

**Eléments de
GESTION DE LA PRODUCTION**

Yves CRAMA

**Ecole d'Administration des Affaires
Université de Liège**

**Notes de cours
Année académique 2002-2003**

TABLE DES MATIERES

	Page
CHAPITRE 1: Introduction	1
1.1. Définitions et concepts de base	1
1.2. Gestion hiérarchique de la production	2
1.3. Développement historique de la GP: Quelques jalons ...	3
1.4. Relations avec les autres fonctions de l'entreprise	3
Bibliographie	4
CHAPITRE 2: Gestion stratégique de la production	5
2.1. Généralités	5
2.2. Impact sur la fonction production	5
2.3. Le rôle stratégique de la gestion de la production	6
2.4. Gestion de la qualité	8
Bibliographie	11
Annexes: Is Boeing out of its spin?	
Building customer partnerships as a competitive weapon ?	
Ford obligé d'organiser le rappel de 13 millions de pneus	
ISO9000 pour responsables affaires	
CHAPITRE 3: Planification à long terme: Les options fondamentales	13
Bibliographie	14
Annexe: How manufacturing can make low-tech products high-tech	
CHAPITRE 4: Planification à long terme: Choix du processus de production	15
4.1. Généralités	15
4.2. Classification des processus de production en fonction des flux de produits	15
4.3. Cycle de vie du produit et organisation de la production	17
4.4. Cellules de fabrication et ateliers flexibles: entre job-shops et lignes de production	19
4.5. Layout (implantation) des centres de production	21
4.5.1. Equilibrage des chaînes d'assemblage	21
4.5.2. Job-shop: Localisation des départements	25
Bibliographie	26
Annexe: Equilibrage d'une chaîne d'assemblage	
CHAPITRE 5: Planification à moyen terme: Le plan agrégé de production	28
5.1. Introduction et concepts de base	28
5.2. Analyse graphique des options simples	31
5.3. Un modèle de programmation linéaire	34
Bibliographie	39
Annexe: Application à une usine de peinture (extrait de Holt, Modigliani, Muth et Simon 1964)	

CHAPITRE 6: Planification à court terme: Le plan directeur de production	40
6.1. Généralités	40
6.2. Schéma d'élaboration du PDP	40
6.3. Commentaires sur l'élaboration du PDP	43
6.4. Le contenu du plan directeur de production	46
Bibliographie	44
CHAPITRE 7: Planification à court terme: La problématique des stocks	48
7.1. Généralités	48
7.2. Objectif de la gestion des stocks	48
7.3. Mesures de l'état des stocks	49
7.4. Le rôle des stocks	49
7.5. La gestion Juste-A-Temps	50
Bibliographie	52
Annexe: Révolte nipponne: Le just-in-time contesté	
CHAPITRE 8: Planification à court terme: La planification des besoins en composants	53
8.1. Exemple introductif	53
8.2. La logique MRP	58
8.3. Systèmes d'information utilisés par les logiciels MRP	59
8.4. Dynamique du système MRP	60
8.5. Techniques de détermination des lots (lot sizing)	60
8.6. Planification des besoins en capacité (CRP) - Calcul des courbes de charge	66
8.7. Epilogue	68
Bibliographie	69
CHAPITRE 9: Planification à court terme: La gestion des stocks pour les articles à demande indépendante	70
9.1. Demande indépendante	70
9.2. Politiques de gestion des stocks: Les choix à effectuer	70
9.3. Politiques de gestion des stocks: Les réponses possibles	72
9.4. Classification ABC	73
9.5. Quelques modèles de gestion des stocks	75
9.5.1. Quantité économique de commande	75
9.5.2. Délai d'obtention	76
9.5.3. Demande aléatoire	77
9.5.4. Modèle du marchand de journaux	79
Bibliographie	80
CHAPITRE 10: Ordonnancement et pilotage des ateliers	81
10.1. Généralités	81
10.2. Quelques concepts utilisés en ordonnancement	82
10.3. Description des tâches et du système de production	83
10.4. Critères de qualité d'un ordonnancement	83
10.5. Ordonnancement sur une machine unique	85

10.6. Pilotage des ateliers: Règles de dispatching	87
10.7. Ordonnancement d'ateliers par les règles de dispatching	88
Bibliographie	89
CHAPITRE 11: Gestion de projets	91
11.1. Généralités	91
11.2. Les méthodes de chemins critiques	92
11.3. Construction d'un ordonnancement et implantation	97
11.4. Modifications du modèle de base	99
11.4.1. Durées incertaines (modèle PERT)	100
11.4.2. Compression des durées (modèle CPM)	102
Bibliographie	107

CHAPITRE 1

INTRODUCTION

*Every time I look at you I don't understand
Why you let the things you did get so out of hand
You'd have managed better if you had it planned
(Jesus Christ Superstar)*

1.1. Définitions et concepts de base.

La *production* est le processus conduisant à la création de produits par l'utilisation et la transformation de ressources. Les *opérations* sont les activités composant le processus de production.

Le terme « transformation » doit être entendu au sens large, puisqu'il recouvre la modification de la l'apparence, des propriétés physico-chimiques, de l'emplacement (transport), etc.

Les « produits » peuvent être des biens (physiques) ou des services.

Les « ressources » consistent principalement en

- capital et équipements
- main d'œuvre
- matières (premières, produits semi-finis)
- information.

Exemples : bois, hommes, atelier de menuiserie → tables

avions, pilotes, hôtesses, systèmes de gestion des réservations → transport aérien

La *gestion de la production (et des opérations)* est la fonction de gestion ayant pour objets la *conception*, la *planification* et le *contrôle* des opérations.

Les activités de *conception* portent sur la définition des caractéristiques

- du système productif (capacité, localisation, technologie, etc)
- des produits.

La *planification* décrit l'utilisation projetée du système productif dans l'objectif de satisfaire la demande. En d'autres termes, elle a pour objectif de *coordonner la capacité disponible avec la demande* (un thème récurrent en gestion de la production !).

L'activité de *contrôle* s'efforce d'évaluer l'adéquation des résultats obtenus par rapport aux plans.

Dans ce cours introductif, nous nous limiterons essentiellement à une discussion de quelques notions de base de la conception des systèmes productifs et à une présentation des principales décisions de planification. Nous articulons notre présentation dans un *cadre de décision hiérarchique* que nous allons maintenant décrire brièvement.

1.2. Gestion hiérarchique de la production.

En gestion de la production, on distingue généralement plusieurs niveaux de décision allant des décisions stratégiques et/ou à long terme aux décisions opérationnelles et/ou à très court terme. Nous nous arrêterons plus particulièrement aux niveaux suivants:

Gestion stratégique et planification à long terme: ce niveau de décision couvre

- la définition de la mission de l'entreprise (« what business are we in ? »), la composition du portefeuille de produits, l'identification des segments de marché visés, les choix de processus de production;
- la définition de la stratégie concurrentielle: domination par les coûts (cost leadership), différenciation (par la qualité, la flexibilité, l'innovation, les délais, ...), concentration de l'activité (focus), ...;

Les décisions ainsi mentionnées, même lorsqu'elles ne relèvent pas sensu stricto de la gestion de la production, présentent néanmoins de profondes implications pour cette fonction. Leur considération nous amènera notamment à présenter une typologie des systèmes productifs et à étudier leur relation avec les caractéristiques du marché et la stratégie concurrentielle. Nous évoquerons encore quelques questions plus spécifiques comme celles de la gestion de la qualité, de l'équilibrage des chaînes de production, de la localisation des départements, etc.

Planification à moyen terme: en gestion de la production, la planification à moyen terme comprend

- la planification agrégée de la production, dont le rôle est de coordonner les décisions en matière de production avec le contenu du plan marketing et la gestion du personnel;
- la gestion de projet.

Nous nous limiterons ici à l'élaboration du plan agrégé de production (quelques aspects de la gestion de projet sont évoqués dans le cours de Recherche Opérationnelle).

Planification à court terme: on retrouve ici

- l'élaboration du plan directeur de production (MPS) qui détaille les quantités à produire sur un horizon restreint;
- la planification des besoins en composants;
- la gestion des stocks.

Nous nous arrêterons assez longuement sur les principes de base de la technique la plus populaire de planification des besoins en composants, la méthode MRP (Material Requirements Planning). Nous évoquerons également quelques notions fondamentales de la gestion des stocks.

Ordonnancement et lancement de la production: affectation des ordres de fabrication aux postes de travail. Nous couvrirons très brièvement ce sujet important, mais souvent mal maîtrisé.

Contrôle: suivi des ordres de fabrication, contrôle de qualité, indicateurs de performance, etc.

Il est important de comprendre que les différents niveaux de décision se distinguent par d'autres caractéristiques que la seule étendue de leur horizon (du long terme au très court terme). Parmi les autres facteurs importants permettant de différencier ces niveaux, on citera par exemple:

- le degré d'irréversibilité des décisions,
- la diversité des fonctions de gestion impliquées,
- le degré d'agrégation des informations, des catégories de produits et de ressources,
- le degré d'incertitude,
- la prise en considération ou l'abstraction de certains détails, etc.

La présentation des décisions de gestion dans un cadre hiérarchique permet d'insister sur leur profonde interdépendance. Les niveaux de décisions supérieurs engendrent des contraintes qu'il est nécessaire de prendre en compte aux niveaux inférieurs et, réciproquement, les informations utilisées lors de toute prise de décision remontent généralement de niveaux inférieurs. Le cadre hiérarchique met en évidence ces nombreuses relations tout en fournissant un fil conducteur logique pour l'exposé des concepts fondamentaux à chaque niveau.

1.3. Développement historique de la GP: Quelques jalons ...

Révolution industrielle (19ème siècle):

- augmentation des volumes de production et de la complexité des organisations; décisions de production souvent laissées aux ouvriers;
- développement de la machinerie et de l'automatisation; ère des ingénieurs mécaniciens (John Cockerill).

Scientific management (~ 1910):

- observation des méthodes de travail: éclatement des tâches complexes en parties simples, sélection de méthodes optimales (parmi celles observées), systématisation des procédures; ère des ingénieurs industriels (Taylor, Gilbreth, Ford).

Recherche opérationnelle, management science (à partir de 1945):

- modélisation mathématique et optimisation ;
- transition des méthodes descriptives vers les méthodes prescriptives (ouvre la voie à l'intervention des gestionnaires).

Développement de l'informatique (à partir de 1950):

- logiciels d'aide à la décision ;
- automatisation de la production ;
- systèmes d'information (MIS) et systèmes intégrés de gestion d'entreprises (ERP).

Développement de la compétition internationale (à partir de 1970):

- importance accrue des coûts, de la flexibilité, de la qualité, des délais (exemple japonais); accent sur la fonction production, dont l'importance stratégique est explicitement reconnue.

1.4. Relation avec les autres fonctions de l'entreprise.

Par sa nature transversale, la gestion de la production comporte des liens avec de nombreux autres aspects de la gestion des entreprises: gestion stratégique, théorie des organisations, marketing, logistique, méthodes quantitatives et recherche opérationnelle, comptabilité, contrôle de gestion, etc. L'étudiant (ou le praticien) pourra trouver dans ces interconnexions l'une des facettes les plus séduisantes de la gestion de la production, ainsi qu'une occasion de redécouvrir certaines questions de gestion sous un angle nouveau.

Bibliographie.

K. Boskma, *Productie en Logistiek*, 2ème édition, Wolters-Noordhoff, Groningen, 1987.

A.C. Hax et D. Candea, *Production and Inventory Management*, Prentice-Hall, Englewood Cliffs, N.J., 1984.

CHAPITRE 2

GESTION STRATEGIQUE DE LA PRODUCTION

2.1. Généralités.

La gestion stratégique de l'entreprise est généralement définie comme couvrant les décisions relatives

- à la composition du portefeuille de produits et aux (segments de) marchés ciblés,
- à l'appropriation et à l'allocation des ressources financières.

Elle place donc principalement l'accent sur les fonctions marketing et finance. Une conséquence de cette vision du management stratégique est que les fonctions marketing et finance constituent souvent les préoccupations centrales des gestionnaires (ce sont les fonctions « nobles » de l'entreprise), alors que ces mêmes gestionnaires ont par ailleurs moins d'intérêt et moins d'expérience dans le domaine de la gestion de la production. En particulier,

- les managers se concentrent sur des stratégies concurrentielles de croissance, diversification, repositionnement basées sur des acquisitions, des fusions ou sur le marketing de leurs produits, plutôt que sur le développement d'une expertise technologique ou manufacturière propres;
- les managers se désintéressent des décisions de production, vues comme purement technologiques et non directement intégrées aux autres décisions de gestion.

Ce désintéressement, cependant, n'est pas totalement neutre. Voyons quel impact *négatif* l'accent mis sur les fonctions marketing et finance peut avoir pour la gestion de la production.

2.2. Impact sur la fonction production.

Marketing: La vision moderne du marketing veut que cette fonction joue un rôle intégrateur et serve d'interface entre le client d'une part et les autres fonctions de l'entreprise (production, finance, personnel, ...) par ailleurs (voir Figure ci-dessous). Selon le « concept marketing », la tâche essentielle de l'organisation est de déterminer les besoins et les désirs des marchés cibles, et de les satisfaire plus efficacement et de façon plus appropriée que la concurrence. En vertu de ce principe, les entreprises soumises à une pression concurrentielle accrue ont été conduites dans un passé récent à

- augmenter constamment le nombre de produits, les options disponibles, les exigences de qualité, et à
- diminuer la durée de vie des produits ainsi que les délais de production et de livraison.

Cette politique orientée vers la satisfaction des désirs du client peut engendrer de nombreuses conséquences néfastes pour la fonction production: coexistence de production massive et en petites séries, redémarrages fréquents de la production, diminution des opportunités d'apprentissage, difficultés de planification (commandes en dernière minute), obsolescence

rapide des procédés, augmentation des stocks, diminution de la qualité, etc. (lire par exemple la description des difficultés rencontrées par Boeing dans l'article annexé à ce chapitre).

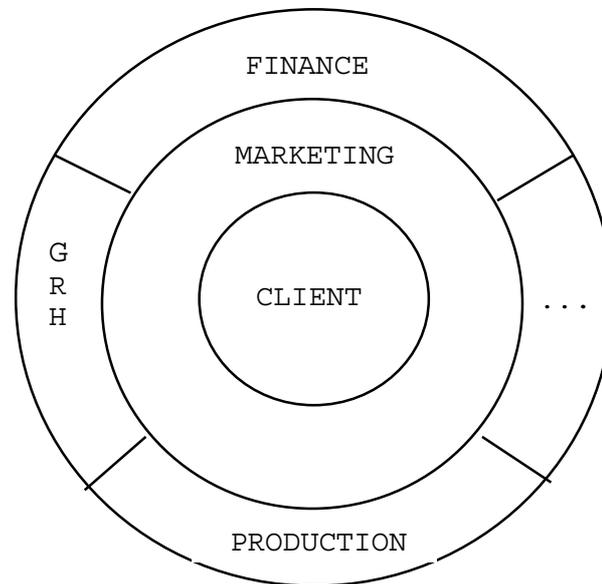


Figure: Le concept marketing

Finance: Hayes et Wheelwright (1984) estiment que les contrôles financiers ont récemment accordé trop d'importance à la rentabilité à court terme des investissements (Return On Investments, rapports trimestriels). Cet accent sur le court terme a été renforcé par la mobilité des cadres et dirigeants. Tout comme le concept marketing, cette politique financière a de profondes implications pour les activités de production: elle tend à décourager les investissements à long terme et/ou dans des technologies innovatrices, et entraîne ainsi une obsolescence plus rapide des équipements et une résistance à l'adoption de technologies avancées de production (par exemple, robotique ou ateliers flexibles; voir l'article de Parker et Lettes 1991 en annexe à ce chapitre).

2.3. Le rôle stratégique de la gestion de la production.

Vers la fin des années 60 et, de façon accélérée, durant les années 70, un nombre croissant d'entreprises (principalement japonaises) commencent à imposer la notion qu'une stratégie concurrentielle basée essentiellement sur l'excellence en production (qualité des produits, raccourcissement des délais, flexibilité, ...) peut se révéler viable et même gagnante.

Vers la même époque, plusieurs gourous académiques (Porter 1990, Skinner 1985, ...) soulignent l'importance pour chaque entreprise de formuler explicitement la stratégie

concurrentielle qu'elle a choisi de suivre et la nécessité de concentrer tous ses efforts vers la réalisation de ces objectifs stratégiques.

Porter, en particulier, distingue trois grandes catégories de stratégies face à la concurrence: domination globale par les coûts, différenciation, concentration de l'activité. Au-delà de ses recommandations quant au choix d'une stratégie adaptée à chaque entreprise, cependant, Porter insiste fortement sur la nécessité d'une mise en oeuvre *cohérente* de tous les *moyens* de l'entreprise (marketing, gestion financière, organisation, production) dans la poursuite des *objectifs* précisés dans sa stratégie. Faute de choix cohérents, l'entreprise risque de se trouver « enlisée dans la voie médiane » (« stuck in the middle ») avec pour conséquence une situation extrêmement faible face à la concurrence.

Skinner, quant à lui, s'intéresse plus particulièrement à la stratégie adoptée par les entreprises en matière de gestion de la production. Sur base de ses nombreux contacts industriels, il formule les observations suivantes (Skinner 1974 et 1985, Chapitres 1 et 6).

1. *Les décisions de production sont de nature stratégique.* Plus précisément, ce sont souvent des décisions déterminantes pour la compétitivité de l'entreprise et intimement liées aux autres décisions stratégiques (en particulier, aux décisions de marketing). Ces décisions portent généralement sur le long terme (investissements lourds), prennent lentement leurs effets et sont difficilement réversibles.
2. *Les critères de performance en production sont variés.* Comme mentionné plus haut, ces critères de performance se sont beaucoup élargis au cours des 20 à 30 dernières années sous la pression concurrentielle et l'allégeance au concept marketing: au seul critère de minimisation des coûts (productivité) sont venus s'ajouter des critères de qualité, de flexibilité, de délais de livraison, de qualité des investissements, etc. (Pensez aux stratégies concurrentielles décrites par Porter).
3. *Un système de production ne peut pas exceller dans tous les domaines.* Un système productif, pour être efficace, exige des choix et des compromis (trade-offs), principalement dans les domaines suivants :
 - technologie: typiquement, une fabrique ne peut pas maîtriser plus de deux ou trois technologies différentes;
 - marchés ciblés: chaque segment est caractérisé par des exigences distinctes en termes de qualité, délais, prix, etc;
 - volumes de production: des volumes différents requièrent des approches différentes de la planification, de la gestion des stocks, du contrôle de qualité, etc;
 - niveaux de qualité;
 - fabrication interne ou sous-traitance (make or buy);
 - degré de spécialisation des équipements;
 - risque technologique;
 - ...

Ne pas poser clairement ces choix conduit à des situations difficilement maîtrisables et empêche la firme de bénéficier des effets d'apprentissage. Pour contrôler ce type de dérive, les entreprises ont alors souvent tendance à augmenter leur staff et les couches de management au prix d'une augmentation de leurs frais fixes - overheads - et d'une perte de compétitivité. (On retrouve les effets pervers de la situation d'enlèvement dans la voie médiane décrite par Porter 1990).

Skinner tire de ces observations quelques leçons essentielles pour l'entreprise. Selon lui, pour rester compétitive, l'entreprise doit se concentrer sur quelques critères d'excellence dans le domaine de la production, formuler explicitement ses choix et mettre les moyens en oeuvre pour les atteindre (en excluant la poursuite d'objectifs non compatibles; concept de focused factory). Si l'entreprise possède des produits ou des marchés impliquant des choix contradictoires, son appareil productif devrait être séparé en usines distinctes, exhibant chacune les caractéristiques de la focused factory (plant within a plant).

Par ailleurs, les choix stratégiques posés dans la sphère marketing doivent être cohérents avec ceux posés dans la sphère production. En particulier, chacune de ces fonctions impose des contraintes dont l'autre doit tenir compte: l'ajustement doit être réciproque plutôt qu'unidirectionnel.

Ces conclusions restent clairement valides et pertinentes aujourd'hui. Le rapport annexé au présent chapitre (De Meyer et al. 1996) décrit la façon dont les grandes entreprises manufacturières européennes, américaines et japonaises envisagent actuellement leurs options stratégiques en matière de production. On y découvrira avec intérêt que les choix qu'elles posent dans ce domaine varient significativement dans les différentes régions. Il reste, cependant, que la plupart des critères de performance mentionnés par les entreprises s'inscrivent dans une des quatre grandes catégories: *coût, qualité, délais, flexibilité*, qui constituent les étalons incontournables en matière de production.

2.4. Gestion de la qualité.

Nous allons discuter brièvement de la gestion de la qualité, comme exemple typique d'une variable dont la gestion est entrée récemment dans le domaine stratégique. (Voir Graves, Rinnooy Kan et Zipkin (1993), Juran (1993), McClain, Thomas et Mazzola (1992)).

Informellement, on appelle *qualité* d'un produit la mesure dans laquelle ce produit est conforme aux attentes et exigences de l'utilisateur. Plus formellement, l'Organisation Internationale de Normalisation (ISO) définit la qualité comme « l'ensemble des caractéristiques d'une entité qui lui confèrent l'aptitude à satisfaire des besoins exprimés ou implicites », où le terme entité peut désigner un produit, une activité, un processus, un système, etc. Il ne faut donc pas confondre qualité et luxe: un crayon ou de l'eau peuvent être de bonne ou de mauvaise qualité.

A ses débuts, la gestion de la qualité se limitait essentiellement à deux aspects :

- le contrôle de la qualité des produits effectué à la sortie du processus de production, souvent par échantillonnage statistique;
- la gestion du service après-vente, éventuellement sous couverture de garantie.

Les inconvénients liés à ces approches sont nombreux. D'une part, elles génèrent souvent des coûts élevés, qu'ils soient directs ou indirects: coût de l'inspection, mise au rebut ou réparation/adaptation des articles ne satisfaisant pas aux spécifications, réparation des articles défectueux sous garantie, etc. (Plus généralement, certains auteurs estiment que le coût de la

qualité - c'est-à-dire les coûts totaux liés au manque de qualité - peuvent s'élever à 20% du chiffre d'affaires d'une entreprise; cité par McClain et al. 1992, p.132).

D'autre part, le contrôle de qualité par échantillonnage est, par sa nature même, imparfait et incomplet. Rappelons brièvement son principe de base. Pour contrôler la qualité d'un lot de production (ou d'un approvisionnement), on en extrait un échantillon et on analyse la qualité de chacun des articles inclus dans l'échantillon. Si X% (où X est un paramètre fixé à l'avance) de ces articles sont jugés satisfaisants, alors le lot est accepté. Sinon, il est rejeté.

Le contrôle par échantillonnage admet donc explicitement la mise en vente de produits de qualité inférieure ou défectueux, tant qu'il n'y en a pas trop ! Mais en pratique, cette tactique peut résulter en une perte de confiance des consommateurs et, dans des contextes de vive concurrence, par une perte de compétitivité et une diminution des parts de marché.

Plus grave encore, le contrôle par échantillonnage peut se révéler tout-à-fait inapproprié dans certaines situations, par exemple lorsque la conformité du produit doit être assurée à 100% (pensons aux normes de sécurité des installations nucléaires ou, plus près de notre expérience quotidienne, aux exigences placées sur la qualité des pneus de notre voiture ; voir le cas « Ford et les pneus Firestone » en annexe à ce chapitre).

Enfin, le contrôle effectué en fin de processus ou les observations glanées lors des réparations après vente ne permettent qu'un feedback tardif et souvent très imparfait (ou même totalement absent) vers la source des problèmes (voir l'expérience de Xerox, citée en exemple par Juran 1993).

Les managers japonais, plus prompts que leurs concurrents occidentaux à réagir au message des consultants (américains !) Deming et Juran, ont été les premiers à appréhender toute l'ampleur de ces problèmes et à prendre des mesures effectives pour y remédier. Cette prise de conscience s'est à présent répandue à l'ensemble du monde industrialisé, où le remède le plus couramment préconisé est celui de l'adoption du principe de la *Gestion de la Qualité Totale* (Total Quality Management). Ce principe vise essentiellement à gérer et à assurer la qualité des produits (c'est-à-dire, la conformité de ces produits à leurs spécifications) tout au long des différentes étapes de la production plutôt qu'à son terme seulement. Le but ultime est donc de ne produire que des unités de qualité parfaite, c'est-à-dire sans défauts: il faut assurer la conformité sans avoir à la tester. Cette approche contraste fondamentalement avec l'approche plus traditionnelle qui vise seulement à ne mettre sur le marché que des produits conformes.

Les méthodes mises en oeuvre par les programmes de Qualité Totale sont très variées. Citons en quelques-unes ...

- Amélioration de la conception du produit, de façon à mieux répondre aux besoins du consommateur d'une part (cf. notre définition de la qualité et le cas « Ford et les pneus Firestone ») et à faciliter la fabrication du produit, donc à diminuer le nombre de défauts d'autre part (conception pour la manufacturabilité). Les méthodes de « déploiement de la fonction qualité » fournissent un outil d'implantation de cette stratégie (voir Cohen 1995).

- Amélioration de la conception du processus de production, de façon à minimiser les risques d'erreurs; par exemple, automatisation ou informatisation de certaines étapes, diminution du nombre de réglages nécessaires, etc. (chez Philips, on minimise le nombre de changements d'outils nécessaires sur les machines d'assemblage automatisé des circuits imprimés, en interdisant les changements pendant des périodes de plusieurs semaines consécutives; ceci entraîne une perte de flexibilité et donc apparemment de productivité, mais diminue aussi fortement les risques d'erreurs). Les améliorations de processus peuvent par exemple être prises sur l'initiative de cercles de qualité (CQ), c'est-à-dire de groupes d'ouvriers ou d'employés discutant ensemble des problèmes qu'ils rencontrent dans leurs activités, et cherchant des solutions à ces problèmes. Les CQ reçoivent une formation à des méthodes simples d'analyse de problèmes et d'analyse statistique. Leur contact quotidien avec le processus et ses défauts leur permet souvent de détecter rapidement les problèmes qui surgissent et d'y trouver des solutions parfois très simples.
- Formalisation et documentation des procédures: « Ecrire ce que l'on va faire. - Faire ce que l'on a écrit. - Ecrire ce que l'on a fait. » Cette pratique vise notamment à imposer le respect des spécifications du produit (dans la mesure, notamment, où les caractéristiques du produit sont largement déterminées par les procédures utilisés lors de sa fabrication). La documentation écrite permet également d'informer le client ou le donneur d'ordre, de manière convaincante, sur la façon dont a été réalisé le produit. La documentation des procédés fait partie intégrante des exigences reprises dans les normes d'assurance qualité adoptées internationalement (telles les normes ISO9000, ...).
- Modification de la mentalité des acteurs de l'entreprise: chacun est amené à se considérer comme le consommateur des produits livrés par les services situés en amont de la chaîne et comme le fournisseur des services situés en aval. Il ne faut donc ni accepter ni livrer des produits de qualité inférieure. Un tel état d'esprit est favorisé par l'adoption de modes de production Juste-A-Temps où les faibles niveaux de stocks mettent immédiatement à jour les défauts éventuels (il n'est plus possible de simplement mettre au rebut les pièces défectueuses livrées par le poste amont, sans risquer de se trouver immédiatement à court de ces pièces).
- La qualité est vue dans un cadre d'amélioration permanente de l'organisation: il n'est donc pas uniquement question de chercher à atteindre un niveau de performance prédéterminé et de s'y maintenir, mais bien de chercher à améliorer ce niveau de façon continue (Permanent Improvement). En d'autres termes, le standard de qualité visé (mais jamais atteint) devrait être celui de la perfection absolue (de là les slogans de type 'Zéro défaut', 'Zéro stock', etc).

Nous tirerons deux conclusions de cette brève discussion de la gestion de la qualité. Tout d'abord, les améliorations de qualité peuvent souvent être atteintes à un coût relativement modeste et par des moyens simples (par exemple, en les gérant dans la phase de conception du produit ou en utilisant l'expertise des ouvriers). De telles améliorations peuvent simultanément résulter en un accroissement de la compétitivité de l'entreprise et en une diminution de ses coûts de production (Cadillac mentionne une réduction de 30% de ses coûts de réparations sous garantie, réduction obtenue en 4 ans malgré l'allongement simultané de la garantie de 1 an à 4 ans; cité par McClain et al. 1992, p.124).

Cependant (et ce sera là notre seconde conclusion), il est essentiel de voir que les remèdes évoqués ci-dessus impliquent des changements fondamentaux dans la façon dont est perçue la gestion de la qualité au sein de l'entreprise. On y retrouve en effet un glissement

- du contrôle (ex post) vers la planification (ex ante),
- de la notion de qualité du produit vers celle de qualité de la conception et des processus,
- d'une activité opérationnelle vers une préoccupation stratégique.

Ces changements ne peuvent être mis en oeuvre qu'au prix d'un engagement total de toutes les forces de l'entreprise (l'amélioration de la conception des produits nécessite par exemple l'implication d'équipes interdisciplinaires - bureau d'études, marketing, production, service après-vente, etc.; les cercles de qualité ou le JAT impliquent toute la chaîne productive). En particulier, la direction de l'entreprise doit jouer un rôle direct dans la gestion de la qualité, qui devient donc un instrument de gestion stratégique (au Japon, la qualité est souvent du ressort direct du Directeur Général).

Bibliographie.

L. Cohen, *Quality Function Deployment: How to Make QFD Work for You*, Addison-Wesley, 1995.

A. De Meyer, H. Kitayama et J.S. Kim, Building customer partnerships as a competitive weapon: The right choice for globalising competition?, Working Paper INSEAD, 1996.

S.C. Graves, A.H.G. Rinnooy Kan et P.H. Zipkin (eds.), *Logistics of Production and Inventory*, North-Holland, Amsterdam, 1993.

J. Greenwald, Is Boeing out of its spin?, *TIME*, July 13, 1998, pp. 55-57.

J.R. Hauser et D. Clausing, The house of quality, *Harvard Business Review*, May-June 1988, pp. 63-73.

R.H. Hayes and D.M. Upton, Operations-based strategy, *California Management Review* 40 (1998) pp. 8-25.

R.H. Hayes et S.C. Wheelwright, *Restoring our Competitive Edge: Competing through Manufacturing*, J.Wiley & Sons, New York, 1984.

J.M. Juran, Made in U.S.A.: A renaissance in quality, *Harvard Business Review* 71(4), July-August 1993, pp. 42-50.

J.O. McClain, L.J. Thomas et J.B. Mazzola, *Operations Management*, 3ème édition, Prentice-Hall, Englewood Cliffs, 1992.

Organisation Internationale de Normalisation, *ISO 8402: Management de la Qualité et Assurance de la Qualité -- Vocabulaire*, 1994 (site Internet: <http://www.iso.ch/indexf.html>).

T. Parker et T. Lettes, Is accounting standing in the way of flexible computer-integrated manufacturing?, *Management Accounting*, January 1991, pp. 34-38.

W.G. Plossl, *La Nouvelle Donne de la Gestion de Production*, AFNOR, Paris, 1993.

M.E. Porter, *Choix Stratégiques et Concurrence*, Economica, Paris, 1990.

W. Skinner, The focused factory, *Harvard Business Review* 52(3), May-June 1974, 113-121.

W. Skinner, *Manufacturing: The Formidable Competitive Weapon*, J.Wiley & Sons, New York, 1985.

ANNEXE : Extraits du site Internet

<http://www.humanite.presse.fr/journal/2001/2001-05/2001-05-24/>

24 Mai 2001

**Ford obligé d'organiser le rappel de 13 millions de pneus.
Ford enlève la gomme**

Le deuxième constructeur automobile mondial lance une gigantesque opération de sécurité en rappelant 13 millions de pneumatiques Firestone équipant son modèle tout-terrain Explorer.

C'est le principe de précaution version US. Mardi après-midi le géant nord-américain de l'automobile a annoncé par la voix de son PDG le rappel de tous les pneus "Wilderness AT" qui équipent le modèle vedette de la marque. "C'est une simple mesure de prévention", a précisé Jacques Nasser, le Big Boss depuis son repère du Michigan à propos de pneus qui "ont montré un taux d'échec total."

Le groupe Ford est donc une nouvelle fois en pleine tourmente, avec ce rappel de millions de pneus potentiellement défectueux. Conséquence immédiate et inévitable, au pays du tout apparence, il faudra que la célèbre marque bataille ferme pour tenter de garder sa bonne image auprès des automobilistes. Et ce d'autant plus que le deuxième constructeur automobile mondial, dont l'ambition est de détrôner General Motors, ne sort décidément pas depuis maintenant un an de ce problème de pneus sur son véhicule vedette, le 4x4 Ford Explorer. Elle se superpose au rappel de 6,5 millions de pneus, principalement montés sur des Ford Explorer, lancé conjointement début août 2000 par Ford et Firestone.

Le PDG de Ford, Jacques Nasser, est même venu lundi au Congrès avertir les parlementaires de la mesure draconienne que son groupe allait prendre : le rappel de 13 millions de pneus Firestone, fabriqués par la filiale américaine du groupe japonais Bridgestone. Comme c'est la sécurité qui est considérée comme la préoccupation principale des consommateurs, Ford, sur un terrain aussi significatif, n'a pas hésité, même si la gestion de l'image pour une affaire de ce type est un exercice difficile, selon les experts en communication.

Pour garder une bonne image ou la restaurer, le constructeur se devait de jouer la transparence. On peut raisonnablement poser la question sur le long terme de ce que sera l'impact de cette affaire mais dans l'immédiat, en tout cas, la décision de la marque a été appréciée, selon les premières analyses d'impact réalisées auprès des consommateurs.

Globalement, le groupe Ford devrait donc plutôt bien s'en tirer, même si il y a un certain nombre de procès en cours à l'heure actuelle de par les poursuites engagées par les familles de victimes d'accident après l'éclatement d'un pneu. Selon la sécurité routière nord-américaine (la NHTSA), 174 décès sont liés à ce problème d'éclatement.

Les experts en marketing ne pensent pas que les ventes de l'Explorer, un véhicule très populaire, soient pour l'instant menacées, sachant que l'indice de satisfaction des propriétaires de ce produit reste très haut. Il y a pourtant des avis divergents. Notamment celui de Wendy Neeham, une analyste automobile de la banque Credit Suisse First Boston, qui voit ce nou-

veau rappel comme le signe de problèmes à répétition pour le constructeur. Elle a précisé également que : "Des pneus endommagés sur une ligne de montage, c'est vraiment un problème spécifique Ford et pas une responsabilité du manufacturier Firestone."

Selon cette analyste les gens seraient fondés à se demander : "Avons-nous réellement besoin d'acheter un Explorer?" Une question angoissante quand on sait que l'Explorer contribue actuellement pour un quart au bénéfice de Ford. Le rappel des 13 millions de pneus demanderait neuf mois en tout pour être mené à bien. Le coût en est évalué à 3 milliards de dollars (3,5 milliards d'euros) et pour l'instant personne ne sait si le fabricant de pneus participera au règlement de la note. Il faut dire que les relations entre les deux entreprises sont au plus mal, chacune rejetant sur l'autre, tout ou partie, de la responsabilité des problèmes rencontrés. Le constructeur reprochant au manufacturier de ne pas avoir ajouté une couche de nylon dans la fabrication, tandis que Firestone n'hésite plus à remettre en cause la conception même, et donc la stabilité du "4X4 de ville de Ford", dans les 174 accidents mortels incriminés. Cette décision de Ford intervient d'ailleurs au lendemain de l'annonce par Firestone, filiale américaine du fabricant de pneumatiques japonais Bridgestone, de mettre fin à ses relations commerciales avec le groupe Ford.



ANNEXE

ISO 9000 pour responsables affairés

Bienvenue à l'itinéraire ISO 9000 de notre Voyage initiatique et démystificateur!

Qu'est-ce qu'ISO 9000?

Les normes de la famille ISO 9000 sont le fruit d'un consensus international sur les bonnes pratiques de management. Leur objectif est d'assurer qu'une organisation peut régulièrement fournir des produits ou services qui répondent aux exigences de la clientèle en matière de qualité. Ces bonnes pratiques ont été distillées en un ensemble d'exigences normalisées pour un système de management de la qualité, indépendamment de ce que fait votre organisation, de sa dimension, de son appartenance au secteur privé ou public.

Pourquoi mon organisation devrait-elle mettre en oeuvre ISO 9000?

Une organisation sans clients, ou dont les clients sont insatisfaits, est en péril! Pour conserver la clientèle - une clientèle satisfaite - votre produit (qui peut fort bien être un service) doit satisfaire à ses exigences. ISO 9000 offre un cadre bien éprouvé qui vous permettra d'adopter une approche systématique pour gérer vos processus (les activités de votre organisation) afin que la production réponde aux attentes de la clientèle - une clientèle qui sera ainsi constamment satisfaite!

Comment fonctionnent les normes de la famille ISO 9000?

Les exigences d'un système qualité ont été normalisées. Or, chacun d'entre nous a tendance à penser que son entreprise est unique. Comment prévoit-on alors, dans ISO 9000, l'écart entre une entreprise familiale et une multinationale de fabrication ayant des composantes de services, ou encore un service public ou une administration gouvernementale?

La réponse est qu'ISO 9000 établit à *quelles* exigences votre système qualité doit répondre, mais ne stipule pas *comment* y répondre dans votre organisation, ce qui crée une grande latitude et une souplesse pour la mise en oeuvre dans divers secteurs et cultures d'entreprises... et dans différentes cultures nationales.

Ainsi, la famille ISO 9000 comprend des normes qui donnent aux organisations des orientations et des exigences sur ce qui constitue un système efficace de management de la qualité. ISO 9004-1 (ainsi que les autres parties d'ISO 9004) donne des lignes directrices sur les éléments du management de la qualité et du système qualité.

La famille comprend également des modèles permettant d'auditer ce système afin de donner à l'organisation et à ses clients une assurance qu'il fonctionne efficacement. Les trois modèles d'assurance de la qualité sont ISO 9001, ISO 9002 et ISO 9003.

Enfin, la famille comprend une norme sur la terminologie et d'autres normes que l'on peut qualifier d'"outils de soutien", qui donnent des orientations sur les aspects spécifiques comme l'audit des systèmes qualité.

Et si mon organisation décide de mettre en oeuvre ISO 9000?

Elle devrait procéder elle-même à un audit de son système qualité fondé sur ISO 9000 afin de vérifier qu'elle gère efficacement ses processus ou, pour le dire autrement, qu'elle maîtrise rigoureusement ses activités.

De plus, l'organisation peut inviter ses clients à auditer le système qualité afin de leur donner confiance dans son aptitude à livrer des produits ou services qui répondent à leurs exigences.

Enfin, l'organisation peut solliciter les services d'un organisme indépendant de certification de systèmes qualité afin d'obtenir un certificat de conformité à ISO 9000. Cette option a gagné toute la faveur des entreprises en raison de la crédibilité dont jouit une évaluation indépendante. Elle peut éviter les audits multiples effectués par les clients de l'organisation, ou diminuer leur fréquence ou leur durée.

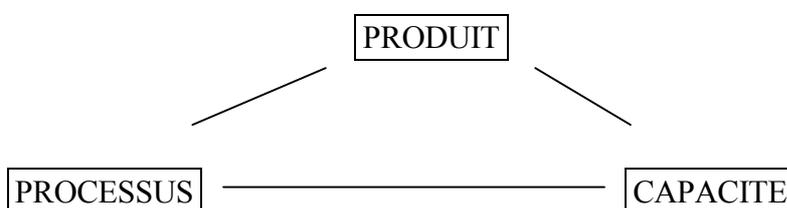
Le certificat peut également servir de référence entre l'organisation et les prospects, en particulier lorsque fournisseur et client ne se connaissent pas ou sont éloignés géographiquement, comme c'est le cas dans un contexte d'exportation.

Document extrait du site de l'Organisation International de Normalisation (ISO) : <http://www.iso.ch/> (09/10/02)

CHAPITRE 3

PLANIFICATION A LONG TERME: LES OPTIONS FONDAMENTALES

Les grandes options fondamentales, c'est-à-dire celles qui déterminent et qui limitent les opportunités ouvertes à l'entreprise dans le long terme, portent principalement sur la composition du portefeuille de produits, sur le processus de production adopté et sur la capacité de production.



Ces options sont soumises à de nombreuses interdépendances: le choix du produit influence la façon de le produire, qui influence les coûts, donc le volume de vente potentiel, donc la capacité, qui affecte le choix de processus, etc.

Avant de nous attarder plus longuement sur le choix du processus de production, évoquons rapidement, et en termes très généraux, les décisions portant sur la conception du produit et sur la capacité du système productif.

Conception du produit. De plus en plus la conception du produit est considérée comme un travail d'équipe réclamant des expertises interfonctionnelles: marketing, bureau d'études, production, etc (voir Section 2.4 ci-dessus). Tout produit doit en effet répondre aux critères du consommateur, mais doit aussi être facilement « manufacturable » afin d'améliorer sa qualité et de réduire les délais de production.

La Conception Assistée par Ordinateur (CAO) permet d'accélérer le processus de développement et de générer facilement des designs différents. Elle peut être couplée à des systèmes permettant de réaliser des prototypes tridimensionnels.

La Conception et Fabrication Assistée par Ordinateur (CFAO, ou Computer Assisted Design/Computer Assisted Manufacturing, CAD/CAM) ajoute aux avantages de la CAO la possibilité de générer automatiquement les séquences d'opérations-machine à effectuer pour fabriquer le produit. Elle aide à réduire considérablement les délais de production totaux (entre la passation de commande et sa livraison au client), et assure une bonne coordination entre le bureau d'études et la production.

Capacité. La capacité d'un système productif est une mesure du nombre d'unités que peut produire le système en un laps de temps donné. La simplicité apparente de cette définition est cependant trompeuse. En effet, la capacité est souvent difficile à estimer, en partie parce qu'elle est déterminée par une multitude de paramètres dont certains sont eux-mêmes difficilement quantifiables:

- l'appareil productif proprement dit,
- la main d'oeuvre,
- les possibilités de sous-traitance,
- les choix technologiques,
- la qualité,
- la motivation des ouvriers,
- le product mix,
- ...

Une partie de cette variabilité s'exprime à moyen ou à court terme (par exemple, les possibilités de sous-traitance ou d'heures supplémentaires), et nous y reviendrons dans ce cadre. Mais la capacité est également fortement limitée par les décisions à long terme. Celles-ci sont de natures diverses et présentent de nombreuses interdépendances.

- Quand faut-il modifier la capacité ? Il n'est pas toujours conseillé de suivre la demande, qui peut montrer des tendances à court terme, ou suivre des cycles conjoncturels (voir l'exemple de la sidérurgie : les producteurs sont obligés de développer des grandes capacités pour diminuer leurs coûts de production en haute conjoncture, et encourent une surcapacité sectorielle en basse conjoncture; une situation similaire se produit dans l'enseignement universitaire qui est soumis aux fluctuations du marché de l'emploi et de la « mode »).
- De combien faut-il modifier la capacité ? De façon progressive (ce qui permet une meilleure adéquation à la demande) ou par « sauts » (qui peuvent s'accompagner d'économies d'échelle) ? Ce problème est évidemment lié au précédent.
- Où faut-il (re)localiser la capacité de production? Les décisions de localisation sont étroitement liées à celles définissant la configuration du *réseau logistique*, c'est-à-dire le réseau d'approvisionnement, de production et de distribution de l'entreprise. Les progiciels de support au Supply Chain Management (du type APS, *Advanced Planning Systems*) intègrent des modèles d'aide à la décision, basés par exemple sur des modèles d'optimisation de localisation similaires à ceux présentés dans l'Annexe « Modèles et méthodes d'optimisation ».
- Quel type de capacité faut-il ajouter ? (Voir les alternatives énumérées plus haut).

Nous ne nous attarderons pas davantage sur ce sujet et renverrons le lecteur à la littérature pour plus d'informations.

Bibliographie.

R.H. Hayes et S.C. Wheelwright, *Restoring our Competitive Edge: Competing through Manufacturing*, J.Wiley & Sons, New York, 1984.

W.G. Plossl, *La Nouvelle Donne de la Gestion de Production*, AFNOR, Paris, 1993.

CHAPITRE 4

PLANIFICATION A LONG TERME: CHOIX DU PROCESSUS DE PRODUCTION

4.1. Généralités.

Le *choix du processus de production* comporte le choix de l'ensemble des équipements, personnel et procédures utilisés pour la production, ainsi que le choix de l'organisation de cet ensemble.

Le choix du processus est bien sûr affecté par la nature du produit et par les contraintes techniques (on ne fait pas de voitures en caoutchouc et on n'utilise pas de foreuses pour produire des hamburgers), mais aussi par des questions d'ordre plus stratégique comme par exemple la relation entre le processus et le type de marché. Nous placerons notre discussion à ce dernier niveau, sur un plan organisationnel plutôt que technique.

4.2. Classification des processus de production en fonction des flux de produits.

On va présenter ici une classification grossière, mais néanmoins utile, des processus de production. La classification est basée sur une analyse du *flux* des produits, c'est-à-dire sur la *séquence de stations de travail* visitées par les produits lors de leur passage à travers le système productif. Sous le terme générique « station de travail », nous désignons ici les différents éléments du système, comme par exemple les machines à commande numérique, les machines opérées manuellement par un ouvrier, les guichets de service, les laboratoires, etc.

- Production en ligne: dans un processus en ligne, il existe un flux dominant de produits en ce sens que (presque) toutes les unités produites parcourent les stations de travail selon la même séquence. On distingue deux grandes sous-classes de processus en ligne :

Production à flux continu: ici, deux unités successives du produit ne peuvent être ni séparées ni distinguées les unes des autres.

Exemples : - industrie pétrochimique,
- sidérurgie (phase à chaud),
- industrie agro-alimentaire (aliments pour bétail, ...)

Chaîne de production ou d'assemblage: les unités produites visitent les stations de travail dans le même ordre et subissent des séquences d'opérations (presque) identiques; mais elles sont physiquement séparées, et peuvent présenter de légères variations les unes par rapport aux autres (par exemple, par l'addition d'options).

Exemples : - construction automobile,
- appareils électroménagers,
- McDonald,
- contrôle technique automobile.

- Job-shop, ou atelier: il n'y a pas de flux dominant de produits; le système produit des articles variés exigeant des séquences d'opérations distinctes.
Exemples : - atelier traditionnel (mécanicien, menuisier),
- sous-traitance de pièces mécaniques,
- cuisine de restaurant traditionnelle,
- laboratoire d'analyses médicales,
- hôpital.
- Production en séries moyennes, ou à flux intermittent (batch flow): intermédiaire entre les précédents; les unités au sein d'une série sont identiques; les différentes séries visitent des séquences de stations similaires, mais chaque série requiert des réglages importants ou des opérations distinctes à chaque station.
Exemples : - cuisson des émaux
- chaussures
- circuits imprimés
- produits exigeant des traitements chimiques (bains).

Un système de production en ligne se caractérise par le fait que les ressources (machines, hommes) sont organisées en fonction de l'article à produire: on dit que le processus est *organisé par produit*. Par contre, dans un job-shop, les ressources sont groupées sur base des opérations qu'elles réalisent: le processus est *organisé par fonction*.

Les différents types de systèmes se distinguent encore par de nombreuses autres caractéristiques. Dans un système de production en ligne, généralement,

- un seul type, ou très peu de types d'articles différents sont produits;
- le nombre d'unités produites (volume de production) est élevé;
- l'équipement est très spécialisé, automatisé et peu flexible;
- les investissements en équipements et en études de conception du système sont importants;
- la main d'œuvre est limitée;
- le taux d'utilisation des équipements est très élevé (souvent plus de 90%);
- la production se fait pour stock, avec par conséquent des stocks de matières premières, de composants et de produits finis élevés et des stocks d'encours relativement faibles.

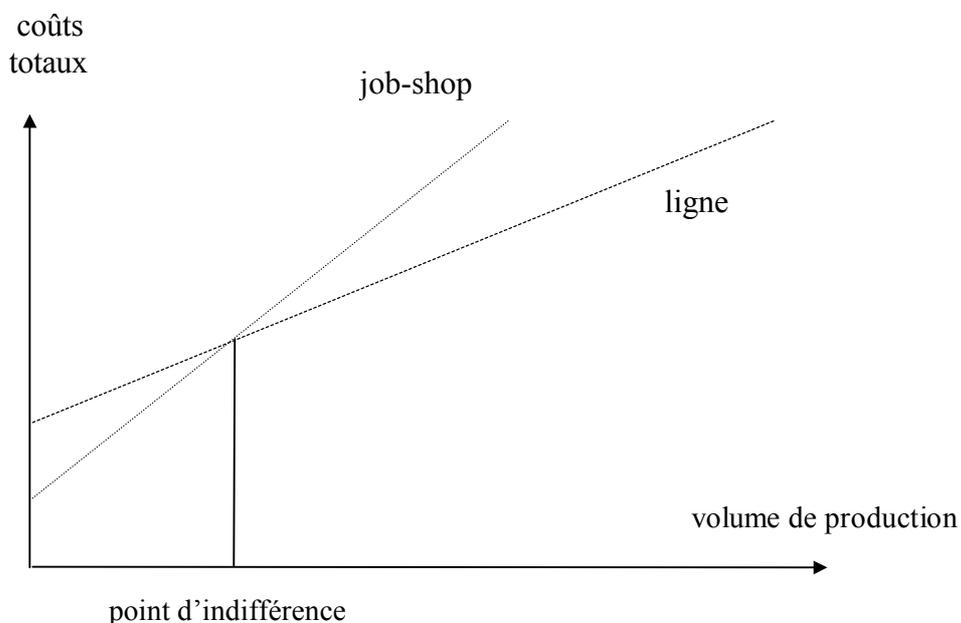
A l'inverse, un job-shop se caractérise souvent par

- une grande variété de produits (adaptés aux exigences spécifiques de chaque client);
- un nombre d'unités de chaque type et un volume de production total peu élevés;
- des équipements non spécialisés, peu automatisés et très flexibles (temps de réglage faibles);
- des investissements en équipements relativement peu élevés;
- une main d'œuvre importante;
- un taux d'utilisation des équipements très faible (souvent moins de 50%); en effet, dans un job-shop, chaque produit ne requiert qu'une minorité des équipements présents; de plus, chaque changement de production entre différents types de produits se traduit par des temps de réglage non productifs; certains observateurs mentionnent ainsi que, dans le secteur de l'industrie mécanique, les produits passent typiquement 95% de leur temps dans l'atelier sous

forme d'encours; 1/3 seulement du temps qu'ils passent sur les machines représente du temps d'usinage effectif, soit moins de 2% du temps de passage total dans l'atelier.

- un mode de production à la commande, et donc des stocks de produits finis relativement faibles.

En termes de coûts, les avantages comparatifs des deux types de systèmes peuvent être esquissés comme suit. Le principe directeur de la production en ligne est que les coûts fixes élevés (investissements) doivent pouvoir être absorbés par de grands volumes de production. A l'opposé, dans un job-shop, les coûts fixes sont maintenus à un niveau plus modeste et les coûts variables, c'est-à-dire liés directement au volume de production, sont prépondérants (par exemple, les coûts salariaux). Le job-shop peut donc être rentable en produisant des quantités relativement faibles (inférieures au *point d'indifférence*), alors que, lorsque le volume de production augmente, les coûts variables plus modestes de la production en ligne se traduisent également par des coûts moyens moins élevés.



4.3. Cycle de vie du produit et organisation de la production.

Dans une vision dynamique, ou évolutive, des caractéristiques d'une entreprise ou d'un produit, la discussion précédente peut être interprétée comme suit: lorsque le volume de vente potentiel et le degré de standardisation d'un produit augmentent, l'entreprise accepte de substituer des coûts variables (main d'oeuvre, temps de réglage) par des coûts fixes (équipements spécialisés) et donc de passer du job-shop à la production en ligne. Hayes et Wheelwright (1984) ont analysé ce phénomène en le replaçant dans le cadre des recommandations formulées par Skinner sur la

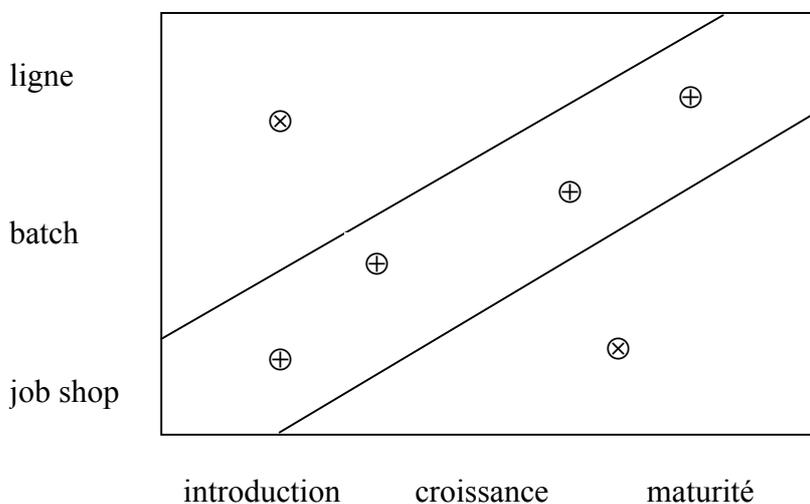
cohérence nécessaire entre les décisions de marketing et de production (voir Chapitre 2 et la confirmation empirique de cette analyse par Safizadeh et al. 1996).

Afin de formuler les conclusions de ces auteurs, rappelons d'abord brièvement les phases successives du cycle de vie d'un produit telles qu'elles sont décrites par la théorie classique du marketing (voir par exemple Kotler et Dubois 1994): introduction (augmentation progressive des ventes), croissance (croissance rapide), maturité (plateau), déclin (décroissance).

La matrice produit-processus de Hayes et Wheelwright (1984) permet de visualiser simultanément la position d'un produit dans son cycle de vie et l'organisation de sa production.

- Dans les premières phases de son cycle, le produit est peu standardisé et sa conception évolue encore, la demande est faible, les firmes qui le produisent sont de petite taille. Un système de production organisé en job-shop est donc généralement le mieux adapté dans ces phases du cycle de vie.
- Dans les phases ultérieures, on constate une standardisation du produit, une augmentation des volumes de production et l'émergence de quelques grandes firmes essayant de dominer par les prix. Parallèlement, une évolution de l'organisation vers la production en ligne se révèle de mieux en mieux adaptée aux besoins rencontrés.

Exemple : évolution de l'industrie automobile depuis le début du siècle.



Matrice produit-processus
(Hayes and Wheelwright 1984)

Cette analyse donne lieu à la notion de *cycle de vie des processus*: idéalement, au cours du cycle de vie d'un produit, le point qui le représente dans la matrice produit-processus devrait se déplacer le long de la diagonale de cette matrice. Insistons sur le fait que cette évolution des processus de production est assez différente de celle des volumes de vente, en ceci qu'elle peut être, en principe, contrôlée par la firme plutôt que d'être simplement subie ou observée. Elle

constitue donc, potentiellement, un important instrument de gestion dans la stratégie concurrentielle d'une entreprise.

D'autres stratégies sont évidemment possibles mais doivent alors être reconnues explicitement comme des stratégies de différenciation non standards. De telles stratégies doivent être soigneusement évaluées pour s'assurer qu'elles ne créent pas d'incohérence entre le plan marketing de l'entreprise et le processus de production adopté (voir Safizadeh et al. 1996 pour quelques exemples, ainsi que l'article décrivant la situation de Boeing, en annexe au Chapitre 2).

4.4. Cellules de fabrication et ateliers flexibles: entre job-shops et lignes de production.

Cellules de fabrication

Dans la vision extrême du job-shop, le nombre de produits différents est énorme, et tous ces produits ont des *gammes opératoires* différentes (séquences de machines visitées, opérations sur ces machines). A l'inverse, la *technologie de groupe* reconnaît explicitement que de nombreux produits peuvent en fait être groupés en *familles* de produits similaires (si même ils ne sont pas parfaitement identiques) sur base des similitudes de leurs caractéristiques physiques et/ou de leurs gammes opératoires.

Exemples : - famille engrenages
- famille tubes

Au sein d'une famille, chaque produit peut alors être défini par ses caractéristiques propres.

Exemple : - engrenages: - nombre de dents
- diamètre
- tubes: - longueur
- diamètre
- épaisseur

La technologie de groupe répertorie chaque produit (et, plus généralement, chaque article connu dans l'entreprise) au moyen d'un *code (alpha)numérique* qui permet de l'identifier de façon unique et qui reflète également ses caractéristiques essentielles. Ce code peut être utilisé pour mettre en évidence les similitudes entre produits (des codes similaires correspondent à des produits similaires) et donc pour définir les familles de produits de façon pertinente.

La classification en familles et la codification des articles permettent de ne pas réinventer perpétuellement la roue: elles facilitent

- la gestion des stocks et des documents techniques,
- la conception d'un nouveau produit au sein d'une famille connue,
- l'élimination de designs superflus,
- la définition des gammes opératoires pour ce produit,
- la fabrication, par la systématisation des procédés (voir ci-dessous).

L'adoption de la technologie de groupe s'accompagne fréquemment d'une réorganisation des ateliers en *cellules de fabrication* autonomes, spécialisées dans la production de familles de produits distinctes. Idéalement, donc, chaque produit devrait être fabriqué entièrement dans une seule cellule, et tous les produits fabriqués dans une cellule devraient être similaires. L'organisation en cellules permet de retirer certains des bénéfices de la production en ligne:

- spécialisation des équipements (tous les produits d'une famille requièrent les mêmes équipements),
- systématisation des procédés (les gammes opératoires sont similaires),
- réduction des temps de lancement,
- simplification de la gestion, en particulier de l'ordonnancement.

(Notez l'analogie avec le Plant Within a Plant de Skinner.)

Ateliers flexibles (FMS).

L'émergence des ateliers flexibles conduit à repenser la distinction traditionnelle entre

- d'une part : production de masse, automatisation poussée, investissements (coûts fixes) élevés, coûts variables faibles, rigidité,
- et d'autre part : petites séries, main d'œuvre importante, investissements faibles, coûts variables importants, flexibilité.

Remarquons que c'est essentiellement le *taux d'utilisation* des équipements qui détermine leur rentabilité, et peu importe que ce taux soit atteint par la fabrication d'un seul produit ou d'un mix de plusieurs produits (c'est-à-dire par addition de petites séries). Cette dernière approche correspond mieux aux exigences des marchés contemporains (voir Chapitre 2). Mais elle ne peut évidemment être mise en oeuvre que sur des équipements suffisamment flexibles.

Les *machines à commande numérique* (Computer Numerical Control, CNC) présentent une flexibilité nettement accrue par rapport aux machines-outils traditionnelles. Sur une machine CNC, les mouvements et opérations des outils sont contrôlés par un micro-ordinateur. La machine possède par ailleurs un dispositif de changement d'outils automatique. Ces caractéristiques permettent à la machine d'enchaîner très efficacement les opérations successives en réduisant le temps consacré aux temps de réglage et aux changements d'outils.

Un *atelier flexible* (Flexible Manufacturing System, FMS) consiste en l'agencement de plusieurs machines CNC desservies par un système de transport et de chargement automatisés des pièces et des outils, l'ensemble étant contrôlé et coordonné par un ordinateur central. Un tel système combine donc la flexibilité individuelle offerte par les machines CNC à une flexibilité accrue du système de manutention.

L'adoption d'un FMS entraîne des coûts d'investissement élevés (frais fixes) mais peu de frais variables (que ce soit sous forme de main d'œuvre ou de coûts de lancement de la production), et fournit une flexibilité élevée. Il permet donc de rentabiliser l'automatisation de la production des petites séries. Ainsi, Jaikumar et Van Wassenhove (1989) ont observé des FMS produisant plus de 100 types de pièces différents chaque année (dont 25 nouveaux produits) et pour lesquels le taux d'utilisation (pour usinage) est supérieur à 75% du temps total.

Notons cependant quelques obstacles à l'utilisation effective des ateliers flexibles:

- investissements très élevés et amortissement lent (5-7 ans) pour certains systèmes;
- bénéfices difficilement quantifiables: flexibilité et qualité accrues, réduction des délais et coûts de lancement, ...;
- complexité des décisions de gestion liées à de tels systèmes (planification, ordonnancement, ...);

(voir Parker et Lettes 1991).

4.5. Layout (implantation) des centres de production.

Après avoir décidé en termes généraux du type de processus de production à utiliser (par exemple, production en ligne ou job-shop) et des équipements à acquérir, la firme se voit encore confrontée à des problèmes plus spécifiques d'*implantation* des centres de production (machines, postes de travail, départements, ...), c'est-à-dire de localisation relative des centres de production.

En pratique, de nombreux éléments peuvent entrer en ligne de compte dans les décisions d'implantation: contraintes historiques (l'implantation évolue souvent de façon progressive, sans être globalement remise en cause), esthétiques, facteurs humains (l'environnement de travail doit rester agréable), sécurité (les ateliers de peinture seront situés loin de ceux de soudure, afin d'éviter l'explosion des matières volatiles), etc. Les modèles mathématiques que nous allons présenter ci-dessous devront donc être considérés comme des modèles d'*aide* à la décision plutôt que comme des systèmes *automatisés* de prise de décision.

Nous considérerons seulement ici des situations idéales où le système doit être conçu à partir de rien et où l'objectif à optimiser peut-être exprimé quantitativement. De telles situations se rencontrent par exemple lors de la construction de nouveaux bâtiments ou lors de la réorganisation de systèmes existants. Les modèles d'optimisation les plus courants tiennent compte d'un objectif de *minimisation des coûts*. Ces coûts peuvent généralement être interprétés comme des coûts de manutention (transfert, transport) ou de déplacement.

Nous considérerons successivement les problèmes qui se posent lors de l'implantation d'une chaîne d'assemblage, puis de celle des départements d'un atelier.

4.5.1. Equilibrage des chaînes d'assemblage.

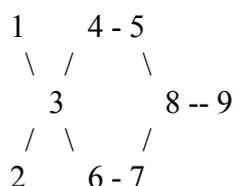
Par nature, une ligne de production est consacrée à la fabrication d'un seul produit (avec variantes éventuelles) et doit donc exécuter des séquences d'opérations (presque) identiques sur chaque unité de ce produit. Il est donc logique que l'implantation d'un tel système soit dictée par les séquences d'opérations nécessaires à la fabrication de ce seul produit. On parle alors d'une *organisation par produit* (*product layout*) du processus de production.

Typiquement, les postes de travail d'un système organisé par produit sont placés en ligne et chaque unité du produit passe de station en station en subissant une transformation à chaque étape. Des exemples extrêmes de ce type d'organisation se trouvent dans les industries à processus de production continu, où la séquence des opérations à effectuer est souvent totalement rigide (comme dans l'industrie chimique ou sidérurgique). Par contre, dans le cas de chaînes d'assemblage, l'ordre des postes de travail, ou même des opérations à effectuer, n'est généralement pas complètement déterminé par les spécifications techniques du produit. Illustrons ce propos par un exemple simple.

Exemple: Liste des opérations à effectuer sur une chaîne d'assemblage automobile:

1. Assemblage de la carrosserie
2. Assemblage du bloc moteur
3. Assemblage carrosserie - bloc moteur
4. Placement transmission
5. Placement roues
6. Placement sièges
7. Portes
8. Installation vitres
9. Contrôle finition

Contraintes de précédence (antériorité) à respecter entre les opérations :



Plusieurs définitions des postes de travail sont compatibles avec les contraintes exprimées: de fait, on pourrait adopter

- 1) un seul poste regroupant toutes les opérations,
- 2) un poste par opération : 1 - 2 - 3 - ... - 9,
- 3) le regroupement de plusieurs opérations par poste: par exemple (1,2,3) - (4,6) - (5,7) - (8) - (9).

Chacune de ces solutions comporte ses inconvénients (ainsi que des avantages absents des autres solutions): difficulté d'apprentissage et lancements répétitifs de la production pour la solution (1), monotonie des tâches à effectuer et lourdeur des manutentions pour la solution (2), etc.

Par ailleurs, les diverses solutions se caractérisent aussi par des taux de production variés. Définissons à présent ce concept de façon précise: le *taux de production* T d'une chaîne d'assemblage est le nombre moyen d'unités qu'elle produit par unité de temps.

Un concept intimement lié au précédent est celui de temps de cycle: le *temps de cycle* C (cycle, cycle time, etc) d'une chaîne est le temps écoulé, en moyenne, entre la production de deux unités successives. On a donc la relation : $T = 1/C$.

Le taux de production et le temps de cycle sont évidemment déterminés par la durée des opérations individuelles à exécuter, mais aussi par l'implantation de la chaîne et en particulier par le regroupement des opérations aux postes de travail successifs.

Exemple: Assemblage automobile (suite).

Durée des opérations (en minutes):

1. Assemblage de la carrosserie : 17
2. Assemblage du bloc moteur : 11
3. Assemblage carrosserie - bloc moteur : 8
4. Placement transmission : 5
5. Placement roues : 4
6. Placement sièges : 6
7. Portes : 8
8. Installation vitres : 5
9. Contrôle finition : 3

On calcule facilement le temps de cycle pour les 3 layouts définis plus haut:

- 1) un seul poste: 67 minutes.
- 2) un poste par opération: 17 minutes.
- 3) regroupement (1,2,3) - (4,6) - (5,7) - (8) - (9) : 36 minutes.

Remarquons que le temps de cycle d'une chaîne de production est déterminé par le temps de travail (la charge) du poste de travail le plus lent, ou *poste goulot* (bottleneck).

Nous sommes maintenant équipés pour définir formellement le *problème d'équilibrage d'une chaîne d'assemblage*: étant donné

- les différentes opérations à effectuer et leurs durées respectives p_1, \dots, p_m
- les relations de précedence entre opérations,
- un temps de cycle « cible » C ,

affecter les opérations à un nombre minimum de postes de travail placés en ligne de telle sorte que les relations de précedence soient respectées et que le temps opératoire total n'excède C pour aucun poste.

Avant d'envisager la solution de ce problème, formulons d'abord quelques remarques sur sa définition.

1. Dans cette formulation, le temps de cycle, et donc le taux de production, sont fixés de façon exogène. Cette hypothèse est raisonnable dans le cadre d'une production de masse pour stock, où la demande est souvent relativement stable. (On considère cependant parfois une formulation alternative dans laquelle le nombre de stations disponibles est fixé a priori et le temps de cycle doit être minimisé).

2. Le nombre de stations peut être interprété comme un substitut approximatif des coûts de manutention: de longues chaînes impliquent de nombreux transferts de matériel. Cependant, une réduction de la longueur de la chaîne implique généralement une augmentation du temps de cycle (voir exemple ci-dessus)!! Ceci explique que l'on incorpore une contrainte sur le temps de cycle dans le modèle.
3. Le *temps mort* (idle time) TM est le temps total durant lequel les postes de travail sont inoccupés pendant un cycle de production. Pour une solution comportant N postes,

$$TM = N \times C - \sum_i p_i.$$

Lorsque le temps de cycle C et les temps opératoires p_i sont considérés comme des données, minimiser N revient à minimiser TM . L'objectif du problème d'équilibrage peut donc également être vu comme un objectif de minimisation des temps morts.

4. Dans les situations pratiques, on rencontre souvent des contraintes additionnelles; par exemple, deux opérations ne peuvent pas être réalisées par le même poste (incompatibilité ou encombrement des équipements, etc), certains postes ne peuvent pas effectuer certaines opérations, etc.

Voyons maintenant comment aborder la solution du problème d'équilibrage.

Tout d'abord, une limite inférieure L sur le nombre de postes optimal peut être calculée par la formule

$$L = \sum_i p_i / C .$$

Mais cette borne inférieure n'est pas toujours réalisable, comme le montre l'exemple suivant.

Exemple: Assemblage automobile (suite).

Si on se donne un temps de cycle de 17 minutes (ce qui correspond à la durée de la plus longue opération), alors:

$$L = \sum_i p_i / C = 67/17 = 3,9.$$

Donc, toute solution impliquera au minimum 4 postes de travail. Mais en fait, pour l'exemple considéré, on peut facilement se persuader qu'il n'y a pas de solution comportant moins de 5 postes.

En pratique, le problème d'équilibrage est un problème difficile à résoudre de façon optimale lorsque le nombre d'opérations devient trop important (il est NP-difficile). Par conséquent, on ressort souvent à des méthodes approchées, ou heuristiques, pour l'attaquer. La plupart des heuristiques d'équilibrage de chaînes d'assemblage appliquent une *stratégie gourmande*. Elles reposent sur le schéma de résolution suivant.

Etape 1. Définir un coefficient de priorité c_i pour chaque opération i (par exemple, la durée de l'opération, ou la somme des durées des opérations qui suivent, ...).

Etape 2. Si les postes 1, ..., j ont déjà été définis, et certaines opérations ont déjà été affectées au poste $j+1$, alors:

i) parmi les opérations disponibles (dont tous les prédécesseurs ont été affectés) et pouvant encore être affectées au poste $j+1$ sans excéder C , affecter celle ayant le plus haut coefficient de

priorité c_i ;

ii) si plus aucune opération ne peut être affectée au poste $j+1$ sans excéder C , créer un poste $j+2$

et lui affecter l'opération disponible ayant le plus haut coefficient de priorité c_i .

Les coefficients de priorité peuvent être définis de différentes façons. Par exemple, on peut définir c_i comme étant le temps opératoire de l'opération i (c'est-à-dire, $c_i = p_i$). Alternativement, on peut choisir c_i égal au nombre d'opérations ultérieures à l'opération i dans le graphe de précedence (l'idée est ici de donner la plus haute priorité aux opérations qui doivent encore être suivies par de nombreuses autres opérations). La méthode du *Ranked Positional Weight* repose sur une variante de cette idée: on y définit c_i comme la somme des durées de toutes les opérations ultérieures à l'opération i . Cette heuristique fournit souvent des solutions de bonne qualité. Enfin, la méthode *COMSOAL* (COMputer Method of Sequencing Operations for Assembly Lines) donne simplement la même priorité à toutes les opérations, ce qui revient à choisir l'affectation de l'opération suivante au hasard! Le manque de sophistication évident de cette méthode est compensé par la force brutale: l'heuristique génère successivement un grand nombre de solutions et ne retient que la meilleure. De nombreuses autres heuristiques ont également été proposées dans la littérature scientifique (voir le logiciel QSOM proposé par Chang 1989).

Exemple: Assemblage automobile (suite).

Pour un temps de cycle $C = 17$ ($L = 4$), on obtient facilement des solutions optimales à 5 postes, impliquant donc 18 minutes de temps mort.

Notons que, pour un temps de cycle de 19 minutes, on peut réduire le nombre de postes à 4 et le temps mort à 9 sec. (dans la pratique, il est très utile d'explorer de telles solutions alternatives).

Avec un temps de cycle de 25 sec ($L = 3$), l'heuristique du « plus grand nombre de successeurs d'abord » génère une solution à 4 postes, tandis que RPW donne une solution optimale à 3 postes.

Un cas pratique d'équilibrage de chaîne, illustrant bien la plupart des propos ci-dessus, est décrit par

Agnetis, Ciancimino, Lucertini et Pizzichella (1992). Le graphe de précedence, le layout de l'usine et la composition des stations de travail avant la réalisation de l'étude de cas sont donnés en annexe à la fin de ce chapitre. En particulier, on observe dans cette étude des paires d'opérations incompatibles, ou 'forcées' dans la même station. L'implantation originale compte 15 stations, mais est mal équilibrée (temps de cycle de 78 secondes, avec 49 secondes de temps mort à la station de traitement des radiateurs, résultant en encours importants). Le taux de production de 692 unités par jour (2 shifts de 7,5 heures) excède légèrement la demande

moyenne, qui est de 625 unités. Une meilleure solution a été obtenue par les auteurs: elle comporte 20 postes de travail, mais ne requiert plus qu'un shift (7,5 heures par jours) pour atteindre le taux de production de 625 unités.

4.5.2. Job-shop: Localisation des départements.

Dans un job-shop, les produits subissent des séquences d'opérations très diverses, et il n'est donc plus possible de se laisser guider par une 'logique de produit' pour décider de l'implantation des centres de production. En pratique, on regroupe souvent les machines ou les équipements en départements sur base de leur fonction. On parle dans ce cas d'*organisation fonctionnelle* (process layout). (Remarquons cependant que, comme nous l'avons brièvement mentionné dans la Section 4.4, la technologie de groupe et les cellules de fabrication offrent une approche alternative - souvent très performante - à l'implantation des job-shops)

Exemples:

Atelier mécanique : département fraiseuses, tours, foreuses ..

Hôpital : radiographie, électroencéphalogrammes, plâtres, ...

Dans une optique de minimisation des coûts de manutention, le *problème de localisation des départements* peut alors être modélisé comme suit: *étant donné*

- *N départements à localiser,*
- *une superficie totale disponible S,*
- *la superficie s_i requise par chaque département i ,*
- *le coût de transport (par unité de matériel et par unité de distance) entre chaque paire de départements i et j , soit c_{ij} ,*
- *le flux de matériel (nombre d'unités) entre les départements i et j , soit f_{ij} ,*

attribuer à chaque département i un emplacement (connexe) de superficie s_i ($i = 1, \dots, N$) dans l'aire disponible, de façon à minimiser les coûts totaux de transport.

La matrice des flux f_{ij} peut être par exemple estimée sur base de données historiques dans le cas d'un système à réorganiser, ou sur base de prévisions dans le cas d'un système à créer. La distance d_{kl} entre les emplacements k et l peut être estimée par la distance (euclidienne ou rectilinéaire) entre les centres de gravité des deux emplacements. Les coûts totaux de transport sont alors calculés par la formule

$$CT = \sum_{i=1}^N \sum_{j=1}^N c_{ij} f_{ij} d_{k(i)l(j)}$$

où $k(i)$ et $l(j)$ sont les centres de gravité des emplacements attribués aux départements i et j respectivement.

Le problème de localisation des départements, tel que nous venons de le formuler, est réputé pour être extrêmement difficile à résoudre de façon optimale (il est NP-difficile). En pratique, on recourt donc à des méthodes de résolution heuristiques. Les méthodes les plus efficaces sont celles procédant par échanges successifs: au départ d'une solution quelconque, on essaye de

déplacer un département vers un emplacement libre ou d'échanger les emplacements de plusieurs départements (généralement, deux à deux ou trois à trois); la modification n'est entérinée que si elle entraîne une diminution du coût total CT. La méthode CRAFT (Computerized Relative Allocation of Facilities Technique) procède de cette façon (voir le logiciel QSOM (Chang 1989) pour une implantation).

Bibliographie.

A. Agnetis, A. Ciancimino, M. Lucertini et M. Pizzichella, Balancing flexible lines for car components assembly, Rapport 20.92, Dipt. di Informatica e Sistemistica, Università degli Studi di Roma 'La Sapienza', 1992.

Y.-L. Chang, *QSOM - Quantitative Systems for Operations Management*, Prentice-Hall, Englewood Cliffs, 1989.

A. Courtois, M. Pillet et C. Martin, *Gestion de Production*, Les Editions d'Organisation, Paris, 1989.

V. Giard, *Gestion de la Production*, 2ème édition, Economica, Paris, 1988.

R. Jaikumar et L.N. Van Wassenhove, A production planning framework for flexible manufacturing systems, *Journal of Manufacturing and Operations Management* Vol. 2 (1989) pp.52-79.

P. Kotler et B. Dubois, *Marketing Management*, 8ème édition, Publi union, 1994.

J.O. McClain, L.J. Thomas et J.B. Mazzola, *Operations Management*, 3ème édition, Prentice-Hall, Englewood Cliffs, 1992.

T. Parker et T. Lettes, Is accounting standing in the way of flexible computer-integrated manufacturing?, *Management Accounting*, January 1991, pp. 34-38.

M.H. Safizadeh, L.P. Ritzman, D. Sharma et C. Wood, An empirical analysis of the product-process matrix, *Management Science*, Vol. 42, November 1996, pp. 1576-1591.

How manufacturing can make low-tech products high-tech, *Harvard Business Review*, September-October 1995, p. 98.

CHAPITRE 5

PLANIFICATION A MOYEN TERME: LE PLAN AGREGÉ DE PRODUCTION

5.1. Introduction et concepts de base.

Nous quittons le domaine des décisions stratégiques, portant sur le long terme et/ou la conception du système productif, pour entrer dans celui de la planification tactique, portant sur le moyen terme. Dans ce cadre de décision, le portefeuille de produits et le processus de production doivent être considérés comme des *données* (même si certaines modifications marginales de capacité peuvent encore être envisagées). La question qui se pose est de savoir comment *utiliser* au mieux le système existant. Plus précisément:

Le rôle de la planification agrégée de la production est de planifier globalement le volume et l'utilisation des ressources ainsi que les niveaux de production et de stocks à atteindre au cours de chaque sous-période afin de satisfaire au mieux la demande prévisionnelle sur un horizon à moyen terme.

Quelques explications s'imposent ...

Horizon du plan.

Comment faut-il fixer les limites temporelles du plan de production agrégé ?

Une première limite supérieure à l'horizon du plan a déjà été évoquée ci-dessus: on s'intéresse à un horizon sur lequel les décisions stratégiques doivent être considérées comme irrévocables. Un autre critère à prendre en compte est lié à la fiabilité des prévisions de demande: cette fiabilité a en effet tendance à décroître lorsque l'horizon du plan s'allonge.

Afin de déterminer une limite inférieure à l'horizon du plan, notons que celui-ci doit être suffisamment long pour permettre un certain lissage de la production obtenu en jouant sur la constitution de stocks et les modifications temporaires de personnel. En particulier, les fluctuations saisonnières (de demande, de disponibilité des ressources, etc) doivent pouvoir être reflétées dans le plan.

→ *En pratique, on utilise souvent un horizon d'un an.*

Cet horizon est découpé en plusieurs sous-périodes. Le plan devra donc spécifier les différentes décisions prises pour chacune de ces sous-périodes. Le choix des sous-périodes est guidé par les principes suivants:

- de trop longues sous-périodes effacent les fluctuations des paramètres (notamment de la demande), et entraînent donc une perte d'information;
- des sous-périodes trop courtes augmentent les erreurs d'estimation et la complexité du problème (en particulier, le nombre de variables de décision).

→ *En pratique, on utilise souvent des sous-périodes de 1, 2 ou 3 mois.*

Comme c'est le cas pour tous les plans d'action établis sur base de données incertaines, la qualité du plan agrégé de production tend à diminuer au fil du temps. Par conséquent, ce plan doit être soumis à des révisions périodiques qui permettent de tenir compte de l'évolution réelle de la situation entre deux révisions (changement de conjoncture, mauvaise évaluation de la demande, changements de productivité, ...).

Demande.

Les produits mis sur le marché étant connus, la demande des sous-périodes successives peut – en principe – être estimée par le département marketing. Remarquons immédiatement que la qualité du plan de production sera très largement influencée par la qualité de ces prévisions, auxquelles une attention toute particulière devrait donc être prêtée avant d'entrer dans le processus de planification proprement dit.

L'obtention des prévisions peut être basée sur différentes sources d'information (données historiques, études de marchés, analyses sectorielles, commandes enregistrées, expérience des distributeurs, etc) et peut faire appel à différentes techniques statistiques (régression, analyse de séries chronologiques, etc.). De nombreux logiciels de gestion de la production offrent des modules d'aide à la prévision de la demande mettant à disposition de l'utilisateur une large panoplie d'outils statistiques. Pour la plupart des entreprises, l'exercice de prévision de la demande reste cependant extrêmement difficile à mettre en œuvre, pour de multiples raisons. En particulier, les séries statistiques obtenues dans le passé ne fournissent pas nécessairement une base fiable pour prédire le futur (la conjecture change, des produits concurrents apparaissent, etc.), la demande peut être très variable, et donc difficile à prévoir avec précision (en particulier pour des produits innovants, ou soumis à des effets de mode), les distributeurs ou les vendeurs peuvent avoir un intérêt « politique » à maquiller leurs prévisions (par exemple pour s'assurer de disposer de stocks suffisants), etc. Souvent, il sera nécessaire pour l'entreprise de développer des relations de partenariat étroit avec ses clients – impliquant typiquement l'échange d'informations sur les chiffres de ventes ou sur l'état des stocks - pour parvenir à surmonter ces difficultés de prévision. Bien qu'il s'agisse là d'un sujet d'une grande importance pratique et d'une grande actualité, nous ne le développerons pas davantage dans le cadre de ce cours introductif (voir par exemple Fisher 1997 pour quelques illustrations).

Par ailleurs, dans le cadre adopté ici, le rôle de la production est en premier lieu de satisfaire la demande, vue comme déterminée de façon exogène. Nous considérerons donc principalement des stratégies *réactives*, qui acceptent la demande comme une contrainte à respecter.

En fait, certaines possibilités d'influencer ou de déplacer la demande sont parfois prises en compte dans la planification à moyen terme: publicité ou rabais dans les périodes de basse demande (compagnies aériennes, chemins de fer, hôtels, ...) ou service sur rendez-vous (cliniques, ...). De telles stratégies *proactives* sont notamment très utilisées par les sociétés de service, qui n'ont pas la possibilité d'utiliser les stocks pour lisser la production sur de longues périodes. Dans certains cas, l'entreprise peut également accepter la possibilité de ne satisfaire qu'une partie de la demande.

Ressources.

La ressource la plus flexible sur l'horizon considéré est celle constituée par le personnel de production (quoique certaines décisions d'acquisition d'équipement ou de sous-traitance peuvent aussi être incluses dans le plan agrégé de production, si les options stratégiques et/ou l'horizon de planification le permettent). Les décisions à prendre concernent donc essentiellement les volumes d'embauche, de licenciement, de prestation d'heures supplémentaires, etc. Chacune de ces décisions peut avoir sa portée limitée a priori par des considérations stratégiques ou d'autres contraintes (par exemple, une politique de maintien de l'emploi, des conventions sectorielles, etc).

Niveaux de production et de stocks.

La demande de chaque sous-période peut être satisfaite de différentes façons: par la production effectuée au cours de cette sous-période, au cours des périodes antérieures (livraisons sur stocks) ou au cours des périodes ultérieures (livraisons en retard). En combinant ces diverses possibilités, la firme peut niveler son niveau d'activité sur l'horizon à moyen terme et donc stabiliser le volume de ressources utilisées. On découvre là un des rôles essentiels des stocks: celui du lissage de la production (remarquons que ce rôle est utile même dans un univers parfaitement déterministe et parfaitement connu; il faut donc bien le distinguer de celui qui consiste à protéger la firme contre les incertitudes de la demande, les aléas des livraisons de fournitures, etc).

Dans ce chapitre, nous supposerons systématiquement que l'entreprise n'a pas recours aux livraisons tardives: la demande doit être satisfaite à la fin de la période où elle est exprimée. Cette contrainte peut par exemple traduire un des objectifs stratégiques de l'entreprise – disponibilité permanente des produits, réduction des délais de livraison – et marquer ainsi la coordination désirée entre deux niveaux de planification (stratégique et long terme). Alternativement, le lecteur pourra n'y voir qu'un simple souci de clarté didactique: des possibilités de livraisons tardives pourraient facilement être prises en considération au prix d'un certain alourdissement de la présentation.

Remarquons enfin que, lorsque la demande est connue avec certitude, il suffit de connaître le nombre d'unités produites au cours de chaque sous-période pour en déduire le niveau des stocks à la fin de chaque sous-période. Réciproquement, l'état des stocks à la fin de chaque sous-période détermine complètement les volumes de production correspondants. Par conséquent, le plan de production agrégé peut être vu comme fixant uniquement l'une des deux quantités – stocks ou production. Cette interprétation peut se révéler utile dans certaines circonstances.

En conclusion: la planification agrégée réalise la coordination à moyen terme entre les départements marketing, personnel et production; le plan agrégé est donc généralement le fruit d'une collaboration entre ces trois départements, qui doivent obligatoirement l'approuver, et l'accepter comme une contrainte à respecter dans leurs activités à plus court terme.

En quoi le plan est-il «agrégé »?

Au niveau de décision considéré, il règne trop d'incertitudes pour vouloir s'encombrer de détails qui compliqueraient inutilement la prise de décision. On cherche donc seulement ici à planifier « dans les grandes lignes », ce qui se traduit par:

- l'agrégation des décisions: les produits sont agrégés en *familles de produits*, les ressources en grandes catégories (de personnel, d'équipements, etc), les sous-périodes sont relativement longues;
- l'omission de nombreux détails: produits peu demandés, ressources peu utilisées, etc.

L'agrégation des décisions permet de simplifier considérablement la formulation, la résolution et l'interprétation des modèles (moins de données à collecter, moins de calculs à effectuer, moins de résultats à analyser). Par ailleurs, elle augmente généralement la qualité des prévisions de demande ainsi que l'estimation d'autres paramètres. Par exemple, si les différents produits agrégés au sein d'une même famille ont des demandes corrélées négativement, alors l'estimation de leur demande agrégée sera plus précise que celle des demandes individuelles (plus précisément, la variance de l'estimation sera inférieure à la somme des variances des demandes individuelles).

5.2. Analyse graphique des options simples.

Dans cette section, nous allons brièvement présenter une méthode graphique d'élaboration du plan de production agrégé. Pour simpliste qu'elle puisse paraître, cette méthode permet de visualiser rapidement certaines des options qui s'offrent à l'entreprise et possède par ailleurs d'intéressantes vertus didactiques. Nous présenterons par la suite quelques approches plus sophistiquées du même problème.

Exemple: Considérons un seul produit (éventuellement après agrégation) et un horizon de 6 mois, où chaque mois comporte 20 jours. On dispose des données reprises dans le tableau ci-dessous.

Mois	Demande	Stock minimum
1	70	15
2	100	20
3	120	30
4	60	15
5	50	10

Plus généralement, l'horizon considéré comporte les périodes 1, 2, ..., T. Pour chaque période, on a estimé la demande prévisionnelle et le stock minimum à posséder en fin de période. Rappelons que la firme n'a pas recours aux livraisons tardives. On supposera par ailleurs, afin de simplifier la discussion, que la position de stock est nulle en début d'horizon. Les *besoins cumulés* à la période t sont alors définis comme suit:

$BC(t) =$ demande cumulée jusqu'à la période t + stock minimum de la période t .

Un plan de production agrégé peut être vu comme une fonction croissante $PC(t)$, où $PC(t)$ désigne la production cumulée jusqu'à la période t ($t = 1, 2, \dots, T$). Le plan $PC(t)$ est *admissible* si, à chaque période t

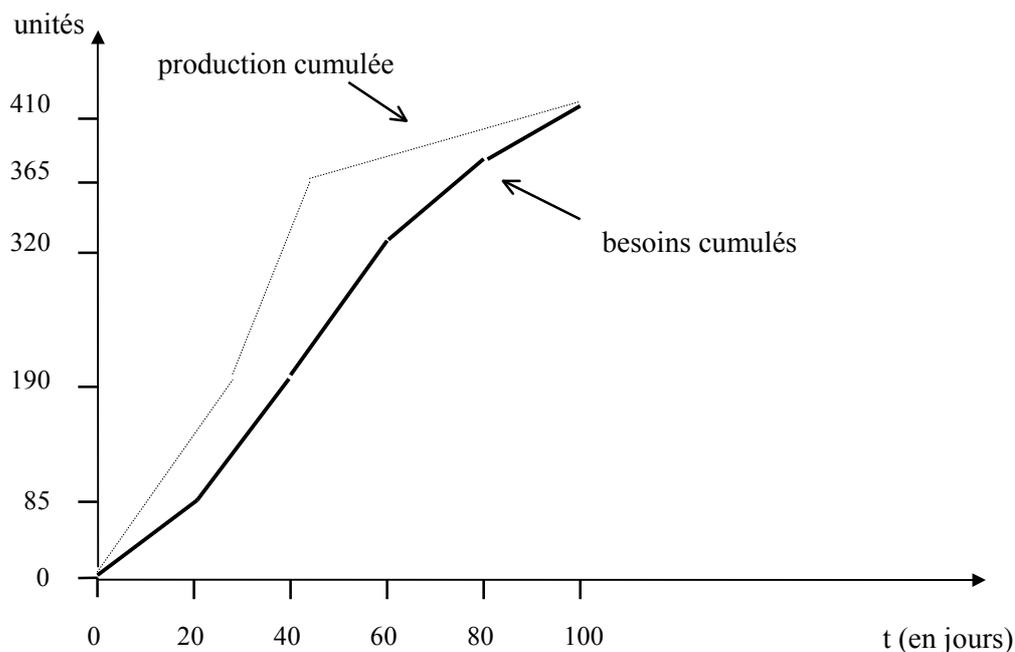
$$BC(t) \leq PC(t) = \text{production cumulée jusqu'à la période } t. \quad (1)$$

Remarquons que cette condition d'admissibilité traduit le fait que la demande doit être entièrement satisfaite sans recourir aux livraisons tardives. Les ruptures de stocks ne sont donc pas autorisées

Exemple (suite): Calculons la demande cumulée et les besoins cumulés pour l'exemple numérique précédent..

Mois	Demande cumulée	Besoins cumulés
1	70	85
2	170	190
3	290	320
4	350	365
5	400	410

Le graphique suivant fournit une représentation de l'évolution des besoins cumulés en fonction du temps (courbe épaisse).



Toute courbe croissante PC partant de l'origine et située au dessus de la courbe BC représente un plan admissible. La différence $PC(t) - BC(t)$ représente l'excédent de stock par rapport au stock minimum requis à la période t.

Dans la pratique, les entreprises se basent souvent sur cette représentation graphique pour développer leur plan de production agrégé. Illustrons cette approche par la discussion de deux stratégies de base simples, mais de natures extrêmes, qui s'offrent à l'entreprise.

Stratégie 1. Suivre la demande en modifiant les effectifs. Le plan de production est établi de telle sorte que

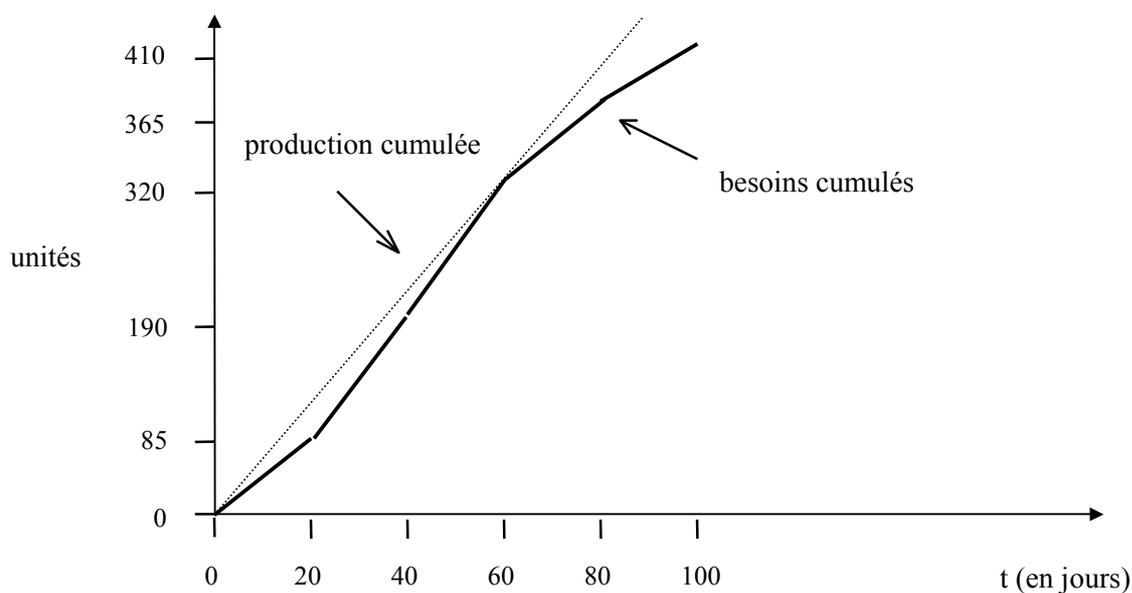
$$PC(t) = BC(t) \text{ pour tout } t, \quad (2)$$

ce qui revient à imposer que la production de chaque période soit exactement égale à la demande de cette période. Le stock en fin de période est donc toujours égal au stock minimum requis.

Ce type de stratégie est largement utilisée dans les services, où les possibilités de stockage sont inexistantes (par exemple, dans le secteur HORECA), mais est également observée dans l'industrie (secteur automobile). En termes de coût, notons qu'elle minimise (évidemment !) les coûts de possession de stock.

Stratégie 2. Minimiser les fluctuations de personnel. Dans sa version la plus simple, cette stratégie implique de produire à taux constant pendant l'horizon du plan. Formellement, l'hypothèse de production à taux constant signifie que $PC(t)$ est une fonction linéaire du temps. On peut donc écrire: $PC(t) = \alpha t$, où α peut être interprété comme un taux de production. Graphiquement, le plan minimisant les fluctuations de personnel correspond alors à un segment de droite situé au-dessus de la courbe BC.

Par exemple, à l'époque 0 (début de l'horizon de planification), on peut choisir un taux de production initial α_1 aussi faible que possible sous réserve que la droite $PC_1(t) = \alpha_1 t$ satisfasse à la première condition d'admissibilité: $BC(t) \leq PC_1(t)$ pour tout t.



Cette stratégie tend évidemment à minimiser les coûts liés aux fluctuations de personnel (coûts monétaires ou sociaux), mais entraîne généralement l'accumulation de stocks importants. De façon alternative, l'entreprise pourrait choisir de produire à taux constant en se contentant d'exiger que la production cumulée en fin d'horizon soit égale aux besoins cumulés: $PC(T) = BC(T)$, sans requérir la condition d'admissibilité (1). Cette stratégie implique d'accepter les ruptures de stocks.

5.3. Un modèle de programmation linéaire.

Comme on l'a remarqué ci-dessus, l'analyse graphique permet de visualiser facilement les différentes options s'offrant à l'entreprise et permet même d'élaborer un plan de production dans des situations relativement simples (par exemple, très agrégées). Ce type d'analyse est donc fréquemment utilisée en pratique.

Il faut bien voir, cependant, que l'analyse graphique ne permet pas l'*optimisation* des coûts entraînés par le plan, ni de tenir facilement compte de produits multiples, d'options mixtes (combinaisons d'heures supplémentaires, d'embauche, de campagnes publicitaires déplaçant la demande, etc) ou de contraintes additionnelles (contraintes budgétaires, capacité, ...). On comprend donc l'intérêt d'une approche plus puissante et plus formalisée de la planification agrégée.

Une approche optimisante célèbre a été développée dans les années cinquante par Holt, Modigliani, Muth et Simon (1964). Ces chercheurs ont élaboré un modèle de planification agrégée pour le compte d'une grande entreprise désirant niveler sa production et réduire ses stocks (une description de cette application se trouve en annexe à la fin du présent chapitre). Les variables considérées dans le modèle HMMS décrivent les niveaux de production, de stocks, de personnel, et d'heures supplémentaires à prester. L'objectif de minimisation des coûts est traduit dans un modèle d'optimisation quadratique sans contraintes. Les conditions d'optimalité du premier ordre donnent lieu à des *règles de décision linéaires*, qui définissent les niveaux de production et d'emploi optimaux à chaque période sous forme de fonctions linéaires de ces niveaux aux périodes précédentes et de la demande prévisionnelle.

Le modèle HMMS permet l'optimisation des coûts, nécessite peu de calculs, et génère des règles d'utilisation facile. Ces avantages ont contribué à sa popularité. Par contre, il connaît aussi une des limitations fondamentales de l'approche graphique: il est peu flexible, en ce sens qu'il n'admet pas l'addition de contraintes (plus précisément, le modèle d'optimisation quadratique qui en résulterait est beaucoup plus difficile à résoudre que le modèle HMMS de base).

Les modèles de programmation linéaire, à l'inverse, fournissent une approche optimisante qui peut être plus facilement enrichie par la considération de facteurs secondaires. Par exemple, les limitations de capacité de certains centres de production peuvent être naturellement intégrées dans le modèle sous forme de contraintes d'inégalité. Ceci autorise un ajustement charge-capacité - 'rough-cut capacity planning' - qui, dans le modèle HMMS ou dans l'approche graphique, doit être réalisé a posteriori. Cette souplesse de modélisation, alliée aux autres avantages bien connus de la programmation linéaire (disponibilité des logiciels, résolution

efficace de programmes de grande taille, analyse de sensibilité) font de la programmation linéaire un instrument de choix pour la planification agrégée de la production.

Décrivons brièvement la structure de base d'un modèle de programmation linéaire pour l'élaboration du plan de production agrégé. Dans cette formulation, nous adoptons une convention de modélisation du temps largement répandue en gestion de la production et qui sera également utilisée dans les autres chapitres de ce syllabus. *Pour chaque période, nous supposons que la production a lieu en début de période et que les livraisons-clients sont effectuées en fin de période. En d'autres termes, toutes les unités produites à la période t sont disponibles pour satisfaire la demande de cette période. Par ailleurs, l'état de stock est déterminé après livraison aux clients, c'est-à-dire à la fin de chaque période t.*

Pour clarifier la présentation du modèle, nous adoptons également les conventions typographiques suivantes: les variables de décision sont notées en MAJUSCULES, les paramètres (c'est-à-dire les données du modèle) en minuscules, et les expressions plus complexes (fonctions, inégalités, etc) en *italiques*.

Hypothèses de base: on considère

- N produits (ou, plus précisément, N familles de produits)
- T périodes de planification
- une seule catégorie de personnel.

Variables de décision: pour $j = 1, \dots, N$ et $t = 1, \dots, T$,

PROD[j,t] :	nombre d'unités du produit j produites <u>en début</u> de période t
STOCK[j,t] :	nombre d'unités du produit j en stock à <u>la fin</u> de la période t
PERS[t] :	nombre d'ouvriers employés à la période t
EMBAUCHE[t] :	nombre de nouveaux ouvriers embauchés au début de la période t
LICENC[t] :	nombre d'ouvriers licenciés au début de la période t
SUPP[t] :	nombre d'heures supplémentaires prestées pendant la période t

Schéma du modèle: le modèle possède la structure de base suivante:

$$\text{minimiser } (\text{coûts de personnel}) + (\text{coûts de possession}) \quad (3)$$

sous les contraintes :

$$\text{admissibilité du plan} \quad (4)$$

$$\text{calcul des stocks} \quad (5)$$

$$\text{calcul des effectifs de personnel} \quad (6)$$

$$\text{contraintes de capacité} \quad (7)$$

$$\text{contraintes additionnelles.} \quad (8)$$

Fonction de coût (3):

coûts de personnel :=

salaires : $\sum_t \text{ salaire}[t] \text{ PERS}[t]$

(où $\text{salaire}[t]$ est le salaire par ouvrier pour la durée de la période t ; par exemple, $\text{salaire}[t] = (\text{salaire horaire}) \times (\text{nombre d'heures dans la période } t)$)

+ *frais d'embauche* : $\sum_t \text{ fraisemb}[t] \text{ EMBAUCHE}[t]$

+ *frais de licenciement* : $\sum_t \text{ fraislic}[t] \text{ LICENC}[t]$

+ *heures supplémentaires* : $\sum_t \text{ tarifsup}[t] \text{ SUPP}[t]$

(où $\text{tarifsup}[t] = \text{coût salarial d'une heure supplémentaire prestée à la période } t$).

coûts de possession des stocks :=

$\sum_j \sum_t \text{ coûtstock}[j,t] \text{ STOCK}[j,t]$

(où $\text{coûtstock}[j,t]$ est le coût encouru pour la détention d'une unité du produit j pendant la période t ; souvent exprimé sous forme d'un pourcentage du prix de revient).

Remarquons que nous n'avons pas incorporé les coûts de production dans la fonction objectif. Cette omission est justifiée si l'on considère que les coûts de production sont indépendants des décisions prises dans le cadre du plan de production à moyen terme (parce que le nombre total d'unités à produire est entièrement déterminé par la demande, qui est elle-même fixée de façon exogène). Dans un tel cas, les coûts de production peuvent être traités comme une constante du modèle. Dans le cas contraire, il faudrait intégrer ces coûts dans l'objectif à minimiser.

Contraintes d'admissibilité du plan (4):

L'admissibilité du plan se traduit par les contraintes *besoins cumulés* \leq *production cumulée* (voir contrainte (1), Section 5.2) ou encore, sous forme algébrique: pour chaque produit $j = 1, \dots, N$ et pour chaque période $t = 1, \dots, T$,

$$\sum_{s=1}^t \text{dem}[j,s] + \text{stock_min}[j,t] \leq \sum_{s=1}^t \text{PROD}[j,s] + \text{stock_init}[j] \quad (9)$$

où $\text{dem}[j,s]$ est la demande estimée pour le produit j durant la période s , $\text{stock_min}[j,t]$ est le stock minimum requis en fin de période t et $\text{stock_init}[j]$ est le stock initial du produit j (au début de la période 1).

Calcul des stocks (5):

Le stock disponible à la fin de chaque sous-période est déterminé par la différence entre la production cumulée (ajoutée au stock initial) et la demande cumulée jusqu'à cette période (voir Section 5.1): pour chaque produit $j = 1, \dots, N$ et pour chaque période $t = 1, \dots, T$,

$$\text{STOCK}[j, t] = \sum_{s=1}^t \text{PROD}[j, s] + \text{stock_init}[j] - \sum_{s=1}^t \text{dem}[j, s] .$$

(10)

De façon équivalente, les contraintes (9)-(10) peuvent être reformulées sous la forme de contraintes d'équilibre des flux (flux entrant = flux sortant) et de contraintes sur les quantités stockées: pour $j = 1, \dots, N$ et $t = 1, \dots, T$,

$$\text{STOCK}[j, t] = \text{STOCK}[j, t-1] + \text{PROD}[j, t] - \text{dem}[j, t],$$

(11)

$$\text{STOCK}[j, 0] = \text{stock_init}[j],$$

(12)

$$\text{STOCK}[j, t] \geq \text{stock_min}[j, t].$$

(13)

L'équation (11) traduit la remarque formulée dans la Section 5.1 selon laquelle les stocks en fin de période sont complètement déterminés par la demande et les quantités produites.

Calcul des effectifs (6):

Le nombre d'ouvriers occupés à la période $t = 0$ (c'est-à-dire en début d'horizon) est donné: notons le pers_init. La contrainte suivante définit alors le nombre d'ouvriers occupés à la période 1:

$$\text{PERS}[1] = \text{pers_init} + \text{EMBAUCHE}[1] - \text{LICENC}[1]$$

et similairement, pour chaque sous-période $t = 2, \dots, T$,

$$\text{PERS}[t] = \text{PERS}[t-1] + \text{EMBAUCHE}[t] - \text{LICENC}[t].$$

Contraintes de capacité (7):

Pour chaque période $t = 1, \dots, N$, on doit avoir :

$$\text{heures productives requises en } t \leq \text{heures disponibles en } t$$

où

$$\text{heures disponibles en } t = \text{heures}[t] \times \text{PERS}[t] + \text{SUPP}[t]$$

(où heures[t] = nombre d'heures prestées par chaque ouvrier à la période t)

$$\text{heures requises en } t = \sum_j \text{temps_opérateur}[j] \text{ PROD}[j,t]$$

(où temps_opérateur[j] = nombre d'heures nécessaires pour produire une unité de j).

Contraintes additionnelles (8):

Il faut remarquer ici que les contraintes (4), (5), (6) et (7) doivent nécessairement être présentes dans le modèle, en ce sens qu'elles expriment ce qu'est un plan admissible (4), ou qu'elles définissent certaines variables en fonction des autres (5)-(6), ou qu'elles expriment les besoins en capacité (7). Par ailleurs, on peut vouloir ajouter aux contraintes (4)-(7) toute une série de contraintes optionnelles limitant:

- l'importance des stocks, par la définition d'un stock maximal autorisé,

$$\text{STOCK}[j,t] \leq \text{stock_max}[j,t],$$

- le volume de personnel, ou d'embauche ou de licenciement à chaque période,
- l'utilisation des heures supplémentaires; celles-ci peuvent par exemple être limitées à un pourcentage du nombre d'heures normales prestées au cours de chaque période par une contrainte du type

$$\text{SUPP}[t] \leq \text{pourcent}[t] \times \text{heures}[t] \times \text{PERS}[t] ;$$

- la capacité utilisée dans certains centres de production; on impose couramment ce type de contraintes pour les « goulots » de production, c'est-à-dire les départements ou les postes de travail où les problèmes de capacité sont les plus aigus.

Par contre, les contraintes d'intégralité sur certaines variables (qui interdiraient, par exemple, aux variables PERS[j,t] de prendre des valeurs fractionnaires) sont généralement omises. En effet, de telles contraintes compliquent considérablement la résolution du modèle sans que leur présence soit indispensable à ce niveau de planification agrégée. On se contente donc, en pratique, d'arrondir la solution optimale calculée.

Notons enfin que le schéma de base présenté ci-dessus est facilement généralisé pour intégrer différentes catégories de personnel, pour autoriser les ruptures de stocks, etc.

Remarques sur l'agrégation des produits.

A quel niveau du modèle l'agrégation des produits peut-elle engendrer des problèmes (en ce sens qu'elle rendrait difficile l'estimation de certains paramètres) ?

- dans la fonction de coût : estimation du paramètre 'coûtstock[j,t]' (3)
- dans les contraintes : d'admissibilité (estimation de 'dem[j,t]') (4)
de capacité (estimation de 'temps_opérateur[j]') (7)
sur les stocks min/max désirés, etc.

Dans les contraintes (4) et (5), l'agrégation de plusieurs produits se traduit simplement par l'addition des paramètres ('dem[j,t]' ou min/max) correspondant aux produits constitutifs de la famille. Par contre, pour (3) et (7), il peut se révéler plus difficile d'évaluer correctement les paramètres 'coûtstock[j,t]' et 'temps_opérateur[j]' pour chaque famille j.

En pratique, on regroupe donc de préférence au sein d'une même famille des produits similaires du point de vue des coûts (de possession) et des temps opératoires. Pour faciliter la désagrégation lors de l'élaboration du plan directeur de production (voir ci-après), il est également recommandé de limiter la composition de chaque famille à des produits présentant des similitudes dans leurs cycles de demande.

Bibliographie.

- K. Boskma, *Productie en Logistiek*, 2ème édition, Wolters-Noordhoff, Groningen, 1987.
- M.L. Fisher, What is the right supply chain for your product?, *Harvard Business Review*, March-April 1997, pp. 105-116.
- V. Giard, *Gestion de la Production*, 2ème édition, Economica, Paris, 1988.
- S.C. Graves, A.H.G. Rinnooy Kan et P.H. Zipkin (eds.), *Logistics of Production and Inventory*, North-Holland, Amsterdam, 1993.
- A. Hax et D. Candea, *Production and Inventory Management*, Prentice-Hall, Englewood Cliffs, N.J., 1984.
- C.C. Holt, F. Modigliani, J.M. Muth et H.A. Simon, *Planification de la Production, des Stocks, de l'Emploi*, Dunod, Paris, 1964.
- L.A. Johnson et D.C. Montgomery, *Operations Research in Production Planning, Scheduling and Inventory Control*, Wiley, New York, 1974.
- J.O. McClain, L.J. Thomas et J.B. Mazzola, *Operations Management*, 3ème édition, Prentice-Hall, Englewood Cliffs, 1992.
- T.E. Vollmann, W.L. Berry et D.C. Whybark, *Manufacturing Planning and Control Systems*, Business One Irwin, Homewood, Illinois, 1992.

CHAPITRE 6

PLANIFICATION A COURT TERME: LE PLAN DIRECTEUR DE PRODUCTION

6.1. Généralités.

Dans ce chapitre, nous présentons brièvement le schéma d'élaboration du *plan directeur de production* (PDP ou Master Production Schedule). Le PDP désagrège le plan à moyen terme afin d'établir

- la planification sur un horizon plus court (souvent 3 à 6 mois),
- une découpe plus fine en sous-périodes (semaines),
- une désagrégation des familles en *références finales*.

Il peut donc être vu comme une version détaillée du plan agrégé de production dans laquelle est intégrée toute l'information actualisée relative au cours terme (prévisions de demande, état des stocks, disponibilité des machines, etc).

Le rôle essentiel du PDP est de planifier et de déclencher la production des références finales. Il fournit également l'information de base pour la gestion des matières et composants nécessaires à la production de ces références (voir chapitres suivants sur la gestion des stocks). Avant de procéder plus avant, précisons quelque peu le sens de ces définitions.

Qu'est-ce qu'une référence ?

Dans le cadre de la gestion des stocks, on appelle *référence* (ou *article*) tout produit fini, ou sous-ensemble, ou composant, ou matière première, dont les unités sont interchangeable et dont on répertorie toutes les unités sous un unique numéro de code – souvent appelé lui-même « référence ».

Qu'est-ce qu'une référence finale ?

Le concept de *référence finale* est souvent synonyme de « produit fini ». Dans un tel cas, c'est donc en termes de produits finis qu'est établi le plan directeur de production. Cette approche (qui peut sembler la plus naturelle) présente certains avantages, mais également des inconvénients potentiels. Nous reviendrons sur ce point ultérieurement: pour l'instant, le lecteur pourra se contenter de lire « produit fini » chaque fois que nous écrirons « référence finale ».

6.2. Schéma d'élaboration du PDP.

La plupart des auteurs présentent le processus d'élaboration du PDP de façon extrêmement informelle. Il semble qu'il existe pour le PDP peu de modèles d'optimisation similaires à ceux utilisés pour le plan agrégé. De nombreuses questions annexes, telles que celles de la périodicité de révision du PDP ou de sa coordination avec le plan agrégé ne sont généralement pas traitées dans la littérature (voir par exemple Vonderembse et White 1988).

Dans le cadre de cet exposé introductif, nous esquisserons simplement le schéma d'une approche simplifiée dans laquelle apparaissent clairement les caractéristiques fondamentales du PDP mentionnées en début de chapitre: désagrégation du plan agrégé et prise en compte de l'information disponible à court terme (voir Berry et al. 1979 ou Hax et Candea 1984 pour une discussion plus approfondie).

La désagrégation du plan agrégé comporte deux aspects: d'une part, le passage des familles de produits aux références finales individuelles, et d'autre part la division des périodes de planification du plan agrégé en sous-périodes plus fines. Par ailleurs, l'information disponible à court terme porte sur la demande (mieux connue que dans le moyen terme, par exemple parce que certaines commandes fermes ont déjà été enregistrées) et sur l'état réel des stocks au début de l'horizon de planification.

Pour chaque famille et chaque période, le plan agrégé a fixé les quantités à produire et à stocker. Dans la logique de court terme qui nous guide actuellement, il est utile de concentrer l'attention sur les quantités à stocker plutôt que sur celles à produire au cours de chaque période. En effet, dans le court terme, les quantités à produire seront fortement influencées par les commandes enregistrées. Par contre, c'est l'état des stocks qui déterminera la situation de l'entreprise à plus long terme et qui assurera donc le lien avec le plan agrégé de production.

Exemple: La firme BCBG est spécialisée dans la confection de prêt-à-porter. Nous sommes à la fin du mois de mars et la firme est occupée à développer son plan directeur de production pour avril et mai, sur base d'un plan agrégé adopté en décembre. La famille de produits considérée dans le plan agrégé comporte 3 références: blue-jeans, pantalons en flanelle, pantalons en lin. La période de planification du plan agrégé est le mois et son horizon est d'un an. Le PDP, quant à lui, est établi pour deux mois sur base de sous-périodes hebdomadaires.

La firme procède d'abord à la *désagrégation de la demande par référence et par semaine*. Le but est ici d'établir des chiffres de *demande prévisionnelle* pour l'horizon du plan, ainsi que les *niveaux de stock à atteindre* à la fin de chaque mois. La demande prévisionnelle est dérivée du carnet de commandes et de toute autre information disponible. Typiquement, le carnet de commandes reflète parfaitement la demande des quelques semaines à venir, alors que les chiffres avancés pour la fin du plan sont de nature plus incertaine (ils peuvent être basés, par exemple, sur une désagrégation des prévisions de demande par famille – $dem[j,t]$ – utilisées pour développer le plan agrégé). La firme obtient ainsi un tableau du type suivant.

Semaines	Avril				Mai			
	1	2	3	4	5	6	7	8
Blue-jeans	210	220	235	250	300	300	300	300
Flanelle	155	135	120	125	150	150	150	150
Lin	105	140	115	135	150	150	150	150
Total	1945				2400			

Table 1: Demande prévisionnelle sur l'horizon du PDP.

Idéalement, la somme des demandes prévisionnelles pour chaque mois – ici, 1945 pour avril et 2400 pour mai – sera proche de la demande agrégée estimée pour la famille de produits au cours de ce même mois – c’est-à-dire $dem[pantalon, 'avril']$ ou $dem[pantalon, 'mai']$ respectivement – utilisée lors de l’élaboration du plan agrégé. En fait, ces quantités seront probablement égales si le premier mois du plan directeur est également le premier mois du plan agrégé de production. Pour les mois ultérieurs, cependant, des écarts dus aux erreurs d’estimation vont généralement se manifester.

Bien entendu, établir le PDP uniquement en fonction de la demande prévisionnelle de l’horizon à court terme révélerait une dangereuse myopie dans le chef des planificateurs. En effet, l’entreprise se doit également de respecter son plan de production à moyen terme (le seul qui permette, par exemple, de tenir compte des variations cycliques, saisonnières ou conjoncturelles de la demande). Par conséquent, la demande à moyen terme sera intégrée dans le PDP par l’intermédiaire des niveaux de stocks à atteindre en fin de mois. Ces niveaux de stocks sont dérivés du plan agrégé: ils sont obtenus par désagrégation des niveaux de stocks requis pour chaque famille de produits en termes des références finales individuelles. La désagrégation peut être basée sur des historiques de ventes ou des prévisions marketing.

Pour l’exemple de la firme BCBG, supposons que le plan agrégé prévoyait 800 pantalons en stock à la fin du mois d’avril et 1200 à la fin mai et que les volumes de ventes pour les trois références considérées soient dans un rapport 2:1:1. Par une simple règle de trois, on obtient alors les quantités suivantes à stocker pour chaque type de pantalon:

fin avril : 400 blue-jeans, 200 flanelle, 200 lin,
fin mai : 600 jeans, 300 flanelle, 300 lin.

Supposons enfin que les stocks disponibles à la fin du mois de mars se répartissent de la façon suivante:

fin mars : 200 jeans, 140 flanelle, 180 lin.

La tâche des planificateurs est à présent de déterminer le nombre d’unités à produire de chaque type de pantalon, pour chacune des semaines de l’horizon du PDP, sous contrainte de

1. pouvoir satisfaire la demande prévisionnelle de chaque semaine (cf. Table 1),
2. assurer la réalisabilité du plan par le département production,
3. se rapprocher autant que possible des niveaux de stocks en fin de mois requis par le plan agrégé.

Ces trois contraintes reflètent respectivement les préoccupations à court terme des départements Ventes (point 1) et Production (point 2) ainsi que les engagements à moyen terme traduits dans le plan agrégé (point 3). Concernant ce dernier point, remarquons qu’il ne sera pas toujours possible d’atteindre *exactement* les niveaux de stocks requis: la réalisabilité de cet objectif sera par exemple fortement influencée par des différences éventuelles entre les prévisions de demande utilisées pour l’élaboration du plan agrégé et les commandes effectivement enregistrées.

Revenons à l'exemple de BCBG et considérons la référence 'pantalons en flanelle'. Au cours du mois d'avril, la firme doit satisfaire à une demande de 535 pantalons flanelle. Elle entame le mois avec 140 pantalons en stock et doit le terminer avec (environ) 200 pantalons en stock. Un décompte facile indique qu'elle doit donc produire: $535 + 200 - 140 = 595$ pantalons flanelle en avril. (Comparez ce calcul avec l'équation (12) du chapitre précédent.) De même, BCBG doit produire: $600 + 300 - 200 = 700$ pantalons flanelle en mai.

Si la firme décide d'étaler sa production uniformément au cours de chaque mois, elle peut par exemple adopter le PDP décrit dans la Table 2 ci-dessous (mais d'autres plans sont également envisageables; voir plus loin).

Semaines	0	1	2	3	4	5	6	7	8
Demande prévisionnelle	–	155	135	120	125	150	150	150	150
PDP (en cours de semaine)	–	150	150	150	145	175	175	175	175
Stock (en fin de semaine)	140	135	150	180	200	225	250	275	300

Table 2. PDP pour les pantalons en flanelle.

6.3. Commentaires sur l'élaboration du PDP.

Réalisabilité du PDP.

Dans la section précédente, nous avons formulé la méthode d'élaboration du PDP de façon très peu précise. Dans la pratique, ce plan est souvent établi par des méthodes heuristiques et par tâtonnements, jusqu'à obtention d'un plan acceptable par les départements production et marketing.

En particulier, l'existence de capacité de production suffisante pour assurer la réalisabilité du PDP est généralement vérifiée a posteriori. Remarquons que la capacité disponible est déterminée par les décisions prises à long terme et à moyen terme (équipements, layout, volume d'emploi, ...). Un certain ajustement charge-capacité a pu être assuré de façon globale lors de l'élaboration du plan agrégé (voir contraintes (7) du Chapitre 5). Cependant, les besoins en capacité de production calculés au niveau agrégé sont nécessairement approximatifs par suite

- de l'identification de plusieurs produits au sein d'une même famille,
- de l'identification de plusieurs ressources,
- des déviations entre demande estimée à moyen terme et à court terme,
- de l'omission de certains « détails », comme par exemple les temps de lancement requis par la production de certains articles.

Ceci explique que l'ajustement charge-capacité réalisé dans le cadre du plan agrégé ne peut garantir à lui seul la réalisabilité du PDP. Il est donc nécessaire de vérifier la réalisabilité de chaque version proposée du PDP. Cette vérification est généralement effectuée par un *calcul de charges global* ou *rough-cut capacity planning* concentré sur les centres de production goulots. D'essais en erreurs, l'élaboration d'un PDP réalisable peut ainsi procéder par ajustements successifs.

Regroupement/étalement de la production dans le PDP.

Entre autres facteurs, la capacité de production requise par le PDP est fortement influencée par le nombre de lancements de production impliqués par le plan, ainsi que par la synchronisation entre les lancements de production relatifs aux diverses références. Par exemple, le PDP de la Table 2 (Section 6.2) nécessite 8 lancements de production, alors que le PDP décrit dans la Table 3 ci-dessous ne requiert que 4 lancements:

Semaines	0	1	2	3	4	5	6	7	8
Demande prévisionnelle	–	155	135	120	125	150	150	150	150
PDP (en cours de semaine)	–	300	–	295	–	350	–	350	–
Stock (en fin de semaine)	140	285	150	325	200	400	250	450	300

Table 3. PDP alternatif pour les pantalons en flanelle.

Remarquons que la Table 3 a été obtenue à partir de la Table 2 en regroupant systématiquement la production de deux semaines successives en une seule semaine: cette approche réduit évidemment le nombre de lancements de production, mais tend à augmenter le niveau des stocks (comme on le constate en comparant la dernière ligne des Tables 2 et 3). Plus généralement, des quantités de production réalisant un compromis « optimal » entre coût de lancement et coût de stockage pourraient être déterminées dans le cadre du PDP par des techniques de regroupement des lots similaires à celles utilisées par la méthode MRP (voir Chapitre 8).

Utilisation du plan directeur de production

Pour être utile, le PDP doit recevoir l'aval des départements Production et Ventes (il doit être satisfaisant pour les Ventes et réalisable pour la Production). Quand le plan a été accepté par les responsables Ventes et Production, il est *gelé* sur un horizon de quelques semaines afin de permettre à la Production de travailler sans encombre. Les quantités entrées dans le PDP permettront ainsi de lancer la production des références finales ainsi que de nombreuses autres références (sous-assemblages, composants, ...; voir Chapitre 8).

Par ailleurs, en plus de son rôle premier de plan de production, le PDP peut être utilisé par le département Ventes pour fixer les dates de livraison de nouvelles commandes. Le lecteur peut aisément se convaincre du bien-fondé de cette affirmation en remarquant simplement que le PDP détermine quelles quantités de références finales seront produites au cours de chaque période. Les vendeurs peuvent donc se servir de ce plan afin de calculer le nombre d'unités disponibles à chaque instant pour satisfaire une commande particulière. Par exemple, la Table 4 indique le calcul des *quantités disponibles pour livraison* (« available to promise », ou ATP) sur base du PDP de la Table 2 et en tenant compte des commandes fermes déjà enregistrées (ce calcul est parfaitement similaire au calcul des stocks dans la Table 2).

Semaines	0	1	2	3	4	5	6	7	8
PDP	–	150	150	150	145	175	175	175	175
Commandes enregistrées		130	105	80	20	75	–	–	–
Disponible pour livraison	140	160	205	275	400	500	675	850	1025

Table 4. Disponible pour livraison sur base du PDP de la Table 2.

On constate ainsi que les vendeurs peuvent encore accepter une commande de 205 unités supplémentaires en semaine 2. Suite à une telle commande, il faudrait cependant modifier la Table 4 de façon appropriée, comme indiqué dans la Table 5. On dit que les vendeurs *consomment le PDP*.

Semaines	0	1	2	3	4	5	6	7	8
PDP	–	150	150	150	145	175	175	175	175
Commandes enregistrées		130	310	80	20	75	–	–	–
Disponible pour livraison	–	–	–	70	195	295	470	645	820

Table 5. Modification de la Table 4 suite à une nouvelle commande de 205 unités en semaine 2.

Nous n'entrerons pas ici dans le détail des calculs liés à la consommation du PDP. Sur le plan des principes de gestion cependant, il est essentiel de comprendre que, dans cette approche de la coordination entre production et ventes, c'est le plan de production qui sert de base fixe (du moins pour l'horizon sur lequel le plan est gelé). Les ventes, quant à elles, doivent suivre le

rythme défini par ce plan. Ceci contraste fortement avec une approche moins rigoureuse (mais malheureusement très répandue) où la Production doit en permanence s'attacher à satisfaire les commandes acceptées par le service Ventes, avec toutes les conséquences néfastes qu'une telle gestion peut avoir pour la planification des activités de production (rappelez-vous la discussion du Chapitre 2).

6.4. Le contenu du plan directeur de production.

Comme promis dans la première section de ce chapitre, revenons à présent brièvement sur la notion de référence finale. Voyons d'abord quels sont les avantages et inconvénients liés à l'élaboration d'un plan directeur de production en termes de produits finis.

Avantages:

- c'est en termes de produits finis que la demande est formulée;
- si l'entreprise dispose d'un plan de production pour les produits finis, elle peut (théoriquement) en déduire un plan de production pour tous les composants et les matières premières.

Inconvénients:

- le nombre de produits finis peut être énorme en cas d'options nombreuses!

Pensons par exemple à la situation rencontrée lors de l'élaboration d'un PDP chez Volkswagen. Quelques-unes des références utilisées pour produire une automobile sont:

- peinture blanche 001 (matière première)
- huile XY40W (matière première)
- boulon n° 8 (composant)
- bloc moteur Golf Diesel 1600 (sous-assemblage)
- Golf D 1600 blanche sans options (sous-assemblage)
- Golf Diesel 1600 blanche, siège chauffant, horloge quartz , ... (produit fini).

Dans une telle situation, où le nombre d'options peut être gigantesque, il est plus aisé de raisonner en termes de sous-assemblages qu'en termes de produits finis. Le PDP sera donc élaboré sur base du sous-assemblage correspondant à un modèle « nu » (Golf sans options) et l'assemblage des produits finis sera planifié le plus tard possible en tenant compte des commandes effectivement enregistrées.

Remarque: même dans le mode de production Juste-A-Temps adopté par de nombreux constructeurs automobiles, les firmes élaborent un PDP afin de faciliter la gestion des matières et la tâche des fournisseurs.

De façon plus générale, les références finales utilisées dans le PDP seront choisies en tenant compte des critères suivants:

- leur nombre ne doit pas être trop élevé, afin de faciliter la planification;
- leur demande (ou utilisation) doit être prévisible avec une bonne précision.

Typiquement, les références finales qui apparaissent dans le PDP sont choisies en fonction du mode de production, puisque celui-ci détermine largement les critères précédents. On obtient ainsi la correspondance suivante entre le mode d'organisation de la production et la nature des références finales:

- production pour stock (chaussures, jouets) → produits finis;
- assemblage sur commande (automobiles, PCs, circuits imprimés) → sous-assemblages;
- production sur commande (type job-shop: imprimerie, sous-traitance mécanique) → composants, matières premières.

Les entreprises produisant des articles uniques sur base de commandes-clients spécifiques (« engineering sur commande ») élaborent souvent leur PDP en tenant compte uniquement des commandes enregistrées (dans ce cas, le PDP peut donc être identifié avec le carnet de commandes).

Bibliographie.

W.L. Berry, T.E. Vollmann et D.C. Whybark, *Master Production Scheduling: Principles and Practice*, Falls Church, Va., American Prod. and Invent. Control Soc., 1979.

V. Giard, *Gestion de la Production*, 2ème édition, Economica, Paris, 1988.

A. Hax et D. Candea, *Production and Inventory Management*, Prentice-Hall, Englewood Cliffs, N.J., 1984.

T.E. Vollmann, W.L. Berry et D.C. Whybark, *Manufacturing Planning and Control Systems*, Business One Irwin, Homewood, Illinois, 1992.

M.A. Vonderembse et G.P. White, *Operations Management: Concepts, Methods and Strategies*, West Publishing Company, St. Paul, Minnesota, 1988.

CHAPITRE 7

PLANIFICATION A COURT TERME: LA PROBLEMATIQUE DES STOCKS

7.1. Généralités.

Les quantités entrées dans le Plan Directeur de Production déterminent le nombre d'unités à produire de chacune des références finales. La production de ces unités aura, à son tour, des conséquences pour la production et/ou l'achat des matières, composants et sous-assemblages entrant dans la fabrication des références finales (dans la suite, le terme «composant» sera utilisé indifféremment pour désigner des matières ou des sous-assemblages).

En principe, lorsque la quantité à produire de chaque référence finale est connue, il est possible de déduire exactement le nombre d'unités requises de chacun de ses composants. Mais en pratique, pour des raisons de facilité, on préfère parfois gérer certains composants de façon statistique, sans tenir compte explicitement de la chaîne d'événements qui génèrent la demande. Les méthodes appliquées dans de tels cas sont alors similaires à celles d'application pour gérer les stocks de produits finis, pour lesquels la demande est générée par le marché.

En conséquence, on distinguera par la suite deux grandes classes de modèles de gestion des stocks: les modèles à *demande dépendante* d'une part et ceux à *demande indépendante* d'autre part. Ces deux classes de modèles seront successivement considérées dans les Chapitres 8 et 9 ci-dessous.

7.2. Objectif de la gestion des stocks.

L'objectif de la fonction «gestion des stocks» est d'assurer la disponibilité des références considérées sur un horizon donné tout en minimisant les coûts encourus. Trois types de coûts sont généralement pris (plus ou moins) explicitement en compte pour évaluer la qualité d'une politique de gestion des stocks:

1. *coûts de possession*: coûts d'opportunité, liés à l'immobilisation du capital; dépréciation; coûts de stockage sous forme de loyer, énergie, assurances, ...;
2. *coûts de rupture*, associés à la non satisfaction de la demande à l'instant où elle s'exprime: ventes manquées, ristournes accordées pour livraisons tardives, ...;
3. *coûts de commande*: frais d'administration, de transport, de réception, ... dans le cas d'un approvisionnement externe; coûts de lancement de la production dans le cas d'un approvisionnement interne: réglages, rebuts, gestion de l'ordre de fabrication, ...

C'est en principe le rôle du *contrôle de gestion* de fournir une évaluation de ces différentes catégories de coûts. En pratique, il faut bien constater qu'il est souvent très difficile de les estimer de manière fiable et appropriée à la gestion des stocks. Faute de mieux, on sera donc souvent obligé de se contenter d'une valorisation approximative des coûts de stockage.

Remarquons également que, dans le cadre traditionnel de la gestion des stocks, ces coûts sont considérés comme des « données », non contrôlables par l'entreprise. Par contre, les approches plus récentes, comme la gestion Juste-A-Temps ou la gestion de la Qualité Totale, regardent les coûts de stockage comme des variables dont la valeur peut (doit) être influencée par la politique de gestion adoptée. Nous reviendrons sur ce point par la suite.

7.3. Mesures de l'état des stocks.

Quelle mesure de l'état des stocks faut-il utiliser dans le cadre de leur gestion? La réponse à cette question varie en fonction du cadre spécifique dans lequel elle se pose. Nous serons ainsi amenés à considérer les concepts suivants (rappelons que, par convention, toutes les mesures de stock sont entendues *en fin de période*):

stock disponible à la période t = nombre d'unités physiquement présentes en t ;

stock net à la période t = stock disponible en t – arriérés-clients en t .

Dans cette définition, les *arriérés-clients* à la période t sont les unités que l'entreprise aurait dû livrer chez le client avant la fin de la période t , mais n'a pas encore livrées à cet instant. Le stock net représente donc le stock utilisable pour satisfaire la demande ultérieure à la période t .

Les *approvisionnements attendus* pour la période t sont ceux correspondant aux commandes-fournisseurs ou aux ordres de fabrication dont la livraison est attendue *au début* de la période t . Les *engagements pris* à la période t sont les commandes passées par les clients et dont la livraison doit avoir lieu après la période t .

Le stock net à la période t ne tient pas compte des approvisionnements attendus après t ni des engagements pris en t . Dans certaines situations, on s'intéressera donc plus naturellement à la mesure suivante de l'état des stocks:

position de stock à la période t =

stock disponible en t + approvisionnements attendus après t – arriérés en t – engagements en t .

7.4. Le rôle des stocks, ou ...

... pourquoi une entreprise constitue-t-elle des stocks, plutôt que d'essayer de disposer simplement de la quantité nécessaire de chaque composant à chaque instant (approche *zéro-stocks*)? Pour différentes raisons. Silver et Peterson (1985, pp. 59-60) mentionnent par exemple cinq fonctions principales des stocks :

- *stocks de sécurité*: contrecarrent les incertitudes de la demande (cas des produits finis) ou des approvisionnements (matières, composants, livraisons à un poste de travail en provenance d'un poste amont);

- *stocks d'anticipation*, ou *saisonniers*: anticipent sur une demande future trop élevée pour pouvoir y répondre par la production de la période où elle s'exprime; le but est de niveler le taux de production au cours du temps, plutôt que de suivre les variations de la demande (c'est un rôle essentiel dans la planification à moyen terme, comme on l'a vu au Chapitre 5);
- *stocks de cycle*: résultent de la mise en production ou de l'achat de certaines références par lots plutôt qu'à l'unité (économies sur les coûts de lancement de la fabrication, rabais sur la commande de grandes quantités, contraintes technologiques nécessitant la production en quantités fixes, ...);
- *encours et stocks de transit*: cette catégorie de stocks est constituée par les produits finis ou semi-finis circulant dans l'atelier (encours) ou entre les différents échelons d'un réseau de distribution (quantités transportées);
- *stocks de découplage*: autorisent le découplage des décisions prises aux différents échelons d'un système de production ou de distribution; par exemple, grâce à de tels stocks, un distributeur peut décider de livrer de façon anticipée une grande quantité à un client sans répercuter immédiatement la décision vers les échelons en amont; ou bien, la maintenance d'une machine peut être effectuée sans interrompre la production sur une machine aval.

On pourrait encore distinguer les stocks de spéculation, ou les stocks liés aux cycles saisonniers de production (agriculture), etc (voir De Bruyn 1991 ou Giard 1988).

7.5. La gestion Juste-A-Temps, ou ...

... pourquoi vouloir éliminer les stocks ? Il y a plusieurs raisons à cela.

- Les stocks coûtent cher: ils immobilisent du capital, se démodent, se détériorent, requièrent de l'espace, etc..
- Les stocks dissimulent certains problèmes: par exemple, des temps de lancement anormalement élevés, des pannes fréquentes, des articles de qualité irrégulière, un mauvais contrôle de la production, des prévisions de demande imprécises, des fournisseurs peu fiables, etc. Dans cette optique, les stocks représentent donc un obstacle à l'obtention de la Qualité Totale. (Ce n'est certainement pas un hasard si les principes de gestion Juste-A-Temps et de gestion de la Qualité Totale ont émergé simultanément, dans les mêmes entreprises).
- Lorsqu'ils sont surévalués (ce qui est fréquemment le cas en pratique), les stocks de sécurité contribuent à augmenter artificiellement la demande, donc à augmenter la charge du système et à créer des difficultés de planification évitables.

Au cours des 30 dernières années, certaines entreprises japonaises (en particulier Toyota) ont développé une philosophie *Juste-A-Temps* (JAT, JIT, Zéro-stocks) de la production, selon laquelle *les produits, composants et matières doivent parvenir à l'endroit où l'on en a besoin au moment où l'on en a besoin et dans les quantités où l'on en a besoin.*

Cette philosophie implique en particulier une réduction (idéalement: la disparition) des stocks. Elle est implantée en pratique par un système de gestion de la production à *flux tirés* (*flux tendus*, pull system). Dans une vision extrême de la gestion à flux tirés, une unité de produit fini ne peut être mise en production que pour satisfaire une commande-client déjà enregistrée (l'ordre de fabrication pour votre voiture ne sera lancé que lorsque votre bon de commande aura été reçu par l'usine). De même, un sous-assemblage ne peut être fabriqué que pour être immédiatement intégré dans une unité du produit fini et un article ne peut être réapprovisionné en stock que si il est immédiatement nécessaire. Un tel mode de gestion résulte évidemment en une diminution drastique de tous les stocks autres que les stocks d'encours indispensables (dont la rotation est elle-même accélérée au maximum).

Il est utile de remarquer que, en réalité, l'approche à flux tirés prend souvent la forme d'une gestion « Juste-Un-Peu-Avant » plutôt que « Juste-A-Temps ». En effet, dans la pratique, une référence n'est pas réapprovisionnée à l'instant précis où elle est nécessaire, car cela créerait inévitablement des délais inacceptables pour le client (pensez à la commande de votre voiture). La solution adoptée consiste donc plutôt à réapprovisionner *en très petite quantité* chaque référence qui vient d'être épuisée ou qui va être épuisée, *en anticipation d'un besoin potentiel futur* (et non pas en réponse à un besoin explicitement formulé, comme le voudrait la théorie pure des flux tirés). Plus précisément, une référence n'est réapprovisionnée que lorsque son stock disponible tombe sous un niveau prédéterminé très faible, par exemple lorsque le dernier container vient d'être entamé. Dans cette optique, la gestion à flux tirés présente certaines similitudes avec les systèmes traditionnels de réapprovisionnement par point de commande, tels ceux que nous étudierons au Chapitre 9.

Au vu de la longue liste de fonctions des stocks énumérée dans la section précédente, une question vient cependant immédiatement à l'esprit: *comment les entreprises qui pratiquent une philosophie Juste-A-Temps peuvent-elles se passer de stocks ?* Plusieurs réponses peuvent être fournies à cette question.

- Les entreprises qui adoptent le JAT opèrent en général sur un marché où la demande est prévisible (produit dans la phase adulte de son cycle de vie, production de masse permettant l'utilisation de méthodes statistiques de prévision) et s'assurent des liens étroits avec leurs fournisseurs pour diminuer l'incertitude du côté des livraisons (en exploitant parfois leur position de force par rapport à ces fournisseurs; voir l'article en annexe à ce Chapitre).
Conséquence: peu ou pas de stocks de sécurité nécessaires.
- La demande est relativement stable, sans grandes fluctuations saisonnières (fluctuations mensuelles de moins de 10% chez Toyota): ces caractéristiques sont liées au type de marché, mais résultent parfois aussi d'une stratégie marketing adaptée (par exemple, peu de modèles et peu d'options); les entreprises JAT maintiennent généralement de la capacité excédentaire, parfois (principalement au Japon) sous forme d'horaires de travail souples (heures supplémentaires disponibles).
Conséquence: peu de stocks d'anticipation.
- Les entreprises JAT n'acceptent pas les temps de lancement de production comme un donné mais essayent au contraire de les réduire à leur strict minimum (implication des ouvriers pour

simplifier les procédures, polyvalence évitant l'intervention de régleurs, réglage des machines en cours d'exécution, conception des pièces pour la manufacturabilité, automatisation, etc).

Exemples:

Canon (fabrication des accessoires): réduction des temps de lancement de 4 min à 15 sec.

Hitachi (centre d'usinage): réduction des temps de lancement de 4 heures à 9 min.

Les coûts liés aux passations de commande sont également diminués par l'utilisation d'ordres d'achat globaux couvrant une quantité donnée sur une période étendue (contrats-cadre, « blanket orders »). Le fournisseur supporte la responsabilité de livrer cette quantité en petits lots, au fur et à mesure des besoins de l'entreprise.

Conséquence: peu de stocks de cycle.

- Les entreprises pratiquant le JAT renoncent de façon explicite aux possibilités de découplage; ceci les amène par exemple à éliminer les goulots dans leur chaînes de production afin d'assurer un flux de production aussi régulier que possible et donc de permettre un « couplage » parfait entre postes de travail.

Conséquence: pas de stocks de découplage.

Conclusion: Malgré les nombreux succès rencontrés et l'engouement qu'elle suscite, l'approche JAT intégrale n'est probablement pas adaptée à toutes les organisations. Sa philosophie de réduction des stocks, des temps de lancement et des défauts de fabrication est très certainement valide universellement. Mais certaines conditions d'implantation ne sont par exemple (typiquement) pas rencontrées dans des entreprises fonctionnant en job-shops et fabriquant une grande variété de produits. Pour nombre d'entreprises, les stocks restent un mal (ou un bien ?) nécessaire, qu'elles doivent gérer au mieux. Remarquons d'ailleurs que même les entreprises pratiquant un Juste-A-Temps orthodoxe continuent à gérer certains de leurs stocks par des méthodes traditionnelles !!

Bibliographie.

C. De Bruyn, *Méthodes Quantitatives de Gestion 1*, FEGSS, 1991.

V. Giard, *Gestion de la Production*, 2ème édition, Economica, Paris, 1988.

S.C. Graves, A.H.G. Rinnooy Kan et P.H. Zipkin (eds.), *Logistics of Production and Inventory*, North-Holland, Amsterdam, 1993.

J.O. McClain, L.J. Thomas et J.B. Mazzola, *Operations Management*, 3ème édition, Prentice-Hall, Englewood Cliffs, 1992.

E. Silver et R. Peterson, *Decision Systems for Inventory Management and Production Planning*, J. Wiley & Sons, New York, 1985.

S.H. Zanakis, L.M. Austin, D.C. Nowading et E.A. Silver, From teaching to implementing inventory management: Problems of translation, *Interfaces* 10 (1980) pp. 103-110.

P.H. Zipkin, Does manufacturing need a JIT revolution? *Harvard Business Review* Jan-Feb. 1991, pp. 40-50.

CHAPITRE 8

PLANIFICATION A COURT TERME: LA PLANIFICATION DES BESOINS EN COMPOSANTS

8.1. Exemple introductif.

Ce chapitre est consacré à la problématique de la gestion des stocks pour les références à *demande dépendante*, c'est-à-dire pour les références intermédiaires (matières premières, composants et sous-assemblages) dont la demande est *explicitement considérée comme étant fonction de la demande pour d'autres références* (typiquement, les références finales qui apparaissent dans le plan directeur de production). Dans de nombreuses entreprises, la gestion des références à demande dépendante est effectuée par des méthodes de type *MRP (Material Requirements Planning, Planification des Besoins en Composants)*. Dans la logique MRP, les calculs de besoins sont réalisés en cascade au départ des besoins exprimés dans le PDP pour les références finales.

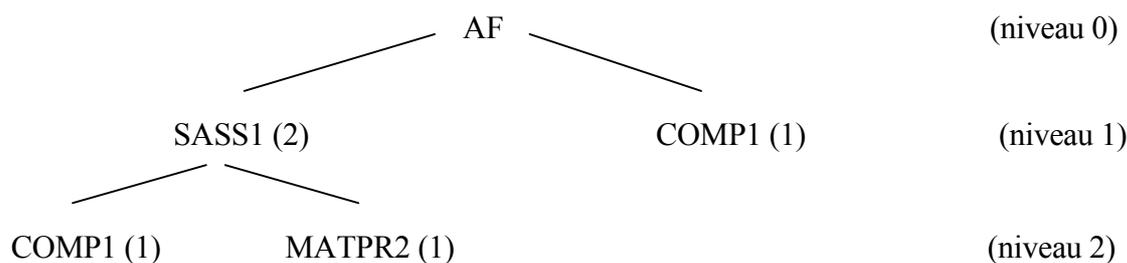
Dans cette section, nous allons développer un exemple simple qui nous permettra d'introduire la terminologie standard de la méthode MRP et d'illustrer les principales caractéristiques de cette méthode.

Un produit fini ou une référence finale est typiquement composé d'un certain nombre de sous-assemblages, eux-mêmes constitués de divers composants. La *nomenclature* (bill of materials, BOM) d'un produit est la liste de toutes les références intermédiaires intervenant dans sa composition, ordonnées de façon à indiquer le processus de production du produit.

Exemple:

AF	(assemblage final)
SASS1 (2)	(sous-assemblage 1)
COMP1 (1)	(composant 1)
MATPR2 (1)	(matière première 2)
COMP1 (1)	(composant 1)

Cette nomenclature nous apprend que la référence AF est obtenue par assemblage de deux unités du sous-assemblage SASS1 et d'une unité du composant COMP1. Chaque unité de SASS1 comporte à son tour une unité de COMP1 et une unité de MATPR2. Une représentation plus 'visuelle' de la nomenclature peut être donnée sous forme d'*arborescence* :



Comment l'entreprise va-t-elle satisfaire ces besoins bruts ? En produisant, bien sûr, mais également en utilisant ses stocks et des unités éventuelles en attente de réapprovisionnement.

Comme mesure de l'état des stocks, on utilise ici la notion de *stock net*: à la fin de la période t ,

$$\text{stock_net}[t] = \text{stock_disponible}[t] - \text{arriérés}[t]$$

où le stock disponible est le stock physiquement présent à la fin de la période t , et les arriérés sont les unités que l'entreprise aurait déjà dû livrer chez le client, mais n'a pas encore livrées à la fin de la période t . (Rappelons que, par convention, toutes les quantités invoquées dans cette définition sont entendues en fin de sous-période). L'intérêt de considérer le stock net, plutôt que le stock disponible, est surtout apparent au début de l'horizon de planification, puisque $\text{stock_net}[0]$ représente le stock effectivement utilisable pour satisfaire la demande exprimée (c'est-à-dire, les besoins nets).

Les *approvisionnements attendus* pour la période t sont ceux correspondant aux commandes déjà passées (ou aux ordres de fabrication déjà lancés) au début de l'horizon de planification et dont la livraison est attendue au début de la période t .

Exemple (suite): Supposons que, pour SASS1, le stock net initial soit de 44 unités et un approvisionnement de 30 unités soit attendu en période 2.

Semaine	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Besoins bruts		0	20	20	20	20	20	20	20	20	20	20	20	20	20	20
Stock net	44															
Appros attendus			30													

Sur base du stock initial et des approvisionnements attendus, il est à présent possible de projeter l'évolution du stock net au cours de l'horizon et d'en déduire les *besoins nets* de chaque période (c'est-à-dire, les quantités pour lesquelles il faudra planifier un approvisionnement). Pour ce faire, calculons successivement, pour $t = 1, 2, \dots, T$, la quantité $Q[t]$ définie comme suit:

$$Q[t] = \text{besoins_bruts}[t] - \text{stock_net}[t-1] - \text{appros_attendus}[t] . \quad (1)$$

Si $Q[t]$ est négatif pour la période t , alors le stock net en fin de période $t-1$, augmenté des approvisionnements de la période t , est suffisant pour couvrir les besoins bruts de la période. On en déduit donc la règle suivante:

$$\text{si } Q[t] \leq 0, \text{ alors } \text{stock_net}[t] = -Q[t] \text{ et } \text{besoins_nets}[t] = 0. \quad (2)$$

Par contre, si $Q[t]$ est positif, alors il sera nécessaire de produire $Q[t]$ unités pour couvrir les besoins de la période t . On conclut donc:

$$\text{si } Q[t] \geq 0, \text{ alors } \text{stock_net}[t] = 0 \text{ et } \text{besoins_nets}[t] = Q[t]. \quad (3)$$

Exemple (suite): Pour la référence SASS1, l'application de ces règles de calcul mène aux conclusions suivantes.

$$\begin{aligned}
 Q[1] &= 0-44-0 = 0, \text{ donc stock_net}[1] = 44 \text{ et besoins_nets}[1] = 0, \\
 Q[2] &= 20-44-30 = -54, \text{ donc stock_net}[2] = 54 \text{ et besoins_nets}[2] = 0, \\
 Q[3] &= 20-54-0 = -34, \text{ donc stock_net}[3] = 34 \text{ et besoins_nets}[3] = 0, \\
 Q[4] &= 20-34-0 = -14, \text{ donc stock_net}[4] = 14 \text{ et besoins_nets}[4] = 0, \\
 Q[5] &= 20-14-0 = 6, \text{ donc stock_net}[5] = 0 \text{ et besoins_nets}[5] = 6, \\
 Q[6] &= 20-0-0 = 20, \text{ donc stock_net}[6] = 0 \text{ et besoins_nets}[6] = 20, \text{ etc.}
 \end{aligned}$$

Semaine	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Besoins bruts			20	20	20	20	20	20	20	20	20	20	20	20	20	
Stock net	44	44	54	34	14											
Appros attendus			30													
Besoins nets						6	20	20	20	20	20	20	20	20	20	

La connaissance des besoins nets conduit à son tour au calcul des *approvisionnements (commandes ou ordres de fabrication)* à planifier pour satisfaire ces besoins. Ce calcul sera guidé par la politique d'approvisionnement que l'entreprise a choisi d'adopter pour la référence considérée: approvisionnements à périodicité fixe, ou par lots de taille prédéterminée, etc. Remarquons qu'il s'agit là de la première et de l'unique *décision* à prendre dans le cadre de la planification MRP telle qu'exposée ici: les calculs précédents (absorption des délais, détermination des besoins bruts, du stock net projeté, des besoins nets) se contentaient en effet de dérouler les conséquences logiques des plans établis pour les références de niveau supérieur et n'impliquaient aucune prise de décision de la part de l'utilisateur. Nous reviendrons plus loin sur les différentes façons d'aborder ce problème important. Pour l'instant, nous supposons (par facilité) que les approvisionnements sont planifiés sur base d'une politique de *périodicité fixe*, ce qui signifie que les besoins nets sont satisfaits par des approvisionnements planifiés à intervalle constant.

Exemple (suite): En supposant que les approvisionnements de SASS1 sont planifiés à 3 semaines d'intervalle, on obtient les regroupements suivants:

Semaine	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Besoins bruts			20	20	20	20	20	20	20	20	20	20	20	20	20	
Stock net	44	44	54	34	14											
Appros attendus			30													
Besoins nets						6	20	20	20	20	20	20	20	20	20	
Appros planifiés						46			60			60			20	

Nous sommes à présent revenus à une situation similaire à celle rencontrée en début de section: afin de recevoir à temps les approvisionnements planifiés pour SASS1, il est nécessaire de tenir compte du délai d'obtention de cette référence et de fixer en conséquence lesancements de production.

Exemple (suite): En supposant que le délai d'assemblage de SASS1 est d'une semaine, on obtient:

Semaine	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Besoins bruts			20	20	20	20	20	20	20	20	20	20	20	20	20	
Stock net	44	44	54	34	14											
Appros attendus			30													
Besoins nets						6	20	20	20	20	20	20	20	20	20	
Appros planifiés						46			60			60			20	
Lancements					46			60			60			20		

Ceci termine l'élaboration du plan MRP pour SASS1. La planification de COMP1 se déroule selon des principes similaires. Pour cette référence, cependant, il est important d'observer que les besoins bruts résultent de la superposition de deux types de besoins: ceux générés par l'inclusion de COMP1 dans le sous-assemblage SASS1 (niveau 2 de l'arborescence), et ceux générés par l'inclusion de COMP1 directement dans l'assemblage final (niveau 1 de l'arborescence). Par conséquent, on n'aurait *pas* pu planifier lesancements de COMP1 avant d'avoir planifié ceux de SASS1. En règle générale, on ne réalise les calculs MRP afférents à une référence particulière qu'après avoir planifié lesancements de tous les articles dans la composition desquels intervient cette référence. Cette technique, qui porte le nom de *low-level coding*, permet donc de regrouper les besoins bruts générés à différents niveaux pour la référence considérée.

Exemple (suite): Pour COMP1, on posera les hypothèses suivantes:

délai d'obtention: 3 semaines;

approvisionnements: toutes les 2 semaines;

stock net initial: 82 unités;

approvisionnement attendu: 25 unités en première semaine.

On commence par calculer les besoins bruts de COMP1. Par exemple, pour la semaine 4, 10 unités de COMP1 sont nécessaires afin de permettre l'assemblage final de 10 unités d'AF, et 46 unités sont nécessaires pour permettre l'assemblage de 46 unités de SASS1. Donc, $\text{besoins_bruts}[4] = 56$. Les besoins bruts des semaines 5 à 14 sont déterminés de façon similaire. Après quelques calculs, on obtient le plan suivant pour COMP1:

Semaine	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Besoins bruts			10	10	56	10	10	70	10	10	70	10	10	30	10	
Stock net	82	107	97	87	31	21	11									
Appros attendus		25														
Besoins nets								59	10	10	70	10	10	30	10	
Appros planifiés								69		80		20		40		
Lancements					69		80		20		40					

Même sur ce petit exemple, on remarque que la superposition des besoins manifestés à différents niveaux de l'arborescence d'une part, et les regroupements de commandes par lots économiques d'autre part, créent des distributions de demande irrégulières pour les composants. Les distributions de ce type (qualifiées de « lumpy » en anglais) sont typiques d'une demande dépendante. C'est précisément leur irrégularité qui rend inutilisables les méthodes classiques de gestion des stocks basées sur des distributions de demande mieux caractérisées.

8.2. La logique MRP.

Résumons à présent la discussion qui précède. Les références finales reprises dans le PDP sont décrites au moyen d'*arborescences* où elles apparaissent au niveau 0. Chacune des références intermédiaires se voit attribuer un *code de niveau* indiquant le niveau le plus bas auquel elle apparaît dans l'ensemble de toutes les arborescences (low-level coding; par exemple, la référence AF considérée dans la section précédente reçoit le code de niveau 0 et la référence COMP1 reçoit le code 2; l'idée est évidemment que les besoins bruts pour une référence résultent des besoins cumulés pour toutes les références de code inférieur). Le but ultime est de planifier les dates d'approvisionnement de chaque référence. Dans ce but, on procédera à un raisonnement en cascade, en planifiant d'abord toutes les références de code de niveau 1, puis toutes celles de code de niveau 2, 3, etc. Pour chaque référence, la méthode MRP détermine successivement les quantités suivantes:

1. Besoins bruts (explosion des nomenclatures). Les besoins bruts pour chaque référence sont dérivés des 'lancements planifiés' de chacune des références dans la composition desquelles entre la référence courante (par exemple, pour COMP1, on a tenu compte desancements planifiés de AF et de SASS1).
2. Besoins nets. Les besoins nets sont déterminés sur base des besoins bruts, des stocks nets et des approvisionnements attendus, comme expliqué dans la section précédente (voir formules (1)-(3)).
3. Approvisionnements planifiés. Les approvisionnements planifiés sont déterminés par regroupement (éventuel) des besoins nets de plusieurs périodes successives.
4. Absorption des délais. Les dates de lancement sont obtenues en retranchant les délais d'obtention des dates d'approvisionnement planifiées.

Quelques commentaires s'imposent sur ces différentes étapes.

Tout d'abord, remarquons que, pour les références finales (références de niveau 0), les approvisionnements planifiés correspondent aux quantités reprises dans le plan directeur de production et ne doivent donc pas être recalculées. Seule l'absorption des délais doit être prise en compte pour ces références.

Au sujet de l'absorption des délais, notons que le délai d'obtention est souvent considéré constant (c'est-à-dire, fixé une fois pour toutes, pour chaque référence) dans les applications de

la méthode MRP. Cette hypothèse peut paraître raisonnable dans le cadre d'un passage de commande, mais est plus discutable si la référence est produite par l'entreprise. En effet, dans un tel cas, on peut imaginer que le délai d'obtention est variable et est fonction, en autres facteurs, de:

- *La taille des lots à produire.* En pratique, il est assez facile de tenir compte de ce facteur, et le délai d'obtention peut être ajusté en conséquence par les logiciels MRP. Alternativement, on peut également calculer un délai d'obtention sur base d'un lot de taille moyenne et le maintenir constant. Remarquons à cet égard que le délai (de livraison ou de fabrication) inclut les temps opératoires, plus les temps d'attente, de transport et de lancement. Si même les temps opératoires sont directement influencés par la taille du lot à produire, ces temps opératoires ne représentent souvent que 5 à 10% du délai d'obtention total, qui peut donc être considéré comme relativement insensible à la taille du lot.
- *La charge de travail de l'atelier.* Cette variabilité n'est *pas prise en compte* dans les calculs MRP et ne pourrait pas l'être de façon simple. Elle ne sera donc considérée qu'a posteriori, lors des calculs de courbes de charge. Il s'agit là d'un des défauts les plus graves de l'approche MRP.

Finalement, il est intéressant d'observer que la logique MRP relève pour une grande part d'une approche à flux tirés, semblable à celle utilisée par les systèmes JAT: la demande de chaque référence déclenche la production de ses composants, et le but est de démarrer cette production 'juste à temps' en tenant compte des délais d'obtention. Une différence essentielle existe cependant entre les deux approches: le JAT utilise la logique 'flux tirés' dans le pilotage même des ateliers, c'est-à-dire dans le très court terme, alors que MRP est principalement un outil de planification à court terme. En particulier, les sous-périodes considérées en JAT sont très courtes (instantanées) et la seule politique de lancement des ordres de fabrication utilisée est celle du lot-pour-lot. A l'inverse, la logique MRP génère des ordres de production destinés à répondre à des besoins futurs, planifiés mais non (encore) concrétisés. C'est la raison pour laquelle on décrit souvent la planification MRP comme un système à *flux poussés* ('push system').

8.3. Systèmes d'information utilisés par les logiciels MRP.

Les systèmes MRP comptent parmi les outils de planification les plus répandus dans la pratique de la gestion des entreprises (on peut supposer que leur simplicité conceptuelle favorise considérablement leur adoption). En particulier, de nombreux logiciels commerciaux performants existent sur le marché. Le principal obstacle à l'adoption de tels systèmes reste probablement que les besoins en information requis pour l'implantation et pour la bonne utilisation de ces systèmes sont extrêmement élevés. Les bases de données nécessaires peuvent être gigantesques (typiquement, plusieurs milliers de références sont concernées) et doivent être établies et maintenues de façon très rigoureuse pour éviter le déraillement du système. Passons brièvement en revue l'essentiel de ces besoins en information.

Plan directeur de production: définit les besoins bruts pour les références finales.

Fichier des références: nom, numéro, description, code de niveau, délai d'obtention, stratégie de groupement des commandes (lot-pour-lot, périodique, ...), etc.

Fichier des arborescences: référence et quantité de chaque composant intervenant dans une référence donnée; souvent établi à un seul niveau de profondeur (par itération, on peut reconstituer toute l'arborescence d'un produit fini).

Systèmes d'information sur l'état des stocks et sur les approvisionnements attendus: à noter que, contrairement aux données précédentes, qui sont de nature relativement statique, les données concernées ici évoluent de façon permanente et leur gestion doit pouvoir s'effectuer de façon dynamique.

8.4. Dynamique du système MRP.

Dans la pratique quotidienne des entreprises, au-delà de son rôle d'outil de planification à court terme, le système MRP est souvent utilisé comme un système *d'information* permettant de gérer les *lancements d'ordres de fabrication* ou de *commande* et de *contrôler* la production des références.

A chaque période, l'utilisateur doit

- lancer les ordres de production ou les commandes, soit sur base des quantités planifiées par le système MRP, soit en fonction d'autres priorités (si la capacité disponible est insuffisante, on préférera par exemple produire pour une commande en attente que pour un stock de sécurité, etc); lesancements calculés par MRP sont donc généralement considérés comme des *propositions* de lancement, qui doivent encore recevoir l'aval du planificateur, plutôt que comme des ordres de fabrication automatisés;
- enregistrer les quantités approvisionnées;
- corriger les données en fonction des événements réels (erreurs de stocks, approvisionnements en retard, etc).

L'approche utilisée est généralement de gérer par exception: le système génère automatiquement des messages indiquant les références pour lesquelles une action doit être entreprise.

Lorsque des changements de données sont introduits, il devient nécessaire de recalculer les ordres deancements. Si ceci se produit trop fréquemment, les plans deviennent instables: on parle de nervosité du système. Une pratique courante est de générer un nouveau plan MRP à intervalles réguliers, par exemple toutes les semaines, sur base d'un 'horizon glissant'.

8.5. Techniques de détermination des lots (lot sizing).

Considérons les données suivantes, extraites d'un tableau de calcul MRP.

Exemple:

Semaine	1	2	3	4	5	6	7	8	9	10	11	12
Besoins nets	10	62	12	130	154	129	88	52	124	160	238	41

La question qui se pose à présent est: comment faut-il satisfaire ces besoins nets? Le groupement des commandes ou des ordres de fabrication entraîne des avantages (moins

d'administration, réduction des coûts et des temps de lancement, rabais éventuels, etc) et des inconvénients (maintien de stocks sur des périodes prolongées, mobilisation de la capacité productive pour une seule référence). La politique de groupement idéale est donc celle qui parvient à équilibrer ces avantages et inconvénients de façon judicieuse.

Pour préciser quelque peu la notion de choix 'judicieux', nous allons considérer un modèle d'optimisation simple, celui de la *détermination des lots économiques* (economic lot sizing model).

Données : pour un horizon de planification subdivisé en T périodes $\{1, 2, \dots, T\}$, on connaît
 d_t = demande (besoin net) à la période t
 c_L = coût de lancement de la production (coût fixe encouru pour chaque lancement)
 c_p = coût de possession d'une unité pendant une période.

Problème : déterminer un plan de production satisfaisant la demande de chaque période et de coût total minimum.

Exemple (suite): Les besoins nets ont été définis plus haut. Nous supposons en outre que $c_L = 54$ et $c_p = 0,40$.

Le problème de détermination des lots économiques fournit un modèle mathématique précis pour les décisions de groupement des approvisionnements, telles qu'elles se posent dans la planification MRP. Par ailleurs, le même modèle pourrait également être utilisé dans le cadre de l'élaboration du plan directeur de production, comme nous l'avons déjà mentionné au Chapitre 6 (Section 6.3).

Nous allons à présent décrire quelques-unes des nombreuses méthodes utilisées pour attaquer le problème de la détermination des lots économiques (comme nous le verrons, certaines de ces approches ne résolvent le problème que de façon très approximative).

1. Politique lot-pour-lot.

Dans une politique d'approvisionnement lot-pour-lot, les besoins nets de chaque période sont satisfaits grâce à des approvisionnements distincts.

Exemple (suite):

Semaine	1	2	3	4	5	6	7	8	9	10	11	12
Besoins nets	10	62	12	130	154	129	88	52	124	160	238	41
Appros	10	62	12	130	154	129	88	52	124	160	238	41

Coût total: $12 \cdot 54 = 648$.

2. Quantité fixe.

Dans cette approche, chaque commande ou ordre de fabrication porte sur un nombre d'unités constant que nous noterons Q (ou sur un multiple de Q , si cela s'avère nécessaire pour couvrir les besoins d'une période). La quantité Q est sélectionnée au préalable sur base de plusieurs critères potentiels. Elle peut par exemple correspondre à une quantité de commande standardisée imposée par le fournisseur, à la capacité d'un récipient ou d'un emballage (container, fût) ou à celle d'un moyen de transport ou de manutention (palette, camion). Dans un objectif de minimisation des coûts, il est également courant de choisir Q égal à la *quantité économique de commande* QEC, définie comme suit:

$$QEC = \sqrt{2c_L D / c_p}$$

où D représente la demande prévisionnelle moyenne par période. (Une justification et une discussion plus approfondie de cette définition seront fournies au Chapitre 9. Pour l'instant, il suffira d'admettre le postulat selon lequel QEC est une taille de lot 'idéale' ou 'optimale' pour la firme. Nous reviendrons au Chapitre 9 sur la validité de cette affirmation.)

Il faut remarquer que, dans l'application de cette méthode, D n'est généralement pas calculé sur base des besoins nets exprimés pour les T périodes considérées, mais plutôt sur base de prévisions effectuées sur un horizon plus long (le but étant de maintenir une taille de lot fixe).

Exemple (suite): si on suppose que $D = 90$, alors $QEC \approx 156$. On obtient ainsi les groupements suivants (notons qu'en période 11, il est nécessaire de planifier l'approvisionnement de $312 = 2 \times 156$ unités pour couvrir les besoins de la période) :

Semaine	1	2	3	4	5	6	7	8	9	10	11	12
Besoins nets	10	62	12	130	154	129	88	52	124	160	238	41
Appros	156			156	156	156		156		156	312	
Stock net	146	84	72	98	100	127	39	143	19	15	89	48

Chaque approvisionnement de taille $Q = 156$ (ou $Q = 312, 468$, etc.) permet de reconstituer les stocks et donc de répondre aux besoins d'une ou plusieurs semaines consécutives. Dès qu'un approvisionnement est planifié, il est donc nécessaire de recalculer les stocks nets de chacune des semaines suivantes, afin de déterminer la première sous-période où les stocks redeviennent négatifs et où il est donc à nouveau nécessaire de réapprovisionner.

Cette politique entraîne un coût total de:

$$7 * 54 + (146 + 84 + 72 + \dots 48) * 0,40 = 770.$$

(Nous avons supposé ici que les 48 unités en stock à la fin de l'horizon seraient écoulées au cours de la semaine suivante.)

3. Nombre de périodes fixe (Period Order Quantity).

Dans cette approche, lorsqu'une commande ou un ordre de fabrication est lancé, il doit suffire à couvrir la demande de P périodes, où P est un paramètre fixé d'avance. L'avantage principal de cette politique d'approvisionnement est d'ordre organisationnel: la régularité des commandes présente en effet des avantages évidents tant pour l'entreprise que pour ses fournisseurs.

Le choix de la période P peut être basé sur différents critères, quantifiables ou non. Dans le but de minimiser les coûts encourus, on recommande souvent de choisir P égal à la *période économique entre commandes* PEC, calculée sur base de la quantité de commande économique QEC. Plus précisément, si la demande prévisionnelle moyenne par période est de D unités, alors PEC est obtenu en arrondissant la valeur calculée

$$QEC/D = \sqrt{2c_L / Dc_p} .$$

(Il est facile de vérifier que des lots de taille QEC devront être commandés toutes les (QEC/D) périodes afin de satisfaire à la demande de D unités par période. Si l'on admet à nouveau que QEC est une taille de lot 'optimale' pour la firme, alors PEC peut donc être vu comme une périodicité de commande optimale. Nous reviendrons sur ce point au Chapitre 9.)

Exemple (suite): en supposant comme avant que $D = 90$, alors $QEC/D = 1,73$ et PEC peut être choisi égal à 2; on obtient ainsi les groupements suivants :

Semaine	1	2	3	4	5	6	7	8	9	10	11	12
Besoins nets	10	62	12	130	154	129	88	52	124	160	238	41
Livraisons	72		142		283		140		284		279	

Cette politique entraîne un coût total de:

$$6 * 54 + (62 + 130 + 129 + \dots 41) * 0,40 = 553,60.$$

Ici encore, la valeur de D ne découle pas des besoins nets exprimés pour les T périodes considérées, mais est calculée sur base de prévisions à plus long terme (le but premier étant d'établir une certaine périodicité des commandes ou des lancements de production).

4. Equilibrage des coûts de lancement et de possession (Part Period Balancing).

Cette heuristique peut être motivée par les observations suivantes. Si l'entreprise produit à la période s pour satisfaire la demande des périodes s, s+1, s+2, ..., t, les coûts totaux encourus ont la forme

$$C(s,t) = c_L + c_p(s,t).$$

Puisque le but est de réaliser un compromis optimal entre coûts de lancement et coûts de possession, il semble raisonnable de choisir t (pour une valeur donnée de s) de telle sorte que c_L soit approximativement égal à $c_p(s,t)$. Intuitivement, tant que $c_L > c_p(s,t)$, la firme préférera

augmenter ses coûts de possession de stocks plutôt que de subir un coût de lancement de production supplémentaire (notons cependant qu'il ne s'agit là que d'un argument heuristique, puisqu'il ne tient aucun compte de l'impact de cette décision sur les sous-périodes ultérieures).

La méthode d'équilibrage des coûts applique donc itérativement la règle heuristique suivante: une commande ou un ordre de fabrication lancé à la période s doit couvrir la demande des périodes s à t , où t est choisi de telle sorte que les coûts de possession $c_p(s,t)$ associés à ce lot soient aussi proches que possible du coût de lancement c_L .

Exemple (suite): Si on produit à la période 1 pour couvrir les besoins des périodes 1 à ...

- ... 1, alors $c_p(1,1) = 0$
- ... 2, alors $c_p(1,2) = 62 * 0,40 = 24,80$
- ... 3, alors $c_p(1,3) = 24,80 + 12 * 0,80 = 34,40$
- ... 4, alors $c_p(1,4) = 34,40 + 130 * 1,20 = 190,40$

Puisque $c_L = 54$ est plus proche de 34,4 que de 190,4, on produira à la période 1 pour couvrir les besoins des périodes 1, 2 et 3, soit 84 unités. En répétant cette approche à partir de la période 4, on obtient finalement le plan suivant, de coût total égal à 600.

Semaine	1	2	3	4	5	6	7	8	9	10	11	12
Besoins nets	10	62	12	130	154	129	88	52	124	160	238	41
Livraisons	84			284		217		176		398		41

5. Méthode de Silver et Meal.

Pour motiver cette méthode, considérons à nouveau les coûts $C(1,t)$ définis ci-dessus. Typiquement, lorsque t augmente, le coût moyen par période, $C(1,t)/t$, est d'abord décroissant (parce que le coût fixe de lancement est amorti sur plusieurs périodes) puis finit par augmenter (parce que les coûts de possession sont de plus en plus élevés).

Pour illustrer cette remarque, considérons le cas (très particulier) où la demande serait constante sur l'horizon de planification: $d_t = d$ pour $t = 1,2,\dots,T$. Dans ce cas, on obtient:

$$C(1,t) = c_L + c_p d + 2c_p d + \dots + (t-1)c_p d = c_L + \frac{1}{2} t (t-1) c_p d$$

et donc $C(1,t)/t = c_L/t + \frac{1}{2} (t-1)c_p d$. On vérifie aisément que cette fonction est d'abord décroissante puis croissante.

L'heuristique de Silver et Meal prescrit de produire pour couvrir les besoins des périodes 1 à t , où t est la dernière période pour laquelle les coûts moyens par période sont décroissants (il est à remarquer que cette règle peut ne conduire qu'à un minimum local de $C(1,t)/t$; pour un exemple, voir Silver et Peterson 1985, p.235).

Exemple (suite): Si on produit à la période 1 pour les périodes 1 à ...

- ... 1, alors $C(1,1)/1 = 54$
- ... 2, alors $C(1,2)/2 = (54 + 62 * 0,40)/2 = 39,40$
- ... 3, alors $C(1,3)/3 = (54 + 34,40)/3 = 29,47$
- ... 4, alors $C(1,4)/4 = (54 + 190,40)/4 = 61,10$

Donc, la méthode de Silver-Meal prescrit de produire lors de la première semaine pour satisfaire aux besoins des semaines 1 à 3. En itérant cette procédure, on obtient :

Semaine	1	2	3	4	5	6	7	8	9	10	11	12
Besoins nets	10	62	12	130	154	129	88	52	124	160	238	41
Livraisons	84			130	283		140		124	160	279	

Cette politique entraîne un coût total de 501,20. Pour notre exemple, il s'agit là du coût total optimal. Mais en règle générale, la méthode de Silver-Meal n'est qu'une heuristique et ne fournit donc aucune garantie d'optimalité.

6. Algorithme de Wagner et Whitin.

Cet algorithme permet de calculer la solution optimale du problème de détermination des lots économiques par *programmation dynamique*. Le principe utilisé peut être décrit simplement: le *coût optimal* $C^*(1,t)$ à encourir pour satisfaire la demande totale des périodes 1 à t (en produisant éventuellement lors de plusieurs périodes) est égal au minimum des expressions

$$\begin{aligned}
 & C^*(1,t-1) + C(t,t) (= C^*(1,t-1) + c_L), \\
 & C^*(1,t-2) + C(t-1,t), \\
 & C^*(1,t-3) + C(t-2,t), \\
 & C^*(1,t-4) + C(t-3,t), \\
 & \dots \\
 & C^*(1,1) + C(2,t), \\
 & C(1,t).
 \end{aligned}$$

En effet, le dernier lancement aura lieu soit à la période t, soit à la période t-1, ..., soit à la période 2, soit à la période 1; pour les périodes précédant ce dernier lancement, le plan de production doit également être optimal.

Donc, si l'on connaît $C^*(1,1)$, $C^*(1,2)$, ..., $C^*(1,t-1)$, on peut aisément calculer $C^*(1,t)$. Par ailleurs, on a évidemment $C^*(1,1) = c_L$. Les relations précédentes permettent donc de calculer récursivement $C^*(1,2)$, puis $C^*(1,3)$, $C^*(1,4)$, ..., $C^*(1,T)$. Il suffit alors de remarquer que $C^*(1,T)$ n'est rien d'autre que le coût optimal recherché.

Exemple (suite):

$$C^*(1,1) = 54.$$

$$C^*(1,2) = \min \{ C^*(1,1) + C(2,2), C(1,2) \}$$

$$= \min \{ 54 + 54, 78.80 \} = 78.80$$

$$C^*(1,3) = \min \{ C^*(1,2) + C(3,3), C^*(1,1) + C(2,3), C(1,3) \}$$

$$= \min \{ 78.80 + 54, 54 + 54 + 4.8, 54 + 24.80 + 9.6 \}$$

$$= 88.40$$

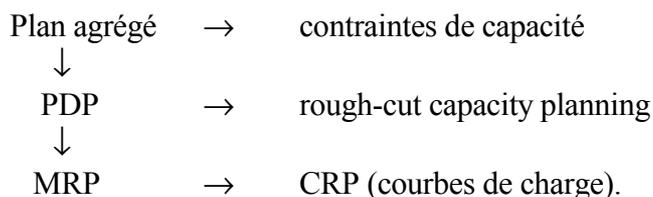
etc.

La méthode de Wagner-Whitin est relativement rapide: le nombre de calcul à effectuer est de l'ordre de T^2 . En pratique, cependant, les méthodes de Silver-Meal ou d'équilibrage des coûts sont plus souvent utilisées. En effet, elles sont encore plus rapides et produisent généralement des solutions de très bonne qualité. Par ailleurs (et c'est probablement là un argument plus fort), il faut bien remarquer que l'algorithme de Wagner-Whitin ne garantit une solution optimale que pour le modèle très simple que nous avons considéré ici. Ce modèle ne représente pas toujours fidèlement la situation réelle; en particulier, il postule un horizon fini alors que, dans la réalité, cet horizon devrait plutôt être considéré comme étant 'infini' (en pratique, les calculs MRP doivent être répétés pour un horizon 'glissant', ou 'déroulant'). Pour des situations de ce dernier type, la méthode de Silver-Meal ou d'autres méthodes peuvent produire – et produisent fréquemment – des solutions de coût moins élevé que Wagner-Whitin. Elles présentent également l'avantage d'être moins sensibles à l'ajout de nouvelles données en fin d'horizon, dans le cadre de la planification glissante évoquée plus haut (puisque les données de fin d'horizon n'affectent que les lots de fin d'horizon; la 'myopie' inhérente aux heuristiques gourmandes produit donc ici des effets stabilisateurs intéressants, en réduisant la nervosité des plans).

8.6. Planification des besoins en capacité (CRP) - Calcul des courbes de charge.

La logique MRP ne tient compte d'aucune contrainte de capacité: elle raisonne à *capacité infinie*. Lorsque MRP est utilisé, il est donc nécessaire de vérifier a posteriori si les plans développés sont réalisables. Ceci peut se faire en pratique par le calcul d'une *courbe de charge* pour chaque centre de production.

Remarquons ici que, dans une approche hiérarchique de la planification de la production, l'ajustement des *besoins* en capacité productive à la capacité *disponible* est pris en compte à chaque niveau de planification. En effet, la coordination charge-capacité s'est traduite dans l'élaboration du plan agrégé de production par l'inclusion de contraintes explicites; celles-ci ne peuvent cependant refléter la coordination qu'à un niveau très global, vu l'agrégation des produits en familles et la longueur des sous-périodes de planification. Une approche similaire à celle qui va être présentée ci-dessous, mais plus grossière, a également été mentionnée lorsqu'il s'agissait de vérifier la réalisabilité du PDP (*rough-cut capacity planning*: voir Section 6.3). Pour le PDP, cependant, il n'était possible de tenir compte que des références finales, mais pas de l'explosion en composants ni des délais d'obtention, des stocks présents, etc.



La plupart des logiciels de MRP contiennent une base de données décrivant les *gammes opératoires* des références (c'est-à-dire, la définition des séquences d'opérations à subir, des centres de production requis, des temps opératoires et de lancement sur chaque poste de travail, ...). Sur base de ces données, il est possible d'estimer le temps requis pour le traitement d'une unité d'une référence quelconque par un centre de production. Par addition de ces temps pour les diverses références, on obtient alors une courbe de charge estimée pour chaque poste de travail, c'est-à-dire une estimation du nombre d'heures requises sur chaque poste au cours de chaque sous-période de planification. Remarquons que, dans ce plan à court terme, il est possible de tenir compte d'un niveau de détail relativement élevé (gammes opératoires, dates des lancements planifiés, délais d'obtention, etc.).

Exemple: La référence A est produite par lots, et chaque lot requiert un délai de production de 2 semaines. Durant la première semaine, le lot est en production au poste de travail M1. Le temps opératoire estimé sur M1 est de 0.6 heure par unité de A. Au cours de la deuxième semaine, le lot passe en production au poste de travail M2. Le temps opératoire sur M2 est de 0.2 heure par unité produite.

Le plan MRP prévoit les lancements suivants pour la référence A :

Semaine	1	2	3	4	5	6
Lancements planifiés	40			30		

Sur base de ces données, on peut estimer comme suit la charge placée sur chaque poste par la référence A.

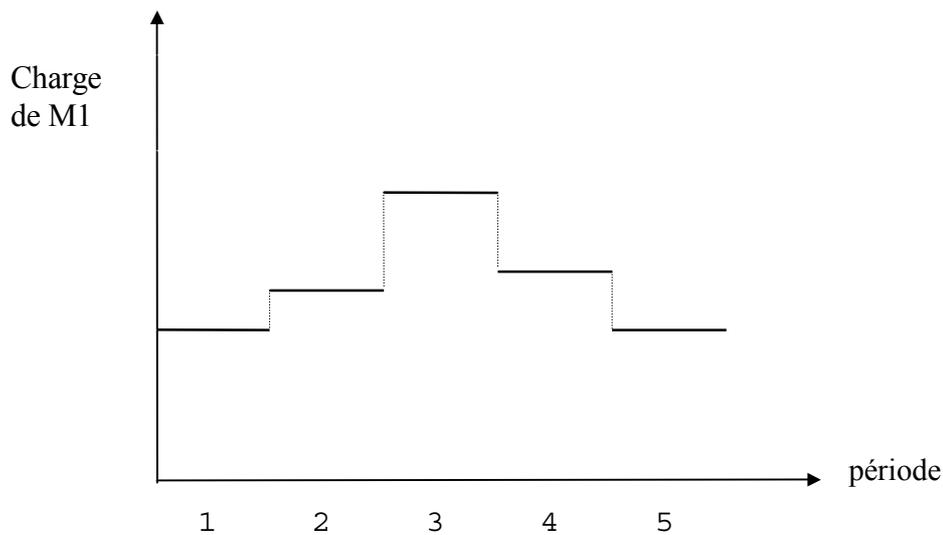
Charge de M1 due à la référence A :

- semaine 1 : $40 * 0.6 = 24$ heures
- semaine 4 : $30 * 0.6 = 18$ heures

Charge de M2 due à A :

- semaine 2 : $40 * 0.2 = 8$ heures
- semaine 5 : $30 * 0.2 = 6$ heures.

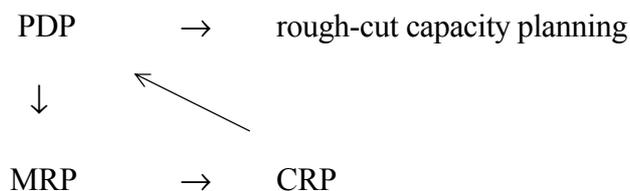
En cumulant les estimations de charge calculées pour chaque référence, on peut ainsi développer des *courbes (profils) de charge* représentant la charge de travail estimée de chaque poste de travail sur l'horizon considéré.



Si la capacité requise pour certaines périodes dépasse la capacité disponible, il faut procéder à des ajustements:

- lancement anticipé de la production (c'est-à-dire modification du plan MRP)
- heures supplémentaires, ... (modifications de capacité)
- en dernier ressort: modification du PDP.

Cette dernière possibilité est prise en compte par certains systèmes qui permettent de « fermer la boucle » de la planification hiérarchique et de modifier le PDP sur base des calculs de CRP (*closed-loop MRP systems*).



Cette possibilité permet également d'établir des PDP plus réalistes en intégrant les modules PDP, MRP, CRP dans un même système informatisé et en utilisant les modules MRP et CRP pour 'simuler' les conséquences des propositions faites par le PDP (voir Section 6.3).

8.7. Epilogue

Les principes de base de la planification MRP ont été développés dans les années 60, avec pour objectif de mieux gérer l'approvisionnement des matières et composants nécessaires à la fabrication des références finales. Au fil du temps, ces concepts ont été progressivement élargis pour englober la planification et le contrôle d'autres ressources de l'entreprise, souvent autour de bases de données communes : capacité des centres de production (CRP), ressources

humaines (qualifications, disponibilité, paie, ...), comptabilité et finance (budgets, facturation et encaissements, comptabilité des coûts, ...), marketing et ventes (bases de données clients, prévisions de demande, ...), logistique (gestion de la chaîné d'approvisionnement, transport, ...), gestion documentaire, etc. On a ainsi vu successivement l'émergence de systèmes de *Manufacturing Resource Planning* (parfois appelés MRPII) puis de systèmes d'*Enterprise Resource Planning* (ERP). Bien que les systèmes ERP, comme leur nom l'indique, soient clairement de lointains descendants des systèmes MRP de première génération, ils couvrent à présent bien d'autres fonctions que celles de gestion de la production. Leur présentation dépasse donc le cadre de ce cours introductif : on renverra l'étudiant intéressé à la littérature florissante sur ce sujet ou à des cours complémentaires pour plus d'informations.

Bibliographie.

A. Courtois, M. Pillet et C. Martin, *Gestion de Production*, Les Editions d'Organisation, Paris, 1989.

V. Giard, *Gestion de la Production*, 2ème édition, Economica, Paris, 1988.

S.C. Graves, A.H.G. Rinnooy Kan et P.H. Zipkin (eds.), *Logistics of Production and Inventory*, North-Holland, Amsterdam, 1993.

E. Silver et R. Peterson, *Decision Systems for Inventory Management and Production Planning*, J. Wiley & Sons, New York, 1985.

J.W. Toomey, *MRPII: Planning for Manufacturing Excellence*, Chapman & Hall, New York, 1966.

T.E. Vollmann, W.L. Berry et D.C. Whybark, *Manufacturing Planning and Control Systems*, Business One Irwin, Homewood, Illinois, 1992.

CHAPITRE 9

PLANIFICATION A COURT TERME: LA GESTION DES STOCKS POUR LES ARTICLES A DEMANDE INDEPENDANTE.

9.1. Demande indépendante.

Ce concept a déjà été brièvement introduit. La demande pour une référence est qualifiée de dépendante si elle est explicitement considérée comme fonction de la demande pour d'autres références. Sinon, elle est appelée indépendante. Pour les références à demande indépendante, on se contente d'une description de la demande sous forme de prévision et/ou de distribution statistique (bien que, même dans ce cas, la demande peut être certaine/déterministe, ou aléatoire/stochastique).

Exemples:

- références finales
- pièces de rechange
- outillage.

On traite souvent de la même façon que les articles à demande indépendante tous ceux dont la demande résulte de l'accumulation de nombreux besoins indépendants (au sens statistique du terme) et pour lesquels la demande peut donc être décrite plus facilement de façon statistique que de façon causale. Il s'agit là, en règle générale, d'articles à faible valeur et à taux d'utilisation élevé.

Exemples:

- clous et vis dans un atelier de menuiserie
- fournitures de bureau.

9.2. Politiques de gestion des stocks: Les choix à effectuer.

Une politique de gestion est déterminée par la réponse à trois questions.

1. *Comment l'entreprise s'informe-t-elle sur l'état des stocks ?*
2. *Quand (sous quelles circonstances) faut-il réapprovisionner ?*
3. *Combien d'unités faut-il réapprovisionner lors de chaque passation de commande ?*

Passons successivement ces trois décisions en revue.

1. Information sur l'état des stocks.

On distingue deux grands principes:

- *inventaire permanent, ou continu:* l'état des stocks est connu en permanence;
- *inventaire périodique:* l'état des stocks est déterminé à intervalles réguliers, par exemple tous les mois.

En pratique, les systèmes d'inventaire permanent sont souvent implantés via l'enregistrement systématique des transactions entrées/sorties. De tels systèmes peuvent se révéler beaucoup plus coûteux que les systèmes périodiques. Mais leur coût a maintenant tendance à décroître, en fonction de l'informatisation croissante qui facilite l'enregistrement des transactions. (Exemple: les grandes surfaces de distribution pratiquent un inventaire permanent lié à l'utilisation des codes-barres.) L'utilisation de tels systèmes ne se justifie cependant pas nécessairement pour tous les articles.

L'inventaire périodique impose l'inspection des stocks à intervalles réguliers, même si aucune transaction n'a eu lieu depuis la dernière inspection. Outre sa simplicité organisationnelle, cette approche permet de détecter des détériorations, erreurs, vols, etc. (en particulier dans le cas d'articles à rotation très lente).

2. Quand faut-il réapprovisionner ?

A nouveau, deux types de réponses sont possibles:

- *gestion par point de commande*: on approvisionne la référence si la position de stock est inférieure à une certaine quantité prédéterminée, appelée *point de commande*;
- *commande périodique*: on approvisionne à intervalles réguliers, toutes les R unités de temps (par exemple, toutes les deux semaines); R est appelé la *période entre commandes*.

Ces deux approches connaissent leurs avantages (+) et leurs inconvénients (-) respectifs.

- **Commande périodique:**
 - + facilité organisationnelle: les commandes d'approvisionnement sont passées à un moment prévisible, qui coïncide typiquement (mais pas nécessairement) avec celui de l'inventaire dans le cas d'un système d'inventaire périodique; cette approche simplifie l'administration, le transport, la réception des commandes, ainsi que la tâche des fournisseurs; elle facilite également le groupement des commandes de différents articles auprès d'un même fournisseur (si les périodicités sont coordonnées pour tous ces articles).
 - + en cas de périodicité de la demande, il existe une possibilité de coordination avec la période de commande.
 - dans une application stricte du système, il n'est pas possible de réagir à des situations inattendues se produisant entre deux instants de commande; la probabilité de rupture de stock s'en trouve donc accrue.
- **Point de commande:**
 - + offre plus de souplesse, et permet donc d'obtenir des coûts inférieurs à ceux du système précédent;
 - + couplée à un système d'inventaire permanent, cette politique de gestion permet de mieux contrôler le niveau de stock, et donc de diminuer les chances de rupture tout en maintenant des stocks de sécurité plus faibles que par la gestion périodique;
 - difficultés organisationnelles, dues à l'irrégularité des époques de passation de commande.

3. Combien d'unités faut-il réapprovisionner ?

Ici encore, deux classes de réponses sont possibles :

- une quantité fixe Q (la *quantité de commande*) à chaque réapprovisionnement;
- la quantité qui ramène la position de stock à un niveau prédéterminé S (le *niveau de rechargement*, ou *up-to-level*); c'est-à-dire, quantité commandée = $S -$ position de stock (si cette quantité est positive).

9.3. Politiques de gestion des stocks: Les réponses possibles.

En théorie, n'importe quelle combinaison de réponses aux alternatives discutées dans le paragraphe précédent détermine une politique de gestion des stocks qui pourrait être adoptée pour la gestion d'un article particulier. En pratique, certaines combinaisons présentent des avantages sur les autres et s'imposent donc plus fréquemment. En particulier, on identifie les politiques suivantes:

- politique (s,Q): inventaire permanent, point de commande (s), quantité de commande (Q);
- politique (s,S) (ou min/max): inventaire permanent, point de commande (s), niveau de rechargement (S);
- politique (R,S): inventaire périodique (R), commande périodique (R), niveau de rechargement (S) (noter que les périodicités d'inventaire et de commande sont égales);
- politique (R,s,S): inventaire périodique (R), point de commande (s), niveau de rechargement (S).

Ayant choisi une des politiques décrites ci-dessus, on doit encore fixer les valeurs des paramètres définissant le système (c'est-à-dire les valeurs de s , S , R et Q) de façon à minimiser la fonction de coût considérée. Ce problème d'optimisation est généralement complexe et on le résout donc de façon heuristique (voir la Section 9.7 ci-dessous pour quelques exemples).

Formulons encore quelques commentaires sur les (dés)avantages respectifs des systèmes mentionnés.

La politique (s,Q) est simple à implanter et permet un bon suivi du stock, mais requiert un système d'inventaire permanent qui peut être onéreux et entraîne des réapprovisionnements à intervalles fluctuants. De plus, lorsque la sortie de stock qui provoque le passage sous le point s est très grande, la quantité de commande Q prédéterminée peut être trop faible pour ramener la position de stock à un niveau satisfaisant.

L'adoption d'un système (s,S) fournit une solution adéquate à cette dernière objection. En fait, on démontre que le système (s,S) optimal entraîne des coûts inférieurs au système (s,Q) optimal. Mais il s'agit là d'un résultat dont la valeur pratique est discutable, puisque les problèmes d'optimisation rencontrés ne sont généralement pas résolus de façon exacte.

La politique (R,S) présente les avantages organisationnels déjà discutés, mais génère souvent des stocks de sécurité importants.

Du point de vue de la minimisation des coûts, le système (R,s,S) permet de développer des politiques moins coûteuses que les autres systèmes (observer que (s,S) et (R,S) n'en sont que des cas particuliers, obtenus lorsque $R \rightarrow 0$ et lorsque $s \rightarrow \infty$, respectivement). Mais les valeurs optimales des paramètres sont, ici encore, très difficiles à calculer.

9.4. Classification ABC.

Quelle politique de gestion de stocks faut-il adopter pour une référence particulière ? Il n'y a pas de réponse simple à cette question. Le critère de coût ne fournit certainement qu'un éclairage très partiel sur cette question. Chaque politique connaît des avantages et des inconvénients dont certains ont été mentionnés ci-dessus et qui sont difficilement quantifiables. Enfin, d'autres aspects, comme par exemple « l'importance stratégique » qu'un article revêt pour la firme, peuvent aussi influencer le choix de la politique.

Pour guider son choix, la firme peut utiliser la notion de classification ABC, qui repose elle-même sur la notion de distribution par valeur. La *valeur d'utilisation annuelle* (Annual Dollar Usage Value) d'un article stocké est le coût total de toutes les unités de cet article utilisées par la firme pendant une année.

Exemple : 5000 unités d'un article coûtant 20 francs/unité sont utilisées par an; la valeur d'utilisation annuelle de cet article est donc égale à 100.000 francs.

Supposons que les articles stockés sont numérotés 1, 2, ..., N par valeur d'utilisation décroissante. Pour chaque $i = 1, 2, \dots, N$, on peut alors calculer la valeur cumulée $VC(i)$, où $VC(i)$ est la somme des valeurs d'utilisation des articles 1, 2, ..., i.

Exemple :

Article	Valeur d'utilisation	VC
1	100.000	100.000
2	80.000	180.000
3	60.000	240.000
...		
50	2.000	400.000

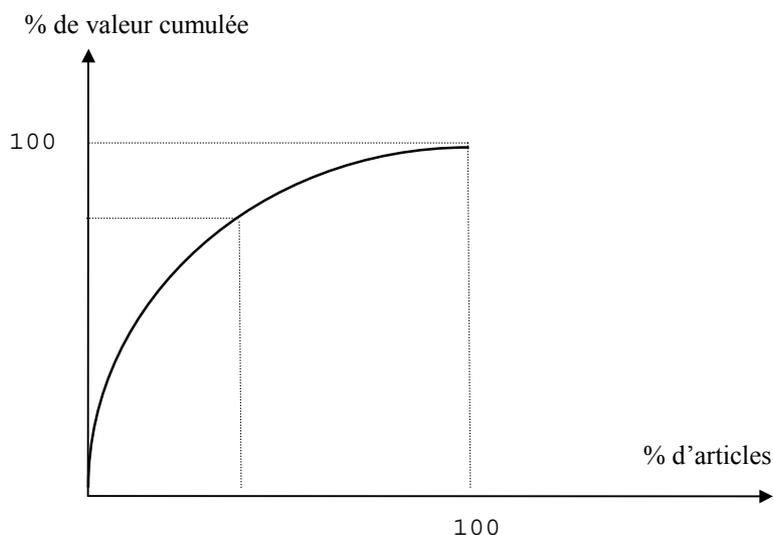
Soit $DV(i) = VC(i)/VC(N)$. La courbe des points $P_i = (i/N, DV(i))$ est appelée *courbe de distribution par valeur*; chaque point P_i indique le pourcentage de la valeur cumulée totale représentée par les i premiers pourcents d'articles.

Exemple :

Article	i/N	$DV(i)$
1	$1/50 = 0,02$	0,25
2	$2/50 = 0,04$	0,45

3	$3/50 = 0,06$	0,60
...		
50	$50/50 = 1,00$	1,00

Une représentation graphique de cette courbe est donnée par la Figure 1 ci-dessous. La forme (concave) de cette courbe est typique. Empiriquement, en effet, on observe souvent que la distribution par valeur suit approximativement une *loi de Pareto*, ou *loi des 80/20*: 80% de la valeur cumulée totale est concentrée sur 20% des articles. Cette observation conduit à consacrer un maximum d'efforts et d'attention à un petit nombre d'articles représentant une valeur cumulée importante et à utiliser des méthodes réclamant moins d'énergie ou de moyens pour les autres articles.



Courbe de distribution par valeur

En pratique, cette observation conduit de nombreuses firmes à utiliser un système de gestion des stocks basé sur une *classification ABC* des articles: les 5 ou 10% d'articles ayant la plus grande valeur d'utilisation sont placés dans une catégorie A, les 50 % suivants dans une catégorie B, et le reste des articles dans une catégorie C. Suivant la loi de Pareto, la catégorie A représente donc typiquement 50% de la valeur d'utilisation totale, la catégorie B environ 40-45%, et la catégorie C quelques pourcents seulement.

Bien sûr, d'autres critères que la seule valeur d'utilisation peuvent également entrer en ligne de compte pour définir la classification ABC. Le principe essentiel à observer est que les articles placés dans la catégorie A doivent être ceux pour lesquels des gains ou des pertes importants peuvent être réalisés au niveau de la gestion des stocks, alors que la gestion des articles de type C ne peut potentiellement conduire qu'à des différences marginales. Ce critère est généralement bien reflété par la valeur d'utilisation. Par exemple, des stocks excessifs dans la catégorie A se traduisent par des surcoûts plus élevés que dans les catégories B ou C. Mais d'autres critères, totalement différents, pourraient également être considérés, comme par exemple les conséquences entraînées par une rupture de stock (l'article est-il ou non substituable par un autre, quelle est son importance pour l'image de marque de la firme, etc), le volume de stockage requis, etc.

En termes de politique de gestion, les articles sont généralement gérés de façon uniforme au sein de chaque catégorie, alors que des différences plus profondes se marquent entre catégories. Par exemple :

- les articles A sont gérés individuellement et de façon continue; ils sont placés sous la responsabilité directe d'un responsable (cadre) qui suit leur évolution de près; les paramètres de gestion (coûts, quantités s , S , etc) sont revus régulièrement et des efforts sont consacrés à leur optimisation;
- les articles B peuvent être gérés par des systèmes demandant moins d'intervention humaine, qui déclenchent par exemple les ordres de commandes de façon automatisée; l'optimisation des paramètres de gestion est moins cruciale;
- les articles C sont généralement contrôlés périodiquement, à longs intervalles (une gestion continue est également possible, pour autant que le système de déclenchement des commandes soit suffisamment simple; par exemple, système à deux casiers); ils sont regroupés en familles de produits similaires plutôt que gérés individuellement, ce qui diminue la quantité d'informations à traiter et simplifie les approvisionnements.

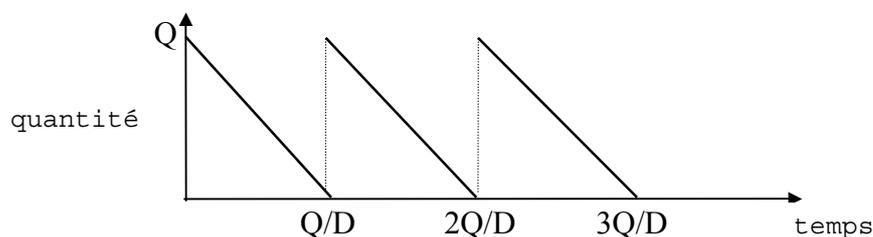
9.5. Quelques modèles de gestion des stocks.

9.5.1. Quantité économique de commande.

Les hypothèses de ce modèle sont les suivantes:

1. il s'agit de gérer *un seul* article, approvisionné extérieurement; toute quantité commandée est livrée en une seule fois, et le *délai d'obtention* est nul (la commande est reçue instantanément);
2. la *demande* est *déterministe* (connue avec certitude), et de *taux constant*: on note D le nombre d'unités consommées par unité de temps;
3. le *coût de possession* s'élève à c_p unités monétaires par unité de l'article et par unité de temps;
4. le *coût de commande* est de c_L unités monétaires par commande;
5. la demande doit être satisfaite au moment où elle est exprimée: les *ruptures de stock* ne sont pas autorisées.

Remarquons que les paramètres D , c_p et c_L sont supposés *stationnaires* (c'est-à-dire constants au cours du temps) et connus avec certitude. Il est très facile de vérifier que, dans ces circonstances, la politique optimale consiste à réapprovisionner une quantité constante Q lorsque le stock s'annule. En effet, il n'y a aucun intérêt à passer commande avant épuisement des stocks, puisque la demande est connue et le délai d'obtention est nul. Par ailleurs, la quantité de commande optimale est nécessairement stationnaire. La position de stock suit donc l'évolution temporelle représentée par le graphique suivant:



Cette politique d'approvisionnement peut être indifféremment interprétée comme une politique (s,Q) avec $s = 0$ ou comme une politique (R,S) avec $R = Q/D$ et $S = Q$. Reste maintenant à déterminer la valeur de Q qui minimise les coûts totaux entraînés par cette politique.

A combien se montent les coûts encourus par unité de temps (disons par année, pour fixer les idées)?

- Coûts de commande: la demande totale s'élève à D unités par année
 \implies il faut passer D/Q commandes par année
 \implies coûts de commande = $c_L D/Q$ par année.
- Coûts de possession: le stock moyen s'élève à $Q/2$ unités
 \implies coûts de possession = $c_p Q/2$.

Donc, le coût total par unité de temps vaut $c_L D/Q + c_p Q/2$.

La valeur de Q qui minimise la fonction de coût est obtenue en annulant la dérivée de cette fonction par rapport à Q. On obtient ainsi le minimum global $Q^* = \sqrt{2c_L D / c_p}$. La quantité Q^* est appelée *quantité économique de commande* (QEC) ou Economic Order Quantity (EOQ). (La formule définissant Q^* est due à Harris, 1913; elle est parfois également attribuée à Wilson, Camp, etc; voir Erlenkotter 1990 pour un historique de cette formule).

L'intervalle de temps écoulé entre deux réapprovisionnements, ou *période économique entre commandes*, est donné par :

$$PEC = Q^*/D = \sqrt{2c_L / Dc_p} .$$

Remarque: Comme on l'a vu au Chapitre 8, la quantité économique de commande et la période économique entre commandes sont souvent utilisées dans le cadre des méthodes MRP, lorsque les groupements de lots sont déterminés par la méthode de la « quantité fixe » ou du « nombre de périodes fixe ». Le choix de la quantité fixe ou de la période fixe est alors fréquemment basé sur le calcul de QEC et de PEC (où une unité de temps = une période MRP). Il est important de comprendre que ceci est une pratique empirique sans aucune justification théorique puisque la demande fortement dépendante, caractéristique des applications MRP, ne satisfait en général nullement aux hypothèses du modèle EOQ.

9.5.2. Délai d'obtention.

Lorsque le délai d'obtention des approvisionnements n'est pas nul, mais est égal à $L \geq 0$, il est à nouveau aisé de se convaincre qu'une politique de type (s,Q*) reste optimale. Le point de commande s doit permettre de satisfaire exactement la demande durant le délai d'obtention, c'est-à-dire $s = D \times L$. Par ailleurs, la fonction de coût est identique à celle calculée dans le cas précédent, et donc $Q^* = \sqrt{2c_L D / c_p}$.

9.5.3. Demande aléatoire.

Une généralisation supplémentaire concerne le cas où la demande est aléatoire, le délai d'obtention étant toujours considéré comme certain. Dans ce cas, il n'est bien sûr plus possible d'exclure totalement l'occurrence de ruptures de stocks. Une modélisation complète de cette situation devrait donc également spécifier le coût associé à ces ruptures et le critère économique à optimiser (par exemple, l'espérance mathématique des coûts sur l'horizon considéré). Ce type d'approche présente cependant plusieurs obstacles. D'une part, les coûts de rupture sont souvent difficilement quantifiables. D'autre part, les modèles d'optimisation ainsi obtenus sont très complexes et ne peuvent pas toujours être résolus de façon exacte. En pratique, l'approche optimisante sera donc réservée à des articles particulièrement importants (de classe A). Pour des articles de classe B, par contre, il est courant d'adopter une approche plus simple basée sur le maintien d'un *niveau de service* prédéterminé.

La notion de niveau de service peut être formalisée de plusieurs façons non-équivalentes: en termes de probabilité de rupture, de nombre attendu de ruptures, de temps écoulé entre ruptures successives, etc. Nous retiendrons ici une des formalisations les plus simples et définirons le niveau de service comme *la probabilité qu'aucune rupture de stock ne se produise entre deux approvisionnements successifs*.

En pratique, le gestionnaire des stocks devra fixer le niveau de service qu'il désire obtenir pour chaque article (ou famille d'articles) particulier, dans le cadre des lignes directrices adoptées par la firme. Ainsi, un niveau de service de 99% pourrait être exigé pour les produits finis si la stratégie concurrentielle de la firme repose sur une haute qualité de service, alors qu'un niveau de 90% pourrait être jugé acceptable par d'autres firmes. Insistons sur le fait que, dans cette approche, le niveau de service est le reflet immédiat des préoccupations stratégiques de la firme et ne doit donc pas être laissé à l'appréciation d'un employé isolé. (La même observation vaut pour l'évaluation des coûts de rupture dans les modèles qui utilisent explicitement ces coûts).

Afin d'illustrer cette discussion, considérons le modèle suivant, qui fournit une généralisation stochastique assez simple du modèle EOQ:

1. il s'agit de gérer *un seul* article, approvisionné extérieurement; toute quantité commandée est *livrée en une seule fois*, et le *délai d'obtention* vaut $L \geq 0$;
2. la *demande* (c'est-à-dire, le nombre d'unités consommées) par unité de temps est une variable aléatoire de moyenne D et d'écart-type σ ;
3. le *coût de possession* s'élève à c_p unités monétaires par unité de l'article et par unité de temps;
4. le *coût de commande* est de c_L unités monétaires par commande;
5. le *niveau de service* ciblé vaut p ($0 \leq p \leq 1$).

L'objectif poursuivi est de *minimiser l'espérance des coûts totaux* (par unité de temps) sous la contrainte imposée par le niveau de service. (Par rapport au modèle EOQ déterministe de base, nous avons généralisé les hypothèses 1, 2 et 5. Plus précisément, le modèle déterministe peut être vu comme le cas limite du modèle stochastique obtenu lorsque $L = 0$, $\sigma \rightarrow 0$ et $p = 1$.)

Le problème stochastique est difficile à résoudre exactement. On se contente donc généralement d'une approche heuristique du type suivant (voir Eppen, Gould and Schmidt 1993 ou Taha 1992).

D'abord, par analogie avec le modèle EOQ, on adopte une politique (s, Q^*) . Sous cette restriction, il est facile de se convaincre que la valeur sélectionnée pour la quantité de commande ne joue pas un rôle déterminant pour le respect du niveau de service. En effet, les ruptures de stock ne peuvent survenir que durant le délai d'obtention, et celui-ci est influencé par s , mais pas par Q^* . Sur base de cette observation, on choisit la quantité de commande comme si la demande était certaine et égale à D , c'est-à-dire qu'on pose $Q^* = \sqrt{2c_L D / c_p}$.

Reste alors à déterminer la valeur du point de commande s : on la prendra juste assez grande pour atteindre le niveau de service requis (bien sûr, toute valeur supérieure permettrait également d'atteindre ce niveau de service, mais au prix d'une augmentation des coûts de possession). Puisque les ruptures de stock peuvent seulement survenir durant le délai d'obtention, on va donc choisir s de telle sorte que

$$\text{Prob}(\text{demande durant le délai d'obtention} \leq s) = p.$$

Pour illustrer ce principe de façon plus précise, supposons par exemple que la demande suive une loi normale de moyenne D et d'écart-type σ , et que la demande dans n'importe quel intervalle de temps soit indépendante de la demande dans tout autre intervalle de temps disjoint. Sous ces hypothèses, la variable aléatoire $X = \ll \text{demande durant le délai d'obtention} \gg$ suit une loi normale de moyenne $D \times L$ et d'écart-type $\sigma \sqrt{L}$. Le point de commande peut donc facilement être déterminé par référence aux tables de la loi normale. Plus précisément, considérons la variable aléatoire $U = (X - DL) / \sigma \sqrt{L}$. Cette variable suit une loi normale standardisée $N(0,1)$. Par référence aux tables de la loi $N(0,1)$, on peut donc trouver un nombre k tel que $\text{Prob}(U \leq k) = p$. On a alors:

$$\begin{aligned} \text{Prob}(\text{demande durant le délai d'obtention} \leq s) &= p \\ \Leftrightarrow \text{Prob}(X \leq s) &= p \\ \Leftrightarrow \text{Prob}(U \leq (s - DL) / \sigma \sqrt{L}) &= p \\ \Leftrightarrow (s - DL) / \sigma \sqrt{L} &= k. \end{aligned}$$

Le point de commande s doit donc être fixé à la valeur $s = DL + k\sigma\sqrt{L}$. Cette valeur admet une interprétation intuitive. Le terme DL peut être vu comme le niveau de stock permettant d'attendre le réapprovisionnement suivant sans encourir de rupture dans des conditions de demande moyennes (de fait, comme on l'a vu plus haut, DL est le point de commande en régime déterministe). Le terme $k\sigma\sqrt{L}$, quant à lui, peut être interprété comme un *stock de sécurité* déterminé d'une part par l'incertitude liée au caractère aléatoire de la demande pendant le délai d'obtention (incertitude mesurée par l'écart-type $\sigma\sqrt{L}$), et d'autre part par le facteur k (qui ne dépend que du niveau de service p , et pas des paramètres D , σ de la distribution de demande). On appelle parfois k le *niveau de sécurité*.

9.5.4. Modèle du marchand de journaux.

Les modèles précédents étaient pertinents dans le cas d'articles approvisionnés de façon répétitive. Considérons maintenant le cas d'un article à *rotation nulle*, pour lequel le stock n'est approvisionné qu'une seule fois. Les unités invendues à la fin de l'horizon pourront être écoulées à prix réduit ou seront considérées comme perdues. On pensera, par exemple, au cas des articles de mode ou périssables. La littérature scientifique illustre habituellement le modèle mathématique correspondant à travers le dilemme d'un marchand de journaux.

Exemple: Un libraire s'approvisionne quotidiennement en journaux auprès d'une maison d'édition. Il paie v francs pour chaque exemplaire du journal, et le revend pour p francs. Les invendus lui sont repris à g francs par l'éditeur. On supposera $g < v < p$. La demande x du quotidien considéré suit une loi de probabilité continue de densité $f(x)$ ($x \geq 0$). Combien de journaux le libraire doit-il commander chaque jour afin de maximiser son bénéfice espéré ? (Le critère du bénéfice espéré semble raisonnable ici, puisque l'on a affaire à une situation très répétitive.)

Si une commande compte Q unités et que la demande se monte à X unités, le bénéfice peut être calculé comme suit:

$$\begin{aligned} b(Q,X) &= pQ - vQ && \text{si } Q \leq X \\ &= pX + g(Q-X) - vQ && \text{si } Q \geq X. \end{aligned}$$

Pour une commande de Q unités, l'espérance du bénéfice vaut donc :

$$\begin{aligned} b(Q) &= E_X(b(Q,X)) \\ &= (p-v) Q \int_0^{\infty} f(x) dx + (g-v) Q \int_0^Q f(x) dx + (p-g) \int_0^Q x f(x) dx \\ &= pQ + (gQ-pQ) \int_0^Q f(x) dx - vQ + (p-g) \int_0^Q x f(x) dx. \end{aligned}$$

La dérivée de $b(Q)$ par rapport à Q vaut (en utilisant simplement la définition d'une primitive):

$$\begin{aligned} b'(Q) &= p - v + (g-p) \int_0^Q f(x) dx + (gQ-pQ) f(Q) + (p-g)Q f(Q) \\ &= p - v + (g-p) \text{Prob}(X \leq Q). \end{aligned}$$

La résolution de l'équation $b'(Q) = 0$ indique que la quantité optimale de commande Q^* est celle pour laquelle :

$$\text{Prob}(X \leq Q^*) = \frac{p - v}{p - g}.$$

Ce résultat peut aussi être interprété plus intuitivement. Soit:

$C_0 = p - v =$ bénéfice unitaire par vente, ou encore coût d'opportunité associé à une vente ratée.
 $C_p = v - g =$ coût de possession associé à une unité stockée mais non vendue.

On a alors :

$$\text{Prob}(X \leq Q^*) = \frac{C_0}{C_0 + C_p} \quad \text{et} \quad \text{Prob}(X \geq Q^*) = \frac{C_p}{C_0 + C_p}.$$

Le marchand accepte donc une probabilité de stocker en excédent ($X \leq Q^*$) directement proportionnelle au bénéfice qu'il retire d'une vente réalisée et une probabilité de rupture de stock ($X \geq Q^*$) proportionnelle au coût d'une unité acquise mais non vendue. Ce résultat peut aussi être déduit d'une analyse marginale de la décision de commande (voir Silver et Peterson 1985, pp. 399-400).

Bibliographie.

C. De Bruyn, *Méthodes Quantitatives de Gestion 1*, FEGSS, 1991.

G.D. Eppen, F.J. Gould and C.P. Schmidt, *Introductory Management Science*, Prentice Hall, Englewood Cliffs, 1993.

D. Erlenkotter, Ford Whitman Harris and the Economic Order Quantity Model, *Operations Research* Vol. 38, 1990, pp. 937-950.

V. Giard, *Gestion de la Production*, 2ème édition, Economica, Paris, 1988.

S.C. Graves, A.H.G. Rinnooy Kan et P.H. Zipkin (eds.), *Logistics of Production and Inventory*, North-Holland, Amsterdam, 1993.

E. Silver et R. Peterson, *Decision Systems for Inventory Management and Production Planning*, J. Wiley & Sons, New York, 1985.

H.A. Taha, *Operations Research – An Introduction*, McMillan Publishing Company, New York, 1992.

T.E. Vollmann, W.L. Berry et D.C. Whybark, *Manufacturing Planning and Control Systems*, Business One Irwin, Homewood, Illinois, 1992.

CHAPITRE 10

ORDONNANCEMENT ET PILOTAGE DES ATELIERS

10.1. Généralités.

Nous abordons dans ce chapitre le domaine de la planification à *très court terme* et celui du contrôle en *temps réel* des opérations de production.

On supposera que, d'une façon ou d'une autre, un plan de production établi sur base de sous-périodes relativement courtes (typiquement, hebdomadaires) est disponible (ceci pourrait être le plan directeur de production, ou un plan MRP y associé). Au début de chaque sous-période, le service production se trouve donc confronté à une collection d'*ordres de fabrication* à exécuter durant cette sous-période. Ces ordres déterminent des listes d'*opérations* à effectuer, mais ne précisent pas nécessairement l'ordre dans lequel ces opérations doivent être exécutées (faut-il intercaler la commande C entre les commandes A et B ? faut-il peindre les pièces bleues avant les rouges ?), ni le moment où elles doivent être entamées (lundi matin ? jeudi après-midi ?), ni les centres de production auxquels elles doivent être affectées (faut-il imprimer ces dépliants sur la vieille rotative, ou vaut-il mieux lui réserver une machine plus rapide ?). En particulier, puisque les contraintes de capacité n'ont été exprimées que de manière approchée jusqu'à ce stade de planification, le plan de production considéré n'est même pas nécessairement réalisable (il pourrait excéder la capacité disponible).

Dans ce cadre, on cherchera donc plus particulièrement à résoudre la problématique de l'*ordonnancement* des opérations, c'est-à-dire de la planification à très court terme:

- de l'affectation des opérations aux postes de production, et
- des dates de lancement et de fin des opérations.

Remarquons, au passage, que, contrairement à ce que son appellation pourrait suggérer, l'activité d'ordonnancement va plus loin que la simple détermination de l'*ordre* d'exécution des opérations sur un centre de production (« sequencing » en anglais). Le terme anglais pour ordonnancement est « scheduling », et suggère, de façon plus adéquate, la construction d'un horaire de production.

Le dernier niveau de décision concernant la mise en route de la production est celui du *contrôle*, ou du *pilotage en temps réel* des opérations. Dans un monde parfaitement déterministe, parfaitement connu, et parfaitement pris en compte par la planification, le contrôle d'un atelier pourrait se limiter à la mise en application des décisions d'ordonnancement prises auparavant. En pratique, ceci n'est (presque) jamais le cas:

- certains éléments ont un caractère aléatoire, non prévisible avec certitude (pannes, temps de réglage, retards de livraison, maladie,);
- même si certains événements sont en réalité prévisibles, ils ne sont pas nécessairement connus par le responsable de la planification;
- tous les éléments déterministes et connus ne sont pas nécessairement pris en compte au moment de l'ordonnancement, pour des raisons de simplicité.

Pour toutes ces raisons, on considère souvent que l'ordonnancement énonce seulement des *propositions de lancement*, qui sont acceptées telles quelles ou modifiées par les décisions de contrôle. Au fil des jours, la réalité de l'atelier s'éloigne ainsi progressivement des plans établis par les responsables de l'ordonnancement. Idéalement, on voudrait donc pouvoir recalculer un nouvel ordonnancement le plus fréquemment possible, par exemple tous les jours, en tenant compte des nouvelles « conditions initiales » dictées par l'état de l'atelier (par exemple, opérations en cours, machines en panne, etc). En pratique, l'utilisation d'un tel *ordonnancement glissant* est rendue difficile par les temps de calcul élevés que requièrent les logiciels d'ordonnancement. On se rabat donc souvent sur des règles de dispatching simples pour le pilotage de l'atelier.

10.2. Quelques concepts utilisés en ordonnancement.

Capacité finie et infinie.

Dans le cadre de l'établissement du plan directeur de production ou du plan MRP, on a travaillé sans se préoccuper de contraintes de capacité: ces plans ont été établis à *capacité infinie*, et leur faisabilité a simplement été contrôlée a posteriori et de façon approximative (rough-cut capacity planning et planification des besoins en capacité; voir Section 8.6).

On ne peut bien sûr plus se contenter d'une telle analyse approximative pour l'établissement d'un plan d'ordonnancement. Il faudra donc à présent tenir compte explicitement de la capacité de chaque poste de travail dans ce plan. Pour cette raison, l'activité d'ordonnancement est parfois appelée *planification à capacité finie*.

Graphiques de Gantt.

Le résultat des analyses à capacité infinie est décrit graphiquement par une courbe, ou profil, de charge pour chaque poste de travail. Ces graphiques permettent de comparer facilement la capacité disponible à la capacité requise.

En ordonnancement, il est utile de disposer d'une représentation graphique plus précise des plans établis. Ce but est atteint par l'utilisation des *graphiques de Gantt*: pour chaque poste de travail, un tel graphique représente simplement, sur un axe temporel, les dates de début et de fin planifiées pour chacune des opérations à effectuer sur ce poste. Plusieurs postes de travail peuvent être pris en compte dans un seul graphique de Gantt.

En plus des dates planifiées, il est également courant de porter sur un graphique de Gantt les dates *réalisées* de début et de fin de chaque opération. Ceci permet une comparaison des plans et des performances effectivement accomplies et fournit donc un outil d'aide au contrôle de la production.

10.3. Description des ordres de fabrication et du système de production.

Le plan d'ordonnancement est établi sur base d'une description des ordres de fabrication (ou OFs, fréquemment appelés *tâches* dans la littérature scientifique) d'une part, et du système de production d'autre part. Ces informations constituent donc la base de données utilisée par tout logiciel d'ordonnancement.

Chaque OF est vu comme un ensemble d'*opérations* à effectuer, où une opération est une activité élémentaire, décrite par des spécifications techniques, l'identification du type de machine et d'outil à utiliser, etc.

Exemple: L'OF U580512 ('usinage d'une pièce de moteur') requiert les opérations suivantes:

1. fraisage sur le poste F12 (temps opératoire: 25 minutes);
2. tournage sur un des postes T01 ou T02 (temps opératoire: 15 min sur T01, 10 min sur T02);
3. fraisage sur le poste F3 (temps opératoire: 10 min);
4. ébarbage (temps opératoire: 10 min);
5. contrôle de qualité final (2 min).

Les informations qui constituent une description complète de chaque OF incluent donc, typiquement:

- la *gamme opératoire*, c'est-à-dire la séquence d'opérations à effectuer et les postes de travail requis, les temps opératoires estimés, le délai de fabrication estimé (qui inclut lui-même les temps opératoires, plus les temps d'attente, de transport et de lancement; voir Section 8.2), et les ressources nécessaires (qualification des ouvriers, outils, matières, ...); à noter que des gammes opératoires alternatives existent parfois (voir l'opération 2 'tournage' dans l'exemple précédent);
- des données non techniques: *date de début au plus tôt* de l'OF (par exemple, mise à disponibilité des matières, date d'achèvement d'une opération entamée sur un poste précédent, ...), *date d'exigibilité* ou *de livraison* de l'OF (fixée par exemple à la passation de commande ou par le plan MRP), *coefficient de priorité* de l'OF (lié à l'importance de la commande, ou au fait que la fabrication est pour stock ou sur commande).

Les informations relatives au système productif ou aux ressources incluent :

- l'identification des postes de travail et une description de leurs caractéristiques techniques (par exemple, vitesse, taux de rebuts, présence ou non d'aire de stockages, ...);
- les caractéristiques du personnel de production: nombre, qualifications, ...
- autres: AGVs, outils, palettes, ...
- disponibilités: horaire normal, maintenances planifiées, pannes, ...

10.4. Critères de qualité d'un ordonnancement.

Les critères utilisés pour juger de la qualité d'un plan d'ordonnancement peuvent être nombreux et parfois même contradictoires. Les critères purement économiques (coût, profit) sont mal adaptés dans ce cadre du très court terme. On utilise donc plutôt des critères d'efficacité organisationnelle: respect des délais, quantité d'en cours, utilisation des postes de travail, ...

Ces critères peuvent être précisés et formulés quantitativement de différentes manières. Pour ce faire, introduisons d'abord quelques notations. Soit un ensemble de n OFs (notés $i = 1, 2, \dots, n$) à ordonnancer. Pour chaque OF, on connaît en particulier:

- r_i : la date de début au plus tôt de l'OF i ,
- d_i : la date d'exigibilité de l'OF i .

De plus, tout plan d'ordonnement détermine (par définition)

- C_i : la date d'achèvement de l'OF i ,
- et donc, par simple déduction,
- $T_i = \max(0, C_i - d_i)$: le retard de l'OF i (le retard est nul si l'OF est fini avant d_i),
 - $F_i = C_i - r_i$: le temps de passage de l'OF i dans le système (flow time).

Quelques-uns des critères de performance couramment utilisés en ordonnancement peuvent à présent être définis en termes des quantités introduites ci-dessus:

- U : nombre d'OFs en retard, c'est-à-dire pour lesquelles $C_i > d_i$;
- T : retard moyen : $T = (1/n) \sum_{1 \leq i \leq n} \max(0, C_i - d_i)$;
- T_{\max} : plus grand retard = $\max_{1 \leq i \leq n} \max(0, C_i - d_i)$;
- C_{\max} : temps d'achèvement du dernier OF = $\max_{1 \leq i \leq n} C_i$ (makespan);
- F : temps de passage moyen des OFs : $F = (1/n) \sum_{1 \leq i \leq n} (C_i - r_i)$.

Tous ces critères connaissent également une version pondérée permettant de prendre en compte les divers degrés de priorité des OFs.

Chacun des critères U , T , T_{\max} donne une mesure du respect des délais, mais ces mesures ne sont pas nécessairement concordantes. De même, C_{\max} et F mesurent de façon différente la vitesse à laquelle les OFs sont évacués du système. En particulier, F peut être vu comme une mesure indirecte de la quantité d'en cours.

Exemple. Considérons 2 OFs à exécuter sur une unique machine. Chaque OF comporte une seule opération et est caractérisée par les paramètres suivants:

OF (opération)	1	2

date de disponibilité r_i		0 0
date d'exigibilité d_i	11	2
temps opératoire p_i	2	10

Deux ordonnancements sont possibles: (1,2) et (2,1). Pour le premier ordonnancement, on obtient

OF	1	2
en retard	non	oui
T_i	0	10
$C_i - r_i$	2	12

et donc: $U = 1$, $T = 5$, $T_{\max} = 10$, $C_{\max} = 12$, $F = 7$.

De même, pour le second ordonnancement,

OF	1	2
en retard	oui	oui
T_i	1	8
$C_i - r_i$	12	10

et donc: $U = 2$, $T = 9/2$, $T_{\max} = 8$, $C_{\max} = 12$, $F = 11$.

On constate que le premier ordonnancement est meilleur que le second sur les critères U et F , mais qu'il est moins bon sur les critères T et T_{\max} .

Les logiciels d'ordonnancement génèrent souvent plusieurs propositions d'ordonnancement pour une situation donnée et fournissent une évaluation de ces propositions selon différents critères. Ceci permet une sélection manuelle du plan jugé globalement le meilleur. L'importance relative accordée aux différents critères doit alors tenir compte des priorités (tactiques ou stratégiques) de la firme: veut-on mettre l'accent sur le respect des délais plutôt que sur la réduction des encours ? peut-on renégocier l'une ou l'autre date de livraison ? etc.

Il faut également remarquer que le calcul d'un ordonnancement optimisant un critère déterminé est en général difficile et se situe bien souvent au-delà des possibilités actuelles des techniques mathématiques d'optimisation. En pratique, on aura donc recours à des heuristiques plutôt qu'à des méthodes d'optimisation exactes pour réaliser le calcul d'ordonnancement. En variant les heuristiques utilisées, on peut assez facilement soumettre plusieurs ordonnancements au décideur, comme évoqué ci-dessus.

10.5. Ordonnancement sur une machine unique.

Afin d'illustrer plus concrètement la discussion précédente, considérons le cas de plusieurs OFs à effectuer sur une seule machine, où chaque OF consiste en une seule opération et tous les OFs partagent une date au plus tôt commune. La considération de ce cas apparemment simpliste peut se justifier par le fait que les problèmes à plusieurs machines sont souvent attaqués, implicitement ou explicitement, par la résolution successive de problèmes à une machine. C'est par exemple le cas lorsque l'on utilise des règles de dispatching pour le pilotage en temps réel (voir plus bas). C'est également le cas pour des approches plus sophistiquées (comme celle

utilisée par le logiciel commercial OPT) qui développent un ordonnancement d'atelier en améliorant de façon itérative l'ordonnancement courant sur chacune des machines-goulots.

Au vu de nos hypothèses, on emploiera indifféremment dans cette section la terminologie « OF » ou « opération » et on supposera que $r_1 = r_2 = \dots = r_n = 0$. Par ailleurs, on notera

- p_i : temps opératoire de l'OF (opération) i .

Exemple : Considérons $n = 5$ OFs.

OFs :	1	2	3	4	5
Temps opératoires p_i :	5	15	8	20	3
Dates au plus tôt r_i :	0	0	0	0	0
Dates d'exigibilité d_i :	10	30	41	40	20.

Remarquons que, pour n'importe quel ordonnancement sans temps mort, C_{\max} est toujours égal à la somme des temps opératoires des différents OFs (ainsi, pour notre exemple, $C_{\max} = 5 + 15 + 8 + 20 + 3 = 51$). Ce critère présente donc peu d'intérêt pour le problème qui nous occupe.

Le critère du temps de passage moyen F est, lui, plus intéressant. On démontre rigoureusement que, pour les problèmes à une machine avec date au plus tôt commune, la règle du *Temps Opératoire Minimum* (TOM; Shortest Processing Time, SPT) permet de minimiser F ; cette règle prescrit simplement d'exécuter les OFs par temps opératoires croissants (sur un plan intuitif, le lecteur se convaincra facilement que, en commençant par les OFs les plus longs, on augmenterait considérablement le temps de passage de *tous* les OFs).

Exemple: Pour l'ordonnancement (1,2,3,4,5) (sélectionné arbitrairement), on calcule: $F = (5 + 20 + 28 + 48 + 51)/5 = 30,4$. La règle TOM fournit l'ordonnancement optimal (5,1,3,2,4), pour lequel $F = 21,8$.

Par ailleurs, on démontre que le critère du plus grand retard (T_{\max}) est minimisé par la règle de la *Date d'Exigibilité Minimum* (Earliest Due Date, EDD), qui consiste à ordonnancer les OFs par date d'exigibilité croissante.

Exemple: Pour l'ordonnancement (5,1,3,2,4) (qui minimisait F), les OFs 2 et 4 sont en retard et $T_{\max} = 11$ (retard encouru par l'OF 4). Par contre, en appliquant la règle EDD, on obtient l'ordonnancement (1,5,2,4,3), pour lequel les OFs 3 et 4 sont en retard et $T_{\max} = 10$.

Mentionnons encore que, même pour une seule machine, les problèmes d'ordonnancement peuvent être d'une grande complexité. Par exemple, décider si il existe un ordonnancement respectant toutes les dates d'exigibilité est un problème difficile (pour autant que les dates au plus tôt ne coïncident pas; sinon, EDD résout la question). Les problèmes à plusieurs machines sont, a fortiori, extrêmement difficile à résoudre de façon exacte. En particulier, dans le cas d'ateliers à plusieurs machines, aucune règle simple n'est connue pour minimiser F ou T_{\max} .

10.6. Pilotage des ateliers: Règles de dispatching.

Dans un job-shop, l'absence ou la détérioration progressive d'un plan d'ordonnancement sont souvent compensées par l'utilisation de règles de priorité simples, dites *règles de dispatching*. Ces règles permettent de décider du prochain OF à lancer en production sur *chaque* poste de travail donné en ne tenant compte que des caractéristiques des OFs disponibles à *ce poste* (et donc, en négligeant les interactions entre ces OFs et les autres OFs en cours dans l'atelier, ainsi que les interactions entre postes de travail). Le problème de pilotage est ainsi ramené, pour l'essentiel, à une série de problèmes d'ordonnancement à une machine, pour lesquels les règles de dispatching fournissent des approches heuristiques simples.

Quelques règles de dispatching souvent utilisées en pratique sont les suivantes: parmi les OFs disponibles au poste considéré, sélectionner

- le premier OF arrivé à ce poste (*premier arrivé premier servi*, FIFO), ou
- l'OF i pour lequel le temps opératoire de l'opération à réaliser est le plus faible (TOM), ou
- l'OF dont la date d'exigibilité d_i est la plus précoce (EDD), ou
- l'OF i dont le *ratio critique* CR_i est le plus faible, avec

$$CR_i = \frac{\text{temps restant jusqu' à } d_i}{\text{temps total requis pour terminer l' OF } i}$$

(en plus des temps opératoires proprement dits, le dénominateur de CR_i inclut les temps d'attente, de transferts, de lancements, ... encore requis par l'OF; un ratio CR_i inférieur à 1 n'indique donc pas nécessairement qu'un OF sera terminé en retard, dans la mesure où la plupart de ces temps sont estimés sur des bases approximatives et peuvent également être réduits en cas d'urgence).

Empiriquement, on constate que la règle TOM présente souvent un comportement très satisfaisant sur différents critères de performance, en dépit de sa simplicité et du fait qu'elle ne tient pas compte explicitement des dates d'exigibilité.

Remarquons encore que les règles de dispatching énoncées ci-dessus se caractérisent par des exigences en information très variables. Par exemple, TOM n'utilise que de l'information locale et relativement statique (temps opératoires sur la machine considérée). Par contre, les dates d'exigibilité constituent de l'information non locale et susceptible de varier au fil du temps (par exemple, lors des révisions périodiques du plan MRP). Il est donc plus difficile de tenir cette information à jour. Des observations similaires s'appliquent à la détermination du ratio critique.

Les règles de dispatching peuvent être utilisées par le management pour imposer un certain contrôle sur le pilotage de l'atelier (contrôle centralisé/décentralisé, standardisation par les procédures de travail; voir Boskma 1987 pour une discussion plus approfondie).

10.7. Ordonnement d'ateliers par les règles de dispatching.

Comme nous l'avons déjà mentionné plus haut, les règles de dispatching peuvent être utiles vues comme des heuristiques de résolution de problèmes d'ordonnement à une machine. En fait, vue de façon plus large, leur utilisation lors du pilotage des ateliers conduit aussi, implicitement, à la construction d'un ordonnement *pour tous les OFs en cours dans l'atelier* (le diagramme de Gantt de cet ordonnement peut facilement être reconstitué, a posteriori, sur base des dates de début et de fin *observées* – plutôt que planifiées – de chacune des opérations).

En prolongeant quelque peu ce raisonnement, et en le transposant dans un cadre de planification plutôt que de pilotage en temps réel, on parvient alors à la conclusion que les règles de dispatching peuvent également être utilisées pour définir des heuristiques d'ordonnement à plusieurs machines. L'idée est ici de *simuler* le passage des OFs à effectuer sur les différents postes de travail. A chaque instant et sur chaque poste, les décisions de lancement sont prises en application d'une (ou de plusieurs) règle(s) de dispatching prédéterminée(s) (voir l'exemple ci-dessous). En faisant varier les règles de dispatching utilisées, on génère facilement des plans alternatifs qui peuvent être évalués sur base des critères de performance jugés pertinents. (C'est ainsi que procèdent certains logiciels d'ordonnement.)

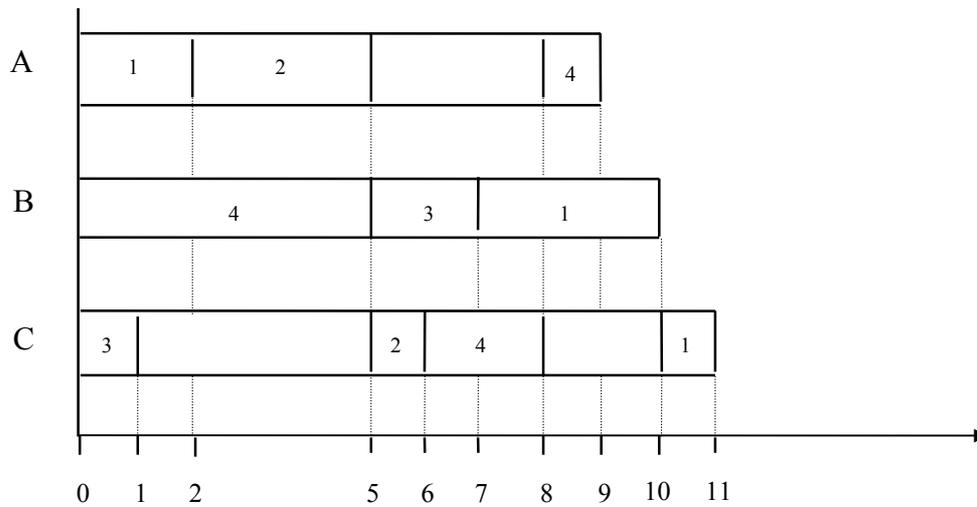
Exemple. Considérons trois machines (A, B, C) et 4 OFs numérotés (1,2,3,4). On supposera que $r_i = 0$ pour chaque OF. Les gammes opératoires des OFs et leurs dates d'exigibilité sont détaillées dans le tableau suivant:

OF	Gamme opératoire (et p_i)	d_i
1	A(2), B(3), C(1)	10
2	A(3), C(1)	5
3	C(1), B(2)	11
4	B(5), C(2), A(1)	9

Nous utiliserons (arbitrairement) la règle de dispatching TOM. La procédure d'ordonnement peut donc être décrite comme suit: *à chaque instant, examiner la file d'attente devant chaque machine et lancer l'OF disponible ayant le temps opératoire le plus faible sur cette machine.*

On obtient ainsi le graphique de Gantt suivant (où le symbole \square indique qu'une machine est inactive) :

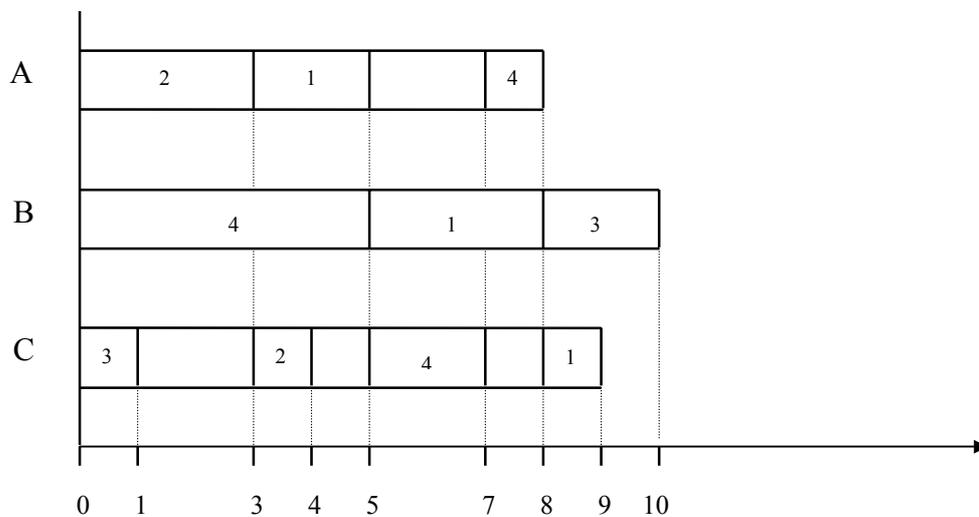
			File d'attente	
A	11222	4	1,2,4	
B	4444433111		4,3,1	
C	3	244	1	3,2,4,1



Remarquons par exemple que, après 5 unités de temps, les OFs 1 et 3 sont disponibles devant la machine B. Puisque les temps opératoires de ces deux OFs sur la machine B valent respectivement 3 et 2, la règle TOM donne priorité à l'OF 3. On obtient ainsi un ordonnancement où le dernier OF est terminé après 11 jours, et où les OFs 1 et 2 sont en retard d'un jour.

En appliquant la règle EDD plutôt que TOM à chaque machine, on aurait obtenu l'ordonnancement suivant:

	File d'attente	
A	22211 4	1,2,4
B	4444411133	4,3,1
C	3 2 44 1	3,2,4,1



A la période 5, contrairement à ce qui se passait avec la règle TOM, la règle EDD donne priorité à l'OF 1 sur l'OF 3. On obtient ainsi un ordonnancement où tous les OFs sont terminés après 10 jours et où aucun OF n'est en retard. (Comme on pouvait s'y attendre, la performance de EDD en termes de respect des délais est donc meilleure sur cet exemple que celle de TOM. Mais il ne s'agit pas là d'une règle absolue!)

Bibliographie.

K. Boskma, *Productie en Logistiek*, 2ème édition, Wolters-Noordhoff, Groningen, 1987.

V. Giard, *Gestion de la Production*, 2ème édition, Economica, Paris, 1988.

S.C. Graves, A.H.G. Rinnooy Kan et P.H. Zipkin (eds.), *Logistics of Production and Inventory*, North-Holland, Amsterdam, 1993.

J.O. McClain, L.J. Thomas et J.B. Mazzola, *Operations Management*, 3ème édition, Prentice-Hall, Englewood Cliffs, 1992.

M. Pinedo and X. Chao, *Operations Scheduling with Applications in Manufacturing and Services*, McGraw-Hill Irwin, 1998.

T.E. Vollmann, W.L. Berry et D.C. Whybark, *Manufacturing Planning and Control Systems*, Business One Irwin, Homewood, Illinois, 1992.

CHAPITRE 11

GESTION DE PROJETS

11.1. Généralités.

Un *projet* est un ensemble d'activités interdépendantes, mobilisant les ressources de l'entreprise sur une période de temps étendue et généralement exécuté une seule fois.

Les exemples concrets de projet sont légions:

- la construction de bâtiments ou de sites industriels (la liste des activités peut inclure, au-delà de la construction proprement dite, la conception des plans et des maquettes, la passation de contrats d'embauche, d'approvisionnement et de sous-traitance, etc);
- les projets techniques de grande envergure, comme les projets spatiaux;
- l'implantation d'un réseau informatique ou d'un système ERP dans une entreprise;
- le lancement d'un nouveau produit;
- la conception et la fabrication de produits uniques sur commande (production unitaire, engineering sur commande);
- la maintenance d'installations industrielles lourdes (nucléaire, sidérurgie), etc.

La production de nombreuses entreprises concerne exclusivement ou principalement la réalisation de projets. Des approches de gestion spécifiques à ce contexte particulier ont donc été développées au fil du temps. Au sein de la *gestion de projets*, on distingue généralement trois problématiques distinctes: la planification, l'ordonnancement et le contrôle.

1) *La planification* du projet comprend le recensement et la définition des activités élémentaires à exécuter, la clarification des relations d'antériorité à respecter entre ces activités, l'estimation de la durée des activités et des ressources requises pour leur exécution (main d'oeuvre, financement, ...). Il s'agit là d'une phase bien plus délicate qu'il n'y paraît (particulièrement lorsqu'un projet est le premier de son espèce, et implique plusieurs groupes d'acteurs pour la réalisation d'un très grand nombre d'activités). Sa bonne conduite est déterminante pour le succès ou l'échec d'un projet.

2) *L'ordonnancement* du projet consiste en l'élaboration d'un calendrier pour son exécution, c'est-à-dire la programmation de dates de début et de fin pour chaque activité. Comme d'habitude en ordonnancement, on y recherche un compromis entre plusieurs objectifs conflictuels: achèvement du projet dans les délais les plus courts, au moindre coût et en nivelant l'utilisation des ressources.

3) *Le contrôle* du projet a pour but de vérifier la correspondance entre les plans établis et leur exécution, et de prendre les recours éventuellement nécessaires en cas de déviation trop importante. Ceci implique des retours fréquents à la phase d'ordonnancement, afin de réactualiser les calendriers en fonction de l'évolution réelle du projet.

Nous nous concentrerons ici plus particulièrement sur la problématique de l'ordonnancement.

11.2. Les méthodes de chemins critiques.

Les principaux outils mathématiques utilisés pour la gestion de projets sont les méthodes PERT et CPM (PERT: Project Evaluation and Review Technique; CPM: Critical Path Method). Ces méthodes permettent de répondre à différentes questions (ordonnancement, probabilité de terminer dans les délais, ...) en *modélisant* le projet à l'aide d'un même graphe appelé *réseau d'activités*. Au vu de cette similitude, on identifie très souvent les deux méthodes sous l'appellation PERT/CPM, ou *méthodes de chemins critiques*. Ces méthodes constituent principalement des outils d'ordonnancement et de contrôle, mais contribuent également à systématiser la phase de planification.

Réseau d'activités.

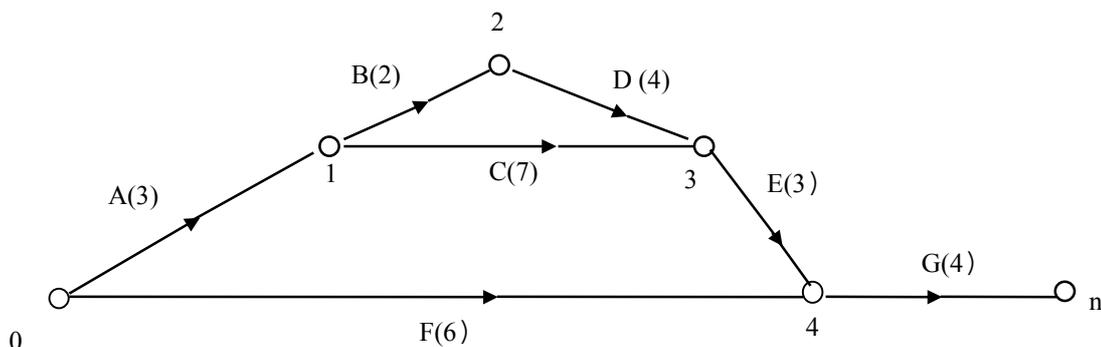
Soit A l'ensemble des activités à exécuter, sur lequel on a défini un certain nombre de contraintes d'antériorité auxquelles tout ordonnancement devra satisfaire. Un réseau d'activités pour A est un graphe G tel que:

- l'ensemble X des sommets de G (appelés *événements*) comporte deux sommets particuliers 0 (début) et n (fin): 0 est le seul sommet où n'entre aucun arc et n est le seul sommet d'où n'émane aucun arc; les autres sommets sont interprétés comme des jalons du projet, représentant l'achèvement ou le lancement de certains groupes d'activités;
- chaque activité est représentée dans G par un arc (i,j) de longueur d_{ij} , où d_{ij} est la *durée* (ou longueur) de l'activité; nous verrons dans un instant que G peut également comporter certains arcs qui ne sont associés à aucune activité réelle; ces arcs fictifs ont alors longueur 0;
- G ne contient pas de circuits (c'est-à-dire de chemins fermés);
- l'activité (i,j) doit absolument précéder l'activité (r,s) dans l'exécution du projet si et seulement si il existe un chemin de G empruntant successivement l'arc (i,j) puis l'arc (r,s) .

Exemple 1. Considérons 7 activités dont la durée (en jours) est indiquée entre parenthèses: A(3), B(2), C(7), D(4), E(3), F(6), G(4). Supposons que ces activités sont soumises aux contraintes d'antériorité suivantes:

- A précède B et C;
- C et D précèdent E;
- B précède D;
- E et F précèdent G.

Un réseau d'activités associé à ce projet est représenté ci-dessous.



En pratique, énumérer toutes les activités et toutes les relations d'antériorité constitue un défi pour le gestionnaire du projet, mais également un travail fructueux en soi. Remarquons en particulier que, dans un projet réel, la notion d'activité n'est pas définie de manière absolue: il existe généralement plusieurs façons de découper un projet en activités élémentaires. La construction du réseau oblige le gestionnaire à s'interroger systématiquement sur la nature des activités à accomplir ainsi que sur leurs relations mutuelles. Ceci justifie la remarque formulée plus haut, selon laquelle les méthodes de chemins critiques fournissent une aide non négligeable dans la phase de planification (en général, *établir* un modèle est au moins aussi utile et instructif que le *résoudre*).

Revenons à présent sur la définition du réseau d'activités pour y ajouter quelques précisions. Tout d'abord, remarquons que chaque sommet peut être interprété comme un jalon marquant l'achèvement de toutes les activités qui 'entrent' dans ce sommet et, par voie de conséquence, la possibilité de lancer chacune des activités qui en émane.

En ce qui concerne la liste des relations d'antériorité, il est d'usage de ne pas reprendre dans cette liste les contraintes impliquées par transitivité. Ainsi, dans l'Exemple 1 ci-dessus, la relation 'A précède D' ne figure pas explicitement dans la liste, mais peut être déduite des relations 'A précède B' et 'B précède D'.

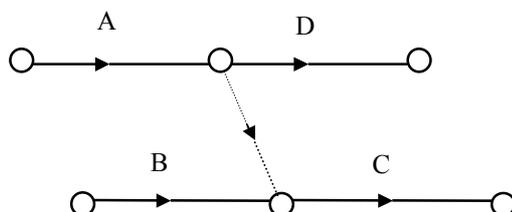
Notons encore que, afin de représenter fidèlement les relations d'antériorité, il est parfois nécessaire d'introduire dans le réseau des activités *fictives* ('dummies') qui ne correspondent donc à aucune activité réelle et dont la durée est par définition égale à zéro. Cette situation se présente notamment lorsque deux activités, disons A et B, doivent être suivies par des ensembles d'activités non disjoints, mais non identiques.

Exemple 2 (construction d'un immeuble).

A: pose des fenêtres; B: rejointoyage des murs; C: peintures extérieures; D: peintures intérieures.

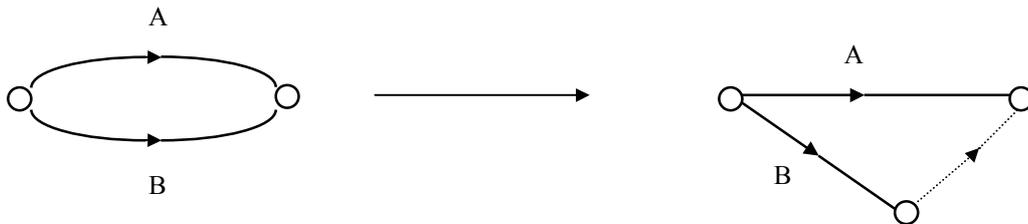
Contraintes: A doit précéder C et D; B doit précéder C.

La démarche 'naturelle', sans introduction d'arcs fictifs, conduit à un réseau où B est forcé de précéder D, alors que cette contrainte n'a pas été posée explicitement. Par contre, en introduisant un arc fictif, on obtient le réseau suivant.



Une autre situation dans laquelle les arcs fictifs sont parfois utilisés se présente lorsque deux activités ont exactement les mêmes prédécesseurs et successeurs, et pourraient donc être

représentées par deux arcs parallèles (c'est-à-dire deux arcs partageant les mêmes extrémités). Afin de mieux distinguer ces activités, il est alors coutumier d'insérer un arc fictif comme indiqué ci-dessous. (Notons que cette convention n'est pas strictement nécessaire pour assurer le traitement mathématique du problème d'ordonnancement. Elle se révèle cependant très pratique dans les implantations informatiques où chaque activité est enregistrée comme une paire de sommets de la forme (i,j)).



Chemins critiques.

Nous appellerons *longueur d'un chemin* de G la somme des durées des arcs situés sur ce chemin.

Supposons maintenant que le projet commence à la date 0.

Observation: la longueur d'un plus long chemin entre le sommet 0 et l'événement i est une borne inférieure sur la date de début de toute activité (i,j) .

En particulier, la longueur d'un plus long chemin (appelé *chemin critique*) entre 0 et n est une borne inférieure sur le temps nécessaire pour exécuter le projet. Il est peut-être moins évident, mais néanmoins vrai, que la longueur d'un chemin critique correspond *exactement* à la *durée minimale nécessaire* pour exécuter le projet. Cette observation fondamentale est à la base de toutes les méthodes de chemins critiques.

Exemple 1 (suite). Le seul chemin critique est A,C,E,G. Sa longueur vaut 17. Il est clair qu'aucune des activités situées sur ce chemin ne peut subir de retard si le projet doit être terminé en 17 jours. Par contre, F pourrait subir un retard de 7 jours sans retarder le projet.

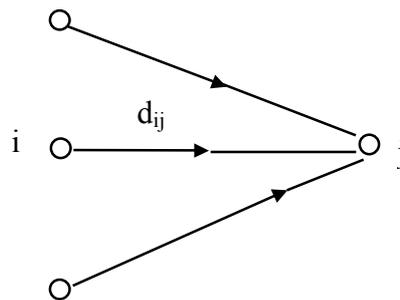
Calcul du chemin critique et des marges.

Le plus long chemin de 0 à n dans G correspond au plus court chemin dans un graphe modifié où l'on aurait remplacé chaque durée d par $-d$. Cette observation implique que l'on peut utiliser un algorithme de plus court chemin (tels ceux étudiés dans le cadre des cours de mathématiques) pour calculer un chemin critique de G . En fait, parce que G ne contient pas de circuits, l'algorithme prend ici une forme particulièrement simple.

Phase 1. Calculer pour chaque événement i (c'est-à-dire chaque sommet de G) une *date au plus tôt* E_i , qui est la date la plus précoce à laquelle l'événement i puisse se réaliser:

- $E_0 = 0$;
- pour chaque événement j autre que 0, si l'on a calculé la date E_i pour tous les événements i qui précèdent j dans G , alors

$$E_j = \max \{ E_i + d_{ij} : (i,j) \text{ est un arc de } G \}.$$



Cet algorithme se justifie aisément de façon intuitive: si l'on ne peut pas atteindre i avant la date E_i (quelque soit i précédant j), alors on ne peut certainement pas atteindre j avant la date $E_i + d_{ij}$.

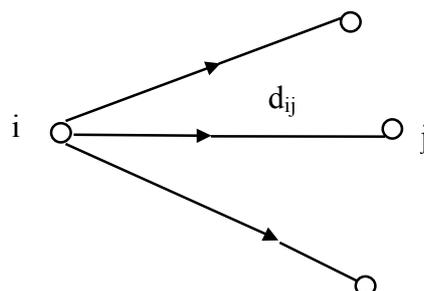
La date E_n est exactement la longueur d'un chemin critique. Dans la suite de notre discussion, nous supposons que le gestionnaire du projet vise à terminer le projet au plus vite, et donc à la date E_n .

Après avoir calculé E_n comme dans la phase 1, il n'est pas difficile de reconstruire le chemin critique lui-même sur base des calculs effectués. L'approche suivante permet de résoudre ce problème et fournit simultanément d'autres informations utiles pour la gestion du projet.

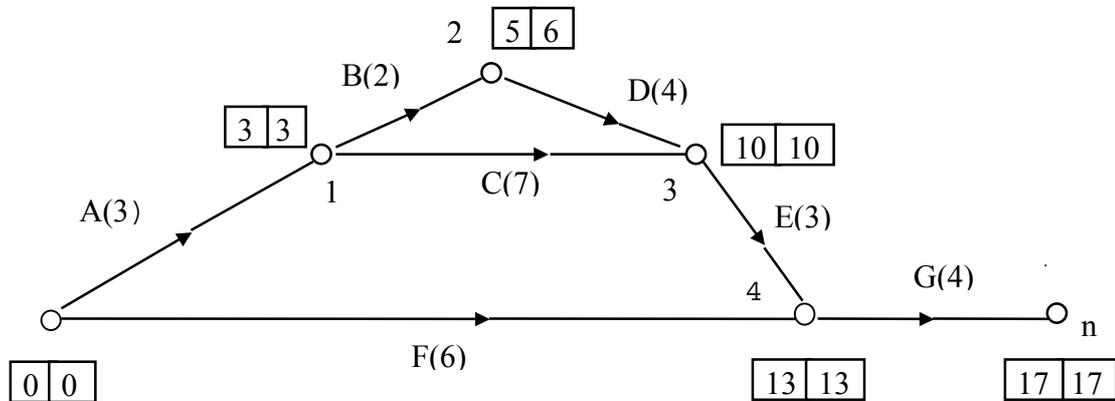
Phase 2. Calculer pour chaque événement i une *date au plus tard* L_i , qui est la date la plus tardive à laquelle l'événement i doit être réalisé pour permettre l'achèvement du projet avant E_n :

- $L_n = E_n$;
- pour chaque événement i autre que n , si l'on a calculé la date L_j pour tous les événements j qui suivent i dans G , alors

$$L_i = \min \{ L_j - d_{ij} : (i,j) \text{ est un arc de } G \}.$$



Exemple 1 (suite). Calculons les dates au plus tôt et au plus tard de ce projet.



La connaissance des dates au plus tôt et au plus tard permet à son tour le calcul des marges, c'est-à-dire de la flexibilité dont dispose le gestionnaire dans l'établissement d'un calendrier relatif à chacune des activités du projet. Différents types de marges sont à considérer et présentent différentes mesures de cette flexibilité.

L'observation fondamentale qui découle directement de la définition des dates au plus tôt et au plus tard est la suivante: une activité (i,j) peut être exécutée n'importe quand dans l'intervalle $[E_i, L_j]$ sans compromettre l'achèvement du projet à la date E_n . La définition suivante paraît donc assez naturelle: la *marge totale* (total float) d'une activité (i,j) est la quantité

$$TF_{ij} = L_j - E_i - d_{ij}.$$

Appelons une activité *critique* si elle se trouve sur (au moins) un chemin critique. On démontre qu'une activité est critique si et seulement si sa marge totale est égale à 0. En particulier, si (i,j) est une activité critique, alors $E_i = L_i$ et $E_j = L_j$: on doit donc entamer l'activité en E_i et la terminer en L_j pour ne pas retarder le projet. Ceci signifie que les activités critiques sont précisément celles dont la réalisation doit être surveillée de près lors du déroulement d'un projet. L'identification des activités critiques permet donc au gestionnaire de concentrer son attention sur quelques activités essentielles plutôt que de la disperser sur l'ensemble du projet.

La définition suivante est motivée par l'observation qu'une activité (i,j) peut être exécutée n'importe quand dans l'intervalle $[E_i, E_j]$ sans influencer en rien le démarrage des autres activités (puisque, de toute façon, l'événement j ne pourra pas être réalisé avant la date E_j): la *marge libre* (free float) d'une activité (i,j) est la quantité:

$$FF_{ij} = E_j - E_i - d_{ij}.$$

La différence entre marge totale et marge libre est bien mise en évidence sur notre exemple numérique.

Exemple 1 (suite).

	TF	FF
A	$3-0-3 = 0$	0
B	$6-3-2 = 1$	$5-3-2 = 0$
C	0	0
D	$10-5-4 = 1$	$10-5-4 = 1$
E	0	0
F	$13-0-6 = 7$	7
G	0	0

Supposons que le calendrier établi pour le projet de l'Exemple 1 demande que chaque activité démarre le plus rapidement possible, et donc à sa date au plus tôt. Considérons à présent les activités B et D. La marge totale de B vaut 1, puisque, en retardant le démarrage de B d'un jour, on pourrait encore terminer le projet en 17 jours. Par contre, la marge libre de B est nulle, puisque tout retard dans l'exécution de B retarderait le démarrage de D par rapport à sa date au plus tôt calculée, c'est-à-dire 10. Le cas de l'activité D est différent: toutes autres choses restant égales, un jour de retard dans le démarrage de D n'influencerait en aucune façon le démarrage de E, ni d'aucune autre activité. La marge libre de D est donc égale à 1.

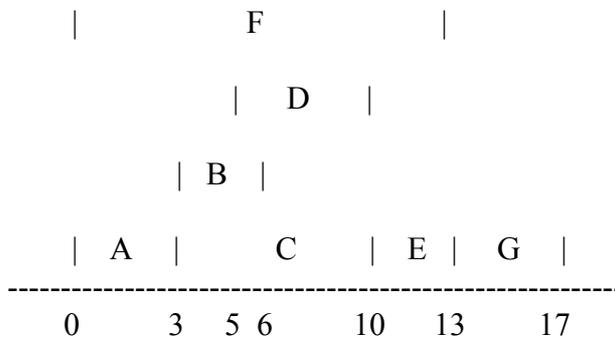
11.3. Construction d'un ordonnancement et implantation.

Le calcul des dates au plus tôt, des dates au plus tard et des marges fournit des informations utiles au gestionnaire mais ne livre pas directement un ordonnancement, c'est-à-dire un calendrier d'exécution du projet. En fait, plusieurs ordonnancements sont généralement compatibles avec les dates calculées, et un choix définitif implique la considération d'autres critères que celui de la date d'achèvement final.

Il est de tradition de résumer par un *graphique de Gantt* l'essentiel de l'information numérique fournie par les méthodes de chemins critiques. Dans le contexte de l'ordonnancement de projets, un graphique de Gantt consiste en un axe temporel sur lequel chaque activité (i,j) est visualisée à l'aide d'un intervalle $[E_i, L_j]$. Si l'activité est critique, sa période d'exécution doit donc nécessairement coïncider avec l'intervalle correspondant. Par contre, si l'activité n'est pas critique, alors sa durée est plus courte que l'intervalle $[E_i, L_j]$ et son exécution doit être programmée entre les bornes de l'intervalle.

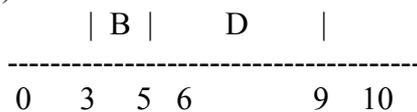
Exemple 1 (suite).

Voici un ordonnancement possible pour le projet de l'Exemple 1 :

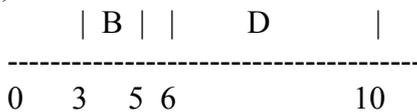


Pour les activités B et D, quelques ordonnancements compatibles avec les dates calculées sont donc (rappelons que B doit précéder D):

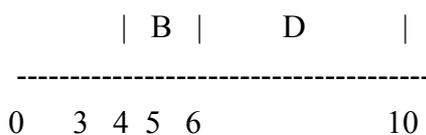
(1)



(2)



(3)



L'ordonnancement (3) est préférable aux ordonnancements (1) et (2) si certaines ressources nécessaires pour l'exécution de B ne sont disponibles qu'à partir du quatrième jour. De même, (2) peut être écarté au profit de (3) si il est préférable de lancer D dans la foulée de B (par exemple, parce que ces deux activités requièrent la présence des mêmes ouvriers sur le chantier).

En règle générale, il est très difficile d'établir un ordonnancement satisfaisant à des contraintes sur les ressources disponibles ou à d'autres critères secondaires. Pour résoudre ce type de problème, les logiciels commerciaux font appel à des méthodes heuristiques qui ne peuvent pas garantir le calcul d'un calendrier optimal et qui doivent donc parfois être guidées 'manuellement'.

Lorsqu'un plan d'ordonnancement a finalement été adopté et que les premières activités ont été lancées, le réseau d'activités devient un précieux outil de contrôle pour le gestionnaire de projet. Les dates et marges calculées lui permettent de surveiller l'évolution du projet et de s'assurer que les déviations enregistrées ne compromettent pas le calendrier établi. Si la réalisation du projet s'écarte exagérément du programme planifié, il ne restera plus qu'à recalculer un nouvel ordonnancement qui tienne compte de l'état d'avancement réel du projet.

11.4. Modifications du modèle de base.

Dans cette section, nous allons brièvement discuter la façon dont le modèle de base de la gestion de projet peut être enrichi lorsque l'on abandonne la vision étroite de la durée des activités comme une donnée déterministe et rigide, pour la considérer plutôt comme un paramètre aléatoire ou encore comme une variable de décision. Quelques remarques s'imposent au sujet de ces modifications.

Tout d'abord, les modèles que nous allons présenter sont plus généraux que le modèle étudié dans les sections précédentes. Il serait faux, cependant, d'en tirer la conclusion que le modèle de base est forcément insuffisant ou incomplet, et peut donc être froidement jeté aux oubliettes. Bien au contraire, il existe de nombreuses situations dont les caractéristiques essentielles sont adéquatement reflétées par le modèle de base. Dans de telles circonstances, la modélisation de détails plus fins sera vue par le gestionnaire comme une nuisance plutôt que comme une option désirable. On retrouve ici l'illustration d'un des principes fondamentaux de la construction de modèles, le principe KISS (Keep It Simple, Stupid; un modèle n'étant jamais qu'une approximation de la réalité, un compromis doit être recherché entre le niveau de détail et la simplicité du modèle).

Les modèles ci-dessous peuvent également être vus comme autorisant une *analyse de sensibilité* des résultats dérivés dans le cadre du modèle de base. Après avoir calculé la durée critique d'un projet et les marges associées à chaque activité, on peut ainsi se poser des questions du type:

- si les durées des activités ne sont pas connues avec certitude, quelle est la probabilité de terminer avant la date critique ?
- si la durée des activités peut être (partiellement) maîtrisée, combien de temps faudrait-il consacrer à chacune d'entre elles pour terminer dans les délais contractuels ?

11.4.1. Durées incertaines (modèle PERT).

Pour certains projets, il est très difficile de fournir une estimation raisonnable de la durée de chaque activité. C'est par exemple le cas lorsque la réalisation de certaines activités est conditionnée par des éléments purement aléatoires (conditions météorologiques, ...) ou échappant au contrôle du gestionnaire de projet (grèves, sous-traitance, ...), ou lorsqu'aucun

projet similaire n'a été réalisé précédemment. Dans de telles conditions, on choisit souvent de traiter la durée de chaque activité comme une variable aléatoire. Ceci a pour conséquence que la durée totale du projet ne peut évidemment plus être fixée de façon déterministe, comme c'était le cas précédemment. On devra donc se contenter, au mieux, d'énoncer des prédictions sous forme probabiliste quant aux dates d'achèvement du projet ou des différentes activités. La méthode PERT s'efforce précisément d'analyser ce type de situations.

Note. La méthode PERT a été conçue dans le cadre du développement du missile Polaris par la U.S. Navy, vers la fin des années cinquante. Ce projet impliquait plus de 9000 sous-traitants et une incertitude considérable entourait la durée des activités, pour lesquelles aucune expérience préalable n'était disponible.

Soit un réseau d'activités G pour lequel la durée d_{ij} de l'activité (i,j) est incertaine. On admettra que chaque durée d_{ij} est une variable aléatoire décrite par trois paramètres numériques:

- la *durée la plus probable* de l'activité, notée m_{ij} ,
- la *durée optimiste* a_{ij} de l'activité, c'est-à-dire le temps requis par cette activité dans des circonstances extrêmement favorables;
- la *durée pessimiste* b_{ij} de l'activité, c'est-à-dire le temps requis par cette activité dans des circonstances adverses.

Ces paramètres doivent bien entendu satisfaire aux contraintes: $a_{ij} \leq m_{ij} \leq b_{ij}$.

Sur base de ces paramètres, on peut obtenir une estimation de l'espérance (μ_{ij}) et de la variance (V_{ij}) de la durée d_{ij} . Nous utiliserons pour cela les formules suivantes:

$$\mu_{ij} = (a_{ij} + b_{ij} + 4m_{ij})/6$$

$$V_{ij} = \left(\frac{b_{ij} - a_{ij}}{6} \right)^2$$

(Ces formules sont justifiées dans la littérature sous l'hypothèse additionnelle que la variable d_{ij} suit une loi bêta. Plus simplement, on peut observer que μ_{ij} est exprimé comme une moyenne pondérée des durées optimiste, probable et pessimiste, et que l'expression de V_{ij} peut être motivée par la considération de l'inégalité de Chebyshev. Cette inégalité assure en effet que, pour une variable aléatoire X quelconque,

$$\text{Prob} \left(|X - E(X)| \leq 3\sqrt{V(X)} \right) \geq 0.89 .$$

Intuitivement, l'écart entre les valeurs observées d'une variable aléatoire et sa moyenne excède donc rarement 3 écarts-types.)

Par la suite, nous traiterons d_{ij} comme une variable aléatoire dont la distribution est partiellement caractérisée par la connaissance de μ_{ij} et V_{ij} .

Puisque la durée de chaque activité est aléatoire, il est clair que la date d'achèvement du projet elle-même, c'est-à-dire E_n , doit être traitée comme une variable aléatoire. En règle générale, il est très difficile d'estimer la distribution de E_n , et donc de calculer les probabilités qui s'y rapportent (par exemple, la probabilité que le projet soit terminé dans des délais prescrits). On se contente souvent d'estimer cette probabilité en formulant une série d'approximations que nous allons décortiquer successivement.

Tout d'abord, appelons G^* le réseau (déterministe) obtenu en affectant une durée μ_{ij} à chaque activité (i,j) de G , et considérons un chemin critique CC de G^* . Nous noterons d_{CC} la variable aléatoire $\sum_{(i,j) \in CC} d_{ij}$ (il s'agit bien là d'une quantité *aléatoire* qu'il convient de distinguer soigneusement de la quantité *certaine* $\sum_{(i,j) \in CC} \mu_{ij}$).

Approximation 1. *On admettra que la durée totale du projet est égale à la somme des durées des activités situées sur le chemin CC , c'est-à-dire qu'on identifiera E_n à d_{CC} .*

Remarquons que ceci constitue bien une approximation de la réalité puisque CC est seulement critique pour G^* et non pas nécessairement pour G . Dans les faits, des activités situées le long d'un autre chemin que CC pourraient déterminer la durée totale du projet (cette observation est bien mise en évidence par la considération d'un projet pour lequel G^* comprendrait deux chemins critiques).

En admettant l'approximation précédente, la variable aléatoire à considérer devient donc d_{CC} plutôt que E_n . Que savons-nous de la distribution de cette variable ? Tout d'abord, l'espérance de d_{CC} est donnée par $\sum_{(i,j) \in CC} \mu_{ij}$. Ensuite, en posant l'hypothèse que les durées des activités de CC sont des variables aléatoires indépendantes, on peut déduire que la variance de d_{CC} est égale à $\sum_{(i,j) \in CC} V_{ij}$. Dans la pratique, il est fréquent de pousser plus loin les hypothèses et de postuler l'approximation suivante:

Approximation 2. *On admettra que la somme des durées des activités situées sur le chemin CC , c'est-à-dire la variable aléatoire d_{CC} , suit une loi normale de moyenne $\sum_{(i,j) \in CC} \mu_{ij}$ et de variance $\sum_{(i,j) \in CC} V_{ij}$.*

Cette approximation est souvent justifiée par un appel au théorème central limite. En effet, ce théorème implique que, si le nombre d'activités situées sur le chemin critique est assez grand et si les durées de ces activités sont mutuellement indépendantes, alors (sous des conditions très générales) la somme des durées des activités suit approximativement une loi normale. Remarquons cependant que, en pratique, ces hypothèses ne sont évidemment pas toujours satisfaites, si bien que l'Approximation 2 doit être utilisée avec prudence.

En utilisant les approximations formulées ci-dessus, il devient possible d'estimer les probabilités relatives à la durée du projet. Considérons par exemple, la probabilité que le projet soit terminé avant une date butoir S :

$$\begin{aligned} \text{Prob}(E_n < S) &\cong \text{Prob}(d_{CC} < S) \\ &= \text{Prob}\left(\frac{(d_{CC} - \sum_{(i,j) \in CC} \mu_{ij})}{\sqrt{\sum_{(i,j) \in CC} V_{ij}}} < \frac{(S - \sum_{(i,j) \in CC} \mu_{ij})}{\sqrt{\sum_{(i,j) \in CC} V_{ij}}}\right) \\ &\cong \text{Prob}\left(Z < \frac{(S - \sum_{(i,j) \in CC} \mu_{ij})}{\sqrt{\sum_{(i,j) \in CC} V_{ij}}}\right) \end{aligned}$$

où Z suit une loi normale standardisée $N(0,1)$ (le symbole \cong indique une approximation).

Pour en terminer avec la question des durées incertaines, observons encore que les multiples approximations décrites ci-dessus se traduisent souvent, dans la pratique, par une estimation trop optimiste de la durée des projets (par exemple, l'approximation 1 néglige la probabilité qu'un chemin autre que CC devienne critique). Ce biais systématique conduit alors à l'élaboration de calendriers trop serrés et au dépassement des délais annoncés.

11.4.2. Compression des durées (modèle CPM).

Revenons à présent en univers certain, comme dans le modèle de base de la Section 11.2. Après avoir déterminé la durée critique du projet (E_n), le gestionnaire peut s'interroger sur la possibilité de réduire cette durée. Pour ce faire, il lui sera évidemment nécessaire de modifier le réseau d'activités, c'est-à-dire la définition des activités, les relations d'antériorité ou les durées des activités.

Pour illustrer le concept de redéfinition des activités, considérons un exemple très simplifié.

Exemple 3. La construction d'un immeuble de 8 étages comporte les activités suivantes:

A: câblage électrique (16 jours).

B: plafonnage (24 jours).

Le plafonnage ne peut être effectué que lorsque le câblage électrique est terminé.

On conclut facilement que la réalisation des activités A et B exige au moins 40 jours.

Supposons maintenant que la définition des activités soit affinée comme suit:

A_i : câblage de l'étage i (2 jours), pour $i = 1, \dots, 8$;

B_i : plafonnage de l'étage i (3 jours), pour $i = 1, \dots, 8$.

Les contraintes d'antériorité peuvent à présent être exprimées étage par étage: par exemple, l'étage i doit être câblé avant l'étage $i+1$ ($i = 1, \dots, 7$) et chaque étage doit être câblé avant

d'être plafonné. Sous ces contraintes, on calcule facilement que la durée critique du projet est de 19 jours.

Cet exemple fournit une nouvelle illustration de l'interaction, déjà soulignée dans les Sections 11.1 et 11.2 ci-dessus, entre les phases de planification et d'ordonnancement d'un projet. En pratique, la redéfinition des activités et des relations d'antériorité constitue un outil fondamental, quoiqu'insuffisamment formalisé, de la gestion de projet.

Une autre façon d'envisager la réduction de la durée du projet est de passer par une compression de la durée des activités. Dans la réalité, la durée des activités peut souvent être réduite à condition de supporter un coût (direct) plus élevé pour leur réalisation: ce coût peut être par exemple celui de la main d'oeuvre additionnelle nécessaire pour accélérer la réalisation, ou le supplément de prix encouru pour une livraison rapide, etc. Notons d'ailleurs que, inversement, une firme sera parfois prête à accepter une augmentation de la durée d'un projet afin d'en diminuer le coût total. Le modèle CPM permet de formaliser rigoureusement cette recherche d'un compromis entre durée et coût du projet.

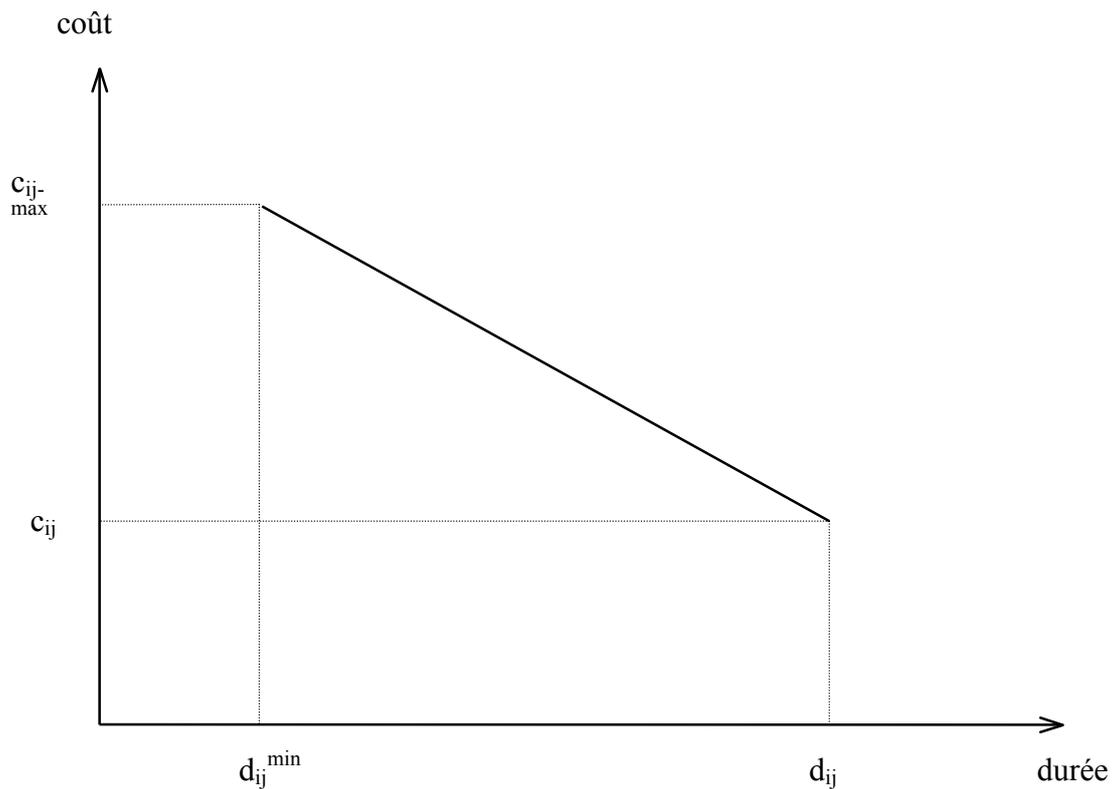
Note. La méthode CPM a été développée dans les années cinquante par la firme DuPont pour planifier et contrôler la maintenance de ses usines chimiques. La durée des activités requises pour ces projets de maintenance était généralement bien maîtrisée et pouvait être affectée de façon prédictible par l'addition de ressources supplémentaires.

Le modèle CPM suppose que chaque activité (i,j) est caractérisée par une *durée normale* d_{ij} et une *durée minimum* (durée accélérée, crash duration) d_{ij}^{\min} , où d_{ij}^{\min} représente la plus courte durée envisageable pour cette activité. Le coût de réalisation de l'activité est égal à c_{ij} pour la durée normale et à c_{ij}^{\max} pour la durée minimum. Entre ces deux extrêmes, on supposera que le coût décroît linéairement en fonction de la durée, comme indiqué sur le graphique ci-dessous.

La pente (en valeur absolue) de la droite représentée sur ce graphique vaut

$$\frac{c_{ij}^{\max} - c_{ij}}{d_{ij} - d_{ij}^{\min}}$$

et peut être interprétée comme le coût marginal (ou unitaire) de réduction de la durée de l'activité (i,j).



Exemple 4. Reprenons une fois encore le projet décrit par le réseau d'activités de l'Exemple 1 (Section 11.2) et associons lui les données numériques suivantes:

Activité	Durée normale	Coût normal	Durée minimum	Coût maximum	Coût marginal
A	3	190	2	200	10
B	2	120	2	120	-
C	7	90	4	150	20
D	4	120	3	150	30
E	3	100	3	100	--
F	6	60	4	70	5
G	4	100	3	140	40

Si l'on ne considère que les durées normales, alors ce projet est identique à celui de l'Exemple 1. Sa durée critique normale est donc de 17 jours, pour un coût normal de 780.

Supposons à présent que la durée normale du projet soit considérée comme excessive. Quelles activités faudrait-il accélérer pour réduire la durée totale du projet ? Une première réponse à cette question, basée sur une analyse marginale de la durée et des coûts, s'exprime à travers deux règles simples et intuitives (cette approche utilise l'hypothèse implicite que la durée de chaque activité comporte un nombre entier d'unités de temps):

Règle 1. *Pour réduire la durée critique d'un projet d'une unité, il faut et il suffit que, sur chaque chemin critique, la durée d'une activité soit réduite d'une unité.*

Cette règle découle évidemment de la définition même des chemins critiques. Permettons-nous cependant d'insister ici sur la nécessité de réduire *simultanément* la durée de tous les chemins critiques.

Exemple 4. Ce projet ne comporte qu'un seul chemin critique (A,C,E,G). Pour réduire sa durée totale à 16 jours, il faudra donc réduire d'une unité la durée d'une des activités A, C ou G (puisque la durée de E est incompressible).

Lorsqu'un choix est possible (comme dans notre exemple), quelles activités faut-il accélérer ? Si le but est de réduire la durée du projet à coût minimum, alors la Règle 1 pourra être complétée de la façon suivante:

Règle 2. *Pour réduire la durée critique d'un projet d'une unité au moindre coût, il faut et il suffit que, sur chaque chemin critique, la durée d'une activité soit réduite d'une unité et que la somme des coûts marginaux associés à ces réductions soit minimisée parmi tous les choix possibles.*

La deuxième partie de cette règle ressemble furieusement à une Lapalissade, bien plus qu'à un énoncé mathématique. Plutôt que de lui chercher une formulation précise, nous allons nous contenter de l'illustrer ci-dessous sur un exemple concret.

Exemple 4 (suite). Nous avons déjà observé que, pour réduire la durée du projet à 16 jours, il était nécessaire d'accélérer une des activités A, C ou G. Puisque le coût marginal de A est le plus faible (parmi ces trois activités), nous commencerons donc par réduire la durée de A de 1 jour. L'activité A est ainsi ramenée à sa durée minimum et le coût total du projet s'élève à 790.

Pour comprimer davantage la durée du projet, il sera nécessaire de réduire la durée de C de 7 à 6 jours, puisque (A,C,E,G) est toujours l'unique chemin critique. Suite à cette réduction, le projet ne demande plus que 15 jours de travail pour un coût de 810.

Il est essentiel de remarquer que le projet ainsi accéléré comporte à présent deux chemins critiques, (A,C,E,G) et (A,B,D,E,G). On observe là un phénomène typique de la compression des activités: les réductions successives imposées aux activités critiques finissent par communiquer ce même statut critique à d'autres activités qui ne jouaient jusque-là aucun rôle déterminant pour la durée du projet.

Tâchons maintenant d'appliquer la Règle 2. Pour ramener la durée du projet à 14 jours, il faut réduire

- soit C, soit G sur le chemin (A,C,E,G),

et simultanément

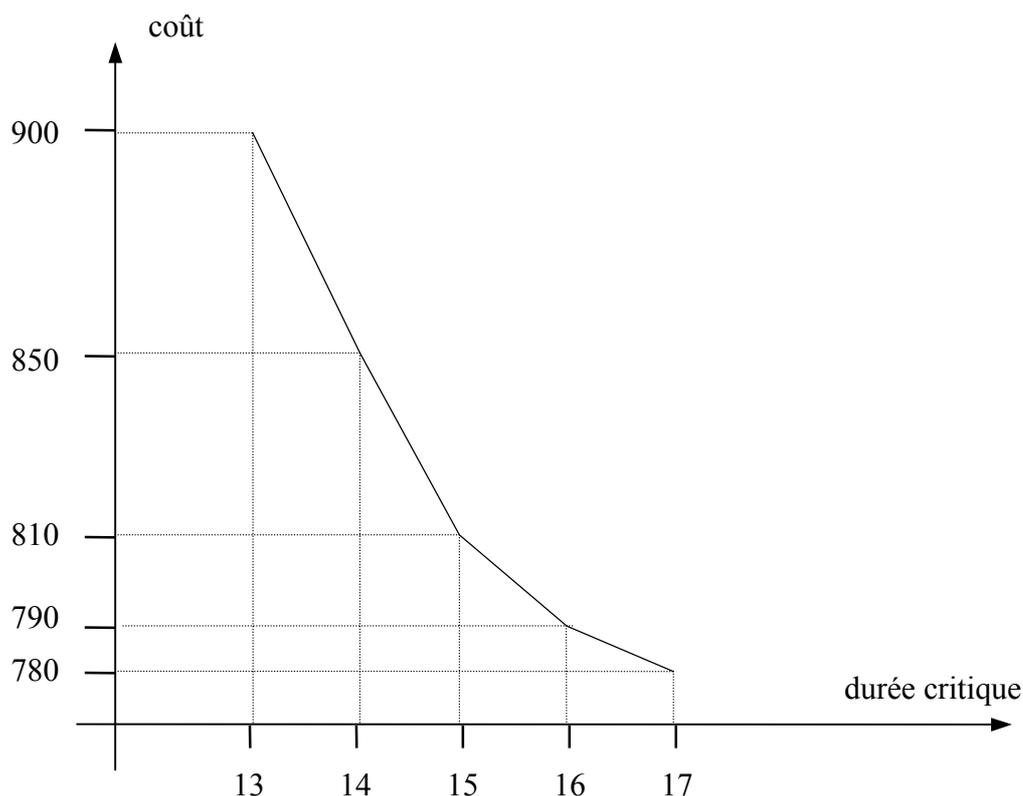
- soit D, soit G sur le chemin (A,B,D,E,G).

Les coûts marginaux de C, D et G étant respectivement de 20, 30 et 40, la solution optimale consiste donc à réduire la durée de G d'un jour (et non pas les durées de C et D, comme un examen superficiel de la situation pourrait le laisser penser).

Si la durée de G est ainsi réduite à 3 jours, le projet pourra être terminé en 14 jours pour un coût de 850. On vérifiera qu'il est encore possible de raccourcir la durée du projet d'un jour supplé-

mentaire contre un coût additionnel de 50. La durée critique de 13 jours qui en résulte représente, elle, le minimum absolu réalisable pour le projet. (Pourquoi ?)

Une analyse marginale du type décrit ci-dessus permet d'établir la courbe de coût du projet en fonction de sa durée totale. Pour l'Exemple 4, cette courbe est représentée sur le graphique suivant:



Dans la pratique, l'analyse marginale fournit un support utile à la prise de décisions relatives à l'ordonnancement du projet, en exprimant de façon explicite les compromis durée-coût.

Exemple 4 (suite). Supposons que le contrat lié au projet décrit à l'Exemple 4 mentionne un délai de 14 jours pour l'achèvement des travaux. En cas de dépassement de ce délai, l'entrepreneur devra verser une pénalité de 25 (unités monétaires) par jour de retard.

Sur base de l'analyse précédente, on calcule facilement le coût total encouru par l'entrepreneur en fonction de la durée du projet:

durée	coût + pénalité
17	$780 + 75 = 855$
16	$790 + 50 = 840$
15	$810 + 25 = 835$
14	$850 + 0 = 850$

D'un strict point de vue de l'analyse des coûts, il est donc dans l'intérêt de l'entrepreneur de terminer le projet avec un jour de retard. Pour cela, il devra réduire d'une unité la durée des activités A et C par rapport à leur durée normale.

Signalons encore, avant de clore ce chapitre, que l'analyse marginale des coûts devient rapidement impraticable pour des problèmes d'ordonnement de grande taille. Dans une telle situation, il est cependant possible d'établir un modèle de programmation linéaire qui exprime l'objectif de minimisation du coût sous des contraintes assurant l'existence d'un ordonnancement de longueur donnée.

Bibliographie

V. Giard, *Gestion de la Production*, 2ème édition, Economica, Paris, 1988.

G.D. Eppen, F.J. Gould and C.P. Schmidt, *Introductory Management Science* (2ème éd.), Prentice Hall, 1993.

H.A. Taha, *Operations Research: An Introduction* (5ème éd.), McMillan Publishing Company, 1992.