

Le modèle de projections démographiques Demosim : mise à jour et nouveaux développements

Par Éric Caron-Malenfant^{1,2}

Résumé

Cet article donne un aperçu des principales nouveautés qui seront intégrées au modèle de projections par microsimulation Demosim, dont une version partant des microdonnées de l'Enquête nationale auprès des ménages de 2011 est en cours de préparation. Dans un premier temps, les ajouts à la population de base sont présentés; il s'agit essentiellement de nouvelles variables, certaines étant ajoutées aux données de l'Enquête nationale auprès des ménages par le biais d'appariement de fichiers. Les méthodes envisagées pour la projection des variables linguistiques, la citoyenneté et la religion sont ensuite brièvement présentées à titre d'exemples de nouveautés relatives aux événements simulés par le modèle.

Mots-clés : projections de population, microsimulation, démographie

Introduction

Demosim est un modèle de projections démographiques par microsimulation dont le développement et le maintien est assuré par Statistique Canada. Il a pour objectif principal d'effectuer des projections selon diverses caractéristiques de la population à des échelons géographiques fins, qui incluent notamment les régions métropolitaines de recensement (RMR). À cette fin, il prend pour point de départ les microdonnées des recensements de la population du Canada et effectue la projection des individus qui s'y trouvent un à la fois en leur faisant "vivre" au fil du temps divers événements, jusqu'à ce qu'ils ne décèdent, n'émigrent ou n'atteignent la fin de la simulation. À cette population présente au départ sont ajoutés en cours de simulation des immigrants et des naissances. Sur la base d'hypothèses relatives aux probabilités de "vivre" les événements simulés, le modèle permet la production de tableaux de population croisant les variables considérées pour des dates de référence futures³.

Développé essentiellement à frais recouvrables, Demosim a permis l'élaboration de divers produits de nature prospective. Ceux-ci comprennent des projections démographiques officielles (notamment les *Projections de la diversité de la population canadienne, 2006-2031* (Caron-Malenfant et al., 2010) et les *Projections de la population selon l'identité autochtone, 2006-2031* (Caron-Malenfant et Morency, 2011⁴),

¹ Statistique Canada, 1708 immeuble principal, 150 Tunney's Pasture Driveway, Ottawa, Canada, K1A 0T6, eric.caron-malenfant@statcan.gc.ca.

² L'auteur tient à remercier le reste de l'équipe de Demosim dont il s'est fait porte-parole lors du Symposium 2014 de méthodologie. Les personnes qui ont travaillé à la mise à jour de Demosim sont (outre l'auteur) : Melanie Abeysondera, Jonathan Chagnon, Simon Coulombe, Karla Fox, Harry François, Nora Galbraith, Dominic Grenier, Chantal Grondin, René Houle, Mark Knarr, Stéphanie Langlois, Laurent Martel, Jean-Dominique Morency, Soumaya Moussa et Martin Spielauer. Des remerciements également à Marc Lachance pour avoir révisé une version préliminaire de ce texte.

³ Pour une présentation générale des méthodes de microsimulation appliquées aux projections démographiques, voir Van Imhoff et Post (1997) et Willekens (2011).

⁴ Le lecteur est invité à consulter ces deux publications pour en savoir davantage sur la méthodologie de la version précédente de Demosim (basée sur 2006), sur laquelle construit la nouvelle version du modèle dont

des articles méthodologiques et d'analyse⁵, des tableaux personnalisés et des projections à court terme utilisées par des enquêtes de Statistique Canada. Ces produits ont été élaborés à ce jour au moyen de versions de Demosim qui se basent sur les données des recensements complets de 2001 et de 2006.

Depuis, des fonds ont été octroyés pour que soit effectuée une mise à jour de Demosim tirant parti des données les plus récentes, notamment celles de *l'Enquête nationale auprès des ménages de 2011* (ENM)⁶. L'objectif du présent texte est de donner un aperçu de certaines des nouveautés qui résulteront de cette mise à jour. Demosim étant l'objet d'une refonte complète, seules certaines des nouveautés qui sont en cours d'intégration au modèle seront présentées. Celles touchant la population de base de Demosim seront d'abord présentées. Suivront des exemples de changements aux méthodes et sources de données servant à certains aspects de la modélisation proprement dite.

2. Nouveautés touchant la population de base

Les premières nouveautés se rapportant à Demosim ont trait à sa population de base. La population de base qui servira de point de départ à la nouvelle version du modèle sera constituée des microdonnées de l'ENM de 2011, lesquelles seront l'objet de certains ajustements, notamment pour tenir compte du sous-dénombrement net au recensement et des réserves partiellement dénombrées.

La nouvelle population de base comprendra toutes les variables que comprenaient les versions précédentes, les principales étant les suivantes : âge, sexe, statut d'immigrant et période d'immigration, statut des générations, lieu de naissance détaillé, groupe de minorités visibles, identité autochtone et statut d'Indien inscrit, religion, langue maternelle, état matrimonial, participation à la population active et lieu de résidence.

De nouvelles variables y seront cependant ajoutées, afin d'accroître le potentiel analytique du modèle. Les deux premières visent à permettre la production de projections linguistiques détaillées. Il s'agit de la langue parlée le plus souvent à la maison (anglais, français, anglais et français, autres⁷) et la connaissance des langues officielles (français seulement, anglais seulement, anglais et français ou ni l'anglais ni le français). La population de base inclura également le statut de chef de ménage et le statut de personne repère dans la famille de recensement, l'objectif étant de permettre des projections dérivées des ménages et familles, selon le statut autochtone. Deux autres nouvelles variables ont trait plus spécialement aux populations immigrantes. La citoyenneté, d'abord, qui permettra de distinguer les personnes ayant ou non la citoyenneté canadienne, en conjonction ou non avec un ou d'autres pays de citoyenneté. La catégorie d'admission des immigrants constitue la seconde de ces variables. Elle permettra de distinguer, parmi les immigrants admis depuis 1980, ceux qui ont été admis en vertu de la composante économique du programme d'immigration, en vertu du programme de réunification familiale, à titre de personnes protégées (réfugiées) et les autres immigrants. La population de base comprendra enfin une nouvelle variable relative aux Indiens inscrits, soit leur type d'inscription au Registre des Indiens (inscription en vertu de l'article 6(1) ou 6(2) de la *Loi sur les Indiens* du Canada de 1985). Cette variable définit les règles de transmission du statut des parents aux enfants, et représente par conséquent un élément clé de la projection de la population ayant le statut d'Indien inscrit⁸.

certaines éléments sont esquissés ici. La première version du modèle (basée sur 2001) est pour sa part documentée dans Bélanger et Caron-Malenfant (2005) et Bélanger et al. (2008).

⁵ Voir par exemple Caron-Malenfant et al. (2011) puis Spielauer (2014).

⁶ La mise à jour bénéficie de l'apport financier d'Affaires autochtones et Développement du Nord Canada puis de Citoyenneté et Immigration Canada.

⁷ Les langues multiples incluant une langue non officielle et une ou les deux langues officielles sont classées parmi les catégories de langues officielles.

⁸ Les Indiens inscrits de type 6(1) peuvent transmettre sans restriction à leurs enfants le statut d'Indien inscrit alors que ceux de type 6(2) ne le peuvent que lorsque l'autre parent a lui aussi le statut d'Indien inscrit. Cette variable avait été imputée dans la version précédente de Demosim pour une petite partie de la population des Indiens inscrits. La nouveauté tient ici au fait que l'information est désormais disponible

On aura remarqué que les deux dernières variables ne font pas partie de la base de données de l'ENM de 2011. Celles-ci y ont été ajoutées par le biais d'appariements de données. La catégorie d'admission des immigrants a été ajoutée en tirant parti d'un appariement pré-existant entre l'ENM de 2011 et les fichiers administratifs de Citoyenneté et Immigration Canada de 1980 à 2011. L'appariement a fonctionné pour 82% des individus ayant déclaré à l'ENM de 2011 avoir été admis à titre d'immigrants depuis 1980. Pour les 18% d'enregistrements pour lesquels l'appariement a échoué, l'information a été imputée. Le type d'inscription au Registre des Indiens, de son côté, a été obtenu d'un appariement pré-existant entre l'ENM de 2011 et le Registre des Indiens, une base de données administratives d'Affaires autochtones et Développement du Nord Canada. L'appariement a fonctionné pour 66% des personnes qui ont déclaré à l'ENM de 2011 avoir le statut d'Indien inscrit. L'information a été imputée pour les 34% restants.

Ajoutons que certaines des variables projetées par le passé le seront, avec la nouvelle version de Demosim, avec davantage de détail. C'est notamment le cas de la géographie. Le modèle comprenait déjà les 33 RMR du pays⁹, les provinces et territoires, puis des géographies propres aux projections autochtones¹⁰. Aux fins d'éventuelles projections de la situation linguistique au pays, de nouvelles régions seront ajoutées. D'abord, la RMR de Montréal sera divisée en deux régions, qui se distinguent notamment par leur profil linguistique : l'île de Montréal et le reste de la RMR. Ensuite, et pour la même raison, une région à concentration francophone sera créée dans la partie hors RMR de l'Ontario puis dans la partie hors RMR du Nouveau-Brunswick, en y regroupant les divisions de recensements comprenant en 2011 au moins 20% de personnes ayant le français comme première langue officielle parlée. Les cartes présentées en annexe montrent, à titre d'exemples, l'infrastructure géographique qui en résultera en Ontario et au Nouveau-Brunswick.

3. Nouveautés relatives à certains des événements projetés

Afin de mettre à jour à l'échelon des individus les variables projetées, Demosim simule de manière dynamique, au moyen de méthodes et sources de données diverses, des événements qui forment ensemble les parcours longitudinaux des individus projetés. Outre la simulation dynamique des événements, le modèle permet également, toujours aux fins de la mise à jour des attributs individuels, l'imputation à date fixe de caractéristiques¹¹. La version de Demosim qui est actuellement en cours d'élaboration continuera de modéliser chacun des événements que modélisaient ses versions précédentes, principalement la fécondité, la mortalité, l'immigration, l'émigration, la migration interne, changement de niveau de scolarité, mobilité religieuse¹², mobilité ethnique des Autochtones¹³, l'inscription au Registre des Indiens et le départ des enfants du foyer parental. Elle continuera également d'imputer annuellement aux individus un état matrimonial et un statut de participation au marché du travail.

L'ajout au modèle de caractéristiques susceptibles de changer au cours de la vie appelle cependant pour leur simulation adéquate le développement de nouvelles composantes. Pour deux d'entre elles, soit celles qui ont trait à la projection des deux nouvelles variables linguistiques puis à la citoyenneté, les méthodes

pour l'ensemble de la population ayant le statut d'Indien inscrit au sein de la population de base de Demosim et qu'elle est, pour l'essentiel, obtenue d'un appariement.

⁹ Il comprenait en réalité 34 régions métropolitaines puisque la RMR d'Ottawa-Gatineau était divisée entre sa partie ontarienne et sa partie québécoise.

¹⁰ Ces régions comprennent l'ensemble des réserves indiennes puis les régions de l'Inuit Nunangat, où résident la majorité des Inuits.

¹¹ La gestion des événements et des autres composantes de Demosim est assurée par Modgen, un langage avec lequel Demosim est programmé. Pour en connaître davantage sur Modgen, voir le site web de Statistique Canada à l'adresse suivante : www.statcan.gc.ca/microsimulation.

¹² On entend par mobilité religieuse les changements de religion qui peuvent survenir d'une génération à l'autre ou au cours de la vie d'un même individu.

¹³ On entend par mobilité ethnique des Autochtones les changements dans l'identité autochtone déclarée qui peuvent survenir d'une génération à la suivante ou au cours de la vie d'un même individu.

que l'on prévoit mettre à l'essai sont décrites brièvement ci-après¹⁴. Une troisième se rapporte à la projection des ménages mais n'est pas présentée ici, le texte de Jean-Dominique Morency dans ce même recueil en explicitant les méthodes.

La mobilité linguistique comprend deux dimensions distinctes mais complémentaires : la mobilité linguistique intragénérationnelle (au cours de la vie) et la mobilité linguistique intergénérationnelle (des parents aux enfants). Les probabilités associées aux changements linguistiques intragénérationnels¹⁵ seront établies au moyen des appariements entre les recensements de 2001 et de 2006 (pour les transitions relatives à la langue parlée à la maison et la langue maternelle) puis entre les recensements de 2006 et de 2011 (pour la connaissance des langues officielles)¹⁶. Des modèles de régressions logistiques spécifiques aux langues d'origine et de destination viseront à estimer à partir de ces sources les probabilités d'avoir vécu un transfert sur une période de 5 ans selon diverses caractéristiques projetées par ailleurs par Demosim. La mobilité linguistique intergénérationnelle, de son côté, consistera à attribuer de manière probabiliste aux enfants qui naissent en cours de simulation un profil linguistique qui sera fonction de celui de leur mère¹⁷. Les probabilités seront estimées sur la base des données de l'ENM de 2011, au sein desquelles on aura lié les plus jeunes enfants aux femmes les plus susceptibles d'en être la mère. Ce lien établi, il deviendra possible de calculer des probabilités de mobilité linguistique entre mères et enfants, ce qui sera fait par le biais de modèles de régressions logistiques qui prendront en considération celles des variables qui sont pertinentes pour un tel calcul parmi les variables que projette Demosim.

Le module de citoyenneté visera essentiellement à permettre aux immigrants qui sont ajoutés à la population au cours de la projection d'acquérir après leur arrivée la citoyenneté canadienne¹⁸. Deux méthodes sont présentement envisagées. Une première pourrait tirer parti des appariements dont il sera fait usage pour le développement du module de mobilité linguistique intragénérationnelle. Il s'agirait alors d'estimer les chances que les immigrants n'ayant pas la citoyenneté en début de période l'aient obtenue cinq ans plus tard, selon certaines caractéristiques-clés. Une seconde pourrait ne se baser que sur les données de l'ENM de 2011. Il s'agirait de dériver des taux d'acquisition de la citoyenneté canadienne d'une estimation de la proportion d'immigrants l'ayant obtenue après 3 ans, 4 ans, 5 ans, 6 ans, etc. selon le lieu de naissance, l'âge et le sexe¹⁹. Il n'est pas impossible que ces deux méthodes soient mises à l'essai puis comparées entre elles.

Si la mise à jour de Demosim donne lieu à l'ajout de composantes, elle est également l'occasion d'une refonte des méthodes touchant certains des événements dont il permettait déjà la modélisation. C'est le cas notamment de la mobilité religieuse, qui était simulée sur la base des données de *l'Enquête sur la diversité ethnique* de 2002, une enquête qui comprenait une question sur la religion du répondant de même qu'une autre sur la religion de sa mère alors que le répondant avait moins de 15 ans. Or, puisque l'enquête n'a plus été tenue après 2002, une méthode alternative sera implantée dans la nouvelle version de Demosim, méthode qui comprendra, comme la mobilité linguistique, une dimension intragénérationnelle et une dimension intergénérationnelle. La mobilité religieuse intragénérationnelle sera modélisée en appliquant

¹⁴ Les méthodes présentées sont celles qui étaient envisagées ou en cours de développement en date du Symposium 2014.

¹⁵ Les changements possibles sont l'apprentissage du français et / ou de l'anglais, l'oubli de la connaissance du français et / ou de l'anglais, le changement de langue parlée le plus souvent à la maison et, bien que cela soit plus rare, le changement de langue maternelle. Le changement de langue maternelle est possible en raison du fait que les recensements du Canada précisent que la première langue apprise au cours de l'enfance doit être encore comprise par le répondant. Voir à ce sujet Lepage (2011).

¹⁶ Le Recensement de 2011 présente des défis de comparabilité avec les recensements antérieurs, notamment celui de 2006, pour ce qui est de la langue maternelle et de la langue parlée le plus souvent à la maison (Houle et Corbeil, 2013). Pour cette raison, l'estimation des paramètres de transition pour ces deux variables est basée sur l'appariement entre 2001 et 2006 plutôt que sur l'appariement entre 2006 et 2011. Les données sur la connaissance des langues officielles du Recensement de 2011 sont comparables avec celles du Recensement de 2006.

¹⁷ Dans Demosim, seul le lien entre les enfants et leur mère (et non leur père) est créé, d'où cette stratégie.

¹⁸ La citoyenneté au moment de l'admission sera attribuée aux immigrants par imputation.

¹⁹ Le calcul de tels taux nécessiterait sans doute certains ajustements pour tenir compte de l'émigration et de la mortalité des immigrants après leur arrivée.

aux données du Recensement de 2001 et de l'ENM de 2011²⁰ ajustées pour le sous-dénombrement net une méthode résiduelle. Celle-ci consistera à estimer la part de la croissance ou de la décroissance de la population des diverses religions qui demeure inexplicée après que l'on ait contrôlé pour les naissances, les décès, l'immigration et l'émigration, cette part inexplicée étant interprétée comme le résultat de la mobilité religieuse. Pour les religions ayant perdu des effectifs par voie de mobilité religieuse, les pertes nettes ainsi estimées seront divisées par la population de départ afin de constituer des taux nets de "sortie" sur 10 ans. Les "sortants" seront distribués parmi les religions "gagnantes" proportionnellement aux gains nets de celles-ci. Des estimations basées sur des recensements précédents pourraient servir à la constitution d'hypothèses alternatives. Pour ce qui est de la mobilité religieuse intergénérationnelle, elle sera modélisée au moyen d'une méthode similaire à celle utilisée dans le cas de la mobilité linguistique intergénérationnelle, en comparant la religion des mères à celle de leur enfant à l'ENM de 2011.

4. Conclusion

Ce texte visait à donner un aperçu des changements qui sont apportés à Demosim dans le cadre de sa mise à jour. S'il a notamment permis d'illustrer le potentiel nouveau qu'offre pour la microsimulation la disponibilité de sources de données issues d'appariements, il n'a pu présenter, et trop brièvement, que certains des changements au modèle. Cette mise à jour est en effet l'occasion d'une ré-estimation de l'ensemble des paramètres de Demosim, puis d'autres développements présentant parfois une certaine complexité, spécialement en ce qui touche les multiples dimensions de la projection du statut d'Indien inscrit. Pour cela, ce texte ne doit être considéré que comme une introduction, préliminaire de surcroît, à certains des aspects de la méthodologie de la version en cours de développement du modèle. Celle-ci sera exposée plus avant dans des publications à venir.

Bibliographie

Bélanger, A. et É. Caron-Malenfant (2005), *Projections de la population des groupes de minorités visibles, Canada, provinces et régions, 2001-2017*, produit no. 91-541 au catalogue de Statistique Canada.

Bélanger, A., É. Caron-Malenfant, L. Martel et R. Gélinas (2008). « Projecting ethno-cultural diversity of the Canadian population using a microsimulation approach », dans le compte-rendu de l'atelier de l'UNECE / Eurostat sur les projections démographiques à Bucarest.

Caron-Malenfant, É., A. Lebel et L. Martel (2010), *Projections de la diversité de la population canadienne, 2006-2031*, produit no. 91-551 au catalogue de Statistique Canada.

Caron-Malenfant, É., P. Dion, A. Lebel et D. Grenier (2011), « Immigration et structure par âge de la population du Canada : quelles relations? », dans *Cahiers québécois de démographie*, vol. 40, no. 2.

Caron-Malenfant, É. et J.-D. Morency (2011), *Projections de la population selon l'identité autochtone au Canada, 2006-2031*, produit no. 91-552 au catalogue de Statistique Canada.

Houle, R. et J.-P. Corbeil (2013), *Document méthodologique sur les données linguistiques du Recensement de 2011*, produit no. 98-314-X2011051 au catalogue de Statistique Canada.

Lepage, J.F. (2011), « L'oubli des langues maternelles : les données du recensement sous-estiment-elles les transferts linguistiques? », dans *Cahiers québécois de démographie*, vol. 40, no. 2.

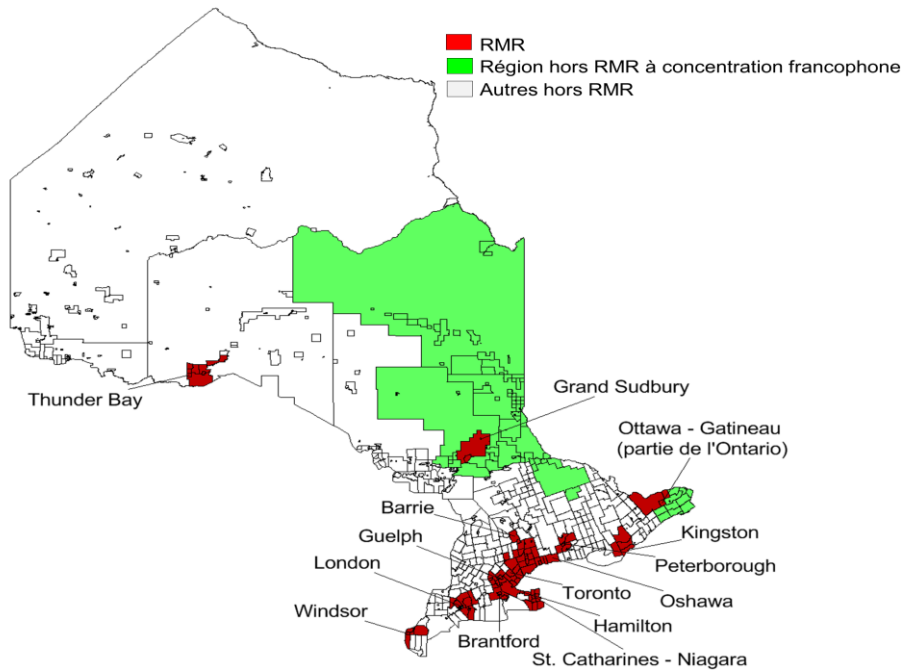
Spielauer, M. (2014), « The relation between education and labour force participation of Aboriginal peoples: A simulation analysis using the Demosim population projection model », dans *Canadian Studies in Population*, 41, no. 1-2.

Van Imhoff, E. et W. Post (1997), « Méthodes de micro-simulation pour des projections de population », dans *Population*, 52^{ième} année, no.4.

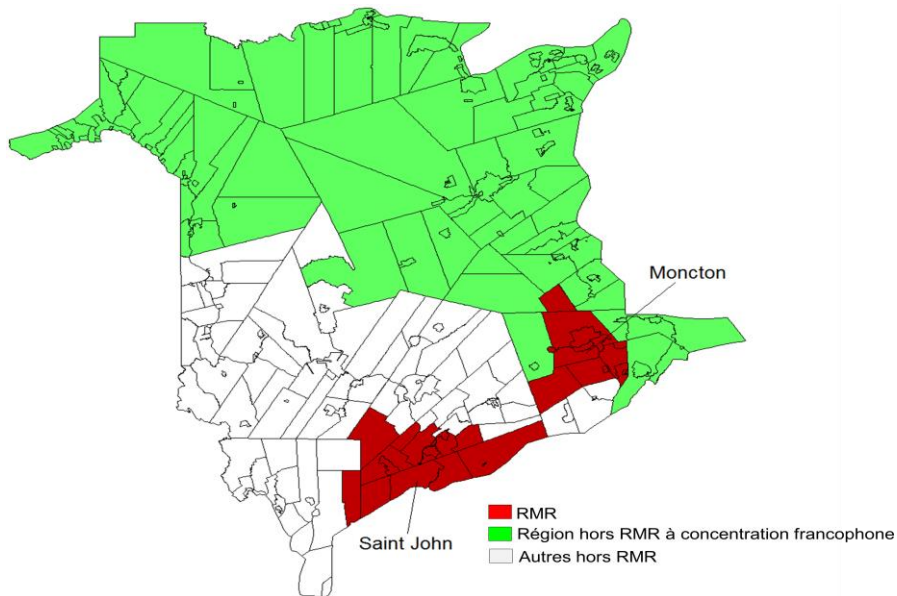
²⁰ Rappelons que la question sur la religion n'est posée qu'aux dix ans dans les recensements canadiens.

Willekens, F. (2011), « La microsimulation dans les projections de population », dans *Cahiers québécois de démographie*, vol. 40, no. 2.

Carte 1. Infrastructure géographique de Demosim : Ontario.



Carte 2. Infrastructure géographique de Demosim : Nouveau-Brunswick.



Note : Les RMR (en rouge) sont toutes projetées distinctement. La partie à concentration francophone de la part hors RMR de chacune des deux provinces (en vert) ne forme, dans ces provinces, qu'une seule région aux fins des projections, de même que son complément non francophone (en blanc). Il est à noter que certaines des régions de l'Ontario et du Nouveau-Brunswick comprennent également une partie sur-réserve, non présentée ici.