S_{e,t}$ la quantité stockée dans l'état e en fin de période t

Ils peuvent alors formuler comme suit les contraintes assurant la cohérence des flux et des stocks :

$$(1) \forall e \in SID, \forall t \in H : S_{e,t} \leq SCapa_e$$

$$(2) \forall e \in SID, \forall t \in H :$$

$$S_{e,t} = S_{e,t-1} + \sum_{(i,e) \in TSA | t - TSDura_{i,e} \geq 1} TSFrac_{i,e} \cdot F_{i,t-TSDura_{i,e}} - \sum_{(e,j) \in STA} STFrac_{e,j} \cdot F_{j,t}$$

Les contraintes (1) respectent les capacités des stocks. Les contraintes (2) très compactes et non évidentes garantissent l'équilibre des stocks, le respect des pourcentages en entrée et en sortie des tâches et le respect des temps d'attente. On peut comprendre ces contraintes en considérant l'état e en période t dans la figure 4.1. Le flux venant de la tâche précédente i a été réalisé par cette tâche $TSDura_{i,e}$ périodes auparavant, à partir d'un flux total $F_{i,t-TSDura_{i,e}}$ de produits consommés. Le flux reçu est donc $TSFrac_{i,e} \cdot F_{i,t-TSDura_{i,e}}$ et il s'ajoute au stock de e . Quant au flux envoyé à la tâche suivante j , il vaut par définition $STFrac_{e,j} \cdot F_{j,t}$ et il se soustrait du stock de e . En sommant les flux reçus sur les arcs entrants et les flux envoyés sur les arcs sortants, on retrouve les contraintes (2).

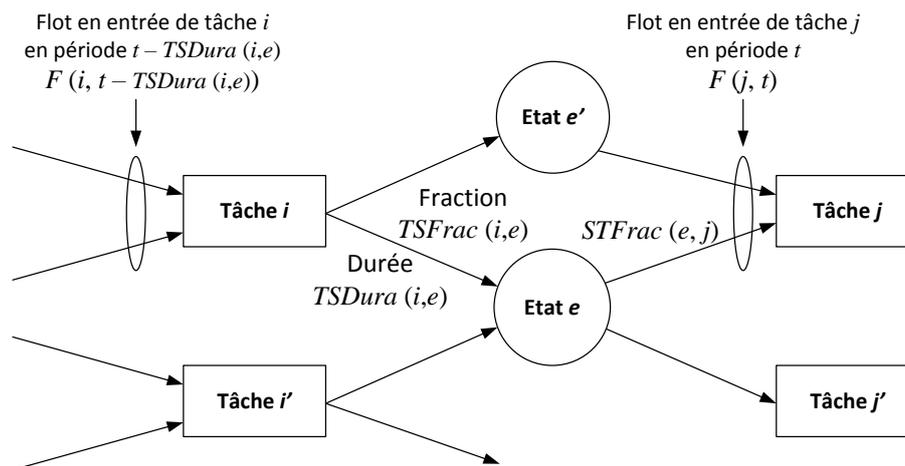


Figure 4.1. Illustration des contraintes (2) de Kondili et al. (1993).

Noter qu'avec ce système on arrive à décrire tous les flux, les stocks, et les durées des processus en utilisant uniquement, pour chaque période, une variable pour la quantité stockée dans chaque état et une variable pour le flux total entrant dans chaque tâche. Cependant, nous avons obtenu des contraintes (2) trop chargées et illisibles quand nous avons voulu introduire toutes les subtilités du modèle mathématique complet (fenêtres de temps, gestion des ressources, etc.).

Nous avons donc préféré "désagréger" les contraintes (2) en plusieurs contraintes plus simples, pour avoir un modèle plus lisible et facilement maintenable. Le modèle résultant paraît plus grand en nombre de contraintes, mais il est équivalent au précédent une fois que le pré-solveur du logiciel d'optimisation a effectué ses simplifications.

Pour chaque période t , on introduit une variable de flot $X_{e,i,t}$ pour chaque arc état-tâche (e,i) et une autre $Y_{i,e,t}$ pour chaque arc tâche-état (i,e) . Les contraintes (2) sont alors équivalentes aux contraintes suivantes :

– Des contraintes exprimant le flux total $F_{i,t}$ à l'entrée de i en fonction des nouvelles variables :

$$(3) \forall i \in TID, \forall t \in H : F_{i,t} = \sum_{(e,i) \in STA} X_{e,i,t}$$

– Des contraintes classiques d'équilibre des stocks (nouveau stock = ancien + entrées – sorties) :

$$(4) \forall e \in SID, \forall t \in H : S_{e,t} = S_{e,t-1} + \sum_{(i,e) \in TSA} Y_{i,e,t} - \sum_{(e,i) \in STA} X_{e,i,t}$$

Notons qu'on peut enlever la première somme pour un état sans prédécesseurs (un champ dans une ferme). Par contre, comme nos chaînes se terminent par la tâche-raffinerie, on n'a pas d'états sans successeurs et les contraintes (4) incluent toujours la seconde somme.

– Des contraintes de conservation des flots pour les tâches :

$$(5) \forall i \in TID, \forall t \in H : F_{i,t} \times \sum_{(i,e) \in TSA} TSFrac_{i,e} = \sum_{(i,e) \in TSA, t+TSDura_{i,e} \leq n} Y_{i,e,t+TSDura_{i,e}}$$

Le flot total en entrée doit être multiplié par la somme des $TSFrac_{i,e}$ car elle peut être inférieure à 1 en cas de pertes (rendement imparfait de la tâche). Il se répartit en différents produits obtenus sur chaque arc (i,e) en sortie, après le temps de fabrication nécessaire à chaque produit. (5) peut aussi s'écrire en remplaçant $F_{i,t}$ par son expression, donnée dans les contraintes (3).

– Pour chaque flot $Y_{i,e,t}$ (quantité entrant dans e via l'arc (i,e)), une contrainte stipulant que ce flot est une fraction $TSFrac_{i,e}$ du flux total consommé par la tâche i , $TSDura_{i,e}$ périodes auparavant :

$$(6) \forall (i,e) \in TSA, \forall t \in [TSDura_{i,e} + 1, n] : Y_{i,e,t} = TSFrac_{i,e} \cdot F_{i,t-TSDura_{i,e}}$$

– Enfin, une contrainte spécifiant que chaque flot $X_{e,i,t}$ (quantité sortant de l'état e par l'arc (e,i)) est une fraction $STFrac_{e,i}$ du flot total consommé par la tâche i . Là encore, on peut remplacer $F_{i,t}$ par son expression des contraintes (3).

$$(7) \forall (e,i) \in STA, \forall t \in H : X_{e,i,t} = STFrac_{e,i} \cdot F_{i,t}$$

Cette formulation est évidemment moins compacte que celle de Kondili et al., mais elle peut être simplifiée si la seule tâche de mélange est la raffinerie, ce qui est le cas actuellement. Dans ce cas, chaque tâche i a un seul prédécesseur e et les variables $F_{i,t}$ deviennent inutiles car $F_{i,t} = X_{e,i,t}$. De plus, les $STFrac_{e,i}$ valent alors tous 100% et les contraintes (7) deviennent redondantes. Sauf pour les états-prédécesseurs de la tâche-raffinerie (ses stocks d'entrée), il reste donc :

$$(1) \forall e \in SID, \forall t \in H : S_{et} \leq SCapa_e$$

$$(4) \forall e \in SID, \forall t \in H \setminus \{1\} : S_{e,t} = S_{e,t-1} + \sum_{(i,e) \in TSA} Y_{i,e,t} - \sum_{(e,i) \in STA} X_{e,i,t}$$

$$(5') \forall i \in TID, \forall t \in H : X_{e,i,t} \times \sum_{(i,e) \in TSA} TSFrac_{i,e} = \sum_{(i,e) \in TSA, t+TSDura_{i,e} \leq n} Y_{i,e,t+TSDura_{i,e}}$$

$$(6') \forall (i, e) \in TSA, \forall (u, i) \in STA, \forall t \in [TSDura_{i,e} + 1, n] : Y_{i,e,t} = TSFrac_{i,e} \cdot X_{u,i,t-TSDura_{i,e}}$$

Les contraintes (6') dérivent de (6) en remarquant que u est le seul état qui précède la tâche i .

2.2 Modéliser les attentes en stock

La figure 4.2 concerne un séchage de 2 mois en stock pour du miscanthus. Le schéma à gauche est incorrect car dans les graphes états-tâches les durées sont associées aux tâches, pas aux états. En fait, on confond un état (le stock) et une tâche de séchage passif. Dans le schéma correct à droite, on a remplacé l'état par deux états séparés par une tâche de séchage. Comme la capacité est partagée entre les deux produits, la contrainte de capacité doit s'appliquer à la somme des deux stocks.

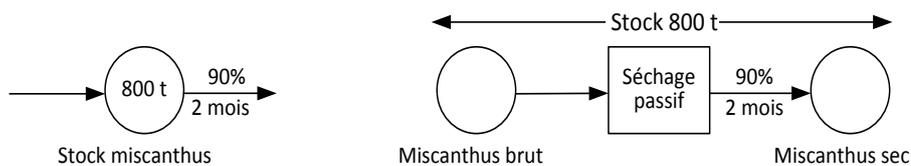


Figure 4.2. Modélisation d'un temps de séchage en stock.

Une attente peut éviter des opérations inutiles. Par exemple, supposons que des balles de menue paille après pressage puissent partir directement à une coopérative ou à un hangar de la ferme : il est inutile de ranger des balles dans le hangar si elles repartent le même jour. On peut donc introduire une tâche "attente" de 1 jour. La modélisation correcte est celle de la figure 4.3.

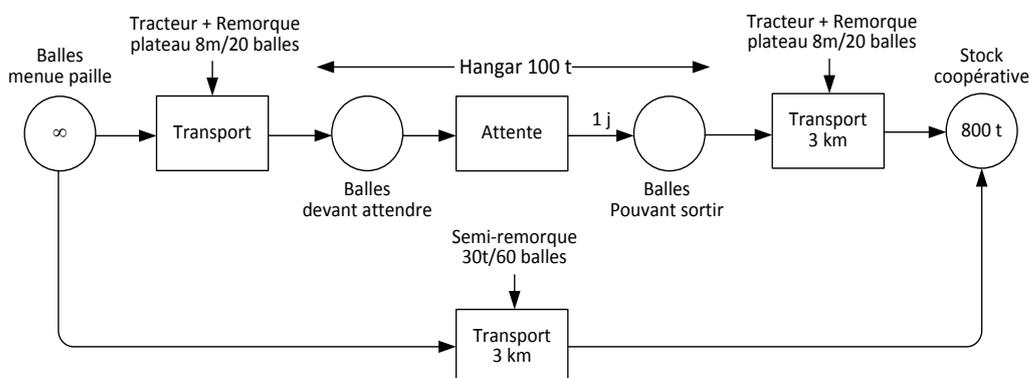


Figure 4.3. Comment éviter de ranger des balles pouvant partir le même jour.

2.3 Modélisation de contraintes "no-wait" pour un produit

Ces contraintes non prévues par Kondili et al. se gèrent simplement en mettant une capacité nulle sur l'état représentant le produit: des flots peuvent y entrer mais ils doivent en ressortir dans la même période (les stocks sont nuls dans les contraintes (4)). C'est cette astuce qui oblige les graines à partir le jour même à la coopérative, dans la figure 3.2 du chapitre 3.

Il est par contre difficile d'imposer qu'une quantité récoltée quitte la ferme sous un mois pour des raisons d'assurance, car les graphes états-tâches ne permettent pas de calculer simplement le nombre de périodes passées par un produit dans les nœuds d'une chaîne de récolte. Mais on peut approcher cette contrainte en forçant les stocks à zéro un mois après la fin de la fenêtre de récolte.

2.4 Modélisation des dégradations des produits stockés

Les experts recommandent d'appliquer à tout état e un coefficient de perte par période $SLoss_e$ (par exemple 0.05 pour 5% de perte) qui s'applique à la quantité stockée dans cet état. Sur un horizon d'un an, ce coefficient est éventuellement nul pour des produits secs très stables comme les balles de paille. Les contraintes (4) d'équilibre des stocks deviennent alors :

$$(9) \forall e \in SID, \forall t \in H \setminus \{1\} : S_{e,t} = S_{e,t-1}(1 - SLoss_e) + \sum_{(i,e) \in TSA} Y_{i,e,t} - \sum_{(e,i) \in STA} X_{e,i,t}$$

2.5 Modélisation de stocks multi-produits

Dans l'exemple de la figure 3.2 du chapitre 3, la coopérative a une plate-forme avec un stock S9 de balles de menue paille et un stock S14 de balles de paille. Comme les graphes états-tâches imposent un seul produit par état, ces deux stocks sont normalement distincts, avec leur propre capacité.

Pour dire qu'un groupe d'états forme un stock partagé, nous ajoutons un attribut $SLink_e$ qui désigne un état choisi pour représenter le groupe. Si $SLink_e = 0$, l'état e ne fait pas partie d'un stock partagé et les contraintes de capacité classiques (1) doivent être vérifiées. Sinon, l'ensemble des états u tels que $SLink_u = e$ forme un stock partagé et la capacité $SCapa_e$ spécifiée pour le représentant concerne l'ensemble du groupe. Pour les états représentant un groupe ($SLink_e = e$), il suffit de remplacer les contraintes (1) par (10).

$$(1) \forall e \in SID \mid SLink_e = 0, \forall t \in H : S_{et} \leq SCapa_e$$

$$(10) \forall e \in SID \mid SLink_e = e, \forall t \in H : \sum_{u \in SID \mid SRep_u = e} S_{ut} \leq SCapa_e$$

Pour conclure la section 2, il est important de remarquer que les contraintes de la section 2.1 sont encore valables pour les graphes obtenus. Ainsi, le modèle mathématique peut toujours être généré automatiquement, à condition que le graphe états-tâches soit réalisé correctement.

3. Description du modèle mathématique complet

Cette section définit des noms utilisés pour les données et les variables, puis décrit la fonction-objectif et les différentes contraintes.

3.1 Comment définir des données à partir des tables ?

Comme notre problème nécessite de nombreux symboles, nous avons renoncé aux symboles mathématiques comme x ou y , rapidement illisibles, pour des identificateurs plus parlants. En fait, les identificateurs utilisés sont tout simplement les noms d'attributs (champs) des entités (tables) du chapitre 3. La table 4.1 rappelle ces tables tout en précisant le préfixe utilisé par leurs noms de champs, le nom du champ-clé, et un exemple de valeur alphanumérique que nous utilisons pour cette clé. Par exemple, la clé de la table LOCATIONS s'appelle LID comme *Location Identifier*. Les lignes (enregistrements) de cette table ont pour valeurs de clé L1, L2, etc. Dans la ligne DEMANDS, H est l'ensemble des périodes de l'horizon, par exemple 1...364 pour 52 semaines de 7 jours.

Nom de table	Rôle	Préfixe	Clé	Exemple
PARAMETERS	Paramètres généraux	Aucun	Aucune	---
LOCATIONS	Sites géographiques	L	LID	L1
CROPS	Plantes cultivées	C	CID	C1
CROP_PARTS	Parties utilisées par plante	CP	CPID	CP1
PRODUCTS	Formes de parties de plantes	P	PID	P1
STATES	Etats du graphe états-tâches	S	SID	S1
TASKS	Tâches du graphe états-tâches	T	TID	T1
ST_ARCS	Arcs état → tâche	ST	Couple de SID x TID	(S1,T3)
TS_ARCS	Arcs tâche → état	TS	Couple de TID x SID	(T2, S4)
DEMANDS	Besoin/période/partie de plante	D	Couple de H x CPID	(2, CP1)
RESOURCES	Equipements	R	RID	R1
COMBINATIONS	Combinaisons de ressources	RC	CGID	RC1
GROUPS	Groupes de combinaisons	CG	CGID	CG1
FAMILIES	Famille de ressources	RF	RFID	RF1
LOCA_FAM	Nb max ressources/famille/site	LF	Couple de LID x RFID	(L2, RF3)

Table 4.1. Rappel des tables avec les préfixes des noms de champs et les noms de clés.

Ce système peut paraître compliqué pour les lecteurs utilisant des indices numériques. Il est en fait très souple et compatible avec les langages de modélisation comme AMPL, GAMS, MOSEL et OPL Studio. Rappelons que chaque table est stockée dans une feuille Excel. La première ligne (titres des colonnes) contient les noms de champs, par exemple SID (colonne des clés) et SCapa (colonne des capacités de stockage) pour la table STATES. En AMPL par exemple, on peut écrire :

```
set SID;
param {e in SID} SCapa;

table STATES IN 'CSV' 'STATES.csv' : [SID], SCapa;

display SID, SCapa, sum{e in SID} SCapa[e];
```

Les deux premières lignes définissent un ensemble SID pour les clés des états, puis un tableau de capacités SCapa indicé par SID. L'instruction "table" dit à AMPL de charger ces données à partir des colonnes de même nom dans le fichier Excel STATES, au format CSV. Les crochets indiquent que les valeurs lues doivent être des clés : AMPL va donc vérifier que toutes les clés sont distinctes. Enfin, le "display" affiche la liste des clés contenues dans SID, le tableau des capacités et la capacité totale.

Cet exemple en AMPL montre la grande souplesse du stockage de données dans des tables Excel :

- Pour chaque table, on déclare dans le modeleur un ensemble E pour les clés alphanumériques puis, pour chaque attribut, un tableau indicé sur cet ensemble.
- L'instruction "table" cherche les colonnes de mêmes noms du fichier Excel pour charger l'ensemble de clés et les tableaux des valeurs d'attributs. Ces tableaux sont donc dimensionnés dynamiquement.
- Les colonnes et les lignes peuvent être dans n'importe quel ordre dans la feuille Excel.
- Pour l'ensemble des clés E, le modeleur vérifie que les valeurs sont distinctes.
- Après chargement de données, on peut écrire des expressions faisant référence à E sans connaître le cardinal de E ni les valeurs des clés. Ainsi, le « display » de l'exemple calcule correctement la somme des capacités et affiche le tableau SCapa sans qu'on connaisse sa taille.
- Les listings de résultats sont plus lisibles grâce aux clés alphanumériques, par exemple le "display" précédent va écrire "SCapa[S1] = 18, SCapa[S2] = 29..." pour les capacités des états.

3.2 Liste des noms de données et de variables utilisés pour les calculs

Nous sommes maintenant prêts à lister les noms des données réellement utilisées dans le modèle mathématique, en les classant selon les tables d'origine. Par rapport au chapitre 3, seul un sous-ensemble de données est nécessaire à l'optimisation. Quelques symboles ont cependant été ajoutés pour améliorer la lisibilité de certaines équations.

Voici quelques exemples de données inutiles pour l'optimisation. Les libellés complets des fermes, des plantes, des équipements, etc. ne servent qu'à rendre plus lisibles les listings de résultats, le modèle identifiant les entités avec leurs clés. Les adresses et coordonnées GPS des sites sont utilisées pour préparer le distancier. Pour les ressources (équipements), la durée d'amortissement RLife et le prix d'achat RPrice sont employés seulement pour pré-calculer les coûts d'utilisation, tandis que la consommation de carburant RFuel permet de déduire les rejets de CO₂ après résolution du modèle.

<i>PerDur</i>	Durée d'une période de base, nombre entier de jours (exemple 1, 7, 14, 28)
<i>NPer</i>	Nombre de périodes de l'horizon (exemple 364 si PerDur = 1, ou 52 si PerDur = 7)
<i>RefPer</i>	Durée d'une période-raffinerie, en périodes de base (exemple 7 avec PerDur = 1)
<i>NA</i>	Chaîne de caractères remplaçant les cellules de texte vide ("not available")
<i>Huge</i>	Grand nombre positif pour initialiser certaines cellules à plus l'infini
<i>KgCO2</i>	Kilogrammes de CO ₂ émis par litre de gazole, pour le calcul des rejets de CO ₂

Table 4.2. Paramètres généraux venant de la table PARAMETERS.

<i>H</i>	Horizon de planification avec les périodes de base dans $\{1,2, \dots, NPer\}$
<i>HR</i>	Idem pour la raffinerie avec des périodes dans $\{1,2, \dots, NumPer/RefPer\}$
<i>LID</i>	Sites L1, L2, ... (clés de la table LOCATIONS)
<i>CID</i>	Plantes C1, C2, ... (clés de CROPS)
<i>CPID</i>	Parties de plantes CP1, CP2, ... (clés de CROP_PARTS)
<i>PID</i>	Produits P1, P2, ... (clés de PRODUCTS)
<i>SID</i>	Etats S1, S2, ... (clés de STATES)
<i>TID</i>	Tâches T1, T2, ... (clés de TASKS)
<i>RID</i>	Ressources R1, R2, ... (clés de RESOURCES)
<i>RFID</i>	Familles de ressources RF1, RF2, ... (clés de FAMILIES)
<i>RCID</i>	Combinaisons de ressources RC1, RC2, ... (clés de COMBINATIONS)
<i>CGID</i>	Groupes de combinaisons de ressources CG1, CG2, ... (clés de GROUPS)
<i>TSA</i>	Arcs tâche-état (i, e), $TSA \subseteq TID \times SID$ (clés de TS_ARCS)
<i>STA</i>	Arcs état-tâche (e, i), $STA \subseteq SID \times TID$ (clés de ST_ARCS)
<i>DPairs</i>	$DPairs \subseteq HR \times CPID$ (clés de DEMANDS)
<i>LFPairs</i>	$LFPairs \subseteq LID \times RFID$ (clés de LOCA_FAM)
<i>RCRPairs</i>	$RCRPairs \subseteq RCID \times RID$
<i>ZHR</i>	Sous-ensemble {Harvest, Rake} des types de tâches
<i>ZB</i>	Sous-ensemble {Bale} des types de tâches
<i>ZT</i>	Sous-ensemble {Transport} des types de tâches
<i>ZR</i>	Sous-ensemble {Refine} des types de tâches
<i>ZHRB</i>	Sous-ensemble des types de tâches concernant la récolte {Harvest, Rake, Bale}

Table 4.3. Ensembles utilisés pour les indexations.

Commentaires sur la table 4.3 :

- RCRPairs correspond aux couples de clés de la relation COMB_RES, que nous n'avons pas décrite dans le chapitre précédent pour ne pas noyer le lecteur dans les détails du modèle entité-relation. Cette relation, visible cependant dans le MCD de la figure 3.8 du chapitre 3, permet de préciser les ressources qui composent chaque combinaison de ressources.
- Les cinq derniers sous-ensembles sont définis pour alléger l'écriture de certaines contraintes qui ne concernent que certains types de tâches. En effet, pour tester le type *SType* d'une tâche, il vaut mieux utiliser ces sous-ensembles (par exemple, est-ce que $SType \in ZHR$?) plutôt que de faire référence aux noms de type (est-ce que $SType = Harvest$ ou $SType = Rake$?) car on pourra changer ces noms dans le futur et ajouter d'autres types de tâches.

La table 4.4 indique les attributs chargés pour chaque table, sauf pour les ressources qui font l'objet de la table 4.5. Nous en profitons aussi pour choisir un indice préféré pour chaque attribut. Le lecteur est invité à se référer au chapitre 3 pour des détails sur chaque attribut.

Pour chaque site $l \in LID$ venant de la table SITES	
$LType_l$	Type de site, valeurs actuelles dans {Farm, Storage, Pretreatment, Refinery}
$LHours_l$	Nombre d'heures d'ouverture, par période
Pour chaque partie de plante $k \in CPID$ venant de la table CROP_PARTS	
$CPCrop_k$	ID de la plante d'origine
Pour chaque produit $p \in PID$ venant de la table PRODUCTS	
$PPart_p$	ID de la partie de plante utilisée (permet aussi de retrouver la plante)
$PDens_p$	Densité
$PHumi_p$	Humidité en % du poids
Pour chaque état $e \in SID$ venant de la table STATES	
$SType_e$	Type d'état, valeurs possibles dans {Field, WIP, Stock}
$SLoca_e$	ID du site géographique de rattachement
$SProd_e$	ID du produit stocké
$SBeg_e$	Première période de la fenêtre d'utilisation
$SEnd_e$	Dernière période de la fenêtre d'utilisation
$SCapa_e$	Capacité en tonnes pour le type Stock (inutile pour Field)
$SQBeg_e$	Quantité initiale dans le stock, en tonnes ($\leq SCapa$)
$SQEndMin_e$	Quantité finale minimale dans le stock, en tonnes ($\leq SCapa$)
$SQEndMax_e$	Quantité finale maximale dans le stock, en tonnes ($\geq SQEndMin, \leq SCapa$)
$SInvCost_e$	Coût de stockage (inventory cost) en €/t et par période
$SInpCost_e$	Coût d'entrée en €/t et par période
$SOutCost_e$	Coût de sortie en €/t et par période
$SThru_e$	Débit maximal de réception, en tonnes par heure
$SLoss_e$	Perte de produit par période en %
$SRep_e$	Pour les états de type Stock, ID du représentant en cas de stock partagé
Pour chaque tâche (task) $i \in TID$ venant de la table TASKS	
$TType_i$	Type dans {Harvest, Dry, Rake, Bale, Pretreat, Transport, Refine}
$TLoca_i$	ID du site (pour les tâches de transport, c'est le site possédant les véhicules)
$TDist_i$	Distance par la route pré-calculée (tâches de transport uniquement)
$TGroup_i$	Groupe de combinaisons de ressources utilisé
Pour chaque arc état-tâche $(e, i) \in STA$ venant de la table ST_ARCS	
$STFrac_{e,i}$	Ratio entre le flux sur l'arc et le flot total consommé par la tâche, en %
$STYield_{e,i}$	Rendement en tonnes/hectare (arc sortant de chaque état Field)
Pour chaque arc tâche-état $(i, e) \in TSA$ venant de la table TS_ARCS	
$TSFrac_{i,e}$	Ratio entre le flux sur l'arc et le flux total en entrée de tâche, en %
$TSDura_{i,e}$	Durée de traitement (en nombre de périodes) pour la sortie vers l'état e

Table 4.4. Attributs utilisés pour chaque table (hors gestion des ressources).

Comme déjà mentionné, de nombreuses données du chapitre 3 ne sont pas listées car elles ne servent à rien dans l'optimisation. Ainsi, la table des plantes (CROPS) n'est pas du tout utilisée.

Pour chaque ressource $r \in RID$ venant de la table RESOURCES	
$RFam_r$	ID de la famille de ressources à laquelle appartient la ressource
$RCCost_r$	Coût fixe (capital cost) en € sur l'horizon considéré
$RWTime_r$	Temps de travail maximum par période, en h
Pour chaque combinaison de ressource $c \in RCID$ venant de la table COMBINATIONS	
$RCGroup_c$	ID du groupe de combinaisons auquel appartient la combinaison de ressources
$RCOCost_c$	Coût horaire d'utilisation (operating cost) en €/h
$RCThru_c$	Débit horaire, exprimé en ha/h pour les combinaisons traitant les tâches <i>Harvest</i> ou <i>Rake</i> mais en balles/h pour celles traitant les tâches <i>Bale</i>
$RCWTime_c$	Temps de travail maximum par période, en h
$RCSpeed_c$	Vitesse moyenne en km/h
$RCWeight_c$	Poids maximal de la charge utile en tonnes
$RCVolume_c$	Volume maximal de la charge utile en m ³
$RCPower$	Puissance en kilowatts
$RCFuel$	Consommation de gazole en litres par heure
Pour chaque couple $(w, k) \in DPairs$ venant de la table DEMANDS	
$DNeed_{w,k}$	Besoin de partie de plante k pour la période-raffinerie w , en tonnes sèches
Pour chaque couple $(l, f) \in LFPairs$ venant de la table LOCA_FAM	
$LFNMax_{l,f}$	Nombre maximum de ressources de la famille f utilisables sur le site l

Table 4.5. Attributs utilisés pour la gestion des ressources.

Variables continues non négatives	
$X_{e,i,t}$	Flux de produit sur l'arc état-tâche (e, i) en période t
$Y_{i,e,t}$	Flux de produit sur l'arc tâche-état (i, e) en période t
$S_{e,t}$	Niveau de stock de l'état e en fin de période t
$Amp_{i,c,t}$	Quantité totale traitée par la tâche i avec la combinaison c en période t
$TCHRB_{i,c,t}$	Temps d'utilisation de la combinaison c en période t pour une tâche $i \in ZHRB$
$TCT_{i,c,t}$	Temps d'utilisation de la combinaison c en période t pour une tâche $i \in ZT$
$TR_{l,r,t}$	Temps d'utilisation de la ressource r sur le site l en période t
Variables entières :	
$RT_{i,c,t}$	Nombre de rotations (allers-retours) réalisées par la combinaison c pour la tâche de transport i en période t ,
$FS_{l,r}$	Nombre de ressources r (fleet size) nécessaires sur le site l

Table 4.6. Variables utilisées.

3.3 Fonction-objectif

La fonction-objectif $TotalCost$ à minimiser est le coût total de l'approvisionnement en biomasse, composé des coûts de récolte, de conditionnement (pressage), de transport, de stockage et de manutention. Le modèle minimise donc la somme des différents coûts suivants :

$$\min TotalCost = FarmOpCosts + TransOpCosts + FixedCosts + InvCosts + HandCosts$$

Les coûts d'utilisation $FarmOpCosts$ concernent les combinaisons de ressources pour les tâches de type Harvest, Rake ou Bale (ensemble $ZHRB$), les transports internes aux fermes étant regroupés avec les autres transports dans $TransOpCosts$. Ils sont calculés à partir des variables $TCHRB_{i,c,t}$ (heures d'utilisation de la combinaison c par la tâche i en période t). L'indexation est assez complexe. D'une part, on considère toute tâche i de l'ensemble des tâches (TID) telle que son type $TType_i$ soit dans $ZHRB$. D'autre part, on considère seulement les combinaisons c de l'ensemble des combinaisons ($RCID$) dont le groupe est celui préconisé pour la tâche ($RCGroup_c = TGroup_i$).

$$FarmOpCosts = \sum_{i \in TID | TType_i \in ZHRB} \sum_{c \in RCID | RCGroup_c = TGroup_i} \sum_{t \in H} TCHRB_{i,c,t} \cdot RCOCost_c$$

Les coûts d'utilisation $TransOpCosts$ des combinaisons de ressources pour les tâches de transport (ZT) sont basés sur les heures d'utilisation $TCT_{i,c,t}$ des véhicules, dont le calcul est spécifique.

$$TransOpCosts = \sum_{i \in TID | TType_i \in ZT} \sum_{c \in RCID | RCGroup_c = TGroup_i} \sum_{t \in H} TCT_{i,c,t} \cdot RCOcost_c$$

Les coûts fixes $FixedCosts$ des ressources à acheter s'obtiennent à partir des variables $FS_{l,r}$ qui indiquent le nombre de ressources de type r nécessaires pour le site l .

$$FixedCosts = \sum_{l \in LID} \sum_{r \in RID | (l, RFam_r) \in LFPairs} FS_{l,r} \cdot RCCost_r$$

Les coûts de stockage $InvCosts$ s'appliquent au tonnage présent dans chaque état e à la fin de chaque période t . En général, les états de type Field et WIP n'ont pas de coût de stockage. Pour les vrais stockages, certains comme les silos sont utilisables seulement une partie de l'année pour l'activité biomasse. Le reste du temps, ils sont réservés à des activités comme la production de céréales. Un silo vide est mis à notre disposition et on doit le rendre vide à la fin. On peut donc compter les coûts de stockage seulement dans la fenêtre de disponibilité $[SBeg_e, SEnd_e]$.

$$InvCosts = \sum_{e \in SID} \sum_{t \in [SBeg_e, SEnd_e]} SInvCost_e \cdot S_{e,t}$$

Enfin, les coûts de manutention (handling costs) $HandCosts$ sont des coûts forfaitaires pour faire entrer des produits dans les stockages. Ils s'appliquent aux flux entrants et sortants et permettent d'ignorer les nombreuses petites ressources de manutention : aspirateurs et convoyeurs à bande des silos à grains, chariots à fourche type Maniscopic pour empiler des balles, etc.

$$HandCosts = \sum_{(i,e) \in TSA} \sum_{t \in [SBeg_e, SEnd_e]} SInpCost_e \cdot Y_{i,e,t} + \sum_{(e,i) \in STA} \sum_{t \in [SBeg_e, SEnd_e]} SOutCost_e \cdot X_{e,i,t}$$

3.4 Contraintes de stocks et de flux

Ces contraintes correspondent aux contraintes proposées par Kondili et al. (1993), que nous avons présentées sous forme simplifiée et que nous avons désagrégées dans la section 2.1. Comme nous avons un grand nombre de contraintes, nous préférons utiliser des noms mnémoniques au lieu de numéros d'équations.

Contraintes de capacité pour les états-stocks non partagés :

Ces contraintes sont générées seulement si une capacité limitée est donnée. Elles ne sont donc pas générées pour les états comme les champs qui ont une capacité égale à *Huge* :

$$\forall e \in SID \mid SCapa_e < Huge \text{ and } SLink_e = NA, \forall t \in [SBeg_e, SEnd_e] : S_{e,t} \leq SCapa_e \quad (\text{InvCap1})$$

Contraintes de capacité pour les états-stocks partagés (avec plusieurs produits) :

Pour les états d'un stock partagé, les contraintes sont générées seulement pour l'état-représentant, tel que $SLink_e = e$ (voir section 2.5).

$$\forall e \in SID \mid SCapa_e < Huge \text{ and } SLink_e = e, \forall t \in [SBeg_e, SEnd_e] : \quad (\text{InvCapN})$$

$$\sum_{u \in SID \mid SLink_u = e} S_{u,t} \leq SCapa_e$$

Equations d'équilibre des stocks (cas général) :

Elles tiennent compte des dégradations pendant le stockage.

$$\forall e \in SID, \forall t \in [SBeg_e + 1, SEnd_e] : \quad (\text{InvBal})$$

$$S_{e,t} = S_{e,t-1} \cdot (1 - SLoss_e/100) + \sum_{(i,e) \in TSA} Y_{i,e,t} - \sum_{(e,j) \in STA} X_{e,j,t}$$

Cas particulier de la première période de la fenêtre :

Le stock initial $SQBeg$ est en général nul, sauf pour les états de type Field et ceux de la raffinerie :

$$\forall e \in SID : \quad (\text{InitBal})$$

$$S_{e,SBeg_e} = SQBeg_e \cdot (1 - SLoss_e/100) + \sum_{(i,e) \in TSA} Y_{i,e,SBeg_e} - \sum_{(e,i) \in STA} X_{e,i,SBeg_e}$$

Contraintes exprimant le flot sur chaque arc sortant d'une tâche :

Actuellement, toute tâche ayant des états-successeurs (toutes, sauf la raffinerie) a aussi un seul prédécesseur u , mais on a mis une somme pour pouvoir traiter des tâches de mélange à l'avenir :

$$\forall (i, e) \in TSA, \forall t \in [SBeg_e, SEnd_e] \mid t - TSDura_{ie} \geq 1 : \quad (\text{GetY})$$

$$Y_{i,e,t} = \left(\frac{TSFrac_{ie}}{100} \right) \cdot \sum_{(u,i) \in STA} X_{u,i,t-TSDura_{ie}}$$

Contraintes exprimant le flot sur chaque arc entrant dans une tâche :

Dans nos tests, les tâches autres que Refine ont un seul état-prédécesseur. Dans ce cas, $STFrac_{e,i}$ vaut 100% et la somme du second membre est égale à $X_{e,i,t}$: ces contraintes sont redondantes mais nous avons vérifié qu'elles sont bien éliminées par le pré-solveur. Nous les conservons dans notre formulation, pour pouvoir traiter des tâches de mélange dans l'avenir.

$$\forall (e, i) \in STA \mid TType_i \notin ZR, \forall t \in [SBeg_e, SEnd_e] : \quad (GetX)$$
$$X_{e,i,t} = \left(\frac{STFrac_{e,i}}{100} \right) \cdot \sum_{(e,j) \in STA \mid j=i} X_{e,j,t}$$

Conservation des flux pour chaque tâche, sauf Refine :

Pour la raffinerie, ces contraintes prennent une forme particulière vue plus loin. La première somme est le flux total consommé par la tâche i , prélevé simultanément sur les arcs (e, i) en entrée. Elle doit être multipliée par la somme des $TSFrac$, qui peut être inférieure à 1 en cas de perte. Le flot obtenu se retrouve sur les arcs de sortie (i, u) , $TSDura_{i,u}$ périodes plus tard.

$$\forall i \in TID \mid TType_i \notin ZR, \forall t \in H : \quad (FlowCons)$$
$$\sum_{(e,i) \in STA \mid t \in [SBeg_e, SEnd_e]} X_{e,i,t} \cdot \sum_{(i,u) \in TSA} \left(\frac{TSFrac_{i,u}}{100} \right) =$$
$$\sum_{(i,u) \in TSA \mid t \in [SBeg_u, SEnd_u - TSDura_{(i,u)}]} Y_{i,u,t+TSDura_{(i,u)}}$$

Contraintes de stock final minimal :

Elles sont relatives à la dernière période de la fenêtre du stock e , générées seulement quand un tel stock est spécifié. Elles servent surtout aux stocks d'entrée de la raffinerie, en permettant d'enchaîner des années successives sans devoir arrêter les activités.

$$\forall e \in SID \mid SQEndMin_e > 0 : S_{e,SEnd_e} \geq SQEndMin_e \quad (MinEndInv)$$

Contraintes de stock final maximal :

Elles sont générées seulement quand un tel stock est spécifié. La valeur $SQEndMax_e = 0$, spécifiée pour les états des fermes et des sites de stockage centralisés, permet de vider les stocks après la dernière période $SEnd_e$ de la fenêtre.

$$\forall e \in SID \mid SQEndMax_e < Huge : S_{e,SEnd_e} \leq SQEndMax_e \quad (MaxEndInv)$$

Contraintes spécifiant la capacité maximale de réception d'un site de stockage :

A cause de capacités de manutention limitées, certains sites de stockage comme les silos ont un débit limité en tonnes/heure ($SThru_e$, en tonnes/heure), à multiplier par le nombre d'heures d'ouverture par période du site géographique $SLoca_e$ qui contient le stock, telle une coopérative.

$$\forall e \in SID \mid SType_e = Stock, \forall t \in [SBeg_e, SEnd_e] : \quad (MaxInput)$$
$$\sum_{(i,e) \in TSA} Y_{i,e,t} \leq SThru_e \cdot LHours[SLoca_e]$$

3.5 Contraintes de satisfaction des demandes de la raffinerie

Il s'agit de contraintes de flots particulières. A part la raffinerie, toutes les activités sont planifiées avec une période dite "de base" (en général 1 jour). Les demandes de la raffinerie utilisent une période plus longue ("période-raffinerie", en général 1 semaine), égale à $RefPer$ périodes de base. De plus, la raffinerie raisonne en tonnes sèches de différentes parties de plantes. Le produit (partie de plante avec un conditionnement et une humidité spécifique) importe peu : un besoin de 185 tonnes sèches de paille de colza peut par exemple être satisfait avec 100 tonnes de balles de paille à 10% d'humidité et 100 tonnes de granulés de paille (pellets) à 5%.

Il faut donc vérifier que les produits livrés pour chaque partie de plante satisfont les demandes, en tenant compte de l'humidité et de la période plus longue utilisée par la raffinerie.

Version 1 :

Comme on ne connaît pas les fluctuations de production dans une période-raffinerie, on suppose que le besoin en tonnes sèches pour une partie de plante k dans une période-raffinerie donnée w est consommé uniformément dans chaque période de base. Par exemple, une demande de 70 tonnes en semaine w sera consommée à raison de 10 tonnes par jour. Attention, ceci n'implique pas de devoir livrer tous les jours les mêmes quantités à la raffinerie, car celle-ci a des stocks d'entrée pour deux semaines de production en moyenne (voir la figure 3.2 du chapitre 3). Les flots sortant des stocks d'entrée de la raffinerie doivent alors respecter les contraintes suivantes :

$$\forall i \in TID | TType_i \in ZR, \forall (w, k) \in DPairs, \forall t \in [(w - 1) \cdot RefPer + 1, w \cdot RefPer] : \quad (\text{SatDem1}) \\ \sum_{(e,i) \in STA | PPart[SProd_e]=k} (1 - PHumi[SProd_e]/100) \cdot X_{e,i,t} = DNeed_{w,k}/RefPer$$

La contrainte est formulée pour chaque tâche-raffinerie. Actuellement, il n'y en a qu'une mais on pourrait gérer plusieurs raffineries. La demande pour la période-raffinerie numéro w et la partie de plante k ($DNeed_{w,k}$) concerne les périodes de base allant de $((w - 1) \cdot RefPer + 1)$ à $(w \cdot RefPer)$. La somme cumule les flots de produits (convertis en tonnes sèches) contenant la partie de plante k et prélevés dans chaque période de base de la période-raffinerie sur les stocks d'entrée. Elle doit être égale à la consommation uniforme $DNeed_{w,k}/RefPer$ de la période-raffinerie.

Version 2 :

Elle correspond au pire cas où la demande de chaque période-raffinerie est consommée dès la première période de base. Une demande de 70 tonnes pour la semaine w est donc consommée le lundi. Cette version plus contrainte peut être infaisable si on n'a pas eu le temps d'amener les produits. En pratique, on réduit ce risque avec des stocks pour deux semaines de consommation.

$$\forall i \in TID | TType_i = Refine, \forall (w, k) \in DPairs, t = (w - 1) \cdot RefPer + 1 : \quad (\text{SatDem2}) \\ \sum_{(e,i) \in STA | PPart[SProd_e]=k} (1 - PHumi[SProd_e]/100) \cdot X_{e,i,t} = DNeed_{w,k}$$

Dans les deux versions, noter que les stocks d'entrée de la raffinerie sont mis à jour correctement grâce aux contraintes (InvBal) de la section 3.4.

3.6 Contraintes de gestion des ressources

Nous les plaçons à part pour pouvoir modifier facilement la gestion des ressources. Les notions de ressources, combinaisons et groupes ont été expliquées dans la section 4 du chapitre 3.

Contraintes répartissant le flot total en entrée de tâche sur les différentes combinaisons :

Un groupe de combinaisons $TGroup_i$ est spécifié pour chaque tâche i , par exemple un tracteur + une remorque, mais la tâche peut utiliser plusieurs combinaisons dans le même groupe, par exemple un tracteur de 200 ch avec une remorque de 6 m et un tracteur de 240 ch avec une remorque de 8 m.

$$\forall i \in TID \mid TType_i \notin ZR, \forall t \in H : \quad (\text{SpanFlow})$$

$$\sum_{c \in RCID \mid RCGroup_c = TGroup_i} Amp_{i,c,t} = \sum_{(e,i) \in STA} X_{e,i,t}$$

Détermination des temps d'utilisation des combinaisons par les tâches de récolte :

Il s'agit de calculer les $TCHRB_{i,c,t}$, nombre d'heures d'utilisation de la combinaison c par la tâche de récolte i (de type Harvest, Rake ou Bale, ensemble $ZHRB$) en période t , qui sont nécessaires pour les coûts d'utilisation des équipements dans la fonction-objectif. Nous avons deux cas car les productivités $RThru$ sont en hectares par heure pour les tâches Harvest et Rake, et en balles par heure pour les tâches Bale. On se ramène dans les deux cas en tonnes par heure.

$$\forall i \in TID \mid TType_i \in ZHRB, \forall c \in RCID \mid RCGroup_c = TGroup_i, \forall t \in H : \quad (\text{GetTCHRB})$$

$$\text{Si } TType_i \in ZHR \text{ alors : } TCHRB_{i,c,t} = \frac{Amp_{i,c,t}}{\sum_{(e,i) \in STA} RCThru_c \cdot STYield_{e,i}}$$

$$\text{Sinon : } TCHRB_{i,c,t} = \frac{Amp_{i,c,t}}{\sum_{(e,i) \in STA} RCThru_c \cdot RCVolume_c \cdot PDens_{SProd(e)}}$$

Détermination des temps d'utilisation des combinaisons par les tâches de transport :

On se base cette fois sur le nombre de rotations (variables entières $RT_{i,c,t}$) et sur la distance $Dist_i$.

$$\forall i \in TID \mid TType_i \in ZT, \forall c \in RCID \mid RCGroup_c = TGroup_i, \forall t \in H : \quad (\text{GetTCT})$$

$$TCT_{i,c,t} = RT_{i,c,t} \left(\frac{2 \cdot TDist_i}{RCSpeed_c} \right)$$

Détermination des temps d'utilisation des ressources :

Ces durées se déduisent des heures d'utilisation des combinaisons où la ressource intervient. Contrairement aux combinaisons, on ne peut pas séparer les tâches *Harvest, Rake, Transport* car un même tracteur peut intervenir sur toutes ces tâches.

$$\forall l \in LID, \forall r \in RID \mid (l, RFam_r) \in LFPairs, \forall t \in H : \quad (\text{GetTR})$$

$$TR_{l,r,t} = \sum_{c \in RCID \mid (c,r) \in RCRPairs} ($$

$$\sum_{i \in TID \mid TType_i \in ZHRB \text{ and } RCGroup_c = TGroup_i \text{ and } TLoca_i = l} TCHRB_{i,c,t} +$$

$$\sum_{i \in TID \mid TType_i \in ZT \text{ and } RCGroup_c = TGroup_i \text{ and } TLoca_i = l} TCT_{i,c,t})$$

Borne inférieure pour le nombre de rotations des véhicules (tâches de transport) :

Il s'agit de borner la variable entière $RT_{i,c,t}$ qui indique le nombre de rotations pour transporter la quantité traitée par la tâche de transport i et la combinaison de ressources c en période t , par exemple un camion-tracteur et une remorque routière. $Amp_{i,c,t}$ est divisée par le minimum entre la capacité du véhicule et le poids maximal résultant du volume du véhicule et de la densité du produit transporté. Comme ce produit est associé aux états, pas aux tâches, la somme permet d'obtenir l'état de départ e , qui est unique pour une tâche de transport.

$$\forall i \in TID | TType_i \in ZT, \forall c \in RCID | RCGroup_c = TGroup_i, \forall t \in H : \quad (LB_RT)$$

$$RT_{i,c,t} \geq \frac{Amp_{i,c,t}}{\min(RCWeight_c, \sum_{(e,i) \in STA} RCVolume_c \cdot PDens_{SProd(e)})}$$

Borne supérieure pour le nombre de rotations :

En fait, $RT_{i,c,t}$ ne doit pas dépasser la borne inférieure précédente, plus 1. Ainsi, si le quotient dans (LB_RT) vaut 6,3, la variable entière $RT_{i,c,t}$ doit être supérieure ou égale à 6,3 et inférieure ou égale à 7,3 : le solveur n'a pas le choix et doit la mettre à 7. Si par contre le quotient est entier et vaut 6, le solveur a le choix, mais entre deux valeurs seulement : 6 et 7.

$$\forall i \in TID | TType_i \in ZT, \forall c \in RCID | RCGroup_c = TGroup_i, \forall t \in H : \quad (UB_RT)$$

$$RT_{i,c,t} \leq \frac{Amp_{i,c,t}}{\min(RCWeight_c, \sum_{(e,i) \in STA} RCVolume_c \cdot PDens_{SProd(e)})} + 1$$

Borne inférieure pour le nombre d'exemplaires de chaque ressource sur chaque site :

Elle résulte du temps maximal d'utilisation de la ressource et du temps de travail maximal sur le site, les deux par période :

$$\forall l \in LID, \forall r \in RID | (l, RFam_r) \in LFPairs, \forall t \in H : \quad (LB_FS)$$

$$FS_{l,r} \geq \frac{TR_{l,r,t}}{\min(RWTime_r, LHours_l)}$$

Borne supérieure pour la taille de la flotte de ressource :

Cette coupe simple obéit au même principe que celui de la contrainte (UB_RT) précédente.

$$\forall l \in LID, \forall r \in RID | (l, RFam_r) \in LFPairs, \forall t \in H : \quad (UB1_FS)$$

$$FS_{l,r} \leq \frac{TR_{l,r,t}}{\min(RWTime_r, LHours_l)} + 1$$

Limitation du nombre total de ressources par famille sur chaque site :

$$\forall (l, f) \in LFPairs : \sum_{r \in RID | (l, RFam_r) \in LFPairs} FS_{l,r} \leq LFNMax_{l,f} \quad (UB2_FS)$$

3.7 Déclaration des variables

On profite des déclarations pour forcer les stocks et les flux à zéro en dehors des fenêtres. Cette précaution est en général inutile si on a un logiciel avec un bon pré-solveur, car il élimine les variables n'apparaissant dans aucune équation.

Flux sortant des états, forcés à zéro en dehors de la fenêtre d'utilisation :

$$\forall (e, i) \in STA, \forall t \in [SBeg_e, SEnd_e] : X_{e,i,t} \geq 0$$

$$\forall (e, i) \in STA, \forall t \notin [SBeg_e, SEnd_e] : X_{e,i,t} = 0$$

Flux entrant dans les états, forcés à zéro en dehors de la fenêtre d'utilisation :

$$\forall (i, e) \in TSA, \forall t \in [SBeg_e, SEnd_e] : Y_{i,e,t} \geq 0$$

$$\forall (i, e) \in TSA, \forall t \notin [SBeg_e, SEnd_e] : Y_{i,e,t} = 0$$

Niveaux des stocks, forcés à 0 en dehors de la fenêtre d'utilisation :

$$\forall e \in SID, \forall t \in [SBeg_e, SEnd_e] : S_{e,t} \geq 0$$

$$\forall e \in SID \mid LType[SLoca_e] = Farm \text{ and } SType_e \neq Field, \forall t \notin [SBeg_e, SEnd_e] : S_{e,t} = 0$$

Définition des variables *Amp* seulement pour les ressources utilisables par la tâche *i* :

$$\forall i \in TID \mid TType_i \notin ZR, \forall c \in RCID \mid RCGroup_c = TGroup_i, \forall t \in H : Amp_{i,c,t} \geq 0$$

Définition des variables *RT* seulement pour les ressources utilisables par la tâche *i* :

$$\forall i \in TID, \forall c \in RCID \mid RCGroup_c = TGroup_i, \forall t \in H : RT_{i,c,t} \in \mathbb{N}$$

Définition des variables *FS* :

$$\forall l \in LID, \forall r \in RID \mid (l, RFam_r) \in LFPairs : FS_{l,r} \in \mathbb{N}$$

3.8 Dédution des consommations d'énergie et rejets de GES

Ces critères environnementaux sont importants pour apprécier une solution. Comme notre modèle n'est pas multi-objectif, nous nous contentons dans le cadre de cette thèse de minimiser le coût total de la chaîne d'approvisionnement. Les consommations d'énergie et les rejets de gaz à effet de serre (GES) sont ensuite déduits à partir des valeurs optimales des variables, donc après la résolution du modèle mathématique.

Pour chaque combinaison de ressources, on connaît la puissance $RCPower$ en kilowatts. C'est celle de l'engin motorisé de la combinaison, par exemple le tracteur pour une combinaison tracteur + presse à balles. Cette puissance pourrait être aussi celle d'un équipement électrique comme un four ou une grosse presse dans un site de prétraitement.

L'énergie totale consommée en kilowatts x heures (kW.h) se déduit simplement des puissances $RCPower$ des combinaisons et de leurs heures d'utilisation, c'est-à-dire $TCHRB_{i,c,t}$ pour les équipements de récolte et $TCT_{i,c,t}$ pour les véhicules de transport :

$$Energy = \sum_{i \in TID} \sum_{c \in RCID \mid RCGroup_c = TGroup_i} \sum_{t \in H} (TCT_{i,c,t} + TCHRB_{i,c,t}) \cdot RPower$$

La formule est presque aussi simple pour les rejets de CO2. En effet, on connaît pour chaque combinaison la consommation horaire de gazole en litres, $RFuel$, tandis que la table des paramètres généraux (PARAMETERS) indique le nombre de kilogrammes de CO2 rejetés par litre de gazole consommé, $KgCO2$ (ce paramètre vaut environ 2,6 kg/litre) :

$$CO2 = \sum_{i \in TID} \sum_{c \in RCID \mid RCGroup_c = TGroup_i} \sum_{t \in H} (TCT_{i,c,t} + TCHRB_{i,c,t}) \cdot RFuel \cdot KgCO2$$

4. Deux études de cas

Nous décrivons maintenant l'évaluation du modèle complet sur deux études de cas. La première, à vocation pédagogique, inclut un horizon de 28 jours, 1 zone de production, 1 plante (colza), 1 site de stockage centralisé et 1 raffinerie. La seconde, plus grosse, implique un horizon de 364 jours, 10 zones de production pouvant livrer du colza et/ou du miscanthus, 3 sites de stockage et 1 raffinerie.

4.1 Etude de cas avec une ferme sur 28 jours

Cette étude de cas correspond au graphe états-tâches de la figure 4.4, similaire à la figure 3.2 du chapitre 3 mais sans stocks partagés. La chaîne comprend une ferme cultivant du colza, une coopérative avec un silo à graines et une plate-forme pour des balles, et une raffinerie. Trois parties de plantes sont récupérées : les graines, la menue paille et la paille. On rappelle que la "ferme" peut être une zone de production plus grande qu'une exploitation agricole, par exemple l'ensemble des fermes d'une commune ou d'un canton. Le but de cette première étude de cas est de faire un test de bon fonctionnement du modèle, sur un exemple simple qui nous permet de vérifier facilement les flux, les niveaux des stocks, les différents coûts, etc.

L'horizon de planification est ici formé de 28 périodes d'un jour. La raffinerie exprime ses demandes avec des périodes d'une semaine, d'où 4 périodes-raffinerie. Le colza a une fenêtre de récolte [1,14] tandis que les états de la ferme ont une fenêtre d'ouverture [1,21]. Ainsi, les stocks de la ferme doivent être vidés au plus tard une semaine après la fin de la fenêtre de récolte. Par contre, la coopérative et la raffinerie sont ouvertes sur tout l'horizon, avec une fenêtre [1, 28].

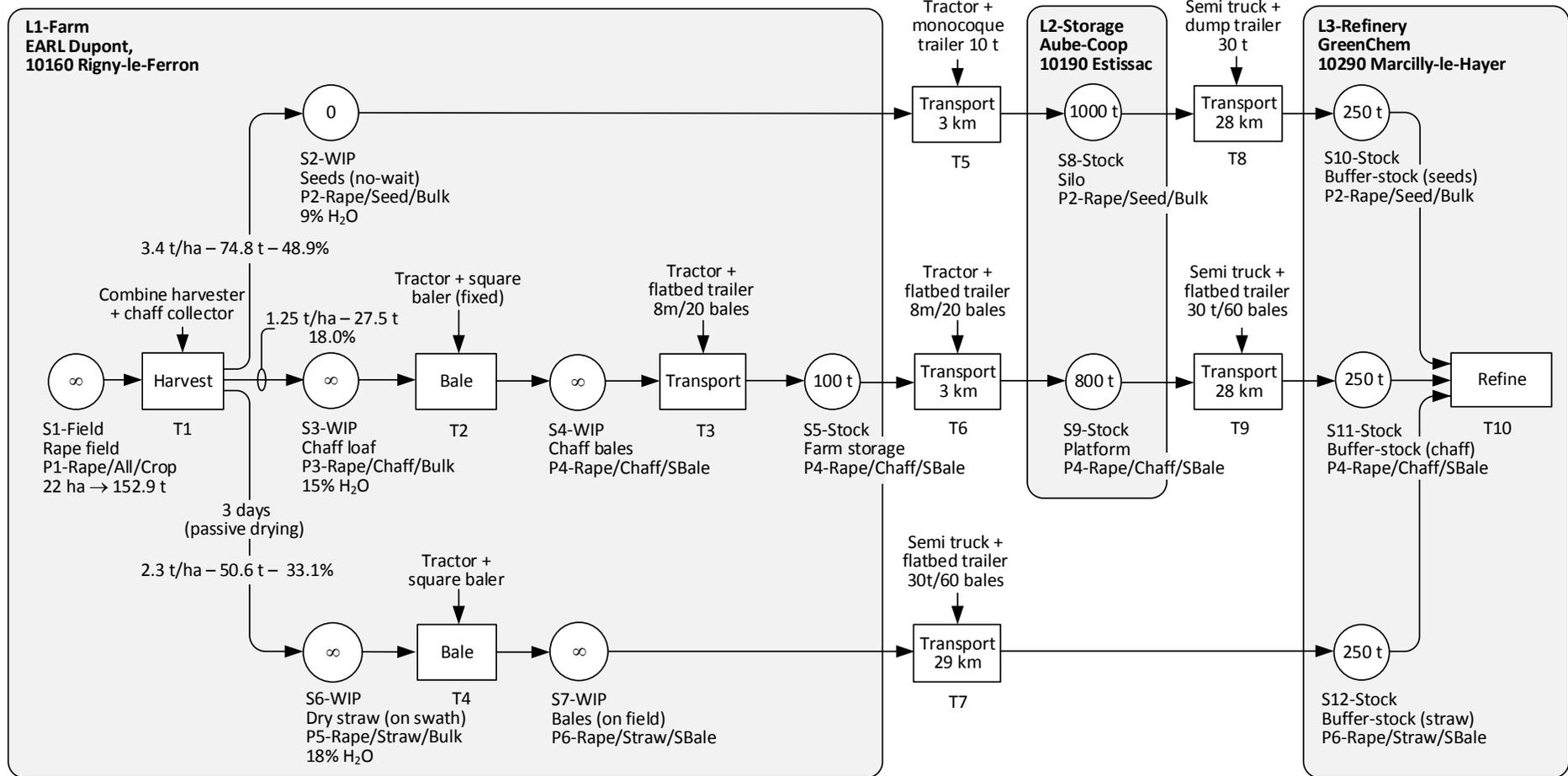


Figure 4.4. Graphe états-tâches du test à une ferme.

Implémentation du modèle de données en Excel

Les contenus des principales tables de données, dont la structure a été décrite dans le chapitre 3, sont montrés ci-dessous et dans les deux pages suivantes. Ces données ont été saisies à la main sous Excel, avec des distances fictives entre sites et de petites demandes pour la raffinerie. Rien que pour une chaîne d’approvisionnement à une ferme, on remarque que le volume de données à renseigner est déjà significatif. Ceci montre l’intérêt de la méthode de génération automatique des états et des tâches pour les fermes qui utilisent la même chaîne de récolte, décrite dans la section 2.3 du chapitre 3. Cette méthode a été implémentée pour les grandes chaînes, nous y reviendrons dans le chapitre 6.

LID	LName	LType	LLon	LLat	LAddr	LCity	LZIP	LHours
L1	EARL Dupont	Farm	3,594	48,189	Le Bois de Rigny	Rigny-le-Ferron	10160	10
L2	Aube-Coop	Storage	3,808	48,266	5 route de Bucey	Estissac	10190	10
L3	GreenChem	Refinery	3,637	48,347	D374	Marcilly-le-Hayer	10290	24

Table 4.7. Table LOCATIONS avec les trois sites (ferme, coopérative, raffinerie).

SID	SName	SType	SLoca	SProd	SBeg	SEnd	SCapa	SArea	SQBeg	SQEndMin	SQEndMax	SInvCost	SInpCost	SOutCost	SThru	SLoss	SLink
S1	Rape field	Field	L1	P1	1	14	999999	22	152,9	0	999999	0	0	0	999999	0	NA
S2	Seeds (no-wait)	WIP	L1	P2	1	21	0	0	0	0	0	0	0	0	999999	0	NA
S3	Chaff loaf	WIP	L1	P3	1	21	999999	0	0	0	0	0	0	0	999999	0	NA
S4	Chaff bales	WIP	L1	P4	1	21	999999	0	0	0	0	0	0	0	999999	0	NA
S5	Farm storage	Stock	L1	P4	1	28	100	0	0	0	0	0	2	2	50	0	NA
S6	Straw (swath)	WIP	L1	P5	1	21	999999	0	0	0	0	0	0	0	999999	0	NA
S7	Straw bales on-field	WIP	L1	P6	1	21	999999	0	0	0	0	0	0	0	999999	0	NA
S8	Silo	Stock	L2	P2	1	28	1000	0	150	0	999999	1	1,5	1,5	100	2	NA
S9	Platform	Stock	L2	P4	1	28	800	0	0	0	999999	0	0	0	50	1	NA
S10	Buffer-stock (seeds)	Stock	L3	P2	1	28	250	0	30	10	999999	2	1,5	1,5	100	2	NA
S11	Buffer-stock (chaff)	Stock	L3	P4	1	28	250	0	30	10	999999	2	3	3	50	1	NA
S12	Buffer-stock (straw)	Stock	L3	P6	1	28	250	0	40	10	999999	0	0	0	50	1	NA

Table 4.8. Table STATES stockant les états.

TID	TName	TType	TLoca	TGroup	TDist
T1	Harvest rape	Harvest	L1	CG1	0
T2	Bale chaff	Bale	L1	CG3	0
T3	Send chaff bales to farm storage	Transport	L1	CG4	1
T4	Bale straw	Bale	L1	CG2	0
T5	Send seeds to silo	Transport	L1	CG5	3
T6	Send chaff bales to platform	Transport	L1	CG4	3
T7	Send straw bales to refinery	Transport	L3	CG6	29
T8	Send seeds to refinery	Transport	L3	CG7	28
T9	Chaff bales to refinery	Transport	L3	CG6	28
T10	Convert biomass	Refine	L3	NA	0

Table 4.9. Table TASKS stockant les tâches.

TSBeg	TSEnd	TSProd	TSFrac	TSYield	TSDura
T1	S2	P2	48,90	3,40	0
T1	S3	P3	18,00	1,25	0
T1	S6	P5	33,10	2,30	3
T2	S4	P4	100,00	0,00	0
T3	S5	P4	100,00	0,00	0
T4	S7	P6	100,00	0,00	0
T5	S8	P2	100,00	0,00	0
T6	S9	P4	100,00	0,00	0
T7	S12	P6	100,00	0,00	0
T8	S10	P2	100,00	0,00	0
T9	S11	P4	100,00	0,00	0

Table 4.10. Table TS_ARCS des arcs tâches-états.

STBeg	STEnd	STProd	STFrac	STYield
S1	T1	P1	100,00	6,95
S2	T5	P2	100,00	0,00
S3	T2	P3	100,00	0,00
S4	T3	P4	100,00	0,00
S5	T6	P4	100,00	0,00
S6	T4	P5	100,00	0,00
S7	T7	P6	100,00	0,00
S8	T8	P2	100,00	0,00
S9	T9	P4	100,00	0,00
S10	T10	P2	0,00	0,00
S11	T10	P4	0,00	0,00
S12	T10	P6	0,00	0,00

Table 4.11. Table ST_ARCS des arcs états-tâches.

CPID	CPCrop	CPName	CPType	CPYield
CP1	C1	Whole rape	All	6,95
CP2	C1	Rape seed	Seed	3,40
CP3	C1	Rape straw	Straw	2,30
CP4	C1	Rape chaff	Chaff	1,25

Table 4.12. Table CROP_PARTS des parties de plantes.

PID	PName	PPart	PForm	PDens	PHumi
P1	Standing rape	CP1	Crop	0,690	00
P2	Rape seeds	CP2	Bulk	0,645	09
P3	Rape chaff (bulk)	CP4	Bulk	0,180	15
P4	Rape chaff (square bales)	CP4	SBale	0,411	15
P5	Rape straw (bulk)	CP3	Bulk	0,180	18
P6	Rape straw (square bales)	CP3	SBale	0,180	18

Table 4.13. Table PRODUCTS des produits.

RID	RName	RFam	RExam	ROCost	RCCost	RPrice	RRate	RLife	RThru	RUtility	RPower	RWTime	RSpeed	RWeight	RVolume
R1	Combine harvester 190 HP 5m	RF1	New Holland CR9080	41,9	10389	131800	10	15	1,33	44	200	8	0	0	0
R2	Square baler 90x120x250	RF3	Presse L340 John Deere	21,3	13888	103000	20	7	14,58	0	0	8	0	0	2,7
R3	Tractor 126 to 140 HP	RF2	Tractor	11,9	6792	68200	11	10	999999	28,6	130	8	25	0	0
R4	Farm flatbed trailer 8m 20 bales	RF5	Maitre pm4r 85	1	502	6800	10	15	999999	0	0	10	0	8	44,44
R5	Semi-truck 460 HP	RF6	Renault Magnum T 460 hp occasion (343kw)	5	9831,7	87500	10	10	999999	35	460	10	80	26	0
R6	Flatbed trailer 24t 60 bales	RF7	Fruehauf 19 t	1,4	1329	18 000	10	15	999999	0	0	10	0	24	133,3
R7	Dump trailer 22t	RF8	Trailor 3 essieux	1,5	4060	55000	7	15	999999	0	0	10	0	22	35
R8	Chaff collector	RF9	Récupérateur menue paille Thiéart	2,1	1596	15000	15	10	0	0	0	8	0	0	0
R9	Square baler accessory	RF10	Entonn'Up Atelier 3T	5	3466,7	17500	25	5	999999	0	0	8	0	0	0
R10	Farm mono-coque trailer	RF4	Maitre ATENA 11	1,7	1775	24050	7	15	0	0	0	10	0	11	11

Table 4.14. Table RESOURCES des équipements utilisables.

RCName	RCGroup	RExam	RCOCost	RCCost	RCThru	RCUtility	RCPower	RCWTime	RCSpeed	RCWeight	RCVolume
Combine harvester 190 HP 5m + Chaff collector	CG1	New Holland CR9080 + Récupérateur menue paille Thiéart	44	11985	1,33	44	200	8	0	0	0
Tractor 126 to 140 HP + Square baler 90x120(x250)	CG2	Tractor + Presse 900 John Deere	26,9	20680	14,58	28,6	130	8	25	0	2,7
Tractor 126 to 140 HP + Square baler 90x120 + Square baler accessory	CG3	Tractor + Presse 900 John Deere + Entonn'Up Atelier 3T	26,9	24146,7	14,58	28,6	130	8	0	0	2,7
Tractor 126 to 140 HP + Farm flatbed trailer 8m 20 bales	CG4	Tractor + Maitre pm4r 85	12,9	7294	999999	28,6	130	8	25	8	44,44
Tractor 126 to 140 HP + Farm mono-coque trailer 10t	CG5	Tractor + Maitre ATENA 11	13,6	8567	999999	28,6	130	8	25	11	11
Semi-truck 460 HP + Flatbed trailer 24t 60 bales	CG6	Renault Magnum T 460 hp occasion (343kw) + Fruehauf 19 t	1,4	11160,7	999999	35	373	10	70	24	133,33
Semi-truck 460 HP + Dump trailer 22t	CG7	Renault Magnum T 460 hp occasion (343kw) + Trailor 3 essieux	1,5	13891,7	999999	35	373	10	70	22	35

Table 4.15. Table COMBINATIONS des combinaisons de ressources utilisables.

DRefPer	DPart	DNeed
1	CP2	5
1	CP3	5
1	CP4	5
2	CP2	5
2	CP3	5
2	CP4	5
3	CP2	5
3	CP3	5
3	CP4	5
4	CP2	5
4	CP3	5
4	CP4	5

Table 4.16. Table DEMANDS des demandes de la raffinerie.

Implémentation du modèle

Le modèle mathématique a été résolu avec l'environnement de modélisation OPL-STUDIO (langage OPL et solveur CPLEX) sur un PC portable avec un processeur Intel Core i7-3740 QM à 2.7 GHz, 16 Go de RAM et Windows 7 Professionnel. En parallèle, le modèle a été implémenté dans l'environnement de modélisation GUSEK (langage GMPL et solveur GLPK). Ce logiciel libre est moins puissant qu'OPL-STUDIO mais la double implémentation a permis un débogage plus rapide.

Résultats et visualisation

La résolution est quasi-instantanée. Le solveur trouve une solution optimale de coût total 59 219 € en 0.16 secondes. Le modèle a 3286 contraintes et 1942 variables dont 178 sont entières. Le pré-solveur le réduit à 1902 contraintes et 1504 variables, dont 140 entières et 38 binaires.

Une vérification manuelle détaillée des résultats montre que toutes les contraintes sont respectées, avec une cohérence totale des niveaux des stocks et des flux de produits.

La solution n'est pas évidente. D'un côté, la table STATES montre que la raffinerie a des stocks initiaux de 30, 30 et 40 tonnes pour les graines, les balles de menue paille et les balles de paille mais doit laisser un stock final de 10 tonnes pour chaque partie de plante. Ces tonnages incluent l'humidité. De l'autre côté, la table DEMANDS montre que la raffinerie a besoin de 5 tonnes sèches de chaque partie de plante, chaque semaine. Bien que la coopérative ait 150 tonnes de stock initial pour les graines, on voit au moins un problème avec la menue paille : la raffinerie a besoin de 20 tonnes sèches mais son stock d'entrée ne contient que 20 tonnes de balles à 15% d'humidité. Le modèle doit donc déclencher la récolte de 44,28 tonnes de colza et le transport à la raffinerie.

La figure 4.5 montre un exemple de listing de résultats détaillés que nous avons réalisé avec GUSEK.

MINIMUM TOTAL COST	:	59218.78	
Harvest/Rake/Bale cost:		296.67	
Transportation cost	:	20.66	
Equipment fixed cost	:	53629.40	
Storage cost	:	5050.47	
Inventory input cost	:	87.20	
Inventory output cost	:	134.37	
RESULTS FOR SITE: EARL Dupont			
Required equipment and utilization times:			
1 unit R1	: Combine harvester 190 HP 5m	(4.79 hours)	
1 unit R2	: Square baler 90x120x250	(3.19 hours)	
1 unit R3	: Tractor 126 to 140 HP	(4.47 hours)	
1 unit R4	: Farm flatbed trailer 8m 20 bales	(0.32 hours)	
1 unit R8	: Chaff collector	(4.79 hours)	
1 unit R9	: Square baler accessory	(1.12 hours)	
1 unit R10	: Farm monocoque trailer	(0.96 hours)	
Activity periods for state: S1 /Rape field			
Periods	:	14	
Last stock	:	152.90	
Inputs	:	0.00	
Outputs	:	44.28	
New stock	:	108.62	
Storage cost:	0.00	Total: 0.00	
Input cost	: 0.00	Total: 0.00	
Output cost	: 0.00	Total: 0.00	
Activity periods for state: S5 /Farm storage			
Periods	:	19	28
Last stock	:	0.00	7.97
Inputs	:	7.97	0.00
Outputs	:	0.00	7.97
New stock	:	7.97	0.00
Storage cost:	0.00	0.00	Total: 0.00
Input cost	: 15.94	0.00	Total: 15.94
Output cost	: 0.00	15.94	Total: 15.94

Figure 4.5. Exemple de listing de résultats détaillés (extrait).

Nous avons aussi réalisé en Visual Basic sous Excel (VBA) un outil pour visualiser les flux, les niveaux des stocks et les ressources employées dans chaque période, directement sur le dessin de la chaîne. La figure 4.6 montre avec cet outil le système à la période 14. 44,27 tonnes de colza sont récoltées pour obtenir 21,65 tonnes de graines de colza et 7,97 tonnes de menue-paille. On remarque qu'il n'y a pas de paille en période 14, ce qui s'explique car la paille de colza doit sécher 3 jours avant d'être disponible pour être mise en balles. En passant à la période 17, on pourrait voir une quantité de 14,65 tonnes de paille passant de T1 à S6.

4.2 Etude de cas à 10 fermes sur 364 jours

Nous considérons cette fois-ci deux plantes, le colza et le miscanthus. Comme dans l'étude précédente, le colza fournit des graines, de la paille et de la menue paille. Le miscanthus est fauché pour donner uniquement des balles de paille, selon la chaîne de récolte de la figure 3.5 du chapitre 3. L'horizon de planification est d'un an découpé en 52 semaines de 7 jours, ce qui donne 364 périodes.

Nous définissons un schéma d'approvisionnement de 50 730 tonnes/an de matières sèches avec 10 zones de production ("fermes"), 3 sites de stockage (coopératives) et une bioraffinerie à Venette près de Compiègne. Chaque ferme peut livrer le stockage le plus proche plus la raffinerie. Les sites sont au maximum à 50 km de la raffinerie et les plus courts chemins par la route sont calculés avec le logiciel MapPoint de Microsoft. Les demandes pour les quatre parties de plantes sont données pour chaque période-raffinerie d'une semaine. Il est hors de question de lister ici toutes les données comme dans l'exemple précédent : nous montrons uniquement la liste des différents sites avec leurs fenêtres de récolte (table 4.17) et la répartition de la demande totale selon les parties de plantes (table 4.18). Ces tonnages annuels correspondent aux raffineries de proximité envisagées dans le futur.

Identifiant	Type de site	Libellé	Commune	Fenêtres de récolte
L1	Refinery	Venette Refinery	Venette	[1,364]
L2	Storage	Acolyance Dammard	Dammard	[1,364]
L3	Storage	Agora Bresles	Bresles	[1,364]
L4	Storage	Agora Guiscard	Guiscard	[1,364]
L5	Farm	Essuiles	Essuiles	[196,271]
L6	Farm	Mouy	Mouy	[69,144]
L7	Farm	Catigny	Catigny	[189,264] , [62,77]
L8	Farm	Villeselve	Villeselve	189,264] , [62,77]
L9	Farm	Lassigny	Lassigny	189,264] , [62,77]
L10	Farm	Ravenel	Ravenel	[196,271]
L11	Farm	Troësnes	Troësnes	[69,144]
L12	Farm	Brumetz	Brumetz	[203,278], [76,91]
L13	Farm	Raray	Raray	[203,278], [76,91]
L14	Farm	Attichy	Attichy	[203,278], [76,91]

Table 4.17. Listes des sites et fenêtres de récolte.

Delete
Mise en forme
Horizon (jours) **28**
Période à saisir **14**
Affichage des résultats

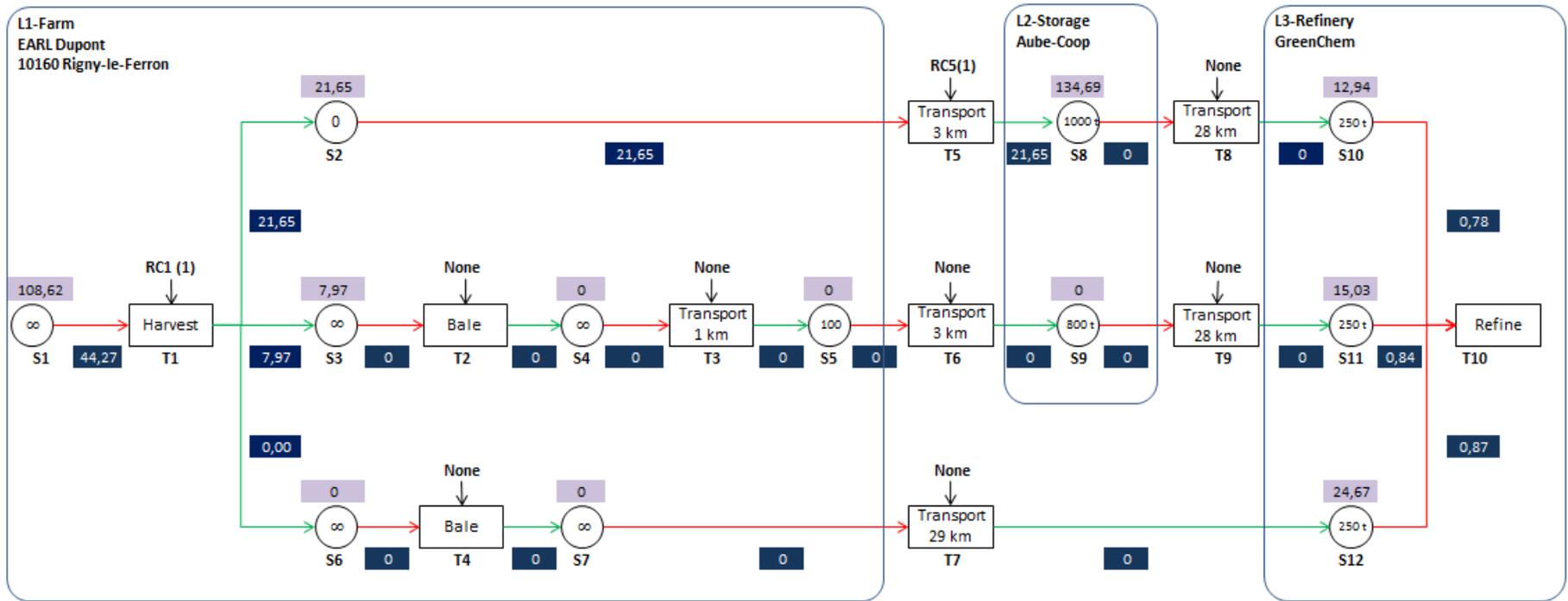


Figure 4.6. Outil en VBA sous Excel montrant les flux, les niveaux des stocks et les combinaisons de ressources utilisées en période $t = 14$.

Plante	Partie de plante	Demande annuelle (tonnes sèches)
Colza	Graines	17690
	Paille	14120
	Menue paille	6920
Miscanthus	Paille	12000

Table 4.18. Demandes annuelles de la raffinerie en parties de plantes.

La raffinerie et les coopératives (sites L1 à L4) sont ouvertes toute l'année. Les fermes cultivent soit du colza, soit du miscanthus, soit les deux. Les fenêtres finissant avant la période 144 concernent le miscanthus, qui est fauché en mars-avril. Celles commençant à partir de la période 189 concernent le colza, moissonné en juillet. Les fenêtres de récolte varient un peu selon les types de sols.

Chaque ferme a une chaîne de récolte prédéfinie. Les sites L5 et L10 (colza) utilisent la chaîne de récolte de la figure 3.4 du chapitre 3. Les sites L6 et L11 (miscanthus) emploient la chaîne de récolte de la figure 3.5. Les autres sites de production (colza + miscanthus) combinent les deux chaînes précédentes avec un stock partagé pour les balles, comme illustré dans la figure 3.6 du chapitre 3.

Le nombre de fermes (10) semble modeste mais rappelons-nous que, rien que pour le colza, chaque ferme utilise 10 tâches et 12 états, avec des variables de stock et de flux indexés sur 364 jours. En fait, le modèle résultant a 588 304 contraintes et 500 094 variables, dont 91 437 entières. Après le pré-solveur, il reste 87 445 contraintes et 94 883 variables, dont 31 575 entières mais non binaires et 9 826 binaires. La résolution est ensuite assez rapide : on trouve un coût total de 4 829 586 € en 164 secondes. En effet, CPLEX peut utiliser jusqu'à 7 threads en parallèle sur notre PC Intel Core i7.

La figure 4.7 montre la répartition des coûts. Les coûts fixes des équipements sont comptés dans les coûts de récolte. Les coûts de stockage représentent 59%, alors qu'ils sont souvent négligés dans la littérature. Ce pourcentage est normal car les périodes de récoltes sont courtes (15 jours environ pour chacune des deux plantes) tandis que la raffinerie consomme les biomasses toute l'année : il est donc nécessaire de constituer des stocks importants. En particulier, il faut des stocks initiaux suffisants en début d'année, avant la récolte de miscanthus et de colza, sous peine d'infaisabilité.

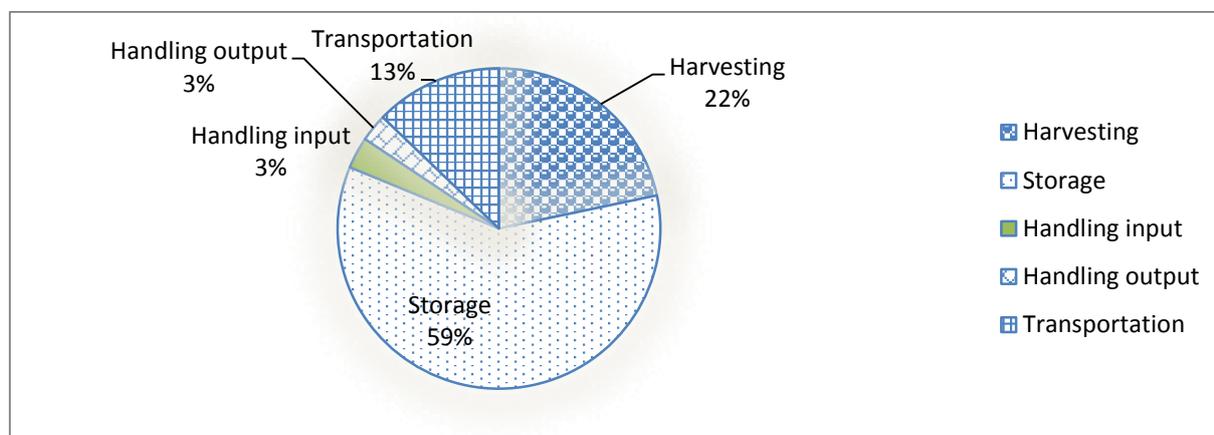


Figure 4.7 Répartition des coûts dans la chaîne d'approvisionnement.

La figure 4.8 montre que CPLEX utilise beaucoup de mémoire pour charger le modèle, exécuter l'algorithme du simplexe pour avoir une première solution, puis traiter les nœuds générés par son algorithme de branch-and-bound. La consommation de mémoire monte alors à 14,7 Go. En fait, 11 Go sont utilisés pour le modèle, le reste pour Windows et l'environnement de développement OPL-STUDIO.

La mémoire presque entièrement utilisée ne signifie pas qu'on a atteint la taille maximale de problème résoluble par le solveur. En effet, CPLEX s'étale dans la mémoire disponible mais peut si nécessaire mettre en attente des nœuds de l'arborescence de recherche sur le disque dur. Cependant, nous avons construit un cas à 100 fermes et 25 sites de stockage qui ne peut pas être résolu, même en utilisant des périodes plus longues d'une semaine (un an décomposé en 52 semaines) : le pré-solveur de CPLEX explose avec une erreur "out of memory".

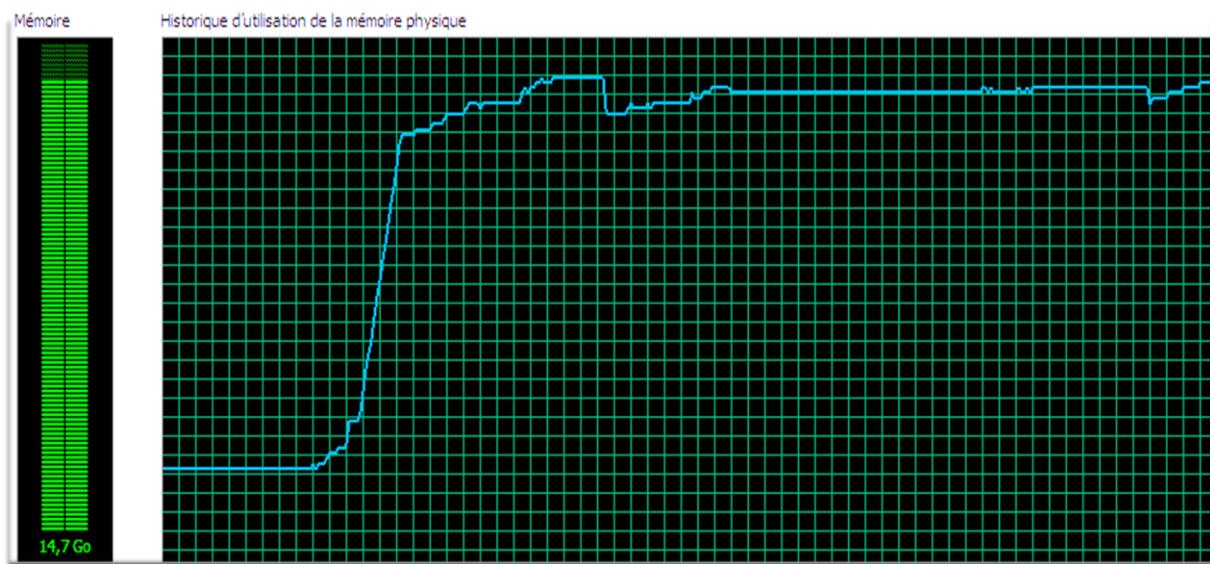


Figure 4.8. Mémoire au chargement du modèle et pendant le branch-and-bound de CPLEX.

Note : 6 s par graduation. Résolution terminée mais mémoire non encore libérée par OPL STUDIO.

5. Simplification de la gestion de ressources et limites atteintes

Le modèle dit "complet" présenté dans les sections précédentes détaille parfaitement l'ensemble des activités de la chaîne d'approvisionnement et, en particulier, les activités de récolte. Cependant, l'étude de cas à 10 fermes sur 364 jours nécessite déjà un grand nombre de variables et de contraintes ainsi qu'une consommation importante de mémoire centrale, même si par ailleurs le temps de calcul est acceptable (environ 3 minutes). Ces difficultés résident en grande partie au niveau de la gestion des équipements et des variables entières associées.

Nous proposons donc ici des simplifications du système de gestion de ressources pour éliminer les variables entières. Par contre, le graphe états-tâches contient toujours les mêmes nœuds. Le chapitre 5 décrit une simplification plus drastique où les états et les tâches de chaque ferme sont condensés en une macro-tâche unique.

5.1 Utilisation de coûts horaires "tout compris"

Le modèle complet précédent suppose que les équipements ne sont pas encore achetés. Il détermine donc le nombre d'équipements $FS_{l,r}$ à acheter pour chaque site. Cette stratégie donne des solutions où les investissements en matériels constituent une partie importante des coûts, voir le listing de la figure 4.5. Or, dans la réalité, l'agriculteur utilise ses machines pour d'autres cultures que l'activité biomasse pour la raffinerie et il devient difficile de calculer le nombre d'équipements à acheter.

Les chambres d'agriculture ont mis au point les barèmes d'entraide, conçus initialement pour des échanges d'équipements entre agriculteurs ou pour des situations où des agriculteurs viennent aider un collègue pour accélérer sa récolte (voir "barème d'entraide" dans la bibliographie). L'intérêt de ces barèmes est de fournir des coûts horaires d'utilisation "tout compris", incluant l'amortissement du matériel. On peut donc optimiser un coût total, composé des heures d'utilisation multipliées par les coûts horaires des barèmes d'entraide. On obtient alors les coûts qu'on peut réellement imputer à la chaîne de biomasse pour la raffinerie.

5.2 Elimination des variables entières

Le modèle complet utilise des variables entières $FS_{l,r}$ pour le nombre de ressources r à acheter pour le site l . De plus, le nombre de ressources utilisables sur chaque site l et chaque famille de ressources f est limité par une constante $LFNMax_{l,f}$ (table LOCA_FAM), par exemple pas plus de 2 tracteurs tous modèles confondus.

Nous proposons de supprimer les variables $FS_{l,r}$ pour les remplacer par $LFNMax_{l,RFam[r]}$. Le solveur va comme d'habitude déterminer pour chaque tâche et chaque période les combinaisons de ressources à mobiliser et leurs heures d'utilisation puis, connaissant les ressources dans chaque combinaison, les heures d'utilisation des ressources. En regroupant les ressources par famille, il s'assurera qu'on ne dépasse pas le nombre d'heures possibles pour les $LFNMax_{l,f}$ équipements. Enfin, les coûts d'utilisation, amortissement inclus, seront déduits à partir des barèmes d'entraide.

Il est important de noter que les équipements peuvent être déjà achetés, mais pas forcément. S'ils ne le sont pas, le nombre minimal de ressources nécessaires, qui peut être inférieur à $LFNMax_{l,f}$, est calculable après la résolution du modèle en utilisant les équations (LB_FS dans la section 3.6) qui permettaient de calculer les $FS_{l,r}$.

Ensuite, nous évitons les variables entières $RT_{i,c,t}$ pour le nombre de rotations des véhicules, en supposant que le transport est assuré par un prestataire qui facture des tonnes x km. Cette option est fréquente en pratique mais il est très difficile d'obtenir des coûts de la part des transporteurs si on n'est pas agriculteur. Nous avons donc estimé le coût par tonne x km en supposant que le camion est chargé au maximum. C'est souvent le cas, vu les tonnages importants qui sont transportés, surtout si au lieu des fermes on considère des zones de production agrégées comme les communes. Notre estimation va modifier un peu le coût optimal puisque la dernière rotation peut s'effectuer en réalité avec une charge partielle. Nous verrons dans les tests que l'erreur est acceptable.

Pour cette simplification, nous n'avons pas voulu modifier la base de données. Pour une tâche de transport i à partir d'un état e et un véhicule c , nous avons à notre disposition le coût d'utilisation $RCOCost_c$ en €/h, la vitesse $RCSpeed_c$ en km/h, le poids maximum $RCWeight_c$ en tonnes, le volume maximum $RCVolume_c$ en m³ et la densité $PDens_{SProd(e)}$ du produit transporté. Le coût en tonnes x km est donc calculable comme suit. Le second terme du minimum est le poids maximal en fonction du volume du véhicule et de la densité du produit. La somme permet de récupérer le prédécesseur e (unique) de la tâche de transport.

$$tkm = \frac{RCOCost_c}{RCSpeed_c \cdot \min(RCWeight_c, \sum_{(e,i) \in STA} RCVolume_c \cdot PDens_{SProd(e)})}$$

Les contraintes (GetTCT) pour le nombre d'heures de transport se réécrivent :

$$\forall i \in TID | TType_i \in ZT, \forall c \in RCID | RCGroup_c = TGroup_i, \forall t \in H :$$

$$TCT_{i,c,t} = \frac{Amp_{i,c,t}}{\min(RCWeight_c, \sum_{(e,i) \in STA} RCVolume_c \cdot PDens_{SProd(e)})} \cdot \frac{TDist_i}{RCSpeed_c}$$

Le terme $TransOpCosts$ de la fonction-objectif du 3.3 (coût total de transport) est théoriquement :

$$TransOpCosts = \sum_{i \in TID | TType_i \in ZT} \sum_{c \in RCID | RCGroup_c = TGroup_i} \sum_{t \in H} Amp_{i,c,t} \cdot tkm$$

Mais en remplaçant tkm par son expression on retrouve le coût de transport de la version non simplifiée :

$$TransOpCosts = \sum_{i \in TID | TType_i \in ZT} \sum_{c \in RCID | RCGroup_c = TGroup_i} \sum_{t \in H} TCT_{i,c,t} \cdot RCOCost_c$$

Donc, il n'est pas nécessaire de modifier l'expression du coût de transport dans la fonction-objectif !

5.3 Modifications résultantes du modèle complet

Les variables entières $FS_{l,r}$ et $RT_{i,c,t}$ sont toutes supprimées.

Comme on utilise maintenant les coûts d'utilisation "tout compris" des barèmes d'entraide, le coût fixe des ressources est enlevé de la fonction-objectif :

$$FixedCosts = \sum_{l \in LID} \sum_{r \in RID | (l, RFam_r) \in LFPairs} FS_{l,r} \cdot RCCost_r$$

Les contraintes de stockage, de conservation des flux, et de satisfaction des demandes de la raffinerie restent inchangées. Par contre, les contraintes de gestion des ressources de la section 3.6 subissent les modifications suivantes :

- Les contraintes permettant de borner les variables $RT_{i,c,t}$ (contraintes LB_RT et UB_RT) sont retirées du modèle du fait de la suppression de ces variables.

- Les contraintes permettant de borner les variables $FS_{l,r}$ (contraintes LB_FS, UB1_FS et UB2_FS) sont également retirées du modèle du fait de la suppression de ces variables. Les contraintes LB_FS sont remplacées par :

$$\forall r \in RID, \forall (l, RFam_r) \in LFPairs, \forall t \in H : \frac{TR_{l,r,t}}{\min(RWTime_r, LHours_l)} \leq LFNMax_{l,RFam[r]}$$

- Les contraintes (GetTCT) pour les heures de transport sont modifiées comme indiqué dans la section précédente.

5.4 Résultats avec les simplifications

Après ces simplifications, CPLEX doit résoudre un programme linéaire à variables continues. Il peut donc utiliser l'algorithme du simplexe au lieu de sa méthode de branch-and-bound.

Nous avons résolu à nouveau le test à 10 fermes et 3 stockages sur 364 jours, pour vérifier l'impact des simplifications. Les nombres de variables et de contraintes sont divisées grosso-modo par deux et, évidemment, il n'y a plus de variables entières. Avant le pré-solveur, le modèle comporte 265 245 variables et 212 760 contraintes. Le pré-solveur le compresse à 26 580 variables et 9 791 contraintes. Quant au temps de résolution, il passe de 164 secondes à moins d'une seconde.

La solution obtenue est très similaire à celle sans simplifications, avec une grosse différence de coût total égale aux coûts fixes des équipements puisqu'ils ne sont plus comptés. Si on regarde les autres coûts, on constate une différence mineure de 0,53% qui vient de notre approximation consistant à considérer des camions pleins pour estimer le coût de transport par tonne x km.

Ces simplifications ont permis aussi de résoudre le cas à 100 fermes, 25 stockages et 52 périodes d'une semaine, qui ne passait pas dans le modèle complet. Le programme linéaire a 162 074 variables et 203 721 contraintes. Après passage au pré-solveur, il n'a plus que 8 339 variables et 5 839 contraintes. Le temps de calcul atteint cependant 8 158 secondes (2 heures et 16 minutes), une durée anormalement longue et inexplicée pour un programme linéaire sans variables entières de cette taille. Après tâtonnement par dichotomie, on a pu atteindre 150 fermes mais pas plus. Cette limite a encouragé le développement d'un modèle dit "compact" dans le chapitre 5.

5.5 Autres simplifications envisagées

A part la simplification drastique présentée dans le chapitre suivant et qui consiste à agréger les tâches d'une chaîne de récolte en une macro-tâche unique, nous avons envisagé des restrictions sur le nombre de combinaisons de ressources utilisables par chaque tâche.

Le modèle complet impose un seul groupe de combinaisons par tâche mais plusieurs combinaisons peuvent être utilisées dans ce groupe. Par exemple, une tâche de transport par l'agriculteur a le groupe "tracteur + remorque". Si la base de données contient pour ce groupe les combinaisons "tracteur 200 ch + remorque-plateau 12 t" et "tracteur 250 ch + remorque-benne 15 t", la tâche peut utiliser ces deux combinaisons au cours de l'horizon, y compris dans une même période.

Un premier niveau de restriction consiste à utiliser une seule combinaison par tâche dans chaque période, même si on peut changer de combinaison d'une période à l'autre. Ce cas correspond par exemple à un agriculteur qui travaille seul sur son exploitation et qui va utiliser le même tracteur et la même remorque toute la journée pour éviter de changer d'attelage. Un niveau de restriction plus fort est de pouvoir choisir une combinaison par tâche, mais la même pour toutes les périodes.

A la réflexion, nous ne voyons pas comment formuler ces contraintes sans variables binaires. Comme visiblement les variables entières sont responsables des temps de calcul élevé du modèle complet, nous avons écarté ces simplifications qui impliquent des variables entières supplémentaires.

Une simplification plus intéressante est d'imposer dans les données une combinaison de ressources à chaque tâche i . Ceci permettrait de supprimer l'indice de combinaison c dans la quantité $Amp_{i,c,t}$ traitée par la tâche en période t , et donc de réduire le nombre de ces variables.

En fait, si on veut imposer une combinaison de ressources à chaque tâche, il n'y a pas de raison de changer cette combinaison selon le site. Par exemple, un utilisateur pressé du modèle préférera saisir une seule combinaison du genre "tracteur 200 ch + remorque 12 t" dans la base de données et l'imposer à toutes les tâches de transport effectuées par les agriculteurs, quelle que soit la ferme.

Nous avons testé cette méthode consistant à saisir dans la base de données une seule combinaison de ressources pour chaque groupe. Que l'on supprime ou pas l'indice c des variables $Amp_{i,c,t}$ dans le modèle, le temps de résolution ne change pas. Notre hypothèse est que le pré-processeur d'OPL STUDIO est si puissant qu'il s'aperçoit que chaque tâche i utilise un groupe $TGroup_i$ mais que ce groupe ne contient en fait qu'une seule combinaison : l'indice c est enlevé de toute façon.

6. Conclusion

Nous avons élaboré un programme linéaire en variables mixtes pour optimiser le coût total d'une chaîne d'approvisionnement en biomasse pour une bioraffinerie. Grâce aux graphes états-tâches et aux extensions réalisées, nous sommes parvenus à un niveau de modélisation très fin et le modèle mathématique devient largement indépendant de la structure de la chaîne. Ce modèle constitue une contribution majeure car, à notre connaissance, il combine plusieurs caractéristiques qu'aucun autre modèle de la littérature ne peut traiter simultanément.

Rappelons ces principales caractéristiques :

- Modèle permettant de gérer plusieurs plantes, plusieurs parties de la même plante, et plusieurs produits (aspects et conditionnements) par partie de plante.
- Plusieurs chaînes de récoltes possibles pour chaque plante.
- Par rapport à la littérature, gestion des équipements de récolte et de transport beaucoup plus fine grâce aux combinaisons de ressources. En particulier, nous pouvons déterminer les équipements nécessaires sur chaque site et leurs coûts.
- Modèle de réseau basé sur les graphes états-tâches, avec des extensions : possibilité de gérer des stocks multi-produits, des états sans attente, des attentes en stock, des fenêtres d'ouverture des états (la littérature ne considère des fenêtres que pour les tâches).

- Génération automatique des équations du modèle à partir du graphe états-tâches et des autres données ("data driven model").
- Grande flexibilité. Par exemple, le modèle est déjà prêt pour inclure plusieurs raffineries et des installations de prétraitement élaboré (fabrication de granulés ou de briquettes, torréfaction...). La granularité spatiale et temporelle est aussi facilement modifiable : on peut ainsi regrouper les fermes par commune ou canton et utiliser des périodes d'une semaine au lieu d'un jour.

Cependant, les limites des solveurs actuels et du PC utilisé, pourtant un Intel Core i7 avec 16 GB de RAM, ne permettent pas d'aller au-delà de 150 fermes sur un an découpé en semaines. Cette taille est déjà impressionnante mais la raffinerie envisagée près de Compiègne va s'alimenter sur plus de 400 communes de l'Oise et de l'Aisne. Nous verrons donc dans le chapitre suivant un modèle plus compact basé sur un système de macro-tâches qui permet de résoudre des instances plus grandes.

Ce chapitre a fait l'objet de trois conférences internationales (dont deux avec actes), de deux conférences nationales et d'un article soumis à une revue.

Références :

Ba, B.H., Prins, C., Prodhon C., 2014. Un modèle tactique pour une chaîne d'approvisionnement de biomasse. 15ème congrès annuel de la Société Française de Recherche Opérationnelle et d'Aide à la Décision (ROADEF), Bordeaux, France, 26-28/02/2014.

Ba, B.H., Prins, C., Prodhon, C., 2014c. An overview of operations research models for biomass supply chains, ICORES 2014 (3rd Int. Conf. on Operations Research and Enterprise Systems), Angers, 06-08/03/2014, 8 pages.

Ba, B.H., Prins, C., Prodhon C., 2015a. Modèle tactique d'approvisionnement en biomasse pour une bio-raffinerie de proximité. 16ème congrès annuel de la Société Française de Recherche Opérationnelle et d'Aide à la décision (ROADEF), Marseille, France, 25-27/02/2015.

Ba, B.H., Prins, C., Prodhon C., 2015b. A new tactical optimization model for bioenergy supply chain. XIII International Conference on Applied Mathematics and Operation Research (ICAMOR 2015), Miami Beach, 9-10/03/2015, USA, 12 pages.

Ba, B.H., Prins, C., Prodhon C., 2015. A general mixed integer linear model for a local biorefinery supply chain. 27th European Conference on Operational Research (EURO 2015) 12-15/07/2015, University of Strathclyde, Glasgow, United Kingdom.

Ba, B.H., Prins, C., Prodhon C., 2015. A generic tactical planning model to supply a bio-refinery with biomass. Soumis à *European Journal of Operational Research*.

Chapitre 5

Modèle mathématique compact

1. Introduction

Dans le chapitre précédent, nous avons conçu un modèle mathématique dit "complet" qui est très précis dans sa prise en compte des activités de récolte au niveau des fermes. Pourtant, même avec les simplifications de la section 5 du chapitre 4 (élimination des variables entières, utilisation de coûts "tout compris"), il présente des limites pour traiter des chaînes logistiques avec un nombre élevé de nœuds. Le développement d'un modèle comportant moins de variables et de contraintes s'impose donc pour pouvoir résoudre des grands cas. Cependant, ce processus implique de trouver un compromis entre finesse de modélisation, temps de calcul et qualité des solutions obtenues. Mais comment parvenir à un tel modèle tout en gardant le paradigme des graphes états-tâches?

En s'appuyant toujours sur le modèle de réseau et le modèle de données vu dans le chapitre 3, nous proposons une modélisation dite "compacte" du problème d'approvisionnement, par opposition au modèle dit "complet" du chapitre 4. Le modèle mathématique qui en résulte est un programme linéaire à variables continues. La réduction de taille est obtenue en agrégeant dans chaque ferme les tâches d'une chaîne de récolte et les états intermédiaires en une tâche unique. Cette compression est importante car la grande majorité des sites des chaînes logistiques étudiées sont des fermes : en comparaison, il y a beaucoup moins de sites de stockage et une seule raffinerie. On peut donc s'attendre à une forte diminution des temps de calcul, pour un nombre de fermes donné, et à une augmentation des tailles de chaînes pouvant être optimisées, pour un temps de calcul fixé.

La section 2 présente de manière pédagogique le principe de la simplification et les conditions de sa validité, puis les modifications à effectuer dans la base de données et dans le modèle complet. La section 3 présente ensuite l'étude d'un cas réel et en vraie grandeur pour l'évaluation de ce modèle. Les résultats sont analysés et commentés. La section 4 propose des perspectives et extensions pouvant être réalisées. Enfin, la section 5 conclut le chapitre en effectuant un bilan.

2. Développement du modèle mathématique compact

Le modèle compact utilise encore le formalisme des graphes états-tâches. Il reprend tout d'abord les simplifications déjà présentées dans la section 5 du chapitre 4 : remplacement des coûts fixes par les coûts "tout compris" des barèmes d'entraide, suppression des variables entières, notamment le nombre de rotations remplacé un calcul des coûts de transport à la tonne x kilomètre.

Pour pousser plus loin la simplification, une agrégation de certains états et de certaines tâches a été introduite au niveau des fermes. Ceci induit ce que nous appelons des *macro-tâches*, qui constituent chacune un regroupement de plusieurs activités fines réalisées au niveau de la ferme (moisson, andainage, séchage, etc.) en une seule tâche. Dans cette section, nous allons d'abord présenter le principe et ensuite décrire l'impact sur la base de données et sur le modèle mathématique complet du chapitre 4.

2.1 Principe de la simplification

L'idée est de regrouper les tâches et les états intermédiaires de chaque chaîne de récolte dans une ferme pour former une macro-tâche, qu'on peut aussi appeler *tâche de récolte*. Cette macro-tâche regroupe et remplace les tâches d'origine concernant la moisson ("harvest"), l'andainage ("raking"), le pressage de balles ("baling") et les transports au sein de la ferme. Elle s'intercale donc entre l'état de type "Field", qui représente le gisement de biomasse au début de la chaîne de récolte, et le stockage local à la fin de cette chaîne. Notons qu'il faut conserver l'état qui modélise le stockage local de la ferme, car sinon la macro-tâche se retrouve reliée à des tâches de transport. Or, deux tâches ne peuvent pas être connectées directement dans le formalisme des graphes états-tâches.

Nous avons vu dans le chapitre 3 que le modèle complet spécifie un seul groupe de combinaisons de ressources *TGroup* pour chaque tâche, comme on peut le vérifier dans le chapitre 3 sur la figure 3.2 (chaîne logistique simple à une ferme) et dans la description de la table des tâches (TASKS). Une conséquence directe des macro-tâches est de nécessiter plusieurs groupes de combinaisons, celles des tâches d'origine. Le lecteur ayant oublié les différences subtiles entre ressources, combinaisons de ressources, familles de ressources et groupes de combinaisons de ressources est invité à relire la section 4 du chapitre 3 où ces concepts sont définis.

Pour fixer les idées, prenons le cas du colza. Les opérations de récolte mises en œuvre sur cette culture demandent plusieurs groupes de combinaisons de ressources, à savoir un groupe "moissonneuse batteuse + collecteur de menue paille" pour couper la plante et récupérer la menue paille, un groupe "tracteur + presse à balles" pour faire des balles de paille et de menue paille, et enfin un groupe "tracteur + remorque" pour le transport des balles au hangar de la ferme.

Il est important de noter à ce stade que, contrairement à une tâche simple qui peut utiliser plusieurs combinaisons de ressources mais dans un même groupe imposé, il apparaît clairement dans cet exemple (et ce sera généralement le cas) que les macro-tâches peuvent utiliser plusieurs groupes de combinaisons et plusieurs combinaisons dans chaque groupe.

Une autre caractéristique importante doit être soulignée. Contrairement aux tâches vues jusqu'à présent, les groupes de combinaisons utilisées dans une macro-tâche ne traitent pas tous la quantité totale en entrée. Reprenons l'exemple du colza : sur la quantité coupée au niveau d'une ferme, 100% est traitée par la moissonneuse, mais seulement 33,1%, en l'occurrence la paille, est mise en balles par le groupe de combinaisons "tracteur + presse à balles". Ainsi, pour bien gérer les quantités de produits fournies par les différentes combinaisons d'une macro-tâche, il faut préciser le pourcentage du flux d'entrée de la tâche qui est traitée par chaque groupe.

Il faut bien comprendre que considérer des macro-tâches à la place des tâches fines de récolte peut rendre la résolution sous-optimale. Par exemple, les attentes de la paille ou de la menue paille avant la mise en balles disparaissent, contrairement au modèle complet. En effet, les états représentant la paille en attente sur le sol sont agrégés avec les tâches de récolte dans la macro-tâche.

En réalité, on peut s'attendre à une faible perte de qualité des solutions. En effet, notre modélisation macroscopique de la récolte correspond au comportement réel des agriculteurs : ces derniers préfèrent éviter d'interrompre le processus une fois la récolte commencée, par peur d'un changement de météo. En considérant une granularité pas trop fine, par exemple hebdomadaire, on peut donc supposer que la quantité récoltée est mise en balles au sein de la même période.

Cela semble contestable si la "ferme" est une grande zone de culture comme une commune : les moissonneuses récoltant une grosse quantité, il semble difficile de tout mettre en balles la même semaine. En fait, bien que les moissonneuses travaillent sur des périodes consécutives et que l'ensemble des champs soit traitée d'affilée, les activités suivantes comme la mise en balles peuvent se réaliser en parallèle. En effet, si la moisson d'un champ doit être terminée pour débiter les étapes suivantes, on n'a pas besoin d'attendre la fin de la moisson complète pour enchaîner les autres activités. Dans la pratique, les agriculteurs s'entraident et concentrent chaque jour quelques moissonneuses sur certains champs. Le lendemain ces moissonneuses vont sur d'autres champs pendant que les presses à balles interviennent tout de suite sur les champs déjà moissonnés.

Ajoutons que certaines plantes nécessitent un temps de séchage au sol. Si ce dernier est plus court que la durée de la période étudiée, comme c'est le cas avec les trois jours requis pour le colza dans la granularité hebdomadaire, cette étape est considérée comme négligeable.

En résumé, la macro-tâche hérite des tâches internes avec un ensemble de groupes de combinaisons. Certains de ces groupes traitent tout le flux d'entrée (moissonneuse) et d'autres travaillent sur des fractions concernant les sous-produits de la moissonneuse.

La flexibilité et le pragmatisme de notre modélisation sont encore clairement mis en évidence ici, car le changement d'échelle et l'introduction de la macro-tâche ne change pas beaucoup le principe du modèle complet présenté dans le chapitre précédent, à l'exception de la gestion des équipements, comme on va le voir. En tout cas, le modèle de réseau obtenu reste un graphe état-tâche valide.

2.2 Modifications dans la base de données

L'implémentation du système de macro-tâche a d'abord un impact sur la base de données. En effet, la table des tâches du chapitre 3 (TASKS) permet de spécifier un seul groupe par tâche. Une nouvelle table que nous appelons TASK_GROUPS est nécessaire pour pouvoir associer à chaque macro-tâche les différents groupes qu'elle utilise, avec des pourcentages.

Dans TASKS, la macro-tâche remplace la tâche "Harvest" d'origine. Les autres tâches regroupées dans la macro-tâche sont supprimées. L'attribut *TGroup* devient inutile pour les macro-tâches, alors qu'il reste utilisé pour les autres tâches qui nécessitent habituellement un seul groupe de combinaisons (transport en dehors des fermes par exemple). Pour éviter cette différence, nous avons préféré uniformiser : même les tâches de transport externes peuvent utiliser plusieurs groupes, elles sont donc mises dans la table TASK_GROUPS avec les macro-tâches et *TGroup* peut être supprimé.

Nous allons définir cette table comme nous l'avons fait pour les autres tables dans le chapitre 3.

Définition. La table TASK_GROUPS permet d'associer, à chaque groupe de combinaisons utilisé dans une tâche, la fraction de la quantité en entrée de tâche qui est traitée par ce groupe. Cette fraction va servir dans le modèle mathématique compact à déterminer les quantités traitées par chaque groupe puis le temps de travail des combinaisons choisies dans les groupes.

- TID Identifiant de la tâche, qui doit exister dans la table des tâches (TASKS).
- CGID Identifiant du groupe de combinaisons, qui doit exister dans la table GROUPS.
- TCGFrac Fraction de la quantité d'entrée traitée par le groupe.

Enfin, pour simplifier l'écriture de certaines équations, nous avons ajouté un attribut à la table des combinaisons de ressources (COMBINATIONS) pour définir le rôle ou la fonction de chaque combinaison. Cet attribut est appelé RCType. Par exemple, il sera égal à "Harvest" pour une moissonneuse et à "Bale" pour une presse à balles.

Exemple d'une ferme produisant du colza

La figure 5.1 représente une chaîne logistique avec une ferme produisant du colza (site L1), un site de stockage (L2) et une raffinerie (L3). Cet exemple est intéressant car il facilite la compréhension du principe des macro-tâches. La ferme utilise pour le colza la chaîne de récolte C1 de la figure 3.4 du chapitre 3. Nous avons ajouté les préfixes de sites et les suffixes de chaîne de récolte aux tâches et aux états, pour montrer les identifiants d'états et de tâches utilisés dans la base de données.

Si on analyse ce qui se passe dans la ferme dans la figure, nous observons différents types d'activités réalisées par différents groupes de combinaisons de ressources.

Plus précisément, lors du passage de la moissonneuse-batteuse équipée d'un collecteur de menue-paille, les grains sont séparés, récupérés et directement transportés vers un site de stockage. La menue-paille récupérée par le collecteur est amassée en bord de champ pour être mise en balles. La paille reste par terre, sèche pendant 3 jours pour être mise en balles par la suite. Les groupes de combinaisons (CG) utilisés dans la table GROUPS et les fractions de la quantité d'entrée traitée sont :

- CG1 – Combine harvester + chaff collector : utilisé pour 100% du colza coupé.
- CG2 – Tractor + square baler (fixed) : mise en balles de la menue paille, 18% du colza coupé.
- CG3 – Tractor + flatbed trailer : transport des balles de menue paille au hangar, 18 %.
- CG4 – Tractor + square baler : mise en balles de la paille, 33,1 % du colza coupé.
- CG3 – Tractor + flatbed trailer : transport des balles de paille au hangar, 33,1 %.

On peut évidemment regrouper les pourcentages de la combinaison CG3 qui assure le transport des balles de pailles et de menue-paille, ce qui donne $18 + 33,1 = 51,1$ %.

La figure 5.2 illustre le graphe état-tâche après introduction de la macro-tâche. Noter que le stock de graines et le stockage couvert pour les balles sont conservés dans la ferme, pour éviter que la macro-tâche soit reliée aux tâches de transport. En effet, en suivant un chemin quelconque dans un graphe état-tâche, on doit traverser une alternance entre états et tâches.

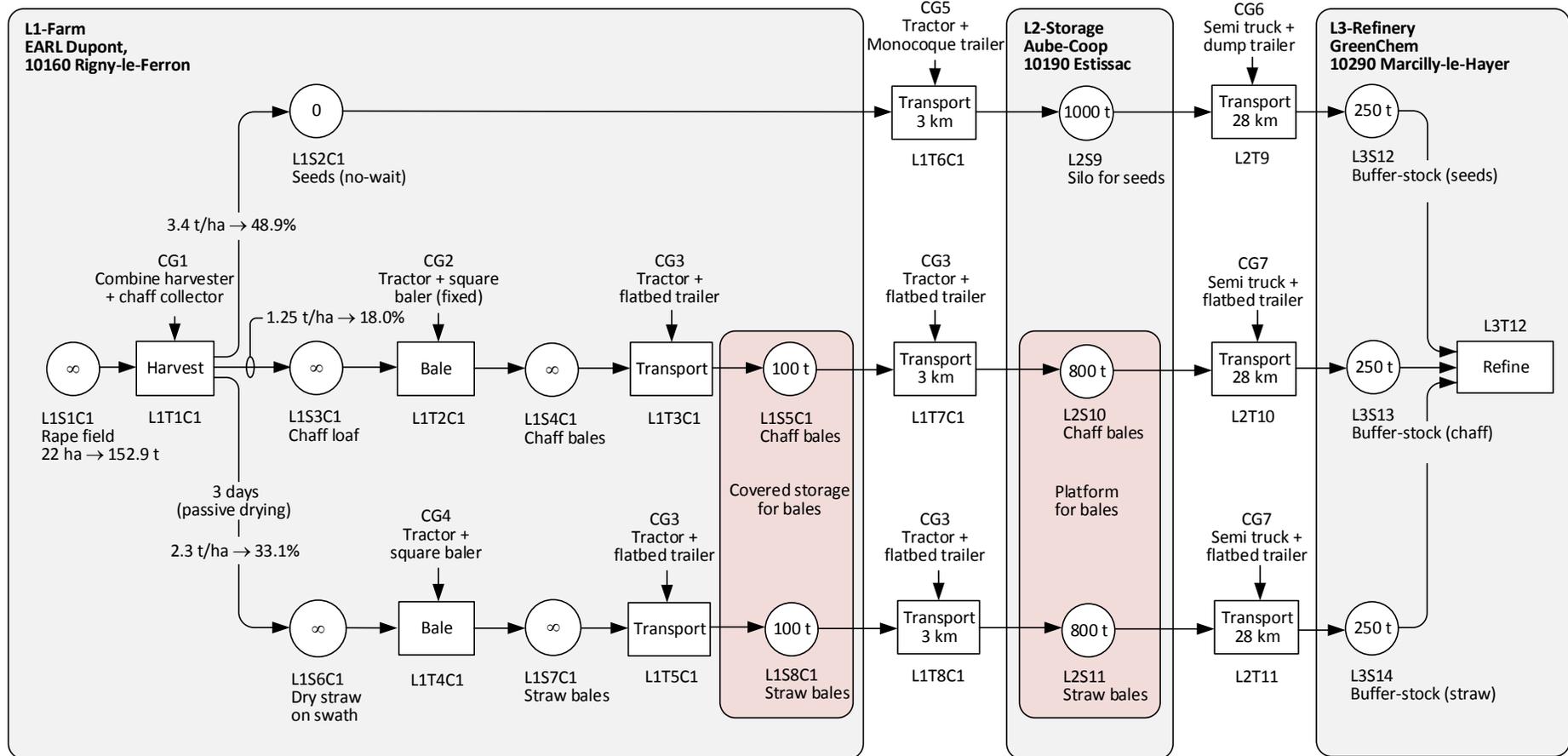


Figure 5.1. Graphe état-tâche avec une ferme produisant du colza avec la chaîne de récolte C1.
 Les combinaisons de ressources nécessaires sont indiquées au-dessus de chaque tâche.

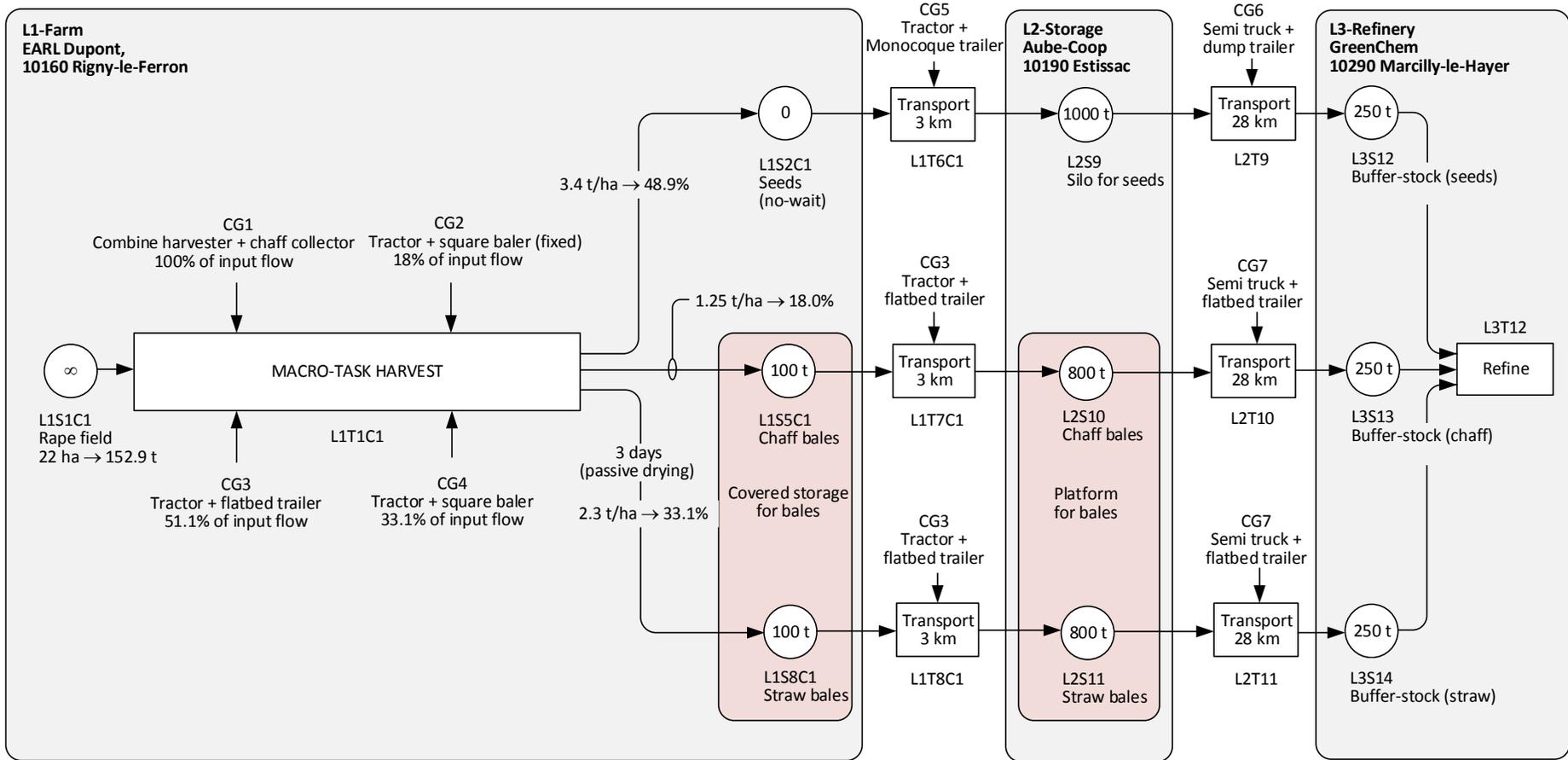


Figure 5.2. Graphe état-tâche après introduction de la macro-tâche.

La table 5.1 montre le contenu de la table TASK_GROUPS pour la chaîne condensée de la figure 5.2. Par rapport à la chaîne de récolte de la figure 5.1, on économise quatre états et quatre tâches qui sont regroupés dans la macro-tâche L1T1C1.

TID	CGID	TCGFrac
L1T1C1	CG1	1
L1T1C1	CG2	0,18
L1T1C1	CG3	0,511
L1T1C1	CG4	0,331
L1T6C1	CG5	1
L1T7C1	CG3	1
L1T8C1	CG3	1
L2T9	CG6	1
L2T10	CG7	1
L2T11	CG7	1

Table 5.1. Exemple de contenu de la table TASK_GROUPS.

Bien que nous ayons choisi le colza comme exemple, la modélisation par macro-tâches peut s'appliquer de la même manière à l'ensemble des chaînes de récolte décrites dans le chapitre 3. Elle est donc très générale et préserve la flexibilité du modèle de réseau.

2.3 Modification du modèle complet

Dans cette section, nous allons présenter la mise en œuvre du principe de la macro-tâche en expliquant comment modifier le modèle complet du chapitre précédent.

2.3.1 Liste des noms de données et de variables utilisés pour les calculs

Certaines variables ont été déjà déclarées dans le chapitre 4, d'autres ont été redéfinies et adaptées au nouveau modèle.

$TCGPairs$	$TCGPairs \subseteq TID \times CGID$ (clés de TASK_GROUPS)
$TCGFrac_{(i,cg)}$	Fraction du flux d'entrée de la tâche i traité par la combinaison cg
Z	Sous-ensemble {Harvesting} des macro-tâches
ZHR	Sous-ensemble {Harvest, Rake} des types de combinaisons
ZB	Sous-ensemble {Bale} des types de combinaisons
ZT	Sous-ensemble {Transport} des types de tâches et des types de combinaisons
ZR	Sous-ensemble {Refine} des types de tâches

Table 5.2. Ensemble utilisés pour l'indexation et nouveaux symboles pour le modèle compact.

Variables continues non négatives	
$X_{e,i,t}$	Flux de produit sur l'arc état-tâche (e, i) en période t
$Y_{i,e,t}$	Flux de produit sur l'arc tâche-état (i, e) en période t
$S_{e,t}$	Niveau de stock de l'état e en fin de période t
$Amp_{i,c,t}$	Quantité totale traitée par la combinaison c dans la tâche i en période t
$TCRH_{i,c,t}$	Temps d'utilisation de la combinaison c de $RCType_c \in ZHR$ en période t pour une macro-tâche $i \in Z$
$TCB_{i,c,t}$	Temps d'utilisation de la combinaison c de $RCType_c \in ZB$ en période t pour une macro-tâche $i \in Z$
$TCT_{i,c,t}$	Temps d'utilisation de la combinaison c en période t pour une tâche $i \in ZT \cup Z$ (macro-tâches + tâches de transport)
$TR_{l,r,t}$	Temps d'utilisation de la ressource r sur le site l en période t

Table 5.3. Variables utilisées.

- TCGPairs correspond aux couples de clés de la relation TASK_GROUPS. Cette relation permet de préciser les fractions de chaque groupe de combinaisons de ressources dans les tâches.
- Les sous-ensembles sont définis pour l'écriture de certaines contraintes qui ne concernent que certains types de tâches ou certaines combinaisons de ressources.

2.3.2 Fonction-objectif

Cette section ressemble beaucoup à la section 5 du chapitre 4, mais elle est adaptée pour refléter les changements supplémentaires apportés au modèle (les macro-tâches).

La fonction-objectif *TotalCost* à minimiser reste le coût total de l'approvisionnement en biomasse, composé des coûts de récolte (l'ensemble des coûts des activités dans les macro-tâches), de transport, de stockage et de manutention. Le modèle minimise la somme des coûts suivants :

$$\min TotalCost = FarmOpCosts + TransOpCosts + InvCosts + HandCosts$$

Les formulations des coûts (*TransOpCosts* + *InvCosts* + *HandCosts*) ressemblent à celles du modèle complet mais quelques changements interviennent dans les coûts de récoltes. Dans le chapitre 4, *FarmOpCosts* incluait les coûts des activités "Harvest", "Bale" et "Rake" mais pas le coût des activités de transport dans les fermes. Ces derniers étaient comptés dans *TransOpCosts*. Désormais, les coûts des transports cachés dans les macro-tâches de récolte sont inclus.

FarmOpCosts =

$$\begin{aligned} & \sum_{i \in TID | TType_i \in Z} \sum_{c \in RCID | (i, RCGroup_c) \in TCGPairs \text{ and } RCType_c \in ZHR} \sum_{t \in H} TCHR_{i,c,t} \cdot RCOCost_c \\ & + \sum_{i \in TID | TType_i \in Z} \sum_{c \in RCID | (i, RCGroup_c) \in TCGPairs \text{ and } RCType_c \in ZB} \sum_{t \in H} TCB_{i,c,t} \cdot RCOCost_c \\ & + \sum_{i \in TID | TType_i \in Z} \sum_{c \in RCID | (i, RCGroup_c) \in TCGPairs \text{ and } RCType_c \in ZT} \sum_{t \in H} TCT_{i,c,t} \cdot RCOCost_c \end{aligned}$$

Les coûts d'utilisation $TransOpCosts$ des combinaisons de ressources pour les tâches de transport ont une expression identique à celle du chapitre 4, mais elles concernent seulement les transports entre les fermes et les stockages, les stockages et la raffinerie, et les fermes et la raffinerie.

$$TransOpCosts = \sum_{i \in TID | TType_i \in ZT} \sum_{c \in RCID | (i, RCGroup_c) \in TCGPairs} \sum_{t \in H} TCT_{i,c,t} \cdot RCOcost_c$$

Le calcul des coûts de stockage n'est pas affecté :

$$InvCosts = \sum_{e \in SID} \sum_{t \in [SBeg_e, SEnd_e]} SInvCost_e \cdot S_{e,t}$$

Enfin, les coûts de manutention d'entrée et sortie dans les stocks (handling costs) ne changent pas :

$HandCosts =$

$$\sum_{(i,e) \in TSA} \sum_{t \in [SBeg_e, SEnd_e]} SInpCost_e \cdot Y_{i,e,t} + \sum_{(e,i) \in STA} \sum_{t \in [SBeg_e, SEnd_e]} SOutCost_e \cdot X_{e,i,t}$$

2.3.3 Contraintes de stocks, de flux et de demandes de la raffinerie

Les contraintes de stocks et de flux ainsi que les contraintes liées aux demandes de la raffinerie restent inchangées par rapport au modèle complet vu au chapitre précédent. Toutes les propriétés énoncées sur ces contraintes restent valables et sont réutilisées dans le nouveau modèle.

2.3.4 Contraintes de gestion des ressources

Ces contraintes, indépendantes des contraintes de stockages et des demandes de la raffinerie, sont mises à part, de manière à pouvoir être modifiées et adaptées à chaque fois que la gestion des ressources change.

Quantités traitées par chaque combinaison de ressources dans une tâche :

Contraintes répartissant le flot total en entrée de tâche sur les différentes combinaisons. Ces contraintes permettent de calculer les quantités de biomasses traitées par les combinaisons dans les tâches.

$\forall i \in TID | TType_i \notin ZR, \forall cg \in CGID | (i, cg) \in TCGPairs, \quad \forall t \in H : \quad (\text{SpanFlow})$

$$\sum_{c \in RCID | RCGroup_c = cg} \frac{Amp_{i,c,t}}{TCGFrac_{(i, RCGroup_c)}} = \sum_{(e,i) \in STA} X_{e,i,t}$$

Temps d'utilisation des combinaisons dans les macro-tâches :

Il s'agit de calculer les nombres d'heures d'utilisation des combinaisons c dans la tâche de récolte i en période t , qui sont nécessaires pour les coûts d'utilisation des équipements dans la fonction-objectif.

$\forall i \in TID \mid TType_i \in Z, \forall c \in RCID \mid (i, RCGroup_c) \in TCGPairs, \forall t \in H :$

$$\text{Si } RCType_c \in ZHR : TCHR_{i,c,t} = \frac{Amp_{i,c,t}}{\sum_{(e,i) \in STA} RCThruc \cdot STYield_{e,i}}$$

$$\text{Si } RCType_c \in ZB : TCB_{i,c,t} = \frac{Amp_{i,c,t}}{\sum_{(e,i) \in STA} RCThruc \cdot RCVolume_c \cdot PDens_{SProd(e)}}$$

$$\text{Si } RCType_c \in ZT : TCT_{i,c,t} = \left(\frac{Amp_{i,c,t}}{\min(RCWeight_c, \sum_{(e,i) \in STA} RCVolume_c \cdot PDens_{SProd(e,i)})} \cdot \frac{TDist_i}{RCpeed_c} \right)$$

La dernière équation concerne le temps d'utilisation de la combinaison de ressources c pour un transport dans la ferme. Elle correspond à l'approximation décrite dans la section 5 du chapitre 4 (utilisation de camions pleins).

Temps d'utilisation des combinaisons par les tâches de transport externes :

$\forall i \in TID \mid TType_i \in ZT, \forall c \in RCID \mid (i, RCGroup_c) \in TGCPairs, \forall t \in H :$ (GetTCT)

$$TCT_{i,c,t} = \left(\frac{Amp_{i,c,t}}{\min(RCWeight_c, \sum_{(e,i) \in STA} RCVolume_c \cdot PDens_{SProd(e,i)})} \cdot \frac{TDist_i}{RCpeed_c} \right)$$

Temps d'utilisation des ressources :

Ils se déduisent des heures d'utilisation des combinaisons où la ressource intervient. Une même ressource, par exemple un tracteur peut intervenir dans différentes combinaisons sur le même site.

$\forall l \in LID, \forall r \in RID \mid (l, RFam_r) \in LFPairs, \forall t \in H :$ (GetTR)

$$TR_{l,r,t} = \sum_{c \in RCID \mid (c,r) \in RCRPairs} \left(\sum_{i \in TID \mid TType_i \in Z \text{ and } (i, RCGroup_c) \in TCGPairs \text{ and } TLoca_i=l} (TCHR_{i,c,t} + TCB_{i,c,t} + TCT_{i,c,t}) + \sum_{i \in TID \mid TType_i \in ZT \text{ and } (i, RCGroup_c) \in TCGPairs \text{ and } TLoca_i=l} TCT_{i,c,t} \right)$$

Respect des temps d'utilisation des ressources sur chaque site :

$$\forall r \in RID, \forall (l, RFam_r) \in LFPairs, \forall t \in H : \frac{TR_{l,r,t}}{\min(RWTime_r, LHours_l)} \leq LFNMax_{l,RFam[r]}$$

2.4 Test sur 10 fermes

Cette section décrit des tests sur 10 fermes pour comparer le modèle complet simplifié et le modèle compact. La résolution se fait toujours avec CPLEX, sur le même ordinateur portable.

Nous considérons deux instances avec 10 fermes de colza et 3 sites de stockage, dans un premier temps sur un horizon de planification d'une année (364 jours) découpé en périodes d'une journée. Nous appelons instance 1 (Inst1), l'instance de test où la chaîne de récolte pour chaque ferme est la chaîne C1 illustrée par la figure 5.1. L'instance 2 (Inst2) est identique, sauf que la chaîne de récolte détaillée est remplacée dans chaque ferme par la macro-tâche de la figure 5.2.

On applique le modèle complet simplifié à l'instance 1. Il donne une solution optimale en 6 secondes. En appliquant le modèle compact à l'instance 2, on obtient une solution optimale de même coût, en moins d'une seconde. Ce résultat était prévisible car la macro-tâche fonctionne comme une boîte noire qui approxime les activités qui se passent à l'intérieur.

Nous avons effectué une agrégation temporelle des deux instances précédentes en passant à des périodes unitaires d'une semaine, ce qui donne deux instances Inst1' et Inst2'. Donc, l'horizon de planification d'un an est maintenant découpé en périodes d'une semaine (7 jours). A noter que ce changement fait disparaître l'attente de 3 jours pour la paille de colza. L'agrégation temporelle se fait à l'aide d'une macro VBA qui modifie la base de données avant la résolution. Le changement d'échelle de temps se fait donc très simplement, au niveau des données, sans affecter les modèles.

La table 5.4 résume les résultats. Le test sur l'instance Inst1' donne une solution en moins d'une seconde avec un coût total augmenté de 3,2 % par rapport au résultat de l'instance Inst1. On obtient la même différence de coût avec le modèle compact.

Modèle	Instances	Temps CPU	Différence
Modèle complet simplifié	Inst1 (jours)	6s	3,2%
	Inst1' (semaines)	< 1s	
Modèle compact	Inst2 (jours)	< 1s	3,2%
	Inst2' (semaines)	< 1s	

Table 5.4. Résultat des tests sur 10 fermes.

En résumé, on n'observe pas de perte de qualité de la solution avec le système des macro-tâches, que ce soit pour l'horizon découpé en 364 jours ou celui divisé en 52 semaines. Par contre, le passage de la granularité journalière à celle hebdomadaire augmente le coût total des solutions de 3,2 %, que les macro-tâches soient utilisés ou non.

Bien entendu, il existe certainement des exemples où les macro-tâches induiraient une petite perte de coût. Il s'agirait par exemple de cas où la solution optimale du modèle complet introduit des périodes d'attente de la paille sur le sol, avant la mise en balles. Heureusement, cette situation, bien que théoriquement possible, ne correspond pas aux pratiques habituelles des agriculteurs.

3. Etude de cas sur une grande instance

Dans cette partie, nous allons présenter des tests sur un jeu de données réels, pour une bioraffinerie prévue à Venette (banlieue de Compiègne).

3.1 Difficultés liées à l'acquisition des données

La question périlleuse des données s'est posée dès le départ du projet. Il nous a fallu avancer sur des hypothèses et des données plus ou moins fiables pour faire les premiers tests. Le "secret statistique" évoqué dans la section 3.2 a également constitué un obstacle.

Nous avons été en contact direct durant la thèse avec des spécialistes de la biomasse, qui ont orienté et guidé certaines décisions et hypothèses :

- Un représentant de la société Coopénergie, M. Cyril Flamin, qui nous a fourni des données sur les sites de stockages que nous allons voir dans la suite. Coopénergie est une union de coopératives pour la fourniture de biomasse agricole et forestière,
- Un représentant du Groupe Avril (ex-Sofiprotéol), M. Francis Valter, qui nous a fourni des données sur les zones de production de biomasse autour de Venette, ainsi que les spécifications détaillées du schéma d'approvisionnement d'une bio-raffinerie de proximité (chapitre 1).
- Un représentant du CETIM, M. Gaël Guégan, avec qui nous avons eu des échanges sur les équipements de récolte dans le cadre de réunions de travail.

Les données sur les équipements (sauf pour le transport entre sites) nous viennent principalement des barèmes d'entraide élaborés par les chambres d'agriculture, déjà évoqués dans la section 5 du chapitre 4. Pour les transports entre sites, la Fédération National du Transport Routier (FNTR) indique des coûts d'utilisation de camions.

3.2 Données sur les fermes de production de biomasse d'un cas réel

Par rapport aux données communales, le Comité du Secret Statistique a donné un avis favorable, le 18 mars 2011, pour une dérogation restreinte à propos de l'application du secret statistique sur les données du recensement agricole 2010. La dérogation porte essentiellement sur les cinq attributs suivants, par commune : les nombres d'exploitations, les surfaces agricoles utilisées (SAU), les nombres d'animaux en "unités gros bétail" (UGB), les productions brutes standards (PBS), les unités de travail annuel (UTA) et les orientations technico-économiques dominantes (OTEX). Le recensement a donné lieu à des données plus fines mais qui restent protégées par le secret statistique : les données ne peuvent donc être rendues publiques.

Bien qu'il soit normal de ne pas avoir de données au niveau parcelle, à cause de la rotation des cultures, nous avons été surpris par l'incroyable opacité qui règne sur les données de production agricole en France. A titre de comparaison, même en tant que citoyens français, nous pouvons accéder gratuitement à de nombreuses bases de données agricoles US, bien plus complètes.

Parmi les attributs dévoilés, il y en a trois qui nous sont principalement utiles dans notre cas d'étude :

- Le nombre d'exploitation agricoles. Cela concerne le nombre total d'exploitations qui ont leur siège sur la commune concernée. Est appelée exploitation toute unité économique qui participe à la production agricole et qui atteint une certaine dimension (1 hectare de superficie agricole).
- La SAU. C'est la surface agricole totale en ha utilisée par les exploitations qui ont leur siège sur la commune concernée.
- La contribution de chaque surface à la production brute standard (PBS). Cette information permet de classer les exploitations en OTEX (orientation technico-économique) selon leur spécialisation ou autrement dit l'orientation principale des productions, par exemple "grandes cultures", "viticulture", etc.

Ces données ont pour source le Recensements Agricole 2010 dans un fichier Excel consultable en ligne sur le site du Ministère français de l'agriculture, de l'agroalimentaire et de la forêt sur ce lien (<http://agreste.agriculture.gouv.fr/chiffres-cles-4/commune/>). Ce fichier contient 2292 communes avec le code postal, le nombre d'exploitations, la surface totale en ha et les autres attributs, pour la région de Picardie. Les données pour chaque exploitation sont inaccessibles à notre niveau.

Le but de l'étude étant de chercher à construire un schéma d'approvisionnement avec du colza, nous avons retenu les exploitations à orientation céréales et oléagineux qui sont compatibles avec le colza, ce qui donne un filtrage du fichier sur les OTEX "15 – Céréales et oléo-protéagineux" et "16 – Grandes cultures de type général"). Ensuite, nous avons fait des hypothèses sur la répartition des types de biomasse. Les surfaces ont été divisées par 3 pour estimer la production annuelle de colza en respectant le système de rotation des cultures colza/blé/orge.

Concernant la granularité spatiale, les "fermes" considérées dans nos modèles correspondent aux communes. Ensuite, parmi l'ensemble des communes issues du filtre "15 - Céréales et oléo-protéagineux" et "16 - Grandes cultures de type général", nous avons sélectionné celles situées à moins de 50 km de Venette, ce qui donne 460 communes.

Dans l'étude de cas à une ferme du chapitre 4, nous avons des transferts des fermes vers la raffinerie. Dans le cas réel, il n'y a pas de tels flux directs. Même si notre modèle le permet, les spécialistes indiquent que dans la réalité l'agriculteur ne livre pas à la raffinerie, car les OS (organisme-stockeurs) et les négociants doivent prendre des échantillons et faire des analyses pour garantir la qualité et assurer une traçabilité dans les sites de stockage. Donc, toutes les quantités passent par ces derniers.

3.3 Données sur les stockages

Les installations de stockages agricoles répondent aux besoins de collecte, de stockage et de conservation en fonction des demandes diverses et variées des marchés des produits agricoles. Elles doivent répondre à une exigence principale qui est de mettre en stock le plus rapidement possible, disposer de volumes de stockage judicieusement dimensionnés, donc économiques, et assurer la conservation et l'expédition dans de bonnes conditions. Il existe différents types d'installation, ayant des caractéristiques et donc des intérêts dépendant de l'utilisation requise.

Les données sur les stocks incluses dans notre cas d'étude nous viennent en grande partie d'un résultat fourni par Coopénergie dans le cadre d'un autre projet de PIVERT, appelé COLLVERT. Ce dernier avait pour but de faire des études relatives aux méthodes de récolte (machinisme agricole et équipements, mise en œuvre de chantiers, manutention), et la collecte des biomasses entre les champs et l'usine pour l'obtention de données. Il a donné lieu à un document de Coopénergie décrivant les principes du stockage et listant les stockages des différentes coopératives, situés à 50 km de Compiègne environ. Nous avons en pratique les types de stockages suivants.

En premier lieu, nous pouvons citer les *hangars agricoles* qui sont des sites de stockage situés chez l'agriculteur. Ils sont de capacités diverses et variées. Ce sont les plus simples mais ils obéissent à quelques règlements spécifiques rappelés dans le document de Coopénergie.

Ensuite, nous avons la *coopérative*, une entreprise créée par des agriculteurs pour mutualiser des outils de production, de conditionnement, de stockage, ainsi que la commercialisation ou la transformation des produits de leurs exploitations, les approvisionnements en engrais et autres intrants. La coopérative agricole collecte et stocke principalement des grains (céréales, oléagineux) à destination de différents marchés. Compte-tenu des importantes quantités de produits à gérer, les centres de stockage sont équipés non seulement de silos, mais aussi d'installations assurant les opérations de réception, de traitement, de contrôle et de déchargement.

Enfin les *organismes stockeurs* appelés communément OS ont pour activités principales la collecte, la conservation, le stockage et la commercialisation de produits agricoles. Pour répondre aux besoins des agriculteurs, ces organismes conseillent et distribuent des agrofournitures (engrais et produits phytosanitaires) nécessaires à la croissance et à la protection des végétaux. Le lecteur intéressé peut se référer au livre *Connaître et faire face aux risques des organismes stockeurs de la filière agricole - Tome 1*, réalisé dans le cadre d'un groupe de travail constitué de plusieurs partenaires, entre autres COOP de France et Champagne-Céréales. Ce livre peut être consulté in extenso en se rendant sur le lien suivant ([http://www.ineris.fr/gesdoc/aida/Guide Incendie Tome 1.pdf](http://www.ineris.fr/gesdoc/aida/Guide%20Incendie%20Tome%201.pdf)).

Les solutions de stockage concernent le plus souvent les graines. Il est donc parfois nécessaire d'avoir recours à un pré-traitement (granulation) pour pouvoir utiliser des silos à grains pour d'autres produits comme la paille, sous réserve de leur disponibilité et de la bonne résistance des granulés lors des opérations de convoyage. Quant aux plateformes de collecte, elles peuvent accueillir de la biomasse en vrac ou conditionnée en balles sur une partie de l'année.

L'étude de COLLVERT a été réalisée par Coopénergie auprès des organismes-stockeurs, à l'aide d'un questionnaire. Les réponses saisies sous Excel nous ont fourni des données relatives à 46 sites de stockages situés dans un rayon de 50 km autour de Compiègne : types de stockage (silo, plateforme...), capacité en tonnes pour les silos et en mètres carrés pour les plates-formes, coût de stockage, coût d'entrée, coût de sortie, fenêtre de disponibilité dans l'année.

Concernant les fenêtres de temps, certains stockages sont en effet fermés en hiver. D'autres ne peuvent être utilisés certains mois pour stocker de la biomasse destinée à la bioraffinerie, car ils sont utilisés pour d'autres débouchés comme la production de blé ou de betteraves.

Les capacités de stockage varient de 200 à 60 000 tonnes. Le format du fichier Excel étant différent de celui de notre base de données, un travail important a été fait pour importer ces données.

3.4 Instance de test d'un cas réel pour le modèle compact

D'après les données récupérées, nous avons pu générer une grande instance de 460 "fermes" de colza (en fait 460 communes, regroupant 2146 exploitations) avec une chaîne de récolte unique identique à celle de la figure 5.2, 46 sites de stockage, et une raffinerie sur un horizon d'une année divisée en 52 semaines. Le modèle résultant contient 588 249 contraintes et 574 795 variables. Après le pré-solveur, il est réduit à 4 801 contraintes et 19 558 variables et est résolu de façon exacte en moins d'une minute (47 secondes).

Ce résultat est intéressant dans la mesure où le modèle de la littérature qui se rapproche le plus de notre travail est celui de Shastri et al. (2011a). Nous rappelons que ces derniers traitent une seule biomasse, avec une chaîne de récolte unique en trois étapes (Harvest, Rake et Bale). Leur modèle est donc beaucoup plus restrictif que le nôtre. Dans Shastri et al. (2011b), une d'étude de cas à seulement 31 fermes, 1 bioraffinerie et 1 site de stockage centralisé nécessite déjà 5 706 secondes en temps de calcul. Bien que les instances utilisées soient très différentes, les comparaisons des tailles de chaînes logistiques pouvant être traitées et des temps de calcul montrent l'apport de notre travail par rapport à la littérature sur des modèles mathématiques d'approvisionnement en biomasse.

La figure 5.3 montre la consommation de mémoire sous Windows 7. Si on déduit l'espace-mémoire occupé par Windows et l'environnement de développement OPL STUDIO, le chargement et la résolution ont nécessité au total 8 Go de mémoire, alors que le modèle complet sur un cas à 10 fermes a besoin de 11 Go. Le PC utilisé, un portable puissant avec un Intel Core i7, permet de monter 32 Go de RAM sur la carte-mère, ce qui montre que nous avons de la marge pour traiter des chaînes d'approvisionnement encore plus grosses.

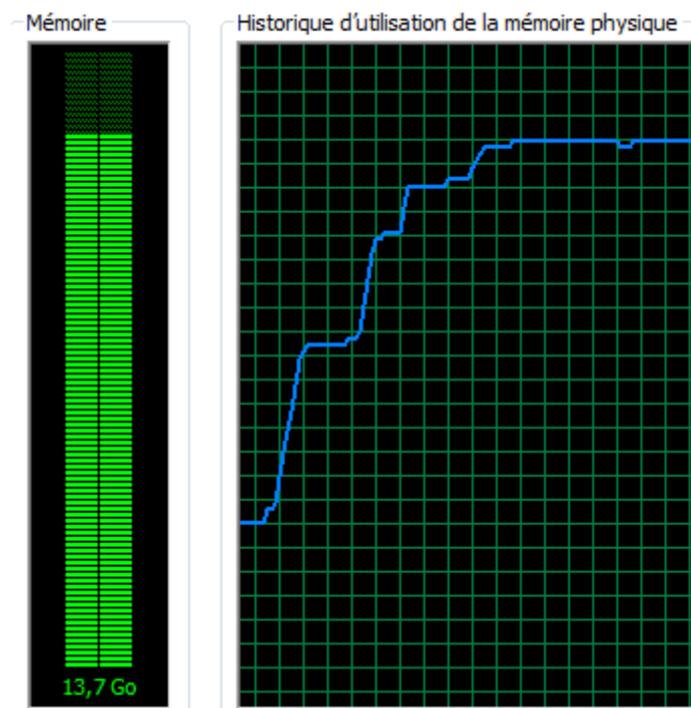


Figure 5.3. Mémoire utilisée par le modèle compact pour le cas d'étude réel.

Note : 1 graduation = 6 s. L'optimisation est finie mais OPL STUDIO n'a pas encore libéré la mémoire.

Nous avons remarqué dans les tests une certaine sensibilité des temps de calcul du modèle aux demandes de la raffinerie : pour une capacité totale donnée des sites de stockage, le temps de calcul a tendance à croître avec la demande totale de la raffinerie. Comme la raffinerie consomme lentement des biomasses toute l'année alors que les différentes plantes sont récoltées massivement à des périodes différentes, les stockages servent de tampons, parfois sur des durées importantes. Nous pensons qu'ils constituent un goulot d'étranglement qui peut rendre le programme linéaire plus difficile à résoudre, en nombre de pivotages pendant l'algorithme du simplexe. Des tentatives avec l'algorithme de point intérieur de CPLEX n'ont rien donné de significatif.

4. Extensions possibles des modèles

Il s'agit essentiellement d'extensions faciles à réaliser des modèles des chapitres 4 et 5.

Options de gestion des ressources

D'autres options de gestion des ressources sont possibles, nous allons en citer deux et leur mise en œuvre ne demande pas beaucoup d'efforts :

- Une option dans la gestion des ressources peut être l'autorisation de n'utiliser qu'une combinaison de ressources par tâche. Cela peut être intégré simplement dans le modèle par l'introduction de variables binaires.
- Une option peut permettre à l'utilisateur de spécifier les équipements disponibles sur chaque site, au lieu de laisser le modèle choisir des combinaisons qui peut-être doivent être achetées. Cela est faisable facilement dans la gestion des données.

Décisions stratégiques

La flexibilité du modèle nous permet de facilement intégrer des décisions stratégiques comme la localisation, la taille et la capacité de sites de stockage et de raffineries. La façon d'intégrer ce genre de décisions dans un modèle peut être trouvée dans l'article de Eksioglu et al. (2009).

Optimisation multi-objectif

L'optimisation *multi-objectif* ou *multicritère* vise à optimiser simultanément plusieurs critères, ce qui peut être interprété mathématiquement de plusieurs manières.

Il existe de nombreux algorithmes pour l'optimisation au sens de Pareto, voir par exemple le livre de Branke et al. (2008). *La méthode ε -contrainte* est très populaire. Sur un problème bi-objectif, elle consiste à convertir un des critères en contrainte, ce qui permet ensuite d'utiliser un algorithme d'optimisation mono-objectif. Le principe peut-être résumé simplement de la façon suivante : Considérons un vecteur x de n variables et deux fonctions-objectifs à minimiser $f(x)$ et $g(x)$. Il est possible de minimiser $f(x)$ sous la contrainte $g(x) \leq \varepsilon$, ε étant une constante choisie, pour obtenir

une solution Pareto-optimale. En faisant varier ε , on peut construire progressivement le front de Pareto.

Dans le cadre de cette étude, on pourrait appliquer ce principe pour intégrer les émissions et les consommations d'énergie dans le processus d'optimisation afin de mieux prendre en compte les aspects liés à la durabilité. La fonction $f(x)$ serait par exemple notre fonction-objectif actuelle (le coût total), tandis que $g(x)$ serait la somme des consommations d'énergie et des rejets de GES, actuellement calculées après la résolution à l'aide des formules (GET_CONS) et (GET_EMISS) du chapitre 4. Nous aurions ainsi un modèle d'optimisation bicritère.

Tâches de prétraitement

Si on souhaite intégrer une activité de prétraitement de la biomasse, le modèle de réseau avec le paradigme des graphes états-tâches permet de définir une tâche *ad hoc*, avec éventuellement des états pour les stocks d'entrée et de sortie. Comme exemples d'activités de prétraitement dans les chaînes d'approvisionnement, nous pouvons citer la granulation ("pelletization") de paille. Le choix de la technologie de prétraitement pourrait même être introduit comme variable dans le modèle.

Optimisation en mode "flux poussé"

Dans les modèles présentés, la raffinerie "tire" les flux en allant puiser des besoins donnés dans les champs. Certains auteurs US considèrent le cas d'une raffinerie de capacité inconnue qui va prendre toute la biomasse sur un territoire donné. Il faut donc collecter un maximum de biomasse et déterminer ensuite la capacité de la raffinerie qu'on peut construire.

Mettre en place un tel "flux poussé" n'est pas compliqué. Il faut supprimer dans les données les besoins de la raffinerie et introduire une variable pour sa capacité de production. Des contraintes doivent cependant être ajoutées pour vider tous les champs. En effet, dans nos modèles actuels, le prélèvement de biomasse dans un champ peut s'interrompre dès que les besoins de la raffinerie sont satisfaits : on peut donc avoir des champs incomplètement moissonnés. Dans la réalité, évidemment, l'agriculteur va récolter le reste, mais pour d'autres usages qui ne nous regardent pas.

Détermination des surfaces à cultiver

Dans le modèle actuellement, nous supposons que les surfaces allouées pour chaque biomasse sont connues. Cependant on aurait pu déterminer les surfaces nécessaires à cultiver pour satisfaire des besoins en biomasse.

Gestion de plusieurs modes de transport

Le modèle actuel considère uniquement du transport routier. Cependant, il est possible d'y intégrer d'autres modes de transport (transport dit "multimodal" ou "intermodal"). Pour ce faire, Il faudrait définir des tâches de transport avec des ressources comme des trains, pour faire du transport ferroviaire, ou des péniches, pour faire du transport fluvial. Il apparaît clairement que cette extension est compatible avec les graphes états-tâches. Cependant elle pourrait engendrer des difficultés au

niveau des sites de transbordements qu'il faudra gérer correctement et finement pour éviter toute incohérence des flux et du système.

Plus généralement, on peut décrire des sous-réseaux pour chaque mode de transport, en utilisant le formalisme des graphes états-tâches, et interconnecter les graphes obtenus par des tâches de transbordement.

Gestion d'incertitudes

Des incertitudes sont présentes tout au long de la chaîne d'approvisionnement, particulièrement au niveau de la disponibilité de la biomasse mais aussi des demandes de la raffinerie. Il serait intéressant d'étudier ces incertitudes et de les intégrer dans un modèle, bien que les méthodes de résolution de programmes mathématiques stochastiques soient très complexes.

5. Conclusion

Un modèle générique a été développé pour permettre de traiter de très grandes instances en des temps de calcul raisonnables. Ce modèle dit "compact" est basé sur l'idée de modéliser et d'approximer une ensemble d'activités dans une ferme comme une boîte noire (macro-tâche).

Les résultats ont montré que ce modèle est très efficace puisqu'il peut résoudre une instance de 460 communes (2 146 exploitations) et 46 stockages en moins d'une minute de calcul. Cette instance en vraie grandeur correspond à une raffinerie prévue près de Compiègne.

Le modèle compact permet donc d'évaluer des chaînes d'approvisionnement de taille réelle, en remédiant ainsi aux limites que présentent le modèle complet et sa version simplifiée, vus dans le chapitre 4. De plus, sur des cas à 10 fermes où les modèles peuvent être comparés, le système des macro-tâches n'induit pas d'augmentations du coût total.

Dans le chapitre 6, nous allons présenter un prototype de logiciel intégrant notre base de données du chapitre 3 et nos trois principaux modèles mathématiques développés (complet, complet simplifié, et compact).

Chapitre 6

Logiciel-prototype BIOSCO

1. Introduction

Ce chapitre a pour but de présenter un logiciel qui intègre la base de données du chapitre 3 et les modèles mathématiques des chapitres 4 et 5. Ce logiciel a été baptisé BIOSCO (Biomass Supply Chain Optimizer). Comme l'objectif de la thèse n'est pas de parvenir à un logiciel professionnel, BIOSCO est actuellement un prototype, certes fonctionnel, mais qui demande encore des perfectionnements.

Le chapitre débute avec la section 2 sur les outils logiciels choisis pour développer BIOSCO. La section 3 montre l'architecture générale, avec un module pour gérer la base de données, un pour calculer les distances, un pour l'optimisation et un pour afficher les résultats. Les sections 4 à 7 commentent ces quatre modules. La section 8 présente l'interface-utilisateur qui chapeaute l'ensemble. Des recommandations sont formulées dans la section 9, avant une conclusion dans la section 10.

2. Outils logiciels utilisés

Tableur Excel

La base de données du chapitre 3 peut être réalisée avec un vrai système de gestion de base de données (SGBD) comme Access ou Oracle. Nous avons choisi d'implémenter les tables sous Excel car ce logiciel est plus répandu, moins coûteux, et permet facilement de visualiser les résultats sous forme de graphiques. Ce choix n'est pas limitatif car nous avons suivi les règles des modèle entités-relations : à tout moment, il sera possible de reprendre l'ensemble des tables dans un vrai SGBD.

Langage de programmation VBA pour Excel

Un autre intérêt d'Excel est la programmation en langage VBA (Visual Basic for Applications). VBA est encapsulé dans Excel et offre des classes pour communiquer avec ce dernier, par exemple pour écrire dans des cellules. Ainsi, l'écriture d'un programme et son exécution peuvent se faire dans Excel, sans recourir à un environnement de développement spécialisé. De plus, il est possible de démarrer automatiquement un programme VBA à l'ouverture d'un classeur Excel, ce qui permet de cacher les menus d'Excel et d'obliger l'utilisateur à choisir parmi les actions proposées. Nous avons sélectionné VBA pour réaliser l'interface-utilisateur de BIOSCO et appeler ses différents modules.

MapPoint pour les calculs de distance

MapPoint de Microsoft a été choisi car il ne nécessite pas de connexion à Internet et il offre des macros VBA pour réaliser des calculs de distance à partir d'Excel.

OPL STUDIO pour l'optimisation

OPL STUDIO d'IBM-ILOG est un environnement de développement de modèles mathématiques, basé sur le langage symbolique OPL et le solveur CPLEX. Ce dernier est considéré actuellement comme le solveur de programmation linéaire mixte le plus rapide, avec GUROBI et XPRESS.

3. Architecture fonctionnelle

L'application BIOSCO peut être schématisée sous forme de quatre modules autour d'une interface-utilisateur, comme indiqué dans la figure 6.1 : DATABASE, DISTANCES, OPTIMIZATION et REPORTING. Ces modules sont décrits dans les sections suivantes. Tous les composants sont installés sur le PC de l'utilisateur et le fonctionnement de l'application ne nécessite pas de connexion Internet. Aux fichiers Excel qui contiennent la base de données et l'interface-utilisateur en VBA s'ajoutent les logiciels MapPoint et OPL STUDIO.

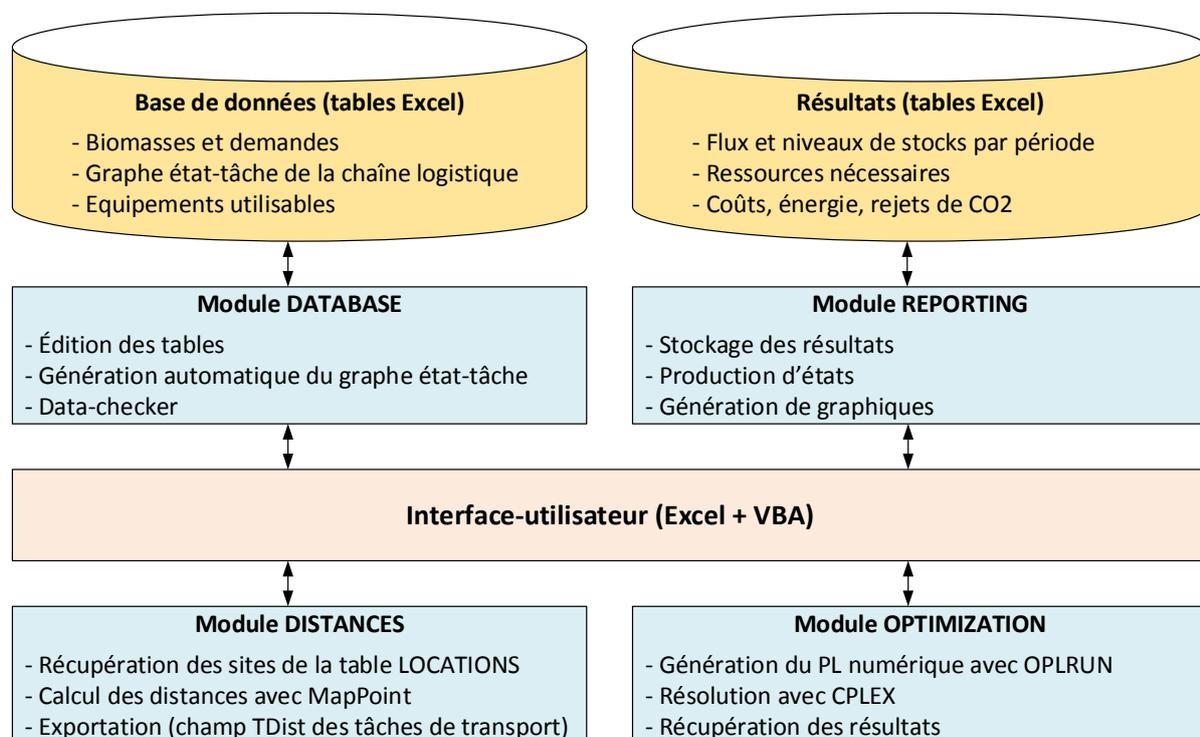


Figure 6.1. Architecture fonctionnelle de BIOSCO.

4. Module DISTANCES

Pour les coûts de transport, il faut calculer la distance entre les différents sites de la chaîne logistique (table LOCATIONS), pour renseigner l'attribut-distance TDist de la table des tâches, voir le chapitre 3. Pour une chaîne réelle de 500 fermes et 40 stockages centralisés, on a 20 540 distances à déterminer, si chaque ferme peut livrer chaque stockage et la raffinerie. Calculer les distances une par une avec un site Internet comme Mappy prend trop de temps dans ces conditions.

Au début de la thèse, nous avons choisi le logiciel MapPoint de Microsoft pour calculer les distances automatiquement. Leader du marché, ce produit inclut une base de données des routes de France et d'Europe et permet de calculer le plus court chemin par la route entre deux lieux dont on connaît les adresses ou les coordonnées GPS. Le logiciel peut aussi géo-localiser des points et afficher des itinéraires sur une carte. Tous ces traitements peuvent être réalisés via l'interface-utilisateur de MapPoint ou à partir d'un programme VBA sous Excel, grâce à des classes fournies.

Il faut noter que les coordonnées GPS sont essentielles en milieu rural, car certaines fermes n'ont pas de numéro de rue ("EARL Durand, route de Grigny") ou ne sont même pas placées en bord de route. Les coordonnées GPS permettent à MapPoint de placer avec précision une ferme le long d'une route ou, si la ferme n'est pas en bord de route, de calculer le point de rabattement le plus proche.

Nous avons développé une macro VBA qui accède aux objets de MapPoint pour calculer les distances entre sites. Ces développements sont réunis dans un classeur Excel "DataOKDist".

La macro commence par constituer dans la feuille principale de ce classeur trois colonnes contenant les sites entre lesquels il faut calculer la distance (fermes, stockages, raffinerie). Pour cela, les sites géographiques et leurs adresses sont extraits de la table LOCATIONS puis placés dans la bonne colonne selon leur type.

Ensuite, les distances sont calculées à l'aide de MapPoint dans trois feuilles : distances fermes-stockages, distances stockages-raffinerie et distances fermes-raffinerie. Pour respecter le modèle entité-relation, les informations sont mises sous formes de triplets (source, destination, distance) au lieu d'une matrice de distance dans chaque feuille du classeur excel.

Enfin, la macro exporte les distances calculées dans le champ TDist des tâches de transport dans la table TASKS.

Nous avons utilisé MapPoint avec succès mais, hélas, Microsoft a abandonné le produit fin 2014. Il existe des solutions de substitution mais la plupart nécessitent d'être connecté à Internet, ce qui est une contrainte forte. Il semble que les sociétés proposant des calculs de distances ne souhaitent plus diffuser des cartographies routières qui doivent être modifiées à chaque création de route ou de pont. C'est pourtant le cas pour les GPS de voitures. Une solution de plus en plus répandue chez ces sociétés est donc de maintenir une base de données centralisée et unique sur un serveur web, les utilisateurs devant se connecter pour calculer des distances.

Nous avons trouvé les solutions comme Maptitude, Map Quest, Google Maps, Maptive, Portatour et Map Business comme alternatives à MapPoint. Seul Map Quest est actuellement 100% gratuit mais il fonctionne uniquement en ligne. Maptitude est le seul qui offre le choix entre un fonctionnement sur PC, après chargement des données cartographiques, et un fonctionnement en ligne. Il est payant mais le prix annuel est raisonnable, autour de 700 dollars. Google Maps est particulier : il est gratuit jusqu'à 2500 requêtes journalières de calcul de distances, ensuite l'accès se bloque et il faut souscrire un abonnement pour faire plus de calculs.

Pour l'instant, le remplaçant de MapPoint dans BIOSCO n'a pas encore été sélectionné.

5. Module DATABASE

Le chapitre 3 a détaillé notre modèle de données et ses tables en Excel. Une première fonction du module DATABASE est de saisir certaines tables à la main. C'est d'abord le cas des tables LOCATIONS et DEMANDS puisqu'il faut bien connaître les sites concernés et les demandes de la raffinerie avant de décrire la chaîne logistique. Il y a aussi les petites tables rarement modifiées : PARAMETERS, CROPS, CROP_PARTS et PRODUCTS. Par exemple, les plantes utilisables ne risquent pas de changer de sitôt. La saisie des tables du système de gestion de ressources (RESOURCES, COMBINATIONS et COMB_RES) est plus fastidieuse car il y a plus de champs à renseigner et il faut déduire les coûts et productivités de chaque combinaison à partir des ressources qui la composent. Cependant, les ajouts sont peu fréquents une fois qu'on s'est constitué une liste standard d'équipements.

En fait, la majeure partie des données réside dans le graphe état-tâche (tables STATES, TASKS, ST_ARCS et TS_ARCS). Nous avons saisi ces tables manuellement pour le test à une ferme du chapitre 4 mais le processus prend trop de temps et présente trop de risques d'erreurs sur les grands cas.

La deuxième fonction du module DATABASE de BIOSCO est donc de générer automatiquement le graphe état-tâche. L'idée, déjà présentée dans la section 2.3 du chapitre 3, consiste à 1) préparer des sous-graphes pour les chaînes de récolte et les stockages centralisés, 2) les recopier sur chaque site concerné, avec tous les états et tâches associés, et 3) ajouter des arcs pour connecter les différents sites. Ce système de génération automatique repose sur des tables supplémentaires qui n'ont pas été décrites dans le chapitre 3 pour éviter de noyer le lecteur.

Un classeur Excel est préparé pour chaque sous-graphe représentant une chaîne de récolte ou un site de stockage centralisé.

Le classeur pour une chaîne de récolte contient des tables qu'on va retrouver dans la base de données du scénario (STATES, TASKS, ST_ARCS et LOCA_FAM), mais seuls les états et tâches du sous-graphe sont décrits, avec leurs connexions internes. Les arcs de connexion avec d'autres sites ne sont pas inclus. Par exemple, la chaîne de récolte standard du colza (figure 3.4 du chapitre 3) comprend 8 états, 5 tâches et 12 arcs. Elle est décrite dans un classeur C1H1 (Crop 1 – Harvesting chain 1).

Le classeur pour un site de stockage standard contient un stock de graines et une plate-forme servant de stock partagé pour les balles de paille et de menue paille, cf. la figure 3.2 du chapitre 3. Le sous-graphe associé ne comporte donc que des états, même si dans l'avenir on pourrait envisager des tâches comme du séchage actif.

Pour réaliser un scénario d'optimisation d'une chaîne logistique, on suppose qu'on a déjà préparé les classeurs contenant les sous-graphes et les tables rarement modifiées, dites "pérennes" : PARAMETERS, CROPS, CROP_PARTS, PRODUCTS, RESOURCES et COMBINATIONS.

La définition du scénario commence par la saisie des tables LOCATIONS et DEMANDS. Par rapport au chapitre 3, deux tables sont ajoutées. La première est la table CHAINS qui décrit, pour chaque site de type ferme et chaque culture, la plante cultivée CID et le nom du classeur décrivant la chaîne, HARV_CHAIN. La seconde est la table STORAGE, qui précise pour chaque site de type stockage centralisé le nom du classeur décrivant les états de ce genre de site, STOR_CHAIN. Les distances peuvent ensuite calculées dans le classeur "DataOKDist" à l'aide du module DISTANCES.

La macro VBA de génération automatique du graphe état-tâche commence par lire la table LOCATIONS et sa colonne LType.

Cas des sites de type Farm

La macro récupère les plantes cultivées et les noms des classeurs des chaînes de récolte dans la table CHAINS. Elle ouvre ces classeurs et recopie les sous-graphes dans les tables STATES, TASKS, ST_ARCS, TS_ARCS et LOCA_FAM du scénario à étudier.

Pour éviter toute ambiguïté, l'identifiant choisi pour chacun des éléments de ces tables est une concaténation de la localisation, de l'identifiant standard, et de la plante en question. Par exemple, le champ cultivé est noté S1 dans la chaîne de récolte du colza (plante C1) et celle du miscanthus (C2). Si deux fermes L1 et L2 cultivent les deux plantes, ces deux états seront codés L1S1C1 et L1S1C2 dans la première, L2S1C1 et L2S1C2 dans la seconde.

Il faut ensuite créer les tâches de transport. La macro ouvre le classeur Excel DataOkDist afin de trouver le stockage le plus proche. Elle crée enfin pour chaque produit les tâches de transport des différents produits entre la ferme et ce site, en ajoutant les distances.

Cas des sites de type Storage

Le processus est similaire sauf qu'on recopie des sous-graphes de sites de stockage et qu'on ajoute des tâches de transport vers la raffinerie. Il est possible d'effectuer un transport vers un autre site de stockage si les transferts de produit entre stockages sont autorisés.

Cas du site de type Refinery

La macro ajoute au classeur du scénario l'ensemble des états et des tâches relatifs à la raffinerie. Il s'agit des stocks-tampons et leur connexion à la tâche de conversion (Refine).

Il faut signaler que certaines données sont remplies par défaut avec des valeurs moyennes, comme les fenêtres d'ouverture de chaque état et la taille maximale de la flotte de véhicules sur chaque site $LFNMax_{l,f}$. Bien entendu, l'utilisateur peut modifier ces valeurs dans les tables générées automatiquement.

La troisième fonction du module est le "database-checker", qu'il est conseillé d'appliquer avant de lancer le module OPTIMIZATION. Cette fonction s'assure de la cohérence des données. Par exemple, elle vérifie que les attributs des tables ont des valeurs dans le domaine de définition spécifié. Elle repère aussi les attributs non renseignés et effectue une correction adéquate en fonction de l'attribut (calcul automatique ou valeur par défaut).

6. Module OPTIMIZATION

Il y a une différence entre le développement des modèles mathématiques à l'aide d'OPL STUDIO et leur intégration dans BIOSCO.

Pour développer et tester un modèle, il faut les saisir en langage OPL dans un fichier-texte, à l'aide de l'éditeur d'OPL STUDIO. Le langage de modélisation OPL permet de définir des tableaux de données et de variables, de charger les données à partir de fichiers Excel et d'écrire des équations avec des sommes indicées sur ces tableaux. On obtient ainsi un modèle générique, largement indépendant des données. Quand on lance la résolution à partir de l'éditeur, le modèle est passé au programme OPLRUN, qui génère le programme linéaire numérique dans un autre fichier puis appelle le solveur CPLEX. Les résultats sont ensuite récupérés pour être affichés dans l'éditeur.

Une fois le modèle générique débogué, le fichier qui le contient est inclus dans les fichiers de BIOSCO. Il n'a plus besoin d'être modifié et il peut être caché à l'utilisateur. Pour résoudre le modèle à partir d'Excel, on n'a plus besoin de l'éditeur d'OPL STUDIO : c'est l'application VBA qui appelle OPLRUN. Elle récupère aussi les résultats et effectue certains calculs après l'optimisation (consommations d'énergie et rejets de CO₂).

Bien qu'OPL STUDIO soit mis gratuitement à disposition des doctorants et enseignants-chercheurs, un industriel à qui on transfère le logiciel-prototype doit acquérir une licence.

Pour avoir aussi une implémentation gratuite de BIOSCO, nous avons traduit les modèles mathématiques dans un environnement du domaine public, GUSEK. Le langage de modélisation est GMPL, un sous-ensemble d'AMPL, et le solveur est GLPK. GLPK ne bénéficie pas d'autant d'années de développement que CPLEX et il est donc moins rapide. En particulier, il n'est pas encore « multi-thread » comme Cplex.

7. Module REPORTING

Ce module stocke les résultats bruts du scénario (valeurs des variables) dans la base de données, sous forme de tables qui sont visibles dans le MCD qui clôture le chapitre 3. Il permet ensuite de produire différents états synthétiques et graphiques :

- répartition des différents coûts,
- équipements utilisés pour chaque site,
- caractéristiques de la chaîne dans chaque période (flux, niveaux des stocks),
- consommations d'énergie et rejets de CO₂.

8. Interface-utilisateur

Etant donné que l'utilisateur final de l'application n'est pas toujours un spécialiste en modélisation mathématique, l'interface-utilisateur doit être suffisamment intuitive, simple et facile d'usage. La figure 6.2 montre celle de BIOSCO. Il s'agit d'un classeur Excel qui permet d'instancier un scénario, d'exécuter le modèle mathématique et de récupérer les résultats. A tout moment, l'utilisateur peut consulter et modifier les tables de la base de données, disponibles sous forme d'onglets.

En cliquant sur Start, on passe à la figure 6.3 qui constitue le menu principal. Avec le bouton "INPUT DATA MANAGEMENT", l'utilisateur peut préparer les tables qui vont être utilisées pour la génération automatique du graphe état-tâche, voir la figure 6.4.



Figure 6.2. Ecran d'accueil de l'application BIOSCO.

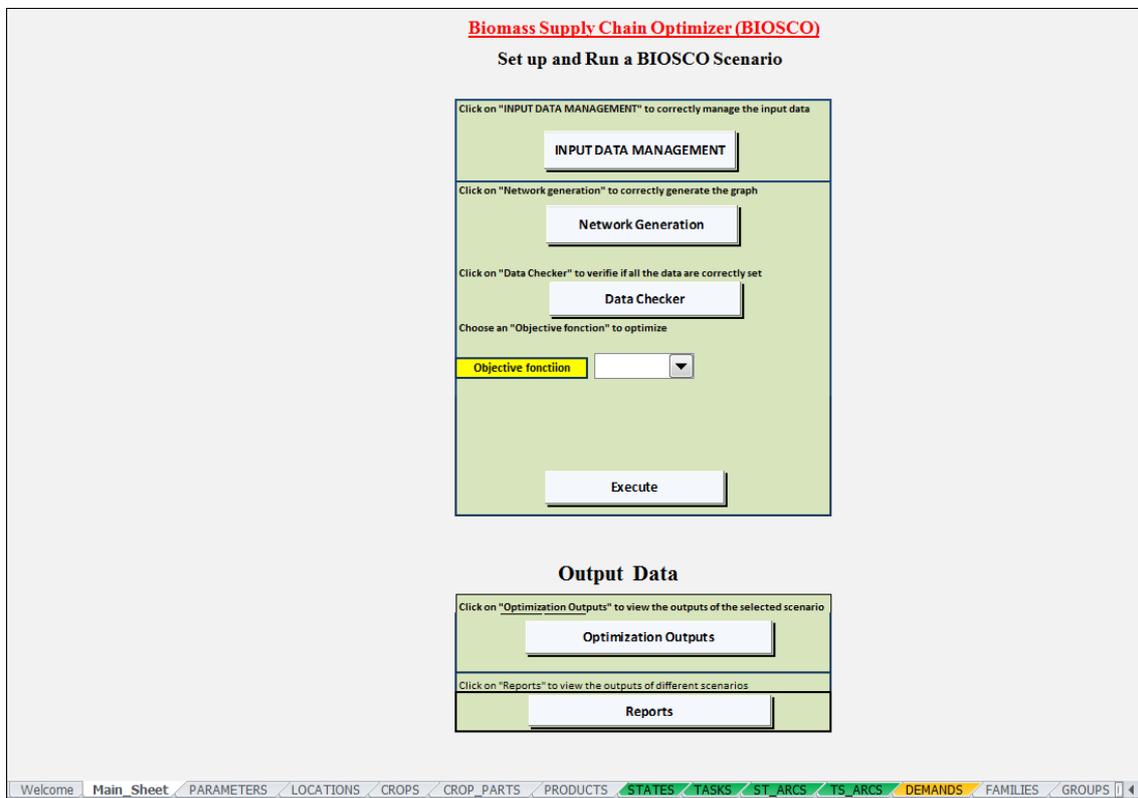


Figure 6.3. Capture d'écran de BIOSCO avec le menu principal.

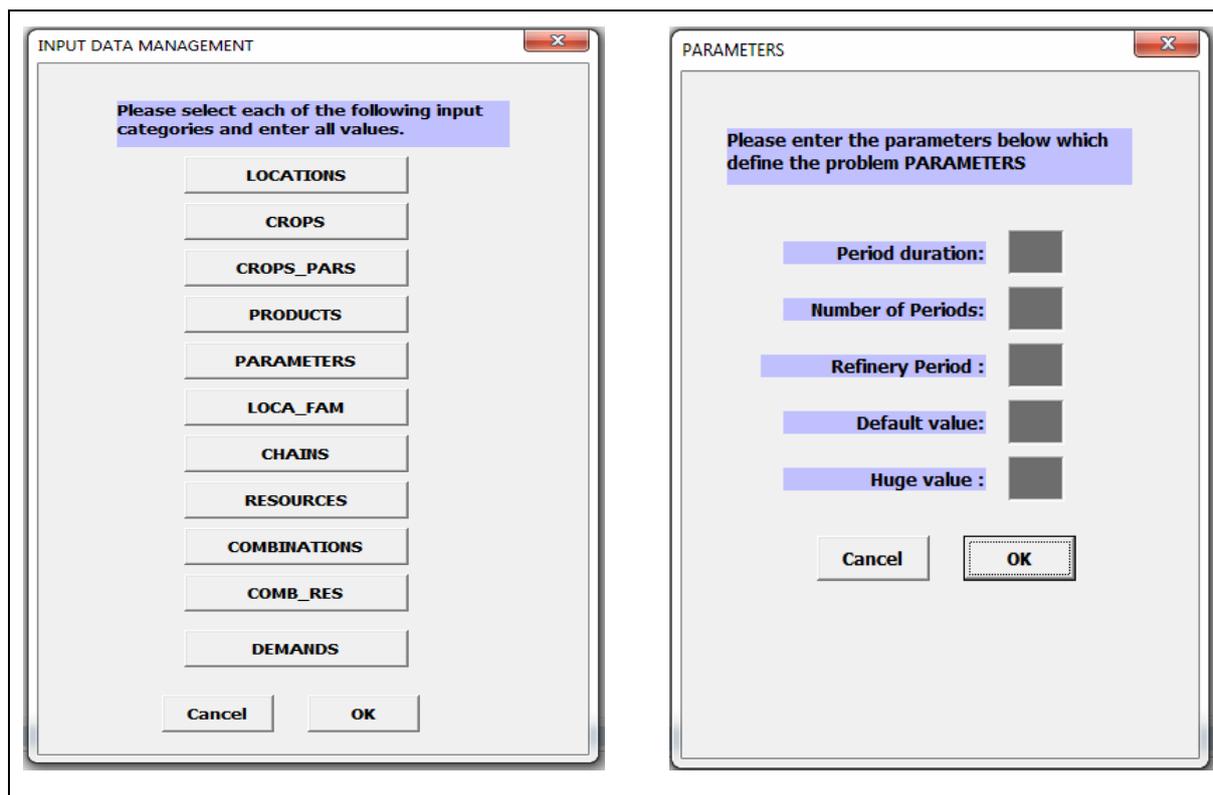


Figure 6.4. Capture d'écran de BIOSCO avec deux sous-menus de gestion de données.

Après préparation de ces données (LOCATIONS, PARAMETERS, etc.), l'étape suivante consiste à générer le graphe automatiquement (copie des chaînes de récolte, établissements des connexions nécessaires) avec le bouton "Network generation".

Une fois que le "data-checker" a été appliqué en répondant "OK" et que la fonction-objectif à minimiser a été choisie (pour le moment, minimisation des coûts uniquement), le module OPTIMIZATION peut être lancé avec le bouton "Execute".

Finalement, les résultats sont disponibles pour l'analyse dans la rubrique "Output Data", sous forme condensé (bouton "Optimization outputs") ou sous forme de rapports détaillés (bouton "Reports").

9. Recommandations

Cette section a décrit le logiciel-prototype en l'état. Le projet prévoit le recrutement d'un ingénieur pour convertir ce prototype en un logiciel diffusable, avec tous les compléments que cela suppose (procédure d'installation, procédures de sauvegarde et de reprise, exemples de scénarios, manuel de référence et manuel pour l'utilisateur).

De par sa construction (application desktop), l'installation, la maintenance et les mises à jour chez un client ne seront pas triviales et soulèveront des problèmes potentiels. Il faudra par exemple installer le modeleur OPL STUDIO, avoir Office avec Excel sur le poste de travail ainsi que le moteur ODBC compatible.

Et si jamais, pour une raison ou une autre, l'utilisateur souhaite utiliser une vraie base de données à la place d'Excel, elle doit être installée et pouvoir ensuite être maintenue facilement en suivant les mises à jour éventuelles de BIOSCO. C'est également le cas pour le logiciel de calcul de distance qui nécessite des mises à jour fréquentes. Etant donné ce manque de flexibilité notoire de l'application, nous recommandons à la place une application "web-based" (figure 6.5).

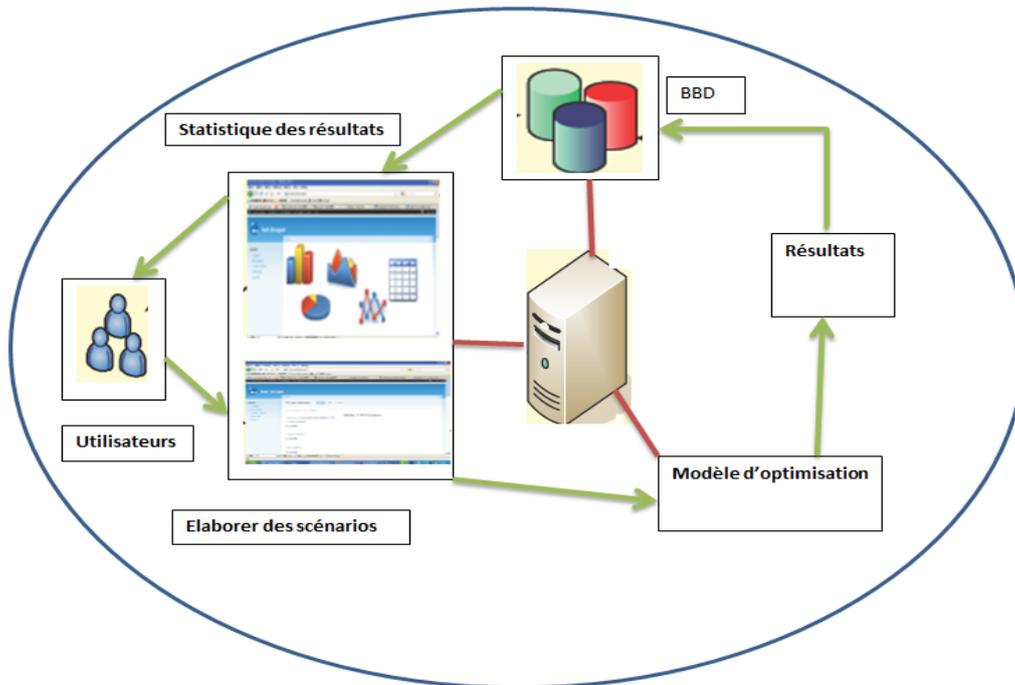


Figure 6.5 Future architecture « web-based » recommandée pour BIOSCO.

Cette proposition consiste en une application divisée en deux parties :

- Un programme-client avec une interface-utilisateur graphique (GUI), qui sera en local sur la machine de l'utilisateur et permettra à ce dernier de travailler à distance;
- Un ensemble de programmes côté serveur incluant le modèle mathématique ainsi qu'une base de données, pour répondre aux demandes de l'interface-utilisateur graphique et fournir des informations aux utilisateurs.

Grâce à cette plateforme, les utilisateurs pourront accéder au système via des navigateurs web et ainsi concevoir des scénarios d'approvisionnement en biomasse, les stocker sur le serveur, les modifier, exécuter les modèles et récupérer les résultats.

Avec un tel système, les utilisateurs pourront tirer parti de la puissance de calcul d'un gros serveur et avoir la garantie que les données cartographiques relatives au réseau routier ainsi que celles sur les plantes et les équipements soient toujours à jour, sans qu'une mise à jour directe de l'application desktop soit nécessaire. La gestion centralisée de l'application permettra aussi une grande flexibilité. Il sera par exemple possible d'ajouter des décisions stratégiques (placement de raffineries), des versions multi-objectifs des modèles et des heuristiques pour les cas de grande taille.

10. Conclusion

Ce chapitre a montré comment intégrer la base de données du chapitre 3 et les modèles mathématiques des chapitres 4 et 5 dans un logiciel-prototype. Ce travail de développement est nécessaire pour permettre à des utilisateurs non spécialistes des logiciels comme OPL STUDIO de résoudre les modèles de manière transparente, en décrivant librement leur chaîne logistique. En particulier, la réalisation du système de génération automatique du graphe état-tâches à partir de sous-graphes de base n'a pas été triviale.

La réalisation du prototype illustre aussi les nombreux aspects périphériques à gérer, comme la réalisation d'une interface-utilisateur conviviale et l'intégration de logiciels auxiliaires (MapPoint, OPL STUDIO). Comme ces dépendances multiples entre de nombreux composants peuvent poser des problèmes d'installation et ensuite de maintenance, une architecture client-serveur sur Internet a été recommandée.

Conclusion générale

Aujourd'hui, il ne fait aucun doute que l'intérêt pour la bio-économie va croissant. Ce concept, en lien étroit avec le développement durable, vise à l'efficacité économique en intégrant d'autres aspects qui sont généralement associés à la qualité environnementale par le biais de la valorisation de la biomasse. L'idée est de recourir en priorité à des ressources dont le cycle biologique est conciliable avec l'échelle humaine, contrairement aux ressources fossiles.

Dans cette mouvance, les biotechnologies ont un rôle central afin d'offrir des solutions dans divers domaines, allant de la production à la santé par exemple, et requiert diverses compétences. Les principales sont fondées sur la chimie et la biologie. Cependant, une vision centrée sur ces domaines n'est pas suffisante. C'est un regard macroscopique qu'il faut porter sur la question si on souhaite réellement viser la durabilité. Il est indispensable d'impliquer divers domaines comme le transport ou le stockage, ainsi que plusieurs acteurs à l'instar des gouvernements, des agriculteurs et des industriels. En effet, la bio-raffinerie, socle de la bio-économie, donne lieu à une logistique complexe, coûteuse et qui nécessite une optimisation pour plus de flexibilité et d'efficacité.

Dans cette thèse, nous avons étudié des chaînes d'approvisionnement de biomasses pour une bio-raffinerie au niveau tactique, sur un horizon d'un an découpé en jours ou semaines. Nous récapitulons nos contributions et évoquons quelques perspectives ouvertes pour l'avenir.

1. Résumé des contributions et travaux effectués

La thèse a l'originalité de se situer à l'intersection de la recherche opérationnelle et du monde de la biomasse. Un premier travail, non visible dans ce document mais qui nous a demandé beaucoup de temps, a consisté à prendre connaissance des concepts et de l'abondante terminologie sur la production et la logistique de biomasses agricoles.

Il a fallu ensuite préciser le type de bioraffinerie envisagée, avec les matières premières nécessaires, les schémas de culture et d'approvisionnement, et formuler diverses hypothèses pour bien délimiter la problématique à étudier. Cette étape a été résumée dans le chapitre 1.

Une revue de la littérature sur la logistique de biomasse a ensuite été préparée et incluse in extenso dans un livrable pour PIVERT (96 pages et 215 références), puis présentée sous forme abrégée dans le chapitre 2. Outre le panorama qu'elle fournit sur la littérature existante, cette revue fait la synthèse des différents modèles et méthodes de résolution des problèmes d'optimisation en logistique de biomasse. Elle montre que des problèmes originaux et intéressants sont soulevés par la conception de chaînes d'approvisionnement de biomasse et qu'un nombre croissant de publications récentes propose des modèles de plus en plus réalistes. Le bilan de cet état de l'art montre cependant des lacunes et limitations dans les modèles actuels. En particulier, aucun travail publié n'englobe simultanément toutes les caractéristiques du problème que nous souhaitons résoudre.

Ensuite, nous nous sommes aperçu qu'une étape de modélisation des activités de la chaîne logistique devait être entreprise avant de réaliser des modèles mathématiques, pour éviter que ces derniers ne dépendent d'une structure figée de chaîne et aussi pour laisser les futurs utilisateurs libres de décrire la chaîne de leur choix. L'idée a donc été de découpler autant que possible la description de la chaîne logistique et les modèles d'optimisation.

Le chapitre 3 a résumé ce travail de modélisation du réseau logistique, basé sur le formalisme des graphes états-tâches. Nous avons montré que ces graphes, conçus initialement pour les processus en pétrochimie, étaient utilisables pour décrire des chaînes de biomasse, moyennant certaines extensions comme la prise en compte de sites géographiques et de ressources de transport. Nous avons aussi proposé un système de modélisation des chaînes de récolte sous forme de sous-graphes prédéfinis, pouvant ensuite être recopiés dans les exploitations agricoles utilisant la même chaîne. Nous pensons qu'il est difficile au niveau tactique d'être plus précis que les graphes états-tâches, à moins de basculer dans des modèles de simulation qui conviennent mal pour l'optimisation.

Le rôle du chapitre 3 a été aussi de voir comment organiser les données codant le graphe état-tâche et les autres données (équipements utilisables par exemple) sous forme d'un modèle entité-relation, pouvant ensuite être traduit sous forme de tables dans un logiciel de bases de données. Ces travaux ont permis d'aboutir à un système flexible de modélisation des données, décrivant sans ambiguïté des chaînes d'approvisionnement en biomasse avec une structure de base de données. Bien que n'étant pas un vrai SGBD, Excel a été sélectionné provisoirement pour sa simplicité.

Après l'élaboration de deux modèles préliminaires qui n'étaient pas assez fins et qui n'ont pas été présentés faute de place, nous avons développé dans le chapitre 4 un modèle mathématique dit "complet", capable d'optimiser des chaînes d'approvisionnement de biomasse prenant en compte simultanément plusieurs périodes, plusieurs biomasses, et plusieurs chaînes de récoltes avec une gestion fine et flexible des équipements. Ce modèle, un programme linéaire à variables mixtes, est suffisamment flexible pour être utilisé et facilement adapté pour la plupart des chaînes d'approvisionnement de biomasse au niveau tactique. Son principal avantage est d'être "data-driven", c'est-à-dire qu'il est généré automatiquement à partir des tables de la base de données.

Le modèle implémenté en OPL et résolu par CPLEX a été testé sur des cas à une puis 10 zones de production de biomasse ("fermes"), deux biomasses différentes, trois sites de stockage et un horizon de 364 jours. Le temps de résolution sur un PC est acceptable, moins de 3 minutes, mais le temps de calcul augmente rapidement avec la taille des instances. La limite est d'environ 100 fermes, non pas à cause du temps de calcul, mais à cause du préprocesseur qui n'a pas assez de mémoire pour générer le programme linéaire.

Des améliorations ont donc été proposées à la fin du chapitre 4 : élimination de variables entières, coûts d'utilisation incluant l'amortissement des matériels, simplification du système de gestion de ressources. Le modèle résultant induit seulement une légère dégradation du coût des solutions (+0.5% sur les exemples testés). Il a permis de résoudre un cas à 100 fermes, 25 stockages et 52 périodes d'une semaine en deux heures et quart de calcul, avec une limite autour de 150 fermes.

La contribution du chapitre 5 repose sur la proposition d'un modèle mathématique plus compact que le précédent, capable de traiter des instances correspondant au système réel envisagé pour une bioraffinerie située près de Compiègne, avec plus de 400 fermes, 46 sites de stockage et un horizon

de planification de 52 semaines. Afin de trouver le bon compromis entre qualité et finesse de la solution, un programme linéaire à variables continues, toujours basé toujours sur le modèle de réseau et le modèle de données vus dans le chapitre 3, agrège certaines activités d'une ferme à travers la gestion de *macro-tâches*. Ces dernières sont utilisées comme des boîtes noires de manière à approximer les coûts inhérents à la récolte au sein d'une exploitation.

La construction du jeu d'essai réel illustre les nombreuses difficultés pour obtenir des données dans une étude de recherche opérationnelle : sources variées et hétérogènes, secret statistique et opacité de certains acteurs, filtrage et mise en forme des données nécessaires. Il a aussi fallu utiliser un logiciel tiers pour calculer les plus courts chemins dans un réseau routier. Des tests et analyses des résultats ont été présentés : on peut résoudre en moins d'une minute un programme linéaire ayant près de 600 000 variables et quasiment autant de contraintes. Le chapitre 5 a finalement listé des extensions possibles des modèles, qui semblent assez faciles grâce à nos choix de conception.

Le chapitre 6 a décrit les principes de conception et de développement d'une application appelée BIOSCO (Biomass Supply Chain Optimizer) avec un ensemble de fonctionnalités et de perspectives. Ce logiciel reprend les contributions des trois précédents chapitres en intégrant la base de données et les modèles mathématiques développés. L'intérêt est de valoriser le travail de thèse en offrant à des utilisateurs non spécialistes la possibilité de résoudre les modèles en décrivant facilement leur propre chaîne logistique. Ceci est permis grâce à un système de génération automatique non triviale du graphe états-tâches à partir de sous-graphes de base prédéfinis.

Cette partie a demandé un travail très important de développement informatique : des compétences pointues sur VBA et OPL Studio, d'une part, de paramétrage et de débogage d'autre part, sans oublier les nombreux aspects périphériques à gérer, comme la réalisation d'une interface-utilisateur conviviale et l'intégration de logiciels auxiliaires (MapPoint, OPL STUDIO).

Ces recherches ont déjà donné lieu à un article publié dans la revue internationale réputée *Renewable Energy*, à un autre article soumis à *European Journal of Operational Research* (EJOR) et à un livrable de PIVERT de 96 pages. Elles ont également fait l'objet de 2 conférences nationales et 3 internationales, dont deux publiées dans des actes, et enfin d'un prototype d'application d'aide à la décision.

2. Perspectives

Au terme de cette thèse, les outils nécessaires à l'étude du coût relatif à une chaîne logistique dédiée à l'approvisionnement d'une bio-raffinerie de proximité ont été développés. Cependant, beaucoup reste à faire. A un niveau pratique, l'application informatique nécessite quelques étapes de tests et de développement supplémentaires pour la rendre plus robuste. Mais, au-delà des contributions mises en lumière ci-dessus, un certain nombre de perspectives peuvent être dégagées.

En premier lieu, nous nous sommes intéressés à une fonction-objectif financière focalisée sur l'approvisionnement. Les rejets de CO2 et les consommations d'énergie sont déduits de la solution. Il pourrait être intéressant d'ajouter au logiciel la possibilité de choisir ou de combiner diverses fonctions-objectifs : minimiser les émissions de gaz à effet de serre (GES), minimiser les coûts d'approvisionnement, maximiser le profil de la raffinerie, etc.

Au niveau de la méthode de résolution proposée, à savoir une résolution exacte d'un modèle mathématique par un solveur, des perspectives sont ouvertes pour des instances de très grandes tailles. Il s'agit d'analyser des algorithmes heuristiques. Plusieurs options sont envisageables, allant d'une heuristique constructive à une métaheuristique. Cependant, aux vues des différents types de décisions à prendre, la combinaison d'approches exactes et approchées semble plus pertinente. On peut penser par exemple à utiliser la programmation linéaire pour la résolution de sous-problèmes, via une matheuristique. Dans cet esprit, cinq idées nous paraissent intéressantes à tester :

- Décomposition en deux sous-problèmes : un problème de production de biomasse qui se concentre sur les activités des fermes, telles que la récolte, et un problème relatif à l'approvisionnement qui intègre les activités de post-production (après récolte), telles que la logistique de transport vers les stocks ou la raffinerie. Cette approche peut être insérée dans un schéma itératif alternant la résolution de ces deux problèmes. Une telle méthode a été mise en évidence par Shastri et al. (2011b).
- Exploiter le fait que les plantes (colza, miscanthus, cameline) ont des fenêtres de récolte disjointes, autrement dit opter pour une décomposition par plante. Il faut alors s'assurer du respect des capacités des ressources en dehors des fermes, en particulier concernant le stockage.
- Relaxation lagrangienne sur les contraintes en lien avec la continuité de la moisson ("no wait") et avec les capacités de stockage des différents sites.
- Planification à la semaine, à travers la granularité hebdomadaire utilisée dans les modèles proposés par exemple, pour ensuite affiner le plan obtenu sur une granularité journalière. La vision macroscopique induite par une agrégation par semaine engendre une sous-optimalité qui pourrait ainsi être réduite à travers schéma itératif.
- Les modèles sont parfois très contraints ou infaisables quand la production est proche de la capacité de stockage dans le système, et la programmation linéaire ne permet pas facilement de voir la cause réelle de l'infaisabilité. Il nous paraît donc également intéressant de tester la programmation par contraintes.

Concernant les données, qui sont actuellement considérées comme déterministes, une prise en compte d'incertitudes semble pertinente. En effet, l'horizon observé est à minima trimestriel, mais souvent annuel, voire pluriannuel. Ceci est inhérent à la combinaison de la saisonnalité des récoltes et de la demande relativement régulière de la bioraffinerie. Cette vision à moyen/long terme est souvent peu compatible avec la gestion précise des activités journalières.

Même si le modèle proposé reste sur un niveau tactique, les informations sur la disponibilité de la biomasse et sur la demande sont essentielles. Cependant, l'estimation actuellement fournie est sujette à évoluer. En effet, la quantité récoltée et la date de maturité sont en grande partie liées aux sols et aux conditions météorologiques. La demande est basée également sur une prévision. Il serait donc intéressant de recourir à la réalisation de variables aléatoires sur ces deux données afin d'étudier la robustesse du modèle.

Une autre perspective, ambitieuse, serait d'étudier une chaîne complète allant de l'allocation des zones de cultures à des biomasses précises jusqu'à la livraison des produits finis aux clients finaux. Cependant le talon d'Achille d'une telle étude est et reste principalement l'acquisition de l'ensemble des données nécessaires. Il n'a déjà pas été toujours facile de disposer des données nécessaires dans le cadre de l'étude présente, mais l'analyse d'une chaîne complète promet des difficultés supplémentaires. Elle implique effectivement différentes parties, ayant des objectifs propres, et des politiques de gestion parfois confidentielles.

Une dernière extension nous paraît également pertinente. Elle concerne l'intégration d'une durée de plusieurs périodes sur l'acheminement de produits, une possibilité inutilisée pour notre étude concernant une raffinerie de proximité. Le fait de permettre l'allongement des temps de transport offrirait l'opportunité de s'intéresser à des cas de plus grande étendue à l'instar du niveau régional ou national.

Au vu des travaux actuellement disponibles sur la logistique de la biomasse, des enjeux qui lui sont liés, et des perspectives énoncées précédemment, la recherche opérationnelle et l'informatique sont des disciplines vouées à un brillant avenir dans le développement économique des bio-raffineries. D'une part, elles permettent l'évaluation des performances d'un système existant ou identifié, en faisant ressortir par exemple des goulets d'étranglement ou les points d'améliorations possibles des systèmes logistiques. D'autre part, comme ce fut le cas dans cette thèse, ces spécialités permettent d'optimiser la logistique afin de la rendre fiable, viable et efficace. Elles font ainsi partie des compétences requises au déploiement de la valorisation de la biomasse et viennent compléter les apports de domaines tels que la chimie végétale et l'étude de processus de conversion.

Pour terminer, signalons qu'un autre projet de PIVERT appelé AMBRE (Approvisionnement Multi-biomasse sur des Bassins Régionaux) a démarré pour généraliser les travaux de cette thèse. Alors que notre thèse s'est concentrée sur une raffinerie de proximité, son objectif est d'étudier la logistique sur un territoire correspondant à deux régions (Picardie plus Champagne Ardenne), des transports intermodaux et plusieurs raffineries, déjà placées ou à placer. Cette étude présentera donc des aspects stratégiques.

Bibliographie

- Ahumada, O., Villalobos, J.R., 2009. Applications of planning models in the agri-food supply chain: A review. *European Journal of Operational Research*, 195, 1–20.
- Akgul, O., Zamboni, A., Bezzo, F., Shah, N., Papageorgiou, L.G., 2011. Optimization-based approaches for bioethanol supply chains. *Industrial and Engineering Chemistry Research*, 50(9), 4927–4938.
- Akgul, O., Shah, N., Papageorgiou, L.G., 2012. An optimization framework for a hybrid first/second generation bioethanol supply chain. *Computers and Chemical Engineering*, 42, 101–114.
- Akhtari, S., Sowlati, T., Day, K., 2013. Optimal flow of regional forest biomass to a district heating system. *International Journal of Energy Research*, 38(7), 954–964.
- Alam, M.B., Shahi, C., Pulkki, R., 2009. Wood biomass supply model for bioenergy production in Northwestern Ontario. *1st International Conference on the Developments in Renewable Energy Technology (ICDRET)*, 1–3.
- Alfonso, D., Perpiñá, C., Pérez-Navarro, A., Peñalvo, E., Vargas, C., Cárdenas, R., 2009. Methodology for optimization of distributed biomass resources evaluation, management and final energy use. *Biomass and Bioenergy*, 33(8), 1070–1079.
- Allen, J., Browne, M., Hunter, A., Boyd, J., Palmer, H., 1998. Logistics management and costs of biomass fuel supply. *International Journal of Physical Distribution and Logistics Management*, 28(6), 463–477.
- Alonso-Ayuso, A., Escudero, L.F., Garin, A., Ortuño, L.T., Pérez, G., 2003. An approach for strategic supply chain planning under uncertainty based on stochastic 0-1 programming. *Journal of Global Optimization*, 26, 97–124.
- Al-Othman, W.B.E., Lababidi, H., Alatiqi, I.M., Al-Shayji, K., 2008. Supply chain optimization of petroleum organization under uncertainty in market demands and prices. *European Journal of Operational Research*, 189(3), 822–840.
- Al-Sharrah, G., Elkamel, A., Almansoor, A., 2010. Sustainability indicators for decision-making and optimisation in the process industry: The case of the petrochemical industry. *Chemical Engineering Science*, 65(4), 1452–1461.
- Amigun, B., Musango, J.K., Stafford, W., 2011. Biofuels and sustainability in Africa. *Renewable and Sustainable Energy Reviews*, 15(2), 1360–1372.
- An, H., Wilhelm, W.E., Searcy, S.W., 2011a. Biofuel and petroleum-based fuel supply chain research: A literature review. *Biomass and Bioenergy*, 35(9), 3763–3774.
- An, H., Wilhelm, W.E., Searcy, S.W., 2011b. A mathematical model to design a lignocellulosic biofuel supply chain system with a case study based on a region in Central Texas. *Bioresource Technology*, 102(17), 7860–7870.
- An, H., Searcy, S.W., 2012. Economic and energy evaluation of a logistics system based on biomass modules. *Biomass and Bioenergy*, 46, 190–202.
- Atchison, J.E., Hettenhaus, J.R., 2004. Innovative methods for corn stover collecting, handling, storing and transporting. Report SR-510-33893, National Renewable Energy Laboratory (NREL), Golden, Colorado, USA.

- Awudu, I., Zhang, J., 2012. Uncertainties and sustainability concepts in biofuel supply chain management: A review. *Renewable and Sustainable Energy Reviews*, 16(2), 1359–1368.
- Awudu, I., Zhang, J., 2013. Stochastic production planning for a biofuel supply chain under demand and price uncertainties. *Applied Energy*, 103, 189–196.
- Ayoub, N., Martins, R., Wang, K., Seki, H., Naka, Y., 2007. Two levels decision system for efficient planning and implementation of bioenergy production. *Energy Conversion and Management*, 48(3), 709–723.
- Ba, B.H., Prins, C., Prodhon, C., 2014a. Revue de la littérature sur les modèles logistiques pour la biomasse. Livrable n°1 du WP1-P3, Institut pour la Transition Énergétique (ITE) PIVERT, 05/02/2014, 96 pages et 215 références.
- Ba, B.H., Prins, C., Prodhon C., 2014b. Un modèle tactique pour une chaîne d'approvisionnement de biomasse. 15ème congrès annuel de la Société Française de Recherche Opérationnelle et d'Aide à la Décision (ROADEF), Bordeaux, France, 26-28/02/2014.
- Ba, B.H., Prins, C., Prodhon, C., 2014c. An overview of operations research models for biomass supply chains, ICORES 2014 (3rd Int. Conf. on Operations Research and Enterprise Systems), Angers, 06-08/03/2014, 8 pages.
- Ba, B.H., Prins, C., Prodhon C., 2015a. Modèle tactique d'approvisionnement en biomasse pour une bio-raffinerie de proximité. 16ème congrès annuel de la Société Française de Recherche Opérationnelle et d'Aide à la décision (ROADEF), Marseille, France, 25-27/02/2015.
- Ba, B.H., Prins, C., Prodhon C., 2015b. A new tactical optimization model for bioenergy supply chain. *XIII International Conference on Applied Mathematics and Operation Research (ICAMOR 2015)*, Miami Beach, 9-10/03/2015, USA, 12 pages.
- Ba, B.H., Prins, C., Prodhon C., 2015c. A general mixed integer linear model for a local biorefinery supply chain. *27th European Conference on Operational Research (EURO 2015)* 12-15/07/2015, University of Strathclyde, Glasgow, United Kingdom.
- Ba, B.H., Prins, C., Prodhon, C., 2015d. Models for optimization and performance evaluation of biomass supply chains: an operations research perspective, *Renewable Energy* (Elsevier), déjà disponible en ligne, DOI : 10.1016/j.renene.2015.07.045.
- Ba, B.H., Prins, C., Prodhon C., 2015e. A generic tactical planning model to supply a bio-refinery with biomass. Soumis à *European Journal of Operational Research*.
- Bai, Y., Hwang, T., Kang, S., Ouyang, Y., 2011. Biofuel refinery location and supply chain planning under traffic congestion. *Transportation Research Part B: Methodological*, 45(1), 162–175.
- Bakos, G.C., Tsioliariidou, E., Potolias, C., 2008. Technoeconomic assessment and strategic analysis of heat and power co-generation (CHP) from biomass in Greece. *Biomass and Bioenergy*, 32(6), 558–567.
- Baños, R., Manzano-Agugliaro, F., Montoya, F.G., Gil, C., Alcayde, A., Gómez, J., 2011. Optimization methods applied to renewable and sustainable energy: A review. *Renewable and Sustainable Energy Reviews*, 15(4), 1753–1766.
- Barème d'entraide 2014-2015 de la Chambre d'Agriculture d'Indre-et-Loire : www.indre-et-loire.chambagri.fr/fileadmin/documents/site_1/internet/Accueil/BaremeEntraide2014-2015.pdf
- Batzias, F.A., Sidiras, D.K., Spyrou, E.K., 2005. Evaluating livestock manures for biogas production: a GIS based method. *Renewable Energy*, 30(8), 1161–1176.
- Bell, D.R., Silalertruksa, T., Gheewala, S.H., Kamens, R., 2011. The net cost of biofuels in Thailand—An economic analysis. *Energy Policy*, 39(2), 834–843.

- Bernardi, A., Giarola, S., Bezzo, F., 2013. Spatially Explicit Multiobjective Optimization for the Strategic Design of First and Second Generation Biorefineries Including Carbon and Water Footprints. *Industrial & Engineering Chemistry Research*, 52(22), 7170–7180.
- Bertsekas, D., 2005. *Dynamic programming and optimal control* (2 volumes), 3ème édition, Athena Scientific.
- Bertsimas, D., Tsitsiklis, J.N., 1997. *Introduction to linear optimization*. Athena Scientific.
- Bowling, I.M., Ponce-Ortega, J.M., and El-Halwagi, M.M., 2011. Facility Location and Supply Chain Optimization for a Biorefinery. *Industrial & Engineering Chemistry Research*, 50 (10), 6276–6286.
- Bradley, P., Fox, B., Schrage, L.E., 1987. *A guide to simulation*. Second edition. Springer.
- Branke, J., Deb, K., Miettinen, K., Slowinski, R., 2008. *Multiobjective Optimization*. Springer.
- Brechbill, S.C., Tyner, W.E., Ileleji, K.E., 2011. The Economics of Biomass Collection and Transportation and Its Supply to Indiana Cellulosic and Electric Utility Facilities. *BioEnergy Research*, 4(2), 141–152.
- Buffett, S., 2005. A Markov model for inventory level optimization in supply-chain management. In : B. Kégl et G. Lapalme (éd.), *Artificial Intelligence 2005*, Lecture Notes in Artificial Intelligence 3501, 133–144, Springer.
- Calvert, K., 2011. Geomatics and bioenergy feasibility assessments: taking stock and looking forward. *Renewable and Sustainable Energy Reviews*, 15(2), 1117–1124.
- Chen., H., Amodeo, L., Chu, F., Labadi, K., 2005. Modelling and performance evaluation of supply chains using batch deterministic and stochastic Petri nets. *IEEE Transactions on Automation Science and Engineering*, 2(2), 132–144.
- Chen, C.-W., Fan, Y., 2012. Bioethanol supply chain system planning under supply and demand uncertainties. *Transportation Research Part E: Logistics & Transportation Review*, 48(1), 150–164.
- Chinese, D., Meneghetti, A., 2005. Optimisation models for decision support in the development of biomass-based industrial district-heating networks in Italy. *Applied Energy*, 82(3), 228–254.
- Chinese, D., Meneghetti, A., 2009. Design of forest biofuel supply chains. *International Journal of Logistics Systems and Management*, 5(5), 525–550.
- Cucek, L., Varbanov, P.L., Klemes, J.J., Kravanja, Z., 2012. Total footprints-based multi-criteria optimisation of regional biomass energy supply chains. *Energy*, 44, 135–145.
- Culioli, J.C., 1994. *Introduction à l'optimisation*. Ellipses, Paris.
- Cundiff, J.S., Dias, N., Sherali, H.D., 1997. A linear programming approach for designing a herbaceous biomass delivery system. *Bioresources Technology*, 59 (1), 47–55.
- Cundiff, J.S., Grisso, R., 2008. Containerized handling to minimize hauling cost of herbaceous biomass. *Biomass and Bioenergy*, 32(4), 308–313.
- Dagnall, S., Hill, J., Pegg, D., 2000. Resource mapping and analysis of farm livestock manures: assessing the opportunities for biomass-to-energy schemes. *Bioresource Technology*, 71(3), 225–234.
- Dal-Mas, M., Giarola, S., Zamboni, A., Bezzo, F., 2011. Strategic design and investment capacity planning of the ethanol supply chain under price uncertainty. *Biomass and Bioenergy*, 35(5), 2059–2071.
- Daskin, M.S., 1995. *Network and discrete location: models, algorithms, and applications*. Wiley.
- Deb, K., 2001. *Multi-objective optimization using evolutionary algorithms*. Wiley.

- Delivand, M.K., Barz, M., Gheewala, S.H., 2011. Logistics cost analysis of rice straw for biomass power generation in Thailand. *Energy*, 36(3), 1435–1441.
- Demirbas, A., 2009. Political, economic and environmental impacts of biofuels: A review. *Applied Energy*, 86, 108–117.
- De Meyer, A., Cattrysse, D., Rasinmäki, J., Van Orshoven, J., 2014. Methods to optimise the design and management of biomass-for-bioenergy supply chains: A review. *Renewable and Sustainable Energy Reviews*, 31, 657-670.
- De Mol, R.M., Jogems, M., Van Beek, P.V, Gigler, J.K, 1997. Simulation and optimization of the logistics of biomass fuel collection. *Netherlands Journal of Agricultural Science*, 45, 219-228.
- Diekema, W.H., De Mol, R.M., Annevelink, E., Elbersen, H.W., 2005. Combining goals in the logistics bio-energy chains, 14th European Biomass Conference, Paris, 495–498.
- Difs, K., Wetterlund, E., Trygg, L., Söderström, M., 2010. Biomass gasification opportunities in a district heating system. *Biomass and Bioenergy*, 34(5), 637–651.
- DOE (US Department of Energy), 2011. Biofuels & greenhouse gaz emissions : myths versus facts, <http://www.energy.gov/>.
- Dornburg, V., Faaij, P.C., 2001. Efficiency and economy of wood-fired biomass energy systems in relation to scale regarding heat and power generation using combustion and gasification technologies. *Biomass and Bioenergy*, 21, 91–108.
- Duku, M.H., Gu, S., Hagan, E.B., 2011. A comprehensive review of biomass resources and biofuels potential in Ghana. *Renewable and Sustainable Energy Reviews*, 15(1), 404–415.
- Dukulis, I., Birzietis, G., Kanaska, D., 2008. Optimization models for biofuel logistic systems. 13th *International Scientific Conference on Engineering for Rural Development*, Jelgava, Lettonie, 29-30/05/2008, 283–289.
- Dunnett, A., Adjiman, C., Shah, N., 2007. Biomass to heat supply chains applications of process optimization. *Transactions of IChemE, Part B: Process Safety and Environmental Protection*, 85(5), 419–429.
- Ebadian, M., Sowlati, T., Sokhansanj, S., Stumborg, M., Townley-Smith, L., 2011. A new simulation model for multi-agricultural biomass logistics system in bioenergy production. *Biosystems Engineering*, 110(3), 280–290.
- Ekşioğlu, S.D., Acharya, A., Leightley, L.E., Arora, S., 2009. Analyzing the design and management of biomass-to-biorefinery supply chain. *Computers & Industrial Engineering*, 57(4), 1342–1352.
- Ekşioğlu, S.D, Li, S., Zhang, S., Sokhansanj, S., Petrolia, D., 2010. Analyzing the Impact of Intermodal Facilities to the Design and Management of Biofuels Supply Chain. *Transportation Research Record: Journal of the Transportation Research Board*, 144-151.
- Elghali, L., Clift, R., Sinclair, P., Panoutsou, C., Bauen, A., 2007. Developing a sustainability framework for the assessment of bioenergy systems. *Energy Policy*, 35(12), 6075–6083.
- Elms, R.D., El-Halwagi, M.M., 2010. The effect of greenhouse gas policy on the design and scheduling of biodiesel plants with multiple feedstocks. *Clean Technologies and Environmental Policy*, 12(5), 547–560.
- Eriksson, L.O., Björheden, R., 1989. Optimal storing, transport and processing for a forest-fuel supplier. *European Journal of Operational Research*, 43(1), 26–33.
- European Commission. Proposal for a directive of the European parliament and of the council on the promotion of the use of energy from renewable sources. COM(2008) 19 final. Brussels.

- Feng, Y., D'Amours, S., LeBel, L., Nourelfath, M., 2010. Integrated bio-refinery and forest products supply chain network design using mathematical programming approach. Research Report 2010-50, CIRRELT, Montréal, Canada.
- Fiedler, P., Lange, M., Schultze, M., 2007. Supply Logistics for the Industrialized Use of Biomass. Principles and Planning Approach. *International Symposium on Logistics and Industrial Informatics (LINDI 2007)*, Wildau, Allemagne, 41–46.
- Forsberg, G., 2000. Biomass energy transport analysis of bioenergy transport chains using life cycle inventory method. *Biomass and Bioenergy*, 19, 17–30.
- Freppaz, D., Minciardi, R., Robba, M., Rovatti, M., Sacile, R., Taramasso, A., 2004. Optimizing forest biomass exploitation for energy supply at a regional level. *Biomass and Bioenergy*, 26(1), 15–25.
- Frombo, F., Minciardi, R., Robba, M., Rosso, F., Sacile, R., 2009. Planning woody biomass logistics for energy production: A strategic decision model. *Biomass and Bioenergy*, 33(3), 372–383.
- Gabbar, H.A., 2009. Engineering design of green hybrid energy production and supply chains. *Environmental Modelling & Software*, 24(3), 423–435.
- Gallis, C., 1996. Activity oriented stochastic computer simulation of forest biomass logistics in Greece. *Biomass and Bioenergy*, 10, 377–382.
- Gan, J., 2007. Supply of biomass, bioenergy, and carbon mitigation: Method and application. *Energy Policy*, 35(12), 6003–6009.
- Gemtos, T.A., Tsircoglou, T., 1999. Harvesting of cotton residue for energy production. *Biomass and Bioenergy*, 16, 51–59.
- Giarola, S., Zamboni, A., Bezzo, F., 2011. Spatially explicit multi-objective optimisation for design and planning of hybrid first and second generation biorefineries. *Computers & Chemical Engineering*, 35(9), 1782–1797.
- Gnansounou, E., 2010. Production and use of lignocellulosic bioethanol in Europe: Current situation and perspectives. *Bioresource Technology*, 101(13), 4842–4850.
- Gold, S., Seuring, S., 2011. Supply chain and logistics issues of bio-energy production. *Journal of Cleaner Production*, 19(1), 32–42.
- Gomes, T.C., Alvelos, F.P., Carvalho, M.S., 2012. Design of Wood Biomass Supply Chains. *Computational Science and Its Applications – ICCSA 2012*. Lecture Notes in Computer Science ,7335, 30-44.
- Graham, R.L., English, B.C., Noon, C.E., 2000. A Geographic Information System-based modeling system for evaluating the cost of delivered energy crop feedstock. *Biomass and Bioenergy*, 18(4), 309–329.
- Graham, R.L., Liu, W., Downing, M., Noon, C.E., Daly, M., Moore, A., 1997. The effect of location and facility demand on the marginal cost of delivered wood chips from energy crops : a case study of the state of Tennessee. *Biomass and Bioenergy* 13(3), 117-123.
- Gronalt, M., Rauch, P., 2007. Designing a regional forest fuel supply network. *Biomass and Bioenergy*, 31(6), 393–402.
- Guillen-Gosalbez, G., Grossmann, I.E., 2009. Optimal design and planning of sustainable chemical supply chains under uncertainty. *AIChE Journal*, 55(1), 99–121.
- Gunnarsson, H., Rönnqvist, M., Lundgren, J.T., 2004. Supply chain modelling of forest fuel. *European Journal of Operational Research*, 158(1), 103–123.
- Haddad, M.A., Anderson, P.F., 2008. A GIS methodology to identify potential corn stover collection locations. *Biomass and Bioenergy*, 32(12), 1097–1108.

- Hamelinck, C.N., Faaij, A.P.C., Den Uil, H., Boerrigter, H., 2003. System Analysis of Biomass Derived FT Liquids; Technical Options, Process Optimization and Development Potential. Research Report, Copernicus Institute, Utrecht University.
- Hamelinck, C.N., Suurs, R. A., Faaij, A.P.C., 2005. International bioenergy transport costs and energy balance. *Biomass and Bioenergy*, 29(2), 114–134.
- Han, S.K., Murphy, G.E., 2012. Solving a woody biomass truck scheduling problem for a transport company in Western Oregon, USA. *Biomass and Bioenergy*, 44, 47–55.
- Haque, M., Epplin, F.M., 2012. Cost to produce switchgrass and cost to produce ethanol from switchgrass for several levels of biorefinery investment cost and biomass to ethanol conversion rates. *Biomass and Bioenergy*, 46, 517–530.
- Higgins, A.J., 2002. Australian Sugar Mills Optimize Harvester Rosters to Improve Production. *Interfaces*, 32(3), 15–25.
- Higuchi, T., Troutt, M.D., 2004. Dynamic simulation of the supply chain for a short life cycle product – Lessons from the Tamagotchi case. *Computers & Operations Research*, 31(7), 1097–1114.
- Hovelaque, V., Duvaleix-Tréguer, S., Cordier, J., 2009. Effects of constrained supply and price contracts on agricultural cooperatives. *European Journal of Operational Research*, 199(3), 769–780.
- Huang, Y., Chen, C.-W., Fan, Y., 2010. Multistage optimization of the supply chains of biofuels. *Transportation Research Part E: Logistics & Transportation Review*, 46(6), 820–830.
- Hugo, A., Rutter, P., Pistikopoulos, S., Amorelli, A., Zoia, G., 2005. Hydrogen infrastructure strategic planning using multi-objective optimization. *International Journal of Hydrogen Energy*, 30(15), 1523–1534.
- Huisman, W., Venturi, P., Molenaar, J., 1997. Costs of supply chains of *Miscanthus giganteus*. *Industrial Crops and Products*, 6(3-4), 353–366.
- Hung, W.Y., Kucherenko, S., Samsatli, N.J., Shah, N., 2004. A flexible and generic approach to dynamic modelling of supply chains. *Journal of the Operational Research Society*, 55(8), 801–813.
- Hutchins, M.J., Sutherland, J.W., 2008. An exploration of measures of social sustainability and their application to supply chains decisions. *Journal of Cleaner Production*, 16, 1688–1698.
- Iakovou, E., Karagiannidis, A., Vlachos, D., Toka, A., Malamakis, A., 2010. Waste biomass-to-energy supply chain management : A critical synthesis. *Waste Management*, 30(10), 1860–1870.
- IEA (International Energy Agency), 2007. Good practice guidelines: bioenergy project development and biomass supply. IEA-OECD, Paris, France.
- Jebaraj, S., Iniyar, S., 2006. A review of energy models. *Renewable and Sustainable Energy Reviews*, 10(4), 281–311.
- Jenkins, B.M., 1997. A comment on the optimal sizing of a biomass utilization facility under constant and variable cost scaling. *Biomass and Bioenergy*, (13), 1–9.
- Johnson, D.M., Jenkins, T.L., Zhang, F., 2012. Methods for optimally locating a forest biomass-to biofuel facility. *Biofuels*, 3(4), 489–503.
- Judd, J.D., Sarin, S.C., Cundiff, J.S., 2012. Design, modeling, and analysis of a feedstock logistics system. *Bioresour. Technol.*, 103 (1) , 209–218.
- Jung, J.Y., Blau, G., Pekny, J.F., Reklaitis, G.V., Eversdyk, D., 2004. A simulation based optimization approach to supply chain management under demand uncertainty. *Computers & Chemical Engineering*, 28(10), 2087–2106.

- Kamm, B., Gruber, P.R., Kamm, M., 2006. An overview: In *Biorefineries-Industrial Processes and Products*. WILEY-VCH, Weinheim, 1, 3-40.
- Kanzian, C., Holzleitner, F., Stampfer, K., Ashton, S., 2009. Regional Energy Wood Logistics – Optimizing Local Fuel Supply. *Silva Fennica* 43(1), 113–128.
- Karagiannidis, A., Wittmaier, M., Langer, S., Bilitewski, B., Malamakis, A., 2009. Thermal processing of waste organic substrates: Developing and applying an integrated framework for feasibility assessment in developing countries. *Renewable and Sustainable Energy Reviews*, 13(8), 2156–2162.
- Keirstead, J., Samsatli, N., Pantaleo, A.M., Shah, N., 2012. Evaluating biomass energy strategies for a UK eco-town with an MILP optimization model. *Biomass and Bioenergy*, 39, 306–316.
- Kerbache, L., Mac Gregor Smith, J., 2004. Queueing networks and the topological design of supply chain systems. *International Journal of Production Economics*, 91(3), 251–272.
- Kim, J., Realff, M.J., Lee, J.H., Whittaker, C., Furtner, L., 2011a. Design of biomass processing network for biofuel production using an MILP model. *Biomass and Bioenergy*, 35(2), 853–871.
- Kim, J., Realff, M.J., Lee, J.H., 2011b. Optimal design and global sensitivity analysis of biomass supply chain networks for biofuels under uncertainty. *Computers & Chemical Engineering*, 35(9), 1738–1751.
- Kinoshita, T., Inoue, K., Iwao, K., Kagemoto, H., Yamagata, Y., 2009. A spatial evaluation of forest biomass usage using GIS. *Applied Energy*, 86(1), 1–8.
- Kondili, E., Pantelides, C.C., Sargent, R.W.H., 1993. A general algorithm for short-term scheduling of batch operations – I. MILP formulations, *Computers and Chemical Engineering*, 17(2), 211-227.
- Kumar, A., Purohit, P., Rana, S., Kandpal, T. C., 2002. An approach to the estimation of the value of agricultural residues used as biofuels. *Biomass and Bioenergy*, 22, 195–203.
- Kumar, A., Cameron, J.B., Flynn, P.C., 2003. Biomass power cost and optimum plant size in Western Canada. *Biomass and Bioenergy*, 24(6), 445–464.
- Kumar, A., Sokhansanj, S., 2007. Switchgrass (*Panicum virgatum* L.) delivery to a biorefinery using integrated biomass supply analysis and logistics (IBSAL) model. *Bioresour. Technol.*, 98(5), 1033–1044.
- Lam, H.L., Klemes, J.J., Kravanja, Z., 2011. Model-size reduction techniques for large-scale biomass production and supply networks. *Energy*, 36(8), 4599-4608.
- Ligon, E., 2003. Optimal risk in agricultural contracts. *Agricultural Systems*, 75(2-3), 265–276.
- Lin, T., Rodríguez, L.F., Shastri, Y.N., Hansen, A.C., Ting, K.C., 2014. Integrated strategic and tactical biomass–biofuel supply chain optimization. *Bioresour. Technol.*, 156, 256-266.
- Lindemann, C. *Performance modeling with deterministic and stochastic Petri nets*. John Wiley and Sons, 1998.
- Linton, J., Klassen, R., Jayaraman, V., 2007. Sustainable supply chains: An introduction. *Journal of Operations Management*, 25(6), 1075–1082.
- Ma, J., Scott, N.R., DeGloria, S.D., Lembo, A. J., 2005. Siting analysis of farm-based centralized anaerobic digester systems for distributed generation using GIS. *Biomass and Bioenergy*, 28(6), 591–600.
- Machani, M., Nourelfath, M., D'Amours, S., 2013. A decision-support tool for evaluating the technical and economic potential of integrating bioenergy production within pulp and paper mills. Rapport technique CIRRELT-2013-07, CIRRELT, Montréal, Canada.

- Mafakheri, F., Nasiri, F., 2013. Modeling of biomass-to-energy supply chain operations: Applications, challenges and research directions. *Energy Policy*, 67, 116-126.
- Mahnam, M., Yadollahpour, M.R., Famil-Dardashti, V., Hejazi, S.R., 2009. Supply chain modeling in uncertain environment with bi-objective approach. *Computers & Industrial Engineering*, 56(4), 1535–1544.
- Mani, S., Tabil, L.G., Sokhansanj, S., 2006. Effects of compressive force, particle size and moisture content on mechanical properties of biomass pellets from grasses. *Biomass and Bioenergy*, 30(7), 648–654.
- Mapemba, L.D., Epplin, F.M., Taliaferro, C.M., Huhnke, R.L., 2007. Biorefinery feedstock production on Conservation Reserve Program land. *Review of Agricultural Economics*, 29(2), 227–46.
- Mapemba, L.D., Epplin, F.M., Huhnke, R.L., Taliaferro, C.M., 2008. Herbaceous plant biomass harvest and delivery cost with harvest segmented by month and number of harvest machines endogenously determined. *Biomass and Bioenergy* 32(11), 1016–1027.
- Marvin, W.A., Schmidt, L.D., Benjaafar, S., Tiffany, D.G., Daoutidis, P., 2012. Economic Optimization of a Lignocellulosic Biomass-to-Ethanol Supply Chain in the Midwest. *Chemical Engineering Science*, 67(1), 68–79.
- Marvin, W.A., Schmidt, L.D., Daoutidis, P., 2013. Biorefinery location and technology selection via supply chain optimization. *Industrial & Engineering Chemistry Research*, 52, 3192–3208.
- Mathews, J.A., 2008. Towards a sustainably certifiable futures contract for biofuels. *Energy Policy*, 36(5), 1577–1583.
- McCormick, K., Kaberger, T., 2007. Key barriers for bioenergy in Europe: Economic conditions, know-how and institutional capacity, and supply chain co-ordination. *Biomass and Bioenergy*, 31(7), 443–452.
- McGowin, C.R., Wiltsee, G.A., 1996. Strategic analysis of biomass and waste fuels for electric power generation. *Biomass and Bioenergy*, 10(95), 167–175.
- Miranda, P.A., Garrido, R.A., 2004. Incorporating inventory control decisions into a strategic distribution network design model with stochastic demand. *Transportation Research Part E: Logistics & Transportation Review*, 40(3), 183–207.
- MirHassani, S.A., Lucas, C., Mitra, G., Messina, E., Poojari, C.A., 2000. Computational solution of capacity planning models under uncertainties. *Parallel Computing*, 26, 511–538.
- Morrow, W.R., Griffin, W.M., Matthews, H.S., 2006. Policy Analysis Modeling Switchgrass Derived Cellulosic Ethanol Distribution in the United States. *Environmental Science & Technology*, 40, 2877–2886.
- Mula, J., Peidro, D., Díaz-Madroñero, M., Vicens, E., 2010. Mathematical programming models for supply chain production and transport planning. *European Journal of Operational Research*, 204(3), 377–390.
- Muttiah, R.S., Engel, B.A., Jones, D.D., 1996. Waste disposal site selection using GIS-based simulated annealing. *Computers & Geosciences*, 22(9), 1013–1017.
- Naik, S.N., Goud, V.V., Rout, P.K. Dalai, A.K., 2010. Production of first and second generation biofuels: A review. *Renewable and Sustainable Energy Reviews*, 14, 578–597.
- Nagel, J., 2000. Determination of an economic energy supply structure based on biomass using a mixed-integer linear optimization model. *Ecological Engineering*, 16, 91–102.
- Nguyen, M.H., Prince, R.G.H., 2006. A simple rule for bioenergy conversion plant size optimisation : bioethanol from sugar cane and sweet sorghum. *Biomass and Bioenergy*, 10(3), 361–365.

- Nikolopoulou, A., Ierapetritou, M., 2012. Optimal design of sustainable chemical processes and supply chains: A review. *Computers & Chemical Engineering*, 44, 94–103.
- Nilsson, D., 1999a. SHAM – A simulation model for designing straw fuel delivery systems. Part 1 : model description. *Biomass and Bioenergy*, 16, 25–38.
- Nilsson, D., 1999b. SHAM – A simulation model for designing straw fuel delivery systems. Part 2 : model applications. *Biomass and Bioenergy*, 16, 39–50.
- Nilsson, D., Hansson, P., 2001. Influence of various machinery combinations, fuel proportions and storage capacities on costs for co-handling of straw and reed canary grass to district heating plants. *Biomass and Bioenergy*, 20, 247–260.
- Nilsson, D., 2000. Dynamic Simulation of Straw Harvesting Systems: Influence of Climatic, Geographical and Biological Factors on Performance and Costs. *Journal of Agricultural Engineering Research*, 76(1), 27–36.
- Osmani, A., Zhang, J., 2013. Stochastic optimization of a multi-feedstock lignocellulosic-based bioethanol supply chain under multiple uncertainties. *Energy*, 59, 157–172.
- Paksoy, T., 2010. Optimizing a supply chain network with emission trading factor. *Scientific Research and Essays*, 5(17), 2535–2546.
- Panichelli, L., Gnansounou, E., 2008. GIS-based approach for defining bioenergy facilities location: A case study in Northern Spain based on marginal delivery costs and resources competition between facilities. *Biomass and Bioenergy*, 32(4), 289–300.
- Papadopoulos, D.P., Katsigiannis, P.A., 2002. Biomass energy surveying and techno-economic assessment of suitable CHP system installations. *Biomass and Bioenergy*, 22(2), 105–124.
- Papapostolou, C., Kondili, E., Kaldellis, J.K., 2011. Modelling biomass and biofuels supply chains. *Computer Aided Chemical Engineering*, 29, 1773–1777.
- Parker, N., Tittmann, P., Hart, Q., Nelson, R., Skog, K., Schmidt, A., Gray, E., Jenkins, B., 2010. Development of a biorefinery optimized biofuel supply curve for the Western United States. *Biomass and Bioenergy*, 34(11), pp.1597–1607.
- Peskett, L., Slater, L., Stevens, C., Dufey, A., 2007. Biofuels, agriculture and poverty reduction. Rapport de projet, Overseas Development Institute, Royaume Uni.
- PRé Consultants, 2000. The Eco-indicator 99, a damage oriented method for life cycle impact assessment. Methodology Report and Manual for Designers. Technical Report, PRé Consultants, Amersfoort, The Netherlands.
- Quariguasi Frota Neto, J., Bloemhof-Ruwaard, J.M., van Nunen, J.A.E.E., van Heck, E., 2008. Designing and evaluating sustainable logistics networks. *International Journal of Production Economics*, 111(2), 195–208.
- Quarigasi Frota Neto, J., Bloemhof-Ruwaard, J.M., van Nunen, J.A.E.E., Spengler, T., 2009. A methodology for assessing eco-efficiency in logistics networks. *European Journal of Operational Research*, 193(3), 670–682.
- Ramachandra, T.V., Krishna, S.V., Shruthi, B.V., 2005. Decision Support System to Assess Regional Biomass Energy Potential. *International Journal of Green Energy*, 1(4), 407–428.
- Rauch, P., Gronalt, M., 2011. The effects of rising energy costs and transportation mode mix on forest fuel procurement costs. *Biomass and Bioenergy*, 35(1), 690–699.
- Ravula, P.P., Grisso, R.D., Cundiff, J.S., 2008a. Comparison between two policy strategies for scheduling trucks in a biomass logistic system. *Bioresour. Technol.*, 99(13), 5710–5721.

- Ravula, P., Grisso, R., Cundiff, J.S., 2008b. Cotton logistics as a model for a biomass transportation system. *Biomass and Bioenergy*, 32(4), 314–325.
- Reche López, P., Jurado, F., Ruiz Reyes, N., García Galán, S., Gómez, M., 2008. Particle swarm optimization for biomass-fuelled systems with technical constraints. *Engineering Applications of Artificial Intelligence*, 21(8), 1389–1396.
- Recio, B., Rubio, F., Criado, J.A., 2003. A decision support system for farm planning using AgriSupport II. *Decision Support Systems*, 36(2), 189–203.
- Rentizelas, A.A., Tatsiopoulos, I.P., Tolis, A., 2009a. An optimization model for multi-biomass tri-generation energy supply. *Biomass and Bioenergy*, 33(2), 223–233.
- Rentizelas, A.A., Tolis, A., Tatsiopoulos, I.P., 2009b. Logistics issues of biomass: The storage problem and the multi-biomass supply chain. *Renewable & Sustainable Energy Reviews*, 13(4), 887–894.
- Rentizelas, A.A., Tolis, A., Tatsiopoulos, I.P., 2009c. Biomass District Energy Trigeneration Systems: Emissions Reduction and Financial Impact. *Water, Air & Soil Pollution: Focus*, 9(1-2), 139–150.
- Rosegrant, M.W., Msangi, S., Sulser, T., Valmonte-Santos, R., 2006. Biofuels and the global food balance; bioenergy and agriculture: promises and challenges. International Food Policy Research Institute (IFPRI), Focus 14. <http://www.ifpri.org/publication/biofuels-and-global-food-balance>
- Sabri, E.H., Beamon, B.M., 2000. A multi-objective approach to simultaneous strategic and operational planning in supply chain design. *Omega*, 28(5), 581–598.
- Santibañez-Aguilar, J.E., González-Campos, J.B., Ponce-Ortega, J.M., Serna-González, M., El-Halwagi, M.M., 2011. Optimal planning of a biomass conversion system considering economic and environmental aspects. *Industrial & Engineering Chemistry Research*, 50(14), 8558–8570.
- Santoso, T., Shabir, A., Goetschalckx, M., Shapiro, A., 2005. A stochastic programming approach for supply chain network design under uncertainty. *European Journal of Operational Research*, 167, 96–115.
- Samsatli, S., Samsatli, N.J., Shah, N., 2015. BVCM: A comprehensive and flexible toolkit for whole system biomass value chain analysis and optimisation – Mathematical formulation. *Applied Energy*, 147, 131–160.
- Schieb, P.A., Clément-Larosière, B., Lescieux-Katir, H., Thénot, M., 2014. *Bioraffinerie 2030. Une question d'avenir*. L'Harmattan, Paris.
- Schmidt, J., Leduc, S., Dotzauer, E., Kindermann, G., Schmid, E., 2009. Using Monte Carlo Simulation to Account for Uncertainties in the Spatial Explicit Modeling of Biomass Fired Combined Heat and Power Potentials in Austria. Rapport de recherche DP-43-2009, University of Natural Resources and Applied Life Sciences, Vienne, Autriche.
- Schmidt, J., Leduc, S., Dotzauer, E., Kindermann, G., Schmid, E., 2010. Cost-effective CO₂ emission reduction through heat, power and biofuel production from woody biomass: A spatially explicit comparison of conversion technologies. *Applied Energy*, 87, 2128–2141.
- Scott, J.A., Ho, W., Dey, P.K., 2012. A review of multi-criteria decision-making methods for bioenergy systems. *Energy*, 42(1), 146–156.
- Shabani, N., Akhtari, S., Sowlati, T., 2013. Value chain optimization of forest biomass for bioenergy production: A review. *Renewable and Sustainable Energy Reviews*, 23, 299–311.
- Shabani, N., Sowlati, T., 2013. A mixed integer non-linear programming model for tactical value chain optimization of a wood biomass power plant. *Applied Energy*, 104, 353–361.

- Shah, N., Pantelides, C.C., Sargent, R.W.H., 1993. A general algorithm for short-term scheduling of batch operations – II. Computational issues, *Computers and Chemical Engineering*, 17(2), 229–244.
- Sharma, B., Ingalls, R.G., Jones, C.L., Khanchi, A., 2013a. Biomass supply chain design and analysis: Basis, overview, modeling, challenges, and future. *Renewable and Sustainable Energy Reviews*, 24, 608–627
- Sharma, B., Ingalls, R.G., Jones, C.L., Huhnke, R.L., Khanchi, A., 2013b. Scenario optimization modeling approach for design and management of biomass-to-biorefinery supply chain system. *Bioresource Technology*, 150, 163–171.
- Shastri, Y., Hansen, A., Rodriguez, L., Ting, K.C., 2011a. Development and application of BioFeed model for optimization of herbaceous biomass feedstock production. *Biomass and Bioenergy*, 35, 2961–2974.
- Shastri, Y., Hansen, A., Rodriguez, L., Ting, K.C., 2011b. A novel decomposition and distributed computing approach for the solution of large scale optimization models, *Computers and Electronics in Agriculture*, 76, 69–79.
- Sheehan, J., Dunahay, T., Benemann, J., Roessler, P., 1998. A look back at the U.S. Department of Energy's aquatic species program: bio-diesel from algae. Rapport technique NREL/TP-580-24190, National Renewable Energy Laboratory.
- Shinners, K., Binversie, B., Muck, R., Weimer, P., 2007. Comparison of wet and dry corn stover harvest and storage. *Biomass and Bioenergy*, 31(4), 211–221.
- Simchi-Levi, D., Kaminky, P., Simchi-Levi, E., 2003. *Designing and managing the supply chain: Concepts, strategies and case studies*. McGraw-Hill, New York.
- Singh, J., Panesar, B., Sharma, S., 2007. Energy potential through agricultural biomass using geographical information system – A case study of Punjab. *Biomass and Bioenergy*, 32, 301–307.
- Skoulou, V., Zabaniotou, A., 2007. Investigation of agricultural and animal wastes in Greece and their allocation to potential application for energy production. *Renewable and Sustainable Energy Reviews*, 11(8), 1698–1719.
- Sodhi, M.S., Tang, C.S., 2009. Modeling supply-chain planning under demand uncertainty using stochastic programming: A survey motivated by asset–liability management. *International Journal of Production Economics*, 121(2), 728–738.
- Sokhansanj, S., Kumar, A., Turhollow, A., 2006. Development and implementation of integrated biomass supply analysis and logistics model (IBSAL). *Biomass and Bioenergy*, 30(10), 838–847.
- Sokhansanj, S., Mani, S., Turhollow, A., Kumar, A., Bransby, D., 2009. Large-scale production, harvest and logistics of switchgrass (*Panicum virgatum* L.) – current technology and envisioning a mature technology. *Biofuels, Bioproducts and Biorefining*, 3(2), 124–141.
- Stephen, J.D., Sokhansanj, S., Bi, X., Sowlati, S., Kloeck, T., Townley-Smith, L., Stumborg, M.A., 2010. The impact of agricultural residue yield range on the delivered cost to a biorefinery in the Peace River region of Alberta, Canada. *Biosystems Engineering*, 105(3), 298–305.
- Svanberg, M., Olofsson, I., Flodén, J., Nordin, A., 2013. Analysing biomass torrefaction supply chain costs. *Bioresource Technology*, 142, 287–296.
- Tan, R.R., Aviso, K.B., Barile, I.U., Culaba, A.B., Cruz Jr., J.B., 2012. A fuzzy multi-regional input–output optimization model for biomass production and trade under resource and footprint constraints. *Applied Energy*, 90(1), 154–160.

- Tatsiopoulos, I.P., Tolis, A.J., 2003. Economic aspects of the cotton-stalk biomass logistics and comparison of supply chain methods. *Biomass and Bioenergy*, 24, 199–214.
- Tembo, G., Epplin, F.M., Huhnke, R.L., 2003. Integrative Investment Appraisal of a Lignocellulosic Biomass-to-Ethanol Industry. *Journal of Agricultural and Resource Economics*, 28(3), 611–633.
- Thorsell, S., Epplin, F.L., Huhnke, R.L., Taliaferro, C.M., 2004. Economics of a coordinated biorefinery feedstock harvest system: Ligno-cellulosic biomass harvest cost. *Biomass and Bioenergy*, 27, 327–334.
- Todorov, V.I., Marinova, D., 2009. Models of Sustainability. *18th World IMACS / MODSIM Congress*, Cairns, Australia 13-17 July 2009 <<http://mssanz.org.au/modsim09>>.
- Van Dam, J.E.G., Klerk-Engels, B., Struik, P.C., Rabbinge, R., 2005. Securing renewable resource supplies for changing market demands in a bio-based economy. *Industrial Crops and Products*, 21(1), 129–144.
- Van Dyken, S., Bakken, B.H., Skjelbred, H.I., 2010. Linear mixed-integer models for biomass supply chains with transport, storage and processing. *Energy*, 35(3), 1338–1350.
- Vera, D., Carabias, J., Jurado, F., Ruiz-Reyes, N., 2010. A Honey Bee Foraging approach for optimal location of a biomass power plant. *Applied Energy*, 87(7), 2119–2127.
- Vlachos, D., Iakovou, E., Karagiannidis, A., Toka, A., 2008. A strategic supply chain management model for waste biomass networks. *3rd International Conference on Manufacturing Engineering (ICMEN)*, 1–3.
- Voivontas, D., Assimacopoulos, D., Koukios, E.G., 2001. Assessment of biomass potential for power production: a GIS based method. *Biomass and Bioenergy*, 20, 101–112.
- Wang, J.J., Jing, Y.Y., Zhang, C.F., Zhao, J.H., 2009. Review on multi-criteria decision analysis aid in sustainable energy decision-making. *Renewable and Sustainable Energy Reviews*, 13(9), 2263–2278.
- Wetterlund, E., Söderström, M., 2010. Biomass gasification in district heating systems – The effect of economic energy policies. *Applied Energy*, 87(9), 2914–2922.
- Wolfsmayr, U.J., Rauch, P., 2013. The primary forest fuel supply chain: A literature review. *Biomass and Bioenergy*, 60, 203–221.
- You, F., Tao, L., Graziano, D.J., Snyder, S.W., 2012. Optimal design of sustainable cellulosic biofuel supply chains : multiobjective optimization coupled with life cycle assessment and input-output analysis. *AIChE Journal*, 58(4), 1157–1179.
- You, F., Wang, B., 2011. Life cycle optimization of biomass-to-liquid supply chains with distributed-centralized processing networks. *Industrial and Engineering Chemistry Research*, 50, 10102–10127, 2011.
- Yu, Y., Bartle, J., Li, C.Z., Wu, H., 2009. Mallee biomass as a key bioenergy source in Western Australia: Importance of biomass supply chain. *Energy & Fuels*, 23, 3290–3299.
- Yue, D., You, F., Snyder, S.W., 2013. Biomass-to-biofuels supply chain optimization: Overview, key issues and challenges. *Computers & Chemical Engineering*, disponible en ligne, DOI: 10.1016/j.compchemeng.2013.11.016.
- Zamboni, A., Bezzo, F., Shah, N., 2009a. Spatially Explicit Static Model for the Strategic Design of Future Bioethanol Production Systems. 1: Cost optimization. *Energy & Fuels*, 23(10), 5121–5133.
- Zamboni, A., Bezzo, F., Shah, N., 2009b. Spatially Explicit Static Model for the Strategic Design of Future Bioethanol Production Systems. 2: Multi-Objective Environmental Optimization. *Energy & Fuels*, 23(10), 5134–5143.

- Zamboni, A., Murphy, R.J., Woods, J., Bezzo, F., Shah, N., 2011. Biofuels carbon footprints: Whole-systems optimisation for GHG emissions reduction. *Bioresource technology*, 102(16), 7457–7465.
- Zhan, F., Chen, X., Noon, C., Wu, G., 2005. A GIS-enabled comparison of fixed and discriminatory pricing strategies for potential switchgrass-to-ethanol conversion facilities in Alabama. *Biomass and Bioenergy*, 28(3), 295–306.
- Zhang, L., Hu, G., 2013. Supply chain design and operational planning models for biomass to drop-in fuel production. *Biomass and Bioenergy*, 58, 238–250.
- Zhang, F., Johnson, D.M., Johnson, M.A., 2012. Development of a simulation model of biomass supply chain for biofuel production. *Renewable Energy*, 44, 380–391.
- Zhang, J., Osmani, A., Awudu, I., Gonela, V., 2013. An integrated optimization model for switchgrass-based bioethanol supply chain. *Applied Energy* 102, 1205–1217.
- Zhou, Z., Cheng, S., Hua, B., 2000. Supply chain optimization of continuous process industries with sustainability considerations. *Computers & Chemical Engineering* 24, 1151–1158.
- Zhu, X., Li, X., Yao, Q., Chen, Y., 2011. Challenges and models in supporting logistics system design for dedicated-biomass-based bioenergy industry. *Bioresource technology*, 102(2), 1344–1351.
- Zhu, X., Yao, Q., 2011. Logistics system design for biomass-to-bioenergy industry with multiple types of feedstocks. *Bioresource technology*, 102(23), 10936–10945.

Birome Holo BA

Doctorat : Optimisation et Sûreté des systèmes

Année 2016

Modélisation et optimisation de chaînes d'approvisionnement en biomasses pour des bioraffineries

Les travaux de cette thèse concernent la modélisation et l'optimisation de chaînes d'approvisionnement en biomasses pour de futures bioraffineries. En effet, des chaînes d'approvisionnement efficaces sont essentielles pour fournir aux installations de conversion, de façon régulière, des quantités suffisantes de biomasse de qualité à des prix raisonnables. Le problème est tout d'abord décrit puis modélisé.

Un modèle de réseau et un modèle de données sont ensuite développés pour permettre de décrire la structure de la chaîne d'approvisionnement et ses données, sans affecter le modèle mathématique sous-jacent. Ce dernier (MILP) combine pour la première fois divers aspects, soit originaux, soit gérés séparément dans la littérature. A partir des demandes de la raffinerie, une résolution exacte précise les activités logistiques dans le réseau et les équipements nécessaires, afin de minimiser le coût total composé des coûts de récoltes, de transport et de stockage. Des études de cas sont décrites pour illustrer ce modèle de planification tactique multi-biomasse et multi-période.

Un modèle plus compact est aussi élaboré pour traiter des instances de très grandes tailles. Il est illustré par une étude de cas réelle pour une bioraffinerie prévue près de Compiègne.

Pour finir, les développements effectués pour la mise en place d'un prototype logiciel d'aide à la décision sont présentés et des recommandations d'un futur logiciel commercial sont proposées.

Mots clés : recherche opérationnelle - optimisation combinatoire - logistique (organisation) - planification - théorie des graphes - biomasse - énergie de la biomasse - installations énergétiques - développement durable.

Modelling and Optimization of Biomass Supply Chains for Biorefineries

The research works of this thesis address the problem of modeling and optimizing biomass supply chains for biorefineries. Indeed, efficient supply chains are essential to provide conversion facilities with sufficient quantities of quality biomass at reasonable prices. The problem is described and modeled.

A network model and a data model are developed to allow to describe the structure of the supply chain and its data, without affecting the underlying mathematical model. The latter is a mixed-integer linear programming that combines for the first time various aspects, either original or tackled separately in the literature. For given refinery needs, its exact resolution by CPLEX specifies the logistic activities in the network (amounts harvested, baled, transported, stored etc.) and the necessary equipment, in order to minimize a total cost including harvesting costs, transport costs and storage costs. Case studies are described to illustrate this multi-biomass and multi-period tactical planning model.

A more compact model is also elaborated to cope with large-scale instances. It is illustrated using a real case study for a bio-refinery planned near Compiègne, France.

Finally, the developments conducted for the implementation of a prototype of decision-support application are presented and recommendations for coming to a commercial software are proposed.

Keywords: operations research - combinatorial optimization - business logistics - planning - graph theory - biomass - biomass energy - energy facilities - sustainable development.

Thèse réalisée en partenariat entre :

