

La durée de résidence des français et l'évaluation des risques liés aux sols pollués

Nedellec V*, Courgeau D**, Empereur-Bissonnet P*

* Service des Études Médicales d'EDF et de Gaz de France

** Institut National des Études Démographiques (INED)

L'intégration des critères sanitaires dans le processus décisionnel de gestion des sols pollués passe par une évaluation quantitative des risques. Pour les polluants à effet cancérigène, la probabilité de survenue du danger est directement proportionnelle à la durée de l'exposition et à la dose reçue. La durée d'exposition est égale au nombre d'années vécues en tant que résident sur le site pollué. Mais les données du recensement donnent un taux de mobilité reflétant très imparfaitement la durée de résidence. L'analyse longitudinale des données de quittancement des abonnés à EDF permet une estimation fiable de la durée de résidence des ménages, sous certaines hypothèses discutées ici. Entre 1977 et 1996, on observe une augmentation de la mobilité. Ainsi, pour les ménages ayant ouvert un abonnement à EDF en 1977, 50 % résidaient au même endroit au bout de 4,25 années, 25 % au bout de 14,5 années, alors que pour la cohorte de 1993 les estimations de ces résultats sont respectivement 3 ans et 9,5 ans. Ces observations infirment l'hypothèse d'une stabilité du comportement migratoire sous-jacente aux estimations transversales de durée de résidence. Une estimation de la variance n'a que peu d'intérêt ici, puisque les données utilisées sont quasiment exhaustives (95 % des ménages français sont inclus dans l'étude). Un modèle paramétrique ouvre la voie aux estimations pour les cohortes récentes. Ces données françaises sont utilisables en évaluation des risques sanitaires de type générique. Elles réduisent l'incertitude liée à la transposition de données étrangères ou à l'usage de valeurs françaises approximatives. L'effet de leur utilisation est discuté en les comparant aux valeurs jusqu'alors disponibles.

introduction

L'intégration des critères sanitaires, dans les décisions de gestion des sols pollués, passe par une évaluation quantitative des risques. Lorsqu'on prévoit une occupation résidentielle du site, il faut considérer notamment les risques chroniques liés aux expositions de longue durée. Dans ce cas, la durée d'exposi-

tion est égale au nombre d'années vécues en tant que résident sur le site (1).

Pour les substances à effet toxique systémique, il existe généralement un seuil en dessous duquel l'effet n'est pas censé se produire quel que soit le nombre d'années d'exposition. Pour la plupart des substances à effet cancérigène, on considère que les doses

s'accumulent pendant la vie entière. La probabilité de survenue du danger est alors directement proportionnelle à la durée de l'exposition et à la dose reçue pendant cette période (2, 3).

En France, les recensements de la population ne permettent pas une mesure de la durée de résidence. Les seules informations fournies portent sur la résidence au précédent recensement, sans que l'on connaisse la date d'installation dans la résidence actuelle. On peut, avec ces données, estimer un taux de mobilité annuel de 8,6 % (4), un calcul simple fournit une durée de résidence moyenne égale à 11,6 ans¹. Cette approximation est peu satisfaisante car elle suppose l'uniformité du comportement migratoire² tout au long de la période intercensitaire.

Il existe en France d'autres fichiers exploitables, tels que celui de l'entreprise Électricité de France (EDF). Elle bénéficie du quasi-monopole de la distribution d'électricité sur le territoire national. Des travaux antérieurs ont déjà permis des estimations transversales des durées de séjour à partir de cette base de données. Ainsi, une étude sur la durée moyenne d'occupation d'un logement, résidence principale ou secondaire selon la définition de l'INSEE³, a été publiée en 1988(5). En comptabilisant le nombre de logements dont le branchement EDF faisait l'objet d'un

abonnement actif en 1988 (envoi d'une quittance), les auteurs ont trouvé une durée moyenne d'occupation de 10,14 ans. En outre, l'étude montre des disparités régionales dont les extrêmes sont la Corse (13,63 ans) et l'Île-de-France (8,6 ans). Ce calcul transversal, basé sur la durée des contrats en cours une année donnée, suppose aussi l'uniformité du comportement migratoire au cours du temps, car on travaille avec des individus de diverses cohortes n'ayant pas encore migré en 1988. Cette hypothèse, peu cohérente avec les connaissances démographiques, nécessite d'être vérifiée.

L'utilisation de données d'enquêtes biographiques rétrospectives permet d'estimer les durées de séjour de façon longitudinale : la possibilité de suivre un individu fournit les durées de séjour dans les divers logements occupés au cours de son existence (6, 7). La richesse de ces enquêtes biographiques permet de déceler les interactions entre les phénomènes démographiques et les migrations, ainsi que l'effet de l'hétérogénéité de la population sur les comportements migratoires. En revanche, devant le coût de telles enquêtes, les effectifs observés sont faibles, de l'ordre de quelques milliers d'individus. Travaillant sur un échantillon de population, il devient nécessaire d'estimer les variances des quotients⁴ calculés et d'utiliser des méthodes d'analyse des biographies (8).

¹ C'est-à-dire $1/0,086 = 11,62$ ans.

² Dans cet article, le « comportement migratoire » désigne l'activité de changement d'adresse de la population.

³ Institut national de la statistique et des études économiques.

⁴ Ils sont en quelque sorte des taux de déménagements annuels.

Face à la difficulté d'obtenir des données fiables sur la durée de résidence des ménages en France, certains auteurs ont préféré utiliser des valeurs nord-américaines dans des études d'évaluation du risque sanitaire (9). Toutefois, l'extrapolation de la durée de résidence américaine à la population française suppose que le comportement migratoire, la structure par âge et par statut d'occupation soient comparables. Nous tenterons d'apprécier la pertinence de la transposition en comparant nos résultats aux données américaines.

L'objectif de cette étude est d'établir une fonction de durée résidentielle de la population française, par une analyse longitudinale des informations issues du système de facturation d'EDF. Nous discuterons la validité de l'hypothèse d'un comportement migratoire constant au cours du temps sur laquelle se fondent les estimations transversales de ce paramètre.

données du système de facturation EDF

La représentativité du fichier de quittance national est de 95 % des consommateurs individuels d'électricité. Environ 5 % des habitants sont alimentés par des régies indépendantes. Le taux de couverture EDF le plus faible est relevé dans les régions suivantes : Alsace (48 %), Poitou-Charentes (79 %), Picardie et Lorraine (86 %), Aquitaine (90 %).

Le tarif bleu, dit de *clientèle ordinaire*, exclut les administrations, les services publics et les entreprises. Il regroupe les

clients domestiques, les clients agricoles, et les clients professionnels (tertiaires) exerçant leur activité à domicile et ne faisant pas la demande d'une facturation séparée. Pour la France entière, les données agrégées ne permettent pas la distinction entre les 3 types de clientèle. Toutefois, nous disposons d'une extraction au centième permettant de les différencier : la durée moyenne des contrats, calculée en transversal au cours de la période 1971-1996, varie en fonction de la clientèle domestique (13,1 ans), agricole (24,2 ans) et professionnelle (10 ans). La moyenne pour les trois clientèles confondues est de 13,1 ans. Il n'y a donc pas de différence avec le résultat obtenu pour la clientèle domestique uniquement.

On dispose d'un état annuel exprimé en nombre de contrats actifs le dernier jour de l'année, classés par année de création. Par différence avec les chiffres de l'année précédente on obtient le nombre de contrats fermés dans l'année. Ces données servent de base aux calculs des quotients (i.e. taux de déménagement annuel). Les informations ne sont connues que pour les contrats ouverts après 1973, date de la mise en place du système informatique. Par ailleurs, il n'est pas possible de connaître pour chaque contrat le statut d'occupation (locataire ou propriétaire, résidence principale ou secondaire), ni le type d'habitat (date de construction ; zone urbaine, périphérique ou rurale), ni les caractéristiques socioprofessionnelles du contractant. De même nous ne disposons pas du nombre de personnes vivant avec le titulaire.

estimation longitudinale versus estimation transversale

Selon les données dont on dispose, on obtient des estimations très différentes de la durée de séjour. Les données transversales correspondent aux observations d'une année. On demande aux individus leur durée de résidence dans le logement qu'ils occupent au moment de l'enquête, la durée *finie* (i.e. temps écoulé entre l'arrivée et le départ définitif d'un logement) n'est donc pas connue. Ainsi les estimations nord-américaines reposent sur des enquêtes transversales (10). La validité des estimations transversales repose sur l'hypothèse que le comportement migratoire est stationnaire au cours du temps, c'est-à-dire que chaque cohorte observée connaît une mobilité identique aux autres cohortes. Sous cette hypothèse un calcul transversal devient équivalent à un calcul longitudinal des migrations de chacune des cohortes. Mais ce type de calcul suppose aussi que tous les ménages migreront (pour respecter la condition que chaque unité statistique est soumise au risque de connaître l'événement étudié). A cette condition seulement on peut calculer une durée moyenne de séjour.

Pour lever l'hypothèse d'un comportement stationnaire, il faut disposer d'un échantillon longitudinal de durées de séjour *fini*. Ainsi en Norvège les données de Registre de Population permettent de dater au jour près toutes les migrations survenues dans la population (11, 12). Cela permet en plus de travailler, non plus sur les seuls chefs de ménage,

dont la définition n'est d'ailleurs valable qu'à un instant donné, mais sur l'ensemble des individus de la population. Dans ce cas il ne sera généralement plus possible de calculer une durée moyenne de résidence puisque tous les individus n'ont pas encore connu l'événement au moment de l'observation et que, sans doute, une partie ne le connaîtra jamais. En revanche, l'estimation de la durée médiane de séjour, celle de la fonction de séjour et des quotients de migration seront possibles, si l'on observe ces cohortes pendant un temps suffisant. Les données dont nous disposons ici, moins parfaites que les données de registre, permettront cependant d'approcher ces résultats.

méthode

Nous allons d'abord utiliser des méthodes non paramétriques pour estimer les fonctions de séjour, selon la durée T écoulée depuis la précédente installation. T est donc une variable aléatoire, dont on va estimer la fonction de séjour à un instant donné t :

$$S(t) = P(T \geq t) \quad [1]$$

et le quotient instantané :

$$h(t) = \lim_{dt \rightarrow 0} \frac{P(T < t + dt / T \geq t)}{dt} \quad [2]$$

Nous ferons ici l'hypothèse que ces quotients restent constants au cours de chaque période annuelle. L'estimation classique nécessite des données plus complètes que celles dont nous disposons ; elle est présentée afin de montrer les hypothèses que nous devons faire par la suite. Des diagrammes de Lexis (*figure 1*) illustrent ces questions : ils sont construits en

prenant pour abscisse les années successives et pour ordonnée les durées finies de séjour.

Nous suivons une cohorte d'individus et observons au cours du temps la répartition de la migration suivante. La population soumise au risque au 1^{er} janvier de l'année j à 0 heure est notée r_j . Au cours de l'année j les individus peuvent soit quitter leur logement (m_j), soit décéder (d_j). Ce type d'observation, parfait pour une étude longitudinale, est présenté sur la figure 1.a : dans chaque parallélogramme se trouvent les effectifs de décès et de migrations, en début se trouvent les populations soumises au risque. On montre que le quotient instantané, h_j , peut être estimé par la méthode du maximum de vraisemblance (8) :

$$h_j = -\frac{m_j}{m_j + d_j} \log\left(1 - \frac{m_j + d_j}{r_j}\right) \quad [3]$$

On a également la possibilité d'estimer la fonction de séjour et le quotient cumulé, ainsi que les variances de tous ces estimateurs. Cependant les variances n'ont guère d'intérêt ici, étant donné que les observations sont quasiment exhaustives.

Les informations issues du fichier EDF se présentent différemment. Elles ne permettent pas de distinguer les sorties par nouvelle migration ou par décès des chefs de ménage. Dans ce cas on ne pourra estimer que des quotients de sortie sans distinguer la cause (migration ou décès) :

$$h_j = -\log\left(1 - \frac{m_j + d_j}{r_j}\right) = -\log\left(1 - \frac{s_j}{r_j}\right) \quad [4]$$

De plus les nombres de départs sont enregistrés par année de survenue de l'événement et

non par durée de séjour. Les départs de l'année initiale ne sont pas fournis : les individus ayant effectué une ou plusieurs migrations pendant cette année ne sont comptés qu'à partir de la dernière. La figure 1.b présente ces données. La population soumise au risque est maintenant fournie à la fin de chaque année d'observation. On voit qu'elles ne permettent pas d'estimer les quotients de sortie de la formule [4], sans hypothèses supplémentaires.

La figure 1.c présente la plus simple de ces hypothèses, qui conduit à travailler sur des intervalles différents des précédents. En effet, on peut supposer que l'effectif qui atteint la durée exacte $j - \frac{1}{2}$, égal à $r'_j + d'_{j-1} - d_{j-1}$, diffère peu de la population r'_j , lorsque d'_{j-1} et d_{j-1} sont approximativement égaux. Quant aux migrations et aux décès entre les durées exactes $j - \frac{1}{2}$ et $j + \frac{1}{2}$, ils sont à peu près égaux à ceux entre le 1^{er} janvier et le 31 décembre, lorsque d_j est voisin de d'_j (13). Il en résulte une estimation approchée du quotient pendant cet intervalle :

$$\hat{h}_{j-\frac{1}{2}} = -\log\left(1 - \frac{s'_j}{r'_j}\right) \quad [5]$$

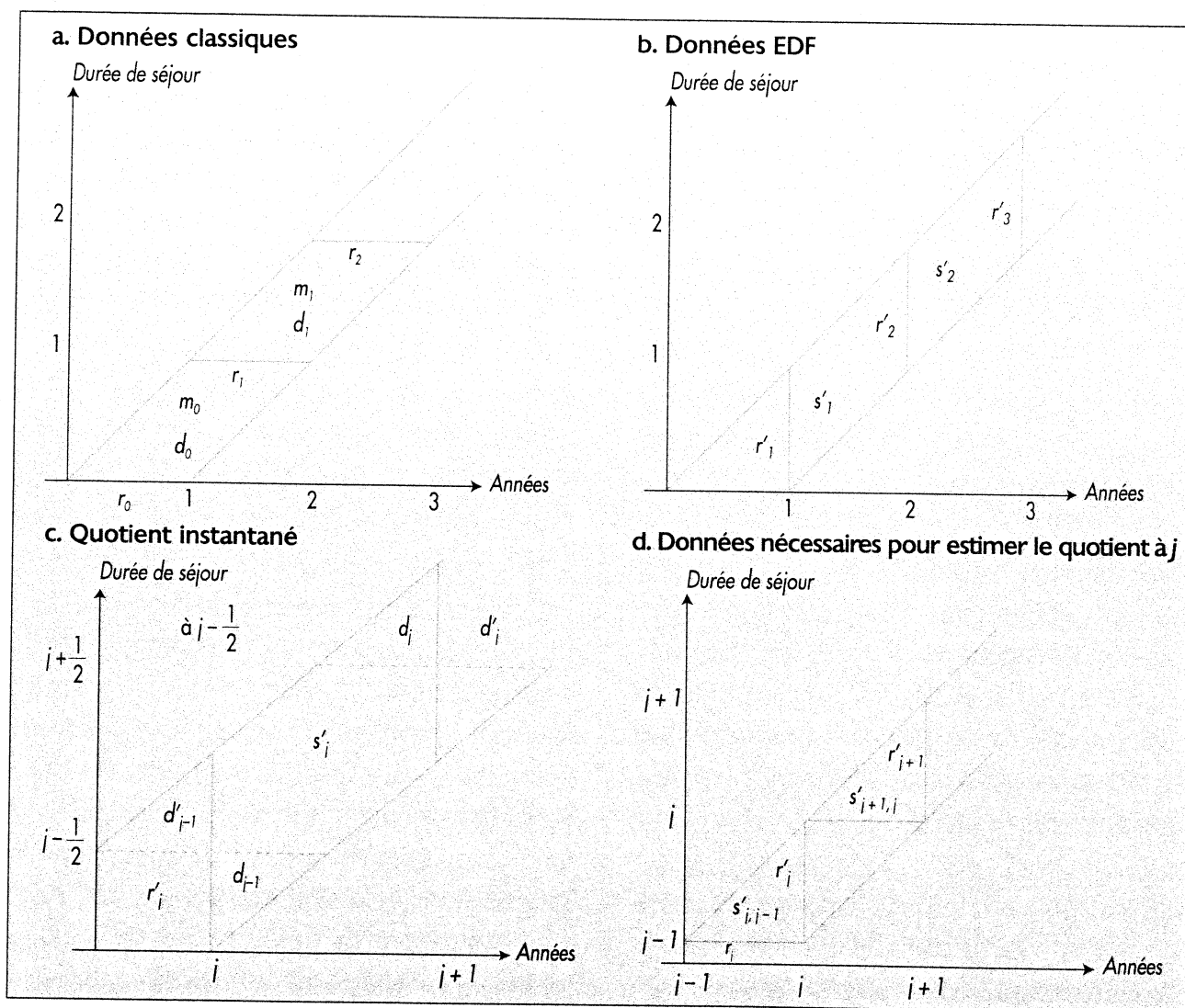
Si les courbes des quotients en fonction de l'âge sont suffisamment régulières, on a la possibilité d'estimer \hat{h}_j par interpolation linéaire entre $\hat{h}_{j-\frac{1}{2}}$ et $\hat{h}_{j+\frac{1}{2}}$. Par contre pour les fonctions de séjour, comme on ignore les événements qui se produisent dans le parallélogramme compris entre les durées 0 et 0,5 an, on devra faire des hypothèses supplémentaires. La plus simple, utilisée ici, consiste à prendre l'instant initial à la durée 0,5 an. Une autre hypothèse serait de supposer que le quotient à appliquer pour ces durées est égal,

par exemple, à $\frac{\hat{h}_{0,5}}{2}$ et d'estimer ainsi $\hat{S}_{0,5}$.⁵ Si l'on désire une estimation plus précise, il faudrait disposer des données dans le découpage plus fin de la figure 1.d. Le calcul d'une vraisemblance correcte nécessite la connaissance de la population initiale et des événements situés dans le premier triangle de la figure 1.b : on les ignore ici. Nous n'irons donc pas plus avant dans cette estimation non paramétrique et retiendrons la plus simple.

résultats de l'estimation non paramétrique

L'utilisation de la formule [5] nous permet d'estimer les quotients en fonction de la durée de séjour pour chacune des cohortes 1977-1995, portés sur la figure 2. Nous avons simultanément estimé les fonctions de séjour, en partant d'une durée initiale, en fait égale à 0,5 an, sur la figure 3.

Figure 1 - Diagrammes de Lexis représentant les différents jeux de données pour l'estimation des quotients instantanés



⁵ Voir Smith (14), pour ces diverses méthodes d'estimation.

Figure 2 - Quotients pour les cohortes installées de 1977 à 1995

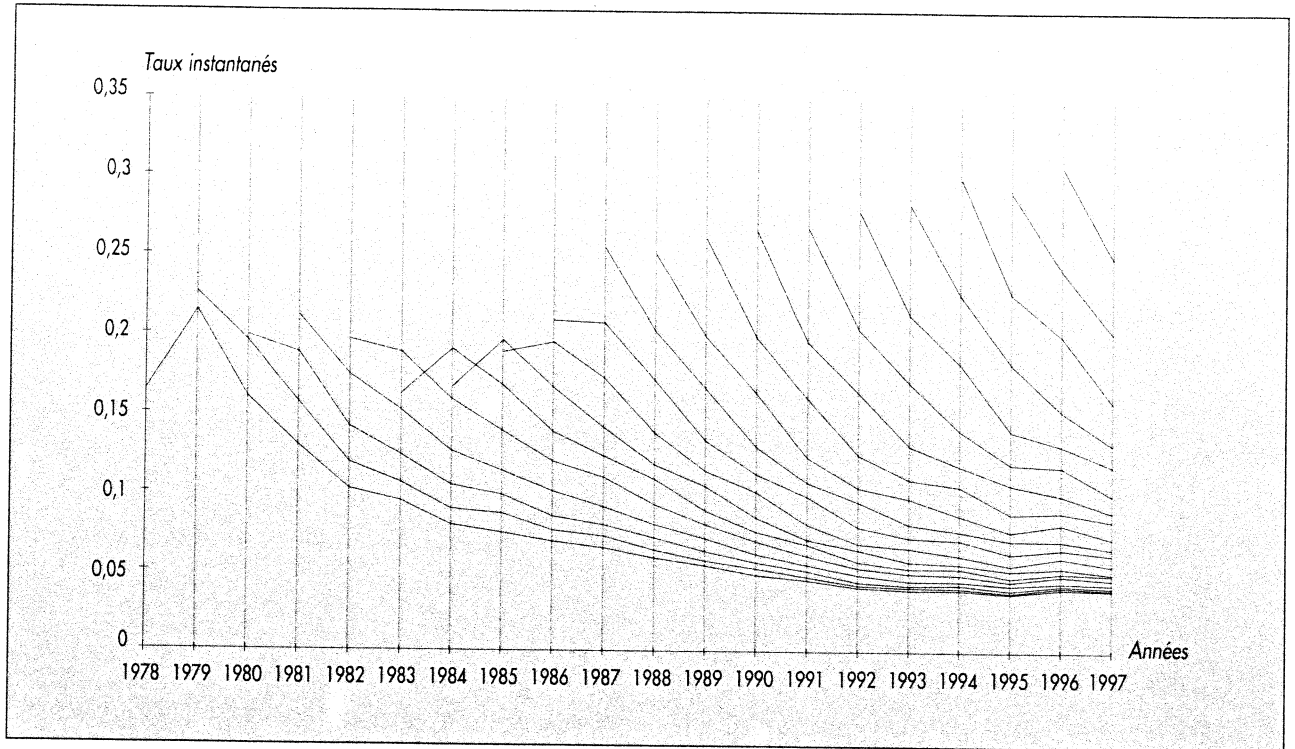
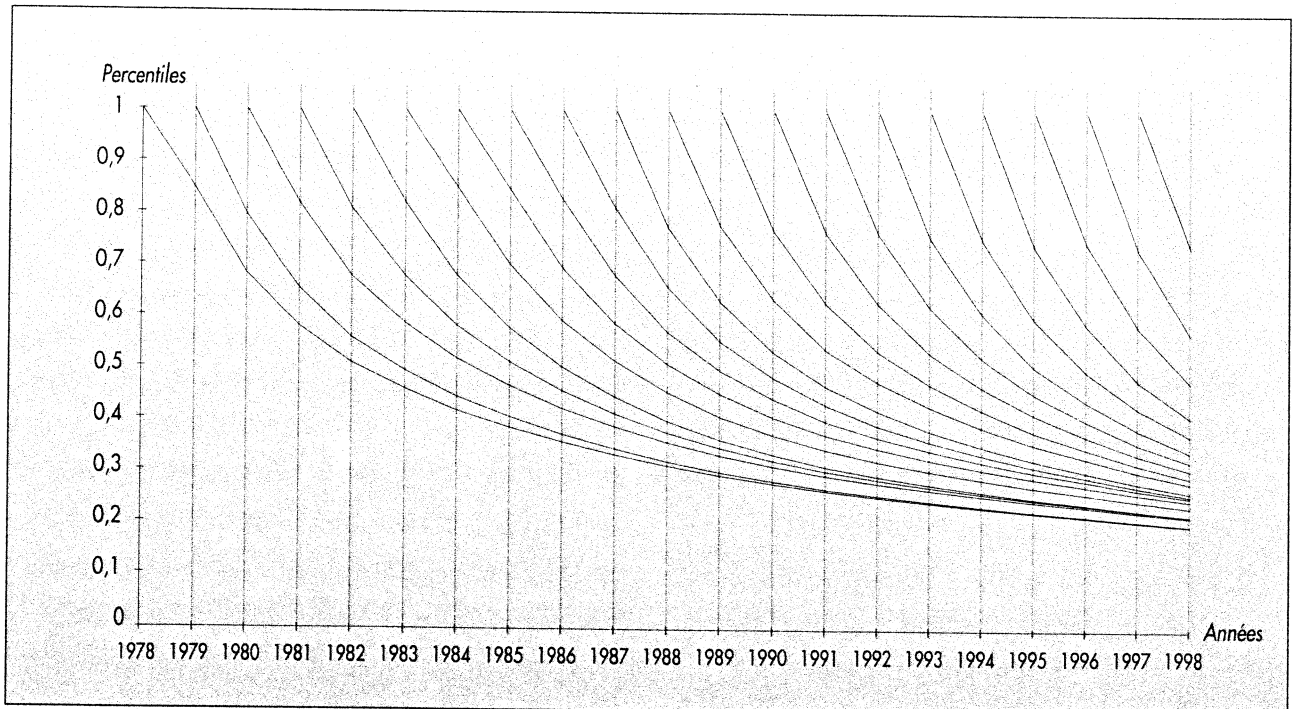


Figure 3 - Fonctions de séjour pour les cohortes installées de 1977 à 1995



Les courbes donnant les quotients (*figure 2*) varient suffisamment pour démentir l'hypothèse d'un comportement identique des diverses cohortes. En premier lieu, on peut distinguer 2 périodes : de 1977 à 1984, où les premiers quotients annuels tendent à diminuer d'une cohorte à l'autre, alors que dans la deuxième période, de 1985 à 1993, la tendance serait plutôt à la croissance. Les quotients initiaux fluctuent autour de 0,2 dans la première période et augmentent jusqu'à atteindre 0,3 lors de la cohorte de 1993 (deuxième période). Ensuite, des fluctuations annuelles importantes sont notées pour les courtes durées de résidence. Par exemple, pour les cohortes 1977, 1982, 1983 et 1984 le quotient instantané de la première année est inférieur à celui de la deuxième année. En revanche, les quotients de toutes les cohortes étudiées décroissent de façon continue quand la durée augmente. Cette décroissance se ralentit, et les quotients semblent atteindre une valeur limite à peu près constante après une durée de 13 ans, proche de 0,05.

L'examen des courbes donnant les fonctions de séjour (*figure 3*) vient compléter ces observations. On voit en premier lieu que la durée médiane de séjour est très courte, mais qu'elle passe d'une valeur proche de 4,25 ans pour les premières cohortes à une valeur inférieure à 3 ans pour les dernières. On voit combien cette durée est différente de la durée moyenne estimée en transversal, qui est autour de 10 ans en 1988 (5). En fait, alors que l'on peut calculer une durée moyenne en transversal, puisque l'on dispose des durées de séjour de chaque migrant d'une année donnée, cette durée moyenne n'a plus aucun sens

en longitudinal, car certains individus d'une cohorte donnée peuvent ne pas avoir encore migré à la date d'observation et pourront même ne jamais migrer après. On voit sur nos observations qu'après une durée de 20 ans, environ 20 % de la population n'a pas encore migré. Cependant, du fait que l'on mélange ici migrations et décès des chefs de ménage, la population d'une cohorte finira par s'éteindre du fait de la mortalité. Si en revanche on avait la possibilité d'éliminer cet effet de la mortalité, avec des données du type de celles de la *figure 1.a*, il est évident que la fonction de séjour pourrait ne jamais atteindre la valeur 0 : la durée moyenne de séjour tendrait alors vers l'infini lorsque la durée d'observation augmente. La durée médiane de séjour, en revanche, est toujours parfaitement définie.

discussion

Ces résultats, bien qu'obtenus avec des données quasiment exhaustives, présentent un certain nombre d'incertitudes.

En premier lieu, si la notion de chef de ménage a un sens précis au moment de l'emménagement initial, celle-ci perd de sa pertinence au fur et à mesure que la durée de séjour va augmenter. Prenons par exemple un ménage constitué au départ d'un couple et de ses enfants. Si au cours du temps le couple se sépare mais que le chef de ménage reste seul dans son logement, le fichier EDF n'enregistrera aucun départ, alors que son ex-compagne et ses enfants ont fait une migration. Si au contraire le chef de ménage quitte ce logement pour le laisser à son ex-compagne et à ses enfants, le fichier EDF enregistrera alors sa migration. Le ménage initial, lui, n'existe plus après l'un ou l'autre de ces départs.

En second lieu, l'incertitude de ces données vient de ce qu'elles ne fournissent pas l'ensemble des effectifs nécessaires au calcul correct d'une fonction de séjour. La méconnaissance de l'effectif du premier triangle de la figure 1.b, et la mesure des effectifs de migrants par année de migration et non par durée de séjour, ne permettent qu'un calcul approché des quotients de migration sur des périodes inhabituelles, allant du milieu d'une année au milieu de la suivante. Des hypothèses sont nécessaires pour mieux ajuster ces résultats, mais la connaissance imparfaite des migrations de chefs de ménage, comparée à celle de la mortalité des populations humaines, rend ces hypothèses hasardeuses.

Nous allons cependant essayer de prévoir la migration des chefs de ménage à l'avenir. Nous baserons cette tentative sur des résultats graphiques sans tenter une recherche plus analytique d'un modèle prévisionnel.

une prévision des migrations à venir

Divers essais graphiques nous ont montré que les modèles qui s'adaptent habituellement bien aux migrations individuelles (modèle de Gompertz, modèle migrants-sédentaires) étaient moins satisfaisants pour les données par ménage. Parmi les meilleurs ajustements empiriques, on constate qu'en portant la racine carrée de l'opposé du logarithme de la fonction de séjour en fonction du logarithme de la durée, on obtient des droites à peu près parallèles pour chaque cohorte. Cela revient à écrire la fonction de séjour sous la forme paramétrique suivante :

$$S(t) = e^{-\beta^2(\log(\alpha) + \log(t))^2} \quad [6]$$

L'estimation des paramètres α et β est faite manuellement sur les courbes elles-mêmes. La figure 4 porte d'une part la fonction de séjour estimée pour la cohorte 1977, que l'on peut entièrement comparer à la fonction observée, déjà donnée dans la figure 3, et celles estimées pour les cohortes 1986 et 1993, qui vont permettre de prévoir le comportement à venir de ces individus. On constate d'abord que l'ajustement est très correct. Ces courbes nous montrent dès lors une très forte baisse des pourcentages de non-migrants, que l'on peut prévoir au bout de 20 ans, qui passent de 19,94 % pour la première cohorte à 14,94 % pour la dernière. Les valeurs de $\log(\alpha)$ et de β pour les cohortes de 1977, 1986 et 1993 sont respectivement : 1,45 et 0,286 ; 1,57 et 0,285 ; 2 et 0,276.

une utilisation en évaluation du risque sanitaire

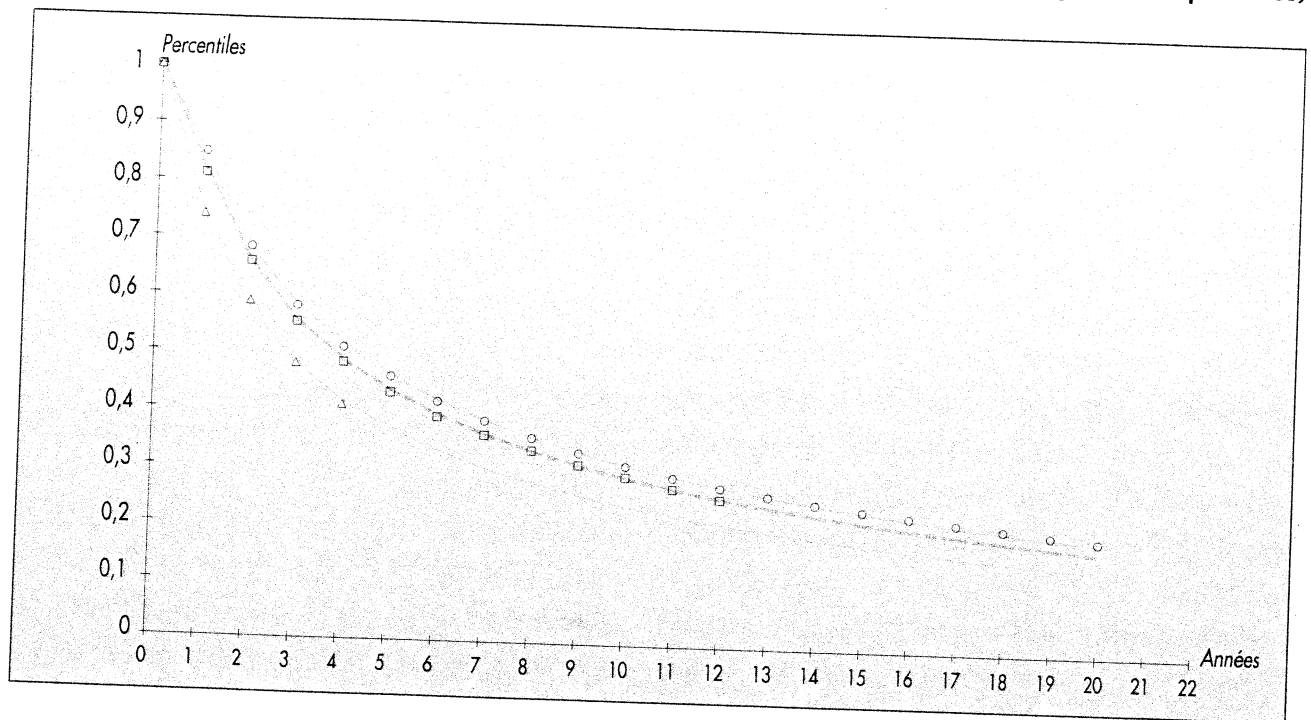
La quantification de l'excès de risque cancérigène, en relation avec une exposition humaine à une substance toxique, est fondée sur l'équation générale suivante :

$$ERI = \left[DJE \times \frac{DE}{PT} \right] \times ERU \quad [7]$$

avec :

- ERI = Excès de risque individuel (vie entière)
- DJE = Dose journalière d'exposition (moyenne annuelle ; en mg/kg.j)
- DE = Durée d'exposition (année)
- PT = Pondération temporelle, par convention : 70 ans (vie entière), pour les effets cancérigènes
- ERU = Excès de risque unitaire (mg/kg.j)⁻¹

Figure 4 - Fonctions de séjour observées (points) et estimées (courbes) pour les cohortes installées en 1977 (points ronds et ligne), 1986 (points carrés et gros pointillés), 1993 (points triangles et fins pointillés)



Pour attribuer une valeur numérique à la durée d'exposition, en France, les évaluateurs disposaient jusqu'à présent soit de l'estimation du ministère du Logement, soit de données d'origine étrangère, en particulier des USA. Les principales informations actuellement disponibles dans la littérature, sélectionnées de manière à prendre en compte des indicateurs de distribution de même nature, sont résumées dans le tableau I. Dans les 3 premières études, les moyennes sont calculées sur la durée actuelle de résidence, c'est-à-dire le temps écoulé entre l'arrivée dans un logement et la date de l'enquête transversale. Israeli et Nelson ont modélisé leurs données pour obtenir une fonction de séjour leur permettant d'estimer la distribution des durées finies de résidence (chiffres en italique dans le tableau I). Les auteurs notent que les résultats sont fortement déséquilibrés avec une moyenne numérique-

ment plus élevée que le 75^e percentile de la distribution. Johnson et Capel (15), avec la même base de données, mais en utilisant une approche probabiliste, obtiennent des résultats moins décalés, la moyenne étant tout de même supérieure à la médiane, l'influence des durées de résidences longues reste encore visible dans ce type de calcul. Nous avons aussi produit une fonction de séjour pour la durée finie de résidence, mais par une analyse longitudinale de données observées. Nous présentons dans le tableau I les résultats observés sur la cohorte 1977 et les données estimées, au moyen de la fonction de séjour paramétrique, de la cohorte 1993.

Nous discuterons ici non pas de la qualité respective de ces informations, mais de l'effet de leur utilisation sur les résultats d'une évaluation de risque. On s'aperçoit ainsi que l'écart

numérique entre la moyenne calculée par le ministère français du Logement et celles établies aux USA est très faible. L'utilisation de l'une ou l'autre valeur est alors sans incidence majeure sur l'ERI, le rapport entre l'estimation française et les estimations nord-américaines n'étant, respectivement, que de 0,86 et 0,95. En revanche, les différences s'accroissent si l'on raisonne cette fois sur les médianes. En prenant dans un premier temps pour référence les données françaises de la cohorte 1977, la description du paramètre « durée de résidence » par la valeur d'Israéli entraîne une sous-estimation de l'ERI d'un facteur 3. La valeur proposée par Johnson, au contraire, conduit à surestimer l'ERI d'un facteur égal à 2,1. La prise en compte du 75^e percentile accentue ce phénomène, car dans ce cas les données de l'étude n° 2 (tableau 1) font estimer un risque de cancer près de 4 fois inférieur à celui que l'utilisation de nos données aurait permis de calculer.

Si nous nous référons maintenant à la fonction de séjour de la cohorte 1993, reflétant la réduction observée de la durée de résidence des Français, les écarts avec les valeurs américaines sont également importants. Ainsi, les valeurs d'Israéli font sous-estimer l'excès de risque d'un facteur compris entre 2 et 3 selon que l'on choisit la médiane ou le 75^e percentile, tandis que les données de l'étude n° 3 (tableau 1) surestiment le risque, respectivement, d'un facteur 3 et 1,7.

conclusions

L'utilisation des données des fichiers EDF nous a permis, sous certaines hypothèses, d'estimer les fonctions de séjour des ménages français dans leur logement. Nous avons ainsi pu voir que cette mobilité a fortement augmenté pour les cohortes ayant fait une migration antérieure après 1986. On peut penser que cette augmentation apparaîtra dans les

Tableau 1 – Synthèse des données françaises et américaines sur la durée résidentielle

Source	Base de données utilisée	Type d'analyse	Valeurs numériques disponibles (en années)		
			MOYENNE*	50 ^e perc.	75 ^e perc.
1. Ministère du Logement (1988, France)	facturation EDF	T	10,1	–	–
2. Israéli et Nelson (1992, USA)	bureau of the US census	T modélisé	10,6 4,5	1,4	3,7
3. Johnson et Capel (1992, USA)	bureau of the US census	T	11,7	9	16
4. Nedellec, Courgeau et Empereur (1998, France) Cohorte 1977 (observés) Cohorte 1993 (75 ^e perc modélisé)	facturation EDF	L	–	4,2	14,5
		L	–	3,0	9,5

T = analyse transversale ; L = analyse longitudinale.

* Moyenne de la durée actuelle de résidence.

résultats du prochain recensement en 1999, même si nous avons vu que les mesures effectuées avec les deux sources étaient très différentes. Cela constituera, de toutes façons, un test de la valeur des données EDF, pour mettre en évidence une évolution dans la mobilité des Français sans avoir à attendre les résultats d'un recensement.

Nous avons également pu voir une assez grande régularité des courbes des fonctions de séjour, qui pourrait permettre des estimations du comportement futur de migrants antérieurs. C'est à nouveau un résultat appréciable, pour juger en particulier de l'effet de politiques du logement, ou de l'effet de crises économiques, sur le comportement migratoire des ménages.

Ces données françaises sont utilisables en évaluation des risques sanitaires de type générique, c'est-à-dire sans tenir compte de spécificité géographique ou socio-économique d'une sous-population définie. Elles réduisent l'incertitude liée à la transposition de données étrangères ou à l'usage de valeurs françaises approximatives. Toutefois, pour que ces données puissent être utilisables dans le cadre d'études plus spécifiques, il faut pouvoir stratifier les cohortes selon des divisions géographiques du territoire national. La mise en place de statistiques plus détaillées, par département par exemple, est possible, à un coût qui ne devrait pas être très élevé. Cela permettrait une meilleure intégration du comportement résidentiel local, dans les décisions de gestion des sites pollués.

Mots clés

Mobilité habitat. Logement. Polluants sol. Évaluation risque. Recensement.

remerciements

Cette étude a été rendue possible grâce au Projet de Développement de Service à la Clientèle Courante de la Direction Générale d'EDF GDF Services, qui a fourni les bilans annuels de facturation EDF. Nous les remercions pour leur collaboration constructive.

références

1. Wood P, Ohilips L, Adenuga A. Population mobility. Chapter 5.4. In : *Exposure Factors Handbook* : US-EPA, 1989 : p. 68-86.
2. Lagoy PK. *Risk assessment : Principles and applications for hazardous waste and related sites*. Park Ridge : Noyes Publications, 1994.
3. US-EPA. *Risk assessment guidance for Superfund*. Washington : USEPA, 1989.
4. Baccaïni B, Courgeau D, Desplanques G. Les migrations internes en France de 1982 à 1990. Comparaison avec les périodes antérieures. *Population* 1993 ; 6 : 1771-90.
5. Ministère de l'Équipement et du Logement. La durée moyenne d'occupation d'un logement est de 10 ans. *Notes vertes économiques* 1988 ; 20 : 1-4.
6. Courgeau D. Migrants et migrations. *Population* 1973 ; 1 : 95-130.
7. Courgeau D. Interaction between spatial mobility, family and career life-cycle : a French survey. *European Sociological Review* 1985 ; 1 : 139-62.
8. Courgeau D, Lelièvre E. *Analyse démographique des biographies*. Paris : INED, 1989.

9. Empereur-Bissonnet P. Usine à gaz de Gennevilliers. Évaluation des risques sanitaires après réhabilitation du site. Rapport final d'étude. Paris : EDF-GDF/SEM, 1996.
10. Israeli M, Nelson CB. Distribution and expected time of residence for US households. *Risk Analysis* 1992 ; 12 : 65-72.
11. Ginsberg R. Timing and duration effects in residence histories and other longitudinal data : I Stochastic and statistical models. *Regional Sciences and Urban Economics* 1979 : 311-31.
12. Ginsberg R. Timing and duration effects in residence histories and other longitudinal data : II Studies of duration effects in Norway, 1965-1971. *Regional Sciences and Urban Economics* 1979 : 369-92.
13. Henry L. Démographie, analyse et modèles. Paris : Larousse, 1972.
14. Smith D. Formal demography. New York : Plenum Press, 1992.
15. Johnson T, Capel J. A Monte Carlo approach to simulating residential occupancy periods and its application to the general US population : Recherche Triangle Park, US-EPA, Office and Air Quality Standards, 1992.