

CUDEP
XI COLLOQUE NATIONAL DE DEMOGRAPHIE

CONSTRUIRE DES TABLES
EVOLUTIVES DE MIGRATION NETTE
PAR SEXE ET PAR AGE

Maks Banens
Membre du CEFRESS
Université de Picardie Jules Verne

Octobre 1999

Les besoins de prévision démographique des populations locales ont conduit à concevoir des outils de projection, qui s'appuient sur des données locales limitées auxquelles s'ajoutent des scénarios-types empruntés à la population nationale. Le logiciel Omphale, développé au sein de l'INSEE (Descours 1992), en est l'un des exemples les plus connus. Dans Omphale, les profils de fécondité par âge et de la mortalité par sexe et par âge sont nationaux, tout comme leur extrapolation future. Les quotients de migration nette par âge et par sexe, par contre, sont estimés sur la base de l'observation locale.

L'observation détaillée de la migration nette intercensitaire est donc la pièce maîtresse de la prévision démographique locale. Les données locales permettent d'observer quatre périodes intercensitaires, plus trois regroupements : 1962-1975, 1968-1982 et 1975-1990.

Nous étudierons dans ce papier la possibilité de substituer au profil migratoire unique, établi sur l'observation d'une seule période intercensitaire, tel que l'utilise Omphale, un profil migratoire évolutif, établi sur l'observation de plusieurs périodes intercensitaires.

L'estimation intercensitaire

Omphale estime les quotients annuels de migration nette par âge (QM_i ; l'âge en années révolues) par une procédure complexe mêlant lissages et délissages. L'hypothèse du départ est la supposition que pour chaque cohorte, le quotient de migration nette est constant sur toute la période d'observation. Pour la cohorte ayant l'âge i au 1^{er} Janvier de l'année de recensement RP1, le quotient annuel de migration nette entre les recensements RP1 et RP2 est approché d'abord par les équations :

$$QM1_i = QM1_{i+1} = \dots = QM1_{i+a-1} = [(P_{i+a}^{RP})/P_{i+a}^{ATT}]^{1/a} - 1$$

où P_{i+a}^{RP} = population d'âge $i+a$ au 1^{er} Janvier de l'année de recensement RP2, estimée d'après RP2 ;

P_{i+a}^{ATT} = population d'âge $i+a$ au 1^{er} janvier de l'année de recensement RP2, attendue sur la base du recensement précédent et de la mortalité intercensitaire ;

a = nombre d'année séparant les années de recensements RP1 et RP2.

De ces équations, seule celle concernant l'âge centrale est retenue :

$$QM1_{i+(a+1)/2} = [(P_{i+a}^{RP})/P_{i+a}^{ATT}]^{1/a} - 1$$

Plus généralement :

$$QM1_i = [(P_{i+(a-1)/2}^{RP})/P_{i+(a-1)/2}^{ATT}]^{1/a} - 1$$

Cette première estimation étant généralement assez chaotique, elle est lissée deux fois :

$$QM2_i = \text{médian de } [QM1_{i-1} ; QM1_i ; QM1_{i+1}]$$

$$QM3_i = [(1+QM2_{i-1}) * (1+QM2_{i+1}) * (1+QM2_i)^2]^{1/4} - 1$$

Ensuite, cette courbe est délissée à l'âge désiré (m) pour reconstituer la concentration migratoire à cet âge.

Jusqu'à l'âge de m ans :

$$QM4_i = [1+QM3_{i-a}] * [(1+QM3_{i-(a-1)/2}) / (1+QM3_{i-(a+1)/2})]^a - 1$$

Après l'âge de m ans :

$$QM5_i = [1+QM3_{i+a}] * [(1+QM3_{i+(a-1)/2}) / (1+QM3_{i+(a+1)/2})]^a - 1$$

Le délissage peut s'effectuer à plusieurs âges. Les différentes parties de la courbe des quotients sont enfin recollées par un dernier lissage autour des bornes de coupures (QM6 et QM7).

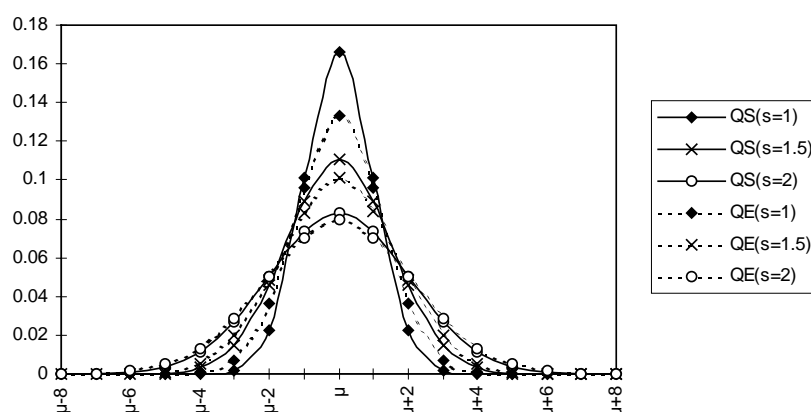
L'ensemble de la procédure est appelé : lissage composite. Son efficacité dépend de plusieurs facteurs que l'on peut étudier séparément : de la durée de la période intercensitaire ; du choix de l'âge de migration maximale ; et enfin du profil même de la migration.

Le cas de la migration nette en sens unique

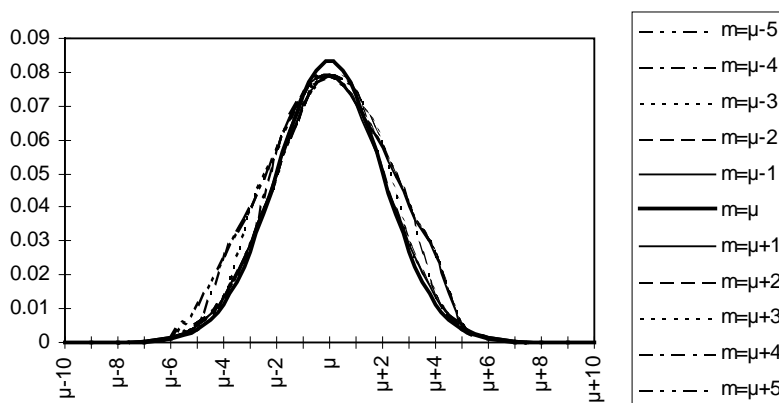
Les graphiques 1 et 2 montrent l'estimation réalisée par le lissage composite d'une migration nette concentrée autour de l'âge μ . La première simulation fait varier la dispersion migratoire : à une intensité migratoire constante de 0,5 (en l'absence de mortalité, chaque cohorte augmente de 50 % en volume au cours des années migratoires), la migration est plus ou moins concentrée autour de l'âge μ selon une distribution normale dont la dispersion est donnée par l'écart type (s). Ensuite, le lissage composite estime à partir de la population du départ et celle projetée les quotients de migration simulés.

Par construction, une migration unique et concentrée autour d'un âge produit des estimations qui sont indépendantes de la durée intercensitaire. Mais elles sont plus ou moins exactes selon la dispersion de la migration. Le graphique 1 montre que si les quotients simulés avec un écart type de 2 années sont très bien estimés, ceux qui résultent d'une dispersion moindre ($s=1$) le sont moins bien. Or, la migration à l'âge de 18 ans est reconnue pour être particulièrement concentrée.

Le graphique 2 étudie le cas où l'âge de concentration maximale (μ), a priori inconnue, serait mal évalué par l'utilisateur du lissage composite. On a varié l'âge (m) introduit dans le calcul, supposant par ailleurs le cas d'une concentration migratoire modérée ($s=2$). Les estimations varient très peu. On peut donc conclure à la robustesse du lissage composite, même dans le cas d'une erreur sur le choix des bornes de délissage.

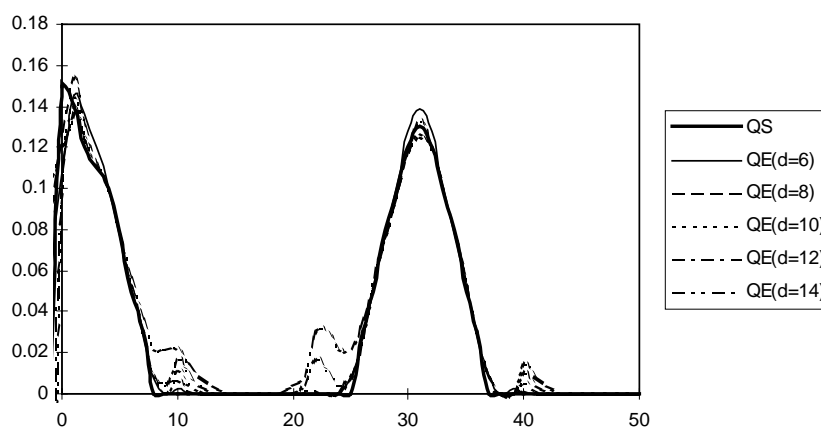


Graphique 1. Simulation 1 : quotients de migration nette simulés (QS, intensité finale : 0,5) et quotients estimés (QE) par la procédure du lissage composite, en fonction de la dispersion migratoire (s)

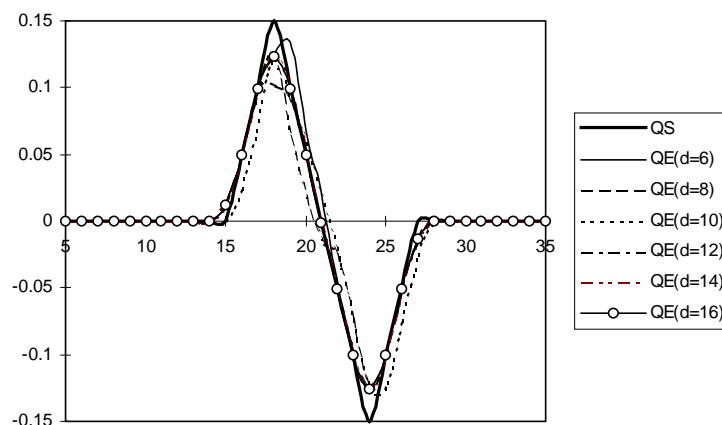


Graphique 2. Simulation 2 : quotients de migration nette simulés ($m=\mu$; $s=1$; intensité finale = 0,5), et quotients estimés par la procédure du lissage composite, en fonction du choix de la borne (m)

Les populations locales montrent souvent des profils de migrations plus complexes que la migration unique autour d'un seul âge. Si les différents âges de migrations sont suffisamment éloignés l'un de l'autre, comme ceux à l'entrée et à la sortie de l'activité, l'efficacité du lissage composite n'est pas affectée. Même quand les âges sont plus rapprochés, comme dans le graphique 3, l'effet perturbateur reste limité. On y montre le profil migratoire classique de la première phase de la périurbanisation. Seule une durée intercensitaire prolongée entraîne des effets indésirés sur l'estimation migratoire non pas au cœur même des âges migratoires, mais à ses marges.



Graphique 3. Simulation 3 : quotients de migration nette simulés (QS), et quotients estimés par la procédure du lissage composite (QE), en fonction de la durée intercensitaire (m)



Graphique 4. Simulation 4 : quotients de migration nette simulés (QS), et quotients estimés par la procédure du lissage composite (QE), en fonction de la durée intercensitaire (m)

Le cas de la migration nette à double sens

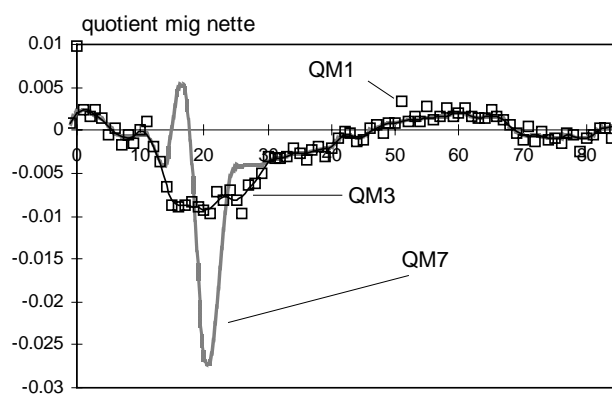
Les populations urbaines sont souvent soumises à des migrations à double sens : avant et après la formation supérieure (18 et 24 ans). Le graphique 4 étudie ce cas, selon la durée de l'observation.

Même dans le cas a priori difficile à observer, notamment quand la durée intercensitaire s'allonge, la procédure du lissage composite semble efficace. Toutefois, il faut noter que le choix des bornes devient plus sensible et donc plus délicat à effectuer dans le cas d'une observation moins lisse que les simulations.

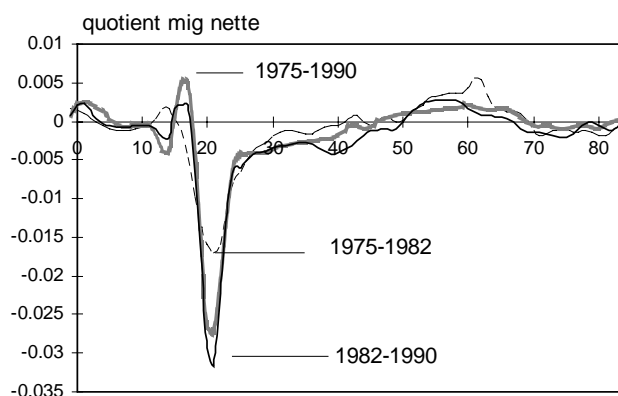
Le cas de la Somme

L'observation d'une population locale réelle ne saurait être aussi fidèle que celle des simulations. D'ailleurs, nous n'avons pas les moyens de mesurer son efficacité. On peut néanmoins juger de sa cohérence, d'une part en comparant la mesure brute avec le résultat du lissage composite ; d'autre part en comparant les estimations des différentes périodes intercensitaires entre elles.

Le département de la Somme est une population suffisamment importante et diversifiée pour que l'observation soit à l'abri d'évolutions trop chahutées. Le graphique 5 montre trois phases dans la mesure des quotients de mortalité de la population féminine de la Somme en 1975-1990. On remarque que dès la première estimation (QM1), la courbe est interprétable : la Somme connaît une émigration nette autour de l'âge de 20 ans, et une immigration nette moins importante autour de l'âge de 60 ans. Le lissage médiane-moyenne (QM3) ne rajoute rien à l'interprétation, mais il est nécessaire pour permettre les délissages ascendant et descendant que l'on retrouve dans le lissage composite (QM7). Le graphique montre, contrairement au cas des simulations, que les délissages concentrent bien la migration à l'intérieur d'un âge déterminé (20-21 ans) qui n'est pas celui choisi par défaut par l'initiation de la procédure (18 ans). Mais, inévitablement, les âges à proximité de l'âge à migration concentrée sont les plus susceptibles d'être perturbés par la procédure du délissage. Ainsi, l'immigration nette à l'âge de 16-17 ans est le résultat « délissé » de la très modeste remontée de QM1 à l'âge de 8-10 ans. A priori, elle n'a rien de vraisemblable.



Graphique 5. Estimation des quotients de migration nette QM1 (estimation brute), QM3 (lissage médiane-moyenne), et QM7 (lissage composite final) ; population féminine du département de la Somme en 1975-1990.



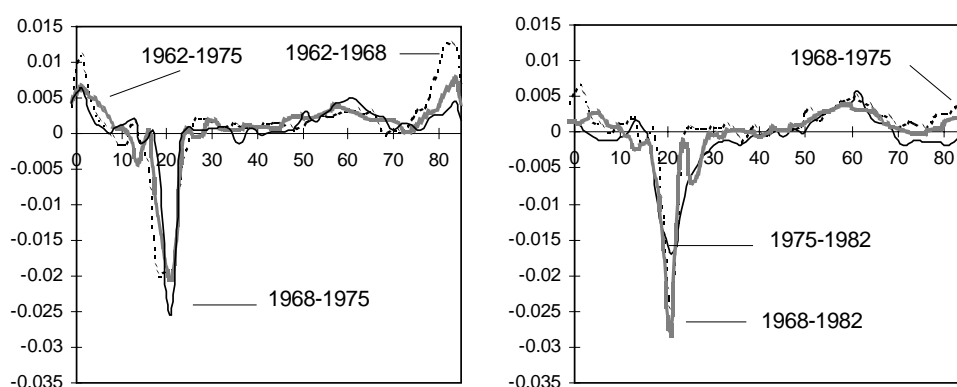
Graphique 6. Lissage composite final des trois périodes d'observation entre 1975 et 1990 ; population féminine du département de la Somme en 1975-1990.

L'observation multiple permet de juger à la fois de la concentration migratoire à 20-21 ans et des perturbations des âges autour (graphique 6). Les trois courbes concernent les trois périodes intercensitaires comprises entre 1975 et 1990 : 1975-1982, 1982-1990 et 1975-1990. Elles confirment l'âge de concentration migratoire (20-21 ans), mais chacune des observations a produit sa propre perturbation de délissage aux âges avant et après l'âge de concentration migratoire. Cela confirme donc le doute qui entoure l'observation aux marges du délissage.

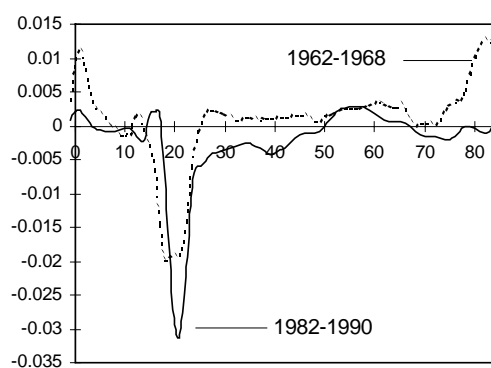
Le graphique 7 montre les observations multiples des périodes 1962-1975 et 1968-1982. Elles montrent les mêmes dessins : émigration nette autour de 20-21 ans (un peu plus tôt en 1962-1968), et une immigration très modérée autour de 60 ans. Aux grands âges (80 ans et plus), il apparaît une immigration nette plus importante, notamment en 1962-1968, mais qui est probablement due à une inadéquation entre données du recensement et

celles de l'état civil¹. Encore une fois, c'est aux marges du déliçage que les estimations sont plus perturbées qu'ailleurs. Et, toujours, la perturbation est plus grande avec les périodes doubles.

Le graphique 8 compare l'observation en 1962-1968 avec celle en 1982-1990. La comparaison confond évolutions réelles avec défauts d'observation, que l'on peut tenter de distinguer avec l'aide des observations intermédiaires des graphiques précédents. Ainsi, la disparition de l'immigration nette aux âges avant 8 ans semble confirmée par les observations intermédiaires, tout comme la transformation de l'immigration nette adulte (25-50 ans) en émigration nette, ce qui est probablement lié à l'évolution aux premiers âges. L'âge de concentration migratoire semble avoir été retardé d'un an, et l'immigration nette autour de 60 ans avoir reculé. Par contre, la disparition de l'immigration nette à 80 ans semble relever des défauts de l'observation, tout comme les variations aux marges du déliçage.



Graphique 7. Lissages composites multiples des périodes 1962-1975 et 1968-1982 ; population féminine du département de la Somme.



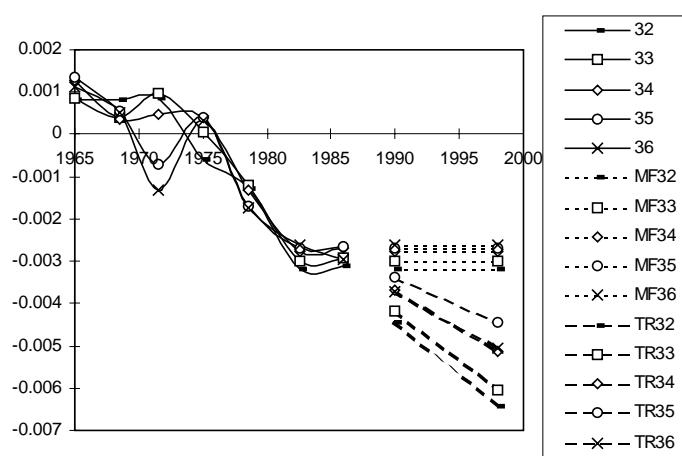
Graphique 8. Lissage composite des périodes d'observation 1962-1968 et 1982-1990 ; population féminine du département de la Somme.

¹ Contrairement au programme Omphale, nous avons utilisé les décès réels plutôt que des décès estimés à l'aide de tables nationales. Toutefois, l'inadéquation constatée concerne probablement le recensement de 1962.

La prospective

L'un des objectifs de l'estimation migratoire intercensitaire est de prévoir l'évolution future de la migration. Si celle-ci a connu une évolution entre 1962-1968 et 1982-1990, alors il est vraisemblable qu'à l'avenir elle ne restera pas immuable. On appellera « migration figée » la reconduction dans l'avenir les quotients de migration nette que l'on a observé au cours d'une période précédente (1982-1990, 1975-1990, ou tout autre période intercensitaire) ; la migration figée correspond donc à l'hypothèse des tendances actuelles utilisée dans Omphale comme hypothèse centrale. Ensuite, on appellera « tendance récente » la reconduction de la tendance linéaire des quotients de migration nette par âge, observée au cours d'une période précédente quelconque, finissant en 1990.

En attribuant les quotients estimés à l'année centrale des périodes d'observation, on peut représenter les quotients estimés de chaque âge par une courbe évoluant dans le temps (graphique 9).



Graphique 9. Quotients de migration nette estimés au cours de la période 1962-1990 et extrapolations jusqu'en 1998 ; MF = Migratoire Figée, TR = Tendance Récente ; population féminine d'âge 32 à 36 ans ; département de la Somme.

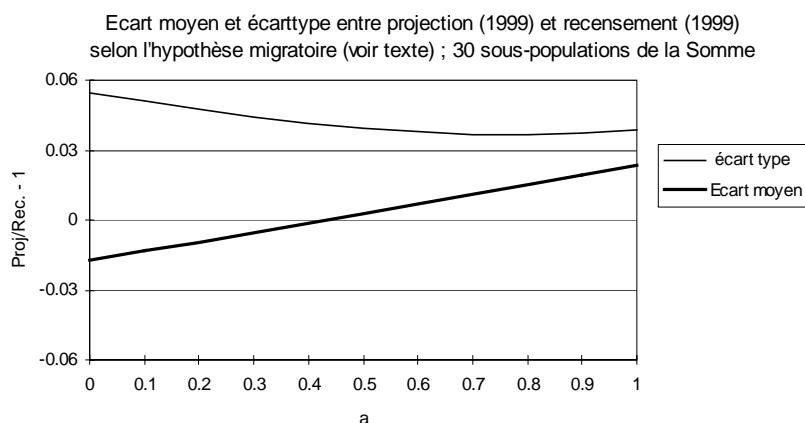
La relative régularité de l'évolution migratoire aux âges représentés se retrouve à un grand nombre d'âges, exception faite pour ceux qui sont aux marges du déliassage. Cela confirme l'intérêt de l'observation temporelle. Le graphique 9 montre à la fois la « migration figée » (1975-1990) et la « tendance récente » (même période) jusqu'en 1998. Pour prévoir la population féminine de la Somme au recensement de 1999, le choix de la migration figée signifierait clairement une rupture avec la tendance récente que l'observation ne justifie pas vraiment. Toutefois, dans d'autres cas l'extrapolation de la tendance récente peut conduire à des valeurs improbables, notamment pour les horizons plus éloignés. Ce sera le cas des populations périurbaines dont le passé récent a été caractérisé par des quotients migratoires exceptionnels, dus aux faibles effectifs présents en début de période.

Si l'on extrapolait les quotients de migration nette par âge selon la tendance récente, la population de la Somme compterait moins de femmes (2,5 % de moins) au RP de 1999 que si l'on reconduisait les quotients observés en 1975-1990 (mêmes hypothèses de mortalité et de fécondité). Or, il ne s'agit pas de choisir entre les deux modèles d'extrapolation. Il s'agit de contrôler l'extrapolation au cas par cas, visuellement, et de la rendre raisonnable, donc raisonnée. La « dose » de tendance récente que nous avons attribuée à l'extrapolation des

quotients départementaux nous a conduit à une différence sur la population totale de 1,5 % en 1999. Ce n'est pas beaucoup. Mais appliquée à la seule population périurbaine de la Somme, la même procédure d'extrapolation raisonnée donne une différence de 9,1 % en 1999.

Nous pensons que la procédure de l'extrapolation raisonnée pourrait améliorer la prospective des collectivités locales dans certaines conditions. Pour tester cette hypothèse, nous allons tirer profit du recensement en cours. Trente sous-populations départementales ont été définies en tirant au sort, pour chaque sous-population, trente communes du département - la commune d'Amiens a été exclue du tirage. Ainsi, les trente sous-populations n'ont pas de caractéristique propre, susceptible d'influencer le choix des paramètres d'extrapolation. Pour chaque sous-population, nous avons calculé la population totale à la date du RP de 1999 de deux façons : par la migration figée (1975-1990) et par l'extrapolation tendancielle (linéaire) des QM mesurés en 1975-1982 et 1982-1990. Les résultats sont comparés aux résultats provisoires du recensement. Le graphique 10 montre la moyenne et l'écart type du rapport entre les 30 populations estimées et recensées en 1999, selon les deux extrapolations, ainsi que les différents mélanges des deux ($a = 0$ pour l'extrapolation tendancielle linéaire ; $a = 1$ pour la migration figée).

Graphique 10.



La meilleure prospective moyenne se situe à $a = 0,42$. C'est-à-dire, presque à mi-chemin entre l'extrapolation tendancielle et la migration figée, un peu plus côté tendanciel. L'écart type le plus réduit, lui, se trouve un peu plus du côté de la migration figée, ce qui est normal, car les extrapolations tendanciennes s'ouvrent vite en éventail. Pour la prospective à un horizon de 9 ans, une extrapolation à mi-chemin entre la migration figée (option d'Omphale) et l'extrapolation tendancielle semble la mieux appropriée.

REFERENCES

DESCOURS Laurence et François POINAT (1992), Le modèle de projection démographique OMPHALE, INSEE Méthodes n° 19, nov. 1992, 64 p.