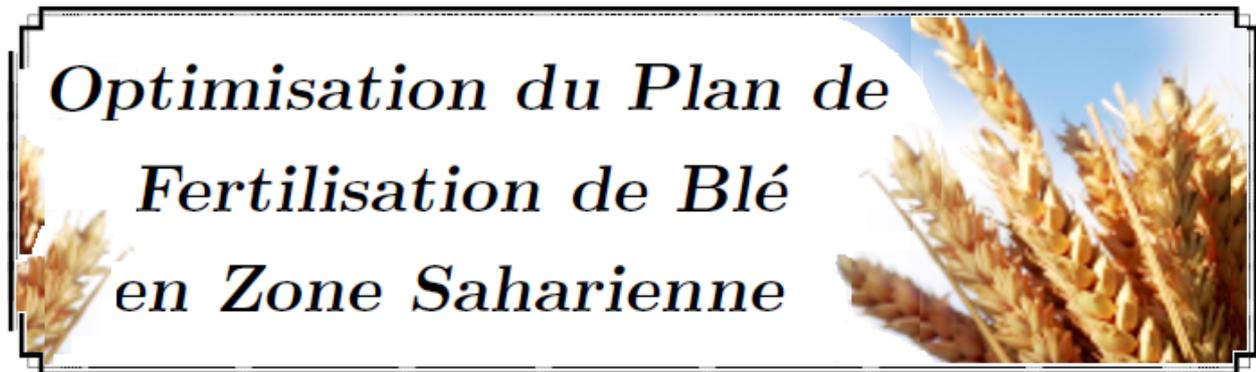


République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université Abderrahmane MIRA - Béjaïa
Faculté des Sciences Exactes
Département de Recherche Opérationnelle

Mémoire de Master en Recherche Opérationnelle
Option Modélisation Mathématique et Techniques de Décision



Présenté par : RACHEDI Karima et TATAH Assia

Soutenu le 25 Juin 2012 devant le jury composé de :

Présidente :	M ^{me} AOUDIA F.	Maître de Conf.	U. A. MIRA - Béjaïa
Rapporteur :	P ^r AISSANI D.	Professeur	U. A. MIRA - Béjaïa
Examineur :	M ^{lle} SAIT R.	Maître Assistant	U. A. MIRA - Béjaïa
Examineur :	M ^r ZOUGAB N.	Maître Assistant	U. M. M. - Tizi Ouzou

Université de Béjaïa - Juin 2012

Remerciements

A travers ce modeste travail, nous tenons à remercier tout particulièrement :

Notre promoteur, Professeur AISSANI Djamil, pour son encadrement et ses précieux conseils,

Les membres de la Commission de Jury, qui ont accepté d'évaluer notre travail, ainsi que l'ensemble des Enseignants du Département de Recherche Opérationnelle, qui ont contribué à notre formation,

Nos remerciements vont également à nos parents et proches, pour leurs encouragements, leur patience et leur présence tout au long de notre cursus.

Sans oublier nos camarades et amis, pour le soutien et l'aide qu'ils nous ont apportés.

Table des matières

Introduction	8
1 La Céréaliculture	11
Introduction	11
1.1 Notions d’Agronomie	11
1.1.1 Les besoins du blé	11
1.1.2 La fertilisation	14
1.1.3 Rendement annuel moyen du blé	14
1.1.4 L’irrigation sous pivot	15
1.2 Développement de la fertilisation en Algérie	16
1.3 La situation du blé en Algérie : ”quelques chiffres”	17
1.4 Extension vers les terres du sud	20
1.4.1 La céréaliculture sous pivot en zones sahariennes	20
1.4.2 Politique agricole	21
1.5 Localisation de la région de l’étude	21
1.6 Position du problème	22
Conclusion	24
2 Outils Statistiques	25
Introduction	25
2.1 Modélisation statistique	25
2.1.1 Un peu de vocabulaire	26
2.2 Dispositifs d’essai expérimental	26
2.2.1 Dispositif en blocs aléatoires complets	27
2.3 L’analyse de la variance	27
2.3.1 Définition	27
2.3.2 Modèles d’analyse de la variance ”ANOVA”	28

2.4	ANOVA à 2 facteurs à effet fixe	28
2.4.1	Modélisation	28
2.4.2	Hypothèses fondamentales	29
2.4.3	Décomposition de la variation totale	30
2.4.4	Tableau de l'ANOVA	31
2.4.5	Tests d'hypothèses	32
2.5	Tests de comparaisons multiples	33
2.5.1	Test de Newman-Keuls	34
	Conclusion	35
3	Optimisation Combinatoire	36
	Introduction	36
3.1	Problèmes d'optimisation combinatoire	36
3.1.1	Formulation d'un problème d'optimisation combinatoire	37
3.1.2	Méthodes de résolution d'un problème d'optimisation combinatoire	37
3.2	Méthodes de résolution heuristiques	38
3.2.1	Approches par voisinage	38
3.2.2	Approches constructives	39
3.3	Métaheuristiques	39
3.3.1	Source d'efficacité des métaheuristiques	40
3.3.2	Place des métaheuristiques dans une classification des méthodes d'optimisation	41
3.3.3	Méthodes hybrides	41
3.4	Méthode GRASP	42
3.4.1	Algorithmes gloutons et gloutons aléatoires	42
3.4.2	Recherche locale	43
3.4.3	Fonctionnement de la méthode GRASP	44
3.5	GRASP pour un PLNE	45
3.5.1	Représentation des solutions	46
3.5.2	Evaluation des solutions	46
3.5.3	Comparaison des solutions	46
3.5.4	Phase de construction	47
3.5.5	Recherche locale	49
	Conclusion	50

4	Analyse Statistique des Données	51
	Introduction	51
4.1	Dispositif expérimental et description des données	51
4.1.1	Outils d'analyse	52
4.1.2	Analyse graphique de l'interaction des deux facteurs	53
4.2	ANOVA à 2 facteurs avec interaction	54
4.2.1	ANOVA sous R	54
4.2.2	Vérification des conditions d'application	55
4.2.3	Synthèse	58
4.3	ANOVA à 2 facteurs sans interaction	58
4.3.1	ANOVA sous R	58
4.3.2	Vérification des conditions d'application	59
4.3.3	Confirmation d'absence d'effet du facteur DOSE : "test de Kruskal-Wallis"	60
4.3.4	Conclusions partielles	60
4.3.5	Test de Student-Newman-Keuls	61
	Conclusion	62
5	Etude des coûts de fertilisation	63
	Introduction	63
5.1	Modélisation du problème de coûts	63
5.2	Estimation des coûts par la métaheuristique GRASP	64
5.2.1	Solution initiale	65
5.2.2	Implémentation de GRASP	67
5.3	Discussion des résultats	70
5.3.1	Méthodes de construction	70
5.3.2	Solutions	70
5.3.3	Coûts de fertilisation pour les solutions trouvées	71
5.4	Application sur les résultats statistiques	72
5.4.1	FOSFACTYL	72
5.4.2	PHOSFERT	73
5.4.3	Comparaisons	73
	Conclusion	74
	Conclusion et Perspectives	75

TABLE DES MATIÈRES

5

Bibliographie

77

Annexe

80

Table des figures

1.1	Epis de blé	12
1.2	Vue aérienne de pivots	15
1.3	Pivot central (El Goléa)	16
1.4	Rendements de blé des pays nord africains (Unité : quintaux/ha)	18
1.5	Superficies réservées à la culture du blé entre 2002 et 2010 (Unité : hectares)	18
1.6	Rendements de blé de certains pays (Unité : quintaux/ha)	19
1.7	Importations de l'Algérie en 2009	20
1.8	Irrigation sous pivot en zone saharienne	21
1.9	Localisation de l'exploitation	22
3.1	Exemple d'allure de la fonction objectif d'un problème d'optimisation	40
3.2	Place des métaheuristiques dans une classification des méthodes d'optimisation	41
3.3	Fonction de densité bi-triangulaire	48
4.1	Graphe d'interaction des facteurs MARQUE et DOSE	54
5.1	Solution initiale pour la fertilisation d'un hectare	66
5.2	Solution initiale pour un pivot de 35 hectares	66
5.3	Résultats renvoyés par le solveur LP-Solve	67
5.4	Solution finale en cas de fertilisation d'un hectare de terre	68
5.5	Solution finale en cas fertilisation de 35 hectares	68
5.6	Solution finale en cas de fertilisation d'un hectare	69
5.7	Solution finale en cas fertilisation de 35 hectares	69
5.8	Liste complémentaire à FOSFACTYL	72
5.9	Liste complémentaire à PHOSFERT	73
10	Engrais solides PROFERT	80
11	Engrais solides TIMAC	81

12	Engrais solides FERTIAL	81
----	-----------------------------------	----

Introduction

Les céréales et leurs dérivées constituent l'alimentation de base dans beaucoup de pays en voie de développement, particulièrement dans les pays maghrébins. En 2008, l'Algérie a importé 6,48 millions de tonnes de céréales. Ce taux a commencé à baisser après la récolte record enregistrée durant la campagne 2008-2009 s'établissant à 6,12 millions de tonnes. Les achats de l'Algérie en céréales ont également baissé en 2010 avant de reprendre leur hausse durant les neuf premiers mois de 2011, une hausse de plus de 100%, incluant les semoules et les farines. Les importations de blé dur et tendre se sont établies à 5,7 millions de tonnes, dont 4,4 millions de tonnes pour le blé tendre, détaille un document du Centre National de l'Informatique et des Statistiques des Douanes (CNIS).

Les premières estimations de la filière céréalière du début de la campagne en cours font ressortir un rendement national moyen de 18 quintaux/ha avec des pointes de rendement dépassant les 85 quintaux/ha, mais ce chiffre reste faible en comparaison avec les rendements moyens de certains pays allant jusqu'à près de 100 quintaux/ha. Selon le président du Comité Interprofessionnel des Céréales (CIC), l'objectif céréalier sur les dix prochaines années, sera d'atteindre les 30 quintaux/ha pour le blé dur et 40 quintaux/ha pour le blé tendre.

Le directeur général de l'Office Algérien Interprofessionnel des Céréales (OAIC) a indiqué que la production céréalière en Algérie ne cessait de s'améliorer en quantité et en qualité enregistrant cette année un deuxième record après celui de 2009, puisqu'une récolte de 5,6 à 5,8 millions de tonnes de blé est attendue. Une production qui couvrira les besoins du pays jusqu'au delà de 2012. L'Algérie ne sera donc plus sur le marché international du blé dur et de l'orge jusqu'à la fin de 2012, mais continuera à importer le blé tendre, a affirmé le même responsable.

Ne plus négliger le potentiel des terres du Sahara algérien serait une solution au renforcement de la production céréalière nationale. En effet, un vaste programme de mise en valeur des terres sahariennes pour l'intensification céréalière en irrigation sous pivot est mené. Grâce à des ressources en eau exceptionnelles, mais non renouvelables, cette politique - quoique controversée - a permis la création de périmètres irrigués pour la production intensive de céréales.

Cependant, les terres du Sud algérien sont généralement pauvres en éléments nutritifs. La fertilisation minérale reste à l'heure actuelle le moyen le plus efficace pour l'obtention de rendements acceptables. De ce fait, il est impératif de raisonner convenablement les apports d'engrais minéraux pour améliorer les rendements et les processus physiologiques de la culture [13]. La question de fertilisation des terres du Sud est donc cruciale. L'agriculteur se pose alors la question : "Quel fertilisant utiliser et à quelles quantités ?"

Dans l'optique de répondre aux préoccupations des décideurs, à savoir l'accroissement des rendements et la minimisation des coûts, des outils statistiques et d'optimisation combinatoire seront utilisés dans notre travail dont la démarche s'inscrit dans cet ordre :

Le premier chapitre est associé à la présentation de quelques chiffres relatifs à la céréaliculture en Algérie. A la fin de ce chapitre nous dégageons notre problématique qui est perçue comme un problème dont l'objectif est de minimiser les coûts relativement aux besoins de l'agriculteur en éléments nutritifs (azote, phosphore, potassium, soufre) et aux catalogues d'engrais solides de trois entreprises leaders sur le marché des fertilisants en Algérie. En particulier, nous souhaitons répondre aux décideurs quant au choix à faire entre deux produits testés dans le cadre d'une expérimentation relativement aux rendements qu'ils offrent et à leur coût d'acquisition.

Le deuxième chapitre est consacré à la présentation des outils statistiques (modélisation et tests) qui seront utilisés pour l'analyse des données. Nous y présentons également quelques notions relatives à l'expérimentation.

Le troisième chapitre présente quelques notions théoriques en rapport avec l'optimisation combinatoire, en particulier les métaheuristiques qui serviront de support pour répondre à la problématique de minimisation du coût de fertilisation.

Le quatrième chapitre est dédié à la discussion des résultats des tests statistiques effectués. Nous tâcherons de présenter des éléments de réponse à la question de la dose et de l'engrais phosphaté, offrant un rendement maximal, sur la base de données expérimentales.

Le cinquième chapitre est consacré à l'application de la métaheuristique choisie, afin de déterminer d'une façon générale, le choix optimal de produits fertilisants à utiliser, de façon à minimiser les coûts à partir des besoins de l'agriculteur en termes d'azote, de phosphate, de potassium et de soufre. Une comparaison des résultats obtenus grâce à la métaheuristique sera faite avec ceux obtenus à l'aide du solveur "LP-Solve" disponible en ligne, basé sur la méthode du simplexe révisé et la méthode Branch and Bound, dans le but de mesurer l'efficacité de la métaheuristique choisie.

Enfin, nous terminons par une conclusion générale, dans laquelle nous synthétisons nos résultats et faisons le lien entre les deux parties d'application et ce, en proposant aux décideurs, au vu des données expérimentales et des catalogues des trois entreprises, un choix optimal de plan de fertilisation offrant un coût minimum et éventuellement un rendement maximum.

1

La Céréaliculture

Introduction

Dans ce chapitre, nous présentons quelques notions d'agronomie, qui faciliteront la compréhension de notre travail, ainsi qu'un bref aperçu du développement de la fertilisation phosphatée en Algérie, suivis de quelques chiffres relatifs à la situation du blé en Algérie et de la céréaliculture en zones sahariennes. La dernière section de ce chapitre est consacrée à la présentation de notre problématique relativement à la situation globale décrite au préalable.

1.1 Notions d'Agronomie

1.1.1 Les besoins du blé

Le rôle de l'eau

L'eau est le constituant essentiel des végétaux, elle représente 70 à 80% de leur poids frais. Elle véhicule les substances, elle maintient l'hydratation et intervient de façon directe dans les réactions métaboliques (respiration et photosynthèse). Une teneur suffisante en



FIGURE 1.1 – Epis de blé

eau et un facteur essentiel à la survie, à la croissance et au développement du végétal [12]. Une irrigation doit être correctement maîtrisée. Un sol sec ou trop humide est néfaste à la germination.

La température

Du semis à la récolte, le blé requiert des températures différentes suivant ses différentes phases physiologiques.

La lumière

L'intensité lumineuse agit directement sur l'intensité de la photosynthèse, dont dépend à la fois la résistance des tiges et le rendement [31].

Le sol

Selon SOLTNER D. [31], trois caractéristiques font une bonne terre à blé :

- Une texture fine limono-argileuse ;
- Une structure stable ;
- Une profondeur du sol.

Selon l'auteur, les moins bonnes terres à la culture du blé sont :

- Les terres très argileuses, moins drainées ;
- Les terres trop caillouteuses ;
- Les terres trop sableuses.

Les éléments fertilisants

Le blé dur est une espèce peu exigeante vis-à-vis de la fertilisation potassique mais moyennement exigeante vis-à-vis de la fertilisation phosphatée [19].

- *L'azote "N"*

L'objectif de la fertilisation azotée est d'aboutir à un rendement optimum et à une teneur en protéines de l'ordre de 14% [7].

- *Le phosphore "P"*

Le phosphore intervient partout où il y a multiplication cellulaire, d'où l'importance du phosphore dans tous les phénomènes de croissance et de reproduction.

- *Le potassium "K"*

Le potassium permet une meilleure utilisation de l'eau, favorise la tolérance à la chaleur et améliore la résistance aux maladies et aux bactéries.

- *Le soufre "S"*

Après N, P, K, le soufre est le quatrième élément fertilisant le plus prélevé par les céréales. La cinétique d'absorption du soufre au cours du cycle de développement est analogue à celle de l'azote.

- *Les oligoéléments*

Nous citons : le fer, le manganèse, le zinc, le cuivre, le bore, le molybdène et le soufre [19].

- Le fer est indispensable à la formation de la chlorophylle ;
- Le manganèse est un activateur des enzymes d'oxydoréduction et de photosynthèse ;
- Le zinc a un rôle important dans la synthèse des acides nucléiques et des protéines ;
- Le cuivre agit sur la viabilité du pollen.

Accidents et ennemis

Les parasites, les maladies des feuilles et de l'épi, les pucerons, la verse, sont les ennemis du blé. Cependant, certains de ces problèmes peuvent être atténués par l'utilisation de variétés résistantes.

1.1.2 La fertilisation

La fertilisation est le processus consistant à apporter à un milieu de culture, les éléments minéraux (les engrais ou les amendements) nécessaires au développement de la plante [32]. Les objectifs finaux de la fertilisation sont d'obtenir le meilleur rendement possible compte tenu des autres facteurs qui y concourent (qualité du sol, climat, apports en eau, potentiel génétique des cultures, moyens d'exploitation), ainsi que la meilleure qualité et ce, à moindre coût.

Plan de fertilisation

En Agronomie, l'établissement d'un plan de fertilisation : fertilisation azotée, phosphatée ou potassique, se fonde sur quatre critères principaux :

- L'exigence de l'espèce et de la variété cultivée ;
- L'analyse du sol ;
- Le passé récent de fertilisation ;
- Le précédent cultural : la restitution ou non des résidus de cultures du précédent.

1.1.3 Rendement annuel moyen du blé

Il s'agit de la quantité moyenne de blé récoltée annuellement dans un pays et ramenée à la superficie cultivée. Cet indicateur est une mesure de l'efficacité agricole d'un pays, mais il n'est pas le seul. Le rendement d'autres productions végétales, comme les fruits et légumes, sont aussi des indicateurs.

L'augmentation des rendements ne se fait pas sans accroître l'utilisation d'engrais, la consommation d'eau et l'utilisation de pesticides. Ce qui peut avoir pour conséquences : une surexploitation des sols, la pollution des nappes d'eau,...

1.1.4 L'irrigation sous pivot

La technologie de l'irrigation mécanisée a littéralement changé la face de la terre. Les photos aériennes d'il y a cinquante ans ne montraient pas les champs comme ils sont aujourd'hui. Aujourd'hui, des régions où les agriculteurs n'auraient jamais pensé pouvoir faire pousser des cultures sont à présent aussi productives que n'importe quel autre endroit.



FIGURE 1.2 – Vue aérienne de pivots

L'irrigation sous pivot central est une méthode d'irrigation de cultures par aspersion, où les buses (asperseurs) tournent autour d'un pivot. Une surface circulaire centrée sur le pivot est alors irriguée, créant un motif circulaire. C'est une forme d'irrigation aérienne, reproduisant une pluie artificielle.

L'irrigation adaptée au Sahara est le pivot car elle n'exige pas de main d'œuvre, elle est automatisée par la régularité de la pression.



FIGURE 1.3 – Pivot central (El Goléa)

1.2 Développement de la fertilisation en Algérie

Durant de nombreuses années, la fertilisation est restée archaïque en Algérie. Peu de produits étaient utilisés. Peu d'analyses des sols étaient faites. Les rares exceptions ont été les exploitations d'agriculteurs leaders dans le nord et dans la région d'Ain Oussera, là où les agriculteurs étaient confrontés à de nombreux problèmes de nutrition, générant de nombreuses carences.

Le peu de produits disponibles, sur un marché monopolisé pendant de nombreuses années par une société nationale, ne correspondaient pas aux besoins des sols et des cultures. Les formes d'éléments (l'azote et le phosphore, par exemple), étaient soit facilement lessivables, soit indisponibles car non assimilables par des sols qui ont des particularités spécifiques en Algérie.

Caractéristiques des sols algériens :

- La plupart des sols algériens sont calcaires avec un pH supérieur à 8. Le calcaire fixe le phosphore, le potassium et les oligo-éléments les rendant indisponibles même si on les apporte ou existent dans le sol.
- Les sols algériens ont un taux de matière organique très faible, ce qui induit une faible activité biologique.
- Dans certaines régions (Relizane, Biskra, M'sila,...), le taux de salinité est élevé, le so-

dium concurrence les autres éléments, les empêchant de se fixer et empêchant le développement des cultures.

La culture des céréales est longtemps restée le parent pauvre de la fertilisation. Les engrais qui étaient proposés étaient principalement du N-P-K équilibrés en concentrations tels que le 15%-15%-15%. L'équilibre de ces formules ne correspond pas aux besoins des cultures. Les éléments qui composent ces engrais (Phosphate peu soluble et Azote nitrique) ne sont pas adaptés aux besoins des sols et climats algériens. Certains produits, aggravent lourdement la salinité.

L'Algérie est également l'un des rares pays à encore utiliser le TSP : triple super phosphate en nutrition phosphatée (46 % en P_2O_5). Cette forme est quasi-totalement abandonnée dans tous les pays du monde, à l'exception du Brésil et de certaines régions du nord de l'Europe dont les sols sont acides, ce qui permet de rendre assimilable toutes formes de phosphate. En plus de contenir un phosphore peu soluble, le TSP apporte beaucoup de calcium à nos sols qui en sont fortement pourvus contribuant ainsi à accentuer l'alcalinité, les rendant plus basiques.

Grâce à l'extension du marché des fertilisants (présence de production nationale et d'importation des engrais), d'autres engrais phosphatés remplacent le SSP : simple super phosphate (16 % en P_2O_5) et le TSP dans la nutrition phosphatée des céréales [23, 24].

1.3 La situation du blé en Algérie : "quelques chiffres"

Sur la base de données émanant de la **FAO** : "Food and Agriculture Organization of the United Nations", nous présentons dans cette partie quelques chiffres relatifs à la production nationale de blé, les importations, le rendement du blé en Algérie et celui de certains pays, ainsi qu'à l'emblavure nationale.

	2002	2003	2004	2005	2006	2007	2008	2009	2010
Algérie	10,74	14,48	13,58	15,06	15,07	12,74	11,04	15,98	16,32
Égypte	64,35	65,00	65,57	64,93	64,30	64,67	65,03	63,83	55,74
Libye	7,56	6,33	6,31	7,58	7,88	7,88	7,88	7,90	7,85
Maroc	12,79	17,22	18,08	10,26	20,37	6,15	13,19	21,40	17,10
Soudan	21,39	19,65	24,17	24,62	38,32	28,26	19,46	16,04	17,94
Tunisie	14,14	21,41	16,63	16,92	14,94	18,76	17,17	21,04	18,93

FIGURE 1.4 – Rendements de blé des pays nord africains (Unité : quintaux/ha)

	2002	2003	2004	2005	2006	2007	2008	2009	2010
Maroc	2625400	2988900	3063700	2965800	3106700	2571900	2858100	2977600	2852400
Algérie	1398460	2047570	2010600	1603740	1783830	1819880	1006570	1848580	1900000
Égypte	1029590	1053020	1094740	1253820	1286750	1140980	1226650	1335300	1287630
Tunisie	298500	926600	1035200	961500	837000	769000	535100	786000	434400
Soudan	115500	169000	180000	169000	174583	284167	301680	400000	224700
Libye	165294	197624	198197	165000	132000	132000	132000	133000	135000

FIGURE 1.5 – Superficies réservées à la culture du blé entre 2002 et 2010 (Unité : hectares)

La figure 1.4 présente les rendements du blé (tendre et dur) enregistrés au niveau national et ceux des pays voisins nord africains. La figure 1.5 présente les superficies réservées à la culture du blé entre 2002 et 2010, de l'Algérie et des pays voisins.

Nous constatons que, bien que la surface consacrée à la culture du blé durant la période citée, soit relativement élevée par rapport à nos voisins, le rendement moyen ne dépasse pas les 17 quintaux/ha, ce qui positionne l'Algérie en avant dernière position, loin derrière l'Égypte qui atteint des rendements moyens allant jusqu'à 65,57 quintaux/ha.

La figure 1.6 quant à elle, présente quelques chiffres relatifs aux rendements de certains pays. Des rendements moyens pouvant atteindre les 99 quintaux/ha, contre seulement un maximum de 16,32 quintaux/ha pour l'Algérie.

	2002	2003	2004	2005	2006	2007	2008	2009	2010
Pays-Bas	78,15	87,44	89,15	85,93	83,94	70,71	87,30	92,91	89,09
Belgique	82,76	85,79	89,81	82,73	82,55	78,91	87,59	94,65	88,27
Irlande	84,44	82,98	99,24	84,32	91,54	84,63	89,70	81,66	85,99
Royaume-Uni	80,03	77,78	77,75	79,61	80,37	72,25	82,81	70,66	76,81
Allemagne	69,06	64,99	81,72	74,65	72,01	69,61	80,87	78,09	73,10
France	74,45	62,50	75,79	69,89	67,41	62,54	71,02	74,47	70,42
Danemark	70,35	70,77	71,41	72,33	69,96	65,61	78,64	80,37	66,26
Namibie	63,72	62,51	41,62	45,26	59,31	51,34	61,42	67,21	65,79
Zambie	62,50	64,29	61,18	62,10	54,95	61,51	99,39	56,99	63,12
Égypte	64,35	65,00	65,57	64,93	64,30	64,67	65,03	63,83	55,74
Mexique	51,00	44,91	44,87	47,52	52,27	50,82	50,13	49,69	54,19
Chine	37,77	39,32	42,52	42,75	45,94	46,08	47,62	47,39	47,49
USA	23,57	29,71	29,03	28,23	25,99	27,05	30,18	29,90	31,18
Espagne	28,35	29,24	32,63	17,71	28,76	35,69	32,48	26,72	29,42
Inde	27,62	26,10	27,13	26,02	26,19	27,08	28,02	29,07	28,30
Canada	18,32	22,57	26,41	27,38	26,10	23,22	28,52	27,86	28,02
Japon	40,01	40,34	40,47	40,97	38,35	43,40	42,20	32,37	27,61
Madagascar	23,18	25,96	27,75	29,66	20,65	24,44	23,81	24,23	25,00
Algérie	10,74	14,48	13,58	15,06	15,07	12,74	11,04	15,98	16,32
Nigéria	9,82	10,36	10,69	11,00	11,27	16,92	16,56	10,52	9,19

FIGURE 1.6 – Rendements de blé de certains pays (Unité : quintaux/ha)

L'Algérie figure parmi les cinq premiers importateurs de blé à l'échelle mondiale. En 2009, l'Algérie a été le premier importateur de blé au monde avec plus de 5,7 millions de tonnes de blé importé.

Position	Produit	Quantité (tonnes)	Valeur (1000\$)
1	Blé	5719730	1830350
2	Maïs	1994810	407494
3	Sucre Brut Centrifugé	811724	337562
4	Tourteaux de Soja	716675	306080
5	Sucre Raffiné	404264	205000
6	Huile de soja	391096	315492
7	Lait Entier Sec	200175	578281
8	Bananes	179578	86848
9	Huile de tournesol	151230	124064
10	Pommes de terre	124919	86960
11	Café vert	122775	230567
12	Pommes	118388	68524
13	Orge	105336	28507
14	Lait écrémé en poudre	92628	220926
15	Huiles et graisses	71624	69884
16	Riz Blanc	66331	39501
17	Préparations alimentaires	63902	193397
18	Viande de bovins, désossée	61339	168615
19	Lentilles	60555	51665
20	Pois chiches	53575	49888

FIGURE 1.7 – Importations de l'Algérie en 2009

1.4 Extension vers les terres du sud

1.4.1 La céréaliculture sous pivot en zones sahariennes

La culture céréalière en générale et le blé en particulier dans les zones sahariennes se pratiquent dans deux systèmes de gestion des exploitations agricoles :

- Le système classique : représenté par l'agriculture d'oasis, où la culture du blé se fait à petite échelle.
- Le système moderne : il traduit une nouvelle agriculture dite stratégique qui vise essentiellement la production de céréales (blé et orge). Elle est pratiquée à grande échelle par l'introduction d'une nouvelle technique d'irrigation "le pivot". La culture de céréales sous pivot au sud a été introduite dans les années quatre-vingt. Elle a connu un essor considérable en matière d'augmentation des superficies. Des rendements de pointes ont même été réalisés au niveau de certaines exploitations agricoles dépassant parfois les 70 quintaux/ha.



FIGURE 1.8 – Irrigation sous pivot en zone saharienne

1.4.2 Politique agricole

La politique du renouveau de l'économie agricole vise l'intensification des productions céréalières pour l'amélioration de la sécurité alimentaire et la réduction des importations. Cette politique encourage l'utilisation de fertilisants, de semences de qualité et l'introduction de technologies d'irrigation d'appoint. Elle inclut des mesures d'accompagnement pour soutenir au mieux l'agriculteur.

Dans un souci de renforcer la production céréalière nationale, le gouvernement encourage l'expansion de l'agriculture vers les terres du sud.

1.5 Localisation de la région de l'étude

Dans le cadre de notre étude, nous avons récolté des données relatives au rendement de blé dur sous conditions sahariennes. L'essai expérimental a été mené au domaine agricole "HADJADJ", qui se situe à 15 Km au Nord de la ville d'El Goléa, au lieu dit Sahab El-Matana à Gardaïa. L'exploitation s'étend sur une superficie totale de 600 ha, la céréaliculture est la principale activité de l'exploitation pour laquelle l'irrigation est assurée

par 12 pivots d'environ 40 ha chacun. Nous y trouvons également des cultures légumières telles que : la pomme de terre, l'oignon et la pastèque, de l'arboriculture : palmiers dattiers, orangers, vigne... et de l'élevage.



FIGURE 1.9 – Localisation de l'exploitation

1.6 Position du problème

L'agriculture en zone saharienne devient possible grâce à la maîtrise de deux éléments fondamentaux : l'irrigation et la fertilisation.

Les agriculteurs s'approvisionnent en engrais nécessaires à la fertilisation auprès de nombreuses entreprises se partageant le marché sur le territoire algérien. Dans le cadre de notre étude portant sur l'acquisition des engrais, nous nous sommes intéressées aux catalogues de produits des leaders du marché, à savoir les entreprises : PROFERT, TIMAC et FERTIAL décrits en annexe. Ces entreprises fournissent un large choix de produits allant des engrais (solides, solubles ou organiques) aux produits phytosanitaires.

Face à une concurrence rude, les firmes font des efforts de recherche pour proposer des catalogues riches en formules, toutes diversifiées en composition en éléments fertilisants. Ces formules -décrites en annexe- aussi nombreuses soient elles, ne peuvent pas toujours correspondre simplement et parfaitement aux besoins de fertilisation de l'agriculteur. Une

combinaison entre les produits des différentes firmes peut s'avérer nécessaire.

Néanmoins, le choix de tel ou tel produit en fonction de sa composition peut s'avérer être une tâche complexe en raison de la présence d'un large choix de produits. Compte tenu du prix relativement important des engrais, un choix mal fondé pèsera lourd sur le budget de l'entreprise agricole.

La question de la fertilisation est cruciale, particulièrement pour l'agriculteur du Sud, puisque les terres sahariennes sont pauvres en éléments nutritifs, les engrais sont apportés en grande quantité c'est pourquoi nous nous proposons de fournir une réponse à la question du choix optimal d'engrais solides à faire au vu des catalogues des trois entreprises citées, sachant la composition de chaque produit et son coût et ce, afin de répondre aux besoins annuels des agriculteurs en engrais solides de façon à minimiser son coût d'acquisition.

En effet, face à une offre variée du marché, l'agriculteur s'interroge sur les conséquences du choix d'un plan de fertilisation. Un plan de fertilisation concerne les doses des éléments nutritifs à apporter à la culture et l'engrais à utiliser.

Au vu de cette problématique, la première étape de notre travail portera sur une étude statistique relative aux données issues d'une expérimentation menée au niveau de l'exploitation "HADJADJ" à El Goléa au cours de la saison agricole 2007-2008, qui va nous permettre d'apporter une réponse quant à l'effet sur le rendement de blé dur, à l'utilisation de deux marques d'engrais phosphatés distinctes, utilisées à différentes doses. Si l'analyse établit l'existence de l'effet d'un facteur sur le rendement, nous tâcherons de dégager quelles sont les modalités des facteurs offrant un rendement maximal et ce à l'aide de tests statistiques.

La seconde étape sera d'essayer de déterminer la combinaison optimale d'engrais solides à acquérir auprès des trois entreprises, de façon à répondre aux exigences de l'agriculteur et ce grâce à l'implémentation d'une métaheuristique.

Enfin, nous tâcherons de lier les deux parties d'application : partie statistique et optimisation combinatoire, en estimant les coûts de fertilisation de l'expérimentation. Nous tâcherons ainsi de répondre aux décideurs quant au choix à faire entre les deux produits testés et les doses appropriées de phosphore, relativement au rendement qu'ils offrent et à leur coût d'acquisition.

Conclusion

Un bilan général de la situation du blé et de la fertilisation en Algérie a été dressé. Il en ressort que beaucoup reste à faire en termes de rendement du blé.

La maîtrise de la fertilisation est un passage obligatoire : fertiliser à bonne dose et avec les bons produits sera une des clés de l'aboutissement des objectifs en terme de rendement. A coté de l'objectif " rendement ", l'objectif de minimisation des coûts de fertilisation est tout aussi primordial pour l'agriculteur. Il peut être atteint avec un choix optimal des produits fertilisants disponibles sur le marché.

2

Outils Statistiques

Introduction

Dans ce chapitre nous présentons les notions théoriques relatives à la partie statistique de notre étude, à savoir les modélisations et les conditions d'application ainsi que les tests statistiques qui seront utilisés pour l'analyse des données.

2.1 Modélisation statistique

Une grande partie des mathématiques appliquées consiste, d'une certaine façon, à faire de la modélisation, c'est-à-dire à définir un (ou plusieurs) modèle(s) de nature mathématique, permettant de rendre compte d'une manière suffisamment générale, d'un phénomène donné, qu'il soit physique, biologique, économique ou autre.

Plus appliquée, la modélisation statistique consiste essentiellement à définir des outils appropriés pour modéliser des données observées, en tenant compte de leur nature aléatoire. Il faut noter que le terme de modélisation statistique est très général et que, à la limite, toute démarche statistique en relève [6].

2.1.1 Un peu de vocabulaire

Afin de faciliter la compréhension des modèles présentés. Il est nécessaire de définir certaines notions qui apparaîtront de façon récurrente dans la suite de notre travail [5].

- **Réponse** : variable dont nous cherchons à comprendre le comportement, généralement notée Y . Nous considérerons dans le cadre de cette étude des réponses quantitatives continues.
- **Facteur** : variable dont nous voulons nous servir pour analyser les variations de la réponse. Les niveaux ou modalités d'un facteur sont notés en indiquant le symbole associé au facteur : A_1, \dots, A_I avec au total I modalités du facteur A .
- **Modèle statistique** : équation reliant les différentes mesures $Y_{i,j}, \dots$ de la réponse Y aux effets des niveaux $A_1, \dots, A_I, B_1, \dots, B_J, \dots$ des facteurs A, B, \dots pour lesquels ont été réalisés ces différentes mesures au travers d'une relation fonctionnelle f et tout en modélisant les fluctuations expérimentales à l'aide de variables aléatoires d'erreur généralement notée $\varepsilon_{i,j}, \dots$, il s'écrit :

$$Y_{i,j}, \dots = \mu + f(A_i, B_j, \dots) + \varepsilon_{i,j}, \dots$$

- **Effet du facteur** : en fonction du type d'effet auquel les termes du modèle sont associés, nous distinguons deux types d'effets :
 1. Effet principal
 2. Effet d'interaction

L'effet d'un facteur peut être fixe, si le facteur est contrôlé. Il peut être aléatoire, s'il ne l'est pas.

2.2 Dispositifs d'essai expérimental

Dans le domaine agronomique, le choix d'un dispositif expérimental dépend des objectifs de l'essai, de son application pratique, des conditions environnementales, des méthodes d'application, des cultures et des traitements. Il faut bien analyser tous ces facteurs avant de choisir le dispositif adéquat qui répondra aux objectifs de l'expérimentation pour démontrer l'efficacité d'un produit.

On utilise habituellement un dispositif en blocs aléatoires complets ou un dispositif en parcelles subdivisées, dans lequel les parcelles traitées sont réparties au hasard dans chaque bloc. D'autres modèles, comme le dispositif complètement aléatoire, peuvent servir à conduire des essais en milieux complètement homogènes, tels des serres ou des chambres de croissance [2].

2.2.1 Dispositif en blocs aléatoires complets

Le dispositif en blocs aléatoires complets est le plus courant dans la recherche agricole. Dans le DBAC, les parcelles sont aléatoires et assignées uniformément dans les blocs. Les traitements sont assignés aux parcelles de manière à ce que chacune apparaisse le même nombre de fois dans un bloc, d'ordinaire une seule fois par bloc, ce qui fait de chaque bloc une répétition. La disposition des blocs dans le modèle dépend des caractéristiques du site de test. L'un des avantages de ce type de modèle est qu'il peut empêcher une source de variation d'entraîner des erreurs expérimentales et ainsi, d'augmenter la précision globale de l'essai. En outre, le nombre de traitements et de blocs utilisés n'a pas d'importance, pourvu que chaque traitement apparaisse le même nombre de fois dans le bloc [2].

2.3 L'analyse de la variance

2.3.1 Définition

L'analyse de la variance (terme souvent abrégé par le terme anglais ANOVA : ANalysis Of VAriance) est un test statistique paramétrique permettant de vérifier que plusieurs échantillons sont issus ou non d'une même population.

Ce test s'applique lorsqu'on mesure une ou plusieurs variables explicatives appelées facteurs de variabilité, qui ont une influence sur la distribution d'une variable continue à expliquer. On parle d'analyse à un facteur, lorsque l'analyse porte sur un modèle décrit par un facteur de variabilité, d'analyse à deux facteurs ou d'analyse multifactorielle.

L'analyse de variance est, à quelques exceptions près, la méthode statistique la plus appropriée pour déterminer s'il existe une différence statistique entre les ensembles de données agronomiques [2].

Lorsque les hypothèses d'une analyse de la variance ne sont pas avérées relativement à la normalité des données et que les données ne peuvent être transformées, il peut être utile

de faire appel à un test non paramétrique. L'équivalent non paramétrique de l'ANOVA à un facteur est le test de Kruskal-Wallis, une analyse de variance à un facteur par classement qui ne fait aucune hypothèse de normalité des échantillons [2, 3].

2.3.2 Modèles d'analyse de la variance "ANOVA"

Dans un modèle d'analyse de la variance à plus de deux facteurs, nous distinguons trois modélisations possibles :

1. Modèle à effet fixe avec ou sans répétition des mesures, où les facteurs sont contrôlés par l'expérimentateur.
2. Modèle à effet aléatoire avec ou sans répétition des mesures, où les facteurs sont aléatoires, non contrôlés par l'expérimentateur.
3. Modèle à effet mixte avec ou sans répétition des mesures, où certains facteurs sont contrôlés par l'expérimentateur, d'autres sont aléatoires.

Nous décrivons dans ce qui suit le modèle d'analyse de la variance à deux facteurs à effets fixes avec répétition, les conditions d'application et les tests d'hypothèses associés.

2.4 ANOVA à 2 facteurs à effet fixe

Un facteur contrôlé A se présente sous I modalités, chacune d'entre elles étant notée A_i . Un facteur contrôlé B se présente sous J modalités, chacune d'entre elles étant notée B_j . Pour chacun des couples de modalités (A_i, B_j) nous effectuons $K > 2$ mesures d'une réponse Y qui est une variable continue. Nous notons $n = I \times J \times K$ le nombre total de mesures ayant été effectuées[3, 4, 5].

2.4.1 Modélisation

Nous introduisons le modèle :

$$Y_{i,j,k} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{i,j} + \epsilon_{i,j,k} \quad (2.1)$$

$$i = 1, \dots, I, j = 1, \dots, J, k = 1, \dots, K$$

Où $Y_{i,j,k}$ est la valeur prise par la réponse Y dans la condition (A_i, B_j) lors de la k ème répétition.

Avec les contraintes :

$$\sum_{i=1}^I \alpha_i = 0, \sum_{j=1}^J \beta_j = 0, \sum_{i=1}^I (\alpha\beta)_{i,j} = 0 \forall j = 1, \dots, J \text{ et } \sum_{j=1}^J (\alpha\beta)_{i,j} = 0 \forall i = 1, \dots, I$$

Nous postulons les hypothèses classiques suivantes pour les erreurs :

$$\forall (i, j, k), i = 1 \dots I, j = 1 \dots J, k = 1 \dots K, \epsilon_{i,j,k} \rightsquigarrow \mathcal{N}(0, \sigma^2)$$

$$\text{Cov}(\epsilon_{i,j,k}, \epsilon_{l,m,n}) = 0 \text{ si } (i, j, k) \neq (l, m, n) \text{ avec } 1 \leq i, l \leq I, 1 \leq j, m \leq J, 1 \leq k, n \leq K$$

Nous regroupons les valeurs que peut prendre la réponse Y dans les conditions (A_i, B_j) lors des K répétitions dans le tableau suivant :

	B_1	...	B_j	...	B_J
A_1	$Y_{1,1,1} \dots Y_{1,1,K}$...	$Y_{1,j,1} \dots Y_{1,j,K}$...	$Y_{1,J,1} \dots Y_{1,J,K}$
\vdots	\vdots	\vdots	...	\vdots	...
A_i	$Y_{i,1,1} \dots Y_{i,1,K}$...	$Y_{i,j,1} \dots Y_{i,j,K}$...	$Y_{i,J,1} \dots Y_{i,J,K}$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
A_I	$Y_{I,1,1} \dots Y_{I,1,K}$...	$Y_{I,j,1} \dots Y_{I,j,K}$	\vdots	$Y_{I,J,1} \dots Y_{I,J,K}$

2.4.2 Hypothèses fondamentales

La forme générale de l'analyse de variance repose sur le test de Fisher et donc sur la normalité des distributions, l'homoscédasticité et l'indépendance des échantillons.

L'indépendance

On suppose que chaque échantillon analysé est indépendant des autres échantillons. En pratique, c'est la problématique qui permet de supposer que les échantillons sont indépendants, ou les conditions de l'expérience.

La normalité

On suppose, que les échantillons sont issus d'une même population et suivent une loi normale. Il est donc nécessaire de vérifier la normalité des distributions, à l'aide de tests statistiques tels que le test de Shapiro-Wilk. Dans le cas contraire, on pourra utiliser les variantes non paramétriques de l'analyse de variance, telle que le test de Kruskal-Wallis, ou l'ANOVA de Friedman.

L'homoscédasticité

Les échantillons doivent être de même variance, on parle d'homogénéité des variances. Plusieurs tests statistiques existent pour tester cette condition, nous citons par exemple le test de Bartlett, ou de Levene. Si l'hypothèse d'homogénéité des variances est rejetée, une transformation non linéaire peut être réalisée sur les données brutes, qui aura pour effet de normaliser les distributions. Notons qu'une transformation est très facile à réaliser avec la plupart des logiciels d'analyses statistiques.

2.4.3 Décomposition de la variation totale

La première étape de l'analyse de la variance, consiste à expliquer la variance totale sur l'ensemble des échantillons en fonction de la variance due aux facteurs (la variance expliquée par le modèle), de la variance due à l'interaction entre les facteurs et de la variance résiduelle aléatoire (la variance non expliquée par le modèle).

Dans ce qui suit :
$$\frac{1}{IJK} \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K Y_{i,j,k} = \bar{Y}$$

La variation théorique due au facteur A est définie par :

$$SC_A = JK \sum_{i=1}^I \left(\frac{1}{JK} \sum_{j=1}^J \sum_{k=1}^K Y_{i,j,k} - \bar{Y} \right)^2$$

La variation théorique due au facteur B est définie par :

$$SC_B = IK \sum_{j=1}^J \left(\frac{1}{IK} \sum_{i=1}^I \sum_{k=1}^K Y_{i,j,k} - \bar{Y} \right)^2$$

La variation théorique due à l'interaction des facteurs A et B est définie par :

$$SC_{AB} = K \sum_{i=1}^I \sum_{j=1}^J \left(\frac{1}{K} \sum_{k=1}^K Y_{i,j,k} - \frac{1}{JK} \sum_{j=1}^J \sum_{k=1}^K Y_{i,j,k} - \frac{1}{IK} \sum_{i=1}^I \sum_{k=1}^K Y_{i,j,k} + \bar{Y} \right)^2$$

La variation résiduelle théorique est quant à elle définie par :

$$SC_R = \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K \left(Y_{i,j,k} - \frac{1}{K} \sum_{k=1}^K Y_{i,j,k} \right)^2$$

Enfin la variation totale théorique est égale à :

$$SC_{TOT} = \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K (Y_{i,j,k} - \bar{Y})^2$$

La relation fondamentale de l'ANOVA :

$$SC_{TOT} = SC_A + SC_B + SC_{AB} + SC_R$$

2.4.4 Tableau de l'ANOVA

Nous résumons les informations précédentes dans le tableau de l'ANOVA ci-dessous :

Source	Variation	DDL	Carré moyen	F	Décision
Facteur A	SC_A	η_A	$s_A^2 = \frac{SC_A}{\eta_A}$	$f_A = \frac{s_A^2}{s_R^2}$	\mathcal{H}_0^A ou \mathcal{H}_1^A
Facteur B	SC_B	η_B	$s_B^2 = \frac{SC_B}{\eta_B}$	$f_B = \frac{s_B^2}{s_R^2}$	\mathcal{H}_0^B ou \mathcal{H}_1^B
Facteur AB	SC_{AB}	η_{AB}	$s_{AB}^2 = \frac{SC_{AB}}{\eta_{AB}}$	$f_{AB} = \frac{s_{AB}^2}{s_R^2}$	\mathcal{H}_0^{AB} ou \mathcal{H}_1^{AB}
Résiduelle	SC_R	η_R	$s_R^2 = \frac{SC_R}{\eta_R}$		
Totale	SC_{TOT}	η_{TOT}			

Où le carré moyen est un estimateur sans biais des variances. Le rapport entre deux estimateurs non biaisés de la variance défini par la statistique f doit donc suivre une Loi de Fisher.

Les degrés de liberté (DDL) associés à chaque ligne du tableau de l'ANOVA sont les suivants :

Source	Degré de liberté
Facteur A	$\eta_A = I - 1$
Facteur B	$\eta_B = J - 1$
Facteur AB	$\eta_{AB} = \eta_A \eta_B$
Résiduelle	$\eta_R = IJ(K - 1)$
Totale	$\eta_{TOT} = IJK - 1$

2.4.5 Tests d'hypothèses

Nous souhaitons faire des tests d'hypothèse, pour montrer l'existence de l'effet d'un facteur sur la variable réponse. Ces tests se basent sur les statistiques $f_{FACTEUR}$ décrites dans le tableau.

Effet du facteur A

$$\begin{aligned} \text{''}\mathcal{H}_0^A\text{''} &: \alpha_1 = \alpha_2 = \dots = \alpha_I = 0 \\ \text{Contre ''}\mathcal{H}_1^A\text{''} &: \text{Il existe } i_0 \text{ } 1 \leq i_0 \leq I \text{ tel que } \alpha_{i_0} \neq 0 \end{aligned}$$

Sous l'hypothèse nulle \mathcal{H}_0^A d'absence d'effet du facteur A et lorsque les conditions de validité du modèle sont respectées, f_A est la réalisation d'une variable aléatoire qui suit une loi de Fisher à $I - 1$ et $IJ(K - 1)$ degrés de liberté. Nous concluons alors à l'aide de la p-valeur, rejet si elle est inférieure ou égale au seuil α du test. Ou bien à l'aide de la table de Fisher, rejet si la valeur f_A est supérieure ou égale à la valeur critique issue de la table. Lorsque l'hypothèse nulle \mathcal{H}_0^A est rejetée, c'est à dire que le facteur A a un effet significatif sur la réponse, nous pouvons procéder à des comparaisons multiples des différents effets des niveaux du facteur.

Effet du facteur B

$$\begin{aligned} \text{''}\mathcal{H}_0^B\text{''} &: \beta_1 = \beta_2 = \dots = \beta_J = 0 \\ \text{Contre ''}\mathcal{H}_1^B\text{''} &: \text{Il existe } i_0 \text{ } 1 \leq i_0 \leq J \text{ tel que } \beta_{i_0} \neq 0 \end{aligned}$$

Sous l'hypothèse nulle \mathcal{H}_0^B précédente d'absence d'effet du facteur B et lorsque les conditions de validité du modèle sont respectées, f_B est la réalisation d'une variable aléatoire qui suit une loi de Fisher à $J - 1$ et $IJ(K - 1)$ degrés de liberté. Nous concluons alors à l'aide de la p-valeur, rejet si elle est inférieure ou égale au seuil α du test. Lorsque l'hypothèse nulle \mathcal{H}_0^B est rejetée, nous pouvons procéder à des comparaisons multiples des différents

effets des niveaux du facteur.

Effet de l'interaction des facteurs A et B

$$\begin{aligned} & \text{''}\mathcal{H}_0^{AB}\text{''} : (\alpha\beta)_{1,1} = (\alpha\beta)_{1,2} = \dots = (\alpha\beta)_{I,J} = 0 \\ & \text{Contre ''}\mathcal{H}_1^{AB}\text{''} : \text{Il existe } (i_0, j_0), 1 \leq i_0 \leq I, 1 \leq j_0 \leq J \text{ tel que } (\alpha\beta)_{i_0, j_0} \neq 0 \end{aligned}$$

Sous l'hypothèse nulle \mathcal{H}_0^{AB} précédente d'absence d'effet du facteur A et B et lorsque les conditions de validité du modèle sont respectées, f_{AB} est la réalisation d'une variable aléatoire qui suit une loi de Fisher à $(I - 1)(J - 1)$ et $IJ(K - 1)$ degrés de liberté. Nous concluons alors à l'aide de la p-valeur, rejet si elle est inférieure ou égale au seuil α du test. Lorsque l'hypothèse nulle \mathcal{H}_0^{AB} est rejetée, nous pouvons procéder à des comparaisons multiples des différents effets des niveaux du facteur.

2.5 Tests de comparaisons multiples

Lors d'une ANOVA, une valeur F significative révèle seulement que les moyennes de traitement ne sont pas égales. Les tests dits post-hoc (c'est-à-dire procédures de comparaison multiple et tests d'écart des moyennes) servent à déterminer les moyennes de traitement qui diffèrent considérablement les unes des autres. Pour identifier les échantillons correspondants, nous utilisons différents tests dits post-hoc ou tests de comparaisons multiples, MCP : Multiple Comparison Test [4, 9].

Ces tests obligent en général à augmenter les risques de l'analyse (en terme de risque statistique). Il s'agit d'une généralisation à k populations du test t de Student de comparaison des moyennes de deux échantillons.

Il existe plusieurs tests de comparaisons multiples, nous citons par exemple : les tests LSD : Least Significant Difference de Fisher, les tests de Student-Newman-Keuls, les tests HSD : Honestly Significant Difference de Tukey, les tests de Bonferroni et de Sheffé [2].

– Question :

Pourquoi ne pas faire directement ces tests, sans passer par une analyse de la variance avant ?

– **Réponse :**

Lorsque l'on analyse plusieurs variables explicatives ayant plusieurs modalités chacune, le nombre de combinaison possible devient rapidement très grand.

2.5.1 Test de Newman-Keuls

Le test de Newman-Keuls est un test de comparaison de moyennes par paires, pratiqué à l'issue d'une ANOVA. Son intérêt principal est qu'il évite, par construction, les situations "paradoxaes" où la moyenne d'un groupe G_1 pourrait être considérée comme significativement différente de celle d'une groupe G_2 , alors qu'elle serait considérée comme non significativement différente de celle d'un groupe G_3 , pourtant plus éloignée d'elle que celle de G_2 .

De fait, ayant déterminé que les moyennes m_i et m_j des groupes G_i et G_j ne sont pas significativement différentes, il n'est pas nécessaire de tester par paires les moyennes comprises entre m_i et m_j (incluses), le test garantissant qu'elles seront, elles aussi, considérées comme non significativement différentes.

Le résultat du test est une série de paires de groupes dont les moyennes sont considérées comme significativement différentes, au niveau de risque α choisi.

Statistique de Newman-Keuls et valeurs critiques

Le test de Newman-Keuls fournit pour chaque paire de comparaison une valeur de statistique de test « $q_{observe}$ » que l'on confronte à une valeur critique théorique issue de la table de Newman-Keuls. Cette valeur critique dépend, pour chaque paire de groupes comparés :

- Du niveau de risque α choisi par l'analyste,
- Des effectifs de ces groupes,
- De la différence des rangs des moyennes des deux groupes comparés. Par exemple, si les moyennes se présentent dans l'ordre $m_1 < m_2 < m_4 < m_3$, si on compare m_2 à m_3 , cette différence (appelée "étendue") est $K = 3$ et si on compare m_2 à m_4 , l'étendue est $K = 2$.

Ainsi, la valeur critique est réajustée à chaque nouvelle comparaison, en fonction du nombre de moyennes comprises dans l'étendue délimitée par les deux moyennes comparées,

et en fonction des effectifs des groupes comparés.

La présence d'un « $q_{observe}$ » supérieur ou égal à la valeur critique « $q_{critique}$ » lue dans la table permet de rejeter l'hypothèse " \mathcal{H}_0 " selon laquelle les moyennes des deux distributions concernées sont égales.

Les paires de moyennes comprises entre deux moyennes jugées non significativement différentes ne sont pas testées.

Conclusion

D'une façon générale, les bonnes pratiques scientifiques dictent que la méthode d'analyse statistique doit être envisagée avant de procéder à l'expérimentation. Le choix du test statistique approprié dépendra fortement des paramètres expérimentaux compris dans le dispositif expérimental, comme le type de paramètre mesuré (par exemple, qualitatif ou quantitatif), le nombre de répétitions et la taille des échantillons.

3

Optimisation Combinatoire

Introduction

Dans ce chapitre, nous donnons l'essentiel de la théorie concernant les problèmes d'optimisation combinatoire et les métaheuristiques, puis nous détaillons la méthode GRASP et son application aux problèmes de programmation linéaire en nombres entiers.

3.1 Problèmes d'optimisation combinatoire

L'optimisation combinatoire occupe une place très importante en recherche opérationnelle, en mathématiques discrètes et en informatique. Son importance se justifie d'une part par la grande difficulté des problèmes d'optimisation [21] et d'autre part par de nombreuses applications pratiques pouvant être formulées sous la forme d'un problème d'optimisation combinatoire [28].

Bien que les problèmes d'optimisation combinatoire soient souvent faciles à définir, ils sont généralement difficiles à résoudre.

On entend par "difficile" le fait que la résolution de ces problèmes nécessite l'examen d'un très grand nombre (exponentiel) de combinaisons [30]. On parle d'explosion combinatoire,

ce phénomène, qui transforme un problème apparemment très simple en un véritable casse-tête dès lors qu'on augmente la taille du problème à résoudre.

3.1.1 Formulation d'un problème d'optimisation combinatoire

Un problème d'optimisation combinatoire est défini [27] par un ensemble fini $E = \{1, \dots, n\}$, un ensemble de solutions réalisables $X \subseteq 2^E$ et une fonction objectif $f : 2^E \rightarrow \mathbb{R}$.

Dans le cas d'une "minimisation", nous cherchons une solution optimale $x^* \in X$ telle que $f(x^*) \leq f(x), \forall x \in X$.

Cas particulier : PLNE

Un programme linéaire en nombres entiers, usuellement noté PLNE, sous forme standard s'écrit comme suit :

$$\begin{aligned} \text{Min } f(x) &= c^t x \\ \text{sc : } Ax &\leq b \\ x &\in \mathbb{N}^n \end{aligned}$$

Où : c, b et x sont des vecteurs de taille n ,
 A est une matrice d'ordre (m, n) .

3.1.2 Méthodes de résolution d'un problème d'optimisation combinatoire

Etant donnée l'importance des problèmes d'optimisation combinatoire, de nombreuses méthodes de résolution ont été développées en recherche opérationnelle. Ces méthodes peuvent être classées sommairement en deux grandes catégories [10, 15] : les méthodes exactes (complètes) qui garantissent la complétude de la résolution et les méthodes approchées (incomplètes) qui perdent la complétude pour gagner en efficacité.

Méthodes exactes

Le principe essentiel d'une méthode exacte consiste généralement à énumérer, souvent de manière implicite, l'ensemble des solutions de l'espace de recherche. Pour améliorer

l'énumération des solutions, une telle méthode dispose de techniques pour détecter le plus tôt possible les échecs (calculs de bornes) et d'heuristiques spécifiques pour orienter les différents choix.

Parmi les méthodes exactes, on trouve la plupart des méthodes traditionnelles (développées depuis une trentaine d'années) telles que les techniques de séparation et évaluation progressive (SEP ou Branch and Bound) ou les algorithmes avec retour arrière.

Les méthodes exactes ont permis de trouver des solutions optimales pour des problèmes de taille raisonnable. Malgré les progrès réalisés (notamment en matière de programmation linéaire en nombres entiers), comme le temps de calcul nécessaire pour trouver une solution risque d'augmenter exponentiellement avec la taille du problème, les méthodes exactes rencontrent généralement des difficultés face aux applications de taille importante [15].

Méthodes approchées

Les méthodes approchées constituent une alternative très intéressante pour traiter les problèmes d'optimisation de grande taille si l'optimalité n'est pas primordiale. En effet, elles contournent le problème de l'explosion combinatoire en faisant délibérément des impasses et en n'explorant qu'une partie des combinaisons [30]. Par conséquent, elles peuvent ne pas trouver de solution, et encore moins prouver l'absence de solution (ou l'optimalité de la solution dans le cas de problèmes d'optimisation) ; en contrepartie, la complexité en temps est généralement faiblement polynomiale.

Ces méthodes sont utilisées depuis longtemps par de nombreux praticiens. On peut citer les méthodes gloutonnes et l'amélioration itérative [15].

3.2 Méthodes de résolution heuristiques

Il existe essentiellement deux grandes familles d'approches heuristiques :

3.2.1 Approches par voisinage

Les approches par voisinage explorent l'espace des combinaisons d'un problème d'optimisation combinatoire en le structurant en terme de voisinage : partant d'une ou de plusieurs combinaisons initiales (généralement prises aléatoirement), l'idée est de générer à chaque étape une ou plusieurs combinaisons voisines de une ou plusieurs combinaisons

courantes [30].

Les approches par voisinage les plus connues sont les algorithmes génétiques, l'optimisation par essaims de particules et la recherche locale que nous détaillerons un peu plus loin.

3.2.2 Approches constructives

Les approches constructives construisent une ou plusieurs combinaisons de façon incrémentale, c'est à dire, en partant d'une combinaison vide et en ajoutant des composants de combinaison jusqu'à obtenir une combinaison complète. Ces approches sont dites "basées sur les modèles" [33], dans le sens où elles utilisent un modèle, généralement stochastique, pour choisir à chaque itération le prochain composant de combinaison à ajouter à la combinaison en cours de construction [30].

Il existe différentes stratégies pour choisir le composant à ajouter à chaque itération, les plus connues étant les algorithmes par estimation de distribution, la métaheuristique d'optimisation par colonies de fourmis, les stratégies gloutonnes et gloutonnes aléatoires sur lesquelles s'appuie notre projet.

3.3 Métaheuristiques

Depuis une dizaine d'années, des progrès importants ont été réalisés avec l'apparition d'une nouvelle génération de méthodes approchées puissantes et générales, souvent appelées métaheuristiques [1, 25].

Les métaheuristiques sont généralement des algorithmes stochastiques itératifs, qui progressent vers un optimum global de la fonction objectif, sans être piégé par les optima locaux. Elles sont représentées essentiellement par les méthodes de voisinage comme le recuit simulé et la recherche tabou, ainsi que les algorithmes évolutifs comme les algorithmes génétiques et les stratégies d'évolution.

3.3.1 Source d'efficacité des métaheuristiques

Considérons un exemple [10] de problème d'optimisation, dont l'allure de la fonction objectif peut être schématiquement représentée comme sur la figure 3.1.

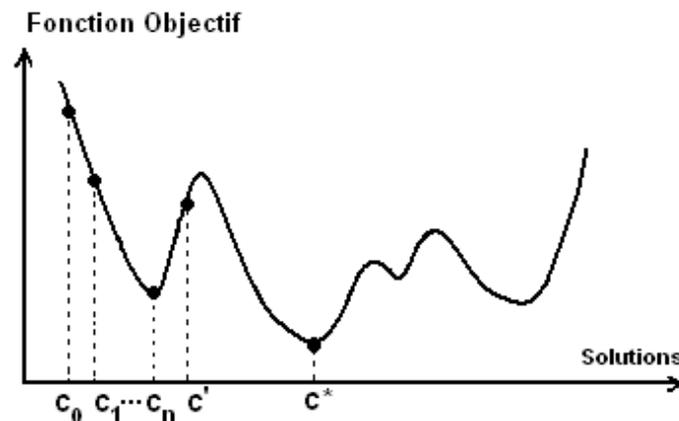


FIGURE 3.1 – Exemple d'allure de la fonction objectif d'un problème d'optimisation

Le principe d'un algorithme classique d'amélioration itérative est le suivant : on part d'une solution initiale c_0 , qui peut être choisie au hasard. On essaie alors une modification élémentaire souvent appelée "mouvement" et on compare les valeurs de la fonction objectif avant et après cette modification. Si le changement conduit à une diminution de la fonction objectif, il est accepté et la combinaison voisine c_1 obtenue sert de point de départ pour un nouvel essai. Dans le cas contraire, on revient à la solution précédente avant de faire une autre tentative et le processus est itéré jusqu'à ce que toute modification rende le résultat moins bon.

La figure 3.1 montre que cet algorithme d'amélioration itérative ne conduit pas, en général, au minimum absolu, mais seulement à un minimum local c_n qui constitue la meilleure des solutions accessibles compte tenu de la solution initiale.

Pour surmonter l'obstacle des minimums locaux, l'idée de base des métaheuristiques dites "de voisinage" s'est montrée fructueuse : il s'agit d'autoriser de temps en temps, des mouvements de remontée (c'est le cas si on passe de c_n à c'), autrement dit d'accepter une dégradation temporaire de la situation. Un mécanisme de dégradation spécifique à chaque métaheuristique permet d'éviter la divergence du procédé.

3.3.2 Place des métaheuristiques dans une classification des méthodes d'optimisation

La figure 3.2 présente une classification générale des méthodes d'optimisation mono-objectif [10].

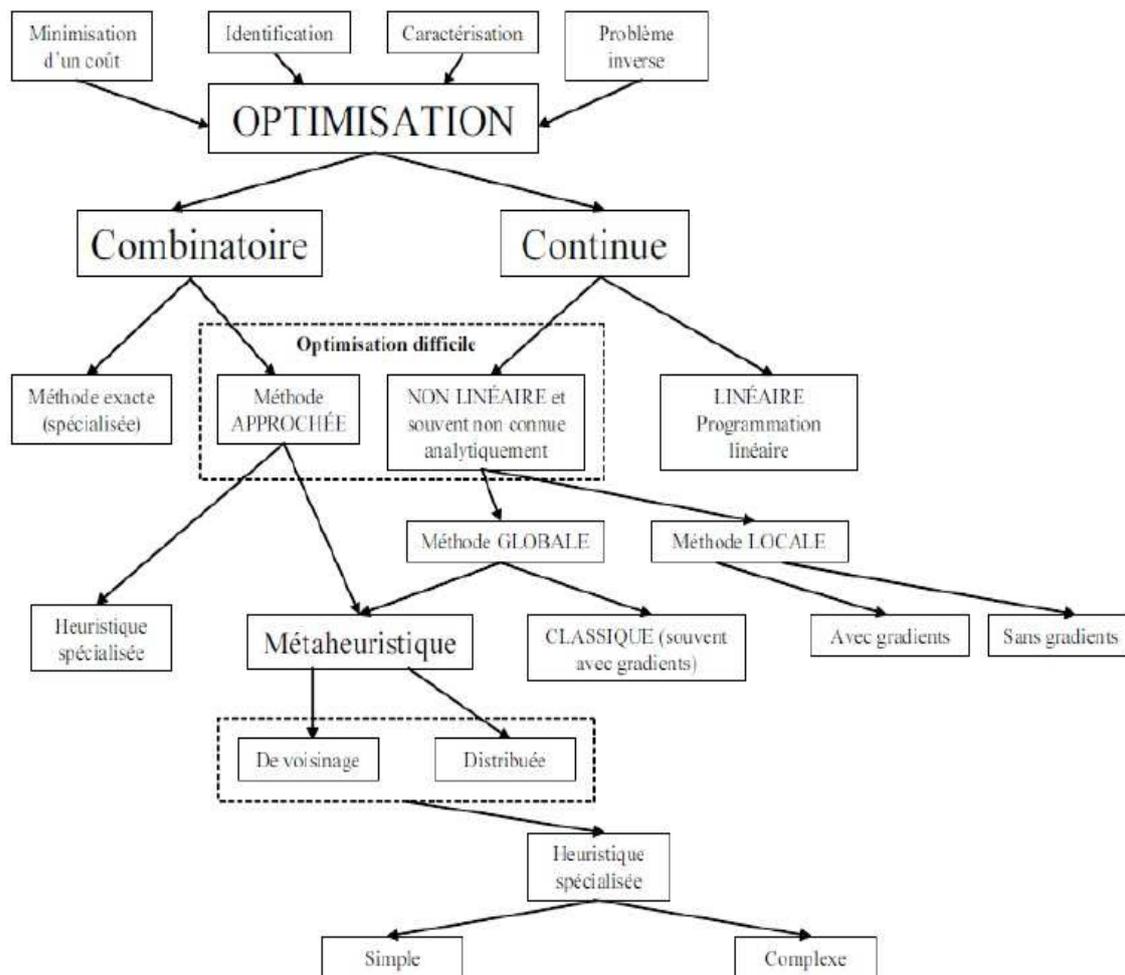


FIGURE 3.2 – Place des métaheuristiques dans une classification des méthodes d'optimisation

3.3.3 Méthodes hybrides

Une pseudo-classe d'algorithmes a émergé ces dernières années pour résoudre les problèmes d'optimisation, dites méthodes hybrides.

On parle d'hybridation quand la métaheuristique considérée est composée de plusieurs méthodes se répartissant les tâches de recherche. Ainsi, les méthodes hybrides tirent profit de la complémentarité de plusieurs méthodes (exactes et/ou approchées) entre elles en vue d'améliorer le comportement global de l'algorithme dédié pour la méthode.

Ces méthodes associent souvent une métaheuristique évolutive et une méthode de recherche locale [17]. Nous citons en exemple les systèmes hybrides IA-RO consacrés aux moyens d'intégrer programmation par contraintes sur domaines finis et programmation linéaire en nombres entiers pour la résolution des problèmes d'optimisation combinatoire [8, 16] et la méthode GRASP [20, 27].

3.4 Méthode GRASP

Développée par Feo et Resende 1989 [11] (Resende and Ribeiro 2001 [26], Pitsouli and Resende 2002 [22]), la méthode GRASP (Greedy Randomized Adaptive Search Procedure) est une métaheuristique adaptée aux problèmes dont les solutions se construisent pas à pas. Son algorithme contient deux phases :

- une phase de construction d'une solution,
- une phase d'amélioration de la solution.

La phase de construction est basée sur une procédure gloutonne aléatoire. Quant à la phase d'amélioration, elle est faite grâce à une procédure de recherche locale. Dans notre travail, nous détaillerons la méthode GRASP et son application à la programmation linéaire, mais d'abord, expliquons le principe d'une méthode gloutonne et celui d'une recherche locale.

3.4.1 Algorithmes gloutons et gloutons aléatoires

Les algorithmes gloutons (greedy) construisent une solution en partant d'une combinaison vide et en choisissant à chaque itération un composant de combinaison pour lequel une heuristique donnée est maximale [30].

Les algorithmes gloutons aléatoires (greedy randomized) construisent plusieurs combinaisons et adoptent également une stratégie gloutonne pour le choix de composants à ajouter aux combinaisons en cours de construction, mais ils introduisent un peu d'aléatoire afin de diversifier les combinaisons [30].

3.4.2 Recherche locale

Une recherche locale explore l'espace des combinaisons de proche en proche, en partant d'une combinaison initiale et en sélectionnant à chaque itération une combinaison voisine de la combinaison courante. Les principales étapes d'une recherche locale sont résumées ci-après [14, 30].

Fonction de voisinage

L'algorithme de recherche locale est paramétré par une fonction de voisinage $v : X \rightarrow \mathcal{P}(X)$ définissant l'ensemble des combinaisons que l'on peut explorer à partir d'une combinaison donnée.

Génération de la combinaison initiale

La combinaison initiale à partir de laquelle le processus d'exploration commence est souvent générée de façon aléatoire. Elle est parfois générée en suivant une heuristique constructive gloutonne.

Choix du voisin

A chaque itération de la recherche locale, il s'agit de choisir une combinaison dans le voisinage de la combinaison courante. Ce choix d'une combinaison voisine est appelé "mouvement". Il existe un très grand nombre de stratégies de choix. On peut par exemple sélectionner à chaque itération le meilleur voisin [29], c'est à dire celui qui améliore le plus la fonction objectif ou bien le premier voisin trouvé qui améliore la fonction objectif.

Critère d'arrêt

Les critères d'arrêt pour la recherche locale sont ceux communs pour toutes les métaheuristiques : Nombre d'itérations maximal ou temps d'arrêt atteint.

Le fonctionnement d'une recherche locale, dans le cas d'une minimisation, est décrit dans l'algorithme 1.

Algorithme 1

RechercheLocale()

débutGénérer une combinaison initiale $x \in X$ $x^* \leftarrow x$ **tant que** critères d'arrêt non atteints **faire**Choisir $x' \in v(x)$ **si** $f(x') < f(x^*)$ **alors** $x^* \leftarrow x'$ **fin si** $x \leftarrow x'$ **fin tant que****retourner** x^* **fin**

3.4.3 Fonctionnement de la méthode GRASP

Ce qui est tout à fait caractéristique de la méthode GRASP, c'est la phase de construction de la solution, où la solution est construite pas à pas (on construit le vecteur solution élément par élément dans la solution en cours de construction). A chaque pas de cette phase, les éléments qui peuvent être incorporés à la solution en cours de construction appelés "candidats", sont rangés relativement à une fonction gloutonne. L'évaluation de ces éléments nous conduit à la création d'une liste de candidats RCL (Restricted Candidate List) où les meilleurs éléments sont par la suite insérés dans la solution. A chaque pas, un éléments de la RCL est aléatoirement sélectionné puis incorporé à la solution en cours de construction. La liste des candidats et l'évaluation de ses éléments sont mis à jours. La taille de la liste RCL est très importante pour la méthode GRASP : si la RCL est un singleton, alors une seule solution purement aléatoire est produite ; si par contre la taille de la liste RCL n'est pas restreinte, GRASP produira donc des solutions aléatoires.

Généralement, les solutions générées dans la phase de construction ne sont pas des optima locaux. Elles peuvent alors être améliorées par le moyen de la procédure de recherche locale. La phase de recherche locale commence avec la solution construite et applique des améliorations itératives jusqu'à ce que l'optimum local soit trouvé.

Les phases de construction et d'amélioration sont répétées un certain nombre de fois

(MaxIterations dans le cas de l'algorithme 2) et la meilleure solution rencontrée durant toutes ces itérations est retournée en tant qu'optimum global.

Le principe général de la méthode est donné par l'algorithme 2 [20, 26].

Algorithme 2

GRASP ()

début

Initialiser x

$x^* \leftarrow x$

tant que critère d'arrêt non atteint **faire**

$x \leftarrow$ ProcédureGloutonneAléatoire(x)

$x \leftarrow$ RechercheLocale(x)

si x est meilleure que x^* **alors**

$x^* \leftarrow x$

fin si

fin tant que

retourner x^*

fin

3.5 GRASP pour un PLNE

Nous détaillons ci-après chaque étape de l'algorithme GRASP, relativement à l'adaptation de la méthode GRASP aux problèmes de programmation linéaire en nombres entiers, telle que présentée par Neto T. et Pedroso J.P. dans [20].

Nous rappelons la formulation d'un PLNE :

$$\begin{aligned} \text{Min } f(x) &= c^t x \\ \text{sc : } Ax &\leq b \\ x &\in \mathbb{N}^n \end{aligned}$$

Où : c , b et x sont des vecteurs de taille n ,
 A est une matrice d'ordre (m, n) .

3.5.1 Représentation des solutions

La solution déterminée par GRASP est un sous ensemble de variables entières, x dans le problème d'optimisation combinatoire précédent. La structure d'information représentant une solution utilisée par GRASP est alors un vecteur d'entiers de dimension n , $\bar{x} = (\bar{x}_1, \dots, \bar{x}_n)$.

3.5.2 Evaluation des solutions

L'idée proposée dans [20] est que l'algorithme GRASP appliqué aux problèmes de programmation linéaire en nombres entiers doit avoir la possibilité de travailler avec, à la fois des solutions réalisables et non-réalisables. Nous décrivons ci-après les stratégies utilisées à cet effet.

Solutions réalisables

Si la solution \bar{x} fixée par GRASP ne viole aucune contrainte, l'évaluation de cette solution est la valeur $z = c\bar{x}$ et la solution est marquée "réalisable".

Solutions non-réalisables

Si \bar{x} n'est pas réalisable, son évaluation est donnée par la somme des contraintes violées. Pour le problème de minimisation défini dans la section 3.1.1, l'évaluation est donnée par la formule ξ telle que :

$$\xi = \sum_{k \in E} s_k, \text{ avec } s_k = \max\{b_k - A^k \bar{x}, 0\} \quad \forall k \in E. \quad (3.1)$$

Où A^k désigne la k ème ligne de la matrice A . La solution est marquée "non réalisable".

3.5.3 Comparaison des solutions

Toutes les solutions, réalisables ou non-réalisables, doivent être comparées de la manière suivante :

- les solutions réalisables sont toujours meilleures que les solutions non-réalisables,
- les solutions réalisables sont rangées entre elles par rapport à l'évaluation de la fonction objectif,

- les solutions non-réalisables sont rangées entre elles par rapport à la mesure de la distance qui les séparent de l'ensemble des solutions réalisables.

On dit que la solution x est meilleure que la solution y si et seulement si :

- $\xi^x < \xi^y$ (x est plus proche de l'ensemble réalisable que y),
- $\xi^x = \xi^y$, et $f(x) < f(y)$ dans le cas de minimisation (x est un meilleur objectif).

3.5.4 Phase de construction

Deux méthodes de construction sont proposées dans [20], qui se distingue chacune par sa fonction gloutonne et la taille de sa liste RCL. Les deux méthodes sont basées sur la solution de la relaxation $x^R = (x_1^R, \dots, x_n^R)$. Le terme RCL tel qu'il est utilisé est légèrement différent de celui qu'on retrouve dans la littérature GRASP. A chaque indice $i = 1, \dots, n$ de variables, nous construisons une liste de valeurs qui peuvent être assignées à la solution. Dans la construction gloutonne usuelle, nous assignons toujours, à la valeur de la relaxation l'entier qui lui est le plus proche ; la RCL de la variable x_i contient alors un seul élément (l'entier le plus proche de x_i). La liste RCL dans le cas d'une construction semi-gloutonne contient plus d'un élément comme le montre les deux méthodes de construction que nous développons ci-après.

L'algorithme 3 présente l'algorithme d'une construction semi-gloutonne.

Algorithme 3

SemiGloutonne (\bar{x}^R)

début

pour $i = 1$ à n **faire**

 RCL = {valeurs autorisées à \bar{x}_i }

 tirer un $r \in$ RCL avec une certaine probabilité

$\bar{x}_i \leftarrow r$

fin pour

retourner \bar{x}

fin

Construction avec arrondissement probabiliste

Cette construction semi-gloutonne est inspirée de l'algorithme proposé par Lengauer dans [18]. La méthode consiste à arrondir chaque variable \bar{x}_i à l'entier le plus proche de

la solution de la relaxation x_i^R , d'une façon aléatoire avec une certaine probabilité. Les probabilités d'arrondissement pour chaque $i = 1, \dots, n$ à l'entier le plus proche de x_i^R sont données par :

$$P(\bar{x}_i = [x_i^R]) = [x_i^R] + 1 - x_i^R,$$

$$P(\bar{x}_i = [x_i^R] + 1) = x_i^R - [x_i^R].$$

La liste RCL est alors construite avec les deux valeurs que \bar{x}_i peut prendre.

Construction bi-triangulaire

Le but de cette stratégie est d'augmenter la diversité des solutions. Dans la construction précédente, nous n'avons que deux possibilités pour arrondir la valeur de la relaxation. La construction bi-triangulaire permet d'assigner à chaque variable x_i , n'importe quelle valeur qui se situe entre sa valeur minimale l_i et sa valeur maximale u_i . La liste RCL est alors construite avec ces valeurs et à chacune de ces valeurs est assignée une probabilité basée sur la solution de la relaxation x_i^R .

La densité de probabilité considérée est composée de deux triangles et est définie par trois paramètres : le minimum $a = l_i - 0.5$, le maximum $b = u_i + 0.5$ et la moyenne x_i^R . Les valeurs a et b ont été considérées de telle manière à avoir la possibilité d'arrondir vers l_i et u_i . La densité bi-triangulaire est représentée dans la figure 3.3.

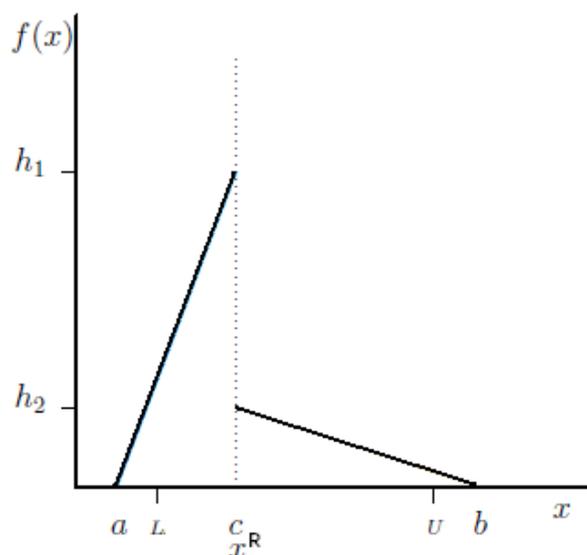


FIGURE 3.3 – Fonction de densité bi-triangulaire

La densité de probabilité est donnée par :

$$f(x) = \begin{cases} \frac{2(b-c)(x-a)}{(b-a)(c-a)^2} & \text{si } a \leq x \leq c, \\ \frac{2(c-a)(b-x)}{(b-a)(b-c)^2} & \text{si } c < x \leq b, \\ 0 & \text{sinon.} \end{cases}$$

et la fonction de répartition est :

$$F(x) = \begin{cases} 0 & \text{si } x < a, \\ \frac{(b-c)(x-a)^2}{(b-a)(c-a)^2} & \text{si } a \leq x \leq c, \\ 1 - \frac{(c-a)(b-x)^2}{(b-a)(b-c)^2} & \text{si } c < x \leq b, \\ 1 & \text{si } x > b. \end{cases}$$

La moyenne de cette distribution est c , qui correspond à la valeur de la relaxation. Avec cette construction, la valeur de chaque variable x_i est obtenue en tirant un nombre aléatoire suivant cette distribution avec $a = l_i - 0.5$, $b = u_i + 0.5$ et $c = x_i^R$, puis en arrondissant ce nombre à l'entier le plus proche de lui.

3.5.5 Recherche locale

La recherche locale essaie d'améliorer la solution en recherchant dans son voisinage. Pour cela, il a été considéré des voisinages où les variables sont augmentées ou diminuées séparément ou simultanément.

Le voisinage de la solution x , noté $N^1(x)$ est composé des solutions qui diffèrent de x en une seule composante x_j dont la valeur diffère d'une unité en plus ou en moins. On dit alors que y est un voisin de x si pour un seul indice i , $y_i = x_i + 1$, ou $y_i = x_i - 1$, avec $y_j = x_j$ pour tous les indices $i \neq j$.

L'idée utilisée pour la construction de ce voisinage peut être élargie, on peut appliquer le changement sur deux composantes ou plus à la fois. De manière générale, on peut déterminer l'ensemble de voisinage à l'ordre k (avec $n > k$) par :

$$N^k(x) = \{y \in \mathbb{N}^n, y \text{ est obtenu en modifiant } k \text{ composantes de } x\}.$$

La méthode de recherche locale implémentée est représentée dans l'algorithme 3, où k_{max} représente le nombre de composantes à changer. Ce nombre, déterminé par l'utilisateur, ne dépasse généralement pas 2.

Algorithme 3

début $k \leftarrow 1$ **tant que** $k \geq k_{max}$ $S \leftarrow N^k(x)$ **tant que** $S \neq \{\}$ $s \leftarrow \text{ChoixAléatoire}(S)$ **si** s est meilleur que x Renvoyer s **fin si** $S = S \setminus \{s\}$ $k \leftarrow k + 1$ **fin tant que**Renvoyer x **fin**

Conclusion

Ce chapitre a été consacré aux problèmes d'optimisation combinatoire et aux méthodes de résolutions. Nous avons examiné la littérature disponible sur les métaheuristiques, dont des algorithmes expliquant le principe d'une fonction gloutonne, d'une recherche locale puis nous nous sommes intéressés particulièrement la méthode GRASP.

4

Analyse Statistique des Données

Introduction

Ce chapitre est consacré à l'étude statistique des données relative à l'expérimentation menée au niveau du domaine agricole HADJADJ. Pour se faire, nous utiliserons les outils statistiques décrits au deuxième chapitre.

4.1 Dispositif expérimental et description des données

Il s'agit d'une expérimentation en blocs aléatoires complets qui a pour objectif de déterminer la dose et la marque d'engrais phosphaté qui maximiseraient le rendement.

Les trois blocs sont partagés en six parcelles chacun, où chaque parcelle a été fertilisée par une combinaison de dose d'engrais associé à une marque. Nous disposons donc de trois répétitions (3 blocs), soient trois observations par combinaison des deux facteurs.

Nous disposons des données suivantes :

– **La dose de phosphore utilisée** : 3 doses : D1, D2 et D3, la variable dose est

considérée comme quantitative avec trois modalités (3 doses différentes). **La dose est un facteur à effet fixe : c'est une variable contrôlée**

- **La marque de l'engrais phosphaté utilisée** : deux marques différentes PHOSFERT et FOSFACTYL -décrites en annexe- issues respectivement des entreprises PROFERT et TIMAC. Cette variable est considérée comme qualitative avec deux modalités (2 marques). **La marque est un facteur à effet fixe : c'est une variable contrôlée**
- **Le rendement** : la variable rendement est considérée comme quantitative, c'est la réponse, la variable à expliquer.

Répartition des traitements au niveau des blocs

Bloc 1	Bloc 2	Bloc 3
PHOSFERT/D1	FOSFACTYL/D2	PHOSFERT/D1
PHOSFERT/D3	FOSFACTYL/D1	PHOSFERT/D2
PHOSFERT/D2	FOSFACTYL/D3	PHOSFERT/D3
FOSFACTYL/D1	PHOSFERT/D1	FOSFACTYL/D3
FOSFACTYL/D3	PHOSFERT/D3	FOSFACTYL/D2
FOSFACTYL/D2	PHOSFERT/D2	FOSFACTYL/D1

Rendements enregistrés au niveau des blocs en quintaux/ha

	BLOC 1	BLOC 2	BLOC 3
PHOSFERT/D1	39,07	35,62	33,34
PHOSFERT/D2	39,87	35,67	46,07
PHOSFERT/D3	46,30	36,54	42,62
FOSFACTYL/D1	42,57	49,27	35,88
FOSFACTYL/D2	50,87	49,73	37,61
FOSFACTYL/D3	45,57	41,16	49,98

4.1.1 Outils d'analyse

La variable mesurée dans un tel schéma expérimental sera notée Y. Nous utilisons trois indices pour les observations :

- Le premier indice i indique la modalité i du facteur MARQUE, $i = 1, 2$;
- Le deuxième indice j indique la modalité j du facteur DOSE, $j = 1, 2, 3$;
- Le troisième indice k indique la position k de l'observation dans l'échantillon, $k = 1, 2, 3$.

Dans le cadre de l'expérimentation réalisée en bloc aléatoire complet sur le site d'El-Goléa, deux outils d'analyse peuvent être utilisés :

1. Le modèle d'ANOVA à effet mixte à trois facteurs sans répétitions : deux facteurs à effets fixes, soient les facteurs DOSE et MARQUE car contrôlés par l'expérimentateur et le facteur BLOC serait à effet aléatoire ;
2. Le modèle d'ANOVA à effet fixe à deux facteurs avec répétitions : deux facteurs à effets fixes, soient les facteurs DOSE et MARQUE car contrôlés par l'expérimentateur et le bloc sert à augmenter le nombre d'observations.

A travers ce choix de dispositif, l'expérimentateur a souhaité tester l'influence des facteurs contrôlés sur le rendement du blé dur, l'effet du bloc ne sera donc pas testé.

Le modèle d'analyse de la variance qui pourrait donc, répondre à la question de l'influence des facteurs contrôlés DOSE et MARQUE sur la variable réponse RENDEMENT, serait celui décrit par l'équation (2.1) :

$$Y_{i,j,k} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{i,j} + \epsilon_{i,j,k} / i = 1 \dots I ; j = 1 \dots J ; k = 1 \dots K$$

4.1.2 Analyse graphique de l'interaction des deux facteurs

Une analyse graphique de l'effet de l'interaction des facteurs DOSE et MARQUE sur la variable RENDEMENT est faite en traçant le graphe d'interaction entre les valeurs du rendement et chacun des facteurs du modèle. Le graphe obtenu est présenté ci-dessous.

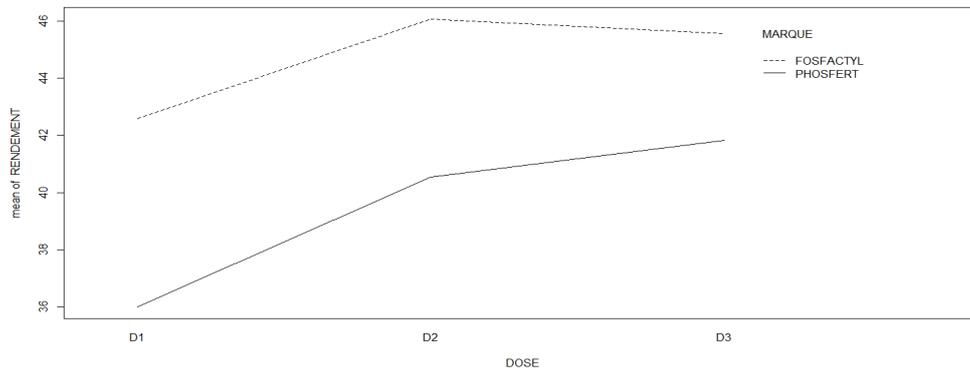


FIGURE 4.1 – Graphe d’interaction des facteurs MARQUE et DOSE

Nous constatons que les deux courbes restent parallèles, il n’y a donc pas d’effet d’interaction entre les deux facteurs. Celui-ci peut donc être négligé dans le modèle d’ANOVA à deux facteurs.

Le modèle d’ANOVA qui correspond à nos données serait un modèle additif de la forme :

$$Y_{i,j,k} = \mu + \alpha_i + \beta_j + \epsilon_{i,j,k}$$

$$i = 1, \dots, I; j = 1, \dots, J; k = 1, \dots, K$$

Ces affirmations seront confirmées par des tests statistiques effectués à l’aide du logiciel libre R, dans ce qui suit.

4.2 ANOVA à 2 facteurs avec interaction

4.2.1 ANOVA sous R

En tenant compte de l’interaction DOSE*MARQUE, le modèle testé est celui décrit par l’équation (2.1).

Les résultats renvoyés par le logiciel en tenant compte de l’interaction sont :

```
DOSE<-factor(DOSE)
MARQUE<-factor(MARQUE)
> DOSE
```

```

Levels: D1 D2 D3
> MARQUE
Levels: FOSFACTYL PHOSFERT

MODELE<-aov(RENDEMENT~DOSE*MARQUE)
> summary(MODELE)
          Df Sum Sq Mean Sq F value Pr(>F)
DOSE      2   71.3   35.64   1.199 0.3351
MARQUE    1  125.6  125.56   4.226 0.0622 .
DOSE:MARQUE 2    6.1    3.04   0.102 0.9035
Residuals 12  356.5   29.71
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Nous constatons que pour un seuil de signification $\alpha = 0.05$ chacun des tests réalisés renvoie une probabilité $Pr(> F)$ supérieure à 0.05. Nous décidons donc de :

- Ne pas rejeter l'hypothèse " \mathcal{H}_0^{DOSE} ", $Pr(> F) = 0.3351 > 0.05$
L'effet du facteur DOSE sur le RENDEMENT n'est pas significatif.
- Ne pas rejeter l'hypothèse " \mathcal{H}_0^{MARQUE} ", $Pr(> F) = 0.0622 > 0.05$
L'effet du facteur MARQUE sur le RENDEMENT n'est pas significatif.
- Ne pas rejeter l'hypothèse " $\mathcal{H}_0^{INTERACTION}$ ", $Pr(> F) = 0.9035 > 0.05$
L'effet de l'interaction DOSE*MARQUE sur le RENDEMENT n'est pas significatif.

4.2.2 Vérification des conditions d'application

Nous tentons de vérifier les conditions d'application de l'ANOVA, à savoir :

1. L'indépendance ;
2. La normalité ;
3. L'homogénéité des variances.

Indépendance

Il n'existe pas, dans un contexte général, de test statistique simple permettant d'étudier l'indépendance. Ce sont les conditions de l'expérience qui nous permettront d'affirmer que nous sommes dans le cas de l'indépendance [5].

L'échantillon dont nous disposons provient d'une expérience en blocs aléatoires complets. Les observations sont issues d'une expérience répartie de façon aléatoire au niveau de chaque bloc (3 répétitions). Donc, sous ces conditions expérimentales, c'est à dire, une répartition aléatoire des mesures, nous pouvons affirmer que les observations sont indépendantes.

Normalité

Le test de normalité se fait sur les résidus du modèle. Grâce au logiciel R, le test de Shapiro-Wilk est réalisé, afin de tester la normalité de notre échantillon. Les résultats renvoyés par le logiciel sont décrits ci-dessous.

```
> shapiro.test(residuals(MODELE))
```

```
Shapiro-Wilk normality test
```

```
data: residuals(MODELE)  
W = 0.9474, p-value = 0.3855
```

Comme la p-value = 0.3855 est supérieure à 0.05, nous décidons donc que l'hypothèse de normalité est satisfaite.

Homogénéité des variances

Le nombre de répétitions étant faible, nous ne pouvons tester l'hypothèse d'homogénéité des variances des erreurs, puisque nous ne disposons que de trois répétitions par combinaison des deux facteurs. Cependant, à titre indicatif, nous pouvons tester l'égalité des

variances des rendements selon les facteurs principaux du modèle [5].

Nous utilisons pour cela le test de Bartlett.

```
> bartlett.test(RENDEMENT~DOSE)
```

```
Bartlett test of homogeneity of variances
```

```
data: RENDEMENT by DOSE
```

```
Bartlett's K-squared = 0.4901, df = 2, p-value = 0.7827
```

```
> qchisq(.95,2)
```

```
[1] 5.991465
```

```
> bartlett.test(RENDEMENT~MARQUE)
```

```
Bartlett test of homogeneity of variances
```

```
data: RENDEMENT by MARQUE
```

```
Bartlett's K-squared = 0.2748, df = 1, p-value = 0.6002
```

```
> qchisq(.95,1)
```

```
[1] 3.841459
```

Pour le premier test la valeur lue sur la table de χ^2 avec un DDL = 2 et un seuil de signification $\alpha = 0.05$ est = 5.991465, celle-ci est supérieure à la statistique $K^2 = 0.4901$ renvoyée par le test de Bartlett.

Nous décidons donc de ne pas rejeter l'hypothèse d'homocédasticité des variances en fonction du facteur DOSE.

Pour le second test la valeur lue sur la table de χ^2 avec un DDL = 1 et un seuil de signification $\alpha = 0.05$ est = 3.841459, celle-ci est supérieure à la statistique $K^2 = 0.2748$ renvoyée par le test de Bartlett.

Nous décidons donc de ne pas rejeter l'hypothèse d'homocédasticité des variances en fonction du facteur MARQUE.

Cependant ces résultats ne nous garantissent pas l'égalité des variances théoriques mais sont de bons indicateurs pour l'homoscédasticité [5].

4.2.3 Synthèse

Les conditions d'application de l'ANOVA à 2 facteurs avec interaction étant vérifiées, à partir des données expérimentales testées sous R, nous concluons que :

- Le rendement observé ne varie pas significativement en fonction des doses D1, D2, D3 de phosphore utilisées lors de l'expérimentation.
- Le rendement observé ne varie pas significativement en fonction des marques de fertilisant utilisées lors de l'expérimentation.
- Le rendement observé ne varie pas significativement en fonction de l'interaction entre les deux facteurs MARQUE et DOSE. L'effet de l'interaction est donc rejeté, nous décidons de tester le modèle additif dans ce qui suit.

4.3 ANOVA à 2 facteurs sans interaction

4.3.1 ANOVA sous R

Nous avons montré que l'effet de l'interaction des facteurs DOSE et MARQUE pouvait être supprimé du modèle de l'ANOVA, nous testons donc le modèle additif suivant :

$$Y_{i,j,k} = \mu + \alpha_i + \beta_j + \epsilon_{i,j,k}$$

$$i = 1 \dots I, j = 1 \dots J, k = 1 \dots K$$

Les résultats renvoyés par le logiciel en ne tenant pas compte de l'interaction sont :

```
> MODELE<-aov(RENDEMENT~DOSE+MARQUE)
> summary(MODELE)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
DOSE	2	71.3	35.64	1.376	0.285
MARQUE	1	125.6	125.56	4.848	0.045 *
Residuals	14	362.6	25.90		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Nous constatons que pour un seuil de signification $\alpha = 0.05$, le test de l'effet du facteur DOSE renvoie une $Pr(> F) = 0.285$ qui est supérieure à 0.05 et que le test de l'effet du facteur MARQUE renvoie une probabilité $Pr(> F) = 0.045$ qui est inférieure à 0.05. Nous décidons donc de :

- Ne pas rejeter l'hypothèse " \mathcal{H}_0^{DOSE} ", $Pr(> F) = 0.285 > 0.05$.
L'effet du facteur DOSE sur le RENDEMENT n'est pas significatif.
- Rejeter l'hypothèse " \mathcal{H}_0^{MARQUE} ", $Pr(> F) = 0.045 < 0.05$.
L'effet du facteur MARQUE est significatif sur le RENDEMENT.

4.3.2 Vérification des conditions d'application

Les conditions d'indépendance et homogénéité des variances ne changent pas pour cette modélisation, elle sont donc vérifiées. Cependant la condition de normalité des résidus doit être vérifiée, car il s'agit d'un autre modèle, donc d'autres résidus.

Normalité

Le test de normalité se fait sur les résidus du modèle. Grâce au logiciel R, le test de Shapiro-Wilk est réalisé, afin de tester la normalité de notre échantillon. Les résultats renvoyés par le logiciel sont décrits ci-dessous.

```
> shapiro.test(residuals(MODELE))
```

```
Shapiro-Wilk normality test
```

```
data: residuals(MODELE)
```

```
W = 0.9575, p-value = 0.5543
```

Comme la p-value = 0.5543 est supérieure à 0.05, nous décidons donc que l'hypothèse de normalité est satisfaite.

4.3.3 Confirmation d'absence d'effet du facteur DOSE : "test de Kruskal-Wallis"

Le test de Kruskal-Wallis, est l'équivalent de l'ANOVA à un facteur en test non paramétrique.

Afin de confirmer que l'effet du facteur DOSE n'est pas significatif sur la variable RENDEMENT, nous effectuons ce test. Le résultat obtenu est décrit ci-dessous :

```
> kruskal.test(RENDEMENT~DOSE)
```

```
Kruskal-Wallis rank sum test
```

```
data: RENDEMENT by DOSE
```

```
Kruskal-Wallis chi-squared = 2.8889, df = 2, p-value = 0.2359
```

Comme la p-value = 0.2359 est supérieure à 0.05, nous décidons donc que, l'effet du facteur DOSE n'est pas significatif sur le RENDEMENT.

4.3.4 Conclusions partielles

Les conditions d'application de l'ANOVA à 2 facteurs sans interaction étant vérifiées, grâce aux résultats renvoyés par le logiciel, nous concluons :

- Le rendement observé ne varie pas significativement en fonction des doses de phosphore utilisées lors de l'expérimentation. Nous avons confirmé ce résultat par le test non paramétrique de Kruskal Wallis.
- Le rendement observé varie significativement en fonction de la marque de fertilisant utilisée lors de l'expérimentation. Nous pourrions donc, procéder à des tests de comparaisons multiples.
- L'effet du facteur MARQUE sur le rendement ayant été démontré, un test de comparaison multiple de Newman et Keuls sera fait afin de classer les marques. Il va de soit

que du fait de n'avoir que deux marques le résultat sera deux groupes distincts. Nous le confirmons dans ce qui suit.

4.3.5 Test de Student-Newman-Keuls

Les résultats renvoyés par le logiciel R pour le test SNK sont présentés ci-dessous :

```
> library(agricolae)
Message d'avis :
le package 'agricolae' a été compilé avec la version R 2.15.0
> SNK.test(MODELE,"MARQUE", main="Effet de la marque sur le rendement")
```

Study: Effet de la marque sur le rendement

Student Newman Keuls Test
for RENDEMENT

Mean Square Error: 25.90054

MARQUE, means

	RENDEMENT	std.err	replication
FOSFACTYL	44.73778	1.892992	9
PHOSFERT	39.45556	1.562911	9

alpha: 0.05 ; Df Error: 14

Critical Range

2
5.145555

Means with the same letter are not significantly different.

Groups, Treatments and means

a	FOSFACTYL	44.7377777777778
b	PHOSFERT	39.4555555555556

Il en ressort deux groupes distincts "a" et "b". L'effet de la marque est significativement différent sur le rendement observé, avec une moyenne de 44,74 quintaux/ha pour le produit FOSFACTYL, issu de l'entreprise TIMAC et de 39,46 quintaux/ha pour le produit PHOSFERT issu de l'entreprise PROFERT.

Conclusion

Les conditions d'application de l'ANOVA étant vérifiées, grâce à l'analyse statistique des données, nous pouvons apporter les conclusions suivantes à l'expérimentateur :

- Le rendement observé ne varie pas significativement en fonction des doses (D1, D2, D3) de phosphore utilisées lors de l'expérimentation. Le test de Kruskal-Wallis confirme ce résultat.
- Le rendement observé varie significativement en fonction des marques de fertilisant utilisées lors de l'expérimentation. Le produit FOSFACTYL se place devant PHOSFERT offrant un rendement moyen de 44,74 quintaux/ha.

Ces conclusions pourront servir de support sur lequel se basera notre étude de coût de fertilisation et ce, en choisissant ces produits pour les besoins de l'entreprise. Cette étude sera présentée au chapitre suivant.

5

Etude des coûts de fertilisation

Introduction

Dans ce dernier chapitre, nous nous penchons sur la question de l'optimisation des coûts de fertilisation. Nous commençons d'abord par donner une formulation mathématique à notre problème, puis pour détourner la difficulté de notre problème de programmation d'une méthode exacte, nous proposons une résolution grâce à la méthode GRASP. Nous implémentons alors l'algorithme GRASP avec, à l'étape de construction, deux méthodes différentes proposées dans [20].

Nous établissons quelques comparaisons concernant les résultats que nous renvoie la métaheuristique GRASP, avec les solutions exactes fournies par le logiciel LP-Solve.

5.1 Modélisation du problème de coûts

Considérons de manière générale I fournisseurs sur le marché, qui offrent chacun un nombre J_i d'engrais solides, $i \in \{1, \dots, I\}$.

A chaque produit P_{ij} est associé un coût d'acquisition c_{ij} , $i \in \{1, \dots, I\}$, $j \in \{1, \dots, J_i\}$.

Dans un produit P_{ij} destiné à la fertilisation d'un hectare de terre, la concentration (exprimée en pourcentage) d'un élément fertilisant $k \in E = \{\text{Azote, Phosphate, Potassium, Soufre}\}$ est a_{ij}^k , $i \in \{1, \dots, I\}$, $j \in \{1, \dots, J_i\}$.

$Q = (q_k)$, $k \in E$, est le vecteur des quantités de chacun des k éléments fertilisants dont le sol a besoin. Les quantités sont fixées par l'agriculteur et calculées relativement à la surface que couvre le pivot en question.

Les engrais solides sont conditionnés dans des sacs de 50 kg, nous cherchons le nombre de sacs x_{ij} à prendre d'un produit P_{ij} de façon à atteindre les quantités données par le vecteur Q et ce avec un coût minimum. Le problème s'écrit alors sous forme d'un programme linéaire en nombres entiers :

$$\text{Min } \{ cx : Ax \geq Q, x \in \mathbb{N}^n \}, \quad (5.1)$$

Avec

$x = (x_{ij})$, $\forall j \in \{1, \dots, J_i\}$, $\forall i \in \{1, \dots, I\}$ vecteur solution de taille n , où les composantes correspondent aux nombres de sacs à prendre de chaque entreprise pour chaque produit, $n = \sum_{i=1}^I J_i$.

$c = (c_{ij})$, $\forall j \in \{1, \dots, J_i\}$, $\forall i \in \{1, \dots, I\}$ vecteur de coûts,

$A = (a_{ij}^k)$, $\forall j \in \{1, \dots, J_i\}$, $\forall i \in \{1, \dots, I\}$, $\forall k \in E$, matrice des concentrations.

5.2 Estimation des coûts par la métaheuristique GRASP

Afin de résoudre le problème, nous avons opté pour la méthode GRAPS appliquée aux problèmes de programmation linéaire en nombres entiers. Notre choix s'est orienté vers la méthode en raison de son efficacité démontrée sur plusieurs problèmes de programmation combinatoire. Nous avons donc implémenté l'algorithme présenté dans la section 3.4.3, avec à la phase de construction, les deux méthodes décrites dans la section 3.5.4.

Pour se faire, une base de données Access a été créée. Le but est de stocker toutes les informations concernant les fournisseurs choisis ainsi que tous les engrais solides proposés par chaque fournisseur, avec le détail de la composition des engrais en éléments fertilisants et bien sûr son prix de vente.

A l'aide de l'outil de programmation Matlab, nous récupérons les données puis appliquons la méthode GRASP au problème 5.1 de minimisation des coûts de fertilisation posé. Une application nous renvoie le nombre de sac à prendre, d'un produit et d'une entreprise donnés. Pour faciliter la lecture, l'application renvoie uniquement les quantités non nulles.

Pour chacune des deux méthodes de construction : avec arrondissement probabiliste et bi-triangular, nous avons déroulé l'algorithme pour différents nombres maximums d'itérations (MaxIterations). Le voisinage dans toute l'expérience est fixé à l'ordre $k = 2$.

Les agriculteurs expriment généralement leurs besoins en engrais pour un hectare de terre à fertiliser. Dans un algorithme itératif comme celui de GRASP, traiter le problème avec une formule de fertilisation destinée à un hectare de terre et multiplier la solution trouvée par la surface que couvre un pivot (surface du pivot variable allant jusqu'à 40 hectares) peut s'avérer inefficace. En effet, pour que l'algorithme gagne en efficacité, il serait meilleur de fournir un vecteur de quantités (\overline{Q}) représentant les besoins en éléments nutritifs pour la totalité de la surface à fertiliser. Nous tenons à vérifier ce fait et ce, en considérant deux cas de figures lors de l'implémentation :

- (a) cas de fertilisation d'un hectare de terre, avec le vecteur quantités que nous cherchons à satisfaire $Q'=(450, 270, 120, 90)$,
- (b) cas de fertilisation d'une surface couverte par un pivot de 35 hectares (atteindre les besoins $\overline{Q} = Q \times 35$).

5.2.1 Solution initiale

Pour l'application de la méthode GRASP, une solution initiale doit être fournie à l'algorithme. Un plan de fertilisation trivial, visant la satisfaction des besoins de la terre en matière de fertilisants sans se soucier du problème de coût est le suivant :

1. Cibler la plus importante quantité à satisfaire parmi les éléments : Azote, Phosphore, Potassium et Soufre,
2. Chercher le produit ayant la meilleure composition en cet élément,
3. Prendre un nombre de sac nécessaire pour satisfaire cette contrainte,

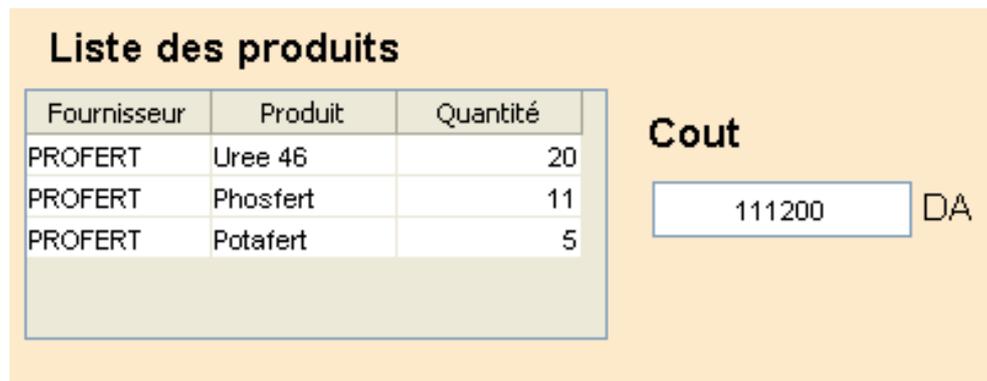


FIGURE 5.1 – Solution initiale pour la fertilisation d'un hectare

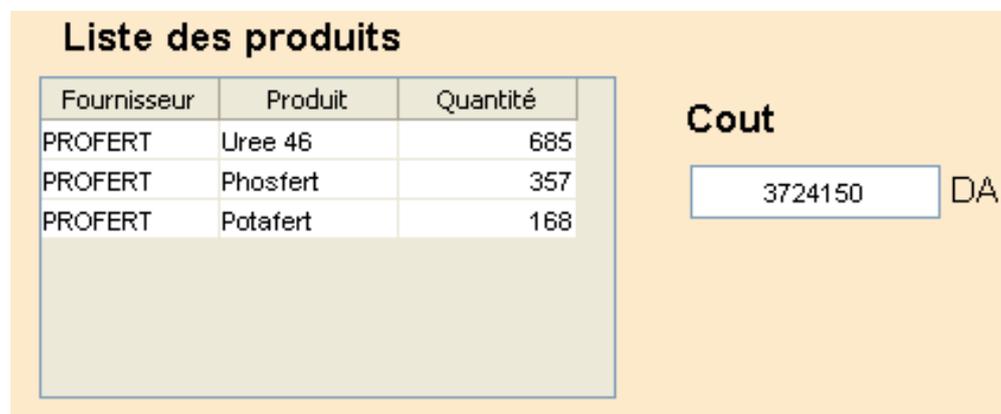


FIGURE 5.2 – Solution initiale pour un pivot de 35 hectares

4. Calculer l'apport de notre variable de décision fixée, x_{ij} en éléments non encore traités,
5. Répéter l'opération pour l'élément fertilisant suivant, jusqu'à satisfaire toutes les contraintes.

A la fin de l'opération, nous obtenons donc une solution réalisable mais pas optimale que nous soumettrons à la procédure de construction semi gloutonne, puis amélioreront par une recherche locale.

Les figures 5.3 et 5.2 représentent les solutions initiales calculées par l'algorithme dans le cas de fertilisation d'un hectare et de 35 hectares de terre, ainsi que leurs coûts respectifs.

5.2.2 Implémentation de GRASP

En vue de mesurer l'efficacité de la méthode GRASP, nous réunissons dans le tableau suivant les résultats renvoyés par un solveur disponible en ligne LP-Solve, basé sur la méthode exacte Branch-Bound. Ces résultats nous seront utiles lors de la comparaison avec la métaheuristique. Précisons que le temps d'exécution du solveur est très réduit.

	1 ha	35 ha
Sans conditions de produit	83 300 DA	2 875 750 DA
Utilisation de Phosfert	100 250 DA	3 387 240 DA
Utilisation de Fosfactyl	142 050 DA	4 911 000 DA

FIGURE 5.3 – Résultats renvoyés par le solveur LP-Solve

Phase de construction : Construction avec arrondissement probabiliste

Cas d'un hectare de terre

MaxIterations	Coût (DA)	Temps machine (secondes)
10	98 300	18.286 558
50	98 500	276.733 167
100	98 500	630.238 537

Cas d'un pivot de 35 hectares de terre

MaxIterations	Coût (DA)	Temps machine (secondes)
100	3 475 550	24.196 138
200	3 460 550	352.042 240
300	3 445 400	646.574 351

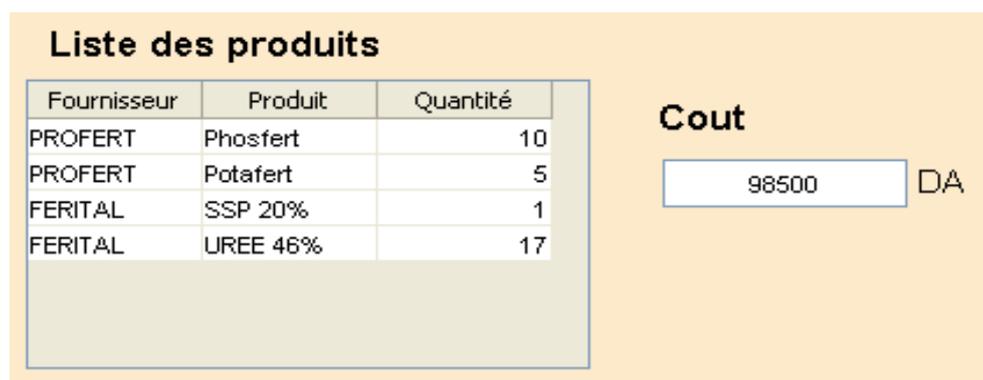


FIGURE 5.4 – Solution finale en cas de fertilisation d'un hectare de terre

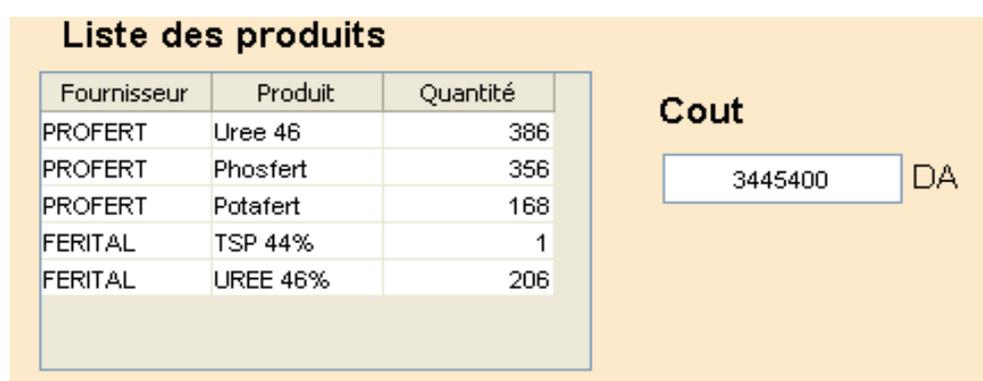


FIGURE 5.5 – Solution finale en cas fertilisation de 35 hectares

Construction bi-triangulaire

Cas d'un hectare de terre

MaxIterations	Coût (DA)	Temps machine (secondes)
50	104 500	0.461 314
100	100 300	3.051 632
500	96 540	39.248 343
1 000	94 440	55.045 281
10 000	94 440	1203.725 844

Cas d'un pivot de 35 hectares de terre

MaxIterations	Coût (DA)	Temps machine (secondes)
3000	3 343 900	153.953
5000	3 486 700	211.199
10 000	3 488 650	331.80

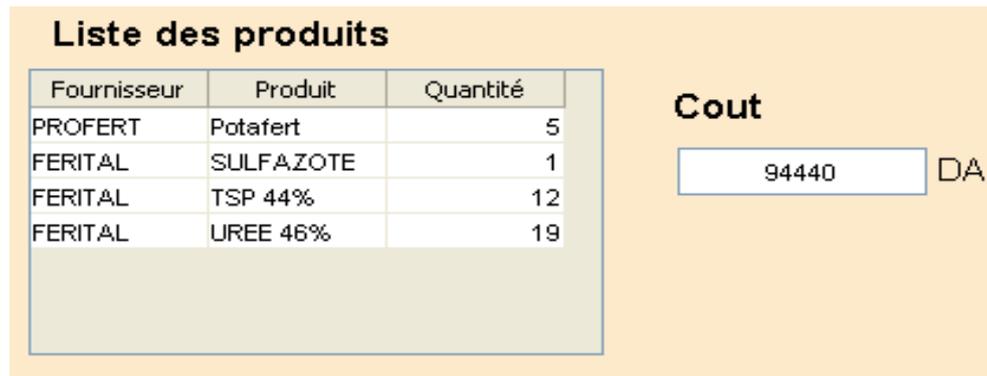


FIGURE 5.6 – Solution finale en cas de fertilisation d'un hectare

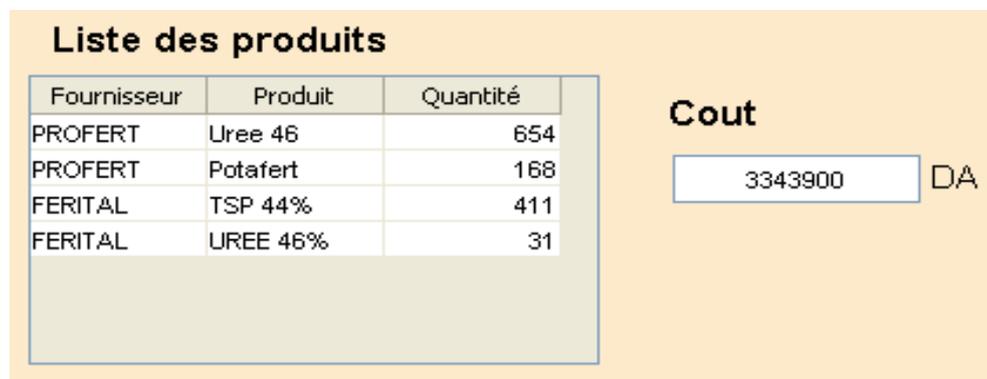


FIGURE 5.7 – Solution finale en cas fertilisation de 35 hectares

5.3 Discussion des résultats

5.3.1 Méthodes de construction

En comparant les résultats renvoyés par GRASP à ceux trouvés par la méthode exacte, il est révélé qu'aucune des deux constructions ne permet d'atteindre l'optimum de manière exacte. Ce qui confirme la nature de la solution de la métaheuristique (solution approchée). Les méthodes s'approchent de l'optimum en un temps qui, pour le même nombre d'itérations peut être variable d'un essai à l'autre. Selon la solution construite à cette phase de l'algorithme, la phase de recherche locale met un temps variable à l'exploration du voisinage. Une différence importante en temps d'exécution et en nombre maximal d'itérations séparent les deux méthodes :

- La méthode avec arrondissement probabiliste s'approche de la solution optimale, à partir des premières itérations ; mais prend un temps d'exécution très important pour l'amélioration. En effet, la liste RCL se retrouve vide à chaque itération, puisque la solution initiale choisie est à composantes naturelles. La fonction de l'algorithme se réduit dans ce cas à une simple procédure de recherche locale.
- La méthode bi-triangulaire prend quant à elle un temps moindre, en comparaison à la première méthode. Néanmoins, cette construction nécessite beaucoup d'itérations pour s'approcher de l'optimum en raison de son caractère aléatoire. D'autant plus que la table de hachage utilisée, mémorise les solutions déjà parcourues, la méthode explore donc, à chaque fois de nouvelles pistes.

5.3.2 Solutions

Dans tous les essais effectués avec l'algorithme GRASP, les produits qui apparaissent à l'optimum sont ceux fournis par PROFERT et FERTIAL. Cela voudrait dire que l'étude montre une différence de coûts significative entre les trois fournisseurs, en faveur des deux marques citées.

Les prix proposés par l'entreprise TIMAC sont relativement élevés, surtout par rapport aux prix proposés par FERTIAL, mais aussi, la composition de ses produits en éléments nutritifs fait que les produits TIMAC sont exclus des solutions trouvées, une simple observation du rapport quantité d'éléments nutritifs, prix induit une telle conclusion. Soulignons que, cette cherté des produits peut être due à une qualité supérieure.

5.3.3 Coûts de fertilisation pour les solutions trouvées

Le tableau suivant permet de comparer les coûts associés aux solutions calculées avec comme besoins de fertilisation :

1. le vecteur Q , une fois trouvée, la solution est ensuite multipliée par la surface du pivot qui est de 35 hectares),
2. le vecteur $\bar{Q} = Q \times 35$ comme donnée initiale à l'algorithme GRASP.

Nous avons relevé dans chaque cas les solutions :

1. initiale,
2. finale, avec la construction à arrondissement probabiliste,
3. finale, avec la construction bi-triangulaire.

	Coût solution avec Q (DA)	Coût $\times 35$	Coût sol. avec \bar{Q} (DA)
Solution initiale	111 200	3 892 000	3 724 150
Solution finale (Arrond. Probabiliste)	98 500	3 447 500	3 445 400
Solution finale (bi-triangulaire)	94 440	3 305 400	3 343 900

Synthèse

Le tableau montre bien qu'il est plus avantageux de dérouler l'algorithme, non pas avec des besoins pour un hectare, mais en offrant au départ à l'algorithme la donnée de la formule de fertilisation pour toute la surface à fertiliser. Néanmoins, la situation est renversée dans le cas de la construction bi-triangulaire (le coût avec \bar{Q} est inférieur au coût avec Q , multiplié à la fin par la taille du pivot). Cela s'explique par le fait que lorsque Q contient de grandes valeurs, l'intervalle de recherche délimité par a et b (phase de construction dans la section 3.5) s'élargit considérablement. L'ensemble des solutions que l'algorithme GRASP doit parcourir et tester à chaque fois s'accroît lui aussi. Il est alors impératif d'autoriser un nombre tout aussi considérable d'itérations afin que l'algorithme puisse s'approcher du mieux de l'optimum, ce qui peut devenir incommode vu le temps d'exécution.

5.4 Application sur les résultats statistiques

Nous avons conclu dans le chapitre concernant l'analyse statistique des données que le produit FOSFACTYL proposé par l'entreprise TIMAC, surclasse PHOSFERT proposé par l'entreprise PROFERT en rendement ; offrant ainsi un rendement moyen de 44,74 quintaux/ha contre 39,46 quintaux/ha pour PHOSFERT, soit 5,28 quintaux/ha de différence. Dans cette section, nous allons essayer de classer ces deux produits par rapport aux coûts de fertilisation qu'ils engendrent.

Dans tout ce qui suit, nous raisonnons par rapport à la fertilisation d'un hectare de terre. Comme nous l'avons déjà démontré, il est plus avantageux pour un algorithme de travailler par rapport à la surface globale à fertiliser. Nos résultats calculés pour un hectare puis généralisés pour la surface globale sont donc considérés comme un gain minimum garanti.

Sur la base d'une formule de fertilisation pour un hectare de terre, $Q = (450, 270, 120, 90)$, nous calculons alternativement, le coût de fertilisation associés à l'utilisation des produits FOSFACTYL et PHOSFERT pour satisfaire la quantité de phosphore requise. Par la suite, à l'aide de l'algorithme GRASP, nous déterminons les quantités optimales des trois autres éléments fertilisants à savoir l'azote, le potassium et le soufre de manière à compléter les besoins exprimés dans le vecteur Q .

5.4.1 FOSFACTYL

Pour satisfaire la quantité de phosphore exprimée par la formule Q , nous prenons 25 sacs (de 50 kg chacun) du produit FOSFACTYL. Grâce à la métaheuristique GRASP, la liste des produits à prendre pour compléter la formule est déterminée dans la figure 5.8.

Liste des produits			Coût
Fournisseur	Produit	Quantité	
PROFERT	Potafert	5	62500 DA
FERITAL	UREE 46%	18	

FIGURE 5.8 – Liste complémentaire à FOSFACTYL

Le coût total de la fertilisation en utilisant FOSFACTYL pour les besoins en phosphore est de 142 500 DA par hectare.

5.4.2 PHOSFERT

Nous allons satisfaire cette fois la quantité de phosphore avec le produit PHOSFERT, il nous faudra donc 11 sacs de ce produit et la liste pour compléter les besoins est donnée dans la figure 5.9.

Liste des produits			Cout
Fournisseur	Produit	Quantité	
PROFERT	Potafert	5	60000 DA
FERITAL	UREE 46%	17	

FIGURE 5.9 – Liste complémentaire à PHOSFERT

Le coût total de la fertilisation en utilisant du PHOSFERT est de 100 700 DA par hectare.

5.4.3 Comparaisons

Contrairement à ce qui était pour le rendement, le produit PHOSFERT de PROFERT surclasse largement le produit FOSFACTYL de TIMAC quant au coût de fertilisation, avec 41 800 DA/ha de plus pour FOSFACTYL.

À première vue, il n'est pas possible de classer les deux produits tenant compte des critères rendement et coût réunis. Pour pouvoir le faire, nous calculons le gain que se ferait l'agriculteur en utilisant chacun des produits.

Considérons que le coût total de l'emblavement d'un hectare de blé C_T qui est égal au coût de fertilisation C_{Fert} , additionné aux charges de l'entreprise agricole F , exprimés pour un hectare.

$$C_T = C_{Fert} + F$$

Sachant que l'Etat rachète le blé aux agriculteurs à hauteur d'environ 4500 DA le quintal, le profit de l'agriculteur ainsi que tous les chiffres relatifs à la situation sont synthétisés

dans le tableau ci-dessous.

	FOSFACTYL	PHOSFERT
Rendement (qtx/ha)	44.75	39.46
Prix de vente (DA)	201 375	177 570
Coût fertilisation (DA)	142 500	100 700
Profit (DA)	58 875 - F	76 870 - F

Le tableau montre que le prix de vente P du blé est supérieur au coût de fertilisation C_{Fert} . Ainsi, pour une valeur des charges F calculée pour chacun des cas (fertilisation avec FOSFACTYL ou PHOSFERT), telle que $F < P - C_{Fert}$, l'agriculteur dégage un bénéfice à l'utilisation de chacun des deux produits.

On remarque que le profit d'un agriculteur à l'utilisation de FOSFACTYL est inférieur à son profit suite à l'utilisation de PHOSFERT, $P_{FOSFACTYL} < P_{PHOSFERT}$, donc si l'agriculteur attribue un poids important à son gain, il serait alors plus intéressant pour lui de fertiliser avec du PHOSFERT, ce qui classerait le produit en premier. En revanche, si le but est de maximiser le rendement du blé, l'agriculteur doit opter pour FOSFACTYL avec un coût de fertilisation supplémentaire de 41 800 DA/hectare.

Conclusion

Nous avons consacré cette dernière partie à l'étude du coût de fertilisation du blé, face à l'offre variée d'engrais disponibles sur le marché national.

Après avoir exprimé le problème sous forme d'un programme linéaire en nombres entiers, nous avons proposé sa résolution avec la métaheuristique GRASP. Une comparaison aux résultats renvoyés par le solveur LP-Solve nous a permis de vérifier et d'affirmer que les solutions renvoyées par la métaheuristiques GRASP sont des solutions approchées mais performantes, notamment quand la liste RCL contient suffisamment d'éléments comme dans le cas de la construction bi-triangulaire.

Conclusion et Perspectives

Face à l'objectif d'amélioration des rendements de blé, nous avons exploité des données concernant le rendement de blé dur, pour élaborer un modèle statistique expliquant l'effet de deux engrais phosphatés sur le rendement, testés à trois doses chacun.

Par la suite, en réponse à la question d'optimisation des coûts de fertilisation, nous avons modélisé le problème sous forme d'un programme linéaire en nombres entiers que nous avons résolu grâce à la métaheuristique GRASP. Le logiciel LP-Solve nous a guidé dans l'évaluation de la performance de la métaheuristique choisie. Nous avons établi un lien entre l'effet de la marque de l'engrais sur le rendement et le coût de fertilisation associé à chacune des marques.

Il est important de souligner que plusieurs facteurs peuvent influencer le rendement de blé. Nous pouvons citer les doses d'azote, l'irrigation, la date de semis, la quantité de semis, plusieurs travaux en font référence dans la littérature agronomique.

Néanmoins, dans le cadre de notre étude, nous avons souhaité étudier l'effet des facteurs dose de phosphore (D1, D2, D3) et marque d'engrais phosphaté (PHOSFERT, FOSFACTYL). Nos conclusions ne sont donc pas une généralisation, mais relatives à l'expérimentation qui a été menée.

Plusieurs conclusions ont été tirées, notamment, à propos de l'effet de l'engrais sur le rendement et sur le coût de fertilisation. En effet, l'analyse statistique a montré une variation significative du rendement en fonction des marques de fertilisants utilisées lors de l'expérimentation. Le produit FOSFACTYL proposé par l'entreprise TIMAC se place devant PHOSFERT proposé par PROFERT, offrant un rendement moyen de 44.74 quintaux/ha, contre seulement 39.46 quintaux/ha.

L'analyse de coûts quant à elle, place le produit PHOSFERT en tête avec une différence

de coûts de fertilisation de 41 800 DA/ha et ce pour les besoins en N,P,K et S décrits dans le cinquième chapitre. Ce coût pouvant s'élever jusqu'à 1 463 000 DA pour chaque pivot couvrant 35 hectares.

Une perte d'un montant de 629 825 DA pour la fertilisation de chaque pivot de 35 hectares, ce qui n'encourage en rien l'agriculteur à opter pour le produit FOSFACTYL qui maximise le rendement. Ainsi l'Etat pourrait intervenir, dans la mesure où le rendement est une question qui n'est pas des moindres et ce en finançant la totalité ou l'écart du coût de fertilisation, pour ces engrais dont l'efficacité sur le rendement est avérée, tel que pour FOSFACTYL.

Nous avons étudié l'influence de la provenance d'un engrais solide sur le rendement, relativement à l'expérience réalisée sur deux marques différentes. Il serait de ce fait intéressant d'étudier l'interaction entre rendements et coûts, d'une manière encore plus vaste, incluant les différents produits émanant de chaque entreprise de fertilisation. L'obstacle auquel nous nous heurtons est que le domaine de l'expérience dépasse le cadre de la recherche opérationnelle, l'intervention d'agronomes dans ce cas serait indispensable, en vue de renforcer et d'élargir notre étude.

D'autres perspectives de recherche, concernent cette fois le domaine de recherche opérationnelle peuvent s'orienter dans le sens suivant :

- Considérer dans l'étape de modélisation, le coût de transport de chaque produit.
- Améliorer la fonction de construction de la méthode GRASP. Par exemple, dans le cas de prise en compte du coût de transport, affecter une grande probabilité aux valeurs que peut prendre une variable de décision correspondant aux capacités de transport. Autrement dit, il serait plus probable d'affecter à la variable de décision une quantité qui coïncide avec, au moins la capacité minimale du moyen de transport.
- Une autre perspective, relativement à la théorie des jeux émerge. En effet, face à la situation de conflit qui ressort de notre étude, entre l'agriculteur souhaitant maximiser ses gains et l'Etat souhaitant augmenter les rendements, un jeu se dégage. Avec pour joueurs, l'agriculteur ayant comme stratégies l'utilisation des engrais, dont certains offrent de hauts rendements, mais sont plus chers, le second joueur serait l'Etat, dont les stratégies seraient de financer les engrais à hauts rendement ou non. Nous pourrions donc trouver l'équilibre de ce jeu.

Bibliographie

- [1] AARTS E., AND LENSTRA J.K. *Local search in combinatorial optimization*. Princeton University Press, 1997.
- [2] ACIA. *Circulaires à la profession : Données relatives à l'efficacité nécessaires pour les engrais et les suppléments réglementés en vertu de la loi sur les engrais*.
- [3] ADJABI S. *Statistique Mathématique*. Editions Universitaires Européennes, 2011.
- [4] AZAIS J.M., AND BARDET J.M. *Le modèle linéaire par l'exemple*. Paris, ,2005.
- [5] BERTRAND F., AND MAUMY M. *Outils élémentaires de statistique appliquée*. IRMA, Université de Strasbourg, Septembre 2010.
- [6] BESSE P. Ressources pédagogiques. [http : //wikistat.fr/](http://wikistat.fr/), Consulté le 12-06-2012.
- [7] BRAUNE P. Qualité de blé dur. . *Perspectives agricoles N°188*, (1996), pp. 30–34.
- [8] DEMASSEY S. *Méthodes hybrides de Programmation par Contraintes et Programmation Linéaire pou le Problème d'Ordonnancement de Projet à Contraintes de Ressources*. PhD thesis, Laboratoire Informatique d'Avignon, 2003.
- [9] DODJE Y., AND ROUSSON V. *Analyse de la régression appliquée*. Dunod, Paris, 2004.
- [10] DREO J., PETROWSKI A., SIARRY P., AND TAILLARD E. *Métaheuristiques pour l'optimisation difficile*. Editions Eyrolles, 2005.
- [11] FEO T.A., AND RESENDE M.G.C. A probabilistic heuristic for a computationally set covering problem. *Operations Research Letters* (Avril 1989), pp. 67–71.
- [12] GHOUAR W. Effet du cumul de pluie hivernale sur la réponse du cultivar waha à la fertilisation azotée. Master's thesis, Université de Batna, 2006.
- [13] HALILAT M.T. Effect of potash and nitrogen fertilization on wheat under saharan conditions. Workshop on potassium and fertigation developpement in West Asia North, Morocco 24-28 November, p. 16.
- [14] HANSAN P., AND MLADENOVIC N. Variable neighborhood search : Principles and applications. *European Journal of Operational Research 130* (2001), pp. 449–467.

- [15] HAO J. K., GALINIER P., AND HABIB M. Métaheuristiques pour l'optimisation combinatoire et l'affectation sous contraintes. *Revue d'Intelligence Artificielle Vol 13*, N° 2 (1999), pp. 283–324.
- [16] HOOKER J. N. *Logic-based methods for optimization : Combining Optimization and Constraint Satisfaction*. Wiley, New-York, 2000.
- [17] IDRISSE AOUAD M. *Conception d'Algorithmes Hybrides pour l'Optimisation de l'Energie Mémoire dans les Systèmes Embarqués et de Fonctions Multimodales*. PhD thesis, UFR STMIA, Université de Nancy, 2011.
- [18] LENGAUER T. Combinatorial algorithms for integrated circuit layout. In *Applicable Theory in Computer Science*. John Wiley & Sons, 1990, ch. 8, pp. 427–446.
- [19] LOUE A. Le potassium et les céréales. In *Revue de la potasse*, vol. 9, N°4. 1993, pp. 1–19.
- [20] NETO T., AND PEDROSO J. P. Grasp for linear integer programming. In *4th Metaheuristics International Conference, Faculdade de Ciência da Universidade do Porto, Portugal* (2001).
- [21] PAPADIMITROU C. H., AND STEIGLITZ K. *Combinatorial optimisation - algorithms and complexity*. Prentice Hall, 1982.
- [22] PITSOULIS L. S., AND RESENDE M. G. C. *Handbook of Applied Optimization*. Oxford University Press, 2002.
- [23] PROFERT. [http : //www.profert - dz.com/cerea - phos.html](http://www.profert-dz.com/cerea-phos.html), Consulté le 06-06-2012.
- [24] PROFERT. [http : //www.profert - dz.com/engrais_solides.html](http://www.profert - dz.com/engrais_solides.html), Consulté le 06-06-2012.
- [25] REEVES C. R. *Modern heuristic techniques for combinatorial problems*. Blackwell Scientific Publications, Oxford, 1993.
- [26] RESENDE M. G. C., AND RIBEIRO C. *State of the art Handbook in Metaheuristics*. Kluwer Academic Publishers, 2001.
- [27] RESENDE M. G. C., AND RIBEIRO C. *Handbook of metaheuristics*. Kluwer Academic Publishers, 2003.
- [28] RIBEIRO C. C., AND MACULAN N. Application of combinatorial optimization. *Annals of Operations Research* 50 (1994).
- [29] SELMAN B., LEVESQUE H., AND MITCHELL D. A new method for solving hard satisfiability problems. In *10th National Conference on Artificial Intelligence (AAAI), San Jose, Canada* (1992), pp. 440–446.

-
- [30] SOLNON C. *Optimisation par colonies de fourmis*. Hermes Sciences, 2008.
- [31] SOLTNER D. *Les grandes productions végétales*, 16ème ed. Phytotechnie spéciale, Paris, 1999.
- [32] YARA. [http : //www.yara.fr/fertilizer/fertilizer_facts/fertilisation_aws/index.aspx](http://www.yara.fr/fertilizer/fertilizer_facts/fertilisation_aws/index.aspx), Consulté le 12-02-2012.
- [33] ZLOCHIN M., BIRATTARI M., MEULEAU N., AND DORIGO M. Model-based search for combinatorial optimization : a critical survey. *Annals of Operations Research* 131 (2004), 373–395.

Annexe

Liste des engrais solides des entreprises de fertilisants et leurs prix

	% AZOTE	% PHOSPHORE	% POTASSIUM	% SOUFRE	PRIX TTC		
PROFERT	Uree 46	46	0	0	53 000	DA/Tonne	
	Sulfate d'Am.	21	0	0	34 000	DA/Tonne	
	TSP	0	46	0	53 000	DA/Tonne	
	Phosfert	12	53	0	74 000	DA/Tonne	
	Potafert	0	0	50	46	70 000	DA/Tonne
	Patfert	8	10	30	36	60 000	DA/Tonne
	NPK	11	15	15	35	53 000	DA/Tonne
	PK	0	20	25	0	60 000	DA/Tonne
	Multifert	15	15	15	13	55 000	DA/Tonne
	Arbofert	5	21	30	27	77 000	DA/Tonne
	Maxifert	16	16	16	0	53 500	DA/Tonne
	Azofert	21	0	0	58	39 000	DA/Tonne

FIGURE 10 – Engrais solides PROFERT

		% AZOTE	% PHOSPHORE	% POTASSIUM	% SOUFRE	PRIX TTC	
TIMAC	Bioactyl	6	8	18	10	75000	DA/Tonne
	Sulfammo	23	5	21	2	68000	DA/Tonne
	Timasprint	10	5	12	16	75000	DA/Tonne
	D-Coder	6	8	18	8	74000	DA/Tonne
	Fosfactyl	3	22	0	18	64000	DA/Tonne

FIGURE 11 – Engrais solides TIMAC

		% AZOTE	% PHOSPHORE	% POTASSIUM	% SOUFRE	PRIX TTC	
FERTIAL	PK 0.20.25	0	20	25	0	66000	DA/Tonne
	SSP 20%	0	20	0	12	30000	DA/Tonne
	Sulfazote	26	0	0	35	36800	DA/Tonne
	Sulfate d'Am.	21	0	0	0	34000	DA/Tonne
	TSP	0	46	0	0	46000	DA/Tonne
	UREE 46	46	0	0	0	50000	DA/Tonne
	NPK 15-15-15	15	15	15	0	37000	DA/Tonne

FIGURE 12 – Engrais solides FERTIAL

Résumé

Dans l'optique de renforcer la production nationale, le gouvernement algérien a lancé un vaste programme de mise en valeur des terres sahariennes pour l'intensification céréalière. Cependant, ces terres sont généralement pauvres en éléments nutritifs. La maîtrise de la fertilisation minérale est le moyen le plus efficace pour l'obtention de rendements acceptables.

Sur la base de données expérimentales, de catalogues d'engrais solides de trois entreprises et grâce aux tests d'analyse de variance et à l'implémentation d'une métaheuristique, nous répondons aux préoccupations des décideurs quant au choix d'engrais à faire entre deux produits différents, testés à trois doses, relativement aux rendements qu'ils offrent et au coût global de fertilisation.

Mots clés :

Principes d'expérimentation, Analyse de la variance (ANOVA), Optimisation Combinatoire, Métaheuristique, GRASP.

Abstract

In order to increase the national production, the government launched an extensive program of development of land for intensification Saharan cereal. However, these areas are usually low in nutrients. Control of mineral fertilizer is the most efficient way to obtain acceptable yields.

Based on experimental data, catalogs of solid fertilizers of three companies and through analysis of variance tests and implementation of a metaheuristic, we answer the concerns of decision makers in the choice of fertilizers to be made between two different products, tested at three doses, relative to the returns they offer and the overall cost of fertilization.

Keywords :

Principles of experimentation, Analysis of variance (ANOVA), Combinatorial Optimization, Metaheuristics, GRASP.