RAIRO. RECHERCHE OPÉRATIONNELLE

DAVID DE ALMEIDA

Agrégation des données pour l'évaluation des performances de systèmes flexibles de production par modèles à réseaux de files d'attente

RAIRO. Recherche opérationnelle, tome 32, n° 2 (1998), p. 145-192

<http://www.numdam.org/item?id=RO_1998__32_2_145_0>

Recherche opérationnelle/Operations Research
(vol. 32, n° 2, 1998, pp. 145-192)

AGRÉGATION DES DONNÉES POUR L'ÉVALUATION DES PERFORMANCES DE SYSTÈMES FLEXIBLES DE PRODUCTION PAR MODÈLES À RÉSEAUX DE FILES D'ATTENTE (*)

par David DE ALMEIDA ([1])

Communiqué par Charles TAPIERO

Résumé. – *Dans cet article, nous nous intéressons à l'évaluation/prédiction quantitative des performances de Systèmes Flexibles de Production (SFP) par modèles à Réseaux de Files d'Attente (RFA) lors des phases de conception et de planification de ces systèmes. L'un des problèmes majeurs posés par de telles études concerne l'obtention objective des caractéristiques d'un modèle à RFA exploitable sur calculateur, à partir des données des industriels (gammes/nomenclatures opératoires, quotas de production). En effet, l'exploitation* ad hoc *de ces données conduit généralement à un nombre prohibitif de chaînes et de classes de clients, excluant* de facto *toute résolution par méthodes analytique et/ou numérique. Nous proposons donc une formalisation de la charge d'un SFP exploitée pour définir un processus systématique d'agrégation des données de la charge, et qui fournit, de manière transparente, les caractéristiques exactes d'un modèle à RFA soluble analytiquement et/ou numériquement. Nous montrons l'applicabilité et la qualité de nos résultats sur plusieurs exemples, dont un SFP de taille industrielle.* © Elsevier, Paris

Mots clés : Systèmes flexibles de production, réseaux de files d'attente, approche stochastique, agrégation des données.

Abstract. – *This paper deals with the use of Queueing Network (QN) models for evaluating/predicting quantitatively the performances of Flexible Manufacturing Systems (FMS) at the strategic and tactical decision levels. One of the main problems in such studies consists in obtaining objectively the characteristics of a computationally tractable QN model from the data provided by industrial experts (process plans, production ratios). Indeed, when exploiting directly these data, one is usually led to considering QN models that cannot be analytically or numerically solved because they exhibit too many customer chains and classes. Thus, we propose a formalization of the workload of an FMS that is further exploited for defining a systematic workload data aggregation process. This process enables a transparent obtention of the exact characteristics of a QN model which can be analytically and/or numerically solved. The application and the quality of our results are shown on several examples, one of which being an industrial FMS.* © Elsevier, Paris

Keywords: Flexible manufacturing systems, queueing network models, stochastic approach, data aggregation.

(*) Reçu en Septembre 1995.

([1]) Laboratoire d'Informatique, de Modélisation et d'Optimisation des Systèmes (LIMOS), ISIMA - Université Blaise Pascal, Clermont-Ferrand-II, Campus des Cézeaux, B.P. 125, 63173 Aubiere Cedex, France. E-mail: david.almeida@isima.fr

TABLE DES MATIÈRES

1. INTRODUCTION

Face à une concurrence internationale de mieux en mieux organisée et à un marché de plus en plus fluctuant, les industriels recherchent des moyens pour améliorer la productivité et la réactivité de leur système, tout en réduisant le plus possible les coûts de production et en augmentant la qualité. Dans ce contexte, les systèmes flexibles de production (SFP) se sont développés significativement. Les problèmes posés par un SFP sont généralement classés selon trois niveaux d'horizons temporels [STE84] – à savoir le niveau stratégique (conception : dimensionnement, types des moyens de production/transport...), le niveau tactique (planification...) et le niveau opérationnel (ordonnancement, pilotage...). Les experts industriels tentent de résoudre ces problèmes en s'appuyant généralement plus sur l'intuition

et l'expérience que sur des méthodes objectives. Or, cette démarche conduit rarement à des résultats satisfaisants. La construction de modèles de ces systèmes offre alors une alternative intéressante, si elle est conduite dans un cadre conceptuel rigoureux qui s'appuie sur une démarche méthodologique, afin que les modèles construits soient cohérents avec les systèmes étudiés et pertinents relativement aux objectifs de l'étude.

La nature des problèmes relevant des niveaux stratégique et tactique est très propice à l'utilisation d'outils permettant d'appréhender le régime permanent d'un SFP. Par ailleurs, compte-tenu du nombre considérable de paramètres de décision, les experts industriels sont davantage intéressés à ces niveaux par des méthodes d'évaluation quantitative de performances offrant un bon compromis qualité/rapidité d'exécution. Une telle approche permet en effet de mener un prototypage rapide [SUR89] où quelques solutions alternatives et pertinentes sont rapidement déterminées analytiquement avant d'être évaluées de manière plus détaillée, par exemple par simulation. Ainsi, [BUZA93, pages 14-16] discute de la complémentarité entre modèles à Réseaux de Files d'Attente (RFA) résolus analytiquement et modèles de simulation, non seulement dans le cadre d'une approche de prototypage rapide, mais aussi dans une perspective d'inter-validation des modèles. Une telle approche a également été adoptée pour la conception d'un atelier pilote de fabrication de composants électroniques chez IBM [BROW88]

Les SFP étant des *systèmes à flux discrets et à partage de ressources* ([DIE91], [MAC93]), la simulation à événements discrets offre une alternative qui n'est pas toujours satisfaisante au prime abord dans ce contexte – au moins pour des raisons de coûts –, en raison notamment du temps et de l'expertise requis pour concevoir, valider/vérifier et exécuter un tel modèle.

Bien que satisfaisant en tant que langage de spécification et largement utilisé dans le cadre des systèmes de production ([MUR89,VALA90, PROTH92, ZHOU93]...), le formalisme des Réseaux de Petri (RdP), munis de ses nombreuses extensions, n'est pas exempt d'inconvénients en termes (i) de lisibilité d'un tel modèle quand la taille du SFP décrit augmente, (ii) d'évolutivité (point crucial par rapport aux SFP), et (iii) de description de l'aspect matériel représentant la partie procédé ([AMAR92]). Concernant l'évaluation quantitative des performances d'un SFP, les RdP – ils sont alors au moins temporisés – offrent une classe d'outils en aval reposant essentiellement sur la simulation (joueur de RdP) et les méthodes numériques [MOL81, AJM86] – ces dernières étant d'ailleurs pratiquement inexploitables si le nombre d'états de la chaîne de Markov isomorphe est très élevé –.

Dans ce contexte, le formalisme des Réseaux de Files d'Attente (RFA) offre une alternative intéressante pour la construction de modèles (voir, par exemple, [DAL86c, HSU93, BUZA93, SUR95a]), en particulier quand ceux-ci sont résolus numériquement ou (surtout) analytiquement (dans cet article, nous ne considérons pas l'analyse opérationnelle [DEN78] [SUR83] [DAL86a] qui constitue une alternative à l'approche stochastique).

Soulignons enfin que les RFA et les RdP sont des formalismes considérés comme plus complémentaires qu'antagonistes.

2. MOTIVATIONS

Malgré de nombreux travaux relatifs à la modélisation par RFA des SFP ([BUZA86], [HSU93]), on constate peu de transferts de résultats des laboratoires de recherche au milieu industriel, à l'exception des logiciels CAN-Q ([SOL80]) et MPX (*voir*, par exemple, [SUR95a]). Cet état de fait est, à notre avis, dû au moins à trois raisons :

1. le formalisme des RFA n'est pas complètement satisfaisant pour appréhender la structure et le fonctionnement d'un SFP. En particulier, l'expert industriel spécifie la charge de son système de manière naturelle avec une *approche transactionnelle* (concept de gammes – séquences d'opérations –). Or, l'évaluation des performances d'un système avec un modèle à RFA nécessite de spécifier le comportement de chaque station du réseau vis-à-vis d'un client susceptible de visiter cette station. Cette approche de spécification est appelée *approche station* ou *approche ressource*. Le passage de l'approche transaction à l'approche station (obtention des caractéristiques d'un RFA modélisant un SFP industriel – temps de service et routage des clients à chaque station –) est une étape incontournable qui pose un problème de transformation adéquate de la connaissance reconnu comme difficile, car cette transformation n'est pas systématisée,
2. les modèles à RFA résolus analytiquement posent des problèmes numériques quand le nombre de types d'entités circulant dans le réseau et/ou le nombre de stations augmentent. Ce cas est caractéristique des SFP qui sont conçus pour produire un très grand nombre de références de produits, et qui comprennent des machines très versatiles,
3. les modèles stochastiques disponibles d'un SFP n'ayant généralement pas de forme produit [PUJ80, DIJ89], un problème de **choix de lois** (crédibilité du modèle) pour les durées de service des stations du RFA se pose donc.

Dans cet article, nous proposons des éléments de solutions à ces trois problèmes ([ALM96a]) en vue d'évaluer quantitativement les critères de performance à l'état stationnaire – *i.e.* dans le cadre des niveaux stratégique et tactique – de SFP de *type job-shop*.

En ce qui concerne le premier point, nous nous inscrivons dans un processus de modélisation développé au LIMOS depuis plusieurs années. Ce processus itératif préconise la construction d'un *modèle de connaissance* du système étudié duquel est déduit un *modèle d'action* exploité pour évaluer les performances du système ([GOUR84]). Pour construire le modèle de connaissance d'un système, nous adoptons une décomposition systématique des SFP en trois sous-systèmes communicants, disjoints deux à deux ([GOUR92, KEL92]) :

- le sous-système logique (SSL) décrit par la charge du système, c'est-à-dire l'ensemble des gammes/nomenclatures des pièces à produire, ainsi que les quotas de production,
- le sous-système physique (SSP) décrit par l'ensemble des unités de production, de stockage et de transport/manutention ainsi que leurs interconnexions (topologie),
- le sous-système décisionnel (SSD) comprenant l'ensemble des règles de gestion/pilotage du système.

Le deuxième problème est un problème typique d'**agrégation de données.** En effet, si le formalisme des RFA offre les concepts de chaîne et de classe de clients pour résoudre le problème du passage de l'approche transaction à l'approche station, l'étude d'un SFP *de dimension industrielle* n'est cependant pas envisageable avec une résolution mathématique d'un modèle à RFA si on exploite, tel quel, le modèle de connaissance (et plus particulièrement le SSL) de ce SFP. Nous proposons un processus systématique d'agrégation de données qui calcule automatiquement, à partir de la description du SSL d'un SFP, les caractéristiques exactes du RFA modélisant ce système : durées moyennes de service agrégées et probabilités de routage des clients pour chaque station du RFA. On *déduit* donc, du modèle de connaissance d'un SFP, un modèle de connaissance dérivé, construit par agrégation des données qui décrivent le SSL.

Enfin, le dernier problème concerne la crédibilité industrielle des modèles analytiques à RFA souvent contestée à cause d'hypothèses simplificatrices peu réalistes, notamment en ce qui concerne les lois des durées de service. Les lignes flexibles de production (SFP en configuration « flow-shop ») ont fait l'objet de nombreuses publications concernant ce problème ([DAL92]). En

revanche, on relève peu de résultats à ce sujet sur les SFP en configuration « job-shop » (voir, par exemple, [SUR95b]). Afin d'obtenir *a priori* une meilleure caractérisation des distributions des durées de service pour chaque station du RFA modélisant un système, nous calculons les Carrés des Coefficients de Variation (CCV) de ces durées à l'état stationnaire. On peut ainsi affiner la modélisation en utilisant, par exemple, des lois coxiennes d'ordre 2 pour les durées de service ([YAO85b]).

3. CONTEXTE ET ÉTUDE DE L'EXISTANT

Le formalisme des RFA propose le concept de *classe* de clients pour regrouper les clients qui ont un comportement statistiquement identique (service et routage) dans une ou plusieurs stations.

Une exploitation directe du SSL conduit à associer, de manière unique, à chaque gamme une chaîne de routage comportant elle-même un ensemble de classes qui représentent l'état d'avancement dans la gamme d'un client. Dans une chaîne donnée, un client change donc de classe après chaque traitement (ou opération). Par nature des SFP, l'exploitation directe du SSL dans un RFA résolu par méthode analytique est impossible pratiquement, un nombre élevé de classes limitant l'exploitation des méthodes mathématiques de résolution des modèles à RFA ([BUZ73, ZAH80, MVA80]). Pour résoudre ce problème, nous proposons deux types d'agrégation des données du SSL, à savoir :

- **une agrégation de type 1**, dite agrégation *intra-gamme*, qui consiste à agréger, pour chaque chaîne de clients associée à une gamme, les classes qui la composent. Cette agrégation a donc pour but d'agréger les classes d'une chaîne associée à une gamme dont la suite d'opérations comporte des passages multiples et statistiquement différents sur les machines,
- **une agrégation de type 2**, dite agrégation *inter-gammes*, consistant à agréger les classes dues à un ensemble de gammes par une seule classe. Cette agrégation a donc pour objectif le remplacement de plusieurs gammes par une gamme commune où les passages multiples sur un même groupe de machines sont statistiquement équivalents.

[ZAH80] aborde cette problématique dans le cadre des RFA *à forme produit* (notamment les réseaux de type BCMP [BCMP75]). Néanmoins, l'auteur étudie davantage les systèmes informatiques qui ne posent pas les mêmes problèmes que les SFP. Par exemple, dans un SFP, une pièce peut

recevoir des services distincts avec des durées différentes sur un même groupe de machines. D'autre part, les ressources qui composent un système informatique ont leur activité décrite généralement de manière statistique – mesures –. L'auteur considère donc une description de la charge sous forme de « gammes » résultant déjà d'une agrégation de type 1, généralement obtenues par mesures sur le système.

[SOL77] et [MVAQ84] proposent des logiciels pour évaluer les performances des SFP par des modèles à RFA fermés. [SOL77] exploite un modèle à RFA mono-classe à forme produit, mais ne précise pas formellement comment les caractéristiques de ce RFA sont obtenues à partir de la description de la charge soumise au SFP étudié. Par ailleurs, cette agrégation se place dans le contexte des hypothèses requises par les méthodes de résolution du RFA. Dans ce contexte, et en particulier, l'auteur ne s'intéresse qu'à l'obtention de durées moyennes agrégées de service, et ne prend pas en compte explicitement une estimation de la variabilité de ces services.

[MVAQ84] exploite l'algorithme MVA pour résoudre des RFA multi-classes à forme produit, chaque classe résultant d'une agrégation de type 1. Cette agrégation exploite une spécification de la charge sous forme de gammes, mais elle n'est pas clairement décrite. Les classes (associées à chaque catégorie de pièces) permettent de prendre en compte les quotas de production de ces catégories et, éventuellement, la présence de palettes dédiées à chaque catégorie de pièces. Néanmoins, on peut penser que l'utilisation de ce logiciel est limitée quand le nombre de catégories de pièces produites simultanément est important, ce qui est généralement le cas des SFP de taille industrielle ; une solution est alors une agrégation de type 2 qui n'est pas prise en compte dans le logiciel. D'autre part, et comme pour [SOL77], l'agrégation de type 1 est effectuée dans le cadre des hypothèses requises par la méthode de résolution du RFA.

Pour évaluer les performances quantitatives d'un SFP, une formalisation du processus d'agrégation (de type 1 ou 2) des données du SSL d'un SFP semble donc intéressante, sinon indispensable. Son automatisation doit permettre de rendre le processus d'agrégation transparent à l'expert industriel qui n'a alors pour charge que de spécifier le SSL. Cette formalisation est bien sûr indépendante des méthodes de résolution d'un RFA.

Cependant, les valeurs des critères de performance obtenues en exploitant le processus d'agrégation peuvent être erronées. Dans [SOL77] et [MVAQ84], les auteurs imputent les erreurs commises sur les valeurs des

critères de performance recherchées à la nature des modèles à RFA utilisés. [ZAH80] étudie davantage les erreurs causées par le processus d'agrégation dans le cadre des RFA à forme produit. L'auteur met en évidence l'existence d'états fictifs (*cf.* paragraphe 7.2), qui sont une source d'erreurs dans le processus stochastique sous-jacent quand une agrégation de type 2 est effectuée, les résultats obtenus avec l'agrégation de type 1 étant exacts pour les RFA à forme produit. L'auteur calcule également des bornes analytiques concernant la qualité du processus d'agrégation. Pour des RFA généraux qui ne sont pas à forme produit, l'obtention de telles bornes ne semble pas aussi immédiate. Une validation expérimentale est donc requise pour étudier la qualité des critères de performance obtenus à partir des caractéristiques calculées par le processus d'agrégation.

Pour éviter la croissance du nombre de ces états fictifs, [ZAH80] propose d'agréger des gammes « homogènes » (ou familles de gammes) entre elles. Le problème revient alors à définir et à quantifier le concept d'*homogénéité* de gammes. De nombreux travaux définissent et quantifient la similarité entre deux gammes ([SHA93]) dans le cadre des SFP. Néanmoins, la majorité des critères proposés repose uniquement sur la description du SSL sans tenir compte des quotas de production.

[ZAH80] propose une heuristique afin de former des groupes de gammes homogènes. Pour chaque gamme, l'auteur forme le vecteur des taux d'utilisation normalisés des groupes de machines (ces taux sont calculés par désagrégation des résultats globaux obtenus par une évaluation des performances exploitant un SSL agrégé mono-classe). L'auteur construit ensuite de manière gloutonne une liste des gammes à partir de la distance euclidienne entre les vecteurs des taux d'occupation précédemment calculés. Cette liste est finalement fractionnée suivant le nombre de classes agrégées que l'on souhaite utiliser dans la modélisation. Cette heuristique donne des résultats intéressants, mais nécessite une évaluation de performances ; en outre, elle ne donne pas de règles précises quant au fractionnement de la liste de gammes construite.

Dans cet article, nous proposons d'abord une formalisation des caractéristiques génériques (*i.e.* logiques) et quantitatives du SSL. Les SFP ayant pour particularité d'avoir un nombre limité de palettes (ressources critiques), ce qui limite *de facto* le nombre de clients présents dans le système à n'importe quel instant, les modèles à RFA *fermés* s'avèrent donc être pertinents pour modéliser de tels systèmes, notamment quand ceux-ci sont étudiés en saturation. Dans notre étude, le modèle d'action d'un

SFP est un RFA fermé composé de stations multi-serveurs à loi de service générale représentant le SSP. Nous donnons ensuite une formalisation du processus d'agrégation des données du SSL reposant sur la formalisation de ce sous-système. Enfin, nous examinons les principales sources d'erreurs sur le calcul des critères de performance introduites par ce processus, et vérifions expérimentalement la robustesse de cette technique d'agrégation sur plusieurs exemples, dont un SFP de la Manufacture MICHELIN ayant fait, au LIMOS, l'objet d'études de simulation et d'ordonnancement antérieures.

4. PROPOSITION D'UNE FORMALISATION DU SOUS-SYSTÈME LOGIQUE

De manière générique, le SSL d'un SFP décrit, pour chaque type de composant, l'enchaînement logique des opérations précisé par la gamme opératoire de ce produit. Les aspects quantitatifs sont liés à la charge soumise au système étudié (quotas de production). La première partie décrit, de manière descendante, les caractéristiques génériques du SSL et introduit, au fur et à mesure, les éléments du formalisme. Le cas où les gammes sont linéaires est tout d'abord traité, puis le formalisme est étendu à des gammes qui comportent des routages alternatifs probabilisés *a priori* (pas de routages dynamiques). Les aspects quantitatifs du sous-système logique sont enfin intégrés pour compléter la formalisation.

4.1. Notations

Soit H un ensemble fini ; $|H| \in \mathbb{N}$ désigne le cardinal (ou nombre d'éléments) de l'ensemble H.

Dans la suite, on nomme groupe de machines un ensemble de machines fonctionnellement identiques – *i.e.* les machines possèdent les mêmes caractéristiques techniques et leur outillage leur permet de réaliser les mêmes opérations à la même vitesse –, et qui partagent la même zone de stockage en entrée des pièces [STE83]. On note $nmac_m$ le nombre de machines appartenant au groupe de machines m. Dans la description du SSP d'un SFP, on note $M = \{m_1, \ldots, m_{|M|}\}$ l'ensemble des groupes de machines du SSP ($|M| < +\infty$).

4.2. Caractéristiques génériques du sous-système logique

4.2.1. Cas des gammes linéaires

Un SFP produit des pièces, chaque type de pièces étant caractérisé par sa **gamme/nomenclature** (*gamme d'ordonnancement* dans [CAMP85]). On

note G l'ensemble des gammes que le système peut accomplir, et on pose $G = \{g_1, g_2, ..., g_{|G|}\}$ avec $|G| < +\infty$.

Chaque gamme $g_j \in G$, $1 \leq j \leq |G|$, est une *séquence* **d'opérations de production** (une opération de production étant une suite *d'opérations élémentaires* exécutées continûment sur la même machine) qui composent le processus de fabrication d'un type de pièce ([CAMP85]) :

$$\forall g_j \in G, \quad g_j = (O_{j,w})_{1 \leq w \leq |g_j|} = ((m, t_{j,w}))_{1 \leq w \leq |g_j|}, \quad |g_j| < +\infty$$

où :

- $|g_j|$ est le nombre d'étapes (*i.e.* d'opérations de production) de la gamme g_j,
- $O_{j,w}$ est la $w^{\text{ième}}$ opération de la gamme j. $O_{j,w}$ est représentée par le couple $(m, t_{j,w})$ avec :
  - $m \in M$ désigne le groupe de la machine requise pour l'opération $O_{j,w}$,
  - $t_{j,w} \in \mathbb{R}^{+*}$ est la durée pour réaliser l'opération $O_{j,w}$.

Par exemple, si $M = \{X, Y, Z\}$, une gamme g_l de G, $1 \leq l \leq |G|$, comportant $|g_l| = 9$ opérations de production est décrite par :

$$g_l = ((X, t_{l,1})(Y, t_{l,2})(X, t_{l,3})...(Z, t_{l,8})(X, t_{l,9})) \text{ où:}$$

- la première opération est effectuée sur une machine du groupe X, et requiert une durée $t_{l,1}$,
- la seconde opération est effectuée sur une machine du groupe Y, et requiert une durée $t_{l,2}$,

  ...
- la neuvième opération est effectuée sur une machine du groupe X, et requiert une durée $t_{l,9}$.

Soit $\mathcal{O}$ l'ensemble des opérations du SSL :

$$\mathcal{O} = \bigcup_{j=1}^{|G|} \left\{ \bigcup_{w=1}^{|g_j|} \{O_{j,w}\} \right\}$$

On partitionne $\mathcal{O}$ par rapport à M de la manière suivante :

$$\mathcal{O} = \bigcup_{m \in M} \mathcal{O}_m \text{ avec } \mathcal{O}_m = \bigcup_{j=1}^{|G|} \left\{ \bigcup_{w=1}^{|g_j|} \{O_{j,w} / O_{j,w} = (m, t_{j,w})\} \right\} \quad (1)$$

En d'autres termes, $\mathcal{O}_m$ est l'ensemble des opérations que réalise une machine du groupe $m \in M$. Cette partition de $\mathcal{O}$ est pertinente quand on se place dans le cadre d'une approche station, comme nous le verrons dans la section 5.3.

$\forall g \in G$, on définit l'ensemble des groupes de machines *distincts* que cette gamme requiert par l'application :

$$\Psi : G \rightarrow \mathcal{P}(M)$$
$$g \mapsto \Psi(g) = \{\text{groupes de machines } \textit{distincts} \text{ qui réalisent } g\}$$

Pour évaluer la flexibilité d'un groupe de machines $m \in M$ relativement à G, on définit l'ensemble des gammes nécessitant le groupe de machines m par l'application :

$$R : M \rightarrow \mathcal{P}(G)$$
$$m \mapsto R(m) = \{g_j \in G \text{ telle que } (\exists w,\ 1 \leq w \leq |g_j|, \text{ et } O_{j,w} \in \mathcal{O}_m)\}$$

m est d'autant plus flexible que l'ensemble $R(m)$ « tend vers » l'ensemble G.

4.2.2. Prise en compte des routages alternatifs

Dans ce paragraphe, on prend en compte le cas où les gammes comportent des routages alternatifs supposés être probabilisés.

$\forall g \in G$, une opération $\tilde{O} \in g$ est une *opération de séparation* si, après $\tilde{O}$, le processus de fabrication d'une pièce de catégorie g peut être réalisé suivant plusieurs *sous-séquences d'opérations*, le choix d'une sous-séquence étant probabilisé. En effet, bien que ces choix soient généralement effectués en fonction de l'état du système (routage dynamique des pièces) au niveau opérationnel, leur prise en compte aux niveaux stratégique et tactique relève davantage d'un fractionnement statique des flux. Comme le souligne [BUZA93, p. 366], on peut utiliser un modèle à RFA pour déterminer les probabilités optimales pour le fractionnement des flux relativement à un critère de performance du système. Ces probabilités peuvent ensuite être utilisées pour guider le contrôle dynamique du routage des pièces dans le système.

Après $\tilde{O}$, plusieurs cas de figures sont envisageables :

- une opération peut être exécutée sur plusieurs groupes de machines, le choix étant exprimé par une probabilité (cas (i) de la figure 1). Il s'agit alors de sous-séquences simples ne comportant qu'une seule opération,
- le processus de fabrication peut être poursuivi suivant plusieurs séquences d'opérations choisies de manière probabiliste. Trois cas de figures sont à considérer :

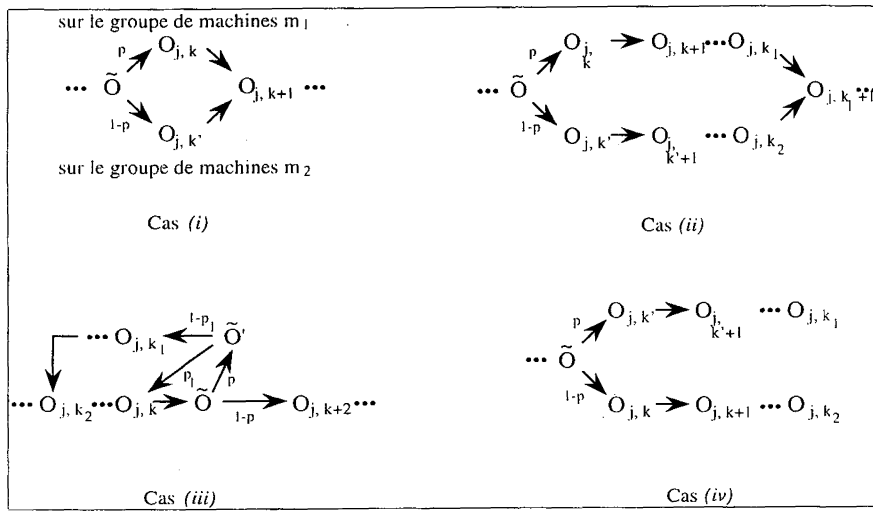

Figure 1. – Exemples de routages alternatifs.

- des sous-séquences d'opérations se rejoignent en aval dans le procédé de fabrication par rapport à $\tilde{O}$ (cas *(ii)*),
- une sous-séquence d'opérations rejoint une opération située en amont de $\tilde{O}$ (cas *(iii)*), ou $\tilde{O}$ elle-même,
- la gamme g se termine par des sous-séquences distinctes d'opérations (cas *(vi)*).

On introduit la notion d'opération fictive pour prendre en compte les deux cas particuliers suivants :

1. si une gamme commence de manière probabiliste par plusieurs sous-séquences d'opérations, on considère une unique opération fictive de séparation en début de gamme notée d, de durée nulle, qui précède chacune de ces sous-séquences,
2. si une gamme se termine de manière probabiliste par des sous-séquences distinctes d'opérations, on considère une unique opération fictive de fin de gamme notée f, de durée nulle, qui succède à chacune de ces sous-séquences.

Pour des commodités de notations, on considère également ces deux opérations fictives quand il n'y a pas ambiguïté sur le début et/ou la fin d'une gamme $g \in G$.

Quelle que soit la gamme $g \in G$ fournie par l'expert industriel, on peut aisément construire le graphe orienté $P_g = (X_g, A_g)$ de précédence des opérations de g où :

$$X_g = \bigcup_{o \in \mathcal{O}} \{o \text{ telle que } o \text{ est une opération (éventuellement fictive) de } g\}$$

$$\forall (o, o') \in X_g^2,\ (o, o') \in A_g \iff o \text{ précède } o' \text{ dans } g$$

Ce graphe est fini, connexe et possède la propriété suivante (par définition d'une gamme) :

$$\forall o \in X_g, \quad \exists \text{ un chemin entre } o \text{ et l'opération de fin de } g$$

Un algorithme polynômial de numérotation des opérations est donné dans [ALM96a]. Pour les graphes considérés, cet algorithme est bien-fondé, converge et numérote tous les sommets du graphe P_g ([ALM96a]).

Définition de la liste principale d'opérations

Pour spécifier aisément les différentes sous-séquences qui peuvent apparaître, on introduit la notion de liste principale d'opérations. $\forall g_j \in G$, la *liste principale d'opérations* est *le* plus court chemin (critère de choix déterministe) [AHU93], au sens « nombre d'opérations du chemin », entre les opérations de début et de fin de la gamme. On note k_j le nombre d'opérations réelles (*i.e.* non fictives) de la liste principale d'opérations associée à g_j.

Sur la figure 2, on montre la liste principale (en gras) ainsi construite. Cette liste comporte $k_j = 15$ opérations.

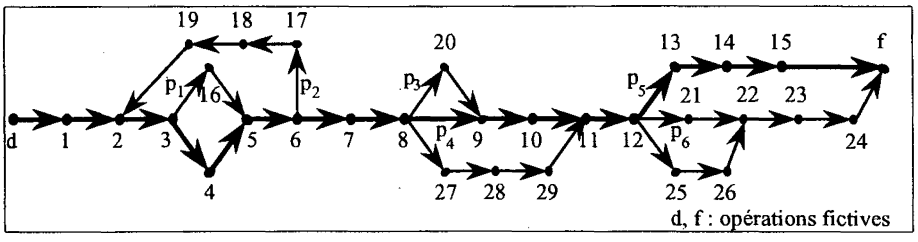

Figure 2. – Exemple de gamme : liste principale et numérotation des opérations.

Spécification des sous-séquences

$\forall g_j \in G$, une sous-séquence alternative d'opérations S_p dans g_j est formalisée par :

$$S_p = (p, (O_{j,w_1}, \ldots, O_{j,w_1+n-1}), \tilde{w_p})$$

où :

- p est la probabilité d'exécution de S_p. p est une donnée relative à l'opération de séparation engendrant S_p,
- $(O_{j,w_1}, \ldots, O_{j,w_1+n-1})$ est la séquence des opérations composant S_p numérotées dans le processus de numérotation (n désignant le nombre d'opérations de S_p),
- $\tilde{w_p}$ est le numéro de l'opération qui succède à la *dernière* opération de S_p. Par convention, on posera $\tilde{w_p} = 0$ pour une séquence alternative de fin de gamme (il n'y a pas d'opération de rebouclage).

Remarque : La séquence $(O_{j,w_1}, \ldots, O_{j,w_1+n-1})$ peut elle-même contenir des opérations de séparation, auquel cas on applique récursivement (le graphe étant fini, il y a arrêt) le principe de spécification aux sous-séquences engendrées par les opérations de séparation de S_p.

Sur l'exemple de la figure 2, la séquence alternative comprenant les opérations 17, 18 et 19 est spécifiée par :

$$(p_2, (O_{j,17}, O_{j,18}, O_{j,19}), 2)$$

tandis que la séquence terminale (21,22,23,24) s'exprime :

$$(p_6, (O_{j,21}, O_{j,22}, O_{j,23}, O_{j,24}), 0)$$

Formalisation d'une gamme

$\forall g_j \in G$, g_j est formalisée par une liste d'opérations construite à partir de la liste principale des opérations. Puis, après chaque opération de séparation $\tilde{O}$ de cette liste, on insère les sous-listes qui décrivent les séquences parallèles : $\forall g_j \in G$,

$$\begin{aligned} g_j = &((O_{j,w})_{1 \leq w \leq (s_{j,1}-1)}, \tilde{O}_{j,s_{j,1}}, (S_p)_{p \in P_{j,s_{j,1}}}, \\ &(O_{j,w})_{(s_{j,1}+1) \leq w \leq (s_{j,2}-1)}, \tilde{O}_{j,s_{j,2}}, \\ &(S_p)_{p \in P_{j,s_{j,2}}}, \ldots, \tilde{O}_{j,s_{j,n_j}}, (S_p)_{p \in P_{j,s_{j,n_j}}}, (O_{j,w})_{(s_{j,n_j}+1) \leq w \leq k_j}) \end{aligned}$$

où :

- $s_{j,1}, s_{j,2}, \ldots, s_{j,n_j}$ est l'ensemble des numéros des opérations de séparation de la liste principale de g_j ($n_j \leq k_j$),
- $\left\{\tilde{O}_{j,s_{j,\ell}}\right\}_{1 \leq \ell \leq n_j}$ est l'ensemble des opérations de séparation,
- $\forall 1 \leq \ell \leq n_j$, $P_{j,s_{j,\ell}}$ est l'ensemble des probabilités des séquences parallèles définies à partir de $\tilde{O}_{j,s_{j,\ell}}$. La probabilité de la séquence incluse dans la liste principale est exclue de cet ensemble car cette séquence fait partie de la liste principale d'opérations.

Pour l'exemple de la figure 2, on obtient la liste suivante :

$$\begin{aligned} g_j =&(O_{j,1}, O_{j,2}, \tilde{O}_{j,3}, (p1, (O_{j,16}), 5), O_{j,4}, O_{j,5}, \\ &\tilde{O}_{j,6}, (p_2, (O_{j,17}, O_{j,18}, O_{j,19}), 2), O_{j,7}, \\ &\tilde{O}_{j,8}, (p_3, (O_{j,20}), 9), (1 - p_3 - p_4, (O_{j,27}, O_{j,28}, O_{j,29}), 11), \\ &O_{j,9}, O_{j,10}, O_{j,11}, \tilde{O}_{j,12}, \ (p_6, (O_{j,21}, O_{j,22}, O_{j,23}, O_{j,24}), 0), \\ &(1 - p_5 - p_6, (O_{j,25}, O_{j,26}), 22), O_{j,13}, O_{j,14}, O_{j,15}) \end{aligned}$$

4.3. Aspects quantitatifs du sous-système logique

Pour évaluer quantitativement les performances d'un système soumis à une charge donnée, il faut prendre en compte les **quotas relatifs** de production associés à chaque type de pièces (un par gamme), c'est-à-dire spécifier le nombre *relatif* de pièces de chaque type qu'il faut produire pendant un horizon temporel fixé et connu. Ici, on suppose que l'horizon temporel est le même pour tous les quotas. Un quota est donc défini par l'application suivante :

$$\begin{aligned} Q : G &\rightarrow \mathbb{R}^+ \\ g &\mapsto Q(g) \end{aligned}$$

4.4. Formalisation du SSL d'un SFP

On formalise donc la charge C d'un SFP de la manière suivante :

$$C = \{(g_1, Q(g_1)), \ (g_2, Q(g_2)), \ \ldots, \ (g_{|G|}, Q(g_{|G|}))\}$$

5. PROPOSITION D'UN CADRE FORMEL POUR L'AGRÉGATION DES DONNÉES

5.1. Introduction

Dans notre étude, le modèle d'action d'un SFP est un RFA fermé dans lequel chaque groupe $m \in M$ de machines est représenté par une station – dite « de type m » – comportant $nmac_m$ serveurs, et à loi de service générale. On désigne, par abus de langage, le groupe de machines OUT comme étant l'extérieur du système.

Une exploitation directe du SSL conduit à considérer $\sum_{m \in \mathcal{M}} |\mathcal{O}_m|$ classes de clients dans le RFA. Une résolution par méthode mathématique peut donc s'avérer impossible pratiquement. En revanche, une agrégation des données réduit considérablement le nombre de classes :

- avec **l'agrégation de type 1** (ou intra-gamme), le RFA comporte $|G|$ classes de clients,

- avec **l'agrégation de type 2** (ou inter-gammes), on agrège *toutes* les classes dues à un ensemble de gammes $G' \subseteq G$ en une seule classe. L'agrégation est dite *totale* si $G' = G$.

5.2. Illustration et justification empirique de l'agrégation de données

5.2.1. Notations

Soient les notations suivantes :

- $\forall m \in M$, on associe un observateur noté Obs_m au groupe m de machines, et qui scrute les $nmac_m$ machines du groupe m,
- $T \in \mathbb{R}^{+*}$ est la période d'observation de Obs_m, $m \in M$,
- Ch_i, $1 \leq i \leq |G|$, est la chaîne de routage associée à la gamme g_i,
- $C_m^{(i)}$ désigne l'ensemble des classes de clients de type i, $1 \leq i \leq |G|$, déduites du SSL, et qui requièrent une machine du groupe m. Ces classes appartiennent toutes à la chaîne de routage Ch_i, et caractérisent des services statistiquement différents sur une machine du groupe m,
- $m.c.t \in \mathbb{R}^{+*}$ (notation pointée) désigne la durée moyenne de service d'un client de classe $c \in \cup_{1 \leq i \leq |G|} C_m^{(i)}$ sur une machine du groupe m,
- $m.c.dest$ désigne la (les) destination(s) possible(s) d'un client de classe $c \in \cup_{1 \leq i \leq |G|} C_m^{(i)}$ en sortie de m, où $m.c.dest$ est :
  - un groupe $m' \in M$ de machines si cette destination est unique,
  - un p-uplet de couples $(dest_j, prob_j)$, $1 \leq j \leq c.p$, si le client a $c.p > 1$ destinations possibles ($0 < c.p \leq |M| + 1$). Dans ce cas, le routage du client est probabilisé, c'est-à-dire que le client de classe c quitte m pour rejoindre le groupe $m.c.dest_j \in M$ de machines, ou bien OUT, avec la probabilité $m.c.prob_j$, $1 \leq j \leq c.p$,
- $m.visite(i, c) \in \mathbb{N}$, $1 \leq i \leq |G|$, $c \in C_m^{(i)}$, désigne le nombre de clients de classe c ayant visité une machine du groupe m durant la période $\tau \in \mathbb{R}^+$, $\forall \tau \leq T$,
- $m.c.destmes_{m'} \in \mathbb{N}$, $m' \in M$, $c \in \cup_{1 \leq i \leq |G|} C_m^{(i)}$, désigne le nombre de clients de classe c qui ont quitté une machine du groupe m pour rejoindre une machine du groupe m' durant la période $\tau \in \mathbb{R}^+$, $\forall \tau \leq T$.

5.2.2. Processus de comptage : définition

Afin de clarifier la présentation de cette section, nous supposons que les durées d'usinage suivent des lois constantes, c'est-à-dire qu'un client de classe $c \in C_m^{(i)}$ visitant une machine du groupe $m \in M$ reçoit un service

de durée *effective* $m.c.t$ (et non pas de durée θ, θ étant une réalisation de la variable aléatoire de moyenne $m.c.t$ distribuée suivant la loi associée à l'opération requise par la visite du client). Nous reviendrons sur cette restriction à la fin de cette section.

$\forall Obs_m$, $m \in M$, on considère le processus de comptage défini par la suite d'instants d'occurrence d'événements notée $(\theta_j^m)_{1 \leq j \leq u}$, $\theta_j^m \in \mathbb{R}^{*+}$, et telle que $\theta_1^m < \theta_2^m < \ldots < \theta_u^m \leq T$. Chaque élément de cette suite est une date de fin de service d'un client sur une machine du groupe m.

A l'instant θ_j^m, $\theta_j^m \leq T$, si un client de classe $c \in C_m^{(i)}$, $1 \leq i \leq |G|$, termine son service de durée m.c.t sur une machine du groupe m, l'observateur :

- incrémente de une unité l'élément $m.visite(i, c)$,
- repère la destination $m' \in M$ où se rend ensuite le client de classe c, et incrémente donc le compteur $m.c.destmes_{m'}$ de une unité.

Obs_m attend ensuite qu'un nouveau client termine son service dans m (instant θ_{j+1}^m) pour réitérer son processus de comptage.

Ces prises de mesures dépendent clairement des proportions relatives de types de pièces circulant dans le système, c'est-à-dire des quotas de production.

5.2.3. Processus de comptage : exploitation

A la fin de la période T, la station de type m ayant été visitée par $m.visite(i, c)$ clients de classe $c \in C_m^{(i)}$, $1 \leq i \leq |G|$, la durée totale de service rendu par cette station pour ces clients de classe c est :

$$m.c.t \times m.visite(i, c)$$

Durant T, la station a travaillé, pour les clients de type i, $1 \leq i \leq |G|$, pendant :

$$s_m^{(i)} = \sum_{c \in C_m^{(i)}} m.c.t \times m.visite(i, c)$$

Plus globalement, et durant T, cette station a donc travaillé au total :

$$s_m = \sum_{i=1}^{|G|} s_m^{(i)} = \sum_{i=1}^{|G|} \sum_{c \in C_m^{(i)}} m.c.t \times m.visite(i, c)$$

En conséquence, pour l'observateur Obs_m, et durant T, la station de type m a assuré, pour un client de type i, un service de durée moyenne :

$$s_m^{(i)} \times 1/\sum_{c \in C_m^{(i)}} m.visite(i,c)$$

et, pour un client de type quelconque, un service de durée moyenne :

$$s_m \times 1/\sum_{i=1}^{|G|} \sum_{c \in C_m^{(i)}} m.visite(i,c)$$

Une démarche identique peut être conduite en ce qui concerne l'étude des transitions entre stations. En effet, durant T, l'observateur Obs_m aura compté $m.c.destmes_{m'}$ transitions entre la station de type m et la station de type m', $m' \in M$, pour les clients de classe $c \in C_m^{(i)}$, $1 \leq i \leq |G|$. En conséquence, durant T, le nombre de transitions entre la station de type m et la station de type m' pour les clients de type i, $1 \leq i \leq |G|$, est :

$$n_{m \to m'}^{(i)} = \sum_{c \in C_m^{(i)}} m.c.destmes_{m'}$$

Globalement, et durant T, le nombre total de transitions entre la station de type m et la station de type m' est donc :

$$n_{m \to m'} = \sum_{i=1}^{|G|} n_{m \to m'}^{(i)}$$

On peut ainsi calculer :

- la fréquence relative observée de transitions entre la station de type m et la station de type m' pour un client de classe $c \in C_m^{(i)}$:

$$\hat{p}_{m,m'}^{(c)} = m.c.destmes_{m'} / \sum_{m'' \in M} m.c.destmes_{m''}$$

Si le système a un état stationnaire, alors, pour une période d'observation T suffisamment longue, la fréquence relative observée $\hat{p}_{m,m'}^{(c)}$ doit tendre vers la probabilité de transitions $m.c.prob_{j(m')}$ entre la station de type m et la station de type m' (principe d'ergodicité), c'est-à-dire :

$$\hat{p}_{m,m'}^{(c)} \longrightarrow m.c.prob_{j(m')} \qquad \text{quand} \qquad T \to +\infty$$

où $j(m')$, $1 \leq j(m') \leq c.p$, est l'index associé à m' dans la liste des stations – groupes de machines – qu'un client de classe c peut visiter après un passage dans la station m,

- la fréquence relative observée de transitions entre la station de type m et la station de type m' pour un client de type i, $1 \leq i \leq |G|$:

$$\hat{p}^{(i)}_{m,m'} = n^{(i)}_{m \to m'} / \sum_{m'' \in M} n^{(i)}_{m \to m''},$$

- la fréquence relative observée de transitions entre la station de type m et la station de type m' :

$$\hat{p}_{m,m'} = n_{m \to m'} / \sum_{m'' \in M} n_{m \to m''}.$$

Toutes les grandeurs mesurées et/ou calculées précédemment sont pertinentes dans la mesure où on peut les relier aux processus d'agrégation intra-gamme et inter-gammes. En effet, $\forall m \in M$:

- $\forall 1 \leq i \leq |G|$, la durée moyenne de service d'un client de type i dans la station de type m et les quantités $\{\hat{p}^{(i)}_{m,m'}\}_{m' \in M}$ sont directement liées aux grandeurs recherchées dans l'agrégation intra-gamme,
- la durée moyenne de service d'un client – quel que soit son type – dans la station de type m et les quantités $\{\hat{p}_{m,m'}\}_{m' \in M}$ sont directement liées aux grandeurs recherchées dans l'agrégation inter-gammes où $G' = G$,
- dans le cas où $G' \subset G$, les calculs précédents s'appliquent en restreignant les indices de sommation aux gammes – et donc aux classes associées – appartenant à G'.

Remarque : Si, pour une classe c de clients, on considère que les durées d'usinage ne sont plus constantes, mais suivent une distribution de probabilité donnée de moyenne $m.c.t$ et de variance $m.c.v$, il faut alors mettre en œuvre un procédé pour répertorier les différentes réalisations de la variable aléatoire associée durant T – on peut utiliser, par exemple, des histogrammes pour compter le nombre de réalisations dans un intervalle de durées d'usinage donné –. *Si le système a un état stationnaire du second ordre*, la moyenne et la variance estimées tendront alors clairement, et respectivement, vers $m.c.t$ et $m.c.v$.

Dans la suite, nous proposons, dans le cas où $G' = G$, une formalisation de l'agrégation de type 2 – et donc du calcul des caractéristiques du RFA – à partir de l'état d'une machine d'un groupe donné active à l'état stationnaire.

Ce concept doit en effet nous permettre d'appréhender formellement et de manière synthétique, à l'état stationnaire, l'état de l'observateur associé, et donc le processus de comptage sous-jacent au processus d'agrégation.

L'agrégation de type 1 et le cas où G' est stritement inclus dans G s'en déduisent de manière immédiate en particularisant (on remplace notamment les ensembles de sommation ou d'appartenance relatifs à G par des ensembles correspondants et relatifs à G').

5.3. Espace des états d'une machine d'un groupe donné à l'état stationnaire

HYPOTHÈSE : A *l'état stationnaire*, le système est traversé par des pièces de type j, $1 \leq j \leq |G|$, suivant les proportions *relatives* $Q(g_j)$.

Définition 1 (État d'une machine à l'état stationnaire)

A l'état stationnaire, une machine du groupe $m \in M$ est soit oisive, soit active. Quand elle est active, elle effectue nécessairement une opération de l'ensemble $\mathcal{O}_m$ (relation (1) *du paragraphe* 4.2.1 *valable dans le cas général).*

Définition 2 (Interaction)

Par analogie avec les systèmes informatiques, on définit une **interaction** *comme étant la prise en charge d'un élément de flux par le système. Pour une pièce d'un type donné, une interaction représente donc le déroulement complet de la gamme de cette pièce.*

$\forall m \in M$, $\forall\, 1 \leq j \leq |G|$, $\forall 1 \leq w \leq |g_j|$, on note $n_{j,w} \in \mathbb{R}^+$ le nombre moyen de fois où l'opération $O_{j,w} = (m, t_{j,w})$ est effectuée *par interaction*.

$\forall 1 \leq j \leq |G|$, les quantités $\{n_{j,w}\}_{1 \leq w \leq |g_j|}$ sont calculées à partir de la matrice $P_o(j)$ des probabilités de transitions entre opérations pour la gamme g_j. Cette matrice est une donnée car elle comporte soit des 1 (opérations consécutives), soit les probabilités de routage alternatif. On résout alors le système linéaire $((|g_j| + 1) \times (|g_j| + 1))$ suivant :

$$N_j = N_j \,.\, \tilde{P}_o(j)$$

où :

- $N_j = (n_{j,OUT}, n_{j,1}, \ldots, n_{j,|g_j|})$. Par définition d'une interaction, on a : $n_{j,OUT} = 1$,
- $\tilde{P}_o(j)$ est obtenue à partir de $P_o(j)$ en rajoutant une première ligne et une première colonne associées à OUT : on y reporte

alors respectivement les probabilités de transitions de OUT vers une opération $O_{j,w}$ (probabilité de début de g_j par $O_{j,w}$) et les probabilités de transitions d'une opération $O_{j,w}$ vers OUT (probabilité d'achèvement de g_j par $O_{j,w}$). Par construction, $\tilde{P}_o(j)$ est une *matrice stochastique.*

On en déduit alors n_m^j le nombre moyen de passages par interaction d'une pièce de type j, $1 \leq j \leq |G|$, sur une machine du groupe $m \in M$:

$$\forall 1 \leq j \leq |G|, \quad \forall m \in M, \quad n_m^j = \sum_{w/O_{j,w} \in \mathcal{O}_m} n_{j,w}.$$

Notons $\mathcal{E}_m$ l'ensemble des états possibles d'une machine du groupe $m \in M$ quand elle est active *à l'état stationnaire du système.* $\mathcal{E}_m$ est déduit de la relation (1) :

$$\mathcal{E}_m = \bigcup_{o \in \mathcal{O}_m} \{s_o\} \text{ où } s_o \text{ est l'état : } \{m \text{ effectue l'opération } o\}.$$

Compte-tenu de l'hypothèse du début de cette section, on attribue un poids relatif $n_{j,w}.Q(g_j)$ à chacun des états $s_{O_{j,w}}$ pour toute opération $O_{j,w} \in \mathcal{O}_m$. Cette remarque permet d'introduire une probabilité $\mathcal{P}$, à l'état stationnaire, sur l'espace discret $\mathcal{E}_m$:

$$\forall O_{j,w} \in \mathcal{O}_m, \mathcal{P}(\{m \text{ effectue l'opération } O_{j,w}\}) = \frac{n_{j,w}.Q(g_j)}{\sum_{j'=1}^{|G|} \left(\sum_{w'=1}^{|g_{j'}|} \mathbb{I}_{\{O_{j',w'} \in \mathcal{O}_m\}}.n_{j',w'}.Q(g_{j'}) \right)}$$

où $\mathbb{I}_{\{O_{j',w'} \in \mathcal{O}_m\}}$ est la fonction indicatrice de l'ensemble $\mathcal{O}_m$.

Cette expression se simplifie de la manière suivante :

$$\forall O_{j,w} \in \mathcal{O}_m, \mathcal{P}(\{m \text{ effectue l'opération } O_{j,w}\}) = \mathcal{P}(\{s_{O_{j,w}}\}) = \frac{n_{j,w}.Q(g_j)}{\sum_{j'=1}^{|G|} n_m^{j'}.Q(g_{j'})}$$

$\mathcal{E}_m$ correspond à un niveau de description fin de l'état d'une machine du groupe m. Il permet de construire d'autres états tels que :

$$\mathcal{B}_m^j = \{\text{une machine du groupe } m \text{ usine une pièce de type } j\} = \bigcup_{O_{j,w} \in \mathcal{O}_m} \{s_{O_{j,w}}\}.$$

$\forall m \in M$, la famille $\{\mathcal{B}_m^j\}_{1 \leq j \leq |G|}$ forme clairement une partition de l'ensemble $\mathcal{E}_m$. Cette remarque est importante pour appliquer, par la suite, le théorème des probabilités totales. On s'intéresse donc maintenant à $\mathcal{P}(\mathcal{B}_m^j)$.

Les états $\{s_{O_{j,w}}\}$ étant disjoints, on obtient :

$$\mathcal{P}(\mathcal{B}_m^j) = \sum_{O_{j,w} \in \mathcal{O}_m} \mathcal{P}(\{s_{O_{j,w}}\})$$

$$= \sum_{w \,/\, O_{j,w} \in \mathcal{O}_m} \left[\left(\sum_{j'=1}^{|G|} n_m^{j'}.Q(g_{j'}) \right)^{-1} .n_{j,w}.Q(g_j) \right].$$

Soit :

$$\mathcal{P}(\mathcal{B}_m^j) = \frac{n_m^j.Q(g_j)}{\sum_{j'=1}^{|G|} n_m^{j'}.Q(g_{j'})}. \tag{2}$$

On définit l'ensemble $\mathcal{T}_m$ des durées moyennes opératoires possibles pour une machine du groupe $m \in M$ à l'état stationnaire du système :

$$\forall m \in M, \quad \mathcal{T}_m = \bigcup_{j=1}^{|G|} \left\{ \bigcup_{w=1}^{|g_j|} \{t_{j,w} / O_{j,w} \in \mathcal{O}_m\} \right\}.$$

Pour chaque groupe $m \in M$ de machines, les ensembles $\mathcal{E}_m$ et $\mathcal{T}_m$ permettent de construire une *variable aléatoire discrète* X_m représentant la durée moyenne d'une opération sur une machine du groupe m :

$$\begin{aligned} X_m : \mathcal{E}_m &\to \mathcal{T}_m \\ s_o = s_{(m,t_{j,w})} &\mapsto X_m(s_o) = t_{j,w} \end{aligned}$$

On a alors :

$$\forall t_m \in \mathcal{T}_m, \quad \mathcal{P}(X_m = t_m) = \sum_{\{o \in \mathcal{O} \,/\, o=(m,t_m)\}} \mathcal{P}(s_o).$$

L'ensemble $\{o \in \mathcal{O} / o = (m, t_m)\}$ est l'ensemble des opérations de $\mathcal{O}_m$ dont la durée est t_m. La décomposition de $\mathcal{O}_m$ donnée dans (1) ainsi que le théorème des probabilités totales nous permettent d'écrire :

$$\forall m \in M, \ \forall t_m \in \mathcal{T}_m, \ \mathcal{P}(X_m = t_m) = \sum_{j=1}^{|G|} \mathcal{P}(X_m = t_m | \mathcal{B}_m^j).\mathcal{P}(\mathcal{B}_m^j) \tag{3}$$

$\mathcal{P}(X_m = t_m|\mathcal{B}_m^j)$ est la probabilité pour qu'une machine du groupe m offre un service de durée moyenne t_m sachant que cette machine sert un client de type j. Il s'agit donc du quotient entre les deux quantités suivantes :

- le nombre moyen de passages de pièces de type j recevant un service de durée t_m sur une machine du groupe m par interaction,
- n_m^j (nombre moyen de passages des pièces de type j sur une machine du groupe m par interaction).

De manière plus formelle, on obtient :

$$\left.\begin{array}{c} \forall m \in M, \quad \forall 1 \leq j \leq |G|, \quad \forall t_m \in \mathcal{T}_m, \\ \mathcal{P}(X_m = t_m|\mathcal{B}_m^j) = \dfrac{\sum_{w/o_{j,w} \in \mathcal{O}_m} \gamma(t_m, t_{j,w}).n_{j,w}}{n_m^j} \end{array}\right\} \quad (4)$$

avec $\gamma : \mathbb{R} \times \mathbb{R} \rightarrow \{0, 1\}$

$$(x, y) \mapsto \gamma(x, y) = \begin{cases} 1 & \text{si } x = y \\ 0 & \text{sinon.} \end{cases}$$

Remarque : Dans le cas de gammes linéaires, la formalisation précédente s'applique et permet de déduire une structure simple de l'espace d'états considéré compte-tenu de la nature des gammes impliquées.

En effet, on a dans ce cas de figure :

$$\forall 1 \leq j \leq |G|, \quad \forall 1 \leq w \leq g_j, \quad n_{j,w} = 1.$$

Cette propriété induit alors des simplifications notables dans toutes les expressions établies dans cette partie.

6. OBTENTION DES CARACTÉRISTIQUES DU RFA AGRÉGÉ MODÉLISANT LE SYSTÈME

L'arrivée d'une pièce sur une machine oisive du groupe m s'accompagne :

* de l'entrée de la machine dans l'état associé à l'opération initialisée par l'arrivée du client. Une durée moyenne opératoire $t_m \in \mathcal{T}_m$ est associée à l'état en question. L'état d'une machine du groupe m est appréhendé par la variable aléatoire X_m définie précédemment,
* de la détermination de la durée de l'opération en réalisant une variable aléatoire Y_{T_m} de loi connue $\mathcal{L}$ – c'est une caractéristique du SSL considéré – et de moyenne t_m liée à l'opération effectuée (et donc à

la classe du client et à l'indice d'avancement dans sa gamme) *i.e.* à l'état du serveur.

Cette description fait clairement apparaître une composition entre deux variables aléatoires : le tirage de la durée moyenne t_m du service, puis la détermination de la durée du service (suivant la loi $\mathcal{L}$ paramétrée par t_m). Ce point est illustré par le schéma suivant :

$$\begin{array}{ccc} \mathcal{E}_m & \xrightarrow{X_m} & \mathcal{T}_m \\ & {\scriptstyle z_m} \searrow & \downarrow {\scriptstyle Y_{t_m}} \\ & & \mathbb{R}^{+*} \end{array}$$

où $\mathcal{E}_m$ et $\mathcal{T}_m$ sont définis dans le paragraphe 5.3.

Pour chaque groupe $m \in M$ de machines, le processus d'agrégation des données conduit à considérer un processus global de service appréhendé par la variable aléatoire Z_m : le service reçu par un client de *classe quelconque* sur une machine du groupe m. L'étude de ce processus nécessite de déterminer sa moyenne et le carré de son coefficient de variation (CCV). En effet, ces deux quantités sont très souvent jugées suffisantes d'un point de vue pratique ([BON86]) pour évaluer à l'état stationnaire les critères de performance d'une file d'attente. Dans le cadre des SFP, on retient donc ces deux indicateurs pour la caractérisation du processus global d'usinage sur une machine ([YAO85a, YAO85b]).

Cette étude nous permet d'obtenir des résultats *paramétrés par la nature de la loi de service $\mathcal{L}$ du SSL considéré*. On peut ainsi aborder le cas où les services suivent une loi générale. Les probabilités de transitions entre les stations dans le RFA agrégé sont ensuite exprimées.

6.1. Calcul de l'espérance et du carré du coefficient de variation

6.1.1. Calcul de $E(Z_m)$

L'expression de $E(Z_m)$ est :

$$E(Z_m) = \sum_{t_m \in \mathcal{T}_m} \mathcal{P}(X_m = t_m).t_m$$

En injectant (3), (4) et (2) dans cette relation, on obtient :

$\forall m \in M$,

$$
\begin{aligned}
E(Z_m) &= \sum_{t_m \in \mathcal{T}_m} t_m . \left[\sum_{j=1}^{|G|} \mathcal{P}(X_m = t_m | \mathcal{B}_m^j) . \mathcal{P}(\mathcal{B}_m^j) \right] \\
&= \sum_{t_m \in \mathcal{T}_m} t_m \\
&\quad \times \left[\sum_{j=1}^{|G|} \frac{\sum_{w \,/\, O_{j,w} \in \mathcal{O}_m} \gamma(t_m, t_{j,w}) . n_{j,w}}{n_m^j} . \frac{Q(g_j) . n_m^j}{\sum_{j'=1}^{|G|} Q(g_{j'}) . n_m^{j'}} \right] \\
&= \frac{1}{\sum_{j'=1}^{|G|} Q(g_{j'}) . n_m^{j'}} . \sum_{t_m \in \mathcal{T}_m} t_m \\
&\quad \times \left[\sum_{j=1}^{|G|} Q(g_j) . \sum_{w \,/\, O_{j,w} \in \mathcal{O}_m} \gamma(t_m, t_{j,w}) . n_{j,w} \right] \\
&= \frac{1}{\sum_{j'=1}^{|G|} Q(g_{j'}) . n_m^{j'}} \\
&\quad \times \sum_{j=1}^{|G|} Q(g_j) . \left[\sum_{t_m \in \mathcal{T}_m} t_m . \sum_{w \,/\, O_{j,w} \in \mathcal{O}_m} \gamma(t_m, t_{j,w}) . n_{j,w} \right]
\end{aligned}
$$

Or, $\sum_{t_m \in \mathcal{T}_m} t_m . \sum_{w \,/\, O_{j,w} \in \mathcal{O}_m} \gamma(t_m, t_{j,w}) . n_{j,w}$ n'est autre que la durée moyenne pour effectuer sur une machine du groupe m *toutes* les opérations nécessaires à la fabrication d'une pièce de type j pendant une interaction. Notons T_m^j ce temps total d'usinage par interaction :

$$
\sum_{t_m \in \mathcal{T}_m} t_m . \sum_{w \,/\, O_{j,w} \in \mathcal{O}_m} \gamma(t_m, t_{j,w}) . n_{j,w} = \sum_{w \,/\, O_{j,w} \in \mathcal{O}_m} n_{j,w} . t_{j,w} = T_m^j .
$$

$\forall m \in M$, notons n_m le nombre moyen de passages d'une pièce de type quelconque sur une machine du groupe m *par interaction* :

$$
n_m = \left(\sum_{j=1}^{|G|} Q(g_j) \right)^{-1} . \sum_{j'=1}^{|G|} Q(g_{j'}) . n_m^{j'} .
$$

D'où l'expression de $E(Z_m)$:

$$
\forall m \in M, \quad E(Z_m) = \frac{\left(\sum_{j=1}^{|G|} Q(g_j) \right)^{-1} . \sum_{j=1}^{|G|} Q(g_j) . T_m^j}{n_m} \qquad (5)
$$

Cette durée moyenne est le quotient entre :

- la durée *moyenne totale* de service reçu par une pièce de type quelconque sur une machine du groupe m par interaction et,
- le nombre moyen de passages d'une pièce de type quelconque sur une machine du groupe m par interaction.

Ce calcul repose donc sur une conservation de la durée totale de service reçu par un client sur une machine d'un groupe donné par interaction.

6.1.2. Calcul de CCV_{Z_m}

Ce calcul est déduit de celui de $E(Z_m^2)$ grâce à la relation suivante :

$$CCV_m = \frac{E(Z_m^2)}{[E(Z_m)]^2} - 1. \tag{6}$$

Or, l'expression de $E(Z_m^2)$ est :

$$E(Z_m^2) = \sum_{t_m \in \mathcal{T}_m} \mathcal{P}(X_m = t_m).E(Y_{t_m}^2).$$

Par ailleurs, et par définition de Y_{t_m}, on a :

$$E(Y_{t_m}^2) = (CCV_{\mathcal{L}} + 1) \times t_m^2.$$

En injectant cette relation, ainsi que (3), (4) et (2), dans l'expression de $E(Z_m^2)$, on obtient : $\forall m \in M$,

$$\begin{aligned} E(Z_m^2) &= \sum_{t_m \in \mathcal{T}_m} (CCV_{\mathcal{L}} + 1).t_m^2.\left[\sum_{j=1}^{|G|} \mathcal{P}(X_m = t_m | \mathcal{B}_m^j).\mathcal{P}(\mathcal{B}_m^j)\right] \\ &= \frac{1}{\sum_{j'=1}^{|G|} Q(g_{j'}).n_m^{j'}} . \sum_{j=1}^{|G|} \left\{ Q(g_j) \right. \\ &\quad \left. \times \left[\sum_{t_m \in \mathcal{T}_m} (CCV_{\mathcal{L}} + 1).t_m^2. \sum_{w \,/\, O_{j,w} \in \mathcal{O}_m} \gamma(t_m, t_{j,w}).n_{j,w} \right] \right\} \end{aligned}$$

D'où : $\forall m \in M$,

$$E(Z_m^2) = \frac{\left\{ \begin{array}{c} (\sum_{j=1}^{|G|} Q(g_j))^{-1} \sum_{j=1}^{|G|} \{Q(g_j).(CCV_{\mathcal{L}} + 1) \\ \times [\sum_{w/O_{j,w} \in \mathcal{O}_m} n_{j,w}.(t_{j,w})^2]\} \end{array} \right\}}{n_m}$$

En injectant cette expression ainsi que (5) dans (6), on obtient l'expression de CCV_{Z_m} :

$$\forall m \in M,\ CCV_{Z_m} = n_m \cdot \frac{\left\{ \begin{array}{l} \sum_{j=1}^{|G|} \{Q(g_j).(CCV_{\mathcal{L}} + 1) \\ \quad \times [\sum_{w/O_{j,w} \in \mathcal{O}_m} n_{j,w}.(t_{j,w})^2] \} \end{array} \right\}}{(\sum_{j=1}^{|G|} Q(g_j))^{-1}.\{\sum_{j=1}^{|G|} Q(g_j).T_m^j\}^2} - 1 \quad (7)$$

6.2. Calcul des probabilités de transitions entre stations

Comme au paragraphe 5.3, on désigne, par abus de langage, le groupe de machine OUT comme étant l'extérieur du système. On rappelle que : $\forall 1 \leq j \leq |G|,\ n_{j,OUT} = 1$.

On pose : $\forall (m,k) \in (M \cup \{OUT\})^2,\ \forall 1 \leq j \leq |G|,\ r_{m,k}^j$ est la probabilité de transitions stationnaire pour qu'une pièce de type j aille, à la fin de son service, d'une machine du groupe m vers une machine du groupe k, ou bien que cette pièce débute (resp. finisse) g_j sur une machine du groupe k (resp. m) si $m = OUT$ (resp. $k = OUT$). On l'obtient par la formule suivante : $\forall (m,k) \in (M \cup \{OUT\})^2,\ \forall 1 \leq j \leq |G|$,

$$r_{m,k}^j = \frac{\text{(nombre moyen de transitions entre } m \text{ et } k \text{ dans } g_j \textit{ par interaction})}{\left\{ \begin{array}{l} \sum_{k' \in M \cup \{OUT\}} \text{(nombre moyen de transitions} \\ \qquad\qquad \text{entre } m \text{ et } k' \text{ dans } g_j \textit{ par interaction}) \end{array} \right\}}.$$

Or, $\forall (m,k) \in (M \cup \{OUT\})^2, \forall 1 \leq j \leq |G|$, le nombre moyen de transitions entre m et k dans la gamme g_j *par interaction* vaut clairement :

$$\begin{array}{ll} \displaystyle\sum_{w\,/\,O_{j,w} \in \mathcal{O}_m} n_{j,w}.\left(\sum_{w'\,/\,O_{j,w'} \in \mathcal{O}_k} (\tilde{P}_o(j))_{w,w'} \right) & \forall (m,k) \in M^2 \\ \displaystyle\sum_{w\,/\,O_{j,w} \in \mathcal{O}_m} n_{j,w}.(\tilde{P}_o(j))_{w,1} & \forall m \in M \text{ et } k = OUT \\ \displaystyle\sum_{w\,/\,O_{j,w} \in \mathcal{O}_k} (\tilde{P}_o(j))_{1,w} & \text{pour } m = OUT \text{ et } \forall k \in M \end{array}$$

On obtient ainsi une expression pour $r_{m,k}^j$.

Notons également $P_{m,k}$ la probabilité de transitions entre les machines des groupes m et k pour une pièce de type quelconque. Par le théorème des probabilités totales, on a :

$$\forall m \in M, \quad \forall k \in M \cup \{OUT\}, \quad P_{m,k} = \sum_{j=1}^{|G|} \mathcal{P}(\mathcal{B}_m^j).r_{m,k}^j$$

En injectant (2) dans cette relation, il vient :

$$P_{m,k} = \sum_{j=1}^{|G|} \left[\frac{Q(g_j).n_m^j}{\sum_{j'=1}^{|G|} Q(g_{j'}).n_m^{j'}} .r_{m,k}^j \right].$$

D'où :

$$\forall m \in M, \quad \forall k \in M \cup \{OUT\}, \quad P_{m,k} = \frac{\sum_{j=1}^{|G|} Q(g_j).n_m^j.r_{m,k}^j}{\sum_{j'=1}^{|G|} Q(g_{j'}).n_m^{j'}}. \quad (8)$$

On a également :

$$\forall k \in M, \quad P_{OUT,k} = \frac{\sum_{j=1}^{|G|} Q(g_j).r_{OUT,k}^j}{\sum_{j'=1}^{|G|} Q(g_{j'})}$$

qui est un cas particulier de la relation (8) car $n_{j,OUT} = 1 \; \forall 1 \leq j \leq |G|$.

7. DISCUSSION

7.1. Respect des quotas de production

Rappelons l'hypothèse faite au début de la section 5.3 :

A l'état stationnaire, le système est traversé par des pièces de type j, $1 \leq j \leq |G|$, *suivant les proportions relatives* $Q(g_j)$.

En d'autres termes, le vecteur des nombres relatifs d'interactions de chaque type de pièces, *à l'état stationnaire*, est proportionnel au vecteur des quotas de production. Cette hypothèse est le fondement majeur du processus d'agrégation, car elle y est exploitée pour définir la mesure de probabilité $\mathcal{P}$ sur l'espace discret $\mathcal{E}_m$ des états d'une machine du groupe $m \in M$ active *à l'état stationnaire*.

Or, cette hypothèse est intimement liée au respect, *à l'état stationnaire*, des quotas de production pour chaque type de pièces fabriqué par le SFP. Pour un SFP donné, l'agrégation des données du SSL de ce SFP n'est donc exacte que si les quotas de production des différents types de pièces produits par ce SFP sont *effectivement* respectés à l'état stationnaire. Ce constat nous amène naturellement à discuter du mécanisme de contrôle des entrées des pièces dans le SFP.

Puisque, dans cet article, nous nous intéressons à des SFP étudiés en saturation, nous considérons donc que le mécanisme de contrôle des entrées

de pièces dans le système est tel que l'admission d'une pièce dans le SFP – début d'une interaction – n'est possible que suite à la fin de la fabrication d'une autre pièce – fin d'une interaction –. Les clients d'un modèle à RFA fermé, noté $\mathcal{M}$, représentent alors des couples (pièces, palettes) ; par conséquent, les chaînes de routage dans $\mathcal{M}$ sont définies par rapport aux types de palettes et au mécanisme de contrôle des entrées des pièces dans le SFP étudié.

Considérons d'abord le cas d'un SFP comprenant N_{Pal} palettes universelles.

On peut alors envisager d'attribuer une palette, qui vient d'être libérée par une pièce de type j, $1 \leq j \leq |G|$, à une pièce de type j. Dans $\mathcal{M}$, ce mécanisme – appelé par la suite mécanisme de contrôle déterministe des entrées des pièces – permet donc de maintenir dans le système un *nombre constant* $ncust_j$ de clients dont la chaîne de routage est celle associée à la gamme g_j. Un tel mécanisme induit donc un modèle à RFA multichaîne. Ce mécanisme ne permet cependant pas, en général, de garantir que les productions des différents types de pièces sont effectuées suivant les quotas de production spécifiés dans le SSL du SFP considéré. En effet, les quotas de production *effectifs* dépendent directement du nombre de clients de chaque type présents *initialement* dans le système (ce point est d'ailleurs clair si $N_{Pal} < |G|$ car, dans ce cas, certains types de pièces ne peuvent être produits puisque l'affectation des palettes a été réalisée une fois pour toutes!). En outre, on peut penser intuitivement que les pièces produites suivant la gamme de durée *moyenne* la plus courte (*i.e.* dont la durée moyenne d'une interaction est la plus courte), auront tendance à circuler plus vite dans le système (*i.e.* interagiront davantage sur le système), au détriment des pièces de types différents qui circulent dans le système.

Comme le souligne [BUZA93, p. 396-397], une alternative consiste à attribuer une palette libre à une pièce dont le type est choisi *aléatoirement* suivant une mesure de probabilité *déduite* des quotas de production du SSL. Plus précisément, et en notant p_j, $1 \leq j \leq |G|$, la probabilité d'attribuer une palette à une pièce de type j, on a :

$$p_j = Q(g_j) / \sum_{j'=1}^{|G|} Q(g_{j'}) \tag{9}$$

Ce mécanisme de contrôle – appelé par la suite mécanisme de contrôle probabiliste des entrées de pièces – permet, à l'état stationnaire, d'assurer une production suivant les quotas de production spécifiés dans le SSL du

SFP. En outre, il autorise l'introduction d'une pièce dans le SFP *sitôt* qu'une palette est libérée par une autre pièce dont la fabrication est terminée. Il constitue donc, par exemple, une approche adéquate pour étudier la capacité de production d'un SFP en fonction du nombre de palettes que comporte ce système. Enfin, d'un point de vue modélisation, nous soulignons que cette approche peut être représentée de la manière suivante : en initialisant le réseau avec autant de clients que de palettes, un client – *i.e.* un couple (pièce, palette) – dont la fabrication est terminée, est *immédiatement* réinjecté dans le système en ayant *éventuellement* changé de classe. Plus précisément, si un client de type j, $1 \leq j \leq |G|$ termine son interaction, il débute *instantanément* une nouvelle interaction en devenant de type j', $1 \leq j' \leq |G|$, avec la probabilité $p_{j'}$ donnée par (9). En d'autres termes, ce processus permet de « casser » les chaînes de routages : on passe donc d'un modèle à RFA $\mathcal{M}$ fermé, multichaîne et multiclasse à un modèle à RFA noté $\mathcal{M}'$ fermé, monochaîne et multiclasse.

Le cas où le SFP étudié comporte des palettes dédiées (les nombres de palettes de chaque type étant alors caractérisés par un vecteur $\vec{N}_{Pal}$) présente, dans une certaine mesure, les mêmes types de problèmes. En effet, il faut de nouveau gérer l'attribution des palettes aux pièces dont le (les) type(s) est (sont) compatible(s) avec celui de la palette en question – mécanismes déterministe ou probabiliste (relativement aux quotas des types de pièces concernés par un type donné de palettes) –. En outre, il est également clair que les nombres relatifs de palettes de chaque type ont une influence sur les quotas des différents types de pièces produits par le SFP. Il faut alors aussi déterminer des nombres de palettes tels que les quotas de production spécifiés dans le SSL du SFP soient respectés. Ce point soulève donc un problème supplémentaire, qui n'est généralement pas trivial à résoudre.

Le tableau I résume les cas où les quotas de production sont respectés, et pour lesquels les caractéristiques agrégées sont donc exactes.

Enfin, notons que les mécanismes de contrôles déterministe et probabiliste présentent également un intérêt indéniable au niveau de leur prise en compte dans des modèles à RFA fermés *solubles par approche stochastique* (*voir* notamment [DAL86c]).

7.2. États fictifs

L'agrégation inter-gammes peut induire, par rapport au système ou à un modèle non agrégé, des états fictifs dans le processus stochastique sous-jacent au modèle à RFA déduit ; ces états sont donc susceptibles de générer

TABLEAU I
Respect des quotas de production. Caractéristiques agrégées exactes.

Mécanisme de contrôle des entrées des pièces	Type de palettes	
	Universelles (N_{Pal})	Dédiées ($\vec{N}_{Pal}$)
Probabiliste	Oui	Oui[1]
Déterministe	Oui[2]	Oui[2]

[1] : si $\vec{N}_{Pal}$ permet de réaliser les $\{Q(g_j)\}_{1\leq j\leq |G|}$
[2] : si les $\{ncust_j\}_{1\leq j\leq |G|}$ permettent de réaliser les $\{Q(g_j)\}_{1\leq j\leq |G|}$ les $\{ncust_j\}_{1\leq j\leq |G|}$ sont également liés à N_{Pal} ou $\vec{N}_{Pal}$

des erreurs sur les valeurs des critères de performance du système, même si les caractéristiques agrégées sont exactes.

A titre d'exemple, considérons un système comprenant un ensemble $M = \{X, Y, Z\}$ de groupes de machines, avec $(nmac_X, nmac_Y, nmac_Z) = (1, 1, 1)$, et deux types de palettes (en nombre N_1 et N_2 respectivement). Soit également l'ensemble de gammes $G = \{g_1, g_2\}$ avec :

$$g_1 = \{(X, 10), (Y, 20), (X, 15)\}$$
$$g_2 = \{(X, 20), (Z, 50), (X, 10)\}$$

dans lesquelles les durées d'usinage suivent des lois constantes. Les palettes de type 1 (resp. de type 2) sont dédiées aux pièces de gamme g_1 (resp. g_2).

N_1 et N_2 sont tels qu'une production suivant des quotas fixés soit possible. Dans ce cas de figure, on adopte donc un mécanisme de contrôle déterministe des entrées des pièces dans le SFP.

Dans un modèle à RFA fermé non agrégé (*i.e.* multichaîne multiclasse), il est alors clair que le nombre maximal de clients que peut comporter la station associée au groupe Y de machines (resp. au groupe Z de machines) est N_1, tous de type 1 (resp. N_2, tous de type 2).

L'obtention d'un modèle de connaissance MC' par application d'une agrégation inter-gammes permet alors de construire un modèle à RFA fermé monochaîne monoclasse, dans lequel les clients partagent $N = N_1 + N_2$ palettes. Le processus stochastique sous-jacent à ce modèle contient des états pour lesquels n clients, $N_1 < n \leq N$ (resp. n, $N_2 < n \leq N$), sont présents dans la station associée au groupe Y de machines (resp. au groupe Z de machines). De tels états sont alors fictifs dans la mesure où ils n'existent pas dans l'espace des états qui caractérise le modèle non agrégé.

Si maintenant le système comprend N_{Pal} palettes *universelles*, on peut de nouveau considérer le mécanisme de contrôle déterministe, pourvu que N_{Pal} permette de réaliser la production souhaitée (en particulier, $|G| \leq N_{Pal}$). Dans ce contexte, une agrégation inter-gammes est de nouveau susceptible d'engendrer, comme précédemment, des états fictifs (ce mécanisme de contrôle, partitionnant les N_{Pal} palettes, est, en effet, équivalent conceptuellement au cas où les palettes sont dédiées).

Par ailleurs, et toujours dans ce cas de figure, on peut déduire un modèle à RFA fermé monochaîne multiclasse intégrant le mécanisme de contrôle probabiliste des entrées des pièces dans le SFP présenté dans la section 7.1. Le modèle agrégé (monochaîne monoclasse) ne présente alors pas d'états fictifs par rapport à ce modèle.

Plus généralement, notons $Pal = \{pal_1, pal_2, \ldots pal_{|Pal|}\}$ l'ensemble des types de palettes appartenant au SSP d'un SFP. Dans cette étude, nous supposons qu'une gamme $g \in G$ ne requiert qu'*un seul* type de palettes dans Pal. On définit alors l'*application surjective* R_{Pal}, de G dans Pal, qui à toute gamme $g \in G$ associe l'unique type de palettes compatible avec les pièces de gamme g. Une telle application induit une partition de G relativement aux types de palettes (classes d'équivalence de R_{Pal}).

On parle alors d'agrégation inter-gammes, intra-palette quand les éléments de $\tilde{G}$ – la partition de G considérée pour l'agrégation inter-gammes – ne contiennent que des gammes appartenant à la même classe d'équivalence de R_{Pal}, c'est-à-dire :

$$\forall \tilde{G}_k \in \tilde{G},\ (g_i, g_j) \in \tilde{G}_k^2,\quad i \neq j, \Longrightarrow\ (R_{Pal}(g_i) = R_{Pal}(g_j)).$$

Dans le cas contraire, on parle d'agrégation inter-gammes, inter-palettes.

Le tableau II résume les cas où des états fictifs peuvent, ou non, être induits par le type d'agrégation choisi. A propos de ce tableau, on constate, en particulier, qu'une agrégation inter-gammes introduit des états fictifs quand les gammes agrégées induisent des chaînes de routage distinctes dans un modèle à RFA. Par exemple, une agrégation inter-gammes inter-palettes introduit des états fictifs par rapport aux stations du RFA, dès que, parmi les gammes agrégées, un groupe de machines qui apparaît dans *au moins* une de ces gammes, n'apparaît pas dans une autre. *Pour les RFA généraux, l'impact de ces états fictifs sur les valeurs des critères de performance est un problème ouvert, car il est difficile d'évaluer a priori leur nombre et la valeur de leur probabilité d'état stationnaire dans le processus stochastique sous-jacent* (voir *la référence sur [ZAH80] en section* 3 *pour les RFA à forme produit*).

TABLEAU II
Cas possibles d'existence d'états fictifs dans le processus stochastique sous-jacent.

Mécanisme de contrôle des entrées des pièces	Palettes universelles	Palettes dédiées	
	Agrégation inter-gammes	Agrégation inter-gammes	
		intra-palette	inter-palettes
Déterministe	Oui* (1)	Oui* (1)	Oui* (1)
Probabiliste	Non (2)	Non (2)	Oui* (1)

* : dépend de l'homogénéité des gammes agrégées
(1) : agrégation inter-chaines
(2) : agrégation intra-chaine

Dans ce contexte, le choix d'une agrégation (quand elle est envisageable – pas de discipline avec priorité suivant les types de pièces... –) nous semble être dicté par l'introduction d'états fictifs et le type de méthode de résolution utilisée en aval. Les approches analytiques (approximatives) existantes (voir par exemple [BAY96]) sont a priori très propices à une agrégation inter-gammes intra-chaîne (agrégation inter-gammes intra-palette avec contrôle probabiliste) – *cf.* tableau II –. Une telle démarche exploite notamment les types de palettes comme critère d'agrégation. Cependant, quand le nombre de chaînes induites est trop important, ou bien quand une approche numérique est requise (prise en compte *exacte* de mécanismes complexes...), une agrégation inter-gammes plus globale – agrégation inter-gammes inter-chaînes – peut s'avérer nécessaire. Il faut alors s'efforcer d'agréger des gammes « homogènes » entre elles pour éviter que l'impact des états fictifs ne biaise trop les résultats. Le problème revient alors à définir et à quantifier le concept d'*homogénéité* de gammes. Notons enfin que les SFP comportent généralement, par nature, des gammes relativement homogènes. *On est donc en droit d'être confiant sur la qualité des résultats fournis par un modèle à RFA exploitant une agrégation inter-gammes inter-palettes.*

8. VALIDATION EXPÉRIMENTALE DU PROCESSUS D'AGRÉGATION

8.1. Étude 1

On considère 3 groupes de machines $M = \{1, 2, 3\}$, avec $(nmac_1, nmac_2, nmac_3) = (1, 1, 1)$, ainsi que les deux gammes montrées par la figure 3, les durées opératoires étant supposées être déterministes.

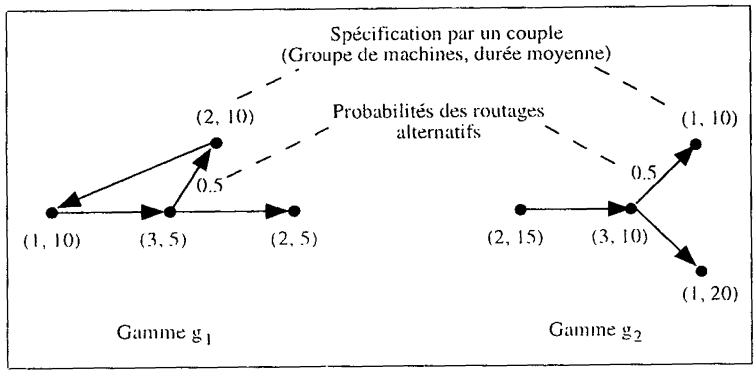

Figure 3. – Exemple de gammes avec routages alternatifs.

Ces deux gammes présentent des routages alternatifs probabilisés. Soit alors le SSL suivant :

$$SSL1 : C_1 = \{(g_1, 1), (g_2, 1)\}$$

Une agrégation intra-gamme pour ce SSL fournit les résultats suivants :

- pour les durées moyennes et les CCV des services dans les 3 groupes de machines :

	Gamme g_1			Gamme g_2		
	Groupe de machines			Groupe de machines		
	1	2	3	1	2	3
Temps moyen	10.0	7.5	5.0	15.0	15.0	10.0
CCV	0.0	0.111	0.0	0.111	0.0	0.0

- pour les probabilités de transitions entre stations :

Gamme g_1				
	OUT	1	2	3
OUT	0	1	0	0
1	0	0	0	1
2	0.5	0.5	0	0
3	0	0	1	0

Gamme g_2				
	OUT	1	2	3
OUT	0	0	1	0
1	1	0	0	0
2	0	0	0	1
3	0	1	0	0

Ces résultats sont conformes à ce que l'on attend intuitivement. Une agrégation inter-gammes donne :

Groupe de machines	Temps moyen	CCV
1	11.667	0.10204
2	10.0	0.16667
3	6.667	0.125

Matrice des probabilités de transition

	OUT	1	2	3
OUT	0	0.5	0.5	0
1	1/3	0	0	2/3
2	1/3	1/3	0	1/3
3	0	1/3	2/3	0

Un autre intérêt de ce SSL réside dans le fait qu'aucun état fictif n'est engendré en effectuant une agrégation inter-gammes, même quand on considère un SFP avec deux types de palettes dédiés respectivement à chaque type de produits. En effet, les deux gammes g_1 et g_2 nécessitent *tous* les groupes de machines de M. Cette propriété est intéressante pour étudier la qualité des résultats fournis par des modèles agrégés comparativement à des modèles qui ne le sont pas.

Dans ce contexte, considérons les 4 configurations de SFP suivantes :

Configuration	SSL	Nombre N_{Pal} de palettes universelles
I	C_1	2
II	C_1	4
III	C_1	6
IV	C_1	8

dans lesquelles les palettes sont universelles, et le respect des quotas de production du SSL est assuré par un mécanisme de contrôle probabiliste des entrées des pièces dans le SFP.

Le tableau III donne, pour chaque configuration, les valeurs des flux des différents groupes de machines – les flux sont identiques pour les trois groupes de machines compte-tenu des gammes et des quotas de production considérés – obtenus en exploitant :

- un modèle de simulation Mod_0 de référence, qui est monochaîne et multiclasse (*i.e.* non agrégé) avec durées de service constantes – résultats avec intervalles de confiance à 95%,
- un modèle Markovien Mod_1, qui est monochaîne et multiclasse (*i.e.* non agrégé) avec durées de service exponentiellement distribuées,

TABLEAU III
Résultats concernant les flux des différents groupes de machines.

Configuration	Flux dans les différents groupe de machines				
	Mod_0	Mod_1	Mod_2	Mod_3	Mod_4
I	0.06153 ±0.00002	0.05152	0.05163	0.05231	0.05863
II	0.08296 ±0.00002	0.06707	0.06730	0.06825	0.07724
III	0.08563 ±0.00001	0.07431	0.07458	0.07538	0.08225
IV	0.08572 ±0.00002	0.07826	0.07853	0.07915	0.08411

- le modèle Markovien Mod_2 exploitant une agrégation intra-gamme, avec des durées de service exponentiellement distribuées,
- le modèle Markovien Mod_3 exploitant une agrégation inter-gammes, avec des durées de service exponentiellement distribuées,
- le modèle Markovien Mod_4 exploitant une agrégation inter-gammes, et dans lequel les durées de service sur les machines des groupes 1, 2 et 3 sont respectivement des lois d'Erlang d'ordre 10, 6 et 8 (on exploite les CCV calculés dans l'agrégation).

L'étude de ce tableau montre clairement l'influence des lois sur les valeurs des critères de performances :

- pour les modèles exploitant des lois de service statistiquement équivalentes – (Mod_0, Mod_4) et (Mod_1, Mod_2, Mod_3) –, les résultats obtenus sont comparables. A ce niveau, il faut souligner que l'exploitation des CCV dans le modèle Markovien Mod_4 a un impact très significatif sur les valeurs des critères de performances estimés (en particulier par rapport à Mod_3). Cette exploitation permet notamment d'appréhender correctement les processus d'usinage reproduits dans le modèle Mod_0,
- on constate des écarts significatifs sur les résultats obtenus par des modèles exploitant des lois de service statistiquement différentes.

Ces résultats sont illustrés sur le graphique de la figure 4. Ils montrent les différentes classes de modèles exhibées par la nature des distributions des durées de service exploitées dans ces modèles. On constate également que l'effet perturbateur des lois diminue lorsque la charge du système augmente – *i.e.* quand on passe de la configuration I à la configuration IV –. Les

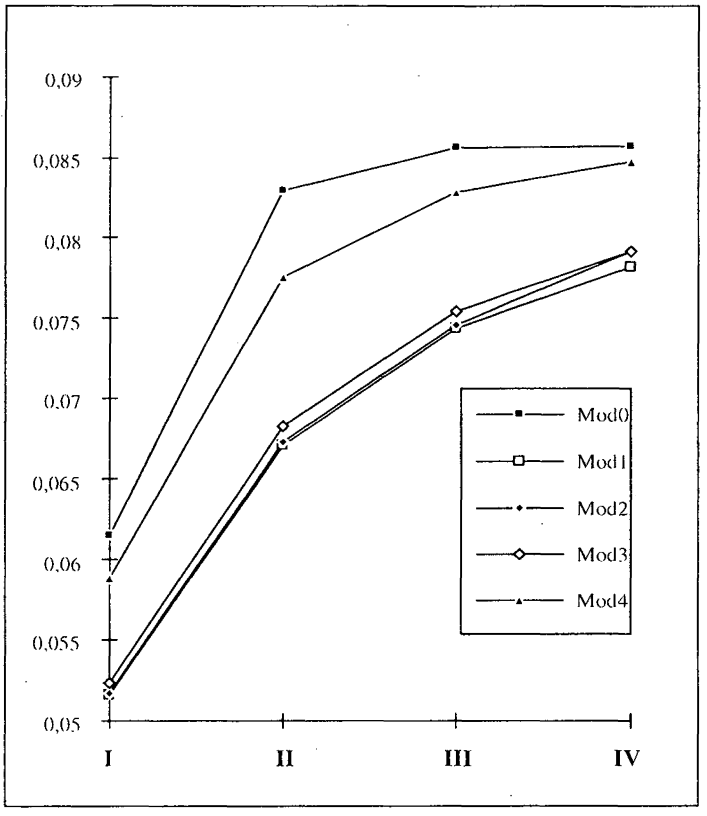

Figure 4. – Flux des différents groupes de machines obtenus avec les différents modèles.

écarts résiduels entre Mod_0 et Mod_4 sont certainement encore dus à la modélisation des lois des durées de service dans Mod_4 : utilisation de distributions erlangiennes, prise en compte seulement des deux premiers moments.

Pour illustrer l'apport des deux types d'agrégations de données, nous donnons, dans le tableau IV, le nombre d'états des différents modèles Markoviens considérés dans cette section. On constate la réduction très significative quand on passe d'un modèle non agrégé – Mod_1 – à des modèles agrégés – partiellement, comme Mod_2, ou totalement, comme Mod_3 –. Cette réduction est d'autant plus importante que le nombre de palettes dans le système (*i.e.* le nombre de clients dans le modèle à RFA) augmente. Ce tableau montre également le coût de la modélisation des lois – dans Mod_4 – par rapport au modèle avec des lois exponentielles – Mod_3 –.

TABLEAU IV
Nombre d'états des différents modèles Markoviens.

Configuration	Nombre d'états dans les modèles			
	Mod_1	Mod_2	Mod_3	Mod_4
I	43	24	6	212
II	793	240	15	2028
III	11191	1792	28	5764
IV	138805	11520	45	11420

8.2. Étude 2

Considérons de nouveau 3 groupes de machines $M = \{1, 2, 3\}$, avec $(nmac_1, nmac_2, nmac_3) = (1, 1, 1)$, ainsi que les 3 gammes suivantes :

$$g_3 = ((1, 30), (2, 50), (1, 70))$$
$$g_4 = ((1, 15), (3, 60), (1, 50))$$
$$g_5 = ((1, 35), (2, 80), (1, 50), (3, 100), (1, 120))$$

On constate que les durées d'usinage sur la machine du groupe 1 présentent une grande variabilité dont l'amplitude dépend globalement des quotas de production des 3 types de pièces. Considérons le SSL suivant :

$$\text{SSL2 : } C_2 = ((g_3, 2), (g_4, 1), (g_5, 1))$$

On constate que, dans le cadre d'un SFP avec palettes dédiées, une agrégation inter-gammes totale de ce SSL génère des états fictifs par rapport aux groupes de machines 2 et 3.

Il est possible de réduire l'influence des états fictifs générés par le processus d'agrégation en effectuant une agrégation inter-gammes partielle (*i.e.* de type 2 avec $G' \subset G$ en reprenant les notations du paragraphe 5.1) : on laisse alors une gamme telle quelle et on agrège les données relatives aux deux autres. L'agrégation de type 1 ne génère pas d'états fictifs.

Nous illustrons ces propos en supposant que les durées d'usinage du SSL sont exponentiellement distribuées, et que le SFP comporte 3 types de palettes (deux palettes de types 1, une de type 2 et une de type 3) dédiés respectivement à chaque type de pièces.

Nous considérons alors un modèle Markovien *non agrégé* c'est-à-dire dans lequel les clients (il y a deux clients dans la chaîne de routage associée à la gamme g_3, un dans celle associée à g_4 et un autre dans celle associée

à g_5) recommencent une interaction dès qu'ils viennent d'en terminer une. Nous soulignons de nouveau qu'un tel mécanisme de contrôle ne permet généralement pas d'atteindre les quotas de production spécifiés dans le SSL2. Aussi, nous mettons en œuvre un processus de calcul *a posteriori* des quotas de production à partir des flux par classe obtenus avec le modèle Markovien non agrégé. Cette manière de procéder *a posteriori* nous évite seulement de rechercher le nombre de palettes dédiées à chaque type de pièces, et qui est nécessaire pour assurer une production suivant les quotas donnés. On calcule alors les quotas *relatifs* de production suivants : 0.55282 pour les pièces de type 3, 0.29761 pour celles de type 4 et 0.14957 pour celles de type 5. Ceci illustre bien les propos de la section 7.1 : on ne satisfait effectivement pas le triplet (2,2,1) de quotas, et la gamme g_5, dont la durée moyenne est la plus longue, est nettement « ralentie » comparativement aux deux autres.

Le tableau V montre les caractéristiques obtenues pour les trois groupes de machines quand on effectue les différentes agrégations de données – intra-gamme, inter-gammes partielle ($G' \neq G$) ou inter-gammes totale ($G' = G$).

TABLEAU V
Caractéristiques des services obtenus par les différentes agrégations.

	Intra-gamme			Inter-gammes			
				partielle			totale
G'	$\{g_3\}$	$\{g_4\}$	$\{g_5\}$	$\{g_3, g_4\}$	$\{g_3, g_5\}$	$\{g_4, g_5\}$	$\{g_3, g_4, g_5\}$
T_1	50.0	32.5	68.33	43.88	55.29	47.90	48.98
CCV_1	1.32	1.58	1.59	1.45	1.49	1.94	1.61
T_2	50.0	–	80.0	50.0	56.39	80.0	56.39
CCV_2	1.00	–	1.0	1.0	1.09	1.0	1.09
T_3	–	60.0	100.0	60.0	100.0	73.38	73.38
CCV_3	–	1.0	1.0	1.0	1.0	1.13	1.13

Quand on effectue ces agrégations, seules les probabilités de transitions en sortie de la station associée au groupe de machines 1 changent. En effet, compte-tenu des gammes considérées, on a toujours :

$$P_{2\to 1} = P_{3\to 1} = P_{OUT\to 1} = 1$$

Le tableau VI donne les probabilités de transitions en sortie de la station associée au groupe de machines 1 obtenues par les différentes agrégations.

TABLEAU VI
Probabilités de transitions en sortie du groupe de machines 1.

G'	$\{g_3\}$	$\{g_4\}$	$\{g_5\}$	$\{g_3, g_4\}$	$\{g_3, g_5\}$	$\{g_4, g_5\}$	$\{g_3, g_4, g_5\}$
$P_{1\to 2}$	0.5	0.0	0.333	0.325	0.452	0.143	0.327
$P_{1\to 3}$	0.0	0.5	0.333	0.175	0.096	0.428	0.208
$P_{1\to OUT}$	0.5	0.5	0.333	0.5	0.452	0.428	0.465

Nous exploitons ensuite toutes ces informations pour construire des modèles Markoviens obtenus :

- par agrégation intra-gamme
- par une agrégation inter-gammes partielle où on agrège deux gammes entre elles en laissant la troisième telle quelle
- par une agrégation inter-gammes totale.

Nous soulignons que les CCV sont pris en compte pour modéliser les lois de service par des coxiennes d'ordre 2 afin de réduire les erreurs de lois. Le choix de modèles Markoviens est justifié ici par le fait que les modèles à RFA étudiés ne sont pas à forme produit.

Le tableau VII montre les résultats concernant les flux et les nombres moyens de clients – lignes $CUSTNB_j$ – pour les trois groupes de machines ; ces résultats sont obtenus avec les modèles Markoviens décrits précédemment.

TABLEAU VII
Flux et nombres moyens de clients obtenus avec les différentes agrégations.

G'	$\times^*$	$\{g_3\}, \{g_4\}, \{g_5\}$	$\{g_3, g_4\}$	$\{g_3, g_5\}$	$\{g_4, g_5\}$	$\{g_3, g_4, g_5\}$
$Flux_1$	0.01968	0.01985	0.01955	0.01966	0.01966	0.01955
$Flux_2$	0.00643	0.00651	0.00639	0.00647	0.00648	0.00639
$Flux_3$	0.00409	0.00410	0.00408	0.00405	0.00403	0.00407
$CUSTNB_1$	3.077	3.101	3.016	3.071	3.058	3.012
$CUSTNB_2$	0.547	0.529	0.564	0.526	0.546	0.560
$CUSTNB_3$	0.376	0.370	0.420	0.403	0.396	0.428
* : aucune agrégation n'est effectuée						

On peut faire plusieurs remarques à partir de ces résultats :

- $\forall m \in \{1, 2, 3\}$, les valeurs de $Flux_m$ obtenus avec les différents modèles sont comparables. Ce point confirme l'idée répandue

selon laquelle une validation sur les flux n'est pas forcément suffisante,

- on constate des écarts notables entre les valeurs des nombres moyens de clients dans la station associée au groupe de machines 3 (ligne $CUSTNB_3$) obtenues avec ces différents modèles. En particulier, les agrégations intra-gamme et inter-gammes partielle avec $G' = \{g_4, g_5\}$ sont les plus proches du modèle non agrégé. Nous pensons que cela est dû au fait que ces deux agrégations *n'engendrent pas*, dans le modèle d'action, d'états fictifs au niveau de ce groupe de machines. En effet, dans ces deux modèles d'action, le nombre maximal de clients qui peuvent être présents dans cette station vaut 3 – la valeur pour le modèle non agrégé –, alors que ce nombre vaut 4 dans les modèles exploitant une autre approche d'agrégation. Par rapport à la station 3, les gammes g_4 et g_5 nous semblent d'ailleurs être, *a priori*, les candidats les plus pertinents dans une agrégation inter-gammes partielle,
- la prise en compte de ces nombres moyens de clients dans les stations est importante si l'on veut étudier les temps de réponse de chaque station du RFA,
- le modèle exploitant une agrégation intra-gamme fournit des résultats par classe très satisfaisants. De manière plus générale, nous pensons, comme [ZAH80], que de meilleurs résultats sont obtenus, pour un type donné de clients, en *n'agrégeant pas* la gamme associée à ce type,
- le modèle Markovien agrégé exploitant une agrégation inter-gammes totale tend à surestimer les nombres moyens de clients présents dans les stations 2 et 3. Nous pensons que les états fictifs engendrés dans ce modèle pour ces deux stations sont responsables de cette surestimation. Il faut également noter que cette surestimation se fait au détriment du nombre moyen de clients présents dans la station 1 (le nombre de clients dans un RFA fermé est constant!).

8.3. Étude 3

Nous présentons maintenant, à titre d'exemple, des résultats obtenus pour un SFP *industriel*. Il s'agit d'une ligne flexible de production de moules de prototypes de pneumatiques de la Manufacture MICHELIN. Le système comporte 3 groupes de machines – $M = \{1, 2, 3\}$ – avec $(nmac_1, nmac_2, nmac_3) = (2, 2, 1)$, un nombre N_{Pal} de palettes universelles et un chariot électrique à rail chargé des opérations de

transport/manutention des pièces au cours de leur usinage. Une description détaillée de ce système est donnée dans [ALM96b]. Ce système fabrique 3 catégories de produits. Les produits de catégories 2 et 3 requièrent la production respective de 8 et 9 types de pièces, tandis que la catégorie 1 ne requiert la production que d'un seul type de pièce. Le tableau VIII décrit la partie logique de ce SSL.

On étudie l'agrégation sur 20 sous-systèmes logiques construits en faisant varier les quotas de production. Ces quotas sont donnés pour chaque catégorie i, $1 \leq i \leq 3$. On en déduit aisément les quotas relatifs par type de pièces (par exemple, un quota q pour la catégorie 2 entraîne un quota de $q/8$ pour les pièces de type k, $2 \leq k \leq 9$). Les valeurs de ces quotas pour chaque SSL étudié sont les suivantes :

SSL	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Catégorie 1	10	10	10	10	10	10	10	10	20	20	20	20	20	20	30	30	30	30	40	40
Catégorie 2	5	10	15	20	25	30	35	40	5	10	15	20	25	30	5	10	15	20	5	10
Catégorie 3	40	35	30	25	20	15	10	5	30	25	20	15	10	5	20	15	10	5	10	5

TABLEAU VIII
Gammes du SSL étudié.

Catégorie 1	
Type 1	((1, 30) (2, 96) (3, 120) (1, 42) (2, 96) (3, 48) (1, 30))
Catégorie 2	
Type 2	((1, 30) (2, 96) (3, 75) (1, 36) (2, 90) (3, 48) (1, 48))
Type 3	((1, 30) (2, 96) (3, 48) (1, 36) (2, 90) (3, 46) (1, 36))
Type 4	((1, 30) (2, 90) (3, 48) (1, 36) (2, 30) (1, 36))
Type 5	((1, 30) (2, 90) (3, 30) (1, 36) (2, 30) (1, 36))
Type 6	((1, 30) (2, 48) (3, 36) (1, 30) (2, 75) (1, 30))
Type 7	((1, 30) (2, 30) (3, 45) (1, 30) (2, 75) (1, 30))
Type 8	((1, 5) (2, 48) (3, 21) (1, 18))
Type 9	((1, 36) (2, 60) (3, 84) (1, 36) (2, 42) (1, 48))
Catégorie 3	
Type 10	((1, 30) (2, 210) (3, 78) (1, 36) (2, 120) (3, 66) (1, 60))
Type 11	((1, 30) (2, 210) (3, 42) (1, 36) (2, 120) (3, 36) (1, 48))
Type 12	((1, 30) (2, 80) (1, 36) (2, 102) (3, 99) (1, 36) (2, 42) (3, 45) (1, 36))
Type 13	((1, 30) (2, 80) (1, 36) (2, 102) (3, 45) (1, 36) (2, 42) (3, 18) (1, 30))
Type 14	((1, 30) (2, 96) (1, 30) (2, 114) (3, 70) (1, 24) (2, 70) (3, 24) (1, 30))
Type 15	((1, 30) (2, 90) (1, 30) (2, 96) (3, 54) (1, 24) (2, 66) (3, 30) (1, 30))
Type 16	((1, 18) (2, 60) (1, 15) (2, 30) (3, 30) (1, 15) (1, 15))
Type 17	((1, 18) (2, 60) (3, 30) (1, 15) (2, 30) (1, 15) (1, 15))
Type 18	((1, 36) (2, 132) (3, 180) (1, 36) (2, 90) (1, 70))

Compte-tenu de certaines contraintes de fonctionnement exprimées dans le SSD de ce SFP (qui induisent notamment des blocages en transfert des machines), l'analyse markovienne offre une alternative pertinente pour modéliser *de manière exacte* ces contraintes. Dans ce contexte, et pour chaque SSL, on compare les résultats obtenus par :

- un modèle de simulation – noté SimDet – non agrégé, avec un mécanisme de contrôle des entrées des pièces probabiliste, et des durées opératoires déterministes,
- un modèle Markovien – noté Mark – exploitant une agrégation inter-gammes totales, et des durées opératoires exponentiellement distribuées,
- un modèle Markovien – noté MarkErl – exploitant une agrégation inter-gammes totales, et des durées opératoires erlangiennes (on utilise les CCV calculés par agrégation),

pour différentes valeurs de N_{Pal} et une durée de transport/manutention constante et égale à 5 unités de temps. Notons que le modèle MarkErl n'est pas exploité pour $N_{Pal} > 4$ en raison de la croissance prohibitive du nombre d'états.

Le tableau IX donne les erreurs relatives entre SimDet et Mark (et entre SimDet et MarkErl suivant les cas) pour le débit des machines du groupe 1 (il s'agit des machines d'entrée/sortie du SFP).

Ces résultats illustrent la qualité du modèle Mark, car celui-ci fournit toujours des résultats pessimistes, avec une erreur relative maximale de l'ordre de 10 %. Par ailleurs, les résultats fournis par MarkErl montrent que les erreurs relatives concernant Mark ne sont pas dues à l'agrégation inter-gammes totales, mais plutôt à la modélisation des lois des durées opératoires (ce point est d'ailleurs corroboré par la diminution des erreurs relatives pour Mark en surcharge – *i.e.* quand N_{Pal} augmente – pour un SSL donné).

Des résultats comparables sont obtenus pour les autres critères de performance. Une étude plus complète de ce SFP industriel est d'ailleurs présentée dans [ALM96a].

Enfin, compte-tenu de la nature des SSL considérés, nous soulignons que l'agrégation inter-gammes est nécessaire ici en vue de l'obtention de modèles à RFA solubles numériquement et/ou par approche analytique approximative ([ALM96a]).

TABLEAU IX

Erreurs relatives (en %) par rapport à SimDet pour le débit des machines du groupe 1.

SSL	Nombre de palettes N_{Pal}						
	3		4		7	9	11
	Mark	MarkErl	Mark	MarkErl	Mark		
1	5.45	0.94	8.26	1.52	9.76	8.56	7.44
2	5.44	0.92	8.13	1.37	9.71	8.39	7.28
3	5.32	0.86	8.24	1.49	9.70	8.39	7.03
4	5.51	1.01	8.47	1.77	9.68	8.41	6.68
5	5.74	1.15	8.60	1.66	9.75	8.17	6.46
6	5.71	0.21	8.81	0.69	9.72	7.90	6.21
7	5.96	0.51	8.99	0.93	9.74	7.70	5.97
8	6.01	0.40	9.35	1.20	9.72	7.34	5.72
9	6.42	1.45	9.32	2.15	7.80	5.66	4.36
10	6.63	1.66	9.10	1.98	7.66	5.23	4.09
11	6.78	0.68	9.19	0.67	7.51	5.03	3.62
12	6.82	0.67	9.28	0.80	7.22	4.60	3.20
13	7.00	0.77	9.31	0.80	6.89	4.25	2.82
14	7.03	0.76	9.25	0.74	6.36	4.02	2.52
15	8.08	1.34	9.57	0.84	5.05	2.99	2.09
16	8.07	1.33	9.65	1.04	4.66	2.56	1.75
17	8.17	1.43	9.58	0.98	4.28	2.18	1.56
18	8.11	1.24	9.47	0.87	3.80	1.94	1.33
19	9.54	1.22	9.51	−0.06	2.59	1.19	0.54
20	9.49	1.15	9.32	−0.11	2.11	0.81	0.32

9. CONCLUSIONS

Dans cet article, nous nous intéressons à l'évaluation des performances à l'état stationnaire de SFP par modèles à RFA, afin d'étudier des problèmes relevant des niveaux stratégique et tactique. La problématique concerne, d'une part, la maîtrise de la complexité d'un SFP pour le recueil fiable des données d'un tel système – en particulier sa charge –, et, d'autre part, l'agrégation de ces données pour obtenir les caractéristiques d'un modèle à RFA effectivement soluble par méthodes mathématiques. Nous nous plaçons dans le cadre d'un processus de modélisation développé au Laboratoire.

Pour le recueil de la connaissance – *i.e.* la construction du modèle de connaissance du système –, nous adoptons une vision systémique en décomposant le système en trois sous-systèmes communicants. Nous proposons ici une formalisation du sous-système logique exploitant une description transactionnelle et naturelle pour les experts du système.

Nous exposons la problématique liée à l'obtention des caractéristiques du RFA modélisant le système. En raison des limitations des techniques de résolutions mathématiques, nous proposons un processus systématique d'agrégation probabiliste des données du SSL à partir de la formalisation de ce sous-système. L'exploitation de cette formalisation permet d'obtenir des modèles agrégés à RFA avec un nombre de classes de clients considérablement réduit.

La formalisation du SSL permet de spécifier un format de saisie pour créer un fichier contenant la description de ce sous-système [ALM96a]. La généricité du processus d'agrégation rend son implantation possible sous la forme d'une procédure de calcul automatique des caractéristiques du RFA modélisant le système à partir des données du SSL préalablement saisies. En outre, grâce à sa rapidité d'exécution, le processus d'agrégation n'est pas contraint par la taille du SSL qui, dans l'industrie, est généralement très volumineuse. Dans notre étude, l'implantation est réalisée avec le langage algorithmique du logiciel d'évaluation de performances QNAP2 (Queueing Network Analysis Package 2) [QNAP91]. En procédant de la sorte, on obtient un processus continu entre la spécification du SSL et l'évaluation des performances proprement dite. Pour un utilisateur, l'obtention des critères de performance de son système est totalement transparente. Les données du système sont exploitées sans interprétation subjective de la connaissance relative au SSL [ALM96a].

La qualité et la robustesse du processus sont testées expérimentalement. Compte-tenu de la qualité des résultats obtenus, nous mettons en avant trois sources d'erreurs concernant les valeurs des critères de performance obtenues consécutivement à la résolution d'un modèle agrégé :

- la nécessité d'introduire des hypothèses simplificatrices – notamment par rapport aux mécanismes de synchronisation – lors de la résolution du modèle à RFA. Ce point souligne l'intérêt des méthodes numériques,
- la caractérisation des lois des durées de service des stations. Dans ce contexte, l'agrégation apporte une réponse par le calcul des CCV de ces lois. L'exploitation de cette information dépend du nombre d'états supplémentaires requis pour l'affinement de la modélisation,
- l'introduction d'états fictifs dans le processus stochastique sous-jacent quand on agrège des classes de clients appartenant à des chaînes de routage différentes. L'exploitation des types de palettes comme critère de regroupement des gammes pour une agrégation inter-gammes nous semble alors être pertinente, notamment dans le cadre d'une résolution

analytique approximative. Le cas échéant, il convient d'agréger des gammes homogènes, ce qui ouvre des perspectives intéressantes quant à la définition d'indicateurs d'homogénéité reposant, par exemple, sur les fonctions Ψ et/ou R.

Notons enfin que ce processus correspond à une *dérivation* – caractérisation d'un nouveau sous-système logique agrégé – du modèle de connaissance initial MC. Le modèle de connaissance MC' ainsi construit peut ensuite être exploité avec l'ensemble des méthodes de résolution disponibles pour les RFA. Cette démarche est d'ailleurs présentée dans [ALM94, ALM96a] et l'utilisation de l'analyse markovienne consécutivement à une agrégation inter-gammes totale d'un SSL de taille industrielle fournit des résultats de qualité très satisfaisante ([ALM96a, ALM96b]).

Remerciements

L'auteur tient à remercier les deux rapporteurs anonymes dont les commentaires ont permis d'améliorer certaines parties de ce papier, ainsi que Monsieur P. Kellert, maître de conférences à l'Université Blaise Pascal, pour sa contribution à ce travail.

RÉFÉRENCES

[AHU93] R. K. Ahuja, T. L. Magnanti et J. B. Orlin, *Networks Flows*. Prentice-Hall, 1993.

[AJM86] M. Ajmone Marsan, G. Balbo et G. Conte, *Performance models of multiprocessor systems*, The MIT Press, USA, 1986.

[ALM94] D. De Almeida, M. Gourgand et P. Kellert, *Manufacturing system modelling using a markovian analysis*. In *IMSE'94*, Grenoble, France, Décembre 1994, p. 347-355.

[ALM96a] D. De Almeida, *Modélisation par réseaux de files d'attente de systèmes de production*, PhD thesis, LIMOS, Université Blaise Pascal, Clermont Ferrand II, 1996.

[ALM96b] D. De Almeida, M. Gourgand et P. Kellert, *Performance evaluation of a job-shop like with transfer blocking flexible manufacturing system using a queueing network model*, International Journal of Computer Integrated Manufacturing (à paraître), 1996.

[AMAR92] S. Amar, E. Craye et J. C. Gentina, *Une méthode hiérachique de spécification et de prototypage des systèmes de production flexibles*. RAIRO-APII, 1992, *26*, p. 483-514.

[BAY96] B. Baynat et Y. Dallery, *A product-form approximation method for general closed queueing networks with several classes of customers*, Performance Evaluation, 1996, *24*, p. 165-188.

[BCMP75] F. Baskett, K. M. Chandy, R. R. Muntz et F. Palacios-Gomez, *Open, closed, and mixed networks of queues with different classes of customers*, Journal of the ACM, April 1975, *22*, (2), p. 335-381.

[BON86] A. B. Bondi et W. Whitt, *The influence of service-time variability in a closed network of queues*, Performance Evaluation, 1986, *6*, p. 219-234.

[BROW88] E. Brown, *IBM combines rapid modeling technique and simulation to design pcb factory-of-the-future*, Industrial Engineering, 1988, *20*, (6), p. 23-26.

[BUZ73] J. P. Buzen, *Computational algorithms for closed queueing networks with exponential servers*. Communications of the ACM, *16*, (9), 1973, p. 527-531.

[BUZA86] J. A. Buzacott et D. D. Yao, *Flexible manufacturing systems: a review of analytical models*, Management Science, 1986, *32*, (7), p. 890-905.

[BUZA93] J. A. Buzacott et J. G. Shanthikumar, *Stochastic models of manufacturing systems*, Prentice-Hall, Englewood Cliffs, New Jersey, 1993.

[CAMP85] J. P. Campagne, J. Peyron et M. Temani, *Structuration des bases de données techniques autour de nomenclatures et gammes-mères en vue de l'élaboration automatique de gammes*, APII, 1985, *19*, (4), p. 343-357.

[DAL86a] Y. Dallery et R. David, *Operational analysis of multiclass queueing networks*, In *25th IEEE Conference on Decision and Control*, 1986, p. 1728-1732.

[DAL86c] Y. Dallery, *On modeling flexible manufacturing systems using closed queueing networks*, Large Scale System, 1986, *11*, (2), p. 109-119.

[DAL92] Y. Dallery et S. B. Gershwin, *Manufacturing flow line systems: a review of models and analytical results*, Queneing Systems, 1992, *12*, p. 3-94.

[DEN78] P. J. Denning et J. P. Buzen, *The operational analysis of queueing network models*, ACM Computing Surveys, 1978, *10*, (3), p. 225-261.

[DIE91] B. L. Dietrich, *A taxonomy of discrete manufacturing systems*, Operations Research, 1991, *39*, (6), p. 886-902.

[DIJ89] N. M. Van Dijk, *Comment on Yao and Buzacott's "modeling a class of flexible manufacturing systems with reversible routing"*. Operations Research, 1989, *37*, (5), p. 845-846.

[GOUR84] M. Gourgand, *Outils logiciels pour l'évaluation des performances des systèmes informatiques*, Thèse d'état, Université Blaise Pascal (Clermont II), 1984.

[GOUR92] M. Gourgand et P. Kellert, *An object-oriented methodology for manufacturing system modelling*, In *Proc. of the* 1992 *Summer Computer Simulation Conference, Reno, USA*, 1992, p. 1123-1128.

[HSU93] L. F. Hsu, C. S. Tapiero et C. Lin, *Network of queues modeling in flexible manufacturing systems: a survey*. RAIRO-RO, 1993, *27*, (2), p. 201-248.

[KEL92] P. Kellert, *Définition et mise en œuvre d'une méthodologie orientée objets pour la modélisation des systèmes de production*, In *Actes du congrès INFORSID* 92, 1992, p. 415-436.

[MACC93] B. L. MacCarthy et J. Liu. *A new classification scheme for flexible manufacturing systems*. International Journal of Production Research, 1993, *31*, (2), p. 299-309.

[MOL81] M. K. Molloy, *On the integration of delay and throughput measures in distributed processing models*, PhD thesis, University of California, Los Angeles, 1981.

[MUR89] T. Murata, *Petri nets: properties, analysis and applications*. Proceedings of the IEEE, 1989, *77*, (4), p. 541-580.

[MVA80] M. Reiser et S. S. Lavenberg, *Mean-value analysis of closed multichain queuing networks*, Journal of the ACM, 1980, *27*, (2), p. 313-322.

[MVAQ84] R. Suri et R. R. Hildebrant, *Modelling flexible manufacturing systems using mean-value analysis*, Journal of Manufacturing Systems, 1984, *3*, (1), p. 27-38.

[PROTH92] J. M. Proth, *Conception et gestion des systèmes de production*, PUF, Paris, 1992.

[PUJ80] G. Pujolle, *Réseaux de files d'attente à forme produit*, Rapport Technique INRIA, 1980.

[QNAP91] *QNAP2 version 8, manuel de référence*, Société SIMULOG, 1991.

[SHA93] S. M. Shafer et D. F. Rogers, *Similarity and distance measures for cellular manufacturing. part i. a survey*, International Journal of Production Research, 1993, *31*, (5), p. 1133-1142.

[SOL77] James J. Solberg, *A mathematical model of computerized manufacturing systems*, In *4th International Conference on Production Research*, Tokyo, Japan, 1977.

[SOL80] J. J. Solberg, *CAN-Q user's guide*, School of Industrial Engineering, Purdue University, West Lafayette, Indiana, 1980.

[STE83] K. E. Stecke, *Formulation and solution of nonlinear integer production planning problems for flexible manufacturing systems*. Management Science, 1983, *29*, (3), p. 273-288.

[STE84] K. E. Stecke, *Design, planning, scheduling and control problems of flexible manufacturing systems*, In *Proceedings of the First ORSA/TIMS Special Interest Conference on Flexible Manufacturing Systems*, 1984.

[SUR83] R. Suri, *Robustness of queuing network formulas*, Journal of the ACM, 1983, *30*, (3), p. 564-594.

[SUR89] R. Suri, *Lead time reduction through rapid modeling*. Manufacturing Systems, p. 66-68, Juillet 1989.

[SUR95a] R. Suri, G. W. Diehl, S. De Treville et M. J. Tomsicek, *From CAN-Q to MPX: evolution of queuing software for manufacturing*, Interfaces, 1995, *25*, (5), p. 128-150.

[SUR95b] R. Suri et R. Desiraju, *Performance analysis of flexible manufacturing systems with a single discrete material handling device*, International Journal of Flexible Manufacturing Systems (à paraître), 1995.

[VALA90] K. P. Valavanis, *On the hierarchical modeling analysis and simulation of flexible manufacturing systems with extended petri nets*, IEEE Transactions on Systems, Man, and Cybernetics, 1990, *20*, (1), p. 94-110.

[YAO85a] D. D. Yao et J. A. Buzacott, *Queueing models for a flexible machining station. part i: The diffusion approximation*, EJOR, 1985, *19*, p. 233-240.

[YAO85b] D. D. Yao et J. A. Buzacott, *Queueing models for a flexible machining station. Part ii: The method of coxian phases*, EJOR, 1985, *19*, p. 241-252.

[ZAH80] J. Zahorjan, *The approximate solution of large queueing network models*, PhD thesis, Department of Computer Science, University of Toronto, 1980.

[ZHOU93] M. C. Zhou et F. Di Cesare, *Petri nets synthesis for discret event control of manufacturing systems*, Kluwer Academics Publishers, 1993.