Balancing cost and flexibility in Supply Chain

Etienne Gaillard de Saint Germain

École doctorale Mathématiques et Sciences et Technologies de l'Information et de la Communication

Thèse de doctorat

Spécialité : Mathématiques

Présentée par

Etienne Gaillard de Saint Germain

Pour obtenir le grade de

Docteur de l'Université Paris-Est

Balancing cost and flexibility in Supply Chain

Soutenance le 17 décembre 2018 devant le jury composé de :

Mme Nadia Brauner	*Université Grenoble Alpes*	Présidente
M. Jean-Philippe Gayon	*ISIMA*	Rapporteur
Mme Safia Kedad-Sidhoum	*CNAM*	Rapporteur
Mme Céline Gicquel	*Université Paris Sud*	Examinateur
M. Fabrice Bonneau	*Argon Consulting*	Encadrant de thèse
M. Vincent Leclère	*École des Ponts ParisTech*	Encadrant de thèse
M. Frédéric Meunier	*École des Ponts ParisTech*	Directeur de thèse

À Papa

Remerciements

Ces trois années de thèse écoulées, je prends conscience de la chance que j'ai eu d'être entouré par chacun d'entre vous.

Je tiens tout d'abord à remercier mes directeurs de thèse Vincent Leclère et Frédéric Meunier. Vous avez su me faire grandir personnellement et scientifiquement, m'aiguiller et m'offrir de l'autonomie sans cesser de veiller sur moi. Surtout, je vous suis reconnaissant pour avoir toujours su me rassurer, calmer mes doutes et me permettre d'effectuer ma thèse dans les meilleures conditions.

J'adresse mes sincères remerciements à Jean-Philippe Gayon et Safia Kedad-Sidhoum pour avoir accepté de rapporter ma thèse et m'avoir montré comment améliorer ce manuscrit. Je remercie également Nadia Brauner et Céline Gicquel d'être présentes pour ma soutenance.

Je souhaiterais adresser un remerciement particulier à Fabrice Bonneau, Directeur Général d'Argon Consulting pour l'intérêt des sujets proposés, sa disponibilité malgré son emploi du temps chargé et sa volonté d'investir et de s'investir dans cette recherche malgré les contraintes industrielles. Comprendre les enjeux humains qui se cachaient dernière les mathématiques fut un réel plaisir. Je remercie également toutes les personnes que j'ai pu côtoyer à Argon Consulting : celles qui m'ont transmis leur passion et leur connaissance de la Supply Chain, les fonctions supports sans qui le quotidien ne serait pas si simple, ceux avec qui j'ai pu tisser des liens d'amitié et particulièrement les membres de l'OSS 117.

Je remercie également tous les chercheurs du CERMICS pour leur disponibilité et en particulier Jean-Philippe Chancelier sans qui l'informatique aurait été beaucoup plus difficile et Bernard Lapeyre et Julien Reygner pour leur aide sur les probabilités. Je remercie également les doctorants qui ont su créer un environnement de travail et de détente en toute circonstance. En particulier, je repense à nos conversations surréalistes à des heures tout aussi improbables, à ce séjour en Auvergne et cette “petite ballade” qui aura finalement duré cinq heures et à notre séjour au ski. Enfin, je souhaiterais adresser un remerciement tout particulier à Isabelle Simunic, la Secrétaire Générale du CERMICS. Arrivé petit stagiaire en césure, tu m'as accueilli quelques semaines dans ton bureau, pendant lesquelles ta réserve de spéculoos a grandement diminuée. Tu t'es toujours pliée en quatre pour nous rendre la vie facile et tu nous montrais toujours ta bonne humeur malgré les difficultés de ton travail.

Je remercie également mes amis qui sont là aujourd'hui et qui me rendent la vie si agréable, que ce soit lors des sorties ou des soirées Donjons et Dragons. Je me libère également aujourd'hui de toute blague sur mon statut étudiant.

Finalement, mes derniers remerciements vont à ma famille. Merci Maman pour ton soutien, ta présence ton écoute. Merci d'être aussi rock-and-roll. Merci à mes frères et sœurs, Aleth, Louise, Simon et Matthias. Merci d'être si déjantés et motivés pour tout (même le pire !) et de supporter ma maniaquerie. Merci à Natacha, la fiancée de Louise, pour nos parties de tennis hebdomadaires et pour être devenue ma troisième sœur. Enfin, tu n'es pas présent aujourd'hui malgré ton importance : merci Papa. Tu as été là à toutes les étapes de ma vie et particulièrement pendant cette thèse. Pendant deux ans et demi, j'ai pu avoir cette relation privilégiée avec toi lorsque je vivais sous ton toit. Tu remplissais tous les rôles : le père, le colocataire, l'ami, le confident. Si cette thèse a pu si bien se dérouler, je te le dois en grande partie.

Abstract

This thesis develops optimization methods for Supply Chain Management and is focused on the flexibility defined as the ability to deliver a service or a product to a customer in an uncertain environment. The research was conducted throughout a CIFRE agreement between Argon Consulting, which is an independent consulting firm in Supply Chain Operations and the École des Ponts ParisTech. In this thesis, we explore three topics that are encountered by Argon Consulting and its clients and that correspond to three different levels of decision (long-term, mid-term and short-term).

When companies expand their product portfolio, they must decide in which plants to produce each item. This is a long-term decision since once it is decided, it cannot be easily changed. More than an assignment problem where one item is produced by a single plant, this problem consists in deciding if some items should be produced by several plants and by which ones. This is motivated by a highly uncertain demand. So, in order to satisfy the demand, the assignment must be able to balance the workload between plants. We call this problem the multi-sourcing of production. Since it is not a repeated problem, it is essential to take into account the risk when making the multi-sourcing decision. We propose a generic model that includes the technical constraints of the assignment and a risk-averse constraint based on risk measures from financial theory. We develop an algorithm and a heuristic based on standard tools from Operations Research and Stochastic Optimization to solve the multi-sourcing problem and we test their efficiency on real datasets.

Before planning the production, some macroscopic indicators, such as the quantity of raw materials to order or the size of produced lots, must be decided at mid-term level. Continuous-time inventory models are used by some companies but these models often rely on a trade-off between holding costs and setup costs. These latters are fixed costs paid when production is launched and are hard to estimate in practice. On the other hand, at mid-term level, flexibility of the means of production is already fixed and companies easily estimate the maximal number of setups. Motivated by this observation, we propose extensions of some classical continuous-time inventory models with no setup costs and with a bound on the number of setups. We used standard tools from Continuous Optimization to compute the optimal macroscopic indicators.

Finally, planning the production is a short-term decision consisting in deciding which items must be produced by the assembly line during the current period. This problem belongs to the well-studied class of Lot-Sizing Problems. As for mid-term decisions, these problems often rely on a trade-off between holding and setup costs. Basing our model on industrial considerations, we keep the same point of view (no setup cost and a bound on the number of setups) and propose a new model. Although these are short-term decisions, production decisions must take future demand into account, which remains uncertain. We solve our production planning problem using standard tools from Operations Research and Stochastic Optimization, test the efficiency on real datasets, and compare it to heuristics used by Argon Consulting's clients.

Key words: Assignment Problem, Heuristics, Lot-Sizing, Operations Research, Risk Measure, Stochastic Optimization, Supply Chain Management.

Résumé

Cette thèse développe des méthodes d'optimisation pour la gestion de la Supply Chain et a pour thème central la flexibilité définie comme la capacité à fournir un service ou un produit au consommateur dans un environnement incertain. La recherche a été menée dans le cadre d'une convention CIFRE entre Argon Consulting, une société indépendante de conseil en Supply Chain, et l'École des Ponts ParisTech. Dans cette thèse, nous étudions trois sujets rencontrés par Argon Consulting et ses clients et qui correspondent à trois différents niveaux de décision (long terme, moyen terme et court terme).

Lorsque les entreprises élargissent leur portefeuille de produits, elles doivent décider dans quelles usines produire chaque article. Il s'agit d'une décision à long terme, car une fois qu'elle est prise, elle ne peut être facilement modifiée. Plus qu'un problème d'affectation où un article est produit par une seule usine, ce problème consiste à décider si certains articles doivent être produits par plusieurs usines et par lesquelles. Cette interrogation est motivée par la grande incertitude de la demande. En effet, pour satisfaire la demande, l'affectation doit pouvoir équilibrer la charge de travail entre les usines. Nous appelons ce problème le multi-sourcing de la production. Comme il ne s'agit pas d'un problème récurrent, il est essentiel de tenir compte du risque au moment de décider le niveau de multi-sourcing. Nous proposons un modèle générique qui inclut les contraintes techniques du problème et une contrainte d'aversion au risque basée sur des mesures de risque issues de la théorie financière. Nous développons un algorithme et une heuristique basés sur les outils standard de la Recherche Opérationnelle et de l'Optimisation Stochastique pour résoudre le problème du multi-sourcing et nous testons leur efficacité sur des données réelles.

Avant de planifier la production, certains indicateurs macroscopiques doivent être décidés à moyen terme tels la quantité de matières premières à commander ou la taille des lots produits. Certaines entreprises utilisent des modèles de stock en temps continu, mais ces modèles reposent souvent sur un compromis entre les coûts de stock et les coûts de lancement. Ces derniers sont des coûts fixes payés au lancement de la production et sont difficiles à estimer en pratique. En revanche, à moyen terme, la flexibilité des moyens de production est déjà fixée et les entreprises estiment facilement le nombre maximal de lancements. Poussés par cette observation, nous proposons des extensions de certains modèles classiques de gestion de stock en temps continu, sans coût de lancement et avec une limite sur le nombre de lancements. Nous avons utilisé les outils standard de l'Optimisation Continue pour calculer les indicateurs macroscopiques optimaux.

Enfin, la planification de la production est une décision à court terme qui consiste à décider quels articles doivent être produits par la ligne de production pendant la période en cours. Ce problème appartient à la classe bien étudiée des problèmes de Lot-Sizing. Comme pour les décisions à moyen terme, ces problèmes reposent souvent sur un compromis entre les coûts de stock et les coûts de lancement. Fondant notre modèle sur ces considérations industrielles, nous gardons le même point de vue (aucun coût de lancement et une borne supérieure sur le nombre de lancements) et proposons un nouveau modèle. Bien qu'il s'agisse de décisions à court terme, les décisions de production doivent tenir compte de la demande future, qui demeure incertaine. Nous résolvons notre problème de planification de la production à l'aide d'outils standard de Recherche Opérationnelle et d'Optimisation Stochastique, nous testons l'efficacité sur des données réelles et nous la comparons aux heuristiques utilisées par les clients d'Argon Consulting.

Mots-clés : Heuristique, Lot-Sizing, Mesure de Risque, Optimisation Stochastique, Problème d'affectation, Recherche Opérationnelle, Supply Chain Management.

Contents

List of Figures

List of Tables

1 Introduction

The research made in this CIFRE thesis deals with Supply Chain Management. It was funded by Argon Consulting which is an independent consulting firm whose mission is to help its clients improve every part of their Supply Chain (from the procurement of raw materials to the delivery of final products) and conducted throughout an industrial partnership with the École des Ponts ParisTech. The goal is to model and develop methods to manage specific parts of the Supply Chain in an optimal way.

The common thread of the three topics developed in this thesis is the *flexibility*. We define the flexibility as the ability to deliver a service or a product to a customer in an uncertain environment. Depending on the level of decision, the flexibility is either a constraint (like the ability of an assembly line to easily switch from the production of one item to another) or a decision variable (like deciding between specialization and versatility). In general, the flexibility of a system relies on long-term (and sometimes mid-term) decisions.

In order to help to the global understanding of these topics, we choose to introduce the three topics beginning by the long-term decisions, then the mid-term decisions, and finally the short-term decisions. However, this manuscript follows a different order prescribed by the introduction of notions and results reused in following parts. The long-term decisions studied in this thesis (Part III) deal with multi-sourcing of production that aims at deciding the flexibility of means of production at a reasonable cost. The mid-term decisions (Part I) and the short-term decisions (Part II) both deal with the reduction of inventories subject to flexibility decisions that were already made. More specifically, the mid-term decisions we are interested in aim at computing indicators that drive several Supply Chain processes whereas short-term decisions aim at deciding the production that must be launched.

The three topics of this thesis and other examples are placed on Figure 1.1 depending on their decision horizon.

1.1 Multi-sourcing of production

Multi-sourcing of production is a strategic decision in Supply Chain Management (i.e., a long-term decision). It consists in deciding if a plant should have the ability to produce an item. For example, in Figure 1.2, the first plant can produce items A, B and C while the second one can produce items B, C and D. Then, items A and D are said to be mono-sourced since each can

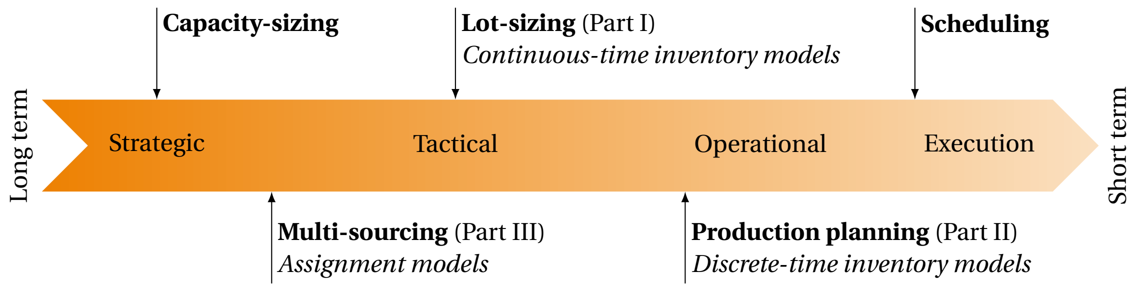


Figure 1.1 – Decision horizon

be produced by only one plant whereas items B and C are said to be multi-sourced since they can be produced by at least two plants. The first characteristic of multi-sourcing decisions is their horizon. They take time to be implemented and have long-term impact on Supply Chain Management. Second, multi-sourcing decisions are taken in a highly uncertain environment. Among others, the future customer demand, the reliability of means of production and the future availability of raw materials are imperfectly known. Finally, multi-sourcing decisions will constrain future production decisions (i.e., mid-term decision). Precisely, they determine the flexibility of the plants and the ability to balance the workload between them.

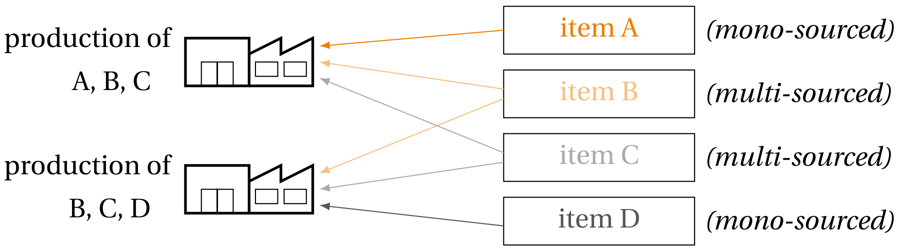


Figure 1.2 – Multi-sourcing of production of four items in two plants

Considering its applications, Argon Consulting chooses to model the demand as the main source of uncertainty with a fixed and known total volume of demand. (In its applications, Argon Consulting is interested by the ability to face variations in the product mix and not in the volume of demand.) In Chapter 9, we model the problem as a *stochastic program* with recourses where first-stage variables are the assignment of items to plants and second-stage variables are the production decisions. In order to deal with randomness and to capture the long-term impact and the risk of assignment decisions, we rely on *risk measures*, which are tools from financial theory used to quantify the risk of a financial position. We choose to use the *Average-Value-at-Risk* (AV@R) applied to the inventory level of items. To the best of our knowledge, it is almost the first time that such a tool is used in Supply Chain applications. High inventory level is expensive but enables an easy satisfaction of the demand. Reducing inventory is then risky and Average-Value-at-Risk aims at quantifying the risk of this decision.

The Average-Value-at-Risk at α% (also known as Expected Shortfall or Conditional-Value-at-Risk) can be interpreted as the expectation restricted to the α% worst cases, i.e., α% lowest values of inventory. It enables the decision maker to have an indicator that captures both the shortfall probability and the undelivered quantity (which are strongly linked to two indicators used to measure service level: the cycle service level and the fill rate service level). Moreover, the parameter α provides a simple way to address the control of the risk level and the Average-Value-at-Risk can be linearized. We apply a classical approximation scheme to solve the stochastic program by doing first a two-stage approximation and then a sampling of scenarios in order to get a *mixed integer linear program (MILP)*, which leads to a tractable formulation.

Real datasets given by Argon Consulting's clients contain only historical values of production and sales. Since we do not know the actual demand, we propose in Chapter 7 a probabilistic model to generate possible realizations of demand from historical values. This model is based on *Dirichlet distributions* and aims at being easy to use while being reasonable. Its only input is a forecast demand (which can be the historical sales or the historical forecast) and a volatility which is a percentage standing for the accuracy of the forecast. (The smallest the volatility the most accurate the forecast.) Our probabilistic model provides scenarios of demand such that the total volume of demand is the same in each scenario, such that the expectation of a realization is equal to the forecast and such that the standard deviation divided by the expectation is close to the volatility. This model meets the assumption made by Argon Consulting on the demand, has few parameters and is easy to simulate (even conditionally to the past).

Finally, on real datasets, computation times may be long. Up-to-date solvers are often unable to find a feasible solution of the problem. Then, we propose a heuristic that enables us to quickly find a feasible solution of the multi-sourcing problem. The returned solution can be directly used by Argon Consulting's clients or as an initial solution of a generic solver.

The experimentations made in Chapter 10 already prove that the company that provides the datasets can reduce its proportion of multi-sourced items (thus reduce its costs) while keeping a good ability to satisfy demand. However, computer performances and real dataset sizes prevent us from dealing with more than a hundred of scenarios. Thus, the method is dependent on the sampling methods and the choice of a representative set of scenarios is critical. We propose a concrete way to reduce this dependence on the sampling methods based on clustering methods (such as K-means).

1.2 Continuous-time inventory models

Argon Consulting uses continuous-time inventory models to compute macroscopic parameters at a tactical level (mid-term decisions). Classical examples are the *lot-size* and the *cover-size*. The lot-size gives the quantity of a same item produced at each production launch. The cover-size gives the number of time units following a production launch during which inventory must be positive. These parameters are used as input for other decision processes in Supply Chain such as the Material Requirement Planning (MRP). For example, having an estimate of the lot-sizes or the cover-sizes allows to decide the quantity of raw materials that must be ordered. They are also used as input to plan the production since they give the sizes of produced lots. (When studying discrete-time models, we will propose to remove this constraint from the models.)

Continuous-time inventory models assume a continuous vision of time. The seminal model known as the *Economic Order Quantity (EOQ)* model from Harris (1913) gives the optimal trade-off between ordering and holding costs. In practice, EOQ formula is hard to use since ordering and holding costs are difficult to compare. Argon Consulting aims at finding the optimal lot-sizes (or cover-sizes) from the flexibility of its assembly lines, which is considered as fixed and defined by previous Supply Chain processes. An assembly line can produce several items but loses time when switching from the production of one item to another. Considering a constant demand rate, Figure 1.3 shows the consequences of several lengths of production cycles. Too short production cycles lead to stock-out since too much time is spent switching between the production of two items whereas too long production cycles lead to unnecessary high inventory and unproductive time. In real datasets, assembly lines produce a lot of items and time lost due to switching between different items is modeled by a maximal number of setups.

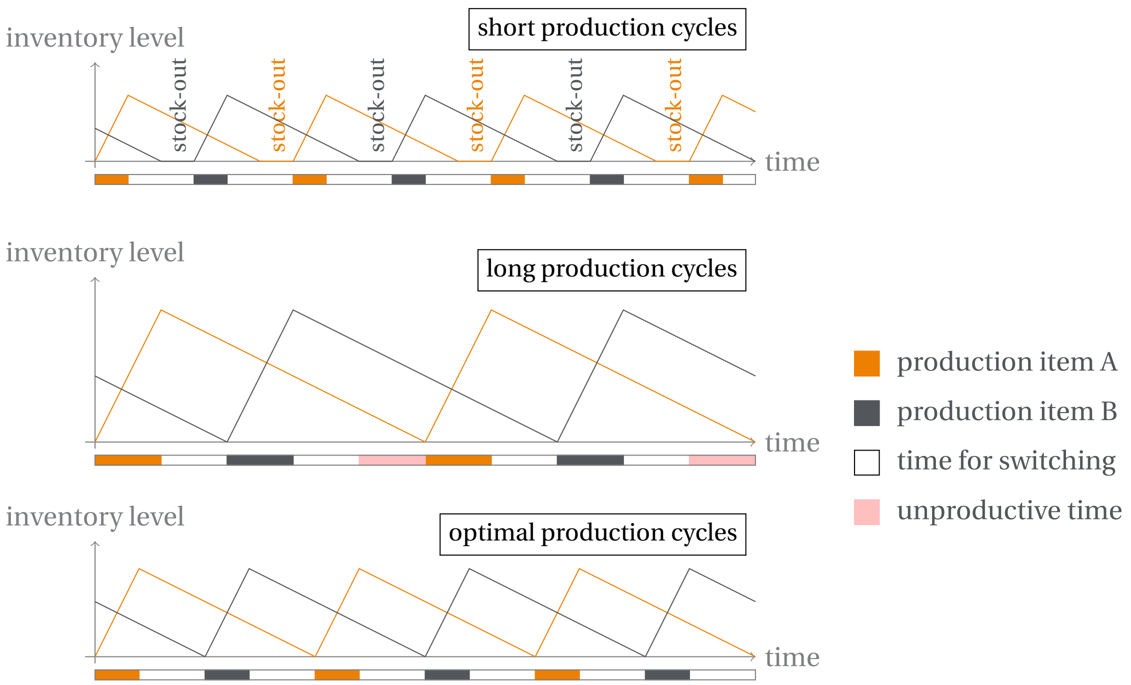


Figure 1.3 – Continuous-time inventory model for a line producing two items

Replacing the ordering costs by an upper bound on the number of setups, we propose in Chapter 3 generalizations of the classical EOQ formula for multiple items. They have already been useful for Argon Consulting. Specifically, we study continuous and integer numbers of setups in deterministic and stochastic settings. Moreover, we also study in Chapter 4 an extension that considers several parallel assembly lines and show that the problem can be stated as a concave minimization problem over a polyhedron (for which it exists an extensive literature).

1.3 Discrete-time inventory models

Discrete-time inventory models (also called *dynamic lot-sizing problems*) assume that time is decomposed into discrete periods. They are used by companies to plan their short-term production. A classical model is the *Capacitated Lot-Sizing Problem (CLSP)*. It considers an assembly line producing several items during a finite number of periods. The demand for each item is deterministic and given for each period. It aims at minimizing the sum of the holding costs (due to inventory carried from a period to the following) and the setup costs (fixed cost due to launch of the production) subject to the capacity of the assembly line.

As already mentioned in Section 1.2, the drawback of this formulation according to Argon Consulting and its clients is the difficulty to estimate the value of the setup costs. On the other hand, estimating the maximal number of setups for a period is an easy task for Argon Consulting's clients. We propose in Chapter 5 a model derived from the Capacitated Lot-Sizing Problem that replaces the setup costs by an upper bound on the number of setups. Figure 1.4 shows an example of production planning of four items when at most two items can be produced during a period. To the best of our knowledge, this model is new and has not been studied in the literature.

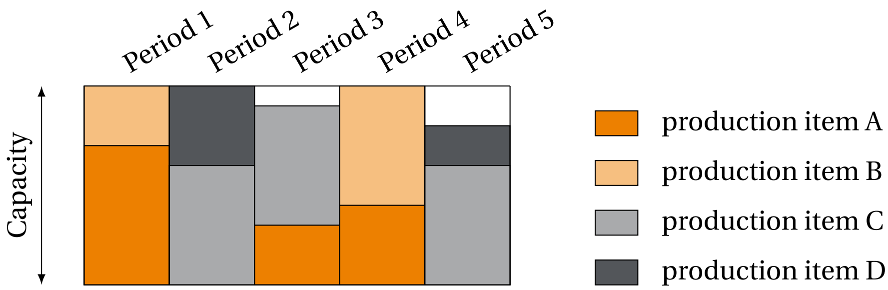


Figure 1.4 – Production planning of four items for five weeks

Our lot-sizing problem can be written as a *mixed integer linear program (MILP)*. We get several theoretical results in the deterministic setting that show the difficulties of the problem. As expected, this problem is NP-hard. A classical method to help solve mixed integer linear programs consists in relaxing some constraints to get a bound on the optimal value of the problem. We show that several natural formulations all yield the same continuous relaxation. Finally, we were left with the following question: what is the complexity status of our lot-sizing problem when there is no capacity constraints and when the maximal number of setups per period is equal to 1? Mathematically, it can be formulated as follows. Consider the problem

$$
\begin{array}{lll}
\min & \sum_{t=1}^{T} \sum_{i \in \mathcal{I}} h^i s_t^i & \\
\text{s.t.} & s_t^i = s_{t-1}^i + q_t^i - d_t^i & \forall t \in [T],\ \forall i \in \mathcal{I}, \\
& q_t^i \leqslant M x_t^i & \forall t \in [T],\ \forall i \in \mathcal{I}, \\
& \sum_{i \in \mathcal{I}} x_t^i \leqslant 1 & \forall t \in [T], \\
& x_t^i \in \{0,1\} & \forall t \in [T],\ \forall i \in \mathcal{I}, \\
& q_t^i,\ s_t^i \geqslant 0 & \forall t \in [T],\ \forall i \in \mathcal{I},
\end{array}
\tag{P}
$$

where M is a big positive number, and for each period t and each item i, the demand d_t^i and the holding cost h^i are given nonnegative real numbers, and the inventory s_t^i, the produced quantity q_t^i and the setup indicator x_t^i are the decision variables.

Open question. What is the complexity status of (P)?

In practice, the demand is not always deterministic. We propose in Chapter 6 a stochastic version of our lot-sizing problem based on *Stochastic Programming* (see Section 1.1). The difference is that we do not use a risk-averse constraint (AV@R) but stick to the classical risk-neutral vision (the expectation). Indeed, production is a repeated decision and a failure at one period is easy to compensate with another period.

Moreover, in a stochastic setting, we must allow backorder because production resources are limited and it may be impossible to cover every single possible realizations of demand. Here, they come with costs in the objective function. Except when they are enshrined through contracts with the customers, backorder costs can be hard to estimate. We adapt from the literature a method based on the *news-vendor problem* (one of the oldest stochastic models) to link the backorder cost and the desired portion of satisfied demand.

As in Section 1.1, our stochastic lot-sizing problem is also solved by doing first a two-stage approximation and then a sampling of scenarios in order to get a mixed integer linear program. Since scenarios were not provided by our partner, we generate them using the probabilistic model mentioned in Section 1.1.

The experimentations made in Chapter 7 seem to show that the company that provides the datasets can reduce its inventory costs while keeping a good ability to satisfy the demand. However, as in Section 1.1, computer performances, real dataset sizes and limited time to return a production planning prevent from dealing with more than twenty scenarios. Since the method is dependent on the sampling methods, we propose again a concrete way to reduce this dependence on the sampling methods based on clustering methods (such as K-means).

1.4 Extensions

In Part IV, we present two extensions of our work. The first is an alternative version of the multi-sourcing problem addressed in Part III. The difference is that the company has a limited budget to invest in flexibility. In this case, the company aims at deciding an assignment which maximizes the demand that can be satisfied. We model this alternative problem and show that it is NP-hard in several simple cases.

The second is an extension of the inventory models addressed in Part I and II. We aim at computing the cover-sizes at mid-term horizon using a model relying on production planning at short-term. We model this alternative problem and experimentally show that it has many drawbacks compared to the continuous-time inventory models.

1 Introduction (version française)

La recherche effectuée dans le cadre de cette thèse CIFRE porte sur la gestion de la Supply Chain. Elle a été financée par Argon Consulting, cabinet de conseil indépendant dont la mission est d'aider ses clients à améliorer l'ensemble de leur Supply Chain (de l'approvisionnement en matières premières à la livraison des produits finis) et est menée dans le cadre d'un partenariat industriel avec l'École des Ponts ParisTech. L'objectif est de modéliser et de développer des méthodes pour gérer de manière optimale certaines fonctions spécifiques de la Supply Chain.

Le point commun des trois sujets développés dans cette thèse est la flexibilité. Nous définissons la flexibilité comme la capacité à fournir un service ou un produit à un client dans un environnement incertain. Selon le niveau de décision, la flexibilité est soit une contrainte (comme la capacité d'une ligne de production à passer facilement de la production d'un article à un autre) soit une variable de décision (comme décider entre spécialisation et polyvalence). En général, décider de la flexibilité d'un système est une décision à long terme (et parfois à moyen terme).

Afin d'aider à la compréhension globale des sujets, nous avons choisi d'introduire les trois sujets en commençant par les décisions à long terme, puis les décisions à moyen terme, et enfin les décisions à court terme. Cependant, ce manuscrit suit un ordre différent dû à l'introduction de notions et de résultats réutilisés d'une partie sur l'autre. Les décisions à long terme étudiées dans cette thèse (Partie III) portent sur le multi-sourcing de la production qui vise à décider de la flexibilité des moyens de production tout en gardant un coût raisonnable. Les décisions à moyen terme (Partie I) et les décisions à court terme (Partie II) traitent toutes deux de la réduction des stocks sous la contrainte des décisions de flexibilité déjà prises. Cependant, les décisions à moyen terme qui nous intéressent visent à calculer des indicateurs qui pilotent plusieurs processus de la Supply Chain alors que les décisions à court terme qui nous intéressent visent à décider de la production qui doit être lancée.

Les trois sujets de cette thèse et d'autres exemples sont placés sur la Figure 1.1 en fonction de leur horizon de décision.

1.1 Multi-sourcing de la production

Le multi-sourcing de la production est une décision stratégique dans la gestion de la Supply Chain (*i.e.* une décision à long terme). Elle consiste à décider si une usine doit avoir la capacité de produire un article. Par exemple, dans la Figure 1.2, la première usine peut produire les

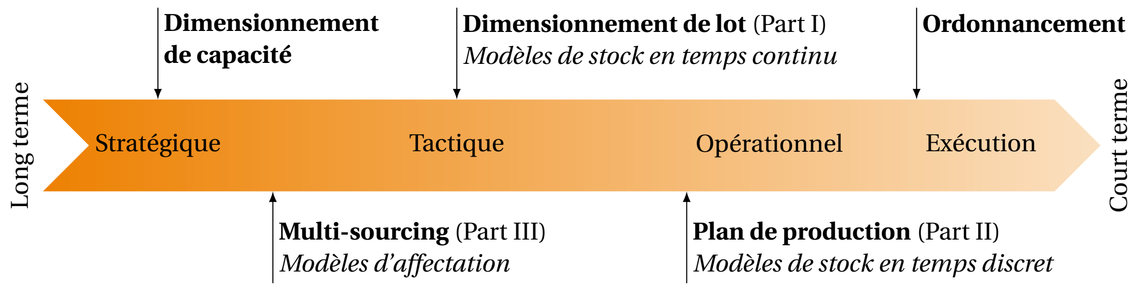


Figure 1.1 – Horizon de décision

articles A, B et C tandis que la seconde peut produire les articles B, C et D. Les articles A et D sont dits mono-sourcés puisque chacun peut être produit par une seule usine alors que les articles B et C sont dits multi-sourcés puisqu'ils peuvent être produits par au moins deux usines. La première caractéristique des décisions de multi-sourcing est leur horizon. Leur mise en œuvre prend du temps et a un impact à long terme sur la gestion de la Supply Chain. Deuxièmement, les décisions de multi-sourcing sont prises dans un environnement très incertain. Entre autres, les demandes futures des clients, la fiabilité des moyens de production ou la disponibilité future des matières premières sont mal connues. Enfin, les décisions de multi-sourcing contraignent les décisions de production futures (décision à moyen terme). Plus précisément, elles déterminent la flexibilité des usines et la capacité à équilibrer la charge de travail entre elles.

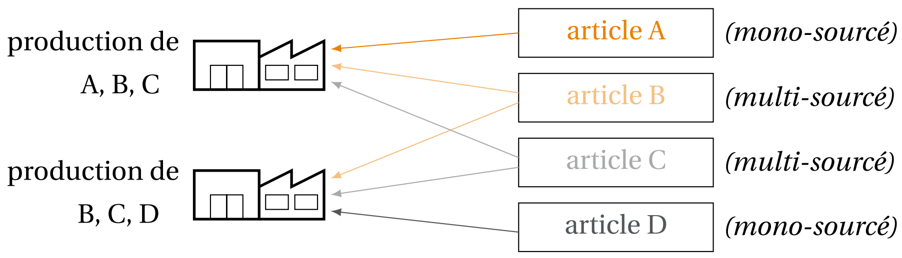


Figure 1.2 – Multi-sourcing de la production de quatre articles sur deux usines

Considérant ses applications, Argon Consulting choisit de modéliser la demande comme principale source d'incertitude avec un volume total de demande fixe et connu. (Dans ses applications, Argon Consulting s'intéresse à la capacité de faire face aux variations du mix produit et non du volume de la demande.) Dans le Chapitre 9, nous modélisons le problème comme un *programme stochastique* avec recours où les variables de première étape sont l'affectation des articles aux usines et les variables de deuxième étape sont les décisions de production. Afin d'intégrer le caractère aléatoire et de prendre en compte l'impact à long terme et le risque des décisions d'affectation, nous nous appuyons sur les *mesures de risque*, qui sont des outils de la théorie financière utilisés pour quantifier le risque lié à une position financière. Nous choisissons d'utiliser l'Average-Value-at-Risk (AV@R) appliquée au niveau de stock des articles. À' notre connaissance, un tel outil a rarement été utilisé dans des applications Supply Chain. Un niveau de stock élevé est coûteux mais permet de satisfaire facilement la demande. La réduction des stocks est alors risquée et l'Average-Value-at-Risk vise à quantifier le risque lié à cette décision.

L'Average-Value-at-Risk à α% (aussi connue sous le nom d'Expected Shortfall ou de Conditionnal-

Value-at-Risk) peut être interprétée comme l'espérance restreinte aux α% pires cas, *i.e.* au α% plus basses valeurs du stocks. Elle permet au décideur de disposer d'un indicateur qui saisit à la fois la probabilité de rupture et la quantité non livrée (qui sont fortement liés à deux indicateurs utilisés pour mesurer le niveau de service : le *cycle service level* et le *fill rate service level*). De plus, le paramètre α fournit un moyen simple de contrôler le niveau de risque et l'Average-Value-at-Risk peut être linéarisée. Nous appliquons un schéma d'approximation classique pour résoudre le programme stochastique en faisant d'abord une approximation en deux étapes, puis un échantillonnage de scénarios afin d'obtenir un *Programme Linéaire en Nombres Entiers (PLNE)* menant à une formulation tractable.

Les jeux de données réelles fournies par les clients d'Argon Consulting ne contiennent que les valeurs historiques de production et de ventes. Comme nous n'avons pas la demande réelle, nous proposons dans le Chapitre 7 un modèle probabiliste pour générer des réalisations possibles de la demande à partir de valeurs historiques. Ce modèle est basé sur les *distributions de Dirichlet* et vise à être facile à utiliser tout en gardant une certaine vraisemblance. Sa seule entrée est une demande prévisionnelle (qui peut être l'historique des ventes ou l'historique des prévisions) et une volatilité qui est un pourcentage représentant l'exactitude de la prévision. (Plus la volatilité est faible, plus la prévision est précise.) Notre modèle probabiliste fournit des scénarios de demande de sorte que le volume total de la demande soit le même dans chaque scénario, de sorte que l'espérance de chaque réalisation soit égale à la prévision et que l'écart type divisé par l'espérance soit proche de la volatilité. Ce modèle répond aux hypothèses faites par Argon Consulting sur la demande, a peu de paramètres et est facile à simuler (même conditionnellement au passé).

Enfin, sur des jeux de données réelles, les temps de calcul peuvent être longs. Les solveurs modernes sont souvent incapables de trouver une solution réalisable au problème. Nous proposons une heuristique qui permet de trouver rapidement une solution réalisable au problème du multi-sourcing. La solution retournée peut être utilisée directement par les clients d'Argon Consulting ou comme solution initiale d'un solveur générique.

Les expérimentations faites dans le Chapitre 10 prouvent d'ores et déjà que l'entreprise qui fournit les jeux de données peut réduire sa proportion d'articles multi-sourcés (réduisant ainsi ses coûts) tout en conservant une bonne capacité à satisfaire la demande. Cependant, les performances des ordinateurs et la taille réelle des jeux de données empêchent de traiter plus d'une centaine de scénarios. La méthode dépend donc de l'échantillonnage et le choix d'un ensemble représentatif de scénarios est essentiel. Nous proposons une façon concrète de réduire cette dépendance vis-à-vis de l'échantillonnage basées sur les méthodes de clustering (telles K-means).

1.2 Modèles de stock en temps continu

Argon Consulting utilise des modèle de stock en temps continu pour calculer des paramètres macroscopiques au niveau tactique (décisions à moyen terme). Les exemples classiques sont la *taille de lot* et la *taille de la couverture*. La taille de lot donne la quantité d'un même article produite à chaque lancement de production. La taille de la couverture indique le nombre d'unités de temps suivant un lancement de production pendant lequel le stock doit être positif.

Ces paramètres sont utilisés comme entrée pour d'autres processus de décision dans la Supply Chain comme pour la planification des besoins en composants (MRP). Par exemple, avoir une estimation des tailles de lot ou des tailles de couverture permet de décider de la quantité de matières premières qui doit être commandée. Ils sont également utilisés comme entrées pour planifier la production puisqu'ils donnent les tailles de lot à produire. (Lors de l'étude des modèles en temps discret, nous proposerons de supprimer cette contrainte des modèles.)

Les modèle de stock en temps continu présupposent une vision continue du temps. Le modèle fondateur connu sous le nom de *Formule de Wilson* et développé par Harris (1913) permet de calculer le compromis optimal entre les coûts de commande et les coûts de stockage. En pratique, la formule de Wilson est difficile à utiliser car les coûts de commande et de stockage sont difficiles à comparer. Argon Consulting cherche à trouver les tailles de lot (ou tailles de couverture) optimales à partir de la flexibilité de ses lignes de production, qui est considérée comme fixe et définie par les processus amont de la Supply Chain. Une ligne de production peut produire plusieurs articles mais perd du temps lorsqu'elle change de la production d'un article à un autre. En considérant un taux de demande constant, la Figure 1.3 montre les conséquences de plusieurs longueurs de cycles de production. Des cycles de production trop courts entraînent des ruptures de stock car trop de temps est perdu dans les changements de production, alors que des cycles de production trop longs entraînent des sur-stocks et des temps improductifs. Dans les jeux de données réelles, les lignes de production produisent beaucoup d'articles et le temps perdu dû au changement entre différents éléments est modélisé par un nombre maximal de changements.

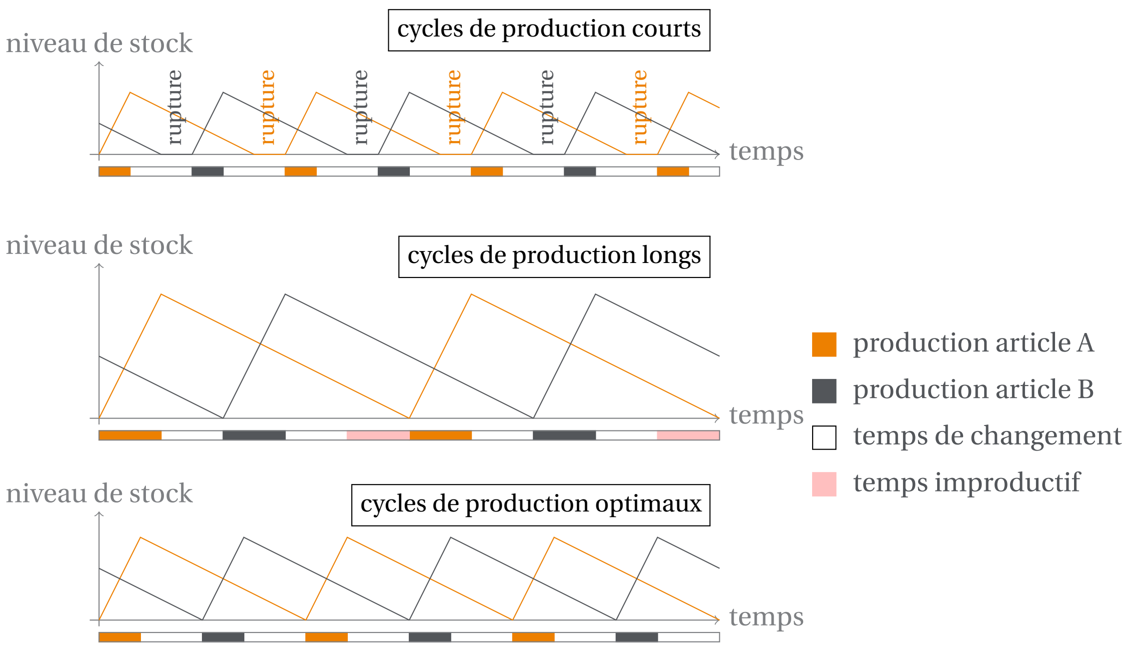


Figure 1.3 – Modèle de stock en temps continu pour une ligne produisant deux articles

En remplaçant les coûts de changement par une borne supérieure sur le nombre de changements, nous proposons dans le Chapitre 3 des généralisations de la formule classique de Wilson pour des cas avec plusieurs articles. Elles sont désormais utilisées par Argon Consulting. En particulier, nous étudions les cas où les tailles de couverture sont continues ou entières dans des contextes déterministes et stochastiques. De plus, nous étudions également dans le Chapitre 4 une extension qui prend en compte plusieurs lignes de production parallèles et montre que le problème peut être décrit comme un problème de minimisation concave sur un polyèdre (pour lequel il existe une vaste littérature).

1.3 Modèles de stock en temps discret

Les modèles de stock en temps discret (également appelés *dynamic lot-sizing problem*) supposent que le temps est décomposé en périodes discrètes. Ils sont utilisés par les entreprises pour planifier leurs productions à court terme. Un modèle classique est le *Capacitated Lot-Sizing Problem (CLSP)*. Il s'agit d'une ligne de production produisant plusieurs articles pendant un nombre fini de périodes. La demande pour chaque article est déterministe et donnée pour chaque période. Il vise à minimiser la somme des coûts de stockage (dus aux stocks reportés d'une période à l'autre) et des coûts de lancement (coûts fixes liés au lancement de la production) sous contrainte de capacité de la ligne de production.

Comme déjà mentionné en Section 1.2, l'inconvénient de cette formulation selon Argon Consulting et ses clients est la difficulté d'estimer la valeur des coûts de lancement. D'autre part, l'estimation du nombre maximal de lancements pour une période donnée est une tâche facile pour les clients d'Argon Consulting. Nous proposons dans le Chapitre 5 un modèle dérivé du Capacitated Lot-Sizing Problem qui remplace les coûts d'installation par une borne supérieure sur le nombre de lancements. La Figure 1.4 montre un exemple de planification de la production de quatre articles sous contrainte qu'au plus deux éléments peuvent être produits pendant une période. À notre connaissance, ce modèle est nouveau et n'a pas été étudié dans la littérature.

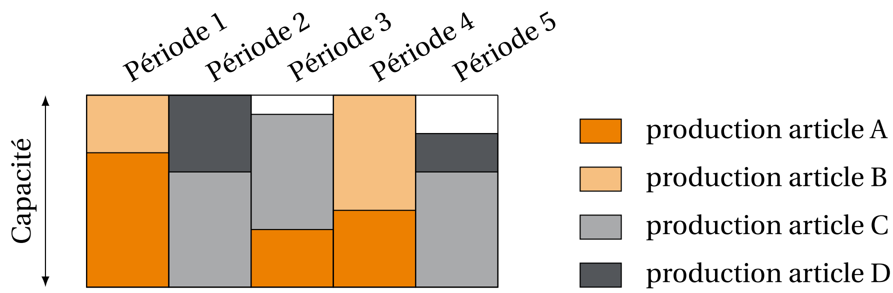


Figure 1.4 – Planification de la production de quatre articles sur cinq semaines

Notre problème de taille de lot peut être écrit comme un *Programme Linéaire en Nombres Entiers (PLNE)*. Nous obtenons plusieurs résultats théoriques dans le cadre déterministe qui montrent les difficultés du problème. Comme prévu, ce problème est NP-hard. Une méthode classique pour aider à résoudre les programmes linéaires en nombres entiers consiste à relâcher certaines contraintes pour obtenir une borne sur la valeur optimale du problème. Nous montrons que plusieurs formulations naturelles produisent toutes la même relaxation continue. Enfin, nous n'avons pas pu répondre à la question suivante : quelle est la complexité de notre problème de Lot-Sizing lorsqu'il n'y a pas de contraintes de capacité et que le nombre maximal de changements par période est égal à 1 ? Mathématiquement, il peut être formulé comme suit.

Considérons le problème

$$
\begin{aligned}
\min \quad & \sum_{t=1}^{T} \sum_{i \in \mathcal{I}} h^i s_t^i \\
\text{s.t.} \quad & s_t^i = s_{t-1}^i + q_t^i - d_t^i && \forall t \in [T],\ \forall i \in \mathcal{I}, \\
& q_t^i \leqslant M x_t^i && \forall t \in [T],\ \forall i \in \mathcal{I}, \\
& \sum_{i \in \mathcal{I}} x_t^i \leqslant 1 && \forall t \in [T], \\
& x_t^i \in \{0,1\} && \forall t \in [T],\ \forall i \in \mathcal{I}, \\
& q_t^i,\ s_t^i \geqslant 0 && \forall t \in [T],\ \forall i \in \mathcal{I},
\end{aligned}
\tag{P}
$$

où M est un grand réel positif et pour chaque période t et chaque article i, la demande d_t^i et le coût de stockage h^i sont des nombres réels positifs et le stock s_t^i, la quantité produite q_t^i et l'indicateur de lancement x_t^i sont les variables de décision.

Question ouverte. Quelle est la complexité de (P) ?

En pratique, la demande n'est pas toujours déterministe. Nous proposons dans le Chapitre 6 une version stochastique de notre problème de taille de lot basée sur la programmation stochastique. (voir Section 1.1). La différence est que nous n'utilisons pas une contrainte d'aversion au risque (AV@R) mais nous nous en tenons à la vision classique de neutralité au risque (l'espérance). En effet, la production est une décision répétée et un échec lors d'une période est facile à compenser par une autre période.

De plus, dans un contexte stochastique, nous devons autoriser le backorder (*i.e.* les commandes livrées en retard) parce que les capacités de production sont limitées et qu'il peut être impossible de couvrir toutes les réalisations possibles de la demande. Ici, le backorder vient avec des coûts dans la fonction d'objectif. À moins qu'ils ne soient contractualisés avec les clients, les coûts de backorder peuvent être difficiles à estimer. Nous adaptons une méthode de la littérature basée sur le problème du *vendeur de journaux* (l'un des plus anciens modèles stochastiques) pour lier le coût de backorder et le niveau souhaité de la demande satisfaite.

Comme dans la Section 1.1, notre problème de taille de lot stochastique est également résolu en faisant d'abord une approximation en deux étapes, puis un échantillonnage de scénarios afin d'obtenir un programme linéaire en nombres entiers. Comme les scénarios de demande n'ont pas été fourni par notre partenaire, nous les générons en utilisant le modèle probabiliste mentionné en Section 1.1.

Les expérimentations faites dans le Chapitre 7 tendent à montrer que l'entreprise qui fournit les jeux de données peut réduire ses coûts de stockage tout en conservant une bonne capacité à satisfaire la demande. Cependant, comme dans la Section 1.1, les performance des ordinateurs, la taille réelle des jeux de données et le temps limité pour retourner un plan de production empêchent de traiter plus de vingt scénarios. Puisque la méthode dépend de l'échantillonnage, nous proposons encore une fois une façon concrète de réduire cette dépendance, fondée sur les méthodes de clustering (comme K-means).

1.4 Extensions

Dans la Partie IV, nous présentons deux extensions de notre travail. La première est une version alternative du problème de multi-sourcing abordé dans la Partie III. La différence est que l'entreprise a un budget limité à investir dans la flexibilité. Dans ce cas, l'entreprise cherche à décider d'une affectation qui maximise la demande qui peut être satisfaite. Nous modélisons ce problème alternatif et montrons qu'il est NP-difficile dans plusieurs cas simples.

La seconde est une extension des modèles de stock abordés dans la Partie I et II. Notre objectif est de calculer les tailles de couverture à moyen terme à l'aide d'un modèle s'appuyant sur la planification de la production à court terme. Nous modélisons ce problème alternatif et montrons expérimentalement qu'il présente de nombreux inconvénients par rapport aux modèles de stock en temps continu.

2 Business context

A Supply Chain is a system of organizations, people, activities, information, and resources involved in producing, transforming or moving a product or a service from suppliers to customers. It can be considered for a subsidiary, a whole company or for multiple companies working together. Its activities extend from the procurement of raw materials to their transformation into finished products and their delivery to the costumer.

Supply Chain can fulfill several functions. Most common are procurement, production, distribution, storage, shipping, sales and billing, and customer relationship. The work of this thesis deals with production and storage optimization within companies.

2.1 Supply Chain objectives

Supply Chain management consists in finding the right balance between conflicting objectives. They can be classified into three categories: *costs*, *stocks* and *service*.

Costs are the expenses paid by companies as a part of their activities. It can be fixed or variable and direct or indirect. Fixed costs, like rent, insurance and investment costs (such as buying machines), do not depend on the volume of the business activity. On the other hand, variable costs, like shipping or energy costs, directly increase with the volume of the business activity. Direct costs concern resources entirely dedicated to the product manufacturing like procurement of raw materials or worker salaries. Indirect costs support functions that are not directly involved in production like marketing or administration.

Stocks, also called *inventory*, induce costs for storage space, insurance, broken or stolen goods, and work time to keep registered stock accurate. The main characteristic of inventory is that it is locked-in money until it is transformed (for raw materials) or sold (for finished products). Thus, it prevents from investing this capital in developments like R&D or geographic expansion. Stocks are more comparable to working capital and because of this specificity, they must form their own category in optimization process.

Service or *service level* measures the ability of the Supply Chain to deliver the right product to the customer while respecting the deadlines. There are many indicators to measure the service level but we can identify two main ones: the *cycle service level* and the *fill rate service level*. Both rely on stock-out (i.e., the unavailability of a product at a given time) and on the replenishment cycle which is the time between two consecutive replenishments of the stock. The cycle service level is the probability of not hitting a stock-out during the next replenishment cycle, and thus, also the probability of not losing sales. The fill rate service level represents the fraction of the demand that is delivered without delays or lost sales. The choice of the indicator depends on the industry. Indeed, in the case of complex products like in aeronautics, if one component is missing in the command, it can delay the whole project. Thus delivering the whole command is more important and this makes the cycle service level more relevant. In case of simple products like in mass distribution, it is more relevant to measure the part of demand delivered at due date and this makes the fill rate service level more relevant.

Automotive and luxury industries are good examples of these conflicting objectives. In automotive industry, high stock is impossible because of depreciation and diversity of products. However, the service level must be high due to high competition. In luxury industry, availability of the right product in the right store is more critical than production costs but high stocks are still impossible.

In general, *industrial agility*, which is the ability to face variations in customers' demand is an efficient way to be competitive if two conditions are satisfied. First, the cost of agility should stay reasonable and second, agility should not be achieved thanks to high stocks. For a company, defining its agility is therefore finding a balance between cost, stock and service.

Industrial agility can be split in two parts: *capacitative reactivity* and *flexibility*. Capacitative reactivity is the ability to face variations of the global volume, by being able to produce more products than the estimated demand for example. This concerns over-capacitated industries or industries being able to outsource part of the production. Capacitative reactivity is more critical in sector like process industry since adding overcapacity is extremely expensive and outsourcing can be extremely hard. Flexibility is the ability to face variations of the product mix. Multi-sourcing the production, which means giving to several plants the ability to produce the same product, enables more flexibility but decreases productivity. Increasing stocks would be another way to increase flexibility. Among other, flexibility enables to adapt to an unexpected cannibalization of a product by another. Luxury industry and mass distribution are probably the sectors where being flexible is the most critical due to the vast number of products.

The exact definition of the agility depends on the industrial context but also on the decision level: strategic (long-term), tactical (mid-term) or operational (short-term). At the strategic level, the executive committee has expressed the desired agility and wants to decide an investment in production capacities. Thus, agility is a constraint, cost is an objective to minimize and stock is not really considered. Just between strategic and tactical levels, capacitative reactivity cannot be changed, S&OP (*Sales and Operations Planning*) process has expressed a desired service level and wants to decide multi-sourcing. Thus, service level is a constraint and cost is an objective. At the tactical and operational levels, service level is an input for the production planning and stock and production are the lever. Thus, service level is a constraint, stock is an objective to minimize and production are the decisions. At very short term horizon, e.g., for scheduling decisions, last-minute optimizations can still be made in order to reach the service level objectives.

2.2 Supply Chain organizations

Supply Chain organizations can be classified into four main models described by Arnold et al. (2007) (among others). They are represented in Figure 2.1.

- *Engineer-To-Order (ETO)*. The customer is involved in the design and gives engineering requirements and specifications which enables a lot of customization and a specific design. Due to purchasing of raw materials and to designing, the consequences are long delivery lead times. Classical domains are aeronautic or aerospace industries.
- *Make-To-Order (MTO)*. Products are made from standard components but with custom-designed components. Therefore, inventory are only composed of raw materials and delivery lead time are still long. For example, marine energy turbines can be produced with a MTO organization.
- *Assemble-To-Order (ATO)*. Customer involvement is limited to selection of component options. Thus, inventory is only composed of semi-finished products ready for assembly components and delivery lead time are short. Production of laptops partially follows this organization.
- *Make-To-Stock (MTS)*. Customer has very little involvement in the design. Products are engineered and manufactured to fill stocks which supply clients demand. This organization enables the shortest delivery lead time. The majority of mass distribution products uses this organization.

For some products, the best organization is obvious due to size of series. For example, a French aeronautic company produces engines of an aircraft carrier with an ETO organization, but turboshafts of choppers with MTO organization. For other industries, identifying the best organization also depends on the commercial strategy (laptops may also be produced with MTS organization) and it is critical to define which decisions are short-term, mid-term or long-term and to know when costs, stock and service can be impacted.

2.3 Presentation of Argon Consulting

Argon Consulting is an international, independent consulting firm whose mission is to help its clients achieve sustainable competitive advantages through operational excellence. It began its consultancy activity in 2001 and employs over 230 consultants in six offices: Paris, London, Atlanta, Singapore, Melbourne and Mumbai.

Argon Consulting has supported many companies in operational transformation projects (R&D, Procurement, Manufacturing, Supply Chain, Distribution, Services, SG&A, Performance Management, Change Management). Industries served cover *Aerospace & Defense* (Latécoère, Safran, Thales) and *Discrete Manufacturing* (Alstom, SNCF, DCNS) which have small-series and large program logic as well as *Retail* (ADEO, Carrefour, Cdiscount) which sells products or services to large number of customers through multiple channels of distribution. In the mid of these extremes, we also find *Automotive* (Michelin, PSA, Valeo) where innovation performance and diversity of products are challenging, *Consumer Packaged Goods* (Bel, Danone, L'Oréal) where decreasing consumption and fluctuation of production costs reduce profitability, and *Textile* (Galeries Lafayette, Camaïeu, Kiabi) where fast product renewal and fast evolution of sourcing are critical at an operational level. Among other sectors with specific rules, *Luxury Goods* has a

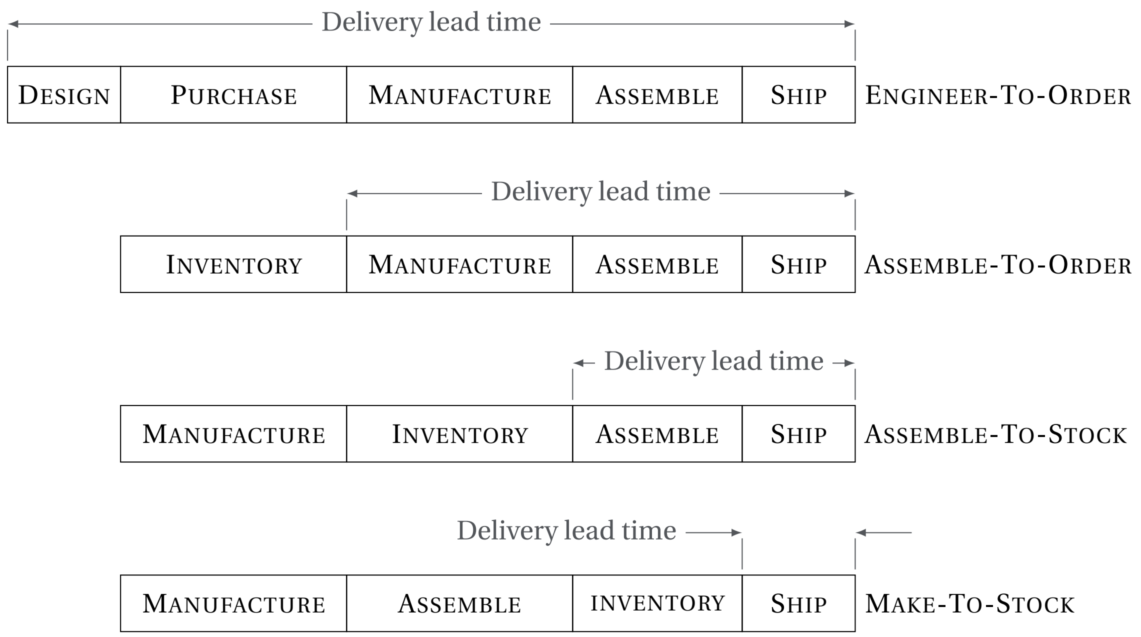


Figure 2.1 – Supply Chain organizations and lead times (from Arnold et al. (2007))

highly erratic demand and is extremely competitive, *Pharmaceutical Industry* (Merck, Sanofi, Servier) has an economic pressure applied by state authorities due to imbalance of the health insurance systems and *Energy & Utilities* (EDF, ENGIE) is capital-intensive, highly cyclical, fully globalized and at the heart of geostrategic issues.

Argon Consulting began as a consultancy company specialized in logistic. Growing, it acquired expertise in every level of the Supply Chain. Argon Consulting describes itself as a multi-specialized firm whose competitive advantage comes from its ability to quantify their studies. Through this thesis, the objective is to find an unified framework for some recurrent problems. Moreover, Argon Consulting wants to test scientifically the accuracy and the efficiency of the developed models.

2.4 Argon Consulting's clients cases

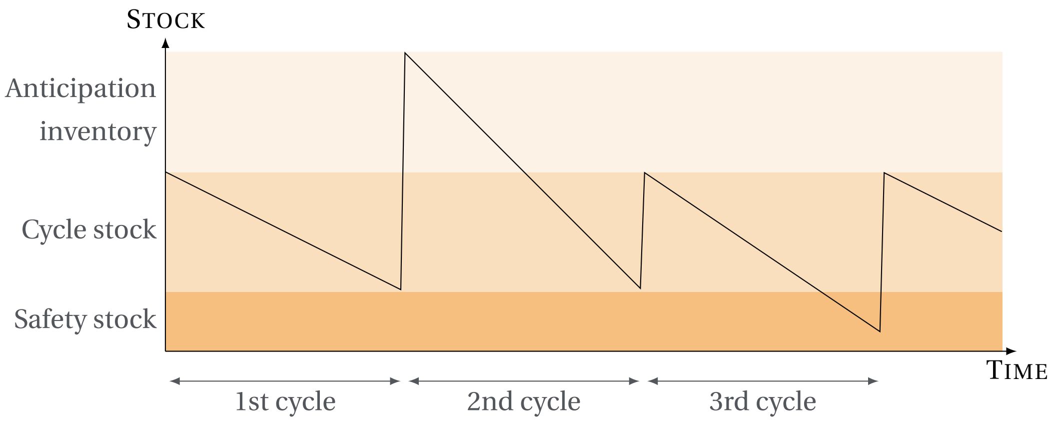
Argon Consulting's clients cases considered in this thesis fall into Assemble-To-Order and Make-To-Stock organizations. We study two cases in this thesis. The first case is both tactical and operational. It occurs when flexibility of means of production is already defined and we aim at deciding which stocks and production levels would enable the company to achieve the desired service level. The second case is more strategic. It occurs when capacity reactivity is already defined and we aims at deciding the multi-sourcing of production which will ensure enough flexibility at tactical and operational levels.

2.4.1 Production planning

Production planning is part of the production function of the Supply Chain and occurs at tactical and operational levels when the flexibility of means of production is already defined. In

Assemble-To-Order and Make-To-Stock organizations, stocks are the last levers on flexibility and there must be enough of them to serve the demand at due dates. Then, production planning problem consists in deciding the orders of production, i.e., when production starts and how much is produced, which defines stocks and allows to achieve the desired service level.

Difficulties of this problem come from the many constraints that prevent last minute production. First, production capacities are limited. Thus, production must be anticipated in order to deliver during peak selling season, promotion program, vacation shutdown, etc. It leads to inventory called *anticipation inventory*. Second, demand and lead time have random variations. If demand or lead time are greater than forecast, a stock-out can occur. To prevent it, *safety stocks* are kept as a reserve. Finally, since the *flexibility* of the means of production is limited, *lot-size inventories* also called *cycle stocks* are needed. They form the portion of stocks that varies over time due to production and demand fulfillment. The different parts of the inventory are represented in Figure 2.2.



In the first cycle, cycle stock enables the company to satisfy perfectly the demand. In the second cycle, an increase in the demand is announced. So, an anticipation inventory is produced in addition to the cycle stock. In the third cycle, the demand is bigger than expected. Thus part of the demand is satisfied using the safety stock.

Figure 2.2 – Inventory decomposition

Production planning aims at minimizing cumulative stocks subject to constraints defined at a higher level.

Service level is the first constraint. In our case, Argon Consulting's clients want to reach a desired service level which depends on the strategy of the company. It is a trade-off between several objectives as loss of reputation, intended costs, risk, benefits, etc. Production planning problems take it as an input since service level is a long-term decision whereas production decisions are mid-term decisions. Since we are interested in Assemble-To-Order and Make-To-Stock models, we will consider mainly the fill rate service level as defined in Section 2.1.

Most of industrial costs are already fixed. In our case, they are modeled by the capacity and the *flexibility*. Indeed, capacities of plants or assembly lines and their flexibility are strategic decisions whereas production decisions are tactical or operational decisions. Thus, when production planning problems occur, they cannot be changed. Capacity constraints are easy to model. Conversely, the literature addresses several models for flexibility. A standard way is setup costs, that is costs induced by the transition from one product type to another. The drawback of this formulation is that it leads to a multi-objective optimization problem since stocks is already an objective. Moreover, discussions with Argon Consulting's clients show that it is difficult to quantify the setup costs. Another option would be to decrease the capacity by the time loss at each new setup. However, the time needed for a new setup often depends on the previous and the next items produced. Thus, scheduling must also be optimized which is not possible since these are short-term decisions. Another way of dealing with scheduling at the tactical level is to define an average setup time. In practice, industrials prefer to define and use a number of setups per period rather than an average setup time although it is computed from the average setup and the loss capacity. We propose to model flexibility by following their recommendation and by adding a constraint on the number of setups.

We address two models for the production planning problem in this thesis. In Part I, we propose continuous-review inventory models whose main objective is to optimize the cycle stock. And in Part II, we propose lot-size models which use a discretization of the time and which simultaneously optimize the three parts of the inventory.

2.4.2 Production multi-sourcing

Production multi-sourcing also concerns the production function of the Supply Chain. Its objective is to define the flexibility level of means of production by deciding if a plant should have the ability to produce a product. Note that several plants may be able to produce the same product which explains why this problem is called *multi-sourcing*. This decision is made between the strategic and tactical levels – depending on the industry – typically when companies expand their product portfolio or when they have expanded their productions capacities.

Multi-sourcing enables the company to face variability of the product mix. When there are many products, there may be a *cannibalization* between the products, i.e., demand for some products may decrease while demand for others increases. In this case, a mono-sourced production (i.e., a product is produced by only one plant) would lead not to use the full capacity of some plants while other plants would be unable to satisfy demand. On the other hand a multi-sourced production allows to produce the exceeding demand in a plant whose activity has decreased. Moreover, even if the production sourcing currently in use by the company allows to satisfy the demand, a higher multi-sourcing may allow to balance the production between plants. Indeed, to face randomness at tactical and operational levels, it is better to have two plants using 80% of their capacities than one plant using 100% and the other only 60%. Thus, multi-sourcing decisions have a high impact on competitiveness once we get into production planning (tactical and operational decisions) and scheduling (very short-term decisions).

Considering all the advantages of production multi-sourcing, one could think than every plant should be able to produce every products. Indeed, this complete flexibility of plants is the safest solution to ensure the future satisfaction of the demand. However, multi-sourcing the production of every products leads to unnecessary high assignment costs. First, giving a plant the ability to produce a product would require to buy new means of production and to pay for employees training. It has a monetary cost. Moreover, it is also a loss of time. Indeed, when employees are trained, they do not produce and after the training, they need time to be fully efficient. This is why specialization of plants also has advantages.

In industries where demand is close to the production capacity, wisely choosing the new assignment is a critical problem since increasing the current production multi-sourcing may decrease short-term production capacity and demand may be hard to satisfy at short term. However, on a long-term horizon, it will help to satisfy the demand. Thus, multi-sourcing decisions must be made while considering long-term impacts.

Due to the characteristics of multi-sourcing, its objective is to minimize the costs and the time loss when giving plants the ability to produce new products, while satisfying the demand on a long-term horizon. Since multi-sourcing decisions have a long-term impact, the worst possible realizations of demand can matter. Indeed, unlike production planning decisions which have a mid-term impact and for which one bad realization of demand can be compensated by every next realizations, there are few multi-sourcing decisions. In this case, values of one bad outcome may have a huge impact on costs and must be controlled.

In this thesis, we propose in Part III a model to decide multi-sourcing while satisfying a long-term service level and controlling the loss in the worst realizations of demand.

2.5 Assumption of the models

In this thesis we propose models that aim at minimizing stocks while taking into account flexibility and capacitative reactivity. However, the latter do not play the same role in each problem. In production planning problems (Parts I and II), both are considered as constraints: flexibility is represented by the number of setups and capacitative reactivity is represented by the production capacity of the assembly line. But in multi-sourcing problems (Part III), flexibility is a decision variable (product assignment to plants) whereas capacitative reactivity remains a constraint (production capacity of plants). Capacity decisions are long-term decisions whereas multi-sourcing decisions are between long and mid-term decisions and production decisions are mid-term or short-term decisions.

Since we do not have any control on capacities in our model, an increase of the demand may lead to infeasible problems. Companies would reconsider their capacity decisions and our models would become irrelevant. In order to avoid this issue in the considered models, the global volume of demand is assumed constant in each possible outcome and the only randomness in demand comes from the product mix.

Even if it seems to be a strong hypothesis, it is realistic for this horizon of decisions. Indeed, in many cases, there is a *cannibalization* between products, i.e., if demand for some products increases, demand for other decreases. Promotions, seasonality or simply offsets between error can cause cannibalization. For example, the average consumption of desserts is the same at any time of the year but ice cream demand will represent a bigger part in Summer.

Part I

Continuous-time inventory models

3 Production on a single line

As explained in Section 2.4.1, the objective is to reduce cumulative inventory. This chapter presents a model which aims at deciding the optimal cycle stocks.

3.1 Motivations

Cycle stocks form the portion of inventory that varies over time due to production and demand satisfaction. Low cycle stocks contribute to globally decrease inventory but it requires a high flexibility of means of production. For mid-term decisions, flexibility is already fixed and companies aim at deciding the values of cycle stocks of each item which minimize the global inventory.

Many production systems are managed using (r, q) policies or similar ones like (s, S) policies as explained for example by Arrow et al. (1951). (r, q) policies are based on a continuous review of inventory. It defines for each item a level r and a quantity q called *lot-size* such that: when inventory level of item reaches level r, a quantity q is produced. A common variation, (s, S) policy are based on a periodic review of the inventory. In this case, replenishments occur at discrete intervals of time and only if inventory level of item is below level s. Then, inventory is completed in order to reach level S.

Argon Consulting and part of its clients also use policy based on *cover-sizes*. This policy is a mix between the two previous policies. When inventory level of item i reaches level r, quantity produced is the cumulative demand of the τ_i next units of time. This quantity τ_i is called the cover-size of item i. Both cover-size and (r, q) policies are used by companies (and are equivalent when demand does not depend on time).

Since we are more interest in the values of the cover-sizes and of the lot-sizes rather than the level r, in the following, we refer to (r, q) policies as lot-size policies to clearly differentiate the two policies. In this chapter, we show how to compute lot-sizes and cover-sizes minimizing inventory subject to a flexibility constraint. We propose a method in both deterministic and stochastic settings.

3.2 Deterministic settings

In this section, data are assumed to be deterministic.

3.2.1 Problem

The problem described by Argon Consulting considers an assembly line producing a set $\mathcal{I}$ of items over an infinite horizon. The internal production time of item i (i.e., the quantity of item i produced in one time unit) is v_i. Inventory of item i must satisfy a demand, known in advance. This is modeled by a continuous decrease of $d_i < v_i$ units per time unit when there is no production and a continuous increase of $v_i - d_i$ units per time unit when item i is produced. Demand d_i is assumed positive for each item i.

We introduce now the decision variables τ_i and t_i of the problem. For each item i, the first time it is produced is t_i. Then, the production is launched at every $t_i + k\tau_i$ where $k \in \mathbb{Z}_+$. Each $t_i + k\tau_i$ is called a *setup*. This τ_i is called the *cover-size* of item i. Each production lasts $\frac{d_i}{v_i}\tau_i$ in order to produce exactly the demand for the next τ_i units of time. The productions of several items can be launched simultaneously (like in model developed by Ohno and Ishigaki (2001) where authors assume the immediate replenishment of the order with lead time). The production decisions and the demand give the inventory of item i which is denoted s_i.

Thus, for each item i, inventory $s_i(t)$ is continuous, right and left differentiable, nonnegative and follows the dynamic

$$\dot{s}_i(t) = \begin{cases} v_i - d_i & \text{if } t \in \bigcup_{k \in \mathbb{Z}_+} \left[t_i + k\tau_i, t_i + \left(k + \dfrac{d_i}{v_i}\right)\tau_i \right), \\ -d_i & \text{otherwise.} \end{cases} \tag{3.1}$$

As explained in Section 2.4.1, in this thesis, the flexibility is modeled by a constraint and not by setup costs. In average over the infinite horizon, the number of setups per unit time for all items must be smaller than N which can be written as

$$\limsup_{T \to +\infty} \frac{1}{T} \sum_{i \in \mathcal{I}} \left\lfloor \frac{T - t_i}{\tau_i} \right\rfloor \leqslant N. \tag{3.2}$$

Indeed, during the interval $[0, T)$, setups occurs with period τ_i from t_i. Then, the number of setups during $[0, T)$ is $\left\lfloor \frac{T-t_i}{\tau_i} \right\rfloor$.

For each item i, there is an initial inventory $s_i(0) \in \mathbb{R}_+$ given in input.

Since inventory varies over times, the cycle stock is measured using its average value over an infinite horizon. Storing one unit of item i incurs a unit holding cost $h_i > 0$ per unit time. The objective is to find the cover-sizes τ_i which minimize the average cycle stock over infinite horizon

$$\limsup_{T \to +\infty} \frac{1}{T} \sum_{i \in \mathcal{I}} h_i \int_0^T s_i(t) dt \tag{3.3}$$

while satisfying nonnegative inventory, constraint (3.1) and constraint (3.2).

The notations and a corresponding inventory profile is representing in Figure 3.1.

Since this model is a variation of the *Economic Production Quantity (EPQ)* model, we call it *Economic Production Quantity model with Bounded number of Setups (EPQ-BS)*. We will consider two versions of this problem:

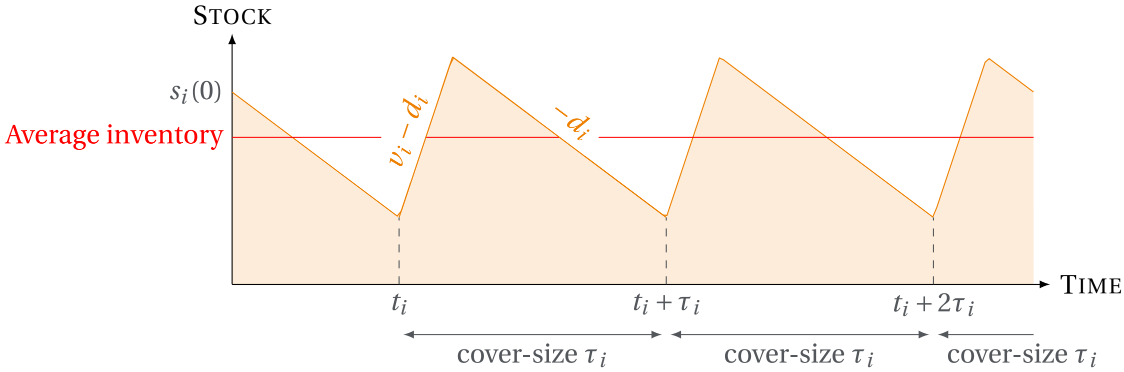


Figure 3.1 – Inventory of item i depending on time for a given cover-size τ_i

- the cover-sizes can be any nonnegative real numbers,
- the cover-sizes have to be inverses of integers.

In other words, since a cover-size is a time period, the production frequencies are unconstrained in the first version, while there are constrained to be integers in the second one. We qualify the first version of being *unconstrained* and the second of being *integer*.

The integer version relies on practical reasons. For decision makers, it is sometimes easier to use frequencies and thus to know that an item is produced once a month, twice a month, etc.

When using lot-sizes instead of cover-sizes, EPQ-BS can be simply adapted. Indeed, the lot-size ℓ_i is the quantity produced to satisfy the demand for the next τ_i units of time. Thus, we have $\ell_i = d_i\tau_i$.

3.2.2 Bibliography

Continuous review inventory models have been studied for more than a century. They aim at deciding when ordering (or producing) and which quantity must be ordered (or produced) making a trade-off between inventory holding costs and ordering (or producing) costs. They are the first models developed for production and assume a continuous vision of time.

These models are mostly developed for ordering. Then, our review is mostly centered on ordering rather than production. However, extensions for production is often simple like the *Economic Production Quantity (EPQ)* (see Taft (1918)) but may lead to new issues like the holding cost of raw materials as proposed by Lin and Su (2013).

The first model was developed by Harris (1913) and gives the famous *Economic Order Quantity (EOQ)* (also known as *Wilson Formula*) which consists in finding a trade-off between average holding cost (proportional to the inventory) and fixed reordering cost. It is a model for a single item and demand is assumed stationary. Multiple extensions of this model have been studied. A first consists in considering time varying demand, as proposed by Resh et al. (1976), Donaldson (1977) or Barbosa and Friedman (1978) where they show polynomial time algorithms for demand of the form $d(t) = \alpha t + \gamma$ or $d(t) = \alpha t^{\beta}$. However, in general, finding the optimal solution is challenging. For example Padmanabhan and Vrat (1990) assumed that the demand is a function of the inventory when stock-out occurs and they proposed a method from numerical analysis to solve the non-linear equations characterizing the optimal solution. Another extension considers a time varying holding cost due to inflation (see Vrat and Padmanabhan (1990)) for which they also propose a method from numerical analysis to find the optimal solution.

Although they are deterministic, many authors models lost sales like Salameh et al. (2003) who introduce a shortage cost or Park (1982) who considers that a prefixed percentage of unsatisfied demand is lost while the remaining is backordered.

The models we just presented all deal with a single item problem. In many real case, an assembly line can produce more than one item. A first multi-item model was proposed by Hadley and Whitin (1963) which also note that perhaps the most important real world constraint is budget restriction on the amount that can be invested in inventory. Since inventory is often controlled by a trade-off with other costs, some models add upper bounds on inventory to prevent it from becoming to high.

As explained in Eynan and Kropp (2007), when many items are produced on a single line, using periodic review instead of continuous review may ease the coordination between the production of various items. (See Section 3.1 for the distinction between the two policies.) A similar approach was also studied by Madigan (1968) who adds scheduling constraints between productions of items and propose a heuristic to find the optimal lot-sizes and when produce them. Finally, Bomberger (1966) and Goyal (1974) impose that every replenishment cycle (cover-size) be a multiple of a unit cover-size (which can be a variable in the optimization process like in Silver (1976)). In real applications, it may enable to reduce ordering costs if many items are in the same batch.

Finally, even if the interpretation of its coefficient is different from our, Ziegler (1982) proposes a production model which has a formulation close to the one found in Section 3.2.3. However, since its objective function keeps a linear term, he was not able to propose a closed-form formula and propose a polynomial approximation algorithm.

3.2.3 Unconstrained EPQ-BS

We address the following alternative mathematical problem. We are going to show that it models the unconstrained EPQ-BS.

$$
\begin{aligned}
\min \quad & \sum_{i\in\mathcal{I}} \frac{1}{2} h_i \tilde{d}_i \tau_i && && \text{(3.4a)}\\
\text{s.t.} \quad & \sum_{i\in\mathcal{I}} \frac{1}{\tau_i} \leqslant N && && \text{(3.4b)}\\
& \tau_i > 0 && \forall i \in \mathcal{I}, && \text{(3.4c)}
\end{aligned}
$$

where $\tilde{d}_i = \left(1 - \frac{d_i}{v_i}\right) d_i$ for each item i.

Theorem 1. *Problem* (3.4) *has a unique optimal solution* $(\tau_i^*)_{i\in\mathcal{I}}$ *with*

$$\tau_i^* = \frac{\sum_{j\in\mathcal{I}} \sqrt{h_j \tilde{d}_j}}{N\sqrt{h_i \tilde{d}_i}} \qquad \forall i \in \mathcal{I}, \tag{3.5}$$

and an optimal cost equal to $\frac{1}{2N}\left(\sum_{i\in\mathcal{I}}\sqrt{h_i\tilde{d}_i}\right)^2$.

Moreover, the optimal solution of Problem (3.4) *is the optimal solution of unconstrained EPQ-BS.*

Formulation (3.4) has many advantages. First, it is much simpler than the original formulation of Section 3.2.1. Second, it removes from the formulation the first production setups $(t_i)_i$ which confirms that the t_i are not relevant to find the optimal cover-sizes.

The link between both problems is however not necessarily immediate. As we will see, the proof will show that the optimal policy is *Zero-Inventory-Ordering (ZIO)*. We recall that a policy is said to be ZIO if an order can only occur when the inventory is zero. Due to flexibility constraint (3.2), production should have to be anticipated before inventory reaches zero.

Lemma 2. *Problem* (3.4) *has a unique optimal solution* $(\tau_i^*)_{i\in\mathcal{I}}$ *given by Equation* (3.5).

Proof. Since Problem (3.4) is a convex problem with qualified constraints, solving it is straightforward using the Karush-Kuhn-Tucker conditions which gives the unique solution given by Equation (3.5) and the optimal cost of Problem (3.4) is $\frac{1}{2N}\left(\sum_{i\in\mathcal{I}}\sqrt{h_i\tilde{d}_i}\right)^2$. □

Lemma 3. *For any nonnegative fixed values of the* t_i*'s, we have*

$$\limsup_{T\to+\infty}\frac{1}{T}\sum_{i\in\mathcal{I}}\left\lfloor\frac{T-t_i}{\tau_i}\right\rfloor=\sum_{i\in\mathcal{I}}\frac{1}{\tau_i}. \tag{3.6}$$

Proof. Let $t_1,\dots,t_{|\mathcal{I}|}$ be $|\mathcal{I}|$ nonnegative real numbers. Let T be a real number greater than every t_i. For each i, we have

$$\frac{1}{T}\left(\frac{T-t_i}{\tau_i}-1\right)\leqslant\frac{1}{T}\left\lfloor\frac{T-t_i}{\tau_i}\right\rfloor\leqslant\frac{1}{T}\left(\frac{T-t_i}{\tau_i}\right) \tag{3.7}$$

and then, $\frac{1}{T}\left\lfloor\frac{T-t_i}{\tau_i}\right\rfloor$ converges to $\frac{1}{\tau_i}$ when T goes to infinity. □

Note that this formulation is independent of the first production setups t_i. Then, the choice of the t_i's is only constrained by nonnegative inventory.

Lemma 4. *Let* $(t_i,\tau_i)_{i\in\mathcal{I}}$ *be a feasible solution of unconstrained EPQ-BS. Then, its cost is at least*

$$\sum_{i\in\mathcal{I}}\frac{1}{2}h_i\left(1-\frac{d_i}{v_i}\right)d_i\tau_i. \tag{3.8}$$

Proof. Let $(t_i,\tau_i)_{i\in\mathcal{I}}$ be a feasible solution of unconstrained EPQ-BS. Using the dynamic (3.1), we have for each item i

$$S_{i,0}=\int_0^{t_i}s_i(t)\,dt=\frac{1}{2}(2s_i(0)-t_id_i)t_i, \tag{3.9}$$

and for each $k\in\mathbb{Z}_+$

$$S_{i,k}=\int_{t_i+k\tau_i}^{t_i+(k+1)\tau_i}s_i(t)\,dt=(s_i(0)-d_it_i)\tau_i+\frac{1}{2}\left(1-\frac{d_i}{v_i}\right)d_i\tau_i^{*2}. \tag{3.10}$$

Let T be a real number greater than t_i. Splitting the integral, we get

$$\frac{1}{T}S_{i,0} + \frac{1}{T}\sum_{k=1}^{\left\lfloor \frac{T-t_i}{\tau_i} \right\rfloor - 1} S_{i,k} \leqslant \frac{1}{T}\int_0^T s_i(t)\,dt \leqslant \frac{1}{T}S_{i,0} + \frac{1}{T}\sum_{k=1}^{\left\lfloor \frac{T-t_i}{\tau_i} \right\rfloor} S_{i,k} \tag{3.11}$$

and the average cycle stock on infinite horizon for item i follows:

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T s_i(t)\,dt = s_i(0) - d_i t_i + \frac{1}{2}\left(1 - \frac{d_i}{v_i}\right) d_i \tau_i. \tag{3.12}$$

Equation (3.1) implies that $s_i(t_i) = s_i(0) - d_i t_i$. Since inventory is nonnegative in a feasible solution, we have $s_i(0) - d_i t_i \geqslant 0$. Finally, the average holding cost of all items over infinite horizon is greater than or equal to $\sum_{i\in\mathcal{I}} \frac{1}{2} h_i \left(1 - \frac{d_i}{v_i}\right) d_i \tau_i$. □

Proof of Theorem 1. Lemma 3 and Lemma 4 prove that every feasible solution of unconstrained EPQ-BS is a feasible solution of Problem (3.4) with greater or equal cost. Conversely, a feasible solution of Problem (3.4) can be completed in a solution of unconstrained EPQ-BS with the same cost setting the first production setup t_i of item i equal to $\frac{s_i(0)}{d_i}$. Then, the unique optimal solution of Problem (3.4) (Lemma 2) is the optimal solution of unconstrained EPQ-BS. □

Note that this model simply adapts to the case where production is considered instantaneous (i.e., $v_i \to \infty$). In this case, just use the real demand d_i instead of the modified demand $\tilde{d}_i = \left(1 - \frac{d_i}{v_i}\right) d_i$.

3.2.4 Integer EPQ-BS

In some cases, it is easier for companies to use integer frequencies. We address the following alternative mathematical problem. As in Section 3.2.3, we are going to show that it models the integer EPQ-BS.

$$\min \quad \sum_{i\in\mathcal{I}} \frac{1}{2} h_i \tilde{d}_i \frac{1}{n_i} \tag{3.13a}$$

$$\text{s.t.} \quad \sum_{i\in\mathcal{I}} n_i \leqslant N \tag{3.13b}$$

$$n_i \in \mathbb{Z}_+^* \qquad \forall i \in \mathcal{I}, \tag{3.13c}$$

where $n_i = \frac{1}{\tau_i}$ is the average number of setups per time unit over the infinite horizon and $\tilde{d}_i = \left(1 - \frac{d_i}{v_i}\right) d_i$.

Proving that the optimal solutions of Problem (3.13) are the optimal solutions of integer EOQ-BS is very similar to the unconstrained case since proofs of Lemma 3 and of Lemma 4 do not rely on the nature of the cover-sizes $(\tau_i)_i$. Moreover, as explained in Section 3.2.3, dealing with finite or infinite internal production time v_i is very similar since, it is sufficient to use the demand d_i in infinite case instead of the modified demand $\tilde{d}_i$ in Problem (3.13).

This formulation is a special case of the integer simple resource allocation problem:

$$\max\left\{\sum_{i\in\mathcal{I}} f_i(n_i) \;\middle|\; \sum_{i\in\mathcal{I}} n_i = N, \quad n \in \mathbb{Z}_+^*\right\} \tag{3.14}$$

where the f_i are concave.

The fastest algorithm known has a $O\left(|\mathcal{I}|\log\frac{N}{|\mathcal{I}|}\right)$ running time and was proposed by Frederickson and Johnson (1982) and then simplified by Hochbaum (1994). Implementation of these algorithms is not easy. Dynamic programming might be used instead, but its complexity is only $O\left(|\mathcal{I}|N^2\right)$ which is pseudo-polynomial.

3.3 Stochastic settings

3.3.1 Problem

However, in real life, many parameters are not known in advance. An obvious example is demand which can change due to forecast errors, marketing promotions, passing fads. Randomness can also comes from production means. Failures, holidays or strikes can affect internal production time.

Here, production means are assumed to be reliable and we only consider randomness on demand. The problem becomes an assembly line still producing a set $\mathcal{I}$ of items over an infinite horizon. The internal production time of item i is v_i. Inventory of item i must satisfy a random demand. This is modeled by a continuous decrease $\boldsymbol{d}_i < v_i$ units per time unit when there is no production and a continuous increase $v_i - \boldsymbol{d}_i$ units per time unit when item i is produced. Moreover, $\boldsymbol{d}_i$ is assumed to be almost surely positive and smaller than v_i. Once d_i is revealed, it is fixed forever.

We introduce now the decision variables τ_i and $\boldsymbol{t}_i$ of the problem. For each item i, the first time item i is produced is $\boldsymbol{t}_i$. Then, the production is launched for every $\boldsymbol{t}_i + k\tau_i$ where $k \in \mathbb{Z}_+$. Each production lasts $\frac{\boldsymbol{d}_i}{v_i}\tau_i$ in order to produce exactly the demand for the next τ_i unit of time. In the stochastic case, cover-sizes $(\tau_i)_i$ are decided before the demand is revealed and thus are called *first-stage decisions* since they are decided before demand realization. On the other hand, first production setups $\left(\boldsymbol{t}_i\right)_i$ and produced quantities are called second-stage decisions (or recourse) since they can be decided knowing the demand realization. The production decisions and the demand give the inventory of item i which is denoted $\boldsymbol{s}_i$.

Thus, for each item i, inventory $\boldsymbol{s}_i(t)$ is continuous, right and left differentiable, nonnegative and follows the dynamic

$$\dot{\boldsymbol{s}}_i(t) = \begin{cases} v_i - \boldsymbol{d}_i & \text{if } t \in \bigcup\limits_{k\in\mathbb{Z}_+}\left[\boldsymbol{t}_i + k\tau_i, \boldsymbol{t}_i + \left(k + \dfrac{\boldsymbol{d}_i}{v_i}\right)\tau_i\right), \\ -\boldsymbol{d}_i & \text{otherwise.} \end{cases} \tag{3.15}$$

As in deterministic settings, flexibility is modeled by a constraint. In average over the infinite horizon, the number of setups per time unit for all items must be smaller than N which can be written as

$$\limsup_{T\to+\infty} \frac{1}{T} \sum_{i\in\mathcal{I}} \left\lfloor \frac{T-\boldsymbol{t}_i}{\tau_i} \right\rfloor \leqslant N \quad \text{almost surely.} \tag{3.16}$$

For each item i, there is an initial inventory $s_i(0) \in \mathbb{R}_+$.

Storing one unit of item i incurs a unit holding cost $h_i > 0$ per time unit. The objective is to find the cover-sizes which minimize the average cycle stock over infinite horizon

$$\mathbb{E}\left[\limsup_{T\to+\infty} \frac{1}{T} \sum_{i\in\mathcal{I}} h_i \int_0^T \boldsymbol{s}_i(t)\,dt\right] \tag{3.17}$$

while satisfying almost surely nonnegative inventory, constraint (3.15) and constraint (3.16).

We call this problem *stochastic Economic Production Quantity model with Bounded number of Setups (stochastic EPQ-BS)* and will consider the same two versions as in deterministic case:

- the cover-sizes can be any nonnegative real numbers (called *unconstrained*),
- the cover-sizes have to be inverses of integers (called *integer*).

When using lot-sizes instead of cover-sizes, one needs to pay more attention to the measurability of variables. Indeed, the lot-size $\boldsymbol{\ell}_i$ is the quantity produced to satisfy the demand for the next τ_i units of time. Thus, in stochastic settings, if cover-sizes are first-stage variables, then lot-sizes are second-stage variables and are given by $\boldsymbol{\ell}_i = \boldsymbol{d}_i\tau_i$. Conversely, using lot-size variables as first-stage variables would lead to use cover-sizes as second-stage variables.

3.3.2 Bibliography

As for deterministic cases, literature about ordering is larger than literature about production. However, in many cases, it is easy to adapt ordering models to production cases. One of the first stochastic ordering problems is the news-vendor problem which is for example addressed by Edgeworth (1888). The goal is to order a quantity q to satisfy an unknown future demand d. The challenge is to find a trade-off between expected remaining inventory and expected shortage.

Many deterministic continuous time inventory models (like those presented in Section 3.2.2) have been extended to a stochastic setting. Yano and Lee (1995) and more recently Ziukov (2015) propose a complete review of these extensions. By essence, stochastic problems in literature often include backorder costs.

A distinction between models can first be made on the source of randomness. For example, Friedman (1984) models lead times as random variables whereas Eynan and Kropp (2007) chose to model demand as a random variable. However, both aim at minimizing holding, shortage and ordering cost. Another example is proposed by Sana (2011) who address extensions of the news-vendor problem adding an uncertainty on sales price. Some of these extensions allow sales price to be correlated to the demand.

Finally, Gallego (1998) shows how to adapt (r,q) policies to the robust case. He models lead-time as random variable but he assumes that the only available information on the distribution is the expectation and the variance. Solving the resulting problem against the worst possible distribution in this class, he shows that the (r, q) policy found in the robust case has a cost which is no more than 6% from the cost of optimal (r, q) policy if the distribution was known.

Another class of continuous time inventory models is based on discrete arrivals of demand. A simple model is proposed by Gavish and Graves (1980) where the demand arrivals of a single item follow a Poisson process. They reformulate their problem as a single-server queueing problem in order to solve it. More recently, Gayon et al. (2009) propose an extension still considering one item but with several customers. Each customer has not the same demand arrival rate neither the same price for lost sales. Thus, in optimal policies it may be better to not satisfy a customer demand in order to satisfy the future unknown demand of another one.

Note that there exists another trend to study stochastic continuous-review inventory models namely fuzzy set theory. A fuzzy set is a set whose elements have degrees of membership. The fuzzy set theory enables to model the partial information and can be understood as a "possibility". Each parameter like ordered quantity, ordered cost, holding cost or annual demand can be represented with a fuzzy set instead of a unique value. Park (1987) was the first to propose fuzzy production inventory models and gives an interpretation of economic order quantity in a fuzzy framework. Among other models, we can cite Hsieh (2002) which introduces two fuzzy production inventory models with fuzzy parameters: the first with crisp production quantity (i.e., a production in a set in the classical sense) and the second with production quantity in a fuzzy set. Contrary to most of the literature, Lee and Yao (1999) propose a model for the economic order quantity in a fuzzy framework but without backorder.

3.3.3 Unconstrained stochastic EPQ-BS

We address the following alternative mathematical problem. We are going to show that it models the stochastic unconstrained EPQ-BS.

$$\min \quad \sum_{i\in\mathcal{I}} \frac{1}{2} h_i \check{d}_i \tau_i \tag{3.18a}$$

$$\text{s.t.} \quad \sum_{i\in\mathcal{I}} \frac{1}{\tau_i} \leqslant N \tag{3.18b}$$

$$\tau_i > 0 \qquad \forall i \in \mathcal{I}, \tag{3.18c}$$

where $\check{d}_i = \left(1 - \frac{\mathbb{E}\left[\boldsymbol{d}_i\right]}{v_i}\right)\mathbb{E}\left[\boldsymbol{d}_i\right] - \frac{1}{v_i}\operatorname{Var}\left[\boldsymbol{d}_i\right]$.

Theorem 5. *Problem* (3.18) *has a unique optimal solution* $(\tau_i^*)_{i\in\mathcal{I}}$ *with*

$$\tau_i^* = \frac{\sum_{j\in\mathcal{I}} \sqrt{h_j \check{d}_j}}{N\sqrt{h_i \check{d}_i}} \qquad \forall i \in \mathcal{I} \tag{3.19}$$

and an optimal cost equal to $\frac{1}{2N}\left(\sum_{i\in\mathcal{I}} \sqrt{h_i \check{d}_i}\right)^2$.

Moreover, the optimal solution of Problem (3.18) *is the optimal solution of unconstrained stochas-*

tic EPQ-BS.

Suppose that program (3.18) is a correct formulation of the unconstrained stochastic EPQ-BS. Then, adaptation of results from deterministic cases to stochastic cases is straightforward. Moreover, when the only randomness comes from the demand, we also show that the optimal solution is completely determined by the expectation and the variance of the demands $\boldsymbol{d}_i$. Note that the optimal solution is not obtained by replacing the demand by its expectation.

Lemma 6. *Problem* (3.18) *has a unique optimal solution* $(\tau_i^*)_{i\in\mathcal{I}}$ *given by Equation* (3.19).

Proof. Since Problem (3.18) is a convex problem with qualified constraints, solving it is straightforward using the Karush-Kuhn-Tucker conditions which give the unique solution given by Equation (3.19) and optimal cost of Problem (3.18) is $\frac{1}{2N}\left(\sum_{i\in\mathcal{I}}\sqrt{h_i\check{d}_i}\right)^2$. □

Lemma 7. *For any almost surely finite random variables $\boldsymbol{t}_i$, we have*

$$\limsup_{T\to+\infty}\frac{1}{T}\sum_{i\in\mathcal{I}}\left\lfloor\frac{T-\boldsymbol{t}_i}{\tau_i}\right\rfloor=\sum_{i\in\mathcal{I}}\frac{1}{\tau_i}\qquad\text{almost surely.}\tag{3.20}$$

Proof. We can apply Lemma 3 to each finite realization of the $\boldsymbol{t}_i$'s. Since $\boldsymbol{t}_i$ is assumed to be almost surely finite, the result follows. □

Lemma 8. *Let $\left(\boldsymbol{t}_i,\tau_i\right)_{i\in\mathcal{I}}$ be a feasible solution of unconstrained stochastic EPQ-BS. Then, its cost is at least*

$$\sum_{i\in\mathcal{I}}\frac{1}{2}h_i\left[\left(1-\frac{\mathbb{E}\left[\boldsymbol{d}_i\right]}{v_i}\right)\mathbb{E}\left[\boldsymbol{d}_i\right]-\frac{1}{v_i}\operatorname{Var}\left[\boldsymbol{d}_i\right]\right]\tau_i.\tag{3.21}$$

Proof. Let $\left(\boldsymbol{t}_i,\tau_i\right)_{i\in\mathcal{I}}$ be a feasible solution of unconstrained EPQ-BS. Let $(d_i)_i$ be a positive and finite realization of $\left(\boldsymbol{d}_i\right)_i$. For each realization of $\left(\boldsymbol{t}_i\right)_i$, according to Lemma 4, the average holding cost over infinite horizon is greater than or equal to $\sum_{i\in\mathcal{I}}\frac{1}{2}h_i\left(1-\frac{d_i}{v_i}\right)d_i\tau_i$. Since demand $\boldsymbol{d}_i$ is assumed to be almost surely finite, the expectation of the average holding cost over infinite horizon is greater than or equal to

$$\mathbb{E}\left[\sum_{i\in\mathcal{I}}\frac{1}{2}h_i\left(1-\frac{\boldsymbol{d}_i}{v_i}\right)\boldsymbol{d}_i\tau_i\right]=\sum_{i\in\mathcal{I}}\frac{1}{2}h_i\left[\left(1-\frac{\mathbb{E}\left[\boldsymbol{d}_i\right]}{v_i}\right)\mathbb{E}\left[\boldsymbol{d}_i\right]-\frac{\operatorname{Var}\left[\boldsymbol{d}_i\right]}{v_i}\right]\tau_i\tag{3.22}$$

□

Proof of Theorem 5. Lemma 7 and Lemma 8 prove that every feasible solution of unconstrained EPQ-BS is a feasible solution of Problem (3.18) with the same cost. Conversely, a feasible solution of Problem (3.18) can be completed in a solution of unconstrained EPQ-BS with the same cost setting the first production setup $\boldsymbol{t}_i$ of item i equal to $\frac{s_i(0)}{\boldsymbol{d}_i}$. (We recall that demand is almost surely finite.) Then, the unique optimal solution of Problem (3.18) (Lemma 6) is the optimal solution of unconstrained EPQ-BS. □

3.3.4 Integer stochastic EPQ-BS

When dealing with integer frequencies, we use the following formulation.

$$\begin{aligned} \min \quad & \sum_{i \in \mathcal{I}} \frac{1}{2} h_i \check{d}_i \frac{1}{n_i} && && (3.23a) \\ \text{s.t.} \quad & \sum_{i \in \mathcal{I}} n_i \leqslant N && && (3.23b) \\ & n_i \in \mathbb{Z}_+^* && \forall i \in \mathcal{I}, && (3.23c) \end{aligned}$$

where $n_i = \frac{1}{\tau_i}$ is the average number of setups per time unit over the infinite horizon and $\check{d}_i = \left(1 - \frac{\mathbb{E}\left[\boldsymbol{d}_i\right]}{v_i}\right) \mathbb{E}\left[\boldsymbol{d}_i\right] - \frac{\mathrm{Var}\left[\boldsymbol{d}_i\right]}{v_i}$.

Proving that the optimal solutions of Problem (3.23) are the optimal solutions of stochastic integer EOQ-BS is very similar to the unconstrained case since proofs of Lemma 7 and of Lemma 8 do not rely on the continuities of the cover-sizes $(\tau_i)_i$. Moreover, like in the deterministic case, dealing with a finite or infinite internal production time v_i is very similar since, it is enough to use demand's expectation $\mathbb{E}\left[\boldsymbol{d}_i\right]$ instead of the modified demand $\check{d}_i$ in Problem (3.23).

Finally, adaptation of results proved in deterministic case is straightforward since program (3.23) is a correct formulation of the stochastic integer EPQ-BS.

4 Production on several lines

Chapter 3 was about to decide the optimal cycle stocks when only a single line was involved. In this chapter, several lines are involved in production and we aim at deciding for each line the part of demand assigned and the cycle stocks for each item.

4.1 Motivations

The context is almost the same as in Chapter 3. A company aims at reducing its cycle stocks (which globally helps decrease the inventory) subject to flexibility constraints of the assembly lines. In many applications, a company has many assembly lines and production of a single item may be done on several lines. This kind of configuration is very common and is the consequence of multi-sourcing decisions previously made (see Section 2.4.2 for a description of multi-sourcing issues). These lines might be in the same plants or in different ones. But, in both cases, it enables to increase flexibility of production and to balance the workload between lines.

As in the single-line case, assembly lines are managed using (r,ℓ) policies or similar ones (like policies using cover-sizes). The challenge of the multi-line case is that the company must decide the part of the demand assigned to each line and the lot-sizes (or cover-sizes) used for the production of each item on each line.

In this chapter, we show how to compute the assignment of the demand to lines and the lot-sizes and cover-sizes minimizing inventory subject to flexibility constraints.

4.2 Problem

The problem described by Argon Consulting considers a set $\mathcal{L}$ of assembly lines producing a set $\mathcal{I}$ of items over an infinite horizon. Each line is managed using the same policy than in Chapter 3. The parameters for each line ℓ are then the same (with an index depending on the line) and we recall them for the sake of completeness. The internal production time of item i on line ℓ (i.e., the quantity of item i produced in one time unit) is v_i^ℓ. Inventory of item i must satisfy a demand, known in advance. The whole demand of item i is modeled by a continuous decrease d_i of the inventory, which is assumed positive.

We introduce now the decision variables d_i^ℓ, τ_i and t_i of the problem. The part of demand for item i assigned to line ℓ is denoted d_i^ℓ. It implies that the inventory of item i produced by line ℓ continuously decreases by $d_i^\ell < v_i^\ell$ units per time unit when line ℓ does not produce item i and continuously increases by $v_i^\ell - d_i^\ell$ units per time unit when item i is produced. If item i is assigned to line ℓ, the first time it is produced on line ℓ is t_i^ℓ and the production is launched for every $t_i^\ell + k\tau_i^\ell$ where $k \in \mathbb{Z}_+$ and τ_i^ℓ is the cover-size of item i for the line ℓ. Each production on line ℓ lasts $\frac{d_i^\ell}{v_i^\ell}\tau_i^\ell$ in order to produce exactly the demand assign to line ℓ for the next τ_i^ℓ unit of time. (Each of these quantities is not defined if item i is not assigned to line ℓ.) Like in the single-line case, the productions of several items can be launched simultaneously. The assignment decisions, the production decisions and the demand give the inventory of item i produced by line ℓ which is denoted s_i^ℓ.

Thus, if item i is assigned to line ℓ, inventory $s_i^\ell(t)$ generated by line ℓ is continuous, right and left differentiable, nonnegative and follows the dynamic

$$\dot{s}_i^\ell(t) = \begin{cases} v_i^\ell - d_i^\ell & \text{if } t \in \bigcup_{k\in\mathbb{Z}_+} \left[t_i^\ell + k\tau_i^\ell, t_i^\ell + \left(k + \frac{d_i^\ell}{v_i^\ell}\right)\tau_i^\ell \right), \\ -d_i^\ell & \text{otherwise.} \end{cases} \tag{4.1}$$

Since the whole demand must be satisfied, we have

$$\sum_{\ell\in\mathcal{L}} d_i^\ell = d_i \qquad \forall i \in \mathcal{I}. \tag{4.2}$$

Contrary to the single-line case, each line has a limited capacity. In average over the infinite horizon, the percentage of time spent for production of all items assigned to line ℓ must be smaller than $C^\ell < 1$ which can be written as

$$\sum_{i\in\mathcal{I}} \frac{d_i^\ell}{v_i^\ell} \leqslant C^\ell \qquad \forall \ell \in \mathcal{L}. \tag{4.3}$$

Indeed, each production of item i on line ℓ lasts $\frac{d_i^\ell}{v_i^\ell}\tau_i^\ell$ (in order to produce exactly the demand for the next τ_i unit of time) and then uses $\frac{d_i^\ell}{v_i^\ell}$ percents of the production time.

Like single-line case, in average over the infinite horizon, the number of setups per unit time for all items must be smaller than N^ℓ which can be written as

$$\limsup_{T\to+\infty} \frac{1}{T} \sum_{i\in\mathcal{I}^\ell} \left\lfloor \frac{T - t_i^\ell}{\tau_i^\ell} \right\rfloor \leqslant N \tag{4.4}$$

where $\mathcal{I}^\ell$ is the set of items produced by line ℓ.

For each item i, there is an initial inventory $s_i(0) \in \mathbb{R}_+$ given in input.

Storing one unit of item i produced by line ℓ incurs a unit holding cost $h_i^\ell > 0$ per unit time. The objective is to find the part d_i^ℓ of the demand for item i assigned to line ℓ and the cover-sizes τ_i^ℓ

which minimize the average cycle stock of every lines over infinite horizon

$$\limsup_{T\to+\infty} \frac{1}{T} \sum_{\ell\in\mathcal{I}} \sum_{i\in\mathcal{I}} h_i \int_0^T s_i^\ell(t)\,dt \tag{4.5}$$

while satisfying nonnegative inventory, constraints (4.2) to (4.4).

We call this model the *multi-line Economic Production Quantity model with Bounded number of Setups (the multi-line EPQ-BS).*

Like for the single-line case, when using lot-sizes ℓ_i^ℓ instead of cover-sizes, the multi-line EPQ-BS can be simply adapted using $\ell_i^\ell = d_i^\ell \tau_i^\ell$.

4.3 Bibliography

In the literature, our problem is often formulated as the decision of the assignment of production between many suppliers, the ordered quantities and the time at which they are ordered. For example, Hong and Hayya (1992) propose a model with a unique item produced by many suppliers which have not the same production costs neither the same production qualities. They aim at minimizing the holding cost subject to an upper bound on production cost and a lower bound on production quality. Their approach consists in formulating a multiple-sourcing model and they solve a mathematical programming problem to obtain the optimal selection of suppliers and the size of the split orders.

More recently, Rosenblatt et al. (1998) propose a model considering one item and multiple suppliers. Each order to a supplier follows the classical assumptions of the EOQ model (as described by Harris (1913)) but they have a capacity constraint. They prove that the problem is NP-hard but that it can be efficiently solved with dynamic programming. Chang (2006) proposes an extension to adapt the model to real world constraints. Other models, which also use ordering costs, are proposed by Kim et al. (2005) or Park et al. (2006).

Another important single item model comes from Chauhan and Proth (2003). It considers multiple suppliers delivering to one or multiple plants. The main difference is that each ordering cost is a concave function of the ordered quantity which models the advantage of mono-sourcing the production. They propose an heuristic for their formulation.

The most recent model which is the closest to ours is proposed by Nobil et al. (2016). They consider a set of lines producing a set of items. Each line is managed using classic EPQ model (as described by Taft (1918)). However, contrary to our case, the production of one item cannot be split between two lines. By deciding the assignment and the cover-sizes, they aim at minimizing total cost of the inventory system, setups, production, holding and disposal costs.

4.4 Solving the multi-line EPQ-BS

We address the following alternative mathematical problem. We are going to show that it models the multi-line EPQ-BS.

$$\begin{aligned}
\min \quad & \sum_{i\in\mathcal{I}} \frac{1}{2N^\ell}\left(\sum_{i\in\mathcal{I}}\sqrt{h_i^\ell\left(1-\frac{d_i^\ell}{v_i^\ell}\right)d_i^\ell}\right)^2 && && (4.6a)\\
\text{s.t.} \quad & \sum_{\ell\in\mathcal{L}} d_i^\ell = d_i && \forall i\in\mathcal{I}, && (4.6b)\\
& \sum_{i\in\mathcal{I}} \frac{d_i^\ell}{v_i^\ell} \leqslant C^\ell && \forall \ell\in\mathcal{L}, && (4.6c)\\
& d_i^\ell \geqslant 0 && \forall i\in\mathcal{I}. && (4.6d)
\end{aligned}$$

Proposition 9 proves that the objective function is concave. Minimization of a concave function over a polyhedron is NP-hard since it contains Zero-One Linear Programming (see Raghavachari (1969)) which is known to be NP-hard (see Garey and Johnson (1979)). However, there are many efficients algorithms to solve this kind of problems (see Tuy (1964) for the first proposed algorithm and Benson (1998) for a review) and it is well known that when the polyhedron is a polytope, at least one optimal solution is an extreme point (see Benson (1985)).

Proposition 10 gives the formulas of the optimal cover-sizes when an optimal assignment $\left(d_i^{\ell*}\right)_{i\in\mathcal{I},\ell\in\mathcal{L}}$ has been found. Note that Equation (4.11) is well-defined. Indeed, C^ℓ is smaller than 1. Thus, the constraint (4.6c) and the definition of the cover-sizes only for the items assigned to a line ensure that $\left(1-\frac{d_i^{\ell*}}{v_i}\right)d_i^{\ell*}>0$.

Proposition 9. *The objective function*

$$\begin{aligned}
g: \quad \prod_{\ell\in\mathcal{L},\ i\in\mathcal{I}}\left[0, v_i^\ell\right] \quad &\to \quad \mathbb{R}_+\\
\left(d_i^\ell\right)_{i,\ell} \quad &\mapsto \quad \sum_{i\in\mathcal{I}}\frac{1}{2N^\ell}\left(\sum_{i\in\mathcal{I}}\sqrt{h_i^\ell\left(1-\frac{d_i^\ell}{v_i^\ell}\right)d_i^\ell}\right)^2
\end{aligned} \tag{4.7}$$

of Problem (4.6) *is concave.*

Proof. Let $F: x\in\mathbb{R}_+^{\mathcal{I}} \mapsto \left(\sum_{i\in\mathcal{I}}\sqrt{x_i}\right)^2 \in \mathbb{R}_+$ be a map. F is continuous on $\mathbb{R}_+^{\mathcal{I}}$ and differentiable on $(\mathbb{R}_+^*)^{\mathcal{I}}$ and we have for all x and y in $(\mathbb{R}_+^*)^{\mathcal{I}}$

$$\frac{\partial F}{\partial x_i}(x) = \sqrt{F(x)}\sqrt{x_i} \quad \forall i\in\mathcal{I}, \tag{4.8}$$

and

$$\begin{aligned}
&\langle \nabla F(y)-\nabla F(x), y-x\rangle\\
&= -\frac{1}{2}\sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{I}}\left[\left(\sqrt[4]{\frac{y_i}{y_j}}\sqrt{x_j}-\sqrt[4]{\frac{y_j}{y_i}}\sqrt{x_i}\right)^2+\left(\sqrt[4]{\frac{x_i}{x_j}}\sqrt{y_j}-\sqrt[4]{\frac{x_j}{x_i}}\sqrt{y_i}\right)^2\right] \leqslant 0
\end{aligned} \tag{4.9}$$

Thus, F is concave on $(\mathbb{R}_+^*)^{\mathcal{I}}$. Since F is continuous on $\mathbb{R}_+^{\mathcal{I}}$, it is concave on $\mathbb{R}_+^{\mathcal{I}}$. For each x and x' in $\mathbb{R}_+^{\mathcal{I}}$, the inequality $x \leqslant x'$ implies that $F(x) \leqslant F(x')$. Since the maps $f_i^\ell : d_i^\ell \in [0, v_i^\ell] \mapsto h_i^\ell\left(1 - \frac{d_i^\ell}{v_i^\ell}\right)d_i^\ell$ are concave and since F is concave and monotone (i.e., $z \leqslant z' \implies F(z) \leqslant F(z')$), we have that

$$\begin{aligned} g^\ell : \quad [0, v_1^\ell] \times \cdots \times [0, v_1^\ell] \quad &\rightarrow \quad \mathbb{R}_+ \\ (d_1^\ell, \ldots, d_{\mathcal{I}}^\ell) \quad &\mapsto \quad \left(\textstyle\sum_{i \in \mathcal{I}} \sqrt{h_i^\ell\left(1 - \frac{d_i^\ell}{v_i}\right)d_i^\ell}\right)^2 \end{aligned} \tag{4.10}$$

is concave for each line $\ell \in \mathcal{L}$. Thus the objective function g of Problem (4.6) is concave as linear combination with nonnegative coefficients of concave functions. □

Proposition 10. *An optimal solution* $\left(d_i^{\ell *}\right)_{i \in \mathcal{I}, \ell \in \mathcal{L}}$ *of Problem* (4.6) *can be completed in an optimal solution of multi-line EPQ-BS by setting*

$$\tau_i^{\ell *} = \frac{\sum_{j \in \mathcal{I}} \sqrt{h_j\left(1 - \frac{d_j^{\ell *}}{v_j}\right)d_j^{\ell *}}}{N\sqrt{h_i\left(1 - \frac{d_i^{\ell *}}{v_i}\right)d_i^{\ell *}}} \tag{4.11}$$

when item i assigned to line ℓ.

Proof. Let $\left(d_i^\ell\right)_{i,\ell}$ be a feasible solution of Problem (4.6). We have now $|\mathcal{L}|$ independent problems (one for each line) that correspond to the single-line case. Consider a line $\ell \in \mathcal{L}$ and the set $\mathcal{I}^\ell$ of the items assigned to ℓ. According to Theorem 1, the cover-sizes which minimize the average holding cost over infinite horizon and satisfy nonnegative inventory constraint (4.1) and constraint (4.4) are given by Equation (4.11) and the corresponding cost is equal to

$$\frac{1}{2N}\left(\sum_{i \in \mathcal{I}^\ell} \sqrt{h_i^\ell\left(1 - \frac{d_i^\ell}{v_i^\ell}\right)d_i^\ell}\right)^2. \tag{4.12}$$

This expression is one term of the sum in the objective function (4.6a) and the cover-sizes are feasible to the multi-line EPQ-BS. Indices in this sum expressing the holding costs (4.12) can be extend to $\mathcal{I}$ (since assigned demand is equal to zero). Thus, a feasible solution of Problem (4.6) can be completed in a feasible solution of multi-line EPQ-BS with same cost.

Conversely, let $\left(d_i^\ell, \tau_i^\ell\right)_{i,\ell}$ be a feasible solution of multi-line EPQ-BS. It is clearly feasible for Problem (4.6). According to Theorem 1, for each line ℓ, average holding cost over infinite horizon is greater than or equal to (4.12). Thus, a feasible solution of multi-line EPQ-BS can be completed in a feasible solution of Problem (4.6) whose cost is greater than or equal to (4.12). □

Discrete-time inventory models Part II

5 Deterministic CLSP-BS

5.1 Introduction

5.1.1 Motivations

Fixing the production level for the forthcoming period is a basic decision to be taken when managing an assembly line. Usually, a demand has to be satisfied at due dates but the limited capacity of the line prevents last minute production. On the other hand, too early productions may lead to unnecessary high inventory costs. The challenge of this kind of problems, known as *lot-sizing problems* in the Operations Research community, consists in finding a trade-off between demand satisfaction and holding costs. When several items can be produced on a same line – the so-called *multi-item* lot-sizing problem –, the capacity of the assembly line is often all the more reduced as the number of distinct items produced over the current period is high. Indeed, changing an item in production stops the line for a moment. This additional capacity reduction is usually modeled by setup costs contributing to the total cost.

As explained in Chapter 2, we choose not to model the capacity reduction due to production setups by setup costs but instead by an explicit upper bound on the total number of items that can be produced over a period. Indeed, according to Argon Consulting, many clients aim at minimizing mainly their inventory costs while keeping the number of distinct items produced over each periods below some threshold. This is essentially because, contrary to inventory costs, setup costs are hard to quantify and a maximal number of possible setups per period is easy to estimate.

We are interested in short-term decisions and consider lot-sizing problem and not scheduling problem which are very short-term decisions. At this level of decision, we do not consider the exact scheduling of the item production and approximate the number of setups by the number of items produced over the period.

To the best of our knowledge, the problem addressed in this chapter is original and such a bound on the number of distinct items produced over a period has not been considered by academics yet, with the notable exception of the model proposed by Rubaszewski et al. (2011) but, contrary to our problem, their bound is an overall bound for the whole horizon and they still consider setup costs.

5.1.2 Problem

The problem considers an assembly line producing a set $\mathcal{I}$ of items over T periods. The number of distinct items produced over a period t cannot exceed $N_t > 0$. There is also an upper bound C_t on the total period production (summed over all items) and an upper bound C_t^i on the production of item i at period t. The capacity needed (in time units) to produce one unit of i in period t is $v_t^i > 0$.

The production and the inventory of item i must satisfy a demand d_t^i at the end of period t. When production of item i is not used to satisfy the demand, it can be stored but incurs a unit holding cost $h_t^i > 0$ per period. For each item i, there is an initial inventory $s_0^i \in \mathbb{R}_+$.

The goal is to satisfy the whole demand at minimum cost.

Since this problem is a variation of the *Capacitated Lot-Sizing Problem (CLSP)* with a new flexibility constraint expressed as an upper bound on the number of setups, we call it the *Capacitated Lot-Sizing Problem with Bounded number of Setups (CLSP-BS)*.

In many of our applications, holding costs h_t^i and internal production times v_t^i do not depend on the period and the upper bounds C_t^i and C_t and the maximal number N_t of setups do not depend on the item nor the period. Then, we get a simpler version called the *Uniform Capacitated Lot-Sizing Problem with Bounded number of Setups (Uniform CLSP-BS)* where

$$h_t^i = h^i, \quad v_t^i = v^i, \quad C_t^i = C_t = C, \quad N_t = N. \tag{5.1}$$

This special case captures the essential part and the difficulty of the problem namely the limited flexibility of the assembly line. Most results are true for the CLPS-BS, but we will give counterexamples when results only stand for the Uniform CLSP-BS.

5.1.3 Main results

This chapter presents in Section 5.2 a short review of the seminal lot-sizing models, in Section 5.3 a formulation of the CLSP-BS as a mixed integer program and in Section 5.4 results showing the theoretical difficulty of the CLSP-BS. We briefly describe in Section 5.5 possible way to solve it.

5.2 Bibliography

While there are many variations of lot-sizing problems in industry, they often minimize a combination of holding costs, production costs and setup costs. However, production can concern a single or multiple items and be made on single or several assembly lines. It can be constrained by capacity or setup times. Backlog, safety stock, minimal productivity, setup times can be considered. Review of models are abundant like those proposed by Geunes (2014) for single-item models, by Gicquel et al. (2008) for multi-item models or by Karimi et al. (2003) for both models. A proposition of classification of lot-sizing models is proposed by Pochet and Wolsey (2006, Chapter 4 and 12).

We do not propose a complete review of the models and only consider some seminal models on a single assembly line.

The first model was proposed by Wagner and Whitin (1958). It is an uncapacitated, single-item model over T periods. The demand d_t is dynamic (i.e., time-dependent) and storage is possible between periods. The objective is to minimize holding and setup costs. This model was solved in polynomial time. Its generalization with production cost (proportional to the quantity produced) depending on period is called *Uncapacitated Economic Lot-Sizing Problem (UELSP)* and was solved in polynomial time by Federgruen and Tzur (1991), Wagelmans et al. (1992) and Aggarwal and Park (1993).

The *Capacitated Economic Lot-Sizing Problem (CELSP)* is the same model with production capacities depending on time. It is one of the simplest lot-sizing problem which is NP-hard as shown by Florian et al. (1980) but has a fully polynomial approximation scheme given by van Hoesel and Wagelmans (2001).

In many applications, a line can produce more than one item. These problems are said to be *multi-item*. Two sub-classes are often considered: big bucket models and small bucket models. In big bucket models, several items can be produced during a period. The *Capacitated Lot-Sizing Problem (CLSP)* is a natural multi-item extension of the CELSP and therefore is NP-hard. In small bucket models, only one item per period can be produced. The problem is said to be *continuous* when production can be a fractional part of capacity and *discrete* when production must be done at full capacity. The *Capacitated Setup Lot-sizing Problem (CSLP)* is an example of continuous small bucket model and *Discrete Lot sizing and Scheduling Problem (DLSP)* an example of discrete one (see the review of Gicquel et al. (2008) for the complete description of these models).

Flexibility of the assembly line is often modeled as a cost and included in setup costs. Another way of modeling flexibility consists in adding capacity reduction to the formulation. Each time a setup is placed, the capacity is reduced by a fixed quantity like in the Kellogg's case described by Pochet and Wolsey (2006, Chapter 4). Some lot-sizing and scheduling problems (see for example Guimarães et al. (2014)) model sequence dependent capacity reduction but keep setup costs.

5.3 Model formulation

In this section, we introduce a mixed integer program that models the problem. We introduce the following decision variables. The quantity of item i produced at period t is denoted by q_t^i and the inventory at the end of the period is denoted by s_t^i. We also introduce a binary variable x_t^i which takes the value 1 if the item i is produced during period t.

The CLSP-BS can be written as

$$\begin{aligned}
\min \quad & \sum_{t=1}^{T}\sum_{i\in\mathcal{I}} h_t^i s_t^i && && (5.2a)\\
\text{s.t.} \quad & s_t^i = s_{t-1}^i + q_t^i - d_t^i && \forall t\in[T],\ \forall i\in\mathcal{I}, && (5.2b)\\
& \sum_{i\in\mathcal{I}} v_t^i q_t^i \leqslant C_t && \forall t\in[T], && (5.2c)\\
& v_t^i q_t^i \leqslant C_t^i x_t^i && \forall t\in[T],\ \forall i\in\mathcal{I}, && (5.2d)\\
& \sum_{i\in\mathcal{I}} x_t^i \leqslant N_t && \forall t\in[T], && (5.2e)\\
& x_t^i \in \{0,1\} && \forall t\in[T],\ \forall i\in\mathcal{I}, && (5.2f)\\
& q_t^i,\ s_t^i \geqslant 0 && \forall t\in[T],\ \forall i\in\mathcal{I}. && (5.2g)
\end{aligned}$$

Objective (5.2a) minimize the holding costs. Constraint (5.2b) is the inventory balance. Capacity of the assembly line is ensured by constraint (5.2c). Constraint (5.2d) is both a “big-M” constraint and a capacity of the production of a single item. Constraint (5.2e) limits the number of setups at each period. Note that without loss of generality, we can suppose that $C_t^i \leqslant C_t$ for each period t and each item i.

In the uniform case, holding costs h_t^i and internal production times v_t^i of item i do not depend on time and production capacities C_t^i and C_t depend neither on time nor on item and are equal to $C > 0$. Then, we normalize production variables setting $\widehat{q}_t^i = \frac{v^i q_t^i}{C}$ and replace accordingly the other variables and parameters setting $\widehat{s}_t^i = \frac{v^i s_t^i}{C}$, $\widehat{d}_t^i = \frac{v^i d_t^i}{C}$ and $\widehat{h}^i = \frac{Ch^i}{v^i}$. For the purpose of notation, the hats are omitted and the optimization problem can then be written as

$$\begin{aligned}
\min \quad & \sum_{t=1}^{T}\sum_{i\in\mathcal{I}} h^i s_t^i && && (5.3a)\\
\text{s.t.} \quad & s_t^i = s_{t-1}^i + q_t^i - d_t^i && \forall t\in[T],\ \forall i\in\mathcal{I}, && (5.3b)\\
& \sum_{i\in\mathcal{I}} q_t^i \leqslant 1 && \forall t\in[T], && (5.3c)\\
& q_t^i \leqslant x_t^i && \forall t\in[T],\ \forall i\in\mathcal{I}, && (5.3d)\\
& \sum_{i\in\mathcal{I}} x_t^i \leqslant N && \forall t\in[T], && (5.3e)\\
& x_t^i \in \{0,1\} && \forall t\in[T],\ \forall i\in\mathcal{I}, && (5.3f)\\
& q_t^i,\ s_t^i \geqslant 0 && \forall t\in[T],\ \forall i\in\mathcal{I}. && (5.3g)
\end{aligned}$$

5.4 Theoretical results

In this section, we show that our problem is theoretically hard and that classical methods do not seem to be easily applicable. After showing the NP-completeness in Section 5.4.1, we show that the number N of setups is not captured by the continuous relaxation (Section 5.4.2) nor by a natural valid inequality (Section 5.4.3) nor by natural extended formulations (Section 5.4.4).

5.4.1 NP-completeness

For any fixed values of the h_t^i's, the CLSP-BS is NP-hard.

Theorem 11. *Deciding if there is a solution of the Uniform CLSP-BS is* NP*-complete in the strong sense.*

Reducing 3-partition problem to the Uniform CLSP-BS, we show that deciding if there is a solution of the Uniform CLSP-BS is NP-complete. We remind that the 3-partition problem consists in deciding whether a given multiset of integers can be partitioned into triples that all have the same sum. This problem is known to be NP-complete in the strong sense (see Garey and Johnson (1979)).

Proof. Let $\{a_1, \ldots, a_{3m}\}$ be an instance of the 3-partition problem. We reduce polynomially this problem to an instance of the Uniform CLSP-BS. Without loss of generality, we can assume that the sum of the a_i's is positive. We set

$$\mathcal{I} = \{1, \ldots, 3m\}, \quad T = m, \quad N = 3, \quad d_t^i = \begin{cases} \frac{m a_i}{\sum_{j=1}^{3m} a_j} & \text{if } t = T, \\ 0 & \text{otherwise,} \end{cases} \quad s_0^i = 0. \tag{5.4}$$

Thus, we have a solution for the 3-partition problem if and only if there is a solution to the Uniform CLSP-BS with these parameters. The conclusion follows from the fact that the 3-partition problem is NP-complete in the strong sense. □

The complexity of the following case is the open problem mentioned in Section 1.3.

Open question. What is the complexity status of the uncapacitated Uniform CLSP-BS with $N = 1$?

The complexity of the case with $N = 2$ is also a challenging question.

5.4.2 Relaxations

Continuous relaxation

The goal of this section is to show that, unless the capacity production of one item is smaller than the production of the line at each period, continuous relaxation is not a good way to get bound on the CLSP-BS. It leaves open the question whether it would be possible to find valid inequalities improving the value of the relaxation.

Proposition 12. *Assume $N_t > 0$ for all period t. If $C_t^i \geqslant C_t$ for all item i and all period t, then the continuous relaxation of formulation* (5.2) *does not depend on N_t.*

The immediate corollary of this proposition is that the continuous relaxation of formulation (5.3) never depends on N.

The following example shows that the conclusion of the proposition does not hold when we relax the condition $C_t^i \geqslant C_t$ for all i, t. Consider the instance of CLSP-BS with

$$\begin{gathered}\mathcal{I} = \{1,2\}, \quad T = 2, \quad N_t = 1, \quad C_t = 2, \quad C_t^i = 1, \\ v_t^i = 1, \quad d_t^i = \begin{cases} 0 & \text{if } t = 1, \\ 1 & \text{if } t = 2, \end{cases} \quad h_t^i = h, \quad s_0^i = 0. \end{gathered} \tag{5.5}$$

This instance enables only one item to be produced at each period ($N_t = 1$) whereas the assembly line capacity at one period is equal to the sum of capacities for each item ($C_t = C_t^1 + C_t^2$). Then, the optimal solution of the continuous relaxation of formulation (5.2) with these parameters is h whereas the optimal solution of the continuous relaxation of the formulation (5.2) without the flexibility constraint (5.2e) is 0.

We now prove Proposition 12.

Proof of Proposition 12. Consider an instance of CLSP-BS where $C_t^i \geqslant C_t$ for all periods t and all items i. Let v denote the optimal value of the continuous relaxation of formulation (5.2) and $\widehat{v}$ the optimal value of the continuous relaxation of formulation (5.2) without the flexibility constraint (5.2e). Obviously, we have $v \geqslant \widehat{v}$. Let us show that $v \leqslant \widehat{v}$.

Let $\left(\widehat{x}_t^i, \widehat{q}_t^i, \widehat{s}_t^i\right)_{t,i}$ be a feasible solution of the continuous relaxation of formulation (5.2) without the flexibility constraint (5.2e). For each period t and each item i, we define

$$x_t^i = \begin{cases} \frac{v_t^i \widehat{q}_t^i}{C_t^i} & \text{if } C_t^i > 0, \\ 0 & \text{if } C_t^i = 0, \end{cases} \quad q_t^i = \widehat{q}_t^i, \quad s_t^i = \widehat{s}_t^i \tag{5.6}$$

We now prove that $\left(x_t^i, q_t^i, s_t^i\right)_{t,i}$ is a feasible solution of the continuous relaxation of formulation (5.2).

By definition of q_t^i and s_t^i, constraints (5.2b) and (5.2c) of formulation (5.2) are satisfied.

If $C_t^i > 0$, by definition of x_t^i and q_t^i constraint (5.2d) of formulation (5.2) is satisfied. If $C_t^i = 0$, since $\left(\widehat{x}_t^i, \widehat{q}_t^i, \widehat{s}_t^i\right)_{t,i}$ is feasible, we have $\widehat{q}_t^i = 0$ and constraint (5.2d) is satisfied.

We have

$$C_t \sum_{i \in \mathcal{I}} x_t^i \leqslant \sum_{i \in \mathcal{I}} C_t^i x_t^i \qquad (C_t^i \geqslant C_t) \tag{5.7a}$$

$$\leqslant \sum_{i \in \mathcal{I}} v_t^i \widehat{q}_t^i \qquad (\text{definition of } x_t^i) \tag{5.7b}$$

$$\leqslant C_t. \qquad (\text{feasibility of } \left(\widehat{x}_t^i, \widehat{q}_t^i, \widehat{s}_t^i\right)_{t,i}) \tag{5.7c}$$

If $C_t > 0$, then $\sum_{i \in \mathcal{I}} x_t^i \leqslant 1 \leqslant N_t$. Otherwise, if $C_t = 0$, since $\left(\widehat{x}_t^i, \widehat{q}_t^i, \widehat{s}_t^i\right)_{t,i}$ is feasible, we have $q_t^i = 0$ for each item i and by definition of x_t^i, we have $\sum_{i \in \mathcal{I}} x_t^i = 0 \leqslant N_t$. In both cases, constraint (5.2e) is satisfied.

If $C_t^i > 0$, we have $x_t^i = \frac{v_t^i \widehat{q}_t^i}{C_t^i} \leqslant \widehat{x}_t^i \leqslant 1$. Otherwise, if $C_t^i = 0$, we have by definition $x_t^i = 0$. In both cases, continuous relaxation of constraint (5.2f) is satisfied.

By definition of q_t^i and s_t^i, constraint (5.2g) is satisfied.

Thus, we get a feasible solution of the continuous relaxation of formulation (5.2). So $v \leqslant \widehat{v}$. □

5.4.3 Valid inequalities

A classical valid inequality relies on the absence of backorder and is given by Proposition 13. The statement for the Capacitated Economic Lot-sizing Problem (CELSP) is given by Geunes (2014).

Proposition 13. *For all period t and all item i, we define $K_t^i = \max_{t' \in [t]} \frac{C_{t'}^i}{v_{t'}^i}$ and assume that it is positive. Then,*

$$\sum_{t'=1}^{t} x_{t'}^i \geqslant \left\lceil \frac{1}{K_t^i} \left(\sum_{t'=1}^{t} d_{t'}^i - s_0^i \right) \right\rceil \tag{5.8}$$

is a valid inequality for Equation (5.2).

Proof. At period t, initial inventory and cumulative production must at least cover the cumulative past demand. So

$$s_0^i + \sum_{t'=1}^{t} q_t^i \geqslant \sum_{t'=1}^{t} d_{t'}^i \tag{5.9}$$

Constraint (5.2d) gives

$$q_{t'}^i \leqslant \frac{x_{t'}^i C_{t'}^i}{v_{t'}^i} \leqslant x_{t'}^i K_t^i \tag{5.10}$$

and then

$$\sum_{t'=1}^{t} x_{t'}^i \geqslant \frac{1}{K_t^i} \sum_{t'=1}^{t} \left(d_{t'}^i - s_0^i \right). \tag{5.11}$$

The conclusion comes from the integrity of $\sum_{t'=1}^{t} x_{t'}^i$. □

Note that we can be more precise when writing valid inequality (5.8). It is easy to show that we can remove from the sum $\sum_{t'=1}^{t} x_{t'}^i$ every index t' such that $C_{t'}^i = 0$.

5.4.4 Extended formulations

Two natural extended formulations come to mind. Unfortunately, their continuous relaxations are equal to the one of the formulation (5.2) as we will show in this section.

Model formulations

The first extended formulation is given by the mixed integer program (5.12). We introduce binary variables y_t^p defined for each period $t \in [T]$ and each choice $p \in \binom{\mathcal{I}}{N_t}$ of items. The variable y_t^p is equal to 1 if items in p are allowed to be produced at periods t.

$$
\begin{aligned}
\min \quad & \sum_{t=1}^{T} \sum_{i \in \mathcal{I}} h_t^i s_t^i && && \text{(5.12a)} \\
\text{s.t.} \quad & s_t^i = s_{t-1}^i + q_t^i - d_t^i && \forall t \in [T],\ \forall i \in \mathcal{I}, && \text{(5.12b)} \\
& \sum_{i \in \mathcal{I}} v_t^i q_t^i \leqslant C_t && \forall t \in [T], && \text{(5.12c)} \\
& v_t^i q_t^i \leqslant C_t^i \sum_{p \in \binom{\mathcal{I}}{N_t} \mid i \in p} y_t^p && \forall t \in [T],\ \forall i \in \mathcal{I}, && \text{(5.12d)} \\
& \sum_{p \in \binom{\mathcal{I}}{N_t}} y_t^p \leqslant 1 && \forall t \in [T], && \text{(5.12e)} \\
& y_t^p \in \mathbb{N} && \forall t \in [T],\ p \in \binom{\mathcal{I}}{N_t}, && \text{(5.12f)} \\
& q_t^i,\ s_t^i \geqslant 0 && \forall t \in [T],\ \forall i \in \mathcal{I}, && \text{(5.12g)}
\end{aligned}
$$

The second extended formulation is given by the mixed integer program (5.13). We introduce binary variables z_ℓ^i defined for each item $i \in \mathcal{I}$ and each choice $\ell \subseteq [T]$ of periods. The variable z_ℓ^i is equal to 1 if item i is produced during the periods of ℓ.

$$
\begin{aligned}
\min \quad & \sum_{t=1}^{T} \sum_{i \in \mathcal{I}} h_t^i s_t^i && && \text{(5.13a)} \\
\text{s.t.} \quad & s_t^i = s_{t-1}^i + q_t^i - d_t^i && \forall t \in [T],\ \forall i \in \mathcal{I}, && \text{(5.13b)} \\
& \sum_{i \in \mathcal{I}} v_t^i q_t^i \leqslant C_t && \forall t \in [T], && \text{(5.13c)} \\
& v_t^i q_t^i \leqslant C_t^i \sum_{\ell \subseteq [T] \mid t \in \ell} z_\ell^i && \forall t \in [T],\ \forall i \in \mathcal{I}, && \text{(5.13d)} \\
& \sum_{i \in \mathcal{I}} \sum_{\ell \subseteq [T] \mid t \in \ell} z_\ell^i \leqslant N_t && \forall t \in [T], && \text{(5.13e)} \\
& \sum_{\ell \subseteq [T]} z_\ell^i \leqslant 1 && \forall i \in \mathcal{I}, && \text{(5.13f)} \\
& z_\ell^i \in \mathbb{N} && \forall t \in [T],\ \ell \subseteq [T], && \text{(5.13g)} \\
& q_t^i,\ s_t^i \geqslant 0 && \forall t \in [T],\ \forall i \in \mathcal{I}, && \text{(5.13h)}
\end{aligned}
$$

It is easy to see that the extended formulations (5.12) and (5.13) also model the CLSP-BS.

Relaxations of extended formulation (5.12)

Proposition 14. *The compact formulation* (5.2) *of CLSP-BS and the extended formulation* (5.12) *of CLSP-BS have the same continuous relaxations.*

In order to prove Proposition 14, we need the following lemma.

Lemma 15. *Let $I \in \mathbb{N}^*$ and let $N \in [I]$. Let $(x_1, \dots, x_I) \in [0,1]^I$ be such that $\sum_{i=1}^I x_i = N$. Then, there exists $(y_p)_{p \in \binom{[I]}{N}} \in [0,1]^{\binom{[I]}{N}}$ such that $\sum_{p \in \binom{[I]}{N}} y_p = 1$ and for each $i \in [I]$, we have $x_i \leqslant \sum_{p \in \binom{[I]}{N} \mid i \in p} y_p$.*

Remark 1. In case $N = 2$, Lemma 15 gives the following result on a graph. If $(x_i)_i$ represents the weights of the vertices of a complete graph such that $0 \leqslant x_i \leqslant 1$ and $\sum_i x_i = 2$, then there exists weights $\left(y_p\right)_p$ of the edges such that the total sum of the weights of the edges is equal to 1 and for each vertex, the sum of the weights of its incident edges is greater or equal to its weight.

Proof of Lemma 15. Let $I \in \mathbb{N}^*$ and let $N \in [I]$. Let $(x_1, \dots, x_I) \in [0,1]^I$ be such that $\sum_{i=1}^I x_i = N$. To prove the lemma, it is sufficient to prove that the linear program (5.14) is feasible and that its optimum is equal to 1.

$$\min \quad \sum_{p \in \binom{[I]}{N}} y_p \tag{5.14a}$$

$$\text{s.t.} \quad \sum_{p \in \binom{[I]}{N} \mid i \in p} y_p \geqslant x_i \qquad \forall i \in [I] \tag{5.14b}$$

$$y_p \geqslant 0 \qquad \forall p \in \binom{[I]}{N} \tag{5.14c}$$

So it is sufficient to prove that the dual (5.15) of program (5.14) is feasible and that its optimum v_D is equal to 1.

$$\max \quad \sum_{i=1}^{I} x_i \lambda_i \tag{5.15a}$$

$$\text{s.t.} \quad \sum_{i \in p} \lambda_i \leqslant 1 \qquad \forall p \in \binom{[I]}{N} \tag{5.15b}$$

$$\lambda_i \geqslant 0 \qquad \forall i \in [I] \tag{5.15c}$$

First, note that $v_D \geqslant 1$. Indeed, $(\lambda_1, \dots, \lambda_I) = \left(\frac{1}{N}, \dots, \frac{1}{N}\right)$ is feasible and provides the value 1 to the objective function. Let us show that $v_D \leqslant 1$.

We assume without loss of generality that $x_1 \geqslant \cdots \geqslant x_I$. Let λ^* be an optimal solution of program (5.15). It is easy to show that we can require

$$\lambda_1^* \geqslant \cdots \geqslant \lambda_I^* \text{ and } \sum_{i=1}^{N} \lambda_i^* = 1. \tag{5.16}$$

Moreover, since $\sum_{i \in p} \lambda_i^*$ is maximal when $p = [N]$, we can also require that

$$\lambda_N^* = \cdots = \lambda_I^*. \tag{5.17}$$

Finally, among all λ^* satisfying these requirements, we choose one with maximal λ_1^*, which exists by compactness. We deal with two cases.

First case: $\lambda_1^* < 1$. Suppose first for a contradiction that $\lambda_2^* > \lambda_I^*$. Then, denote ℓ the largest index such that $\lambda_\ell^* = \lambda_2^*$. The fact that $\lambda_N^* = \cdots = \lambda_I^*$ implies that $\ell < N$. Increasing λ_1^* by a small $\varepsilon = \lambda_\ell^* - \lambda_{\ell+1}^* > 0$ and decreasing λ_ℓ^* by the same amount provides a new optimal solution (feasible because $\sum_{i \in p} \lambda_i^*$ is maximal when $p = [N]$), with a larger λ_1. This is in contradiction with the definition of λ^*. We have thus $\lambda_2^* = \cdots = \lambda_I^*$ and v_D is upper bounded by the optimal value of

$$\begin{aligned}
\max \quad & x_1\lambda_1 + \left(\sum_{i=2}^{I} x_i\right)\lambda_2 && (5.18a)\\
\text{s.t.} \quad & \lambda_1 + (N-1)\lambda_2 = 1 && (5.18b)\\
& \lambda_2 - \lambda_1 + z = 0 && (5.18c)\\
& \lambda_1,\ \lambda_2,\ z \geqslant 0 && (5.18d)
\end{aligned}$$

The variable z is introduced so that the program is in standard form. Let $(\tilde\lambda_1, \tilde\lambda_2, \tilde z)$ be an optimal basic solution. If the optimal basis is $\{\lambda_1, \lambda_2\}$, then $\tilde z = 0$ and we have $\tilde\lambda_1 = \tilde\lambda_2 = 1/N$, which gives a value 1 to the objective function of problem (5.18). If the optimal basis is $\{\lambda_1, z\}$, then $\tilde\lambda_2 = 0$ and we have $\tilde\lambda_1 = 1$, which gives a value x_1 to the objective function of problem (5.18). The basis $\{\lambda_2, z\}$ being not feasible, we get $v_D \leqslant \max(x_1, 1) = 1$.

Second case: $\lambda_1^* = 1$. We have obviously $\lambda_2^* = \cdots = \lambda_I^* = 0$, and thus $v_D = x_1 \leqslant 1$.

In both cases, $v_D \leqslant 1$, as required. Hence $v_D = 1$ which concludes the proof. □

Proof of Proposition 14. Let $\left(s_t^i, q_t^i, y_t^p\right)_{t,i,p}$ be a solution of the continuous relaxation of (5.12). For each period t and each item i, we set

$$x_t^i = \sum_{p \in \binom{\mathcal{I}}{N_t} \mid i \in p} y_t^p. \tag{5.19}$$

Since $\left(s_t^i, q_t^i, y_t^p\right)_{t,i,p}$ is a feasible solution of the continuous relaxation of (5.12), we have

$$x_t^i = \sum_{p \in \binom{\mathcal{I}}{N_t} \mid i \in p} y_t^p \leqslant \sum_{p \in \binom{\mathcal{I}}{N_t}} y_t^p \leqslant 1 \qquad \forall t \in [T],\ \forall i \in \mathcal{I}, \tag{5.20}$$

and

$$\sum_{i \in \mathcal{I}} x_t^i = \sum_{i \in \mathcal{I}} \sum_{p \in \binom{\mathcal{I}}{N_t} \mid i \in p} y_t^p = \sum_{p \in \binom{\mathcal{I}}{N_t}} N_t y_t^p = N_t \sum_{p \in \binom{\mathcal{I}}{N_t}} y_t^p \leqslant N_t \qquad \forall t \in [T]. \tag{5.21}$$

So $\left(s_t^i, q_t^i, x_t^i\right)_{t,i}$ is a solution of the continuous relaxation of (5.2) with the same cost as $\left(s_t^i, q_t^i, y_t^p\right)$.

Let $\left(s_t^i, q_t^i, x_t^i\right)_{t,i}$ be a solution of the continuous relaxation of (5.2). Without loss of generality, we can suppose that for each period t, we have $\sum_{i \in \mathcal{I}} x_t^i = N_t$. For each period t, according to Lemma 15, there exists $\left(y_t^p\right)_{p \in \binom{\mathcal{I}}{N_t}} \in [0,1]^{\binom{\mathcal{I}}{N_t}}$ such that

$$\sum_{p \in \binom{\mathcal{I}}{N_t}} y_t^p = 1 \text{ and for each } i \in [I],\ x_t^i \leqslant \sum_{p \binom{\mathcal{I}}{N_t} \mid i \in p} y_t^p. \tag{5.22}$$

Since $\left(s_t^i, q_t^i, x_t^i\right)_{t,i}$ is a solution of the continuous relaxation of (5.2), we have

$$v_t^i q_t^i \leqslant C_t^i x_t^i \leqslant C_t^i \sum_{p \in \binom{\mathcal{I}}{N_t} \mid i \in p} y_t^p \qquad \forall t \in [T],\ \forall i \in \mathcal{I}. \tag{5.23}$$

So $\left(s_t^i, q_t^i, y_t^p\right)_{t,i,p}$ is a solution of the continuous relaxation of (5.12) with the same cost as $\left(s_t^i, q_t^i, x_t^i\right)_{t,i}$. □

Relaxations of extended formulation (5.13)

Proposition 16. *The compact formulation* (5.2) *of CLSP-BS and the extended formulation* (5.13) *of CLSP-BS have the same continuous relaxations.*

Proof. Let $\left(s_t^i, q_t^i, z_\ell^i\right)_{t,i,\ell}$ be a solution of the continuous relaxation of (5.13). For each period t and each item i, we set

$$x_t^i = \sum_{\ell \subseteq [T] \mid t \in \ell} z_\ell^i. \tag{5.24}$$

Since $\left(s_t^i, q_t^i, z_\ell^i\right)_{t,i,\ell}$ is a feasible solution of the continuous relaxation of (5.13), we have

$$x_t^i = \sum_{\ell \subseteq [T] \mid t \in \ell} z_\ell^i \leqslant \sum_{\ell \subseteq [T]} z_\ell^i \leqslant 1 \qquad \forall t \in [T], \forall i \in \mathcal{I}, \tag{5.25}$$

and

$$\sum_{i \in \mathcal{I}} x_t^i = \sum_{i \in \mathcal{I}} \sum_{\ell \subseteq [T] \mid t \in \ell} z_\ell^i \leqslant N_t \qquad \forall t \in [T]. \tag{5.26}$$

So $\left(s_t^i, q_t^i, x_t^i\right)_{t,i}$ is a solution of the continuous relaxation of the compact formulation (5.2) with the same cost as $\left(s_t^i, q_t^i, z_\ell^i\right)_{t,i,\ell}$.

Let $\left(s_t^i, q_t^i, x_t^i\right)_{t,i}$ be a solution of the continuous relaxation of the compact formulation (5.2). For each item i, we define the sequence $\left(\ell_p^i\right)_p$ of subsets of $[T]$ and the associated sequence $\left(z_{\ell_p^i}^i\right)_p$ of integers as follow

$$\begin{aligned} \ell_1^i &= \left\{ t \in [T] \;\middle|\; x_t^i > 0 \right\}, \\ z_{\ell_1^i}^i &= \min \left\{ x_t^i \;\middle|\; t \in \ell_1^i \right\}, \end{aligned} \tag{5.27}$$

and for $p \geqslant 1$, while $\ell_p^i \neq \emptyset$,

$$\begin{aligned} \ell_{p+1}^i &= \left\{ t \in [T] \;\middle|\; x_t^i - \sum_{q=1}^{p} z_{\ell_p^i}^i > 0 \right\}, \\ z_{\ell_{p+1}^i}^i &= \min \left\{ x_t^i - \sum_{q=1}^{p} z_{\ell_p^i}^i \;\middle|\; t \in \ell_p^i \right\}. \end{aligned} \tag{5.28}$$

The sequence $(\ell_p^i)_p$ is finite and strictly decreasing. Note that z_ℓ^i is now defined for each ℓ occurring in the sequence $(\ell_p^i)_p$. For the others, we set $z_\ell^i = 0$. An example for the definition of the $\left(z_\ell^i\right)_\ell$ is given on Figure 5.1.

Then, we have $x_t^i = \sum_{\ell \subseteq [T] \mid t \in \ell} z_\ell^i$ and $\left(s_t^i, q_t^i, z_\ell^i\right)_{t,i,\ell}$ is a solution of continuous relaxation of the extended formulation (5.13). □

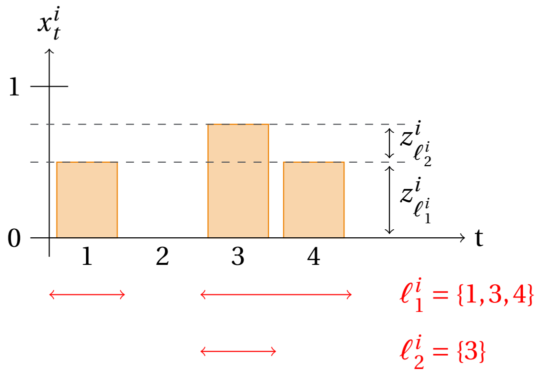


Figure 5.1 – Example of definition of the variables $\left(z_\ell^i\right)_\ell$ for four periods

5.5 Solving the CLSP-BS

We present in this section different ways to solve the CLSP-BS.

5.5.1 Off-the-shelf solver

In the data provided by Argon Consulting from its clients, we have at most 30 items and 13 periods. So the number of binary variables is always smaller to 400, which is very small for the up-to-date solvers. We succeed in getting the optimal solution in less than 20 seconds.

5.5.2 Dynamic programming with fixed $\mathcal{I}$

For a fixed set $\mathcal{I}$ of items, CLSP-BS with infinite capacities is polynomial.

Theorem 17. *For a fixed set $\mathcal{I}$ of items, CLSP-BS with infinite capacities ($C_t = C = +\infty$) is polynomial.*

We first prove that, when there are no initial inventory, there exists an optimal solution such that a production can occurs only if inventory is equal to zero. Then, using dynamic programming, we prove that CLSP-BS with infinite capacities can be solved in $O\left(T^{2|\mathcal{I}|+1}\right)$.

Lemma 18. *Consider an instance of CLSP-BS with infinite capacities ($C_t = C = +\infty$) and no initial inventory ($s_0^i = 0,\ \forall i \in \mathcal{I}$). Then, there exists an optimal solution that is ZIO (Zero-Inventory-Ordering).*

Proof. Let (q, s, x) be an optimal solution of the CLSP-BS with infinite capacity and no initial inventory. Suppose that there exists a period t and an item j such that $s_{t-1}^j q_t^j \neq 0$. We choose the smallest t such that $s_{t-1}^j q_t^j \neq 0$. Since inventory at time t is positive and since there is no initial inventory, there exists a period t_0 in $\{1, \dots, t-1\}$ such that $q_{t_0}^j > 0$ and we choose t_0 maximal. Then, decreasing $q_{t_0}^j$ by s_{t-1}^j and increasing q_t^j by s_{t-1}^j give a feasible solution (since capacities are infinite) whose cost is at least as good as the cost of the initial solution (q, s, x). Using this transformation for each period t and item j such that $s_{t-1}^j q_t^j \neq 0$, we get a solution which is ZIO and at least as good as the initial optimal solution. □

Proof of Theorem 17. Let us first note that solving the problem with initial inventory is the same as solving the problem with no initial inventory and with the demand decreased from the first period to the last by the amount that can be satisfied by the initial inventory. It can be done in $O(T|\mathcal{I}|)$ operations. Thus, we can assume than there is no initial inventory.

Lemma 18 implies that there exists an optimal solution that is ZIO and we can use the dynamic programming as follow. At period t, the inventory level s_t^i of item i is necessarily of the form $\sum_{t'=t}^{\tilde{t}} d_t^i$ for some $\tilde{t}$. We define the set $\mathcal{S}_t$ of possible inventory levels of items as

$$\mathcal{S}_t = \left\{ \left(\sum_{t'=t}^{\tilde{t}_i} d_{t'}^i \right)_{i\in\mathcal{I}} \middle| \tilde{t} \in \{t, \ldots, T\}^{\mathcal{I}} \right\} \qquad \forall t \in [T]. \tag{5.29}$$

For a period $t \in [T-1]$, a transition from $s_t \in \mathcal{S}_t$ to $s_{t+1} \in \mathcal{S}_{t+1}$ is feasible if and only if $s_{t+1}^i \neq s_t^i - d_{t+1}^i$ for at most N indices i. Then, the cost of a transition from $s_t \in \mathcal{S}_t$ to $s_{t+1} \in \mathcal{S}_{t+1}$ is

$$c_t(s_t, s_{t+1}) = \begin{cases} \sum_{i\in\mathcal{I}} h_t^i s_t^i & \text{if the transition is feasible,} \\ +\infty & \text{otherwise.} \end{cases} \tag{5.30}$$

Let $V_t(s_t)$ be the minimal holding cost for times t to T with inventory s_t at period t. We have

$$V_t(s_t) = \begin{cases} \min\limits_{s_{t+1}\in\mathcal{S}_{t+1}} \{c_t(s_t, s_{t+1}) + V_{t+1}(s_{t+1})\} & \text{if } t \in [T-1], \\ \sum_{i\in\mathcal{I}} h_t^i s_t^i & \text{otherwise.} \end{cases} \tag{5.31}$$

There exists an optimal solution such that inventory is equal to zero at the end of the horizon T. Thus, we can compute the cost $V(0)$ of an optimal solution in $O\big(T^{2|\mathcal{I}|+1}\big)$. $\square$

6 Stochastic CLSP-BS

6.1 Motivations and problem

In Chapter 5, data are deterministic. In practice, part of data is uncertain. Depending on the industry or on the planning horizon, uncertainty can be on demand forecast, production rate, capacities or any other part of the supply chain. When making a production planning, the main uncertainty in cases given by Argon Consulting comes from demand forecast. Thus, in this chapter, we only consider this source of randomness in the model to get the stochastic counterpart of the Uniform CLSP-BS.

As for the Uniform CLSP-BS in Chapter 5, we consider an assembly line producing a set $\mathcal{I}$ of items over T periods. The number of distinct items produced over a period cannot exceed N. There is also an upper bound on the total period production (summed over all items). We remind that we normalize all quantities so that this upper bound is equal to 1.

The production must satisfy a random demand as well as possible. However, because of uncertainty, backorder (i.e., late delivery) is allowed. Part of demand satisfied on time must meet a fill rate service level β as defined in Chapter 2. The demand of item i over period t is a random parameter $\boldsymbol{d}_t^i$, whose realization is known at the end of period t. When production of an item i is not used to satisfy the demand, it can be stored but incurs a unit holding cost $h^i > 0$ per period. For each item i, there is an initial inventory $s_0^i \in \mathbb{R}_+$.

Regarding randomness, we assume that for any i and t, realizations of $\left(\boldsymbol{d}_t^i, \ldots, \boldsymbol{d}_T^i\right)$ have finite expectation and can be efficiently sampled, knowing a realization of $\left(\boldsymbol{d}_1^j, \ldots, \boldsymbol{d}_{t-1}^j\right)_{j \in \mathcal{I}}$.

Decisions can be made at the beginning of each period knowing past realizations of the demand. In particular, decisions for the current period are definitive decisions, but decisions for the following periods may change depending on the realization of the demand at the end of the period. This kind of formulation is called *multi-stage*.

Finally, the model must always be feasible. Indeed, in real applications, the model must always return a production planning. That's why some constraints are "soft constraints". For example, in case of capacity issue, production must be planned in such a way as to best approach service level even if it cannot reach it.

Adding uncertainty in the formulation makes the description of the problem incomplete. We propose two closely related formulations of the problem in Section 6.3. We discuss some of the modeling assumptions in Section 6.5.

6.2 Bibliography

Most of deterministic models – like those presented in Section 5.2 – have been extended to deal with uncertainty. A review of old models is presented by Yano and Lee (1995) in the second part of their paper that looks at the models where the yields (e.g., ordered quantity may differ from received quantity) are uncertain and where the others parameters may depend on these randomness. More recently, we can cite the reviews of Mula et al. (2006) and of Díaz-Madroñero et al. (2014). One of the most complete reviews for the post 2000 publications dealing with stochastic lot-sizing is proposed by Aloulou et al. (2014). This review shows that uncertainty is mostly represented in a first approach with probabilistic and stochastic programming formulations and to a lesser extend, with fuzzy logic or scenario formulations. Likewise, we see that the first source of uncertainty in the models is mostly the demand, then the costs and finally other parameters like yields or lead-times. Latest works present other sources of uncertainty like uncertainty on setup-times (Taş et al. (2018)).

Integrating uncertainty in models allow to apply lot-sizing models to new fields. For example, Mukhopadhyay and Ma (2009) and Macedo et al. (2016) deal with remanufacturing where the main uncertainty comes from the quantity and the quality of returned pieces which impact the quantity of required raw materials. We also find it in automotive industry where composition of the product mix is highly uncertain (see Gyulai et al. (2015)). Stochastic production planning also deals with huge size problems like electricity production (see Rozenknop et al. (2013)) and it requires to develop special heuristics to solve them.

Since we consider uncertainty, demand may not be satisfied. In order to control backorder, many methods exist. A classical one consists in using backorder costs (see Zangwill (1969) and Absi and Kedad-Sidhoum (2009)) that is paid for undelivered quantities. However, as explained by Tempelmeier (2013), backorder costs may be hard to quantify due to indirect consequences of undelivered quantities and he propose several service level measures that can be used as a substitute. Chance constraints, which are quite close to cycle service levels, are also used (see Tarim and Kingsman (2004) and Gicquel and Cheng (2018)). Non-satisfied demand is also controlled using robust optimization like in the models proposed by Gyulai et al. (2017) or by Minoux (2018) who obtains better results than classical stochastic optimization when targeting very low risk levels. Since backorder costs remain in many cases the easiest way of dealing with undelivered quantity, van Houtum and Zijm (2000) propose for the stochastic single-item inventory systems a method to link backorder costs to several service level measures. They provide conditions for which pure cost models can be transform into to service models and give examples as for the single-stage periodic-review inventory system. A similar approach is used in Section 6.3.2 to compute backorder costs in case of fill rate service level.

Finally, to the best of our knowledge, we observe than, like in deterministic cases, problems use setup costs and not a bound on the number of setups.

6.3 Model

In order to formalize the stochastic counterpart of the Uniform CLSP-BS, we introduce the following decision variables. The quantity of item i produced at period t is denoted by $\boldsymbol{q}_t^i$ and the inventory at the end of the period is denoted by $\boldsymbol{s}_t^i$. We also introduce a binary variable $\boldsymbol{x}_t^i$ which takes the value 1 if the item i is produced during period t. All these variables are random and may depend on the past realizations of the random demand $\left(\boldsymbol{d}_1^j, \dots, \boldsymbol{d}_{t-1}^j\right)_{j\in\mathcal{I}}$.

6.3.1 Model with service level constraint

For each period t and each item i, we introduce the decision variable $\tilde{\boldsymbol{d}}_t^i$ which is the part of demand $\boldsymbol{d}_t^i$ satisfied at the end of period t. We decide to model the service level constraint for all items by

$$\mathbb{E}\left[\sum_{i\in\mathcal{I}} w^i \frac{\sum_{t=1}^T \tilde{\boldsymbol{d}}_t^i}{\sum_{t=1}^T \boldsymbol{d}_t^i}\right] \geqslant \beta \quad \text{where} \quad w^i = \frac{h^i}{\sum_{j\in\mathcal{I}} h^j}. \tag{6.1}$$

The fill rate service level is taken on average over items and each item is weighted by its holding cost.

Then, we can write the mathematical program corresponding to our problem at time t

$$\begin{aligned}
\min \quad & \mathbb{E}\left[\sum_{t'=t}^{T}\sum_{i\in\mathcal{I}} h^i \boldsymbol{s}_{t'}^i\right] && && \text{(6.2a)}\\
\text{s.t.} \quad & \boldsymbol{s}_{t'}^i = \boldsymbol{s}_{t'-1}^i + \boldsymbol{q}_{t'}^i - \tilde{\boldsymbol{d}}_{t'}^i && \forall t' \in \{t,\dots,T\}, \forall i \in \mathcal{I}, && \text{(6.2b)}\\
& \sum_{i\in\mathcal{I}} \boldsymbol{q}_{t'}^i \leqslant 1 && \forall t' \in \{t,\dots,T\}, && \text{(6.2c)}\\
& \boldsymbol{q}_{t'}^i \leqslant \boldsymbol{x}_{t'}^i && \forall t' \in \{t,\dots,T\}, \forall i \in \mathcal{I}, && \text{(6.2d)}\\
& \sum_{i\in\mathcal{I}} \boldsymbol{x}_{t'}^i \leqslant N && \forall t' \in \{t,\dots,T\}, && \text{(6.2e)}\\
& \mathbb{E}\left[\sum_{i\in\mathcal{I}} w^i \frac{\sum_{t=1}^T \tilde{\boldsymbol{d}}_t^i}{\sum_{t=1}^T \boldsymbol{d}_t^i}\right] \geqslant \beta && && \text{(6.2f)}\\
& \tilde{\boldsymbol{d}}_{t'}^i \leqslant \boldsymbol{d}_{t'}^i && \forall t' \in \{t,\dots,T\}, \forall i \in \mathcal{I}, && \text{(6.2g)}\\
& \boldsymbol{x}_{t'}^i \in \{0,1\} && \forall t' \in \{t,\dots,T\}, \forall i \in \mathcal{I}, && \text{(6.2h)}\\
& \boldsymbol{q}_{t'}^i,\ \boldsymbol{s}_{t'}^i,\ \tilde{\boldsymbol{d}}_{t'}^i \geqslant 0 && \forall t' \in \{t,\dots,T\}, \forall i \in \mathcal{I}, && \text{(6.2i)}\\
& \boldsymbol{q}_{t'}^i \text{ is } \sigma\Big(\big(\boldsymbol{d}_1^i,\dots,\boldsymbol{d}_{t'-1}^i\big)_{i\in\mathcal{I}}\Big)\text{–measurable} && \forall t' \in \{t,\dots,T\}, \forall i \in \mathcal{I}. && \text{(6.2j)}
\end{aligned}$$

Objective (6.2a) minimizes the future expected holding costs. Constraints (6.2b), (6.2c), (6.2d) and (6.2e) have the same meaning than their deterministic counterparts (5.3b), (5.3c), (5.3d) and (5.3e). Constraint (6.2f) ensures the service level. Constraint (6.2g) means that we cannot satisfy more than the demand. Last constraint (6.2j) of the program, written as a *measurability constraint*, means that the values of the variables $\boldsymbol{q}_{t'}^i$ can only depend on the values taken by the demand before time t' (the planner does not know the future). Every constraint of the problem, except the service level constraint (6.2f), holds almost surely.

One may expect that the satisfied part $\tilde{\boldsymbol{d}}_t^i$ of demand is also bounded by the sum of the previous inventory and the production of the period, but such a constraint can be deduced from constraints (6.2b) and (6.2i).

This formulation perfectly matches Argon Consulting objectives. Indeed, the objective is simple (only holding costs) and does not weigh hardly comparable indicators. Moreover, every parameter is easily provided by clients.

The drawback of this formulation is the feasibility. Because of the service level constraint (6.2f), program (6.2) may be infeasible. In this cases, heuristics must be developed to create a production planning which violates as few as possible the constraints.

6.3.2 Model with backorder costs

Service level constraint is a main issue since it may lead to infeasibility of the model. Thus, we remove this constraint and penalize backorder quantities. We introduce new decisions variables. When a demand for item i is not satisfied by the production of the current period or by inventory, it can be satisfied later but incurs a unit backorder cost γ^i per period for some coefficient $\gamma^i > 0$ and the backorder of item i at the end of the period t is denoted by $\boldsymbol{b}_t^i$. We also used the inventory level $\tilde{\boldsymbol{s}}_t^i$ which is the relative value of the inventory of item i at the end of period t (i.e., the inventory minus the backorder).

The problem at time t can be written as follow.

$$
\begin{aligned}
\min \quad & \mathbb{E}\left[\sum_{t'=t}^{T}\sum_{i\in\mathcal{I}}\left(h^i \boldsymbol{s}_{t'}^i + \gamma^i \boldsymbol{b}_{t'}^i\right)\right] & & \text{(6.3a)}\\
\text{s.t.} \quad & \tilde{\boldsymbol{s}}_{t'}^i = \tilde{\boldsymbol{s}}_{t'-1}^i + \boldsymbol{q}_{t'}^i - \boldsymbol{d}_{t'}^i & \forall t' \in \{t,\dots,T\},\ \forall i \in \mathcal{I}, & \quad \text{(6.3b)}\\
& \sum_{i\in\mathcal{I}} \boldsymbol{q}_{t'}^i \leqslant 1 & \forall t' \in \{t,\dots,T\}, & \quad \text{(6.3c)}\\
& \boldsymbol{q}_{t'}^i \leqslant \boldsymbol{x}_{t'}^i & \forall t' \in \{t,\dots,T\},\ \forall i \in \mathcal{I}, & \quad \text{(6.3d)}\\
& \sum_{i\in\mathcal{I}} \boldsymbol{x}_{t'}^i \leqslant N & \forall t' \in \{t,\dots,T\}, & \quad \text{(6.3e)}\\
& \tilde{\boldsymbol{s}}_{t'}^i = \boldsymbol{s}_{t'}^i - \boldsymbol{b}_{t'}^i & \forall t' \in \{t,\dots,T\},\ \forall i \in \mathcal{I}, & \quad \text{(6.3f)}\\
& \boldsymbol{x}_{t'}^i \in \{0,1\} & \forall t' \in \{t,\dots,T\},\ \forall i \in \mathcal{I}, & \quad \text{(6.3g)}\\
& \boldsymbol{q}_{t'}^i,\ \boldsymbol{s}_{t'}^i,\ \boldsymbol{b}_{t'}^i \geqslant 0 & \forall t' \in \{t,\dots,T\},\ \forall i \in \mathcal{I}, & \quad \text{(6.3h)}\\
& \boldsymbol{q}_{t'}^i \text{ is } \sigma\left(\left(\boldsymbol{d}_1^i,\dots,\boldsymbol{d}_{t'-1}^i\right)_{i\in\mathcal{I}}\right)\text{–measurable} & \forall t' \in \{t,\dots,T\},\ \forall i \in \mathcal{I}. & \quad \text{(6.3i)}
\end{aligned}
$$

In contrast to model (6.2), objective (6.3a) minimizes the sum of inventory and backorder at the end of each period and service level constraint (6.2f) is replaced by constraint (6.3f) which links real inventory and backorder quantities.

An interesting feature of this model is that there always exists a feasible solution, which makes it more amenable to real-world applications. However, it cannot guarantee a specified service level. Moreover, parameters of the first model were easy to get whereas in practice, except when they are enshrined through contracts with the clients, backorder costs can be hard to estimate.

Finally, note that the models (6.2) and (6.3) are not equivalent since the service level constraint does not make a difference between a one-period delay and a two-period delay whereas the backorder costs increase with the length of the delay.

When backorder costs are not given by the clients, we propose a way to “price” backorder coefficients γ^i for each item i before the first period, with the idea to heuristically entice it to choose solutions satisfying service level constraint (6.2f). We set

$$\gamma^i := \frac{\mathbb{P}\left[\boldsymbol{d}^i \leqslant q^i(\beta)\right]}{\mathbb{P}\left[\boldsymbol{d}^i > q^i(\beta)\right]} h^i \tag{6.4}$$

with

$$q^i(\beta) := \inf\left\{ q \in \mathbb{R}_+ \;\middle|\; \mathbb{E}\left[\frac{\min(\boldsymbol{d}^i, q)}{\boldsymbol{d}^i}\right] \geqslant \beta \right\} \tag{6.5}$$

where $\boldsymbol{d}^i = \sum_{t=1}^T \boldsymbol{d}_t^i$ is the demand of item i aggregated over time. Since $\boldsymbol{d}^i$ is non-negative, $q^i(\beta)$ is well-defined (we set $\frac{0}{0} = \beta$ so that items with no demand would not impact the constraint). Computing an approximate value of $q^i(\beta)$ at an arbitrary precision can easily be performed by binary search.

To justify this choice, consider the second problem (6.3) with only one item and for a horizon of one period. Assuming no initial inventory, it takes then the form of the famous *newsvendor problem* (see e.g., Shapiro et al. (2009, Chapter 1))

$$\min_{q \geqslant 0} \quad \mathbb{E}\left[h^i (q - \boldsymbol{d}^i)^+ + \gamma^i (\boldsymbol{d}^i - q)^+\right], \tag{6.6}$$

where γ^i is a unit backorder cost specific to item i. The next proposition means that with the right choice for γ^i, the formulation Equation (6.2) (which takes the form of Problem (6.5) since $h^i > 0$) is equivalent to the second one (6.6). Of course, it holds only for the case with one item and a horizon of one period.

Proposition 19. *Define γ^i as in* (6.4)*. Then $q^i(\beta)$ is the smallest optimal solution to* (6.6).

Proof. The aggregated production problem (6.6) is known to have the optimal solution

$$q^{i*} = F^{-1}_{\boldsymbol{d}^i}\left(\frac{\gamma^i}{\gamma^i + h^i}\right) \tag{6.7}$$

where $F^{-1}_{\boldsymbol{d}^i}$ is the left-inverse of the cumulative distribution function of $\boldsymbol{d}^i$, i.e., , $F^{-1}_{\boldsymbol{d}^i}(\kappa) = \inf\{q \mid \mathbb{P}(\boldsymbol{d}^i \leqslant q) \geqslant \kappa\}$. Since we have set $\gamma^i = \frac{\mathbb{P}[\boldsymbol{d}^i \leqslant q^i(\beta)]}{\mathbb{P}[\boldsymbol{d}^i > q^i(\beta)]}$, we have $q^{i*} = \inf\{q \mid \mathbb{P}(\boldsymbol{d}^i \leqslant q) \geqslant \mathbb{P}(\boldsymbol{d}^i \leqslant q^i(\beta)\}$, which implies that $q^{i*} \leqslant q^i(\beta)$.

Now if this inequality were strict, then it would mean that $\mathbb{P}(\boldsymbol{d}^i \in (q^{i*}, q^i(\beta))) = 0$, which contradicts the minimality assumption in the definition of $q^i(\beta)$ (Equation (6.5)). □

Note that this formulation does not take into account the capacity constraint. It is therefore probably better suited to overcapacited production lines.

Remark 2. If instead of controlling the *fill rate service level*, we want to control the *cycle service level*, defined as the probability of satisfying the whole demand, then we can choose

$$\gamma^i = \frac{\beta}{1-\beta} h^i. \tag{6.8}$$

Indeed, in this case, the optimal solution q^{i*} of (6.6) satisfies $\mathbb{P}(q^{i*} \geqslant \boldsymbol{d}^i) = \beta$. This corresponds to the formula given by van Houtum and Zijm (2000). Interestingly, Equation (6.8) does not depend on the distribution of the demand, which contrasts with Equation (6.4).

6.4 Solving method and theoretical results

6.4.1 Solving method

As explained in Section 5.4, the deterministic version of CLSP-BS is hard. Therefore, we cannot expect a quick algorithm solving exactly the problem, and this holds especially for the full stochastic version. Indeed, consider than $|\mathcal{I}| = 10$, $T = 10$ and that at each period t, the demand d_t^i of item i can take 2 different values. Then, we get $2^{|\mathcal{I}|T} \simeq 10^{30}$ possible realizations of the demand, which gives as many new indices. Thus, a frontal solving is not tractable.

Moreover, one of the requests of the partner was to have an easy to understand method, which can be used and maintained in practice, with short computation times. We propose a two-stage approximation consisting in replacing the measurability constraint (6.3i) by

$$\begin{cases} \boldsymbol{q}_t^i \text{ is } \sigma\big((\boldsymbol{d}_1^i,\ldots,\boldsymbol{d}_{t-1}^i)_{i\in\mathcal{I}}\big)\text{–measurable} & \forall i \in \mathcal{I} \\ \boldsymbol{q}_{t'}^i \text{ is } \sigma\big((\boldsymbol{d}_1^i,\ldots,\boldsymbol{d}_T^i)_{i\in\mathcal{I}}\big)\text{–measurable} & \forall t' \geqslant t+1, \forall i \in \mathcal{I}, \end{cases} \tag{6.9}$$

which provides a relaxation of the initial program: the production decisions for the current period t can still not depend on the future, but now the subsequent production decisions depend on the future demand. We denote this relaxation by (2SA).

This approximation is a *two-stage approximation* as we distinguish between two levels of information over the uncertainty: production decisions for the first period are the *first stage* variables, while all other decisions are *second stage* variables. Three-stages or more generally multistage approximation would give better approximations of stochastic Uniform CLSP-BS but increases exponentially the number of variables. We chose for practicability reasons to stick to the two-stage approximation.

The (2SA) relaxation is then solved by a classical *sample average approximation* (see Kleywegt et al. (2002) for a presentation of the method). We build a set Ω of m scenarios sampled uniformly at random. Each of these scenarios is a possible realization of $(\boldsymbol{d}_t^i, \boldsymbol{d}_{t+1}^i, \ldots, \boldsymbol{d}_T^i)$ for each item i. The parameter m is fixed prior to the resolution.

Figure 6.1 gives a scheme of the two steps involved in the reduction of the scenario space. The complete multi-stage problem is represented by a scenario tree. In this case, even if the realization 1 occurs, there remain random events which can take the values 1.1 or 1.2. The two-stage approximation consists in removing these remaining random events and considering that, if the realization 1 occurs, the remaining events are deterministic. The sampling consists in choosing a finite number of first stage realizations (two in Figure 6.1).

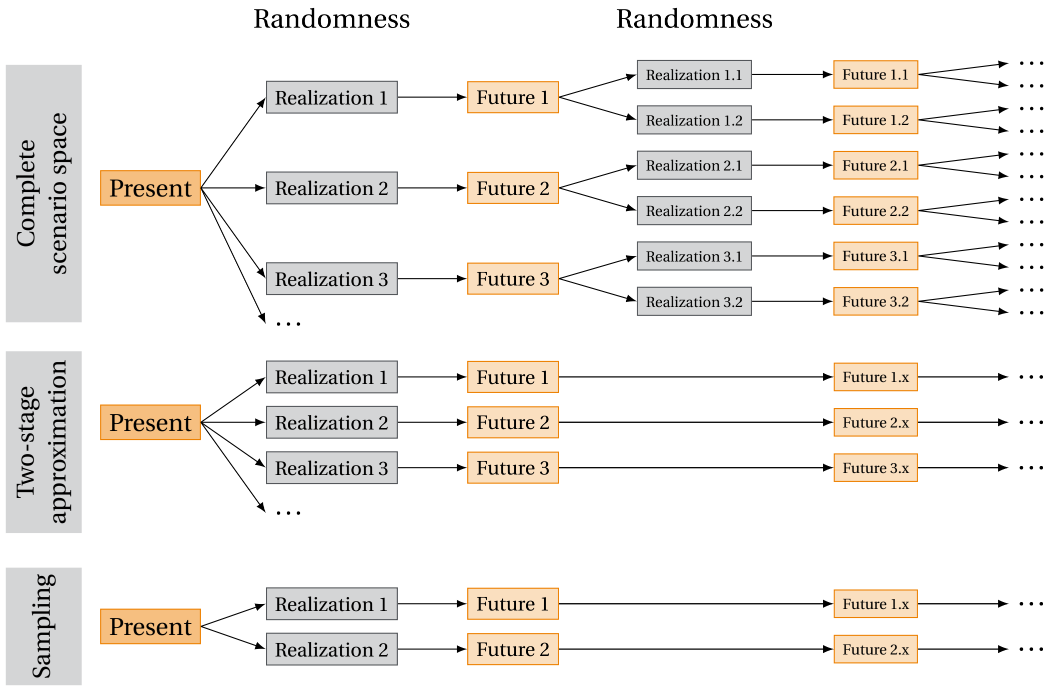


Figure 6.1 – Scheme of the reduction of the scenario space

We get the following mixed integer program (2SA-m), solved by any standard MIP solver.

$$\min \quad \frac{1}{m}\sum_{\omega\in\Omega}\sum_{t'=t}^{T}\sum_{i\in\mathcal{I}}\left(h^i s^i_{t',\omega}+\gamma^i b^i_{t',\omega}\right) \tag{6.10a}$$

$$\text{s.t.} \quad \tilde{s}^i_{t',\omega}=\tilde{s}^i_{t'-1,\omega}+q^i_{t',\omega}-d^i_{t',\omega} \qquad \forall\omega\in\Omega,\ \forall t'\in\{t,\dots,T\},\ \forall i\in\mathcal{I}, \tag{6.10b}$$

$$\sum_{i\in\mathcal{I}} q^i_{t',\omega}\leqslant 1 \qquad \forall\omega\in\Omega,\ \forall t'\in\{t,\dots,T\}, \tag{6.10c}$$

$$q^i_{t',\omega}\leqslant x^i_{t',\omega} \qquad \forall\omega\in\Omega,\ \forall t'\in\{t,\dots,T\},\ \forall i\in\mathcal{I}, \tag{6.10d}$$

$$\sum_{i\in\mathcal{I}} x^i_{t',\omega}\leqslant N \qquad \forall\omega\in\Omega,\ \forall t'\in\{t,\dots,T\}, \tag{6.10e}$$

$$\tilde{s}^i_{t',\omega}=s^i_{t',\omega}-b^i_{t',\omega} \qquad \forall\omega\in\Omega,\ \forall t'\in\{t,\dots,T\},\ \forall i\in\mathcal{I}, \tag{6.10f}$$

$$x^i_{t,\omega}=x^i_t \qquad \forall\omega\in\Omega,\ \forall i\in\mathcal{I}, \tag{6.10g}$$

$$q^i_{t,\omega}=q^i_t \qquad \forall\omega\in\Omega,\ \forall i\in\mathcal{I}, \tag{6.10h}$$

$$x^i_t,\ x^i_{t',\omega}\in\{0,1\} \qquad \forall\omega\in\Omega,\ \forall t'\in\{t,\dots,T\},\ \forall i\in\mathcal{I}, \tag{6.10i}$$

$$q^i_t,\ q^i_{t',\omega},\ s^i_{t',\omega},\ b^i_{t',\omega}\geqslant 0 \qquad \forall\omega\in\Omega,\ \forall t'\in\{t,\dots,T\},\ \forall i\in\mathcal{I}, \tag{6.10j}$$

At period t, the production is then set to be the solution $(q^i_t)_{i\in\mathcal{I}}$ found by the solver.

The validity of this method for solving (2SA) is supported by the following proposition.

Proposition 20. *The following three properties hold when m goes to infinity.*

(i) *The optimal value of* (2SA-m) *converges almost surely to the optimal value of* (2SA).

(ii) *For every m, we consider the values $(\hat{q}^i_{t,m}, \hat{x}^i_{t,m})_{i\in\mathcal{I}}$ of the decision variables for period t of an optimal solution of* (2SA-m). *Any limit point of these values is an optimal solution of* (2SA).

(iii) *Let $\varepsilon > \delta > 0$. Assume that the random demand $(\boldsymbol{d}^i_{t'})_{t'\geqslant t, i\in\mathcal{I}}$ is such that*

$$\exists C, K, \quad \forall u \in \mathbb{R}, \qquad \mathbb{E}[e^{u\|\boldsymbol{d}\|}] \leqslant Ce^{u^2K}. \tag{6.11}$$

Denote by $\mathcal{Q}^\delta_m$ (resp. $\mathcal{Q}^\varepsilon$) the set of all possible values of $(\hat{q}^i_{t,m})_{i\in\mathcal{I}}$ in a δ-optimal solution of (2SA-m) *(resp. in an ε-optimal solution of* (2SA)*). Then for every $\alpha \in (0,1)$, we have $\mathbb{P}\big[\mathcal{Q}^\delta_m \subseteq \mathcal{Q}^\varepsilon\big] > 1-\alpha$ for m large enough.*

Item (i) is a result on the convergence of the optimal values of (2SA-m) toward the optimal value of (2SA).

Item (ii) is a result on the convergence of the optimal solutions of (2SA-m) toward an optimal solution of (2SA). However, it requires the existence of the limits of the optimal solutions of (2SA-m).

Item (iii) is a result on the convergence of the set of δ-optimal solutions of (2SA-m) toward the set of ε-optimal solutions of (2SA). Equation (6.11) is a technical condition. It just means than distributions are not heavy-tailed. If the random demand $\boldsymbol{d}$ is bounded or positive part of Gaussian then it satisfies (6.11). Conversely, exponential distributions does not satisfied (6.11).

The proof of Item (iii) relies on the following technical lemma.

Lemma 21. *Consider $g(d) = \inf_{y\in Y} G(y,d)$, where Y is non-empty and where the function G is non-negative and κ-Lipschitz with respect to d. If the random variable $\boldsymbol{d}$ satisfies* (6.11)*, then $g(\boldsymbol{d})$ also satisfies* (6.11) *with $C' = \max\{1, e^{G(y_0,0)}C\}$ and $K' = K\kappa + G(y_0,0)$, for $y_0 \in Y$.*

Proof. For $u \geqslant 0$ we take $y_0 \in Y$ yielding $g(\boldsymbol{d}) \leqslant G(y_0,0) + \kappa\|\boldsymbol{d}\|$. We then have $\mathbb{E}[e^{ug(\boldsymbol{d})}] \leqslant \mathbb{E}[e^{uG(y_0,0)}e^{\kappa u\|\boldsymbol{d}\|}] \leqslant C'e^{K'u^2}$. For $u \leqslant 0$, by non-negativity of G we have $g(\boldsymbol{d}) \geqslant 0$, hence $\mathbb{E}[e^{ug(\boldsymbol{d})}] \leqslant 1$. □

Proof of Proposition 20. Let Q_t be the (bounded) set of feasible values for the first-stage variables $q_t = (q^i_t)_{i\in\mathcal{I}}$ in (2SA). Denote by $F(q_t, d)$ the minimal cost of (2SA) that can be reached when the first stage-variables are fixed to $q_t \in Q_t$ and the realization of the demand is $d = (d^i_{t'})_{t'\geqslant t, i\in\mathcal{I}}$. We introduce the map $f\colon q_t \mapsto \mathbb{E}\big[F(q_t,\boldsymbol{d})\big]$, which associates to a given choice of $q_t \in Q_t$ the expected minimal cost, and similarly the map $\hat{f}_m$, which associates to a $q_t \in Q_t$ the minimal cost of (2SA-m) when the first-stage variables are set to this q_t.

The map f is continuous and Q_t is compact. Thus f is bounded, and, by Shapiro et al. (2009, Theorem 7.48), we have that $(\hat{f}_m(q_t))_{m\in\mathbb{Z}_+}$ converges to $f(q_t)$ uniformly on Q_t. Then Item (i) and Item (ii) are direct consequences of Shapiro et al. (2009, Theorem 5.3).

By Lemma Lemma 21, there exist K and C such that for any $q_t \in Q_t$, $F(q_t,\boldsymbol{d})$ satisfy (6.11). Consequently there exists $\sigma > 0$ such that for all $q_t, q'_t \in Q_t$, the random variable $[F(q_t,\boldsymbol{d}) - f(q_t)] - [F(q'_t,\boldsymbol{d}) - f(q'_t)]$ is σ-subgaussian, (see e.g., Vershynin (2010)). Furthermore, for any demand d, the map $F(\cdot, d)$ is Lipschitz-continuous on Q_t. Then, according to Shapiro et al. (2009, Theorem 5.18), for every $\alpha \in (0,1)$, there exists $M \in \mathbb{Z}_+$ such that for $m \geqslant M$, we have $\mathbb{P}(\mathcal{Q}_m^\delta \subseteq \mathcal{Q}^\varepsilon) > 1 - \alpha$. □

6.4.2 Bounds

The following proposition proposes bounds on formulation Equation (6.3) in order to evaluate the quality of a solution. It relies on classical stochastic results (see for example Shapiro et al. (2009)).

Proposition 22. *For each $m \in \mathbb{Z}_+^*$, we have the following inequalities on the optimal values:*

$$\mathbb{E}\left[v_{2SA-m}^*\right] \leqslant \mathbb{E}\left[v_{2SA-(m+1)}^*\right] \leqslant v_{2SA}^* \leqslant v^*, \tag{6.12}$$

where v_{2SA-m}^* is the optimal value of (2SA-m), v_{2SA}^* is the optimal value of (2SA) and v^* is the optimal value of formulation Equation (6.3).

Proof. The two-stage approximation (2SA) being a relaxation of the original problem, we get $v_{2SA}^* \leqslant v^*$. The other inequalities are proved in Shapiro et al. (2009, Proposition 5.6). □

6.5 Discussion about modeling

In the way we formulate the stochastic counterpart of Uniform CLSP-BS, we made some choices. The first one is the information structure i.e., the time when randomness is revealed. Relying on the denominations used by Carpentier et al. (2015), there can in principle be two versions for each period: decision-hazard where decisions are made before the information is revealed and hazard-decision which is the opposite. Both make sense in industrial applications but we choose the decision-hazard version, which suppose that the demand of period t is revealed at the end of the period after the production decisions were made. This version is more pessimistic than the hazard-decision version. However, the methods developed in this thesis can be easily adapted to the hazard-decision case where the demand is revealed at the beginning of the period.

The second one concerns the feasibility of the planning. In Argon's applications, covering every possible realizations of the demand often leads to too expensive production planning or to infeasible planning. For example, if the demand is the positive part of a Gaussian distribution, then the set of feasible solutions is empty because of the possible high demand realizations. When it occurs, a part of the demand is not satisfied and the model must return a production planning satisfying every other constraints and reaching an objective service level.

Thus we have to model the undelivered quantities. In many cases, serving only 95% of the demand is less critical than serving only 80%. Thus, Argon aims at reaching a *service level* for delivered quantities. The fill rate service level, which is the proportion of demand satisfied on time, is considered. As previously explained in Chapter 2, the cycle service level, which is the proportion of command entirely satisfied, could have also been considered but in most of our application, it is less relevant. When used in practice, the decision-maker can choose the service level requirement depending on the criticality of the item. (100% service level is the robust case.)

Another non-obvious question is the mesh considered for the service level constraint. It can be extremely focused as a service level on each pair (item, period) or extremely general as a global service level for all items and all periods simultaneously. The aggregation of the KPI's can also be done in many ways (e.g., average of the service level over items, sum of the satisfied demand over all items and all periods divided by sum of demand over all items and all periods). We were not able to get a precise formulation from our partner and from its clients since it depends on the industry or the strategic objectives. However, the methods developed in this thesis can be easily adapted to the different modeling of the fill rate service level. We choose to model the fill rate service level taking the average of the service level of each item. Each service service level is weighted by the holding cost of the item.

7 Numerical experiments

In this chapter, we test the practical efficiency of our approach. We will simulate realizations of the demand and planning decisions, updated week after week with the help of various heuristics.

7.1 Simulation

In this chapter, each method used to compute a production planning is called a *heuristic* since it cannot guarantee to reach the optimum. Moreover, we are not able to prove that one heuristic is theoretically better than another. Thus, in order to compare their practical efficiency, we use simulation.

A simulation requires:

- the characteristics of the items (holding costs h^i, initial inventory) and the characteristics of the assembly line (capacity C and upper bound N on the number of setups),
- the forecast demand (value and reliability). This forecast demand may depend on time and on past realizations of the demand,
- the choice of the heuristic used to compute the production planning,
- the number n of runs.

A run of the simulation starts at period $t = 1$. Then, for each period t, we observe the current inventory level of each item, i.e., the inventory at the end of the previous period. Using the forecast demand function and the chosen heuristic, we compute the production planning and fix the decision for the current period t. At the end of the period, we observe the demand outcome for period t and update the inventory level. At the end of a run, we get several *Key Performance Indicators (KPI)* measuring the efficiency of the heuristic for a particular realization of the demand. A scheme of a typical run of the simulation is given in Figure 7.1.

After the n runs of the simulation, we compute other KPI's from the KPI of all runs like the average over all runs or the standard deviation.

C++11 has been chosen for the implementations and Gurobi 6.5.1 (see Gurobi Optimization, LLC (2018)) was used to solve the model on a PC with 8 processor Intel® Xeon™ E5-2667 @ 3.30GHz and 48Go RAM.

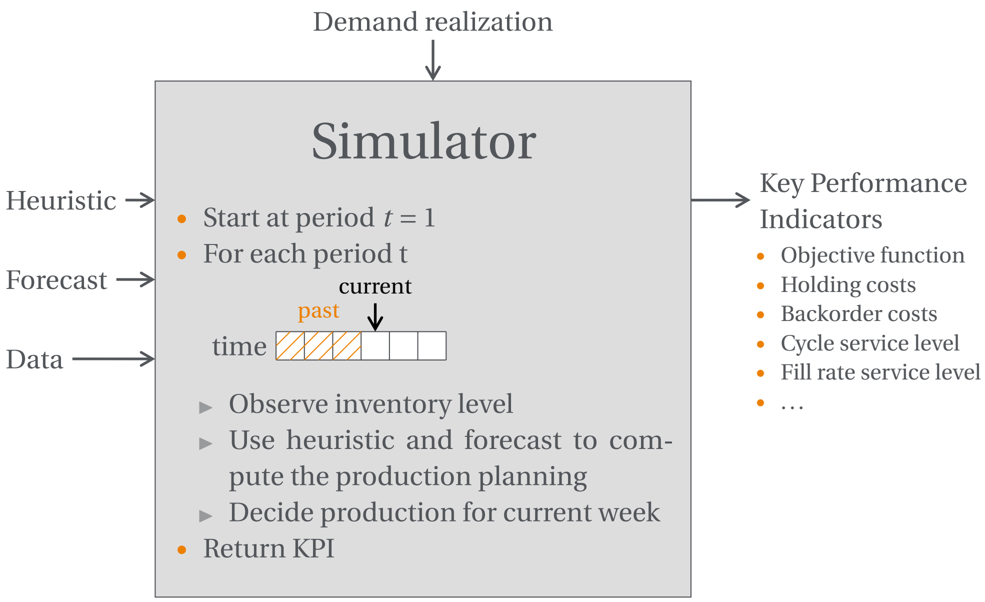


Figure 7.1 – Scheme of the run of the simulation

7.2 Heuristics

Our heuristic consists drawing m new scenarios of demand knowing the past realizations of the demand. Then we solve (2SA-m), as defined in Problem (6.10), at the beginning of each period t with a MIP solver. We compare it with three other heuristics:

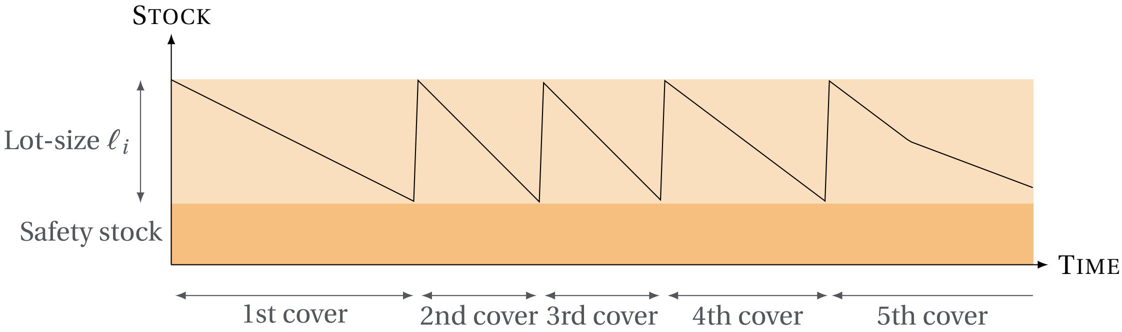
- deterministic approximation,
- lot-size,
- cover-size.

The first heuristic is the deterministic version of Problem (6.3), where the random demand is replaced by its expectation. Note that it is not the CLSP-BS described in Chapter 5 since there are backorder costs. When backorder costs are not given by the client, we used the method described in Section 6.3.2 to compute them.

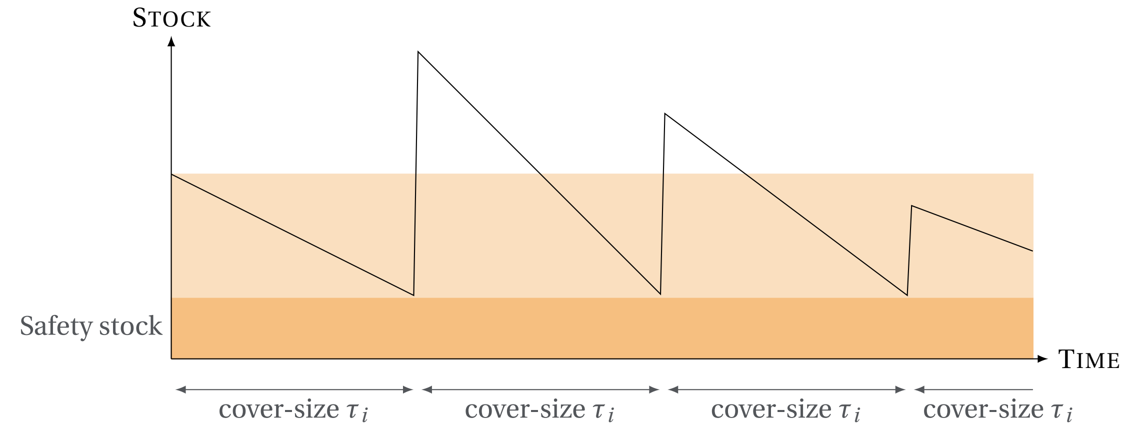
The second one, the *lot-size heuristic*, consists in determining before the first week once and for all a value ℓ_i^* for each item $i \in \mathcal{I}$. At time t, if the inventory of item i is below a precomputed safety stock (see Equation (7.1)), the quantity q_t^i is chosen so that the inventory of item i exceeds the safety stock of exactly ℓ_i^*. In case of capacity issues, the production is postponed and thus backorder costs appear. In addition, if some capacity issues are easily anticipated, the production of an item i can be activated even if the inventory is not below the safety stock.

The third one, the *cover-size heuristic*, is almost the same, but instead of precomputing a fixed quantity for each item, a duration τ_i^* is fixed before the first week. When the inventory of item i is below the safety stock, the quantity q_t^i is computed so that the inventory of item i exceeds the safety stock of the expected demand for the next τ_i^* weeks. An example is given on Figure 7.2b.

The values ℓ_i^* and τ_i^* are determined using results of Chapter 3. $(\tau_i^*)_{i \in \mathcal{I}}$ is actually chosen to be the optimal solution to Program (3.4), which somehow considers the problem at a “macroscopic”



(a) lot-size



(b) Cover-size

Figure 7.2 – Computation of produced quantities using lot-size and cover-size heuristics

level where the demand d^i is chosen equal to $\mathbb{E}\left[\sum_{t=1}^{T} \boldsymbol{d}_t^i\right]$.

For each item i, the parameter ℓ_i^* of the lot-size heuristic is then set to $d^i \tau_i^*$.

There is no universal formula for the safety stocks $\left(s_{\min}^i\right)_i$. Thus we turn to the formula used in practice by most of Argon Consulting's clients (see Arnold et al. (2007, Chapter 11)) where the underlying model assumes that the demand has a Gaussian distribution. Production being decided at the beginning of the period and the demand being revealed at the end of the period, the lead time is one period (and is certain). Thus, for each item i, the safety stock used to reach cycle service level α is

$$s_{\min}^i = z_\alpha \sqrt{\mathrm{Var}\left[\boldsymbol{d}^i\right]} \tag{7.1}$$

where $\boldsymbol{d}^i = \sum_{t=1}^{T} \boldsymbol{d}_t^i$ and z_α is the inverse distribution function of a standard normal distribution with cumulative probability α. When fill rate service level β is used, an abacus gives the coefficient k_β which must be used instead of z_α.

The cover-size heuristic adapts the production to the realization of the demand, contrary to the lot-size heuristic. According to Argon Consulting, it makes the cover-size heuristic more suitable for situations with low short term volatility of demand or for overcapacitated lines, while the lot-size heuristic is expected to behave better with high short term volatility of demand or for undercapacitated lines. Lot-sizes and cover-sizes heuristics are used in practice.

Note that the backorder costs are not taken into account at all for determining the values of the parameters ℓ_i^* and τ_i^*. But playing with safety stocks allows to prevent too large backorder costs. However, in real life it is usually the other way round: the company does not associate costs to backorder and plays with the safety stock to address directly the service level.

7.3 Instances and probabilistic model

7.3.1 Datasets

The instances used are realistic and have been provided by a client of Argon Consulting. The client gave the data of seven assembly lines and the demand for each week over a full quarter. The horizon T is the typical one used in practice by this client, namely $T = 13$ weeks.

The historical demands are denoted by $\bar{d}_t^i$.

The initial inventory is set to $s_0^i = \frac{1}{3}(\bar{d}_1^i + \bar{d}_2^i + \bar{d}_3^i)$.

The other parameters are provided in Table 7.1. The parameter C is the capacity of the line before the normalization leading to formulation (5.3). (Recall that the problem and the model have been formulated in Section 6.3 after normalization.) In the column $\tilde{h}^i$, we indicate the range of the holding costs before normalization. We obtain the h^i's by dividing these costs by C since internal production times v^i are unitary.

We also add the loading characteristic κ_t at period t, only given as an indication of the hardness of the instance, and defined as follows.

$$\kappa_t = \frac{\sum_{i\in\mathcal{I}}\left(\sum_{t'=1}^{t}\bar{d}_{t'}^i\right) - s_0^i}{\sum_{t'=1}^{t} C_{t'}} \tag{7.2}$$

It is the ratio of cumulative forecast demand up to period t minus the initial inventory over cumulative capacity up to period t. If there were no flexibility constraints and if the holding costs were not an issue, then for a period t, the inequality $\kappa_t \leqslant 100\%$ would imply that it is possible to supply the whole demand. As shown in Table 7.1, the lines L_0, L_1, L_2, L_3 and L_4 experience overcapacity: the loading indicator is smaller than 100%. The lines L_5 and L_6 experience undercapacity: the loading indicator is larger than 100% at some periods.

The number m of scenarios used to solve (2SA-m) is fixed to 20, determined by preliminary experiments showing that it is a good trade-off between accuracy and tractability. The time limit of the solver has been set to 120 seconds.

7.3.2 Demand distribution

For the purpose of notation, in this section, we use index $k \in [K]$ to represent a pair $(t,i) \in [T] \times \mathcal{I}$. We choose to model the distribution of demand $\left(\tilde{\boldsymbol{d}}_1,\ldots,\tilde{\boldsymbol{d}}_K\right)$ using the “expanded Dirichlet” distribution defined as follow.

Let $\gamma \in (0,1)$ and let $\tilde{d}_1,\ldots,\tilde{d}_K$ be K positive real numbers. We denote by $\mathcal{D}\left(\gamma,\tilde{d}_1,\ldots,\tilde{d}_K\right)$ the

Instances	Instance characteristics						
	$\lvert\mathcal{I}\rvert$	$\max(\tilde{d}_t^i)$	C	N	$\tilde{h}^i$	$\max(\kappa_t)$	κ_T
L_0	21	3780	8518	7	45–88	91%	74%
L_1	30	4122	13326	12	52–82	66%	52%
L_2	21	4992	10562	7	35–61	61%	61%
L_3	13	6220	10394	5	22–30	80%	65%
L_4	18	10584	62164	8	12–14	40%	35%
L_5	12	11772	7902	6	15–17	126%	98%
L_6	22	8640	13299	8	16–23	118%	98%

Table 7.1 – Instance characteristics

probability distribution of $\left(\tilde{\boldsymbol{d}}_1,\dots,\tilde{\boldsymbol{d}}_K\right)$ on $\mathbb{R}_+^K$ when

$$\begin{gathered}\frac{1}{\tilde{d}_0}\left(\tilde{\boldsymbol{d}}_1,\dots,\tilde{\boldsymbol{d}}_K\right) \sim \mathrm{Dir}\left(\alpha_1,\dots,\alpha_K\right)\\ \text{with} \quad \tilde{d}_0 = \textstyle\sum_{k=1}^K \tilde{d}_k, \quad \alpha_0 = \frac{1}{\gamma^2} - 1, \quad \text{and} \quad \alpha_k = \frac{\tilde{d}_k}{\tilde{d}_0}\alpha_0,\ \forall k \in [K].\end{gathered} \tag{7.3}$$

(We remind that $\mathrm{Dir}\left(\alpha_1,\dots,\alpha_K\right)$ designates the Dirichlet distribution and recall its definition and the main results in Appendix A.1.) We call it the "expanded Dirichlet" distribution.

The "expanded Dirichlet" distribution has the following properties.

Proposition 23. *Let γ be a real number in $(0,1)$ and let $\tilde{d}_1,\dots,\tilde{d}_K$ be K positive number (with $K \geqslant 2$). Let $\left(\boldsymbol{d}_1,\dots,\boldsymbol{d}_K\right)$ be a random vector following an "expanded Dirichlet" distribution $\mathcal{D}\left(\gamma,\tilde{d}_1,\dots,\tilde{d}_K\right)$. Then, the following properties hold.*

1. *Expectation*
$$\mathbb{E}\left[\boldsymbol{d}_k\right] = \tilde{d}_k \qquad \forall k \in [K]. \tag{7.4}$$
2. *Variance*
$$\mathrm{Var}\left[\boldsymbol{d}_k\right] = \gamma\, \tilde{d}_k\left(\tilde{d}_0 - \tilde{d}_k\right) \qquad \forall k \in [K]. \tag{7.5}$$
3. *Sum*
$$\sum_{k=1}^K \boldsymbol{d}_k = \tilde{d}_0 \qquad \textit{almost surely.} \tag{7.6}$$
4. *Let k be an integer in $\{1,\dots,K-1\}$. We denote by $\boldsymbol{d}_{(1)} = \left(\boldsymbol{d}_1,\dots,\boldsymbol{d}_k\right)$, $\boldsymbol{d}_{(2)} = \left(\boldsymbol{d}_{k+1},\dots,\boldsymbol{d}_K\right)$ and $d_{(1)} = (d_1,\dots,d_k)$ a fixed realization of $\boldsymbol{d}_{(1)}$. Then, $\left(\boldsymbol{d}_{(2)}\middle|\boldsymbol{d}_{(1)} = d_{(1)}\right)$ has an "expanded Dirichlet" distribution $\mathcal{D}\left(\gamma',\tilde{d}_{k+1},\dots,\tilde{d}_K\right)$ where γ' is the unique positive solution of*
$$\frac{1}{\gamma'^2} - 1 = \left(\frac{1}{\gamma^2} - 1\right)\left(1 - \frac{\sum_{i=1}^k d_i}{\tilde{d}_0}\right). \tag{7.7}$$

Proposition 23 is proved Appendix A.2.

The parameters of the "expanded Dirichlet" distribution used to generate the demand are the historical forecast demands $\tilde{d}_t^i$ for the $\tilde{d}_k$ and γ is set equal to

$$\gamma = v \frac{\sum_{k=1}^{K} \sqrt{\frac{\tilde{d}_0}{\tilde{d}_k} - 1}}{\sum_{k=1}^{K} \left(\frac{\tilde{d}_0}{\tilde{d}_k} - 1 \right)}. \tag{7.8}$$

The parameter γ is defined from another parameter v which we call *volatility* and whose interpretation is easier. Volatility gives the ratio between standard deviation and expectation of a random variable. Defining γ as in Equation (7.8) enables the volatility of each demand to be as close to v as possible and is justified in Appendix A.3.

Using "expanded Dirichlet" distributions has many advantages. First, Item 1 ensures that demand expectation is equal to the historical forecast demand and Item 3 ensures that, as explained in Chapter 2, total volume $\sum_{t,i} \boldsymbol{d}_t^i$ of the forecast demand is constant over every possible outcomes. Second, Item 4 ensures that the demand distribution keeps its properties when part of the demand is revealed.

Moreover, random variables with an "expanded Dirichlet" distribution can easily be generated using Algorithm 1 since most of modern programming languages have a generator of gamma distributed variables.

Algorithm 1: Generator of "expanded Dirichlet" distribution $\mathcal{D}(\gamma, \tilde{d}_1, \ldots, \tilde{d}_K)$

Input: $\gamma, \tilde{d}_1, \ldots, \tilde{d}_K$
Output: Random realization $(d_1, \ldots, d_K)$ of $\mathcal{D}(\gamma, \tilde{d}_1, \ldots, \tilde{d}_K)$

$\tilde{d}_0 := \sum_{k=1}^{K} \tilde{d}_k$

$\alpha_0 := \frac{1}{\gamma^2} - 1$

for $k = 1, \ldots, K$ **do** $\alpha_k := \frac{\tilde{d}_k}{\tilde{d}_0} \alpha_0$

for $k = 1, \ldots, K$ **do**
 | draw a number y_k from Gamma distribution $\Gamma(\alpha_k, 1)$
end

for $k = 1, \ldots, K$ **do** $d_k := \frac{y_k}{\sum_{\ell=1}^{K} y_\ell} \tilde{d}_0$

Return $(d_1, \ldots, d_K)$

Proposition 24. *For any* $\gamma \in (0,1)$ *and any* $\tilde{d}_1, \ldots, \tilde{d}_K$ *positive real numbers, Algorithm 1 samples the "expanded Dirichlet" distribution* $\mathcal{D}(\gamma, \tilde{d}_1, \ldots, \tilde{d}_K)$.

Proposition 24 is proved Appendix A.2.

Finally, Algorithm 1 and Item 4 of Proposition 23 enable to generate easily a realization of the demand when part of the demand is already revealed.

7.4 Use of K-means algorithm

As explained in Section 7.3.1, we use a small number of scenarios to represent every possible outcomes. Thus, using only Algorithm 1 gives results which may be very dependent of the generating scenarios. In this section, we propose to use a K-means algorithm on a bigger set of scenarios to get a small set of representative scenarios.

We refer to Hastie et al. (2009, Chapter 14) for a detailed presentation of the method. Roughly, K-means clustering aims at partitioning n observations into K clusters in which each observation belongs to the cluster with the nearest mean. The mean of each cluster can be used as a prototype (i.e., a representative element) of the cluster.

K-means algorithm is easy to implement but its simplest version have some well-known limitations. First, it uses Euclidean distance as a metric and variance is used to measure cluster scatter. In some cases, other distances might be more relevant and choosing a more appropriate one may not be an easy task. Second, the number of clusters K is an input parameter. Thus, an inappropriate choice of K may yield poor results. Last, results depend of the initialization of the first K prototypes and may lead algorithm to converge to local optimum. Some variations of K-means algorithms exist to deal with these points (see Jain (2010) for a review of K-means algorithm and its extensions).

Due to time constraints, we have not implemented K-means algorithm. However, a straightforward way to use it would consists in:

- using Algorithm 1 to generate a sample of M scenarios which can be seen as real vectors of $\mathbb{R}^{T\times\mathcal{I}}$,
- initializing K with $m = 20$,
- using K-means algorithm to classify the M scenarios into m clusters.

Then, our set Ω of representative scenarios used in Equation (6.10) would be composed of the prototype of each cluster, with a probability equal to the number of scenarios in its cluster divided by M.

7.5 Numerical results

We run tests on lines described in Table 7.1. Complete tables of results are given in Appendix B and the results for line L_1 is given in Table 7.2. (We choose line L_1 because it does not suffer from undercapacity like L_5 and L_6, and it is not a line that is clearly overcapacitated like L_4.)

The table is composed as follows. In the column *Input*, we give the characteristic of the inputs. The sub-column *Vol.* gives the volatility of the demand as defined in Section 7.3.2. The sub-column *Heur.* gives the heuristic used to make production decisions (see Section 7.2). The abbreviation *det* stands for the deterministic heuristic, *sto* for (2SA-m) heuristic, *cover* for cover-size and *lot* for lot-size. Finally, the sub-column β gives the service level used to compute backorder costs as explained in Section 6.3.2.

The outputs are composed of the following KPI's. They are all averaged over the 30 runs of the simulation. In the column *Obj.*, we write the objective value (6.10a) of the solution found by the heuristic (in k€).

In the column *LB*, we write the best lower bound on the objective value (6.10a) (in k€). This lower bound is computed by solving Program (6.10) at time $t = 1$ with $m = 5000$ scenarios and a computing time limit equal to 24 hours.

The column *Inventory* gives several KPI relative to stocks. The sub-column *Av.* gives the average of the positive part of the inventory over periods, i.e.,

$$\langle s \rangle_T = \frac{1}{T} \sum_{t=1}^{T} \sum_{i \in \mathcal{I}} \left(h^i s_t^i \right)^+. \tag{7.9}$$

The sub-column *Max* gives the maximum of the positive part of the inventory over periods, i.e.,

$$\max_T (s) = \max_{t \in [T]} \left(\sum_{i \in \mathcal{I}} \left(h^i s_t^i \right)^+ \right). \tag{7.10}$$

The sub-column *Cover* gives the average cover-size of items, i.e., the average number of periods covered when a production occurrs. This KPI is taken on average over items and each item is weighted by its holding cost. For each item i, denoting by $\mathcal{T}^i$ the subset of $[T]$ of periods where item i is produced, the KPI is computed by

$$\langle \tau \rangle_{\mathcal{I}} = \sum_{i \in \mathcal{I}} \frac{w^i}{|\mathcal{T}^i|} \sum_{t \in \mathcal{T}^i} \frac{q_t^i}{\langle \bar{d}^i \rangle_T} \quad \text{where} \quad w^i = \frac{h^i}{\sum_{j \in \mathcal{I}} h^j} \quad \text{and} \quad \langle \bar{d}^i \rangle_T = \frac{1}{T} \sum_{t=1}^{T} \bar{d}_t^i. \tag{7.11}$$

The column *Service* gives the KPI relative to service level with the two classical measure. The sub-column *Fill rate* gives the average fill rate service level of items. This KPI is taken on average over items and each item is weighted by its holding cost. It is given by

$$\langle \beta \rangle_{\mathcal{I}} = \sum_{i \in \mathcal{I}} w^i \frac{\sum_{t=1}^{T} \min\left(\left(s_{t-1}^i \right)^+ + q_t^i,\ d_t^i \right)}{\sum_{t=1}^{T} d_t^i} \quad \text{where} \quad w^i = \frac{h^i}{\sum_{j \in \mathcal{I}} h^j}. \tag{7.12}$$

The sub-column *Cycle* give the average cycle service level of items. This KPI is taken on average over items and each item is weighted by its holding cost. It is given by

$$\langle \alpha \rangle_{\mathcal{I}} = \sum_{i \in \mathcal{I}} w^i \frac{\sum_{t=1}^{T} y_t^i}{T} \quad \text{where} \quad w^i = \frac{h^i}{\sum_{j \in \mathcal{I}} h^j} \quad \text{and} \quad y_t^i = \begin{cases} 1 & \text{if } \left(s_{t-1}^i \right)^+ + q_t^i \geqslant d_t^i \\ 0 & \text{otherwise.} \end{cases} \tag{7.13}$$

Note that these two KPI give priority to satisfy current demand rather than backorder.

In the column *Workload*, we compute the average workload of the assembly line. This KPI is taken on average over periods. It is given by

$$\langle \text{Workload} \rangle_T = \frac{1}{T} \sum_{t=1}^{T} \frac{\sum_{i \in \mathcal{I}} q_t^i}{C}. \tag{7.14}$$

In the column *Flex.*, we compute the average number of setups used on the assembly line. This

KPI is taken on average over periods. It is given by

$$\langle\text{Flexibility}\rangle_T = \frac{1}{T}\sum_{t=1}^{T}\frac{\sum_{i\in\mathcal{I}} x_t^i}{N}. \tag{7.15}$$

Input			Obj.	LB	Inventory			Service		Work-	Flex.
Vol.	Heur.	β			Av.	Max	Cover	Fill rate	Cycle	load	
volatility $\upsilon = 20\%$	lot	85%	12320	495	948	1125	7.6	94%	95%	62%	92%
		95%	14087	504	1080	1353	7.2	94%	95%	64%	94%
		98%	15516	560	1161	1520	7.2	94%	95%	65%	96%
	cover	85%	10719	495	824	1024	4.4	100%	100%	58%	95%
		95%	11833	504	910	1051	4.3	100%	100%	59%	98%
		98%	12397	560	953	1075	4.3	100%	100%	60%	99%
	det	85%	709	495	54	313	2.8	78%	62%	51%	99%
		95%	839	504	54	313	2.8	81%	63%	44%	99%
		98%	1485	560	69	313	3.0	86%	68%	51%	99%
	sto	85%	506	495	39	300	2.6	68%	55%	49%	96%
		95%	725	504	42	300	2.6	77%	59%	43%	97%
		98%	1451	560	62	305	3.0	84%	67%	51%	98%
volatility $\upsilon = 50\%$	lot	85%	12316	518	946	1142	8.9	93%	94%	61%	92%
		95%	14450	634	1082	1353	8.6	94%	95%	64%	94%
		98%	16454	759	1171	1515	8.8	94%	95%	65%	95%
	cover	85%	11083	518	851	1045	4.7	99%	99%	58%	94%
		95%	12403	634	948	1092	4.6	99%	99%	59%	98%
		98%	13148	759	996	1137	4.6	99%	99%	60%	98%
	det	85%	1089	518	79	340	3.1	76%	64%	40%	98%
		95%	2006	634	102	340	3.4	84%	70%	51%	99%
		98%	3179	759	154	346	3.7	90%	78%	52%	100%
	sto	85%	696	518	43	307	2.7	62%	56%	37%	95%
		95%	1869	634	78	318	3.0	79%	66%	51%	98%
		98%	3140	759	151	345	3.5	90%	77%	52%	99%
volatility $\upsilon = 80\%$	lot	85%	12597	629	960	1160	48.1	92%	93%	61%	91%
		95%	15173	1081	1091	1381	46.8	93%	94%	64%	94%
		98%	18567	1330	1196	1556	47.6	93%	94%	65%	95%
	cover	85%	11068	629	848	1054	12.1	99%	97%	58%	94%
		95%	12497	1081	943	1100	12.1	99%	97%	59%	97%
		98%	13525	1330	992	1158	12.1	99%	97%	60%	99%
	det	85%	1743	629	114	360	8.5	77%	69%	48%	98%
		95%	3290	1081	163	361	9.0	84%	75%	52%	99%
		98%	5059	1330	213	365	9.2	88%	80%	53%	99%
	sto	85%	1185	629	55	323	2.9	62%	60%	47%	94%
		95%	3123	1081	126	342	5.9	79%	71%	52%	98%
		98%	4947	1330	220	387	7.3	88%	81%	53%	100%

Table 7.2 – Results for L_1

Table 7.2 and the others presented in Appendix B first show that the best objective values for the problem are always reach by the (2SA-m) heuristic. Then the deterministic heuristic also gives good objective values. On the other hand, the objective values returned by the lot-size and cover-size heuristics are about 10 times greater than the objective values returned by the (2SA-m) heuristic. Regarding the real inventory, we see that it represents the main part of the objective value for each heuristic. Naturally, the average cover-sizes for each item are greater for lot-size and cover-size heuristics than for the deterministic and the (2SA-m) heuristics.

Backorder is strongly linked to fill rate service level. As expected for heuristics with such huge inventory, lot-size and cover-size heuristics have the best fill rate service level. However, it is the same for any value of the tested desired service level. On the other hand, deterministic and (2SA-m) heuristics adapt to the desired service level but they are too optimistic about their ability to satisfied uncertain demand.

By looking at the values of the backorder cost, we can see that the values computed by method proposed in Section 7.3.2 are very small. It is especially visible for small desired service level like $\beta = 85\%$. Thus, the method to compute the backorder cost is too optimistic and must be improved. A simple way might be to skew the method. A more sophisticated one might be to use extensions of the news-vendor problem like the multi-stage extension.

Comparing results between undercapacitated lines (L_0 to L_4) and overcapacitated lines (L_5 and L_6) does not show clear difference on the behaviors of the heuristics. However, we simply see that the GAP between the different KPI is smaller in undercapacitated case than in overcapacitated case.

Finally, we recall that the lot-size and cover-size heuristics are used in practice. However, we do not know every technical detail of the implementation. Thus, the results that can be obtained by the planners may be slightly better that those we have obtained.

Production multi-sourcing Part III

8 Deterministic multi-sourcing

8.1 Introduction

8.1.1 Motivations

Deciding which plants should have the ability to produce a (new) item is a long-term decision that companies have to answer when expanding their product portfolio. Note that several plants may be able to produce the same item that is why this problem is called *multi-sourcing*. Because of their costs and of the time needed to implement them, multi-sourcing decisions cannot be easily changed. Thus, they have a high impact on competitiveness when deciding production planning (mid-term decision) and scheduling (short-term decision). One could think that giving to each plant the ability to produce every item is the safest solution to ensure future high service level. However, multi-sourcing the production of every item leads to unnecessary high assignment costs. The challenge is to find a trade-off between assignment cost and demand satisfaction.

In multi-sourcing problems, we consider a centralized inventory. In practice, this can take several forms. One consists in many plants located in different places but whose productions go to a unique warehouse before being sent to the stores. This is the case for one client of Argon Consulting in the luxury industry. Another one consists in several assembly lines in one big plants. In this cases, it is the lines which are given the ability to produce an item.

Moreover, logistic costs are not considered in the problem. In practice, in the cases we just presented, logistic costs are not a big issue compared to multi-sourcing costs. In more complex cases, deciding the best logistic network relies on many more constraints and due to size of the problem, it has to be optimized in its own process.

Argon Consulting meets multi-sourcing problems with clients from different industries. Thus, Argon Consulting aims at getting a versatile model which can be used in many situations and should be not to too dependent on the industry. We propose a model which captures the essence of the multi-sourcing problem and which can be easily adapted to particular client cases.

8.1.2 Problem statement

We consider a set $\mathcal{P}$ of plants producing a set $\mathcal{I}$ of items over T periods. There is an upper bound C_{pt} on the total period production of plant p at period t (summed over all items). This upper bound is expressed in time unit since it correspond to available working hours.

Giving a plant p the ability to produce an item i has a cost c^i_p. This cost is paid once and for all for the whole horizon. When a plant p is able to produce item i, there is an upper (resp. lower) bound $u^i_{pt} \geqslant 0$ (resp. $\ell^i_{pt} \geqslant 0$) on the production of item i in plant p at period t. The capacity needed (in time units) to produce one unit of item i in plant p in period t is $v^i_{pt} > 0$.

The production and the inventory of item i (summed over all plants) must satisfy a demand d^i_t at the end of period t. When production of an item i is not used to satisfy the demand, it can be stored and incurs no cost. For each item i, there is an initial inventory $s^i_0 \in \mathbb{R}_+$.

The goal is to satisfy the whole demand at minimum cost.

We call this problem the *deterministic multi-sourcing problem.*

Let's give some explanations. First, when a company gives a plant p the ability to produce an item i, some of them impose that production of item i must not be under some threshold ℓ^i_{pt} at period t. It is often to train the workers and in this particular case, these thresholds are usually constant and positive during the first periods after assignment and equal to zero during the last periods. Second, the capacity needed to produce one item may depends on time. It is often the case when people are involved in production since they become more skilled as time goes by. Finally, since multi-sourcing is a long-term decision and since deciding inventory levels is a mid-term decision, comparing assignment costs to holding costs does not make sense. Thus, we do not consider holding costs although it is easy for an industrial to extend the model we propose here after by adding a constraint keeping inventory below some threshold.

8.1.3 Main results

We model the deterministic multi-sourcing problem as a mixed integer program in Section 8.3 and prove that deterministic multi-sourcing problem is NP-hard in the strong sense and give some polynomial cases in Section 8.4.

8.2 Bibliography

Production multi-sourcing can be viewed as a lot-sizing problem on several parallel machines. Although it is typically a long-term decision and for this reason, uncertainty is often taken into account, a vast literature exists for the deterministic case. Reviews are proposed by Sethi and Sethi (1990), Koste and Malhotra (1999) and Stevenson and Spring (2007) even if they deal with more topics than the only multi-sourcing problem. They are interested in every types of flexibility whereas we focus our research on mix flexibility (ability to produce different combinations of products given certain capacity).

As explained by Fiorotto et al. (2018), deterministic settings may appear in real life. For example, in semi-conductor manufacturing, pharmaceutical production, packing for yogurt or textile industry, demand is know but machines must be qualified to produce items. They develop a deterministic lot-sizing problem on multiple parallel machines aiming at minimizing the sum of setup, production, holding and backlog costs. One main constraint is the maximal flexibility of the configuration. Their work is based on "chains", which is a group of products and plants which are all connected, directly or indirectly, by product assignment decisions. The key idea behind chaining is that excess capacity can be shifted along the chain and hence decrease amount of lost sales. They show than "long chains" configurations give results almost as good as complete flexibility. A previous work by Ignizio (2009) aiming at minimizing the balance between the use of the machines shows similar results. An extensive study of this chains based on simulations was made by Muriel et al. (2006).

In previous papers, flexibility is upper bounded and this bound is an input of the model. The following models address a complete flexibility but the trade-off is made thanks to the different costs in the objective function. In the model proposed by Rezaei and Davoodi (2008), each supplier have a part of imperfect items and a trade-off must be found between the revenue of selling goods and imperfect items and the other costs (like purchase cost or holding cost).

Finally, Snyder et al. (2006) do not directly address a multi-sourcing problem. However, their work aims at anticipating a disruption in the Supply Chain and controlling bad cases values. They consider multiple facilities and multiple clients. Each facility has a known probability of disruption. They aims at minimizing the weighted sum of assignment and reassignment if one or many disruption occurs.

Our model is very close to the existing models. The main difference is that our model often has much less parameters and constraints. However, to the best of our knowledge, it is not a special case of existing models since our minimal production constraint due to assignment decisions is not considered in the literature.

8.3 Model formulation

In this section, we introduce a mixed integer program which models the deterministic multi-sourcing problem. We introduce the following decision variables. The quantity of item i produced at period t by plant p is denoted by q^i_{pt} and the inventory at the end of the period is denoted by s^i_t. We also introduce a binary variable y^i_p which takes the value 1 if plant p is given the ability to produce item i.

The deterministic multi-sourcing problem can be written as

$$
\begin{aligned}
\min \quad & \sum_{i\in\mathcal{I}}\sum_{p\in\mathcal{P}} c^i_p y^i_p && && \text{(8.1a)}\\
\text{s.t.} \quad & s^i_t = s^i_{t-1} + \sum_{p\in\mathcal{P}} q^i_{pt} - d^i_t && \forall t\in[T],\ \forall i\in\mathcal{I}, && \text{(8.1b)}\\
& \sum_{i\in\mathcal{I}} v^i_{pt} q^i_{pt} \leqslant C_{pt} && \forall t\in[T],\ \forall p\in\mathcal{P}, && \text{(8.1c)}\\
& \ell^i_{pt} y^i_p \leqslant v^i_{pt} q^i_{pt} \leqslant u^i_{pt} y^i_p && \forall t\in[T],\ \forall p\in\mathcal{P}, \forall i\in\mathcal{I}, && \text{(8.1d)}\\
& s^i_{pt},\ q^i_{pt} \geqslant 0 && \forall t\in[T],\ \forall p\in\mathcal{P}, \forall i\in\mathcal{I}. && \text{(8.1e)}\\
& y^i_p \in \{0,1\} && \forall p\in\mathcal{P}, \forall i\in\mathcal{I}, && \text{(8.1f)}
\end{aligned}
$$

Objective (8.1a) minimizes the assignment costs. Constraint (8.1b) is the inventory balance. Capacity of each plant is ensured by constraint (8.1c). Constraint (8.1d) is both a "big-M" constraint and a bound on the production of each item in each plant. Note that without loss of generality, we can suppose that $u^i_{pt} \leqslant C_{pt}$ for each period t, each plant p and each item i.

8.4 NP-completeness

For any fixed values of the c^i_p's, the deterministic multi-sourcing problem is NP-hard. Moreover, looking at the special cases which remain NP-hard and at the polynomial cases, we see that the difficulty of the multi-sourcing problem is due either to the lower bounds ℓ^i_{pt} on production or to the assignment cost c^i_p.

Theorem 25. *Deciding if there is a solution of the deterministic multi-sourcing problem is* NP*-complete in the strong sense.*

Reducing the 3-partition problem to the deterministic multi-sourcing problem, we show that deciding if the deterministic multi-sourcing problem has a solution is NP-hard in the strong sense. We remind that the 3-partition problem consists in deciding whether a given multiset $\{a_1, \dots, a_{3m}\}$ of integers can be partitioned into triples that all have the same sum. This problem is known to be NP-complete in the strong sense (see Garey and Johnson (1979)) even if $\frac{B}{4} < a_i < \frac{B}{2}$ for each i with $B = \frac{1}{m}\sum_{i=1}^{3m} a_i$.

Proof. Let $\{a_1, \dots, a_{3m}\}$ be an instance of the 3-partition problem such that $\frac{B}{4} < a_i < \frac{B}{2}$ for each i with $B = \frac{1}{m}\sum_{i=1}^{3m} a_i$. We reduce polynomially this problem to an instance of the deterministic multi-sourcing problem. We set

$$\begin{gathered} T = 1, \quad \mathcal{P} = \{1, \dots, m\}, \quad \mathcal{I} = \{1, \dots, 3m\} \\ v^i_{p,1} = 1, \quad d^i_1 = \tfrac{a_i}{B}, \quad \ell^i_{p,1} = \tfrac{a_i}{B}, \quad u^i_{p,1} = C_{p,1} = 1. \end{gathered} \tag{8.2}$$

Thus, if we have a solution for the 3-partition problem, finding a solution of the deterministic multi-sourcing problem is straightforward assigning each triple to a plant.

Conversely, if we have a solution of the deterministic multi-sourcing problem, since $d^i_1 = \frac{a_i}{B} > 0$ for each item i, each item is assigned to at least one plant (otherwise, solution is infeasible). Since $\frac{1}{4}C_{p,1} = \frac{1}{4} < \frac{a_i}{B} = \ell^i_{p,1}$ ensures that there are at most three items per plant, then each plant produces exactly 3 items. Plants having the same capacity and sum of plant capacities being equal to sum of demands, each triple has the same sum. Thus, we get a solution of the 3-partition problem.

The conclusion follows from the fact that the 3-partition problem is NP-complete in the strong sense even if $\frac{B}{4} < a_i < \frac{B}{2}$ for each i. □

Theorem 25 prove the strong NP-hardness but the proof requires instances of the deterministic multi-sourcing problem with many plants and uses the lower bound on production which may be equal to zero in some real instances. One can ask for simpler cases. Proposition 26 gives some special cases which remain NP-hard and Proposition 27 gives some polynomial cases.

Proposition 26. *The following special cases of the deterministic multi-sourcing problem are* NP*-hard:*

1. *the deterministic multi-sourcing problem without lower bounds on production,*

2. *the deterministic multi-sourcing problem with only one period and only two plants.*

Proof of Proposition 26 (Item 1). Let $\{a_1, \ldots, a_{3m}\}$ be an instance of the 3-partition problem such that $\frac{B}{4} < a_i < \frac{B}{2}$ for each i with $B = \frac{1}{m}\sum_{i=1}^{3m} a_i$. We reduce polynomially this problem to an instance of the deterministic multi-sourcing problem without lower bounds on production. We set

$$\begin{gathered} T = 1, \quad \mathcal{P} = \{1, \ldots, m\}, \quad \mathcal{I} = \{1, \ldots, 3m\}, \\ c_p^i = 1, \quad d_1^i = \frac{a_i}{B}, \quad v_{p,1}^i = 1, \quad \ell_{p,1}^i = 0, \quad u_{p,1}^i = C_{p,1} = 1. \end{gathered} \tag{8.3}$$

Thus, if we have a solution for the 3-partition problem, finding a solution with cost $|\mathcal{I}|$ of the deterministic multi-sourcing problem is straightforward assigning each triple to a plant.

Conversely, suppose that we have a solution with cost $|\mathcal{I}|$ of the deterministic multi-sourcing problem. Since $d_1^i = \frac{a_i}{B} > 0$ for each item i, each item is assigned to at least one plant (otherwise, solution is infeasible). Assignment costs being equal to 1 and solution cost being equal to the number of item, each item is assigned to at most one plant, hence to exactly one plant. Since $\frac{1}{4}C_{p,1} = \frac{1}{4} < \frac{a_i}{B} = d_1^i$ ensures that there are at most three items per plant, then each plant produces exactly 3 items and we get a collection of m triples. Plants having the same capacity and sum of plant capacities being equal to sum of demands, each triple has the same sum. Thus, we get a solution of the 3-partition problem.

The conclusion follows from the fact that the 3-partition problem is NP-complete in the strong sense even if $\frac{B}{4} < a_i < \frac{B}{2}$ for each i. □

Reducing the partition problem to the deterministic multi-sourcing problem, we show that the deterministic multi-sourcing problem with only one period and only two plants remains NP-hard. We remind that the partition problem is the task of deciding whether a given set of positive integers can be partitioned into two subsets that have the same sum. This problem is known to be NP-complete (see Garey and Johnson (1979)).

Proof of Proposition 26 (Item 2). Let $\{a_1, \ldots, a_m\}$ be an instance of the partition problem. We reduce polynomially this problem to an instance of the deterministic multi-sourcing problem. We set

$$\begin{gathered} T = 1, \quad \mathcal{P} = \{1, 2\}, \quad \mathcal{I} = \{1, \ldots, m\}, \\ c_p^i = 1, \quad d_1^i = \frac{2a_i}{\sum_{i=1}^m a_i}, \quad v_{pt}^i = 1, \quad \ell_{pt}^i = 0, \quad u_{pt}^i = C_{pt} = 1. \end{gathered} \tag{8.4}$$

Thus, if we have a solution for the partition problem, finding a solution with cost $|\mathcal{I}|$ of the deterministic multi-sourcing problem is straightforward.

Conversely, if we have a solution with cost $|\mathcal{I}|$ of the deterministic multi-sourcing problem, positivity of the a_i ensures that each item is assigned to at least one plant. Cost of the solution being equal to $|\mathcal{I}|$, each item is assigned to exactly one plant. Plants having the same capacity and sum of plant capacities being equal to sum of demands, each subset define by the assignment has the same sum. Thus, we get a solution of the partition problem.

The conclusion follows from the fact that the partition problem is NP-complete. □

We now gives the polynomial cases.

Proposition 27. *The following special cases of the deterministic multi-sourcing problem are polynomial:*

1. *deterministic multi-sourcing problem with a single plant ($\mathcal{P} = \{1\}$),*
2. *deterministic multi-sourcing problem without lower bound on production and without assignment cost ($\ell^i_{pt} = 0$ and $c^i_p = 0$),*
3. *deterministic multi-sourcing problem with infinite capacities ($u^i_{pt} = C_{pt} = +\infty$).*

Proof. Case 1: deterministic multi-sourcing problem with a single plant.

For each item i, we set

$$y^i_1 = \begin{cases} 1 \text{ if there exists } t \in [T] \text{ such that } d^i_t > 0, \\ 0 \text{ otherwise.} \end{cases} \tag{8.5}$$

Then, we solve the resulting linear program to get an optimal solution of the deterministic multi-sourcing problem. (It may return that the problem is infeasible.)

Case 2: deterministic multi-sourcing problem without lower bound on production and without assignment cost.

For each plant p and each item i, we set $y^i_p = 1$. These decisions do not affect the cost in the objective function and do not lead to infeasibility since they relax the constraints on inventory and production variables. Then, we solve the resulting linear program to get a feasible solution of the deterministic multi-sourcing problem. (It may return that the problem is infeasible.)

Case 3: deterministic multi-sourcing problem with infinite capacity.

For each item i we choose a plant $p(i)$ among $\arg\min_{p\in\mathcal{P}} \left(c^i_p\right)$. Then, we set

$$y^i_p = \begin{cases} 1 \text{ if } p = p(i), \\ 0 \text{ otherwise.} \end{cases} \tag{8.6}$$

Then, we solve the resulting linear program to get an optimal solution of the deterministic multi-sourcing problem. (It may return that the problem is infeasible.) □

9 Stochastic multi-sourcing

9.1 Introduction

9.1.1 Motivations

In Chapter 8, data are deterministic. In practice, part of data is uncertain. Indeed, multi-sourcing decisions are long-term decisions and it is unlikely that forecast demand is perfectly accurate. As for production planning decisions, uncertainty may come from efficiency of machines (and affect time needed for the production of one item), capacities (due to breakdown, strike) or any other part of the supply chain like the procurement of raw materials. For multi-sourcing decisions, Argon Consulting has identified uncertainty on forecast demand as the main issue for its clients. Since the multi-sourcing problem addresses the question of flexibility (see Chapter 2), the relevant uncertainty comes from variation in the product mix and neither from variations of the global volume of demand nor from other sources.

Dealing with uncertainty leads to consider possible unsatisfied demand. Methods to face uncertainty goes from the most conservative, which does not allow stock-out in any outcome, to methods controlling stock-out probability, stock-out quantity or other relevant indicators. Discussions with Argon Consulting raise two characteristics of unsatisfied demand. First, multi-sourcing decisions should lead to "high long term service level". Indeed, multi-sourcing decisions take time to be implemented. Thus, service level on the first periods can be low due to augmentation of capabilities and should be less relevant than service level in steady state. The measure used for service level should not be too sensitive to sparse failure at the beginning. Second, worst cases must be "not too bad". Multi-sourcing decisions are long-term decisions which impact the company competitiveness for long periods since they cannot be easily changed. Thus, while remaining not too expensive, multi-sourcing decisions should be efficient enough to control the loss when bad outcomes occur.

9.1.2 Problem statement

For the sake of completeness, we present the problem in full details but it is very close to the deterministic one. We consider a set $\mathcal{P}$ of plants producing a set $\mathcal{I}$ of items over T periods. There is an upper bound C_{pt} on the total period production of plant p at period t (summed over all items). This upper bound is expressed in time units since it correspond to available working hours.

Giving a plant p the ability to produce an item i has a cost c_p^i. This cost is paid once and for all for the whole horizon. When a plant p is able to produce item i, there is an upper (resp. lower) bound $u_{pt}^i \geqslant 0$ (resp. $\ell_{pt}^i \geqslant 0$) on the production of item i in plant p at period t. The capacity needed (in time units) to produce one unit of item i in plant p in period t is $v_{pt}^i > 0$. For each item i, the average value of v_{pt}^i over plants and periods is denoted v^i.

The production must satisfy a random demand as well as possible. However, because of uncertainty, backorder (i.e., late delivery) is allowed. We introduce a parameter $\beta \in (0,1)$ which controls the part of demand delivered on time. It first defines the proportion of cases for which demand has to be satisfied. Second, in the remaining cases (i.e., the "worst cases"), the expectation of the sum of inventory and backorder must be nonnegative.

The demand of item i over period t is a random parameter $\boldsymbol{d}_t^i$, whose realization is known at the end of period t. When production of an item i (summed over all plants) is not used to satisfy the demand, it can be stored and incurs no cost. For each item i, there is an initial inventory $s_0^i \in \mathbb{R}_+$.

Regarding randomness, we assume that for any i and t, realizations of $(\boldsymbol{d}_t^i, \dots, \boldsymbol{d}_T^i)$ have finite expectation and can be efficiently sampled, knowing a realization of $(\boldsymbol{d}_1^j, \dots, \boldsymbol{d}_{t-1}^j)_{j \in \mathcal{I}}$. Multi-sourcing decisions are made before knowing any realizations of the demand. Production decisions can be made at the beginning of each period knowing past realizations of the demand. This kind of formulation is called *multi-stage*.

The goal is to satisfy the service level constraint at minimum cost.

We call this problem the *stochastic multi-sourcing problem*.

9.1.3 Main results

We propose a model for the stochastic multi-sourcing problem in Section 9.3 using the *Average-Value-at-Risk* denoted by the acronym AV@R. In Section 9.4, we develop a method to solve it by linearizing the AV@R, and a heuristic to speed up the resolution on big dataset. In Section 9.5, we discuss the choices we made to model the uncertainty.

9.2 Bibliography

Deciding multi-sourcing is a long term decision and often comes with stochastic considerations. An extensive review is proposed by Yao and Minner (2017).

A seminal model about stochastic multi-sourcing is given by Jordan and Graves (1995). In automotive industry, they aim at deciding which plant (or which line) should have the ability to produce each item. This is a single-period model which aims at minimizing the shortfall subject to limited flexibility. As in the deterministic case, they conclude that a long chain configuration enables to have results almost as good as a complete flexibility. A similar model is proposed by Tomlin and Wang (2005) that consider risk-neutral and risk averse objectives. In order to deal with more complicated cases (multi-item and multi-market), Garavelli (2003) uses simulations to quantify impact of flexibility configurations. Again, a limited flexibility seems sufficient to achieve very good results.

One goal of stochastic multi-sourcing is to deal with risk. In literature, production, ordering or multi-sourcing problems taking risk into account are often single-item problems. Risk aversion extension of the news-vendor problem was extensively studied (see Agrawal and Seshadri (2000) and Choi and Ruszczyński (2008)). More recently, Arikan and Fichtinger (2017) propose a classification of existing extensions (included those with AV@R). Ahmed et al. (2007) are one of the rare authors using coherent risk measure for a multi-period inventory problem. Their objective is to minimize a risk averse function (like AV@R) instead of a risk neutral function (like expectation). However, these are production decisions and not multi-sourcing decisions.

A single-item single-period multi-sourcing problem is also addressed in Meena et al. (2011). In this paper, each supplier has a different rate of failure (i.e., to produce nothing) but demand is equally divided between suppliers. They studied different cases and in particular, the minimization of the expected total cost for a pre-determined service level.

9.3 Model formulation

To model the constraint on the backorder, we first introduce the Average-Value-at-Risk and then present the model. Choices made to model the risk will be discussed in Section 9.5.

9.3.1 Average-Value-at-Risk

In order to measure and control the service level, we use the *Average-Value-at-Risk (*AV@R*)* which is a widely studied *coherent risk measure* (see for example Artzner et al. (1999) and Rockafellar and Uryasev (2000, 2002)) and chose the definition given in Föllmer and Schied (2004). In financial context, risk measures are used to quantify the risk of a position.

Fix some level $\lambda \in (0,1)$. For a random variable $\boldsymbol{X} : \Omega \to \mathbb{R}$, its *Value-at-Risk at level* λ is defined as

$$\mathrm{V@R}_{\lambda}\left(\boldsymbol{X}\right) = \inf\left\{m \in \mathbb{R} \middle| \mathbb{P}\left\{\boldsymbol{X} + m < 0\right\} \leqslant \lambda\right\}. \tag{9.1}$$

For the inventory, $\mathrm{V@R}_{1-\beta}\left(\boldsymbol{s}\right)$ might correspond to the safety stock in the case of a cycle service level β. More precisely, adding this quantity to the inventory keeps the probability of a stock-out below $1-\beta$. The main drawback of the Value-at-Risk is that it only controls the probability of a stock-out and not its size. Thus, we use the following coherent risk measure defined from the Value-at-Risk.

Fix some level $\lambda \in (0,1)$. For a random variable $\boldsymbol{X} : \Omega \to \mathbb{R}$, its *Average-Value-at-Risk at level* λ is defined as

$$\mathrm{AV@R}_{\lambda}\left(\boldsymbol{X}\right) = \frac{1}{\lambda}\int_0^{\lambda} \mathrm{V@R}_{\gamma}\left(\boldsymbol{X}\right) \mathrm{d}\gamma. \tag{9.2}$$

For the inventory, $\mathrm{AV@R}_{1-\beta}\left(\boldsymbol{s}\right)$ looks like the safety stock in the case of a fill rate service level β. More precisely, $\mathrm{AV@R}_{1-\beta}\left(\boldsymbol{s}\right) \leqslant 0$ means that the average of safety stocks needed to ensure cycle service level from β to 1 is negative (i.e., no safety stocks is required).

Figure 9.1 represents the Value-at-Risk and the Average-Value-at-Risk in the case of a Gaussian distribution of the stock $\boldsymbol{s}_t^i$.

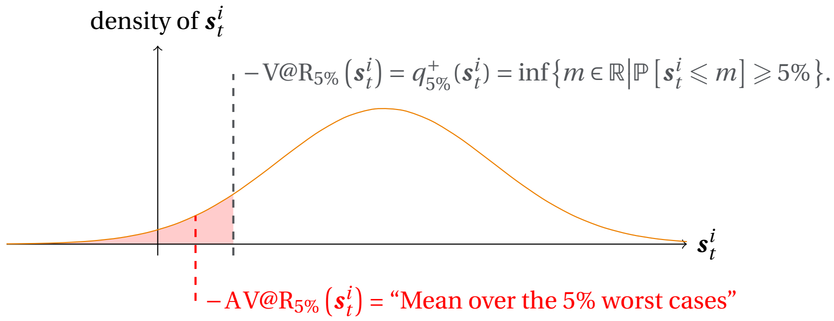


Figure 9.1 – Example of risk measures for the inventory

9.3.2 Model

In order to solve the stochastic multi-sourcing problem, we introduce the following decision variables. The quantity of item i produced at period t by plant p is denoted by $\boldsymbol{q}_{pt}^i$ and the inventory at the end of the period is denoted by $\boldsymbol{s}_t^i$. We also introduce a binary variable y_p^t which takes the value 1 if plant p is given the ability to produce item i. All these variables are random and may depend on the past realizations of the random demand $(\boldsymbol{d}_1^j, \dots, \boldsymbol{d}_{t-1}^j)_{j \in \mathcal{I}}$.

We model the stochastic multi-sourcing problem as

$$
\begin{aligned}
\min \quad & \sum_{i \in \mathcal{I}} \sum_{p \in \mathcal{P}} c_p^i y_p^i && && (9.3a)\\
\text{s.t.} \quad & \boldsymbol{s}_t^i = \boldsymbol{s}_{t-1}^i + \sum_{p \in \mathcal{P}} \boldsymbol{q}_{pt}^i - \boldsymbol{d}_t^i && \forall t \in [T],\ \forall i \in \mathcal{I}, && (9.3b)\\
& \sum_{i \in \mathcal{I}} v_{pt}^i \boldsymbol{q}_{pt}^i \leqslant C_{pt} && \forall t \in [T],\ \forall p \in \mathcal{P}, && (9.3c)\\
& \ell_{pt}^i y_p^i \leqslant v_{pt}^i \boldsymbol{q}_{pt}^i \leqslant u_{pt}^i y_p^i && \forall t \in [T],\ \forall p \in \mathcal{P}, \forall i \in \mathcal{I}, && (9.3d)\\
& \mathrm{AV@R}_{1-\beta}\left(\boldsymbol{s}_t^i\right) \leqslant 0 && \forall t \in [T],\ \forall i \in \mathcal{I}, && (9.3e)\\
& y_p^i \in \{0,1\} && \forall p \in \mathcal{P},\ \forall i \in \mathcal{I}, && (9.3f)\\
& \boldsymbol{q}_{pt}^i \text{ is } \sigma\Big(\big(\boldsymbol{d}_1^i, \dots, \boldsymbol{d}_{t-1}^i\big)_{i \in \mathcal{I}}\Big)\text{–measurable} && \forall t \in [T],\ \forall p \in \mathcal{P},\ \forall i \in \mathcal{I}. && (9.3g)
\end{aligned}
$$

Objective (9.3a) still minimizes the assignment costs. Constraints (9.3b), (9.3c), (9.3d) have the same meaning than their deterministic counterpart (8.1b), (8.1c) and (8.1d). Constraint (9.3e) enables to control the service level using Average-Value-at-Risk. Last constraint (9.3g) of the program, written as a *measurability constraint*, means that the values of the variables $\boldsymbol{q}_{pt}^i$ can only depend on the values taken by the demand before time t (the decision maker does not know the future). Every constraint of the problem, except the Average-Value-at-Risk constraint (9.3e), holds almost surely.

9.4 Solving method and theoretical results

Due to the obvious hardness of program (9.3), we turn to an approximate method.

9.4.1 Linearization of Average-Value-at-Risk

Linearization of AV@R constraint can be efficiently achieved in case of a finite probability space Ω. Lemma 28, based on results by Rockafellar and Uryasev (2000), gives the complete definition of the polyhedron defined by AV@R constraint in case of a finite probability space.

Lemma 28. *Consider a random variable* $\boldsymbol{X}$ *taking its values in a finite set* $\{X_\omega : \omega \in \Omega\}$ *with probability* $\{p_\omega : \omega \in \Omega\}$. *For* $\lambda \in (0,1)$, *the inequality* $\mathrm{AV@R}_\lambda(\boldsymbol{X}) \leqslant 0$ *holds if and only if the following polyhedron is nonempty:*

$$\left\{ (\alpha,\mu) \in \mathbb{R}_+^\Omega \times \mathbb{R} \;\middle|\; \frac{1}{|\Omega|} \sum_{\omega\in\Omega} p_\omega \alpha_\omega \leqslant \lambda\mu \quad \text{and} \quad \alpha_\omega \geqslant \mu - X_\omega \text{ for all } \omega \in \Omega \right\}.$$

Proof. We have the following equalities.

$$\mathrm{AV@R}_\lambda(\boldsymbol{X}) = \frac{1}{\lambda} \inf_{\mu\in\mathbb{R}} \left(\mathbb{E}\left[(\mu - \boldsymbol{X})^+ \right] - \lambda\mu \right) \tag{9.4a}$$

$$= \frac{1}{\lambda} \inf_{\mu\in\mathbb{R}} \left(\sum_{\omega\in\Omega} p_\omega (\mu - X_\omega)^+ - \lambda\mu \right) \tag{9.4b}$$

Equality (9.4a) is given by Föllmer and Schied (2004, Proposition 4.37). Equality (9.4b) holds because the space probability Ω is finite. Since Ω is finite, $\mathrm{AV@R}_\lambda(\boldsymbol{X}) \leqslant 0$ is equivalent to the existence of a $\mu \in \mathbb{R}$ such that

$$\frac{1}{\lambda} \left(\sum_{\omega\in\Omega} p_\omega (\mu - X_\omega)^+ - \lambda\mu \right) \leqslant 0, \tag{9.5}$$

which in turn is equivalent to the existence of $(\alpha,\mu) \in \mathbb{R}_+^\Omega \times \mathbb{R}$ such that

$$\begin{cases} \frac{1}{\lambda} \left(\sum_{\omega\in\Omega} p_\omega \alpha_\omega - \lambda\mu \right) \leqslant 0, \\ \alpha_\omega \geqslant \mu - X_\omega \quad \forall \omega \in \Omega. \end{cases} \tag{9.6}$$

□

9.4.2 Solving method

As shown in Section 8.4, the deterministic multi-sourcing problem is hard. Therefore, we cannot expect a quick algorithm solving exactly the problem, and this holds especially for the full stochastic version. Like stochastic production planning problems (see Part II), stochastic multi-sourcing problems are hard to solve.

We use the same approximation method than for solving Stochastic CLSP-BS (see Section 6.4.1). We propose a two-stage approximation consisting in replacing the measurability constraint (9.3g) by

$$\boldsymbol{q}_{pt}^i \text{ is } \sigma\Big(\big(\boldsymbol{d}_1^i, \ldots, \boldsymbol{d}_T^i \big)_{i\in\mathcal{I}} \Big)\text{–measurable} \quad \forall t \in [T],\ \forall i \in \mathcal{I} \tag{9.7}$$

which provides a relaxation of the initial program: the assignment decisions can still not depend on the future, but now the production decisions depend on the future demand. (Relaxing the measurability constraint means that once the y_p^r has been chosen, the demand over the whole horizon is supposed to be revealed.) We denote this relaxation by (2SA).

This approximation is a *two-stage approximation* as we distinguish between two levels of information over the uncertainty: assignment decisions are the *first stage* variables, while all other decisions are *second stage* variables.

The (2SA) relaxation is then solved by a classical *sample average approximation* (see Kleywegt et al. (2002) for a presentation of the method). We build a set Ω of m scenarios sampled uniformly at random. Each of these scenarios is a possible realization of $(\boldsymbol{d}_1^i, \dots, \boldsymbol{d}_T^i)$ for each item i. The parameter m is fixed prior to the resolution.

In order to linearize AV@R constraint (9.3e), we introduce for each period t and each item i auxiliary variables $\mu_t^i \in \mathbb{R}$ and $\left(\alpha_{t,\omega}^i\right)_\omega \in \mathbb{R}_+^\Omega$. Thanks to Lemma 28, we get the following mixed integer program (2SA-m).

$$
\begin{aligned}
\min \quad & \sum_{i \in \mathcal{I}} \sum_{p \in \mathcal{P}} c_p^i y_p^i && && \text{(9.8a)}\\
\text{s.t.} \quad & s_{t,\omega}^i = s_{t-1,\omega}^i + \sum_{p \in \mathcal{P}} q_{pt,\omega}^i - d_{t,\omega}^i && \forall \omega \in \Omega,\ \forall t \in [T],\ \forall i \in \mathcal{I}, && \text{(9.8b)}\\
& \sum_{i \in \mathcal{I}} v_{pt}^i q_{pt,\omega}^i \leqslant C_{pt} && \forall \omega \in \Omega,\ \forall t \in [T],\ \forall p \in \mathcal{P}, && \text{(9.8c)}\\
& \ell_{pt}^i y_p^i \leqslant v_{pt}^i q_{pt,\omega}^i \leqslant u_{pt}^i y_p^i && \forall \omega \in \Omega,\ \forall t \in [T],\ \forall p \in \mathcal{P}, \forall i \in \mathcal{I}, && \text{(9.8d)}\\
& \frac{1}{|\Omega|} \sum_{\omega \in \Omega} \alpha_{t,\omega}^i \leqslant (1-\beta)\mu_t^i && \forall t \in [T],\ \forall i \in \mathcal{I}, && \text{(9.8e)}\\
& \alpha_{t,\omega}^i \geqslant \mu_t^i - s_{t,\omega}^i && \forall \omega \in \Omega,\ \forall t \in [T],\ \forall i \in \mathcal{I}, && \text{(9.8f)}\\
& q_{pt,\omega}^i,\ \alpha_{t,\omega}^i \geqslant 0 && \forall \omega \in \Omega,\ \forall t \in [T],\ \forall i \in \mathcal{I}, && \text{(9.8g)}\\
& y_p^i \in \{0,1\} && \forall p \in \mathcal{P}, \forall i \in \mathcal{I}. && \text{(9.8h)}
\end{aligned}
$$

This mixed-integer program can be solved by any linear solver for small datasets but needs heuristics or decomposition methods for big datasets.

9.4.3 Heuristic to solve mixed integer program (9.8)

When working on big datasets, LP solvers are unable to find feasible solutions (see Section 10.3). Thus, we propose a simple heuristic to find a solution in reasonable time.

As explained in Chapter 2, randomness comes from the variations of product mix and not from the global volume which is assumed to be constant. The basic idea of the heuristic is to scale the demand of every item at every period with a factor $\alpha \geqslant 1$ and to solve the deterministic version (8.1) obtained with this demand. We obtain an assignment $(\tilde{y}_p^i)_{p,i}$. By binary search, we look for a small α in order to minimize the assignment costs but large enough so that Problem (9.8) with the assignment $(\tilde{y}_p^i)_{p,i}$ fixed is feasible.

The details of the heuristic are given by Algorithm 2. It has two parameters:

- N, which is the number of iterations of the binary search,

- τ, which is the time limit to solve the Problem (8.1) with its demand scaled by α.

Algorithm 2: Heuristic to solve mixed integer program (9.8)

Input: Sets $\mathcal{I}$, $\mathcal{P}$ and Ω, parameters c_p^i, v_{pt}^i, C_{pt}, ℓ_{pt}^i, u_{pt}^i and β, and demand $d_{t,\omega}^i$
Output: Assignment $(y_p^i)_{p,i}$
Parameters: N, τ

```
for p ∈ P and i ∈ I do  y_p^i := 1                       // current assignment
cost := +∞                                                // cost of current assignment

for t ∈ [T] and i ∈ I do  d̂_t^i := (1/|Ω|) Σ_{ω∈Ω} d_{t,ω}^i      // average demand
W := Σ_{i∈I} Σ_{t=1}^T v^i d̂_t^i.                          // cumulative sum of working hours
C := Σ_{p∈P} Σ_{t=1}^T C_{pt}.                              // cumulative capacity of plants

if C < W then
    Return assignment (y_p^i)_{p,i}
end

Binary search on α ∈ [1, C/W] with N iterations
    for t ∈ [T] and i ∈ I do  d_t^i := α d̂_t^i             // demand in Problem (8.1)
    Solve Problem (8.1) with a time limit τ
    if Problem (8.1) is infeasible then
        Decrease α
    else
        (ỹ_p^i)_{p,i} := solution of Problem (8.1)
        new_cost := cost of assignment (ỹ_p^i)_{p,i}
        Solve Program (9.8) with assignment (ỹ_p^i)_{p,i} fixed
        if Program (9.8) with assignment (ỹ_p^i)_{p,i} is infeasible then
            Increase α
        else
            if new_cost ⩽ cost then
                (y_p^i)_{p,i} = (ỹ_p^i)_{p,i}
                cost := new_cost
            end
            Decrease α
        end
    end
end
Return assignment (y_p^i)_{p,i}
```

The rational idea of the heuristic is the following. Solving Problem (8.1) with a higher demand enables to find a multi-sourcing which faces an augmentation of the demand of every item. In practice, it is too conservative since not every demand can increase in the same scenario (the cumulative sum of working hours is constant). For the binary search, the upper bound on α is chosen equal $\frac{C}{W}$ where W is the cumulative sum of working hours which is assumed constant over every scenarios and C is the cumulative capacity of plants over periods. Scaling the demand by $\frac{C}{W}$ gives a model where capacity is almost equal to working hours needed to satisfy demand.

Thus, every α greater than $\frac{C}{W}$ is unlikely to produce a feasible program. The lower bound on α is chosen equal to 1. If we cannot satisfy the demand in the case where it is equal to its average on scenarios, then program (9.8) is likely to be infeasible.

A binary search algorithm on α is justified by the following proposition.

Proposition 29. *Denote by* (P_α) *the Problem* (8.1) *with its demand scaled by* α. *Let* α_1 *and* α_2 *be two positive real numbers such that* $\alpha_1 \leqslant \alpha_2$. *Then, the optimal value of* (P_{α_1}) *is lower than or equal to the optimal value of* (P_α).

Proof. Let α_1 and α_2 be two positive real numbers such that $\alpha_1 \leqslant \alpha_2$. If (P_{α_2}) is feasible, it has an optimal solution $({y_p^i}^*)_{p,i}$, which is also a solution of program (P_{α_1}) and has the same cost. □

In Algorithm 2, the condition comparing the value of current multi-sourcing to the value of the new one is necessary in the case where the solver has not enough time to find the optimal solution of program (P). That is why, finding the smallest α such that (P_α) is feasible and such that Problem (9.8) with assignment equal to the optimal solution of (P_α) is feasible gives the best assignment that the heuristic can return.

We end this section with limits on Algorithm 2. The algorithm may not find a feasible solution even if the problem is feasible. Indeed, in general, the set of α such that (P_α) is feasible and such that problem (9.8) with fixed assignment equal to the optimal solution of (P_α) is not an interval.

Consider the following example. We have two items, two plants with capacities $C_1 = 100$ and $C_2 = 55$, one period, and two scenarios with same probability. The parameter β is chosen equal to 100% preventing from any stock-out. The other parameters of the multi-sourcing problem are given in Table 9.1.

	plant 1	plant 2
item 1	1	20
item 2	1	10

(a) Assignment costs c_p^i

	scenario 1	scenario 2
item 1	60	40
item 2	40	60

(b) Demand $d_{1,\omega}^i$

$v_{p,1}^i = 1$
$\ell_{p,1}^i = 0$
$u_{p,1}^i = 100$

(c) Other parameters

Table 9.1 – Parameters of the counterexample

As shown in Table 9.2, program (P_α) is feasible for $\alpha \in \{1, 1.1, 1.2\}$. However, if Problem (9.8) with the optimal assignments of (P_1) and of $(P_{1.2})$ is feasible, it is infeasible with the optimal assignments of $(P_{1.1})$.

9.4.4 Bender decomposition

Formulation (9.8) has a structure which fit with Bender decomposition. We refer to Birge and Louveaux (2011, Chapter 5) for a presentation of the method. Indeed, introducing the vector z of first-step variables (assignment variables $(y_p^i)_{p,i}$ and auxiliary variables $(\mu_t^i)_{t,i}$) and for each scenario ω, the vector w_ω of second-step variables (inventory variables $(s_{t,\omega}^i)_{t,i,\omega}$, production

	Optimal assignment of (P_α)			Cost	Feasible MIP (9.8)
		plant 1	plant 2		
$\alpha = 1$	item 1	1	0	2	true
	item 2	1	0		
		plant 1	plant 2		
$\alpha = 1.1$	item 1	1	0	11	false
	item 2	0	1		
		plant 1	plant 2		
$\alpha = 1.2$	item 1	1	0	12	true
	item 2	1	1		

Table 9.2 – Solutions returned by an iteration of Algorithm 2 for $\alpha \in \{1, 1.1, 1.2\}$

variables $(q^i_{pt,\omega})_{p,t,i}$ and auxiliary variables $(\alpha^i_{t,\omega})_{t,i,\omega}$, we get the following program

$$\begin{aligned}
\min \quad & f(z) && && (9.9a)\\
\text{s.t.} \quad & A(z,u) = a && \forall \omega \in \Omega, && (9.9b)\\
& Tz + W u_\omega = h_\omega && \forall \omega \in \Omega, && (9.9c)\\
& z \in \mathcal{Z}, && && (9.9d)
\end{aligned}$$

where constraint (9.9b) is the coupling constraint (9.8e), constraint (9.9c) regroups constraints (9.8b), (9.8c), (9.8d), (9.8f) and (9.8g) and $\mathcal{Z}$ is binary constraint on part of first-step variables. Thus, the constraints have the following block structure.

$$\begin{pmatrix} \multicolumn{5}{|c|}{A} \\ T & W & & (0) & \\ T & & W & & \\ \vdots & & & \ddots & \\ T & (0) & & & W \end{pmatrix} \qquad (9.10)$$

We did not try to implement this method, but it might be a promising way to speed up the resolution of formulation (9.8).

9.5 Discussion on expectation, robust, probabilistic and Average-Value-at-Risk constraints

As explained in Section 9.1.1, multi-sourcing decisions must lead to "high long term service level" and to good controls on "worst cases". These constrains are ill-defined. Many way of modeling them are possible. We chose to use the Average-Value-at-Risk. In this section, we compare its pros and cons to the ones of expectation, robust and probabilistic constraints.

In the case of multi-sourcing problems, these three constraints are defined as follows. For each item i and each period t, the expectation constraint on the inventory $\boldsymbol{s}^i_t$ is written as "$\mathbb{E}\left[\boldsymbol{s}^i_t\right] \geqslant 0$", the probabilistic constraint as "$\mathbb{P}\left[\boldsymbol{s}^i_t \geqslant 0\right] \geqslant \beta$" and the robust constraint as "$\boldsymbol{s}^i_t \geqslant 0$, almost surely". We consider the most rough version of robustness. Note moreover that this robust constraint coincide with the extreme cases of probabilistic and AV@R constraints with $\beta = 1$.

9.5.1 Small example

To illustrate discussions about the different constraints, we propose the following example. Consider a company producing a single item for one period. Consider that the company may open up to 4 identical plants to produce the item and when it opens one, it must use its full capacity which is equal to 100. The random demand for the item is uniformly drawn in $\{50, 100, 200, 250, 350\}$ and the parameter β to control the service level is equal to 60%. Opening a plant (i.e., assigning the item to a plant) has a unit cost. The parameters of the multi-sourcing problem are then set as follow

$$\begin{gathered} T = 1, \quad \mathcal{P} = \{1, \ldots, 4\}, \quad \mathcal{I} = \{1\}, \quad \boldsymbol{d}_1^i \sim \mathcal{U}\{50, 100, 200, 250, 350\}, \quad \beta = 60\%, \\ c_p^i = 1, \quad v_{p,1}^i = 1, \quad \ell_{p,1}^i = u_{p,1}^i = C_{p,1} = 100. \end{gathered} \tag{9.11}$$

Since the plants are identical, we simply use the number n of open plants as a decision variable and denote by $\boldsymbol{s}$ the inventory at the end of the period. Table 9.3 gives the values of the left part of each constraint depending on the number of open plants. Orange values correspond to cases where the constraint is satisfied. The two last lines correspond to the expected fill rate service level

$$\sum_{d \in \{50,100,200,250,350\}} \frac{1}{5} \times \frac{\min(d, 100n)}{d} \tag{9.12}$$

and to the expected number of lost sales

$$\sum_{d \in \{50,100,200,250,350\}} \frac{1}{5} \max(d - 100n, 0) \tag{9.13}$$

depending on the number n of open plants.

Number n of open plants	0	1	2	3	4
$\mathbb{E}[\boldsymbol{s}_t^i]$	-190	-90	10	110	210
$\mathbb{P}[\boldsymbol{s}_t^i \geqslant 0]$	0%	40%	60%	80%	100%
$\text{AV@R}_{1-\beta}(\boldsymbol{s})$	300	200	100	0	-100
$\boldsymbol{s}_t^i \geqslant 0$, almost surely	false	false	false	false	true
Expected fill rate service level	0%	64%	87%	97%	100%
Expected number of lost sales	190	100	40	10	0

Table 9.3 – Indicator values depending the number of open plants

To satisfy expectation or probabilistic constraints, the company must open at least 2 plants. To satisfy AV@R constraint, the company must open at least 3 plants. To satisfy robust constraint, the company must open at least 4 plants. As expected by the theory, robust constraint is more conservative than AV@R which is more conservative than probabilistic constraint.

9.5.2 Discussion

Table 9.4 summarizes the comparison between the expectation, robust, probabilistic and Average-Value-at-Risk constraints. We do not pretend to state general truth but rather general trend on common cases. $\oplus$ means that the constraint formulation satisfies the property, $\ominus$ means that the constraint formulation has difficulties to satisfy the property and ~ means that the constraint formulation is between these two extremes.

Properties	Constraints			
	AV@R	Expectation	Robust	Probabilistic
Easy linearization	$\oplus$	$\oplus$	$\oplus$	$\ominus$
Easy interpretation	~	$\ominus$	$\oplus$	$\oplus$
Control proportion of worst cases	$\oplus$	$\ominus$	$\ominus$	$\oplus$
Control value of worst cases	$\oplus$	$\ominus$	$\oplus$	$\ominus$
Control increase of objective function	$\oplus$	$\ominus$	$\ominus$	$\oplus$
Few feasibility issues	~	$\oplus$	$\ominus$	~
Computational complexity	~	$\oplus$	$\oplus$	$\ominus$

Table 9.4 – Comparison of constraint properties

We propose seven criteria to compare the constraints. First, the ease to linearize the constraint may be a good property of a formulation since it enable to use linear programming. Expectation constraint is already linear making it easy to use. On the other hand, probabilistic constraint may be hard to linearize since it may not be convex and thus may require additional binary variables. Robust and AV@R constraints are convex. Then, using scenarios often enables to get linear formulations of robust constraints of reasonable size and as shown in Section 9.4.1, AV@R can also be linearized.

Interpretation of robust and probabilistic constraints are quite easy. Robust constraint, in its vanilla form, consists in satisfying the constraint in every possible outcome. In multi-sourcing, demand distribution is bounded. Thus satisfying the robust constraint consists in having enough stock to satisfy the highest possible demand. Probabilistic constraint at level $\beta\%$ consists in satisfying the constraint in $\beta\%$ of the cases. In multi-sourcing, this means that in $\beta\%$ of the cases, stock is high enough to satisfy the demand but in the remaining cases demand can as well be fully, partially or not satisfied at all. It perfectly matches with the cycle service level. Even though expectation constraint seems easy to interpret, in practice, knowing the behavior of the mean does not give much information on the behavior of the outcomes. For example, if the demand has an all-or-nothing distribution, expectation is far from any outcome. More generally, as soon as variance of distribution is high, it is hard to interpret from expectation behavior. Understanding AV@R constraint at level $\lambda\%$ is not straightforward. However, it means that $(100-\lambda)\%$ of the cases satisfies the constraint and that the mean over the remaining cases also satisfies the constraint. Thus, it includes the interpretations of a probabilistic constraint at $(100-\lambda)\%$ and of an expectation constraint on the bad cases. If there are not too many bad cases (i.e., if λ is small), we might expect a small variance in their distribution (while nothing ensure it). Note that, for a same level, AV@R constraint is more strict than the probabilistic constraint since AV@R constraint implies that the probabilistic constraint is satisfied (see Rockafellar and Uryasev (2000, 2002)).

Thus, we easily see that the AV@R constraint enables to control the proportion of bad cases and their value. On the other extreme, expectation constraint does not control any of them. Being very conservative, the robust constraint does not enable to control the proportion of bad case: you must cover every outcome but it ensures that every outcome value satisfies the constraint. On the other hand, the probabilistic constraint enables to choose the proportion of worst cases but it is done at costs of control on their value.

Another advantage of the AV@R and probabilistic constraints is the possibility to choose the level of conservativeness. Indeed, the company can easily see the price implied by the service level constraint and possibly relax partially this constraint with a single parameter. On the other hand, there is no way to control the level of risk acceptance with the robust and expectation constraints.

Feasibility issues of each of these constraints are directly linked to their levels of conservativeness. The robust constraint is the most conservative and a single very bad outcome lead to an infeasible model. At the other extreme, the expectation constraint is much easier to satisfy. Choice of the type of constraints for this criterion is strongly related to the distribution of the random variable.

Finally, as expected, the flexibility offered by the AV@R and probabilistic constraints leads to computational complexity. On the other hand, the expectation and robust constraints are often easier to deal with.

We choose AV@R constraint because of its linearization properties. Moreover, as wanted by Argon Consulting, AV@R enables to control the acceptable proportion of bad cases and their values. For Argon Consulting, the real issue remains the computational complexity.

10 Numerical experiments

We test our method on datasets from Argon Consulting's clients. Moreover, we compare the multi-sourcing currently in use by the client to the multi-sourcing computed from our method. In addition to comparing their cost, several Key Performance Indicators (KPI) will be used as the service level (fill rate and cycle), the proportion of multi-sourced items, etc.

10.1 Simulation

Simulation is based on real datasets from Argon Consulting's clients. We first retrieve the historical data and the current multi-sourcing to establish a comparison point.

First, we define plants data (capacities, internal production times, eligibility or not to be given the ability to produce items) and a forecast demand which is a function of the current period and of the past outcomes. Then, we construct our model defined in Chapter 9. Finally, we solve it and get a multi-sourcing decision.

The question of whether this solution is feasible or not (i.e., satisfies th AV@R constraint) remains open, since, due to dataset sizes, the number of scenarios is small (100 in our experiments). In order to show the validity of the obtained assignment, we fix it and compute Problem (9.8) with a larger number of scenarios (typically 10000). This problem is a big linear program but has no integer variables. If the solver succeeds in finding a feasible solution, it is likely that the AV@R constraint is satisfied and we then compute other KPI's like the fill rate or the cycle service level, or the multi-sourcing level. If the solver proves the problem is infeasible, it is likely that the AV@R constraint is not satisfied and that the assignment is infeasible. Finally, no conclusion can be drawn if time limit is reached before finding a solution.

10.2 Instances

10.2.1 Datasets

We use two instances provided by Argon Consulting's clients: Mel and Lux.

The Mel dataset is a toy problem. It is a simplified dataset created by a client to enable the students from École des Ponts to train and propose new approaches to multi-sourcing.

The Lux dataset comes from the luxury industry. In the original dataset, there were 500 different items. However, we keep only 100 items of Lux dataset, which represent 80% of the production (and of the sales). Due to limited computational resources, the client prefers that the 400 items excluded from Lux dataset be mono-sourced to concentrate our effort on potential multi-sourced items.

Table 10.1 summarizes the characteristics of the two instances used for tests.

Instances characteristics		Instances					
		Mel			Lux		
Number of plants	$\lvert\mathcal{P}\rvert$			4			27
Number of items	$\lvert\mathcal{I}\rvert$			11			100
Number of time steps	T			3			6
Historical demand	$\bar{d}_t^i$	4884	–	19959	0	–	16996
Capacity of plants	C_p	6156	–	28969	2213	–	43416
Capacity for an item	u_p^i	1539	–	7242	721	–	12753
Internal production time	v_{pt}^i	0.4	–	1.1	0.4	–	9.0
Assignment cost	c_p^i	8392	–	82196	3516	–	62234

Table 10.1 – Instance characteristics

The parameter β, which controls the service level is chosen in $\{70\%, 85\%, 95\%\}$. The number m of scenarios used to solve (2SA-m) is fixed to 100 due to tractability restrictions. The time limit to return a solution has been set to 20 minutes because the client wants to try several service levels. Thus, the parameters for the heuristic defined in Section 9.4.3 are $N = 10$ and $\tau = 120$ seconds.

10.2.2 Demand distribution

As explained in Chapter 2, we study the ability to face variations of the product mix and assume that we already know the cumulative working hours of production $\sum_{t=1}^{T}\sum_{i\in\mathcal{I}} v^i \boldsymbol{d}_t^i$.

We generate randomness from historical data using "expanded Dirichlet" distributions as in Chapter 7. But this time, we assume that $(v^i \boldsymbol{d}_t^i)$ follows an "expanded Dirichlet" distribution. We refer to Section 7.3.2 for the properties of this distribution and to Algorithm 1 to sample it.

The parameters of the "expanded Dirichlet" distribution used to generate the working hours of production are computed from the historical forecast demand and γ is set equal to

$$\gamma = v \frac{\sum_{t=1}^{T}\sum_{i\in\mathcal{I}} \sqrt{\frac{\bar{p}_0}{\bar{p}_t^i} - 1}}{\sum_{t=1}^{T}\sum_{i\in\mathcal{I}} \left(\frac{\bar{p}_0}{\bar{p}_t^i} - 1\right)}. \tag{10.1}$$

The parameter γ is defined from another parameter v which we call *volatility* and which represents the ratio between standard deviation and expectation of a random variable. It can be easily interpreted. Justification of such a choice for γ can be found in Appendix A.3.

As in Chapter 7, results may depend on the number of generated scenarios. If we had the time to implement method of Section 7.4, we would like to apply the same scheme, that is:

- use Algorithm 1 to generate a sample of M scenarios,

- initialize K with $m = 100$,
- use K-means algorithm to classify the M scenarios into m clusters.

Then, our set Ω of representative scenarios used in Equation (9.8) would be composed of the prototype of each cluster, with a probability equal to the number of scenarios in its cluster divided by M.

10.3 Numerical results

C++11 has been chosen for the implementations and Gurobi 6.5.1 Gurobi Optimization, LLC (2018) was used to solve the model on a PC with 32 processor Intel® Xeon™ E5-2667 @ 3.30GHz and 192Go RAM.

There are two ways of measuring the KPI's. First, they can be measured *in-sample* (i.e., with the $m = 100$ scenarios used to decide multi-sourcing). In this case, the obtained solution of Problem (9.8) is used to compute the KPI as if it stood for the whole possible realizations of the randomness. On the other hand, they can be measured *out-of-sample*. We generate $m = 10000$ new realizations of the demand. Then, we solve Problem (9.8) with the assignment obtained from this optimization and these m new realizations of the demand.

Simulation results for Mel dataset with a frontal resolution with a solver are given in Table 10.2 and with the use of the Algorithm 2 in Table 10.3. Simulation results for Lux dataset with the use of the Algorithm 2 are in Table 10.4.

The inputs are decomposed as follows. The parameter β is the value used to control the desired service level, i.e., the risk level used in the AV@R constraint. The volatility v is the value used in the "expanded Dirichlet" distribution that generates the random realizations of the demand (see Section 10.2.2).

The outputs are decomposed as follows. The row *Solver Objective* (resp. *Heuristic Objective*) gives the assignment costs returned by the solver (resp. the heuristic) at the end of the time allowed to find a solution. This is the objective (9.8a) and it is expressed in k€. The row *Multi-sourcing* gives the percentage of multi-sourced families (i.e., families produced on two plants or more) and the row *Multi-sourcing (max)* gives the maximal number of plants producing one item.

The row $\mathbb{P}[s \geqslant 0]$ measures in-sample the probability (in %) that inventory of a random item at a random time is positive.

$$\mathbb{P}[s \geqslant 0] = \frac{1}{T \times |\mathcal{I}| \times |\Omega|} \sum_{t=1}^{T} \sum_{i \in \mathcal{I}} \sum_{\omega \in \Omega} \mathbb{1}_{\{s^i_{t,\omega} \geqslant 0\}} \tag{10.2}$$

The rows *GAP*, *Solver LB Cont. relax.*, and *Time for bound* are measured in-sample. The row *GAP* is the gap (in %) between the objective value and the best known lower bound on the problem. The row *Solver LB* is the lower bound (in k€) found by the solver when the solver is frontally used to solve Problem (9.8). The row *Cont. relax.* is the value (in k€) of the continuous relaxation of Problem (9.8). The row *Time for bound* is the time (in seconds) needed to get the bound written on the previous row of the table.

The row *Feasibility* says whether Problem (9.8) with the fixed obtained assignment and $m =$

10000 scenarios is feasible (*Yes*), infeasible (*No*), or whether time limit was reached *Limit*).

The rows *Cycle service level* and *Fill rate service level* are measured out-of-sample. They are computed using Equation (7.12) and Equation (7.13) by replacing w^i by

$$w^i = \frac{v^i}{\sum_{j\in\mathcal{I}} v^j}. \tag{10.3}$$

In this case, the item are weighted by their internal production time. These two KPI's are averaged over the $m = 10000$ scenarios and are expressed in %.

The row *Evaluation time* gives the time (in seconds) needed to solve Problem (9.8) with the fixed obtained assignment and $m = 10000$ scenarios. The time limit of the solver is set equal to 24 hours.

β	Volatility v						Output
	1%	5%	10%	30%	50%	80%	
70%	953	967	1011	1034	1134	1381	Solver Objective
	45.5	54.5	54.5	63.6	100.0	72.7	Multi-sourcing
	2	2	3	3	4	3	Multi-sourcing (max)
	93.8	93.5	93.3	91.2	90.1	91.4	$\mathbb{P}[s \geqslant 0]$
	Yes	Yes	Yes	Yes	No	Yes	Feasibility
	71.2	74.4	73.4	78.2	–	82.8	Cycle service level
	90.5	96.0	94.8	94.1	–	92.4	Fill rate service level
	22	28	86	34	61	38	Evaluation time
85%	957	957	1000	1087	1188	∞	Solver Objective
	54.5	54.5	54.5	72.7	63.6	–	Multi-sourcing
	3	2	2	3	3	–	Multi-sourcing (max)
	98.5	98.1	97.6	96.5	95.9	–	$\mathbb{P}[s \geqslant 0]$
	17.5	15.0	18.1	17.7	11.3	–	GAP
	790	813	819	895	1054	–	Solver LB
	1200	1200	1200	1200	1200	–	Time for bound
	Yes	Yes	Yes	Yes	Yes	–	Feasibility
	75.1	67.6	61.9	80.4	–	–	Cycle service level
	94.1	94.2	94.3	95.3	–	–	Fill rate service level
	17	21	40	25	43	–	Evaluation time
95%	954	956	997	1173	1371	∞	Solver Objective
	54.5	54.5	72.7	72.7	81.8	–	Multi-sourcing
	3	3	2	3	3	–	Multi-sourcing (max)
	99.6	99.6	99.3	99.0	99.0	–	$\mathbb{P}[s \geqslant 0]$
	Yes	Yes	Yes	No	No	–	Feasibility
	72.1	76.6	inf.	–	–	–	Cycle service level
	95.5	95.6	inf.	–	–	–	Fill rate service level
	16	18	1204	863	1880	–	Evaluation time

Table 10.2 – Results for Mel dataset (frontally solved with the solver)

β	Volatility v						Output
	1%	5%	10%	30%	50%	80%	
70%	1152	1176	1156	1136	1541	1640	Heuristic Objective
	90.9	100.0	100.0	100.0	100.0	100.0	Multi-sourcing
	4	3	4	4	4	4	Multi-sourcing (max)
	95.5	93.8	92.7	89.7	88.8	90.1	$\mathbb{P}[s \geqslant 0]$
	Yes	Yes	Yes	Yes	Yes	Yes	Feasibility
	83.6	83.4	72.1	82.7	88.7	85.3	Cycle service level
	96.5	97.5	96.2	95.9	96.6	94.0	Fill rate service level
	77	25	40	29	32	36	Evaluation time
85%	1152	1134	1146	1112	1230	∞	Heuristic Objective
	90.9	81.8	81.8	63.6	81.8	100.0	Multi-sourcing
	4	4	4	3	4	4	Multi-sourcing (max)
	97.7	97.2	96.6	96.2	95.7	–	$\mathbb{P}[s \geqslant 0]$
	31.4	28.3	28.5	19.5	14.3	–	GAP
	673	707	686	797	900	1144	Cont. relax.
	3.8	13.8	3.5	2.4	13.7	3.6	Time for bound
	Yes	Yes	Yes	Yes	No	Yes	Feasibility
	79.4	75.0	91.0	82.7	–	86.5	Cycle service level
	95.2	96.3	98.4	95.7	–	95.1	Fill rate service level
	42	35	33	27	55	27	Evaluation time
95%	1152	968	1021	1210	1396	∞	Heuristic Objective
	90.9	63.6	54.5	72.7	72.7	100.0	Multi-sourcing
	4	2	2	4	3	4	Multi-sourcing (max)
	99.2	99.8	99.8	99.0	99.1	–	$\mathbb{P}[s \geqslant 0]$
	Yes	Yes	Yes	No	No	No	Feasibility
	76.2	76.7	72.7	–	–	–	Cycle service level
	95.0	97.1	93.9	–	–	–	Fill rate service level
	25	30	24	2333	2394	6739	Evaluation time

Table 10.3 – Results for Mel dataset (solved with Algorithm 2)

10.4 Discussion

First, note that we do not provide results for Lux dataset without the heuristic. Indeed, we were not able to find a solution in the allowed time (20 minutes). Moreover, for the Lux dataset, it is even impossible to compute the continuous relaxation in less than 24 hours for a volatility equal to 1%, 50% and 80%.

Second important point, on small datasets, the frontal use of a solver always gives better results than Algorithm 2 even if this gap decreases with the increase of volatility. However, on huge datasets, the solver was never able to get a feasible solution.

Although it is not a proof, the values returned by $\mathbb{P}[s \geqslant 0]$ show on some examples that the satisfaction of the AV@R constraint implies the satisfaction of the probabilistic constraint.

As shown by the computational times, evaluating the quality solution with fixed assignments is hard. It may take up to two hours while the problem has no integer variables.

β	Volatility v						Output
	1%	5%	10%	30%	50%	80%	
70%	1392	1409	1420	1476	1624	1861	Heuristic Objective
	18.0	23.0	18.0	19.0	22.0	25.0	Multi-sourcing
	3	3	3	4	4	3	Multi-sourcing (max)
	99.4	99.3	99.2	98.7	98.6	98.5	$\mathbb{P}[s \geqslant 0]$
	Yes	Yes	Yes	Yes	Yes	Yes	Feasibility
	77.9	77.3	82.9	72.0	88.2	92.0	Cycle service level
	86.2	85.9	88.7	80.6	91.9	94.0	Fill rate service level
	4703	4293	3892	2212	2363	2154	Evaluation time
85%	1392	1409	1407	1516	1708	1997	Heuristic Objective
	18.0	23.0	16.0	21.0	24.0	31.0	Multi-sourcing
	3	3	3	3	3	4	Multi-sourcing (max)
	99.9	99.8	99.8	99.6	99.5	99.4	$\mathbb{P}[s \geqslant 0]$
	–	24.1	22.3	24.7	–	–	GAP
	–	1070	1084	1142	–	–	Cont. relax.
	86400	28728	31277	51399	86400	86400	Time for bound
	Yes	Limit	Yes	Yes	Yes	Yes	Feasibility
	77.8	–	82.0	85.2	90.5	92.7	Cycle service level
	86.1	–	87.7	89.4	94.0	94.6	Fill rate service level
	4262	86402	5148	3692	3652	3044	Evaluation time
95%	1392	1420	1421	1624	1900	2382	Heuristic Objective
	18.0	18.0	19.0	22.0	26.0	34.0	Multi-sourcing
	3	3	3	4	4	5	Multi-sourcing (max)
	100.0	100.0	99.9	99.9	99.9	99.9	$\mathbb{P}[s \geqslant 0]$
	Yes	Yes	Yes	Yes	Yes	Yes	Feasibility
	77.8	80.3	81.1	86.4	90.6	92.7	Cycle service level
	85.9	87.2	87.1	90.1	93.1	94.5	Fill rate service level
	4025	5146	5209	5448	5716	5428	Evaluation time

Table 10.4 – Results for Lux dataset (solved with Algorithm 2)

On Mel instances, we see that solving on a small set of scenarios is too optimistic when computing the multi-sourcing. Indeed, several assignments prove to be infeasible of the larger set of scenarios.

Regarding results on Lux instances, we see that Algorithm 2 is not that bad. When we succeed in finding the continuous relaxation of Problem (9.8) with $m = 100$ scenarios, the GAP to optimality is at most 25%. For small values of the volatility v, playing with the value of β has little impact on the objective value and on service level. For greater values of the volatility, playing with the values of β enables us to reach the desired service levels. However, when β becomes close to 100%, our method becomes very sensitive to the sampling and approaches like robust ones may be more appropriate.

10.5 Comparison with the current multi-sourcing of Argon Consulting's client

Regarding the whole items list (i.e., the 500 families), Argon Consulting's client has almost 25% of multi-sourced items. In the 100 considered items, it is about 80% of the items. With our approach, we show that we can drastically reduce the number of multi-sourced items (and thus the multi-sourcing costs) while keeping a high service level. As mentioned in Section 10.4, some complementary implementations can be made in order to improve the method and ensure the validity of the approach.

Extensions Part IV

11 Production planning using cover-sizes

11.1 Introduction

11.1.1 Motivations

As seen in Part I and Part II, assembly lines may be managed using (r,ℓ) policies (i.e., production when the safety stock is reached) or directly making the production planning for the next week. The choice depends partially on the complexity of the production, on the variation of demand over time. Sometimes, companies compute the cycle stocks of items and use them as an input of their production planning tools.

(r,ℓ) policies are easy to compute and to use but they do not take into account part of the available information when making their optimization process. On the other hand, production planning algorithms are more sophisticated and sometimes less convenient to use but they take into account much more information to find the optimal production (variation of demand over time, uncertainty). Thus, their output is likely to be more accurate.

In this chapter, we aim at considering the variations of demand in our computation of the optimal cover-sizes defined in Chapter 3. Context and modeling choices are quite the same as in Chapter 5. A company aims at minimizing its inventory subject to flexibility constraint. As explained in Chapter 2, when deciding cover-sizes, flexibility is already fixed and many companies express it with an upper bound on the number of items produced during a period. At mid-term, it matches with the number of setups since scheduling is not taken into account. However, contrary to Chapter 5, we do not look for production levels but to find the optimal cover-sizes.

11.1.2 Problem statement

For the sake of completeness, we recall every parameters including those which are identical to those of the Uniform CLSP-BS (see Section 5.1.2). The problem considers an assembly line producing a set $\mathcal{I}$ of items over T periods. The number of distinct items produced over a period t cannot exceed $N > 0$. There is also an upper bound C on the total period production (summed over all items). The capacity needed (in time units) to produce one unit of i is $v^i > 0$.

The production of item i must satisfy a demand d_t^i at the end of period t. When production of a item i is not used to satisfy the demand, it can be stored but incurs a unit holding cost

$h^i > 0$ per period. For each item i, there is an initial inventory $s_0^i \in \mathbb{R}_+$. Since we use cover-sizes, the production of item i is made periodically and thus, the time interval between two consecutive setups of i is identical over the planning horizon and belong to a set $\mathcal{T}_i \subseteq [T]$ of possible cover-sizes. These cover-sizes are integers, which makes them easier to use for planners.

The goal is to satisfy the whole demand at minimum cost.

Since this problem is a variation of the Uniform CLSP-BS with the use of cover-size instead of production levels, we call it the *Uniform Capacitated Cover-Sizing Problem with Bounded number of Setups (Uniform CCSP-BS)*.

11.1.3 Main results

In this chapter, we model the Uniform CCSP-BS as a mixed integer program and show that this problem is NP-hard (Section 11.2). Then, we present the method used to test the efficiency and the reasons for which we do not make further investigations (Section 11.3).

11.2 Model formulation

In this section, we introduce a mixed integer program which models the problem. We introduce the following decision variables. The quantity of item i produced at period t is denoted by q_t^i and the inventory at the end of the period is denoted by s_t^i. We introduce a binary variable x_t^i which takes the value 1 if the item i is produced over week t. Finally, we introduce a binary variable y_τ^i which takes the value 1 if the cover-size τ is chosen for item i.

We normalize the production variables by setting $\widehat{q}_t^i = \frac{v^i q_t^i}{C}$ and replace accordingly the other variables and parameters by setting $\widehat{s}_t^i = \frac{v^i s_t^i}{C}$, $\widehat{d}_t^i = \frac{v^i d_t^i}{C}$ and $\widehat{h}^i = \frac{Ch^i}{v^i}$. For the purpose of notation, hats are omitted and the Uniform CCSP-BS can be written as

$$
\begin{aligned}
\min \quad & \sum_{t=1}^{T} \sum_{i \in \mathcal{I}} h^i s_t^i && && (11.1a)\\
\text{s.t.} \quad & s_t^i = s_{t-1}^i + q_t^i - d_t^i && \forall i \in \mathcal{I},\ \forall t \in [T], && (11.1b)\\
& \sum_{i \in \mathcal{I}} q_t^i \leqslant 1 && \forall t \in [T], && (11.1c)\\
& q_t^i \leqslant x_t^i && \forall i \in \mathcal{I},\ \forall t \in [T], && (11.1d)\\
& \sum_{i \in \mathcal{I}} x_t^i \leqslant N && \forall t \in [T], && (11.1e)\\
& \sum_{\tau \in \mathcal{T}_i} y_\tau^i = 1 && \forall i \in \mathcal{I}, && (11.1f)\\
& 2x_{t+\tau}^i = x_t^i + y_\tau^i && \forall i \in \mathcal{I},\ \forall \tau \in \mathcal{T}_i,\ \forall t \in [T-\tau] && (11.1g)\\
& \sum_{t=1}^{T} x_t^i \leqslant 1 + M \sum_{(\tau' \in \mathcal{T}_i | \tau' < \tau)} y_{\tau'}^i && \forall i \in \mathcal{I},\ \forall \tau \in \mathcal{T}_i, && (11.1h)\\
& y_\tau^i \in \{0,1\} && \forall i \in \mathcal{I},\ \forall \tau \in \mathcal{T}_i, && (11.1i)\\
& x_t^i \in \{0,1\} && \forall i \in \mathcal{I},\ \forall t \in [T], && (11.1j)\\
& q_t^i,\ s_t^i \geqslant 0 && \forall i \in \mathcal{I},\ \forall t \in [T], && (11.1k)
\end{aligned}
$$

where M is a big positive real number.

Inventory balance (11.1b), capacity constraints (11.1c) and (11.1d), and the upper bound on the number of setups (11.1e) have the same meaning than their counterparts (5.3b), (5.3c), (5.3d) and (5.3e) in the Uniform CLSP-BS. We add three other constraints to the model. Constraint (11.1f) ensures that we choose exactly one cover-size τ for each item i. Constraint (11.1g) ensures that production of item i at period $t+\tau$ is allowed only if production of item i is allowed at period t and if the cover-size of item i is equal to τ. Constraint (11.1h) ensures that, if cover-size τ is chosen for item i, there is at most one setup of the production of i during the first τ periods. (Note that with constraint (11.1g), it ensures that there is exactly one setup.)

Like the Uniform CLSP-BS, the Uniform CCSP-BS is NP-hard for any fixed values of the h_t^i's.

Theorem 30. *Deciding if there is a solution of the Uniform CCSP-BS is* NP*-complete in the strong sense.*

The proof is very close to proof of NP-hardness of CLSP-BS since it uses reduction of 3-partition problem to the Uniform CCSP-BS.

Proof. Let $\{a_1,\ldots,a_{3m}\}$ be an instance of the 3-partition problem. We reduce polynomially this problem to an instance of the Uniform CCSP-BS. Without loss of generality, we can assume that sum of the a_i's is positive. We set

$$\mathcal{I}=\{1,\ldots,3m\},\quad T=m,\quad \mathcal{T}_i=\{m\},\quad N=3,\quad d_t^i=\begin{cases}\frac{m\,a_i}{\sum_{j=1}^{3m}a_j} & \text{if } t=T,\\ 0 & \text{otherwise,}\end{cases}\quad s_0^i=0. \tag{11.2}$$

Thus, we have a solution for the 3-partition problem if and only if there is a solution to the Uniform CCSP-BS with these parameters. The conclusion follows from the fact that the 3-partition problem is NP-complete in the strong sense. □

11.3 Numerical experiments

The efficiency of this approach is also tested using simulations. (See Chapter 7.)

This new heuristic, called *MIP-cover-size heuristic*, is very close to the cover-size heuristic. It also consists in determining before the first week once and for all a value τ_i^* for each item $i\in\mathcal{I}$, from Problem (11.1) where the deterministic demand is taken equal to the demand expectation. Then, like in the cover-size heuristic, at time t, if the inventory of item i is below a precomputed safety stock (see Equation (7.1)), the quantity q_t^i is computed so that the inventory of item i exceeds the safety stock of the expected demand for the next τ_i^* weeks. In addition, if some capacity issues are easily anticipated, the production of an item i can be activated even if the inventory is not below the safety stock.

Tests with MIP-cover-size heuristic show that this approach does not seem to improve cover-size heuristic described in Section 7.2. Inventory holding costs and service level are not better than with the cover-size heuristic. Moreover, computing cover-sizes from Problem (11.1) requires a solver contrary to the closed-form formula of Section 3.2.3. Since this method uses a solver, one might as well use the (2SA-m) heuristic which also relies on a solver and give better results.

12 Multi-sourcing problem with budget constraint

12.1 Introduction

12.1.1 Motivations

As explained in Part III, multi-sourcing decisions consist in deciding which plants should have the ability to produce an item. They are long-term decisions and are piloted by company's strategy. When making multi-sourcing decisions, companies have already defined the desired service level and want to reach it at a minimal cost. However, because they deal with long-term decisions, we use a risk constraint to control stock-out and ensure service level. Finally, company's strategy is a trade-off between service level and costs to ensure it.

In this chapter, we propose an alternative version of the multi-sourcing problem described in Chapter 8 reversing the roles of service level and cost of multi-sourcing. Like in Fiorotto et al. (2018), we suppose that we have a limited budget for the flexibility (i.e., to multi-source production of items) and that the demand is deterministically known. Thus, companies are likely to not satisfy the whole demand and aim at minimizing their backorder.

12.1.2 Problem statement

For the sake of completeness, we recall every parameters including those which are identical to those of the multi-sourcing problem. We consider a set $\mathcal{P}$ of plants producing a set $\mathcal{I}$ of items over T periods. There is an upper bound C_{pt} on the total period production of plant p at period t (summed over all items). This upper bound is expressed in time units since it corresponds to available working hours.

Giving a plant p the ability to produce an item i has a cost c^i_p. This cost is paid once and for all for the whole horizon and the cumulative assignment cost must not exceed a quantity K. When a plant p is able to produce item i, there is an upper (resp. lower) bound $u^i_{pt} \geqslant 0$ (resp. $\ell^i_{pt} \geqslant 0$) on the production of item i in plant p at period t. The capacity needed (in time units) to produce one unit of item i in plant p in period t is $v^i_{pt} > 0$.

The production and the inventory of item i (summed over all plants) should satisfy a demand d^i_t at the end of period t. When production of an item i is not used to satisfy the demand, it can be stored and incurs no cost. When a demand for item i is not satisfied by the production of the current period or by inventory, it can be satisfied later but incurs a unit backorder cost $\gamma^i > 0$ per period. For each item i, there is an initial inventory $s_0^i \in \mathbb{R}_+$.

The goal is to minimize the backorder cost.

We call this problem the *multi-sourcing problem with budget constraint.*

12.1.3 Main results

In this chapter, we model the multi-sourcing problem with budget constraint as a mixed integer program (Section 12.2). Then, we show that the multi-sourcing problem with budget constraint is NP-hard even in many simple cases (Section 12.3).

12.2 Model formulation

In this section, we introduce a mixed integer program which models the multi-sourcing problem with budget constraint. We introduce the following decision variables. The quantity of item i produced at period t by plant p is denoted by q_{pt}^i and the inventory at the end of the period is denoted by s_t^i. The backorder of item i at the end of the period t is denoted by b_t^i. We also use the inventory level $\tilde{s}_t^i$ that is the relative value of the inventory of item i at the end of period t (i.e., the inventory minus the backorder). We also introduce a binary variable y_p^i which takes the value 1 if plant p is given the ability to produce item i.

The multi-sourcing problem with budget constraint can be written as

$$
\begin{aligned}
\min \quad & \sum_{t=1}^{T} \sum_{i\in\mathcal{I}} \gamma^i b_t^i && && (12.1a)\\
\text{s.t.} \quad & \tilde{s}_t^i = \tilde{s}_{t-1}^i + \sum_{p\in\mathcal{P}} q_{pt}^i - d_t^i && \forall t\in[T],\ \forall i\in\mathcal{I}, && (12.1b)\\
& \sum_{i\in\mathcal{I}} v_{pt}^i q_{pt}^i \leqslant C_{pt} && \forall t\in[T],\ \forall p\in\mathcal{P}, && (12.1c)\\
& \ell_{pt}^i y_p^i \leqslant v_{pt}^i q_{pt}^i \leqslant u_{pt}^i y_p^i && \forall t\in[T],\ \forall p\in\mathcal{P},\ \forall i\in\mathcal{I}, && (12.1d)\\
& \sum_{i\in\mathcal{I}} \sum_{p\in\mathcal{P}} c_p^i y_p^i \leqslant K && && (12.1e)\\
& \tilde{s}_t^i = s_t^i - b_t^i && \forall t\in[T],\ \forall i\in\mathcal{I}, && (12.1f)\\
& s_t^i,\ b_t^i,\ q_{pt}^i \geqslant 0 && \forall t\in[T],\ \forall p\in\mathcal{P},\ \forall i\in\mathcal{I}, && (12.1g)\\
& y_p^i \in \{0,1\} && \forall p\in\mathcal{P},\ \forall i\in\mathcal{I}. && (12.1h)
\end{aligned}
$$

Objective (12.1a) minimizes the cumulative backorder over periods. Constraint (12.1b) is the inventory balance. Capacity of each plant is ensured by constraint (12.1c). Constraint (12.1d) is both a “big-M” constraint and a bound on the production of each item in each plant. Capital invested in assignment is limited by constraint (12.1e). Constraint (12.1f) is the decomposition of the inventory in its positive and negative parts.

12.3 NP-completeness

The multi-sourcing problem with budget constraint is NP-hard in the strong sense but contrary to the deterministic multi-sourcing problem introduced in Section 8.3, it always has a feasible solution. (It is sufficient to give to no plants the ability to produce any item.) Moreover, as in the deterministic multi-sourcing problem, the combinatorial difficulty of the multi-sourcing problem with budget constraint is due either to the lower bounds ℓ^i_{pt} on production, or to the upper bound K on the cumulative assignment cost.

Theorem 31. *The multi-sourcing problem with budget constraint is* NP*-hard in the strong sense.*

The proof is very close to the proof of NP-hardness of the deterministic multi-sourcing of problem Chapter 8 since it uses a reduction of 3-partition problem to the multi-sourcing problem with budget constraint.

Proof. Let $\{a_1,\ldots,a_{3m}\}$ be an instance of the 3-partition problem such that $\frac{B}{4} < a_i < \frac{B}{2}$ for each i with $B = \frac{1}{m}\sum_{i=1}^{3m} a_i$ (it is known to be NP-complete according to Garey and Johnson (1979)). We reduce polynomially this problem to an instance of the multi-sourcing problem with budget constraint. We set

$$\begin{gathered} T=1, \quad \mathcal{P}=\{1,\ldots,m\}, \quad \mathcal{I}=\{1,\ldots,3m\}, \quad K=3m, \\ \gamma^i=1, \quad c^i_p=1, \quad v^i_{p,1}=1, \quad d^i_1=\frac{a_i}{B}, \quad \ell^i_{p,1}=0, \quad u^i_{p,1}=C_{p,1}=1. \end{gathered} \tag{12.2}$$

Thus, if we have a solution for the 3-partition problem, finding a solution of the multi-sourcing problem with budget constraint that satisfies the whole demand is straightforward.

Conversely, suppose that we have a solution of the multi-sourcing problem with budget constraint which satisfies the whole demand. Since $d^i_1 = \frac{a_i}{B} > 0$ for each item i, each item is assigned to at least one plant (otherwise, there would be backorder). Assignment costs being equal to 1 and the cumulative assignment cost being upper bounded by the number of items, each item is assigned to exactly one plant. Moreover, $\frac{1}{4}C_{p,1} = \frac{1}{4} < \frac{a_i}{B} = d^i_1$ ensures that there are at most three items per plant. Then, we get a collection of m triples. Plants having the same capacity and the sum of plant capacities being equal to the sum of demands, each triple has the same sum. Thus, we get a solution of the 3-partition problem.

The conclusion follows from the fact that the 3-partition problem is NP-complete in the strong sense even if $\frac{B}{4} < a_i < \frac{B}{2}$ for each i. □

Proposition 32 gives some special cases that are still NP-hard.

Proposition 32. *The following special cases of the multi-sourcing problem with budget constraint are* NP*-hard:*

1. *with only one period and only two plants ($T = 1$ $\mathcal{P} = \{1,2\}$),*
2. *with no budget constraint ($K = +\infty$),*
3. *with one plant and infinite capacities ($C_{pt} = +\infty$).*

Proof of Proposition 32 (Item 1). Let $\{a_1, \ldots, a_m\}$ be an instance of the partition problem (it is known to be NP-complete according to Garey and Johnson (1979)). We reduce polynomially this problem to an instance of the multi-sourcing problem with budget constraint. We set

$$\begin{gathered} T = 1, \quad \mathcal{P} = \{1,2\}, \quad \mathcal{I} = \{1,\ldots,m\}, \quad K = m, \\ \gamma^i = 1, \quad c_p^i = 1, \quad v_{pt}^i = 1, \quad d_1^i = \frac{2a_i}{\sum_{i=1}^m a_i}, \quad \ell_{pt}^i = 0, \quad u_{pt}^i = C_{pt} = 1. \end{gathered} \tag{12.3}$$

Thus, if we have a solution for the partition problem, finding a solution of the multi-sourcing problem with budget constraint that satisfies the whole demand is straightforward.

Conversely, suppose that we have a solution of the multi-sourcing problem with budget constraint that satisfies the whole demand. The positivity of the a_i ensures that each item is assigned to at least one plant. (Otherwise, there would be positive backorder.) The assignment cost being equal to 1 and the sum of assignment cost being upper bounded by the number of items, each item is assigned to exactly one plant. Plants having the same capacity and sum of plant capacities being equal to sum of demands, each subset define by the assignment has the same sum. Thus, we get a solution of the partition problem.

The conclusion follows from the fact that the partition problem is NP-complete. □

Looking at the proofs of Theorem 31 and of Proposition 32 (Item 1), we see that they can be easily adapted to the case of infinite upper bound on the sum of assignment cost.

Proof of Proposition 32 (Item 2). In the proofs of Theorem 31 and of Proposition 32 (Item 1), the lower bounds ℓ_{pt}^i on production must be taken equal to d_1^i and the capacities of plants and demand satisfaction imply that each item is produced in exactly one plant. Hence, we have the desired result. □

Reducing the knapsack problem to the multi-sourcing problem with budget constraint and infinite capacities, we show that this latter is NP-hard. We remind that the knapsack problem consists in choosing a subset of a set of n items, each with a positive value v_i and a positive weight $w_i >$, which maximizes the sum of their cumulative value and whose cumulative weight is lower than $W > 0$. This problem is known to be NP-complete (see Garey and Johnson (1979)).

Proof of Proposition 32 (Item 3). Let $\big((v_i, w_i)_{i\in[n]}, W\big)$ be an instance of the knapsack problem. We reduce polynomially this problem to an instance of the multi-sourcing problem with budget constraint. We set

$$\begin{gathered} T = 1, \quad \mathcal{P} = \{1\}, \quad \mathcal{I} = \{1,\ldots,n\}, \quad K = W, \\ \gamma^i = 1, \quad c_p^i = w_i, \quad d_1^i = \max_j \{v_j\} - v_i, \quad v_{pt}^i = 1, \quad \ell_{pt}^i = 0, \quad u_{pt}^i = C_{pt} = +\infty. \end{gathered} \tag{12.4}$$

Thus, the multi-sourcing problem with budget constraint can be written as

$$\max \quad \sum_{i=1}^{n} v_i y^i \tag{12.5a}$$

$$\text{s.t.} \quad \sum_{i=1}^{n} w_i y^i \leqslant W \tag{12.5b}$$

$$y^i \in \{0,1\} \qquad \forall i \in \mathcal{I}, \tag{12.5c}$$

and $\left(y^{*i}\right)_i$ is an optimal solution solution of the multi-sourcing problem with budget constraint if and only if $\left(y^{*i}\right)_i$ is an optimal solution of the knapsack problem. The conclusion follows from the fact that the knapsack problem is NP-hard. □

Conclusion

This thesis develops optimization methods for Supply Chain Management and is focused on flexibility. Argon Consulting was our partner during this thesis and we applied our work to the cases of its clients.

The first part deals with mid-term decisions. It aims at computing lot-sizes and cover-sizes which are used in several Supply Chain processes as MRP (to decide the quantity of raw materials to order) or production planning (to decide the size of produced lots). At this level of decision, the objective for Argon Consulting is to reduce future inventory and the main constraint is the flexibility. Motivated by industrial considerations, we propose extensions of classical continuous-time inventory models by replacing setup costs by a bound on the maximal number of setups. We find closed-form formulas in the single-line cases and proposed efficient methods for multi-line cases. Argon Consulting was very enthusiastic about this formula because unlike the setups costs, the bound on the number of setups captures the interaction between items and is easy to estimate. Moreover, the closed-form formula makes it easier to use and is already used by practitioners.

The second part deals with short-term decisions. It aims at deciding the production for the following periods in an uncertain environment. (Indeed, even if the data may be reliable for the current period, it is often uncertain for the next periods and must be anticipated.) As in mid-term decisions, our study was motivated by industrial considerations. Argon Consulting aims at minimizing the inventory subject to flexibility constraint modeled by an upper bound on the number of setups per period. We propose a new model based on the classical Capacitated Lot-Sizing Problem in both deterministic and stochastic settings. In order to solve it in reasonable time, we use a classical approximation scheme (two-stage approximation and a sampling method). However, Argon Consulting's clients were not able to provide us with a set of scenarios for the demand. Thus, we develop a probabilistic model to generate the set of scenarios, which is easy to use while being reasonable. Our experimentations already prove that the company that provides the datasets can reduce its inventory costs while keeping a good service level.

The last part deals with long-term decisions. Its aims at deciding the multi-sourcing of production, i.e., if some items should be produced by several plants and by which ones. These decisions are made in a highly uncertain environment and contrary to short-term decision, they are not repeated. In order to take into account the risk in multi-sourcing decision, we introduce the Average-Value-at-Risk in our model to quantify the risk of an assignment. This risk measure can be interpreted by the decision makers since it captures some characteristics of cycle service level and of fill rate service level and it has good computational properties. However, the size of the problem prevents us from frontally using up-to-date solvers. Instead, we develop a heuristic to find a feasible solution which can be used by practitioners or as an initial solution for a solver. Our experimentations already prove that the company that provides the datasets can reduce its multi-sourcing (thus its costs) while keeping a good service level.

First extensions of our work must be done on the sampling method. We propose a concrete way to reduce the dependence on the sampling based on clustering methods but, for lack of time, we have not implemented it yet. Secondly, to accelerate the solving of the mixed integer linear program, we propose decomposition methods (such as Bender decomposition) but again, for lack of time, we have not implemented it yet. Finally, our work uses expectation for the short-term decisions and Average-Value-at-Risk for the long-term decisions. A natural evolution would be to compare them to robust optimization to quantify the savings made and the price of robustness.

Bibliography

Absi, N. and Kedad-Sidhoum, S. (2009). The multi-item capacitated lot-sizing problem with safety stocks and demand shortage costs. *Computers & Operations Research*, 36(11):2926 – 2936.

Aggarwal, A. and Park, J. K. (1993). Improved algorithms for economic lot size problems. *Oper. Res.*, 41(3):549–571.

Agrawal, V. and Seshadri, S. (2000). Impact of uncertainty and risk aversion on price and order quantity in the newsvendor problem. *Manufacturing & Service Operations Management*, 2(4):410–423.

Ahmed, S., Çakmak, U., and Shapiro, A. (2007). Coherent risk measures in inventory problems. *European Journal of Operational Research*, 182(1):226–238.

Aloulou, M. A., Dolgui, A., and Kovalyov, M. Y. (2014). A bibliography of non-deterministic lot-sizing models. *International Journal of Production Research*, 52(8):2293–2310.

Arikan, E. and Fichtinger, J. (2017). The risk-averse newsvendor problem under spectral risk measures: A classification with extensions. *European Journal of Operational Research*, 256(1):116 – 125.

Arnold, J., Chapman, S., and Clive, L. (2007). *Introduction to Materials Management*. Pearson Education. Pearson Prentice Hall.

Arrow, K. J., Harris, T., and Marschak, J. (1951). Optimal inventory policy. *Econometrica*, 19(3):250–272.

Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. (1999). Coherent measures of risk. *Math. Finance*, 9(3):203–228.

Barbosa, L. C. and Friedman, M. (1978). Deterministic inventory lot size models—a general root law. *Management Science*, 24(8):819–826.

Benson, H. P. (1985). A finite algorithm for concave minimization over a polyhedron. *Naval Research Logistics Quarterly*, 32(1):165–177.

Benson, H. P. (1998). Deterministic algorithms for constrained concave minimization: A unified critical survey. *Naval Research Logistics (NRL)*, 43(6):765–795.

Birge, J. R. and Louveaux, F. (2011). *Introduction to stochastic programming.* Springer Series in Operations Research and Financial Engineering. Springer, New York, second edition.

Bomberger, E. E. (1966). A dynamic programming approach to a lot size scheduling problem. *Management Science*, 12(11):778–784.

Carpentier, P., Chancelier, J.-P., Cohen, G., and De Lara, M. (2015). *Stochastic multi-stage optimization*, volume 75 of *Probability Theory and Stochastic Modelling.* Springer, Cham. At the crossroads between discrete time stochastic control and stochastic programming.

Chang, C.-T. (2006). An acquisition policy for a single item multi-supplier system with real-world constraints. *Applied Mathematical Modelling*, 30(1):1 – 9.

Chauhan, S. S. and Proth, J.-M. (2003). The concave cost supply problem. *European Journal of Operational Research*, 148(2):374 – 383. Sport and Computers.

Choi, S. and Ruszczyński, A. (2008). A risk-averse newsvendor with law invariant coherent measures of risk. *Operations Research Letters*, 36(1):77 – 82.

Delmas, J.-F. and Jourdain, B. (2006). *Modèles aléatoires*, volume 57 of *Mathématiques & Applications (Berlin) [Mathematics & Applications].* Springer-Verlag, Berlin. Applications aux sciences de l'ingénieur et du vivant. [Applications to engineering and the life sciences].

Donaldson, W. A. (1977). Inventory replenishment policy for a linear trend in demand—an analytical solution. *Journal of the Operational Research Society*, 28(3):663–670.

Díaz-Madroñero, M., Mula, J., and Peidro, D. (2014). A review of discrete-time optimization models for tactical production planning. *International Journal of Production Research*, 52(17):5171–5205.

Edgeworth, F. Y. (1888). The mathematical theory of banking. *Journal of the Royal Statistical Society*, 51(1):113–127.

Eynan, A. and Kropp, D. H. (2007). Effective and simple EOQ-like solutions for stochastic demand periodic review systems. *European Journal of Operational Research*, 180(3):1135 – 1143.

Federgruen, A. and Tzur, M. (1991). A simple forward algorithm to solve general dynamic lot sizing models with n periods in 0(n log n) or 0(n) time. *Management Science*, 37(8):909–925.

Fiorotto, D. J., Jans, R., and de Araujo, S. A. (2018). Process flexibility and the chaining principle in lot sizing problems. *International Journal of Production Economics*, 204:244 – 263.

Florian, M., Lenstra, J. K., and Rinnooy Kan, A. H. G. (1980). Deterministic production planning: algorithms and complexity. *Management Sci.*, 26(7):669–679.

Frederickson, G. N. and Johnson, D. B. (1982). The complexity of selection and ranking in x + y and matrices with sorted columns. *Journal of Computer and System Sciences*, 24(2):197 – 208.

Friedman, M. F. (1984). On a stochastic extension of the EOQ formula. *European Journal of Operational Research*, 17(1):125–127.

Frigyik, B. A., Kapila, A., and Gupta, M. R. (2010). Introduction to the dirichlet distribution and related processes.

Föllmer, H. and Schied, A. (2004). *Stochastic Finance: An Introduction in Discrete Time*, chapter 4, pages 157–206. De Gruyter studies in mathematics. Walter de Gruyter.

Gallego, G. (1998). New bounds and heuristics for (q, r) policies. *Management Science*, 44(2):219–233.

Garavelli, A. (2003). Flexibility configurations for the supply chain management. *International Journal of Production Economics*, 85(2):141 – 153. Supply Chain Management.

Garey, M. and Johnson, D. (1979). *Computers and Intractability: A Guide to the Theory of NP-Completeness*. Series of books in the mathematical sciences. W. H. Freeman.

Gavish, B. and Graves, S. C. (1980). A one-product production/inventory problem under continuous review policy. *Operations Research*, 28(5):1228–1236.

Gayon, J.-P., Benjaafar, S., and de Véricourt, F. (2009). Using imperfect advance demand information in production-inventory systems with multiple customer classes. *Manufacturing & Service Operations Management*, 11(1):128–143.

Geunes, J. (2014). *Operations Planning: Mixed Integer Optimization Models*. Operations Research Series. Taylor & Francis.

Gicquel, C. and Cheng, J. (2018). A joint chance-constrained programming approach for the single-item capacitated lot-sizing problem with stochastic demand. *Annals of Operations Research*, 264(1):123–155.

Gicquel, C., Minoux, M., and Dallery, Y. (2008). Capacitated lot sizing models: a literature review. working paper or preprint.

Goyal, S. K. (1974). Determination of optimum packaging frequency of items jointly replenished. *Management Science*, 21(4):436–443.

Guimarães, L., Klabjan, D., and Almada-Lobo, B. (2014). Modeling lotsizing and scheduling problems with sequence dependent setups. *European Journal of Operational Research*, 239(3):644–662.

Gurobi Optimization, LLC (2018). Gurobi Optimizer Reference Manual.

Gyulai, D., Kádár, B., and Monosotori, L. (2015). Robust production planning and capacity control for flexible assembly lines. *IFAC-PapersOnLine*, 48(3):2312 – 2317. 15th IFAC Symposium on Information Control Problems in Manufacturing.

Gyulai, D., Pfeiffer, A., and Monostori, L. (2017). Robust production planning and control for multi-stage systems with flexible final assembly lines. *International Journal of Production Research*, 55(13):3657–3673.

Hadley, G. G. and Whitin, T. M. (1963). *Analysis of inventory systems*. Englewood Cliffs, N. J. : Prentice-Hall. Prentice-Hall quantitative methods series.

Harris, F. W. (1913). How many parts to make at once. *Factory, The Magazine of Management*, 10(2):135–136.

Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The elements of statistical learning*. Springer Series in Statistics. Springer, New York, second edition. Data mining, inference, and prediction.

Hochbaum, D. S. (1994). Lower and upper bounds for the allocation problem and other nonlinear optimization problems. *Mathematics of Operations Research*, 19(2):390–409.

Hong, J.-D. and Hayya, J. C. (1992). Just-in-time purchasing: Single or multiple sourcing? *International Journal of Production Economics*, 27(2):175 – 181.

Hsieh, C. H. (2002). Optimization of fuzzy production inventory models. *Information Sciences*, 146(1):29 – 40.

Ignizio, J. P. (2009). Cycle time reduction via machine-to-operation qualification. *International Journal of Production Research*, 47(24):6899–6906.

Jain, A. K. (2010). Data clustering: 50 years beyond k-means. *Pattern Recognition Letters*, 31(8):651 – 666. Award winning papers from the 19th International Conference on Pattern Recognition (ICPR).

Jordan, W. C. and Graves, S. C. (1995). Principles on the benefits of manufacturing process flexibility. *Management Science*, 41(4):577–594.

Karimi, B., Fatemi Ghomi, S. M. T., and Wilson, J. M. (2003). The capacitated lot sizing problem: a review of models and algorithms. *Omega*, 31(5):365 – 378.

Kim, T., Hong, Y., and Lee, J. (2005). Joint economic production allocation and ordering policies in a supply chain consisting of multiple plants and a single retailer. *International Journal of Production Research*, 43(17):3619–3632.

Kleywegt, A. J., Shapiro, A., and de Mello, T. H. (2002). The sample average approximation method for stochastic discrete optimization. *SIAM Journal on Optimization*, 12(2):479–502.

Koste, L. L. and Malhotra, M. K. (1999). A theoretical framework for analyzing the dimensions of manufacturing flexibility. *Journal of Operations Management*, 18(1):75 – 93.

Kotz, S., Balakrishnan, N., and Johnson, N. L. (2000). *Continuous Multivariate Distributions. Volume 1: Models and Applications*. Wiley Series in Probability and Statistics. New York: Wiley.

Lee, H.-M. and Yao, J.-S. (1999). Economic order quantity in fuzzy sense for inventory without backorder model. *Fuzzy Sets and Systems*, 105(1):13 – 31.

Lin, S.-D. and Su, S.-M. (2013). The optimal inventory policy of production management. *Engineering*, 5(5).

Macedo, P. B., Alem, D., Santos, M., Junior, M. L., and Moreno, A. (2016). Hybrid manufacturing and remanufacturing lot-sizing problem with stochastic demand, return, and setup costs. *International Journal of Advanced Manufacturing Technology*, 82(5):1241–1257.

Madigan, J. G. (1968). Scheduling a multi-product single machine system for an infinite planning period. *Management Science*, 14(11):713–719.

Meena, P., Sarmah, S., and Sarkar, A. (2011). Sourcing decisions under risks of catastrophic event disruptions. *Transportation Research Part E: Logistics and Transportation Review*, 47(6):1058 – 1074.

Minoux, M. (2018). Robust and stochastic multistage optimisation under markovian uncertainty with applications to production/inventory problems. *International Journal of Production Research*, 56(1-2):565–583.

Mukhopadhyay, S. K. and Ma, H. (2009). Joint procurement and production decisions in remanufacturing under quality and demand uncertainty. *International Journal of Production Economics*, 120(1):5 – 17. Special Issue on Operations Strategy and Supply Chains Management.

Mula, J., Poler, R., García-Sabater, J. P., and Lario, F.-C. (2006). Models for production planning under uncertainty: A review. *International Journal of Production Economics*, 103(1):271–285.

Muriel, A., Somasundaram, A., and Zhang, Y. (2006). Impact of partial manufacturing flexibility on production variability. *Manufacturing & Service Operations Management*, 8(2):192–205.

Nobil, A. H., Sedigh, A. H. A., and Cárdenas-Barrón, L. E. (2016). A multi-machine multi-product epq problem for an imperfect manufacturing system considering utilization and allocation decisions. *Expert Systems with Applications*, 56:310 – 319.

Ohno, K. and Ishigaki, T. (2001). A multi-item continuous review inventory system with compound poisson demands. *Mathematical Methods of Operations Research*, 53(1):147–165.

Padmanabhan, G. and Vrat, P. (1990). Inventory model with a mixture of back orders and lost sales. *International Journal of Systems Science*, 21(8):1721–1726.

Park, K. S. (1982). Inventory model with partial backorders. *International Journal of Systems Science*, 13(12):1313–1317.

Park, K. S. (1987). Fuzzy-set theoretic interpretation of economic order quantity. *IEEE Transactions on Systems, Man, and Cybernetics*, 17(6):1082–1084.

Park, S. S., Kim, T., and Hong, Y. (2006). Production allocation and shipment policies in a multiple-manufacturer–single-retailer supply chain. *International Journal of Systems Science*, 37(3):163–171.

Pochet, Y. and Wolsey, L. (2006). *Production Planning by Mixed Integer Programming*. Springer Series in Operations Research and Financial Engineering. Springer New York.

Raghavachari, M. (1969). On connections between zero-one integer programming and concave programming under linear constraints. *Operations Research*, 17(4):680–684.

Resh, M., Friedman, M., and Barbosa, L. C. (1976). On a general solution of the deterministic lot size problem with time-proportional demand. *Operations Research*, 24(4):718–725.

Rezaei, J. and Davoodi, M. (2008). A deterministic, multi-item inventory model with supplier selection and imperfect quality. *Applied Mathematical Modelling*, 32(10):2106 – 2116.

Rockafellar, R. and Uryasev, S. (2002). Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance*, 26(7):1443 – 1471.

Rockafellar, R. T. and Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of Risk*, 2:21–41.

Rosenblatt, M. J., Herer, Y. T., and Hefter, I. (1998). Note. an acquisition policy for a single item multi-supplier system. *Management Science*, 44(11):S96–S100.

Rozenknop, A., Wolfler Calvo, R., Alfandari, L., Chemla, D., and Létocart, L. (2013). Solving the electricity production planning problem by a column generation based heuristic. *Journal of Scheduling*, 16(6):585–604.

Rubaszewski, J., Yalaoui, A., Amodeo, L., Yalaoui, N., and Mahdi, H. (2011). A capacitated lot-sizing problem with limited number of changes. In *International Conference on Industrial Engineering and Systems Management, IESM, Metz, France, May 25-27*.

Salameh, M., Abboud, N., El-Kassar, A., and Ghattas, R. (2003). Continuous review inventory model with delay in payments. *International Journal of Production Economics*, 85(1):91 – 95. Planning and Control of Productive Systems.

Sana, S. S. (2011). The stochastic EOQ model with random sales price. *Appl. Math. Comput.*, 218(2):239–248.

Sethi, A. K. and Sethi, S. P. (1990). Flexibility in manufacturing: A survey. *International Journal of Flexible Manufacturing Systems*, 2(4):289–328.

Shapiro, A., Dentcheva, D., and Ruszczyński, A. (2009). *Lectures on Stochastic Programming*. Society for Industrial and Applied Mathematics.

Silver, E. A. (1976). A simple method of determining order quantities in joint replenishments under deterministic demand. *Management Science*, 22(12):1351–1361.

Snyder, L. V., Scaparra, M. P., Daskin, M. S., and Church, R. L. (2006). *Planning for Disruptions in Supply Chain Networks*, chapter 9, pages 234–257.

Stevenson, M. and Spring, M. (2007). Flexibility from a supply chain perspective: definition and review. *International Journal of Operations & Production Management*, 27(7):685–713.

Taft, E. W. (1918). The most economical production lot. *Iron Age*, 101(18):1410–1412.

Tarim, S. and Kingsman, B. G. (2004). The stochastic dynamic production/inventory lot-sizing problem with service-level constraints. *International Journal of Production Economics*, 88(1):105 – 119.

Taş, D., Gendreau, M., Jabali, O., and Jans, R. (2018). A capacitated lot sizing problem with stochastic setup times and overtime. *European Journal of Operational Research*.

Tempelmeier, H. (2013). *Stochastic Lot Sizing Problems*, pages 313–344. Springer New York.

Tomlin, B. and Wang, Y. (2005). On the value of mix flexibility and dual sourcing in unreliable newsvendor networks. *Manufacturing & Service Operations Management*, 7(1):37–57.

Tuy, H. (1964). Concave programming under linear constraints. *Soviet Mathematics*, 5:1437–1440.

van Hoesel, C. P. M. and Wagelmans, A. P. M. (2001). Fully polynomial approximation schemes for single-item capacitated economic lot-sizing problems. *Math. Oper. Res.*, 26(2):339–357.

van Houtum, G.-J. and Zijm, W. (2000). *On the relation between cost and service models for general inventory systems*. Number 1522 in Memorandum / Faculty of Mathematical Sciences. Department of Applied Mathematics, University of Twente. Imported from MEMORANDA.

Vershynin, R. (2010). Introduction to the non-asymptotic analysis of random matrices. *arXiv preprint arXiv:1011.3027*.

Vrat, P. and Padmanabhan, G. (1990). An inventory model under inflation for stock dependent consumption rate items. *Engineering Costs and Production Economics*, 19(1):379 – 383.

Wagelmans, A., Van Hoesel, S., and Kolen, A. (1992). Economic lot sizing: An o(n log n) algorithm that runs in linear time in the wagner-whitin case. *Operations Research*, 40:S145–S156.

Wagner, H. M. and Whitin, T. M. (1958). Dynamic version of the economic lot size model. *Management Science*, 5(1):89–96.

Yano, C. A. and Lee, H. L. (1995). Lot sizing with random yields: A review. *Operations Research*, 43(2):311–334.

Yao, M. and Minner, S. (2017). Review of multi-supplier inventory models in supply chain management: An update. Available at http://dx.doi.org/10.2139/ssrn.2995134.

Zangwill, W. I. (1969). A backlogging model and a multi-echelon model of a dynamic economic lot size production system-a network approach. *Management Science*, 15(9):506–527.

Ziegler, H. (1982). Solving certain singly constrained convex optimization problems in production planning. *Operations Research Letters*, 1(6):246–252.

Ziukov, S. (2015). A literature review on models of inventory management under uncertainty. *Business Systems & Economics*, 5:26–35.

Appendix Part

A Probabilistic model for scenarios generation

A.1 Reminders on Gamma and Dirichlet distributions

A.1.1 Gamma distributions

The *gamma distribution* denoted by $\Gamma(k,\theta)$ is a family of continuous two-parameter probability distributions. Its definition can be found in Delmas and Jourdain (2006, Appendix A). It is parametrized by a shape parameter $k > 0$ and a scale parameter $\theta > 0$. Its support is $\mathbb{R}_+$.

The gamma distribution has a probability density function with respect to the Lebesgue measure on $\mathbb{R}_+$ given by

$$f(x,k,\theta) = \frac{x^{k-1}e^{-\frac{x}{\theta}}}{\theta^k \Gamma(k)} \tag{A.1}$$

where the gamma function is expressed for any $z \in \mathbb{R}_+^*$ by

$$\Gamma(z) = \int_0^\infty x^{z-1} e^{-x} \, \mathrm{d}x. \tag{A.2}$$

A.1.2 Dirichlet distributions

We refer to Kotz et al. (2000, Chapter 49) for the details of the proofs of classical results about Dirichlet distributions.

Definition

The *Dirichlet distribution* denoted by $\mathrm{Dir}(\alpha)$ is a family of continuous multivariate probability distributions. It is parametrized by a number $K \geqslant 2$ of categories and a vector $\alpha = (\alpha_1,\dots,\alpha_K) \in \left(\mathbb{R}_+^*\right)^K$ of positive real numbers called concentration parameters. Its support is the $(K-1)$-dimensional simplex $\left\{(x_1,\dots,x_K) \in \mathbb{R}_+^K \;\middle|\; \sum_{k=1}^K x_k = 1\right\}$.

The Dirichlet distribution has a probability density function with respect to the Lebesgue measure on the Euclidean space $\mathbb{R}^{K-1}$ given by

$$f(x,\alpha) = \frac{1}{\mathrm{B}(\alpha)} \prod_{k=1}^K x_k^{\alpha_k - 1}. \tag{A.3}$$

The normalizing constant is the multivariate beta function, which can be expressed in terms of the gamma function

$$\mathrm{B}(\alpha) = \frac{\prod_{k=1}^{K} \Gamma(\alpha_k)}{\Gamma\left(\sum_{k=1}^{K} \alpha_k\right)}. \tag{A.4}$$

Moments

Let $\boldsymbol{x} = (\boldsymbol{x}_1, \dots, \boldsymbol{x}_K) \sim \mathrm{Dir}(\alpha)$. Setting $\alpha_0 = \sum_{k=1}^{K} \alpha_k$, the moments of the Dirichlet distribution are

$$\mathbb{E}[\boldsymbol{x}_k] = \frac{\alpha_k}{\alpha_0} \qquad \forall k \in [K], \tag{A.5a}$$

$$\mathrm{Var}[\boldsymbol{x}_k] = \frac{\alpha_k(\alpha_0 - \alpha_k)}{\alpha_0^2(\alpha_0 + 1)} \qquad \forall k \in [K], \tag{A.5b}$$

$$\mathrm{Cov}[\boldsymbol{x}_k, \boldsymbol{x}_\ell] = \frac{-\alpha_k \alpha_\ell}{\alpha_0^2(\alpha_0 + 1)} \qquad \forall k \neq \ell. \tag{A.5c}$$

Aggregation property

The Dirichlet distribution has the aggregation property given by Proposition 33.

Proposition 33. *Let $\boldsymbol{x} = (\boldsymbol{x}_1, \dots, \boldsymbol{x}_K) \sim \mathrm{Dir}(\alpha_1, \dots, \alpha_K)$. Suppose that the random variables with indices k and ℓ are dropped from the vector and replaced by their sum. Then, this new random vector satisfies*

$$\boldsymbol{x} = (\boldsymbol{x}_1, \dots, \boldsymbol{x}_k + \boldsymbol{x}_\ell, \dots, \boldsymbol{x}_K) \sim \mathrm{Dir}(\alpha_1, \dots, \alpha_k + \alpha_\ell, \dots, \alpha_K). \tag{A.6}$$

Conditional Dirichlet distribution

The distribution of part of the variables of a Dirichlet conditionally to the other variables is a scaled Dirichlet distribution.

Proposition 34. *Let $\boldsymbol{x} = (\boldsymbol{x}_1, \dots, \boldsymbol{x}_K) \sim \mathrm{Dir}(\alpha_1, \dots, \alpha_K)$ and $m \in \{1, \dots, k-1\}$. We denote $\boldsymbol{x}_{(1)} = (\boldsymbol{x}_1, \dots, \boldsymbol{x}_m)$ and $\boldsymbol{x}_{(2)} = (\boldsymbol{x}_{m+1}, \dots, \boldsymbol{x}_K)$. Then, we have*

$$\frac{1}{1 - \sum_{i=1}^{m} x_i} \left(\boldsymbol{x}_{(2)} | \boldsymbol{x}_{(1)} = x_{(1)}\right) \sim \mathrm{Dir}(\alpha_{m+1}, \dots, \alpha_K). \tag{A.7}$$

Generating Dirichlet distributed variables

Realizations of the random vector $\boldsymbol{x} = (\boldsymbol{x}_1, \dots, \boldsymbol{x}_K) \sim \mathrm{Dir}(\alpha_1, \dots, \alpha_K)$ can be obtained from a source of Gamma distributed random variables using Algorithm 3 described in Frigyik et al. (2010).

Proposition 35. *For any $\alpha_1, \dots, \alpha_K$ positive, Algorithm 3 samples the Dirichlet distribution* $\mathrm{Dir}(\alpha)$.

Other methods exist to generate random variables following a Dirichlet distribution. But this one is easy to understand and to implement because the C++ standard library has a gamma generator since the 2011 edition.

Algorithm 3: Generator of Dirichlet distribution $\mathrm{Dir}(\alpha_1, \dots, \alpha_K)$

Input: $\alpha_1, \dots, \alpha_K$

Output: Random realization $(x_1, \dots, x_K)$ of $\mathrm{Dir}(\alpha_1, \dots, \alpha_K)$

for $k = 1, \dots, K$ **do**

  draw a number y_k from Gamma distribution $\Gamma(\alpha_k, 1)$

end

for $k = 1, \dots, K$ **do** $x_k := \dfrac{y_k}{\sum_{\ell=1}^K y_\ell}$

Return $(\alpha_1, \dots, \alpha_K)$.

A.2 Proofs of Section 7.3.2

In this section, we prove Proposition 23 and that Algorithm 1 gives a vector with an "expanded Dirichlet" distribution.

Proof of Proposition 23. Let γ be a real number in $(0,1)$ and let $\tilde{d}_1, \dots, \tilde{d}_K$ be K positive numbers (with $K \geqslant 2$). Let $(\boldsymbol{d}_1, \dots, \boldsymbol{d}_K)$ be a random variable with an "expanded Dirichlet" distribution $\mathcal{D}(\gamma, \tilde{d}_1, \dots, \tilde{d}_K)$. We recall that $\tilde{d}_0 = \sum_{k=1}^K \tilde{d}_k$, that $\alpha_0 = \frac{1}{\gamma^2} - 1$ and that $\alpha_k = \frac{\tilde{d}_k}{\tilde{d}_0}\alpha_0$, for each $k \in [K]$.

Item 1. Since $\frac{1}{\tilde{d}_0}(\boldsymbol{d}_1, \dots, \boldsymbol{d}_K)$ has a Dirichlet distribution, we have for each $k \in [K]$

$$\begin{aligned}
\mathbb{E}[\boldsymbol{d}_k] &= \tilde{d}_0 \mathbb{E}\left[\frac{\boldsymbol{d}_k}{\tilde{d}_0}\right] = \tilde{d}_0 \frac{\alpha_k}{\alpha_0} && \text{(expectation of Dirichlet distribution)} \quad \text{(A.8a)} \\
&= \tilde{d}_k. && \text{(definition of } \alpha_k) \quad \text{(A.8b)}
\end{aligned}$$

Item 2. Since $\frac{1}{\tilde{d}_0}(\boldsymbol{d}_1, \dots, \boldsymbol{d}_K)$ has a Dirichlet distribution, we have for each $k \in [K]$

$$\begin{aligned}
\mathrm{Var}[\boldsymbol{d}_k] &= \tilde{d}_0^2 \mathrm{Var}\left[\frac{\boldsymbol{d}_k}{\tilde{d}_0}\right] = \tilde{d}_0^2 \frac{\alpha_k(\alpha_0 - \alpha_k)}{\alpha_0^2(\alpha_0 + 1)} && \text{(variance of Dirichlet distribution)} \quad \text{(A.9a)} \\
&= \tilde{d}_k^2 \left(\frac{\alpha_0}{\alpha_k} - 1\right) \frac{1}{\alpha_0 + 1} && \text{(definition of } \alpha_k) \quad \text{(A.9b)} \\
&= \tilde{d}_k^2 \left(\frac{\tilde{d}_0}{\tilde{d}_k} - 1\right) \gamma^2 && \text{(definition of } \gamma) \quad \text{(A.9c)} \\
&= \gamma^2 \tilde{d}_k (\tilde{d}_0 - \tilde{d}_k). && \text{(A.9d)}
\end{aligned}$$

Item 3. Since $\frac{1}{\tilde{d}_0}(\boldsymbol{d}_1, \dots, \boldsymbol{d}_K)$ has a Dirichlet distribution, we have $\sum_{k=1}^K \frac{\boldsymbol{d}_k}{\tilde{d}_0} = 1$ almost surely and the result follows.

Item 4. Let k be an integer in $\{1, \dots, K-1\}$ and let $d_{(1)}$ be a vector of k positive real numbers such that $\sum_{i=1}^k d_i < \tilde{d}_0$. We denote $\boldsymbol{d}_{(1)} = (\boldsymbol{d}_1, \dots, \boldsymbol{d}_k)$ and $\boldsymbol{d}_{(2)} = (\boldsymbol{d}_{k+1}, \dots, \boldsymbol{d}_K)$.

Since $\frac{1}{\tilde{d}_0}\big(\boldsymbol{d}_1,\ldots,\boldsymbol{d}_K\big)$ has a Dirichlet distribution and thanks to Proposition 34, we have

$$\frac{1}{1-\frac{\sum_{i=1}^k d_i}{\tilde{d}_0}}\left(\frac{\boldsymbol{d}_{(2)}}{\tilde{d}_0}\middle|\frac{\boldsymbol{d}_{(1)}}{\tilde{d}_0}=\frac{d_{(1)}}{\tilde{d}_0}\right)\sim \mathrm{Dir}\,(\alpha_{m+1},\ldots,\alpha_K) \tag{A.10}$$

and then

$$\frac{1}{\tilde{d}_0-\sum_{i=1}^k d_i}\big(\boldsymbol{d}_{(2)}\big|\boldsymbol{d}_{(1)}=d_{(1)}\big)\sim \mathrm{Dir}\,(\alpha_{m+1},\ldots,\alpha_K). \tag{A.11}$$

Wet set γ' as the unique positive solution of Equation (7.7), i.e., $\gamma'=\sqrt{\frac{\gamma^2}{(1-a)+a\gamma^2}}$ with $a=\frac{\sum_{i=1}^k d_i}{\tilde{d}_0}$. The parameter γ' is well-defined and belong to $(0,1)$ since $\gamma^2, a\in(0,1)$. Thus, the random vector $\left(\boldsymbol{d}_{(2)}\middle|\boldsymbol{d}_{(1)}=d_{(1)}\right)$ follows an "expanded Dirichlet" distribution $\mathcal{D}\big(\gamma',\tilde{d}_{m+1},\ldots,\tilde{d}_K\big)$. Indeed, we have

$$\tilde{d}_0'=\sum_{k=m+1}^K \tilde{d}_k=\tilde{d}_0-\sum_{i=1}^k d_i. \tag{A.12}$$

Moreover,

$$\alpha_0'=\left(\frac{1}{\gamma'^2}-1\right)=\left(\frac{1}{\gamma^2}-1\right)\left(1-\frac{\sum_{i=1}^k d_i}{\tilde{d}_0}\right)=\left(\frac{1}{\gamma^2}-1\right)\frac{\tilde{d}_0'}{\tilde{d}_0}=\frac{\tilde{d}_0'}{\tilde{d}_0}\alpha_0 \tag{A.13}$$

and

$$\alpha_k'=\frac{\tilde{d}_k}{\tilde{d}_0'}\alpha_0'=\alpha_k \quad \forall k\in\{m+1,\ldots,K\}. \tag{A.14}$$

□

We now prove that Algorithm 1 samples the "expanded Dirichlet" distribution.

Proof of Proposition 24. Algorithm 3 is a subroutine of Algorithm 1 and Proposition 35 gives that, if $y_k\sim\Gamma(\alpha_k,1)$, then the random vector $\frac{1}{\sum_{k=1}^K y_k}\big(y_1,\ldots,y_K\big)\sim \mathrm{Dir}\,(\alpha_1,\ldots,\alpha_K)$. Thus, the last step produces a vector d such that $\frac{1}{\tilde{d}_0}(d_1,\ldots,d_K)\sim\mathrm{Dir}\,(\alpha_1,\ldots,\alpha_K)$ with $\tilde{d}_0=\sum_{k=1}^K\tilde{d}_k$ and $\alpha_0=\frac{1}{\gamma^2}-1$ and $\alpha_k=\frac{\tilde{d}_k}{\tilde{d}_0}\alpha_0$, for $k\in[K]$. So Algorithm 1 returns a vector with an "expanded Dirichlet" distribution $\mathcal{D}\big(\gamma,\tilde{d}_1,\ldots,\tilde{d}_K\big)$. □

A.3 Fitting the parameters to the probabilistic model

When proposing a probabilistic model for demand distribution, we face many objectives. First it is desirable that the demand expectation be equal to its historical forecast, which is ensured by Item 1 of Proposition 23. Second, a main hypothesis of the model is that the total sum of demand does not depend of the realization. The chosen distribution ensures this properties thanks to Item 3. Last, every company does not have the same uncertainty on forecast demand. For example, some companies may have more historical data to rely on to produce forecast or some sectors may rely on firm command whereas other may not. Thus our model should be able to take into account this reliability.

For the sake of simplicity, we choose a unique parameter v which represents the ratio between standard deviation and expectation of a random variable and which can be easily interpreted by companies. For $(\boldsymbol{d}_1, \ldots, \boldsymbol{d}_K) \sim \mathcal{D}(\gamma, \tilde{d}_1, \ldots, \tilde{d}_K)$, we decide to look for $\gamma \in (0,1)$ which minimizes

$$\gamma \mapsto \left\| \left(\frac{\sigma_1(\gamma)}{e_1(\gamma)}, \ldots, \frac{\sigma_K(\gamma)}{e_K(\gamma)} \right) - (v, \ldots, v) \right\|_2 \tag{A.15}$$

where $e_k(\gamma) = \mathbb{E}[\boldsymbol{d}_k]$ and $\sigma_k(\gamma) = \sqrt{\mathrm{Var}[\boldsymbol{d}_k]}$. Note that the norm $\|\,.\,\|_2$ has been chosen. An interesting feature of $\|\,.\,\|_2$ is the closed-form formula for the value of γ (see Proposition 36). If $\|\,.\,\|_1$ or $\|\,.\,\|_\infty$ had been chosen, finding the best γ would be done with classical linearization methods and the use of a linear solver.

Proposition 36. *For any $v \in (0,1)$ and any $\tilde{d}_1, \ldots, \tilde{d}_K$ positive,*

$$\gamma = v \frac{\sum_{k=1}^K \sqrt{\frac{\tilde{d}_0}{\bar{d}_k} - 1}}{\sum_{k=1}^K \left(\frac{\tilde{d}_0}{\bar{d}_k} - 1 \right)} \tag{A.16}$$

minimize the quantity $\left\| \left(\frac{\sigma_1(\gamma)}{e_1(\gamma)}, \ldots, \frac{\sigma_K(\gamma)}{e_K(\gamma)} \right) - (v, \ldots, v) \right\|_2$ where $e_k(\gamma) = \mathbb{E}[\boldsymbol{d}_k]$ and $\sigma_k = \sqrt{\mathrm{Var}[\boldsymbol{d}_k]}$.

Proof. Let v be a real number in $(0,1)$ and $\tilde{d}_1, \ldots, \tilde{d}_K$ be positive real numbers. For each index k, Proposition 23 gives

$$\frac{\sigma_k}{e_k} = \frac{\sqrt{\mathrm{Var}[\boldsymbol{d}_k]}}{\mathbb{E}[\boldsymbol{d}_k]} = \gamma \sqrt{\frac{\tilde{d}_0}{\bar{d}_k} - 1}. \tag{A.17}$$

Then, finding $\gamma' \in (0,1)$ which minimizes $\left\| \left(\frac{\sigma_1(\gamma)}{e_1(\gamma)}, \ldots, \frac{\sigma_K(\gamma)}{e_K(\gamma)} \right) - (v, \ldots, v) \right\|_2$ is equivalent to look for the minimizer on $(0,1)$ of the function

$$f(\gamma') = \sum_{k=1}^K \left(\gamma' \sqrt{\frac{\tilde{d}_0}{\bar{d}_k} - 1} - v \right)^2 \tag{A.18a}$$

$$= \sum_{k=1}^K \left(\frac{\tilde{d}_0}{\bar{d}_k} - 1 \right) \gamma'^2 - 2v \left(\sum_{k=1}^K \sqrt{\frac{\tilde{d}_0}{\bar{d}_k} - 1} \right) \gamma' + K v^2 \tag{A.18b}$$

f is a quadratic function and reach its minimum for

$$\gamma = v \frac{\sum_{k=1}^K \sqrt{\frac{\tilde{d}_0}{\bar{d}_k} - 1}}{\sum_{k=1}^K \left(\frac{\tilde{d}_0}{\bar{d}_k} - 1 \right)} \tag{A.19}$$

$\left(\frac{d_1}{\tilde{d}_0}, \ldots, \frac{d_K}{\tilde{d}_0} \right)$ being in the interior of the simplex, according to Lemma 37, we have

$$0 < \frac{\sum_{k=1}^K \sqrt{\frac{\tilde{d}_0}{\bar{d}_k} - 1}}{\sum_{k=1}^K \left(\frac{\tilde{d}_0}{\bar{d}_k} - 1 \right)} \leqslant \frac{1}{\sqrt{K-1}}. \tag{A.20}$$

Since $v \in (0,1)$, we have $\gamma \in (0,1)$ and the result follows. $\square$

Lemma 37. *Let K be an integer non-smaller than 2. We define*

$$\delta^{K-1} = \left\{ (x_1, \ldots, x_K) \in (0,1)^K \,\middle|\, \sum_{k=1}^{K} x_k = 1 \right\} \tag{A.21}$$

and

$$\begin{array}{rccc} g: & \delta^{K-1} & \to & \mathbb{R}_+ \\ & x & \mapsto & \frac{\sum_{k=1}^{K} \sqrt{\frac{1}{x_k} - 1}}{\sum_{k=1}^{K} \left(\frac{1}{x_k} - 1\right)} \end{array} \tag{A.22}$$

Then, g reaches its supremum, which is $\frac{1}{\sqrt{K-1}}$

Proof. For each index $k \in [K]$, we set

$$y_k(x) = \frac{\frac{1}{x_k} - 1}{\lambda(x)} \quad \text{with} \quad \lambda(x) = \sum_{\ell=1}^{K} \left(\frac{1}{x_\ell} - 1 \right), \quad \text{and} \quad \mu(y) = \sum_{\ell=1}^{K} \sqrt{y_\ell}. \tag{A.23}$$

Then we get

$$g(x) = \frac{\sum_{k=1}^{K} \sqrt{\lambda(x) y_k(x)}}{\lambda(x)} = \frac{\mu(y(x))}{\sqrt{\lambda(x)}} \leqslant \frac{\sup_{y \in \delta^{K-1}} \mu(y)}{\sqrt{\inf_{x \in \delta^{K-1}} \lambda(x)}}. \tag{A.24}$$

The map μ being concave and symmetric (i.e., its value does not depend of the order of its arguments), we get

$$\sup_{y \in \delta^{K-1}} \mu(y) = \mu\left(\frac{1}{K}, \ldots, \frac{1}{K} \right) = \sqrt{K}. \tag{A.25}$$

λ being convex and symmetric, we get

$$\inf_{x \in \delta^{K-1}} \lambda(x) = \lambda\left(\frac{1}{K}, \ldots, \frac{1}{K} \right) = K(K-1). \tag{A.26}$$

Then, we have for each $x \in \delta^{K-1}$

$$g(x) \leqslant \frac{1}{\sqrt{K-1}}. \tag{A.27}$$

Since $g\left(\frac{1}{K}, \ldots, \frac{1}{K}\right) = \frac{1}{\sqrt{K-1}}$, the result follows. $\square$

B Complete computational experiments on discrete-time inventory models

In this chapter, we give the complete results of the simulations made in Chapter 7. The characteristics of each line is given in Section 7.3.1 and the meaning of each column of Tables B.1 to B.7 as well as the way of computing the KPI's are given in Section 7.5.

Input			Obj.	LB	Inventory			Service		Work-	Flex.
Vol.	Heur.	β			Av.	Max	Cover	Fill rate	Cycle	load	
volatility $v = 20\%$	lot	85%	6668	158	512	1011	5.3	79%	83%	89%	96%
		95%	7644	237	576	1279	5.6	81%	84%	93%	98%
		98%	8481	638	603	1397	5.6	81%	85%	95%	100%
	cover	85%	5277	158	404	550	4.1	87%	85%	79%	88%
		95%	5654	237	419	562	4.1	87%	85%	79%	89%
		98%	6265	638	427	575	4.2	87%	85%	80%	89%
	det	85%	310	158	24	137	3.2	68%	57%	73%	100%
		95%	476	237	25	136	2.8	72%	59%	57%	100%
		98%	1331	638	39	150	3.4	78%	65%	73%	100%
	sto	85%	172	158	13	126	3.3	59%	51%	72%	100%
		95%	400	237	16	128	2.8	67%	55%	56%	100%
		98%	1345	638	38	147	3.5	76%	63%	73%	100%
volatility $v = 50\%$	lot	85%	6949	200	528	1000	5.1	81%	85%	89%	96%
		95%	8237	655	584	1235	5.2	81%	86%	92%	99%
		98%	9841	1109	621	1359	5.2	82%	87%	94%	100%
	cover	85%	5560	200	416	562	4.2	86%	86%	79%	88%
		95%	6530	655	441	589	4.2	86%	86%	80%	88%
		98%	8000	1109	450	602	4.3	86%	86%	80%	88%
	det	85%	659	200	46	164	2.8	68%	59%	58%	100%
		95%	1679	655	63	172	3.5	75%	66%	73%	100%
		98%	3147	1109	129	244	3.8	86%	74%	74%	100%
	sto	85%	324	200	16	133	2.8	55%	52%	55%	98%
		95%	1616	655	50	159	3.4	72%	62%	73%	100%
		98%	3188	1109	139	275	4.0	85%	74%	74%	100%
volatility $v = 80\%$	lot	85%	7565	314	566	1048	5.7	82%	86%	89%	96%
		95%	9127	973	621	1258	5.8	82%	87%	93%	98%
		98%	11296	1330	660	1376	5.9	83%	87%	94%	99%
	cover	85%	6056	314	440	580	4.4	83%	84%	79%	88%
		95%	7480	973	460	608	4.4	83%	85%	80%	89%
		98%	9958	1330	468	628	4.5	83%	85%	80%	89%
	det	85%	1179	314	71	179	3.3	71%	63%	61%	100%
		95%	2847	973	118	214	4.3	79%	71%	74%	100%
		98%	4648	1330	211	399	4.6	88%	80%	75%	100%
	sto	85%	743	314	24	135	3.0	55%	55%	59%	97%
		95%	2710	973	96	207	3.6	74%	68%	74%	100%
		98%	4610	1330	209	415	4.0	86%	79%	75%	100%

Table B.1 – Results for L_0

Vol.	Input Heur.	β	Obj.	LB	Inventory Av.	Inventory Max	Inventory Cover	Service Fill rate	Service Cycle	Work-load	Flex.
volatility $v = 20\%$	lot	85%	12320	495	948	1125	7.6	94%	95%	62%	92%
		95%	14087	504	1080	1353	7.2	94%	95%	64%	94%
		98%	15516	560	1161	1520	7.2	94%	95%	65%	96%
	cover	85%	10719	495	824	1024	4.4	100%	100%	58%	95%
		95%	11833	504	910	1051	4.3	100%	100%	59%	98%
		98%	12397	560	953	1075	4.3	100%	100%	60%	99%
	det	85%	709	495	54	313	2.8	78%	62%	51%	99%
		95%	839	504	54	313	2.8	81%	63%	44%	99%
		98%	1485	560	69	313	3.0	86%	68%	51%	99%
	sto	85%	506	495	39	300	2.6	68%	55%	49%	96%
		95%	725	504	42	300	2.6	77%	59%	43%	97%
		98%	1451	560	62	305	3.0	84%	67%	51%	98%
volatility $v = 50\%$	lot	85%	12316	518	946	1142	8.9	93%	94%	61%	92%
		95%	14450	634	1082	1353	8.6	94%	95%	64%	94%
		98%	16454	759	1171	1515	8.8	94%	95%	65%	95%
	cover	85%	11083	518	851	1045	4.7	99%	99%	58%	94%
		95%	12403	634	948	1092	4.6	99%	99%	59%	98%
		98%	13148	759	996	1137	4.6	99%	99%	60%	98%
	det	85%	1089	518	79	340	3.1	76%	64%	40%	98%
		95%	2006	634	102	340	3.4	84%	70%	51%	99%
		98%	3179	759	154	346	3.7	90%	78%	52%	100%
	sto	85%	696	518	43	307	2.7	62%	56%	37%	95%
		95%	1869	634	78	318	3.0	79%	66%	51%	98%
		98%	3140	759	151	345	3.5	90%	77%	52%	99%
volatility $v = 80\%$	lot	85%	12597	629	960	1160	48.1	92%	93%	61%	91%
		95%	15173	1081	1091	1381	46.8	93%	94%	64%	94%
		98%	18567	1330	1196	1556	47.6	93%	94%	65%	95%
	cover	85%	11068	629	848	1054	12.1	99%	97%	58%	94%
		95%	12497	1081	943	1100	12.1	99%	97%	59%	97%
		98%	13525	1330	992	1158	12.1	99%	97%	60%	99%
	det	85%	1743	629	114	360	8.5	77%	69%	48%	98%
		95%	3290	1081	163	361	9.0	84%	75%	52%	99%
		98%	5059	1330	213	365	9.2	88%	80%	53%	99%
	sto	85%	1185	629	55	323	2.9	62%	60%	47%	94%
		95%	3123	1081	126	342	5.9	79%	71%	52%	98%
		98%	4947	1330	220	387	7.3	88%	81%	53%	100%

Table B.2 – Results for L_1

Input			Obj.	LB	Inventory			Service		Work-load	Flex.
Vol.	Heur.	β			Av.	Max	Cover	Fill rate	Cycle		
volatility $v = 20\%$	lot	85%	7764	241	597	796	6.2	93%	95%	77%	92%
		95%	9243	299	708	991	6.4	93%	95%	80%	92%
		98%	9998	589	757	1133	6.5	94%	95%	82%	93%
	cover	85%	6784	241	521	666	4.6	94%	96%	71%	89%
		95%	7418	299	566	720	4.5	94%	96%	73%	91%
		98%	7983	589	597	784	4.5	94%	96%	75%	92%
	det	85%	415	241	32	180	3.1	75%	64%	61%	100%
		95%	514	299	32	180	3.0	78%	67%	59%	100%
		98%	1095	589	39	190	3.3	82%	70%	63%	100%
	sto	85%	252	241	19	173	2.9	63%	58%	60%	97%
		95%	419	299	22	175	2.8	73%	63%	58%	99%
		98%	1076	589	37	190	3.3	82%	70%	63%	100%
volatility $v = 50\%$	lot	85%	7669	273	588	783	7.0	93%	95%	76%	91%
		95%	9280	579	701	994	7.1	93%	95%	79%	92%
		98%	10257	813	748	1119	7.2	94%	95%	82%	94%
	cover	85%	6868	273	523	666	5.0	92%	94%	71%	89%
		95%	7820	579	571	732	4.9	92%	95%	73%	91%
		98%	8928	813	594	761	4.9	92%	95%	74%	92%
	det	85%	760	273	55	199	3.2	73%	66%	45%	100%
		95%	1481	579	65	205	3.6	80%	71%	63%	100%
		98%	2622	813	121	262	4.2	89%	79%	65%	100%
	sto	85%	387	273	22	183	2.5	58%	59%	42%	96%
		95%	1367	579	48	204	3.3	75%	69%	63%	100%
		98%	2553	813	116	279	4.1	88%	79%	64%	100%
volatility $v = 80\%$	lot	85%	7999	371	610	840	9.9	93%	94%	77%	91%
		95%	9633	782	713	1003	10.0	93%	95%	80%	92%
		98%	10780	1023	751	1129	10.1	93%	95%	82%	95%
	cover	85%	7265	371	545	698	6.8	91%	92%	72%	88%
		95%	8532	782	594	771	6.8	91%	93%	73%	91%
		98%	10221	1023	622	823	6.8	91%	93%	74%	91%
	det	85%	1316	371	86	215	4.5	75%	71%	62%	100%
		95%	2570	782	116	230	5.3	82%	77%	64%	100%
		98%	4114	1023	186	370	6.3	89%	83%	65%	100%
	sto	85%	739	371	29	195	3.3	56%	63%	60%	95%
		95%	2352	782	88	230	4.5	75%	73%	64%	100%
		98%	3924	1023	181	379	5.4	87%	83%	65%	100%

Table B.3 – Results for L_2

Input			Obj.	LB	Inventory			Service		Work-	Flex.
Vol.	Heur.	β			Av.	Max	Cover	Fill rate	Cycle	load	
volatility $v = 20\%$	lot	85%	3991	78	307	392	4.9	97%	97%	64%	91%
		95%	4492	119	345	434	5.0	98%	97%	66%	90%
		98%	4850	244	370	460	5.1	98%	98%	67%	91%
	cover	85%	3391	78	261	345	4.2	100%	99%	66%	92%
		95%	3755	119	289	384	4.2	100%	99%	68%	94%
		98%	4088	244	314	416	4.1	100%	99%	68%	96%
	det	85%	135	78	10	60	3.6	65%	53%	59%	97%
		95%	229	119	10	60	2.8	67%	57%	40%	97%
		98%	562	244	21	61	3.2	79%	66%	61%	97%
	sto	85%	84	78	6	57	3.5	57%	47%	57%	95%
		95%	215	119	8	58	2.8	65%	55%	39%	97%
		98%	553	244	19	59	3.4	78%	63%	60%	97%
volatility $v = 50\%$	lot	85%	4008	105	308	400	6.5	97%	95%	64%	89%
		95%	4481	290	341	436	6.5	97%	96%	66%	90%
		98%	4883	392	363	466	6.6	97%	96%	67%	91%
	cover	85%	3321	105 s	255	351	5.4	99%	97%	66%	93%
		95%	3804	290	292	387	5.4	99%	98%	68%	94%
		98%	4218	392	321	425	5.5	99%	97%	68%	94%
	det	85%	285	105	18	66	4.2	66%	58%	38%	97%
		95%	741	290	31	71	4.5	77%	66%	61%	98%
		98%	1285	392	60	96	4.6	89%	77%	61%	100%
	sto	85%	187	105	8	60	2.8	53%	51%	37%	95%
		95%	707	290	25	64	4.0	73%	63%	60%	97%
		98%	1256	392	56	91	4.9	88%	75%	61%	100%
volatility $v = 80\%$	lot	85%	3981	163	305	404	24.3	95%	93%	63%	89%
		95%	4598	358	344	443	24.3	96%	95%	66%	91%
		98%	5188	428	374	487	24.4	97%	96%	68%	91%
	cover	85%	3401	163	260	356	19.7	97%	94%	66%	94%
		95%	3982	358	298	403	19.8	97%	95%	68%	94%
		98%	4543	428	327	440	19.8	98%	96%	68%	94%
	det	85%	524	163	29	73	18.6	66%	62%	51%	97%
		95%	1263	358	60	94	18.8	80%	74%	61%	100%
		98%	2036	428	83	129	19.1	87%	80%	62%	100%
	sto	85%	397	163	13	65	9.6	50%	55%	48%	94%
		95%	1181	358	42	82	10.3	73%	68%	60%	100%
		98%	1964	428	82	129	28.9	86%	80%	62%	100%

Table B.4 – Results for L_3

Input			Obj.	LB	Inventory			Service		Work-	Flex.
Vol.	Heur.	β			Av.	Max	Cover	Fill rate	Cycle	load	
volatility $v = 20\%$	lot	85%	6383	241	491	587	4.6	96%	96%	40%	90%
		95%	8168	252	627	722	5.2	96%	96%	41%	89%
		98%	9284	388	706	853	5.8	96%	96%	43%	91%
	cover	85%	6420	241	494	579	3.2	100%	100%	39%	93%
		95%	8044	252	619	741	3.7	100%	100%	40%	91%
		98%	8807	388	677	789	4.4	100%	100%	41%	91%
	det	85%	339	241	26	206	2.2	77%	62%	34%	98%
		95%	353	252	25	206	2.1	77%	61%	31%	98%
		98%	657	388	31	206	2.8	85%	67%	35%	98%
	sto	85%	248	241	19	201	2.1	72%	56%	34%	98%
		95%	292	252	20	201	2.1	75%	58%	30%	98%
		98%	632	388	27	204	2.7	83%	65%	35%	98%
volatility $v = 50\%$	lot	85%	6377	244	490	587	6.2	95%	96%	40%	91%
		95%	8292	364	630	731	6.8	96%	96%	41%	90%
		98%	9481	409	699	838	7.5	96%	96%	43%	90%
	cover	85%	6322	244	486	592	4.4	100%	99%	38%	92%
		95%	8125	364	625	764	4.9	100%	99%	40%	91%
		98%	8931	409	686	814	5.4	100%	99%	41%	90%
	det	85%	468	244	35	216	2.5	72%	62%	28%	98%
		95%	830	364	44	216	3.7	83%	68%	35%	98%
		98%	1420	409	62	218	3.9	89%	74%	35%	99%
	sto	85%	284	244	20	203	2.5	63%	52%	26%	95%
		95%	755	364	32	208	3.3	80%	65%	34%	98%
		98%	1404	409	63	227	3.9	89%	74%	35%	98%
volatility $v = 80\%$	lot	85%	6474	270	496	596	6.9	95%	95%	40%	91%
		95%	8372	516	625	730	7.4	95%	96%	41%	90%
		98%	9946	658	692	822	8.3	95%	96%	43%	90%
	cover	85%	6110	270	470	601	4.9	98%	97%	38%	93%
		95%	7942	516	609	764	5.2	99%	98%	40%	92%
		98%	8809	658	669	807	5.7	99%	98%	40%	92%
	det	85%	714	270	50	227	2.9	76%	66%	33%	98%
		95%	1364	516	65	228	4.3	83%	71%	35%	98%
		98%	2316	658	105	258	4.3	90%	78%	35%	100%
	sto	85%	416	270	22	209	2.8	62%	57%	32%	95%
		95%	1287	516	50	220	3.5	77%	68%	35%	98%
		98%	2269	658	105	253	4.5	89%	79%	35%	100%

Table B.5 – Results for L_4

Input			Obj.	LB	Inventory			Service		Work-	Flex.
Vol.	Heur.	β			Av.	Max	Cover	Fill rate	Cycle	load	
volatility $v = 20\%$	lot	85%	711	80	55	131	2.3	81%	78%	98%	100%
		95%	749	93	54	131	2.3	81%	78%	97%	100%
		98%	1015	204	54	132	2.3	81%	78%	97%	100%
	cover	85%	373	80	29	78	3.4	69%	59%	90%	78%
		95%	561	93	29	78	3.4	69%	59%	90%	78%
		98%	1233	204	29	78	3.4	69%	59%	90%	78%
	det	85%	126	80	10	70	2.2	74%	59%	95%	95%
		95%	155	93	10	70	2.0	81%	63%	62%	97%
		98%	331	204	12	74	2.1	86%	65%	96%	97%
	sto	85%	82	80	6	69	2.0	66%	52%	93%	97%
		95%	127	93	7	69	1.8	77%	57%	61%	97%
		98%	321	204	11	73	2.1	85%	63%	96%	98%
volatility $v = 50\%$	lot	85%	823	87	61	136	2.4	81%	79%	98%	100%
		95%	1093	192	62	138	2.4	81%	79%	98%	100%
		98%	2005	515	63	139	2.4	81%	79%	98%	100%
	cover	85%	472	87	29	81	3.4	67%	60%	91%	78%
		95%	1213	192	29	81	3.4	67%	60%	91%	78%
		98%	3225	515	29	81	3.4	67%	60%	91%	78%
	det	85%	239	87	17	76	2.0	78%	65%	64%	97%
		95%	433	192	20	82	2.2	83%	68%	94%	99%
		98%	988	515	28	85	2.2	88%	74%	98%	99%
	sto	85%	122	87	7	70	1.7	63%	54%	58%	96%
		95%	400	192	15	79	2.0	79%	63%	94%	99%
		98%	974	515	27	91	2.1	87%	72%	98%	99%
volatility $v = 80\%$	lot	85%	947	108	66	139	2.4	80%	79%	97%	100%
		95%	1532	322	68	145	2.4	80%	78%	98%	100%
		98%	2846	797	68	144	2.4	80%	79%	97%	100%
	cover	85%	763	108	36	86	3.3	65%	61%	91%	77%
		95%	2111	322	36	86	3.3	65%	61%	91%	77%
		98%	4628	797	36	86	3.3	65%	61%	91%	77%
	det	85%	413	108	28	82	2.2	76%	68%	66%	97%
		95%	833	322	34	89	2.3	82%	73%	97%	99%
		98%	1668	797	39	91	2.4	84%	75%	98%	99%
	sto	85%	232	108	10	74	1.7	60%	56%	59%	97%
		95%	760	322	26	89	2.1	78%	68%	96%	99%
		98%	1592	797	36	96	2.2	82%	74%	98%	100%

Table B.6 – Results for L_5

Input			Obj.	LB	Inventory			Service		Work-	Flex.
Vol.	Heur.	β			Av.	Max	Cover	Fill rate	Cycle	load	
volatility $v = 20\%$	lot	85%	1499	140	115	312	3.0	76%	72%	97%	97%
		95%	1729	244	115	312	3.0	76%	72%	97%	97%
		98%	2583	833	117	317	3.0	77%	72%	97%	98%
	cover	85%	1409	140	107	322	3.4	73%	67%	94%	83%
		95%	1753	244	108	322	3.4	73%	67%	94%	83%
		98%	2921	833	108	322	3.4	73%	67%	94%	84%
	det	85%	233	140	18	94	2.6	67%	58%	90%	100%
		95%	400	244	18	94	2.3	72%	61%	91%	100%
		98%	1194	833	46	243	2.3	80%	65%	95%	100%
	sto	85%	145	140	11	86	2.6	59%	51%	87%	99%
		95%	364	244	13	87	2.3	68%	58%	90%	100%
		98%	1170	833	42	235	2.4	79%	64%	95%	100%
volatility $v = 50\%$	lot	85%	1759	181	125	314	3.1	75%	73%	97%	98%
		95%	2692	818	127	315	3.1	75%	73%	98%	98%
		98%	4594	1382	130	322	3.1	76%	73%	98%	98%
	cover	85%	1811	181	122	330	3.6	72%	68%	94%	84%
		95%	3107	818	122	330	3.6	72%	68%	94%	84%
		98%	5750	1382	123	330	3.6	72%	68%	94%	84%
	det	85%	492	181	33	114	2.4	69%	62%	82%	100%
		95%	1429	818	54	216	2.5	77%	67%	95%	100%
		98%	2576	1382	92	319	2.6	84%	73%	97%	100%
	sto	85%	291	181	14	93	2.2	56%	54%	79%	98%
		95%	1365	818	48	233	2.4	74%	64%	94%	100%
		98%	2504	1382	87	318	2.6	83%	72%	97%	100%
volatility $v = 80\%$	lot	85%	2087	367	135	309	3.3	75%	74%	97%	98%
		95%	3594	1187	135	310	3.3	76%	74%	97%	98%
		98%	6624	2083	141	314	3.3	76%	74%	98%	98%
	cover	85%	2300	367	134	323	3.7	71%	70%	93%	84%
		95%	4547	1187	135	323	3.7	71%	70%	93%	84%
		98%	8983	2083	135	323	3.7	71%	70%	93%	84%
	det	85%	992	367	54	137	2.6	68%	65%	92%	100%
		95%	2424	1187	103	313	2.7	78%	72%	97%	100%
		98%	4298	2083	115	318	2.8	80%	74%	98%	100%
	sto	85%	680	367	21	110	2.4	55%	58%	90%	98%
		95%	2262	1187	89	316	2.6	75%	70%	96%	100%
		98%	4063	2083	107	316	2.8	78%	73%	98%	100%

Table B.7 – Results for L_6