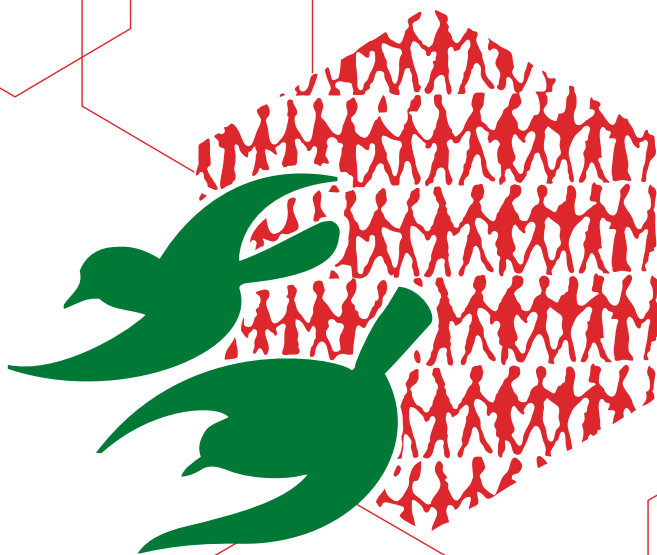


Croissance démographique et urbanisation

Politiques de peuplement et aménagement du territoire

Séminaire international de Rabat (15-17 mai 1990)



ASSOCIATION INTERNATIONALE DES DÉMOGRAPHES DE LANGUE FRANÇAISE

AIDELF

Un modèle de projection démographique par quartiers

Jacques MENTHONNEX

Service Cantonal de Recherche et d'Information Statistiques, Lausanne, Suisse

Contexte et objectifs

Pour élaborer des perspectives de population à un niveau spatial fin, il peut être judicieux de procéder en deux étapes. Après avoir établi des perspectives par âge à un niveau peu désagrégué (pour une ville, par la méthode des composantes par exemple), on répartit les résultats obtenus au niveau le plus fin (le quartier). Cette démarche se justifie pour plusieurs raisons :

- il faut utiliser une méthode qui assure une cohérence globale des résultats (la somme des prévisions doit être plausible);
- la connaissance ou la fiabilité des données statistiques est dépendante du territoire considéré;
- l'inertie des phénomènes démographiques est d'autant plus faible que la région retenue est petite;
- notre connaissance relative à certains phénomènes est également sensible au niveau spatial et conditionne notre aptitude à imaginer des scénarios.

Ces deux dernières raisons expliquent que le choix des variables utilisées pour formuler un scénario sera sans doute dépendant du niveau spatial.

Cette démarche à deux niveaux a été utilisée par le Bureau de la Statistique du Québec⁽¹⁾ à l'aide d'une méthode astucieuse et simple qui fait cependant dépendre, dans une certaine mesure, la croissance d'une région de sa structure par âge de manière incohérente.

Dans cette communication, je propose une méthode un peu plus complexe, mais qui offre deux aspects intéressants.

Premièrement, elle permet de formuler distinctement des hypothèses sur l'évolution du poids démographique des quartiers par rapport à la ville. En effet, expérience faite, il est indispensable de pouvoir infléchir les tendances observées lorsque l'on établit des perspectives pour des régions suffisamment petites pour ne pas offrir d'inerties démographiques importantes. Pour que les prévisions soient crédibles, il est parfois nécessaire de tenir compte de facteurs qui ont eu, ou auront, un impact important sur le développement d'un quartier (construction d'un grand ensemble, changement d'un plan de quartier, aménagement important dans le réseau de transport).

⁽¹⁾ *Perspectives démographiques infrarégionales 1981-2001*, Québec, BSQ, 1984. 498 p.

Deuxièmement, elle décompose certains mécanismes de manière à rendre intelligible le procédé de ventilation par quartier de la population calculée pour la ville. Il sera alors possible d'interpréter la signification des paramètres propres aux quartiers, voire de les lier à certaines caractéristiques quantifiables, telles celles de la structure du parc de logements qui intervient dans le processus de vieillissement démographique, plus ou moins marqué selon le quartier.

Description du modèle

Compte tenu de ces considérations, la population des quartiers sera déterminée en tenant compte des calculs prospectifs relatifs à la population de la ville obtenus par la méthode classique des composantes⁽²⁾.

$$P_i^a$$

Nous déterminerons aussi a priori le poids relatif de chaque quartier i ($i = 1, 2, \dots, I$) à l'époque t :

$$\left\{ \frac{P_i}{P} \right\}_t$$

Ces proportions peuvent être estimées au moyen d'une extrapolation des tendances, mais elles devront parfois être corrigées pour tenir compte des éléments connus susceptibles de les infléchir.

La somme des poids relatifs doit évidemment être égale à 1.

Le choix de la méthode permettant de ventiler les perspectives de population de la ville (par âge et sexe) selon les quartiers découle de l'observation de l'évolution de la structure par âge de quelques quartiers. Parfois on constate clairement un vieillissement de la pyramide alors que, d'autres fois, on a plutôt l'impression que les caractéristiques typiques du quartier restent stables (figure 1). Ces deux situations peuvent aussi s'observer simultanément sur la même pyramide pour des groupes d'âges différents.

Une deuxième considération a été prise en compte pour la formalisation du modèle. Nous avons en effet supposé qu'il existe une liaison entre l'accroissement de la population totale d'un quartier et l'évolution de sa structure par âge.

En définitive, le modèle retenu est le suivant :

$$P_{i,t}^a = l_{i,t} \cdot (\alpha_i \cdot k_a \cdot PS_{i,t}^a + (1 - \alpha_i \cdot k_a) \cdot PG_{i,t}^a) \cdot c_{a,t}$$

L'estimation de la population $P_{i,t}^a$ des quartiers (i) par âge (a) pour l'année t résulte d'un compromis entre deux populations-types : la première, PS , présuppose une stabilité des caractéristiques de la structure par âge des quartiers comparativement à celle de la ville ; la deuxième, PG , privilégie les effets de génération (la pyramide « vieillit ») et tient compte si nécessaire d'effets d'âge dus aux migrations.

Le calcul et l'interprétation de ces deux populations-types sont présentés plus loin.

⁽²⁾ Notation : a pour l'âge, t pour l'année, i pour le quartier. Si i n'est pas indiqué, il s'agit de la ville. Par souci de simplification, le sexe n'est pas indiqué dans la formulation.

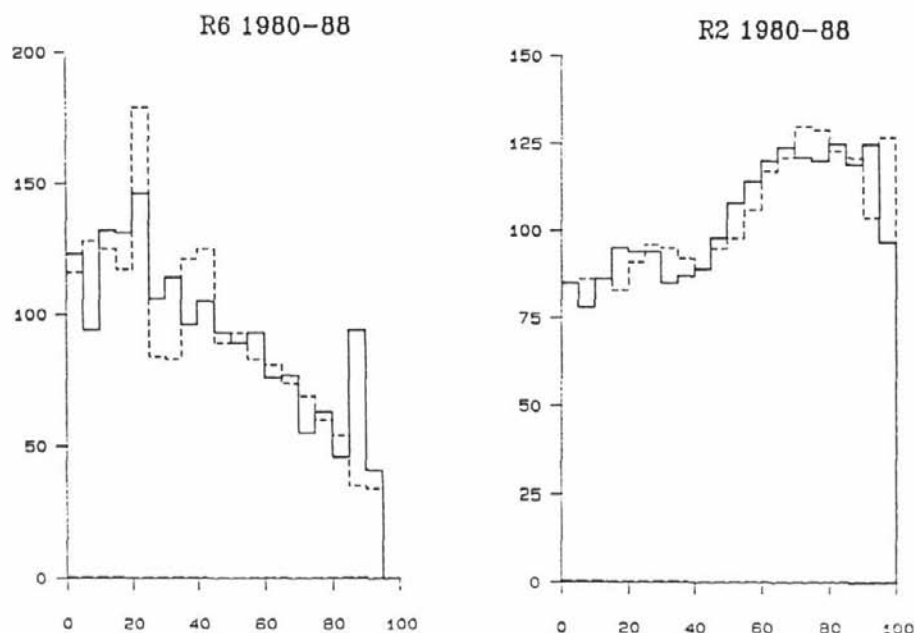


Figure 1.- Structures par âge de deux quartiers de la ville de Lausanne (1980 et 1988).
Indice 100 = structure de la ville entière

Le compromis est déterminé par les valeurs α_i et k_a qui permettent de prendre en compte un effet propre à chaque quartier (α_i) qui se combine à un effet d'âge (k_a).

Autrement dit, on considère qu'une part ($\alpha_i \cdot k_a$) de la population du quartier i et d'âge a est directement liée aux infrastructures typiques du quartier qui conditionnent une certaine stabilité de la structure par âge de la population. Logiquement, le produit ($\alpha_i \cdot k_a$) doit être compris entre 0 et 1, quel que soit le quartier i et l'âge a .

$l_{i,t}$ et $c_{a,t}$ sont des coefficients, en pratique proches de l'unité, qui réajustent légèrement les résultats, afin que les valeurs de $P_{i,t}^a$ soient parfaitement cohérentes avec les perspectives établies pour la population de la ville et satisfassent aux hypothèses sur le poids relatif des quartiers.

Signification et détermination de la population-type PS

Le point de vue choisi sous-entend que la structure par âge des quartiers est pratiquement déterminée par ses infrastructures : on imagine que l'habitat explique et « stabilise » l'âge moyen des habitants : la structure du parc de logements, la présence de logements subventionnés, d'internats, d'établissements médico-sociaux concernent chacun des âges spécifiques. La proximité de places de travail ou d'écoles peut aussi avoir une influence dans le même sens. Dans ces conditions nous avons défini la population-type PS correspondante de la manière suivante :

$$PS_{i,t}^a = \left(\frac{P_i^a / P_i}{P^a / P} \right)_{t_0} \cdot \left(\frac{P_i}{P} \right)_t \cdot P_i^a$$

Le premier terme, constant, correspond à la structure par âge relative à celle de la ville à l'époque t_0 (dépendant des dernières statistiques disponibles). Cette méthode présuppose que la structure relative de la population de chaque quartier est stable et que l'évolution d'un groupe d'âges dans un quartier est identique à celle de la ville, en terme de taux.

On doit encore remarquer, que cette formule n'assure rigoureusement

l'égalité $\sum_i PS_{i,t}^a = P_t^a$ que pour $t = t_0$;

les écarts, de fait très petits, seront pris en compte ultérieurement par $l_{i,t}$ et $c_{a,t}$.

Signification et détermination de la population-type PG

On prend ici un point de vue plus «démographique» puisque l'on tient compte du vieillissement des résidents et du fait que les flux migratoires influencent la structure par âge de certains quartiers.

La population-type PG a été calculée au moyen de la formule suivante :

$$PG_{i,t}^a = (PG_{i,t-1}^{a-1} / PG^{a-1})_{t-1} \cdot P_t^a + d_{i,t} \cdot w_a$$

pour ⁽³⁾ $a \geq 1$ et pour ⁽⁴⁾ $t > t_0 + 1$

Le premier terme sous-entend que les taux par âge de mortalité et de migrations nettes sont identiques pour le quartier et pour la ville (et donc déterminés par les hypothèses du modèle par composantes). Comme la population résultant de ce premier terme ne prend pas en compte certaines migrations liées à la croissance relative de chaque quartier, on rajoute $d_{i,t}$ migrants (ou on soustrait) que l'on répartit par âge selon le vecteur w . Ces migrants correspondent, en quelque sorte, aux mouvements extraordinaires par rapport à la tendance générale déjà prise en compte dans le premier terme; par définition leur somme est nulle ($\sum_i d_{i,t} = 0$).

On signalera encore que les deux marginales de PG sont rigoureusement conformes aux perspectives de la ville et aux poids des quartiers.

Ce modèle permet de simuler des cycles sur la structure d'âge de certains quartiers (les quartiers «vieux» deviennent «jeunes» puis revieillissent...).

⁽³⁾ Pour $a = 0$, on a utilisé

$$PG_{i,t}^0 = F_i \cdot \left(\frac{P_i^{25-35}}{P^{25-35}} \right)_t + d_{i,t} \cdot w_0$$

où F_i est un indicateur de fécondité générale de i relative à celle de la ville, observé en t_0 .

⁽⁴⁾ Si $t = t_0 + 1$, on a

$$\left(\frac{P_i^{a-1}}{P^{a-1}} \right)_{t_0} \cdot P_t^a$$

comme premier terme.

Test du modèle

Cadre du test

Pour vérifier l'intérêt pratique du modèle, les valeurs des paramètres α_i , k_a et w_a ont été estimées à partir de données lausannoises⁽⁵⁾. Non seulement les écarts entre les populations calculées et celles observées doivent être suffisamment faibles, mais, en plus, il est souhaitable que les valeurs des paramètres soient relativement stables dans le temps et permettent une interprétation plausible.

Concrètement, c'est sur la base des statistiques⁽⁶⁾ de la population résidente par quartier en 1980, 1984 et 1988 que nous avons déterminé les estimateurs de α_i , k_a et w_a . Les variantes testées ont été nombreuses afin d'établir des comparaisons dans le temps (1980-84 ou 1984-88). Différents découpages spatiaux ont été utilisés et, dans certains cas, quelques paramètres ont été fixés a priori. De ce fait, seule une synthèse des résultats est présentée dans cette communication.

Alors que pour l'élaboration de perspectives par quartier, on doit utiliser, comme on l'a vu, des perspectives déjà établies au niveau de la population de la ville (P_i^a) et au niveau du poids relatif des quartier (P_i/P), on a utilisé pour tester le modèle les valeurs réelles.

Estimation des paramètres du modèle

L'estimation des α_i , k_a et w_a , qui minimisent l'écart entre les populations par âge estimées et observées des quartiers, est a priori complexe. Non seulement le modèle n'est pas linéaire par rapport aux paramètres, mais il existe des contraintes sur les paramètres : la somme des w_a doit être égale à un, et les valeurs des produits $\alpha_i \cdot k_a$ doivent être compris entre 0 et 1.

La difficulté a été contournée en utilisant un procédé itératif⁽⁷⁾ qui estime les paramètres en plusieurs étapes (chaque étape correspond à une régression linéaire avec contraintes). Expérience faite, le processus converge vers une solution sans être sensible aux valeurs initiales choisies.

Le calcul des paramètres « d'ajustement comptable » $l_{i,t}$ et $c_{a,t}$ a été effectué par la méthode classique de doubles ajustements proportionnels (appelé aussi RAS). Ces derniers paramètres ne sont intéressants dans la phase de test du modèle que dans la mesure où l'on doit vérifier qu'ils ont un rôle pratiquement mineur (ils doivent être très proches de 1).

⁽⁵⁾ Lausanne, la 5^e ville de Suisse, avec ses 120 000 habitants, est située au bord du lac Léman et est le chef lieu du canton de Vaud. Cette commune, qui constitue le noyau d'une agglomération de 250 000 habitants, a une population qui décroît depuis 1973.

⁽⁶⁾ Fournies par Catherine Schaffner de l'Office d'études socio-économiques et statistiques de la Ville de Lausanne.

⁽⁷⁾ L'annexe fournit quelques indications supplémentaires.

Découpage spatial utilisé

Dans le cadre de cette communication, nous utilisons un découpage en six quartiers, obtenus par agrégation de plus petits secteurs, définis par l'Office d'études socio-économiques de la Ville de Lausanne. La figure 2 illustre le choix effectué, qui ne correspond pas à un découpage administratif, mais qui réunit des régions qui ont certaines caractéristiques communes et qui correspondent à peu près à des phases de croissances de la ville.

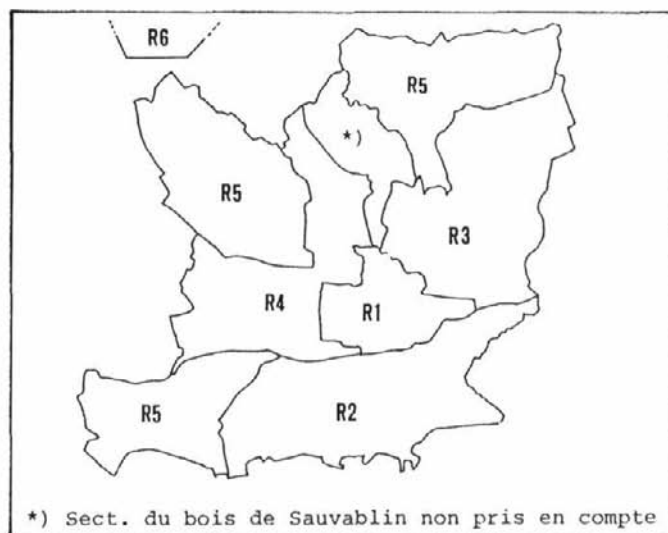


Figure 2.- Les six quartiers lausannois retenus

TABLEAU 1.- QUELQUES POINTS DE REPÈRE

	Population en 1984	Croissance en %		Proportion de 65 ans et +
		1980-84	1984-88	
R1	11 000	-5	-7	0,21
R2	26 500	-2	-6	0,22
R3	16 400	0	-1	0,20
R4	36 800	-2	-4	0,18
R5	27 000	0	5	0,12
R6	1 500	5	8	0,13

Présentation de quelques résultats*Valeurs obtenues pour w_a*

Lorsque l'on calcule PG , on détermine d'abord la population théorique du quartier en supposant que les taux de migrations nettes et de mortalité par âge sont identiques d'un quartier à l'autre. Ce premier résultat doit être corrigé, afin que le poids relatif du quartier soit conforme à notre attente (à la réalité dans le cas du test). La correction pour le quartier i est répartie par âge par les coefficients w_a ; elle correspond à des migrations « extraordinaires » par rapport à la situation générale.

Le nombre des d_{it} migrants à répartir est directement lié à la structure par âge et à la croissance du quartier. Pratiquement, on constate que, dans certains cas (R2, R3),

la structure est suffisamment « vieille » pour que d_i soit positif malgré le dépeuplement du quartier.

L'estimation de w_a montrant quelques variations curieuses, dues au fait que w_a n'est pas totalement indépendant du quartier, nous avons été conduits à utiliser deux vecteurs de répartition : w_p lorsque d_i est positif et w_n lorsque d_i est négatif. A titre illustratif, la figure 3 montre les résultats obtenus sur la période 1984-88 (hommes et femmes réunis).

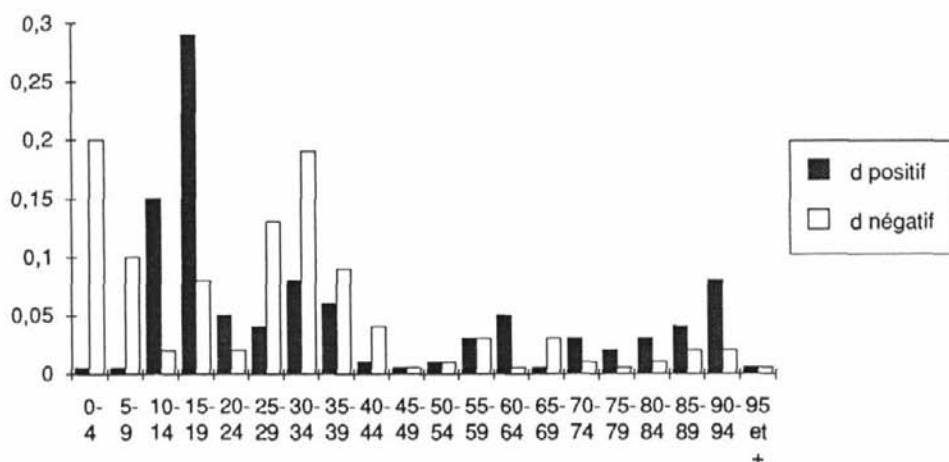


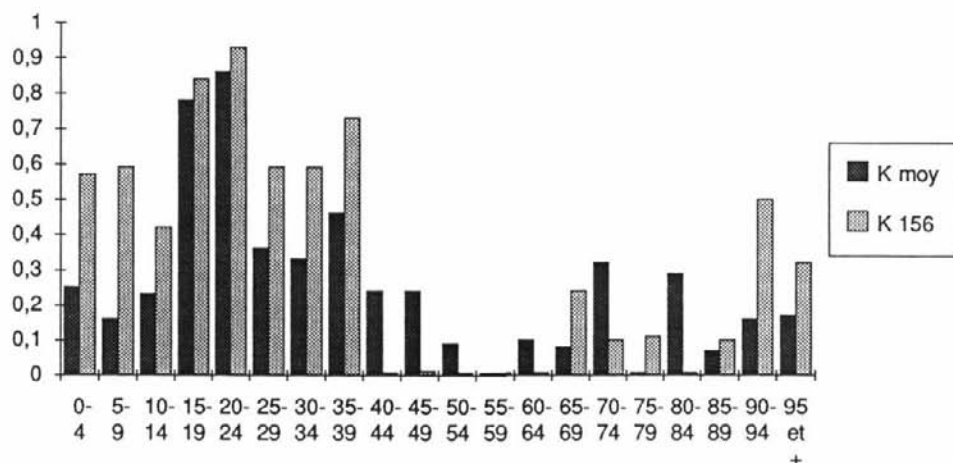
Figure 3.- Estimation de w_a sur la base de la période 1984-88

On remarque que les arrivées « extraordinaires » sont fortement liées aux apprentis et étudiants alors que les départs sont plutôt le fait de jeunes familles avec enfants ou d'étudiants qui ont terminé leur formation. Ces caractéristiques se retrouvent dans la majorité des variantes (par sexe ou sur la période 1980-84). Aux âges plus élevés, on remarque quelques instabilités d'une période à l'autre, qui s'expliquent parfois par la période de la retraite, ou le départ dans un établissement médico-social, mais qui sont aussi le fait d'effets aléatoires liés à des petites populations (comme par exemple les valeurs relativement élevées de w_p au-dessus de 80 ans dans la variante présentée sur la figure 3).

Valeurs obtenues pour k_a et α_i

Il faut préalablement rappeler que le modèle prendra en compte, pour le calcul de $P_{i,n}^a$, une proportion $k_a \cdot \alpha_i$ de la population $PS_{i,n}^a$, qui correspond à la population à structure relative par âge constante.

A partir des résultats obtenus pour l'estimateur de k_a , un élément se dégage clairement : les effets de structure relative stable sont plus fréquemment apparents vers 15-

Figure 4.- Estimation de k_a (2 variantes)

35 ans et, dans une mesure un peu moindre, à partir de 60 ans. La figure 4 illustre succinctement ce constat pour deux variantes.

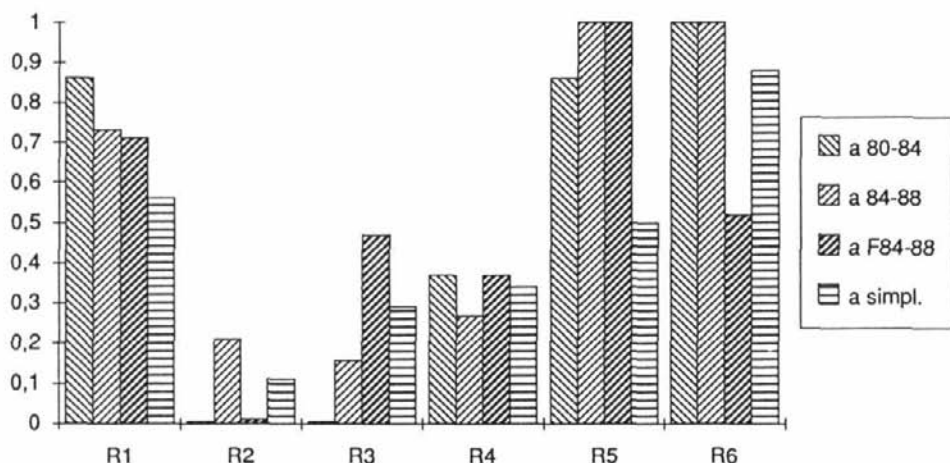
La première (k_{moy}) correspond à la moyenne des estimations effectuées sur les deux périodes (hommes et femmes réunis), alors que la seconde (k_{156}) n'a été calculée qu'à partir des populations des quartiers N° 1, 5 et 6, c'est-à-dire des quartiers qui ont une structure par âge relativement stable (α_i élevé).

On a pu remarquer que la construction d'un grand ensemble⁽⁸⁾ « perturbait » dans certains cas les estimations de k_a , qui prenait une valeur étonnamment élevée aux âges 55-65 ans. Il est clair que la mobilité élevée des jeunes, ainsi que l'existence d'établissements spécialisés (pour les jeunes ou pour les vieux), expliquent assez bien l'allure générale de la courbe.

La valeur prise par les coefficients α_i est particulièrement importante, tant au niveau de l'intérêt descriptif qu'au niveau de son rôle sur les résultats de perspectives démographiques : lorsqu'ils sont proches de l'unité, les caractéristiques de la structure par âge du quartier (relativement à celles de la ville) sont conservées aux âges où le coefficient k_a est élevé.

Pratiquement on a pu constater que le quartier centre (R1) et les quartiers périphériques (R5 et R6) semblaient avoir nettement ces caractéristiques. La figure 5 met en évidence les résultats obtenus dans quatre situations ; (a_{80-84}) correspond à l'estimation des coefficients calculés sur base de la période 1980-84, les hommes et les femmes étant réunis, alors que (a_{84-88}) a été déterminé sur la période 1984-1988. Pour illustrer les instabilités des résultats, la figure montre encore le résultat de l'estimation pour la population féminine sur la 2^e période (a_{F84-88}), ainsi que dans le cas où on a simplifié le modèle (a_{simpl}) : vecteur w unique et k_a fixé à 1 pour tous les âges. On constate que, malgré quelques écarts entre les variantes, l'allure générale des résultats est stable.

(8) Praz-Séchaud et les Grangettes dans la région 5.

Figure 5.- Estimation des α_i

Signification des paramètres α_i

Il est intéressant de comprendre pourquoi les quartiers R1, R5 et R6 ont un coefficient proche de l'unité, c'est-à-dire une structure par âge de leurs résidents nettement plus stable que celle des autres quartiers.

Une analyse des caractéristiques des quartiers donne la réponse spécifique à chaque quartier.

Le quartier R1 (le centre-ville) est caractérisé par un parc de logements anciens (89% ont été construits avant 1960 contre 68% pour la ville), avec plus de la moitié de logements d'une ou deux pièces. On y trouve alors une proportion particulièrement élevée de personnes vivant seules (31% contre 22% pour la ville), de nombreux étrangers (34% contre 27%) et une part d'étudiants plus élevée que la moyenne. On imagine aisément que la structure du parc va encourager une forte mobilité qui stabilisera, dans une certaine mesure, la forme de la pyramide des âges.

La structure par âge du quartier R5 est stable pour d'autres raisons : ce quartier en croissance est constitué d'un parc de logements plus récents avec de nombreux trois pièces (37% contre 31% pour la ville); ce parc comprend une proportion importante de logements subventionnés qui ne sont attribués qu'à des locataires répondant à certains critères; on y trouve, typiquement, des familles traditionnelles avec enfants et des familles monoparentales.

Le quartier R6 (1 500 habitants), qui comprend l'ensemble des zones dites foraines, a des caractéristiques très différentes du reste de la commune : outre le fait qu'il possède peu de logements de une ou deux pièces (25% contre 47% pour la ville), de nombreux habitants logent dans un ménage collectif (20% contre 3% pour la ville), il s'agit surtout d'étudiants et de pensionnaires d'établissements médico-sociaux spécialisés. Dans ces conditions, il n'est pas étonnant que la structure par âge des résidents conserve certaines spécificités.

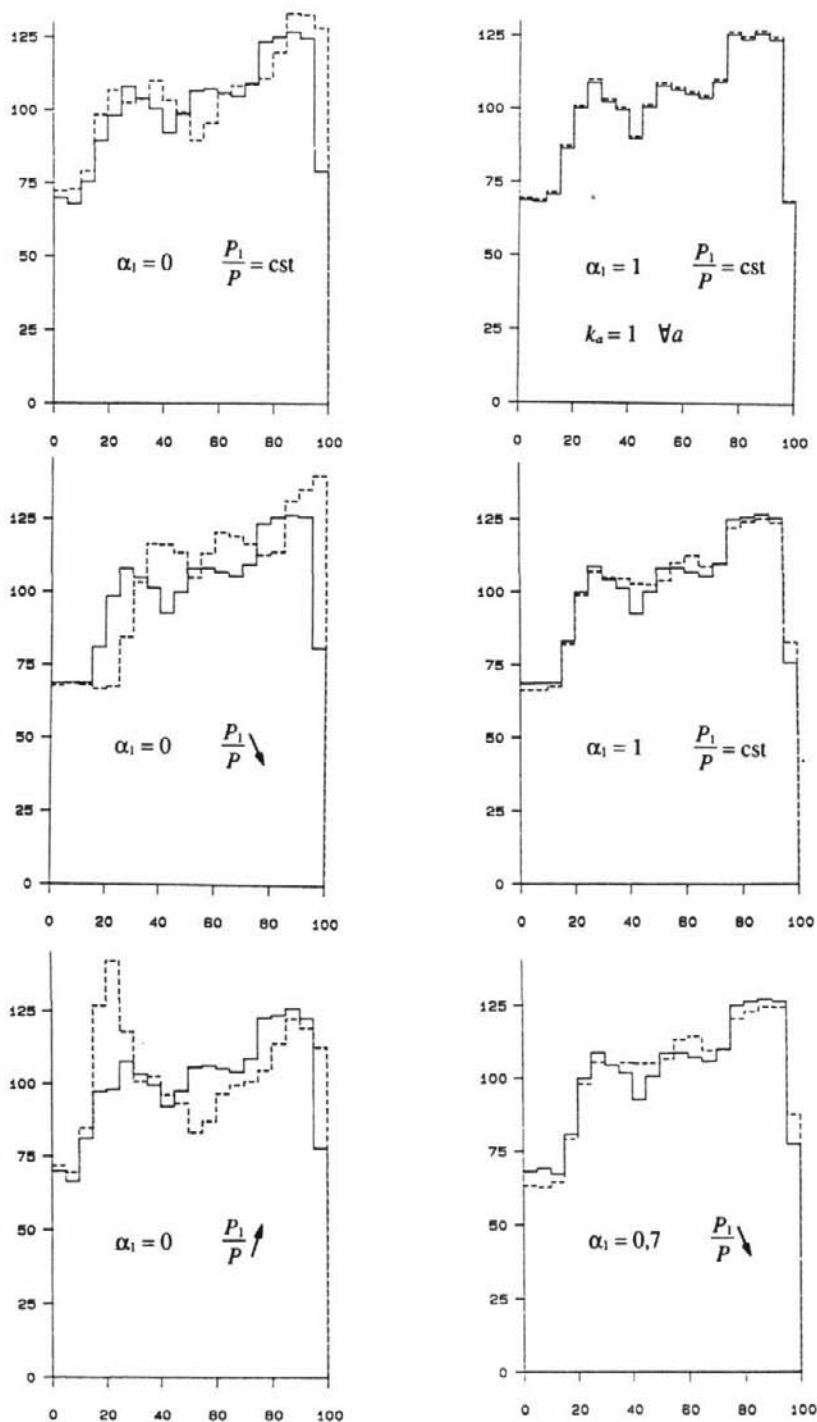


Figure 6.— Structure relative par âge de la région 1 en 1989 et 2000 ; simulation de quelques situations typiques (indice 100 = structure de la ville)

Sans entrer dans les détails on signalera que les caractéristiques du parc de logements et celles des résidents des quartiers R2, R3 et R4 sont plus proches de la moyenne de la ville que celles des quartiers R1, R5 et R6.

Simulation de l'impact de α_i sur les perspectives

Afin de visualiser l'évolution de la structure par âge des quartiers et de mesurer sur un cas concret le rôle de α_i , les perspectives lausannoises établies par la méthode des composantes ont été ventilées par la méthode exposée dans cette communication. Comme il s'agissait d'effectuer une simulation, et non pas des perspectives démographiques, les calculs ont été effectués jusqu'en l'an 2000, avec des variantes sur l'évolution du poids des quartiers et sur le niveau de certains α_i .

Afin de dissocier le rôle de la méthode présentée ici de l'influence des perspectives établies au niveau de la ville, on a présenté la structure par âge d'un quartier relativisée à celle de la ville. Les résultats obtenus pour le quartier R1 sont présentés sur la figure 6 (cas où $i = 1$).

Les graphiques de gauche illustrent les résultats obtenus lorsque $\alpha_i = 0$, c'est-à-dire lorsque la méthode privilégie les effets de vieillissement. Outre un décalage de la courbe de structure relative vers la droite, on remarque clairement l'impact de l'hypothèse formulée sur l'évolution du poids relatif du quartier dans le temps $(P_i/P)_t$ qui favorise plus ou moins le poids des jeunes (effet des migrants $d_{i,t}$, ventilés par âge selon w_a) et crée un effet de rajeunissement de la population du quartier, dans la mesure où celui-ci croît.

Le premier graphique de droite correspond au cas où la structure relative est stable (modèle PS lorsque $\alpha_i \cdot k_a = 1$). Si l'effet d'âge (selon la figure 4) est introduit, on remarque sur le deuxième graphique de droite qu'un effet de vieillissement apparaît aux âges où k_a est faible.

Le dernier graphique, en bas à droite, simule la situation la plus probable, compte tenu des observations effectuées sur la période 1980-1988.

Conclusion

On a pu constater que ce modèle est satisfaisant dans la mesure où l'estimation des paramètres conduit à des résultats plausibles, qui offrent une certaine stabilité dans le temps. Dans la situation du test, les écarts entre les calculs et la réalité sont faibles à une exception près : certaines particularités de la structure par âge de la région 5 n'ont pas été correctement déterminées par le modèle, à cause de la construction d'un grand ensemble.

L'estimation des paramètres sur une période connue offre quelques difficultés techniques mais permet de mieux déceler certaines spécificités des quartiers et offre donc un intérêt descriptif évident.

Les paramètres du modèle étant spécifiés, il est aisé de rendre le procédé opérationnel pour élaborer des perspectives. Comme les paramètres ont une signification et un impact sur les résultats explicites, il est possible de les adapter pour tenir compte d'éléments connus qui changeront probablement certaines caractéristiques des quartiers.

ANNEXE

Quelques indications utiles pour estimer les paramètres du modèle

Notation

$\hat{a}^0$: valeur initiale de a choisie

$\hat{a}^n$: estimateur de a , au sens des moindres carrés, calculé à la $n^{\text{ème}}$ itération

pour une année t donnée on écrit pour alléger l'écriture :

$$y_i^a = P_i^a - PC_i^a$$

$$x_i^a = PS_i^a - PC_i^a \quad \text{où} \quad PC_{i,t}^a = \left(\frac{PG_i^{a-1}}{PG^{a-1}} \right)_{t-1} \cdot P_i^a$$

Estimation de α_i , w_a et k_i

On connaît pour une année passée :

P_i^a par observation,

PS_i^a par calcul,

PC_i^a par calcul, y_i^a et x_i^a sont donc directement déduits

$$d_i = P_i - PC_i$$

α_i , w_a et k_a sont estimés par itérations successives ; chaque étape est constituée par la résolution de deux régressions linéaires avec contraintes⁽¹⁾. Pour la $n^{\text{ème}}$ itération on a :

$$y_i^a = \hat{k}_a^{n-1} \cdot x_i^a \cdot \alpha_i^n + (1 - \hat{\alpha}_i^{n-1} \cdot \hat{k}_a^{n-1}) \cdot d_i \cdot w_a^n$$

$$\text{avec } \sum_a w_a^n = 1$$

$$\alpha_i^n = 1 \quad \text{si } \hat{\alpha}_i^{n-1} > 1$$

$$\alpha_i^n = 0 \quad \text{si } \hat{\alpha}_i^{n-1} < 0, \quad i = 1, 2, \dots$$

Ce système permet de calculer $\hat{\alpha}_i^n$ et $\hat{w}_a^n$ puis, en réarrangeant l'équation, on a :

$$y_i^a = d_i \cdot \hat{w}_a^n + \hat{\alpha}_i^n \cdot (x_i^a - d_i \cdot \hat{w}_a^n) \cdot k_a^n$$

$$\text{avec } k_a^n = 1 \quad \text{si } \hat{k}_a^{n-1} > 1$$

$$k_a^n = 0 \quad \text{si } \hat{k}_a^{n-1} < 0$$

on détermine alors $\hat{k}_a^n$

Tous calculs faits, il est possible de regrouper les quartiers en deux sous-groupes et d'imaginer une variante pour estimer des w_a spécifiques à chaque sous-groupe.

⁽¹⁾ Voir, par exemple : Johnston, *Econometric methods*, Mc Graw-Hill, 1972. p. 157