

Modèle d'affectation de la demande journalière dans une grille de trains

par Albert AURIGNAC
et Maryvonne BENATTAR

Département Planification et Recherche Economique de la SNCF.

I - SITUATION ET OBJECTIFS

L'un des problèmes de l'exploitation des chemins de fer et des transports publics de voyageurs en général, est la mise en place d'une offre de transport discontinue dans le temps qui réponde le mieux possible aux souhaits de la clientèle (effective et potentielle), tout en étant réalisée au moindre coût.

Sur une relation entre deux agglomérations le problème présente déjà des difficultés; mais il se complique en exploitation ferroviaire où la programmation des horaires et de la capacité de chaque train sur un axe doit tenir compte des exigences des clientèles de plusieurs courants de trafic, exigences bien souvent non compatibles. Le placement horaire des trains se trouve alors être un compromis entre des besoins divergents de plusieurs catégories de clientèles ainsi que des impératifs techniques et économiques.

Habituellement, les transports publics procèdent à l'amélioration des grilles d'offre par des méthodes empiriques: chaque

nouvelle grille est une retouche plus ou moins profonde de la grille existante, effectuée après examen de l'occupation des différents trains et qui se traduit par le dédoublement, le déplacement ou la suppression de certains sillons.

Dans certains cas, lorsque se pose un problème particulier de placement horaire, il peut être utile de procéder à une enquête auprès de la clientèle, dans le but de connaître directement ses desiderata en matière d'horaires et lui faire exprimer sa préférence sur des projets de grilles.

Pour l'étude d'un projet de desserte à moyen ou long terme, une seule retouche, même substantielle, d'une grille existante ne suffit pas pour obtenir une bonne adéquation de l'offre à la demande. Les deux méthodes citées ci-dessus se révélant, en général, peu adaptées pour l'étude de tels projets (pour des raisons de coûts ou de délais, notamment lorsque le projet comporte un grand nombre de variantes à étudier), le développement de méthodes d'étude plus systématiques est apparu nécessaire.

Cet article présente une méthode d'analyse de l'interaction offre-demande en transport de voyageurs à moyenne et longue distances. Elle a pour objet de déterminer, pour un parcours origine-destination donné, la répartition théorique de la demande de voyage par fer au cours de la journée (1); la journée étant divisée en intervalles de temps courts, cette répartition est une fonction du temps donnant le pourcentage de voyageurs souhaitant partir (ou arriver (2)) durant chaque intervalle par rapport à la demande totale de la journée.

Si l'offre de transport était quasi continue, c'est à dire les départs très fréquents, et suffisante en capacité à chaque départ, le profil théorique recherché serait exactement celui de la demande constatée dans les trains. Mais, dans les faits, sur les liaisons interurbaines, la fréquence n'est pas très élevée et les voyageurs doivent adapter leur heure de départ souhaitée à la grille des trains qui leur est offerte.

(1) L'étude est limitée au trafic utilisant les trains de jour.

(2) Si l'on étudie la demande selon l'heure d'arrivée à destination au lieu de l'examiner selon l'heure de départ de l'origine.

Le profil que l'on peut effectivement constater sur une relation déterminée est donc différent du profil théorique cherché. Toute la difficulté de l'étude consiste à trouver les liens entre le profil théorique et le profil estimé, c'est à dire à éclairer le mécanisme du choix du train de départ par l'utilisateur; ce dernier a des contraintes personnelles d'horaires de départ ou d'arrivée; d'autre part, il a le choix entre plusieurs trains qui ont évidemment des horaires différents mais qui peuvent également avoir des caractéristiques de prix et de confort différentes. Comme ce mécanisme de choix, certainement très complexe, n'est pratiquement pas connu, on a été conduit à poser diverses hypothèses dont on a vérifié la validité a posteriori en testant la capacité du modèle à reproduire l'occupation effective des trains.

Dès lors que l'on sait formaliser les relations permettant de passer du profil réellement constaté train par train, sur la base de l'offre actuelle, au profil théorique, on est également en mesure de passer du profil théorique obtenu à l'occupation des trains en cas de mise en place de telle ou telle grille d'offre nouvelle. Les utilisations du modèle d'affectation peuvent alors s'envisager dans deux directions:

- ayant établi le profil théorique de demande sur une relation donnée, on peut rechercher la grille de trains qui permet de le suivre au mieux et répond donc au plus près aux souhaits des voyageurs. On peut, à la limite, sélectionner des grilles de trains selon un critère comme, par exemple, celui de la minimisation du coût généralisé moyen pour l'ensemble des voyageurs de la relation, sous telle ou telle contrainte technico-économique;
- à partir d'un couple demande globale journalière-grille de trains offerte, on peut simuler le choix des voyageurs entre les trains de la grille, en tenant compte des contraintes horaires des usagers et des caractéristiques des trains (placement horaire, temps de parcours, suppléments de prix éventuels) et déterminer les capacités les mieux adaptées à la demande pour chacun de ces trains. On a ainsi un moyen d'action sur les coûts d'exploitation au travers des ajustements de capacité.

Il convient de bien préciser, à ce stade, que les outils élaborés dans cette étude ne concernent que les profils de demande et en aucune façon les valeurs absolues de ces demandes. Ainsi, en cas d'adjonction de trains dans une grille horaire, ils permettent de trouver la nouvelle répartition relative du trafic entre les trains, mais ne fournissent pas d'indication sur l'accroissement éventuel du volume de cette demande.

Il est cependant possible d'utiliser simultanément le modèle d'affectation et un modèle d'induction de trafic. On peut dès lors envisager de déterminer de bonnes grilles au sens économique du terme (à défaut des grilles optimales) en intégrant les volumes de trafic, leur répartition dans le temps, les variations de trafic liées à l'amélioration de la qualité du service offert ainsi que les variations des coûts de production correspondants.

On a parlé jusqu'à présent du profil théorique de la demande sur une relation. Il est intuitif qu'un tel profil existe bien pour chaque relation mais l'échantillon de données utilisé pour cette étude n'était pas suffisante pour rechercher une telle répartition théorique relation par relation. On a donc déterminé un nombre limité de groupes rassemblant des relations présentant des profils assez semblables pour pouvoir être résumés en un profil type.

On a procédé au regroupement des relations en se fondant sur le type de relation : PARIS - Province, Province - PARIS et Province - Province, ainsi que sur la distance de la relation. Cette dernière variable rend compte à la fois du temps de parcours et de l'existence ou non de concurrence aérienne.

On précisera enfin que les méthodes présentées ne simulent pas la répartition entre classes. On a supposé, en effet, qu'il existait un profil horaire par classe. Cette hypothèse d'indépendance entre classes n'est pas gênante parce qu'il existe une certaine stabilité des clientèles 1ère et 2ème classes et que, d'autre part, si les grilles 1ère et 2ème classes sont différentes l'un de l'autre, le modèle permet d'estimer les parts du trafic de 2ème classe se reportant vers la 1ère classe (1) par suite des meilleurs temps de trajet et des meilleurs placements des trains n'offrant que des voitures de 1ère classe.

II - FORMALISATION DU MODELE

Les facteurs qui influencent le choix individuel d'un horaire de train sont très vraisemblablement de deux ordres :

- la nature des activités avant et après le déplacement; par exemple, pour un voyage d'affaire, l'heure de début des travaux à destination commandera le choix de l'heure de départ (à noter d'ailleurs que, dans certains cas, c'est l'inverse qui se produit : on fixe l'heure d'une réunion en fonction des horaires de transport les plus convenables). Au contraire, pour des voyages personnels, les contraintes d'horaires sont probablement plus souples,

(1) Usagers habituels de la 2ème classe voyageant dans un train n'offrant que des voitures de 1ère classe

- les caractéristiques des trains : temps de parcours, prix, confort. Il est probable que ces facteurs influencent différemment le comportement des usagers selon le motif du déplacement.

Etant donné la diversité des comportements dans le domaine du choix de l'horaire, on n'a pas cherché à simuler des comportements individuels mais plutôt à décrire, par des lois analytiques et des hypothèses sur le comportement "moyen", ce qui résulte au total des choix exprimés. Malheureusement, les données nécessaires au calibrage du modèle n'ont pas pu, dans cette première phase de développement, être obtenues par motif; il est très vraisemblable cependant qu'une approche par motif devrait améliorer, de façon significative, la précision du modèle.

Considérons une relation R donnée desservie par n trains et les usagers désirant voyager en classe c sur cette relation. Les contraintes horaires et pécuniaires de chaque voyageur seront représentées par son heure désirée de départ (1) et sa valeur du temps. Soit:

\dot{A} : la courbe de répartition de la demande dans la journée, c'est à dire la courbe telle que, pour tout instant t de la journée, $\dot{A}(t)$ est la part (par rapport à la demande totale journalière) des usagers désirant partir à cet instant,

f : la courbe de densité de la distribution de la valeur du temps dans la population des usagers habituels de la classe c sur la relation R.

Pour simuler le choix du voyageur V, on associe à chaque train un temps dit "généralisé" qui mesure la pénibilité du voyage. Le temps généralisé du train i pour le voyageur V tient compte de ses contraintes (t son heure désirée de départ et λ sa valeur du temps) et des caractéristiques du train i. C'est la somme pondérée du temps de trajet principal, du délai entre l'heure de départ souhaitée et l'heure de départ effective du train, dénommé "temps d'attente", ainsi que de l'équivalent en temps du prix du billet, calculé au moyen de la valeur du temps. La formule citée ci-après n'intègre ni pénalisation pour rupture de charge, ni bonification pour supplément de confort (comme par exemple dans les TEE), mais le temps généralisé peut fort bien incorporer de tels paramètres.

Le temps généralisé du train i pour l'utilisateur V caractérisé par le couple (t, λ) est défini par la formule suivante:

$$T_i(t, \lambda) = \theta_i + \xi_i + \delta |t_i - t| + \frac{p_i}{\lambda} \quad (I)$$

(1) Ou d'arrivée à destination.

<u>Notation</u> :	P_i	prix du billet dans le train i ; $P_i = pd + S_i$
	p	produit moyen du voyageur x km (hors suppléments) dans la classe étudiée ou dans l'autre classe si le train ne comporte pas la classe étudiée (1)
	S_i	supplément de prix éventuel dans le train i
	d	distance de la relation
	θ_i	temps de trajet du train i
	ξ_i	temps de parcours terminaux
	t_i	instant de départ du train i
	δ	paramètre de pondération du "temps d'attente" $ t_i - t $ par rapport au temps de trajet principal et au prix.

A noter que le temps de parcours terminal ξ n'a pas été pris en compte, car il n'existait pas dans les enquêtes qui ont permis de calibrer le modèle.

Plus le temps généralisé d'un train est grand par rapport à celui des autres trains, moins l'utilisateur aura "intérêt" à choisir ce train. A partir du temps généralisé on est donc passé à la notion classique d'utilité. Celle-ci a été définie comme une fonction de l'inverse du temps généralisé et on a retenu une formulation simple, la fonction puissance. L'utilité du train i pour l'utilisateur V est donc définie par:

$$U_i(t, \lambda) = \left(\frac{1}{T_i(t, \lambda)} \right)^\alpha \quad (\text{II})$$

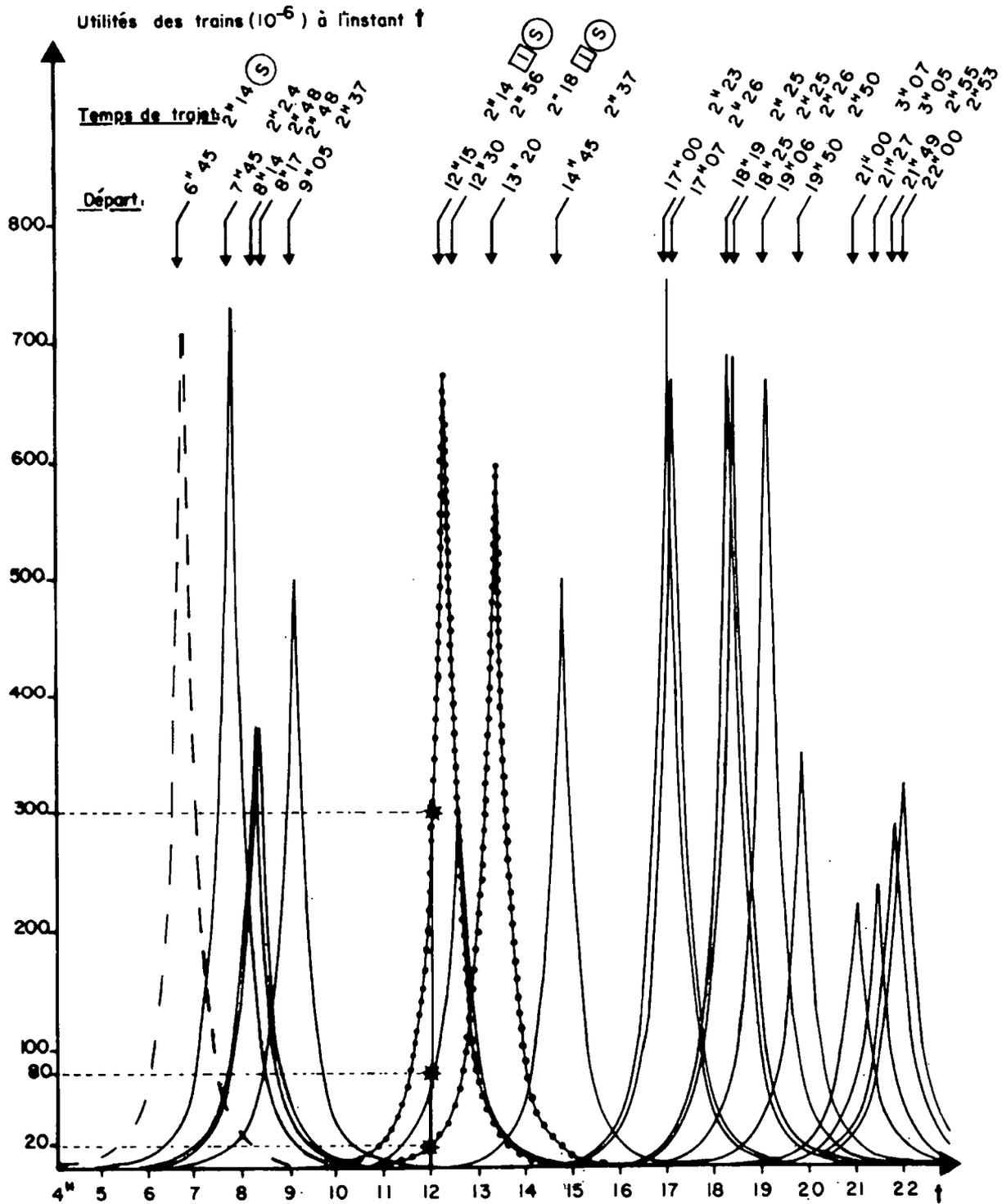
La fonction d'utilité U_i du train i est totalement indépendante du nombre et des caractéristiques des autres trains de la grille. Elle dépend seulement des caractéristiques du train i , de l'heure de départ t souhaitée par l'utilisateur et de sa valeur du temps.

Le graphique n°1 permet de visualiser les 19 fonctions d'utilité des 19 trains offerts aux voyageurs habituels de 1ère classe sur la relation PARIS - DIJON en 1974. La fonction d'utilité d'un

(1) Le temps généralisé d'un voyageur habituel de 2ème classe dans un train uniquement 1ère classe sera nettement supérieur à son temps généralisé en 2ème classe dans un train 1ère classe-2ème classe, à moins que le gain de temps de trajet ou le gain de "temps d'attente" ne viennent compenser le supplément de prix.

COURBES D'UTILITÉ DES TRAINS DE LA GRILLE PARIS-DIJON EN 1974 POUR LES VOYAGEURS DE 1^{ère} CLASSE

Graphique I



train donné étant fonction des deux variables t et λ , on ne pouvait la représenter dans le plan qu'en maintenant l'une des variables constante. Chaque courbe montre comment, pour l'utilisateur de valeur du temps 60F/heure, l'utilité varie en fonction de l'heure de départ souhaitée (1). Sur ce graphique on peut vérifier que :

- l'utilité d'un train est maximale lorsque l'heure souhaitée de départ coïncide avec l'heure de départ du train. Lorsque l'heure souhaitée de départ s'éloigne de l'heure de départ du train, l'utilité de ce train diminue; la courbe décroît symétriquement de chaque côté de l'instant de départ du train.

α est un paramètre de calibrage du modèle; il permet de moduler l'intervalle d'attraction d'un train.

- deux trains qui ont les mêmes caractéristiques de temps de trajet, de prix et de confort ont des courbes d'utilité identiques à une translation près; par exemple:

. les trains de 8 h 14 et 8 h 17, de même les trains de 18 h 19 et 18 h 25.

- par exemple, pour un client de valeur horaire du temps 60 F qui désire partir en 1ère classe à 12 h :

le train 1ère classe de 12h15 a une utilité de 300×10^{-6}

le train 1ère et
2ème classes de 12h30 a une utilité de 80×10^{-6}

le train 1ère classe de 13h20 a une utilité de 20×10^{-6}

les autres trains ont une utilité pratiquement nulle.

Pour passer de la notion d'utilité au choix des trains, deux types de raisonnements étaient envisageables:

- l'un dans lequel on postule que l'utilisateur choisit le train ayant la plus grande utilité pour lui,
- l'autre dans lequel on postule que l'utilisateur choisit en fonction de l'utilité relative de chaque train et donc chaque train a une chance, même faible, d'être choisi.

(1) La i ème courbe est la fonction qui à t associe $U_i(t, 60)$.

La première formule conduit à un modèle par tout ou rien; celui-ci revient, pour une valeur du temps donnée, à découper la journée en n intervalles de temps disjoints, chacun étant l'"intervalle d'influence" de l'un des n trains. Cette partition est effectuée de telle sorte que pour un usager (ayant la valeur du temps donnée) désirant partir durant l'intervalle d'influence du train i , l'utilité la plus grande est celle du train i et, par conséquent, l'usager choisit le train i .

Le graphique n°2 montre le mécanisme très simple de la construction des intervalles d'influence. Il représente, pour chaque train d'une grille, la courbe du temps généralisé du train (repéré en ordonnée) en fonction de l'heure de départ souhaitée par l'usager (repérée en abscisse). Chaque courbe est constituée par deux demi-droites inclinées selon un certain angle x (tel que $\text{tg } x = \delta$) par rapport à l'horizontale et constituant une forme de V. L'ordonnée minimale de la courbe, c'est-à-dire le temps généralisé minimal du train, est atteinte à l'instant de départ du train et c'est le point d'intersection des deux demi-droites.

L'usager choisit, en fonction de son heure de départ souhaitée, le train qui a le temps généralisé minimal; ce choix est donc visualisé par la ligne brisée continue (repérée en trait épais sur le graphique) constituée par les segments de droite d'ordonnée minimale. Mathématiquement :

. si la courbe du train i est la fonction T_i qui à t associe $T_i(t)$

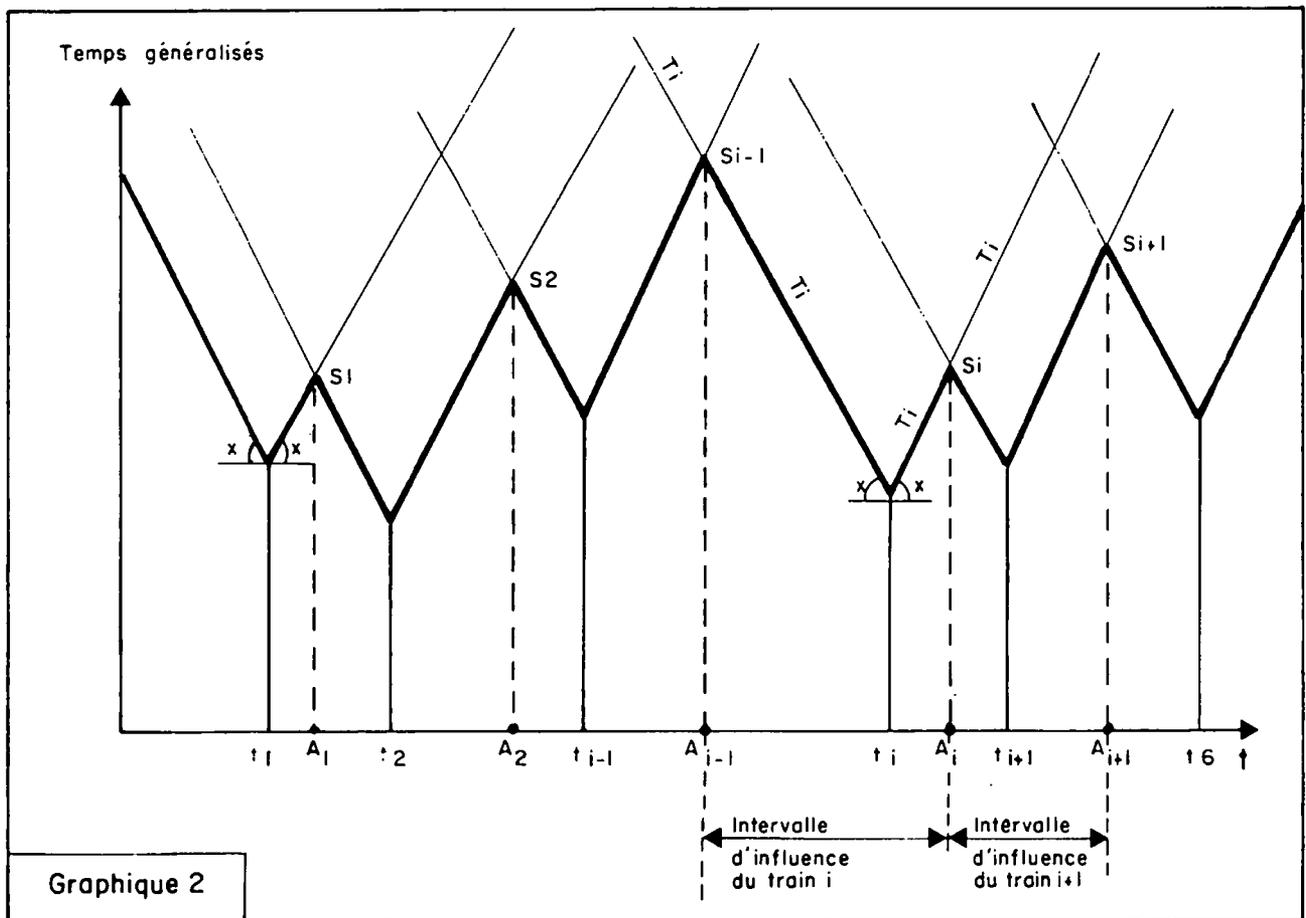
la ligne brisée du choix de l'usager est la fonction qui à t associe $\text{Min}_{i=1 \text{ à } n} T_i(t)$

Cette ligne brisée permet de déterminer immédiatement les intervalles d'influence des trains. L'intervalle d'influence du train i s'étend autour de t_i , heure de départ du train i , et ses extrémités sont les abscisses A_{i-1} et A_i des deux sommets S_{i-1} et S_i de la ligne brisée.

Ainsi, selon ce modèle, tout usager désirant partir durant l'intervalle d'influence du train i sera supposé choisir le train i .

Le premier modèle a été testé mais il n'a pas été retenu car avec la formulation adoptée, il est apparu trop rigide pour représenter la diversité des situations réelles.

En effet, il existe certainement une dispersion importante dans l'appréciation de la valeur du temps entre heure souhaitée de départ et heure effective de départ (paramètre δ). Il serait certainement intéressant de poursuivre la recherche sur le modèle par tout ou rien en introduisant une dispersion dans le paramètre δ .



Graphique 2

C'est donc le deuxième type de modèle qui a été retenu. On a défini la probabilité de choisir un train i comme le rapport de l'utilité de ce train à la somme des utilités de tous les trains. On a donc supposé que la probabilité pour que l'utilisateur V choisisse le train i parmi les n trains était:

$$X_i(t, \lambda) = \frac{U_i(t, \lambda)}{\sum_{j=1}^n U_j(t, \lambda)} \quad (III)$$

La fonction de probabilité de l'utilisateur V (c'est à dire la fonction qui à i associe $X_i(t, \lambda)$) est bien une fonction de probabilité au sens mathématique du terme puisque:

$$\forall t, \forall \lambda \quad \sum_{i=1}^n X_i(t, \lambda) = 1 \quad (IV)$$

Le graphique n°3 est, comme le graphique n°1, relatif à la grille de trains offerte en 1ère classe sur PARIS - DIJON en 1974. Il permet de visualiser, pour chacun des 19 trains, la façon dont, pour l'utilisateur de valeur du temps 60F/heure, sa probabilité de choix du train varie selon son heure de départ souhaitée (1). Sur ce graphique on peut vérifier que:

- toutes les courbes sont dépendantes les unes des autres, car la probabilité de choisir un train dépend des caractéristiques de ce train mais également de celles des autres trains. A chaque instant t , la somme des ordonnées de toutes les courbes est égale à 1 (vérification de l'équation IV);
- la probabilité du choix d'un train est maximale lorsque l'heure souhaitée de départ coïncide avec l'heure de départ du train. Les courbes ne sont pas symétriques;
- deux trains qui ont les mêmes caractéristiques de temps de trajet, de prix et de confort, n'ont pas des courbes de probabilité identiques; tout dépend de leur placement horaire;
- par exemple, pour un client de valeur du temps 60F/heure qui désire partir en 1ère classe à 12 h :

le train 1ère classe de 12h15 a une probabilité d'environ
0,75 d'être choisi

le train 1ère et de 12h30 a une probabilité d'environ
2ème classes 0,20 d'être choisi

le train 1ère classe de 13h20 a une probabilité d'environ
0,05 d'être choisi

$0,75 + 0,20 + 0,05 = 1$; les autres trains ont une probabilité pratiquement nulle d'être choisis.

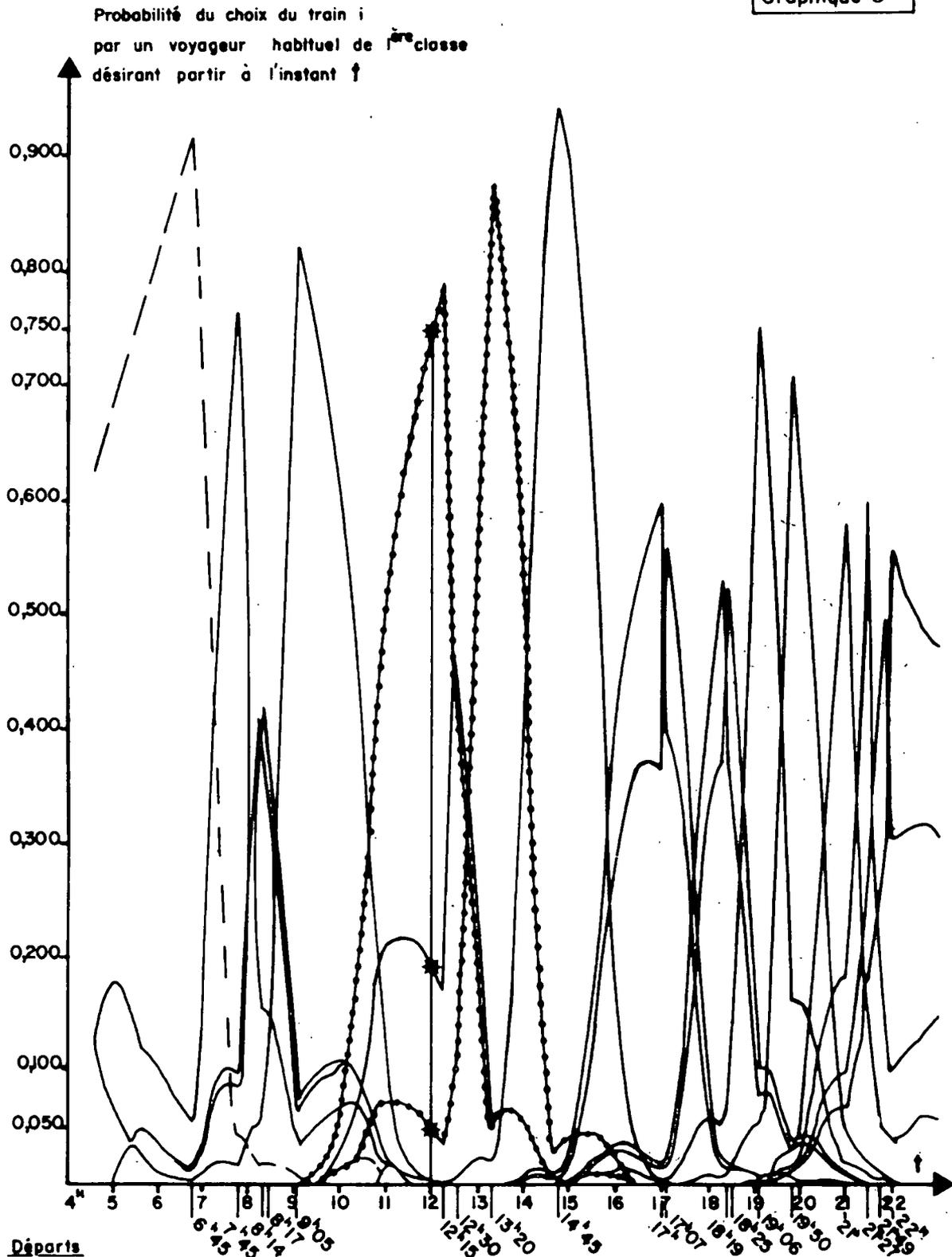
L'enchevêtrement des courbes au niveau des faibles ordonnées provient du fait que le modèle retenu suppose que, dans son choix, le voyageur perçoit l'intérêt de tous les trains, même ceux qui ont une utilité très faible. On peut penser, qu'en fait, l'utilisateur restreint son éventail de choix et qu'il faudrait "filtrer" les probabilités pour éliminer les plus faibles. Etant donné sa faible incidence sur les résultats chiffrés (2), cette correction n'a pas été testée mais

(1) La i ème courbe est la fonction qui à t associe $X_i(t, 60)$

(2) A cause des compensations entre trains.

COURBES DE PROBABILITÉ DES TRAINS DE LA GRILLE PARIS-DIJON EN 1974 POUR LES VOYAGEURS DE 1^{ère} CLASSE

Graphique 3



elle le sera dans une phase ultérieure de développement du modèle.

Par ailleurs, dans son état actuel, le modèle suppose que le voyageur qui choisit de partir dans le train i y trouve effectivement de la place, ce qui est très largement le cas pendant les jours ordinaires, seuls jours pour lesquels le modèle est actuellement complètement utilisable.

Les lois de probabilité étant formalisées, on peut établir le lien entre :

- le trafic emporté par chaque train,
- les choix des voyageurs en fonction de leurs contraintes, décrits par les probabilités X_i ,
- et la distribution temporelle de la demande \dot{A} .

En effet, si f est la fonction de densité de la distribution des valeurs du temps des usagers, et si $\dot{A}(t)$ est la part (1) des voyageurs désirant partir à l'instant t (2), alors:

$\ddot{A}(t, \lambda) = \dot{A}(t) \cdot f(\lambda)$ est la part (1) des voyageurs désirant partir à l'instant t et de valeur du temps λ

$\ddot{A}_i(t, \lambda) = \ddot{A}(t, \lambda) \cdot X_i(t, \lambda)$ est la part (1) des voyageurs désirant partir à l'instant t , de valeur du temps λ et qui choisissent le train i

$A_i = \int_{t=6h}^{t=24h} \int_{\lambda=0}^{\lambda=+\infty} \ddot{A}_i(t, \lambda) \cdot d\lambda \cdot dt$ est la part (1) des voyageurs qui choisissent le train i

$A_i = \int_{t=6h}^{t=24h} \int_{\lambda=0}^{\lambda=+\infty} \dot{A}(t) \cdot f(\lambda) \cdot X_i(t, \lambda) \cdot d\lambda \cdot dt$ (V)

(1) Par rapport à la demande totale journalière.

(2) On a bien évidemment $\int_{6h}^{24h} \dot{A}(t) dt = 1$

Cette équation est à la base de la résolution des deux problèmes énoncés au début de cet article:

- connaissant la part du trafic journalier dans chaque train d'une grille d'offre, déterminer une estimation de la répartition temporelle de la demande \hat{A} ,
- connaissant une approximation de la répartition temporelle de la demande \hat{A} , déterminer les parts de trafic emportées par chacun des trains d'une grille offerte.

Pour les applications ultérieures du modèle, il n'est pas nécessaire d'estimer la loi \hat{A} par une courbe définie pour chaque instant de la journée. Nous l'avons approchée par une courbe en escalier \hat{A} définie sur un certain nombre d'intervalles de temps suffisamment petits pour que l'approximation soit correcte pour des utilisations ultérieures, mais pas trop nombreux pour des raisons d'ajustements statistiques.

La journée est divisée en p intervalles disjoints J_1, J_2, \dots, J_p , appelés intervalles de référence.

\hat{A}_j est la part de voyageurs désirant partir au cours de l'intervalle J_j et la fonction \hat{A} doit respecter l'équation suivante:

$$\sum_{j=1}^P \hat{A}_j = 1 \quad (VI)$$

Les usagers souhaitant partir au cours de la période J_j sont supposés vouloir partir à l'instant t_j , milieu de J_j , et les probabilités de choix de ces voyageurs sont calculées pour ces seuls instants t_j .

La part des voyageurs qui choisissent le train i (équation V) est alors estimée par l'équation suivante:

$$A_i \approx \sum_{j=1}^P \int_0^{\infty} \hat{A}_j \cdot f(\lambda) \cdot X_i(t_j, \lambda) \cdot d\lambda \quad (VII)$$

De même, les probabilités de choix varient très peu pour des usagers dont les valeurs du temps sont très proches. On divise donc l'espace des valeurs du temps en r intervalles (suffisamment petits pour que l'approximation soit correcte) : I_1, I_2, \dots, I_r . La probabilité de choix d'un voyageur ayant une valeur du temps située dans I_k est alors mesurée pour λ_k , milieu de l'intervalle I_k .

Si on note $\hat{f}(I_k)$ la part des voyageurs dont la valeur du temps est située dans I_k (1) alors la part de voyageurs qui choisissent le train i (équation VII) est estimée par:

$$A_i \approx \sum_{j=1}^p \hat{A}_j \cdot \left[\sum_{k=1}^r \hat{f}(I_k) \cdot X_i(t_j, \lambda_k) \right] \quad (\text{VIII})$$

On voit ainsi se dessiner un modèle linéaire. En effet, notons:

$$X_{i,j} = \sum_{k=1}^r \hat{f}(I_k) \cdot X_i(t_j, \lambda_k)$$

$X_{i,j}$ correspond à la moyenne des probabilités de choix du train i pour les usagers désirant partir durant J_j . L'équation VIII devient:

$$A_i \approx \sum_{j=1}^p \hat{A}_j \cdot X_{i,j} \quad (\text{IX})$$

III - CALIBRAGE DU MODELE

Le modèle linéaire décrit par l'équation (IX) permet:

- 1°) Si on dispose d'un fichier de données fournissant les parts de trafic emportées par chacun des trains ayant circulé sur une OD, d'estimer, par la méthode mathématique de régression linéaire multiple sans constante et sous contraintes (coefficients de régression positifs et de somme 1), le profil temporel de la demande sur l'OD,
- 2°) Si on connaît une approximation \hat{A} de la courbe de répartition temporelle de la demande sur une OD, de prévoir la part de trafic qui sera emportée par chacun des trains d'une grille que l'on se propose d'offrir pour desservir l'OD.

Mais dans le premier cas, comme on l'a déjà précisé il était exclu dans cette étude, pour des raisons d'échantillonnage et de qualité des ajustements, de rechercher un profil temporel de la demande pour chacune des origines-destinations desservies par le chemin de fer. Il était donc préférable d'appliquer ce modèle à la recherche d'un profil unique pour un ensemble de relations homogènes. C'est pourquoi le calibrage du modèle n'a été effectué qu'après regroupement des origines-destinations (du fichier à notre disposition) présentant des profils semblables.

(1) $\hat{f}(I_k) = \int_{I_k} f(\lambda) \cdot d\lambda$

a) Base de données et regroupement des relations

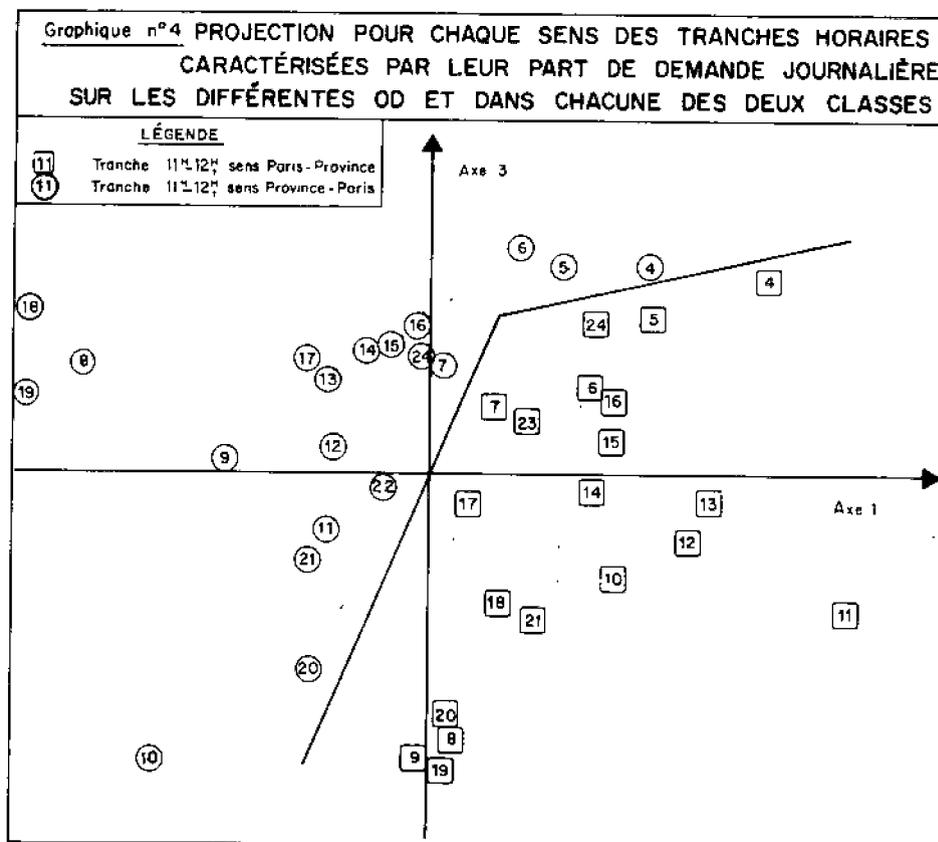
Un système de sondage par cartes prédécoupées dans tous les trains rapides et express a été jusqu'en 1975 l'une des sources de données statistiques sur les voyageurs de la SNCF. Les informations relatives à la période 1972-1973 et au JOB (1) ont été extraites de ce sondage, fournissant ainsi un fichier d'environ 75 grilles horaires (pour chaque classe), soit 1 500 données d'occupations de trains (pour chaque classe).

Une analyse factorielle des correspondances effectuée sur ce fichier a mis très nettement en évidence deux variables de segmentation des origines-destinations selon leur profil temporel de demande: le sens et la distance.

Le graphique n°4 montre très nettement la distinction entre les tranches horaires selon qu'elles sont caractérisées par la part de demande du sens PARIS - Province (carré) ou celle du sens Province-PARIS (cercle).

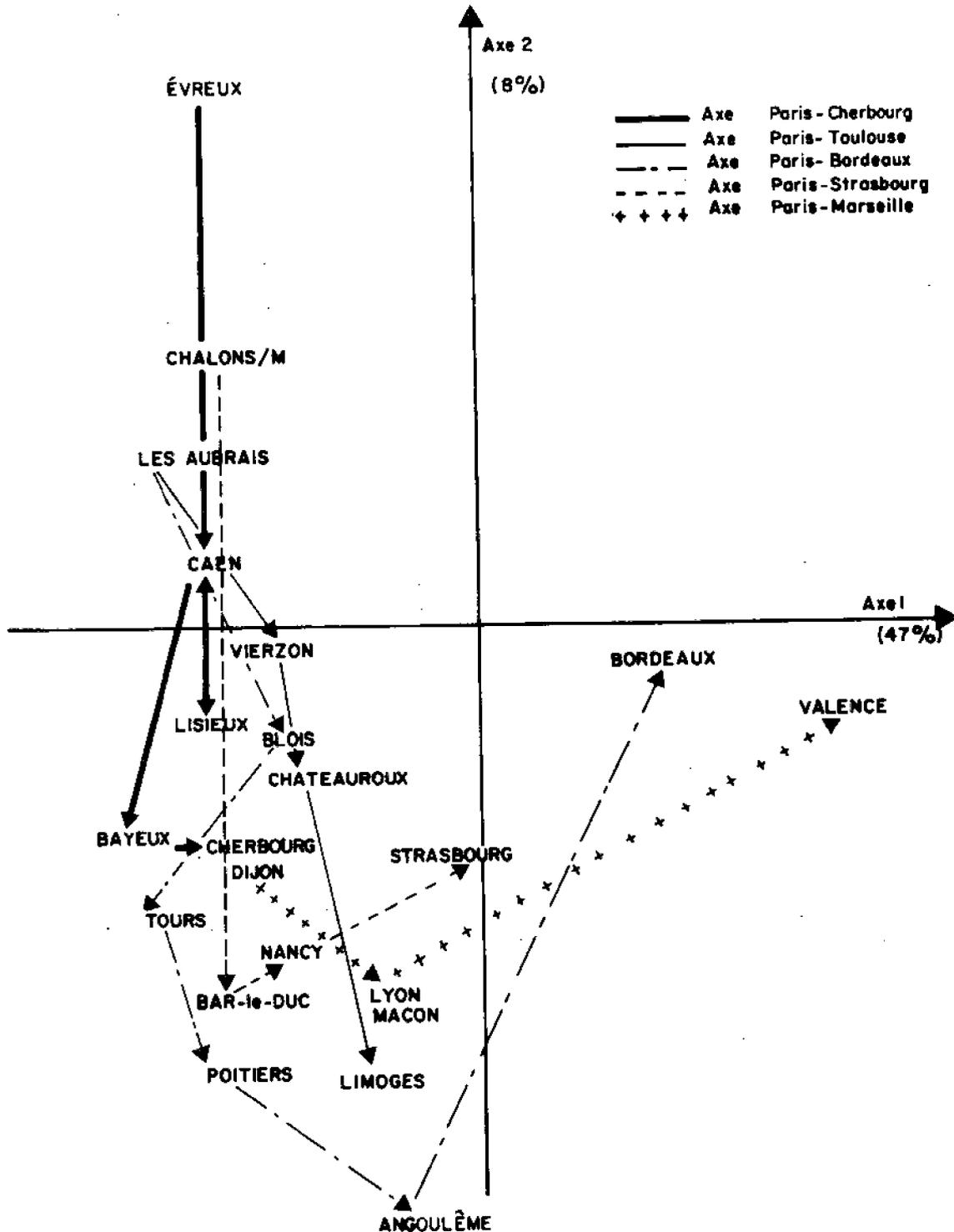
Le graphique n°5 représente les projections de relations PARIS - Province caractérisées par leur profil de demande en 1ère classe. On constate que le deuxième axe factoriel décrit la distance. En effet, hormis les relations comportant du trafic de nuit (PARIS - BORDEAUX, PARIS - VALENCE), plus la distance de la relation est élevée, plus celle-ci se projette vers le bas du 2ème axe factoriel.

(1) Période hors week-end : mardi, mercredi, jeudi



PROJECTIONS DE VILLES CARACTÉRISÉES PAR
LEUR PROFIL DE DEMANDE 1^{ère} CLASSE SUR LEUR
LIAISON AVEC PARIS

Graphique 5



Ces deux variables de segmentation ont permis de déterminer un groupe de relations homogènes sur lequel on a procédé au calibrage du modèle. Ce sous-fichier est constitué d'une vingtaine d'origines-destinations reliant PARIS à des villes de province (dans le sens PARIS - Province) et de distance comprise entre 100 et 350 km. Il fournit environ 300 données d'occupations de trains (pour chaque classe) pour 35 grilles de trains.

b) Les paramètres δ et α

Le calibrage du modèle consiste à déterminer les paramètres δ et α intervenant dans la formulation de l'utilité. Il s'effectue par la technique du balayage; c'est-à-dire qu'on effectue un certain nombre de régressions linéaires selon le modèle décrit par l'équation (IX), chacune correspondant à un couple (δ_i, α_i) de valeurs possibles des paramètres. Le couple qui correspond au meilleur ajustement est alors sensé être celui avec lequel la formulation de l'utilité (équation II) reproduit le mieux l'utilité réellement ressentie par les voyageurs.

c) Le choix des intervalles de référence et la technique des quatre régressions.

Les p intervalles de référence J_j qui partagent la journée sont choisis disjoints et de durée identique appelée "pas". Si la journée est définie par l'intervalle (t_D, t_F) et si le "pas" égale k minutes, l'ensemble des p intervalles de référence constitue "l'échelle horaire de pas k sur (t_D, t_F) ".

Plus le pas des intervalles de référence est petit, plus la courbe qui approxime un profil de demande est proche de ce dernier.

Cependant, si le pas est trop petit, certaines variables $X_{i,j}$ sont très corrélées entre elles et la technique de régression n'est plus applicable.

Il faut donc utiliser un pas optimal qui respecte les deux contraintes. Dans cette étude plusieurs pas ont été testés. Le pas de 60 mn est le pas minimal qui permette d'utiliser la régression.

Le pas de 60 mn donnant une approximation trop grossière du profil de demande, on se ramène à un pas de 15 mn par l'intermédiaire de quatre régressions:

- la 1ère sur l'échelle horaire (6h00 - 24h00) de pas 60 mn
- la 2ème sur l'échelle horaire (5h15 - 23h15) de pas 60mn (1)

(1) Les bornes de l'intégrale de l'équation (V) vont alors de 5h15 à 23h15.

- la 3ème sur l'échelle horaire (5h30-23h30) de pas 60mn (1)
- la 4ème sur l'échelle horaire (5h45-23h45) de pas 60mn (2)

Chaque régression fournit une estimation du profil de demande. La moyenne des quatre courbes fournit alors une estimation beaucoup plus fine du profil de la demande.

d) La valeur du temps

La nécessité de connaître la valeur du temps des usagers peut poser des problèmes car fréquemment on la connaît peu ou même pas du tout.

Si on ne connaît pas la fonction f de densité de la distribution des valeurs du temps dans la population étudiée, mais seulement sa valeur moyenne $\bar{\lambda}$, on peut quand même appliquer le modèle. Les ajustements et les estimations seront sans doute moins bons mais les formules sont plus simples.

$$T_i(t) = \theta_i + \delta |t_i - t| + \frac{P_i}{\lambda} + \xi_i \quad (\text{I})$$

$$A_i = \int_{6h}^{24h} \dot{A}(t) \cdot X_i(t) \cdot dt \quad (\text{V})$$

$$A_i \approx \sum_{j=1}^P \hat{A}_j \cdot X_i(t_j) \quad (\text{IX})$$

Si on ne connaît pas $\bar{\lambda}$, alors la technique de balayage permet, en fonction de la qualité des ajustements, de déterminer une plage de valeurs du temps probables.

(1) Les bornes de l'intégrale de l'équation (V) vont alors de 5h30 à 23h30.

(2) Les bornes de l'intégrale de l'équation (V) vont alors de 5h45 à 23h45.

IV - RESULTATS DU CALIBRAGE ET VALIDATION DU MODELE SUR QUELQUES RELATIONS

a) Résultats du calibrage

a.1) Les paramètres δ et α

Pour chacune des deux classes, les paramètres δ et α ont été testés sur le groupe des origines-destinations de PARIS vers la PROVINCE, de distance comprise entre 100 et 350 km.

On constate pour chacune des deux classes que, quand δ croît de 1 à 2 et que α croît de 1 à 6, les paramètres permettant de juger de la qualité de l'ajustement s'améliorent; au-delà ils ne s'améliorent plus.

C'est donc le couple $\delta = 2$ et $\alpha = 6$ qui a été retenu tant en 1ère classe qu'en 2ème classe.

a.2) La valeur du temps

Ne connaissant pas exactement la moyenne des valeurs du temps des usagers de l'échantillon à notre disposition, nous avons testé trois valeurs moyennes horaires en 2ème classe : 5F, 10F, 15F (1) et cinq valeurs moyennes horaires en 1ère classe : 24F, 30F, 50F, 60F, 100F (1).

La qualité de l'ajustement varie assez peu en fonction de ces valeurs du temps, cependant elle est légèrement supérieure pour 10F/heure en 2ème classe et 100F/heure en 1ère classe.

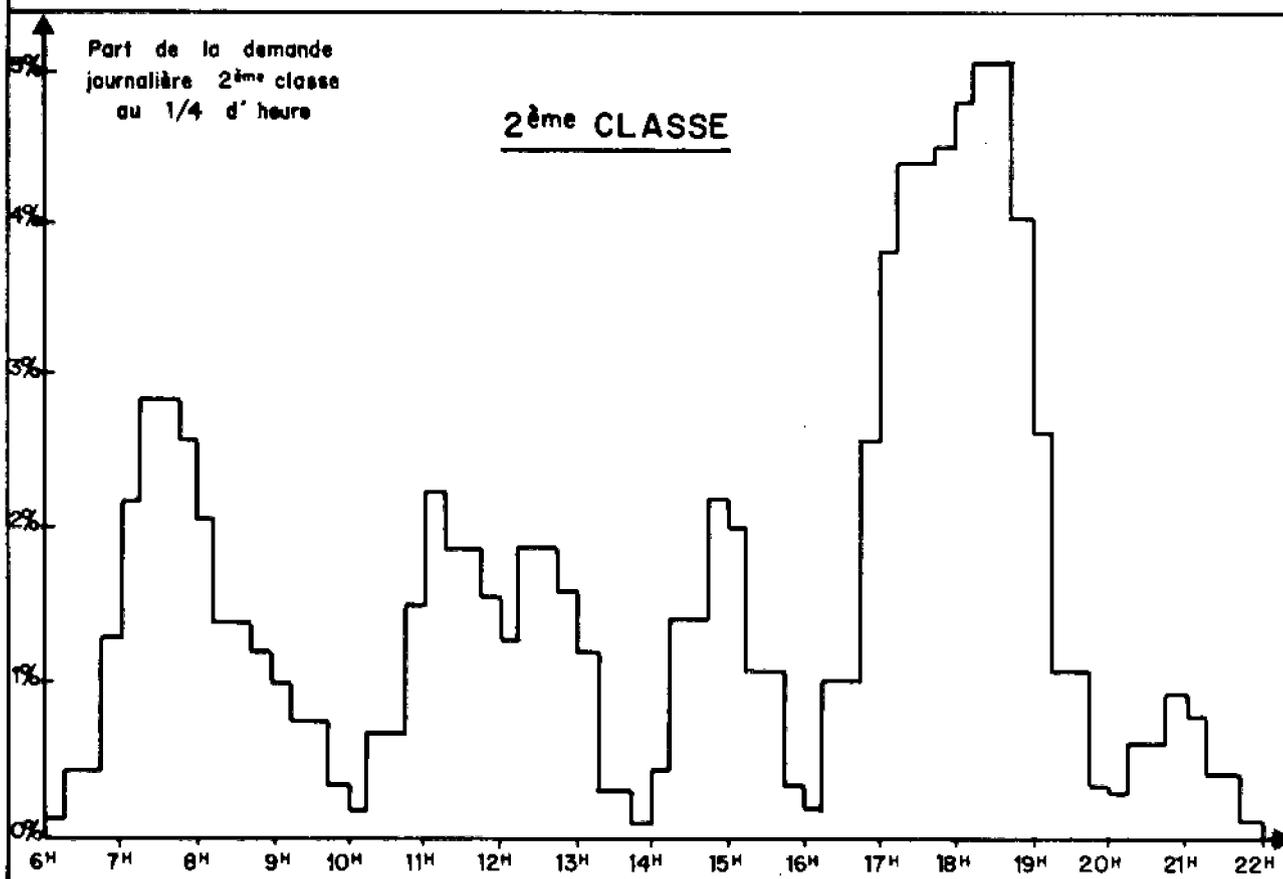
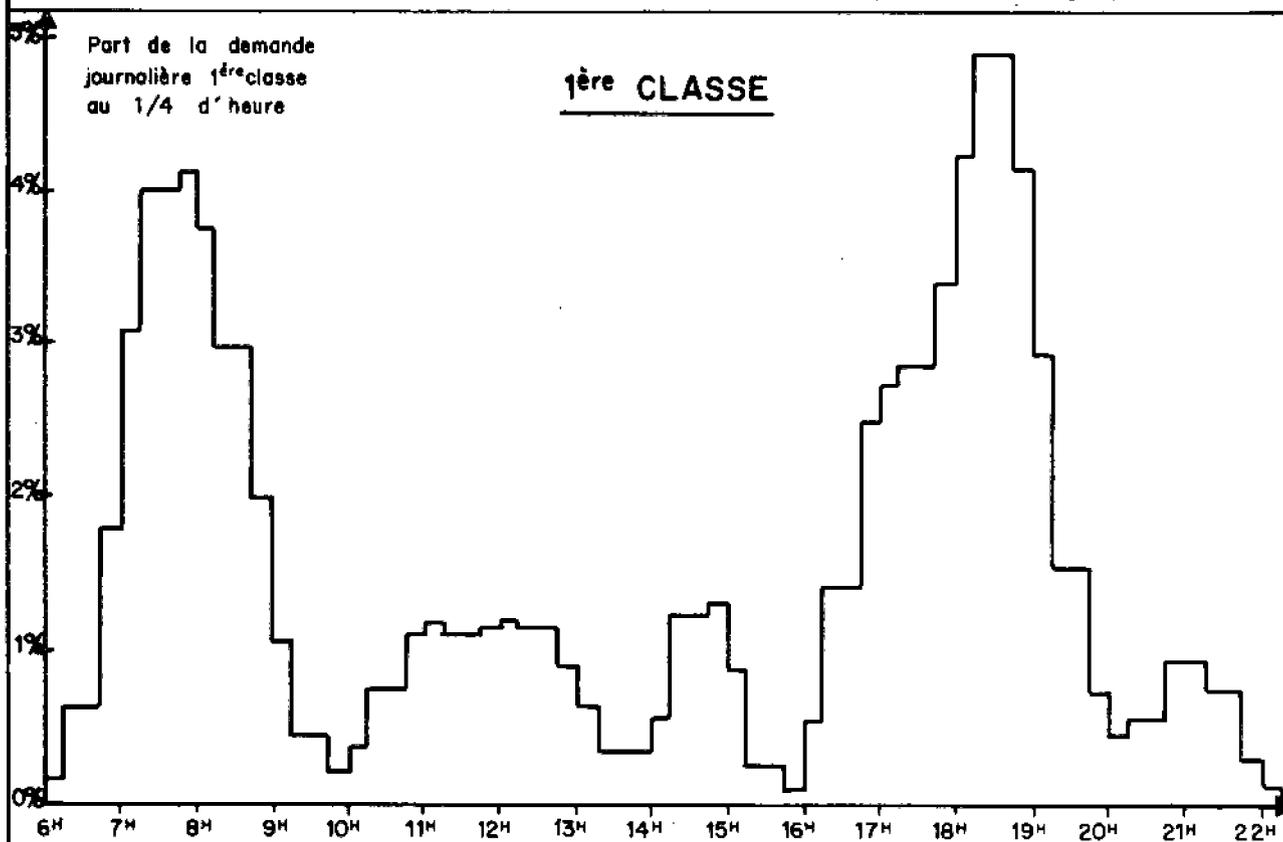
a.3) Profils théoriques de la demande

Des séries de quatre régressions effectuées sur des fichiers de données relatives à certains groupes de relations, ont permis de déterminer les répartitions temporelles probables de la demande sur ces relations.

Le graphique n°6 montre les estimations des profils de demande en JOB en 1972/73 sur les relations de sens PARIS vers PROVINCE et de distance comprise entre 100 et 350 km.

(1) Ces valeurs sont exprimées en Francs 1975

**GRAPHIQUE n°6 ESTIMATION DES PROFILS THÉORIQUES DES DEMANDES
SUR LES RELATIONS PARIS → PROVINCE 100 $d \leq 350\text{km}$**



Le premier schéma visualise le profil de la demande 1ère classe en estimant à 100F/heure (1) la valeur moyenne du temps des usagers de la 1ère classe. Le deuxième schéma représente le profil de la demande 2ème classe en estimant à 10F/heure (1) la valeur moyenne du temps des usagers de 2ème classe. La différence assez nette entre les deux profils s'explique probablement par la nature des motifs de déplacement assez différente selon les classes. La clientèle de 1ère classe, à dominante voyages d'affaires, présente un profil temporel plus accusé que celui de la clientèle de 2ème classe composée, en majorité, de voyages personnels.

L'échantillon étudié est constitué par 25 grilles de trains.

Les estimations sont de bonne qualité; en effet la part de variance expliquée est de 92% en 1ère classe et de 90% en 2ème classe.

Le modèle de recherche d'un profil de demande ainsi calibré s'ajuste donc de façon très satisfaisante sur l'échantillon. On peut penser que les hypothèses formulées sur le mécanisme du choix des usagers sont correctes.

b) Validation du modèle sur quelques relations

Pour tester la validité du modèle de choix des voyageurs, on a, sur certaines grilles qui ont réellement été en service et pour lesquelles on connaît la part de demande réellement emportée par chaque train, comparé la répartition observée à la répartition probable estimée par le modèle.

Les grilles étudiées étaient en service durant l'hiver 1974 -1975.

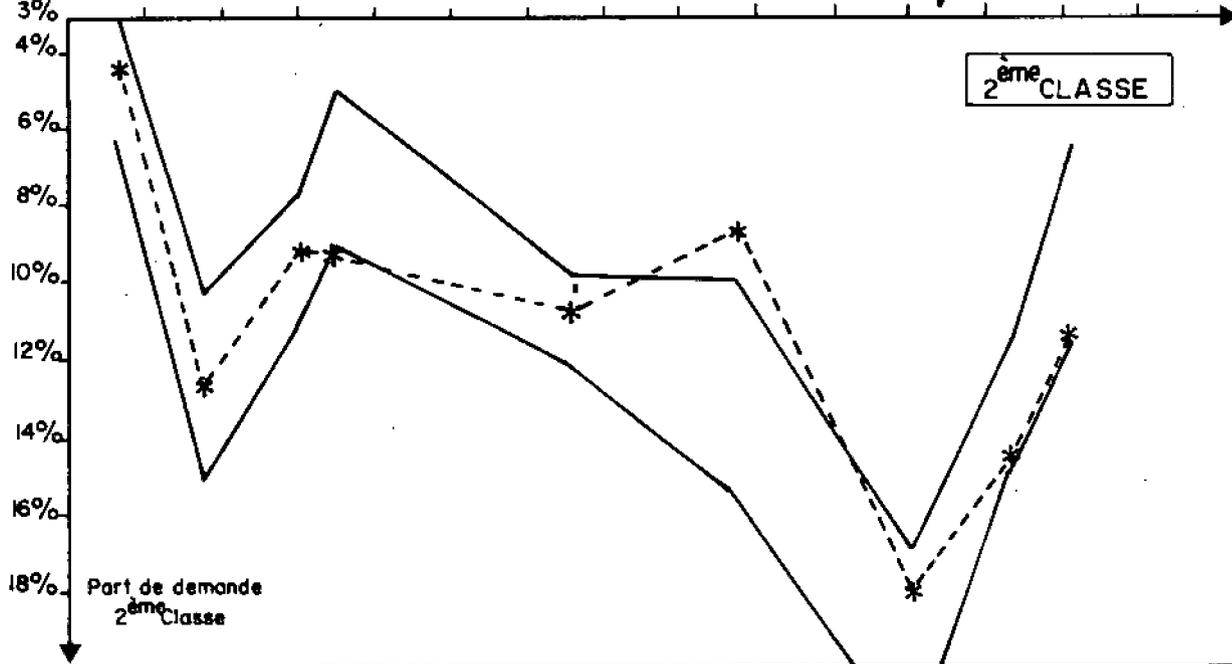
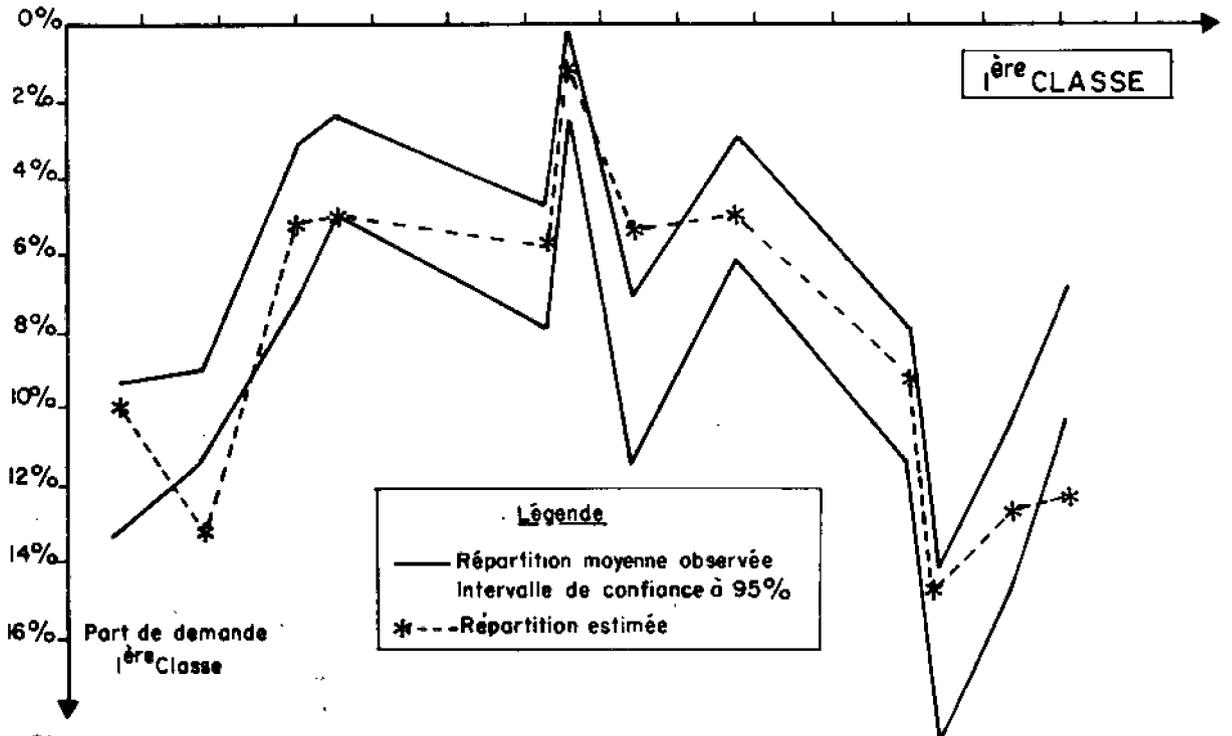
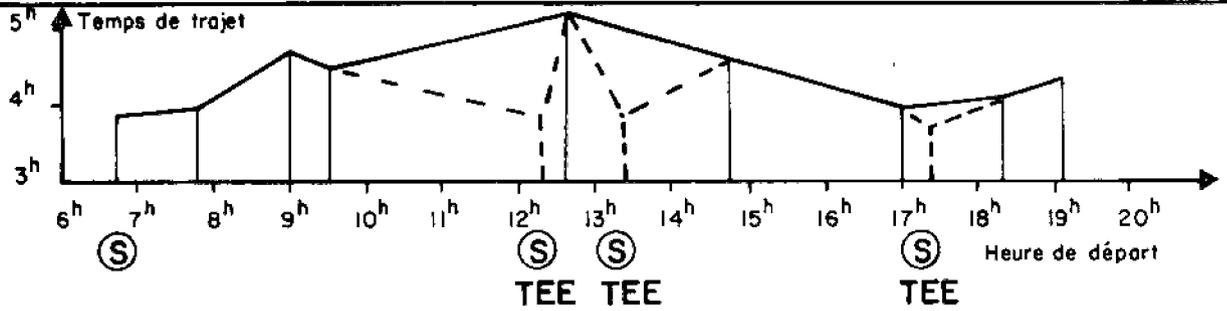
Les graphiques n°7 à 9 montrent les résultats de cette étude comparative.

Le trafic des trains varie selon les jours; il en est donc de même pour la répartition observée qui ne peut être représentée que par un estimateur de sa moyenne; l'estimateur choisi est un intervalle de confiance à 95% (délimité par deux traits pleins) (2).

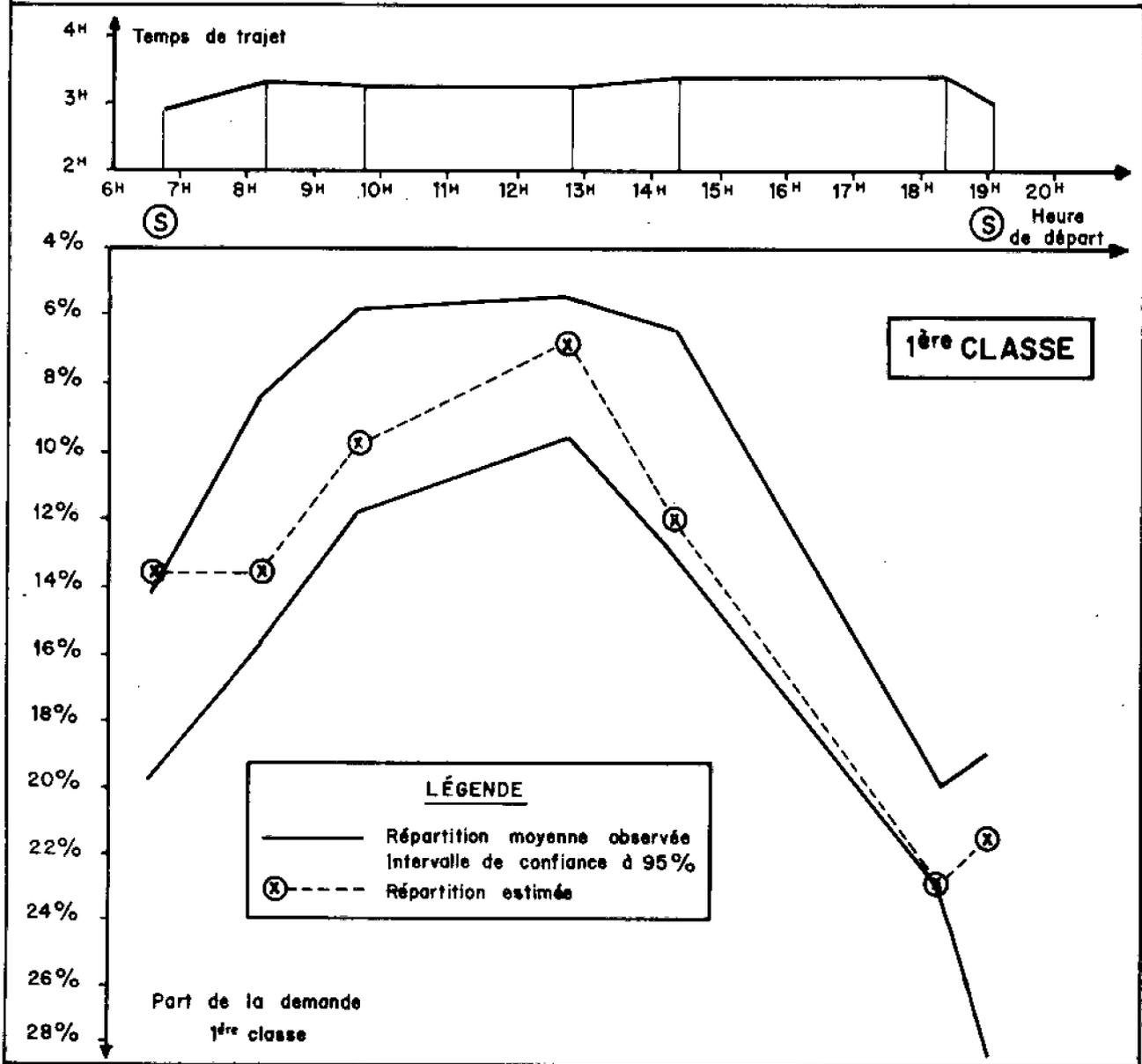
(1) Ces valeurs sont exprimées en Francs 1975

(2) Bien que le modèle permette d'estimer les reports probables de la 2ème classe dans les trains uniquement 1ère classe (repérés en pointillé sur le schéma supérieur), les reports réels n'étant pas connus on a, pour la 2ème classe, tracé uniquement la répartition entre les trains ayant des voitures de 2ème classe.

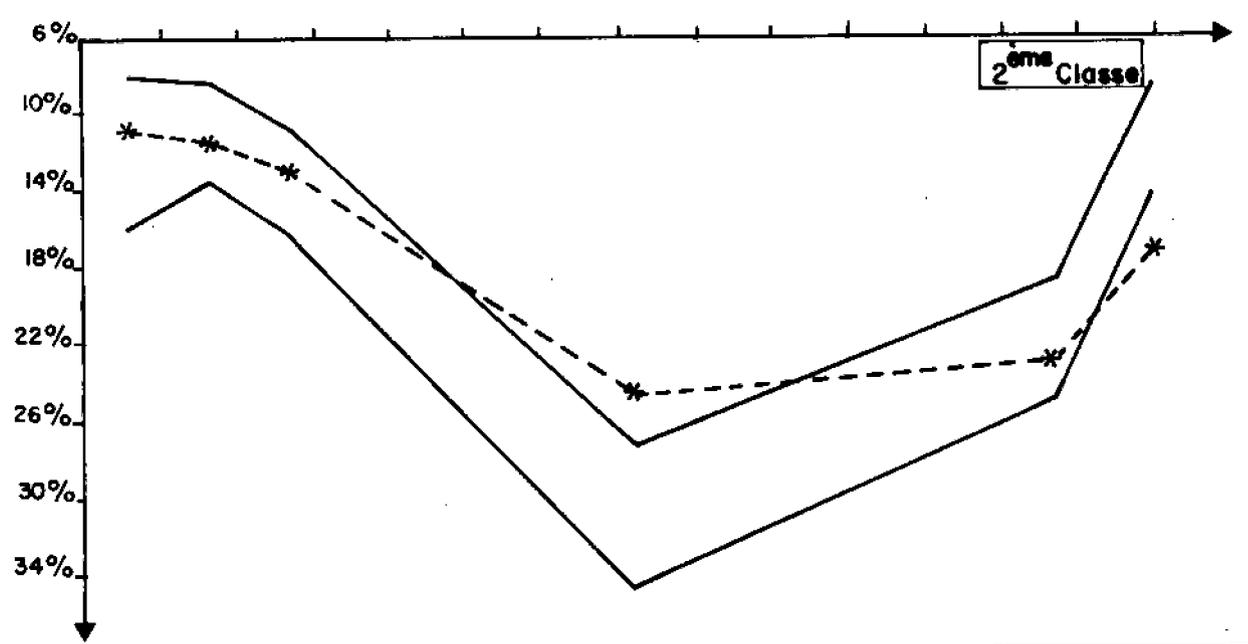
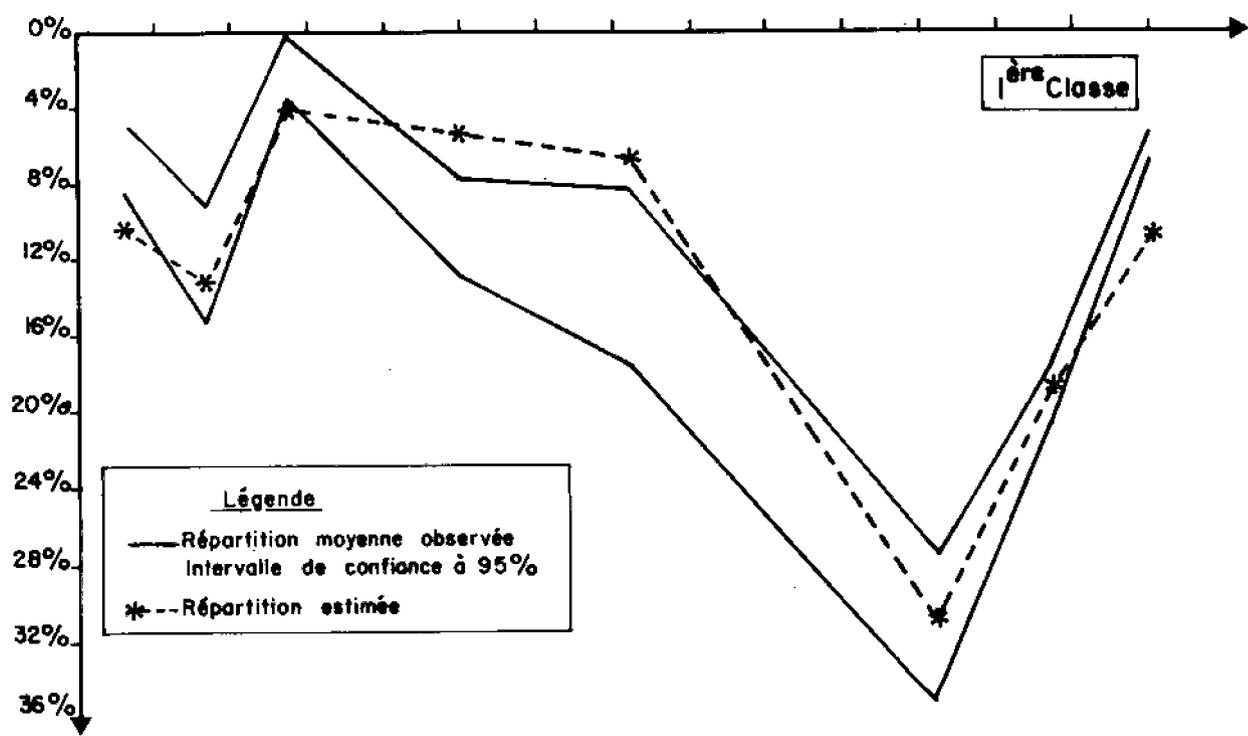
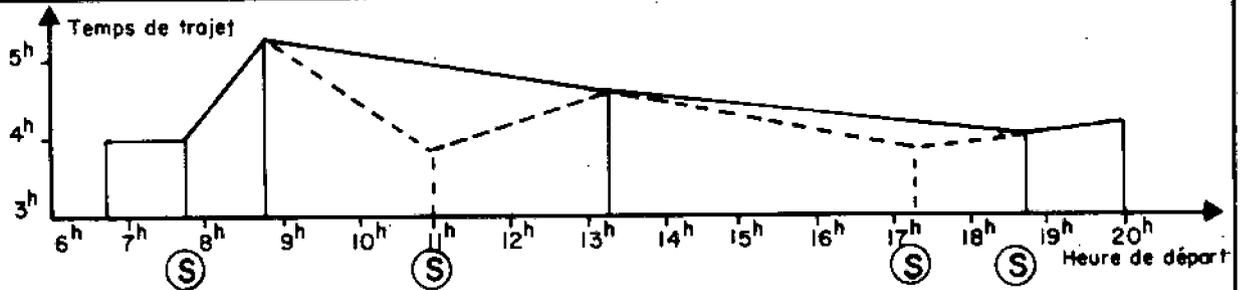
Graphique 7 RÉPARTITIONS RÉELLES ET ESTIMÉES DES DEMANDES DANS LES TRAINS OFFERTS AU SERVICE D'HIVER 74-75 SUR L'OD PARIS-LYON



**GRAPHIQUE n°8 RÉPARTITIONS RÉELLE ET ESTIMÉE DE LA DEMANDE
1^{ère} CLASSE DANS LES TRAINS OFFERTS AU SERVICE
D'HIVER 74-75 SUR L'OD PARIS → RENNES**



Graphique 9 RÉPARTITIONS RÉELLES ET ESTIMÉES DES DEMANDES DANS LES TRAINS OFFERTS AU SERVICE D'HIVER 74-75 SUR L'OD PARIS→STRASBOURG



La répartition probable (représentée par un trait pointillé) a été calculée par le modèle en estimant la moyenne des valeurs du temps à respectivement: 100F/heure (Francs 1975) pour les usagers de la 1ère classe et 10F/heure (Francs 1975) pour les usagers de la 2ème classe.

Le schéma situé en haut de chaque graphique repère les trains offerts par leur heure de départ en abscisse et leur temps de trajet en ordonnée. C'est pour visualiser le lien entre le temps de trajet d'un train et son remplissage que, sur le graphique des répartitions, l'axe des ordonnées est dirigé vers le bas.

L'ensemble des graphiques montre que, bien que les courbes de répartition entre les trains des différentes grilles aient des profils différents, le modèle reproduit bien les répartitions observées. Il ne surestime ni ne sous-estime pendant les périodes de pointe ou de creux de demande.

V - DEVELOPPEMENT DU MODELE

L'étude se poursuit actuellement dans deux directions:

- la généralisation du modèle à la recherche d'un profil théorique de la demande fer 1ère classe + avion et à l'affectation de la demande fer 1ère classe + avion dans une grille composée de trains et d'avions.

Outre les informations apportées sur le remplissage des trains et des avions, ce développement a l'avantage de supprimer l'incertitude qui pèse sur le profil de la demande potentielle fer quand la fréquence des trains offerts est faible. En effet, dans ce cas, il est possible que des voyageurs potentiels du fer empruntent l'avion. Dans l'étude fer 1ère classe + air l'incertitude est levée puisque c'est la demande potentielle fer 1ère classe + air qui est prise en compte;

- la généralisation du modèle au trafic empruntant tous les trains (trains de jour et trains de nuit). Le temps de trajet d'un train de nuit n'est certainement pas ressenti par le voyageur de façon identique à celui d'un train de jour; en effet, la possibilité de passer une nuit de sommeil en voyageant constitue un bonus pour l'utilisateur, bonus qui est probablement variable en fonction du temps de trajet des trains de jour sur la relation.