

Perspectives de population et de ménages des communes belges.

Rapport II

Lissage gaussien et typologie des communes

Décembre 2010

Jean-Paul Sanderson, Luc Dal, Thierry Eggerickx et Michel Poulain
Centre de Recherche en Démographie et Sociétés (DEMO)-Université Catholique de Louvain

1. Introduction

Les perspectives locales à une échelle fine, comme celle des communes, sont un exercice rare (cf. Rapport I) et posent principalement le problème de variations aléatoires liées aux petits effectifs. Pour les 589 communes belges, seules 27 (5%) comptent au moins 50.000 habitants¹ en 2006 et 447 (76%) en compte moins de 20.000. La problématique des petits effectifs va donc concerner la majorité des communes belges, si on se réfère par exemple au seuil minimal imposé par le modèle Omphale, dans le cas de la France.

Tableau 1. Répartition des communes belges selon la taille de leur population en 2006

Taille des communes	Nombre de communes	
- de 2.000 habitants	6	1%
2.000-4.999 habitants	79	13%
5.000-9.999 habitants	163	28%
10.000-19.999 habitants	199	34%
20.000-49.999 habitants	115	20%
50.000 habitants	27	5%

L'objectif de ce rapport est de présenter les pistes envisagées pour répondre au problème des petits nombres, de décrire la méthode finalement préconisée en justifiant les choix posés. D'emblée, certaines solutions ont été écartées, tels que :

- appliquer aux communes des indicateurs calculés au niveau national, régional ou à l'échelle des arrondissements (méthode utilisée en Flandre). Cela revenait à nier les spécificités locales et à renoncer à exploiter toute l'information disponible ;
- appliquer aux communes des indicateurs calculés au niveau national en fonction de profils types d'individus (méthode utilisée en Suède). Cela revenait également à nier les spécificités locales ainsi que le poids de l'environnement (milieu d'habitat) sur les comportements.

La piste qui semble à priori la plus prometteuse réside dans la combinaison d'un lissage gaussien avec une typologie de communes permettant d'identifier des entités voisines ayant des profils similaires, afin de n'appliquer les lissages gaussiens qu'aux communes semblables.

2. Lissage gaussien

L'objectif est d'exploiter de manière optimale les données disponibles afin de conserver les spécificités locales, pour autant qu'elles soient représentatives. Dans le cadre de ce travail, c'est la conjonction de deux éléments qui entraîne une perte de signification statistique des probabilités calculées :

- d'une part, des événements, comme la fécondité et la mortalité, peuvent être qualifiés de « rares » pour certains âges, car leur probabilité d'occurrence est proche de 0 ;
- d'autre part, dans certaines communes, les effectifs soumis au risque sont également très faibles. Dès lors, le risque de sous(sur)-évaluation du phénomène est important.

¹ Le seuil de 50.000 habitants est cité ici car utilisé notamment en France comme seuil critique pour l'utilisation du modèle OMPHALE (cf. Rapport I.).

Prenons un exemple concret :

La fécondité des femmes de 15-19 ans est un événement rare. Pour la Belgique, le taux de fécondité de ce groupe d'âges (période 2001-2005) est de 24,5‰. Pour Namur, on observe 131 naissances de mères âgées de 15 à 19 ans révolus, soit un taux de fécondité de 39,3‰. Dans le cas de Vresse-sur-Semois, on observe pour le même groupe d'âges aucune naissance, soit un taux de 0,0‰. Toutefois, dans ce cas précis, la population soumise au risque (les femmes âgées de 15 à 19 ans) représente à peine un effectif de 66 personnes. On peut dans ce cas légitimement supposer que le nombre de naissances observées (0) est simplement dû à une insuffisance des effectifs, plutôt qu'à une particularité locale.

Tableau 2. Taux de fécondité des 15-19 ans et population soumise au risque, l'exemple des communes de Namur et de Vresse-sur-Semois

Entité	Naissances observées	Population soumise au risque (dénominateurs)	Taux observés
Vresse-sur-Semois	0	66	0,0‰
Namur	131	3332	39,3‰
Royaume	7535	302157	24,9‰

Afin d'identifier les communes dont la fécondité est significativement différente de celle du pays, pris dans son ensemble (avec un seuil de signification de 0,05), classiquement, on procède à un test de χ^2 en comparant les naissances observées à celles attendues sous l'hypothèse nulle d'identité des distributions du pays et de la commune considérée. On sait que le nombre minimal de naissances observées, par commune et par groupe d'âges, pour procéder valablement à un test doit être de 5. Dès lors, on peut estimer pour chaque taux de fécondité par groupe d'âges quinquennal un seuil minimal de population soumise au risque pour obtenir 5 naissances (si la commune avait le même niveau de fécondité que la Belgique).

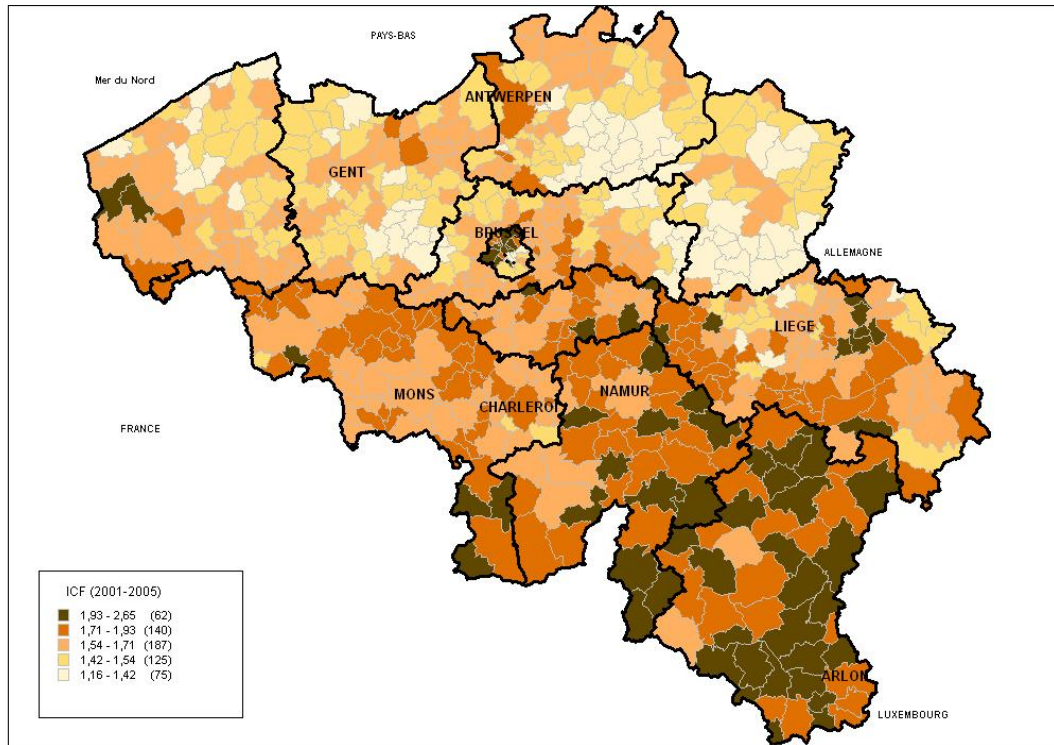
Tableau 3. Taux de fécondité quinquennaux pour l'ensemble du pays et identification des seuils minimaux de population soumise au risque

Période 2001-2005	15-19 ans	20-24 ans	25-29 ans	30-34 ans	35-39 ans	40-44 ans	45-49 ans
Taux de fécondité observés (pays)	24,5‰	236,6‰	626,4‰	537,5‰	188,4‰	32,4‰	1,5‰
Seuils minimaux (dénominateurs)	205	22	9	10	28	155	3263
Nombre de communes qui satisfont le critère	437	588	588	588	588	529	9

On observe que, pour les 15-19 ans, on pourra utiliser les taux de fécondité calculés à l'échelle communale pour 437 communes sur 589. Pour les 152 communes restantes, la population soumise au risque est insuffisante.

Par ailleurs, l'examen de la distribution spatiale de l'indice conjoncturel de fécondité, calculé pour la période 2001-2005, indique clairement l'existence d'une cohérence spatiale. On observe notamment dans le Limbourg des zones de sous-fécondité et dans les provinces de Namur et du Luxembourg, de larges zones de sur-fécondité (par rapport à la moyenne nationale) (figure 1).

Figure 1. Indice conjoncture de fécondité (2001-2005)



Z

Aussi, afin de lever l'obstacle statistique (atteindre le seuil minimal de population soumise au risque), il s'agirait de remplacer l'estimation discrète du phénomène (pour chaque commune) par une mesure continue (en chaque point de l'espace) du phénomène, celle-ci faisant appel à une fonction de lissage par potentiel gaussien qui inclura un paramètre de portée spatiale². En d'autres termes, il s'agira de remplacer une estimation ponctuelle par une estimation au moyen d'un noyau (« kernel density estimation »). Il s'agit d'une méthode classique en statistique : on n'estime plus une probabilité mais une densité de probabilité au moyen d'une convolution en dimension 2³ visant à éliminer le « bruit de fond » d'un signal et prenant en compte le voisinage des mesures ponctuelles. L'estimateur est le produit de convolution du signal observé et d'un « noyau », une fonction gaussienne dans le cas présent.

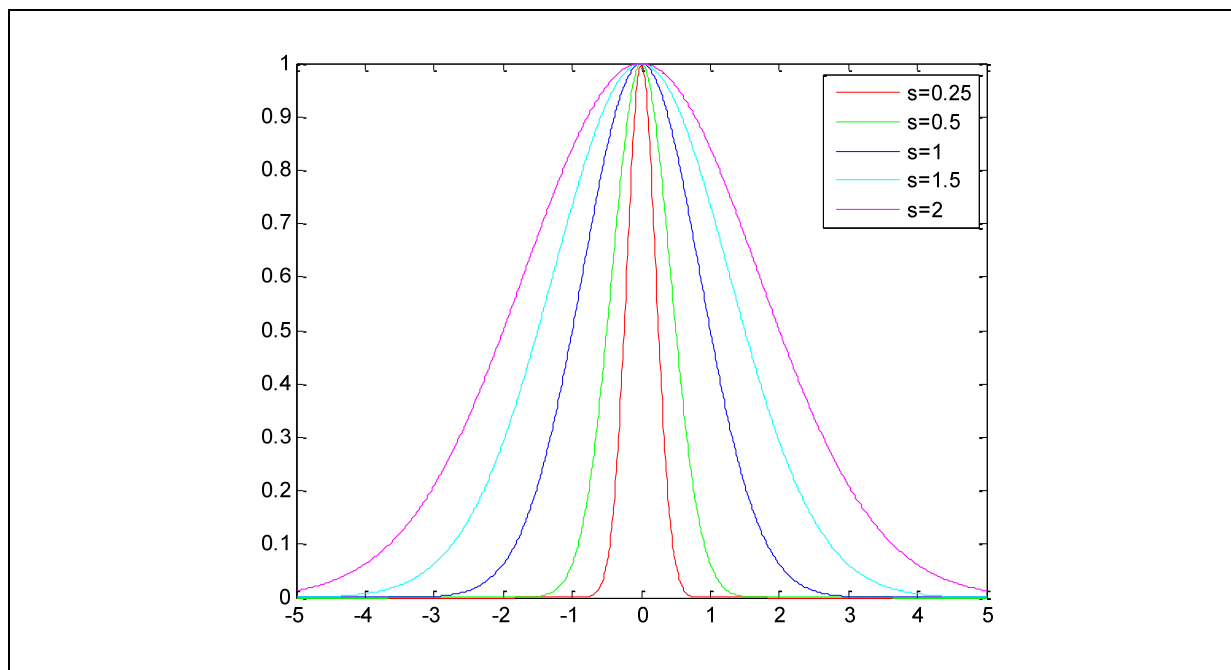
En chaque point « i », et pour un voisinage gaussien de portée δ on calculera les fonctions suivantes :

- Potentiel observé de femmes : $N^*(i, \delta) = \sum_j N_j e^{\frac{\ln(0.5)}{\delta^2} D_{i,j}^2}$
- Potentiel observé de naissances : $P^*(i, \delta) = \sum_j P_j e^{\frac{\ln(0.5)}{\delta^2} D_{i,j}^2}$
- Potentiel théorique de femmes : $A^*(i, \delta) = P(i, \delta) \frac{N^*_{tot}}{P^*_{tot}}$
- Proportion moyenne de naissances : $p(i, \delta) = \frac{N^*(i, \delta)}{P^*(i, \delta)}$
- Chi2 de déviation : $\chi^2(i, \delta) = \frac{(N^*(i, \delta) - A^*(i, \delta))^2}{A^*(i, \delta)}$

² La portée spatiale est la distance à partir de laquelle les contributions de chaque commune est de 50%. Elle peut-être interprétée comme étant l'écart-type de la fonction gaussienne.

³ La convolution de deux fonctions p et q en dimension 2 est donnée par $\iint_{R^2} p(x-u, y-v) q(u, v) du dv$

Figure 2. Fonctions non normalisées de lissage gaussien de portées variables



Pratiquement, pour chacune des 152 communes n'atteignant pas le seuil minimal de population soumise au risque, on complètera l'information en allant rechercher des données dans les communes voisines et en pondérant l'apport de celle-ci par un facteur exprimant la distance entre les deux communes. Cette intégration portera sur les communes les plus proches et s'étendra progressivement jusqu'à ce que l'on atteigne le seuil minimal.

Afin d'illustrer la méthode, reprenons le cas de Namur et de Vresse-sur-Semois pour le taux de fécondité des 15-19 ans, pour la période 2001-2005.

Tableau 4. Taux de fécondité des 15-19 ans et application du lissage gaussien, l'exemple des communes de Namur et de Vresse-sur-Semois

	Naissances observées	Dénominateurs observés	Taux observés	Naissances lissées	Dénominateurs lissés	Taux lissés
Vresse-sur-Semois	0	66	0‰	3.02	207.35	14,6‰
Namur	131	3332	39,3‰	131	3332	39,3‰

Dans le cas de Namur, le lissage n'est pas nécessaire, le nombre de femmes âgées de 15 à 19 ans étant suffisant ($3332 > 205$). Ce n'est pas le cas de Vresse-sur-Semois, car la population soumise au risque est largement inférieure à 205. On doit donc recourir au lissage gaussien.

Concrètement, pour chaque commune, en fonction de sa distance à Vresse-sur-Semois et d'un pas de lissage (variant par bonds de 0.25, cette valeur est déterminée de manière heuristique afin ne pas alourdir inutilement le processus de calcul), on calcule au moyen des formules données précédemment le nombre de femmes soumises au risque ainsi que le nombre de naissances

correspondantes. Ces quantités sont cumulées jusqu'au moment où le nombre de femmes est supérieur ou égal à 205.

On diminue ainsi l'importance de la commune de Vresse-sur-Semois en lui ajoutant de l'information régionale. Dans ce cas, il est nécessaire de prendre une portée de 9,50 afin d'obtenir le nombre requis de femmes.

Tableau 5. Itérations sur les pas (Vresse-sur-Semois)

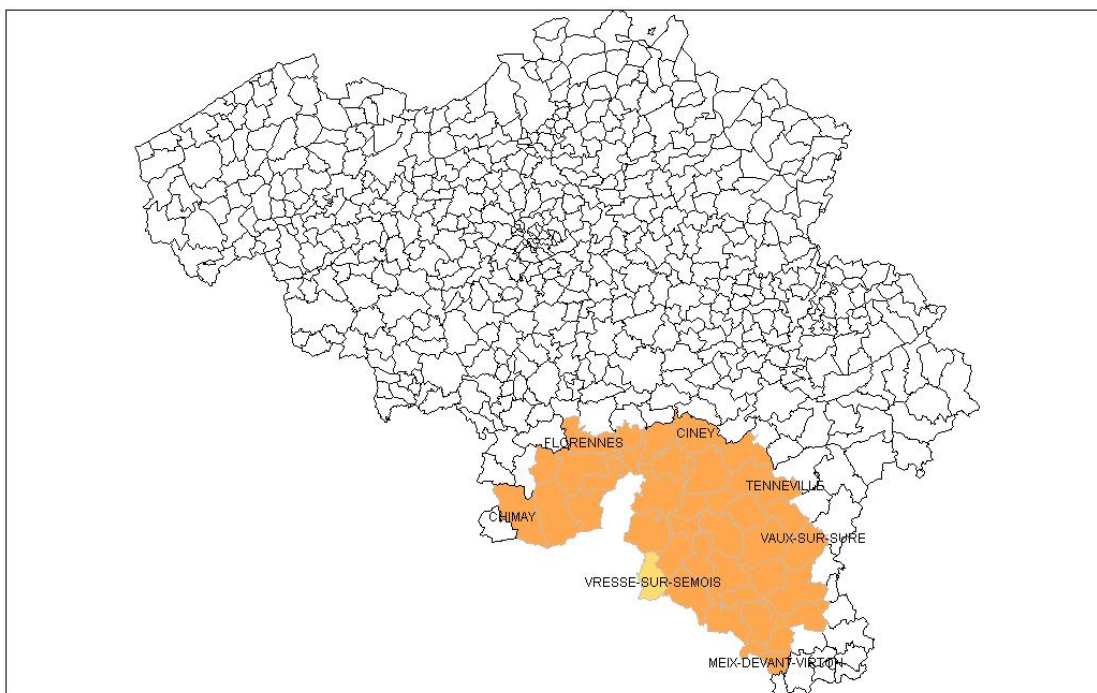
Portée	Femmes gaussiennes	Naissances gaussiennes	Taux gaussien
0,25	66,00	0,00	0,00%
0,50	66,00	0,00	0,00%
0,75	66,00	0,00	0,00%
1,00	66,00	0,00	0,00%
1,25	66,00	0,00	0,00%
1,50	66,00	0,00	0,00%
1,75	66,00	0,00	0,00%
2,00	66,00	0,00	0,00%
...			
8,50	163,36	1,98	1,21%
8,75	173,80	2,22	1,28%
9,00	184,62	2,48	1,34%
9,25	195,81	2,74	1,40%
9,50	207,35	3,02	1,46%
9,75	219,35	3,30	1,51%
10,00	231,49	3,60	1,56%

Tableau 6. Estimation du taux gaussien

	Distance	Femmes	Naissances	Taux	Poids	Femmes gaussiennes	Naissances gaussiennes
VRESSE-SUR-SEMOIS	0,00	66	0	0,00%	1	66	0
BIEVRE	10,93	102	0	0,00%	0,39964	40,5634019	0,0000000
BOUILLON	13,02	174	6	3,25%	0,27218	47,2236411	1,5329812
GEDINNE	14,95	156	5	3,52%	0,17950	27,9129861	0,9833738
PALISEUL	17,13	168	3	1,73%	0,10490	17,6228841	0,3049721
DAVERDISSE	20,26	44	0	0,00%	0,04269	1,8572117	0,0000000
BERTRIX	23,77	293	10	3,42%	0,01302	3,8154388	0,1302996
WELLIN	26,78	99	0	0,00%	0,00405	0,3993483	0,0000000
BEAURAING	27,14	294	13	4,27%	0,00349	1,0270107	0,0438571
LIBIN	27,18	163	3	1,74%	0,00344	0,5585729	0,0097419
HERBEUMONT	28,98	38	0	0,00%	0,00158	0,0593723	0,0000000
...							
CERFONTAINE	50,20	138	5	3,28%	3,92339E-09	5,39466E-07	1,76897E-08
MARCHE-EN-FAMENNE	50,36	628	24	3,80%	3,48333E-09	2,18579E-06	8,3108E-08
VAUX-SUR-SURE	50,51	139	0	0,00%	3,10005E-09	4,30908E-07	0
HABAY	52,10	222	3	1,16%	8,8238E-10	1,95447E-07	2,2726E-09
Total des autres communes						4,39139E-07	1,27698E-08
TOTAL						207,3514014	3,018178408

Les femmes habitant dans les communes précitées interviendront donc dans le calcul du taux lissé de la commune de Vresse-sur-Semois : il a donc été nécessaire de prendre en considération l'information (pondérée avec un poids décroissant) des communes dans un rayon de 52 km (figure 3) afin d'obtenir une taille de population suffisante.

Figure 3. Fonctions de lissage gaussien de portées variables



3. Typologies des communes

Si le recours au lissage gaussien est acquis, il semble néanmoins pertinent de ne pas regrouper des communes voisines sans s'être préalablement assuré qu'elles se ressemblaient en termes de comportements démographiques. En effet, certaines communes bien que voisines, peuvent avoir des profils sociodémographiques totalement différents. Ainsi, dans le Hainaut, Gerpennes se caractérise par une espérance de vie masculine de 74,5⁴ ans (période 2001-2005), tandis que Charleroi, qui se trouve à proximité, affiche à peine 71,5 ans (période 2001-2005). Si on devait recourir au lissage gaussien dans le cas de Gerpennes, il serait inopportun de lui associer Charleroi qui a un profil de mortalité totalement différent.

Dans un premier temps, la solution envisagée fut de réaliser des typologies des communes belges afin de ne considérer dans le lissage gaussien que des communes ayant des comportements et structures démographiques similaires. Au vu des choix méthodologiques posés, quatre phénomènes sont à envisager :

- La fécondité.
- La mortalité.
- Les migrations.
- Les transitions de ménages

Dans le contexte post-transitionnel actuel, les phénomènes démographiques sont relativement indépendants les uns des autres. Dès lors, l'option retenue était de réaliser une typologie pour chacun de ces phénomènes pris séparément. L'objectif de ce travail est simplement d'identifier les communes semblables de manière à ne pas prendre en compte, lors du lissage gaussien, des communes

⁴ Source : Cytise communes.

fondamentalement différentes ; la typologie n'était pas une fin en soi. Nos travaux ont dans un premier temps porté sur la fécondité.

3.1. Premiers essais de typologie sur la fécondité : longitudinal versus transversal

3.1.1. Typologie sur la fécondité : approche longitudinale⁵

A partir des données rétrospectives sur la fécondité extraites des recensements de 1981 et 2001, on a reconstitué la fécondité des générations 1902-1906 à 1962-1966. On a calculé pour chacun de ces groupes quinquennaux de génération trois indicateurs :

- L'âge à la première naissance.
- L'âge à la dernière naissance.
- La descendance finale.

Ce sont donc 36 indicateurs qui ont été considérés. Préalablement à la construction d'une typologie, une analyse en composante principale (ACP) a été réalisée de manière à ne procéder à la classification que sur des facteurs indépendants. En effet, un examen de la matrice de corrélations révèle, d'une part, des associations assez fortes entre générations voisines et, d'autre part, une corrélation très forte entre les descendances finales et l'âge à la dernière naissance des générations les plus anciennes. Le résultat de cette ACP est une réduction de l'information à 4 facteurs (dont la valeur propre dépasse 2 et qui expliquent 80 % de la variance) :

- Le premier facteur rassemble l'âge à la dernière maternité pour les générations 1902 à 1931 et la descendance finale de ces mêmes générations. Il s'agit d'un facteur « intensité ancien » puisqu'il se réfère aux générations les plus vieilles, des femmes ayant vécu l'essentiel de leur vie génésique avant 1970.
- Le second facteur rassemble l'âge à la dernière naissance pour les générations 1907 à 1916 et 1922 à 1951, ainsi que l'âge à la première naissance de ces mêmes générations. Il s'agit ici d'un facteur « calendrier ».
- Le troisième facteur reprend l'âge à la dernière naissance et la descendance finale de la génération 1937-1941 ainsi que la descendance finale des générations 1942 à 1961. Ce troisième facteur porte surtout l'empreinte de la génération 1937-1941, soit celle qui vit la révolution contraceptive.
- Le quatrième facteur reprend l'âge à la première naissance et à la dernière naissance des générations 1942-1956 ainsi que les trois indicateurs de la génération 1957-1961. Il reprend les principales caractéristiques des générations post baby-boom et s'oppose ainsi aux deux premiers facteurs plus marqués par la fécondité des générations anciennes.

Le résultat de cette ACP consacre surtout une opposition entre les comportements des générations anciennes et ceux des générations plus récentes. A partir de cette analyse factorielle, nous avons procédé à une analyse de classification de type hiérarchique (méthode de Ward), afin de constituer des groupes homogènes. Le résultat est la partition de l'espace belge en six groupes (figure 4).

- Le premier groupe est de type urbain-périurbain. Il se compose de grandes villes (Bruxelles, Bruges, Gand, Anvers et Liège) et des communes périurbaines qui les ceignent. Il se distingue des autres par une fécondité très faible tout au long du 20^e siècle. Très basse pour les générations du début du siècle, la descendance finale augmente progressivement mais ne dépasse guère le seuil de 2 enfants par femme pour les générations du baby-boom. En résumé, ce groupe « urbain-périurbain » se caractérise par une descendance faible, un calendrier tardif qui a subi un vieillissement rapide et un modèle familial qui privilégie les familles avec très peu d'enfants.
- Le second groupe « Flamande » se compose essentiellement de communes de Flandres occidentale et orientale, ainsi que des communes Wallonnes voisines du Tournaisis. Figurent

⁵ Cette partie fait l'objet d'un article soumis pour publication à la revue Espace, population et société (Costa, Eggerickx et Sanderson, 2010).

aussi dans ce groupe, des communes de la province de Liège (notamment le long de la frontière avec l'Allemagne). L'évolution de la fécondité de ce groupe se calque assez fidèlement à la moyenne nationale, tant pour les indicateurs d'intensité que de calendrier et de parité. Il s'agit donc d'un modèle « moyen ».

- Pour le groupe '**Campine ouest**', l'évolution de l'intensité et du calendrier de la fécondité des générations ressemble à celle du groupe précédent, mais de manière moins brutale. Néanmoins, pour les générations récentes, ce groupe présente des traits distinctifs forts. C'est pour ce groupe que la descendance finale est la plus faible et la durée de vie génésique la plus réduite (moins de 3,5 années), compte tenu d'un âge à la première maternité tardif et d'un âge à la dernière maternité précoce. Les familles nombreuses sont faiblement représentées, à l'inverse des familles à un et deux enfants. On relèvera également que la proportion de femmes sans enfant figure parmi les plus élevée.
- Pour le groupe '**Campine est**', la fécondité des premières générations du 20^e siècle était particulièrement élevée, proche de 4 enfants en moyenne par femme. Cette région se caractérise par une transition de la fécondité tardive, mais brutale lorsqu'elle s'amorce. Dès la génération 1942-1946, le niveau de fécondité a rejoint la moyenne nationale. D'une manière générale, c'est ce groupe qui a connu les évolutions les plus radicales au cours du 20^e siècle. Pour les générations les plus récentes, ce groupe se distingue par un âge moyen à la première maternité proche de la moyenne nationale et par un âge à la dernière maternité peu élevé. En d'autres termes, la période de procréation est très brève et largement concentrée avant 30 ans. Les femmes sans enfant et à enfant unique sont peu représentées, alors que les familles à deux enfants constituent de loin le modèle dominant.
- Le cinquième groupe peut être baptisé « Ardennes ». Il reprend surtout des communes de l'Ardenne, de la Lorraine belge, du Condroz ainsi que le Westhoek en Flandre. L'accroissement de la fécondité des générations du baby boom a été particulièrement important pour les communes de ce groupe. Entre les générations 1902-1906 et 1932-1941, la descendance finale a augmenté de 1,8 à 2,4 enfants par femme. Mais, contrairement aux autres groupes, la baisse de la fécondité qui caractérise les générations suivantes fut très lente, les descendes finales se maintenant au-dessus du seuil de remplacement. En ce qui concerne le calendrier de la fécondité, ce groupe a connu une évolution radicale au fil des générations.
- Le sixième groupe « Hainaut » comprend les communes qui constituaient les anciens bassins industriels de Charleroi, du Centre (près de La Louvière) et du Borinage (près de Mons) ainsi que des communes rurales de l'Entre-Sambre-et-Meuse. Les communes de ce groupe se distinguent par de très faibles descendes finales pour les premières générations du 20^e siècle. Néanmoins, l'intensité de la fécondité va rapidement s'accroître pour culminer à 2,4 enfants par femme pour la génération 1932-1936. Pour les générations suivantes, la fécondité a chuté tout en restant supérieure à la moyenne nationale. Mais la principale particularité de ce groupe de commune concerne la précocité de son calendrier de fécondité. Pour toutes les générations, l'âge moyen à la première maternité est le plus bas, avec en permanence, un écart important par rapport aux autres groupes.

Figure 4. La typologie communale de l'évolution de la fécondité des générations de femmes nées entre 1902 et 1961.

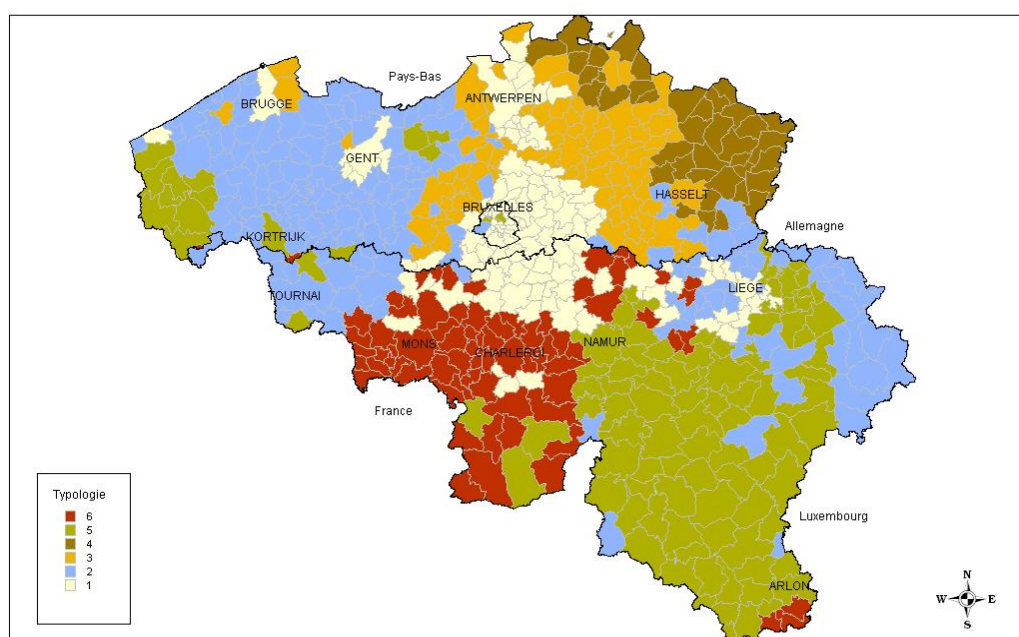


Tableau 4. Les principales caractéristiques des 6 groupes de communes

Nom	N	Population 1/01/2008	Descendance finale (enfant/femme)				Âge à la première naissance (année)				Âge à la dernière naissance (année)			
			1902-06	1937-41	1957-61	Test F	1902-06	1937-41	1957-61	Test F	1902-06	1937-41	1957-61	Test F
Urbain-périurb.	120	3.436.708	1,56	1,96	1,85	0,64	25,4	24,4	25,7	1,41	30,1	28,6	29	0,89
Flandre	172	2.657.790	2,09	2,16	1,87	0,64	26	24,4	24,6	2,36	32,2	29,1	29,8	1,50
Campine ouest	78	1.425.757	2,67	2,18	1,75	0,50	25,2	24,1	24,8	1,10	32,9	28,6	28,5	0,65
Campine est	40	700.735	4,01	2,6	1,95	0,60	25,8	24,2	24,6	2,11	35,7	29,6	28,2	0,69
Ardenne	104	1.269.475	1,79	2,38	2,13	0,95	26	24,2	24,5	1,73	31,6	29,5	28,5	1,06
Hainaut	74	1.176.317	1,65	2,34	1,91	0,57	24,2	23,1	24,2	1,14	28,8	28,4	29,5	0,68
Total	588	10.666.782	1,92	2,17	1,89	0,65	25,7	24,2	24,9	1,64	31,2	28,8	28,8	0,91

Cette typologie permet de mettre en évidence la disparité du territoire belge, mais aussi de vérifier l'adage « qui se ressemble, s'assemble » puisqu'on observe une assez grande cohésion spatiale de la fécondité selon cette approche longitudinale. Toutefois, ce premier exercice se base sur la longue tendance et on peut supposer que les changements récents dans les comportements de fécondité peuvent avoir modifié la géographie de la fécondité.

3.1.2. Typologie sur la fécondité : approche transversale⁶

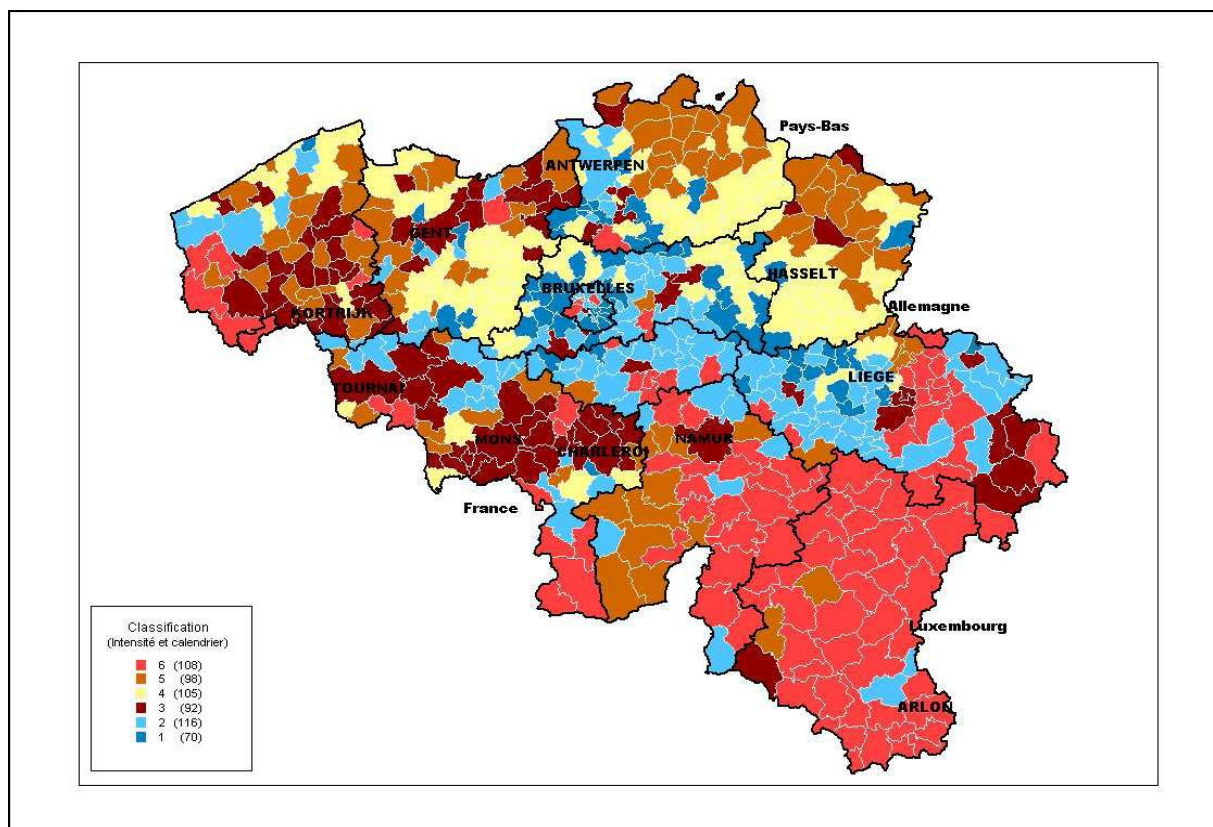
Pour cette deuxième typologie, on a utilisé les indices conjoncturels de fécondité observés de 1970 à 2005 et des indicateurs de calendriers de 1991-95 et 2001-05. Six types de communes ont été dégagés (figures 5). L'analyse de classification se base sur 6 indicateurs : 4 indices conjoncturels de

⁶ Cette partie fait l'objet d'un article présenté au colloque de la CUDEP de 2010 (Costa, Eggerickx et Sanderson, 2010).

fécondité pour les périodes quinquennales 1971-75, 1981-85, 1991-95 et 2001-05 et 2 indices de calendriers, la contribution des femmes de moins de 30 ans à la fécondité totale en 1991-95 et en 2001-05. Pour les périodes 1971-75 et 1981-85, l'indicateur conjoncturel de fécondité a été estimé sur base de l'indice standardisé de fécondité. La confrontation de ces deux indicateurs pour les périodes récentes a montré une corrélation quasi parfaite entre eux (pour 1991-95, $r = 0,99$ et pour 2001-05, $r = 0,98$).

- *Le groupe 1* est composé de 70 communes. La plupart d'entre elles sont localisées en Flandre, principalement dans la périphérie ouest de Bruxelles et au sud d'Anvers. Les communes situées au sud, sud-est de l'agglomération bruxelloise, les plus riches de la capitale, figure également dans ce groupe. Il s'agit d'un groupe de commune de « périurbanisation ancienne ». Il se caractérise par de faibles niveaux de fécondité tout au long de la période. Aujourd'hui, la fécondité est de 1,45 enfant par femme, alors que le calendrier est particulièrement tardif avec en moyenne 51% de la fécondité réalisée après 30 ans.
- *Le groupe 2* est composé de 116 communes, pour la plupart localisées en Wallonie, dans la partie centrale du pays. Ce groupe correspond au vaste espace périurbain qui, d'une part, encercle Bruxelles et qui, d'autre part, s'étire d'est en ouest, au sud de la Capitale. Ce groupe englobe également les grandes villes d'Anvers et de Liège, ainsi que leurs espaces périurbains. Il se caractérise par une fécondité assez faible en début de période, mais qui remonte dès les années 1990 pour atteindre un des niveaux les plus élevés en fin de période. En moyenne, ce groupe se distingue aussi par l'un des calendriers les plus tardifs, avec en moyenne 46 % de la fécondité réalisée après 30 ans.
- *Le groupe 3* est composé de 92 communes. Il concerne principalement des villes de grande taille et leur proche banlieue, telles que Gand et Charleroi, des villes de taille moyenne, comme Courtrai, Tournai, Mons ou encore Namur et les communes socio-économiquement déshéritées des anciens bassins industriels de Charleroi et du Borinage (Mons). La fécondité de la population de ce groupe assez élevée en début de période connaît une baisse brutale pour ensuite se stabiliser à un niveau proche de la moyenne nationale. Le calendrier de ce groupe est précoce ; 58,4 % de la fécondité étant réalisée avant l'âge de 30 ans.
- *Le Groupe 4* est composé de 105 communes, en très grande majorité flamandes. Un grand nombre se situe au nord-est du pays, en Campine, mais aussi plus à l'ouest dans l'ancien triangle linier coince entre Anvers, Gand et Bruxelles. Ce groupe se caractérise par une chute brutale de sa fécondité en début de période et il atteint aujourd'hui la valeur la plus faible, à peine supérieur à 1,4 enfant par 15 femmes. Le calendrier de la fécondité y est également précoce avec en moyenne 57 % de la fécondité réalisée avant l'âge de 30 ans.
- *Le groupe 5* est composé de 98 communes, la plupart localisée en Flandre. L'évolution de la fécondité de ce groupe est similaire au précédent. Partant d'un niveau élevé de 2,5 enfants par femme, c'est pour ces communes que la chute fut la plus brutale entre 1970 et 1980. Actuellement, elles se caractérisent par un niveau faible, inférieur à 1,6 enfant par femme et par le calendrier le plus précoce, puisque plus de 60 % de la fécondité est réalisée avant 30 ans.
- *Le groupe 6* est composé de 108 communes, en très grande majorité située en Wallonie et plus précisément au sud de l'axe Mons-Namur-Liège. Cette vaste région rurale – les Ardennes – se caractérise hier, comme aujourd'hui, par les niveaux les plus élevés de fécondité et par un calendrier précoce, la contribution des femmes de moins de 30 ans à la fécondité totale étant de 58 %.

Figure 5. La typologie des communes belges selon l'intensité et le calendrier de la fécondité



Ces deux typologies présentent une spatialisation des comportements assez différente. Elles montrent l'importance du choix des indicateurs, mais posent également la question de la méthode et de la nécessité d'inclure un facteur de contiguïté spatiale des groupes permettant d'éviter de mettre ensemble des communes comme Anvers ou Vresse-sur-Semois. Dès lors, deux démarches ont été menées parallèlement :

- la première consistait à tester différents indicateurs et à comparer les résultats obtenus ;
- la seconde visait à travailler en introduisant une contrainte de contiguïté spatiale.

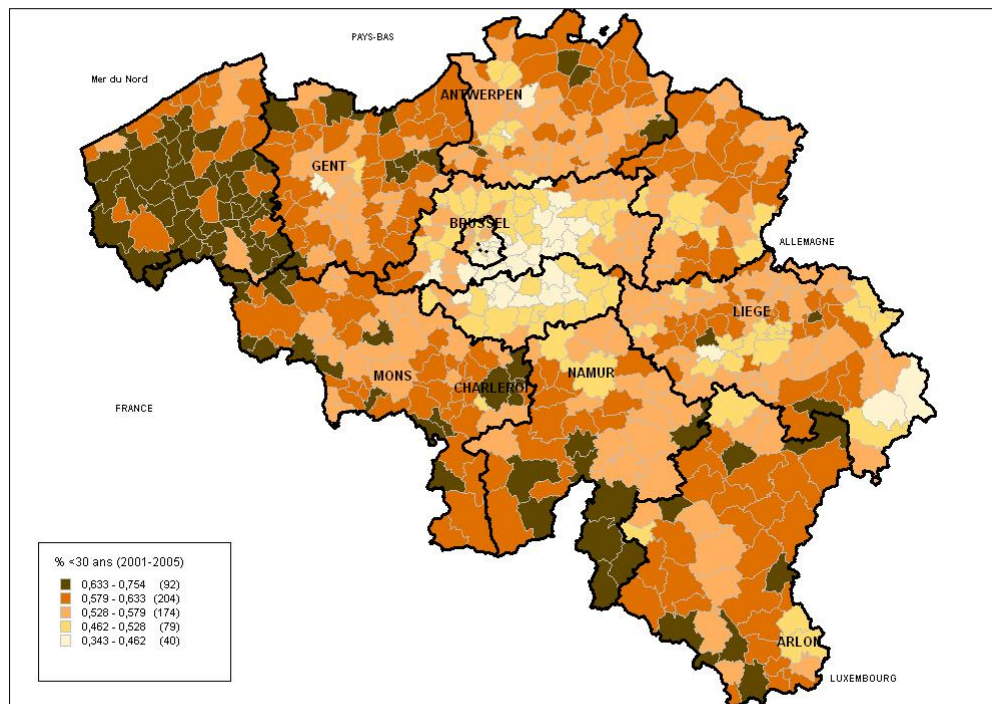
Les deux approches se voulaient complémentaires et devaient se rejoindre dans une troisième étape.

3.1.3. Typologie sur la fécondité : le choix des indicateurs

Les deux typologies précédentes ont mis en évidence la grande cohérence de la géographie de la fécondité. De même, elles soulignent l'importance du choix des indicateurs puisqu'elles aboutissent à des découpages spatiaux qui présentent quelques différences.

Si on examine la distribution spatiale de l'intensité de la fécondité (indice conjoncturel de fécondité (figure 1)) et celle du calendrier de la fécondité (part de la fécondité réalisée avant 30 ans (figure 6)), on constate que ces deux composantes de la fécondité présentent des configurations spécifiques. Très schématiquement, la première présente un clivage nord-sud et la seconde une opposition centre-périphérie.

Figure 6. Part de la fécondité réalisée avant 30 ans (201-2005)



Dès lors, il est nécessaire d'intégrer ces dimensions d'intensité et de calendrier dans l'élaboration de la typologie. C'est pourquoi, il a été décidé de construire la typologie directement à partir des taux de fécondité quinquennaux puisqu'ils traduisent immédiatement le calendrier et l'intensité.

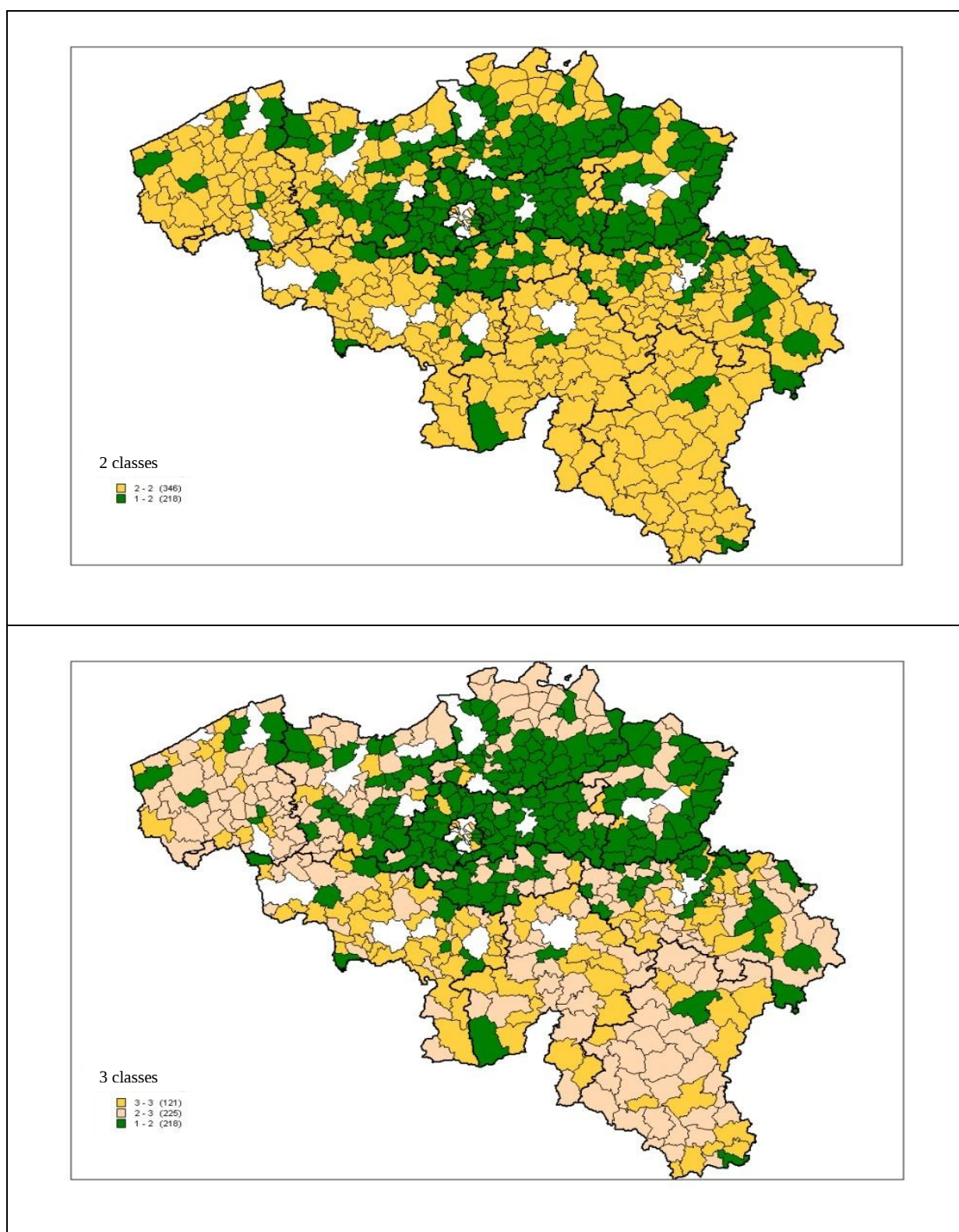
Par ailleurs, sachant que pour certaines communes, l'effectif des femmes en âge de procréer dans certains groupes d'âge est insuffisant pour produire des taux de fécondité valides et fiables, on a choisi de tester deux approches :

- la première consistait à travailler à partir des taux de fécondité calculés sur l'ensemble de la période d'observation (1991-2005) ;
- la seconde utilisait les taux de fécondité calculés pour chacune des périodes quinquennales (1991-1995, 1996-2000 et 2001-2005).

Dans les deux cas, les communes de plus de 50.000 habitants ont été exclues dans la mesure où, d'une part, ces communes n'ont pas besoin des communes voisines pour permettre le calcul des transitions et, d'autre part, leur maintien dans le processus d'élaboration de la typologie influencerait le résultat. Dans les deux cas, on a observé comment les groupes se formaient, la typologie optimale étant une typologie à 4 groupes.

Ces deux typologies présentent, comme les précédentes, une assez grande cohésion spatiale. Le résultat obtenu par l'utilisation des taux calculés par période quinquennale fournit une géographie plus homogène. Mais dans les deux cas, des communes, en ce compris des communes de petites tailles, s'isolent par rapport aux communes voisines. Cette situation est susceptible de générer des problèmes lors du recours au lissage gaussien. La solution pouvait venir de l'introduction d'une contrainte de contiguïté.

Figure 7. Typologie des communes belges élaborée à partir des taux de fécondité calculés sur 15 ans



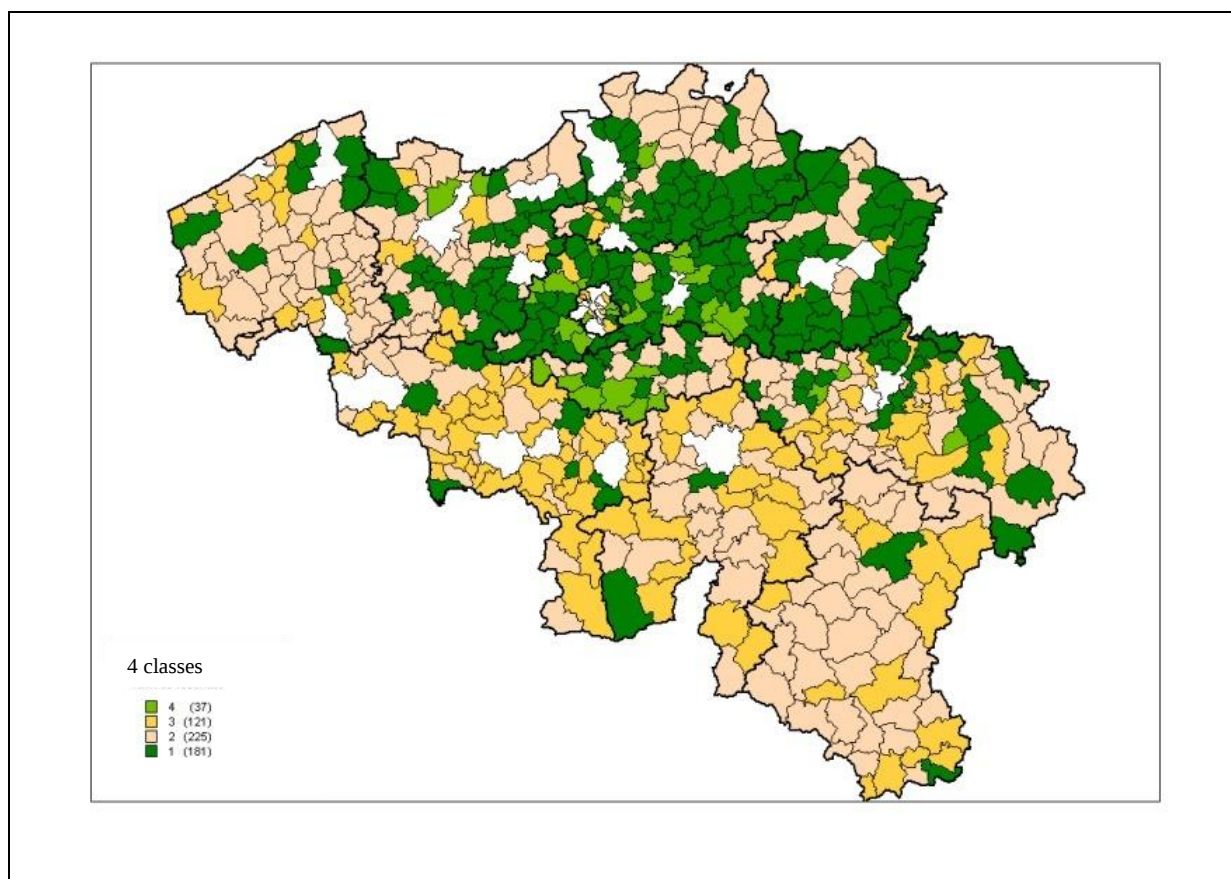
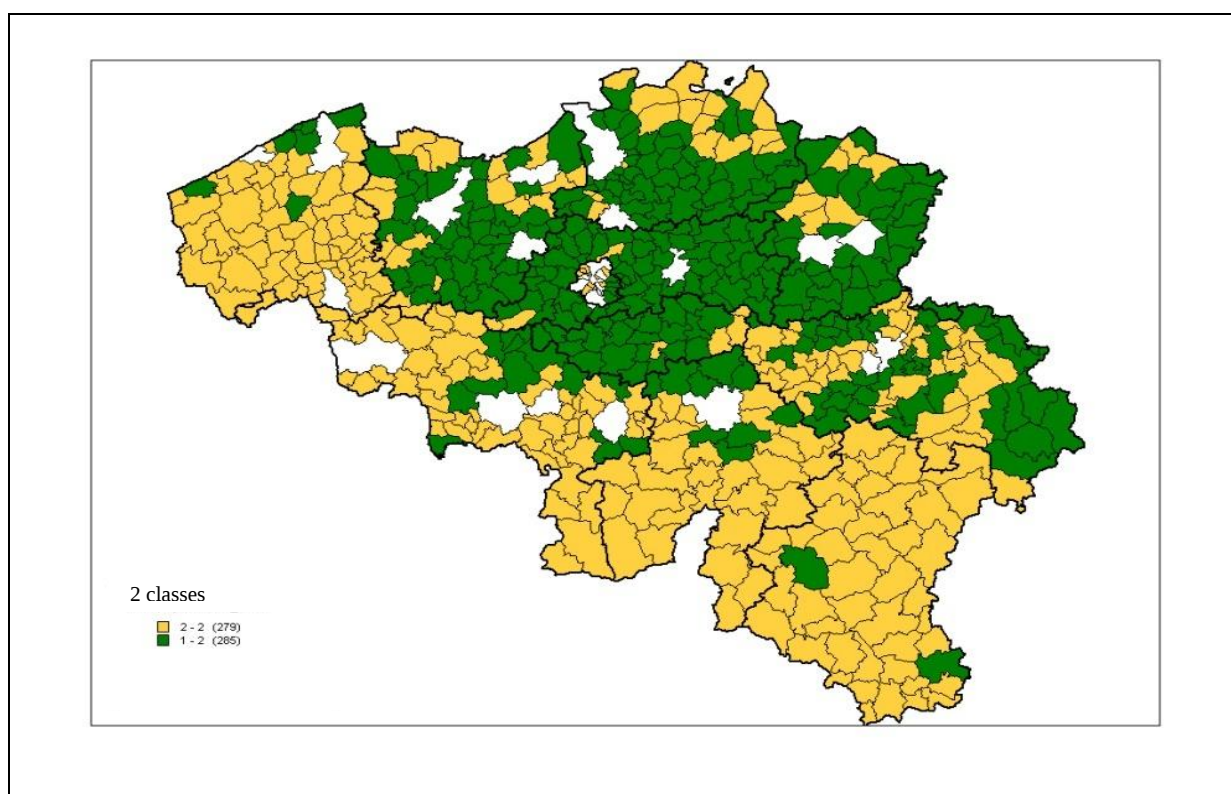
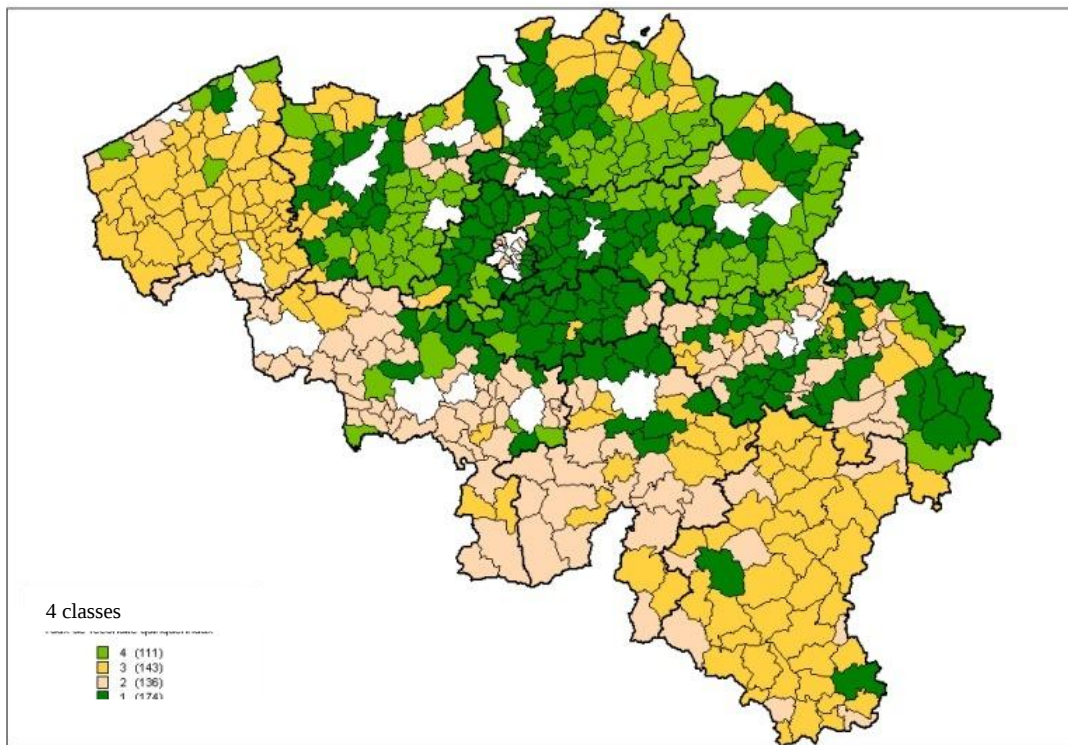
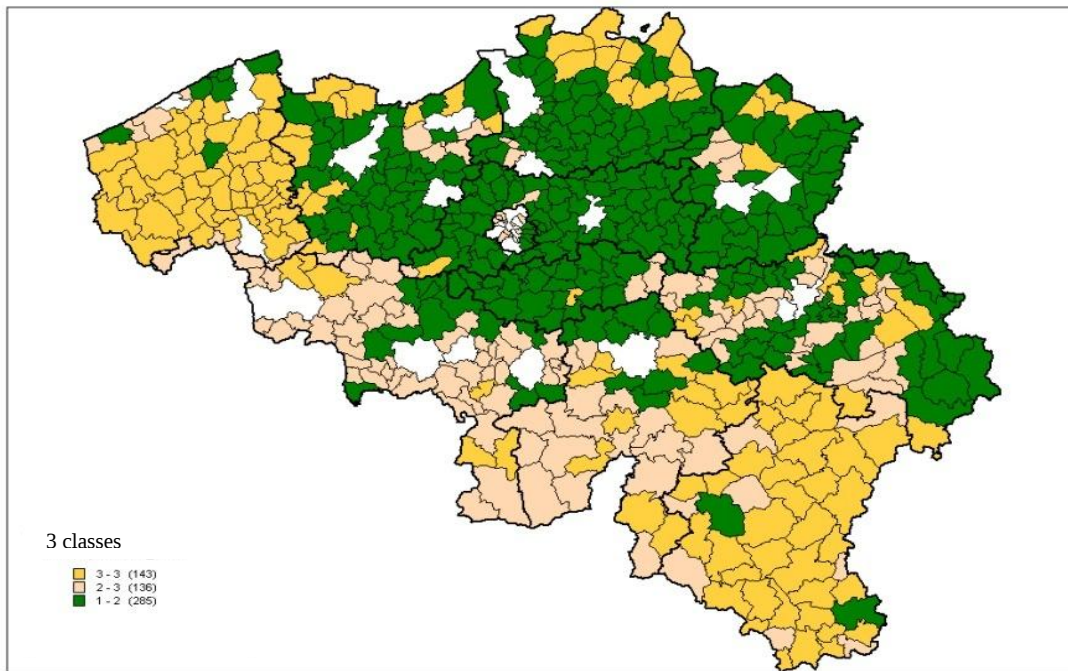


Figure 8. Typologie des communes belges élaborée à partir des taux de fécondité calculés par période quinquennale





3.1.4. Typologie sur la fécondité : avec contrainte de contiguïté spatiale

La méthode proposée vise à regrouper les communes de manière itérative en prenant en compte la contiguïté ou la proximité spatiale de ces communes. Il s'agit d'un processus qui regroupe deux par deux les communes jusqu'à l'obtention d'une régionalisation optimale au sens d'un critère. Ce dernier est la perte minimale de variance obtenue lors des agrégations successives : partant de la variance initiale, on cherche parmi tous les regroupements possibles celui qui minimise la diminution de variance résiduelle.

Cette méthode fonctionne parfaitement dans le cas de proportions simples : pour chaque commune on considère une proportion notée $t_i = \frac{N_i}{P_i}$, et pour l'ensemble du pays on note cette proportion $\bar{t} = \frac{\sum N_i}{\sum P_i}$. La variance totale des N_i est $V = \frac{\sum (t_i - \bar{t})^2}{\sum P_i} = \sum V_i$. Afin de déterminer le regroupement optimal au sens défini ci-dessus, on calcule pour chaque paire de zones i et j la quantité suivante : $M_{i,j} = V_i + V_j - V_{i \cup j}$ avec $V_i = \frac{(t_i - \bar{t})^2}{\sum P_i}$ et $V_{i \cup j} = \frac{(t_{i \cup j} - \bar{t})^2}{\sum P_i}$. La variance itérée = $V - M_{im,jm}$ où im et jm sont les indices optimaux pour M . Le processus est itéré jusqu'à l'obtention d'une seule zone.

L'introduction d'une fonction objectif permet de déterminer un nombre optimal de zones. Elle est définie par $W_i = \frac{V_i}{V * (n+i)}$ avec $V = V_0$ et est strictement décroissante vers 0.

Il est possible de contraindre le processus en introduisant un indice de contiguïté : l'agrégation ne pourra se faire que si les zones considérées sont contiguës, ce qui permettra de former des zones simplement connexes (« d'un seul tenant »). Plusieurs indices de contiguïté sont proposés : un indice élémentaire qui vaut 1 si les zones i et j sont contiguës, ou 0 dans le cas contraire, un indice plus

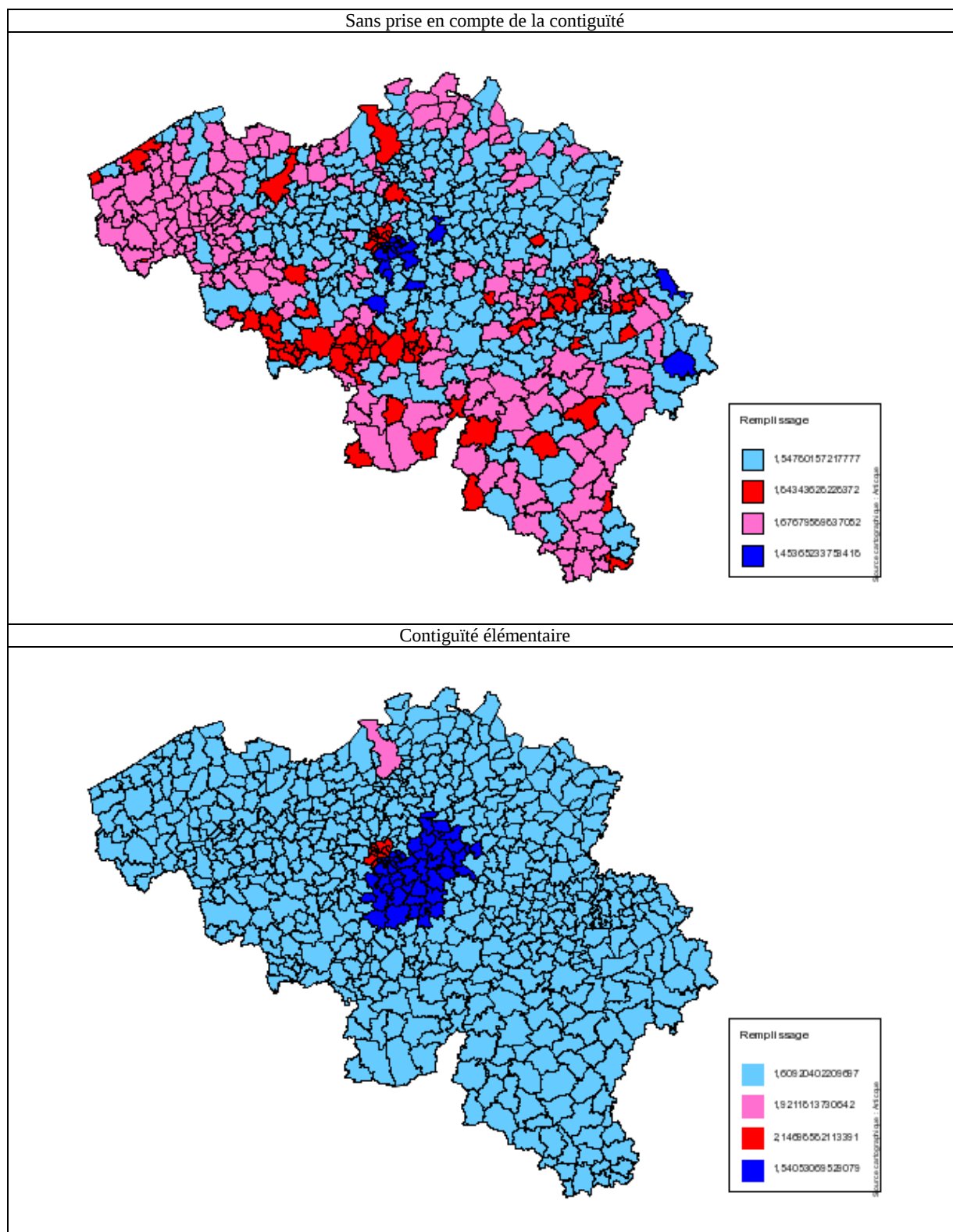
élaboré défini par $\delta_{ij} = \frac{F_{i,j}}{\max(P_i, P_j)}$ où $F_{i,j}$ est la longueur de la frontière commune des communes

i et j , P_i étant le périmètre de la commune i , ou enfin $\delta_{ij} = \frac{F_{i,j}}{\sqrt[4]{A_i * A_j}}$, A_i étant l'aire de la commune i .

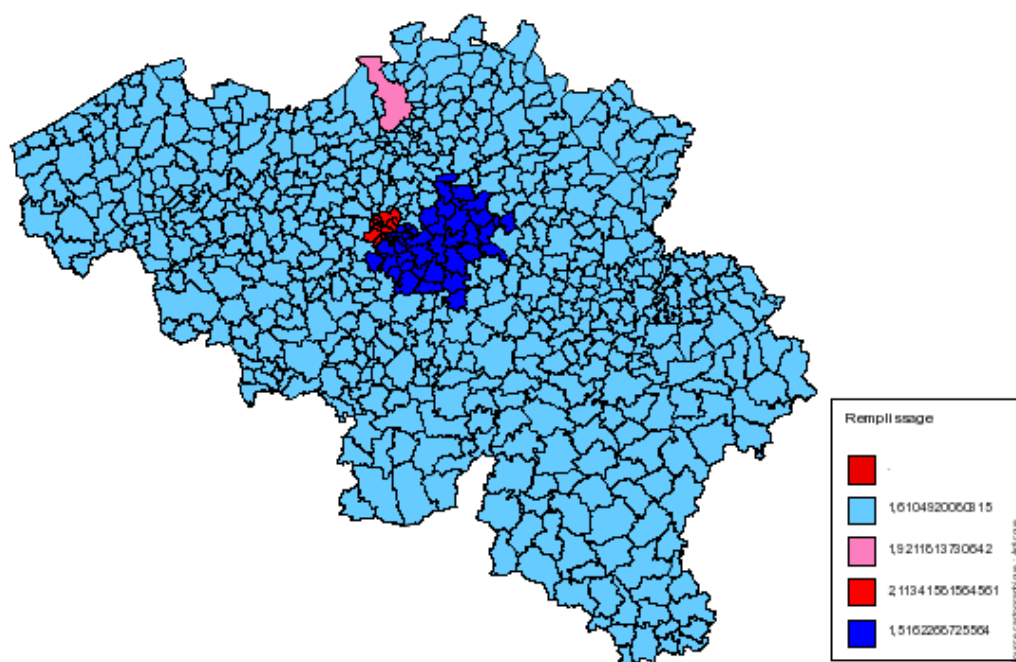
Le premier indice met toutes les communes sur pied d'égalité dans le processus, qu'elles soient contiguës par un point ou que l'une soit complètement enclavée dans l'autre, ce qui biaise la réalité de terrain. Le second indice permet de mieux prendre en compte la géométrie locale et il varie de manière continue de 0 (cas non contigu) à un maximum (0.43 à l'étape initiale dans le cas de Stavelot et Trois-Ponts). Le dernier indice permet également la prise en compte de la géométrie locale et varie continûment entre 0 et 3.56 (Bruxelles et Ixelles)

La méthode décrite ci-dessus est facilement applicable à l'indice synthétique de fécondité, ce dernier n'étant qu'une somme de taux quinquennaux ; à chaque étape, on recalcule facilement l'indice exact pour les communes i et j regroupées.

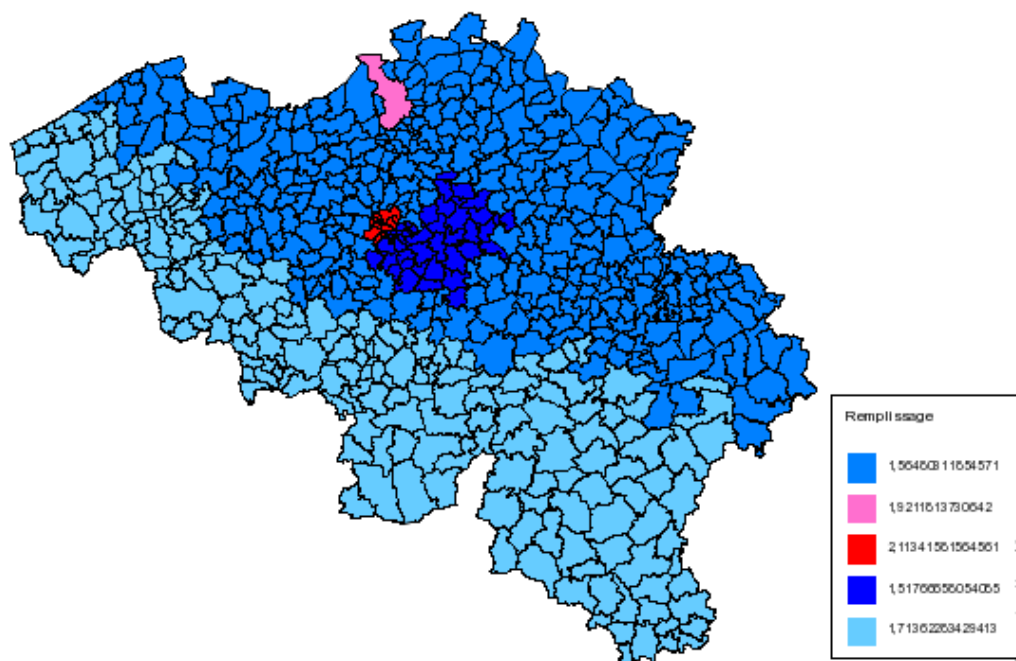
Figure 9. Typologie des communes belges élaborée à partir de l'indice synthétique de fécondité (2001-2005) avec contrainte de contiguïté



Contiguïté faisant intervenir le maximum des périmètres



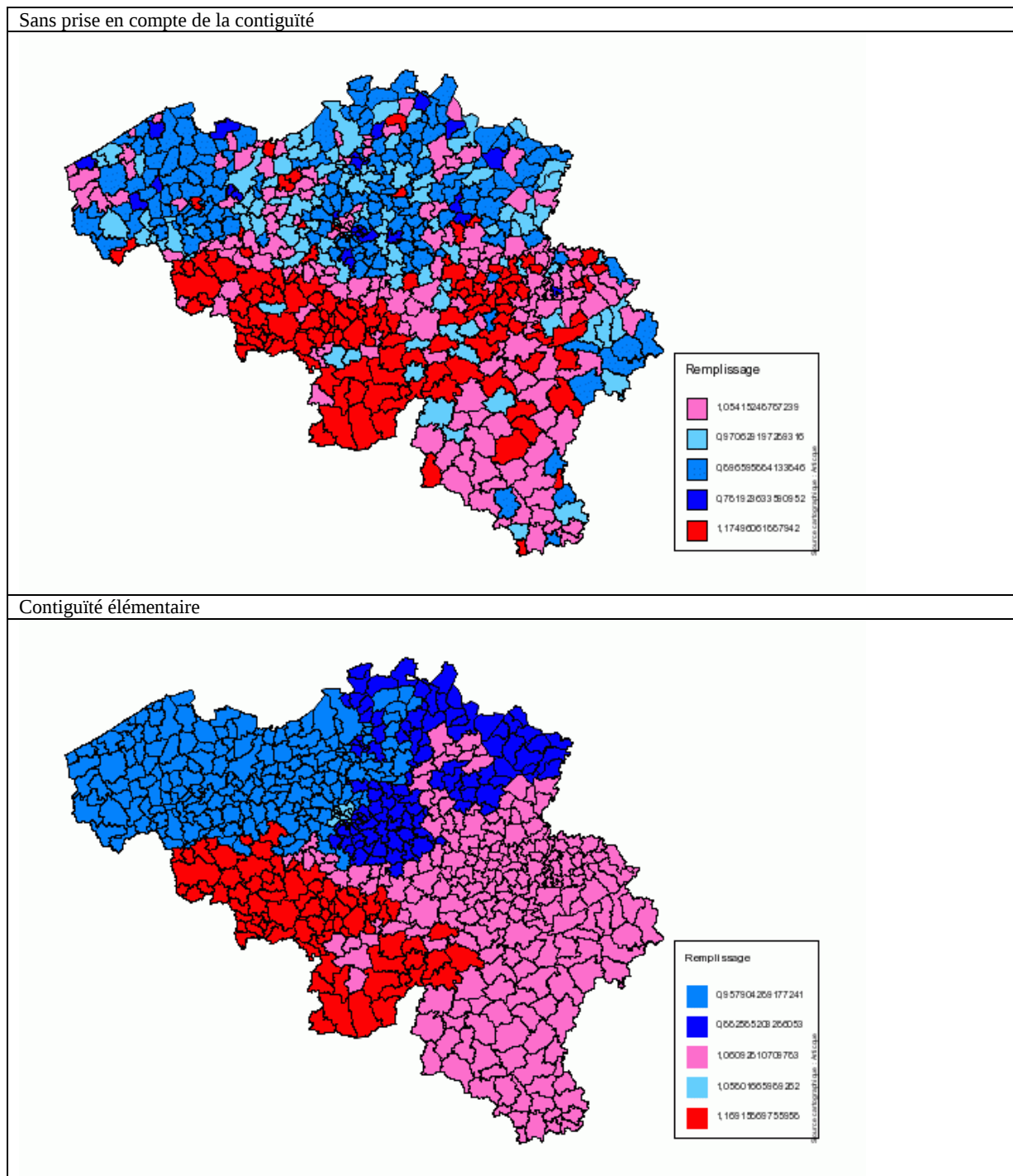
Contiguïté faisant intervenir la racine quatrième des aires



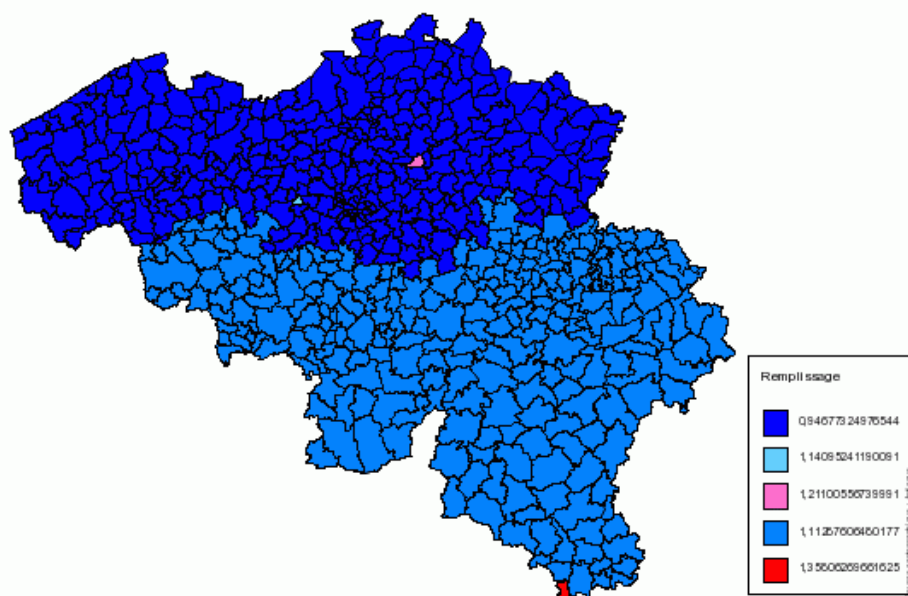
Dans le cas de la mortalité, l'indicateur « standard » est l'espérance de vie à la naissance, mais ce dernier ne peut se ramener à une fonction de proportions (nombre de décès divisés par une population soumise au risque) et il est non additif. L'indice retenu par la suite est l'indice synthétique de mortalité, qui est le rapport de la somme des décès communaux observés divisée par la somme des décès communaux attendus, sous la distribution standard observée au niveau national. Il a l'avantage d'être fortement corrélé à l'espérance de vie et est aisé à calculer.

Au vu des cartes des figures 9 et 10, il est malaisé de dresser une typologie des communes répondant aux objectifs fixés : pouvoir produire des perspectives pour des « petites communes » dans lesquelles on observe des événements rares. En effet, l'introduction de la contrainte de contiguïté aboutit à une fusion quasi complète des communes belges, niant ainsi les disparités régionales mises en évidence précédemment. Compte tenu des conclusions de cette étape, il faut se tourner vers une approche différente qui est celle du lissage gaussien.

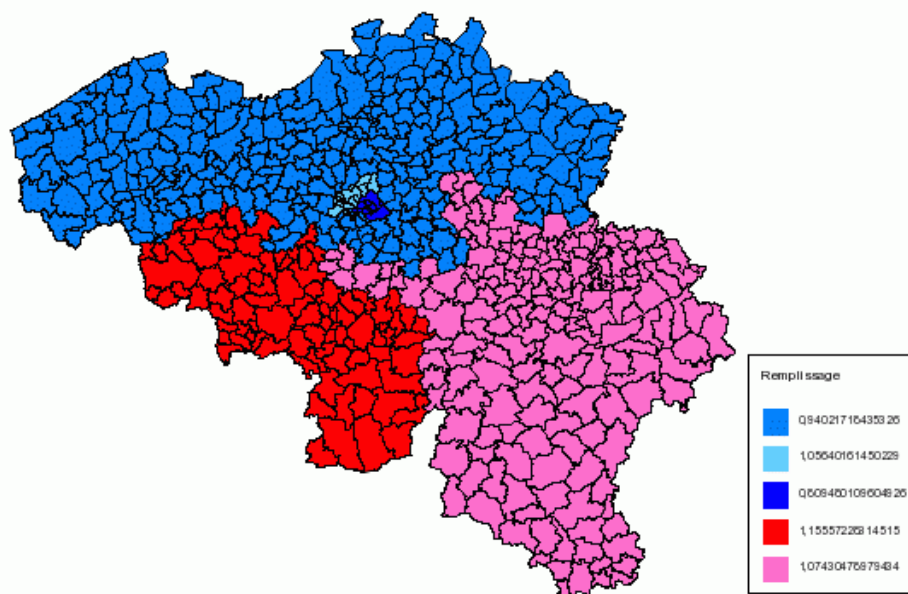
Figure 10. Typologie des communes belges élaborée à partir de l'indice synthétique de mortalité (2001-2005)



Contiguïté faisant intervenir le maximum des périmètres



Contiguïté faisant intervenir la racine quatrième des aires



4. En guise de conclusion : Méthode des communes semblables et introduction d'un nouveau paramètre

Les typologies livrent des résultats intéressants, mais ne permettent pas de résoudre l'entière du problème. En effet, quelle que soit la typologie utilisée, certaines communes demeurent isolées, car parmi les communes environnantes, on n'en trouve aucune appartenant au même groupe. Ces typologies révèlent néanmoins une grande cohésion des espaces belges confortant l'idée « qui se ressemble s'assemble » et donc l'existence d'un « déterminisme spatial » des comportements démographiques.

Dès lors, afin de palier les limites des typologies tout en intégrant l'existence de disparités locales, la procédure de lissage gaussien sera appliquée à chaque fois qu'il en sera nécessaire. On procédera pas à pas, sans exclusion d'aucune commune même dissemblable, mais la distance considérée entre deux communes sera une mesure combinant à la fois la distance spatiale et la distance socio-économique et culturelle entre les communes concernées. Ce faisant, le poids qui sera affecté à une commune proche mais socio-économiquement et culturellement différente sera faible, au même titre que le poids attribué à une commune lointaine semblable. Pour la composante socio-économique et culturelle de la distance, des tests seront effectués dans le courant du premier semestre 2011 afin de déterminer la mesure la plus adéquate.

5. Bibliographie

- Costa R., Eggerickx T., Sanderson J.-P. (2010), « Permanence et changement des territoires de la fécondité en Belgique. Approches longitudinale et communale », *Espace population et société* (soumis pour publication)
- Costa R., Eggerickx T., Rizzi E., Sanderson J.-P. (2010), « Analyse spatiale et temporelle de la fécondité en Belgique: une approche communale », Actes du colloque de la CUDEP, *La fécondité : représentation, causalité, prospective*, Strasbourg, 2010 (à paraître).