

CHAPITRE 5 : LE CALAGE DES PROJECTIONS

5.1. Un modèle non additif

Si l'on fait la projection de la population d'une région selon un scénario déterminé, et la projection selon le même scénario des départements de celle-ci, la somme des populations départementales ne donnera pas la population régionale. Ceci résulte évidemment de la nature non linéaire de plusieurs équations du modèle.

Cette propriété n'est pas indispensable dans l'utilisation courante des projections car les écarts sont faibles, si les hypothèses sont les mêmes. Toutefois, il peut être nécessaire de présenter des résultats strictement additifs selon les usages qui en seront faits. C'est la fonction du calage de projections de restituer cette propriété *a posteriori* aux résultats.

Par extension, cette procédure peut rendre additives des projections issues de scénarios différents, c'est ce qui se passe pour les projections standards de l'Insee. Les projections de la France métropolitaine sont élaborées de manière totalement externe à Omphale. Les projections des régions peuvent être réalisées en suivant au plus près mais pas exactement le scénario métropolitain. Ces projections seront calées sur la projection nationale. Ce calage a un effet plus important qu'un calage pour des scénarios identiques (correction des populations de l'ordre de 2 % pour 2030). Cette correction se répercute sur les projections infra-régionales par l'intermédiaire du calage sur les projections régionales elles-mêmes calées.

5.2 Le calage des données

Les données rendues additives sont les populations, les décès, les naissances et les migrations nettes pour chaque année projetées par sexe et par âge détaillé. Les trois premières données sont calées par une procédure de calage sur marges puis arrondies à l'unité, les migrations nettes sont ensuite recalculées par différence.

La procédure de calage sur marge assure l'égalité entre la somme des effectifs des zones z par sexe et âge et les effectifs de la zone dite « englobante » (Z).

Soit E la matrice des effectifs des zones par sexe et âge et TZ le vecteur des effectifs de la zone « englobante » par sexe et âge.

1. Tout d'abord, on cale les effectifs totaux des zones :

$$TE_z = \sum_i E_{z,i} \cdot \sum_i TZ_i / \sum_i \sum_z E_{z,i}$$

2. On cale ensuite les effectifs par sexe et âge de chaque zone sur le nouveau total des zones :

$$E'_{z,i} = E_{z,i} \cdot TE_z / \sum_i E_{z,i}$$

3. On cale ces nouveaux effectifs par sexe et âge des zones sur le total des zones :

$$E''_{z,i} = E'_{z,i} \cdot TF_i / \sum_z E'_{z,i}$$

La procédure est itérée au point 2, en remplaçant E par E'' , jusqu'à la convergence.

Dans une seconde étape, les valeurs de E sont arrondies à l'unité par défaut et les résidus sont répartis sur les cases ayant les plus gros écarts entre la valeur calée et la valeur arrondie.