

DERRIERE LES DISPARITES REGIONALES DU P.I.B. PAR HABITANT DANS L'UNION EUROPEENNE

ANALYSE STATISTIQUE

Laure CHARLEUX : Laboratoire SEIGAD, Institut de Géographie Alpine, Université Joseph Fourier, Grenoble.

Email : charleux@lamasig.ujf-grenoble.fr

RÉSUMÉ. *L'article présente une analyse statistique spatiale des disparités régionales du P.I.B. par habitant, indicateur phare de la politique régionale de l'Union Européenne. Elle se développe classiquement en trois temps : analyse univariée, modélisation par régression linéaire, puis classification en fonction des variables sélectionnées par la régression. En guise de conclusion, une évaluation de l'éligibilité à l'Objectif 1 est effectuée à partir des mêmes variables.*

ABSTRACT. *This article presents a statistical and spatial analysis of the regional disparities in the GDP/inh., which is the primary index for the regional policy of the European Union. There are three phases to be developed : the first involves univariate analysis, the second, linear regression modeling and the third, clustering based on the variables selected through the regression. To conclude, the same variables will be used to evaluate eligibility for Objective 1.*

MOTS-CLÉS : *analyse de données, disparités régionales, Objectif 1 des Fonds structurels, P.I.B. par habitant, Union européenne.*

KEY WORDS: *data mining, European Union, GDP/inh., Objective 1 of Structural Funds, regional disparities.*

Même si les pays membres de l'Union européenne comptent parmi les plus prospères de la planète, il subsiste entre eux des différences de niveau de développement et l'hétérogénéité est encore plus grande lorsque l'on considère l'échelle régionale. Cette disparité a été jugée nuisible, au-delà d'un certain seuil, à la « cohésion » de l'Union, envisagée comme résultante de la similarité¹. Une politique régionale a donc été progressivement mise en place depuis 1975, qui vise principalement à aider les régions les moins développées à rattraper leur "retard". Le budget alloué à cette politique *via* les « fonds structurels » européens représente actuellement plus d'un tiers du budget total de l'Union (juste derrière la PAC qui en représente la moitié), dont 70% sont réservés aux interventions en faveur de ces régions en retard de développement (dans le cadre de l'Objectif 1 des fonds structurels). C'est assez dire l'enjeu qui préside à leur désignation.

Depuis la réforme des fonds structurels de 1988, qui marque la véritable naissance d'une politique régionale coordonnée et ambitieuse, les régions en retard de développement sont définies comme les NUTS 2² dont le P.I.B. par habitant en S.P.A.³ est inférieur à 75% de la moyenne de l'Union⁴. Cet indicateur a l'avantage d'être simple et synthétique. Il a pourtant l'inconvénient de masquer des différences socio-économiques fondamentales du point de vue du développement. Par exemple, un faible P.I.B. par habitant peut caractériser une région où beaucoup d'habitants travaillent avec une faible productivité ou, au contraire, une région à forte productivité mais dont beaucoup d'habitants travaillent dans la région voisine. La première est clairement dans une situation plus préoccupante que la seconde à moyen et long termes. Pourtant toutes deux sont égales du point de vue des fonds structurels.

Il nous a donc paru important, en première approche de la politique régionale de l'Union européenne, de pousser plus loin l'analyse des disparités régionales du P.I.B. par habitant. Nous avons pour cela employé des méthodes classiques d'analyse statistique et spatiale sur des données socio-économiques issues de la base REGIO éditée par Eurostat.

1 Autour des données

1.1 Récriminations d'usage et sélection des données Eurostat

Eurostat a pour rôle de rassembler les données produites par les offices statistiques des Etats membres. De cette organisation, les données Eurostat tirent la réputation de ne pas être bien comparables pour des pays différents. En fait, d'importants progrès ont été faits ces dernières années dans le domaine de l'harmonisation, que ce soit en aval – redressement des données –, ou en amont – harmonisation des méthodes de production des offices nationaux. S'il convient donc de demeurer attentif, les données Eurostat sont largement utilisables et nous sommes d'autant plus fondés à les utiliser qu'elles sont à la base des critères d'éligibilité aux aides de la politique régionale.

Cependant, ces progrès sont récents, tout comme la demande de statistiques régionales au niveau européen, qui date de la volonté d'affermir une politique régionale dans les années 80. Il est donc difficile de rassembler, au niveau régional, des séries cohérentes et à peu près complètes avant 1988. A partir de cette date, par contre, la plupart des données sont disponibles annuellement dans la base REGIO. Mais il ne nous a pas semblé utile de travailler sur un pas de temps annuel et nous avons choisi un intervalle de trois ans pour la constitution de notre base d'étude (années 1988, 91, 94 et 97, auxquelles nous espérons pouvoir ajouter 2000 d'ici 2002).

Du point de vue thématique, REGIO est bien renseignée pour les indicateurs démographiques et socio-économiques classiques, mais de nombreuses lacunes apparaissent lorsqu'on considère des variables sur les infrastructures ou la production des régions, sur la santé ou l'éducation des populations régionales. Pratiquement, ces tables sont inutilisables pour des comparaisons inter-régionales à l'échelle européenne. Les indicateurs à partir desquels

¹ En soi, cette conception de la cohésion prête à discussion. Le problème a notamment été brillamment posé par Laurent Davezies (DAVEZIES, 1997). Mais ce n'est pas notre propos ici, où nous adoptons les hypothèses fondamentales de la politique régionale européenne.

² La NUTS est la Nomenclature des Unités Territoriales Statistiques de l'Union Européenne, c'est-à-dire le maillage utilisé par Eurostat. Son niveau 2 correspond par exemple aux régions françaises et aux provinces autonomes espagnoles.

³ Standard de Pouvoir d'Achat, unité de mesure utilisée par Eurostat pour gommer les différences de coût de la vie entre les Etats membres. Cf. 2.1.

⁴ Pour la période de programmation 2000-2006, ont été adjointes de droit, sans référence au P.I.B. par habitant, les « zones très peu peuplées » (6 régions scandinaves) et les « zones ultra-périphériques » (DOM, Canaries, Açores et Madère).

nous pouvons mener nos analyses se restreignent donc à la population, la densité, les parts des personnes âgées de moins de 25 ou de plus de 65 ans, le taux d'activité par sexe, le taux de chômage par sexe, le nombre de personnes occupant un emploi par sexe et selon qu'il s'agit d'un emploi à temps plein ou partiel, la répartition des emplois en trois secteurs d'activité et, enfin, le P.I.B. en euros ou en S.P.A..

1.2 *Composer avec un maillage instable*

Les « régions » auxquelles se réfère la politique régionale européenne sont les unités correspondant au deuxième niveau de la Nomenclature des Unités Territoriales Statistiques de l'Union Européenne (NUTS). Ce maillage, toutefois, n'est pas européen à proprement parler. Il résulte de la juxtaposition des maillages nationaux des Etats membres, sur lesquels l'Union n'a pas à se prononcer. En soi, c'est un véritable problème, car la taille des unités d'un même niveau NUTS est très différente d'un Etat à l'autre, et l'on peut soupçonner cette hétérogénéité d'avoir une influence sur la variabilité des indicateurs compilés à travers un tel maillage⁵. Ce n'est toutefois pas le point que nous souhaitons aborder ici, puisque nous voulons mener nos analyses sur les régions telles que considérées par la politique régionale européenne.

Mais la nature disparate du maillage NUTS a une autre conséquence : il est instable. Chaque Etat membre est libre de modifier son maillage administratif à sa convenance et Eurostat doit reporter ces changements dans la NUTS : avec 15 membres, ils sont relativement fréquents. Souhaitant évaluer des évolutions temporelles, nous avons préféré utiliser le même maillage pour toutes les dates de notre étude, même si cela se traduit par des valeurs manquantes aux dates les plus anciennes pour certains pays (notamment pour le Royaume-Uni qui vient de refondre quasi-complètement sa NUTS 2). Ce maillage est fondé sur la NUTS 2 de 1999, que nous avons légèrement altérée pour mieux "coller" aux données disponibles : nous avons fusionné les deux NUTS 2⁶ d'Irlande (la division est une nouveauté de la NUTS 1999) ainsi que les trois NUTS 2 de Saxe (la division apparaissait dans la nomenclature de 1989, disparaissait dans celle de 1995 et apparaît à nouveau dans celle de 1999).

2 Le P.I.B. par habitant dans les régions de l'Union européenne

2.1 *Les défauts connus d'un indicateur simple*

Le P.I.B. par habitant est l'indicateur phare de la politique régionale européenne. Il a été choisi pour sa simplicité et sa robustesse. Mais, outre les difficultés de fond qu'on peut trouver à l'utiliser pour mesurer le niveau de développement d'une région ou de bien-être de ses habitants, sa mesure elle-même pose quelques problèmes connus. Parmi eux, la surestimation de l'indicateur pour les grandes régions urbaines, notamment du fait des navettes en provenance des régions voisines, ou encore l'occultation des productions et activités non-monétarisées.

Puisqu'il faut néanmoins l'utiliser, reste à choisir l'unité de mesure dans laquelle travailler. Dans la base REGIO, le P.I.B. par habitant est disponible en euros par habitant (à prix courant) et en S.P.A. par habitant, et c'est sur cette dernière mesure qu'est fondée l'éligibilité des régions à l'Objectif 1. Il s'agit d'un standard qui efface les disparités du coût de la vie entre les Etats membres. Certes, calculé sur une base nationale, il n'élimine pas les disparités interrégionales. Néanmoins, beaucoup d'éléments influençant le coût de la vie, les services publics ou les systèmes de taxation par exemple, ont effectivement une base nationale. C'est pourquoi, présumant que les variations du coût de la vie étaient beaucoup plus importantes entre les Etats qu'entre les régions d'un même Etat, nous n'avons pas vu d'inconvénient majeur à utiliser cette mesure.

2.2 *Les caractéristiques statistiques de la distribution et leur évolution*

Les distributions du P.I.B. par habitant au niveau NUTS 2 aux quatre dates de référence suivent toutes approximativement une loi normale, ce qui autorise à les soumettre aux modélisations statistiques courantes. De ce fait également, leur moyenne et leur médiane sont approximativement identiques comme le montre le tableau 1.

⁵ Voir à ce sujet les travaux du groupe de travail Hypercarte, notamment l'article de C. GRASLAND, H. MATHIAN ET J.-M. VINCENT, « Multiscalar analysis and map generalisation of discrete social phenomena : statistical problems and political consequences », *Statistical Journal of UNECE*, 17, 2000, pp. 157-188.

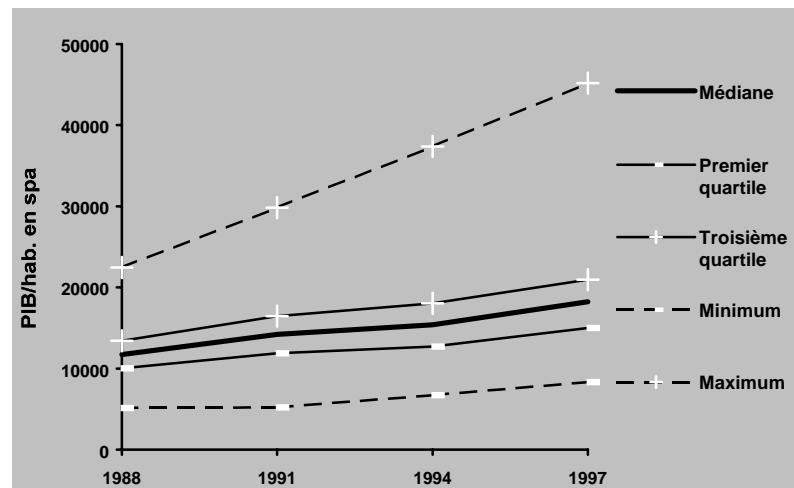
⁶ *Stricto sensu*, le terme « NUTS » renvoie à la nomenclature complète (« la NUTS 2 »). Mais un usage métonymique généralisé l'a étendu aux unités territoriales définies par la nomenclature (« les NUTS 2 européennes » = les régions européennes).

Tableau 1 : Indicateurs statistiques univariés

	1988	1991	1994	1997
Nombre d'individus pris en compte ⁷	184	192	208	204
Minimum	5117	5192	6698	8342
Maximum	22471	29819	37384	45202
Amplitude	17354	24627	30686	36860
Médiane	11715	14225	15393	18236
Premier quartile	10005	11870	12726	14986
Troisième quartile	13412	16460	18030	20964
Coefficient interquartile	29.1%	32.3%	34.5%	32.8%
Moyenne	11813	14262	15737	18546
Ecart-type	3089	4173	4541	5105
Coefficient de variation	26.1%	29.3%	28.9%	27.5%

La moyenne et la médiane augmentent de 1988 à 1997, même si nos chiffres ne tiennent pas compte de l'inflation globale dans l'Union. Notons que la croissance qu'on constate entre 1994 et 1997 n'est que très légèrement exagérée par le changement de composition de l'Union en 1995.

Mais à l'intérieur de l'Union, toutes les régions ne bénéficient pas de la même croissance, comme en témoignent les indicateurs de dispersion du tableau. De significativité certes relative, l'amplitude montre que l'écart entre la région la plus pauvre de l'Union et la région la plus riche n'a cessé de croître, comme on peut l'apprécier visuellement sur la figure 1. Plus significatif, le coefficient interquartile croît jusqu'en 1994, puis ne diminue que légèrement de 1994 à 1997. Le coefficient de variation, quant à lui, augmente considérablement entre 1988 et 1991, largement parce que la réunification allemande a adjoint à la distribution 7 nouvelles NUTS 2, qui ont alors les 7 plus faibles P.I.B. par habitant d'Europe. Il ne diminue ensuite que faiblement de 1991 à 1994 puis un peu plus franchement de 1994 à 1997, notamment grâce à l'élargissement, sans lequel le rythme resterait constant.

Figure 1

En première approche, et d'un point de vue encore aspatial, on ne peut donc pas dire que la mise en œuvre de la politique régionale réformée ait coïncidé avec une nette convergence du niveau de P.I.B. par habitant des régions européennes.

2.3 Les caractéristiques spatiales de la distribution et leur évolution

Paradoxalement, la politique régionale de l'Union Européenne se soucie relativement peu de la dimension spatiale, au moins au premier abord. Si son but avoué est de réduire les disparités de développement interrégionales au sein de l'Union, aucun objectif n'est exprimé quant à la distribution dans l'espace communautaire des disparités qui subsistent forcément. Or, à variabilité statistique égale et si l'on considère que la cohésion est engendrée par la similarité, quelques gradients en pente douce semblent préférables à la juxtaposition de régions très dissemblables.

⁷ L'année 1994 est la seule pour laquelle nous disposons du P.I.B. par habitant pour toutes les NUTS 2 de la nomenclature 1999. Pour 1997, il manque les données sur les DOM français. Pour 1991, il manque les DOM, deux NUTS finlandaises, deux NUTS suédoises et huit NUTS britanniques. Pour 1988, il manque en outre les huit Länder d'Allemagne de l'Est.

Une simple cartographie thématique choroplèthe permet d'apprécier visuellement quelques grandes caractéristiques de la distribution spatiale du P.I.B. par habitant pour une année donnée, mais les limites de classes peuvent être trompeuses localement. Nous avons donc quantifié les écarts relatifs entre régions contiguës. Une telle démarche⁸ a des limites : la portée des discontinuités ainsi mises en valeur est faible – l'augmenter en calculant des écarts relatifs sur des contiguïtés de niveau 2 serait compliqué – et elle est variable dans l'espace, dans la mesure où la taille des NUTS 2 varie elle-même fortement. Mais les cartes ainsi produites permettent néanmoins de dégager quelques faits.

La figure 2 reproduit ces cartes pour 1988 et 1997. Il s'en dégage une impression de plus forte concentration des discontinuités importantes en Allemagne et au Royaume-Uni. Sans doute faut-il surtout y voir un effet du maillage : ces deux pays ont les NUTS 2 les plus petites en superficie et des variations locales qui seraient moyennées ailleurs peuvent mieux s'exprimer ici. Au-delà, les deux cartes mettent en **lumière un fort effet « capitale »** : les régions dont font partie les capitales d'Etat ou les villes les plus importantes d'Europe ont toutes un P.I.B. par habitant très supérieur à celui de leurs voisines et apparaissent entourées de rouge. Cela traduit certes une réelle concentration de richesse, mais il faut se souvenir que le P.I.B. par habitant a tendance à être surestimé pour les grandes régions urbaines.

Visuellement, il est difficile de juger de l'évolution globale de 1988 à 1997 : quelques discontinuités s'estompent, d'autres se renforcent. Mais le calcul d'un coefficient d'auto-corrélation (rapport de la variance des NUTS 2 contiguës sur la variance totale), à composition constante⁹, montre que l'auto-corrélation diminue sur la période et que cette diminution s'accélère (cf. tableau 2). Même s'il est dangereux d'affirmer une tendance sur un temps aussi court, **tout se passe donc comme si la cohésion spatiale entre les régions d'Europe diminuait.**

Tableau 2

Coefficient d'auto-corrélation de Geary	1988	1991	1994	1997
<i>Sur toutes les données disponibles à chaque date</i>	0.40	0.43	0.42	0.45
<i>Sur les mêmes NUTS 2 pour toutes les dates</i>	0.40	0.42	0.45	0.49

Cependant, un même P.I.B. par habitant peut recouvrir des situations économiques forts différentes et il convient de pousser plus loin les investigations. C'est pourquoi nous proposons, en deux temps, d'identifier les principales composantes du P.I.B. par habitant parmi les variables dont nous disposons, puis d'établir des typologies régionales en fonction de ces composantes.

3 Identification des composantes du P.I.B. par habitant

Nous nous proposons d'établir, pour chaque date, un modèle de régression linéaire multiple sur le P.I.B. par habitant avec pour candidats-régresseurs les séries précédemment extraites de la base REGIO plus quelques variables issues de leur combinaison arithmétique. Mais parmi elles, à l'examen des matrices des corrélations, deux ressortent. Il s'agit du [P.I.B./nombre de personnes occupées], que nous appellerons « productivité », qui est très fortement corrélée avec le P.I.B. par habitant et du [nombre de personnes occupées/population moyenne], que nous appellerons « participation », qui est fortement corrélée avec la plupart des autres variables. Nous avons donc décidé de traiter ces deux variables en « méta-variables » et de les examiner en profondeur avant de passer à la modélisation du P.I.B. par habitant lui-même.

3.1 Productivité et participation

En effet, la mesure du P.I.B. par habitant peut-être décomposée en productivité, entendue comme le P.I.B. par personne occupant un emploi, et en participation, entendue comme le nombre de personnes occupant un emploi

dans la population totale¹⁰ :
$$\frac{\text{P.I.B.}}{\text{habitants}} = \frac{\text{P.I.B.}}{\text{occupés}} \times \frac{\text{occupés}}{\text{habitants}} .$$

⁸ Cette démarche a déjà été utilisée pour le même sujet par Claude Grasland pour la contribution de l'équipe P.A.R.I.S au « Study Programme on European Spatial Planning » (cf. rapport final, figures 14 et 15). Mais il a employé le P.I.B. par habitant mesuré en euros. Ses cartes mettent donc en évidence des disparités assez fortes sur certaines frontières nationales qui ont tendance à ne pas apparaître sur les nôtres.

⁹ C'est-à-dire en éliminant toutes les NUTS 2 pour lesquelles les données manquaient à une des quatre dates.

¹⁰ Nous n'avons pas retenu de distinction entre emplois à temps plein et emplois à temps partiel dans la mesure où la proportion d'emplois à temps partiel n'est pas corrélée avec la productivité et où son pouvoir explicatif envers la proportion d'occupés dans la population totale est épuisé par d'autres variables (cf. infra). Par ailleurs, les occupés sont comptabilisés sur leur lieu de résidence et non sur leur lieu d'emploi.

Un faible P.I.B. par habitant s'explique, selon les régions, plutôt par une faible productivité ou plutôt par une faible participation. Il importe de distinguer les situations dans la mesure où l'assistance aux régions devrait être différenciée : dans le premier cas, on attend plutôt une intervention visant la modernisation de l'appareil productif, dans le second cas une intervention visant à favoriser l'emploi de la main-d'œuvre. Les cartes de la figure 3 récapitulent la situation des régions européennes par rapport à ces deux indicateurs. Spatialement, c'est plutôt la participation qui définit les grands ensembles à l'échelle européenne (grandes zones claires ou foncées), tandis que la productivité rend compte de variations plus locales (juxtaposition de petites zones de couleur différente, avec le même effet « capitale » que pour la distribution du P.I.B. par habitant).

D'un point de vue statistique, il importe également de connaître leur poids respectif, en sachant que leur interaction explique bien entendu la totalité de la variation du P.I.B. par habitant, et en notant que leur colinéarité est négligeable (leur corrélation n'est significative à aucune des quatre dates). Des deux composantes, c'est cette fois-ci la productivité qui explique le mieux la variabilité du P.I.B. par habitant, comme le montrent les coefficients du tableau 3. Les r^2 associés à la productivité sont de trois à huit fois supérieurs à ceux associés à la participation.

Tableau 3 : Coefficients de détermination simples avec le P.I.B. par habitant

	1988	1991	1994	1997
R ² pour la productivité	51.9%	62.7%	67%	64.6%
Nombre d'individus pris en compte	152	160	160	190
R ² pour la participation	15.6%	7.6%	15.9%	20.1%
Nombre d'individus pris en compte	152	160	166	190

Cependant, productivité et participation demeurent des catégories larges, intégrant elles-mêmes les effets de nombreuses variables qu'il convient de rechercher.

3.2 Les composantes de la participation

Par définition, trois ou quatre facteurs doivent suffire à expliquer la variabilité du nombre de personnes occupant un emploi dans la population totale :

- la part de population en âge de travailler,
- le taux d'activité,
- et le taux de chômage.

Nous ne disposons pas de données directes sur la part de population en âge de travailler, notamment parce que les définitions varient d'un Etat membre à l'autre. Mais nous pouvons approcher la notion indirectement à l'aide de deux variables relatives à la structure par âge de la population : la part de jeunes de moins de 25 ans et la part de personnes âgées de plus de 65 ans. Par contre, nous disposons de données détaillées par sexe sur les taux d'activité et de chômage.

Pour les quatre années, nous avons procédé à une régression pas à pas ascendante sur la participation, en entrant à chaque étape dans le modèle la variable la plus corrélée avec les résidus de l'étape précédente. Cette procédure a abouti à la sélection des mêmes variables pour les quatre dates, seule leur hiérarchie variant (donné par le test T^{11}).

¹¹ Il eût sans doute mieux valu employer les coefficients de corrélation partielle, qui n'étaient malheureusement pas disponibles sur le logiciel utilisé.

Tableau 4 : Valeurs du test T^{12} par variable pour les régressions linéaires sur la participation

	1988	1991	1994	1997
Nombre d'individus pris en compte	151	155	155	183
Taux d'activité féminin	25.99	32.17	28.17	25.67
<i>Rang</i>	1	1	1	2
Taux d'activité masculin	20.47	23.51	19.25	20.51
<i>Rang</i>	2	2	3	3
Taux de chômage	-18.47	-23.05	-26.45	-31.09
<i>Rang</i>	3	3	2	1
Part des moins de 25 ans	-16.64	-16.05	-15.65	-16.26
<i>Rang</i>	4	4	4	4
R ² général du modèle	97.6%	98.4%	98.2%	98.1%

Les taux d'activité féminin et masculin conservent jusqu'au bout des pouvoirs explicatifs indépendants. En première approche, les variations du taux d'activité féminin, principal régresseur pour les trois premières dates, semblent liées à des différences culturelles classiques, les plus faibles taux se retrouvant dans des régions méditerranéennes et les plus forts dans des régions britanniques et les Länder de l'Est. Les faibles taux d'activité masculins semblent eux largement dûs à des prises de retraite anticipées, surtout pour les régions françaises et belges¹³. Le pouvoir explicatif du taux de chômage croît, au point qu'il devient le premier régresseur en 1997. La part des moins de 25 ans, qui peut être considérée comme l'effet de la structure par âge de la population, joue un rôle mineur mais constant aux quatre dates.

3.3 Les composantes de la productivité

En ce qui concerne la productivité, exceptée la ventilation des emplois en trois secteurs d'activité, notre base est pauvre en données susceptibles d'expliquer logiquement sa variabilité. Nous avons donc décidé de tester toutes les variables à notre disposition afin de repérer d'éventuels effets inattendus. De nouveau, nous avons abouti à des modèles portant sur les mêmes variables pour les quatre dates.

Tableau 5 : Valeurs du test T par variable pour les régressions linéaires sur la productivité

	1988	1991	1994	1997
Nombre d'individus pris en compte ¹⁴	148	155	155	186
Part d'emploi agricole	-9.45	-8.14	-7.39	-6.6
<i>Rang</i>	1	2	2	1
Taux d'activité	-6.06	-9.34	-7.88	-5.15
<i>Rang</i>	2	1	1	2
Part des moins de 25 ans	-4.87	-3.11	-3.27	-4.54
<i>Rang</i>	3	4	4	3
Population moyenne	3.86	3.25	3.99	3.61
<i>Rang</i>	4	3	3	4
R ² général du modèle	58.4%	52%	48.6%	40.3%

Comme attendu, la structure des emplois, traduisant plus généralement la structure économique, joue un rôle important. Plus la part de l'agriculture, activité globalement peu productive, est forte, plus la productivité globale est faible. Mais de manière plus surprenante, le taux d'activité, sans distinction de sexe, est une variable tout aussi importante : plus il est faible, plus la productivité est forte. Certes les économies productives, réclamant des employés qualifiés, engendrent un allongement des études, mais l'ampleur du phénomène évoque une stratégie

¹² Pour les trois tableaux suivants, les valeurs de T sont significatives au moins à 0.005.

¹³ cf. la communication de la Commission Européenne de décembre 2000 sur l'avenir de la protection sociale, p.11.

¹⁴ La distribution du P.I.B./occupés comporte quelques individus aberrants. Nous avons donc volontairement éliminé la Corse et Bruxelles pour 1988 ; la Corse, Bruxelles et Hambourg pour 1991 ; Bruxelles, Hambourg et Luxembourg pour 1994 et 1997.

plus systématique de **mise à l'écart de la main-d'oeuvre la moins productive**. On peut faire l'hypothèse que des données sur les qualifications de la population se seraient substituées au taux d'activité dans le modèle.

Sans qu'il s'agisse d'un effet de structure évident mais sans qu'une explication causale soit facile à établir, la part de jeunes dans la population varie avec la productivité : les régions productives d'Allemagne de l'Ouest et d'Italie du Nord ont très peu de jeunes, tandis que des régions moins productives du Sud de l'Espagne, du Portugal, d'Irlande du Nord ou du Nord de la France en comptent de nombreux. Le rôle de la population est encore plus obscure du fait de la variabilité du maillage NUTS : le coefficient est tiré vers le haut par quelques grandes régions espagnoles, italiennes, françaises et allemandes qui sont les pôles économiques de leurs pays, mais les régions les plus productives du Nord-Ouest de l'Europe, prises dans un maillage plus serré, ont plutôt une faible population en valeur absolue.

Nos modèles aux quatre dates n'expliquent qu'autour de 50% de la variabilité de la productivité et la tendance est à la baisse. Il est clair que des variables rendant compte des infrastructures, de l'équipement des entreprises ou encore de la formation de la main-d'oeuvre seraient utiles, mais elles ne sont pas disponibles ou peu utilisables à l'échelle européenne pour le moment. Cependant, la cartographie des résidus ne révèle pas de structure spatiale évidente.

3.4 Le modèle final

Le taux d'activité joue en sens inverse vis-à-vis de la participation et de la productivité. Logiquement, il sort donc du modèle lorsque l'on régresse les variables sélectionnées aux étapes précédentes sur le P.I.B. par habitant – ne se maintenant, en dernière position, que pour 1991. Apparaissant comme une variable cruciale mais ambivalente à l'analyse, son rôle est masqué lorsqu'on modélise directement cet indicateur synthétique qu'est le P.I.B. par habitant. Son effacement pouvant permettre l'intégration de nouveaux éléments au modèle, nous avons testé le reste de nos variables pour établir le modèle final du tableau 6.

Tableau 6 : Valeurs du test T par variable pour les régressions linéaires sur le P.I.B. par habitant

	1988	1991	1994	1997
Nombre d'individus pris en compte	150	155	155	189
Part d'emploi agricole	-9.2	-7.65	-7.18	-6.89
Rang	1	1	1	2
Taux de chômage	-3.04	-4.67	-4.75	-8.4
Rang	5	2	3	1
Part des moins de 25 ans	-7.1	-3.55	-5.34	-5.73
Rang	2	3	2	3
Population moyenne	4.71	3.42	3.26	3.55
Rang	3	4	5	5
Densité moyenne	4.11	3.29	4.07	5.05
Rang	4	5	4	4
Taux d'activité	Non intégré	-2.87	Non intégré	Non intégré
Rang		6		
R ² général du modèle	71.6%	59.5%	61.3%	60.5%

Effectivement, à jeu égal avec la population en valeur absolue, apparaît la densité, qui traduit le léger effet « capitale » des cartes du P.I.B. par habitant et dont nous savons qu'il résulte peut-être d'une surestimation du P.I.B. par habitant des régions urbaines. Mais les principales variables sont celles des modèles précédents : d'abord l'emploi agricole, qui rend compte de l'effet productivité, le taux de chômage, qui rend compte de celui de la participation, puis la part des moins de 25 ans, qui ne peut être considérée totalement comme un effet de structure. Les résidus du modèle sont normaux mais il reste une forte part de variance à expliquer. Une nouvelle fois, la cartographie des résidus ne met pas en évidence de structure spatiale stable dans le temps, comme le montrent les cartes de la figure 4.

Sept variables ressortent donc de ces trois étapes de modélisation (on considère le taux d'activité dans sa composition par sexe), auxquelles on peut ajouter les résidus du modèle final. Elles vont maintenant être utilisées pour construire une typologie des situations économiques.

4 Typologie de la situation économique des régions de l'Union européenne

Nous avons choisi d'employer l'algorithme des nuées dynamiques pour classer les régions en fonction des variables sélectionnées. Appliqué sur les variables standardisées, il donne des résultats stables sur 5 occurrences du traitement de 250 itérations chacune (seules quelques régions changent de classe et peuvent être affectées manuellement à la classe dans laquelle elles apparaissent le plus souvent).

Les cartes résultantes (figure 5), montrent une tendance à l'auto-corrélation spatiale, permettant de dégager des ensembles qui débordent souvent les frontières nationales. Pour 1997, la probabilité que deux régions appartiennent à la même classe est 2,33 fois plus importante si les régions sont contiguës que si elles sont choisies au hasard (et cette proportion ne varie pas significativement aux autres dates).

Tableau 7 : profil des classes, année 1997

	Classes	A	B	C	D	E
	Effectif	75	67	23	16	7
Taux d'activité féminin (%)	Moyenne	44.0	52.2	34.8	38.1	46.8
	<i>Ecart-type</i>	5.8	5.3	5.8	7.3	6.2
Taux d'activité masculin (%)	Moyenne	63.3	70.5	60.5	64.7	66.4
	<i>Ecart-type</i>	3.2	3.0	2.7	4.0	3.6
Part d'emploi agricole (%)	Moyenne	5.0	3.8	10.7	28.0	0.4
	<i>Ecart-type</i>	3.0	2.6	4.9	10.6	0.2
Taux de chômage	Moyenne	10.2	6.8	20.9	7.3	12.2
	<i>Ecart-type</i>	3.8	2.9	3.9	3.0	6.8
Part des moins de 25 ans (%)	Moyenne	28.5	30.9	32.3	32.1	31.7
	<i>Ecart-type</i>	2.9	2.8	3.0	3.6	4.9
Population (milliers)	Moyenne	1939	1788	2744	580	1818
	<i>Ecart-type</i>	1516	1595	1999	506	1121
Densité de population (hab./km ²)	Moyenne	232.4	295.3	157.5	74.3	3576
	<i>Ecart-type</i>	256.8	330.9	146.4	72.7	1353
Résidus de la modélisation du P.I.B. par habitant	Moyenne	-1002	1140	1131	-885	-2360
	<i>Ecart-type</i>	2409	3409	967	3013	4747

Seuls les profils de la typologie de 1997 figurent dans le tableau 7, car ils sont qualitativement identiques aux quatre dates. Les régions de la classe A couvrent la majeure partie de la France et de la Belgique, l'Allemagne du Nord et l'Italie du Nord. Ce sont des régions dans une position moyenne pour la plupart des indicateurs, se distinguant avant tout par une faible part de jeunes – traduisant une population vieillissante – et un taux d'activité masculin assez faible. Leur P.I.B. par habitant tend à être surestimé par le modèle. Leur situation économique n'est pas mauvaise mais elles semblent manquer de dynamisme.

Les régions de la classe B s'étendent surtout sur l'Allemagne du Sud et l'Autriche du Nord, les Pays-Bas, l'Angleterre et la Scandinavie. Elle sont avant tout caractérisées par de forts taux d'activité, un faible emploi agricole et un faible taux de chômage. Leur P.I.B. par habitant a tendance à être sous-estimé par le modèle. Tous les indicateurs sont favorables.

Les régions de la classe C, concentrées principalement en Europe méditerranéenne ont de faibles taux d'activité, un très fort taux de chômage et comptent beaucoup de jeunes. Leurs difficultés se concentrent donc du côté de la participation, mais leur jeune population peut être une chance d'avenir si elle est bien formée. Et leur P.I.B. par habitant a tendance à être plus fort que prévu par le modèle.

La classe D est dans une meilleure situation du point de vue de l'emploi (taux d'activité moyens et taux de chômage faible) mais elle regroupe des régions très agricoles (28% de l'emploi en moyenne). Il leur faut convertir leur structure économique pour améliorer leur productivité, d'autant que leur P.I.B. a tendance à être surestimé par le modèle.

La classe E, enfin, regroupe quelques petites régions très urbanisées et dispersées, avec un emploi agricole quasi-inexistant et d'assez forts taux d'activité.

Cette typologie pourrait directement servir de guide pour l'action. Elle distingue notamment **deux types de régions à problème**, qui devraient être soutenues par des programmes de nature différente. D'ailleurs, l'extension des classes C et D n'est pas très éloignée de l'aire d'éligibilité à l'Objectif 1 pour la période 1994-1999. Mais des différences existent, qui justifient un retour conclusif sur le sujet.

5 Retour sur l'Objectif 1

Pour boucler la boucle, il nous faut finalement confronter la liste des régions éligibles à l'Objectif 1 aux variables que nous avons sélectionnées. Pour ce faire, nous avons employé l'analyse discriminante, qui permet de repérer d'éventuelles régions mal classées. La part d'emploi agricole, le taux de chômage et le taux d'activité féminin sont les variables qui contribuent le plus à différencier les groupes. Les cartes de la figure 6 illustrent les résultats pour 1997 et 1991, date des dernières statistiques disponibles lorsqu'ont été choisies les régions éligibles pour la période 1994-1999.

Le premier constat est qu'il y a peu d'erreurs de classement – et moins en 1991 qu'en 1997. Le second est que ces erreurs sont ciblées : en 1991, il s'agit surtout de régions qui ne sont en fait que partiellement éligibles (Berlin, Nord-Pas-de-Calais, Abruzzes), des régions capitales de pays entièrement éligibles (ex : Lisbonne) ou de cas ayant fait grand bruit à l'époque (ex : Flevoland). En 1997, de nouvelles « erreurs » apparaissent, marquant l'amélioration de la situation de régions aidées (Irlande dans son ensemble) ou l'aggravation de la situation de régions non-éligibles (Madrid, Aragon, Catalogne et Languedoc-Rousillon). Une anomalie notable disparaît : le Nord-Pas-de-Calais (qui figure constamment dans la classe C de notre typologie), rendu partiellement éligible en 1994 pour des raisons politiques, aurait vu sa situation se dégrader suffisamment pour justifier *a posteriori* son éligibilité !

Si l'on confronte maintenant l'éligibilité pour la période de programmation 2000-2006 aux variables de 1997, la fonction discriminante change (c'est le taux d'activité masculin qui devient le plus discriminant, devant le taux de chômage et la part d'emploi agricole) et les erreurs de classement – dont la proportion est stable – également. Elles se concentrent sur les régions qui changent de statut : bon nombre de celles qui ont perdu leur éligibilité (mais qui bénéficient d'un soutien transitoire) sont classées éligibles, et la plupart des « zones très peu peuplées » sont classées non-éligibles.

Bien entendu, les quelques résultats présentés ici ne se suffisent pas et la méthode d'analyse elle-même est susceptible d'évoluer. Ils ne sont qu'une des premières briques d'un vaste chantier de thèse, qui vise à produire *in fine*, à l'échelle européenne, un bilan spatialisé, critique et – qui sait – prospectif de l'action européenne en matière de politique régionale. Dans ce but, il nous faudra mener ce type d'exploration quantitative ou d'autres formes d'analyse spatiale, en les appuyant sur de sérieux décryptages conceptuels et contextuels.

BIBLIOGRAPHIE

CHADULE (1994), Initiation aux pratiques statistiques en géographie, 3^{ème} édition, Paris, Masson.

COMMISSION DES COMMUNAUTÉS EUROPÉENNES (2000), L'évolution à venir de la protection sociale dans une perspective à long terme : des pensions sûres et viables, *Communication de la Commission au Conseil, au Parlement Européen et au Comité Economique et Social*, Bruxelles, COM(2000) 622 final.

COMMISSION DES COMMUNAUTÉS EUROPÉENNES (2000), Onzième rapport annuel sur les Fonds structurels (1999), *Rapport de la Commission*, Bruxelles, COM(2000) 698 final.

COMMISSION DES COMMUNAUTÉS EUROPÉENNES (1999), Sixième rapport périodique sur la situation et l'évolution socio-économique des régions de l'Union européenne, Bruxelles.

DAVEZIES L. (1997), Intégration économique internationale et cohésion interrégionale en Europe, le problème de la gestion des échelles géographiques, *Les Découpages du Territoire, Dixièmes entretiens Jacques Cartier*, Lyon, 9-10 décembre 1997, Paris, INSEE, p. 267-278.

GRASLAND C. (1997), L'analyse géographique des discontinuités territoriales : l'exemple de la structure par âge des régions européennes vers 1980, *Espace géographique*, n°4, 30 p.

SANDERS L. (1989), L'analyse statistique des données en géographie, Montpellier, Collection Alidade, GIP Reclus.

Study Programme on European Spatial Planning, Final Report (2000), disponible en ligne à l'adresse www.nordregio.se

Figure 2a

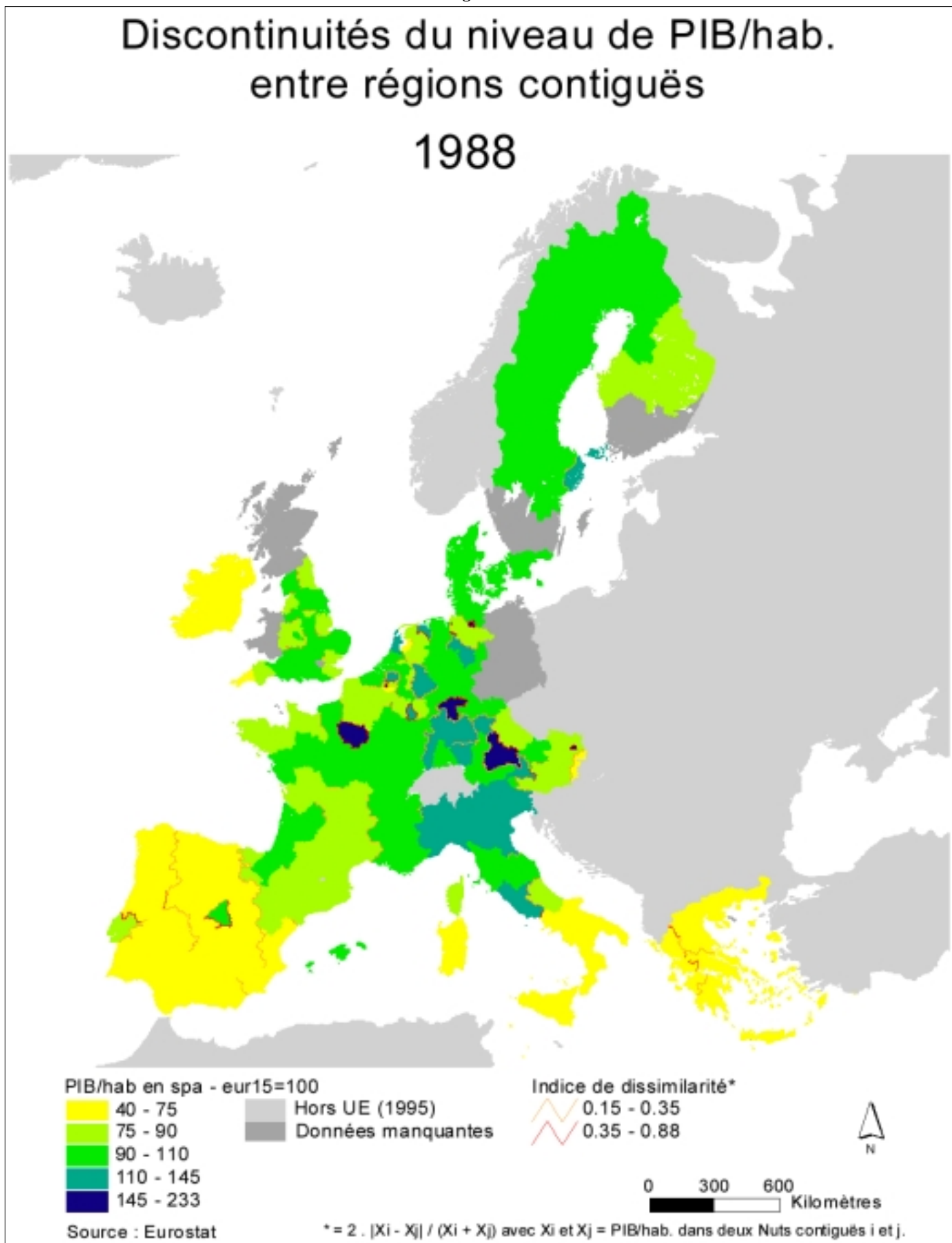


Figure 2b

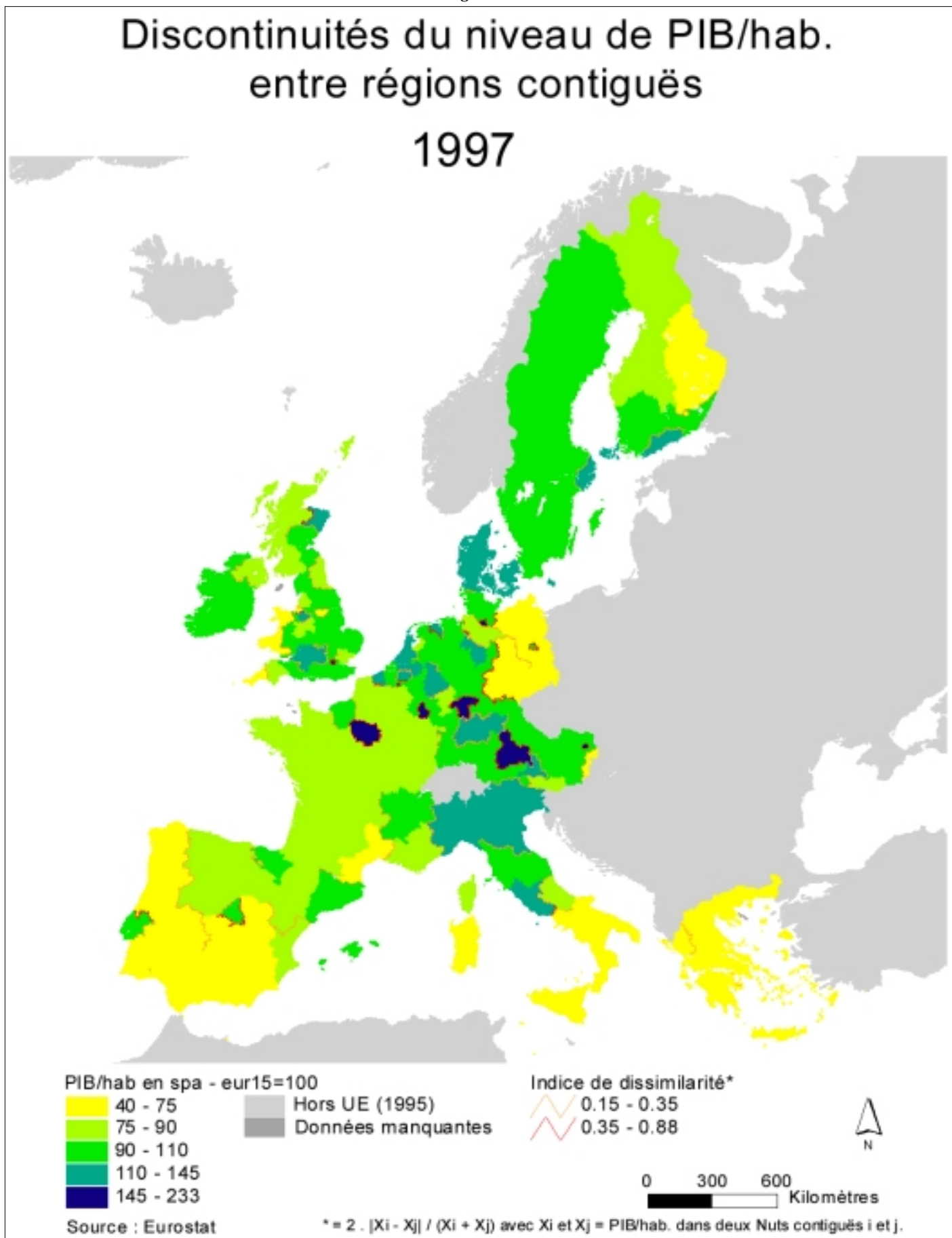


Figure 3a

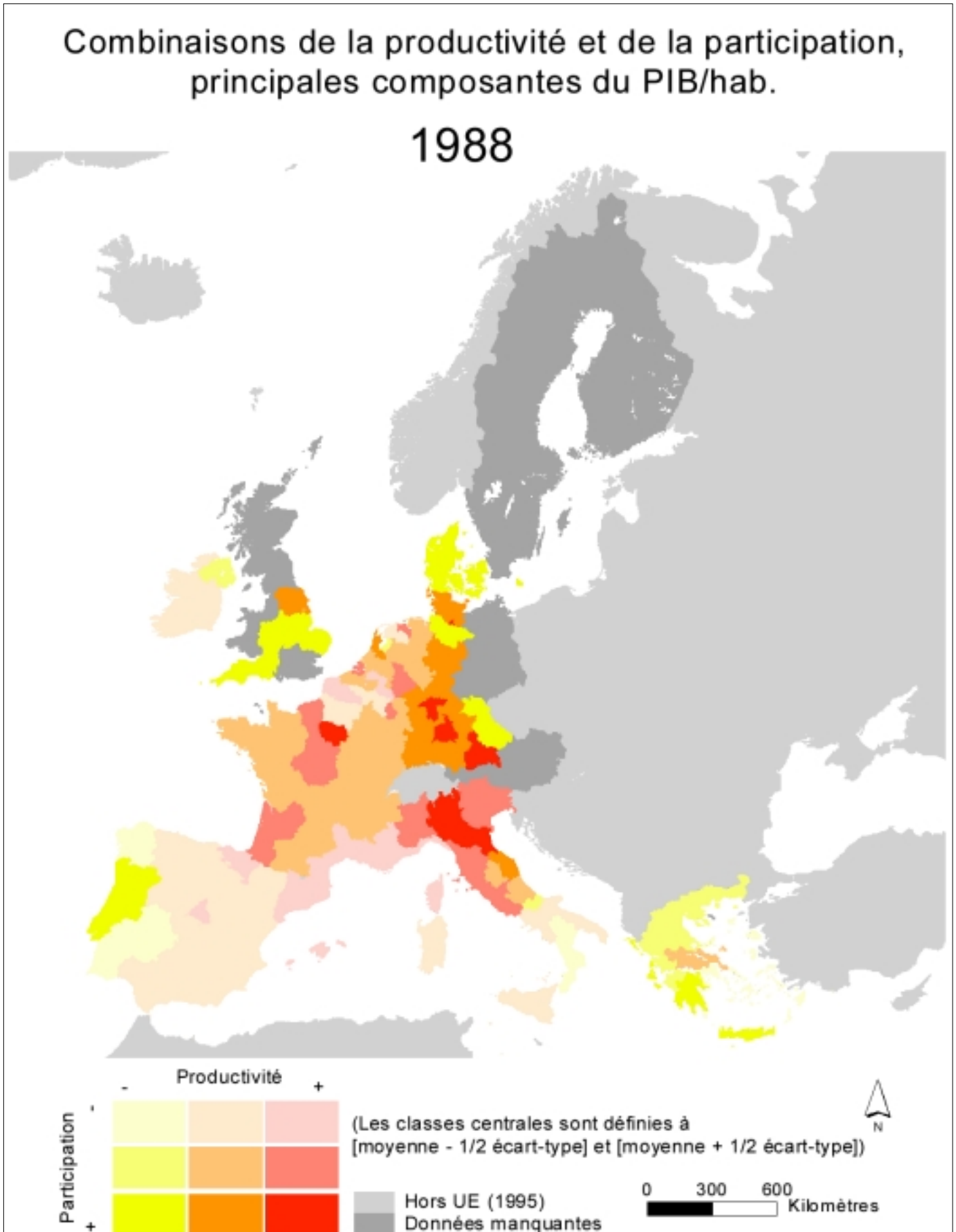


Figure 3b

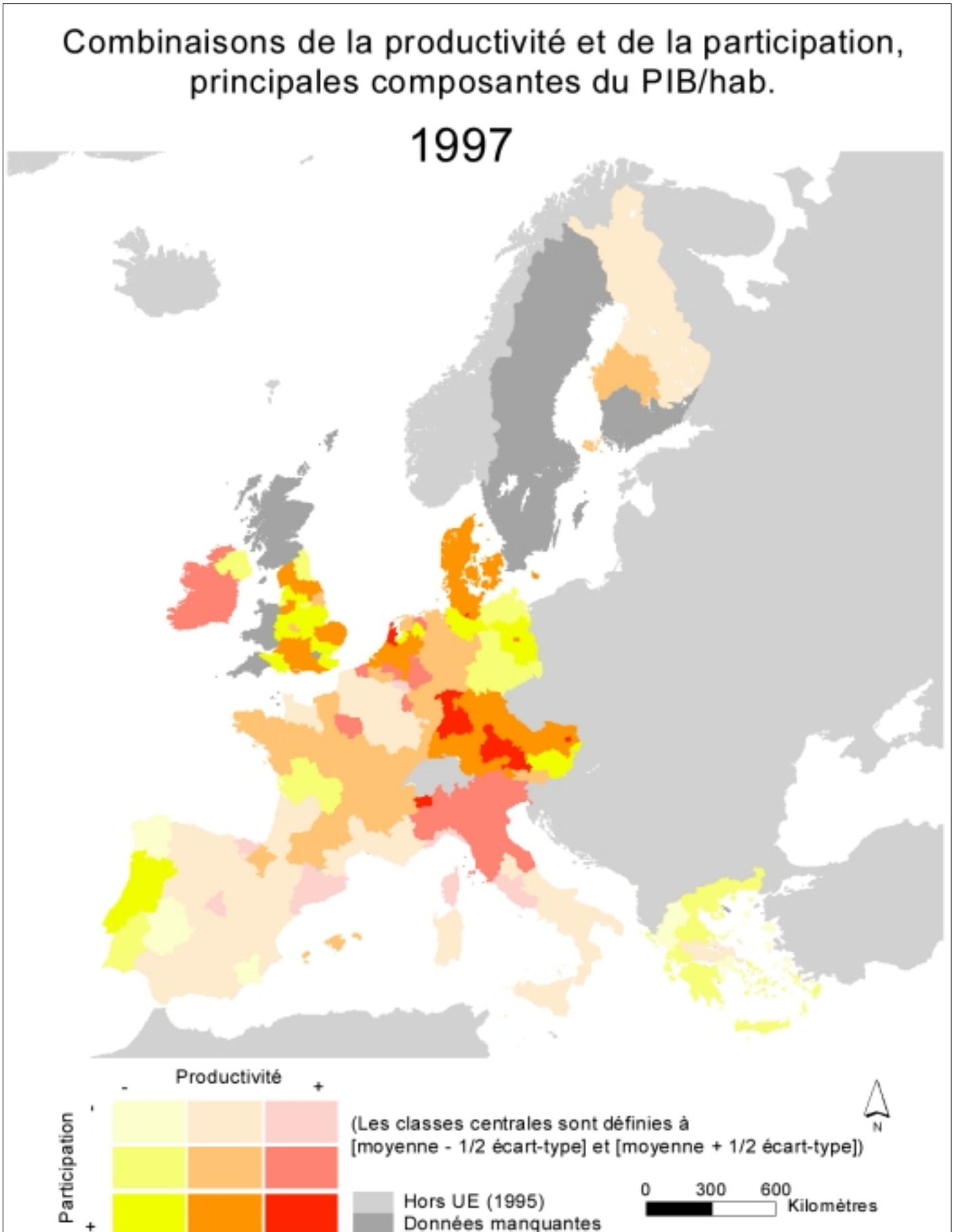


Figure 4a

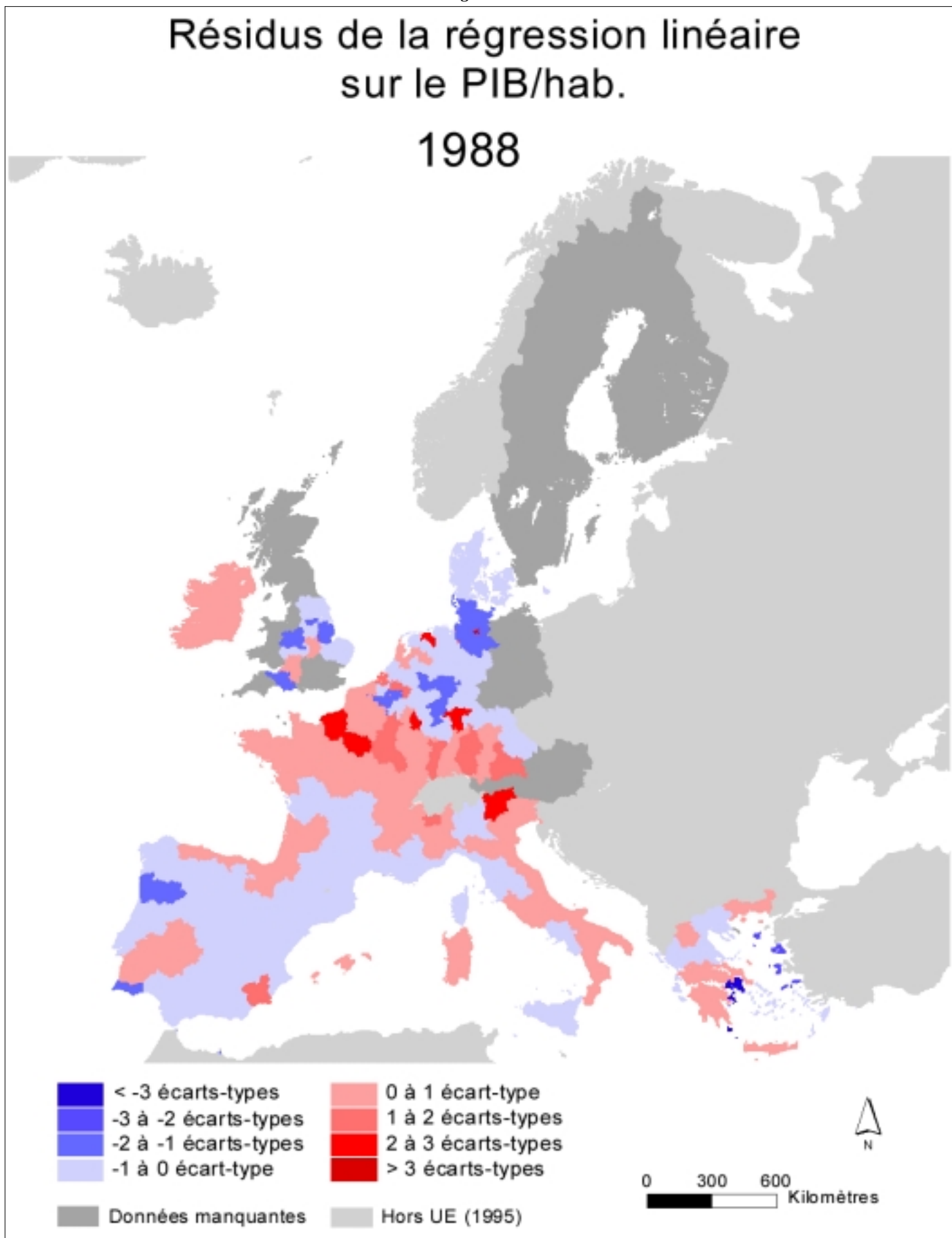


Figure 4b

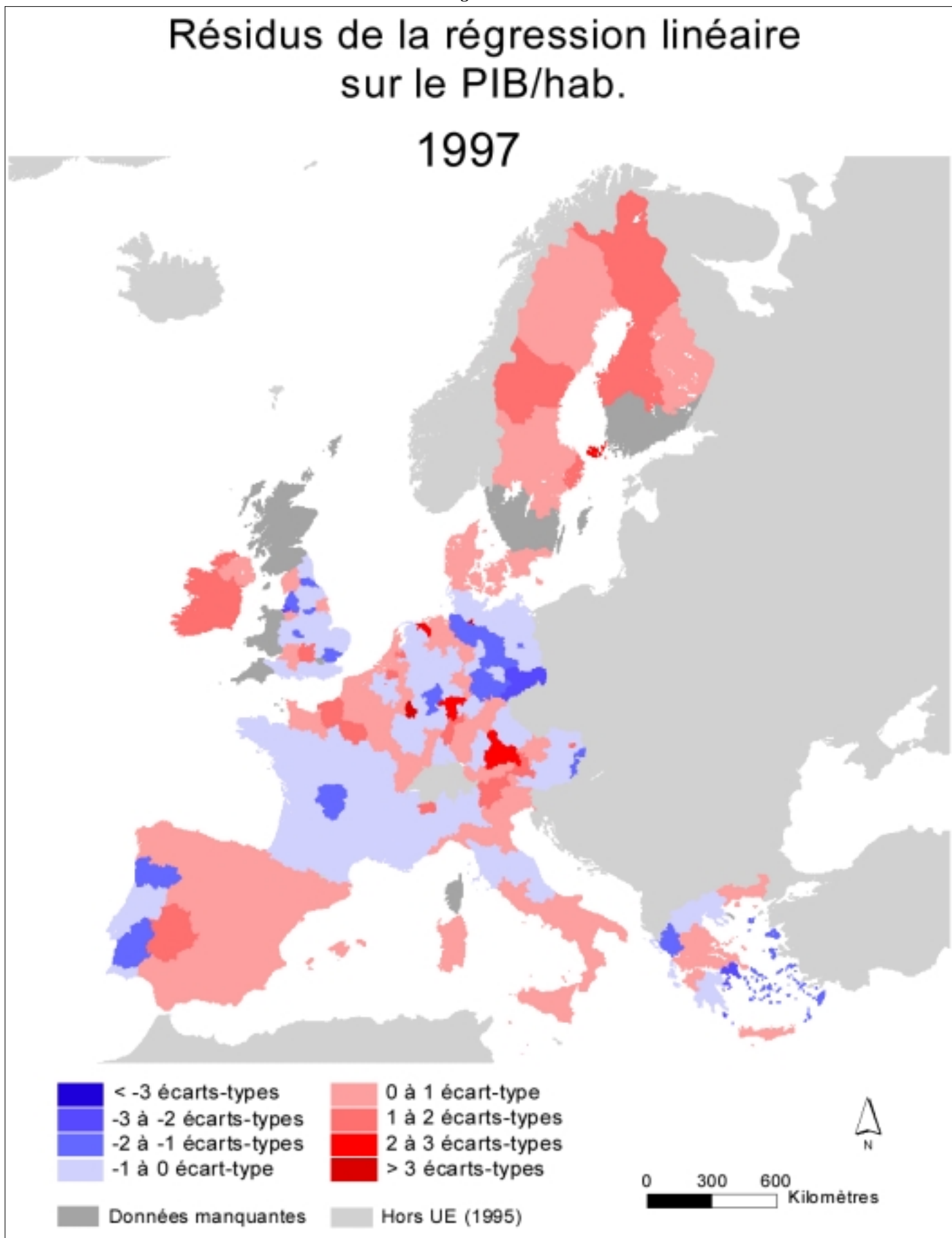


Figure 5a

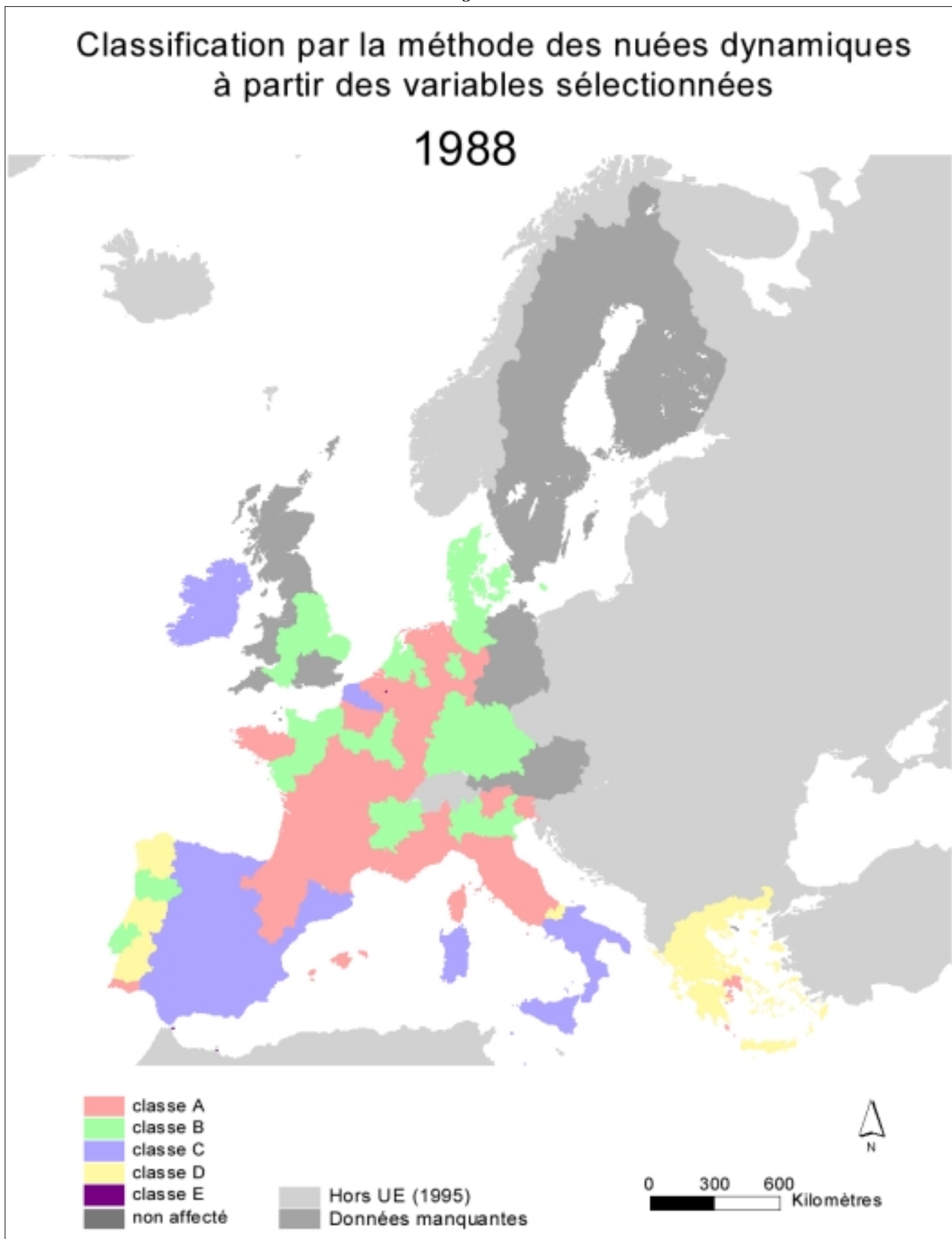


Figure 5b

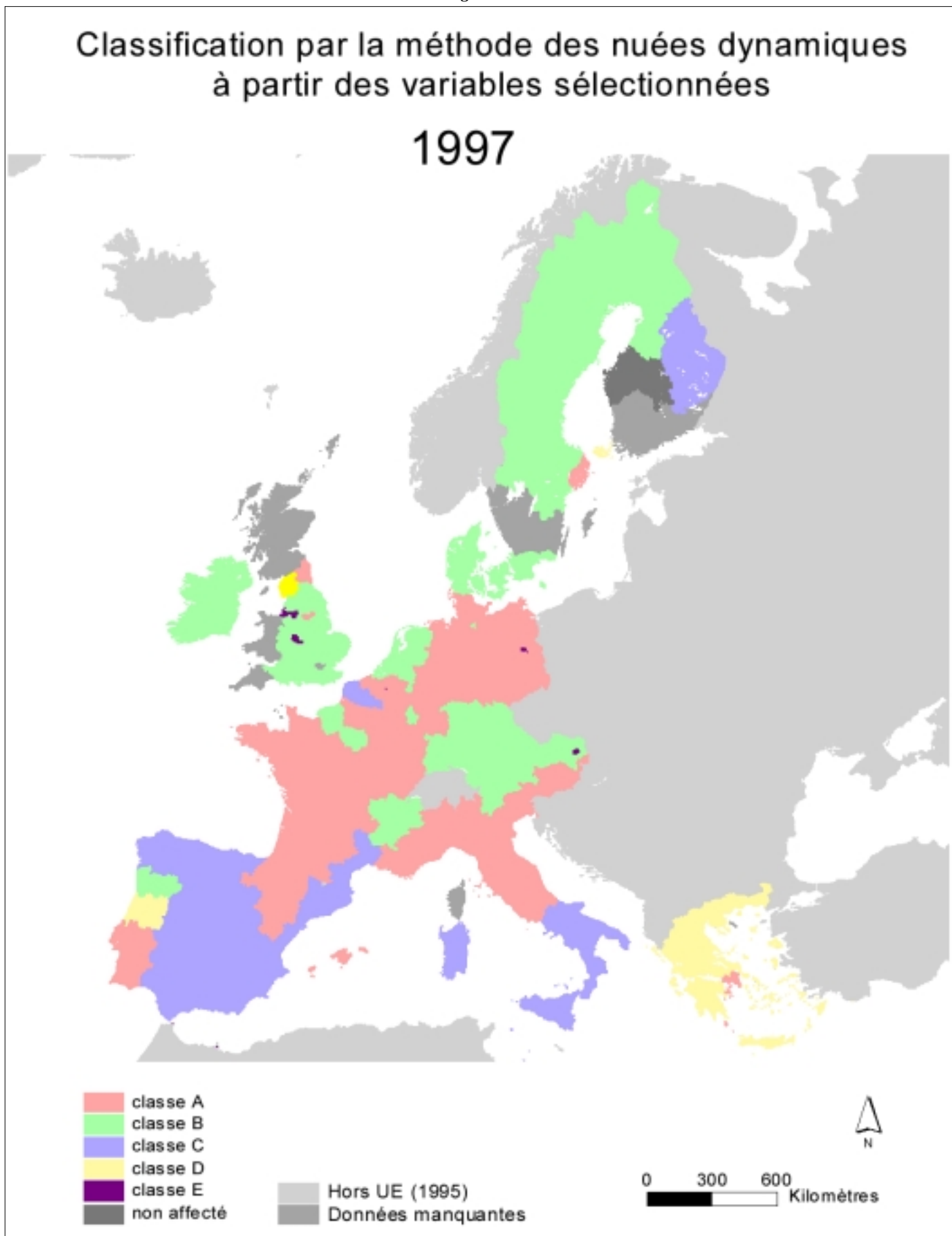


Figure 6a

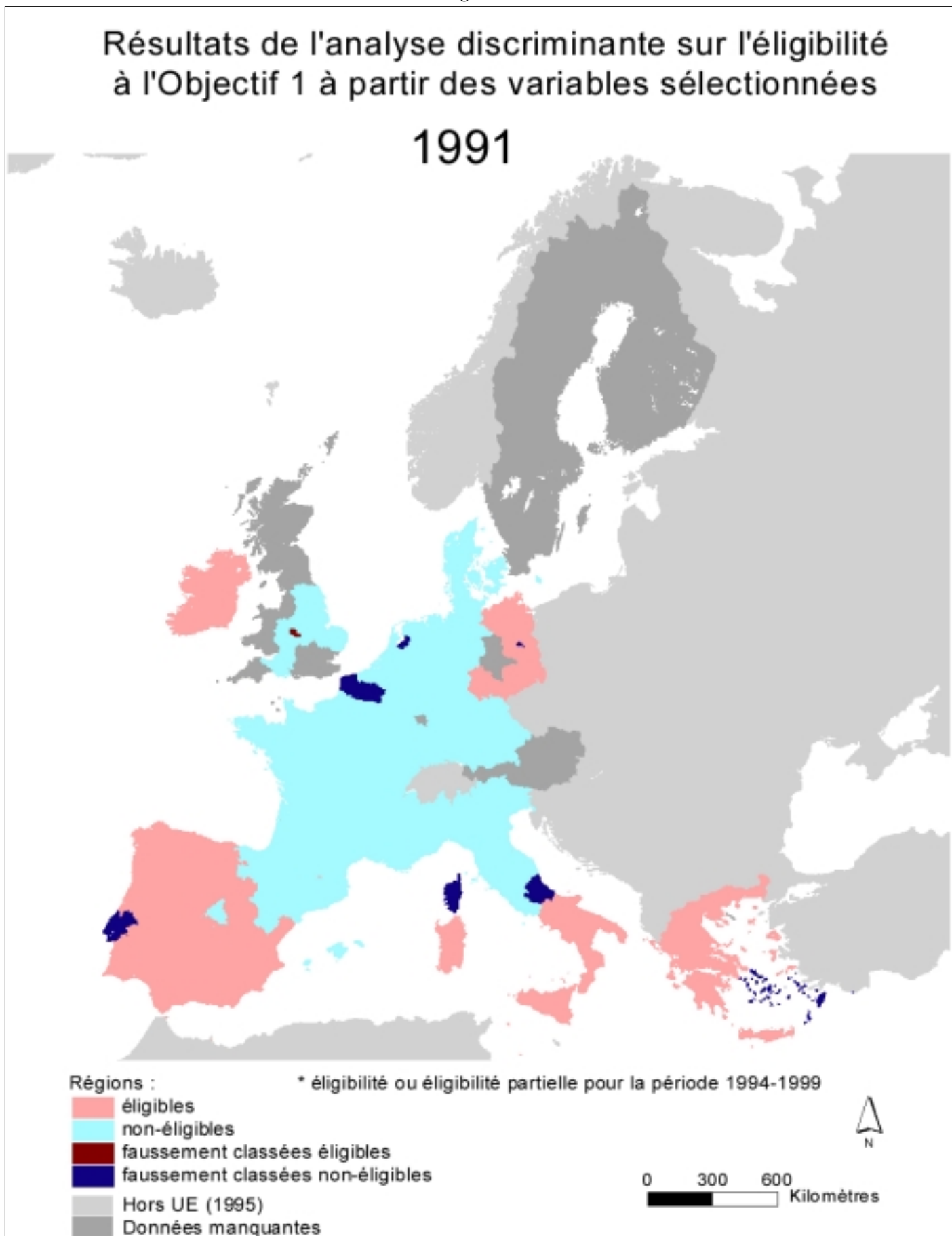


Figure 6b

