

La modélisation des ménages complexes et des communautés – Exemple de demo4 : module démographique du projet SustainCity

Sophie Pennec, Élisabeth Morand et Laurent Toulemon¹

Publié dans Recueil du Symposium 2014 de Statistique Canada : Au-delà des méthodes traditionnelles d'enquêtes : l'adaptation à un monde en évolution

Résumé

Dans le cadre du projet européen SustainCity, un modèle de microsimulation des individus et des ménages a été créé en vue de simuler la population de différentes métropoles européennes. L'objectif du projet était d'unir plusieurs modèles de microsimulation de transports et d'occupation des sols (land use modelling)^o, de leur adjoindre un module dynamique de population et d'appliquer ces approches de microsimulation à trois zones géographiques d'Europe (la région d'Île de France et les agglomérations de Bruxelles et Zurich).

Mots Clés : Microsimulation dynamique, population.

1. Introduction

1.1 Le projet SustainCity

Le projet européen SustainCity (<http://www.sustaincity.org/>) a pour objectif d'unir plusieurs modèles de microsimulation de transports et d'occupation des sols (land use modelling) comme URBASIM, MATSIM, METROPOLIS et de leur adjoindre un module dynamique de population pour créer URBASIM-E, adapté aux métropoles européennes, et d'appliquer ces approches de microsimulation à trois zones géographiques d'Europe (la région d'Île-de-France et les agglomérations de Bruxelles et Zurich).

La présentation s'intéresse au modèle démographique construit : demo4. Le modèle a été conçu comme un module dynamique autonome au sein du projet SustainCity et peut donc être utilisé pour des applications purement démographiques.

1.2 Les microsimulations

Par rapport aux modèles de simulation plus traditionnels, les modèles de microsimulation se caractérisent tout d'abord par le fait qu'ils se situent au niveau individuel et non à un niveau agrégé et donc les résultats au niveau macro résultent de la somme des comportements des individus micro. Ces comportements micro sont basés sur des probabilités de transition d'une situation vers une autre en fonction de ces caractéristiques individuelles et d'un tirage aléatoire pour déterminer si chaque événement survient ou non.

À la fin des années 1950, pour éviter les erreurs liées aux modèles de simulations utilisant des agrégats et pour mieux prendre en compte la complexité de la réalité, Orcutt (1957) propose un forme de modélisation of « various sorts of interacting units which receive inputs and generate outputs », inputs being « anything which enters into, acts upon, or is taken into account of, by the unit » and outputs « anything which stems from, or is generated by, the unit ». Son objectif principal était de mieux simuler les variantes de politiques sociales et en particulier la diversité des situations individuelles. Avec ce type de modèle, il est possible de vérifier si les politiques touchent bien les personnes qu'elles visent et s'il n'y a pas d'effets contreproductifs (une prestation qui par exemple fait perdre

¹ Sophie Pennec, Institut national d'études démographiques (Ined), 133 boulevard Davout, 75980 Paris Cedex 20, France (pennec@ined.fr); Élisabeth Morand, Ined (elisabeth.morand@ined.fr); Laurent Toulemon, Ined (toulemon@ined.fr).

d'autres prestations ou augmente des prélèvements et entraîne involontairement une dégradation de la situation de certaines personnes), permet de connaître les gagnants et les perdants d'un changement de politique. Cette forme de modélisation permet de produire des résultats en distribution et non en moyenne.

Cette même logique s'applique lorsque l'on s'intéresse à la répartition des ménages et des individus en vue d'une localisation. Il est nécessaire de connaître non seulement les caractéristiques moyennes des ménages, mais aussi leur distribution selon différentes caractéristiques déduites de celles des membres qui les composent.

Cette communication est consacrée à la présentation du modèle démographique construit dans le cadre du projet Sustaincity, notamment les choix effectués pour répondre aux contraintes du projet. Nous présentons aussi la manière dont est faite la simulation des ménages collectifs et des ménages complexes, formes de ménage souvent peu simulés dans les modèles.

2. Le modèle démographique

2.1 Caractéristiques du modèle

Le modèle démographique du projet SustainCity a été développé sous un certain nombre de contraintes : le module devait pouvoir s'articuler avec le modèle d'aménagement du territoire Urbansim, il devait être suffisamment simple d'utilisation pour être utilisé par les autres équipes du projet et suffisamment flexible pour intégrer les contraintes en termes de données des autres partenaires des agglomérations de Zurich et Bruxelles, après son calibrage sur la région parisienne d'Île-de-France. Enfin il devait être gratuit et basé sur des logiciels libres. Nous avons souhaité que le modèle soit indépendant de la plateforme d'urbansim, écrite en python. Deux raisons ont guidées ce choix : cela permettait de tester et modifier le modèle plus facilement ; nous souhaitions aussi pouvoir utiliser le modèle pour des usages ultérieurs sans lien avec Urbansim.

Le langage Modgen développé par Statistiques Canada a été choisi car il offre un environnement de programmation de modèle de microsimulation stable, mis à jour régulièrement. Il permet de créer une interface facile d'utilisation, ce qui nous a paru très utile pour que les partenaires du projet puissent utiliser le modèle sans avoir de connaissance en informatique, y compris saisir les données initiales de simulation et changer les paramètres de la simulation.

Le modèle démographique pour SustainCity est :

- Un modèle transversal : il prend comme population initiale une population représentative de la population étudiée et fait évoluer cet échantillon au cours du temps de manière à avoir des états la population à différents points du temps ;
- un modèle fermé : La taille et la structure de la population ne se modifie qu'en fonction des événements qui surviennent aux membres de la population : on entre dans la population par naissances ou immigration et on en sort par décès ou émigration ; de plus, les liens de parenté entre les membres des ménages sont explicitement créés et conservés au cours de la simulation. Cela entraîne que les mises en union se font à l'intérieur de la population et qu'il faut donc créer une forme de marché matrimonial pour former les couples ;
- un modèle à pas discret. Le pas de simulation est annuel, on utilise donc des transitions annuelles pour simuler les comportements ;
- un modèle qui simule à la fois les individus et les ménages : les changements de structure des ménages résultent des événements survenus aux membres du ménage ; les nouveaux ménages proviennent de départ des ménages existants (enfants qui quittent le foyer parental, séparation de couple), la disparition d'un ménage suit le départ de l'ensemble des membres vers un autre ménage lors d'une union, ou le décès d'une personne seule dans son ménage

2.2 Événements simulés

Le modèle simule deux types d'acteurs : les personnes et les ménages. Chaque personne de l'échantillon appartient à un ménage et à un seul (sauf les personnes en ménage collectif qui ne sont reliés à aucun ménage spécifique). Trois

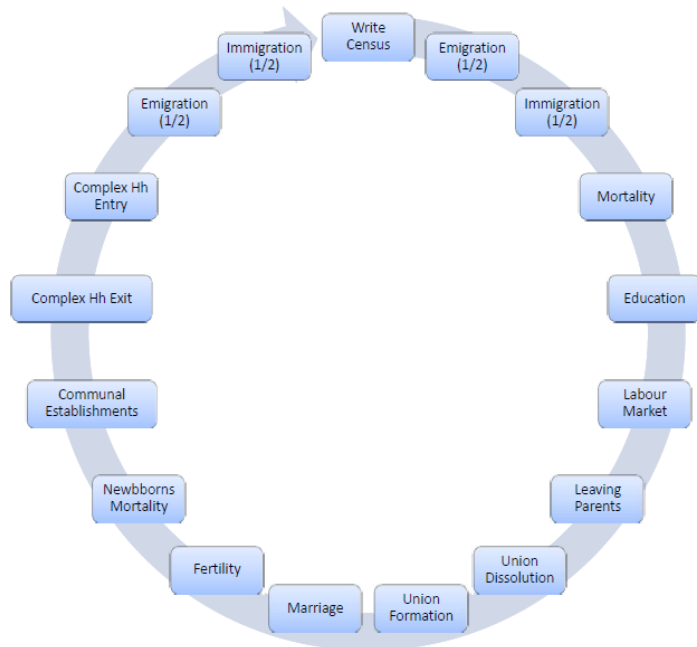
formes de liens entre personnes sont créées : partenaire, mère et père. Une personne de référence du ménage est définie. Le vieillissement de la population, c'est-à-dire tous les changements que peuvent connaître chaque personne ou ménage, dépendent de transitions annuelles à pas discret. L'âge étant la seule variable évoluant de manière déterministe. Ces probabilités varient selon différentes caractéristiques de la personne, restreintes pour certains événements au sexe et à l'âge. Certaines caractéristiques varient au cours du temps et sont donc estimées pour chaque année de la simulation en fonction des évolutions connues. Les utilisateurs du modèle peuvent choisir de conserver les probabilités constantes au cours du temps ou d'intégrer des estimations pour chaque année. De plus, certains événements ne dépendent pas uniquement de la probabilité de le connaître mais aussi d'événements affectant d'autres personnes (le veuvage résulte du décès du partenaire) ou de caractéristiques d'autres personnes (comme pour la mise en union).

Même si les modèles de microsimulation utilisent des chaînes de Markov et les transitions basées sur la situation actuelle, cette limite potentielle est en général dépassée avec l'utilisation de caractéristiques biographiques et des variables intégrant des informations de durée. C'est ainsi que ce type de modèle prend en compte le comportement passé des personnes et certaines des variables d'intérêt (ainsi la probabilité d'avoir un enfant dépend du nombre antérieur d'enfants de la femme ainsi que de l'âge du dernier enfant, ce qui permet de contrôler les écarts entre les naissances).

Le choix des événements simulés reflète deux contraintes ; l'une est de disposer des variables nécessaires pour simuler l'évolution démographique des personnes et des ménages; l'autre tient à la disponibilité des données dans toutes les régions devant être simulées dans le cas du projet Sustainability.

Deux variables permettent de tenir compte de l'hétérogénéité sociale, chacune divisée en trois classes : le niveau d'éducation – bas (primaire), moyen (secondaire) et élevé (universitaire), et le lieu de naissance : dans la région, dans une autre région du pays, à l'étranger. Pour les migrants, la durée depuis l'arrivée peut être utilisée comme variable complémentaire. La participation au marché du travail est binaire : est sur le marché du travail (actif ou au chômage) ou n'est pas sur marché du travail (étudiant, inactif ou retraité).

Figure 2.2-1
Modèle démographique : ordre de simulation des événements



Certains événements qui impliquent plusieurs personnes sont dans ce modèle « gérés » par les femmes (women driven) ; c'est le cas des mises en union ou de la fécondité. Ainsi, pour l'entrée en union, on détermine parmi les femmes hors union celles qui vont se mettre en union (par une probabilité d'entrée en union variant selon l'âge et la

situation dans le ménage, et un tirage stochastique pour déterminer parmi les femmes éligibles celles qui vont effectivement entrer en union), et ensuite on crée un marché matrimonial pour appairer ces femmes avec un conjoint parmi l'ensemble des hommes hors union à l'aide de techniques d'optimisation sous contraintes. On sélectionne un ensemble de conjoints potentiels, on calcule la similarité avec chacun de ces conjoints potentiels, dépendant des écarts d'âge, du niveau d'éducation, de lieu de naissance, puis on sélectionne au hasard un conjoint en fonction des similarités.

Les migrations se font au niveau du ménage, tous les membres du ménage immigreront ou émigreront ensemble. Les immigrations se font selon le profil observé des migrants récents dans la région.

Les événements sont simulés dans l'ordre illustré par la figure 2.2-1. Cet ordre a été élaboré de manière à réduire l'effet d'événements concurrents. Cette limite, liée à l'utilisation de transitions discrètes, est réduite par un choix judicieux dans l'ordre des événements. Ainsi, par exemple, nous simulons les ruptures d'union puis les mises en union ; cela veut dire que les personnes qui connaissent une rupture d'union durant une année peuvent se remettre en union la même année par contre les ruptures d'union ne peuvent se faire qu'à partir de l'année suivant la mise en union.

3. Les ménages complexes et des ménages collectifs

En plus des ménages simples composés d'une famille (couples ou adulte avec ou sans enfants), le modèle inclut également des ménages complexes et des ménages non ordinaires.

3.1 Ménages complexes

Un ménage est dit « complexe » s'il n'est pas constitué uniquement d'un noyau familial. Cela englobe donc l'ensemble des différents modes de cohabitation en dehors du noyau familial ou qui s'ajoutent à un noyau familial : ménages à plusieurs générations, ménages comprenant une famille et une autre personne, apparentée ou non (un parent âgé, un autre membre de la famille, un locataire), ménages constitués de personnes non apparentées de type colocation, etc.

Dans le modèle, nous considérons comme ménage complexe :

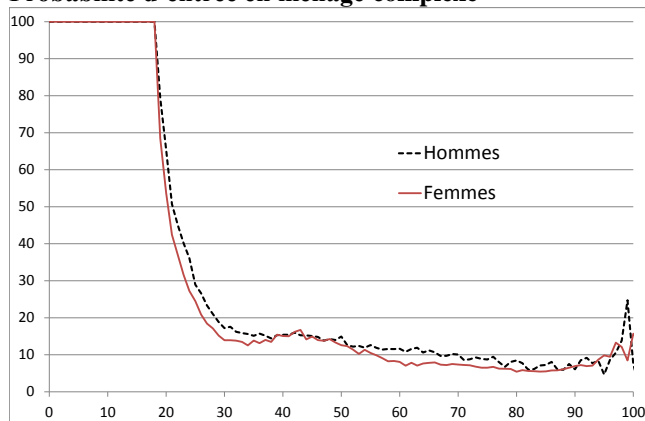
- un ménage avec des personnes isolées sans liens familiaux ;
- un ménage avec un noyau familial et d'autres personnes pouvant être ou non apparentées avec certains membres du noyau familial (parent, neveu ou étudiant louant une chambre de type « pensionnaire ») ;

Pour simplifier la simulation, nous avons considéré qu'il ne pouvait y avoir qu'un seul couple dans un ménage. On n'a donc pas de cas où le couple de parents vit avec l'un de ses enfants lui-même en couple. Ce n'est pas une hypothèse lourde car très peu de ménages comprennent deux couples ou même deux noyaux familiaux.

Pour simuler les ménages complexes, nous avons affecté à toute personne vivant seule une probabilité de rejoindre un ménage déjà existant (figure 3.1-1). Cette probabilité est issue d'une estimation par sexe et âge faite à partir des données du recensement de la population. La probabilité est augmentée pour les personnes ayant des enfants dans la région simulée. On obtient ainsi un ensemble de personnes devant rejoindre un autre ménage. L'appariement avec un ménage existant se fait de manière assez proche de l'appariement entre conjoints : 50 ménages potentiels sont tirés au hasard, un indicateur de similarité basé sur la différence d'âge entre la personne seule et la personne de référence du ménage (deux types de différence sont considérées comme préférable : moins de 5 ans et entre 25 et 35 ans pour simuler respectivement les colocations de jeunes et les ménages multigénérationnels) ainsi que la taille du ménage « receveur ». Le binôme personne seule et ménage receveur avec l'indice de similarité le plus grand est retenu.

La probabilité de sortie de ménage complexe est estimée à 0.5 mais il est possible de re-renter en ménage complexe dans la même année et donc de rester en ménage complexe. En cas de sortie de ménage complexe, la personne vit dans un nouveau ménage d'une personne.

Figure 3.1-1
Probabilité d'entrée en ménage complexe



3.2 Ménages collectifs

D'après la définition de l'Institut français de statistique (Insee), une communauté, ou ménage collectif, est définie comme un ensemble de locaux d'habitation, relevant d'une autorité gestionnaire et dont les habitants partagent à titre habituel un mode de vie en commun. Cela concerne donc les maisons de retraite, les institutions médico-légales, les pensionnats, les cités universitaires, les établissements hospitaliers, établissements pénitentiaires, communautés religieuses, établissements militaires.

En 2011, 2,5% des personnes recensées vivent dans les communautés soit 1.3 million de personnes. Elles vivent pour environ 30% dans une maison de retraite, 20% dans un établissement médico-social, 30% dans un internat et résidence universitaire, 8% dans un foyer de travailleurs, 4% dans un établissement militaire, 4% dans un établissement pénitentiaire et enfin 2% dans communauté religieuse ou autre type de communauté.

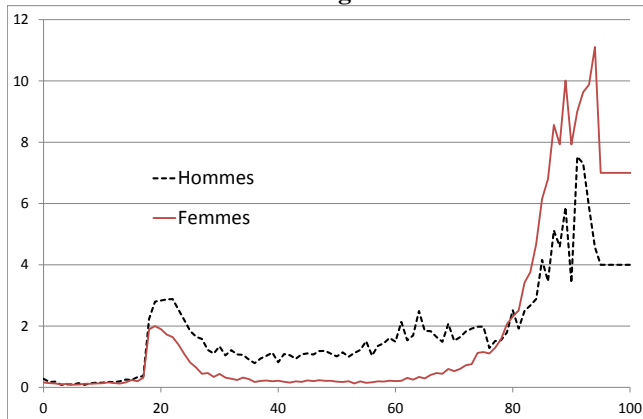
Le modèle considère comme un même type de ménages toutes les formes de ménages collectifs. L'âge discrimine fortement le type des ménages collectifs (maison de retraite pour les personnes les plus âgées, pensionnats pour les enfants, cité universitaire pour les jeunes adultes). Les personnes en ménage collectif ne sont pas rattachés à un ménage particulier et l'entrée en ménage collectif détruit le lien avec son ménage précédent. Par simplicité, nous avons considéré les personnes en ménages collectifs comme sans partenaire (si la personne était en union avant son entrée en ménage collectif, on considère que l'union est rompue).

Les probabilités d'entrer dans un ménage collectif sont estimées à partir des données du recensement. À partir d'un profil par sexe et âge, les personnes vivant en couple ont une probabilité plus faible (multipliée par 0,3) et pour ceux non en couple elle est double. La probabilité ne dépend ni du niveau d'éducation ni du fait d'être étudiant ou non.

Les personnes en ménages collectifs peuvent connaître le décès, le changement de niveau d'éducation, un changement de statut vis-à-vis du marché du travail et peuvent quitter le ménage collectif. Elles ne peuvent pas former une union, avoir des enfants, émigrer ou entrer directement dans un ménage complexe.

Comme pour les ménages complexes, la probabilité de sortie de ménage collectif est estimée à 0,5 avec la possibilité pour une même personne de sortir et re-renter dans la même année, ce qui revient aussi à rester en ménage collectif. Après 90 ans la probabilité de sortie diminue : rester en ménage collectif est particulièrement fréquent aux grands âges ou de plus le taux d'entrée est très fort. Si la personne sort du ménage collectif, elle vit seule dans un nouveau ménage d'une personne.

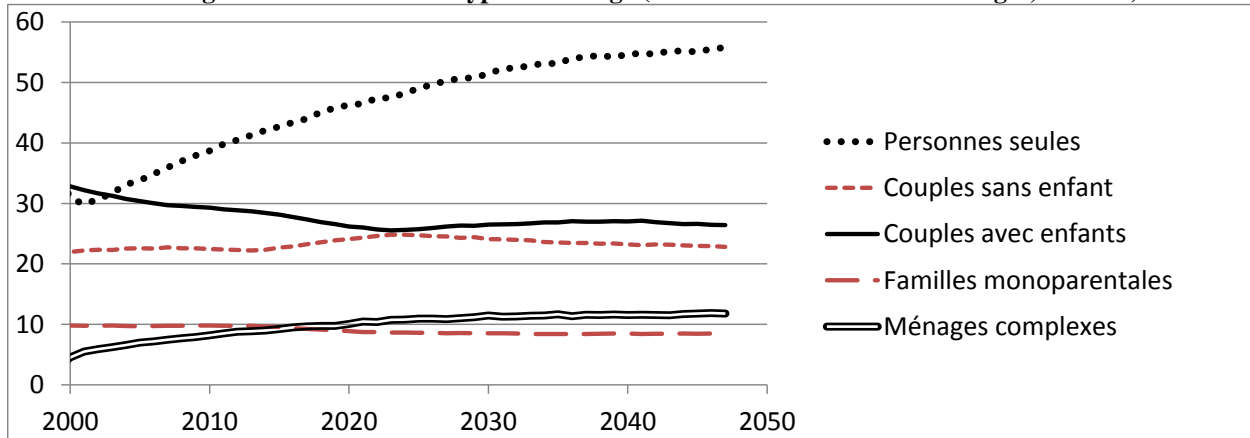
Figure 3.2-1
Probabilité d'entrer en ménages collectifs



4. Quelques résultats

Les probabilités de transition sont pour la plupart estimées d'après l'enquête Étude de l'histoire familiale, réalisée dans le cadre du recensement de la population de 1999, et qui contient une information rétrospective riche sur un échantillon de grande taille (Cassan et al., 2000 ; Lefèvre et Filhon, 2005). Le premier résultat du modèle porte sur la distribution des ménages par taille et type. Dans le cas de l'Ile-de-France, le nombre de ménages augmente plus vite que la population, en raison de la diminution de la taille moyenne des ménages. Les ménages d'une seule personne sont ceux dont l'augmentation est la plus marquée (figure 4-1) : leur nombre double presque entre 2000 et 2050. La seconde catégorie de ménages dont l'effectif augmente est celle des ménages complexes, tandis que les ménages hébergeant une famille restent pratiquement constants : couples avec ou sans enfant, ménage monoparental.

Figure 4-1
Nombre de ménages ordinaires selon le type de ménage (base 100 = ensemble des ménages, an 2000)

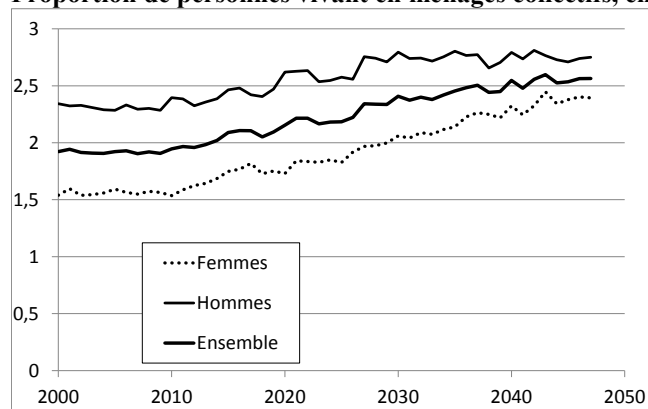


La hausse du nombre de ménages complexes s'explique surtout par la propension croissante des jeunes à vivre en co-location, après avoir quitté le domicile parental et avant de vivre en couple. Ainsi les ménages complexes sont plus souvent composés de personnes partageant un logement, sans qu'aucun ne déclare vivre en couple. La hausse est surtout marquée en début de période, en raison du retard des premières unions.

L'évolution du nombre de ménages collectifs suit une autre logique (figure 4-2). C'est d'abord le vieillissement de la population qui explique la hausse la proportion de personnes vivant en ménages collectifs. La part de la

population vivant en ménages collectifs augmente tout au long de la période de simulation, et la hausse est surtout marquée pour les femmes.

Figure 4-2
Proportion de personnes vivant en ménages collectifs, en %, selon l'année



5. Conclusion

Les résultats du module démographique ont permis au projet SustainCity d'appuyer les projections de ménages dans les métropoles européennes sur une microsimulation des comportements individuels. (Turci et al. 2010, 2011, 2012). Le recueil de données détaillées sur les personnes vivant en ménage collectif permettra de valider les évolutions observées. Le module démographique est très riche et permet de modéliser finement les comportements démographiques à l'origine de l'évolution du nombre et de la structure des ménages.

Une prochaine étape consistera à estimer des probabilités de sortie des ménages complexes et des ménages collectifs. Les estimations actuelles fondées sur des probabilités de sortie très élevées, utiles pour estimer les probabilités d'entrée à partir des données du recensement, conduisent à des estimations transversales cohérentes, mais à des parcours individuels composés d'allers-retours trop nombreux. Les données du recensement annuel (Lefranc, 2010), intégrant des informations sur la résidence un an avant le recensement depuis 2011, seront très utiles pour estimer plus précisément les mouvements des personnes entre les ménages.

Bibliographie

Cassan F., Héran F., Toulemon L., 2000, "Study of Family History. France's 1999 Family Survey", *Courrier des statistiques*, English series, n° 6, 2000 annual issue, p. 7-19.

Lefèvre C., Filhon A., (eds.), 2005, *Histoires de familles, histoires familiales. Les résultats de l'enquête Famille de 1999*, Les Cahiers de l'INED, N° 156. Paris: INED, 642 pages.

Christophe Lefranc C, 2010, « La diffusion des résultats du recensement rénové de la population », *Courrier des statistiques*, n° 129, p. 1-6.

Turci, L., A. Bringé, E. Morand, S. Pennec, L. Toulemon, R. Baggio and F. Billari (2010) Provisional demographic outline, SustainCity Working Paper, 4.1, INED, Paris. <http://www.sustaincity.org/publications>

Turci L., Pennec S., Toulemon L., Bringé A., Morand E. 2011. Demo4: SustainCity's demographic model. CEPS/INSTEAD, MiDal Final Conference, Esch-sur-Alzette, 28-30/11/2011.

Turci L., Pennec S., Toulemon L., Bringé A., Morand E. Baggio R. (2012) Demographic Model User Guide, SustainCity Working Paper, 4.4, INED, Paris. <http://www.sustaincity.org/publications>