

Des projections nationales aux projections locales

Documents de travail

N° H2021-01 - Février 2021



L'Insee réalise régulièrement des projections de population, le dernier exercice régional et local pour la période 2013-2050 a été mis à disposition via le modèle Omphale en juin 2017. La mise à jour du modèle Omphale a fait suite à la publication en décembre 2016 des projections de population au niveau France entière. Les projections nationales sont réalisées de manière apériodique par la Direction des Statistiques Démographiques et Sociales de l'Insee. Le dernier exercice a pour point de départ la population par sexe et âge au 1er janvier 2013, issue du bilan démographique 2013 et pour horizon 2070. Ces projections simulent jusqu'à l'horizon de projection le nombre d'hommes et de femmes de chaque âge sur la base d'hypothèses sur l'évolution des trois composantes des variations de population (méthode des composantes) : fécondité, mortalité et migrations.

Du côté des projections infra-régionales, le modèle développé par l'Insee, nommé *Omphale*, permet de réaliser des projections de population à l'échelle locale en s'appuyant, pour des raisons de cohérence, sur les hypothèses et les résultats des projections nationales. Tout regroupement de communes d'un seul tenant et dont la population est au moins égale à 50 000 habitants peut faire l'objet de projections. Ces dernières sont diffusées par différents canaux : insee.fr, prestations sur mesures, études en partenariat entre les Directions Régionales de l'Insee et des collectivités locales ou des services déconcentrés de l'État. Omphale est donc un outil central en matière de prospective démographique pour de nombreux acteurs publics en charge de l'aménagement du territoire. On peut par exemple citer les Schémas de cohérence territoriaux (Scot) qui doivent nécessairement intégrer un volet prospectif. Les projections peuvent également servir de base à l'évaluation de besoins sur des populations spécifiques (élèves, personnes âgées dépendantes ...) ou venir appuyer des exercices d'évaluation des besoins à venir en logements. L'exploitation des projections réalisées à partir du modèle pour le calibrage des politiques publiques nécessite de prendre en considération ses limites intrinsèques afin de s'assurer de leur pertinence :

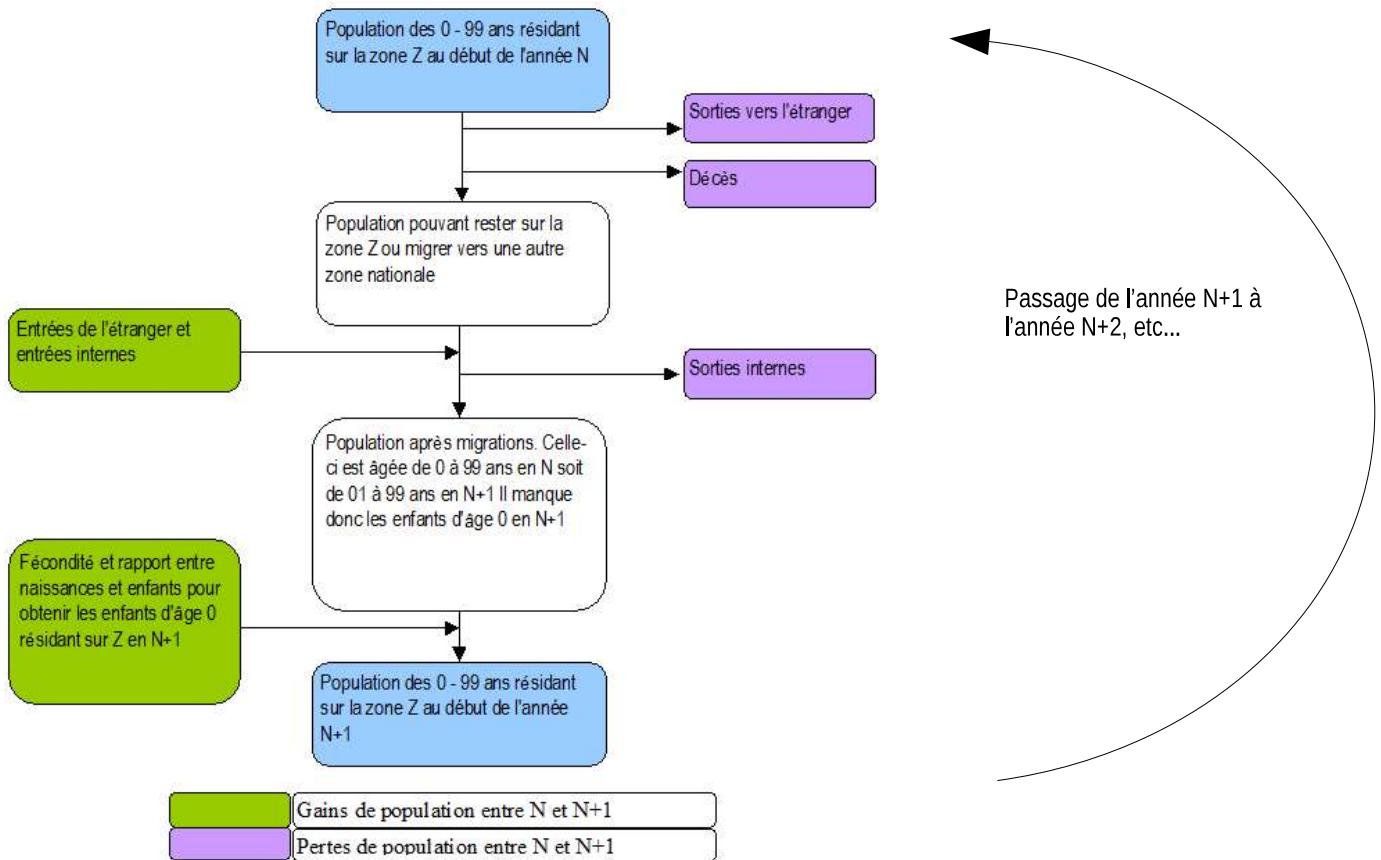
- la compréhension par l'utilisateur de la nature d'une projection et notamment de la différence avec une prévision. Il existe nécessairement des écarts entre la projection et la réalité observée *ex-post*. Loin d'être problématiques, ces écarts sont des éléments de compréhension du territoire et de son évolution. Un retour en arrière sur les projections de l'ancien modèle Omphale 2010 comparées aux dernières données du recensement et à Omphale 2017 témoigne de l'intérêt de ces analyses ;
- la notion de robustesse des projections qui dépend de la fiabilité des données en entrées et de l'adaptation des hypothèses à la réalité du territoire. Dans ce document, la robustesse d'une projection est mesurée par la capacité d'Omphale à prolonger les tendances passées.

Une méthode itérative pour faire évoluer une pyramide des âges

La méthodologie retenue pour répondre aux besoins des acteurs publics en matière de prospective démographique est appelée la méthode des composantes : la population d'une zone varie d'une année sur l'autre en fonction de la fécondité, de la mortalité et des migrations sur le territoire. Pour chaque année de projection, la population en début d'année par sexe et âge augmente sous l'effet des naissances et des entrées sur le territoire et diminue avec les décès et les sorties du territoire, ce qui permet d'obtenir la population par sexe et âge au début de l'année suivante (figure I). Les composantes (fécondité, mortalité, migrations) font l'objet d'hypothèses d'évolution des comportements jusqu'à l'horizon de projection. La modélisation est réalisée à partir des quotients de fécondité par âge, les quotients de décès par sexe et âge et les quotients migratoires par sexe, âge, lieu d'origine et de destination. Les quotients utilisés au point de départ de la projection sont propres au territoire projeté. Seuls les flux avec l'étranger sont mobilisés en volume dans le modèle et ne font pas appel à un quotient, ils ne sont donc pas directement liés à l'évolution de la population (tableau 1). Les évolutions des quotients démographiques proposées dans Omphale sont en cohérence avec les hypothèses nationales sur les composantes. Les quotients migratoires internes sont calculés pour l'année de départ, ils sont ensuite maintenus constants en projection dans les scénarios standard.

La combinaison de ces différentes hypothèses détermine les scénarios démographiques. Le scénario central s'appuie sur la prolongation des tendances passées au niveau national. Des variantes hautes et basses sont disponibles en standard pour chacune des composantes. Aussi, un scénario de « population haute » par exemple combine les hypothèses de chacune des composantes qui auront tendance à faire augmenter la population. Des scénarios standards issus des hypothèses nationales sont proposés largement aux utilisateurs d'Omphale. Dans le cadre d'études et de partenariats, le modèle peut également être paramétré pour construire des scénarios à façon pour des besoins de simulation des hypothèses (encadré).

Figure I : La mécanique de projection dans Omphale entre l'année N et l'année N+1



Source : Insee, Omphale 2017.

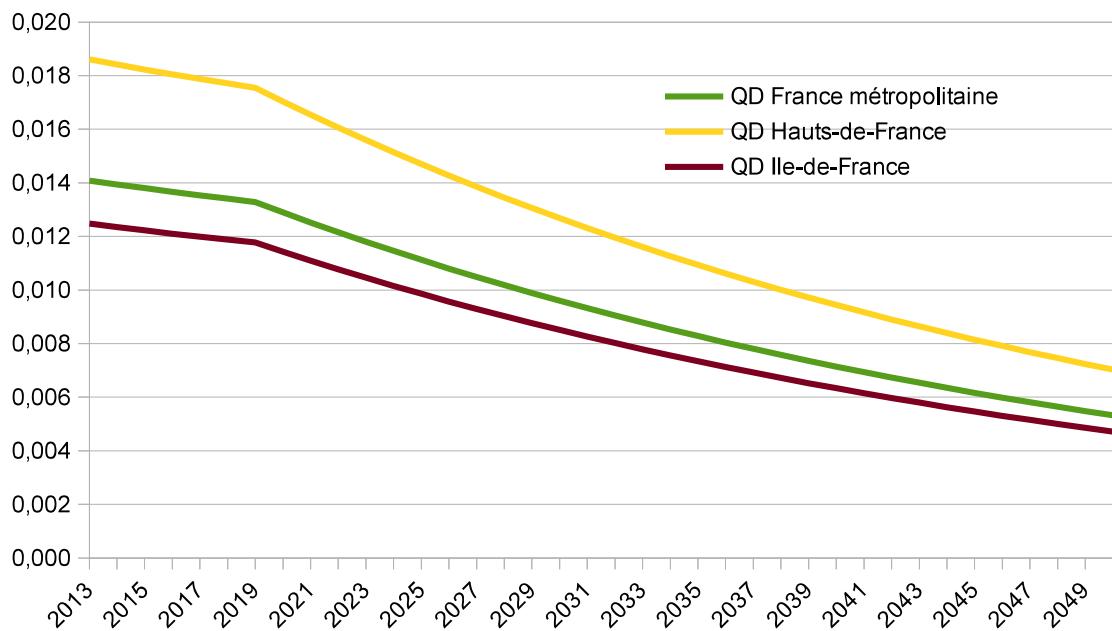
Tableau 1 : définition des quotients démographiques et des volumes retenus dans la méthode des composantes

Composante	Quotients ou méthode retenue	Commentaire
Fécondité	$Qf_n = \frac{\text{Nombre de naissances vivantes par des femmes d'âge } n}{\text{Nombre de femmes d'âge } n}$	Le quotient de fécondité rapporte les naissances selon l'âge et la zone de résidence de la mère à la population des femmes du même âge de cette zone. Au point de départ des projections, ce quotient est spécifique à chaque zone projetée.
Mortalité	$Qd_n = \frac{\text{Nombre de décès de personnes de sexe } s \text{ et d'âge } n}{\text{Nombre de personnes de sexe } s \text{ et d'âge } n}$	Le quotient de décès rapporte le nombre de décès pour un sexe et un âge donné sur une zone à la population correspondante sur la zone. Au point de départ des projections, ce quotient est spécifique à chaque zone projetée.
Migrations internes	$Qe_{\text{origine} \rightarrow \text{destination}} = \frac{\text{Flux zone origine} \rightarrow \text{zone destination}}{\text{Population de la zone origine avant les flux}}$	Le quotient d'émigration interne rapporte un flux migratoire (de la zone d'origine vers la zone de destination) à la population de la zone d'origine de ces migrants. Les flux sont modélisés par sexe et âge. Ces quotients permettent de modéliser uniquement les migrations entre zones de France métropolitaine (migrations internes).
Migrations avec l'étranger	Les flux avec l'étranger sont modélisés en volume (en nombre de personne) par sexe et par âge. Les flux d'entrée et de sortie de la métropole sont déterminés par les hypothèses nationales de projection. La répartition de ces flux sur le territoire est déterminée au début de la projection par une clé de répartition qui dépend du poids de l'immigration de la zone dans les immigrations nationales.	

Dans la méthode des composantes, ces scénarios et hypothèses ne sont pas probabilisés. Les différents scénarios doivent permettre à l'utilisateur d'appréhender le champ des possibles en termes d'évolution de la population et de simuler l'impact de la déviation d'une composante l'évolution sur la population. Les projections constituent une simulation de l'avenir dans le cadre théorique des hypothèses formulées.

Toutes les projections issues d'un scénario standardisé dans Omphale tiennent compte des spécificités démographiques du territoire l'année de départ de la projection ; en revanche, les évolutions des quotients sont semblables quel que soit le territoire (Figure II). Ces projections doivent servir de cadre au débat public et ne constituent pas des prévisions de population.

Figure II : Evolution des quotients de décès des hommes de 65 ans en Hauts-de-France, Île-de-France et en France métropolitaine



Source : Insee, Omphale 2017.

La mécanique de projection traite successivement chacune des composantes de projection pour passer de la population au début de l'année N à la population au début de l'année N+1. La boucle est répétée jusqu'à l'horizon de la projection.

Passer des projections nationales aux projections locales

L'introduction des flux internes dans le processus de projection

La différence essentielle entre les projections nationales et le modèle Omphale vient de l'ajout d'une dimension dans les composantes d'évolution de la population : les flux migratoires internes. En effet, au niveau national, seuls les flux entre la France et l'étranger sont modélisés. Sur un découpage territorial plus fin, les flux bilocalisés de population influencent fortement la dynamique démographique des territoires. Les données mobilisées dans Omphale pour mesurer les flux internes de population sont issues du recensement de la population. La bilocalisation des flux induit que les entrées d'une zone A vers B constituent les sorties de B vers A et qu'il suffit alors de projeter les flux dans un sens. Omphale modélise les sorties d'un territoire plutôt que les entrées, le nombre de sorties d'un territoire étant proportionnel à la zone de départ et non pas à celle d'arrivée.

En particulier un territoire connaissant des heurts dans sa pyramide des âges fournira toujours des effectifs proportionnels de sortants en rapport avec la population projetée de la zone de départ.

Localiser les flux avec l'étranger

Le niveau du solde migratoire avec l'étranger dans le scénario central de projection a été estimé à +70 000 personnes par an dans les projections nationales en s'appuyant sur un solde migratoire apparent, issu de la comparaison des populations entre deux dates et du solde naturel (naissances – décès). Par nature, ce type de mesure intègre tous les effets non mesurés par ailleurs, et en partie les effets d'échantillonnage ou les incohérences entre le recensement et l'état civil. Les projections nationales proposent une hypothèse de solde migratoire avec l'étranger par sexe et âge jusqu'à l'horizon de projection. Cette hypothèse nationale n'est pas décomposée entre entrées et sorties du territoire.

Dans la mécanique de projection, Omphale fait intervenir pour chaque zone d'intérêt des immigrants en provenance de l'étranger et des émigrants vers l'étranger par sexe et âge détaillés. Le recensement de la population permet, via sa question sur les migrations, d'estimer l'immigration en provenance de l'étranger par zone ; il n'existe par contre pas d'information locale sur les émigrants vers l'étranger. Les deux exercices de projection devant être cohérents, les échanges locaux avec l'étranger constituent une ventilation de l'hypothèse de solde migratoire national. Aussi, pour produire une hypothèse utilisable dans Omphale, le solde migratoire national est dans un premier temps décomposé en nombre d'émigrants vers l'étranger par sexe et âge et un nombre d'immigrants. Les émigrants vers l'étranger sont déduits du solde migratoire national et des immigrants calculés avec le recensement. Cette estimation au niveau national des entrées et des sorties est différente de celles des travaux proposés et publiés par l'Insee, qui font appel à des traitements des données beaucoup plus fins (Brutel, 2014). Ce calcul est réalisé uniquement pour répondre aux besoins de localiser les composantes d'émigration et d'immigration dans le processus de projection d'Omphale. Dans un second temps, une clé de répartition est utilisée pour ventiler les entrées et sorties nationales au niveau des zones d'intérêt locales. Cette clé consiste à calquer la répartition de ces émigrants sur l'intensité d'immigration de chaque zone (tous sexes et âges confondus) : une zone qui accueille X % des immigrants métropolitains va concentrer X % des émigrants métropolitains de chaque âge.

Toutefois, l'hypothèse de cohérence géographique entre arrivées et départs avec l'étranger n'est pas toujours vérifiée. Il existe deux types d'émigrés :

- des « natifs » quittant le pays dont il n'y a pas de raison particulière qu'ils soient originaires des zones d'arrivées massives d'immigrés ;
- des immigrés quittant quelques années après leur arrivée le pays : pour ces derniers, l'hypothèse conduit à ne pas prendre en considération leurs éventuelles mobilités au sein du pays entre leur arrivée et leur départ.

Les conséquences de ces limites sont prégnantes pour des territoires où les flux avec les étrangers sont déséquilibrés. A titre d'exemple, certains départements d'Île-de-France accueillent de nombreux arrivants de l'étranger sans pour autant que les départs soient élevés. La conséquence de l'hypothèse réalisée est une émigration trop élevée en projection et donc une sous-estimation de la population conduisant à une rupture à la baisse des tendances démographiques récentes.

Les limites de cette hypothèse étaient bien connues au moment de l'implémentation d'Omphale cependant aucune autre méthode alternative n'a donné satisfaction. À titre d'exemple, une méthode basée sur l'estimation des sorties via une ventilation du solde national utilisant le solde migratoire apparent du territoire a été testée en retenant les évolutions suivantes :

$$\text{Population } T+1 = \text{Population } T + \text{Solde naturel} + \text{Solde Migratoire Interne} + \text{Entrées de l'étranger} - \text{Sorties vers l'étranger} \quad (1)$$

$$\text{Sorties vers l'étranger} = \text{Population } T + \text{Solde naturel} + \text{Solde Migratoire Interne} + \text{Entrées de l'étranger} - \text{Population } T+1 \quad (2)$$

Néanmoins, cette méthode testée en projection posait davantage de problème que la solution retenue. En effet, comme expliqué dans le cadre national, le solde apparent, estimé comme un résidu, intègre outre son objet d'intérêt, l'ensemble des erreurs de mesures inhérentes à la source. Ce problème déjà réel à l'échelle nationale devient encore plus important à l'échelle locale qui plus est sur des petites zones où ce solde peut s'avérer très erratique dans le temps ou proposer une vision faussée des flux avec l'étranger du territoire. In fine, il est apparu que localiser les sorties vers l'étranger via le solde migratoire apparent était souvent problématique, y compris sur des zonages de taille importante comme les départements.

Les naissances et les décès

La fécondité et la mortalité ont moins d'impact sur les projections locales que les flux migratoires. Les évolutions pour ces composantes se basent sur les caractéristiques observées du territoire cohérentes avec les hypothèses de l'exercice national :

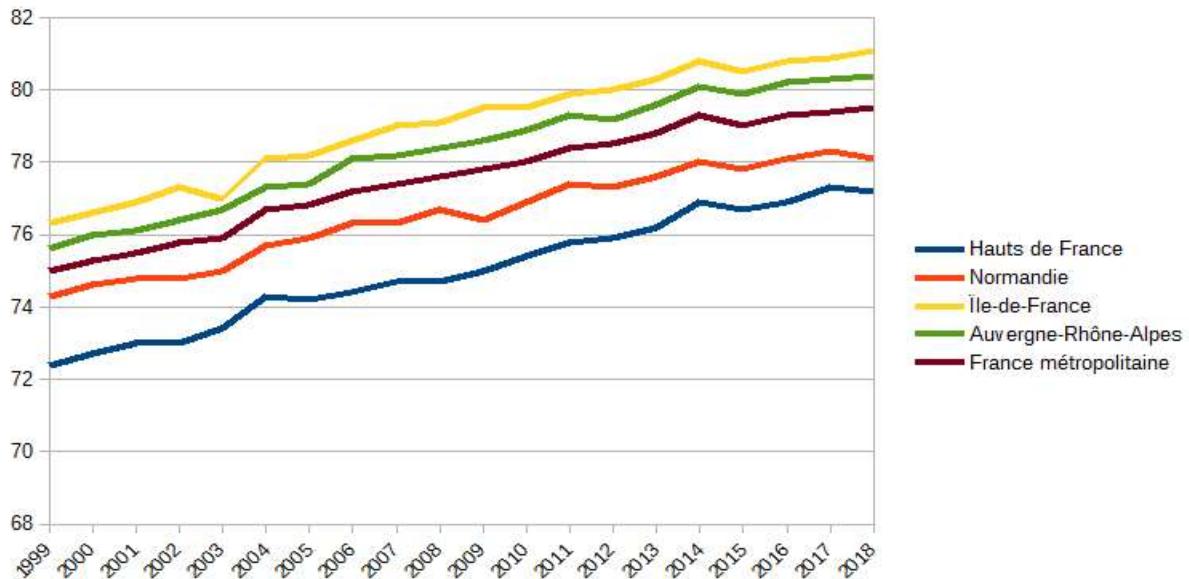
- l'indicateur conjoncturel de fécondité baisse légèrement, de 0,04, jusqu'en 2016, puis il est maintenu constant jusqu'en 2050 ;
- l'espérance de vie progresserait en passant au niveau national de 78,7 en 2013 à 86,8 ans pour les hommes et 85 ans à 90,3 ans pour les femmes en 2050.

En d'autres termes, un territoire avec une mortalité élevée verra son espérance de vie progresser dans les mêmes proportions que l'ensemble du pays sans rattrapage ou décrochage. Ce choix s'explique par 2 raisons :

- la fécondité et la mortalité présentent plus d'inertie que les migrations dans leur évolution et dans leurs effets sur les résultats des projections ;
- s'il existe des phénomènes de rattrapage ou de décrochage, il est impossible de les anticiper et de les systématiser dans un outil qui fonctionne sur des milliers de zonages à façon potentiels.

Les chroniques des espérances de vie à la naissance des hommes (figure III) illustrent bien ces difficultés. Si l'on considère les 2 régions avec l'espérance de vie la plus basse (Hauts-de-France et Normandie) et celles avec la plus haute (Île-de-France et Auvergne-Rhône-Alpes), toutes ont une évolution de leur espérance de vie proche de la moyenne nationale (+4,8 ans contre +4,5 ans) à l'exception de la Normandie dont le décrochage s'accentue (+3,8 ans). Aucune donnée démographique facilement mobilisable ne permet d'anticiper la spécificité normande. Au final, bien qu'il existe des différences de tendance entre territoires, l'application des hypothèses nationales semble la solution la plus simplement systématisable dans Omphale.

Figure III : Evolution de l'espérance de vie à la naissance des hommes entre 1999 et 2018



Source : Insee, États – Civils

Des projections locales cohérentes avec les nationales : le calage

Bien que s'appuyant sur les hypothèses de l'exercice France entière, les résultats d'Omphale sur une partition du territoire national ne convergeront pas spontanément avec la projection nationale pour plusieurs raisons :

- la population de départ utilisée à l'échelle locale est celle du recensement de la population alors qu'au niveau national, il s'agit du bilan démographique. Bien que globalement cohérentes, il existe des différences telles que le concept d'âge¹ ;

- des légères adaptations ont été apportées aux hypothèses nationales lors de leur implémentation dans Omphale comme la ventilation par âge du solde migratoire avec l'étranger ;

- enfin la principale cause de ces divergences provient de la méthode des composantes, qui n'est pas additive entre différentes échelles géographiques. Un cas d'école fictif peut illustrer ce phénomène (tableau 2) : un territoire peut être découpé en deux zones de populations équivalentes mais aux comportements en matière de fécondité très différents. S'il existe des flux importants de femmes en âge de procréer de la zone à forte fécondité vers celle à faible fécondité, la projection à l'échelle globale ne va pas en tenir compte. À l'inverse, une projection sur chaque zone va affecter aux femmes ayant migré le comportement de fécondité de leur nouvelle zone de résidence générant un déficit de naissance par rapport à la projection englobante. Au final, la projection au niveau géographique le plus fin (partie 1 et partie 2 de la zone X) conduit dans l'exemple à une estimation plus faible que la projection directe de la zone X.

¹ Le recensement considère l'âge révolu à la date de référence du recensement (troisième jeudi de janvier). Omphale assimile ces données à des populations au 1^{er} janvier sans y apporter de correction.

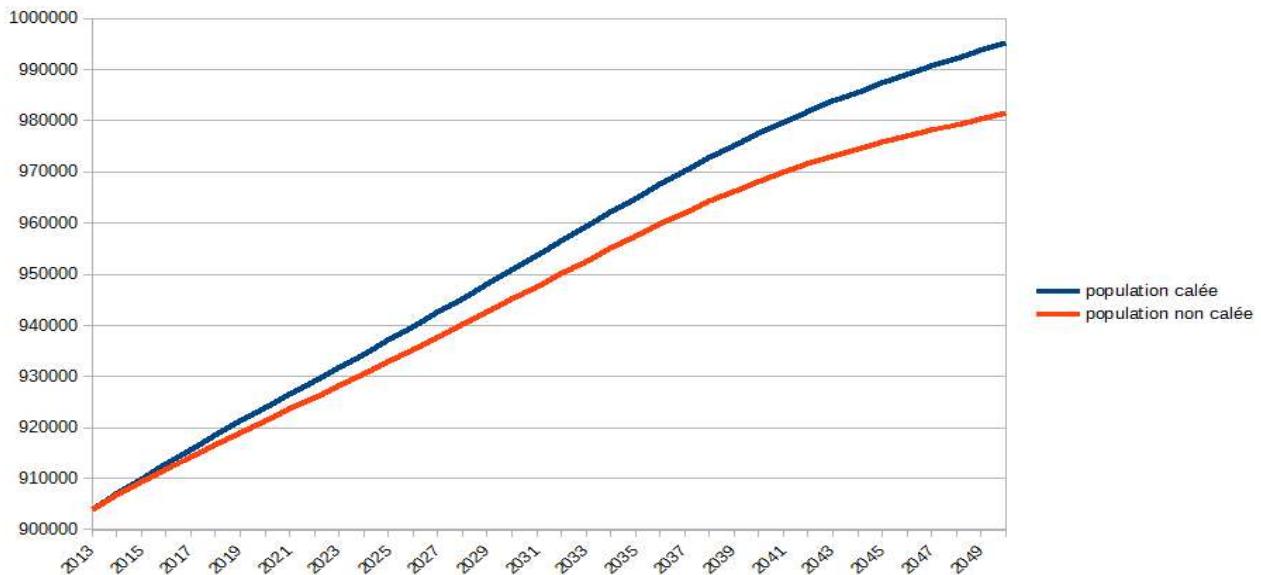
Tableau 2 : exemple de non additivité de la méthode des composantes

		Zone X	Zone X – partie 1	Zone X – partie 2	Zone X = partie 1 + partie 2
T	Population	1000	500	500	1000
	Nombre de femmes 15-49 ans	300	200	200	300
	Taux de fécondité	6,00 %	11,00 %	1,00 %	6,00 %
	Solde migratoire interne	0	-100	100	0
	Solde des femmes de 15-49 ans	0	-100	100	0
T+1	Femmes 15-49 ans	300	100	300	300
	Taux de fécondité	6,00 %	11,00 %	1,00 %	
	Naissances	18	11	3	14
	Décès	0	0	0	0
	Population en T+1	1018	411	603	1014

Ces divergences ne sont toutefois pas majeures, mais elles posent un problème de cohérence pour les utilisateurs des données de l'Insee, notamment ceux qui commandent des projections sur l'ensemble des régions, départements ou zones d'emploi de France par exemple. L'ultime étape dans la mécanique de projection d'Omphale consiste donc à caler les populations par sexe et âge sur les populations projetées au niveau national. Ainsi, les projections territoriales peuvent être sommées et correspondre à la projection nationale pour une année donnée. Le calage consiste donc à répartir entre les territoires par sexe et âge le différentiel de population entre la projection nationale et la somme des projections locales sur le territoire national. L'effet conduit donc à faire augmenter ou diminuer la population non calée d'Omphale d'autant plus qu'on s'éloigne du point de départ de la projection (figure IV). Le sens du calage est donc le même quel que soit le territoire mais son ampleur peut varier du fait de sa structure démographique. Sur l'ensemble des départements français, la hausse de croissance sur la période 2013 – 2050 liée au calage est en moyenne de +1,6 point (+10,9 % contre +12,6%) et varie de +1,2 point (Ardennes) à +1,9 point (Haute-Savoie).

Figure IV: Evolution de la population avant et après calage du Finistère

Source : Insee, Omphale 2017.



Il est à noter que seule la population peut être calée sur la projection nationale par ce biais. Les naissances et les décès ne le sont pas². De plus pour assurer la cohérence entre les composantes de la projection, le solde naturel est finalement estimé comme la différence entre la population calée et le solde migratoire. Il ne correspond donc plus parfaitement au solde entre naissances et décès.

Projection n'est pas prévision

Si Omphale est un outil au service des acteurs publics, il est nécessaire que ces derniers en comprennent la logique et les enjeux : compte tenu de la méthodologie explicitée ci-dessus, une projection démographique n'est ni une prévision, ni un outil de prospective. L'objectif d'Omphale est de fournir une ou des visions du futur possibles selon différents jeux d'hypothèses appelés scénarios. Parmi ces scénarios, le central joue un rôle majeur dans le sens où il vise à prolonger les tendances démographiques récentes. Il ne cherche donc pas à anticiper une rupture démographique majeure sur un territoire (par exemple l'arrivée de migrants, la fermeture d'une caserne, la création d'une université ...) ni à évaluer les effets ou les probabilités de réussite de politiques publiques (telle que la création de lotissements ou la lutte contre la périurbanisation). Le scénario central d'Omphale propose donc, en théorie, une vision du futur à comportements des habitants et à politique publique constants.

En l'absence de certaines informations (la localisation des flux vers l'étranger par exemple), des hypothèses du modèle peuvent être mal adaptées à certains territoires et générer des projections non satisfaisantes : il convient donc d'être au clair sur ce qui fait qu'une projection peut être jugée robuste ou pas. Une projection dans le cadre du scénario central pour laquelle il advient plusieurs années après sa réalisation qu'elle n'est pas conforme à ce qu'est advenu le territoire n'est pas, pour autant, problématique si elle a été prise pour ce qu'elle est : un futur possible dans le cadre d'une prolongation des tendances récentes. Les écarts illustrent simplement que des changements sont survenus sur le territoire de manière « spontanée » (par exemple une saturation du foncier qui limite les arrivées) ou suite à une politique publique. Une projection pourra davantage être soumise à critique si les rythmes de croissance observés sur la période passée et celle projetée divergent fortement. Cette prolongation des tendances passées semble donc le principal critère pour juger de la robustesse d'une projection. L'évaluation de la robustesse du modèle est primordial pour les utilisateurs d'Omphale, d'une part évidemment pour éviter d'induire en erreur des acteurs publics, mais d'autre part aussi parce que Omphale est un outil entièrement paramétrable.

2 Dans le cadre de la méthode des composantes, l'équation démographique (1) doit rester vraie à chaque pas de projection. Si l'ensemble des données des composantes (naissances, décès, solde migratoire) sont calées sur le niveau national, l'équation ne peut plus être vraie sans faire un arbitrage sur l'une des composantes. Aussi, seule la population est calée, afin de permettre de communiquer sur des résultats cohérents avec le niveau national. Les analyses sur les autres composantes de l'équation doivent donc être réalisées sur des données non calées, en communiquant sur des évolutions ou des contributions par exemple.

Le scénario central n'est qu'un scénario conçu comme une base de travail selon des hypothèses communes à tous les territoires. De ce fait, l'expertise de sa robustesse peut conduire à la création de scénarios à façon .

Dans deux régions, il existe des écarts forts entre les projections Omphale 2010 et le réalisé

Si projection n'est pas prévision, il est toutefois nécessaire pour l'utilisateur d'Omphale de comprendre la logique de cet outil : il est tout de même intéressant d'en appréhender les capacités prédictives, l'ampleur des écarts, pour adapter son usage du modèle. Pour cela, il est possible d'effectuer un retour en arrière sur le modèle précédent d'Omphale pour observer la réalisation ou non des projections. Les possibilités de zonages dans Omphale étant infinies, les contours des régions ont été privilégiées pour cet exercice.

Les régions pour lesquelles la projection proposée dans Omphale 2010 est très différente de la population observée en 2013 dans le recensement sont PACA (-1,6 points) et la Corse (+2,7 points) (tableau 3). Pour toutes les autres régions, l'écart entre l'évolution projetée et l'évolution réalisée est inférieur en valeur absolue à 1 point. Les écarts les plus faibles se trouvent dans les régions Auvergne Rhône-Alpes (0,2 point) et Occitanie (-0,3 point).

Tableau 3 : comparaison des projections Omphale 2010 et du recensement 2007 et 2013 par région

Région	Population 2007 Omphale 2010	Population 2013 Omphale 2010	Population observée 2013	Ecart en 2013 en effectifs	Evolution projetée avec Omphale 2010	Evolution Observée	Ecart entre les évolutions projetées et constatées (en points de pourcentage)
Provence-Alpes-Côte d'Azur	4 864 000	5 030 000	4 954 000	-77 000	3,4%	1,9%	1,56
Bourgogne-Franche-Comté	2 793 000	2 846 000	2 820 000	-27 000	1,9%	1,0%	0,93
Centre-Val de Loire	2 527 000	2 593 000	2 571 000	-22 000	2,6%	1,7%	0,87
Bretagne	3 120 000	3 283 000	3 259 000	-24 000	5,2%	4,5%	0,77
Grand Est	5 507 000	5 589 000	5 552 000	-37 000	1,5%	0,8%	0,67
Normandie	3 278 000	3 345 000	3 328 000	-17 000	2,0%	1,5%	0,52
Pays de la Loire	3 483 000	3 679 000	3 661 000	-18 000	5,6%	5,1%	0,52
Nouvelle Aquitaine	5 628 000	5 872 000	5 844 000	-28 000	4,3%	3,8%	0,50
Hauts-de-France	5 922 000	6 010 000	5 988 000	-22 000	1,5%	1,1%	0,37
Occitanie	5 371 000	5 693 000	5 684 000	-9 000	6,0%	5,8%	0,17
Auvergne Rhône-Alpes	7 405 000	7 737 000	7 758 000	20 000	4,5%	4,8%	-0,28
Ile-de-France	11 599 000	11 897 000	11 960 000	63 000	2,6%	3,1%	-0,54
Corse	299 000	312 000	320 000	8 000	4,3%	7,0%	-2,68

Source : Insee, Omphale 2010, recensement de la population 2013

En PACA, le taux de croissance annuel moyen observé entre 1999 et 2007 est de 0,6 point supérieur à celui observé sur la période 2007 et 2013. Le modèle Omphale projette des populations en fonction des tendances de moyen terme observées, les données disponibles lors de la réalisation de l'exercice de projection ne permettaient pas d'anticiper cette rupture de croissance due à l'évolution des flux migratoires.

La projection 2010 a donc surestimé la population. L'évolution observée entre 2007 et 2013 aura finalement été près de deux fois moins forte que celle projetée dans le modèle.

En Corse, Omphale 2010 a sous-estimé la population en 2013. L'évolution de la population observée entre les recensements 2007 et 2013 est de +7,0 % contre une projection de +4,3% sur la même période. Les données migratoires introduites dans le modèle 2010 conduisaient en effet à un fort ralentissement de la croissance de la population par rapport à la tendance observée entre 1999 et 2007 ; le taux de croissance annuel moyen sur la période 1999-2007 est de 1,8 %. Les données observées depuis ont bien conclu à un ralentissement de l'évolution de la population de la région après 2007 ; le taux de croissance annuel moyen sur la période 2007 à 2013 est de 1,1%, mais dans une moindre mesure par rapport à la projection (le taux de croissance annuel moyen en projection est de seulement +0,42% en projection sur la même période). En Corse, l'évolution de la population est essentiellement due au solde migratoire, le solde naturel présentant une contribution quasi nulle (voire négative en projection sous l'effet du vieillissement de la population), ce qui contribue à la fragilité des projections.

Comparaison théorique des modèles 2010 et 2017

Les principales divergences entre Omphale 2010 et le recensement en 2013 s'expliquent donc par des ruptures de la dynamique de la population dans les régions concernées. L'étape suivante conduit donc à s'interroger sur la capacité d'Omphale lors de ses mises à jour à tenir compte de ces derniers changements observés. Il convient donc de comparer les projections entre les modèles Omphale datant de 2010 et 2017. Cette comparaison s'avère en réalité difficile car outre l'espace de 7 ans et les changements comportementaux des habitants des territoires, certaines données en entrée du modèle ont changé tant dans leur champ que dans leur concept.

Une première difficulté tient au fait que les exercices nationaux de projection 2010 et 2016 sont difficilement comparables entre eux car le champ géographique des projections nationales a évolué : les projections 2010 sont réalisées sur le champ France métropolitaine tandis que celles de 2016 ont intégré les cinq Dom (Martinique, Guadeloupe, Guyane, la Réunion et Mayotte). Les hypothèses nationales de 2016 ayant été construites directement sur le champ France entière, sans passer par la construction d'une projection métropolitaine, aucune reconstitution d'une projection au niveau France métropolitaine à méthode analogue n'est possible.

Par ailleurs, entre 2010 et 2017, le modèle Omphale a connu une refonte conséquente : la modification de la mesure des flux internes issus du recensement de la population. Les flux migratoires mesurés dans le recensement sont passés d'une mesure quinquennale à une mesure annuelle.

La décomposition des effets des changements d'hypothèses sur les projections de population est dès lors difficile à reconstituer à la fois pour les projections au niveau national et au niveau local.

Les flux internes

Les données mobilisées dans Omphale pour mesurer les flux internes de population sont issues du recensement de la population. Depuis 2013, la mesure des flux a changé, en lien avec le questionnaire individuel de l'enquête. La question permettait jusque-là de mesurer des flux quinquennaux (« Où habitez-vous au 1^{er} janvier N-5 ? ») ; elle porte depuis 2013 sur une mesure annuelle des migrations (« Où habitez-vous au 1^{er} janvier N-1 ? »). La période d'observation des flux étant plus courte, la probabilité de migrer pour un individu est plus faible. Les flux mesurés à partir du recensement 2013 sont donc moins nombreux en volume que dans le modèle précédent. Si l'on observe plus précisément les années de bascule entre les 2 questions (2010 et 2011) dans les enquêtes annuelles de recensement (EAR), les flux ont été divisés par 3,3 pour ceux de communes à communes jusqu'à 3,6 pour ceux de départements à départements (tableau 4). La mesure prend en revanche davantage en compte des migrations multiples et des allers-retours ce qui explique que le nombre de migrations sur 1 an est plus important que le cinquième de la mesure quinquennale.

Tableau 4 : comparaison du volume de flux selon le niveau géographique et le millésime du recensement

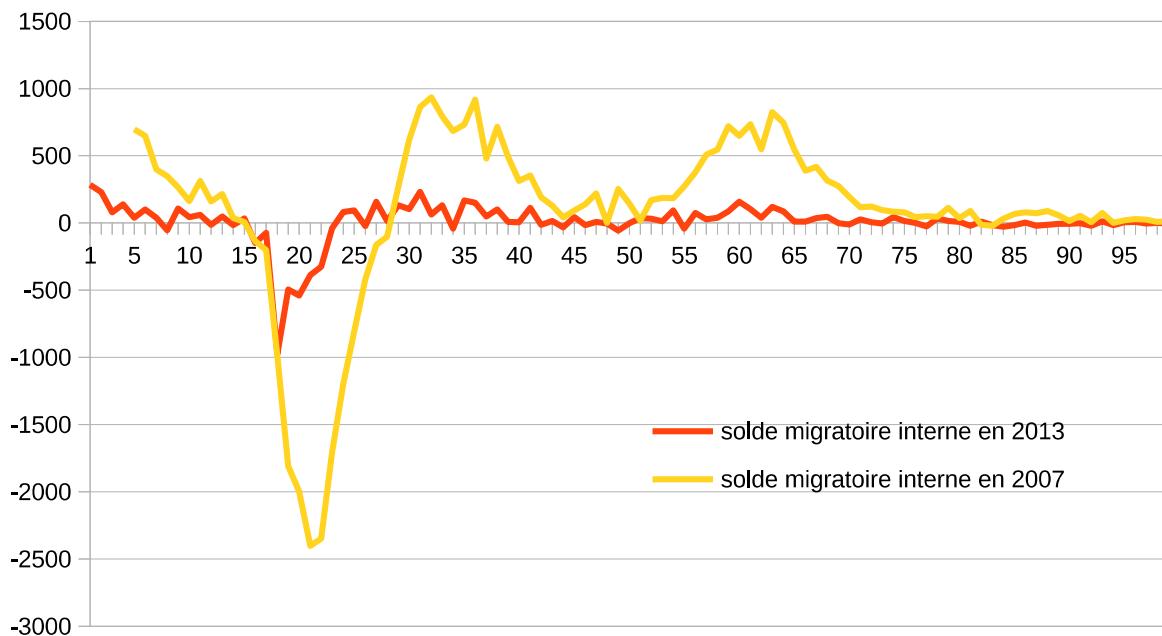
	De communes à communes	De zones d'emploi à zones d'emploi	De départements à départements
EAR 2010 (flux entre 2005 et 2010) (1)	13 947 000	7 168 000	5 933 000
EAR 2011 (flux entre 2010 et 2011) (2)	4 222 000	2 068 000	1 663 000
Ratio (1)/(2)	3,3	3,5	3,6

Champ : Flux internes à la France métropolitaine sur les 5 ans et plus.

Source : Insee, Enquêtes annuelles de recensement 2010 et 2011.

Outre une modification du nombre de migrants, on observe un resserrement des tranches d'âge concernées par les flux pour les étudiants et les jeunes actifs avec la mesure annuelle. En effet, les flux mesurés sur 5 ans ne permettent pas de connaître précisément la date de la migration et donc l'âge du migrant au moment de la migration. Par exemple, le solde migratoire de la région Centre Val de Loire avec le reste de la France (c'est-à-dire le solde migratoire interne) présente un profil différent selon qu'il est estimé sur 5 ans dans l'EAR 2010 ou sur un an dans l'EAR 2011 (figure V). En particulier un solde positif est estimé sur les 0 à 10 ans sur cinq ans ce qui n'est pas le cas sur un an. A l'inverse, le fort déficit entre 16 et 27 ans autre qu'il est largement réduit du fait du changement de référentiel temporel est bien plus concentré sur les 17 – 21 ans.

Figure V: Solde migratoire de la région Centre Val de Loire avec le reste de la France



Sources : Insee, EAR 2010 et 2011.

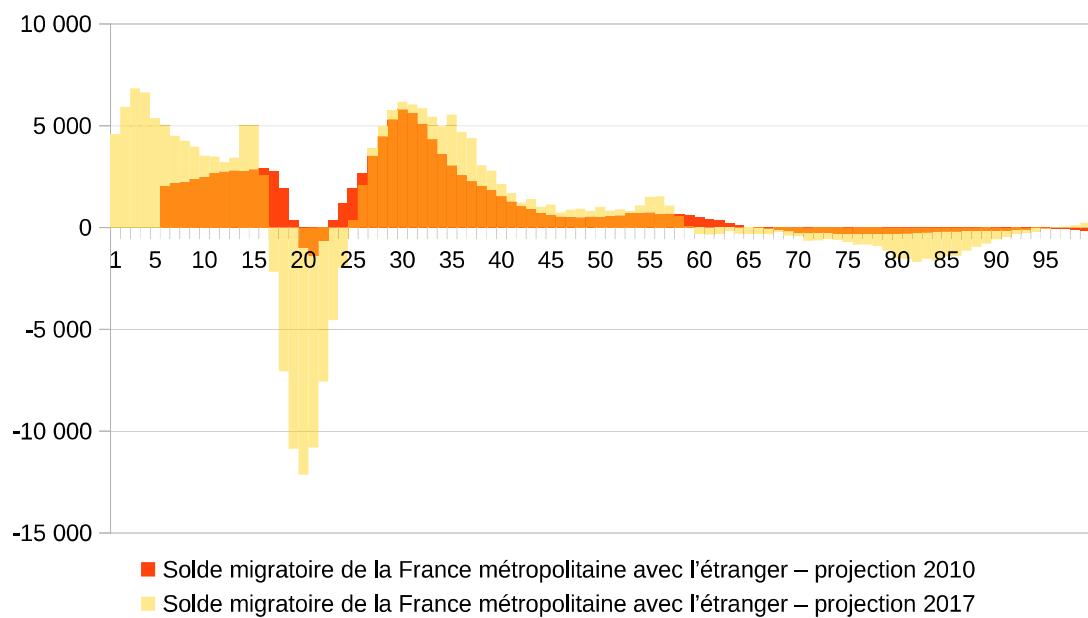
Le changement dans la mesure des flux a également impacté la mécanique du modèle, dans la version 2010, les projections étaient, du fait de la mesure des flux, réalisées sur un pas quinquennal, les données projetées directement par la méthode des composantes correspondaient aux années 2012, 2017, 2022, etc. La pyramide des âges était donc progressivement déformée sur un pas de 5 ans. Entre deux années charnières, les points intermédiaires étaient estimés par interpolation. Le passage aux flux annuels dans le modèle 2017 a permis de lever cette contrainte de pas quinquennal et de travailler directement sur des projections annuelles. L'interpolation linéaire générait parfois des profils heurtés sur certaines zones. Le passage au pas annuel génère des évolutions plus lisses. Cependant, la mesure des flux sur un an amène potentiellement plus de fluctuations, réelles ou liées à la collecte des Enquêtes annuelles du recensement 2013. Dans les deux modèles, la mesure des flux est conditionnée par les limites inhérentes à l'échantillonnage du recensement, à son caractère déclaratif et à l'absence de mesure des flux en direction de l'étranger.

Les flux avec l'étranger

Le niveau du solde migratoire avec l'étranger dans le scénario central de projection a été revu à la baisse dans les projections nationales (+70 000 en 2010 contre +100 000 pour le dernier exercice). La révision de cette hypothèse est issue d'une réflexion nationale à partir des dernières estimations de population disponibles et des travaux récents de l'Insee sur les entrées et les sorties du territoire (Brutel, 2014).

La répartition du solde migratoire par sexe et âge a également été modifiée entre les deux exercices nationaux (figure VI). Le déficit migratoire des 18-25 ans est exacerbé dans l'hypothèse 2017 par exemple.

Figure VI : Hypothèses de distribution du solde migratoire avec l'étranger par âge selon les projections



Champ : France métropolitaine.
 Sources : Omphale 2010 et Omphale 2017.

Les naissances et les décès

Au niveau national, la révision de l'hypothèse de mortalité a conduit entre 2010 et 2016 à relever l'espérance de vie à la naissance, essentiellement pour les hommes. De fait, l'écart d'espérance de vie à la naissance entre les hommes et les femmes est réduit par rapport à l'exercice précédent. Les projections 2016 révisent les quotients de décès à la baisse pour tous les âges chez les hommes et seulement pour les 65 ans et plus pour les femmes. Dans Omphale, les quotients de décès locaux par sexe et âge évoluent parallèlement à la tendance nationale à partir de 2016. Aussi, si l'espérance de vie des hommes et des femmes augmente en projection dans toutes les régions, le classement des zones par espérance de vie reste inchangé entre le début et la fin de la projection.

Les hypothèses de fécondité nationales retiennent un indicateur conjoncturel de fécondité (ICF) de 1,95 enfants par femme, constant entre 2016 et 2070³ dans le scénario central ; dans la projection 2010, l'ICF était également stabilisé à 1,95 à partir de 2015.

³ Les évolutions des premières années de la projection nationale sont basées sur les taux provisoires de fécondité par âge issus de l'état civil. Ces données ne sont pas mobilisées localement dans les projections Omphale. Une interpolation des taux de fécondité par âge a donc été réalisée pour définir une évolution applicable à tous les zonages à façon dans Omphale.

L'âge moyen à la maternité augmente progressivement en projection pour atteindre 32 ans en 2040. Dans la projection 2010, l'âge moyen à la maternité augmentait également, mais plus légèrement et il se stabilisait en 2015 à 30,4 ans. Dans Omphale, seule l'hypothèse de maintien de l'ICF a pu être intégrée dans le modèle 2010 comme dans le modèle 2017, Le modèle ne permet en effet pas de paramétriser des hypothèses valables sur tous les territoires où à la fois l'âge moyen à la maternité augmente et l'ICF est constant. La priorité a été donnée au maintien de l'ICF constant au niveau des territoires. Pour réaliser cette hypothèse, les quotients de fécondité sur la zone d'intérêt sont maintenus constants à partir de 2016 et jusqu'à l'horizon de projection. Ainsi, les indices conjoncturels de fécondité sont également constants et dépendent du niveau de départ calculé pour la zone d'intérêt. Dans les modèles 2010 et 2017, étant donné les hypothèses retenues sur les quotients de fécondité, l'âge moyen à la maternité reste stable (tableau 5).

Tableau 5 : Comparaison des mécanismes et des hypothèses des modèles Omphale 2010 et Omphale 2017

	Omphale 2010	Omphale 2017
Point de départ	Population par sexe et âge détaillés mesurée dans le recensement 2007	Population par sexe et âge détaillés mesurée dans le recensement 2013
Horizon	2042 (7 cycles de 5 ans)	2050
Pas de projection	Pas quinquennal (« Où habitez-vous au 1 ^{er} janvier N-5 ? »)	Pas annuel (« Où habitez-vous au 1 ^{er} janvier N-1 ? »)
Géographie de référence	France métropolitaine + hypothèses à façon pour 4 DOM	France entière (y compris Mayotte)
Migrations internes	Maintien des quotients d'émigration quinquennaux constants	Maintien des quotients d'émigration annuels constants
Migrations avec l'étranger	Construite à partir des hypothèses France métropolitaine de la projection nationale 2010	Construite à partir des hypothèses France entière de la projection nationale 2016
Fécondité	Mesure des quotients de fécondité par les naissances détaillées par âge de la mère sur la période 2011-2015 dans l'état civil et de la population des femmes dans le recensement. Evolution de ce quotient : maintien des quotients de fécondité par âge et de l'ICF constants au niveau de l'année de départ de la projection.	Mesure des quotients de fécondité par les naissances détaillées par âge de la mère sur la période 2011-2015 dans l'état civil et de la population des femmes dans le recensement. Evolution de ce quotient : légère baisse de l'ICF jusqu'en 2016, conformément aux hypothèses nationales retenues en début de période ⁴ , puis maintien des quotients de fécondité par âge et de l'ICF constants jusqu'en 2050.
Mortalité	Mesure par les quotients de décès au niveau France métropolitaine.	Mesure des quotients par les décès détaillés par sexe et âge sur la période 2011-2015 dans l'état civil et de la population par sexe et âge dans le recensement. Evolution de ce quotient : légère hausse des quotients de décès jusqu'en 2016 conformément aux hypothèses nationales retenues en début de période, puis évolution parallèle à la tendance nationale jusqu'en 2050.
Calage des données	Projections France métropolitaine 2010 par sexe et âge détaillés.	Projections France entière 2016 par sexe et âge détaillés.

Source : Insee, Omphale 2010 et 2017.

⁴ En effet, dans les projections nationales, les premiers pas de projection sont basés sur des données observées alors dans l'état civil sur les naissances et les décès entre 2013 et 2016, le modèle Omphale reproduit la tendance globale de ces données pour pouvoir définir une hypothèse de projection utilisable quel que soit le zonage projeté.

Comparaisons Omphale 2010 – Omphale 2017 sur la période 2013 - 2018

Comparer les projections entre les deux modèles d'Omphale se révèle un exercice difficile étant donné les changements apparus dans la méthode de projection au niveau national (passage du champ France métropolitaine au champ France entière) et la révision de la mécanique de projection au niveau local (passage des flux quinquennaux aux flux annuels). Il n'existe pas de quantification simple des différences entre les projections de 2010 et 2017. Pourtant, il est possible de comparer les résultats des deux modèles sur la période 2013 – 2018, qui constitue le deuxième pas de projection d'Omphale 2010 et le début de la projection d'Omphale 2017 (tableau 6). Cette comparaison ne vise pas à s'intéresser aux niveaux des populations projetées mais aux rythmes de croissance. Elle est particulièrement pertinente pour les 2 régions où Omphale 2010 s'est écarté du recensement de manière accentuée.

Tableau 6 : Comparaison des évolutions 2013 – 2018 dans Omphale 2010, Omphale 2017 et le recensement

Région	Écart entre les évolutions projetées et constatées (en points de pourcentage)	Evolution 2013-2018 projetée avec Omphale 2010	Evolution 2013-2018 projetée avec Omphale 2017	Écart entre les évolutions Omphale 2017 et Omphale 2010 sur 2013-2018 (en points de pourcentage)
Provence-Alpes-Côte d'Azur	1,6	2,5%	1,5%	-1,0
Bourgogne-Franche-Comté	0,9	1,3%	1,0%	-0,3
Centre-Val de Loire	0,9	1,8%	1,2%	-0,6
Bretagne	0,8	3,7%	3,4%	-0,3
Grand Est	0,7	0,8%	0,9%	0,1
Normandie	0,5	1,3%	0,8%	-0,5
Pays de la Loire	0,5	4,0%	4,0%	0,0
Nouvelle Aquitaine	0,5	3,1%	3,0%	-0,1
Hauts-de-France	0,4	0,9%	1,0%	0,1
Occitanie	0,2	4,4%	3,8%	-0,6
Auvergne Rhône-Alpes	-0,3	3,2%	3,8%	0,6
Île-de-France	-0,5	1,7%	1,9%	0,2
Corse	-2,7	3,2%	4,4%	1,2

Source : Insee, Omphale 2010 et 2017, recensement de la population 2013

En Paca, la population entre les recensements 2007 et 2013 a marqué un ralentissement non anticipé par Omphale 2010. Les projections 2017 intègrent désormais des données fraîches notamment sur les échanges migratoires moins bénéficiaires sur les dernières années : les flux d'arrivées observés aux âges de la retraite sont désormais moins marqués. La croissance de la population s'avère donc réduite par rapport à Omphale 2010 (+1,5 % sur 2013 – 2018 contre +2,5 %) et semble désormais cohérente avec les dernières tendances disponibles.

À l'inverse, Omphale 2010 sous-estimait la croissance de la population corse entre 2013 et 2017 et prolongeait cette tendance de progression sur la suite de la projection. L'introduction de nouveaux flux migratoires encore plus favorables permet de redynamiser la croissance démographique corse dans Omphale 2017 sur 2013 - 2018 (+4,4 % contre +3,2% dans Omphale 2010).

Dans d'autres régions (Pays-de-la-Loire, Grand Est, Nouvelle-Aquitaine, Hauts-de-France voire Bourgogne Franche Comte et Île-de-France), malgré des écarts de niveau plus ou moins importants sur la population 2013, les changements intervenus dans le modèle de projection semblent avoir relativement peu d'impact sur l'évolution de la population projetée. En Grand-Est et Bourgogne-Franche-Comté, le vieillissement de la population et les hypothèses de mortalité jouent à plein pour expliquer les évolutions de populations d'où une importante stabilité du modèle. Ces zones sont en effet relativement peu concernées par les flux internes ou les flux avec l'étranger, le solde migratoire étant quasi nul.

En Nouvelle Aquitaine, en Occitanie, dans les Pays de Loire, dans les Hauts de France ou encore en Île-de-France, les écarts constatés entre les populations projetées en 2010 et en 2017 concernent plutôt la fin de période de projection, qui est plus impactée par les hypothèses sur la mortalité. Par exemple, pour la Nouvelle Aquitaine et l'Occitanie, les hypothèses 2017 sur la mortalité réduisent l'espérance de vie des femmes par rapport à la projection 2010, en lien avec les évolutions observées au niveau national.

D'une manière générale, pour les femmes l'espérance de vie est plus souvent revue à la baisse entre les projections du modèle 2010 et du modèle 2017. En effet, au niveau régional, l'impact des hypothèses nationales sur les quotients de décès et sur l'espérance de vie va dépendre de la structure de population par sexe et âge du territoire. Pour les hommes, quelle que soit la région, l'espérance de vie en 2013 est révisée à la hausse. Cette révision est de l'ordre de 5,5 mois pour les régions de Bretagne, des Hauts de France et de Corse. Pour ce qui concerne l'ICF, dans les projections 2017, l'indicateur est révisé à la hausse dans 6 des 13 régions et la révision la plus importante concerne la région PACA. Sur cette région, l'écart sur les ICF projetés en 2013 avec les modèles 2010 et 2017 s'explique essentiellement par la fécondité des femmes de 30 à 40 ans.

Le changement de méthodologie sur les flux internes (passage d'une mesure quinquennale à une mesure annuelle) et sur les flux externes (changement du solde national et de sa répartition par âge) se répercute différemment au niveau des régions. Dans certaines régions (Île-de-France, Occitanie, Bretagne,...), on retrouve un décalage d'âge sur les pics et les creux associés aux mouvements migratoires en raison du changement de mesure sur les flux internes, en particulier aux âges étudiants, des jeunes actifs ou à l'âge de la retraite.

Par exemple en Île-de-France, les mouvements migratoires correspondant aux départs à la retraite sont centrés autour de 60 ans dans Omphale 2017 contre 57 ans dans Omphale 2010 ; le pic d'arrivée des étudiants ou des jeunes actifs est estimé à 22 ans, contre 20 ans avec la mesure annuelle des flux.

Sans surprise au regard de l'hypothèse nationale, le solde migratoire avec l'étranger négatif pour des étudiants impacte fortement certaines régions, en particulier ce solde externe devient négatif pour certaines d'entre elles (Grand-Est, Hauts-de-France) ou se creuse plus nettement (Auvergne-Rhône-Alpes). Aux âges des jeunes actifs, le pic d'entrées en France disparaît ou s'atténue fortement dans certaines régions, comme en Normandie ou en Nouvelle Aquitaine ; ce qui peut contribuer à une moindre augmentation de la population.

Ainsi, les changements introduits dans Omphale 2017 conduisent à réviser les messages d'Omphale 2010 en tenant compte des dernières tendances démographiques. Pourtant cela ne signifie pas pour autant qu'Omphale 2017 « prédirait » mieux l'avenir qu'Omphale 2010. Les hypothèses introduites dans le modèle Omphale 2017 proposent comme en 2010 des tendances de moyen terme qui peuvent être rapidement contredites sans que le principe de l'exercice de projection ne soit remis en question. Par exemple, dans toutes les régions métropolitaines, on observe dans l'état civil une baisse de l'ICF à partir de 2013. Cette baisse de la fécondité concerne tous les âges avant 35 ans, elle est d'autant plus forte sur les plus jeunes. Ces nouvelles tendances ne sont pas introduites dans les scénarios standard d'Omphale 2017, avec un peu de recul temporel (pour confirmer qu'il s'agit d'une nouvelle tendance de la fécondité et pas d'un épiphénomène), une comparaison de la fécondité observée et projetée pourra nous donner des éléments d'analyse de l'évolution des populations sur les territoires concernés par la rupture de tendance. La simulation Omphale peut s'apparenter à une population « toutes tendances conservées ».

Évaluer la robustesse d'une projection locale

Les écarts entre les populations issues du recensement et les projections Omphale à l'échelle des régions ont tendance à s'accroître en descendant à des niveaux géographiques fins. Le modèle étant plus volatile sur la question des flux migratoires, il est nécessairement moins robuste sur des niveaux géographiques fins où les effets conjoncturels mesurés dans les flux internes ont plus d'impact sur la population projetée. Aussi, au niveau des départements de France métropolitaine, pour la période 2007-2013, l'écart d'évolution de la population entre la projection 2010 et le recensement 2013 varie entre -4,3 points pour la Corse du Sud et +4,7 points pour la Lozère sur l'évolution de la période 2007-2013.

Au niveau des zones d'emploi de France métropolitaine, l'écart d'évolution fluctue entre +7,2 points (Corte) et -10,2 points (Sartène) sur l'évolution de la période 2007 – 2013 ; la population de ces zones d'emploi étant de l'ordre de 10 000 habitants. La projection et la réalisation s'écartent d'autant plus que les territoires en jeu sont petits, le sens de l'évolution de la population peut également être inversé entre la projection et le réalisé. Ces écarts sont souvent interprétés par les utilisateurs d'Omphale comme des « erreurs » du même ordre qu'un résidu dans un modèle économétrique. C'est alors se méprendre sur l'objectif d'une projection de population : il n'est pas question de prédire l'avenir mais de proposer une vision simulatoire du futur dans le cadre de certaines hypothèses. De ce fait dès lors qu'un territoire connaît un changement significatif dans le comportement de ses habitants, Omphale divergera mécaniquement, quelques années plus tard, de la population observée sans pour autant remettre en cause la qualité de la projection.

Dès lors que la puissance prédictive n'est pas un critère de qualité d'une projection de population, peut-on dire que toutes les projections issues d'Omphale se valent et présentent chacune une robustesse satisfaisante ? En réalité, il arrive que des projections issues d'Omphale puissent être considérées comme non satisfaisantes mais ce jugement se fait sous un autre prisme que celui de la capacité à prévoir l'avenir. Comme Omphale ne traduit que l'effet d'hypothèses démographiques sur un territoire il s'agit alors essentiellement d'interroger les hypothèses pour voir si elles sont adaptées au territoire et donc si elles permettent d'atteindre l'objectif du scénario central : prolonger les dernières tendances locales.

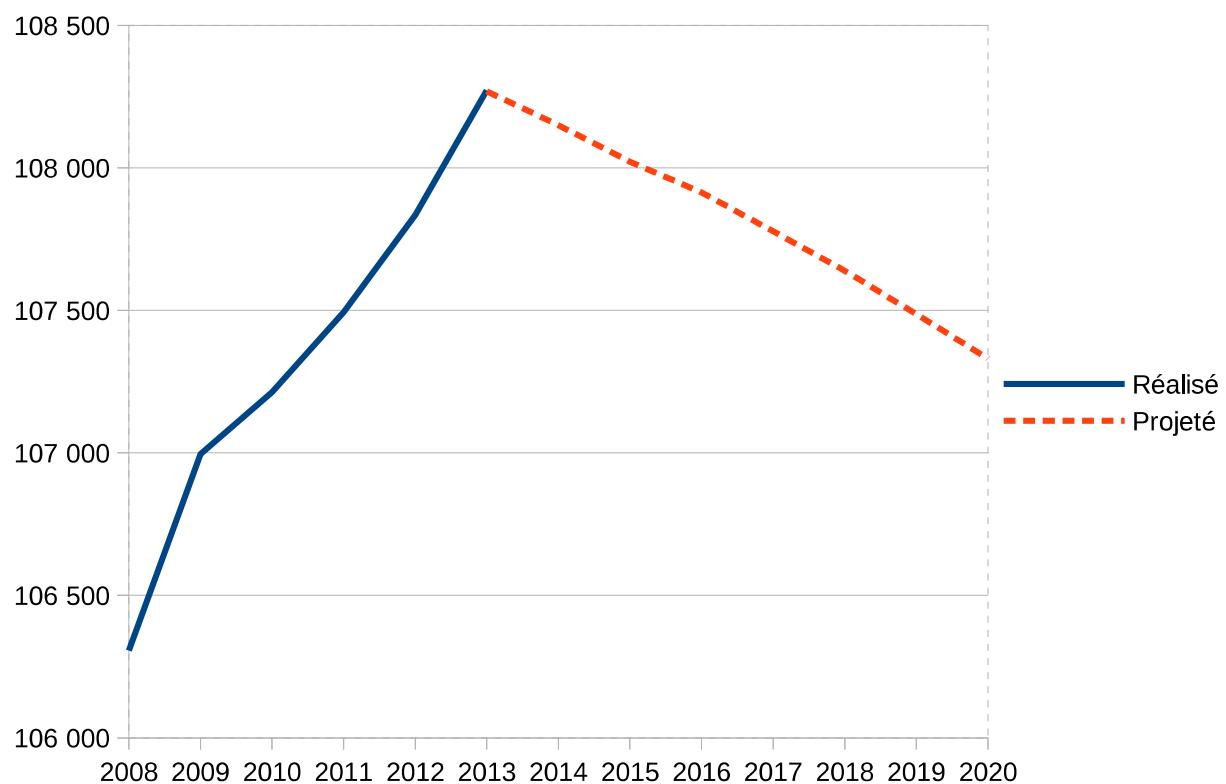
En partant du principe qu'une projection dite robuste dans le cadre du scénario central est une projection qui prolonge de manière cohérente les dernières tendances observées sur le territoire, il est alors primordial de s'interroger sur la manière d'en juger. Deux livrables d'Omphale sont alors à la disposition du statisticien pour cela : les évolutions de population et les pyramides des âges.

Une approche empirique

Les statisticiens de l'Insee réalisent d'ores et déjà des analyses de robustesse de manière très empirique, « à dire d'experts » en observant les courbes de population sur le passé et en projection pour évaluer d'éventuelles ruptures. Des discontinuités trop fortes entre le passé et la projection peuvent attester d'un manque de robustesse d'Omphale sur un territoire. Si sur un territoire, il existe un déséquilibre fort entre les entrées et les sorties de France, Omphale pourrait sous-estimer sa population en générant tout de même des départs importants : c'est le cas, par exemple en Île-de-France mais aussi sans doute dans l'arrondissement de Soisson (figure VII). On observe sur ce territoire une forte rupture de tendance au début de la projection.

Les hypothèses de mortalité, de fécondité ou de flux avec le reste de la France suivent les tendances passées. Par déduction, ce sont les migrations avec l'étranger qui sont à l'origine de cette défaillance du modèle. Il est probable que le rapport entre entrées et sorties vis-à-vis de l'étranger y soit largement supérieur à celui du niveau national. Mais sans moyen de mesurer cet état de fait, on attribue un poids des sorties de Soissons vers l'étranger parmi les départs nationaux équivalent à celui mesuré dans les entrées. Il est donc probable qu'Omphale sur-estime les sorties et génère cette rupture de tendance en 2013.

Figure VII : Réalisation et projection de population dans l'arrondissement de Soissons

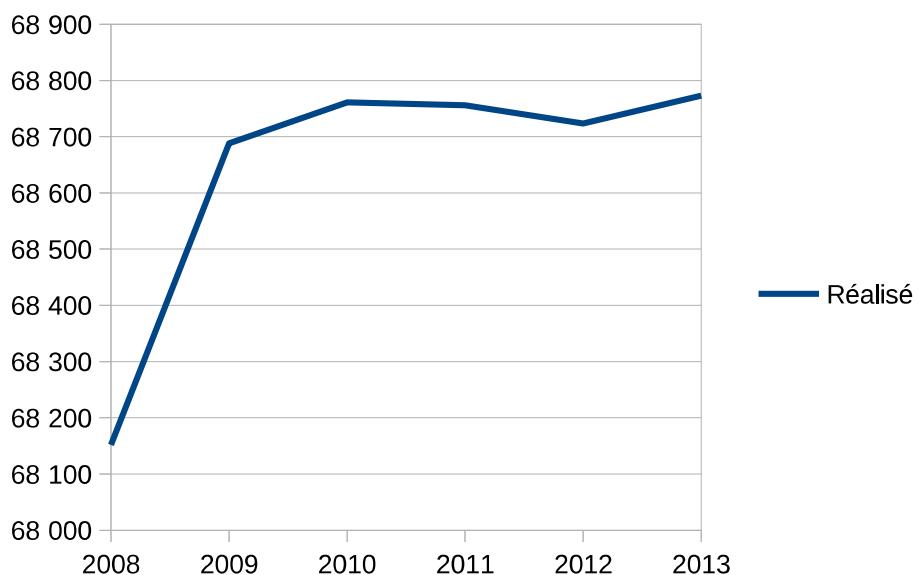


Source : Insee, Omphale 2017 et recensements 2008 à 2013.

Ainsi, un premier indice de la robustesse d'Omphale est bien la prolongation des tendances passées en comparant les rythmes de croissance des populations observées dans le recensement et celles projetées. Pour autant, cette première méthode présente au moins trois limites :

- il est difficile de déterminer quel recul historique retenir pour estimer la tendance démographique passée du territoire. En toute rigueur, la mécanique du recensement devrait conduire à privilégier des évolutions quinquennales. Dans ce cas, on fait fi des ruptures éventuelles au cours des cinq dernières années comme pour l'arrondissement de Château-Thierry (figure VIII). Le mode de collecte glissant du recensement devrait en théorie conduire à éviter de dégager des tendances à partir d'évolutions annuelles : pourtant la population a-t-elle réellement augmenté puis stagné ou a-t-elle augmenté en continu au fil de la période ? Selon que l'on choisisse l'une ou l'autre, l'évaluation que l'on pourra faire de la projection Omphale associée sera différente.

Figure VIII : Evolution de la population dans l'arrondissement de Château-Thierry

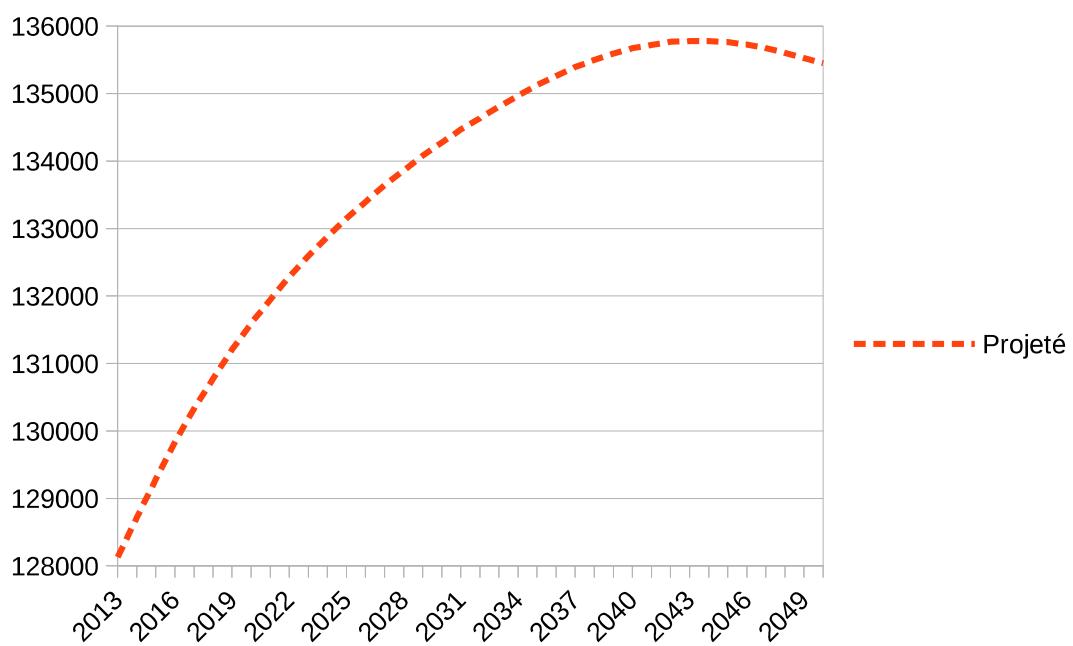


Source : Insee, recensements 2008 à 2013.

- il n'existe pas de seuil qui permettrait de juger si le différentiel entre l'évolution récente observée et celle projetée peut s'avérer problématique. Dans le cas de Soissons, la rupture est franche mais pour d'autres cas elle est plus légère et la décision pourra être considérée comme arbitraire. Qui plus est, il faut avoir en tête que la mécanique d'Omphale a pour objectif de prolonger les tendances passées mais que le résultat n'est pas forcément une adéquation parfaite entre taux de croissance projetés et observés. Un territoire où les comportements démographiques restent inchangés verra tout de même sa croissance de population affectée par des phénomènes tels que le vieillissement de la population ou la dynamique démographique des zones avec lesquelles il échange.

Une projection robuste ne reproduit donc pas à l'identique les derniers taux de croissance observés : si les évolutions trop brutales sont suspectes, des changements assez profonds peuvent intervenir sur des périodes plus longues. La projection du scénario central sur l'arrondissement de Saint-Omer (figure IX) témoigne des infléchissements progressifs à la dynamique démographique du fait du vieillissement de la population : l'importance des décès des générations du baby boom réduit progressivement le taux de croissance de la population. A partir des années 2040, les décès seraient tels que la population diminuerait sans pour autant modifier les hypothèses d'attractivité résidentielle.

Figure IX: Projection de population dans l'arrondissement de Saint-Omer



Source : Insee, Omphale 2017.

- l'intérêt d'Omphale n'est pas uniquement de donner des informations sur la population totale mais également de la ventiler par sexe et âge. Les politiques d'aménagement du territoire à mener ne seront pas les mêmes selon que la croissance de la population est portée par des étudiants, des couples avec enfants ou des seniors. De plus, les projections de population et en particulier celles issues d'Omphale permettent de réaliser des projections sur d'autres populations d'intérêt : actifs, ménages, élèves, personnes âgées dépendantes ... Par conséquent, la robustesse d'une projection ne se limite pas à l'étude de la population totale mais doit également valider les évolutions par âge. Cette analyse est plus complexe car la présence de classes d'âge pleines et de classes d'âges creuses conduisent à des évolutions parfois très erratiques du nombre de personnes au sein d'une classe d'âge.

Une approche statistique : analyser de manière systématique les effets âges et de génération

Pour tenter de répondre aux trois critiques énoncées précédemment de l'analyse à dire d'expert de la robustesse d'une projection, une méthode visant à détecter des projections potentiellement problématiques est proposée aux utilisateurs d'Omphale, en intégrant notamment une analyse par âge. Il va de soi qu'une méthode systématique ne peut se substituer à la connaissance territoriale d'un statisticien local : le but de cette méthode est de traiter rapidement un nombre important de zones afin de cibler celles sur laquelle pousser prioritairement l'expertise. Le principe de base d'Omphale est de déformer séquentiellement une pyramide des âges dans le temps. Deux grands types de phénomènes entrent en jeu pour expliquer les formes constatées et projetées d'une pyramide des âges :

- les effets de générations sont constitués par des différences d'effectifs entre les classes d'âge liés à la génération à laquelle appartiennent les individus. Ils sont donc la résultante de phénomènes relativement structurant, nationaux ou internationaux dont un bon exemple est le baby boom. Dans la pyramide des âges de la France en 2013 et dans la plupart des pyramides des âges locales, le baby boom se traduit par des effectifs importants entre 40 et 65 ans ;
- les effets âge tiennent aux caractéristiques par âge du territoire observé, notamment en lien avec les différents comportements en matière de migrations résidentielles au long du cycle de vie. Les effets âge peuvent exister au niveau national mais ils sont bien plus visibles à l'échelle locale : les pôles universitaires ont une pyramide des âges très caractéristique avec des effectifs nombreux depuis l'entrée dans l'enseignement supérieur jusqu'aux premières années d'activité et des classes plus creuses ensuite liées notamment au phénomène de périurbanisation.

Chacun de ces deux phénomènes a des répercussions différentes en projection : les effets de générations sont soumis au vieillissement de la population (les baby boomers vont continuer à vieillir dans les prochaines années sur tout le territoire national) tandis que les effets âges sont plus constants dans le temps (un territoire étudiant le reste quand bien même les étudiants de 2013 vieillissent, ils sont remplacés par ceux des générations suivantes). Cette distinction entre effet âge et effet de générations se fait dans Omphale par le biais des quotients de migration : les classes pleines liées au baby boom ne sont pas liées à des arrivées sur la zone d'intérêt, elles vont continuer à s'élever dans la pyramide des âges. À l'inverse, celles liées aux étudiants s'expliquent par des entrées et des sorties modélisées dans Omphale ce qui permet un maintien de la classe pleine à partir de 17 ans.

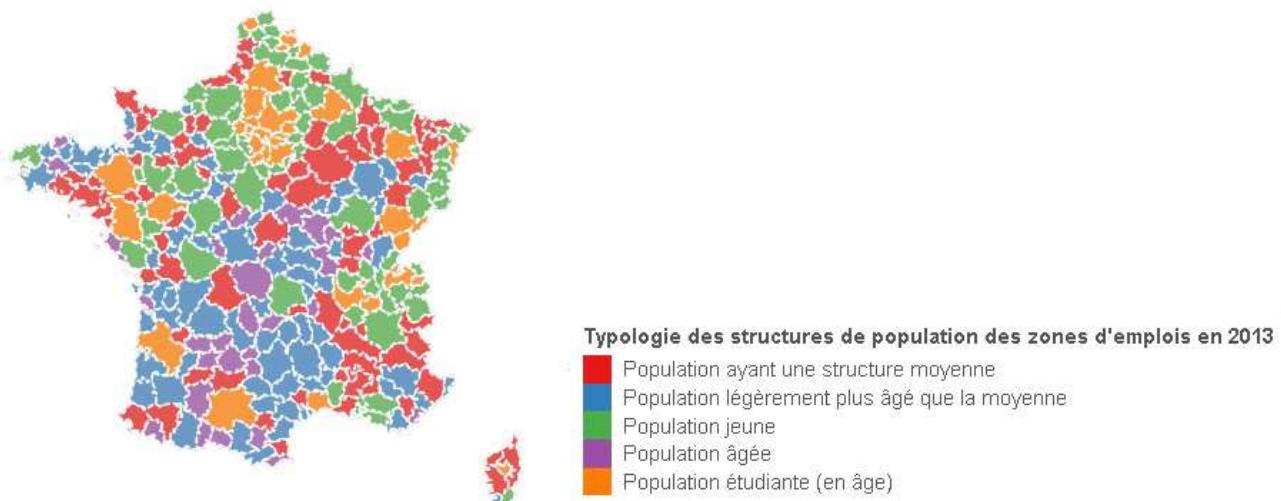
Par conséquent, de manière plus poussée que l'observation de la seule évolution de l'ensemble de la population, un critère de robustesse peut tenir au fait que les tendances passées sont prolongées lorsqu'en projection les effets âges et les effets générations ont été respectés.

Il n'est pourtant pas simple de distinguer de manière systématique et automatique les effets de générations et les effets âge. Un territoire comme Paris connaît un déficit migratoire important des jeunes retraités par conséquent, l'effet âge lié à ce phénomène va influencer l'effet lié au vieillissement des baby boomers. Comment dès lors qu'on ne peut pas déterminer *a priori* pour toute zone les effets âges et de générations s'assurer que la déformation de la pyramide des âges respecte les caractéristiques du territoire ?

La méthode retenue ici vise à déterminer des types ou classes de pyramides des âges en 2013 par le biais d'une typologie : une classification par la méthode k-means⁵ a été réalisée une fois pour toute sur l'ensemble des zones d'emploi de France métropolitaine (figure X). Pour chacun de ces types de pyramide des âges sont associées les déformations les plus fréquentes entre 2013 et 2050 : par exemple, cela revient à associer aux territoires étudiants un maintien dans le temps du pic de population chez les 17 – 25 ans. Ainsi pour toutes les projections issues d'Omphale, le territoire projeté peut être regroupé avec le type de pyramide des âges 2013 dont il est le plus proche. Les évolutions de sa population par âge peuvent alors être comparées à celles des zones du même type. Dès lors, des éventuels problèmes de robustesse pourront être détectés dans les cas où ces évolutions s'avéreraient « extrêmes » au regard de la distribution des autres zones pourtant comparables en termes de structure en 2013. Plus concrètement on peut réaliser un premier test qui calcule les différences d'évolutions de population par tranche d'âge entre le territoire testé et la moyenne de la classe. Ces différences par âge sont ensuite sommées en valeur absolue. La même opération est effectuée pour l'ensemble des zones d'emploi de la classe. Il est donc alors possible de positionner le territoire dans la distribution de sa classe. Lorsqu'il se situe dans les 5 % des zones ayant des écarts les plus forts avec la moyenne de la classe, on peut considérer que ses évolutions sont extrêmes ou atypiques au regard de territoires ayant une pyramide des âges comparable en 2013. Un second test peut alors être réalisé selon le même principe mais par tranche d'âge afin de cibler laquelle ou lesquelles sont concernées au premier chef.

⁵ L'algorithme des k-means est un algorithme de classification des données permettant d'obtenir une partition de la base de donnée en k groupes. Contrairement à la CAH, l'algorithme des k-means nécessite de connaître *a priori* le nombre de classes souhaité.

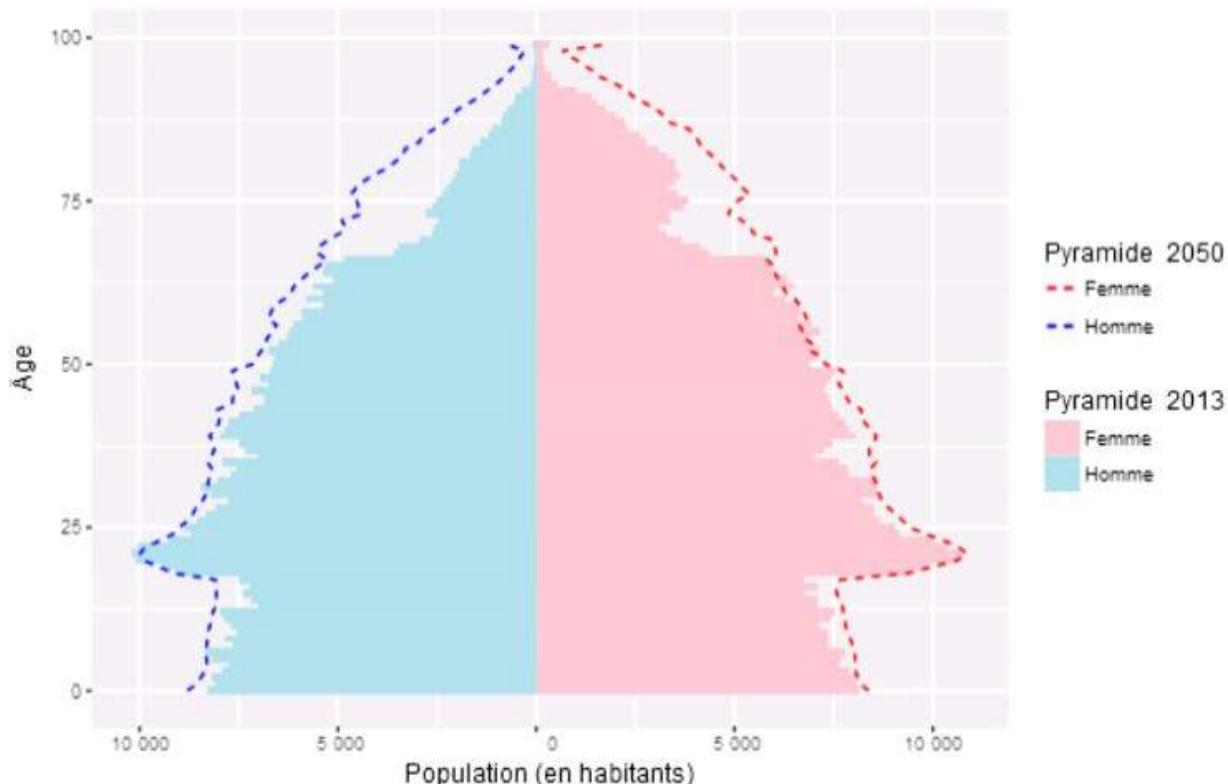
Figure X : Typologie des structures de population des zones d'emploi en 2013



Source : recensement 2013.

Prenons l'exemple de la métropole européenne de Lille (MEL). Les contours de cette métropole ne correspondant pas à ceux d'une zone d'emploi, la zone est proche en termes de structure par âge de la population de la classe « population étudiante ». En effet, la pyramide des âges 2013 présente un important pic entre 18 et 23 ans (figure XI). Le même constat semble, à l'observation des pyramides des âges, vérifié en 2050.

Figure XI : Pyramides des âges en 2013 et en 2050 pour la MEL

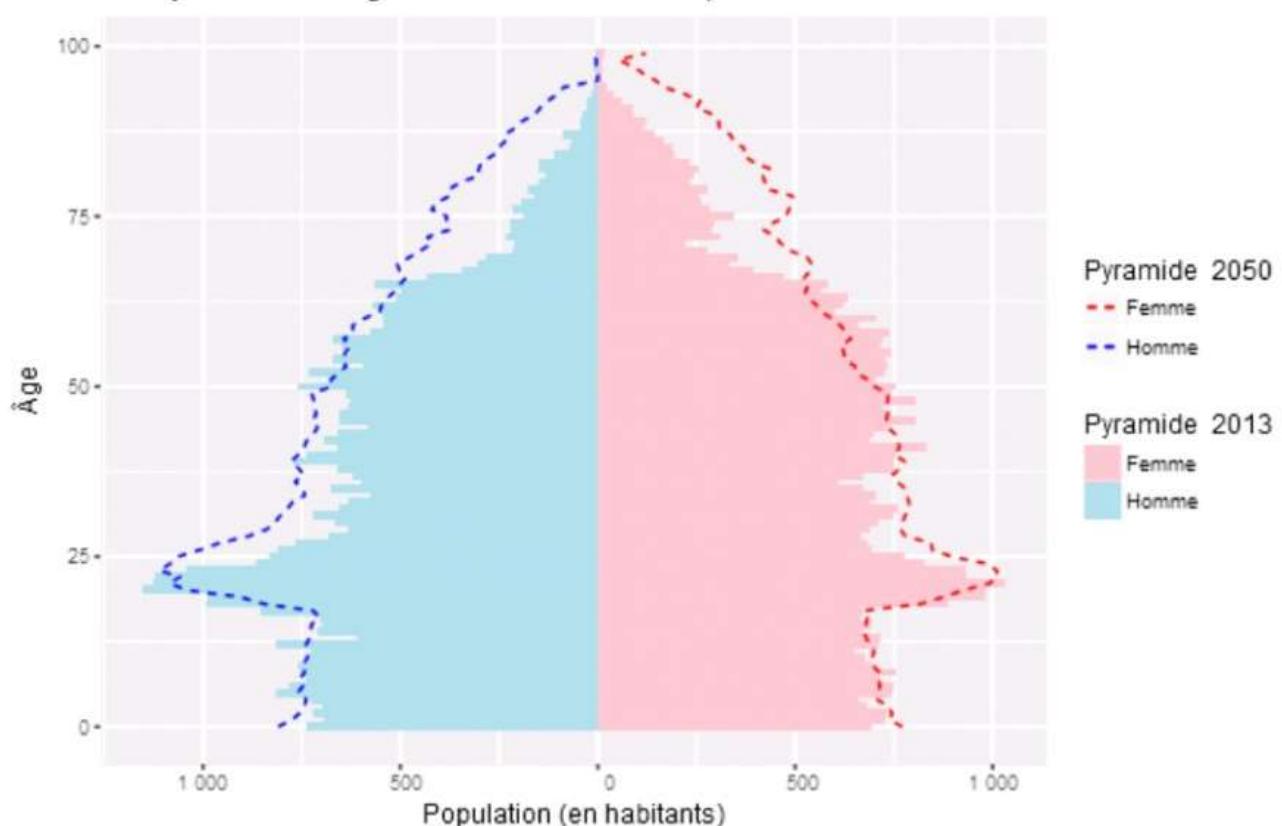


Source : Omphale 2017.

Si l'on compare les évolutions des populations par âge de la MEL à celle de l'ensemble des zones d'emploi étudiantes de France métropolitaine, il apparaît qu'elles sont dans la norme. Ce résultat statistique conforte l'intuition visuelle née de l'examen des pyramides des âges.

Un sous-territoire de la MEL appelé territoire de l'Est intègre la commune de Villeneuve d'Ascq où siègent les principales universités de la région des Hauts-de-France. Sans surprise, en reproduisant la mécanique précédente, il est également associé à la classe « population étudiante ». Pour autant, on observe à l'œil nu une déformation de la pyramide des âges puisqu'en 2050 le pic de population étudiante a été translaté de 2 à 3 ans vers le haut (figure XII), probablement la conséquence d'une non adéquation de l'hypothèse sur le solde migratoire avec l'étranger pour cette zone. Au final, le test comparant les évolutions de population par âge entre le territoire est et l'ensemble des zones d'emploi dites étudiantes permet bien de mettre en avant le caractère atypique de la zone étudiée qui ressort en « extrême » et donc de cibler un éventuel problème de robustesse de la projection

Figure XII : Pyramides des âges en 2013 et en 2050 pour le territoire de l'Est de la MEL



Source : Omphale 2017.

L'intérêt de cette méthode est double :

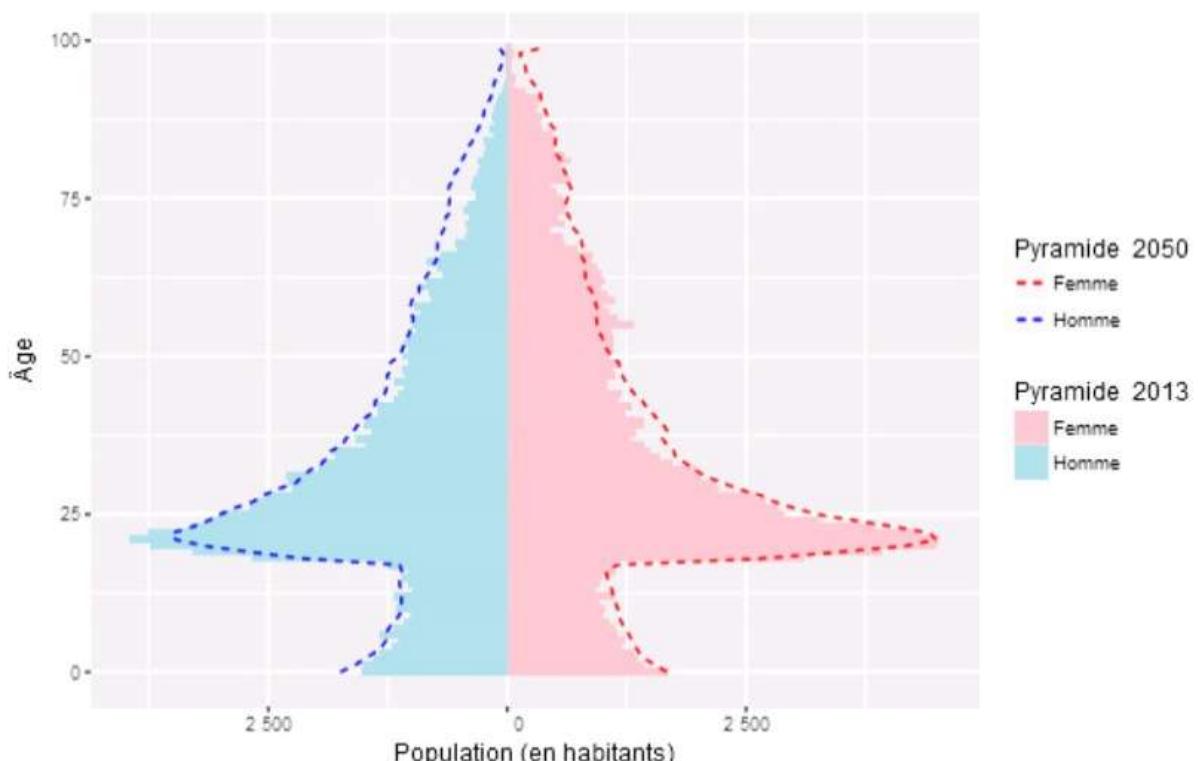
- elle ne présuppose pas les effets âge et les effets de générations. Elle postule que sur un grand nombre de zones du même type, ces effets sont respectés. Dès lors des déformations de la pyramide des âges différentes de la déformation typique peuvent permettre d'isoler des cas potentiellement problématiques ;
- en se basant sur l'observation de distributions, elle ne nécessite pas de déterminer a priori des seuils au-delà desquels des particularités sont à prendre en considération.

Une approche qui ne se substitue pas complètement à l'analyse empirique

Il faut garder en tête qu'en ciblant les individus « extrêmes » au sens de la distribution, la méthode conduit, par définition, à cibler pour chaque type de territoire des cas potentiellement problématiques, quand bien même toutes les projections seraient parfaitement robustes. De plus, des évolutions atypiques peuvent très bien se justifier par des facteurs non intégrés dans les critères de constitution des groupes de territoire. Être atypique ne signifie donc pas être automatiquement non robuste : il ne s'agit que d'une aide à l'analyse de la robustesse venant compléter une expertise plus qualitative des projections.

Ainsi, le territoire formé de la seule commune de Lille - présente une très forte sur-représentation des âges étudiants, y compris au sein même de la classe des zones d'emploi dites étudiantes. Si à l'œil nu, la pyramide des âges ne paraît pas se déformer de manière problématique, le test lui affecte néanmoins des évolutions extrêmes, probablement parce que cette zone est dès 2013 au sein de sa classe (figure XIII).

Figure XIII : Pyramide des âges en 2013 et en 2050 pour le territoire de Lille



Source : Omphale 2017.

Encadré : Les différentes utilisations d'Omphale

Omphale est un outil de projections de populations à destination principale des directions régionales de l'Insee. Il existe trois modes de restitution de ses résultats :

- des données sur zonages (départements et régions) et scénarios standards sont mis à disposition du grand public sur insee.fr (<https://www.insee.fr/fr/statistiques/2859843>) et ont été valorisées dans un Insee Première (voir Desrivierre, 2017) et une série d'études régionales ;
- des produits normalisés peuvent être commandés sur des zonages à façon et des scénarios standards (<https://www.insee.fr/fr/information/1303412>) ;
- des études locales en partenariat entre l'Insee et un acteur public en région sur des zonages à façon et si besoin des scénarios spécifiques (voir Mary-Portas, F & Bertrand, 2017).

Bibliographie

- Léon, O. & Desrivierre, D. (2011).** Le modèle de projection démographique Omphale 2010. Insee. *Insee Méthodes* n° 124.
- Brutel, C. (2014).** Estimer les flux d'entrées sur le territoire à partir des enquêtes annuelles de recensement. Insee. *Document de travail* n° F1403.
- Desrivierre, D. (2017).** D'ici 2050, la population augmenterait dans toutes les régions de métropole. Insee. *Insee Première* n° 1652.
- Mary-Portas, F & Bertrand, D. (2017).** Vieillissement et attractivité conditionnent l'avenir démographique des villes moyennes. Insee. *Insee Analyses Auvergne-Rhône-Alpes* n° 45.
- Fabre, J. & Larbi, K. (2018).** Évaluation de la robustesse d'une projection de population. Insee. *Actes des journées de la méthodologie statistique 2018*