

POLYTECHNIQUE MONTRÉAL
affiliée à l'Université de Montréal

**Planification tactique du cargo aérien : comparaison entre deux formulations
en programmation linéaire**

AHMED NABIL OUAKIL
Département de mathématiques et de génie industriel

Mémoire présenté en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*
Mathématiques appliquées

Août 2019

POLYTECHNIQUE MONTRÉAL
affiliée à l'Université de Montréal

Ce mémoire intitulé :

**Planification tactique du cargo aérien : comparaison entre deux formulations
en programmation linéaire**

présenté par **Ahmed Nabil OUKIL**
en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*
a été dûment accepté par le jury d'examen constitué de :

M. EL HALLAOUI Issmaïl, Ph. D., président
M. SOUMIS François, Ph. D., membre et directeur de recherche
M. LEGRAIN Antoine, Ph. D., membre

RÉSUMÉ

L'étude menée durant cette maîtrise porte sur la planification à moyen terme, aussi appelée planification tactique, de la filiale cargo de la compagnie aérienne Air Canada. Le but de cette étude est de tester la capacité à répondre à la demande du marché cargo, et ce, en utilisant le réseau de vols d'Air Canada, mais aussi de déterminer un plan d'acheminement des marchandises à travers ce même réseau. En effet, l'une des hypothèses clés dans ce projet est que toute la demande cargo se doit de transiter à travers la capacité résiduelle des vols commerciaux de transport de passagers.

À ce niveau de planification, la demande est agrégée de façon dans le but de voir l'impact de larges variations de la demande globale du marché. En effet, cette agrégation se fait pour deux raisons majeures. La première étant que la taille des données brutes est colossale, et nécessite d'être réduite pour avoir des temps de résolution raisonnables. Concernant la deuxième raison, elle est en rapport avec l'aspect prévisionnel de la demande. Les résultats expérimentaux des travaux de recherche menés ont démontrés que le regroupement des paires Origine-Destination de la demande diminuait leurs variations. Pour ce faire, l'agrégation se fait dans un processus en deux étapes durant lequel il s'agit de regrouper des terminaux dans des hubs logistiques, et ce, à l'aide d'un modèle mathématique de groupement de terminaux, implémenté à travers un programme linéaire en nombres entiers, pour ensuite regrouper les paires Origine-Destination ensembles.

Dans la première phase de recherche qui porte sur l'agrégation des demandes, il a été remarqué que le regroupement de paires Origine-Destination introduisait de l'erreur, car de façon générale, le coût de transport de deux différentes paire Origine-Destination n'est pas forcément le même. En considérant cette erreur sur le coût, ceci a mené vers la recherche d'un bon compromis entre le coût et la variation de la demande.

L'agrégation de la demande étant faite, il a été temps d'étudier le problème d'acheminement de la demande cargo à travers le réseau de la compagnie. Deux modèles mathématiques de programmation linéaire ont été développés afin de résoudre ce problème, le but est de les confronter aux mêmes tests, et analyser les résultats obtenus pour chaque modèle afin de conclure sur leurs efficacités. Dans la première formulation, les variables de flot sont définies sur des arcs, tandis que dans la deuxième, elles sont définies sur des chemins.

Il est ressorti de cette comparaison que le modèle à base de chemin était plus prometteur sur plusieurs aspects dont la génération à priori des chemins qui permet de modéliser des contraintes compliquées, les temps de résolution, ainsi que l'espace mémoire nécessaire à la résolution. De plus, pour l'application éventuelle de cet outil d'aide à la décision, Air Canada Cargo dispose déjà d'un générateur de chemins qui prend en compte toutes les contraintes du problème.

ABSTRACT

Cargo companies, just like any other company, are striving to do better and be more profitable. In order to do so, this continuous improvement has to include strategic decisions about evaluating and adapting the network. The company has to study the optimal shipment of the forecasted demand in the network for the following season. Given a set of demands to be transported from origins to destinations and a set of flights, the objective is to deliver the freight to destination through the network as efficiently as possible. However, operations are constrained by the physical characteristics of the infrastructure and the operational policies of the airports. To this end, we describe two different formulations of the problem and their implementations.

At this level of planning, the aim is to determine an efficient allocation of physical and human resources to improve the transportation system. As we are not concerned with day-to-day operations, data is aggregated to decrease the size of the problem and also to improve the demand forecasting.

Once the demand is aggregated, it needs to be transported through the existing service network. To do so, we created two linear programming mathematical models. The main goal is to compare their efficiency. The first formulation, flow variables are defined on arcs, whereas in the second one, flow variables are defined on paths.

This study showed that path-based models are more efficient when it comes to solving large-scale problems such as this industrial case. Also, an a priori generation of paths is more convenient to formulate complex non-linear constraints. In addition to that, Air Canada Cargo has an entirely functional router that takes into consideration all the problem constraints.

TABLE DES MATIÈRES

RÉSUMÉ	iii
ABSTRACT	iv
TABLE DES MATIÈRES	v
LISTE DES TABLEAUX	vii
LISTE DES FIGURES	viii
LISTE DES ANNEXES	xi
LISTE DES SIGLES ET ABRÉVIATIONS	xii
CHAPITRE 1 INTRODUCTION	1
1.1 Contexte	1
1.2 Objectifs de recherche	3
1.3 Plan du mémoire	3
CHAPITRE 2 REVUE DE LITTÉRATURE	5
2.1 Agrégation de données	5
2.1.1 Niveaux hiérarchiques des terminaux	5
2.1.2 Problème de p-hub médian	5
2.2 Acheminement des demandes de transport dans le réseau	7
2.2.1 Différent niveaux de planification	7
2.2.2 Planification tactique	8
2.3 Positionnement et contribution	14
CHAPITRE 3 AGRÉGATION DE LA DEMANDE ET PRÉVISION DE LA DEMANDE	15
3.1 Description du problème	15
3.1.1 Données :	15
3.1.2 Étude statistique sur les terminaux	16
3.2 Agrégation de terminaux	17
3.2.1 Nombre de clusters	17
3.2.2 Modèle mathématique d'agrégation	17
3.2.3 Premiers résultats sur la variabilité de la demande et discussion	18
3.3 Agrégation des paires Origine-Destination	20
3.3.1 Sélection des OD	20
3.3.2 Regroupement des ODs	20
3.3.3 Premiers résultats sur la variabilité de la demande et discussion	22
3.3.4 Une erreur sur les coûts	25
3.3.5 Meilleur compromis entre erreur sur les coûts et variance de la demande	25

3.3.6	Résultats expérimentaux et discussion	26
CHAPITRE 4 ÉTUDE COMPARÉE DE DEUX FORMULATIONS DU PROBLÈME D'ACHEMINEMENT DES DEMANDES		
		29
4.1	Description du problème	29
4.1.1	Données	29
4.1.2	Contraintes et fonction d'optimisation	29
4.1.3	Implémentation du réseau	30
4.2	Deux formulations du problème	32
4.2.1	Formulation par arc	33
4.2.2	Formulation par chemin	33
4.2.3	Algorithmes pour la génération des chemins à priori	34
4.3	Comparaison expérimentale des modèles	37
4.3.1	Instances	37
4.3.2	Résultats sur les instances	38
CHAPITRE 5 CONCLUSION ET RECOMMANDATIONS		
		44
5.1	Synthèse des travaux	44
5.2	Limitations de la solution proposée et améliorations futures	45
ANNEXES		
		48

LISTE DES TABLEAUX

Tableau 2.1	Symboles utilisés dans le modèle d'optimisation et leurs représentations	9
Tableau 3.1	Demandes de transport	16
Tableau 3.2	Demande de transport regroupée par OD par jour	16
Tableau 4.1	Symboles utilisés dans le problème d'acheminement des demandes	32
Tableau 4.2	Résultats pour $p = 15$	39
Tableau 4.3	Résultats pour $p = 20$	39
Tableau 4.4	Résultats pour $p = 25$	40
Tableau 4.5	Résultats pour $p = 30$	40
Tableau 4.6	Résultats pour $p = 35$	41
Tableau 4.7	Résultats pour $p = 40$	41
Tableau 4.8	Résultats pour $p = 45$	42
Tableau 4.9	Résultats pour $p = 50$	42

LISTE DES FIGURES

Figure 3.1	Exemple réseau physique avec 5 terminaux	16
Figure 3.2	Matrice des distances	16
Figure 3.3	Variation de cv sur les revenus en fonction du nombre de clusters p	19
Figure 3.4	Variation de cv sur les coûts en fonction du nombre de clusters p	19
Figure 3.5	Exemple d'agrégation d'OD secondaire	20
Figure 3.6	Variation du revenu perdu par jour	21
Figure 3.7	Variation du revenu perdu en moyenne	21
Figure 3.8	Variation du coût perdu par jour	21
Figure 3.9	Variation du coût perdu en moyenne	21
Figure 3.10	Variation de cv sur les revenus en fonction du nombre de clusters p avec et sans agrégation des OD	22
Figure 3.11	Variation de cv sur les coûts en fonction du nombre de clusters p avec et sans agrégation des OD	22
Figure 3.12	Variation de cv pour le lundi avec $p = 15$	23
Figure 3.13	Variation de cv pour le lundi avec $p = 20$	23
Figure 3.14	Variation de cv pour le lundi avec $p = 25$	23
Figure 3.15	Variation de cv pour le lundi avec $p = 30$	23
Figure 3.16	Variation de cv pour le lundi avec $p = 35$	24
Figure 3.17	Variation de cv pour le lundi avec $p = 40$	24
Figure 3.18	Variation de cv pour le lundi avec $p = 45$	24
Figure 3.19	Variation de cv pour le lundi avec $p = 50$	24
Figure 3.20	Variation de cv pour le lundi sans clustering	24
Figure 3.21	Moyenne des variations sur une semaine	27
Figure 3.22	Variations pour le lundi	27
Figure 3.23	Moyenne pour le mardi	27
Figure 3.24	Variations pour le mercredi	27
Figure 3.25	Moyenne pour le jeudi	27
Figure 3.26	Variations pour le vendredi	27
Figure 3.27	Moyenne pour le samedi	28
Figure 3.28	Variations pour le dimanche	28
Figure 4.1	Réseau espace-temps : Arcs d'attente seulement	30
Figure 4.2	Réseau espace-temps : Arcs de service seulement	31
Figure 4.3	Réseau espace-temps : Arcs de service et d'attente	31
Figure 4.4	Réseau espace-temporel : Tous les arcs	31
Figure 4.5	Réseau espace-temps du problème	32
Figure 4.6	Chemins entre la source et le puits de la commodité	34
Figure A.1	Variation de cv pour le lundi avec $p = 15$	48
Figure A.2	Variation de cv pour le lundi avec $p = 20$	48
Figure A.3	Variation de cv pour le lundi avec $p = 25$	48
Figure A.4	Variation de cv pour le lundi avec $p = 30$	48

Figure A.5	Variation de cv pour le lundi avec $p = 35$	49
Figure A.6	Variation de cv pour le lundi avec $p = 40$	49
Figure A.7	Variation de cv pour le lundi avec $p = 45$	49
Figure A.8	Variation de cv pour le lundi avec $p = 50$	49
Figure A.9	Variation de cv pour le lundi sans clustering	49
Figure A.10	Variation de cv pour le mardi avec $p = 15$	49
Figure A.11	Variation de cv pour le mardi avec $p = 20$	50
Figure A.12	Variation de cv pour le mardi avec $p = 25$	50
Figure A.13	Variation de cv pour le mardi avec $p = 30$	50
Figure A.14	Variation de cv pour le mardi avec $p = 35$	50
Figure A.15	Variation de cv pour le mardi avec $p = 40$	50
Figure A.16	Variation de cv pour le mardi avec $p = 45$	50
Figure A.17	Variation de cv pour le mardi sans clustering	51
Figure A.18	Variation de cv pour le mercredi avec $p = 15$	51
Figure A.19	Variation de cv pour le mercredi avec $p = 20$	51
Figure A.20	Variation de cv pour le mercredi avec $p = 25$	51
Figure A.21	Variation de cv pour le mercredi avec $p = 30$	51
Figure A.22	Variation de cv pour le mercredi avec $p = 35$	51
Figure A.23	Variation de cv pour le mercredi avec $p = 40$	52
Figure A.24	Variation de cv pour le mercredi avec $p = 45$	52
Figure A.25	Variation de cv pour le mercredi avec $p = 50$	52
Figure A.26	Variation de cv pour le mercredi sans clustering	52
Figure A.27	Variation de cv pour le jeudi avec $p = 15$	52
Figure A.28	Variation de cv pour le jeudi avec $p = 20$	52
Figure A.29	Variation de cv pour le jeudi avec $p = 25$	53
Figure A.30	Variation de cv pour le jeudi avec $p = 30$	53
Figure A.31	Variation de cv pour le jeudi avec $p = 35$	53
Figure A.32	Variation de cv pour le jeudi avec $p = 40$	53
Figure A.33	Variation de cv pour le jeudi avec $p = 45$	53
Figure A.34	Variation de cv pour le jeudi avec $p = 50$	53
Figure A.35	Variation de cv pour le jeudi sans clustering	54
Figure A.36	Variation de cv pour le vendredi avec $p = 15$	54
Figure A.37	Variation de cv pour le vendredi avec $p = 20$	54
Figure A.38	Variation de cv pour le vendredi avec $p = 25$	54
Figure A.39	Variation de cv pour le vendredi avec $p = 30$	54
Figure A.40	Variation de cv pour le vendredi avec $p = 35$	54
Figure A.41	Variation de cv pour le vendredi avec $p = 40$	55
Figure A.42	Variation de cv pour le vendredi avec $p = 45$	55
Figure A.43	Variation de cv pour le vendredi avec $p = 50$	55
Figure A.44	Variation de cv pour le vendredi sans clustering	55
Figure A.45	Variation de cv pour le samedi avec $p = 15$	55
Figure A.46	Variation de cv pour le samedi avec $p = 20$	55
Figure A.47	Variation de cv pour le samedi avec $p = 25$	56

Figure A.48	Variation de cv pour le samedi avec $p = 30$	56
Figure A.49	Variation de cv pour le samedi avec $p = 40$	56
Figure A.50	Variation de cv pour le samedi avec $p = 50$	56
Figure A.51	Variation de cv pour le samedi sans clustering	56
Figure A.52	Variation de cv pour le dimanche avec $p = 15$	56
Figure A.53	Variation de cv pour le dimanche avec $p = 20$	57
Figure A.54	Variation de cv pour le dimanche avec $p = 25$	57
Figure A.55	Variation de cv pour le dimanche avec $p = 30$	57
Figure A.56	Variation de cv pour le dimanche avec $p = 35$	57
Figure A.57	Variation de cv pour le dimanche avec $p = 40$	57
Figure A.58	Variation de cv pour le dimanche avec $p = 45$	57
Figure A.59	Variation de cv pour le dimanche avec $p = 50$	58
Figure A.60	Variation de cv pour le dimanche sans clustering	58

LISTE DES ANNEXES

Annexe A	RÉCAPITULATIF DES GRAPHIQUES SUR LA VARIATION DE CV EN FONCTION DE p	48
----------	--	----

LISTE DES SIGLES ET ABRÉVIATIONS

OD Origine-Destination
CA Chiffre d'affaire

CHAPITRE 1 INTRODUCTION

1.1 Contexte

Le transport des fret par voie aérienne est le processus où des marchandises sont transportées dans des avions cargo spécialisés ou dans la soute à bagage des avions effectuant des vols commerciaux de transport de passagers. Ce mode de transport est considéré comme étant le plus rapide et aussi le plus sûr.

Enjeux du transport

Le transport de fret est un maillon crucial au sein des chaînes d’approvisionnement afin d’assurer un acheminement efficace et la disponibilité en temps voulu des matières premières ou aussi des produits finis T. G. CRAINIC 2003. En effet, le coût du transport dans une chaîne logistique est non négligeable, il peut constituer un pourcentage relativement important du coût global du produit dépendamment du type de marchandise. La quantité des volumes de transport ne cesse d’augmenter au fil des années, ce qui est notamment dû au fait que les producteurs et les consommateurs peuvent être à deux points géographiques complètement différents dans cette ère de mondialisation FENG, LI et SHEN 2015. Couplé à ce phénomène de libre-échange qui est en expansion rapide, le monde de l’industrie se veut aussi très exigeant en terme de délais d’approvisionnement et de livraison avec les productions de plus en plus en juste-à-temps. L’internationalisation de l’économie engendre de nouvelles réglementations commerciales qui ouvrent la possibilité à beaucoup d’industries de dé-localiser leurs activités à des pays où il est possible de réduire les coûts de main d’œuvre.

Les compagnies de transport de fret opèrent dans un marché toujours plus ouvert à l’international, et donc, de plus en plus compétitif compte tenu du grand nombre d’acteurs présents dans ce secteur. Ce qui les force à se conformer aux normes et prix du marché afin de capter le plus de clients potentiels. Par conséquent, il leur est nécessaire d’avoir une structure très efficiente et à l’affût de toute économie possible, là où la concurrence s’en va vers la guerre des prix bas, dans le but de suivre la tendance SCIENCE 2018 et délivrer une meilleure qualité de service et aussi de maximiser leurs chiffres d’affaires (CA).

Compte tenu du contexte du secteur d’activité des compagnies de transport, la gestion des ressources physiques et humaines devient très vite extrêmement complexe et nécessite une optimisation assez poussée afin de subsister, voire croître. En effet, l’introduction de certains outils d’aide à la décision semble crucial afin de tendre vers cet objectif d’amélioration des performances. Dans ce mémoire, l’outil d’aide à la décision en question se basera sur les principes de la recherche opérationnelle.

Un problème de transport en deux parties

Dans ce mémoire, le problème global de planification tactique sera décomposé en deux parties majeures. La première est en rapport avec l’aspect prévisionnel de la demande du marché tandis que la deuxième se focalise sur l’acheminement des marchandises dans le réseau.

Ce projet de recherche est issu de la collaboration entre le laboratoire du GERAD, Polytechnique Montréal, ainsi que la compagnie de fret aérien Air Canada Cargo. L’objectif de ce travail est de fournir un outil d’aide à la décision à l’entreprise partenaire dans le but de mieux répondre à la demande du marché.

La première partie de ce travail est consacrée à la prédiction de cette demande du marché. La prédiction se doit d'être faite par paire Origine-Destination (OD) des demandes, et ce, sur une semaine typique qui est considérée comme étant représentative d'une saison donnée. En effet, le but de cette partie est de pouvoir obtenir une estimation sur le trafic (masse, volume et chiffre d'affaire) à venir sur la saison en question sur chacune des OD de notre problème et pour chaque jour de notre semaine typique. D'une part, l'un des problèmes rencontrés durant le traitement des données est la forte fluctuation de la demande observée, et ce pour la plupart des paires OD. D'autre part, le volume très important des données de demandes de transport de marchandises, qui n'est pas sans conséquence sur les temps de calcul des modèles mathématiques d'acheminement de fret. Dans les faits, en se basant sur les données reçues d'Air Canada Cargo, il s'est avéré que la volatilité de la demande décroît quand le chiffre d'affaire d'une OD est relativement important. En d'autres termes, les prédictions sont beaucoup plus précises dans les aéroports qui génèrent beaucoup de revenus que pour les aéroports qui ont moins de trafic. Si bien qu'il a été remarqué que 10% des aéroports du réseau physique, à eux seuls, comptabilisaient pour à peu près 80% du chiffre d'affaire de l'entreprise, (et 20% des aéroports représente 92% du CA). Cette concentration de volume d'activité sur des endroits bien précis donnent la possibilité de pouvoir agréger des aéroports entre eux sans trop perdre d'information. En revanche, afin que cette agrégation aille dans le sens de nos intérêts, il est nécessaire qu'elle soit faite de façon à minimiser la variance de la demande, et par conséquent les prévisions sur la demande se retrouvent être plus précises.

La deuxième partie de ce travail, quant à elle, consiste à trouver des réponses aux questions suivantes : Quels sont les itinéraires optimaux que peuvent emprunter les marchandises sur le réseau de vols d'Air Canada ? Quels sont les parts du marchés de l'entreprise, et comment les augmenter ? Comment se situe l'offre d'Air Canada Cargo par rapport à la demande sur chacun de leurs marchés ? Étant donné que l'ensemble du volume cargo transite à travers la capacité résiduelle des vols commerciaux d'Air Canada, le but est d'obtenir un plan de transport des marchandises qui minimise les coûts d'acheminement, tout en veillant à ce que les contraintes physiques (capacités des vols, être au bon endroit pour emprunter un vol,...) et temporelles (respecter les dates de disponibilité et livraison au plus tard des marchandises, les temps de manutention,...) du processus de transport soient vérifiées. Par la même occasion, il est nécessaire d'analyser les informations sur les OD en sous-capacité, où la demande surpasse l'offre, afin d'entreprendre des actions de correction dans un procédé de boucle de rétroaction. Et parallèlement, il est tout aussi essentiel de se pencher sur les marchés où le contraire se passe. Être en surcapacité signifie qu'il y a de l'espace vide inutilisé, il y a des mesures commerciales à prendre afin de rentabiliser cette situation et tirer le meilleur parti de la situation.

Dans une optique de bonne préparation, il est utile de créer plusieurs scénarios de demande afin de mettre à l'épreuve le réseau d'Air Canada. Le but de ces scénarios est de faire varier la demande sur différents endroits du problème et tester par la même occasion l'ensemble des scénarios d'offres où l'on modifie, ajoute ou supprime des vols, pour voir la robustesse des solutions et aussi quelle configuration de l'offre répond le plus efficacement à la demande pour minimiser les coûts.

Vocabulaire de transport

Services : Les vols des avions cargo et des avions commerciaux pour le transport des passagers sont considérés comme des services, auxquels viennent s'ajouter les trajets de camionnage. Chaque service

est caractérisé par un point d'origine, un point de destination, une date de départ, une date d'arrivée, ainsi qu'une capacité de transport.

Commodités : La demande est définie en terme de commodités. Chaque commodité a un certain volume à transporter, un terminal d'origine, un terminal de destination, une date de mise à disposition et une date de livraison au plus tard.

Chemin : Pour chaque commodité, un chemin réalisable est un chemin qui commence au terminal d'origine de la commodité à une date postérieure à la date de mise à disposition. Ce chemin contient un ou plusieurs services, et s'achève au terminal de destination de la commodité en question avant la date limite de livraison.

Transport multimodal : Le transport multimodal utilise au moins deux modes de transport (chemins de fer, terrestre, aérien ou maritime) dans les chemins utilisés pour acheminer les commodités vers leurs destinations.

1.2 Objectifs de recherche

L'objectif global de cette recherche est la création de modules qui permettront d'utiliser des algorithmes et modèles mathématiques dans le but de résoudre le problème de planification tactique chez Air Canada Cargo. En effet, ces modules sont au nombre de deux.

Pour le premier qui est en relation avec la première partie du projet, le but est d'établir une bonne agrégation des aéroports et des paires OD de façon à avoir un problème de taille raisonnable et avec une meilleure prévision de la demande. Pour ce faire, un modèle mathématique de "clustering" est créé et est capable de regrouper des terminaux sous formes de clusters sur critères de proximité ainsi que de chiffre d'affaire généré.

En ce qui concerne la deuxième partie, l'objectif de la recherche est de créer un modèle mathématique de multiflots qui permet de résoudre un problème d'acheminement de chacun des flots de son point origine à son point destination de façon à minimiser le coût total de transport.

1.3 Plan du mémoire

Ce présent mémoire se découpe en trois chapitres. Le premier est une revue de littérature qui illustre quelques problèmes liés au cargo. Ensuite vient la première partie de l'étude du problème qui est l'étude de l'agrégation des demandes de transport. Et pour finir, la deuxième et dernière partie qui est consacrée à la résolution du problème tactique d'acheminement de la demande globale en fonction de plusieurs scénarios.

Revue de littérature

La première section de la revue de littérature est dédiée à donner un aperçu des travaux réalisés, dans le cadre des problèmes d'agrégation, dans la littérature. Il est question d'évoquer la structure des réseaux de transport, où il existe des terminaux qui font office de plate-forme de correspondance. Ce qui engendre un trafic très élevé. Il existe des modèles qui aident à trouver où les localiser, ou éventuellement comment organiser le trafic autour d'eux. La deuxième section est consacrée à l'acheminement des demandes. Cette section commence d'abord par une définition des différents niveaux de planification existant dans la littérature pour ensuite se concentrer sur le niveau pertinent pour ce problème.

Première partie

La première section de chapitre contient une description du problème général d'agrégation rencontré dans ce projet ainsi que quelques résultats statistiques sur l'état du réseau. La seconde section traite la première partie de l'agrégation qui est le regroupement de terminaux dans des *clusters*. La troisième section, quant à elle, traite des regroupements de paires OD. Finalement la dernière section regroupe les résultats obtenus dans cette partie.

Deuxième partie

Dans cette deuxième partie, le but est de développer les outils d'aide à la décision qui vont exploiter les résultats de la première partie. La première section vient décrire le problème d'acheminement des demandes rencontré chez Air Canada Cargo, une implémentation du réseau est proposée. La seconde section contient deux différentes formulations du problème qui vont être confrontées de façon expérimentale dans la dernière section de ce chapitre.

Conclusion

La conclusion, quant à elle, reprend essentiellement les résultats des deux parties précédentes dans une synthèse qui résume le travail effectué. Nous évoquons les limitations de la solution proposée pour finir sur les perspectives de recherches futures.

CHAPITRE 2 REVUE DE LITTÉRATURE

Deux points essentiels sont abordés dans cette revue de littérature. D'un côté, il est question de faire un état de l'art spécifique de la problématique d'agrégation existant dans le milieu du transport. D'un autre côté, le but est de décrire d'un point de vue général l'activité cargo dans la littérature pour ensuite décrire les opérations du milieu dans leur ensemble. Pour ce faire, il faut décrire les différentes strates où les décisions sont prises, en d'autres termes, il est d'usage de parler de différents niveaux de planification. Éventuellement, chacun de ces niveaux a ses propres spécificités, qui sont montrés dans cette revue de littérature du point de vue de la recherche opérationnelle

2.1 Agrégation de données

Cette section est consacrée à la partie qui porte sur l'agrégation des données liées aux demandes de transport. Dans un premier lieu, un exemple sur le pays chinois montre la pertinence d'une agrégation de terminaux dans le milieu du transport cargo. Par la suite, une formulation mathématique sera illustrée et expliquée.

2.1.1 Niveaux hiérarchiques des terminaux

De façon assez générale, à l'échelle des régions ou des pays, il se trouve que les systèmes de transports sont beaucoup plus développés sur certaines villes que d'autres, ce qui crée une concentration du trafic sur ces quelques points de la carte géographique. On parle assez souvent dans la littérature, de "hub cities" pour désigner les villes où le transit de marchandise est relativement plus important que le reste de la région géographique. YARU et LINA (2012) expliquent, pour l'exemple de la Chine, que les résultats de leurs analyses montrent en effet que le développement de l'activité cargo a une structure hiérarchique où les "villes-hub" forment un certain squelette sur lequel l'activité se développe et où les vols sont concentrés. Ceci engendre le fait que la taille et la capacité des terminaux se retrouvent être donc plus importantes.

En effet, une structure de transport comme celle-ci favorise les économies d'échelles. La grande partie des flux de transports sont groupés dans les hubs logistiques qui ont les capacités nécessaires pour les traiter. Par la suite, il est possible de faire la distribution de façon locale. L'inconvénient de cette méthode est que, souvent, entre deux terminaux qui ne sont pas des hubs, l'itinéraire peut s'avérer être plus long, car en plusieurs étapes.

2.1.2 Problème de p-hub médian

Dans l'article de C.-c. LIN, J.-y. LIN et CHEN (2012), une formulation générale d'un modèle de p-hub médian avec contraintes de capacité est proposé. Les points de ramassage ou de livraison sont définis comme des centres, tandis que les hubs sont les points de consolidation. Comme expliqué, auparavant, le but de cette méthodologie est d'éliminer le flux entre deux centres et de favoriser les flux entre les hubs, ce qui réduit le coût total de transport.

Le problème p-hub médian avec contraintes de capacité peut être défini de la manière suivante. Il faut sélectionner un ensemble de p hubs parmi un ensemble de candidats dans le but de minimiser le coût total, qui est la somme des coûts fixes des hubs ainsi que les coûts de transport.

Soit H l'ensemble des points candidats à devenir des hubs. Chacun des hubs $h \in H$ a une certaine capacité de consolidation qui est notée par U_h ainsi qu'un coût fixe d'établissement noté C_h . En effet, chaque centre est représenté par deux nœuds, un pour le ramassage et un autre pour la livraison représentant ainsi l'origine de la demande $o \in O$ et aussi sa destination $d \in D$. Il y a donc un ensemble dénoté $\hat{O}D$ qui regroupe toutes les paires OD du problème, et où chaque demande est déterminée par un volume q^{od} . On note finalement l'ensemble des nœuds du graphe $N = \{O \cup H \cup D\}$ qui sont donc reliés par un ensemble d'arcs orientés $ij \in A$ qui connectent tous les nœuds origine vers tous les hubs, ensuite tous les hubs vers tous les nœuds de destination, et finalement chaque hub vers tous les autres hubs sans aucun nœud intermédiaire. Chacun de ses arcs $ij \in A$ a une capacité notée U_{ij} ainsi qu'un coût d'utilisation C_{ij} . Un facteur de rabais α_{ij} pour les transports entre hubs est introduit pour prendre en compte les économies de transport.

Le problème de p-hub médian est modélisé par un problème linéaire en nombres entiers. Il existe deux types de variables de décision, $y_h, h \in H$ vaut 1 si le candidat h est désigné comme étant un hub, 0 sinon. Ensuite x_{ij}^{od} vaut 1 si l'arc $ij \in A$ est choisi pour la paire $od \in OD$. Le modèle développé est le suivant.

$$\min \sum_{h \in H} C_h y_h + \sum_{od \in \hat{O}D} \sum_{ij \in A} \alpha_{ij} C_{ij} x_{ij}^{od} q^{od} \quad (2.1)$$

sujet à :

$$\sum_{j \in N} x_{ji}^{od} - \sum_{j \in N} x_{ij}^{od} = \begin{cases} -1 & \text{si } i = o \\ 1 & \text{si } i = d \\ 0 & \text{sinon} \end{cases} \quad \forall i \in N, \forall od \in \hat{O}D \quad (2.2)$$

$$\sum_{od \in \hat{O}D} \sum_{j \in N} q^{od} x_{hj}^{od} \leq U_h y_h \quad \forall h \in H \quad (2.3)$$

$$\sum_{od \in \hat{O}D} q^{od} x_{ij}^{od} \leq U_{ij} \quad \forall ij \in A \quad (2.4)$$

$$\sum_{h \in H} y_h \leq p \quad (2.5)$$

$$x_{ij}^{od} \in \{0, 1\} \quad \forall ij \in A, \forall od \in \hat{O}D \quad (2.6)$$

$$y_h \in \{0, 1\} \quad \forall h \in H \quad (2.7)$$

La fonction objectif (2.1) minimise la somme totale des coûts fixes liés aux hubs ainsi que les coûts sur les opérations de transport entre chaque deux nœuds. Le coût de transport entre deux nœuds est par unité de volume, cette quantité reste constante et le coût de transport dépend du volume. Tandis que le transport entre deux hubs a un facteur de rabais qui vaut $\alpha_{ij} \leq 1$ pour $i, j \in H$, autrement $\alpha_{ij} = 1$.

Les contraintes (2.2) assurent une conservation du flot de son origine jusqu'à sa destination, tandis que les contraintes (2.3) sont des contraintes de capacités pour les hubs. De la même manière, les contraintes (2.4) sont des contraintes sur les capacités des arcs, ensuite vient la contrainte (2.5) qui vient assurer le fait d'avoir au plus p hubs parmi tous les candidats possibles du problème. Finalement les contraintes (2.6) et (2.7) sont des contraintes d'intégrité pour les variables de décision du problème.

2.2 Acheminement des demandes de transport dans le réseau

Dans cette section, le but est de voir les différents niveaux de décision, ainsi qu'un ensemble de problèmes de transport découlant de ces strates de planification. Ceci dans l'objectif de situer le problème traité dans la littérature et d'en ressortir les potentielles contributions.

2.2.1 Différents niveaux de planification

T. G. CRAINIC et LAPORTE (1997) ont classifié les différents niveaux de planification en trois niveaux majeurs.

Niveau stratégique

À ce niveau de planification, les décisions portent généralement sur d'importants investissements en capitaux, et mobilisent des comités de haut niveau chargés des questions de gestion de l'entreprise. Ces décisions sont assez souvent en rapport avec l'infrastructure du réseau, et traitent des problèmes du genre :

- acquisition de nouvelles ressources ;
- localisation et installation de nouveaux équipements ;
- Mise en place de nouvelles lignes (exemple de rails en chemins de fer)
- Désaffectation de certaines structures qui ne sont plus désirées.

Niveau tactique

Ce niveau est dédié à la planification à moyen terme où le but est de gérer et contrôler les ressources présentes de façon optimale afin de garantir une satisfaction des demandes clients, tout en offrant de bonnes prestations, et donc offrir une meilleure qualité de service. Le problème général de la planification tactique consiste à concevoir un réseau de service capable de répondre à la demande en s'assurant que les détails cités ci-dessus sont pris en compte. Pour ce faire, il faut :

- localiser les services offerts dans l'espace et dans le temps ;
- déterminer la fréquence des services, les ressources à affecter afin de les mobiliser ;
- chercher l'acheminement des marchandises optimal dans le réseau des services.

Le but est de réussir à obtenir un plan de vols ainsi que les acheminements optimaux afin d'avoir la bonne qualité de service pour attirer les clients mais aussi une meilleure utilisation des ressources afin de maximiser les profits de l'entreprise.

Niveau opérationnel

La planification à court terme, quant à elle, est associée aux opérations très dynamiques où les problèmes apparaissent et sont résolus de façon très rapide et fréquente. Il est possible de retrouver ce genre de problèmes :

- gestion des opérations au sein d'un terminal : de la manutention des marchandises, ordonnancement des temps de travail des grues dans un port ;
- problème de tournées de véhicules, avec ou sans fenêtres de temps.

2.2.2 Planification tactique

Le problème de transport de ce mémoire se situe de façon globale au niveau tactique. Dans la suite de cette section, quelques problèmes sur le niveau de planification tactique présents dans la littérature seront illustrés.

Problème de flot multicommodités

Étant donné un réseau déjà existant, où chaque arc est doté d'une certaine capacité, le but est de faire acheminer un ensemble de commodités à travers le réseau de la façon la plus optimale. Dans le cas où il y a une seule commodité, il s'agit d'un problème de flot classique, relativement facile à résoudre car il est polynomial. Cependant dès qu'il y a au moins deux commodités, si les variables sont entières, le problème devient *NP*-complet, ASSAD 1978. Une formulation classique du problème est donnée dans ce qui suit.

Soit $G = (N, A)$ le réseau déjà mis en place, chaque arc $(i, j) \in A$ du réseau a une capacité notée u_{ij} , ainsi qu'un coût d'utilisation par unité de demande k noté c_{ij}^k . Chacune des commodités $k \in K$ doit être transportée de son point d'origine o_k vers son point de destination d_k dans sa totalité w_k . Les variables de décision sont sur les arcs, elles sont notées x_{ij}^k et représentent la quantité de flot $k \in K$ transitant à travers l'arc $(i, j) \in A$ du réseau. Le modèle est comme suit.

$$\min \sum_{k \in K} \sum_{(i,j) \in A} c_{ij}^k x_{ij}^k \quad (2.8)$$

sujet à :

$$\sum_{j \in N} x_{ij}^k - \sum_{j \in N} x_{ji}^k = \begin{cases} w_k & \text{si } i = o_k \\ -w_k & \text{si } i = d_k \\ 0 & \text{sinon} \end{cases} \quad \forall i \in N, \forall k \in K \quad (2.9)$$

$$\sum_{k \in K} x_{ij}^k \leq u_{ij} \quad \forall (i, j) \in A \quad (2.10)$$

$$x_{ij}^k \geq 0 \quad \forall (i, j) \in A, \forall k \in K \quad (2.11)$$

La fonction objectif (2.8) minimise les coûts de transport des commodités à travers le réseau. Les contraintes (2.9) sont des contraintes de conservation de flot dans le réseau qui assurent que les commodités sont acheminées de leurs origines vers leurs destinations. Les contraintes (2.10) assurent que les capacités sur les arcs ne sont pas violées et finalement, les contraintes (2.11) sont les contraintes sur les domaines des variables.

Problème de flot multicommodités et multimodal à base de services programmés

Le transport multimodal est significativement différent du transport unimodal traditionnel qui n'emploie qu'un seul mode de transport. Le but du problème est de sélectionner les routes à coût minimal pour transporter un certain ensemble de commodités de leurs origines à leurs destinations en utilisant des services déjà mis en place. Ce problème est relativement plus sensible au temps étant donné que cette fois les commodités ont des dates de mise à disposition R_k ainsi que des dates de livraison au plus tard D_k . Par conséquent, les routes trouvées doivent être réalisables dans le sens où aucune commodité ne peut emprunter un service si

elle n'est pas à son aéroport de départ avant l'heure limite d'embarquement. Les autres contraintes sont que la commodité ne peut emprunter un service qu'après le moment de sa mise à disposition, et la livraison doit obligatoirement se faire avant la date limite.

Plusieurs articles dans la littérature se sont proposés de résoudre ce type de problème, dont SUN et LANG (2015), AYAR et YAMAN (2012) et ARCHETTI et PEIRANO (2019). Afin d'éviter d'être redondant, seule la méthode de SUN et LANG (2015) sera illustrée dans cette revue de littérature.

Afin de résoudre ce problème, SUN et LANG (ibid.) ont établi un modèle de programmation mixte non-linéaire en nombre entiers avec une méthode de linéarisation, qui a été testée sur un réseau de 78 arcs avec 42 services inspiré du réseau intérieur de transport Chinois. Les temps de résolution sont aux alentours d'une minute et demi.

Dans cette étude, $G = (N, A, S)$ représente un réseau de transport multimodale où N , A , et S représentent respectivement les terminaux physiques, les arcs orientés, et l'ensemble des services dans le réseau. Il y a trois types de variables de décision.

- X_{ij}^{ks} vaut 1 si la commodité k est transportée du terminal i au terminal j avec le service s , et vaut 0 sinon ;
- Y_i^k Date d'arrivée de la commodité k au terminal i ;
- Z_{ij}^{ks} Temps d'attente de la commodité k au terminal i avant d'être transportée sur vers j avec le service s .

Tableau 2.1 Symboles utilisés dans le modèle d'optimisation et leurs représentations

Indices	Représentations
h, i, r	Indices pour des terminaux
s, r	Indices pour des services
(i, j)	Arc orienté du terminal $i \in N$ vers le terminal $j \in N$
Services	Représentations
S_{ij}	Ensemble des services qui s'en vont du terminal $i \in N$ au terminal $j \in N$
$\delta^-(i)$	Ensemble des prédécesseurs du terminal $i \in N$
$\delta^+(i)$	Ensemble des successeurs du terminal $i \in N$
Commodités	Représentations
o_k	Terminal d'origine de la commodité $k \in K$
d_k	Terminal de destination de la commodité $k \in K$
q_k	Volume de la commodité $k \in K$
R_k	Date de mise à disposition de la commodité $k \in K$
D_k	Date de livraison au plus tard de la commodité $k \in K$
Paramètres	Représentations
l_i^s	Date de début d'embarquement pour le service $s \in S$ au départ du terminal $i \in N$
u_i^s	Date limite d'embarquement pour le service $s \in S$ au départ du terminal $i \in N$
SD_i^s	Date de départ du service $s \in S$ au départ du terminal $i \in N$
SA_i^s	Date d'arrivée du service $s \in S$ au départ du terminal $i \in N$
Q_{ij}^s	Capacité du service $s \in S$ sur l'arc $(i, j) \in A$
t_{ij}^s	Temps de transport du service $s \in S$ sur l'arc $(i, j) \in A$
C_{ij}^{ks}	Coût de transport unitaire de la commodité $k \in K$, sur l'arc $(i, j) \in A$, par le service $s \in S$
c_s	Coût unitaire de chargement et de déchargement pour le service $s \in S$
c_{store}	Coût unitaire d'inventaire
M	Un nombre positif très grand dont sa valeur n'est pas spécifiée dans l'article

Ci-dessous le modèle en question. L'ensemble des symboles utilisés dans le modèle d'optimisation ainsi que leurs représentations sont illustrés dans le Tableau 2.1.

$$\min \sum_{k \in K} \sum_{(i,j) \in A} \sum_{s \in S_{ij}} C_{ij}^{ks} X_{ij}^{ks} \quad (2.12)$$

$$+ \sum_{k \in K} \sum_{i \in N} \left(\sum_{h \in \delta^-(i)} \sum_{r \in S_{hi}} c_r q_k X_{ij}^{kr} + \sum_{j \in \delta^+(i)} \sum_{s \in S_{ij}} c_s q_k X_{ij}^{ks} \right) \quad (2.13)$$

$$+ \sum_{k \in K} \sum_{(i,j) \in A} \sum_{s \in S_{ij}} c_{store} q_k Z_{ij}^{ks} \quad (2.14)$$

Dans cette formulation, la fonction objectif minimise le coût total d'acheminement des différentes commodités du problème à travers le réseau. Ce coût total englobe la somme des coûts d'utilisation des services (2.12), à laquelle se rajoute les coûts liés au chargement et déchargement des commodités (2.13), et finalement le coût d'entreposage dans les terminaux (2.14).

$$\sum_{j \in \delta^+(i)} \sum_{s \in S_{ij}} X_{ij}^{ks} - \sum_{h \in \delta^-(i)} \sum_{r \in S_{hi}} X_{hi}^{kr} = \begin{cases} 1 & \text{si } i = o_k \\ -1 & \text{si } i = d_k \\ 0 & \text{sinon} \end{cases} \quad \forall k \in K, \forall i \in N \quad (2.15)$$

$$\sum_{s \in S_{ij}} X_{ij}^{ks} \leq 1 \quad \forall k \in K, \forall (i,j) \in A \quad (2.16)$$

Les contraintes (2.15) sont des contraintes de conservation de flot, tandis que les contraintes (2.16) assurent qu'il y a au plus un service utilisé pour transporter une commodité sur un arc. Regrouper les deux contraintes ensemble assure que chaque commodité reste indivisible et suit un seul chemin de son terminal d'origine jusqu'à celui de sa destination.

$$\sum_{k \in K} q_k \times X_{ij}^{ks} \leq Q_{ij}^s \quad \forall (i,j) \in A, \forall s \in S_{ij} \quad (2.17)$$

Les contraintes (2.17) sont des contraintes de capacités sur les services.

$$Y_{o_k}^k = R_k \quad \forall k \in K \quad (2.18)$$

Les contraintes (2.18) impliquent que la date d'arrivée de la commodité à son terminal d'origine est la date de mise à disposition de la commodité.

$$(\max\{Y_i^k, SD_i^s\} + t_{ijs} - Y_j^k) \times X_{ij}^{ks} = 0 \quad \forall k \in K, \forall (i,j) \in A, \forall s \in S_{ij} \quad (2.19)$$

Les contraintes (2.19) sont des contraintes liantes entre les variables de date d'arrivée ainsi que les variables de flot sur les arcs.

$$Y_i^k \leq u_i^s \times X_{ij}^{ks} + M \times (1 - X_{ij}^{ks}) \quad \forall k \in K, \forall (i,j) \in A, \forall s \in S_{ij} \quad (2.20)$$

Les contraintes (2.20) assurent que la date d'arrivée de la commodité k au terminal i ne peut être plus tard que la date limite d'embarquement du service s si la commodité compte être embarqué dans ce dernier.

Autrement, quand $X_{ij}^{ks} = 0$, alors $Y_i^k \leq M$. Cette inégalité est toujours vérifiée.

$$Y_{d_k}^k \leq D_k \quad \forall k \in K \quad (2.21)$$

Les contraintes (2.21) assurent que chaque commodité arrive à destination avant sa date limite de livraison.

$$(\max\{0, l_i^s - Y_i^k\} - Z_{ij}^{ks}) \times X_{ij}^{ks} = 0 \quad \forall k \in K, \forall (i, j) \in A, \forall s \in S_{ij} \quad (2.22)$$

Les contraintes (2.22) mettent en lien les variables d'entreposage, les variables d'arrivée au terminaux, ainsi que les variables de flot. Elles aident à calculer les coûts d'entreposage de chaque commodité à chaque terminal par lequel elle transite.

$$X_{ij}^{ks} \in \{0, 1\} \quad \forall k \in K, \forall (i, j) \in A, \forall s \in S_{ij} \quad (2.23)$$

$$Y_i^k \geq 0 \quad \forall k \in K, \forall i \in N \quad (2.24)$$

$$Z_{ij}^{ks} \geq 0 \quad \forall k \in K, \forall (i, j) \in A, \forall s \in S_{ij} \quad (2.25)$$

Les contraintes (2.23), (2.24) et (2.25) sont des contraintes de domaine des variables.

Linéarisation du modèle

Le modèle proposé est non-linéaire du fait qu'il contient des fonctions du type $max()$, ainsi que des multiplications de variables. Par conséquent, dans le but de rendre ce modèle résolvable par un solveur de programmation linéaire en nombres entiers, une linéarisation des contraintes (2.19) et (2.22) est proposée ci-dessous.

Les contraintes non-linéaires (2.19) peuvent être linéarisées grâce aux contraintes suivantes.

$$SD_i^s + t_{ijs} - Y_j^k \geq M \times (X_{ij}^{ks} - 1) \quad \forall k \in K, \forall (i, j) \in A, \forall s \in S_{ij} \quad (2.26)$$

$$SD_i^s + t_{ijs} - Y_j^k \leq M \times (1 - X_{ij}^{ks}) \quad \forall k \in K, \forall (i, j) \in A, \forall s \in S_{ij} \quad (2.27)$$

$$Y_i^k + t_{ijs} - Y_j^k \geq M \times (X_{ij}^{ks} - 1) \quad \forall k \in K, \forall (i, j) \in A, \forall s \in S_{ij} \quad (2.28)$$

$$Y_i^k + t_{ijs} - Y_j^k \leq M \times (1 - X_{ij}^{ks}) \quad \forall k \in K, \forall (i, j) \in A, \forall s \in S_{ij} \quad (2.29)$$

Et pour les contraintes (2.22).

$$Z_{ij}^{ks} \geq M \times (X_{ij}^{ks} - 1) + (l_i^s - i - Y_i^k) \quad \forall k \in K, \forall (i, j) \in A, \forall s \in S_{ij} \quad (2.30)$$

$$Z_{ij}^{ks} \leq M \times X_{ij}^{ks} \quad \forall k \in K, \forall (i, j) \in A, \forall s \in S_{ij} \quad (2.31)$$

Concernant le modèle ci-dessus, il a été testé sur un jeu de données où il faut faire transiter un ensemble de 25 commodités à travers un réseau composé de 40 terminaux et de 120 services, représenté par un graphe constitué de 40 nœuds et de 118 arcs. La solution optimale est atteinte en 1 minute et 35 secondes. Le modèle est implémenté grâce au logiciel Lingo 12 sur un ordinateur portable Lenovo Core i5, 2.6GHz de processeur et 4Gb de RAM.

Conception de réseau de service

Le problème de conception de réseau de service, aussi appelé *service network design* dans la littérature, est un problème où il s'agit de faire transiter les marchandises grâce à des services qui doivent être créés par les ressources (camions, avions, trains,...) du problème. Pour qu'un service soit créé, il faut opérer ce service à l'aide d'un véhicule. Ce qui implique l'introduction de contraintes assurant que le choix des services opérés puisse permettre de définir un itinéraire pour les différents véhicules du problème, nommées aussi *design balance constraints*.

La résolution de ce type de problème permet en effet d'apporter une aide à la décision quant à la programmation des horaires de vols ainsi que la gestion des ressources physiques. Si bien qu'on pourrait apporter une réponse à la question : comment déterminer la bonne fréquence des services T. CRAINIC, FERLAND et ROUSSEAU (1984)? Ou encore quel type de véhicule faut-il mettre sur quel trajet afin d'optimiser l'usage des capacités? Et finalement comment faire transiter la demande de façon optimale?

Jardar ANDERSEN, GABRIEL et CHRISTIANSEN (2009) proposent un état de l'art sur la conception de réseau de service avec quatre différentes formulations du problème. Il sera ensuite question de les tester sur différents jeux de données dans le but de les comparer. Il est possible de définir les variables de conception de réseau soit des sur arcs, soit sur des itinéraires (aussi appelés cycles car il y a des rotations de véhicules sur l'horizon de planification). De la même façon, il est possible de définir les variables de flot de marchandises soit sur des arcs, soit sur des chemins. Ce qui donne naissance aux formulations arc-arc, arc-chemin, cycle-arc, et cycle-chemin. Dans cette sous-section, afin d'éviter de surcharger la revue de littérature, seulement la formulation arc-arc est décrite, suivie d'une synthèse sur la comparaison des quatre formulations.

Cependant J. ANDERSEN et al. (2011) et BEKTA, CHOUMAN et T. G. CRAINIC (2010) ont proposés des approches différentes de résolution plus efficaces afin de résoudre des problèmes plus grands. La première J. ANDERSEN et al. 2011 approche se base une résolution du type branch-and-price, tandis que BEKTA, CHOUMAN et T. G. CRAINIC 2010 propose une relaxation lagrangienne où les contraintes pénalisantes sont sur les capacités.

Le problème est défini sur un réseau espace-temps où le temps est discrétisé par périodes. Chaque terminal physique est copié autant de fois qu'il y a de périodes. Donc chaque nœud représente un terminal physique à une certaine période. Les notations de la sous-section "*Problème de flot multicommodités*" sont reprises. Auxquelles se rajoutent les coûts fixes d'ouverture d'arc C_{ij}^v où le véhicule $v \in V$ avec la capacité u_v est affecté à la création du nouvel arc $(i, j) \in A$. Ainsi une nouvelle variable binaire y_{ij}^v est créée, elle vaut 1 si le service $(i, j) \in A$ est opéré par le véhicule $v \in V$, et vaut 0 sinon. La formulation suivant est notée *Formulation arc-arc*, car les variables de décision du design, ainsi que les variables de flot, sont sur les arcs.

$$\min \sum_{(i,j) \in A} C_{ij}^v y_{ij}^v + \sum_{k \in K} \sum_{(i,j) \in A} c_{ij}^k x_{ij}^k \quad (2.32)$$

sujet à :

Contraintes de conservation de flot :

$$\sum_{j \in N} x_{ij}^k - \sum_{j \in N} x_{ji}^k = \begin{cases} w_k & \text{si } i = o_k \\ -w_k & \text{si } i = d_k \\ 0 & \text{sinon} \end{cases} \quad \forall i \in N, \forall k \in K \quad (2.33)$$

Contraintes de capacité :

$$\sum_{k \in K} x_{ij}^k \leq \sum_{v \in V} u_v y_{ij}^v \quad \forall (i, j) \in A \quad (2.34)$$

Contraintes de conservation de véhicules :

$$\sum_{v \in V} \sum_{j \in N} y_{ij}^v - \sum_{v \in V} \sum_{j \in N} y_{ji}^v = 0 \quad \forall i \in N \quad (2.35)$$

Contraintes de domaine des variables :

$$\begin{aligned} x_{ij}^k &\geq 0 && \forall (i, j) \in A, \forall k \in K \\ y_{ij}^v &\in \{0, 1\} && \forall (i, j) \in A, \forall v \in V \end{aligned} \quad (2.36)$$

La fonction objectif (2.32) minimise le coût global qui est somme des coûts fixes de mouvement de véhicules et des coûts variables d'acheminement des commodités. Les contraintes (2.33) demeurent inchangées par rapport au modèle de flot multicommodités. Les contraintes (2.34) sont des contraintes liantes de capacité. (2.35) assure que le nombre de véhicules qui entre dans un terminal soit égal au nombre de véhicules qui sort de ce même terminal afin qu'il n'y ait ni perte ni création de véhicules.

Afin de tester les différentes formulations, 21 instances ont été générées par Jardar ANDERSEN, GABRIEL et CHRISTIANSEN (2009), où le nombre de nœuds physiques (terminaux) est de 5, le nombre d'arcs par période de temps est soit de 10, soit de 15. Le nombre de périodes varie entre 15 et 25 par saut de 5, et finalement le nombre de commodités varie entre 15 et 25. Trois exemples sont pris à la fin où le nombre de commodités est de 50, 100 et finalement 200 pour 15 périodes et 15 arcs par périodes, dans le but de voir l'influence du nombre de commodités.

Les formulations à base de cycle (où les variables des itinéraires des véhicules sont sur des cycles) ont de meilleurs résultats que les formulations à base d'arc pour les variables de conception de réseau. Cependant, tous les différents modèles sont très sensibles à l'augmentation du nombre de commodités à transporter. Les deux formulations à base de cycles ont pu résoudre toutes les instances jusqu'à l'optimalité, avec un temps de calcul limité à une heure, sauf pour les deux dernières qui ont respectivement 100 et 200 commodités à transporter *ibid.*

Cependant, l'une des limites de ces formulations est qu'elles n'offrent pas la possibilité de traiter des instances de plus grandes tailles, voire des problèmes industriels. En effet, ces formulations utilisent une génération à priori complète des cycles pour les véhicules, et il s'avère que le nombre de cycles augmente drastiquement avec le nombre de périodes.

2.3 Positionnement et contribution

Le problème d'acheminement de marchandises pour le transport de fret étudié s'inscrit dans le cadre de la planification tactique. Cependant, il comporte une spécificité propre à lui. Il n'est pas possible d'affréter d'avion cargo pour acheminer des demandes. La globalité des demandes doit transiter à travers les vols commerciaux, qui sont planifiés en fonction du transport passager et non de la demande cargo.

En effet, les travaux statistiques sur les données obtenues de chez Air Canada Cargo ont démontré la véracité de la déclaration YARU et LINA (2012) concernant la hiérarchie des terminaux dans un réseau de transport. En effet, une petite proportion des terminaux renferme la grande majorité des revenus de l'activité Cargo.

Deux défis majeurs ont été rencontrés durant ce projet quand il a été question de traiter les données de demande. D'une part, la taille très volumineuse des fichiers de demandes qui comptent, en moyenne, pas loin de 80,000 demandes de transport par mois. Environ 20,000 demandes par semaine à acheminer avec un modèle de multiflots. Très vite, il s'est avéré que c'était trop ambitieux. D'autre part, la difficulté de pouvoir effectuer des prévisions sur la demande future étant donné les très grandes variations que connaissent les volumes de transport.

Le problème d'agrégation rencontré s'inspire de l'idée générale qu'est le p-hub médian dans C.-c. LIN, J.-y. LIN et CHEN 2012. Cependant, l'idée est de ne conserver que lesdites OD de façon à ce que ces dernières renferment la globalité des revenus cargo de l'entreprise (aux alentours de 90% du chiffre d'affaires) et d'agréger les autres OD. Pour ce faire, il a été jugé utile de découper le problème de transport général en deux parties afin que le travail d'agrégation soit réalisé de façon pointilleuse.

La contribution générale est que l'agrégation des demandes utilisée pour le problème d'acheminement des marchandises se fait en se basant sur une étude statistique relativement poussée, où le but est de faire un regroupement de paires OD qui, non seulement diminue la taille du problème, mais qui s'assure aussi de prendre un regroupement qui minimise les variations de la demande. Et ceci afin d'avoir une meilleure prédiction sur les flux à transporter entre les différents hubs. Pour le modèle d'optimisation, le fait d'avoir une formulation par chemins permet aussi une meilleure flexibilité sur les contraintes. Elle permet d'intégrer plus facilement des contraintes spécifiques comme des embargos, ou sinon, au besoin, des contraintes avec des non-linéarités (comme des politiques de rabais sur certains itinéraires où le coût global du chemin est inférieur à la somme des coûts sur chaque arc). Sachant aussi qu'Air Canada Cargo dispose de son propre routeur capable de générer des chemins en total accord avec toutes les contraintes dictées par leurs règlements, il est facile d'imaginer que ces routes peuvent être utilisées dans le modèle par chemin développé.

CHAPITRE 3 AGRÉGATION DE LA DEMANDE ET PRÉVISION DE LA DEMANDE

Dans ce chapitre est présenté le travail de recherche réalisé sur un problème d'agrégation de données liées à des demandes de transports réelles enregistrées par le transporteur Air Canada Cargo sur une certaine fenêtre de temps.

Les caractéristiques du problème seront détaillées dans la première section de ce chapitre. La section 3.2, quant à elle, traite de la première partie de l'agrégation qui consiste à regrouper des aéroports entre eux. Et finalement, la section 3.3 traite de l'agrégation des OD.

3.1 Description du problème

Cette section a pour but d'illustrer les principales caractéristiques du problème d'agrégation des données brutes. le but est d'obtenir en sortie des demandes groupées par OD, qui sont agrégées, et qui ont une meilleure prévisibilité

3.1.1 Données :

- *Demande* : Le problème d'agrégation de demande est considéré comme un problème spatio-temporel où l'information sur toutes les demandes de transport d'une saison est considérée comme connue. En effet la demande est définie en terme de commodités, et dans tout le reste de cette section, le travail d'agrégation ne se basera que sur ces commodités.
En partant d'un jeu de données avec toutes les commodités sur une saison, les commodités qui ont le même terminal d'origine, le même terminal de destination, ainsi que la même date de disponibilité sont fusionnées en une seule grande commodité qui représente la demande de cette OD pour ce même jour qui est le jour de disponibilité de ces commodité au terminal d'origine. Un exemple est illustré dans Table 3.1 où les demandes sont prises une par une, ensuite elles sont regroupées dans Table 3.2. o_k représente l'origine de la demande, od_k la destination, m_k la masse, v_k le volume, et rd_k la date de mise à disposition.
- *Réseau* : Les terminaux quant à eux sont définis sur un réseau physique qui peut être représenté comme un graphe complet $G = (V, E)$ (exemple Figure 3.1). Chaque sommet $v \in V$ de ce graphe représente un terminal, et chacun de ces sommets a un revenu global qui lui est associé. Afin d'estimer le revenu d'un terminal, l'hypothèse prise en considération est que : le revenu d'une demande allant d'un terminal α vers un terminal β , et qui génère revenu σ , générera alors pour chacun des deux terminaux (α et β) un revenu de $\frac{\sigma}{2}$. Ce revenu est donc la moitié de la somme de tous les revenus de toutes les commodités qui ont ce terminal soit pour origine ou destination.

$$\text{Revenu}_{(T)} = \frac{\sum_{(O,D):\{T=O \text{ ou } T=D\}} \text{Revenu}_{(O,D)}}{2}$$

T étant le terminal dont on souhaite calculer le revenu, revenu_{OD} représente le revenu généré par la paire Origin-Destination OD, tandis qu'une arête $e \in E$ reliant deux sommets $v, w \in V$ représente la distance à vol d'oiseau $d(v, w) \geq 0$ qui les sépare. (Exemple Figure 3.2)

Tableau 3.1 Demandes de transport

Demande k	Origine o_k	Destination d_k	Masse m_k	Volume v_k	Jour rd_k
1	A	B	5	2	1
2	C	D	3	3	3
3	C	E	1	4	1
4	A	B	4	1	1
5	D	C	5	2	4
6	D	A	3	3	2
7	B	C	1	4	5
8	C	D	4	1	3
9	A	C	4	2	4
10	A	C	4	2	4
11	D	C	4	1	4
12	B	C	4	1	5
13	C	E	4	3	1
14	D	A	4	5	2
15	C	D	1	1	2

Tableau 3.2 Demande de transport regroupée par OD par jour

Demande k	Origine o_k	Destination d_k	Masse m_k	Volume v_k	Jour rd_k
1	A	B	9	3	1
2	C	D	7	4	3
3	C	D	1	1	2
4	C	E	5	7	1
5	D	C	9	3	4
6	D	A	7	8	2
7	B	C	5	5	5
8	A	C	8	4	4

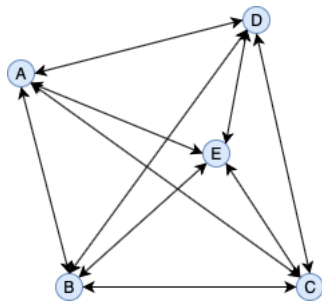


Figure 3.1 Exemple réseau physique avec 5 terminaux

	A	B	C	D	E
A	0	2	5	2	3
B	2	0	3	3	5
C	5	3	0	4	1
D	2	3	4	0	1
E	3	5	1	1	0

Figure 3.2 Matrice des distances

3.1.2 Étude statistique sur les terminaux

Le but de cette sous-section est d'étudier l'impact de chacun des terminaux de notre problème sur le chiffre d'affaire global généré par l'activité cargo sur la saison sélectionnée. D'après le calcul effectué plus haut, il est possible d'avoir une idée représentative sur le trafic d'un terminal qui est mesuré en dollars (revenu par terminal). L'idée générale qui se cache derrière l'agrégation de terminaux est que le regroupement des terminaux se fasse de façon à avoir des terminaux principaux qui ont un revenu relativement élevé. Et en

parallèle des terminaux secondaires, avec un revenu moins élevé, qui sont affiliés aux terminaux principaux dans ce qui est appelé un groupement ou cluster. Pour la suite, l'appellation privilégiée est cluster. Cette première partie contribue à trouver la séparation entre terminaux principaux et terminaux secondaire.

Cependant, la question qui demeure encore sans réponse est : quel est le nombre de clusters idéal à considérer pour ce problème. Tout le reste de ce chapitre est dédié à trouver la réponse à cette question.

3.2 Agrégation de terminaux

Dans cette section, il sera question de voir quels sont les différents scénarios de clustering possibles, ainsi que le modèle mathématique nécessaire pour cette première partie de l'agrégation de la demande.

3.2.1 Nombre de clusters

Il y a en effet plusieurs possibilités pour le nombre final de terminaux principaux. Soit p un entier naturel qui représente le nombre de clusters. Dans le but de mesurer la qualité de clustering, p prendra plusieurs valeurs différentes, et ce seront les résultats de la partie agrégations des OD, qui mesureront la variabilité de la demande et les erreurs d'information générées par l'agrégation, et qui trancheront sur la valeur optimale de p .

Pour l'instant, les valeurs que p prend sont comprises entre 15 et 50 clusters dans le jeu de données utilisé, où il y a initialement 300 terminaux. Il a été décidé ainsi car les 15 plus grands terminaux englobent à eux seuls 70% du CA global, et plus de 90% du CA global de la compagnie se fait sur seulement 50 terminaux du réseau physique.

3.2.2 Modèle mathématique d'agrégation

Comme expliqué précédemment, l'objectif est d'affilier des terminaux secondaires à des terminaux principaux pour obtenir différentes configurations de clustering. Ce regroupement se fait sur la base de deux critères principaux qui sont les distances entre les terminaux ainsi que les revenus de ces derniers. Le modèle utilisé est une modèle de programmation linéaire en nombres entiers qui a pour objectif de minimiser la somme totale des distances pondérées entre chaque terminal principal et les terminaux secondaires qui lui sont affiliés. La pondération se fait avec les revenus des terminaux secondaires dans le but de construire des clusters relativement centrés sur des terminaux importants.

Pour ce modèle, le graphe ainsi que les notations de la sous-section 3.1.1 sont reprises. L'ensemble $W \subset V$ représente l'ensemble des sommets candidats à devenir un terminal principal. Il existe deux types de variables binaires de décision ici, la première étant x_w pour $w \in W$ qui vaut 1 si le sommet candidat w est considéré comme terminal principal, et qui vaut 0 autrement. Le deuxième type de variable est y_{vw} pour $v \in V$ et $w \in W$ qui vaut 1 si le terminal v est affilié à w , et qui vaut 0 autrement. Le modèle est montré ci-dessous.

$$\min \sum_{v \in V, w \in W} \text{revenu}_v \times d(v, w) \times y_{vw} \quad (3.1)$$

s.a. :

$$\sum_{w \in W} x_w = p \quad (3.2)$$

$$\sum_{w \in W} y_{vw} = 1 \quad \forall v \in V \quad (3.3)$$

$$y_{vw} \leq x_w \quad \forall v \in V, \forall w \in W \quad (3.4)$$

$$x_w \in \{0, 1\} \quad \forall w \in W \quad (3.5)$$

$$y_{vw} \in \{0, 1\} \quad \forall v \in V, \forall w \in W \quad (3.6)$$

La fonction objectif (3.1) a pour but de minimiser la somme des distances entre les terminaux secondaires et les terminaux principaux auxquels ils sont affiliés pondérés avec les revenus des terminaux secondaires. Ce choix a pour but d'éviter au mieux de regrouper de grands terminaux dans le même clusters. La contrainte (3.2) assure qu'il y a exactement p clusters au total, les contraintes (3.3) quant à elles font en sorte que tout terminal est affilié à un seul cluster et pas plus. Les contraintes (3.4) interdisent qu'un terminal soit affilié à un autre terminal qui n'est pas considéré comme terminal principal, tandis que les contraintes (3.5) et (3.6) sont des contraintes d'intégralité des variables.

À l'issue de cette première partie de l'agrégation, les demandes par OD se voient modifiées étant donné que seuls les clusters sont considérés à présent. Donc supposons qu'une demande était préalablement entre un terminal A et un terminal B , et que les terminaux font partie des clusters dont les centres sont les terminaux X et Y respectivement.

1. Si $X \neq Y$, alors toutes les demandes de A vers B sont agrégées avec les demandes dont l'origine est X et la destination est Y .
2. Si $X = Y$, alors toutes les demandes de A vers B sont ignorées.

3.2.3 Premiers résultats sur la variabilité de la demande et discussion

Après l'implémentation du modèle, le fichier de sortie contient un clustering par valeur initiale de p . Ici, le but est de voir l'incidence du clustering sur la variabilité de la demande sur les différents OD, qui est considérée comme une variable aléatoire suivant une distribution normale. Afin de mesurer cette dernière, l'indicateur statistique choisi est le coefficient de variation $cv = \frac{\sigma}{\mu}$ où σ représente l'écart-type et μ la moyenne de la demande. Cet indicateur mesure le pourcentage de dispersion des valeurs par rapport à la moyenne. Il faut préciser que le but de l'exercice est de faire des prédictions sur une semaine type, donc pour chaque OD, l'observation des séries temporelles porte sur chaque jour de la semaine de façon distincte. S'il est question du Lundi par exemple, la moyenne du lundi pour une certaine paire Origine-Destination $O_i D_j$ $\mu_{Lundi}^{O_i D_j}$ se calcule en faisant la moyenne de toutes les demandes $O_i D_j$ de la saison arrivées un lundi, c'est-à-dire comme suit.

$$\mu_{Lundi}^{O_i D_j} = \frac{1}{\#(\text{Lundis de la saison})} \times \sum_{\text{Lundi} \in \text{Lundis de la saison}} \text{revenu}_{\text{Lundi}}^{O_i D_j} \quad (3.7)$$

Par la même occasion, l'écart-type $\sigma_{Lundi}^{O_i D_j}$ est défini comme ci-dessous.

$$\sigma_{Lundi}^{O_i D_j} = \sqrt{\frac{1}{\#(\text{Lundis de la saison})} \times \sum_{Lundi \in \text{Lundis de la saison}} (\text{revenu}_{Lundi}^{O_i D_j} - \mu_{Lundi}^{O_i D_j})^2} \quad (3.8)$$

De la même façon, les coefficients de variation sont calculés pour les coûts de transport des OD en se basant sur les mêmes données, il est possible de retrouver les coûts, car proportionnels au tonnage des demandes. Le calcul des coefficients de variation se fait pour les autres jours de la semaine pour toutes les OD du problème. Les graphiques suivants Figures 3.3 et 3.4 illustrent la variation du coefficient de variation en fonction du nombre de clusters dans le modèle, cependant comme il y a un coefficient par jour de semaine, ici le coefficient de variation est la moyenne des coefficients de variation sur les sept jours de la semaine.

$$cv = \frac{1}{7} \times (cv_{lundi} + cv_{mardi} + cv_{mercredi} + cv_{jeudi} + cv_{vendredi} + cv_{samedi} + cv_{dimanche}) \quad (3.9)$$

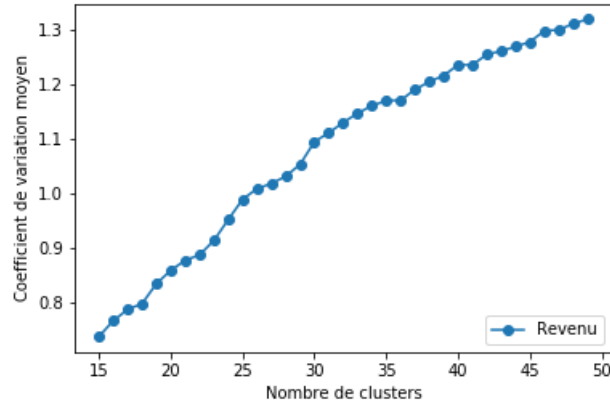


Figure 3.3 Variation de cv sur les revenus en fonction du nombre de clusters p

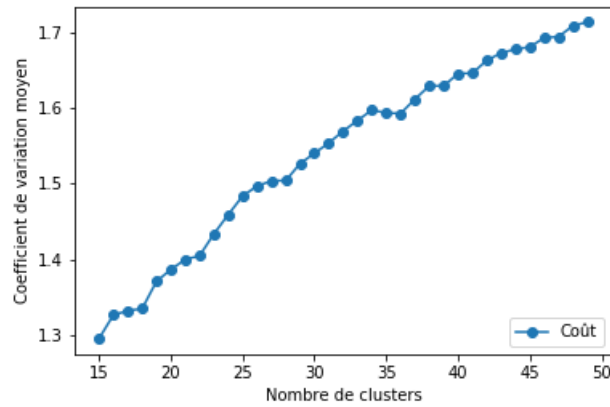


Figure 3.4 Variation de cv sur les coûts en fonction du nombre de clusters p

À noter qu'avant les opérations de clustering, les coefficients de variation sur le revenu et le coûts étaient de $CV_{revenu} : 2.713$ et $CV_{cout} : 2.799$ respectivement. À ce stade, la demande est groupée par clusters, et il est possible de remarquer une nette amélioration sur la variabilité de la demande quand le nombre de clusters diminue. Ce qui prouve que plus on regroupe des terminaux entre eux, plus la variation est faible.

3.3 Agrégation des paires Origine-Destination

Dans cette section, l'objectif est de voir la méthode d'agrégation des paires OD dans le but d'améliorer les prédictions de la demande, son implémentation, ainsi que les résultats finaux obtenus afin de constater s'il y a une amélioration réelle de la variabilité de la demande.

3.3.1 Sélection des OD

Après le traitement sur les OD brutes du problème, le nombre d'OD a été réduit à au plus $p(p - 1)$ où la demande reste encore concentrée sur une proportion de ce nombre. En effet, pour être sûr de capter la totalité de la demande importante, comme fait précédemment, les OD principales sont sélectionnées de façon à prendre les plus rentables qui totalisent 90% du CA global. Pour ce faire, un classement des OD par CA est effectué. Les OD principales sont les prises dans ce classement dans l'ordre décroissant des CA jusqu'à parcourir un CA cumulatif de 90% du CA globa. Pour les autres OD secondaires, dans la même philosophie de regroupement, le but est de les affilier à des OD principales. Pour la suite, les demandes intra-clusters¹ sont négligées, seules celles qui sont inter-clusters² sont prises en compte.

3.3.2 Regroupement des ODs

Dans le but de regrouper les OD secondaires avec des OD principales, il est nécessaire de s'immerger dans les données historiques sur les demandes, et en particulier les itinéraires observés sur les trajets des marchandises. Le principe est qu'il faut déterminer une liste d'OD de remplacement qui font partie des OD principales conservées.

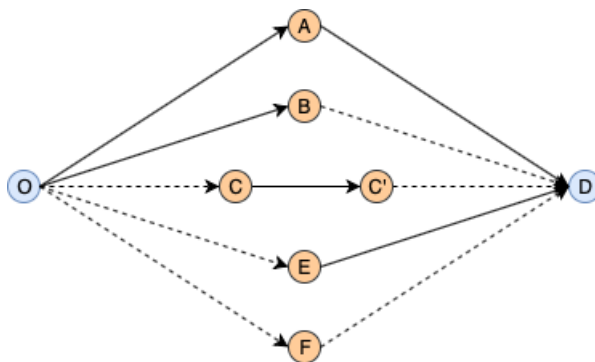


Figure 3.5 Exemple d'agrégation d'OD secondaire

En pratique, dans le travail qui a été fourni, pour chaque OD secondaire, tous les trajets utilisés historiquement et qui ont une ou plusieurs correspondances sont dénombrés. Le but est que le volume ainsi que le revenu de

1. Demandes entre différents terminaux au sein du même cluster
2. Demandes entre différents terminaux présents dans deux clusters différents

la demande de l'OD à agréger soient répartis uniformément sur toutes les ODs principales présentes dans les trajets utilisés. La Figure 3.5 illustre comment la totalité de la demande OD est distribuée.

Dans la Figure 3.5, il s'agit de trouver une ou des OD de remplacement pour "OD". Les arcs en pointillés représentent des ODs secondaires, tandis que les arcs continus représentent des OD principales. Il y a donc 5 OD de remplacement qui sont les OA, AD, OB, CC' et ED. La demande (CA, ainsi que la masse et le volume) de OD est donc répartie équitablement sur ces 5 OD.

Cependant, si aucune OD de remplacement n'a été trouvée, la demande est négligée. Ce qui veut dire que son revenu est donc perdu, et que le coût de son transport disparaît par la même occasion. Afin de mesurer l'impact de cette hypothèse, les Figure 3.6 et 3.7 montrent, en pourcentage du revenu global, les revenus perdus en fonction du nombre de clusters p , par jour, et ensuite en moyenne sur une semaine afin d'avoir un ordre de grandeur global. Et de la même manière pour les Figures 3.8 et 3.9 où cette fois-ci, c'est pour les coûts.

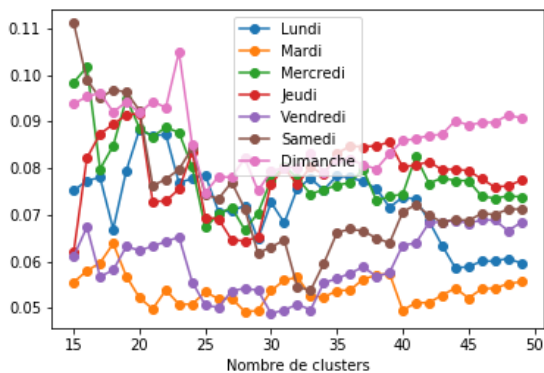


Figure 3.6 Variation du revenu perdu par jour

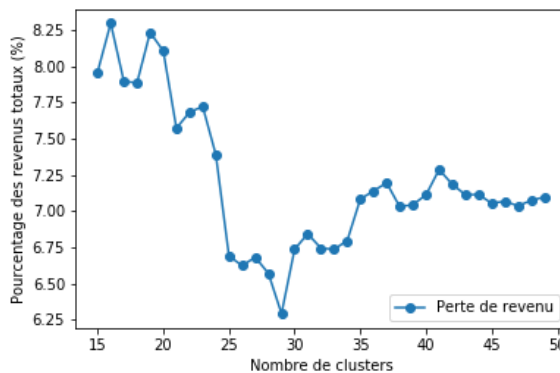


Figure 3.7 Variation du revenu perdu en moyenne

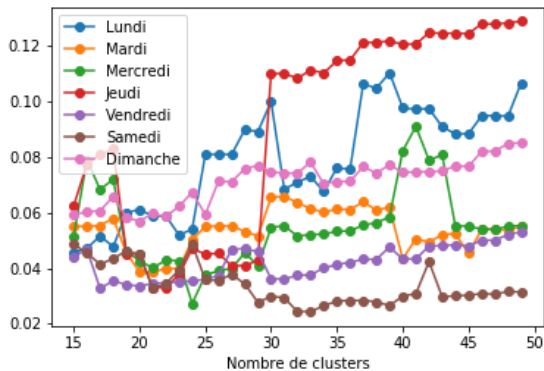


Figure 3.8 Variation du coût perdu par jour

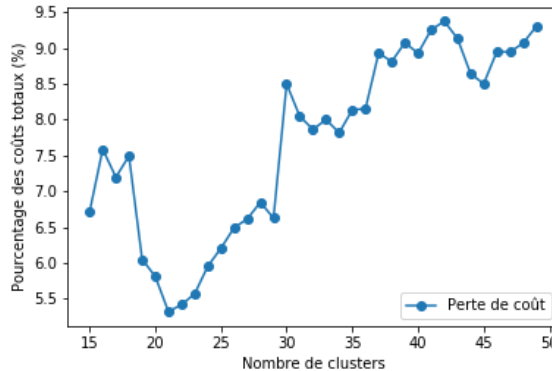


Figure 3.9 Variation du coût perdu en moyenne

Il est possible de voir, à travers les figures ci-dessus, que la proportion des revenus et des coûts négligés est

relativement faible. Elle varie entre 6.3% et 8.3% pour les revenus, et entre 5.3% 9.4% pour les coûts. Il est effet possible de compenser ce manque quand les scénarios de demandes sont générés. Cependant, il est dur de trouver une tendance car l'analyse de la configuration des clusters montre qu'il est possible, qu'entre deux valeurs de p consécutives, les changements dans le regroupement des terminaux soient significatifs.

3.3.3 Premiers résultats sur la variabilité de la demande et discussion

Après avoir utilisé les résultats du clustering illustré dans la sous-section 3.2.2 pour regrouper les OD, les coefficients de variation sur les différentes OD par jour de la semaine sont calculés de la même façon qui est expliqués dans la sous-section 3.2.3.

Les Figures 3.10 et 3.11 permettent de comparer les coefficients de variation moyens sur le coût et le revenu entre l'agrégation des terminaux et l'agrégation des OD.

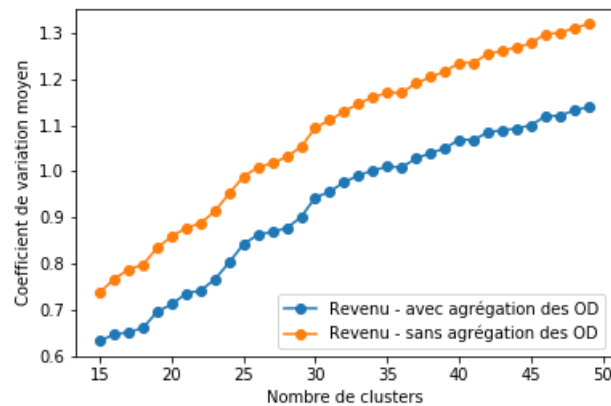


Figure 3.10 Variation de cv sur les revenus en fonction du nombre de clusters p avec et sans agrégation des OD

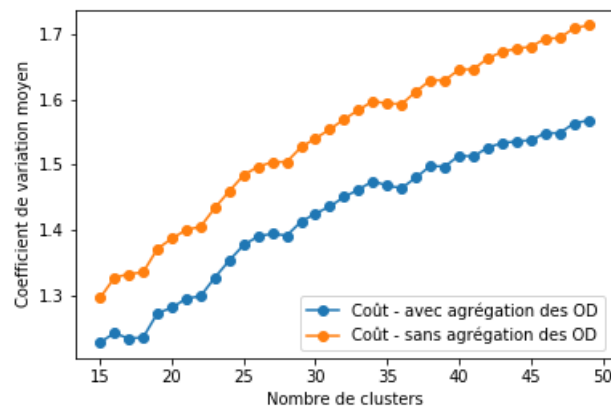


Figure 3.11 Variation de cv sur les coûts en fonction du nombre de clusters p avec et sans agrégation des OD

Les deux graphiques ci-dessus montrent que l'agrégation des OD diminue les coefficients de variation sur le coût et le revenu. Afin d'avoir une vision plus précise de l'ordre de grandeur des coefficients de variation par

OD et non pas en moyenne, les figures suivantes (Figure 3.12 jusqu'à Figure 3.20) montrent les variations du coefficient de variation du revenu et du coût en fonction de revenu de l'OD, pour le premier jour de la semaine, et ce pour plusieurs valeurs de p . Il est possible de retrouver les autres graphiques pour les autres jours de la semaine dans l'Annexe A. Il y a, en effet, un phénomène très intéressant qui se produit. Il est possible de remarquer à travers les graphiques que plus le revenu d'une OD est important, plus son coefficient de variation sur la demande est faible. Par la même occasion, il est possible de vérifier que le coefficient de variation moyen augmente avec p . Il y a, en effet, moins d'agglutination de points sur la partie où les coefficients de variation sont très élevés.

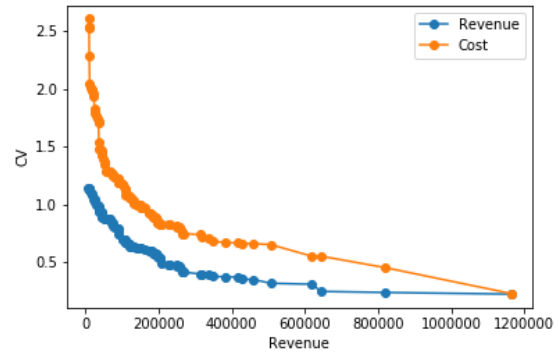
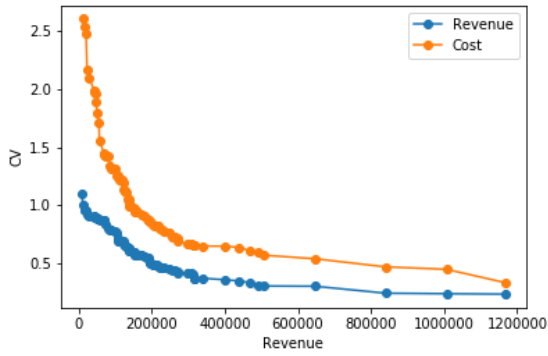


Figure 3.12 Variation de cv pour le lundi avec $p = 15$ Figure 3.13 Variation de cv pour le lundi avec $p = 20$

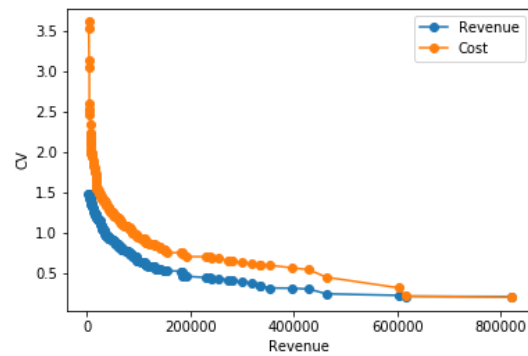
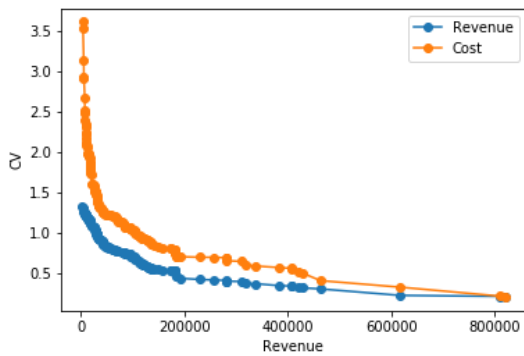


Figure 3.14 Variation de cv pour le lundi avec $p = 25$ Figure 3.15 Variation de cv pour le lundi avec $p = 30$

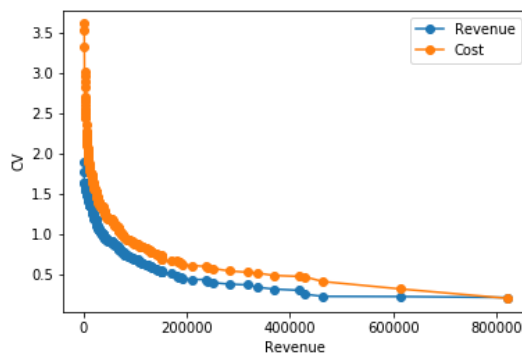
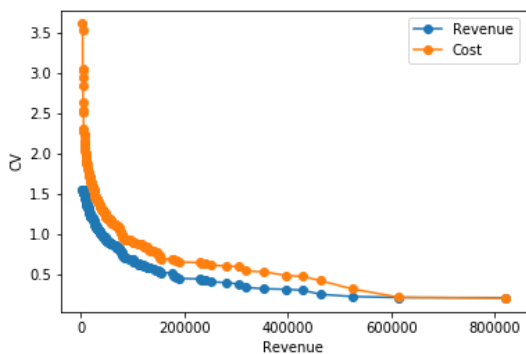


Figure 3.16 Variation de cv pour le lundi avec $p = 35$ Figure 3.17 Variation de cv pour le lundi avec $p = 40$

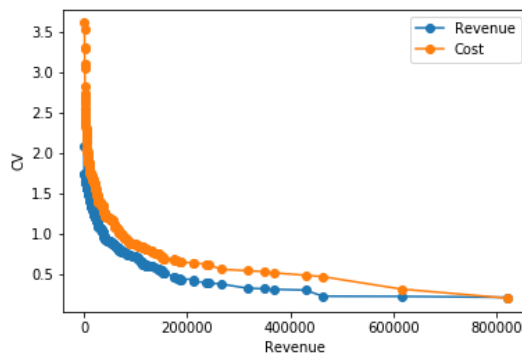
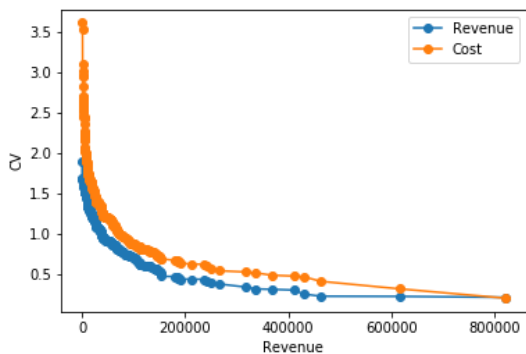


Figure 3.18 Variation de cv pour le lundi avec $p = 45$ Figure 3.19 Variation de cv pour le lundi avec $p = 50$

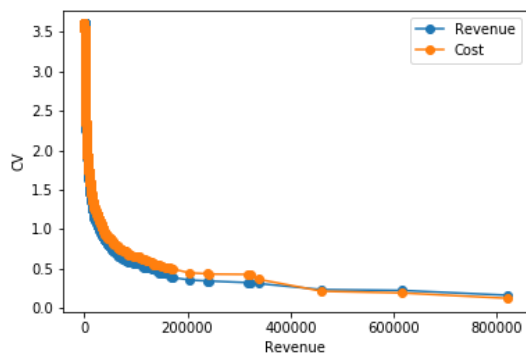


Figure 3.20 Variation de cv pour le lundi sans clustering

3.3.4 Une erreur sur les coûts

Agréger des OD sous-entend forcément qu'une partie de l'information devient approximative et inexacte. En effet, l'une des facettes du problème qui est affectée est le coût de transport des marchandises entre deux points. Les coûts sont généralement définis sur des services avec une origine et une destination spécifique, dépendamment des ressources utilisés pour ce service, le prix pour un service allant d'une origine à une destination n'est pas forcément le même pour un autre trajet. Et faire des regroupements d'OD ne vient clairement pas résoudre le problème. En effet, afin de choisir une bonne agrégation, il est nécessaire de faire un bon compromis entre la variabilité de la demande ainsi que l'erreur sur les coûts engendrée par l'agrégation des OD.

Afin de pouvoir faire ce choix, il faut d'abord commencer par calculer l'erreur. Le calcul ci-dessous estime l'erreur générée par l'agrégation des OD dépendamment des OD de remplacement qui ont été trouvées.

$$\text{erreur} = M \times \left(\frac{\sum_{i=1}^n (\text{coût}_{\text{OD}_{\text{rempl}}^{(i)}} - \text{coût}_{\text{OD}_{\text{org}}})}{n} \right) = M \times \left(\frac{\sum_{i=1}^n \text{coût}_{\text{OD}_{\text{rempl}}^{(i)}}}{n} - \text{coût}_{\text{OD}_{\text{org}}} \right),$$

où M est la masse quotidienne totale de marchandises passant par l'OD originale OD_{org} , $\text{coût}_{\text{OD}_{\text{rempl}}^{(i)}}$, $i = 1, \dots, n$ est le coût par kg des n OD de remplacement, et $\text{coût}_{\text{OD}_{\text{org}}}$ est le coût par kg de l'OD originale.

3.3.5 Meilleur compromis entre erreur sur les coûts et variance de la demande

Toujours dans l'optique de déterminer le bon nombre de clusters p , le critère déterminant est de trouver le minimum entre la somme de la variance de la demande et l'erreur sur les coûts (les deux quantités sont homogènes car elles représentent de l'argent en dollars).

À supposer que le profit par OD est donné par une variable aléatoire normale X_{OD} de moyenne $\mu_{X_{\text{OD}}}$ et d'écart type $\sigma_{X_{\text{OD}}}$. Le profit sur l'ensemble du réseau (agrégé) correspond donc à la variable aléatoire suivante.

$$X_{\text{profit}} = \sum_{\text{OD}} X_{\text{OD}}.$$

Il s'ensuit que X_{profit} est une variable aléatoire normale de moyenne

$$\mu_{\text{profit}} = \sum_{\text{OD}} \mu_{X_{\text{OD}}},$$

et d'écart type

$$\sigma_{\text{profit}} = \sqrt{\sum_{\text{OD}} \sigma_{X_{\text{OD}}}^2}, \quad (3.10)$$

De façon similaire, à supposer que l'erreur sur le coût par OD est donnée par une variable aléatoire normale R_{OD} de moyenne $\mu_{R_{\text{OD}}}$ et d'écart type $\sigma_{R_{\text{OD}}}$, alors l'erreur sur le coût pour l'ensemble du réseau correspond

à

$$R = \sum_{OD} R_{OD}.$$

Il s'ensuit que R est une variable aléatoire normale de moyenne

$$\mu_R = \sum_{OD} \mu_{R_{OD}}, \quad (3.11)$$

et d'écart type

$$\sigma_R = \sqrt{\sum_{OD} \sigma_{R_{OD}}^2}, \quad (3.12)$$

Afin de pouvoir minimiser la variance de la demande ainsi que l'erreur générée par le clustering (somme de la moyenne et de l'écart-type). En effet, σ_{profit} représente la variance de la demande, et $\mu_R + \sigma_R$ représente l'erreur générée par le clustering.

Il faut choisir le clustering qui minimise la quantité suivante.

$$\sigma_{\text{profit}} + \mu_R + \sigma_R, \quad (3.13)$$

3.3.6 Résultats expérimentaux et discussion

Les résultats de cette sous-section est le fruit du travail sur des jeux de données constitués des demandes historiques de la compagnie de transport de fret Air Canada Cargo sur toute une saison. Pour avoir un ordre de grandeur sur le nombre de demandes traitées, il faut compter en moyenne vers les 20,000 demandes par semaine.

Après avoir agrégé tous les terminaux sous forme de clusters, puis ensuite toutes les OD secondaires avec les OD principales comme expliqué plus haut, il a été question de mesurer la variation, d'une part du profit (3.10), et d'autre part de l'erreur sur le coût qui est définie comme étant la somme de la moyenne des erreurs (3.11) et de l'écart-type total sur l'erreur (3.12), en fonction du nombre de clusters. Ce travail a été fourni sur les sept jours de la semaine afin d'être en mesure de faire des prédictions par jour pour le modèle d'acheminement à venir dans le prochain chapitre.

Les résultats expérimentaux de cette section sont illustrés ci-dessous. Les figures qui suivent montrent les variations de l'écart-type des profits (triangles bleus), ainsi que l'erreur sur les coûts (étoiles oranges), et puis la somme des deux 3.13, qui est à minimiser, pour différentes valeurs de p allant de 15 à 50 clusters. Le but étant de déterminer la valeur optimale de clusters qui permet de faire le meilleur compromis.

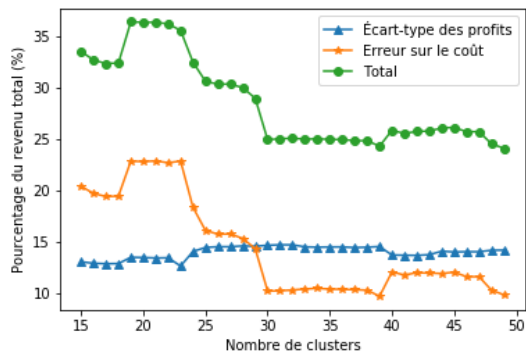


Figure 3.21 Moyenne des variations sur une semaine

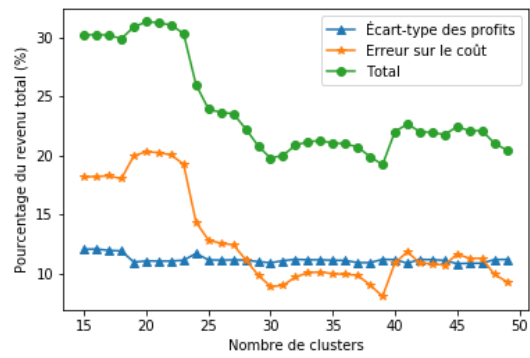


Figure 3.22 Variations pour le lundi

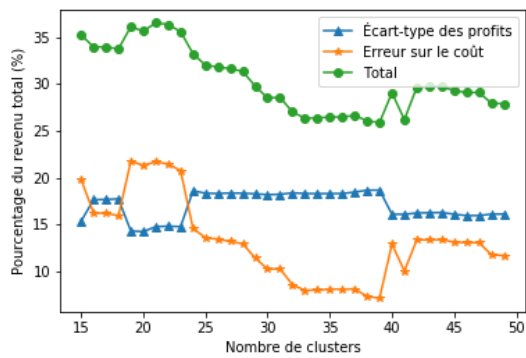


Figure 3.23 Moyenne pour le mardi

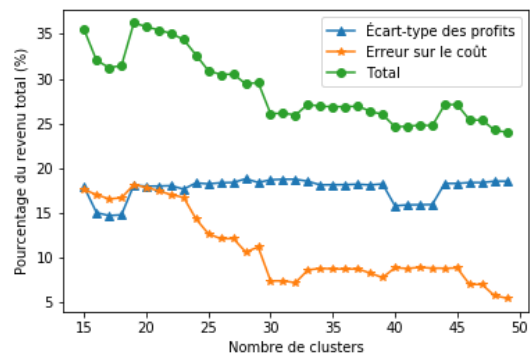


Figure 3.24 Variations pour le mercredi

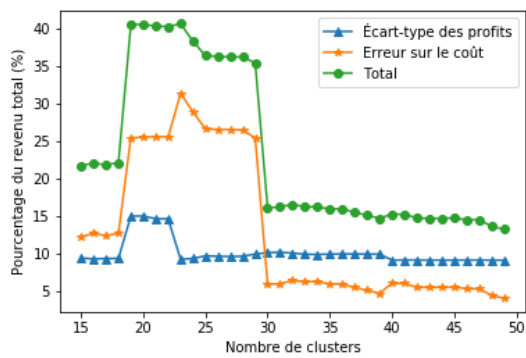


Figure 3.25 Moyenne pour le jeudi

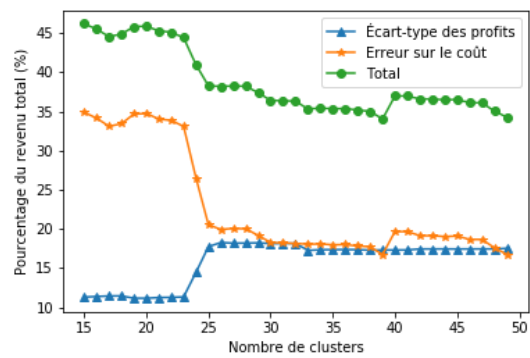


Figure 3.26 Variations pour le vendredi

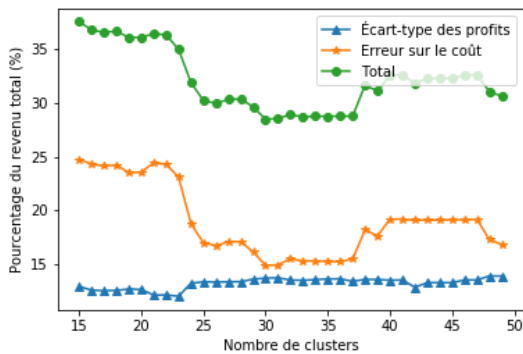


Figure 3.27 Moyenne pour le samedi

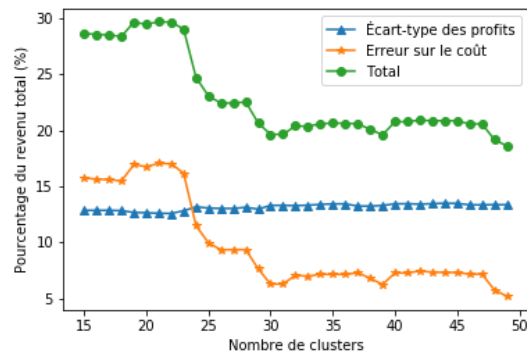


Figure 3.28 Variations pour le dimanche

Il est possible de remarquer que l'erreur sur le coût baisse quand le nombre de clusters augmente. En effet, ceci peut être expliqué par le fait que plus il y a de clusters, moins il y a de terminaux par cluster. Étant donné que le regroupement des terminaux se fait en fonction de la distance, les terminaux au sein d'un cluster sont de plus en plus rapprochés. Comme le coût de transport entre deux terminaux dépend généralement de la distance qui les séparent, il est possible d'en conclure que l'erreur diminue.

Concernant le profit maintenant, l'écart-type du profit augmente légèrement en fonction du nombre de clusters. D'une part, comme vu précédemment, la variation croît avec le nombre de clusters. Donc plus le nombre de clusters augmente, plus il y a de la variation, plus l'écart-type des profits est important. Cependant, cette augmentation est lente car elle représente le rapport de l'écart-type de tous les profits sur le revenu total. Or, comme vu dans la section 3.3.3, les OD avec les plus grands profits ont des variations de profits très basses, tandis que les OD avec des petits profits ont une grande volatilité de la demande. En reprenant la définition (3.10), ce qui justifie une augmentation aussi faible de l'écart-type sur le profit est que les OD qui ont une très grande variance ont un revenu relativement faible. Quand on normalise donc les écart-types des différentes OD par rapport au revenu global, l'écart-type des profits augmente lentement.

L'objectif est de déterminer une agrégation globale des demandes par OD qui propose un bon compromis entre la variance de la demande et l'erreur engendrée par l'agrégation. D'une part, la Figure 3.72 démontre une stabilisation de la courbe qui représente la somme de l'erreur sur le coût et de l'écart-type sur le profit à partir de 30 clusters, et d'autre part, les courbes dans les Figure 3.10 et 3.11 montrent que la moyenne des coefficients de variation sur le coût et le revenu augmentent en fonction du nombre de clusters. Avoir donc un nombre de clusters aux alentours du nombre 30 représente une assez bonne solution.

Concernant le chapitre à venir, tous les fichiers d'instance pour les modèles d'acheminement sont créés sur la base de cette agrégation.

CHAPITRE 4 ÉTUDE COMPARÉE DE DEUX FORMULATIONS DU PROBLÈME D'ACHEMINEMENT DES DEMANDES

Le but de ce chapitre, comme son nom l'indique, est de comparer deux formulations de programmes linéaires du problème d'acheminement des demandes. Tout d'abord, la première section offre une description détaillée du problème et de sa modélisation. Concernant les formulations, dans la première formulation, les variables de décision des flots portent sur les arcs, tandis que pour la deuxième, elles sont sur des chemins qui ont été générés à priori. La comparaison de ces modèles se fait de manière expérimentale sur des instances créées à travers le processus d'agrégation traité dans le chapitre précédent.

4.1 Description du problème

Cette section a pour objet de décrire les principales caractéristiques du problème d'acheminement des demandes agrégées à travers le réseau de service d'Air Canada. Une revue des données utilisées est proposée afin de mieux expliquer l'implémentation du réseau global.

4.1.1 Données

- *Terminaux* : Le transport des demandes ainsi que l'utilisation des services se fait autour du cadre des terminaux physiques du problème. Chaque terminal a un coût d'entreposage c_{store} par unité de masse pour les demandes qui y transitent et reste un certain temps, cependant les capacités des terminaux sont considérées comme infinies.
- *Demande* : La demande est agrégée à ce niveau-ci, chacune des OD sur un jour est assimilée à une commodité. Le transporteur se doit de satisfaire le maximum possible des commodités. Ici, l'horizon temporel est sur une semaine typique d'une saison, ce qui veut dire que les demandes sont périodiques de semaine en semaine. Chacune de ces commodités k est définie par un terminal d'origine o_k , un terminal de destination d_k , un volume v_k , une masse m_k , une date de mise à disposition rd_k , qui représente le jour de la semaine sur lequel se fait l'agrégation dans le CHAPITRE 3, ainsi qu'une date de livraison au plus tard dd_k . L'ensemble de demandes est noté K . L'ensemble des symboles et ensembles est repris dans Tableau 4.1
- *Services* : Le réseau est constitué d'une multitude de services qui sont les vols opérés par la compagnie Air Canada où Air Canada Cargo peut utiliser les capacités résiduelles des compartiments bagage. Comme l'horizon de la demande est sur une semaine, il est nécessaire d'avoir un horizon roulant pour les services afin que les demandes en fin d'horizon puissent être acheminées à destination. Chacun de ces services s est défini par un terminal d'origine o_s , un terminal de destination d_s , une date d'embarquement limite l_s , une date d'arrivée au terminal de destination u_s , des capacités massique M_s et volumique V_s , ainsi qu'un coût de transport par unité de masse c_s .

4.1.2 Contraintes et fonction d'optimisation

- *Contraintes* : Il s'agit de satisfaire au maximum les demandes de transport en utilisant le réseau de service. Dans ce processus d'acheminement, il n'est pas possible de dépasser les différentes capacités des services, et il faut veiller à ce que les livraisons aient lieu durant la fenêtre de temps correspondante.

- *Fonction d'optimisation* : La fonction d'optimisation intègre tous les coûts liés au plan de transport des marchandises. Le but est de minimiser ces coûts de transport qui sont la somme du coût variable linéaire dépendant de la masse des marchandises et des coûts d'entreposage quand les marchandises sont en attente dans les terminaux.

4.1.3 Implémentation du réseau

Le système de transport de la compagnie Air Canada est structuré comme un réseau espace-temps afin de représenter la dimension temporelle du problème qui est nécessaire pour sa modélisation. Étant donné l'horizon de planification, le réseau espace-temporel est représenté à l'aide d'un graphe orienté $G = (N, A)$. N étant l'ensemble des nœuds, cet ensemble est subdivisé en trois sous-ensembles distincts.

- $N_{\text{Départ de service}}$: Il y a exactement un nœud pour chaque service qui a des dates de départ et d'arrivée incluses dans l'horizon temporel du problème. Ce nœud représente le départ du service.
- $N_{\text{Source de commodité}}$: Il y a exactement un nœud pour chaque commodité du problème, ce nœud représente l'origine, ou plus communément appelée, la source de flot de la commodité en question.
- $N_{\text{Puits de commodité}}$: Il y a exactement un nœud pour chaque commodité du problème, ce nœud représente la destination, ou plus communément appelée, le puits de flot de la commodité en question.

L'ensemble des arcs quant à lui est subdivisé en quatre sous-ensembles distincts.

- $A_{\text{Source de commodité}}$: Il y a un arc pour chaque commodité k , et comme son nom l'indique, cet arc relie le nœud source de la commodité au premier nœud de départ de service s qui a la même origine $o_k = o_s$, et une date de départ plus tard que la date de mise à disposition de la commodité $rd_k \leq l_s$.

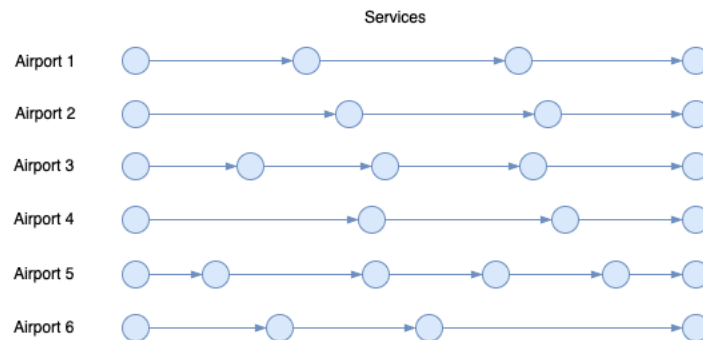


Figure 4.1 Réseau espace-temps : Arcs d'attente seulement

- A_{Attente} : Relie chaque deux nœuds de départ de service avec des départs consécutifs et qui ont le même terminal de départ. Voir la Figure 4.1.
- A_{Service} : Il y a un arc pour chaque service. Cet arc relie la nœud de départ de service s_1 avec un autre nœud de service s_2 qui est le premier service où le terminal de départ de s_2 est celui de destination de s_1 , $d_{s_1} = o_{s_2}$, tel que la date d'arrivée de s_1 soit antérieure à la date de départ s_2 , $u_{s_1} \leq l_{s_2}$. Voir la figure 4.2.
- $A_{\text{Puits de commodité}}$: Il y a un ensemble d'arcs pour chacune des commodités k du problème. Chaque arc de service s avec la même destination que d_k qui a une date de départ l_s après la mise à disposition de la commodité rd_k et qui arrive avant dd_k . En d'autres termes, ce sont tous les arcs de service pouvant être utilisés pour atteindre les destination de k . Voir la Figure 4.4.

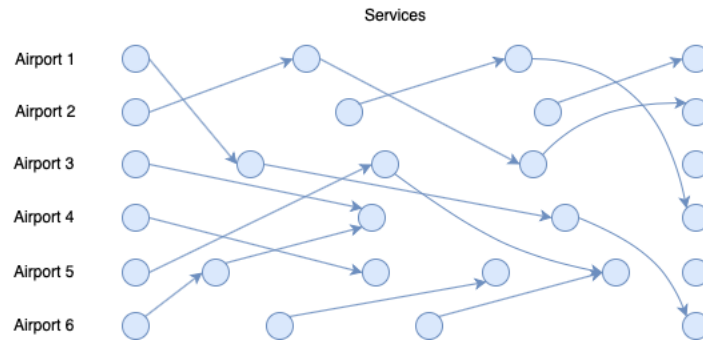


Figure 4.2 Réseau espace-temps : Arcs de service seulement

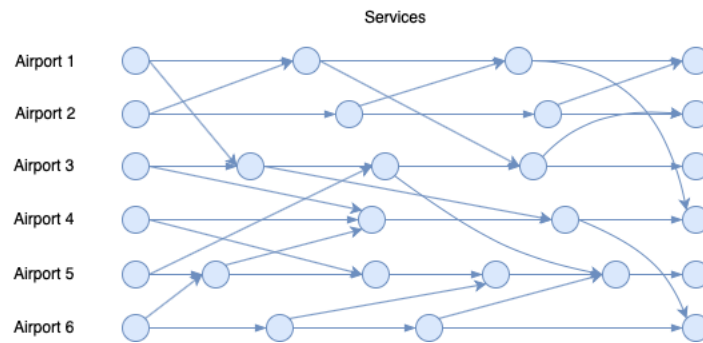


Figure 4.3 Réseau espace-temps : Arcs de service et d'attente

Chaque arc $(i, j) \in A$ a des capacités massiques M_{ij} et volumiques V_{ij} qui lui sont associées, ainsi qu'un coût d'utilisation c_{ij} . S'il s'agit d'un arc de service, alors éventuellement ces capacités sont les capacités du service en soit, en revanche si c'est un autre type d'arc, les capacités sont supposées infinies. De la même façon, les coûts sur les arcs de service sont les coûts d'utilisation du service, et les coûts sur les arcs d'attente représentent les coûts d'entreposage.

Cette modélisation, en elle même, prend en compte un certain nombre de contraintes et informations liées au réseau comme les correspondances possibles, les temps de manutention où encore les coûts de stockage.

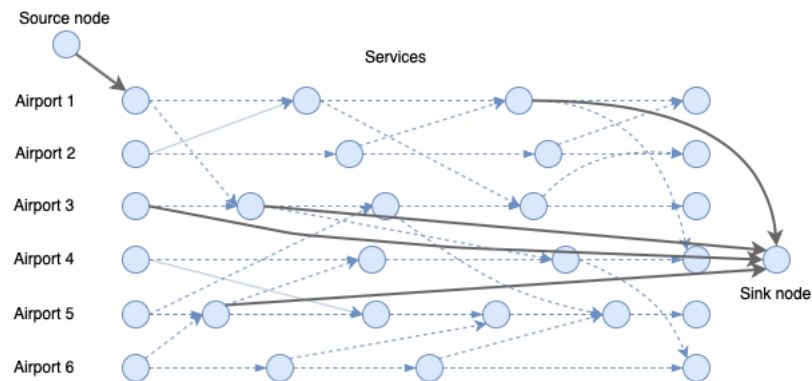


Figure 4.4 Réseau espace-temporel : Tous les arcs

4.2 Deux formulations du problème

Cette section est dédiée à présenter les différentes formulations du problème de transport. La première formulation est une formulation par arc, tandis que la deuxième est par chemin. La dernière sous-section illustre les méthodes de génération de chemins priori utilisée afin de résoudre le problème formulé par chemin.

Dans le réseau espace-temps $G = (N, A)$ Figure 4.5, étant donné un nœud $i \in N$, les ensembles des nœuds successeurs et prédécesseurs à i sont noté respectivement comme suit $\delta^+(i) = \{j \in N : (i, j) \in A\}$ et $\delta^-(i) = \{j \in N : (j, i) \in A\}$.

Le Tableau 4.1 reprend tous les symboles utilisés pour les services et les commodités dans le problème.

Tableau 4.1 Symboles utilisés dans le problème d'acheminement des demandes

Commodités	Représentations
K	Ensemble des commodités
o_k	Origine de la commodité $k \in K$
d_k	Destination de la commodité $k \in K$
v_k	Volume de la commodité $k \in K$
m_k	Masse de la commodité $k \in K$
rd_k	Date de mise à disposition de la commodité $k \in K$
dd_k	Date de livraison au plus tard de la commodité $k \in K$
Services	Représentations
S	Ensemble des services
o_s	Origine du service $s \in S$
d_s	Destination du service $s \in S$
l_s	Date d'embarquement limite du service $s \in S$
u_s	Date d'arrivée du service $s \in S$
V_s	Capacité volumique du service $s \in S$
M_s	Capacité massique du service $s \in S$
c_s	Coût d'utilisation du service $s \in S$

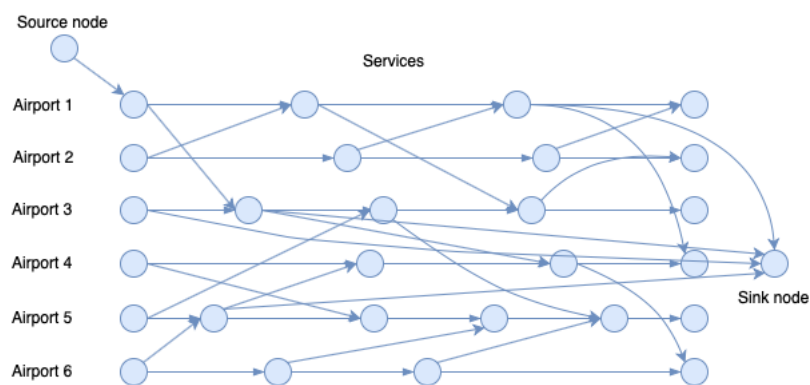


Figure 4.5 Réseau espace-temps du problème

4.2.1 Formulation par arc

La variable x_{ij}^k continue représente le flot de la commodité $k \in K$ qui transite à travers l'arc $(i, j) \in A$ du réseau espace-temps. Donc, pour chaque commodité, il y a autant de variables que d'arcs dans le graphe.

Le reste des symboles a été illustré plus haut. La formulation par arcs s'écrit donc comme suit.

$$\min \sum_{k \in K} \sum_{(i,j) \in A} c_{ij} x_{ij}^k m_k \quad (4.1)$$

sujet à :

$$\sum_{j \in \delta^+(i)} x_{ij}^k - \sum_{j \in \delta^-(i)} x_{ji}^k = \begin{cases} 1 & \text{si } i = o_k \\ -1 & \text{si } i = d_k \\ 0 & \text{autrement} \end{cases} \quad \forall k \in K, \forall i \in N \quad (4.2)$$

$$\sum_{k \in K} x_{ij}^k m_k \leq M_{ij} \quad \forall (i, j) \in A \quad (4.3)$$

$$\sum_{k \in K} x_{ij}^k v_k \leq V_{ij} \quad \forall (i, j) \in A \quad (4.4)$$

$$x_{ij}^k \in [0, 1] \quad \forall k \in K, \forall (i, j) \in A \quad (4.5)$$

La fonction objectif (4.1) minimise le coût global d'acheminement des commodités à travers le réseau espace-temps, Les contraintes (4.2) assurent la conservation du flot à chaque nœud. Il apparaît ici que la totalité de la demande est transportée, cependant afin d'assurer la faisabilité du problème, des arcs artificiels sont introduits entre les nœuds source et puits des commodités. Ces arcs ont des capacités infinies afin d'assurer que la contrainte soit tout le temps vérifiée, cependant le coût d'utilisation de ces arcs artificiels est très grand, ceci a pour but de pénaliser leurs utilisations. Les contraintes (4.3) et (4.4) sont des contraintes de capacité massique et volumique, respectivement. Elles englobent les capacités sur les services ainsi que sur les terminaux quand il s'agit d'entreposage. Et finalement les contraintes (4.5) sont les contraintes de domaine des variables.

4.2.2 Formulation par chemin

Pour cette deuxième approche, la formulation du problème est basée sur des variables définies sur des chemins. En effet, Air Canada Cargo dispose d'un générateur de chemins qui contient toutes les règles à satisfaire pour un chemin. Pour cette expérimentation, dont le but est d'étudier les temps de calcul des différentes formulations, les chemins sont générés grâce à des algorithmes d'exploration de graphe qui ne contiennent pas la totalité des contraintes (étant donné que cette liste de contraintes ne nous a pas été communiquée). Les méthodes utilisées sont expliquées dans la prochaine sous-section.

Soit P_k l'ensemble des chemins qui peuvent transporter la commodité k de son nœud source à son nœud puits dans le réseau espace-temps. Les variables sont maintenant indicées par des chemins, $h_p^k \in [0, 1]$ représente la fraction de flot de la commodité k qui emprunte le chemin p . Un nouveau paramètre apparaît, a_p^{ij} qui vaut 1 si le chemin $p \in P_k$ contient l'arc $(i, j) \in A$, et qui vaut 0 sinon. Chaque chemin p a un certain coût d'utilisation qui est égal à la somme des coûts des services et de l'entreposage qu'il contient. Un chemin est

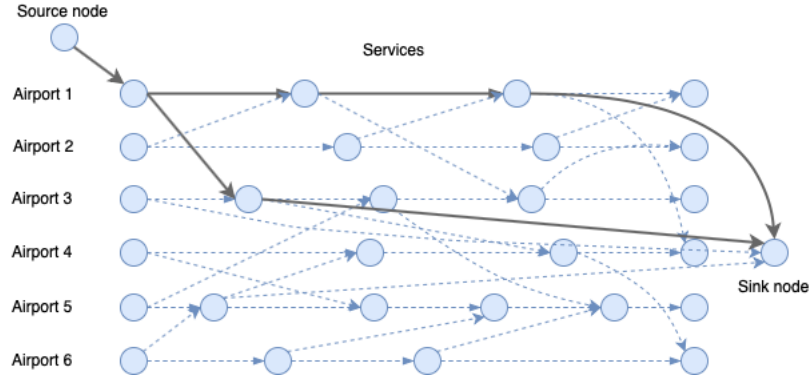


Figure 4.6 Chemins entre la source et le puits de la commodité

une suite d'arcs, de ce fait son coût est la somme des coûts des arcs $K_p = \sum_{(i,j) \in A} c_{ij} a_p^{ij}$.

La formulation par chemin du problème s'écrit comme suit.

$$\min \sum_{k \in K} \sum_{p \in P_k} K_p h_p^k m_k \quad (4.6)$$

sujet à :

$$\sum_{p \in P_k} h_p^k = 1 \quad \forall k \in K \quad (4.7)$$

$$\sum_{k \in K} \sum_{p \in P_k} a_p^{ij} \times h_p^k \times m_k \leq M_{ij} \quad \forall (i, j) \in A \quad (4.8)$$

$$\sum_{k \in K} \sum_{p \in P_k} a_p^{ij} \times h_p^k \times v_k \leq V_{ij} \quad \forall (i, j) \in A \quad (4.9)$$

$$h_p^k \in [0, 1] \quad \forall k \in K, \forall p \in P_k$$

La fonction objectif (4.6) minimise, tout comme avant, les coûts de transport des marchandises dans le réseau. Les contraintes (4.7) stipule que pour chaque commodité, la somme des flots sur tous les chemins empruntés doit être égale à 1 afin d'assurer que la totalité du flot est acheminée. Ces contraintes viennent remplacer les contraintes classique de conservation de flot. Les contraintes (4.8) et (4.9) représentent les contraintes de capacité sur les arcs du réseau.

4.2.3 Algorithmes pour la génération des chemins à priori

Trois différents algorithmes sont présentés pour la génération à priori des chemins de flot pour les demandes du problèmes. Ils sont illustrés dans Algorithm 1, 2 et 3. Tous les chemins générés pour une demande $k \in K$ se doivent de commencer exactement au nœud source de la demande o_k et finir au nœud puits de la même demande d_k , et ce, en utilisant un certain nombre d'arcs intermédiaires. La nature acyclique du graphe prévient de faire des cycles dans l'exploration du graphe. Pour une commodité k , l'algorithme considère tous les arcs sortants au nœud source de cette dernière. Pour chacun de ces arcs, une séquence d'arcs est

créée et elle contient cet arc, ces séquences constituent un ensemble initial de débuts de chemins candidats. L'exploration de proche en proche vers les nœuds successeurs continue à travers le réseau espace-temps, tout en faisant des mise à jours sur les coûts des chemins, les arcs et les nœuds visités. Une fois le puits de la commodité est atteint par l'une des branches de l'arbre d'énumération, la branche est coupée et un chemin réalisable reliant l'origine de la commodité à sa destination est trouvé. Afin d'éviter de générer un nombre exponentiel de chemins, des règles d'exploration sont introduites. La plus importante est qu'il n'est pas possible d'emprunter un service qui emmène vers un terminal dans lequel le chemin candidat est déjà passé, étant donné qu'il existe des arcs d'attente et d'autant plus qu'ils coûtent beaucoup moins cher que d'emprunter un service.

Les différents algorithmes proposés ont la même structure étant donné que ce sont des algorithmes d'exploration de graphe, cependant l'élément clé qui les différencie les uns des autres est la structure de données qui gère la file d'attente dans l'exploration des nœuds.

Pour l'Algorithme 1, une file est utilisée. L'algorithme 1 est aussi appelé exploration en largeur, car le premier successeur stocké est le premier à être traité (First In First Out). L'avantage de cette exploration c'est qu'elle permet de balayer l'ensemble du réseau graduellement sans rater de l'information.

Result: Ensemble de chemins utilisables pour acheminer les demandes

initialization;

for p *in* *Commodity* **do**

 Création d'une file;

 Ajout du nœud source de p dans la file;

while *file non vide* **do**

 Premier élément de la file devient le nœud actuel;

for *successeurs nœud actuel* **do**

if *successeur est la destination* **then**

 Un chemin a été trouvé;

 Ajouter à la liste des chemins;

else

if *successeur ne fait pas revenir à un terminal déjà visité* **then**

 Ajout du successeur à la file d'attente;

else

 Ne rien faire;

end

end

end

 Suppression du premier élément de la file;

end

end

Algorithm 1: Pseudo-code d'un algorithme d'exploration en largeur pour générer des chemins.

Concernant l'Algorithme 2, l'utilisation d'une pile fait en sorte que l'exploration devient en profondeur, car le premier successeur stocké est le dernier à être traité (Last In First Out). L'avantage de cette méthode est qu'elle permet d'avoir des chemins assez rapidement si la chance fait que l'exploration s'en va dans la bonne région dès le début.

Result: Ensemble de chemins utilisables pour acheminer les demandes

initialization;

for p *in* *Commodity* **do**

 Création d'une pile;

 Ajout du nœud source de p dans la pile;

while *pile non vide* **do**

 Dernier élément de la pile devient le nœud actuel;

for *successeurs nœud actuel* **do**

if *successeur est la destination* **then**

 Un chemin a été trouvé;

 Ajouter à la liste des chemins;

else

if *successeur ne fait pas revenir à un terminal déjà visité* **then**

 Ajout du successeur à la pile;

else

 Ne rien faire;

end

end

end

 Suppression du premier élément de la pile;

end

end

Algorithm 2: Pseudo-code d'un algorithme d'exploration en profondeur pour générer des chemins.

Et finalement pour le dernier algorithme, les successeurs sont placés dans une file de priorité où les chemins candidats sont triés en fonction de leurs coûts. Donc le successeur qui a le chemin candidat le plus bas est celui qui est traité en premier. Cette méthode présente beaucoup d'avantages dans le cas où il est décidé de résoudre le problème de façon heuristique. Il est possible d'obtenir des chemins de très bonne qualité qui donnent un résultat très proche de l'optimal sans pour autant générer tous les chemins possibles. L'algorithme en question est illustré ci-dessous.

Result: Ensemble de chemins utilisables pour acheminer les demandes

initialization;

for p *in* *Commodity* **do**

 Création d'une file de priorité ;

 Ajout du nœud source de p dans la pile;

while *pile non vide* **do**

 L'élément de la file avec le coût le plus bas devient le nœud actuel;

for *successeurs nœud actuel* **do**

if *successeur est la destination* **then**

 Un chemin a été trouvé;

 Ajouter à la liste des chemins;

else

if *successeur ne fait pas revenir à un terminal déjà visité* **then**

 Ajout du successeur à la file;

else

 Ne rien faire;

end

end

end

 Suppression du nœud actuel de la file;

end

end

Algorithm 3: Pseudo-code d'un algorithme d'exploration utilisant une file de priorité pour générer des chemins.

4.3 Comparaison expérimentale des modèles

Dans cette section, il est question de présenter les résultats numériques obtenus. L'étude expérimentale qui a été menée a pour but de comparer la qualité des deux formulations du modèle d'acheminement des demandes qui ont été implémentées. Cette comparaison se fait en terme de temps de résolution, lorsqu'une solution est trouvée.

4.3.1 Instances

Les instances utilisées pour ces tests sont issues du CHAPITRE 2 où les données brutes ont été agrégées suivant plusieurs valeurs de p , qui est le nombre de clusters, de la manière qui a été expliquée dans la section 3.2 du même chapitre.

Une des motivations de l'agrégation des demandes est la taille des instances qui est importante. Donc, afin d'observer la variation des tailles des modèles et des temps de résolution, les tests ont été effectués avec les différentes valeurs de p suivantes : 15, 20, 25, 30, 35, 40, 45, et 50.

Pour chacun des p , 7 scénarios de demande ont été créés où le nombre de demandes monte graduellement jusqu'à atteindre la demande globale d'une semaine. Ajouté à cela, dans un but purement expérimental, 3

scénarios de livraison possible ont été créés : livraison à un jour et demi, deux jours et demi ou trois jours et demi pour toutes les demandes. Dans la réalité des faits, le processus de livraison est quelque peu différent. Si un client souhaite envoyer un bien en utilisant les services d’Air Canada Cargo, la compagnie de fret lui propose un itinéraire détaillé de livraison. Par conséquent la date de livraison n’est plus laissée au hasard, ni même le chemin d’acheminement du bien en question. Cependant, ici la demande est traitée d’une façon globale et non au détail. Les délais choisis sont donc des estimations sur les délais de livraison qui vont d’un délai relativement correct à un délai un peu exagéré.

Pour un total de 21 scénarios par valeur de p , et sachant que p varie entre 15 et 50 par saut de 5, le nombre total des instances à résoudre est de 168. Pour chacune de ces instances, les modèles ont été implémentés en Java sur eclipse avec ILOG Cplex 12.8.0 sur un ordinateur de bureau doté d’un processeur Intel(R) Core(TM) i7-6700 CPU à 3.40GHz possédant 32Go de RAM et sous Linux.

4.3.2 Résultats sur les instances

Avant d’évoquer les temps de résolution des deux formulations, le modèle par chemin dispose déjà d’un avantage non négligeable, qui est la possibilité d’utiliser le générateur de chemins d’Air Canada Cargo qui prend en compte toutes les contraintes du problème. Ces contraintes peuvent être plus durs à modéliser dans la formulation par arcs et peuvent aussi le rendre plus dur à résoudre.

Les résultats expérimentaux sont sous forme de tableaux. Afin de faciliter la lecture, les tableaux sont organisés en fonction de nombre de clusters p . Et pour chaque instance présentée $inst_{n,m}$, n représente un scénario de demande parmi les 7 créés, et m l’un des 3 délais de livraison choisis. (1 pour un jour et demi, 2 pour deux jours et demi, et 3 pour trois jours et demi).

Pour chacune des instances, il est possible de retrouver :

- le nombre de demandes à acheminer ;
- la taille du graphe associé : nombre de nœuds, nombre d’arcs ;
- le temps de résolution du modèle qui utilise la formulation par arcs (en secondes) ;
- le temps de résolution du modèle qui utilise la formulation par chemins (en secondes).

Tableau 4.2 Résultats pour $p = 15$

Instances	Demandes	Taille Graphe		Temps de résolution	
		Nœuds	Arcs	Modèle par arcs	Modèle par chemins
inst _{1,1}	132	1872	15825	2.9	0.9
inst _{1,2}	132	2711	28006	4.5	1.4
inst _{1,3}	132	3510	39571	5.3	2.1
inst _{2,1}	264	2971	31259	8.5	2.8
inst _{2,2}	264	3770	53238	12.6	3.9
inst _{2,3}	264	4599	74491	19.3	5.7
inst _{3,1}	377	3996	46349	14.7	5.3
inst _{3,2}	377	4825	77910	21.5	7.9
inst _{3,3}	377	5638	109204	27.3	10.8
inst _{4,1}	481	5033	60134	25.9	10.2
inst _{4,2}	481	5846	101091	31.4	13.7
inst _{4,3}	481	6698	141467	34.1	16.6
inst _{5,1}	598	6080	75083	36.1	13.0
inst _{5,2}	598	6932	125615	43.0	21.9
inst _{5,3}	598	7715	175592	60.1	29.6
inst _{6,1}	716	7168	89735	63.2	27.9
inst _{6,2}	716	7951	149833	70.9	29.2
inst _{6,3}	716	8774	208873	78.3	40.4
inst _{7,1}	839	8197	104683	58.3	31.6
inst _{7,2}	839	9020	173934	81.2	39.8
inst _{7,3}	839	9874	243187	92.9	45.3

Tableau 4.3 Résultats pour $p = 20$

Instances	Demandes	Taille Graphe		Temps de résolution	
		Nœuds	Arcs	Modèle par arcs	Modèle par chemins
inst _{1,1}	179	1964	18967	4.35	1.2
inst _{1,2}	179	2803	33782	7.0	2.2
inst _{1,3}	179	3602	47867	9.4	3.0
inst _{2,1}	342	3127	36919	12.2	3.6
inst _{2,2}	342	3926	63394	16.6	5.5
inst _{2,3}	342	4755	88942	24.0	9.2
inst _{3,1}	476	4194	54099	22.2	8.1
inst _{3,2}	476	5023	91607	30.6	11.7
inst _{3,3}	476	5836	128761	36.2	15.3
inst _{4,1}	589	5249	68496	35.8	12.6
inst _{4,2}	589	6062	115889	40.6	18.6
inst _{4,3}	589	6914	162540	54.5	22.4
inst _{5,1}	732	6348	86427	56.3	17.8
inst _{5,2}	732	7200	145501	66.5	27.9
inst _{5,3}	732	7983	203947	77.1	34.0
inst _{6,1}	877	7490	103307	70.1	31.1
inst _{6,2}	877	8273	173631	98.6	44.5
inst _{6,3}	877	9096	242700	108.5	47.1
inst _{7,1}	1027	8573	120468	99.1	39.8
inst _{7,2}	1027	9396	201435	124.9	52.2
inst _{7,3}	1027	10250	282376	117.4	60.0

Tableau 4.4 Résultats pour $p = 25$

Instances	Demandes	Taille Graphe		Temps de résolution	
		Nœuds	Arcs	Modèle par arcs	Modèle par chemins
inst _{1,1}	284	2586	27952	10.3	3.0
inst _{1,2}	284	3661	50034	16.8	4.9
inst _{1,3}	284	4688	71295	20.8	7.5
inst _{2,1}	528	4147	53951	30.8	9.2
inst _{2,2}	528	5174	93376	36.1	15.1
inst _{2,3}	528	6238	131817	60.6	18.8
inst _{3,1}	722	5562	77810	54.7	19.6
inst _{3,2}	722	6626	132819	71.6	25.8
inst _{3,3}	722	7658	187347	97.6	32.9
inst _{4,1}	889	6960	98709	86.8	28.6
inst _{4,2}	889	7992	168084	108.1	41.0
inst _{4,3}	889	9083	236500	121.0	49.4
inst _{5,1}	1112	8438	123971	141.1	42.3
inst _{5,2}	1112	9529	209965	116.3	51.4
inst _{5,3}	1112	10511	295012	163.3	71.9
inst _{6,1}	1345	9995	149203	158.5	59.6
inst _{6,2}	1345	10977	252145	OUT OF MEMORY	89.3
inst _{6,3}	1345	12012	352977	OUT OF MEMORY	105.3
inst _{7,1}	1569	11425	173060	OUT OF MEMORY	82.9
inst _{7,2}	1569	12460	290689	OUT OF MEMORY	100.8
inst _{7,3}	1569	13548	408078	OUT OF MEMORY	106.7

Tableau 4.5 Résultats pour $p = 30$

Instances	Demandes	Taille Graphe		Temps de résolution	
		Nœuds	Arcs	Modèle par arcs	Modèle par chemins
inst _{1,1}	351	2760	32132	10.0	4.4
inst _{1,2}	351	3853	57430	19.7	8.4
inst _{1,3}	351	4898	81829	25.5	11.8
inst _{2,1}	647	4445	61176	37.8	13.1
inst _{2,2}	647	5490	105940	59.1	22.4
inst _{2,3}	647	6572	149545	78.3	31.9
inst _{3,1}	871	5936	87505	59.9	16.6
inst _{3,2}	871	7018	149517	87.9	33.1
inst _{3,3}	871	8068	210964	102.3	35.5
inst _{4,1}	1067	7410	110715	101.4	30.9
inst _{4,2}	1067	8460	188823	OUT OF MEMORY	41.4
inst _{4,3}	1067	9567	265849	OUT OF MEMORY	51.12
inst _{5,1}	1338	9002	139060	111.15	43.7
inst _{5,2}	1338	10109	235928	OUT OF MEMORY	63.8
inst _{5,3}	1338	11111	331797	OUT OF MEMORY	70.9
inst _{6,1}	1627	10687	168040	OUT OF MEMORY	86.0
inst _{6,2}	1627	11689	284601	OUT OF MEMORY	83.6
inst _{6,3}	1627	12744	398853	OUT OF MEMORY	91.4
inst _{7,1}	1895	12225	194914	OUT OF MEMORY	81.3
inst _{7,2}	1895	13280	328212	OUT OF MEMORY	116.9
inst _{7,3}	1895	14384	461154	OUT OF MEMORY	129.0

Tableau 4.6 Résultats pour $p = 35$

Instances	Demandes	Taille Graphe		Temps de résolution	
		Nœuds	Arcs	Modèle par arcs	Modèle par chemins
inst _{1,1}	373	2804	34556	11.8	3.8
inst _{1,2}	373	3897	61893	21.7	6.2
inst _{1,3}	373	4942	88272	25.4	9.2
inst _{2,1}	694	4539	66087	42.0	10.0
inst _{2,2}	694	5584	114732	56.1	15.9
inst _{2,3}	694	6666	162117	81.7	22.4
inst _{3,1}	927	6050	93941	67.7	26.0
inst _{3,2}	927	7132	160931	106.5	40.3
inst _{3,3}	927	8182	227335	OUT OF MEMORY	55.3
inst _{4,1}	1134	7546	118798	OUT OF MEMORY	32.1
inst _{4,2}	1134	8596	203181	OUT OF MEMORY	46.6
inst _{4,3}	1134	9703	286385	OUT OF MEMORY	56.9
inst _{5,1}	1421	9170	149379	OUT OF MEMORY	45.3
inst _{5,2}	1421	10277	254137	OUT OF MEMORY	63.5
inst _{5,3}	1421	11279	357854	OUT OF MEMORY	79.6
inst _{6,1}	1734	10903	180856	OUT OF MEMORY	65.5
inst _{6,2}	1734	11905	307216	OUT OF MEMORY	90.0
inst _{6,3}	1734	12960	431072	OUT OF MEMORY	108.8
inst _{7,1}	2022	12481	210474	OUT OF MEMORY	87.1
inst _{7,2}	2022	13536	355464	OUT OF MEMORY	121.3
inst _{7,3}	2022	14640	500095	OUT OF MEMORY	137.6

Tableau 4.7 Résultats pour $p = 40$

Instances	Demandes	Taille Graphe		Temps de résolution	
		Nœuds	Arcs	Modèle par arcs	Modèle par chemins
inst _{1,1}	416	2894	37655	17.8	4.3
inst _{1,2}	416	3989	67543	29.8	7.7
inst _{1,3}	416	5036	96464	32.9	11.2
inst _{2,1}	762	4681	70335	45.2	13.7
inst _{2,2}	762	5728	122420	75.7	18.9
inst _{2,3}	762	6810	173226	128.6	23.9
inst _{3,1}	1016	6236	100380	85.9	21.9
inst _{3,2}	1016	7318	172379	OUT OF MEMORY	31.0
inst _{3,3}	1016	8368	243901	OUT OF MEMORY	39.2
inst _{4,1}	1234	7754	126874	OUT OF MEMORY	32.5
inst _{4,2}	1234	8804	217601	OUT OF MEMORY	48.1
inst _{4,3}	1234	9911	307147	OUT OF MEMORY	60.3
inst _{5,1}	1553	9442	160743	OUT OF MEMORY	52.3
inst _{5,2}	1553	10549	274154	OUT OF MEMORY	68.6
inst _{5,3}	1553	11553	386609	OUT OF MEMORY	81.9
inst _{6,1}	1891	11225	194671	OUT OF MEMORY	72.5
inst _{6,2}	1891	12229	331578	OUT OF MEMORY	92.5
inst _{6,3}	1891	13286	465893	OUT OF MEMORY	113.1
inst _{7,1}	2197	12841	226252	OUT OF MEMORY	92.2
inst _{7,2}	2197	13898	383118	OUT OF MEMORY	122.4
inst _{7,3}	2197	15004	539699	OUT OF MEMORY	150.0

Tableau 4.8 Résultats pour $p = 45$

Instances	Demandes	Taille Graphe		Temps de résolution	
		Nœuds	Arcs	Modèle par arcs	Modèle par chemins
inst _{1,1}	470	3046	41326	23.3	4.1
inst _{1,2}	470	4153	74026	25.9	10.5
inst _{1,3}	470	5214	105539	46.6	15.9
inst _{2,1}	853	4919	77120	67.6	13.9
inst _{2,2}	853	5980	134153	87.9	22.4
inst _{2,3}	853	7085	189752	OUT OF MEMORY	32.0
inst _{3,1}	1119	6512	109020	OUT OF MEMORY	25.4
inst _{3,2}	1119	7617	185622	OUT OF MEMORY	44.7
inst _{3,3}	1119	8691	262624	OUT OF MEMORY	46.7
inst _{4,1}	1360	8099	137085	OUT OF MEMORY	42.5
inst _{4,2}	1360	9173	235461	OUT OF MEMORY	57.8
inst _{4,3}	1360	10304	332462	OUT OF MEMORY	58.6
inst _{5,1}	1708	9869	173033	OUT OF MEMORY	50.8
inst _{5,2}	1708	11000	295535	OUT OF MEMORY	75.2
inst _{5,3}	1708	12026	416866	OUT OF MEMORY	82.8
inst _{6,1}	2075	11734	209701	OUT OF MEMORY	76.2
inst _{6,2}	2075	12760	357667	OUT OF MEMORY	105.6
inst _{6,3}	2075	13841	502764	OUT OF MEMORY	119.1
inst _{7,1}	2410	13430	242898	OUT OF MEMORY	102.6
inst _{7,2}	2410	14511	412053	OUT OF MEMORY	140.5
inst _{7,3}	2410	15641	580632	OUT OF MEMORY	158.3

Tableau 4.9 Résultats pour $p = 50$

Instances	Demandes	Taille Graphe		Temps de résolution	
		Nœuds	Arcs	Modèle par arcs	Modèle par chemins
inst _{1,1}	503	3112	43322	19.0	4.9
inst _{1,2}	503	4219	77659	29.2	9.5
inst _{1,3}	503	5280	110761	46.8	11.2
inst _{2,1}	907	5027	80713	58.4	13.0
inst _{2,2}	907	6088	140557	112.5	21.4
inst _{2,3}	907	7193	198877	135.6	33.1
inst _{3,1}	1186	6646	113589	121.5	26.3
inst _{3,2}	1186	7751	195424	OUT OF MEMORY	36.2
inst _{3,3}	1186	8825	276636	OUT OF MEMORY	48.5
inst _{4,1}	1436	8251	143884	OUT OF MEMORY	35.4
inst _{4,2}	1436	9325	247480	OUT OF MEMORY	53.8
inst _{4,3}	1436	10456	349620	OUT OF MEMORY	63.0
inst _{5,1}	1802	10057	180790	OUT OF MEMORY	56.8
inst _{5,2}	1802	11188	309193	OUT OF MEMORY	82.5
inst _{5,3}	1802	12214	436349	OUT OF MEMORY	91.9
inst _{6,1}	2186	11956	218827	OUT OF MEMORY	80.4
inst _{6,2}	2186	12982	373723	OUT OF MEMORY	107.9
inst _{6,3}	2186	14063	525599	OUT OF MEMORY	128.9
inst _{7,1}	2544	13698	254681	OUT OF MEMORY	108.2
inst _{7,2}	2544	14779	432440	OUT OF MEMORY	145.0
inst _{7,3}	2544	15909	609742	OUT OF MEMORY	161.0

Qualité des solutions et discussion

Les tableaux ci-dessus présentent les résultats sur les temps de résolution des différentes instances résolues. En effet, il est possible de remarquer que plus le nombre de clusters grandit, plus le modèle devient volumineux et plus les temps de résolution s'allongent. Si bien qu'à partir de $p = 25$ quelques instances sur le modèle par arc saturent la mémoire de l'ordinateur et n'aboutissent pas à une solution. Ce qui peut être expliqué par le fait que les graphes ont un nombre d'arcs très important, et pour chaque nouvelle demande introduite, le nombre de variables augmente à peu près du nombre d'arcs dans le graphe. Bien qu'il y ait eu un nettoyage du réseau pour chaque commodité avec des optimisations dans le code qui ont pour but de ne pas créer des variables inutiles, ce n'est juste pas suffisant pour faire en sorte que le modèle soit résolu.

Concernant le modèle qui utilise la formulation par arc, toutes les générations à priori ont été faites en utilisant l'algorithme 3. Ceci est dû au fait que la taille des graphes est très importante, et que des algorithmes de type exploration en profondeur ou en largeur ne sont bénéfiques que s'il y a une génération complète de tous les chemins existants. Cependant, dans le cadre de cette expérimentation, ce nombre de chemins augmente tellement vite qu'il sature la mémoire de l'ordinateur. Il a donc été convenu de générer au maximum le nombre de 100 chemins par demande. Il s'est avéré qu'avec ce nombre, les solutions étaient équivalentes aux solutions du modèle avec la formulation par arcs.

Concernant la comparaison entre les deux formulations du modèle, d'une manière générale, les résultats mettent en avant la qualité de la formulation par chemins. Elle aboutit tout le temps à une solution optimale sur toutes les instances du problème, et les temps de résolution sont 2 jusqu'à 4 fois plus courts sur les instances résolues. L'autre point essentiel de cette comparaison est que la génération des chemins à priori permet une intégration facile de contraintes assez compliquées qui peuvent contenir des non-linéarités difficiles voire impossible à formuler dans le modèle par arc. Ceci valide donc, de façon expérimentale, la dominance du modèle formulé par chemins sur le modèle par arcs.

CHAPITRE 5 CONCLUSION ET RECOMMANDATIONS

5.1 Synthèse des travaux

Dans ce mémoire, l'étude a porté sur un problème de planification tactique chez la filiale cargo de la compagnie canadienne Air Canada. Dans ce problème, il est question de considérer une semaine type représentative d'une saison donnée. Le but est de tester la capacité à répondre à la demande du marché grâce au réseau d'Air Canada, et aussi de déterminer un acheminement des marchandises optimal en terme de coûts. En effet, la globalité de la demande doit être acheminée dans le compartiment bagage des vols commerciaux proposés, il n'y a pas de possibilité d'affréter des avions cargo pour transporter de la demande. Tout d'abord, la première étape a été de traiter les données industrielles brutes. À ce niveau de planification, la demande est agrégée de façon à conserver l'essence du problème. Cette agrégation s'est faite pour deux raisons majeures. La première étant que la taille colossale des instances de demande de transport met à rude épreuve les modèles d'acheminement des demandes. La deuxième est en rapport avec la prévisibilité de la demande, il s'est avéré avec les résultats expérimentaux que le regroupement des paires Origine-Destination pour la demande baissait leurs variations. Ce qui aide à améliorer les estimations de la demande par jour de la semaine. La deuxième étape a été de créer un modèle d'optimisation qui permet d'acheminer les demandes agrégées à travers le réseau de service d'Air Canada Cargo. En effet, il est question de créer un plan de transport de la demande cargo qui minimise la somme des coûts liés au transport des marchandises, tout en s'assurant de ne pas dépasser les capacités des services.

La première partie de ce mémoire était consacrée à déterminer la direction que devait prendre l'agrégation de la demande, ainsi que définir son procédé. Dans les faits, la première étape était de déterminer quels étaient les terminaux qui regroupaient la partie la plus représentative du chiffre d'affaire global afin de bâtir les clusters autour d'eux, ces terminaux sont appelés terminaux principaux. Par la suite, le reste des terminaux, appelés terminaux secondaires, ont été affiliés aux principaux au sein de clusters. De premiers résultats expérimentaux ont prouvés que plus les regroupements sont grands (moins il y a de clusters), plus les variations sont faibles. Cependant, même avec ce regroupement de terminaux, la plupart des OD ont gardé une variation très grande. Ce qui a engendré une deuxième étape dans l'agrégation : l'agrégation des demandes par OD. Il a été mis en évidence qu'il y avait quelques scénarios d'agrégation qui offraient un très bon compromis entre la variation de la demande et la perte d'information liée à l'agrégation.

La deuxième partie quant à elle, a été dédiée à l'étude du problème d'acheminement de la demande cargo à travers un réseau de service qui modélise le réel problème industriel de planification tactique de la compagnie de transport de fret. En effet, plusieurs flots de marchandises doivent être transportés à travers le réseau d'Air Canada, en passant par des terminaux et en empruntant des services. Deux formulations du problème ont été présentées et confrontées. Dans la première, les décisions concernant le flot des demandes dans le réseau sont déterminées sur les arcs, tandis que pour la deuxième celles-ci se font sur des chemins pré-définis. Suite à l'implémentation des deux modèles, il a été question de les comparer de façon expérimentale sur différents types d'instance allant de réseaux moyennement grands à très grands. Il est ressorti de cette comparaison que les modèles à base de chemins sont plus prometteurs, spécialement sur les grands réseaux où les modèles à base d'arcs saturer les mémoires des ordinateurs.

5.2 Limitations de la solution proposée et améliorations futures

L'une des limitations majeures de ce problème étant le manque de données reflétant la réalité des choses sur le terrain. En effet, dans le cadre de cette étude, il a été question, des fois, de combler ce manque d'informations avec les quelques données à disposition. L'un des enjeux est de pouvoir déterminer la réelle capacité sur un vol, sachant que les compartiments de bagages sont aussi utilisés pour transporter les affaires des voyageurs. Pour cette question, Air Canada Cargo a en effet un outil qui permet de mesurer avec une bonne précision, un mois à l'avance, la capacité qui sera utilisée par les clients des vols commerciaux, cependant cet outil n'était pas en mesure d'être utilisé pour cette étude.

Concernant les délais de livraison, comme il a été mentionné dans la section 4.3.1, les scénarios de livraison ont été créés en se basant sur le bon sens. Cependant, il est possible de remarquer dans les résultats des instances de demande que cette donnée a une énorme incidence sur la taille des graphes et à fortiori sur les tailles des modèles. Une des améliorations possibles serait d'effectuer un travail de recherche, pour chaque OD, et d'observer les itinéraires utilisés dans le passé en détails afin de pouvoir quantifier des délais moyens de livraison, et s'ajuster sur ces derniers.

Pour la question de regroupement de données, dès lors que deux marchandises sont combinées, elles deviennent une seule demande, où sa masse et son volume sont respectivement la somme des masses et la somme des volumes des deux marchandises initiales. Cependant même si la masse et le volume sont corrélés, la densité des marchandises peut varier de façon assez considérable. Par exemple, une demande groupée composée de métaux lourds et de fruits et légumes. Au cas où la demande suit deux itinéraires différents, il se peut que tout le volume passe sur un des itinéraire et toute la masse sur l'autre. Ce qui complique la tâche, c'est qu'une fois le regroupement fait, il s'agit juste d'une grande demande avec une masse totale et un volume total et donc une densité moyenne. Il serait donc très intéressant de se pencher sur la question de densité sur ce niveau de planification.

Une dernière piste d'amélioration possible serait la segmentation du marché. Jusqu'à présent, la demande est estimée en fonction des données passées d'Air Canada Cargo seulement. Cependant, afin d'estimer la demande du marché, il serait judicieux d'utiliser des données sur la demande globale du marché, afin d'avoir une première idée sur les parts de marché de la compagnie, pour mieux fixer les objectifs de croissance sur les différents marchés. Aussi, il serait intéressant de prendre en considération des données macro-économiques afin de mieux prédire la tendance de la demande. Avoir un modèle de prédiction qui prend en compte toutes ces variables pourrait être très challengeant du point de vue de la recherche, mais les résultats seraient certainement de meilleure qualité.

RÉFÉRENCES

- [1] Jardar ANDERSEN, Teodor GABRIEL et Marielle CHRISTIANSEN. “Service network design with asset management : Formulations and comparative analyses”. In : *Transportation Research Part C* 17.2 (2009), p. 197–207. ISSN : 0968-090X. DOI : 10.1016/j.trc.2008.10.005. URL : <http://dx.doi.org/10.1016/j.trc.2008.10.005>.
- [2] J. ANDERSEN et al. “Branch and Price for Service Network Design with Asset Management Constraints”. In : *Transportation Science* 45.1 (2011), p. 33–49. ISSN : 0041-1655. DOI : 10.1287/trsc.1100.0333.
- [3] Claudia ARCHETTI et Lorenzo PEIRANO. “PT US CR”. In : *Omega* (2019). ISSN : 0305-0483. DOI : 10.1016/j.omega.2019.02.009. URL : <https://doi.org/10.1016/j.omega.2019.02.009>.
- [4] A A ASSAD. “Multicommodity network flows- A survey”. In : *Networks* 8 (1978), p. 37–91. ISSN : 00283045. DOI : 10.1002/net.3230080107.
- [5] Burak AYAR et Hande YAMAN. “An intermodal multicommodity routing problem with scheduled services”. In : (2012), p. 131–153. DOI : 10.1007/s10589-011-9409-z.
- [6] Tolga BEKTA, Mervat CHOUMAN et Teodor Gabriel CRAINIC. “Lagrangian-Based Decomposition Algorithms for Multicommodity Network Design Problems with Penalized Constraints”. In : December 2009 (2010). DOI : 10.1002/net.
- [7] Teodor Gabriel CRAINIC. “Long-haul freight transportation”. In : *Handbook of transportation science* (2003), p. 66. DOI : 10.1007/0-306-48058-1_13. URL : http://link.springer.com/chapter/10.1007/0-306-48058-1%7B%5C_%7D13.
- [8] Teodor Gabriel CRAINIC et Gilbert LAPORTE. “Planning models for freight transportation”. In : *European Journal of Operational Research* 97.3 (1997), p. 409–438. ISSN : 03772217. DOI : 10.1016/S0377-2217(96)00298-6.
- [9] Teodor CRAINIC, Jacques-a FERLAND et Jean-marc ROUSSEAU. “A Tactical Planning Model for Rail Freight Transportation”. In : August 2015 (1984).
- [10] Bo FENG, Yanzhi LI et Zuo-jun Max SHEN. “Air cargo operations : Literature review and comparison with practices”. In : *TRANSPORTATION RESEARCH PART C* 56 (2015), p. 263–280. ISSN : 0968-090X. DOI : 10.1016/j.trc.2015.03.028. URL : <http://dx.doi.org/10.1016/j.trc.2015.03.028>.
- [11] Cheng-chang LIN, Jr-yung LIN et Yin-chieh CHEN. “The capacitated p -hub median problem with integral constraints : An application to a Chinese air cargo network”. In : *Applied Mathematical Modelling* 36.6 (2012), p. 2777–2787. ISSN : 0307-904X. DOI : 10.1016/j.apm.2011.09.063. URL : <http://dx.doi.org/10.1016/j.apm.2011.09.063>.
- [12] Environmental SCIENCE. “Revenue management of air cargo service in theory and practice Revenue management of air cargo service in theory and practice”. In : (2018).
- [13] Yan SUN et Maoxiang LANG. “Modeling the Multicommodity Multimodal Routing Problem with Schedule-Based Services and Carbon Dioxide Emission Costs”. In : 2015 (2015). DOI : 10.1155/2015/406218.

- [14] Dang YARU et Peng LINA. “Hierarchy of Air Freight Transportation Network Based on Centrality Measure of Complex Networks”. In : *Journal of Transportation Systems Engineering and Information Technology* 12.3 (2012), p. 109–114. ISSN : 1570-6672. DOI : 10.1016/S1570-6672(11)60207-9. URL : [http://dx.doi.org/10.1016/S1570-6672\(11\)60207-9](http://dx.doi.org/10.1016/S1570-6672(11)60207-9).

ANNEXE A RÉCAPITULATIF DES GRAPHIQUES SUR LA VARIATION DE CV EN FONCTION DE p

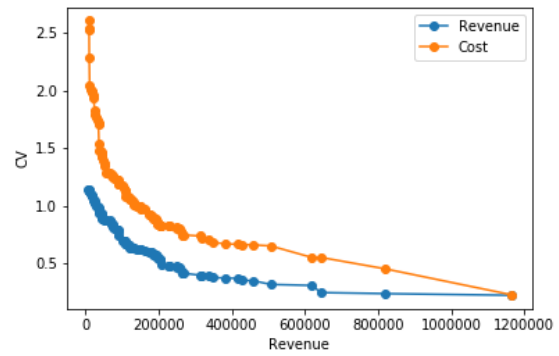
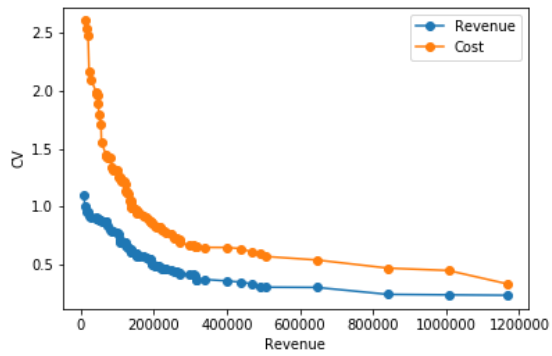


Figure A.1 Variation de cv pour le lundi avec $p = 15$ Figure A.2 Variation de cv pour le lundi avec $p = 20$

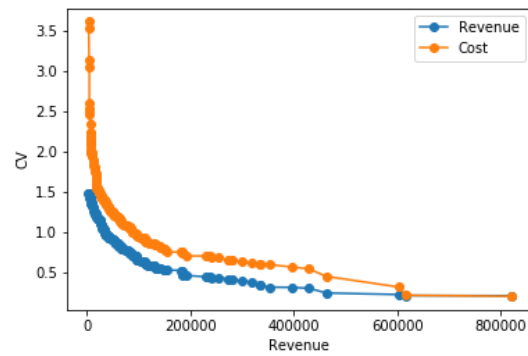
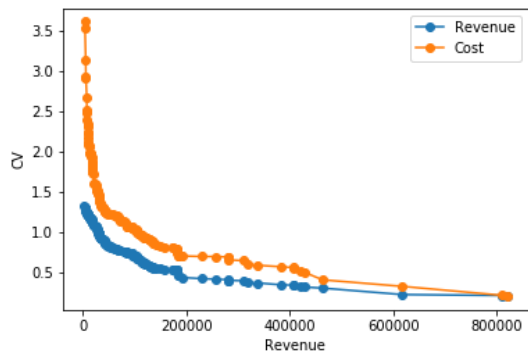


Figure A.3 Variation de cv pour le lundi avec $p = 25$ Figure A.4 Variation de cv pour le lundi avec $p = 30$

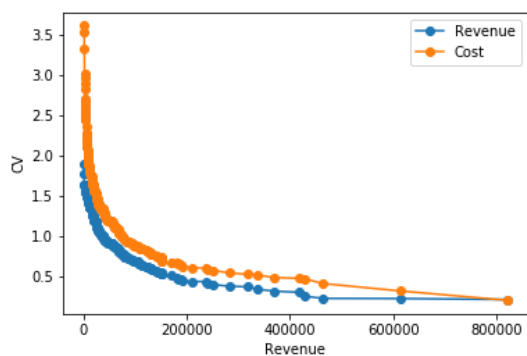
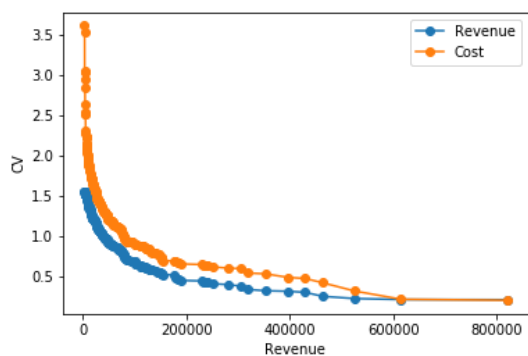


Figure A.5 Variation de cv pour le lundi avec $p = 35$ Figure A.6 Variation de cv pour le lundi avec $p = 40$

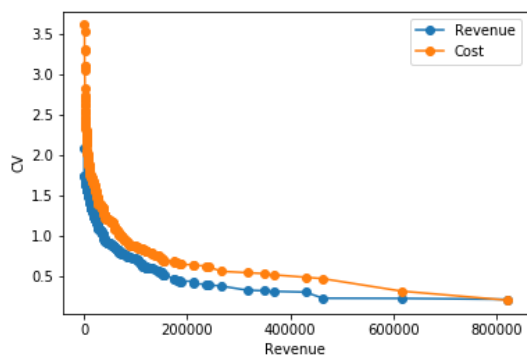
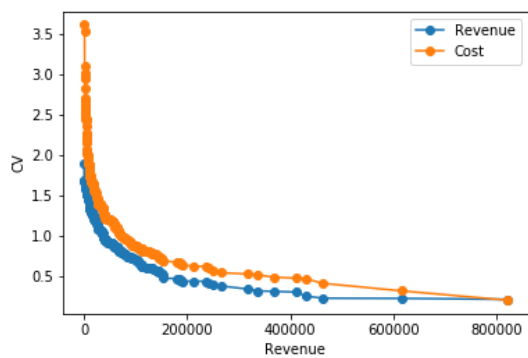


Figure A.7 Variation de cv pour le lundi avec $p = 45$ Figure A.8 Variation de cv pour le lundi avec $p = 50$

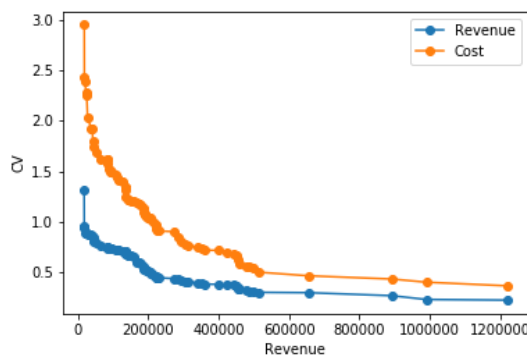
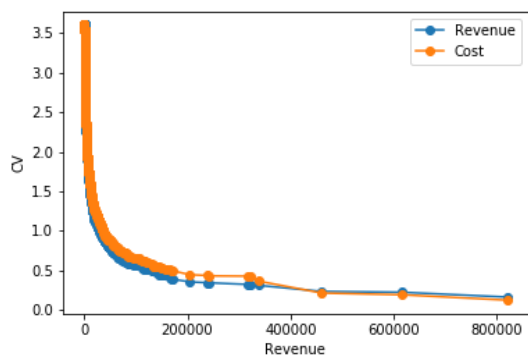
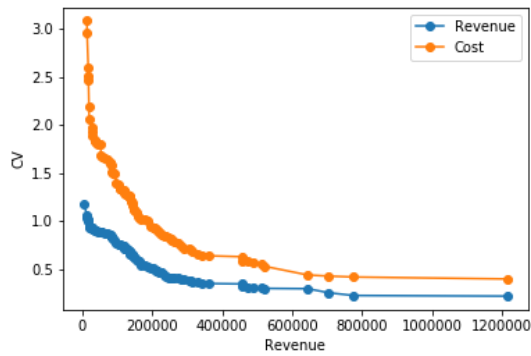
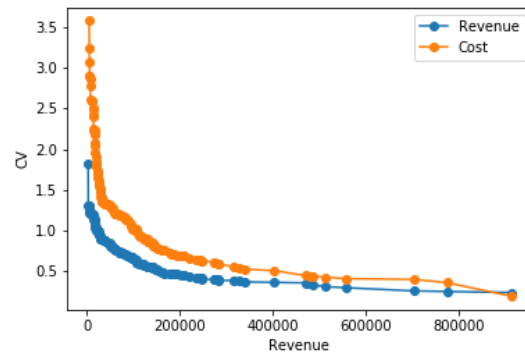
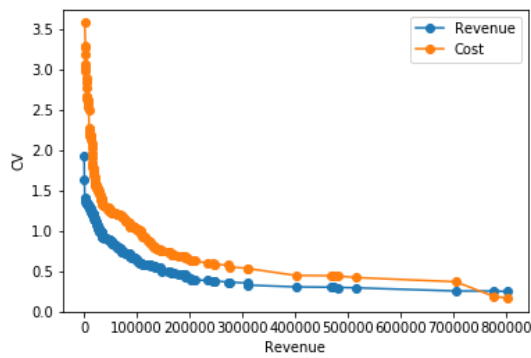
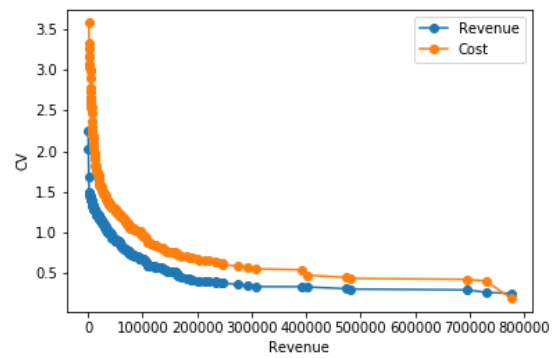
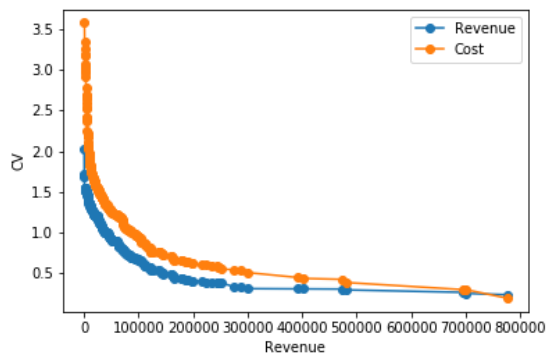
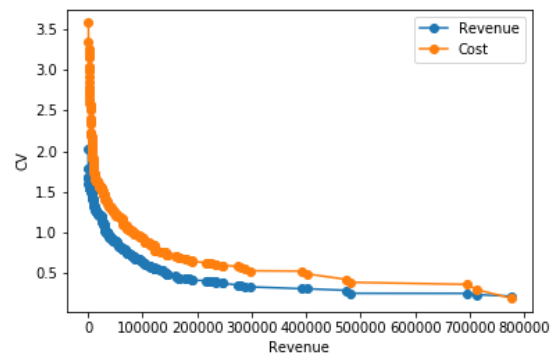


Figure A.9 Variation de cv pour le lundi sans clustering Figure A.10 Variation de cv pour le mardi avec $p = 15$

Figure A.11 Variation de cv pour le mardi avec $p = 20$ Figure A.12 Variation de cv pour le mardi avec $p = 25$ Figure A.13 Variation de cv pour le mardi avec $p = 30$ Figure A.14 Variation de cv pour le mardi avec $p = 35$ Figure A.15 Variation de cv pour le mardi avec $p = 40$ Figure A.16 Variation de cv pour le mardi avec $p = 45$

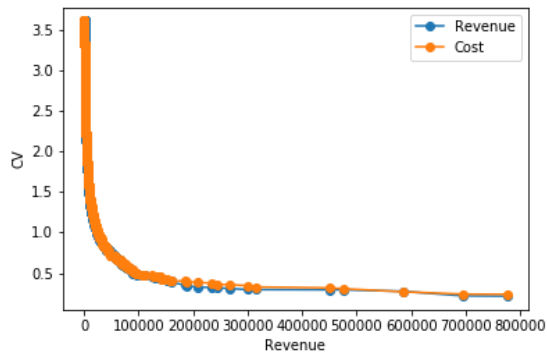


Figure A.17 Variation de cv pour le mardi sans clustering

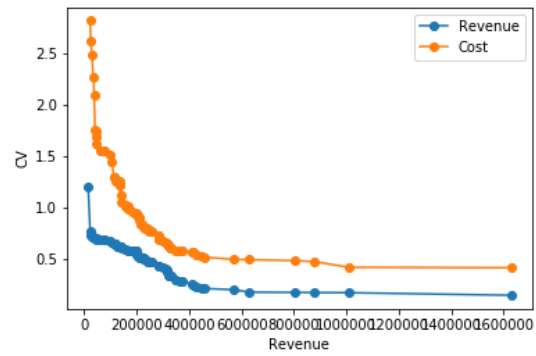


Figure A.18 Variation de cv pour le mercredi avec $p = 15$

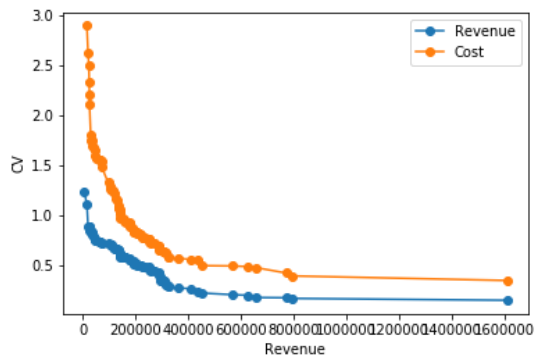


Figure A.19 Variation de cv pour le mercredi avec $p = 20$

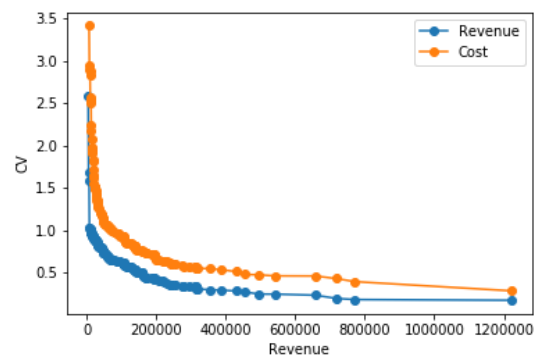


Figure A.20 Variation de cv pour le mercredi avec $p = 25$

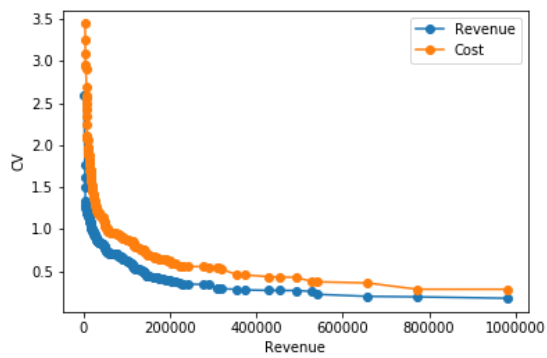


Figure A.21 Variation de cv pour le mercredi avec $p = 30$

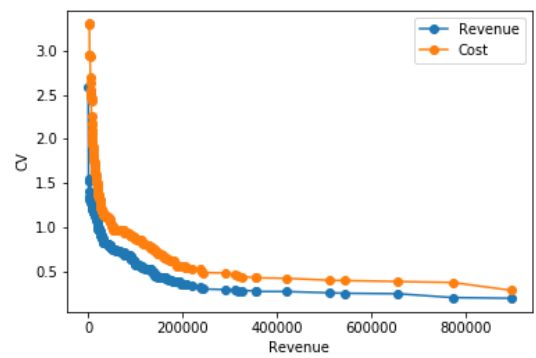


Figure A.22 Variation de cv pour le mercredi avec $p = 35$

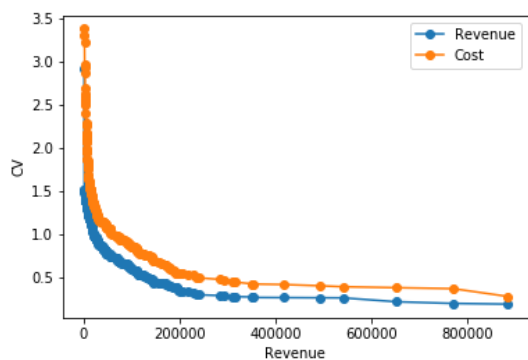


Figure A.23 Variation de cv pour le mercredi avec $p = 40$

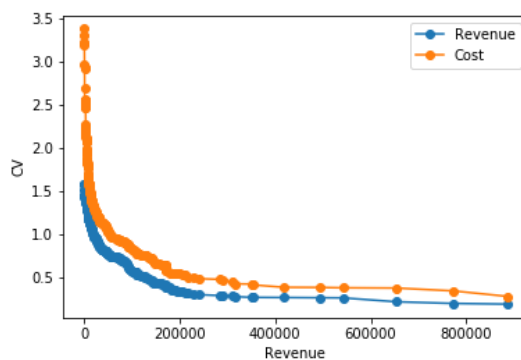


Figure A.24 Variation de cv pour le mercredi avec $p = 45$

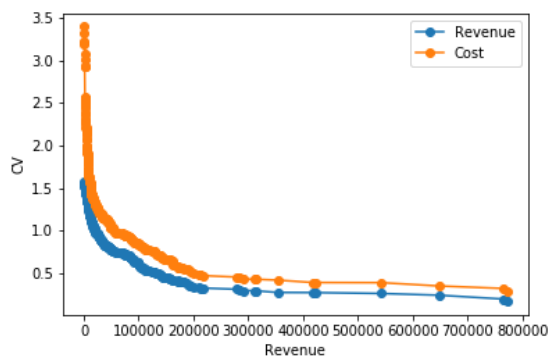


Figure A.25 Variation de cv pour le mercredi avec $p = 50$

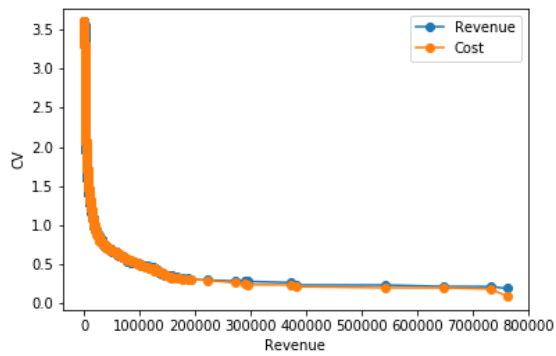


Figure A.26 Variation de cv pour le mercredi sans clustering

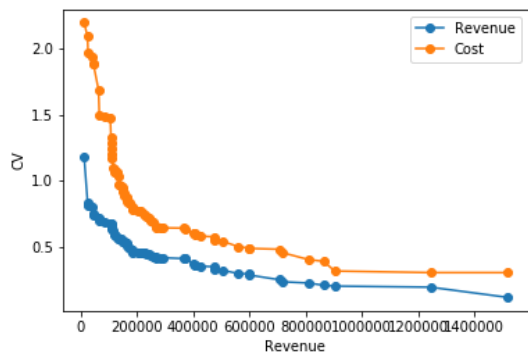


Figure A.27 Variation de cv pour le jeudi avec $p = 15$

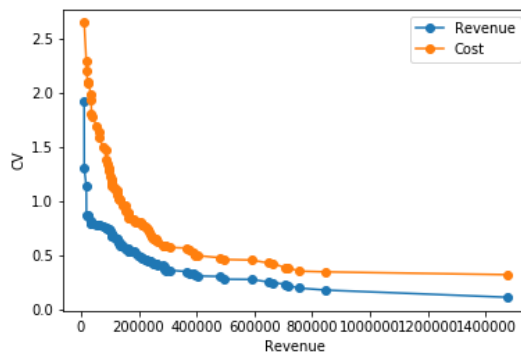


Figure A.28 Variation de cv pour le jeudi avec $p = 20$

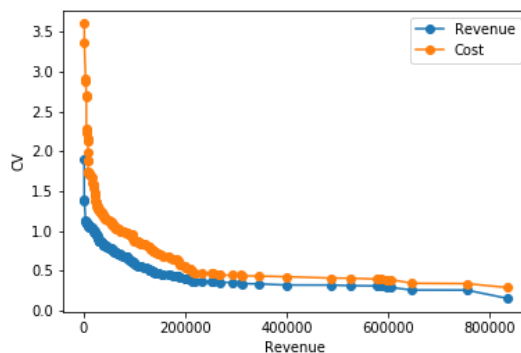
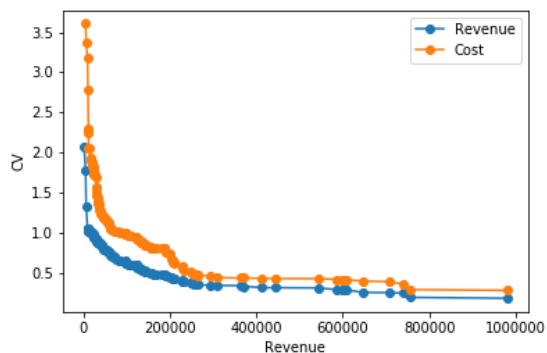


Figure A.29 Variation de cv pour le jeudi avec $p = 25$ Figure A.30 Variation de cv pour le jeudi avec $p = 30$

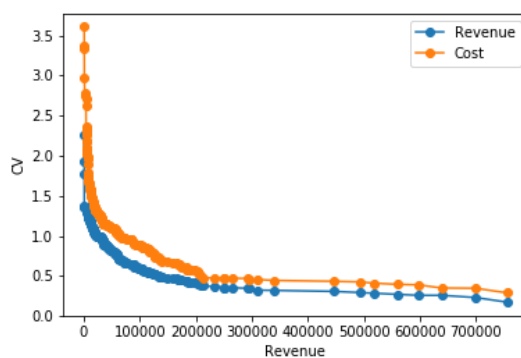
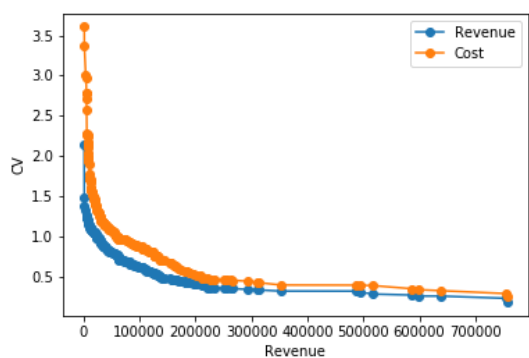


Figure A.31 Variation de cv pour le jeudi avec $p = 35$ Figure A.32 Variation de cv pour le jeudi avec $p = 40$

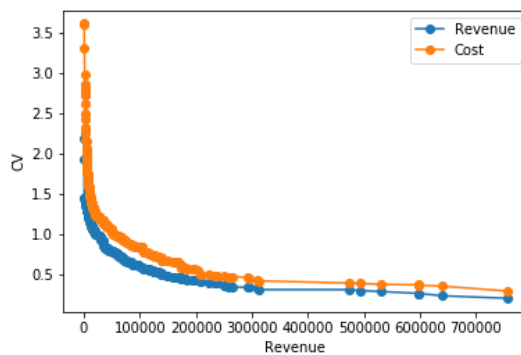
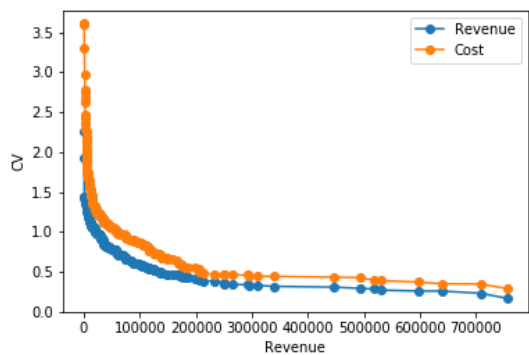


Figure A.33 Variation de cv pour le jeudi avec $p = 45$ Figure A.34 Variation de cv pour le jeudi avec $p = 50$

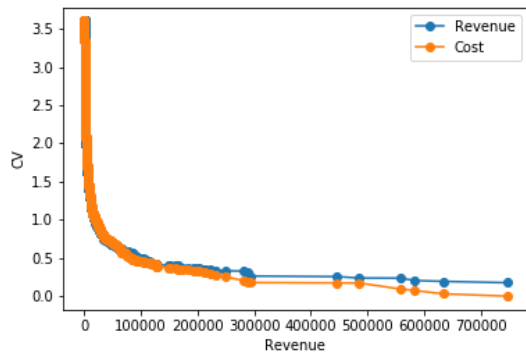


Figure A.35 Variation de cv pour le jeudi sans clustering

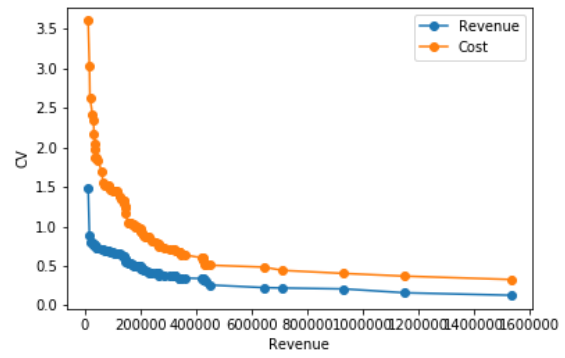


Figure A.36 Variation de cv pour le vendredi avec $p = 15$

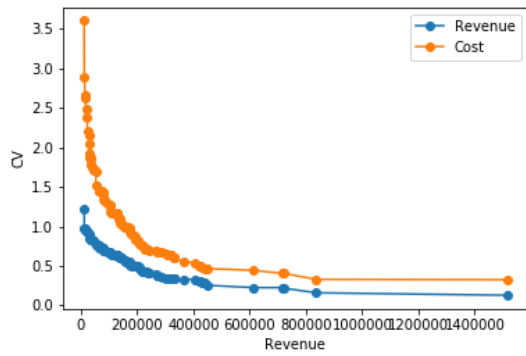


Figure A.37 Variation de cv pour le vendredi avec $p = 20$

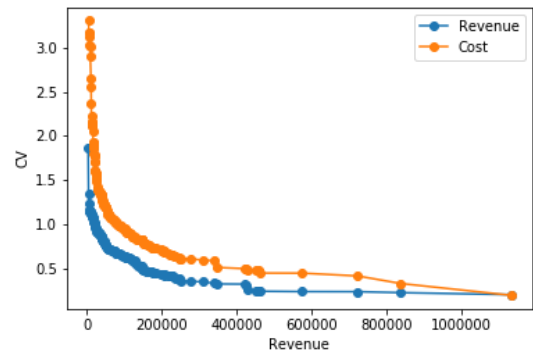


Figure A.38 Variation de cv pour le vendredi avec $p = 25$

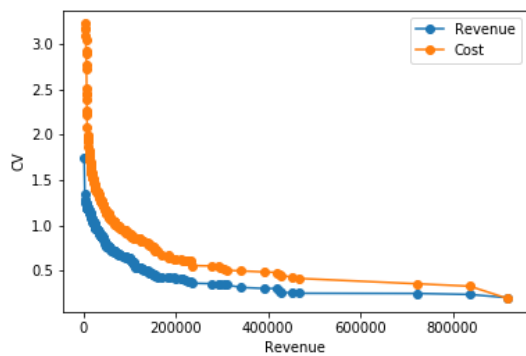


Figure A.39 Variation de cv pour le vendredi avec $p = 30$

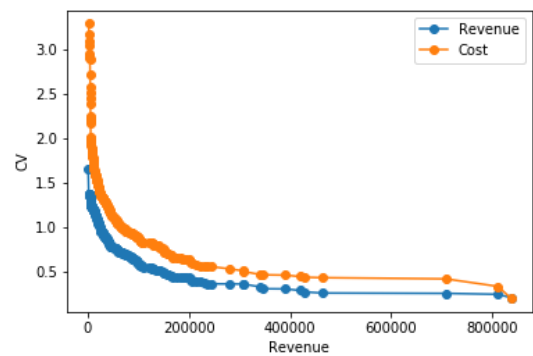


Figure A.40 Variation de cv pour le vendredi avec $p = 35$

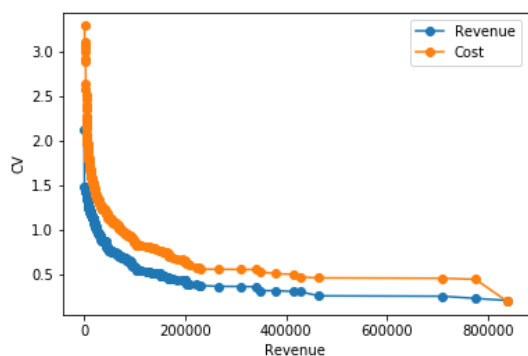


Figure A.41 Variation de cv pour le vendredi avec $p = 40$

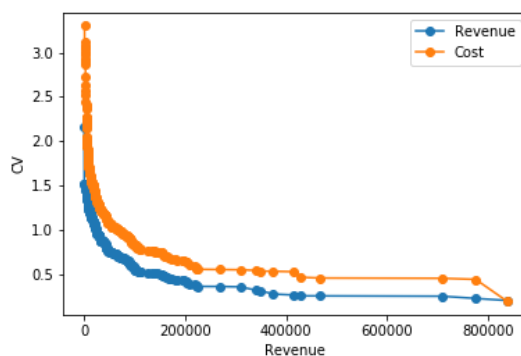


Figure A.42 Variation de cv pour le vendredi avec $p = 45$

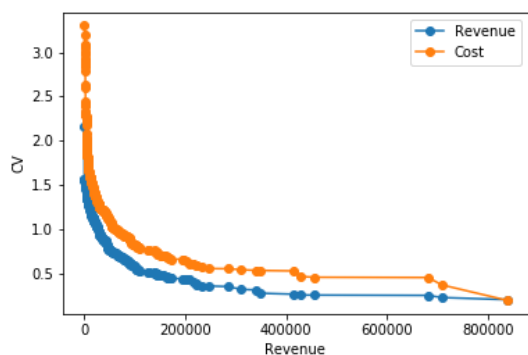


Figure A.43 Variation de cv pour le vendredi avec $p = 50$

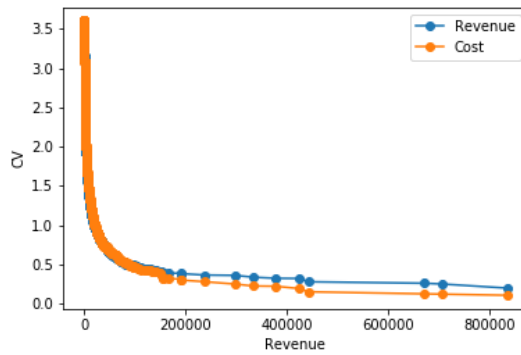


Figure A.44 Variation de cv pour le vendredi sans clustering

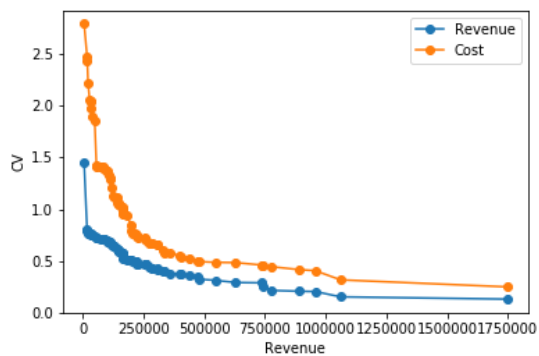


Figure A.45 Variation de cv pour le samedi avec $p = 15$

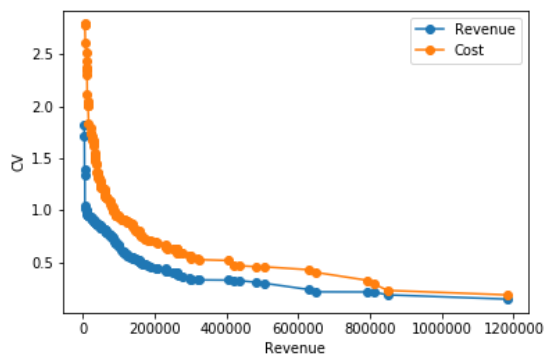


Figure A.46 Variation de cv pour le samedi avec $p = 20$

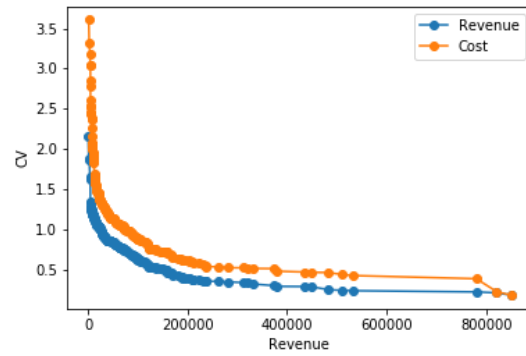
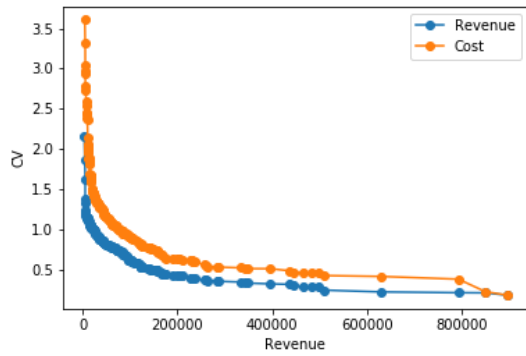


Figure A.47 Variation de cv pour le samedi avec $p = 25$ Figure A.48 Variation de cv pour le samedi avec $p = 30$

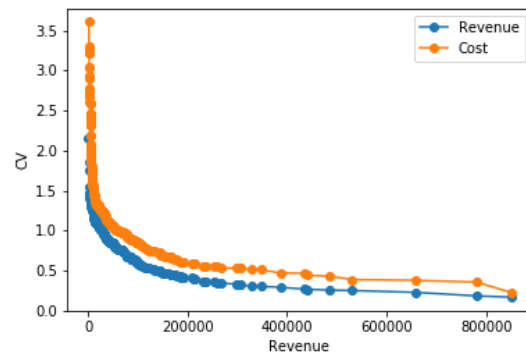
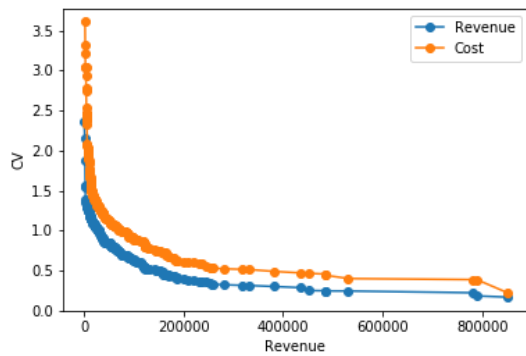


Figure A.49 Variation de cv pour le samedi avec $p = 40$ Figure A.50 Variation de cv pour le samedi avec $p = 50$

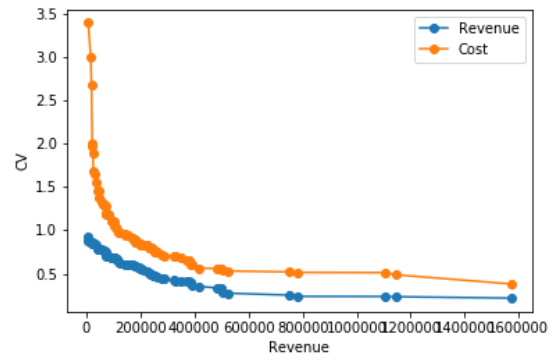
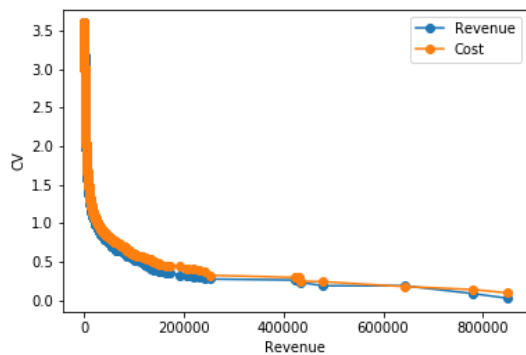


Figure A.51 Variation de cv pour le samedi sans clustering Figure A.52 Variation de cv pour le dimanche avec $p = 15$

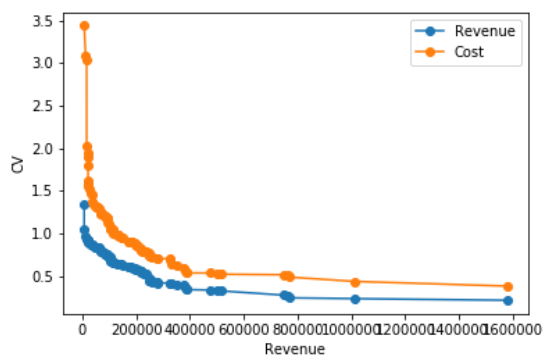


Figure A.53 Variation de cv pour le dimanche avec $p = 20$

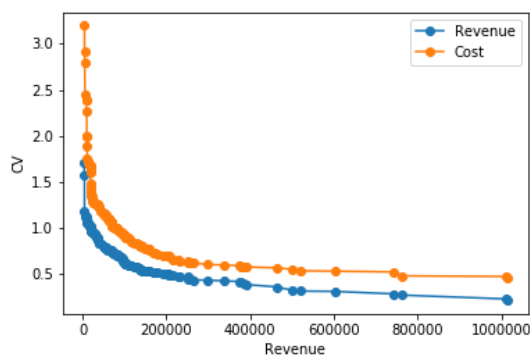


Figure A.54 Variation de cv pour le dimanche avec $p = 25$

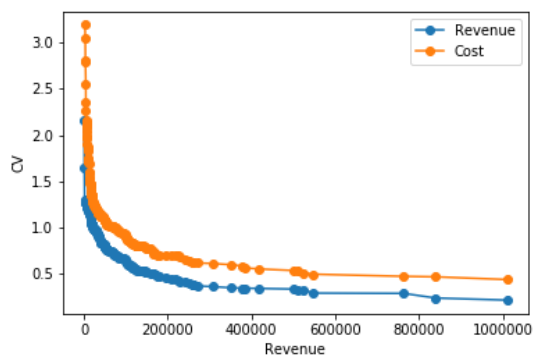


Figure A.55 Variation de cv pour le dimanche avec $p = 30$

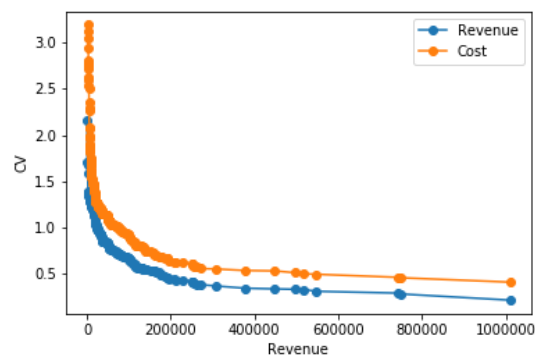


Figure A.56 Variation de cv pour le dimanche avec $p = 35$

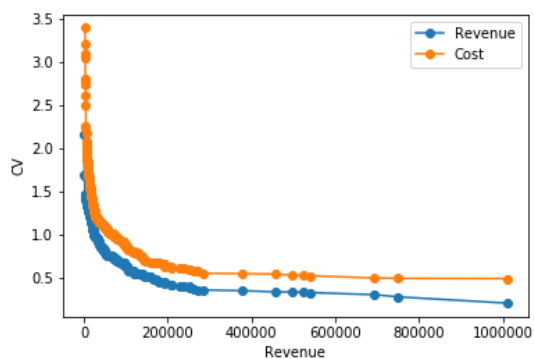


Figure A.57 Variation de cv pour le dimanche avec $p = 40$

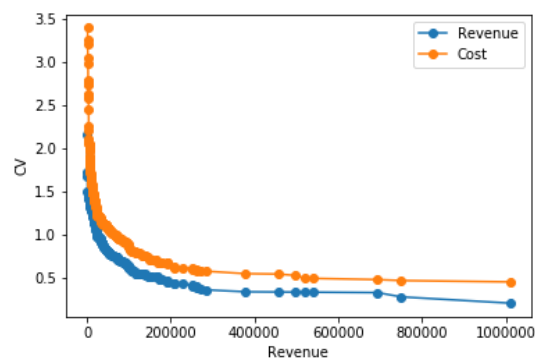


Figure A.58 Variation de cv pour le dimanche avec $p = 45$

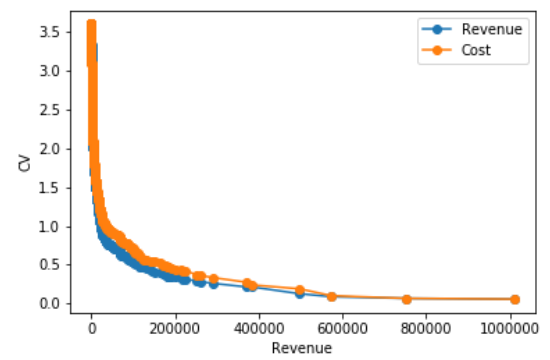
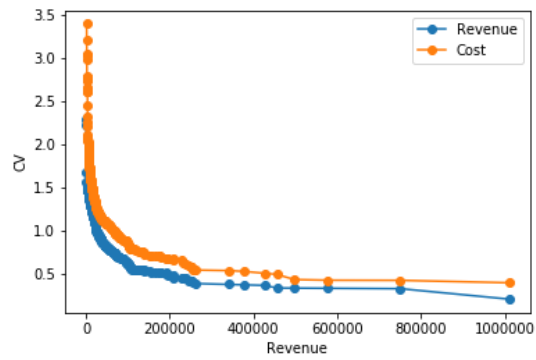


Figure A.59 Variation de cv pour le dimanche avec $p = 50$ Figure A.60 Variation de cv pour le dimanche sans clustering