

Philippe Waniez

***Cartographie thématique
et
Analyse des Données***

avec Philcarto 5.xx pour Windows

Philippe Waniez

***Cartographie thématique
et
Analyse des Données***

avec Philcarto 5.xx pour Windows

Les **DOCS de GRANIT** sont publiés par l'UMR 5185 ADES. Ils contribuent à la réalisation des objectifs du groupe par la diffusion de documents didactiques portant sur les méthodes de traitement de l'information territoriale.

Les **DOCS de GRANIT** s'adressent aux étudiants de Master et de Doctorat, aux ingénieurs, chercheurs, enseignants-chercheurs et plus largement, à tous ceux qui ont à analyser des informations portant sur les territoires (chargés d'études, élus...). Ils facilitent la compréhension et la mise en pratique des techniques d'analyse spatiale et de cartographie géographique, en évitant le jargon, sans sacrifier la rigueur méthodologique.

Les **DOCS de GRANIT** présentent aussi des logiciels gratuits, téléchargeables sur Internet, qui permettent d'appliquer les méthodes exposées.

Les auteurs qui souhaiteraient offrir leur contribution aux **DOCS de GRANIT** sont priés de se faire connaître. Leur travail sera évalué, avec bienveillance, par le Conseil éditorial et une réponse leur sera donnée dans les meilleurs délais.

Directeur de la publication

Guy Di Méo, Directeur de l'UMR ADES

Directeur de la rédaction

Philippe Waniez, Professeur à l'Université Bordeaux 2

Secrétariat de rédaction

Marie-Bernadette Darignac, Secrétaire d'édition de l'UMR ADES, CNRS

Marie-Louise Penin, Cartographe de l'UMR ADES, CNRS

Charles Cheung, Cartographe de l'UMR ADES, Université Bordeaux 2

Olivier Pissot, Cartographe de l'UMR ADES, CNRS

Conseil éditorial

Xavier Amelot, géographe, Maître de Conférences à l'Université Bordeaux 3

Véronique André-Lamat, géographe, Maître de Conférences à l'Université Bordeaux 3

Charles-Antoine Arnaud, Statisticien de l'UMR ADES, CNRS

Laurent Couderchet, géographe, Maître de Conférences à l'Université Bordeaux 3

Pierre Drevet, géographe, Attaché d'enseignement et de recherche à l'Université Bordeaux 3

Marina Duféal, géographe, Maître de Conférences à l'Université Bordeaux 3

Maryse Gaimard, démographe, Maître de Conférences à l'Université Bordeaux 2

David Leyle, géographe, Attaché d'enseignement et de recherche à l'Université Bordeaux 3

Anne-Marie Meyer, géographe, Maître de Conférences à l'Université Bordeaux 3

Pierre-Yves Trouillet, géographe, Attaché d'enseignement et de recherche à l'Université Bordeaux 3

Sandrine Vaucelle, Professeure agrégée de géographie, chercheuse associée à l'UMR ADES

Tout courrier concernant les DOCS de GRANIT doit être adressé à :

Philippe Waniez

Université Victor Segalen Bordeaux 2

Laboratoire Sociétés, Santé, Développement (SSD)

146, rue Léo-Saignat, case 71

33076 Bordeaux CEDEX

philgeo@club-internet.fr

Table des matières

PREFACE POUR LES « DOCS DE GRANIT »	7
L'AUTEUR	9
LE GROUPE DE RECHERCHE EN ANALYSE DE L'INFORMATION TERRITORIALE - GRANIT	10
INTRODUCTION	11
LICENCE POUR L'UTILISATION DE <i>PHILCARTO</i> 5.XX	17
1. QUELS TYPES DE CARTES AVEC <i>PHILCARTO</i> ?	19
1.1. LES CARTES CHOROPLETHES	20
1.2. LES CARTES EN CERCLES PROPORTIONNELS	25
1.3. LES CARTES EN CERCLES PROPORTIONNELS COLORES	28
1.4. LES CARTES EN CERCLES PROPORTIONNELS SUR SURFACES COLOREES	30
1.5. LES CARTES EN DEMI-CERCLES AFFRONTES	32
1.6. LES CARTES EN DIAGRAMMES A SECTEURS	34
1.7. LES CARTES EN PLAGES DE NIVEAUX ET EN BLOC-DIAGRAMME	36
1.8. LES CARTES EN SURFACES DE TENDANCES	39
1.9. LES CARTES EN SEMIS DE POINTS	43
1.10. LES CARTES EN SEMIS DE POINTS COLORES	45
1.11. LES CARTES DE LIENS DITES AUSSI « EN OURSINS »	47
1.12. LES CARTES DE RESEAUX	49
1.13. LES CARTES DE GRAVITATION ISSUES DU MODELE DE HUFF	53
2. INSTALLER ET UTILISER <i>PHILCARTO</i>	57
2.1. TELECHARGEMENT ET INSTALLATION	57
2.1.1. INSTALLATION DU <i>.NET FRAMEWORK</i>	58
2.1.2. INSTALLATION <i>PHILCARTO</i> VERSION 5	59
2.2. EXECUTION DE <i>PHILCARTO</i>	60
2.2.1. OUVERTURE DES FICHIERS FOND DE CARTE ET DONNEES STATISTIQUES	61
2.2.2. LE BUREAU DE <i>PHILCARTO</i>	68
2.2.3. LA REALISATION D'UNE CARTE	68
3. LES CARTES CHOROPLETHES	71
3.1. SELECTIONNER UNE VARIABLE DE COULEUR	71
3.2. AFFICHER UNE CARTE CHOROPLETHE	71
3.3. LE DIALOGUE COULEURS	72
3.4. LES DIFFERENTES METHODES DE DISCRETISATION	72
3.5. MODIFIER LES BORNES DES CLASSES SUR L'HISTOGRAMME	75
3.6. CHANGER LE NOMBRE DE CLASSES	77

3.7. CHOISIR UNE GAMME DE COULEURS	78
3.8. UTILISER UNE GAMME DE COULEURS EN OPPOSITION.	79
3.9. AJOUTER UN HABILLAGE	82
3.10. IDENTIFIER LES ELEMENTS CARTOGRAPHIQUES	84
3.11. ENREGISTRER LA CARTE DANS UN FICHIER	84
3.12. EVALUER L'AUTOCORRELATION SPATIALE	89
3.13. COMPARER DES CARTES CHOROPLETHES « A VUE »	91
3.14. LA COADJACENCE DES CARTES CHOROPLETHES	95
3.15. LES CARTES CHOROPLETHES DE VARIABLES NOMINALES	97
4. LES CARTES EN CERCLES PROPORTIONNELS	101
4.1. SELECTIONNER UNE VARIABLE DE SYMBOLE	101
4.2. AFFICHER UNE CARTE EN CERCLES PROPORTIONNELS	101
4.3. LE DIALOGUE CERCLES	101
4.4. LE CALIBRAGE DES CERCLES	103
4.5. MODIFIER L'APPARENCE DES CERCLES	104
4.6. MODIFIER LA COULEUR DU FOND DE LA CARTE	105
4.7. CHOISIR LES CERCLES DE LA LEGENDE	106
4.8. DISQUE OU CIRCONFERENCE ?	108
4.9. TRACER PLUSIEURS CARTES EN CERCLES COMPARABLES ENTRE ELLES	110
4.10. CERCLES POUR REPRESENTER DES EFFECTIFS NEGATIFS ET POSITIFS	112
4.11. CARTES D'IDENTIFIANTS ET/OU DE NOMS.	113
5. LES CARTES EN CERCLES PROPORTIONNELS SUR CARTES CHOROPLETHES	115
5.1. SELECTIONNER UNE VARIABLE DE COULEUR ET UNE VARIABLE DE SYMBOLE	115
5.2. AFFICHER UNE CARTE EN CERCLES SUR UNE CARTE CHOROPLETHE	115
6. LES CARTES EN CERCLES PROPORTIONNELS COLORES	117
6.1. SELECTIONNER UNE VARIABLE DE COULEUR ET UNE VARIABLE DE SYMBOLE	117
6.2. AFFICHER UNE CARTE EN CERCLES PROPORTIONNELS COLORES	117
7. LES CARTES EN CERCLES PROPORTIONNELS AFFRONTES OU EN DIAGRAMMES A SECTEURS	121
7.1. SELECTIONNER DEUX VARIABLES DE SYMBOLE	121
7.2. AFFICHER LA CARTE EN CERCLES PROPORTIONNELS AFFRONTES	121
7.3. AFFICHER LA CARTE EN DIAGRAMMES A SECTEURS PROPORTIONNELS	122
8. LES CARTES EN SEMIS DE POINTS	125
8.1. SELECTIONNER UNE VARIABLE DE SYMBOLE	125
8.2. AFFICHER UNE CARTE EN SEMIS DE POINTS	125
8.3. LE DIALOGUE SEMIS DE POINTS	125
8.4. LE CALIBRAGE DU POINT	126
8.5. MODIFIER LA COULEUR DES POINTS ET LA COULEUR DU FOND	128
8.6. POINTS POUR DES EFFECTIFS NEGATIFS ET POSITIFS	129
9. LES CARTES EN SEMIS DE POINTS COLORES	131
9.1. SELECTIONNER UNE VARIABLE DE COULEUR ET UNE VARIABLE DE SYMBOLE	131
9.2. AFFICHER UNE CARTE EN SEMIS DE POINTS COLORES	131

10. LES CARTES DE LIENS DITES « EN OURSINS »	133
10.1. SELECTIONNER UNE VARIABLE DE LIEN	133
10.2. AFFICHER UNE CARTE DE LIENS	133
11. LES CARTES DE LIENS COLORES	135
11.1. SELECTIONNER UNE VARIABLE DE LIENS COLORES	135
11.2. AFFICHER UNE CARTE DE LIENS	135
12. LES CARTES EN LIGNES PROPORTIONNELLES	139
12.1. SELECTIONNER UNE VARIABLE DE SYMBOLE	139
12.2. AFFICHER UNE CARTE EN LIGNES PROPORTIONNELLES	139
12.3. LE DIALOGUE LIGNES	139
12.4. LE CALIBRAGE DES LIGNES	141
12.5. MODIFIER L'APPARENCE DES LIGNES	142
12.6. MODIFIER LA COULEUR DE FOND DE LA PAGE	143
12.7. CHOISIR LES LIGNES DE LA LEGENDE	145
12.8. TRACER PLUSIEURS CARTES EN LIGNES COMPARABLES ENTRE ELLES	146
13. LES CARTES EN LIGNES PROPORTIONNELLES COLOREES	147
13.1. SELECTIONNER UNE VARIABLE DE COULEUR ET UNE VARIABLE DE SYMBOLE	147
13.2. AFFICHER UNE CARTE EN LIGNES PROPORTIONNELLES COLOREES	147
14. LES CARTES LISSEES	149
14.1. SELECTIONNER UNE VARIABLE DE COULEUR	149
14.2. AFFICHER UNE CARTE CHOROPLETHE NON LISSEE	149
14.3. AFFICHER UNE CARTE LISSEE PAR TRIANGULATION	150
14.4. AFFICHER UNE CARTE LISSEE PAR KRIGEAGE	158
14.5. AFFICHER UNE CARTE DE SURFACE DE TENDANCE	161
15. LES AIRES D'INFLUENCES URBAINES : MODELES DE REILLY ET HUFF	165
15.1. SELECTIONNER UNE VARIABLE DE MASSE (D'EFFECTIF)	166
15.2. AFFICHER LA CARTE DES PROBABILITES D'INFLUENCES	166
15.3. LE PARAMETRAGE DU MODELE DE GRAVITATION	169
15.4. LE TRACE DES POLYGONES DE THIESSEN	171
16. L'ANALYSE FACTORIELLE ET LA CLASSIFICATION AUTOMATIQUE	173
16.1. CHOISIR UNE METHODE D'ANALYSE MULTIVARIEE	174
16.2. REALISER UNE ACP	175
16.3. REALISER UNE CAH	186
16.4. COMBINER CAH ET ACP	191
17. L'ANALYSE EXPLORATOIRE DES DONNEES	195
17.1. REALISER UNE ANALYSE EXPLORATOIRE	196
17.2. L'ANALYSE EXPLORATOIRE BIVARIEE	198
17.3. LE DIAGRAMME TRIANGULAIRE ET L'EXPLORATION TRIVARIEE	209

18. L'ANALYSE TERRITORIALE MULTISCALEAIRE	215
19. LES DONNEES STATISTIQUES	227
19.1. LES MATRICES D'INFORMATION SPATIALE	227
19.1.1. LES TABLEAUX DE CORRESPONDANCES	227
19.1.2. LES TABLEAUX DE MESURES QUANTITATIVES DE RAPPORTS OU D'INTERVALLES	228
19.1.3. LES TABLEAUX DE VALEURS NOMINALES	228
19.1.4. LES DONNEES MANQUANTES	229
19.1.5. LES FICHIERS DES MATRICES D'INFORMATION SPATIALE	229
19.2. LES MATRICES DE LIENS	232
19.2.1. LE FORMAT DE LA MATRICE POUR LES CARTES DE LIENS NON COLORES	233
19.2.2. LE FORMAT DE LA MATRICE POUR LES CARTES DE LIENS COLORES	233
19.3. L'ENREGISTREMENT DES VARIABLES CREEES PAR <i>PHILCARTO</i>	234
19.4. LE CALCUL D'INDICES	236
19.4.1. CALCULER DES POURCENTAGES	237
19.4.2. CALCULER DES DISTANCES AU POURCENTAGE MOYEN	239
19.4.3. CALCULER UN TAUX DE VARIATION GEOMETRIQUE	240
19.4.4. CALCULER UN INDICE DE CONCENTRATION DE GINI	242
19.4.5. CALCULER DES INDICES DE SPECIALISATION OU DE DIVERSIFICATION	244
20. LES FONDS DE CARTES	247
BIBLIOGRAPHIE	249

Préface pour les « DOCS de GRANIT »

Avec la série de documents en ligne des « DOCS de GRANIT », l'UMR 5185 ADES du CNRS et des Universités de Bordeaux 2 et 3 souhaite affirmer son intérêt majeur pour la réflexion dans le vaste domaine des méthodes de la recherche en sciences humaines et sociales. À leur propos, les chercheur(e)s et ITA de notre laboratoire partagent la conviction que le déroulement méthodologique constitue véritablement un tout dont les différentes étapes s'enchaînent, de façon inexorable et rigoureuse. Ils pensent que la validité globale, la réussite finale d'une recherche dépendent de la qualité des procédures et de l'innovation des pratiques jalonnant chacun de ses stades, comme d'ailleurs de la fluidité et de la cohérence de leur liaison.

En amont de la méthode figurent les grands corps théoriques et leurs cortèges de concepts, lesquels se déclinent des plus abstraits aux plus opératoires ; ces derniers étant ceux qui permettent d'encadrer les procédures très concrètes d'une investigation de terrain chargée de tester telle ou telle hypothèse de travail, de l'accréditer, de l'infirmer ou, plus généralement, de la nuancer. On sait combien, dans la tradition française, les géographes furent longs à se préoccuper de cet aspect théorique, pourtant crucial, de la méthode. S'ils proclamaient que l'induction est irremplaçable pour qui veut cerner une réalité, ils oubliaient qu'elle ne saurait se concevoir sans un mouvement itératif permanent entre l'observation de terrain, les pistes interprétatives qu'elle suggère, mais aussi des principes théoriques plus ou moins masqués, implicites, qui habitent immanquablement tout chercheur et qu'il se doit d'éclaircir afin d'annoncer sans fard d'où il parle et selon quels schémas interprétatifs il opère (passage inévitable par la méthode déductive). Cependant, ces matériaux qui vont servir à la vérification de nos hypothèses, qui sont modelés par le parti théorique (implicite ou explicite) que nous prenons, il nous faut d'abord les forger, les construire et les critiquer, indiquer leurs biais et leurs limites, que nous travaillions à partir de données de première ou de seconde main. Il s'agit ensuite de les traiter, d'en tirer une information plus élaborée que celle qu'ils contiennent à l'état brut ; ceci à l'aide d'outils mathématiques, statistiques ou probabilistes, eux-mêmes charpentés par de solides hypothèses. Puis, cette information affinée, nous devons la rendre géographiquement plus expressive, plus parlante ; c'est là qu'intervient toute la tâche de représentation graphique et cartographique.

Finalement, des fondements conceptuels et théoriques jusqu'à la carte, jusqu'au croquis, jusqu'à l'histogramme ou autres effets de l'art et des techniques graphiques, le cheminement méthodologique imprègne la recherche. Il en dresse l'armature, il en campe le squelette ; il en malaxe les chairs, jusqu'à engendrer parfois cette forme qui couronne l'ouvrage : celle de l'idéal type ou, mieux encore, du modèle (spécifique, régional, général, etc.). D'un bout à l'autre de ce processus, l'informatique demeure la compagne obligée du chercheur : banque de données laborieusement réalisées, innombrables outils informatiques permettant leur traitement quantitatif et/ou qualitatif, SIG confectionnés à l'aide de logiciels professionnels ou plus artisanaux, logiciels infographiques et cartographiques autorisant la production de spectaculaires images, plus vraies que nature ! Mais attention aux effets pervers d'une telle objectivation/réification... Nous le savons bien, tous, aujourd'hui, nous autres géographes ; même si le public, même si le (les) pouvoir(s) avides de ces représentations léchées et plaisantes, voire artistiques, nous en demandent toujours plus, conquis par la capacité de séduction, par la vertu performative de ces réalisations graphiques, de ces formes dont les qualités esthétiques permettent bien souvent de faire passer le fond, éventuellement moins reluisant, moins idyllique, auprès de l'opinion... Restons prudents. Méfions-nous que dans nos méthodes de géographes le signifiant ne s'impose pas arbitrairement au signifié ! Ainsi, tout au long de cette chaîne de la méthode, l'esprit technique et scientifique certes, mais aussi sa dimension critique doivent impérativement rester en éveil, refuser le « bluff technologique » pour satisfaire l'exigence suprême : celle d'une mise en évidence et d'un décryptage, même éphémère, de la complexité qui préside à l'organisation de ce monde... Celui de la géographie comme celui de toutes les sciences humaines et sociales dans lequel le premier se fond et se confond.

C'est à cette longue chaîne que, modestement mais sûrement, cette série « DOCS de GRANIT » s'attaque et va se consacrer, au fil des semaines et des mois, voire des années ; en tout cas on l'espère. Les chercheurs, les ITA et les doctorants du laboratoire ADES qui ont décidé d'en assumer la charge éditoriale souhaitent ardemment qu'un nouvel esprit souffle sur la méthode. Certes, ils ne sont pas vraiment les premiers, loin de là, à émettre ce vœu ! Comme leurs glorieux précurseurs, on pense en particulier aux chercheurs du GIP RECLUS dans les années 1980 – 1990, ils veulent absolument que les outils méthodologiques de la géographie sortent de leur « boîte noire », que le géographe ou autre spécialiste des sciences sociales de l'espace, ou impliquant l'espace des sociétés, voire la dimension spatiale de celles-ci (nous avons désormais l'embarras du choix quant aux formulations), accède à la pleine connaissance des instruments méthodologiques qu'il choisit et qu'il utilise... Qu'il soit également apte à en fabriquer de nouveaux, plus performants, plus adaptés aux problèmes qu'il aborde ; en un mot que la créativité des chercheurs porte autant sur la méthode que sur les thèmes d'investigation retenus, les sujets abordés et les hypothèses faites en vue de l'élucidation des problématiques et de la résolution des questions posées.

La mise en ligne, par le groupe de recherche GRANIT, de cette série de documents fait écho à un principe éthique majeur, celui de la mise à disposition libre et gratuite des outils de la méthode (informatiques en particulier) que le chercheur appartenant à un établissement, à une institution publique de recherche (le CNRS et l'Université en l'occurrence), construit au sein de son laboratoire. Les « DOCS de GRANIT » reflètent donc le souci des chercheurs de GRANIT et d'ADES de transmettre et de partager une connaissance, des savoir-faire forgés au cours des années ; savoirs en devenir dans la mesure où cette série s'ouvre volontiers au débat, aux participations extérieures que la formule souple d'édition adoptée autorise quasiment sans limite. Précisons que les « DOCS de GRANIT » émaneront pour partie de recherches originales menées par les membres de l'UMR ADES. Ils seront aussi l'expression de séances d'ateliers et de séminaires plus collectifs auxquels des chercheurs extérieurs au laboratoire ADES sont et seront régulièrement conviés. Les « DOCS de GRANIT » s'ouvrent ainsi à toute la communauté scientifique.

Guy Di Méo
Directeur de l'UMR 5185 ADES
CNRS – Universités de Bordeaux 2 et 3

L'auteur

Philippe Waniez, 52 ans, est Professeur des Universités à l'Université Victor Segalen Bordeaux 2.

Ancien Directeur de Recherche de l'Institut de Recherche pour le Développement (IRD ex-ORSTOM), il a exercé sa profession de géographe au Brésil et dans quelques autres pays d'Amérique du Sud. Ses problématiques de recherche se rapportaient alors aux fronts pionniers, aux mutations des systèmes agraires, aux processus de démocratisation et aux transformations religieuses.

Les principaux résultats de ces recherches ont fait l'objet d'une synthèse sous la forme d'un mémoire intitulé « les données et le territoire au Brésil » présenté pour l'Habilitation à Diriger des Recherches, HDR, obtenue en 2002 à l'Université Paris-X Nanterre (ce mémoire est accessible sur son site internet personnel)

De 1987 à 1995, il a participé aux travaux du Groupement d'Intérêt Public RECLUS à la Maison de la Géographie de Montpellier. Dans ce cadre, il a conçu plusieurs logiciels de cartographie (Cabral 1500, Choroscope...) et s'est intéressé à l'Analyse Exploratoire des Données.

Professeur à l'UFR Sciences et Modélisation de l'Université Bordeaux 2 depuis 2006, il y enseigne le traitement des données statistiques. Il dispense également des enseignements de géographie de la santé dans les UFR de Médecine et de démographie au Département de Sociologie.

Membre de l'équipe Sociétés, Santé, Développement de l'UMR 5185 ADES qui rassemble des chercheurs du CNRS et des enseignants-chercheurs des universités Bordeaux 2 et Bordeaux 3, il s'est attaché aux questions de santé au Brésil, problématique qui l'a conduit à établir récemment une géographie de l'abus des césariennes comme méthode d'accouchement dans ce pays.

Il participe à l'animation du Groupe de Recherche en Analyse de l'Information Territoriale, GRANIT, de l'UMR ADES par des séminaires, des ateliers et des publications méthodologiques.

Parallèlement à ses enseignements et ses recherches, il poursuit une activité de développement de logiciels d'analyse des données territoriales : Philcarto, Phildigit... qui connaissent un grand succès tant pour l'apprentissage de la cartographie thématique que pour son application scientifique l'aide à la décision.

On trouve sur son site internet personnel (<http://philgeo.club.fr>) certains de ses travaux scientifiques et les logiciels dont il est l'auteur, en téléchargement gratuit.

Le Groupe de Recherche en Analyse de l'Information Territoriale - GRANIT

Le Groupe de Recherche en Analyse de l'Information Territoriale, GRANIT, de l'UMR 5185 ADES, fédère diverses activités transversales de recherche menées au sein du laboratoire pour développer les compétences dans les domaines de la cartographie thématique, de l'analyse des données (classification, analyse factorielle), l'analyse d'images photographiques et satellitaires, des systèmes d'information géographique en général et des systèmes experts... Dans cette perspective, les activités du groupe portent sur :

- L'innovation dans le domaine des méthodes de traitement et de représentation de l'information territoriale ; toutes les sources d'information sont concernées, comme les fichiers administratifs, les images satellitaires, les cartes à toutes les échelles, etc. ;
- La capitalisation des savoir-faire disponibles dans l'UMR et leur confrontation, dans une perspective méthodologique sur certains programmes de recherche de l'UMR choisis parce qu'ils permettent de mobiliser chercheurs et ingénieurs participant à GRANIT.
- L'organisation des séminaires de recherche internes à l'UMR destinés aux chercheurs et doctorants impliqués dans cette démarche qui peuvent ainsi débattre de points techniques trop souvent relégués en annexe.
- La formation des doctorants et des professionnels sous la forme d'ateliers de haut niveau qui visent à rendre autonomes les participants afin qu'ils puissent réaliser eux-mêmes l'analyse des données rassemblées dans le cadre de leurs recherches.

Le manuscrit de cet ouvrage a été relu, évalué et amélioré par Marie-Louise Penin, Charles Cheung et Olivier Pisscoat, cartographes de l'UMR ADES et membres du GRANIT.

Introduction

La cartographie thématique peut être considérée comme une composante des Systèmes d'Information Géographique (SIG), à partir du moment où elle met en œuvre des logiciels ayant pour objectif de faciliter l'enregistrement de toute information relative à des éléments localisés dans l'espace afin de faciliter l'étude et la représentation graphique de leurs caractéristiques. Le champ d'application des SIG est extrêmement vaste et varié puisqu'il comprend virtuellement l'ensemble de la terre et même de l'Univers !

Il existe sur le marché du logiciel plusieurs SIG généralistes permettant de traiter (parfois simultanément) des données aussi différentes que des cartes topographiques, des photographies aériennes, des images satellitaires, et les attributs qui s'y rapportent (les altitudes, les parcelles agricoles ou la croissance de la végétation). Peu nombreux, relativement chers, nécessitant un apprentissage lourd pour être correctement maîtrisés, ces systèmes sont principalement destinés aux corps de métiers ayant une action concrète sur l'espace géographique : ingénieurs de l'équipement, urbanistes, agronomes, gestionnaires des collectivités locales, mais aussi pompiers, prospecteurs de ressources naturelles, logisticiens ... Dans la plupart des cas, ces professionnels n'ont pas pour objectif premier d'analyser les structures et les dynamiques des territoires. Il s'agit pour eux de répondre à des questions pratiques dans des domaines spécialisés et cela de manière à la fois rapide et précise, en particulier pour ce qui a trait aux localisations.

La cartographie thématique des données statistiques correspond à une classe de problèmes plus circonscrite puisqu'elle cherche à exploiter au mieux la grande masse de données statistiques localisées (dites aussi « à références spatiales » ou « géo-référencées ») qu'engendre la vie des sociétés. En règle générale, ces données ne se rapportent pas à des observations individuelles (comme une parcelle forestière, un poteau électrique, ou bien encore une arrivée de gaz...), mais à des circonscriptions au sein desquelles des observations, homogènes par nature, sont dénombrées (des personnes physiques, des exploitations agricoles...), formant ainsi des agrégats (ou si l'on préfère des « sous-populations » au sens statistique du terme « population ») décrits par une série d'attributs (sexe, âge, profession, types de cultures, nombre d'animaux...).

Les SIG généralistes proposent, bien entendu, des fonctions de cartographie des données statistiques, et un spécialiste de tel ou tel autre système *leader* du marché aura beau jeu de montrer qu'avec le logiciel dont il assure la promotion, « on peut tout faire ». L'expérience montre que les choses ne sont pas aussi simples que les agents commerciaux ont évidemment intérêt à le laisser penser. Lorsqu'on examine les SIG du marché sous l'angle qui nous intéresse ici, on peut être surpris par plusieurs de leurs caractéristiques. La première est leur surcharge en fonctions ou options rarement employées ; ce caractère encyclopédique s'explique avant tout par la concurrence effrénée à laquelle se livrent les sociétés spécialisées avec pour conséquence la nécessité de rallier le plus grand nombre de clients en leur offrant les produits les plus complets. Ce qui peut être considéré comme une fuite en avant, loin d'être un atout de mon point de vue, conduit dans la pratique, à des logiciels confus qui demeurent hermétiques à ceux qui souhaitent seulement utiliser un outil d'analyse sans se transformer en ingénieur spécialiste des SIG.

De même qu'on ne se transforme pas en expert comptable en achetant un logiciel de comptabilité, on ne devient pas *geo-analyst* seulement en achetant, au prix fort, *Machins Info*, *Cartes et trucs* ou *Bidules View* : une fois les boîtes ouvertes,

les premiers essais laborieux ne conduiront-ils pas à un constat d'impuissance et à remiser lesdites boîtes au placard ?

La seconde observation qui découle de l'examen des possibilités des SIG généralistes en matière de cartographie des données statistiques apparaît à l'exacte symétrie de la précédente. On est étonné, en effet, qu'au delà de la variété des techniques qu'ils proposent, demeurent des lacunes incompréhensibles. Par exemple, on apprend (il faudrait dire on apprenait), dans le cours de cartographie thématique (vocabulaire employé autrefois, à l'époque de l'encre de chine sur du papier calque) de première année de géographie, à cartographier des données ternaires (du genre : pourcentage de la population active dans les secteurs primaire, secondaire et tertiaire) en recourant à une partition en classes d'un diagramme triangulaire. Il s'agit d'une méthode simple, robuste, largement connue des cartographes et pourtant absente de nombre de SIG commerciaux. Et l'on pourrait multiplier les exemples ! L'existence de telles lacunes a une conséquence pernicieuse : la désaffection des géographes-cartographes pour des méthodes certes éprouvées, mais que les fabricants de logiciels ne jugent pas utile de proposer ; on pourrait même se demander si les informaticiens concepteurs de ces savants systèmes sont réellement épaulés par des cartographes thématiciens !

Il ne serait pas difficile de multiplier les critiques, mais il ne s'agit pas ici de développer un réquisitoire, sans doute un peu stérile, à l'encontre des SIG généralistes qui sont d'une grande utilité dans de nombreux domaines. On cherche seulement à faire pressentir au lecteur que le système le plus complet, et souvent le plus cher, n'est pas obligatoirement celui qui lui convient, la « richesse fonctionnelle » n'étant pas nécessairement le bon critère de choix. C'est en faisant ce constat dans le cadre de ma propre pratique scientifique que j'ai été amené à concevoir et à programmer *Philcarto* pour *Windows*. On y trouve les méthodes et techniques d'analyse les plus courantes, celles habituellement utilisées pour analyser des données statistiques spatialisées, dont l'archétype est le recensement de la population. Sur un fond de carte représentant les unités de recensement (des polygones pour les circonscriptions administratives, des points pour les chefs-lieux), et en fonction des valeurs statistiques enregistrées sur ces unités, le cartographe fait varier les couleurs des polygones en fonction d'une gamme de couleurs appropriée ou fait varier la surface des symboles placés sur les chefs-lieux dans le but de découvrir des arrangements spatiaux révélateurs de l'organisation des phénomènes étudiés dans un cadre territorial donné. On peut ainsi représenter des populations tout en étudiant leurs caractéristiques socio-économiques, ce qui demeure l'un des objectifs majeurs de la géographie et de bien d'autres sciences sociales.

Le champ d'application de la cartographie des données statistiques ne se limite pas aux données démographiques, mais concerne l'ensemble des sources statistiques agrégées dans des circonscriptions géographiques et présentant de ce fait une dimension spatiale. Par exemple, les recensements ou enquêtes agricoles, les résultats électoraux, les enquêtes sur la fréquentation des équipements collectifs... sont des domaines pour lesquels ce genre de cartographie s'avère d'une grande efficacité pour extraire de la confusion des données, quelques grands axes de structuration des territoires. Bien entendu, les possibilités d'application de ces techniques s'avèrent extrêmement vastes car elles ne se limitent pas aux seules activités humaines.

En effet, nombreuses sont les données environnementales pouvant aussi être traitées de la même façon pour analyser la variation spatiale d'un ou plusieurs

attributs de populations du monde animal (poissons, crevettes...), végétal (platanes...) ou minéral (sols...).

En réalisant *Philcarto* et en permettant le téléchargement du logiciel sur mon site internet (<http://philgeo.club.fr>), je ne cherche pas à concurrencer les logiciels de cartographie du commerce dont on continuera à nous vanter, à tort ou à raison, les grandes qualités et la « richesse fonctionnelle ». Il s'agit pour moi de rendre accessibles des méthodes de cartographie et d'analyse des données géographiques éprouvées qui devraient faire partie de la boîte à outils de tous les géographes, ce qui ne semble toujours pas être le cas aujourd'hui. Dans cette perspective, *Philcarto* est avant tout destiné à appuyer un enseignement de géographie dans l'enseignement secondaire et supérieur afin d'offrir aux élèves et étudiants un outil facile d'accès, simple d'emploi, tout en étant efficace et si possible juste sur le plan scientifique.

Philcarto pour Windows® bénéficie de l'expérience acquise lors de la réalisation des versions pour Macintosh de *Philcarto* et de *Philexplo* (pour l'analyse exploratoire). J'ai entrepris la réalisation de la version Windows® à la fin de l'année 2000, et une première version a été disponible au début de l'année 2001. Ce n'est qu'après un important effort de développement que la version 2, caractérisée par de très nombreuses améliorations dans l'économie générale du logiciel (par exemple l'accélération de la jonction des données avec le fond de carte) a proposé quatre nouveaux modules : *Explo* pour l'analyse exploratoire, *Multiv* pour l'analyse multivariée, *AST* comme Analyse en Surfaces de Tendances, et *Huff* pour le modèle gravitationnel de même nom. Ces méthodes directement connectées aux techniques cartographiques courantes ont fait de *Philcarto* un puissant outil d'analyse des données géographiques. La version 3 de *Philcarto*, diffusée dans le courant de l'année 2003, ajoute les cartes en plages de niveaux obtenues par triangulation ou par krigeage. Une représentation en perspective (3D) dotée de possibilités d'animation complète l'ensemble. La version 4 disponible au début de l'année 2004 améliore beaucoup les cartes de liens de point à point (oursins) en permettant la coloration des liens en fonction des valeurs d'une variable statistique. C'est également le cas des cartes en semis de points. Mais la nouveauté la plus visible de la version 4 réside dans l'affichage simultané de la carte et de la fenêtre d'options pouvant lui être appliquées ; l'effet du calibrage des cercles ou de la sélection d'un mode de discrétisation est immédiatement perceptible facilitant ainsi la réalisation de la meilleure carte possible.

Pour la version de 5 de *Philcarto*, j'ai entièrement réécrit le logiciel en adoptant la nouvelle plate-forme de programmation (gratuite !) de Microsoft dénommée *Visual Studio Express 2005* sous *Windows Vista*. L'intérêt principal de cette démarche réside dans l'existence d'une nouvelle bibliothèque de fonctions graphiques d'une grande efficacité (GDI+) qui permet d'améliorer nettement les performances du programme (qui semblaient déjà satisfaisantes dans les versions précédentes). Comme cette version 5 de *Philcarto* exige, pour s'exécuter, le .Net Framework, elle fonctionne seulement sous *Windows 2000 Service Pack 3*, *Windows XP Service Pack 2* et sous *Windows Vista*. J'ai profité de cette « refondation » pour retrancher quelques fonctions peu utilisées ou donnant des résultats discutables qui alourdisaient inutilement le programme. J'ai également ajouté de nouvelles possibilités comme, par exemple l'affichage simultané de 4 cartes différentes, la méthode d'Analyse Territoriale Multiscalaire, la technique de comparaison des cartes choroplèthes à l'aide d'un graphe dérivé, etc. L'interface a été entièrement repensée

afin de permettre une plus grande facilité dans la succession des opérations de construction interactive des cartes.

De même qu'on ne pilote pas une *Ferrari* comme un *Solex* (une image un peu flatteuse pour *Philcarto*, j'en conviens !), l'utilisation de cette nouvelle version demande quelques connaissances de base à la fois en cartographie, mais aussi en traitement des données. L'utilisateur doit avoir conscience que le logiciel peut produire un résultat inutilisable si la méthode employée n'est pas applicable aux données à traiter ; pour limiter ce risque, j'ai introduit dans la version 5 de *Philcarto*, la notion de type de donnée (nominale, rapports, quantités) qui réduit, sans les éliminer totalement, les pièges liés à une mauvaise utilisation de telle ou telle autre méthode. Mais dans tous les cas, c'est à l'utilisateur de veiller à ce que les bonnes données soient soumises aux bonnes techniques d'analyse ; et comment pourrait-il le faire en ignorant les conditions de leur application ?

Il ne s'agit pas d'exiger du géographe-cartographe qu'il se transforme en statisticien, mais seulement de lui demander de faire l'effort nécessaire au bon emploi des méthodes que *Philcarto* met à sa disposition. Le risque est d'autant plus grand que *Philcarto* est devenu très performant et largement utilisé pour l'enseignement. Limiter le risque de produire très rapidement des résultats erronés (et qui plus est, en couleur !) a été mon souhait permanent en concevant cette nouvelle mouture du logiciel. Avec cette même perspective, je recommande la lecture des bons ouvrages suivants (presque tous écrits par des géographes) :

Beguín M., Pumain D., 2000, *La représentation des données géographiques. Statistique et cartographie*. Paris, Armand Colin, Col. Cursus, 192 p.

Dumolard P. L'outil informatique en géographie. Fichier au format pdf (nécessite Acrobat Reader) à télécharger à l'adresse suivante (gratuit) :
http://iga.ujf-grenoble.fr/cours-en-ligne/book_geoinfo.pdf

Dumolard P. Analyse multivariée de données géographiques. Fichier au format pdf (nécessite Acrobat Reader) à télécharger à l'adresse suivante (gratuit) :
http://iga.ujf-grenoble.fr/cours-en-ligne/bk_multiv.pdf

Groupe Chadule, 1997, *Initiation aux pratiques statistiques en géographie*. Paris, Armand Colin, Col. U, 192 p.

Le Fur A., 2007, *Pratiques de la cartographie* (2^e édition). Paris Armand Colin, Col. 128, 128 p.

Minvielle E., Souiah S.-A., 2003, *L'analyse statistique et spatiale*. Nantes, Editions du Temps, 284 p.

Pumain D., Saint-Julien T., 1997. *L'analyse spatiale. Tome 1 - Localisations dans l'espace*. Paris, Armand Colin, Col. Cursus, 192 p.

Pumain D., Saint-Julien T., 2001. *Les interactions spatiales*. Paris, Armand Colin, Col. Cursus, 192 p.

Toutain S., Souiah S.-A., 2005, *L'Analyse démographique et spatiale*, Nantes, Editions du Temps. Téléchargeable en version électronique (pdf) sur <http://www.numilog.com> (payant)

Zanin C., Trémélo M.-L., 2003, *Savoir faire une carte*. Paris, Belin, Col. Sup., 200 p.

In memoriam

Philcarto est dédié à la mémoire de deux géographes-cartographes disparus prématurément :

Jean Piwnik, Maître de conférences à l'Université Paris-X Nanterre qui m'a enseigné les bases de la cartographie thématique dans les années 1975-1980.

Thérèse Panouillères, Ingénieur au CNRS qui a assuré le secrétariat de rédaction de la revue *Mappemonde* depuis sa création.

Remerciements à ...

V. Brustlein, ingénieur cartographe pour ses nombreux conseils tant pendant la conception de *Philcarto* que lors de la rédaction de ce manuel.

C. Cauvin, professeur à l'Université Louis Pasteur de Strasbourg, pour m'avoir communiqué le programme source de la discrétisation de Jenks.

C. Cheung, cartographe de l'UMR ADES pour son soutien indéfectible pour instiller un peu d'information géographique dans l'Université Bordeaux 2.

G. Di Méo, Directeur l'UMR ADES qui ne ménage pas ses efforts pour faire de l'UMR un lieu de recherche d'excellence, de convivialité et de valorisation scientifique individuelle et collective.

J.-C. Filleron, professeur à l'Université de Toulouse Le Mirail, pour son soutien indéfectible à mes efforts de diffusion de *Philcarto*.

C. Grasland, professeur à l'Université Paris VIII, pour sa description précise de l'Analyse Territoriale Multiscale.

C. Lefebvre, cartographe à l'INRA, pour ses remarques et critiques toujours bienvenues.

M.-L. Penin, cartographe de l'UMR ADES, pour la relecture détaillée et critique du présent document.

O. Pissot, cartographe de l'UMR ADES, pour le travail en commun sur l'Analyse Territoriale Multiscale.

R. Gambini pour avoir publié le programme source de recherche des limites d'aires d'influences urbaines que j'ai adapté à *Philcarto*.

A. Goodman pour le programme source d'analyse en surfaces de tendances qu'il met à la disposition des chercheurs sur Internet et que j'ai adapté à *Philcarto*.

F. Guerrazzi pour sa DLL de géostatistique *Terrain 5* que j'ai intégrée à *Philcarto*.

M. Jambu et **M.-O. Lebeaux** pour avoir publié le programme source de la routine AGREGA utilisée en CAH et que j'ai adaptée à *Philcarto*.

F. Murtagh, professeur associé à l'Observatoire Astronomique de Strasbourg, pour les programmes sources d'ACP et d'AFC qu'il met à la disposition des chercheurs sur Internet et que j'ai adaptés à *Philcarto*.

F. Schindler pour son module de classe d'ajustements polynomiaux.

Ainsi qu'aux utilisateurs de *Philcarto* qui, grâce à leurs encouragements, suggestions, conseils, remarques m'ont incité à poursuivre le développement de *Philcarto* et à en améliorer les performances.

Adobe Illustrator, Postscript, PDF, Acrobat et Acrobat Reader sont des marques déposées de Adobe Inc.
Arc Info, Arc View sont des marques déposées de ESRI. Mapinfo est une marque déposée de Mapinfo.
Windows 95 / 98 / ME / 2000 Pro, XP, Vista, Excel sont des marques déposées de Microsoft Corporation.

Licence pour l'utilisation de *Philcarto* 5.xx

Veuillez lire attentivement les conditions d'utilisation du logiciel *Philcarto* dénommé ci-après PRODUIT LOGICIEL, et les limitations de responsabilité. Notez que certaines de ces conditions ne sont pas applicables dans certains pays, en fonction de leur propre législation.

Toute utilisation du PRODUIT LOGICIEL implique votre acceptation pleine et entière de l'ensemble des clauses de la licence d'utilisation.

Publication de cartes réalisées avec le PRODUIT LOGICIEL : Vous vous engagez à porter sur toute publication de cartes réalisées avec le PRODUIT LOGICIEL (quel que soit le support, imprimé, vidéo, CD, web, etc.), la mention suivante :

« Réalisé avec *Philcarto* : <http://philgeo.club.fr> »

Distribution : Vous êtes autorisé à télécharger librement le PRODUIT LOGICIEL pour Windows à partir du site Internet suivant : <http://philgeo.club.fr>

Cette autorisation est accordée seulement pour la version 5.0. à 5.99 de *Philcarto* et ne s'applique donc pas aux éventuelles futures versions du PRODUIT LOGICIEL.

Vous n'êtes pas autorisé à communiquer des copies du PRODUIT LOGICIEL à d'autres personnes ou institutions sans l'accord préalable demandé par écrit à l'Auteur (philgeo@club-internet.fr). En revanche, il vous est recommandé de transmettre l'adresse de téléchargement du PRODUIT LOGICIEL indiquée ci-dessus à toute personne ou institution souhaitant l'utiliser.

Participation à la mise au point du PRODUIT LOGICIEL : En contrepartie des conditions très avantageuses qui vous sont offertes pour l'utilisation du PRODUIT LOGICIEL, vous vous engagez à signaler toute difficulté d'utilisation ou tout dysfonctionnement par e-mail à l'Auteur (philgeo@club-internet.fr). Ces remarques sont faites à titre gracieux sans qu'aucune rémunération ou indemnité d'aucune sorte puisse être exigée par leurs auteurs. Cette démarche n'engage nullement l'Auteur du PRODUIT LOGICIEL à procéder aux corrections ou améliorations indiquées.

Mentions de droits d'auteur : Vous n'êtes pas autorisé à supprimer ou modifier les mentions de droits d'auteur sur le PRODUIT LOGICIEL.

Services d'assistance : Aucune assistance ne peut être exigée de l'Auteur, de son employeur ou de son fournisseur d'accès pour le téléchargement. Vous pouvez toutefois contacter l'Auteur (qui vous répondra dans les limites de ses disponibilités de temps) par e-mail à l'adresse : philgeo@club-internet.fr

Respect des lois applicables : Vous devez respecter toutes les lois applicables en ce qui concerne l'utilisation du PRODUIT LOGICIEL.

Propriété : Tous les droits de propriété et notamment les droits d'auteurs relatifs au PRODUIT LOGICIEL et à toute copie de celui-ci, sont détenus par M. Philippe WANIEZ. Tous les droits de propriété intellectuelle et autres droits relatifs aux contenus auxquels le PRODUIT LOGICIEL peut donner accès, sont détenus par les propriétaires respectifs de ces contenus et peuvent être protégés par la réglementation et les traités internationaux en matière de droit d'auteur ou de propriété intellectuelle.

Exclusion de garantie : L'Auteur, son employeur et son fournisseur d'accès pour le téléchargement excluent expressément toute garantie relative au PRODUIT LOGICIEL. Le PRODUIT LOGICIEL est fourni « en l'état », sans garantie d'aucune sorte, expresse ou

implicite, notamment sans aucune garantie implicite de qualité, d'adéquation à un usage particulier ou d'absence de contrefaçon.

**Vous assumez l'ensemble des risques découlant de l'utilisation
ou des performances du PRODUIT LOGICIEL.**

Limitation de responsabilité : Dans toute la mesure permise par la réglementation en vigueur, l'Auteur, son employeur et son fournisseur d'accès pour le téléchargement ne pourront en aucun cas être tenus pour responsables de tout dommage, de quelque nature que ce soit, (notamment et de manière non limitative, toute perte de bénéfices, interruption d'activité, perte d'informations commerciales ou toute autre perte pécuniaire) résultant de l'utilisation ou de l'impossibilité d'utiliser le PRODUIT LOGICIEL ou de la fourniture ou du défaut de fourniture de services d'assistance, même si l'Auteur, son employeur et son fournisseur d'accès pour le téléchargement ont été prévenus de l'éventualité de tels dommages. En tout état de cause, la responsabilité totale de l'Auteur au titre de toute stipulation du présent contrat ne saurait excéder le montant que vous avez effectivement payé pour le PRODUIT LOGICIEL.

1. Quels types de cartes avec *Philcarto* ?

Avant d'exposer en détail les fonctions assurées par *Philcarto*, il apparaît utile de délimiter son domaine d'application. Il s'agit d'une partie seulement de l'information géographique telle que la définit J. Dangermond (fig. 1.1) dans un article fort pédagogique [Dangermond ;1990]. *Philcarto* prend en compte les seules « localisations x,y » c'est-à-dire des points, des lignes et des polygones enregistrés à l'aide de leurs coordonnées ; les localisations dites « topologiques » ne sont pas reconnues. Les données statistiques pouvant être analysées avec *Philcarto* sont de type « classes » et « valeurs ». La prise en compte directe du temps est possible si le maillage de l'espace est considéré comme invariant (ce qui pose des problèmes parfois insolubles lorsque ces maillages se transforment dans le temps). En revanche, la cartographie de taux ou d'indices de variation dans le temps ne présente aucune difficulté sur un maillage géographique constant dans le temps.

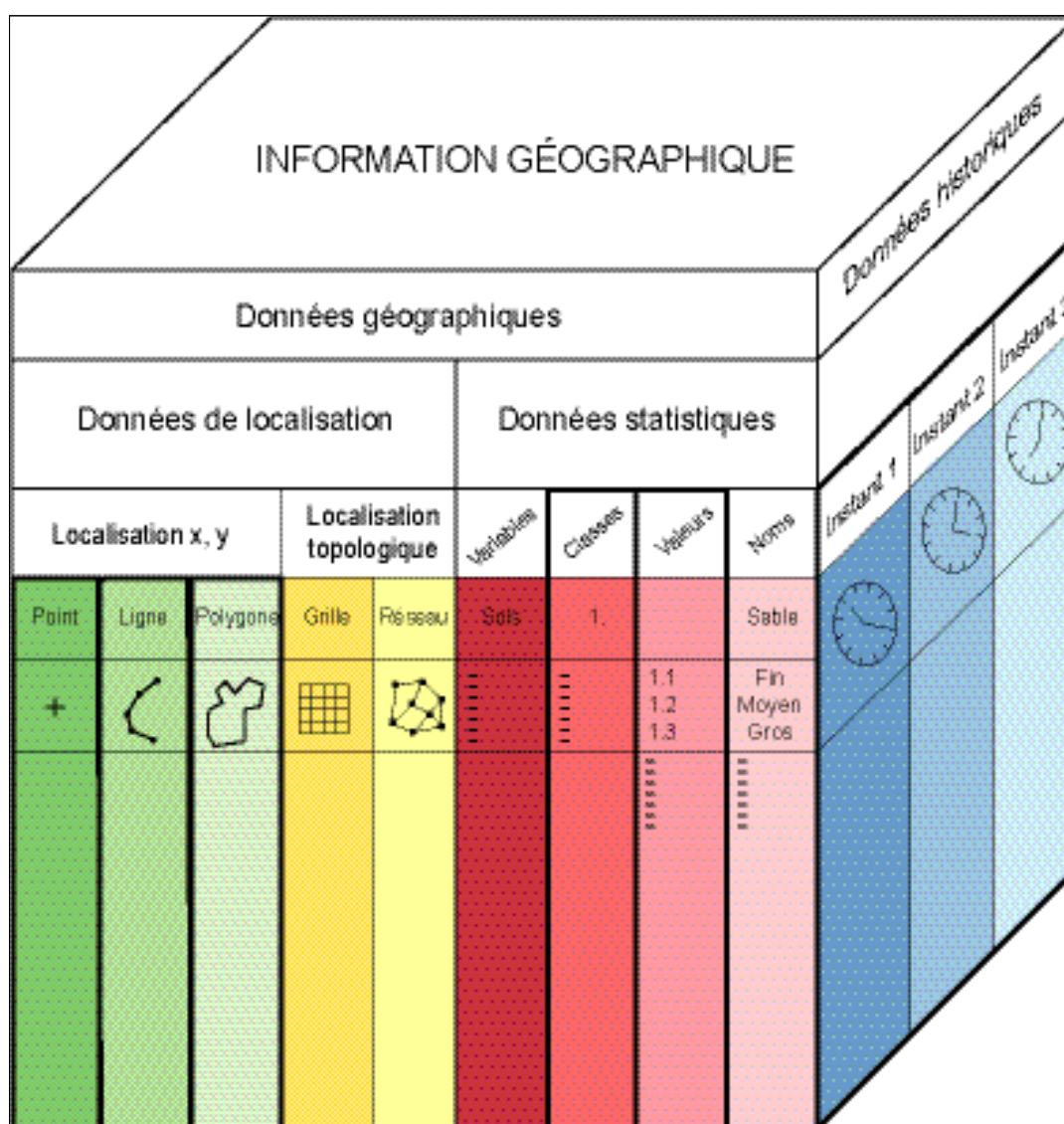


Figure 1.1 : L'information géographique selon J. Dangermond.

Comme tout SIG, *Philcarto* exige que deux types d'informations lui soient communiquées en entrée : un fond de carte représentant les unités spatiales (ou unités de recensement, ou unités statistiques) et un tableau de données statistiques se rapportant à ces unités spatiales. Celles-ci sont figurées soit par des surfaces, soit par des lignes, soit par des points. Les données statistiques représentent des effectifs, des mesures, des catégories ou des classes. Les objectifs du travail cartographique étant définis, les caractéristiques du fond de carte et celles des données statistiques qui lui sont associées conduisent à faire appel à tel ou tel autre type de carte. L'exposé qui suit présente au moins un exemple pour chaque type de carte pouvant être réalisée avec *Philcarto*. Les fichiers ayant permis de réaliser ces cartes sont fournis (sauf lorsqu'un *copyright* en limite la diffusion) ; les noms des dossiers les contenant sont indiqués sur fond gris clair. Le lecteur peut ainsi reproduire lui-même ces cartes, et même en réaliser d'autres avec *Philcarto*.

Outre leur qualité graphique, ces documents montrent aussi la capacité de *Philcarto* à produire des cartes à différentes échelles, depuis le pays (Afrique du Sud, Brésil, Allemagne, France), à la région (Région Rhône-Alpes), et jusqu'à la ville (New-York).

1.1. Les cartes choroplèthes

Les cartes en surfaces colorées (ou en niveaux de gris) scientifiquement dénommées « cartes choroplèthes » permettent de représenter une variable numérique continue ou discrète, à l'exclusion des quantités ou des effectifs. Ce type de carte nécessite un maillage de la surface de la région étudiée en unités spatiales surfaciques représentées par des polygones. Afin de traduire, sur le plan géographique, les variations des données étudiées, chaque polygone est colorié. Dans ce but, différentes méthodes de discrétisation permettent le découpage en classes des variables continues afin d'affecter à chaque classe une couleur sélectionnée dans une gamme continue de couleurs chaudes, froides ou opposées. Une gamme « perroquet » traduit les valeurs des variables nominales. Il est fort utile de faire figurer en légende un histogramme qui présente l'étendue des classes vis-à-vis de la distribution statistique.

La carte de la population de race noire dans les districts sud-africains en 1991 (fig. 1.2) recourt à une gamme de couleurs variant du bleu clair (faibles pourcentages de noirs) au marron (très fort pourcentages). Elle montre que, à l'échelle du pays considéré dans son ensemble, l'apartheid n'était pas seulement un système politique et social, mais aussi un fait géographique fondamental de l'organisation du territoire.

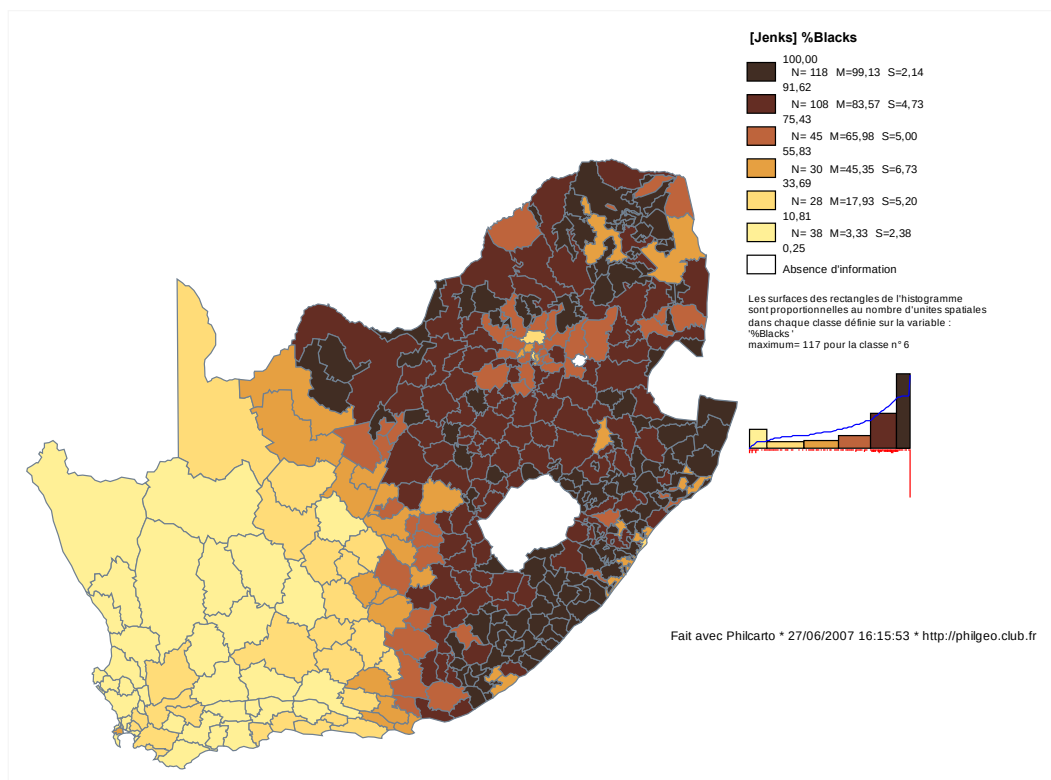


Figure 1.2 : Afrique du Sud (districts).
 % de la population de race noire dans la population totale en 1991.
 Fichiers : \Philcarto exemples\Afrique du Sud\

La carte du taux de variation moyen annuel de la population des communes de la région Rhône-Alpes (fig. 1.3) utilise une gamme de couleurs opposées pour exprimer les pertes ou les gains de population sur la période 1975-1990 (Source : INSEE, RGP 1975 et 1990) : le bleu foncé traduit de fortes diminutions relativement à la population de chaque commune, alors que le rouge vif exprime d'importants gains relatifs. On observe sur la carte l'exode rural qui affecte principalement les communes montagnardes de l'Ouest de l'Ardèche, de la Loire et des contreforts des Alpes. La croissance de l'agglomération lyonnaise est particulièrement nette : alors que la ville centre et les communes de la première couronne (Vénissieux, Bron...) perdent une partie de leurs habitants, l'accroissement est particulièrement fort à l'Est, sur les départements de l'Ain et de l'Isère. Un phénomène analogue, quoique moins marqué est visible à Roanne, au Nord de la Loire, à Chambéry en Savoie, à Annecy en Haute-Savoie, ou à Grenoble en Isère.

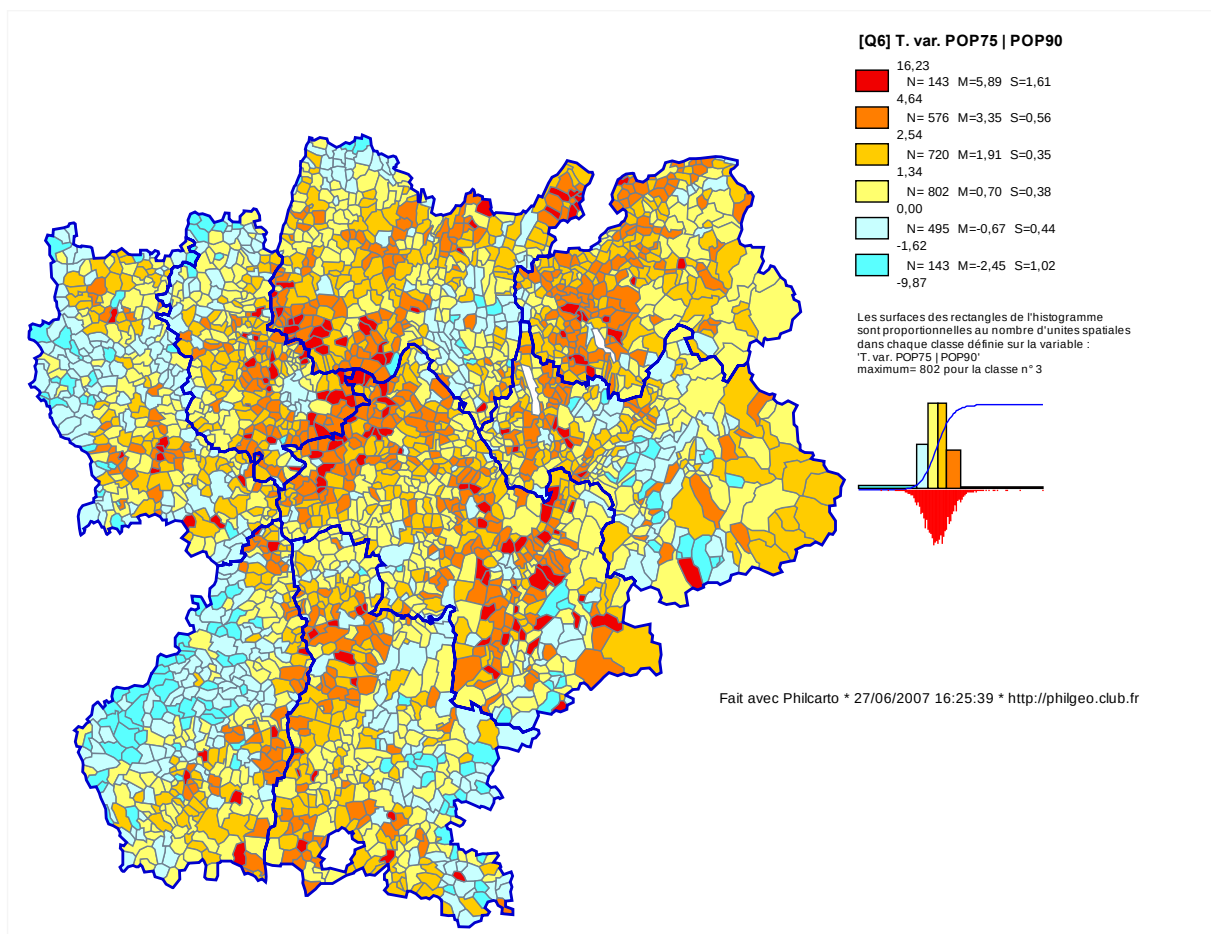


Figure 1.3 : Région Rhône-Alpes (communes).
Variation moyenne annuelle de la population pour la période 1975-1990.
Fichiers : IGN / GéoFLA, fichiers non fournis.

La carte « ethnique » de New-York (île de Manhattan, fig. 1.4) utilise une gamme de couleurs « perroquet » pour la cartographie des classes issues d'une classification ascendante hiérarchique ; ces classes ne sont pas ordonnées mais traduisent divers « dosages » de populations différentes, permettant ainsi de qualifier les quartiers en fonction de leur profil (blanc, noir, hispanique...).

Réalisée sur 6 variables directement par *Philcarto* (aucun logiciel de traitement des données statistiques supplémentaire n'a été utilisé pour réaliser ce travail) la classification prend en compte les populations blanche, noire, amérindienne, asiatique, hispanique et non hispanique. Les données et le fond de carte correspondent au recensement démographique de 1990 diffusé sur Internet par CIESIN / SEDAC à l'adresse : <http://sedac.ciesin.columbia.edu/plue/>

Les profils moyens des classes calculés et dessinés par *Philcarto* montrent la nette prédominance des blancs non hispaniques, dans les trois quarts Sud de la ville (classes 9 et 3), sauf au Sud-Est où les asiatiques (classes 2 et 6) sont mieux implantés (Chinatown). Au Nord de l'île, Harlem, le quartier noir (classe 1) s'individualise nettement, flanqué au Nord, à l'Est et à l'Ouest de quartiers hispaniques.

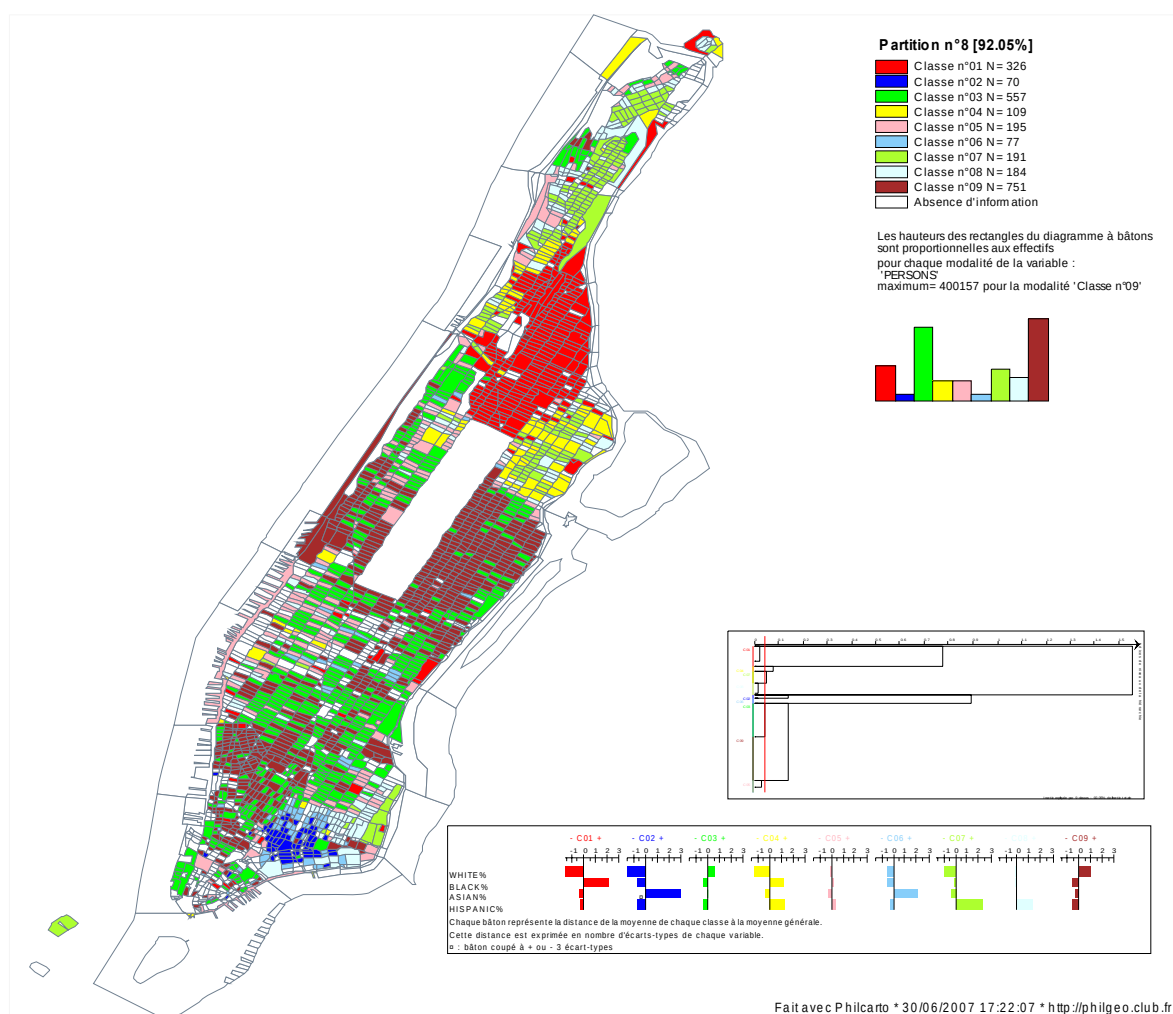


Figure 1.4 : New-York (îlots urbains de Manhattan). Classification « ethnique » en 1990.

Fichiers : \Philcarto exemples\Manhattan\

Les trois exemples ci-dessus montrent l'importance du choix d'une gamme de couleurs appropriée au type de données à cartographier. Il en est de même pour les autres types de cartes lorsqu'ils font appel à une gamme de couleurs destinée à traduire une variation dans l'espace.

Les cartes choroplèthes

Le terme « choroplèthe » est forgé à partir du grec *chorê* qui évoque l'espace, et de *plethos*, qui exprime la multitude, le nombre...

Une carte choroplèthe représente la variation d'une variable continue (pourcentages, taux, mesures...) ou discrète (catégories numériques ou nominales), à l'exclusion des quantités ou des effectifs.

Le maillage est surfacique ; les limites de ses éléments sont enregistrées sous forme de polygones. Un second maillage surfacique ou linéaire peut être superposé au premier pour former un habillage qui complète la carte et en facilite la lecture.

Les variables continues doivent être discrétisées afin d'affecter une couleur à chaque classe. Plusieurs techniques de discrétisation permettent de prendre en compte la forme des distributions statistiques.

Les gammes de couleurs ne comprennent qu'un petit nombre de paliers (une dizaine au plus). Elles sont ordonnées pour exprimer le chaud, le froid, ou l'opposition du chaud et du froid lorsque le phénomène étudié diverge à partir d'une certaine valeur. La gamme de couleurs dite « perroquet », qui n'est pas ordonnée, permet de représenter les modalités des variables discrètes. Pour l'impression noir et blanc, on dispose de gammes ordonnées en niveaux de gris ou en trames (lignes ou points).

Chaque polygone composant le maillage surfacique est colorié, ce qui a pour résultat de traduire graphiquement la variation géographique de la variable représentée.

La légende se compose de caissons colorés permettant de qualifier chaque couleur par un intervalle de valeurs ou une modalité ; un histogramme (pour les variables continues), ou un diagramme à bâtons (pour les variables discrètes) complété d'une courbe des fréquences cumulées permettent d'apprécier la distribution des valeurs.

L'inconvénient principal de ce type de carte réside dans le risque d'une mauvaise interprétation du phénomène cartographié lorsqu'existent de grandes différences de surfaces entre les polygones du fond de carte, différences qui peuvent conduire à surévaluer le poids des plus grands de ces polygones.

1.2. Les cartes en cercles proportionnels

Les cartes en cercles proportionnels sont destinées à représenter des quantités ou des effectifs. L'emploi de ce mode de représentation est plus délicat qu'il n'y paraît au premier abord : un calibrage des cercles s'avère nécessaire dans la plupart des cas, et plusieurs essais sont souvent indispensables pour obtenir un résultat satisfaisant. Une partie de la difficulté cartographique réside parfois dans la trop forte densité des unités spatiales provoquant des chevauchements qu'il faut gérer au mieux : les cercles les plus petits sont situés devant les cercles plus grands et un détournage (une circonférence tracée en blanc) permet de distinguer les petits cercles placés sur les grands. Ce mode de représentation n'est donc véritablement efficace que lorsque les unités spatiales ne sont pas en trop grand nombre (quelques centaines au maximum), et qu'elles sont réparties dans l'espace de manière relativement homogène.

Un exemple de ce mode de représentation, classique en géographie de la population, est celui de la population du Viêt-Nam en 1989 (fig. 1.5).

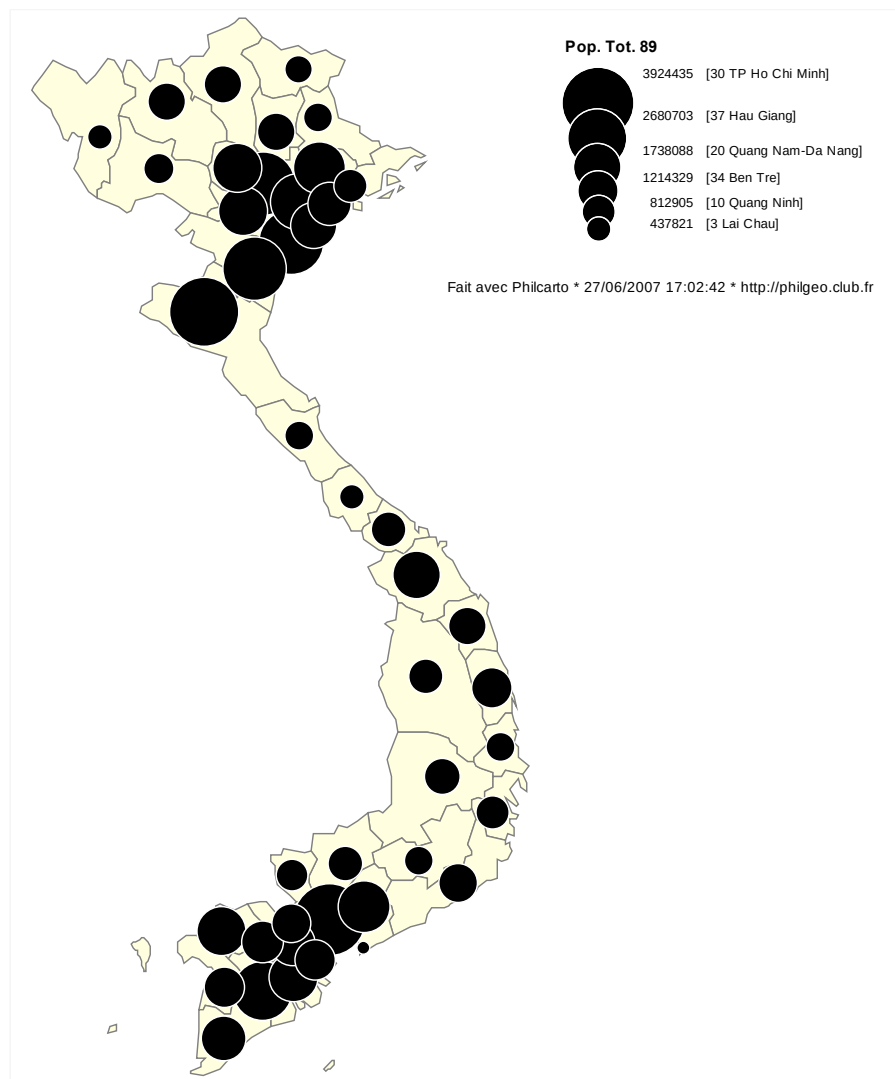


Figure 1.5 : Viêt-Nam (provinces). Population totale en 1989.

Fichiers : \Philcarto exemples\Viêt-Nam\

Dans le cas de la carte des provinces du Viêt-Nam, les centres des cercles sont des points placés à l'intérieur de chaque polygone représentant une province. Il y a donc un point et un seul destiné à représenter les effectifs dans chaque province. La localisation des points peut avoir une raison géographique (chef-lieu...), ou simplement géométrique (au centre du polygone). Mais *Philcarto* peut parfaitement réaliser des cartes en cercles proportionnels en l'absence d'éléments cartographiques surfaciques, ou avec de tels éléments réduits à un rôle d'habillage.

La carte de la population des villes européennes (et de quelques villes d'Afrique du Nord) de 200 000 habitants et plus contient bien plus de cercles représentant la population des villes, qu'il n'y a de polygones représentant les États (fig. 1.6). Ces derniers ne jouent d'ailleurs aucun rôle dans la construction de la carte par *Philcarto* ; ils facilitent simplement la lecture en apportant des éléments contextuels complémentaires.

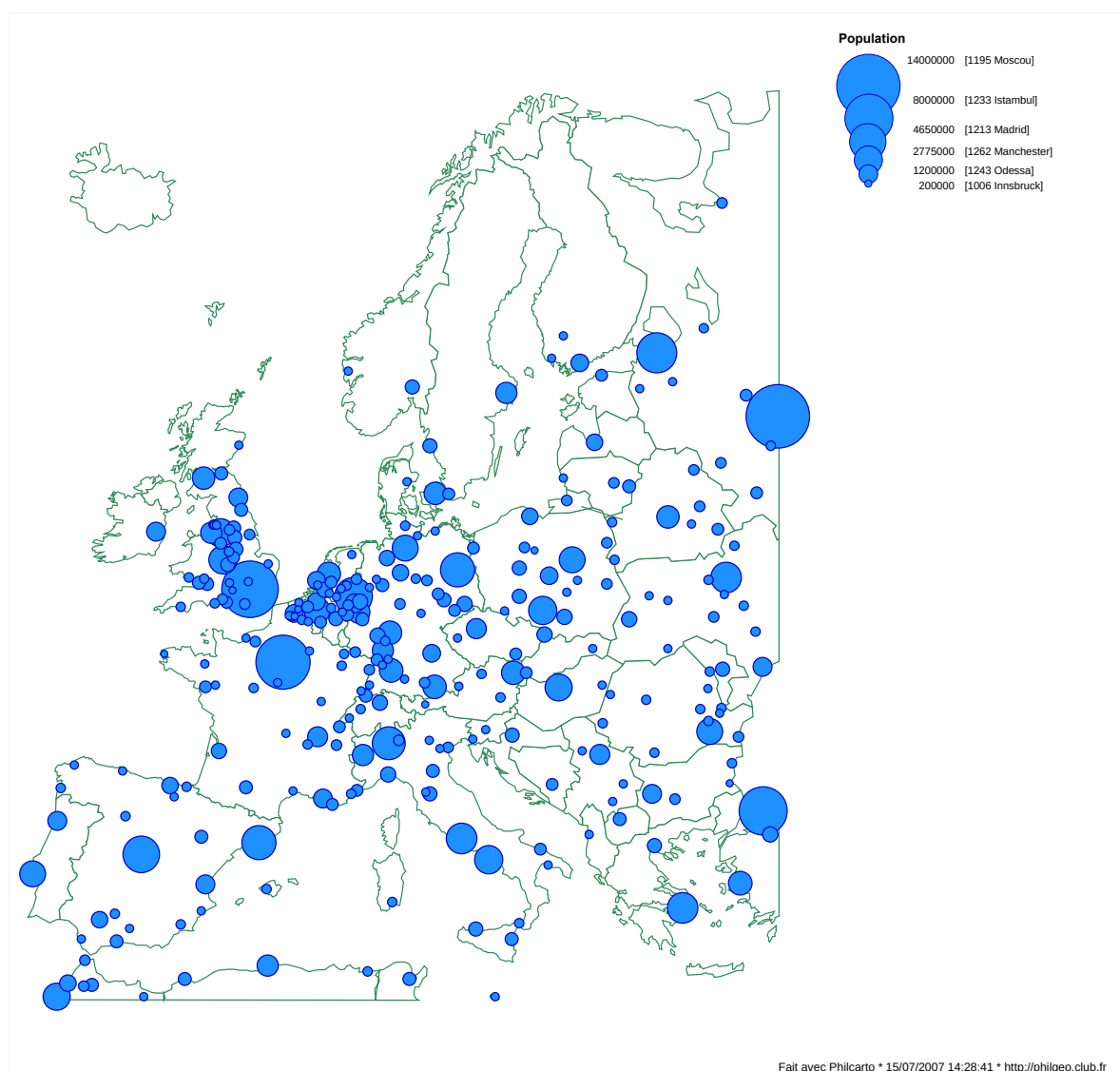


Figure 1.6 : Europe (villes). Population totale durant les années 1990-1995.

Fichiers : \Philcarto exemples\ Europe villes\

Les cartes en cercles proportionnels

Une carte en cercles proportionnels représente des quantités ou des effectifs par un cercle plein (disque) dont la surface est proportionnelle à ces quantités ou à ces effectifs, à l'exclusion des pourcentages, taux, mesures ou catégories.

Le maillage est ponctuel ; il s'agit de lieux de coordonnées X,Y, ou bien d'un point caractéristique de chaque polygone (centroïde, chef lieux...) d'un maillage surfacique. Un second maillage surfacique ou linéaire peut être superposé au maillage ponctuel pour former un habillage qui complète la carte et en facilite la lecture.

Les cercles doivent être préalablement calibrés ; cette opération consiste à choisir le diamètre du plus grand cercle : souvent, plusieurs essais sont nécessaires pour obtenir un résultat satisfaisant, c'est-à-dire pour que le plus grand cercle n'occupe pas une surface trop étendue, risquant en cela de masquer les autres cercles, sans pour autant que le plus petit cercle disparaisse.

Les cercles peuvent être remplis par une couleur unique (disques colorés) ou demeurer vides (seules les circonférences sont tracées). Ils sont dessinés à partir du plus grand jusqu'au plus petit afin que ces derniers soient situés au-dessus des plus grands pour ne pas être occultés. Dans le cas courant des disques, une circonférence blanche (parfois colorée) permet de détourner chacun d'eux afin de les individualiser les uns par rapport aux autres dans le cas de superpositions.

La légende se compose d'un échantillon de cercles qui correspondent à des valeurs effectivement présentes sur la carte, ceci afin de permettre au lecteur d'évaluer les quantités représentées en chaque point.

L'inconvénient principal de ce type de carte réside dans le fréquent chevauchement des cercles qui doit être géré au mieux par un calibrage adéquat, par le recours au détourage et même, exceptionnellement, par le déplacement de certains points du fond de carte lorsque cela ne risque pas d'entraîner des erreurs de localisation (ceci est possible sans grand risque lorsque les points sont des centroïdes de polygones).

1.3. Les cartes en cercles proportionnels colorés

Les cartes en cercles proportionnels peuvent donner lieu à diverses variations qui permettent d'en enrichir le contenu, sans pour autant en lever les limitations. Les cartes en cercles proportionnels colorés permettent de représenter simultanément des quantités ou des effectifs (cercles) et des valeurs numériques continues ou discrètes exprimant des pourcentages, des mesures, etc. (couleurs des cercles), mais pas des effectifs.

La proportion de résidences principales équipées de WC extérieurs dans les départements français en 1991 (fig. 1.7) est un bon indicateur du degré de vétusté des logements (mais aussi de certains traits culturels qui se traduisent par une forte proportion de maisons individuelles dotées d'un édicule situé au fond du jardin...). C'est dans le Nord du pays que cette situation est la plus nette avec plus de 13% des logements sans WC intérieurs. Le grand Ouest n'apparaît pas non plus en très bonne situation, même si dans le détail certains départements touristiques (Vendée, Loire-Atlantique, Finistère) sont mieux équipés que les autres. En revanche, la partie centrale du Bassin Parisien, l'Est et le pourtour méditerranéen sont mieux équipés.

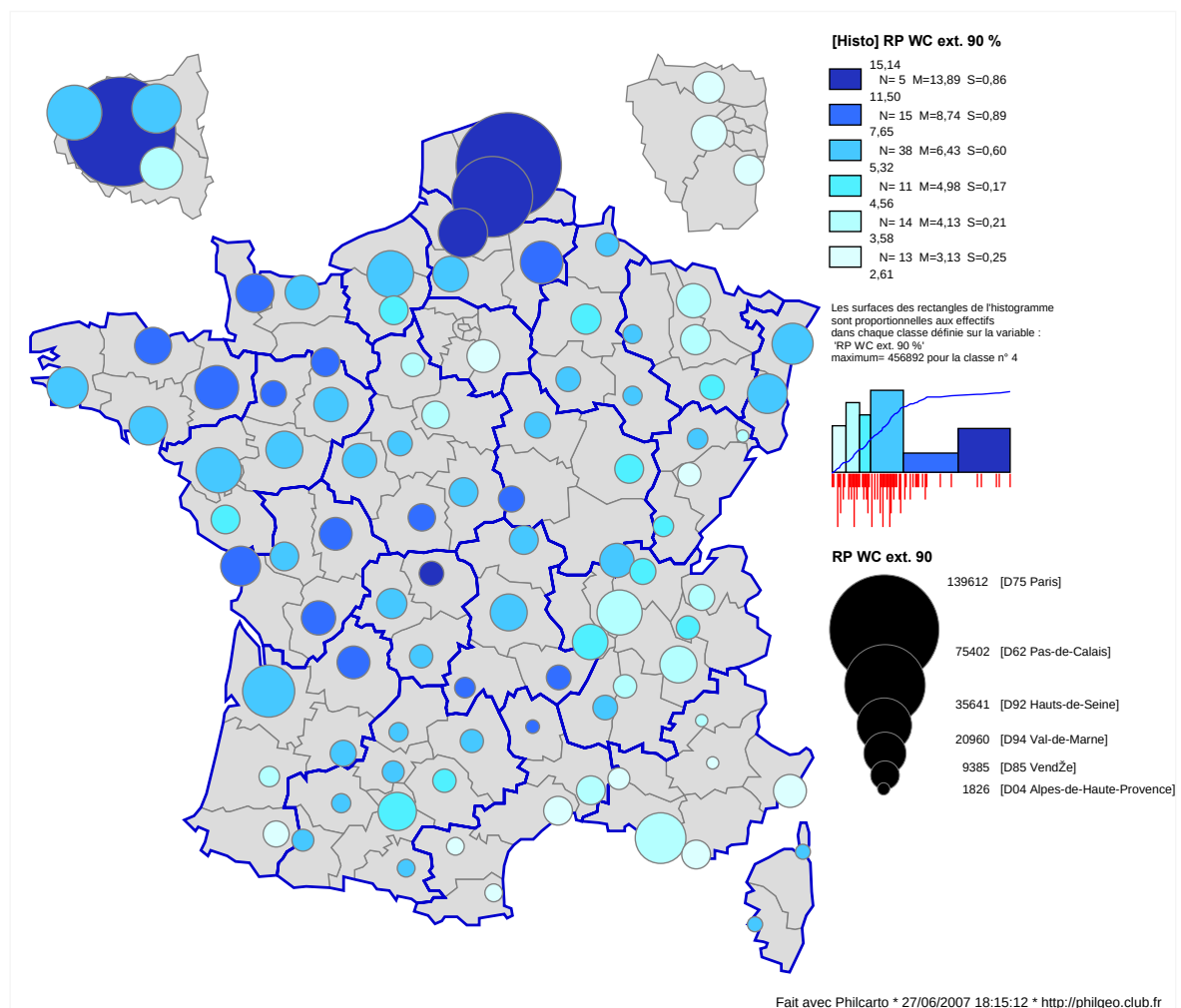


Figure 1.7 : France (départements).
Résidences principales avec WC extérieurs en 1991.
Fichiers : \Philcarto exemples\ France départements\

Les cartes en cercles proportionnels colorés

Une carte en cercles proportionnels colorés représente simultanément deux variables. Les surfaces des cercles (disques) sont proportionnelles aux valeurs d'une première variable exprimant des quantités ou des effectifs. Les couleurs des disques correspondent aux valeurs de la seconde variable, soit continue (pourcentages, taux, mesures...) et donc préalablement discrétisée, soit discrète (catégories numériques ou nominales).

Le maillage est ponctuel ; il s'agit de lieux de coordonnées X,Y, ou bien d'un point caractéristique de chaque polygone (centroïde, chef lieux...) d'un maillage surfacique. Un second maillage surfacique ou linéaire peut être superposé au maillage ponctuel pour former un habillage qui complète la carte et en facilite la lecture.

Le calibrage des cercles et la discrétisation des variables doivent être réalisés dans les mêmes conditions que pour les cartes en cercles proportionnels et choroplèthes.

La légende comprend deux parties. Dans sa partie supérieure, des caissons colorés permettent de qualifier chaque couleur par un intervalle de valeurs ou une modalité ; un histogramme (pour les variables continues), ou un diagramme à bâtons (pour les variables discrètes) complété d'une courbe des fréquences cumulées permettent d'apprécier la distribution des valeurs. Au-dessous, figure un échantillon de cercles qui correspondent à des valeurs effectivement présentes sur la carte, ceci afin de permettre au lecteur d'évaluer les quantités représentées en chaque point.

L'inconvénient principal de ce type de carte provient du chevauchement des cercles qui gêne parfois la lecture des couleurs; de plus, les couleurs pâles rendent les cercles peu visibles (cet inconvénient peut être contourné en partie par l'emploi d'un fond coloré sombre).

1.4. Les cartes en cercles proportionnels sur surfaces colorées

Comme les précédentes, les cartes en cercles proportionnels sur surfaces colorées permettent aussi de représenter simultanément des quantités ou des effectifs (cercles) et des valeurs numériques continues ou discrètes (couleurs des surfaces). Leur construction est encore plus sensible à la dimension des unités spatiales et à leur répartition dans l'espace. En effet, il faut éviter d'une part que les cercles se chevauchent, et d'autre part qu'ils recouvrent entièrement les surfaces colorées situées au-dessous d'eux (il est aussi possible de tracer seulement les circonférences de cercles, mais la lecture de la carte s'avère plus difficile). Dans la plupart des cas, une intervention du graphiste sur le placement des cercles est inévitable lors de la mise en page finale.

La carte de la population « hispanique » aux États-Unis en 1996 (fig. 1.8) souligne l'importance de la Californie et du Texas, mais c'est au Nouveau-Mexique, sans doute pour des raisons historiques, que la proportion d'hispaniques est la plus élevée, toutefois avec un effectif beaucoup plus faible que les précédents. On observe d'autres implantations importantes, quoique secondaires, en Floride (immigration cubaine), et dans les régions industrielles de l'Illinois, de l'État de New-York et du New-Jersey.

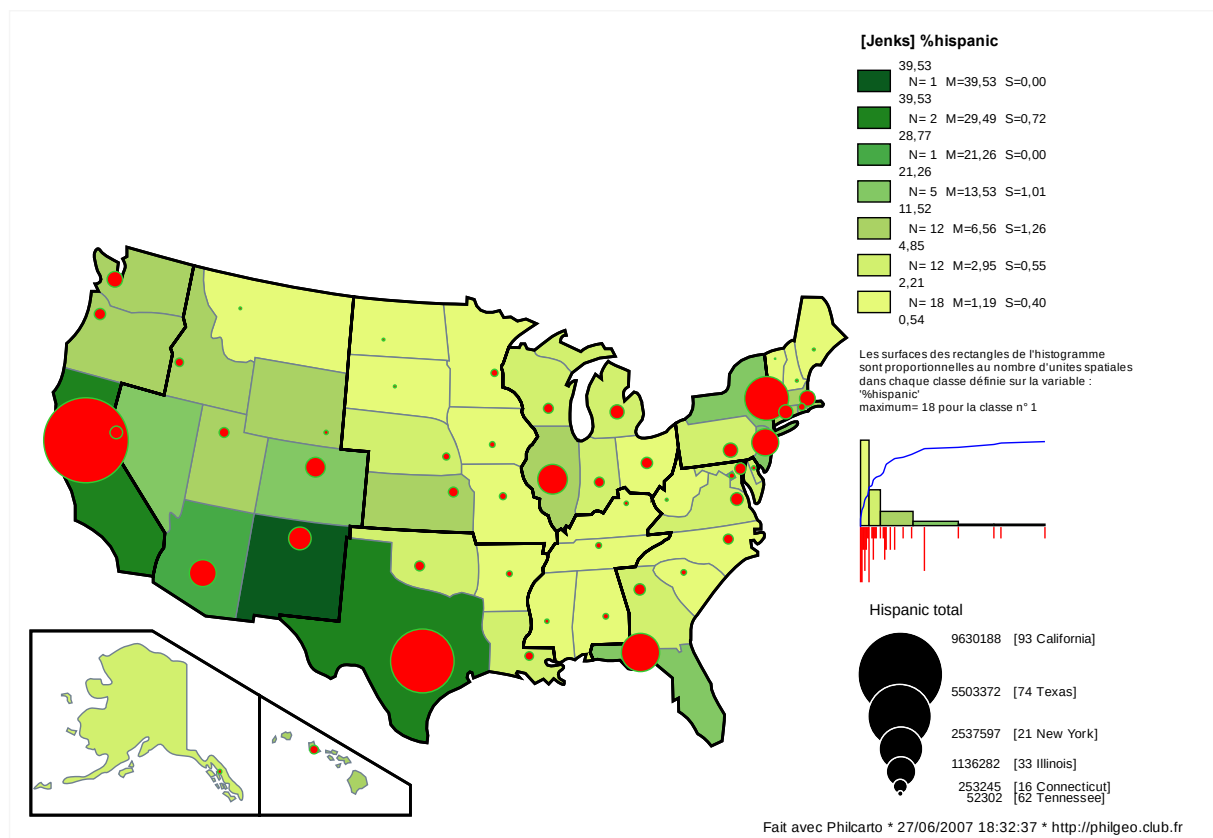


Figure 1.8 : États-Unis (États). Population « hispanique » en 1996.

Fichiers : \Philcarto exemples\Etats Unis\

Les cartes en cercles proportionnels sur surfaces colorées

Il s'agit de la superposition d'une carte en cercles proportionnels à une carte choroplèthe.

Une carte en cercles proportionnels sur surfaces colorées représente simultanément deux variables. Les surfaces des cercles (disques) sont proportionnelles aux valeurs d'une première variable évaluant des quantités ou des effectifs. Les couleurs des polygones correspondent aux valeurs de la seconde variable, soit continue (pourcentages, taux, mesures...) et donc préalablement discrétisée, soit discrète (catégories numériques ou nominales).

Les cercles sont localisés sur un maillage est ponctuel ; il s'agit de lieux de coordonnées X,Y, ou bien d'un point caractéristique de chaque polygone (centroïde, chef lieux...) d'un maillage surfacique. Les surfaces sont délimitées par les polygones formant le maillage surfacique. Un troisième maillage surfacique ou linéaire peut être superposé aux deux maillages précédents pour former un habillage qui complète la carte et en facilite la lecture.

Le calibrage des cercles et la discrétisation des variables doivent être réalisés dans les mêmes conditions que pour les cartes en cercles proportionnels et que pour les cartes choroplèthes.

La légende comprend deux parties. Dans sa partie supérieure, des caissons colorés permettent de qualifier chaque couleur par un intervalle de valeurs ou une modalité ; un histogramme (pour les variables continues), ou un diagramme à bâtons (pour les variables discrètes) complété d'une courbe des fréquences cumulées permettent d'apprécier la distribution des valeurs. Au-dessous, figure un échantillon de cercles qui correspondent à des valeurs effectivement présentes sur la carte, ceci afin de permettre au lecteur d'évaluer les quantités représentées en chaque point.

Les inconvénients principaux de ce type de carte sont naturellement ceux, cumulés, des cartes choroplèthes et des cartes en cercles proportionnels ; à ceux-ci s'ajoute la difficulté de lire les couleurs situées sous les disques lorsqu'ils occultent la majeure partie de la surface des polygones (problème pouvant être contourné par le recours à des circonférences au lieu des disques).

1.5. Les cartes en demi-cercles affrontés

Les cartes en demi-cercles affrontés permettent de comparer deux effectifs ou deux quantités se rapportant à deux catégories d'une même population, ou bien à la même catégorie à deux moments donnés de son histoire. La perception des différences impose aux cercles un diamètre assez grand, ce qui limite leur nombre de manière drastique.

La carte de l'emploi agricole et industriel dans les régions italiennes en 1998 (fig. 1.9) est un modèle du genre : le nombre de régions est petit (19 au total) et les différences au sein d'une même région sont marquées, même si l'emploi industriel est partout plus important que l'emploi agricole. En Lombardie, le rapport industrie/agriculture s'élève à 14,5 alors qu'en Sicile, il n'atteint que 1,5. La carte montre sans ambiguïté la diminution de ce rapport du Nord du pays au Sud tout en soulignant le déséquilibre géographique du nombre total des emplois.

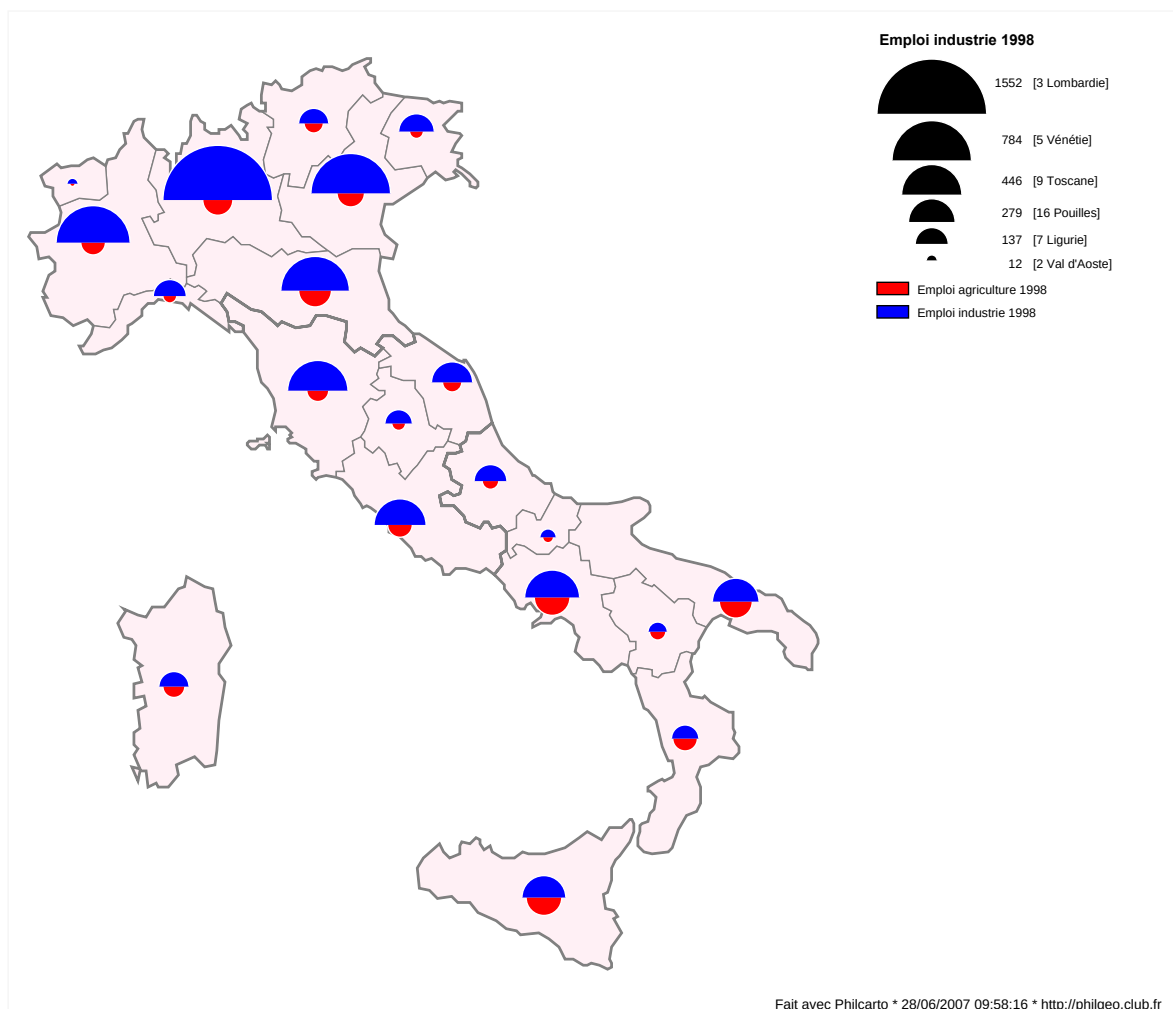


Figure 1.9 : Italie (régions). Emploi agricole et industriel en 1998.

Fichiers : \Philcarto exemples\Italie\

Les cartes en demi-cercles affrontés

Une carte en demi-cercles affrontés représente simultanément deux variables portant sur des quantités ou des effectifs, à l'exclusion des pourcentages, taux, mesures ou catégories.

Les cercles sont localisés sur un maillage ponctuel ; il s'agit de lieux de coordonnées X,Y, ou bien d'un point caractéristique de chaque polygone (centroïde, chef lieux...) d'un maillage surfacique. Un second maillage surfacique ou linéaire peut être superposé au maillage précédent pour former un habillage qui complète la carte et en facilite la lecture.

Les cercles doivent être préalablement calibrés; cette opération consiste à choisir le diamètre du plus grand demi-cercle correspondant au maximum des deux variables : souvent, plusieurs essais sont nécessaires pour obtenir un résultat satisfaisant, c'est-à-dire pour que le plus grand demi-cercle n'occupe pas une surface trop étendue, risquant en cela de masquer les autres demi-cercles, sans pour autant que le plus petit demi-cercle disparaisse.

La légende comprend deux parties. En haut figure un échantillon de demi-cercles qui correspondent à des valeurs effectivement présentes sur la carte et relatives à la variable contenant le maximum, ceci afin de permettre au lecteur d'évaluer les quantités représentées en chaque point. Au-dessous, deux caissons colorés permettent d'identifier chacune des deux variables.

Les inconvénients principaux de ce type de carte sont naturellement ceux des cartes en cercles proportionnels ; à ceux-ci s'ajoute l'obligation d'un diamètre assez grand pour permettre la perception des différences entre les demi-cercles affrontés, ce qui limite leur nombre de manière drastique. Ce type de carte ne doit donc pas être réalisé sur des fonds de cartes composés d'un grand nombre d'éléments cartographiques.

1.6. Les cartes en diagrammes à secteurs

Le lecteur doit être d'emblée mis en garde : l'utilisation abusive de ce type de carte, c'est-à-dire lorsque le nombre d'éléments cartographiques est grand (au-delà d'une trentaine), ne peut conduire qu'à des résultats peu intelligibles car la multiplication des diagrammes n'autorise pas la perception globale du phénomène cartographié. Dans le même ordre d'idée, il est souhaitable de ne pas dépasser 4 à 5 secteurs par cercle (4 à 5 catégories).

Les cartes en diagrammes à secteurs permettent de visualiser la proportion dans l'effectif total de catégories composant une population. Les variables correspondent à chaque catégorie dont on connaît la population ou l'effectif ; les effectifs de ces populations doivent donc pouvoir être additionnés. Lorsqu'on a affaire à deux catégories seulement, il est préférable de réaliser une carte en cercles affrontés. Le diamètre des cercles correspond à la somme des effectifs dans chaque catégorie.

La carte de l'emploi agricole, industriel et autres dans les régions italiennes en 1998 (fig. 1.10) montre clairement à la fois le volume d'emplois plus élevé dans le Nord que dans le Sud, et la composition de ces emplois proche de 50% dans l'industrie au Nord, alors qu'au Sud, les autres emplois (tertiaires essentiellement) apparaissent largement majoritaires. Partout, l'agriculture ne représente plus qu'une petite fraction, plus importante au Sud qu'au Nord.

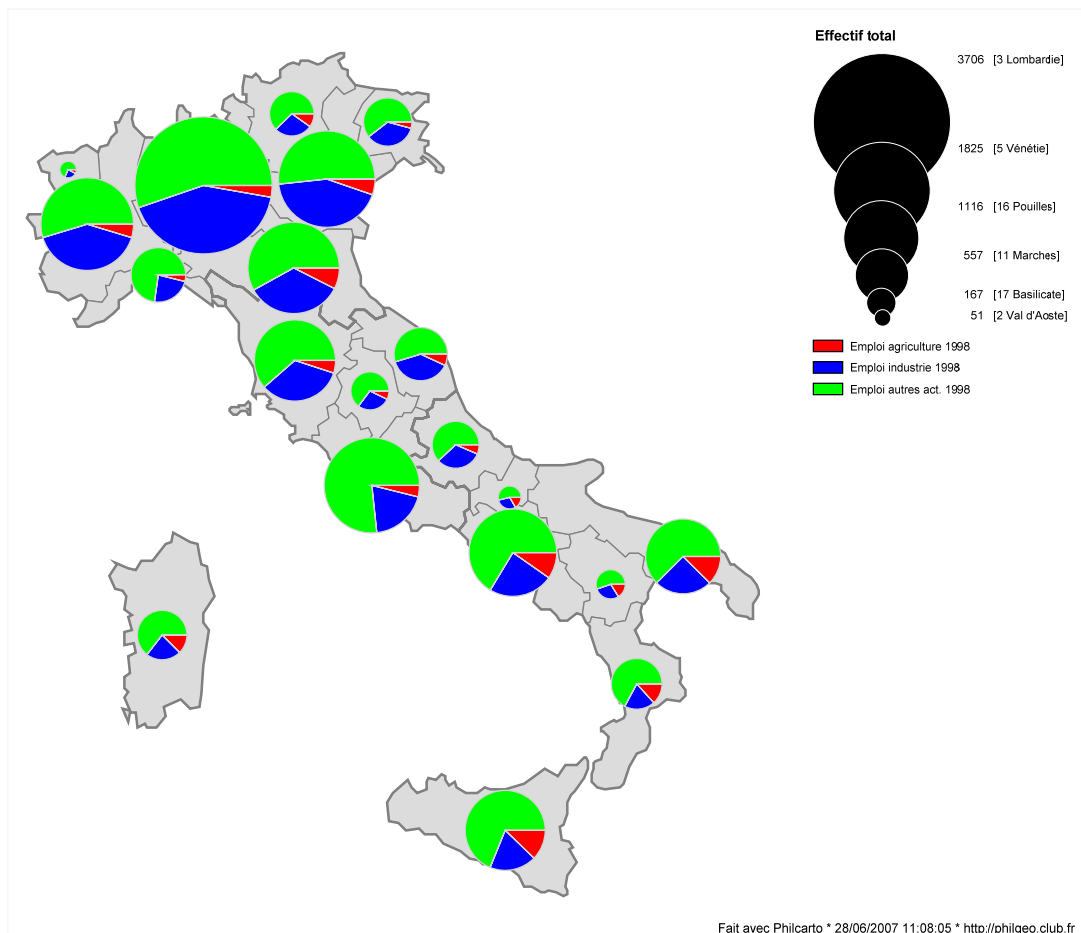


Figure 1.10 : Italie (régions). Emploi agricole, industriel et tertiaire en 1998.

Fichiers : \Philcarto exemples\Italie\

Les cartes en diagrammes à secteurs

Une carte en diagrammes à secteurs (figure désignée aussi par les noms « camembert » ou « pizza ») permet de représenter la composition d'une population en fonction de plusieurs catégories, chacune d'elles, enregistrée dans une variable, portant sur des quantités ou des effectifs, à l'exclusion des pourcentages, taux, mesures ou catégories. Le cercle représente la somme des effectifs des catégories de populations relatives à chaque élément cartographique ; pour que le résultat soit valide, il faut donc que ces effectifs puissent être sommés (ne pas additionner des kilos de carottes et de pommes de terre...).

Les cercles sont localisés sur un maillage ponctuel ; il s'agit de lieux de coordonnées X,Y, ou bien d'un point caractéristique de chaque polygone (centroïde, chef lieux...) d'un maillage surfacique. Un second maillage surfacique ou linéaire peut être superposé au maillage précédent pour former un habillage qui complète la carte et en facilite la lecture.

Les cercles doivent être préalablement calibrés ; cette opération consiste à choisir le diamètre du plus grand cercle : souvent, plusieurs essais sont nécessaires pour obtenir un résultat satisfaisant, c'est-à-dire pour que le plus grand cercle n'occupe pas une surface trop étendue, risquant en cela de masquer les autres cercles, sans pour autant que le plus petit cercle disparaisse.

La légende comprend deux parties. En haut figure un échantillon de cercles qui correspondent à des valeurs effectivement présentes sur la carte et relatives à la somme des catégories de population, ceci afin de permettre au lecteur d'évaluer les quantités représentées en chaque point. Au-dessous, plusieurs caissons colorés permettent d'identifier chacun des secteurs.

Les inconvénients principaux de ce type de carte sont naturellement ceux des cartes en cercles proportionnels ; à ceux-ci s'ajoute la quasi-impossibilité d'une lecture globale de la carte et de la recherche d'une organisation générale ; chaque diagramme doit être lu l'un après l'autre, ce qui limite leur nombre de manière drastique. Ce type de carte ne doit donc pas être réalisé sur des fonds de cartes composés d'un grand nombre d'éléments cartographiques.

1.7. Les cartes en plages de niveaux et en bloc-diagramme

Les cartes en plages de niveaux sont réalisées à l'aide d'un ensemble de points cotés permettant d'estimer des valeurs en tout point de la surface cartographiée, en faisant l'hypothèse de la continuité du phénomène étudié dans l'espace. Le résultat est une carte lissée donnant une image souvent suggestive des principaux traits de la structuration spatiale de la variable représentée.

D'un point de vue technique, il s'agit de cartes choroplèthes où chaque élément cartographique est un carreau sur lequel sont estimées des valeurs à partir des points de mesure. Plus la résolution du carroyage est fine, meilleurs sont les lissages obtenus, et plus long est le temps de calcul...

Deux techniques peuvent être utilisées pour estimer les valeurs pour tout carreau du carroyage. Dans la « triangulation de Delaunay », les points cotés permettent d'établir un maillage triangulaire dans lequel chaque point est le sommet d'un ou de plusieurs triangles. Les niveaux sont ensuite interpolés linéairement sur la longueur de chaque côté de chaque triangle. Puis des plages de niveaux sont tracées, en fonction de l'espacement choisi dans l'unité de mesure de la variable cartographiée (par exemple de 10 en 10), ceci de manière à recouper les côtés des triangles à l'intersection correspondant aux valeurs interpolées sur ces côtés.

Le « krigeage » est une autre manière d'estimer les valeurs en tout point de la surface. Cette méthode géostatistique, qui fait appel à des fonctions mathématiques assez complexes, est plus délicate à manier, mais donne souvent des résultats très expressifs.

Dans les deux cas, les estimations peuvent aussi être visualisées sous la forme d'un bloc diagramme qui montre, en perspective, les iso-surfaces, ce qui conduit à un « paysage statistique », parfois suggestif. Attention ! Il ne faut pas abuser de ce genre de représentation, notamment en raison de sa très grande dépendance de l'angle de vue choisi (fig. 1.13).

Pour illustrer cette méthode, examinons la carte choroplèthe du pourcentage d'abstention au référendum du 24 septembre 2000 dans les arrondissements de Paris et les communes des départements de la première couronne. Rappelons qu'il s'agissait d'une modification de la Constitution de la Vème République réduisant le mandat du Président de la République de 7 à 5 ans. Ce référendum n'avait pas soulevé l'enthousiasme des électeurs et l'abstention avait été très forte.

La carte de l'abstention (fig. 1.11) montre les clivages habituels de l'expression électorale dans la capitale : la moitié Ouest, politiquement à droite, et favorable à Jacques Chirac, s'est peu abstenue. A l'opposé, les électeurs de la périphérie (sauf à l'Ouest), se sont bien peu rendus aux urnes, avec des pourcentages d'abstention dépassant 75%, notamment dans la banlieue Nord.

Une autre manière de représenter ces données est de considérer que les pourcentages traduisent un continuum spatial (ce qui, dans ce cas, est évidemment un abus de langage, les données dont on dispose étant communales, et non pas celles des bureaux de vote). En affectant au centre de chaque commune son pourcentage d'abstention, on obtient une carte en plages de niveaux (fig. 1.12). Pour réaliser ce type de carte, on utilise un carroyage plus ou moins fin sur lequel on estime, ici par krigeage, les valeurs en chaque carreau. On obtient ainsi une carte très expressive, qui s'affranchit des limites communales.

Enfin, il est possible de représenter les valeurs du carroyage sous la forme d'un bloc diagramme (fig. 1.13).

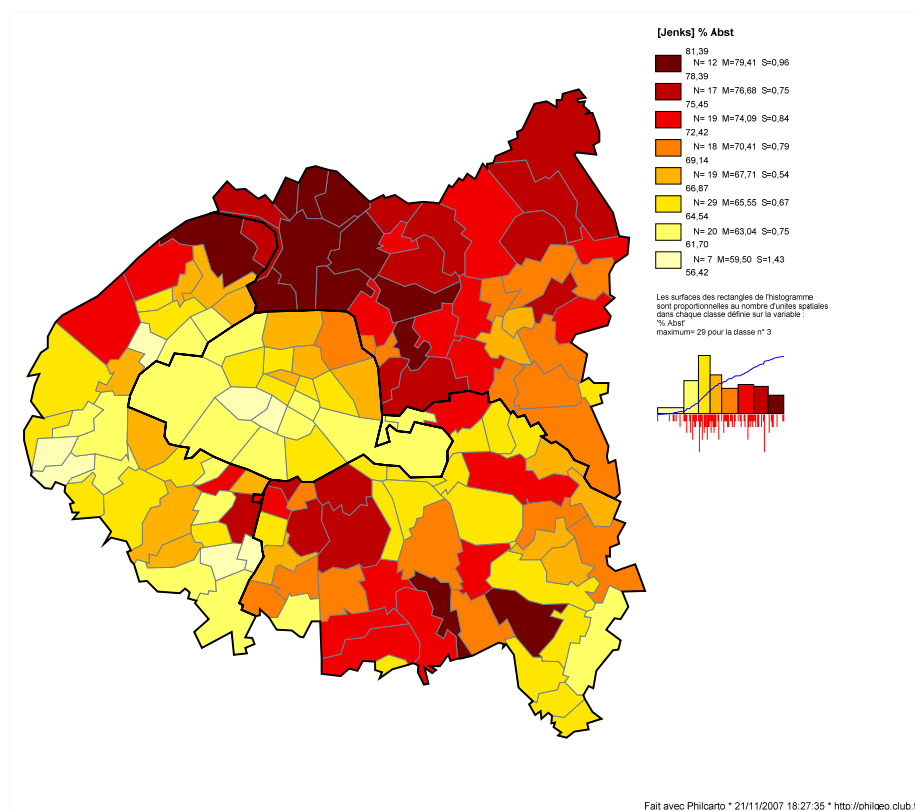


Figure 1.11 : France, arrondissements de Paris et communes de la première couronne.
Carte choroplèthe du pourcentage d'abstention au Référendum du 24/09/2000.

Fichiers : \Philcarto exemples\ Paris & 1° Couronne\

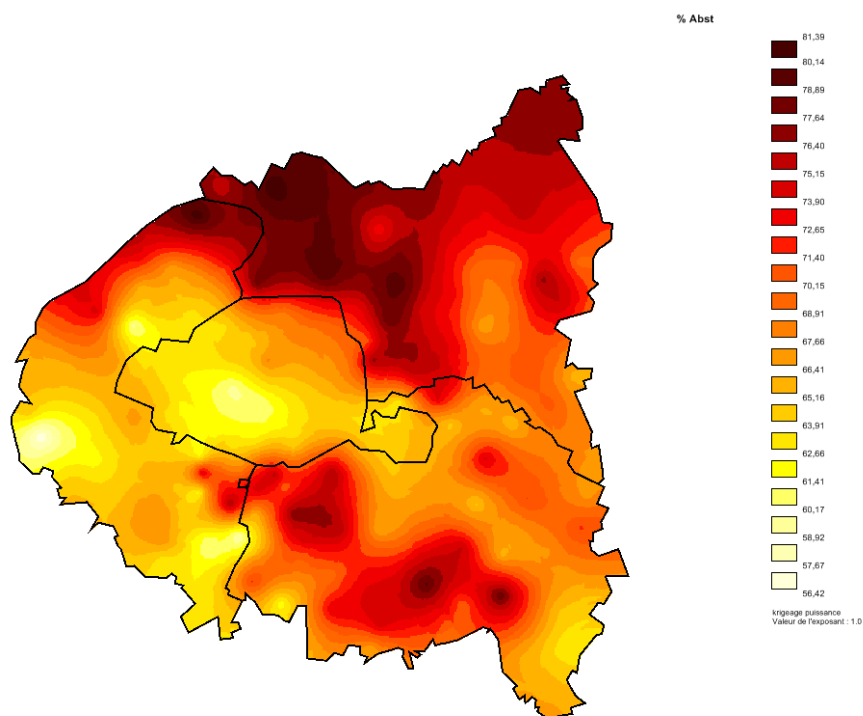


Figure 1.12 : France, arrondissements de Paris et communes de la première couronne.
Carte en plages de niveaux colorées du pourcentage d'abstention au Référendum du 24/09/2000.
Lissage par krigeage.

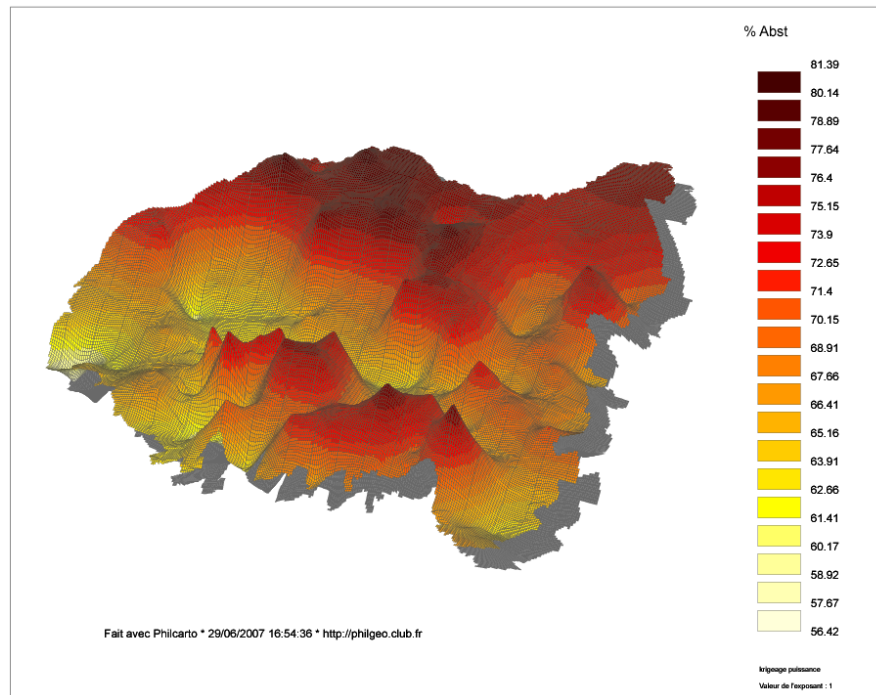


Figure 1.13 : France, arrondissements de Paris et communes de la première couronne.
Bloc diagramme du pourcentage d'abstention au Référendum du 24/09/2000.
Lissage par krigeage vu vers le Nord.

Les cartes en plages de niveaux et en bloc-diagramme

Les cartes en plages de niveaux reposent sur l'hypothèse que la variation des mesures statistiques présente une certaine continuité dans l'espace. C'est le cas, en météorologie pour les cartes de précipitations. En géographie humaine, une telle hypothèse est souvent trop forte, et la visualisation des données par des plages de niveaux n'est que la représentation des traits les plus saillants de la distribution.

Les points de mesure forment un maillage ponctuel, régulier ou non ; il s'agit de lieux de coordonnées X,Y, ou bien d'un point caractéristique de chaque polygone (centroïde, chef lieux...) d'un maillage surfacique. Un second maillage surfacique ou linéaire peut être superposé au maillage précédent pour former un habillage qui complète la carte et en facilite la lecture.

A partir du maillage ponctuel est dérivé un carroyage plus ou moins fin formant un maillage surfacique carré. Pour chaque centre de chaque carreau est estimée une valeur. Deux méthodes sont couramment employées pour ces estimations : la triangulation de Delaunay et le krigeage. Lorsque toutes les valeurs sont estimées pour chaque carreau, la cartographie en plages de niveaux devient un simple problème de carte choroplèthe, sachant que les limites des classes sont les limites des niveaux basés sur un espacement régulier, dans l'unité de mesure de la variable cartographiée.

Le maillage coloré peut aussi faire l'objet d'une représentation par un bloc-diagramme, selon un angle de vue choisi par l'utilisateur.

Les cartes de niveaux donnent de bons résultats si les valeurs ne sont pas spatialement indépendantes, caractéristique pouvant être évaluée par les coefficients d'autocorrélation spatiale de Moran et de Geary. La longueur des temps de calcul est l'inconvénient principal de ce type de carte.

1.8. Les cartes en surfaces de tendances

Les lissages par triangulation ou voisinage ne sont pas les seules techniques cartographiques permettant de détecter une organisation générale en réduisant l'importance des aspérités d'une distribution spatiale. L'analyse en surfaces de tendances utilise la méthode des moindres carrés pour ajuster des polynômes bidimensionnels et en visualiser la forme : les coordonnées géographiques (X,Y) sont les variables explicatives d'une variable Z à représenter, soit $z_i = f(x_i, y_i)^P$ où P représente l'ordre du polynôme. On obtient ainsi une série de surfaces de plus en plus complexes quand s'accroît l'ordre du polynôme. Pour représenter les surfaces tridimensionnelles sur le plan cartographique, on recourt à des plages de niveaux colorées.

En appliquant cette technique à la carte du nombre de naissances pour 1000 habitants en Allemagne (fig. 1.14), on détecte deux tendances significatives.

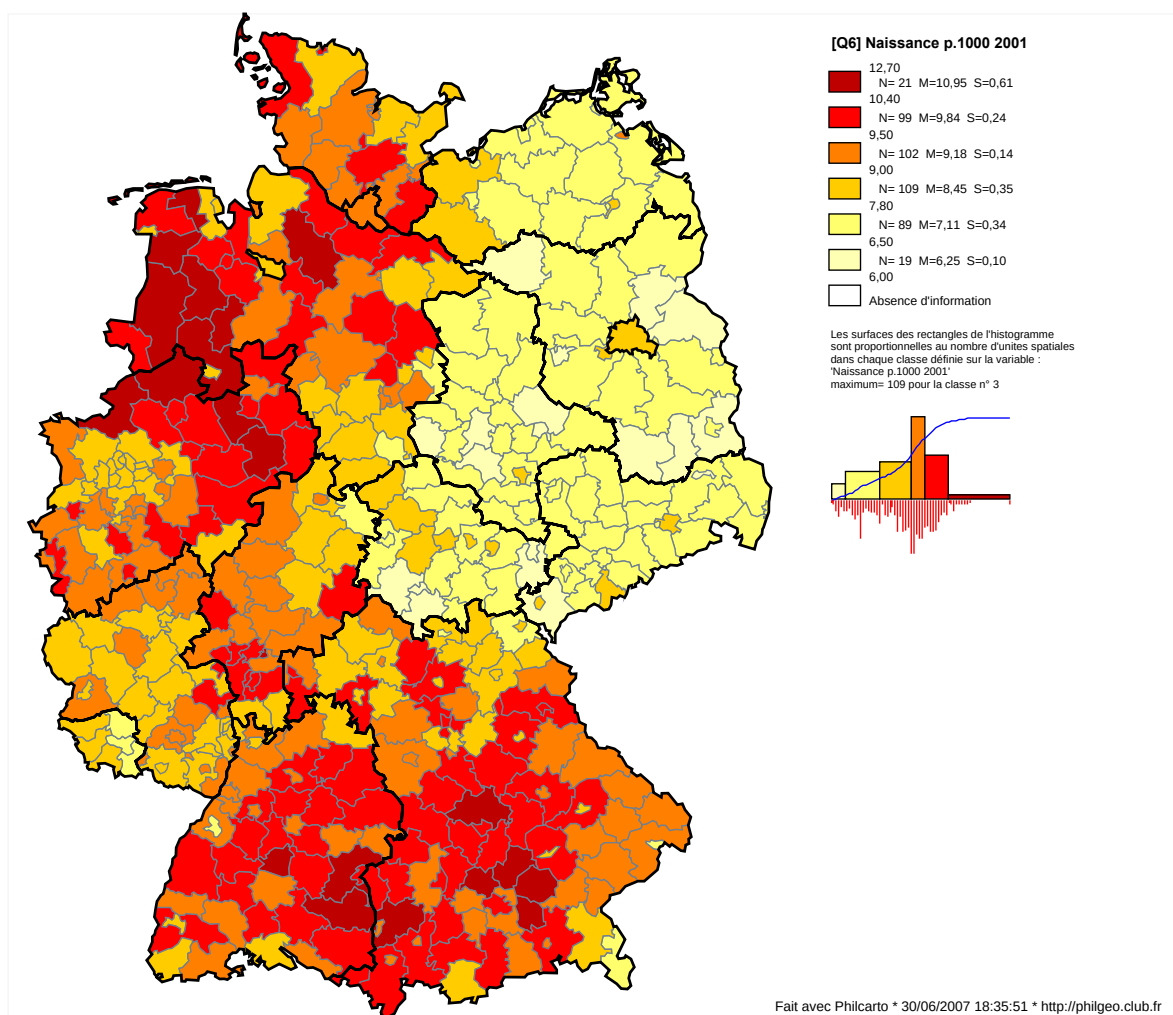


Figure 1.14 : Allemagne (kreise). Nombre de naissances pour 1000 habitants en 2001.

Fichiers : \Philcarto exemples\ Allemagne\

La première correspond à un ajustement par un polynôme d'ordre 1 qui absorbe près de 30% de la variance totale (fig. 1.15). Cette tendance, de forme plane, exprime l'accroissement de la natalité selon une orientation Nord-Est / Sud-Ouest.

En ajustant un polynôme d'ordre 4, le pourcentage de variance expliquée grimpe à 66%. Cette tendance, plus complexe que la précédente, traduit l'existence de deux maximums de natalité, l'un localisé sur la Bavière au Sud, l'autre sur la Basse Saxe. Cette surface de tendance révèle aussi l'existence d'une « cuvette » centrée sur Berlin, couvrant l'ensemble de l'ex-Allemagne de l'Est (fig. 1.16).

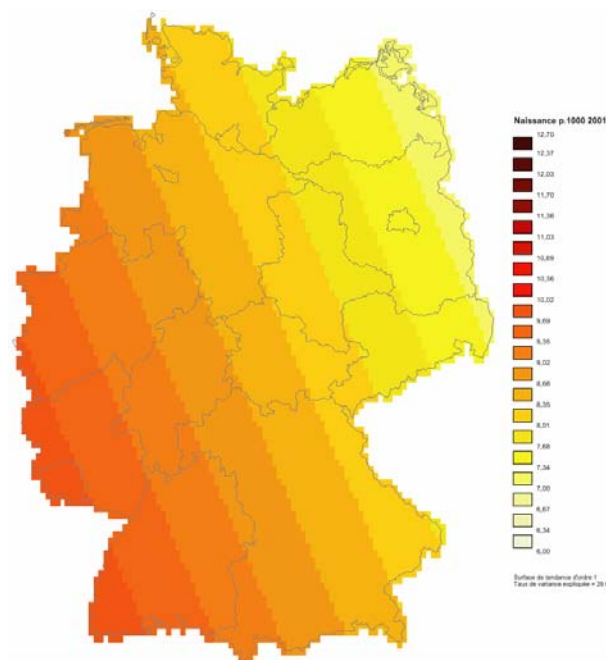


Figure 1.15 : Allemagne (kreise).
Nombre de naissances pour 1000 habitants en 2001.
Surface de tendance d'ordre 1 ($R^2 = 29,60\%$).

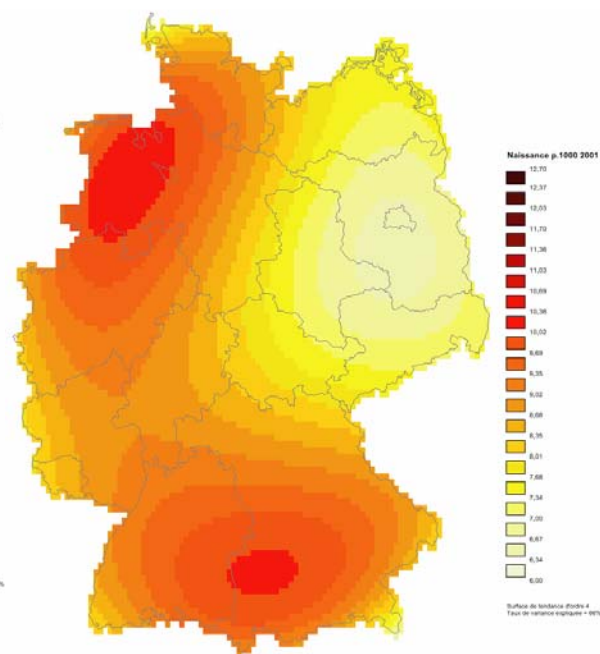


Figure 1.16 : Allemagne (kreise).
Nombre de naissances pour 1000 habitants en 2001.
Surface de tendance d'ordre 4 ($R^2 = 66\%$).

Le tracé du bloc-diagramme montre clairement les monts et dépression de la carte de la natalité, ici selon un angle de vue vers le Nord (fig. 1.16).

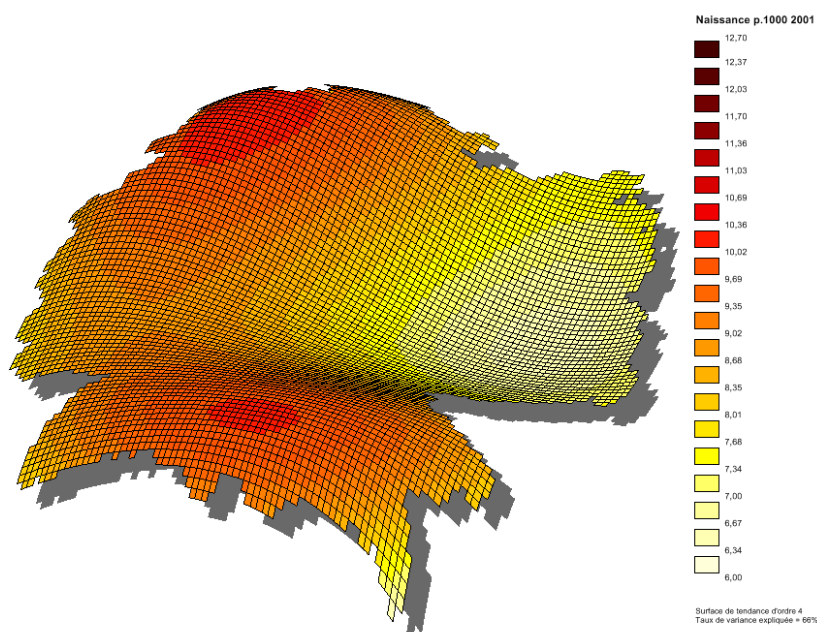


Figure 1.16 : Allemagne (kreise). Nombre de naissances pour 1000 habitants en 2001.
Surface de tendance d'ordre 4 ($R^2 = 66\%$). Bloc-diagramme vu vers le Nord.

Enfin, les résidus de l'ajustement par la surface d'ordre 4 peuvent être cartographiés de manière à détecter les localités dont la natalité s'ajuste mal au modèle général révélé par la surface de tendance (fig. 1.17). Noter, sur cette carte, l'utilisation d'une gamme de couleurs en opposition : du jaune au rouge pour les résidus positifs, du cyan au bleu foncé pour les résidus négatifs.

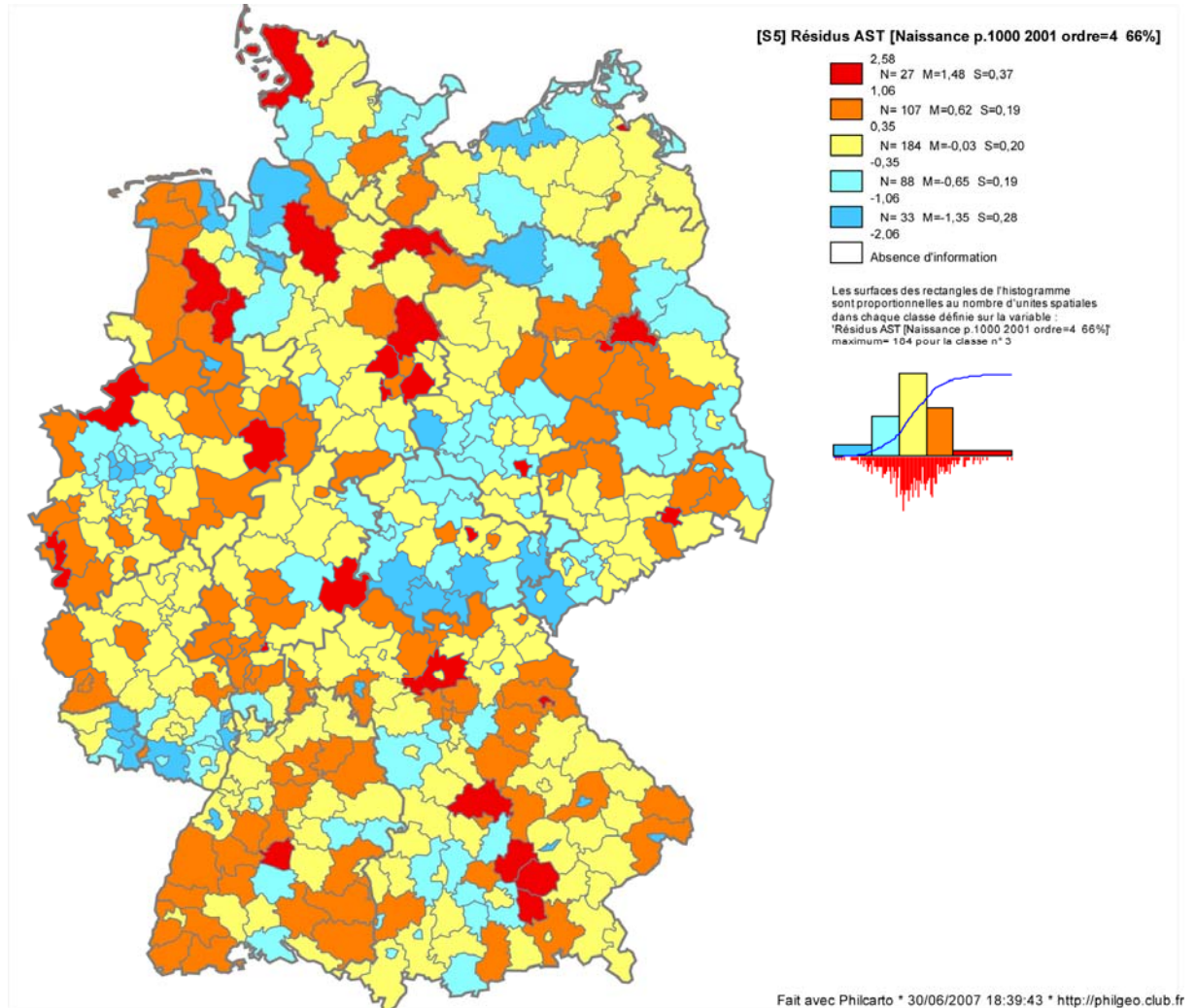


Figure 1.17 : Allemagne (kreise).
 Nombre de naissances pour 1000 habitants en 2001.
 Résidus de la surface de tendance d'ordre 4 ($R^2 = 66\%$).

Les cartes en surfaces de tendances

Les cartes en surfaces de tendances reposent sur l'hypothèse que la variation des mesures statistiques présente une certaine continuité dans l'espace, hypothèse également à l'origine des cartes en plages de niveaux. La différence avec ces dernières réside dans la décomposition du phénomène supposé continu dans l'espace en composantes d'échelles se rapportant successivement à des espaces de plus en plus réduits. Aux valeurs calculées avec ces composantes s'ajoute un résidu local.

Les points de mesure forment un maillage ponctuel, régulier ou non ; il s'agit de lieux de coordonnées X,Y, ou bien d'un point caractéristique de chaque polygone (centroïde, chef lieu...) d'un maillage surfacique. Un second maillage surfacique ou linéaire peut être superposé au maillage précédent pour former un habillage qui complète la carte et en facilite la lecture.

A partir du maillage ponctuel est dérivé un carroyage plus ou moins fin formant un maillage surfacique carré. Pour chaque centre de chaque carreau sont estimées plusieurs valeurs à l'aide de fonctions polynomiales du type $Z = f(X,Y)^n$ où n est une puissance entière variant de 1 à 4 en pratique. Avec un polynôme de degré 4, on estime 4 valeurs pour chaque centre de chaque carreau. A ces fonctions est associé un taux de variance expliquée qui rend compte du plus ou moins bon ajustement des données par ces fonctions.

Lorsque toutes les valeurs sont estimées pour chaque carreau, la cartographie en plages de niveaux devient un simple problème de carte choroplèthe, sachant que les limites des classes sont les limites des niveaux basés sur un espacement régulier, dans l'unité de mesure de la variable cartographiée. Le nombre de cartes de tendances est égal au plus grand degré du polynôme.

Le maillage coloré peut aussi faire l'objet d'une représentation par un bloc-diagramme, selon un angle de vue choisi par l'utilisateur.

Les cartes de tendances spatiales donnent de bons résultats si les valeurs ne sont effectivement pas spatialement indépendantes, caractéristique pouvant être évaluée par les coefficients d'autocorrélation spatiale de Moran et de Geary.

La difficulté d'attribuer une signification géographique aux degrés successifs des polynômes constitue le principal inconvénient de cette technique. En pratique, on ne retient que les cartes absorbant un taux élevé de variance et présentant des configurations géographiques différentes. La carte des résidus permet ensuite d'évaluer l'importance de l'échelle locale dans la structuration spatiale du phénomène étudié.

1.9. Les cartes en semis de points

Les cartes en semis de points représentent des quantités ou des effectifs. Chaque point est placé de manière aléatoire sur la surface de l'unité spatiale à laquelle il appartient. Après différentes tentatives de calibrage, on obtient souvent un heureux effet de densité, surtout lorsque le maillage du fond de carte est assez fin pour rendre compte de la localisation du phénomène (le caractère aléatoire du positionnement de chaque point est alors contrôlé au moins en partie). La mise en œuvre de cette technique est plus complexe que celle des cercles proportionnels. Il faut en effet calibrer le point élémentaire de manière à ce que, au total, le nombre de points ne soit pas trop élevé afin d'éviter les fichiers graphiques trop volumineux. Par ailleurs, lorsqu'il y a trop de points, des phénomènes de saturation se produisent : certaines unités spatiales sont trop petites pour recevoir tous les points correspondant à leur effectif. Mais cela n'est pas très gênant d'un point de vue sémiologique car la saturation parfaitement visible, exprime bien un effectif élevé.

La carte en semis de points de la population de Bolivie en 1992 (fig. 1.18) montre tout l'intérêt de cette technique. «L'arc andin» y apparaît en effet avec une grande netteté. D'orientation NO-SE, puis N-S, il rassemble l'essentiel de la population du pays. Les régions les plus densément peuplées forment la partie orientée NO-SE qui relie Cochabamba à La Paz. En dehors de cet arc, la population est principalement localisée dans la région de Santa Cruz de la Sierra qui connaît un *boom* démographique grâce à la mise en valeur agricole de la région.

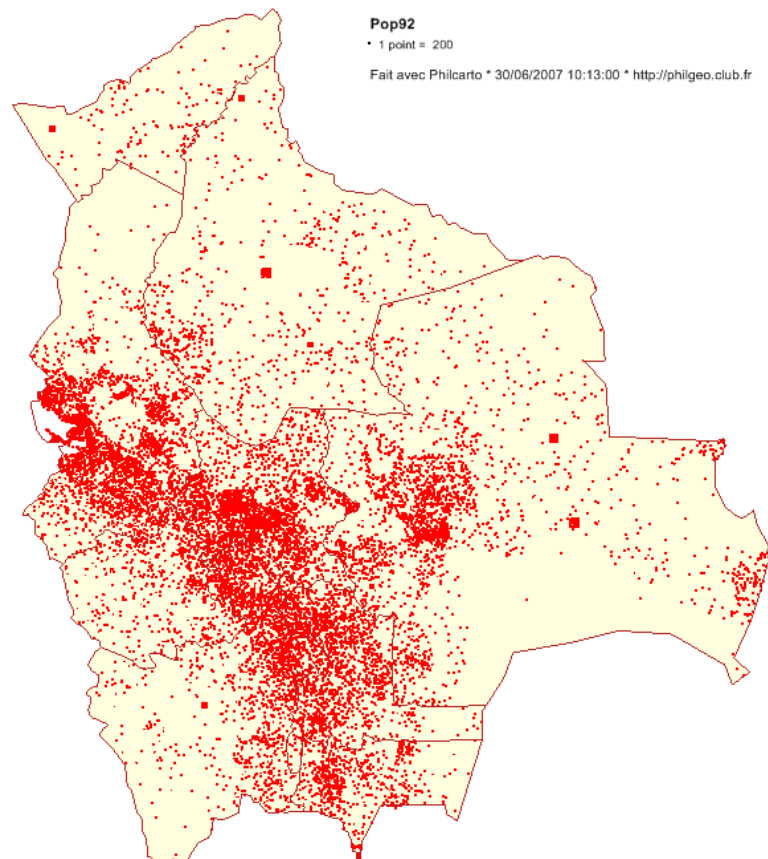


Figure 1.18 : Bolivie (cantons). Population totale en 1992.

Fichiers : \Philcarto exemples\ Bolivie\

Les cartes en semis de points

Une carte en semis de points représente des quantités ou des effectifs par un nombre de points proportionnel à ces quantités ou à ces effectifs, à l'exclusion des pourcentages, taux, mesures ou catégories.

Le maillage est surfacique ; les limites de ses éléments sont enregistrées sous forme de polygones. Un second maillage surfacique ou linéaire peut être superposé au premier pour former un habillage qui complète la carte et en facilite la lecture.

A l'intérieur de chaque polygone sont placés au hasard des points en nombre proportionnel à l'effectif enregistré pour ce polygone. La valeur du point doit être préalablement calibrée, soit en choisissant de manière arbitraire cette valeur, soit en choisissant le nombre maximum de points qu'un polygone peut contenir ; dans ce cas, la valeur du point est égale au plus grand effectif à cartographier (le maximum de la variable) divisé par le nombre maximum des points choisi.

Pour obtenir l'effet de densité souhaité, on peut aussi jouer sur la taille du point (de 1 à 5 pixels au plus), sur sa couleur (dans ce cas, tous les points de la carte ont la même couleur), et sur le fond sur lequel sont placés les points en maintenant ou non les contours des polygones, et en leur appliquant ou non une couleur. Tout l'art du cartographe consiste à choisir le résultat le plus conforme à la géographie du phénomène représenté. Mais le résultat dépend aussi du maillage surfacique qui doit être suffisamment détaillé pour que la géographie soit bien traduite, malgré l'abus lié au caractère aléatoire du placement des points sur la carte.

Plusieurs tentatives sont souvent indispensables pour obtenir un résultat satisfaisant, mais le principal inconvénient de cette technique réside dans la nécessité de rappeler au lecteur que les points ne sont pas localisés au lieu précis d'enregistrement des quantités cartographiées. Des malentendus peuvent résulter de la mauvaise compréhension de la méthode par des non-cartographes. Un second inconvénient est lié au nombre total de points figurant sur la carte conduisant à des fichiers graphiques très lourds et donc difficiles à manipuler ensuite (corrections, habillage, impression).

1.10. Les cartes en semis de points colorés

Les cartes en semis de points colorés permettent de représenter simultanément des quantités ou des effectifs (densité du semis) et des valeurs numériques continues ou discrètes exprimant des pourcentages, des mesures, etc. (couleurs des points). Pour améliorer la perception des densités de couleurs, les points peuvent être grossis.

La carte en semis de points colorés de la population des cantons de Bretagne (fig. 1.19) met en évidence la croissance des principaux pôles urbains (nombreux habitants en 1999 et fort taux de croissance) de la région au cours du XXe siècle, qui, hormis la capitale régionale, Rennes, sont situés sur la côte, ainsi que le dépeuplement des communes de l'intérieur (peu d'habitants en 1999 et forte décroissance).

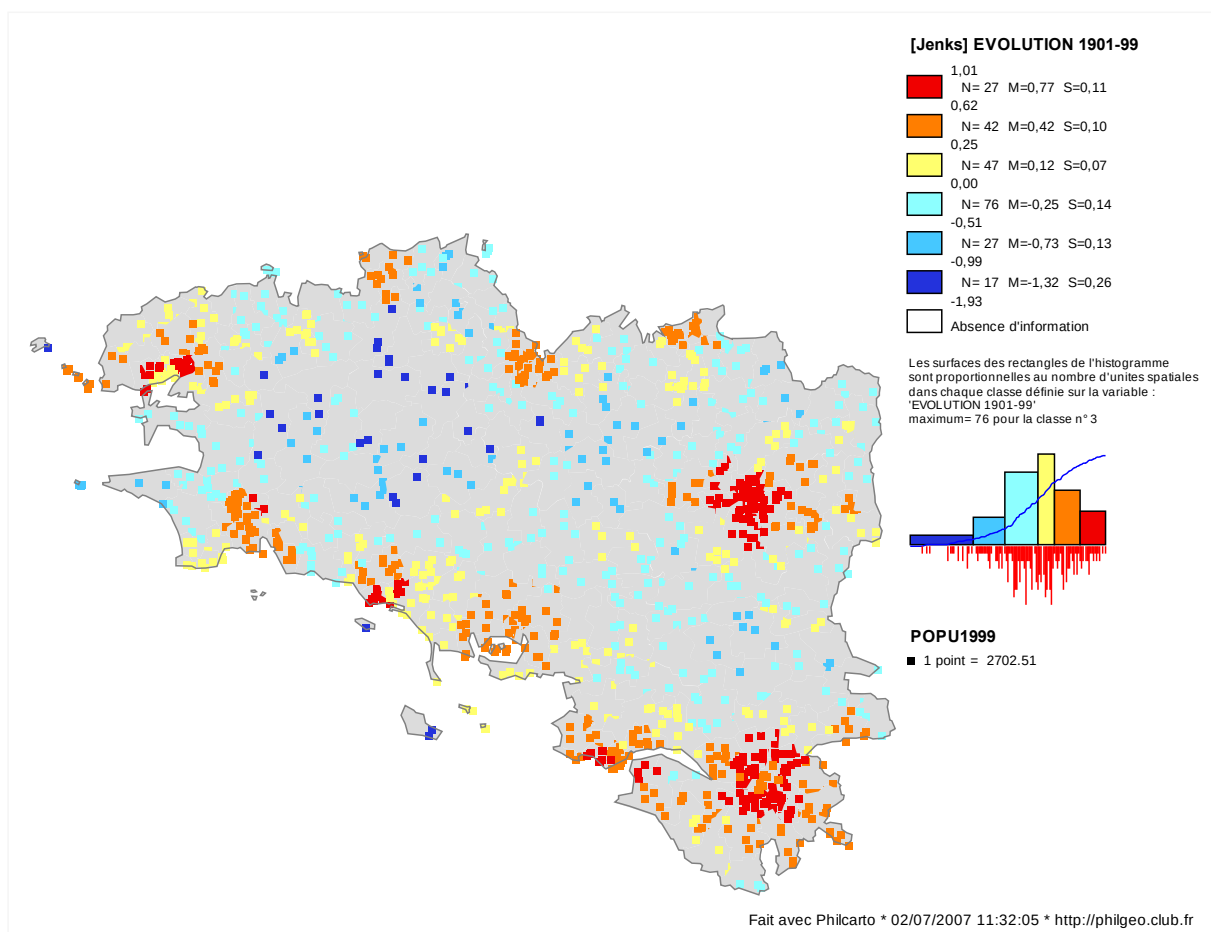


Figure 1.19 : Bretagne (cantons). Evolution de la population 1901-1999.
Fichiers : IGN / GéoFLA, fichiers non fournis.

Les cartes en semis de points colorés

Une carte en semis de points colorés représente simultanément deux variables. Le nombre de points est proportionnel aux valeurs d'une première variable exprimant des quantités ou des effectifs. Les couleurs des points correspondent aux valeurs de la seconde variable, soit continue (pourcentages, taux, mesures...) et donc préalablement discrétisée, soit discrète (catégories numériques ou nominales).

Le maillage est surfacique ; les limites de ses éléments sont enregistrées sous forme de polygones. Un second maillage surfacique ou linéaire peut être superposé au premier pour former un habillage qui complète la carte et en facilite la lecture.

A l'intérieur de chaque polygone sont placés au hasard des points en nombre proportionnel à l'effectif enregistré pour ce polygone. La valeur du point doit être préalablement calibrée, soit en choisissant de manière arbitraire cette valeur, soit en choisissant le nombre maximum de points qu'un polygone peut contenir. La taille du point peut également varier.

La discrétisation des variables de couleurs est réalisée de la même façon que pour les cartes choroplèthes.

La légende comprend deux parties. Dans sa partie supérieure, des caissons colorés permettent de qualifier chaque couleur par un intervalle de valeurs ou une modalité ; un histogramme (pour les variables continues), ou un diagramme à bâtons (pour les variables discrètes) complété d'une courbe des fréquences cumulées permettent d'apprécier la distribution des valeurs. Dans la partie inférieure de la légende figure un point avec sa valeur.

Les inconvénients de ce type de carte sont ceux des cartes en semis de points auxquels s'ajoute un manque de lisibilité des points de couleur pâle (inconvenient pouvant être contourné en partie avec un fond coloré sombre).

1.11. Les cartes de liens dites aussi « en oursins »

Les cartes de liens, aussi appelées cartes en « oursins », permettent de visualiser des relations entre deux points. Ce mode de représentation est extrêmement efficace lorsqu'on cherche montrer une organisation spatiale polarisée car les liens s'arrangent alors à la manière des piquants d'un oursin.

C'est ainsi que l'on peut apprécier les formes engendrées par les liens fonctionnels qui unissent les lieux centraux de l'État du Paraná au Brésil pour tous les niveaux de la hiérarchie urbaine (fig. 1.20) tels qu'ils apparaissent dans l'enquête REGIC de l'Institut Brésilien de Géographie et de Statistique, IBGE.

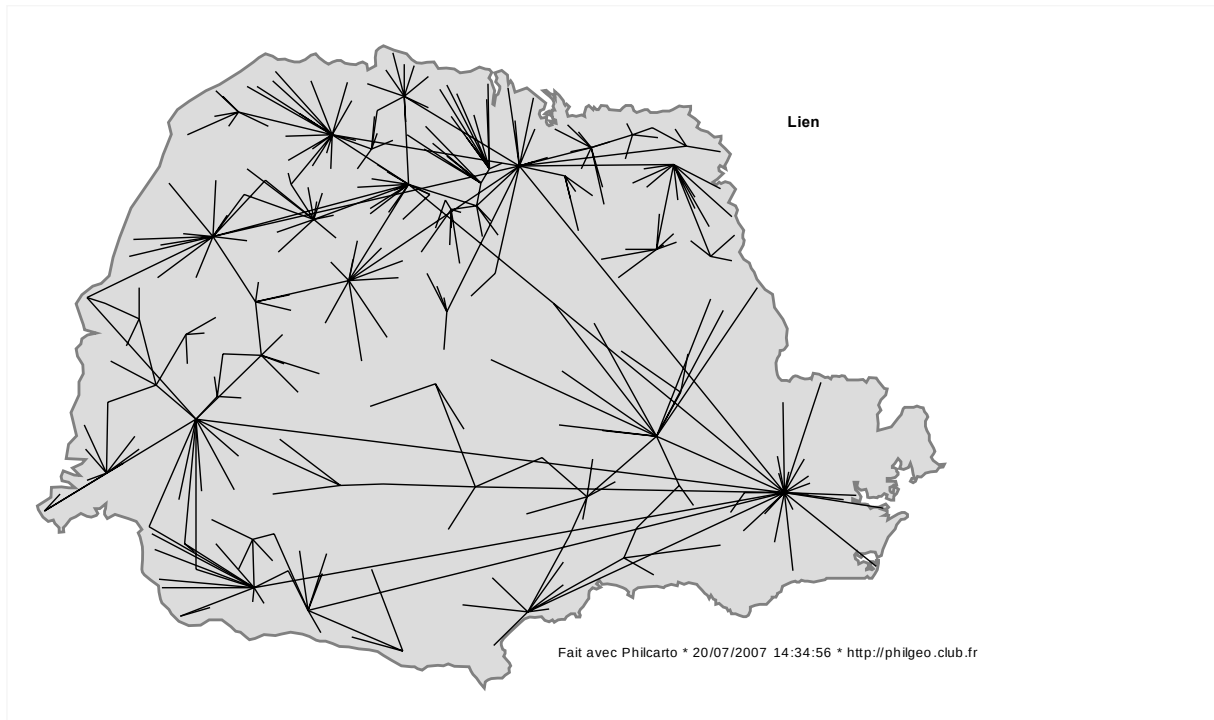


Figure 1.20 : Paraná, Brésil (municípios). Liens entre lieux centraux.
Fichiers : \Philcarto exemples\Parana\

De manière à rendre compte de la hiérarchie des centres, il est possible de colorier les liens en fonction du niveau hiérarchique du centre dont le lien est originaire, c'est à dire du centre subordonné à celui de niveau supérieur auquel il est relié. Ces niveaux sont des modalités d'une échelle discrète ordonnée qui prennent les valeurs 1, 2, 3, 4, 5, 6, 7, 1 correspondant au niveau le plus bas, 7 au niveau le plus élevé dans la hiérarchie urbaine. Une gamme de couleurs « perroquet » arrangée avec logique permet d'obtenir un résultat satisfaisant (fig. 1.21).

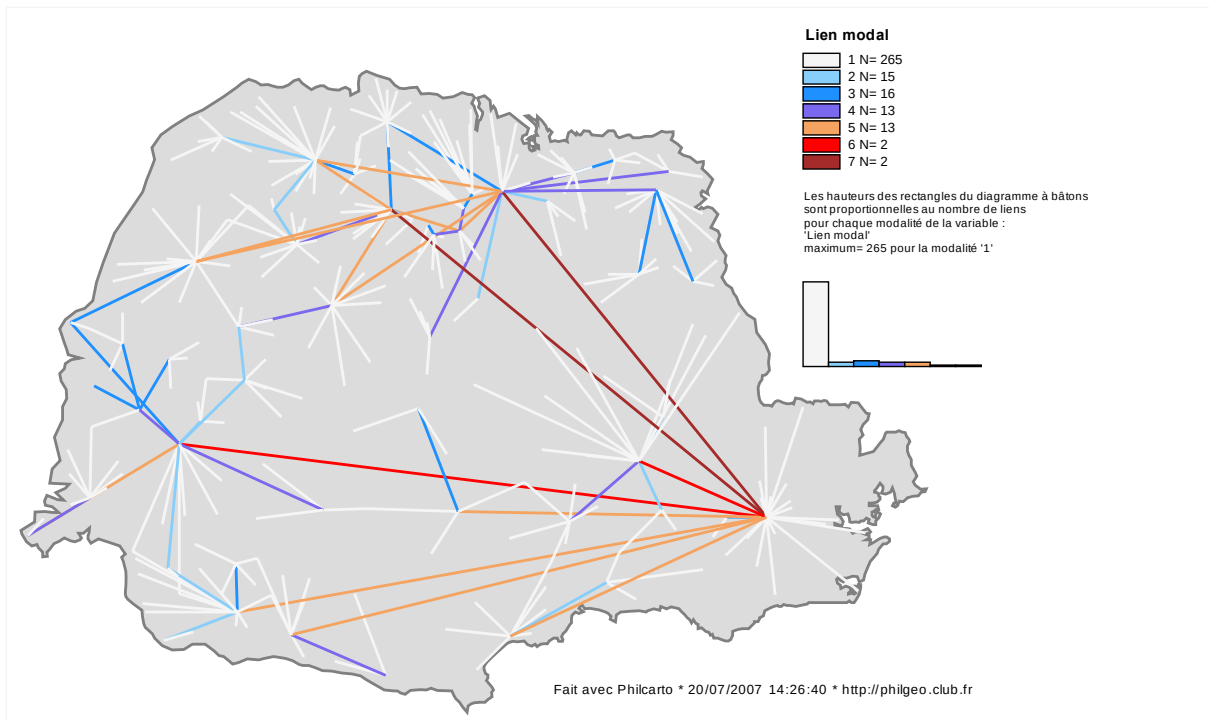


Figure 1.21 : Paraná, Brésil (municípios).
Liens entre lieux centraux.

Les cartes de liens dites aussi en « oursins »

Une carte de liens représente des relations entre des points de l'espace formant un maillage régulier ou non ; il s'agit de lieux de coordonnées X,Y, ou bien d'un point caractéristique de chaque polygone (centroïde, chef lieux...) d'un maillage surfacique. Un second maillage surfacique peut être superposé au maillage précédent pour former un habillage qui complète la carte et en facilite la lecture.

Les relations sont figurées par des lignes droites qui joignent les points deux à deux. Ces lignes peuvent être coloriées en fonction des modalités d'une variable nominale.

Ce mode de représentation est efficace lorsqu'on cherche à exprimer une organisation spatiale polarisée, car les liens s'arrangent alors à la manière des piquants d'un oursin.

L'emploi de ce genre de carte présente de nombreuses limitations : risque d'enchevêtrement des liens en l'absence d'une forte organisation concentrique ; difficulté de lecture des couleurs des lignes.

1.12. Les cartes de réseaux

Les cartes présentées jusqu'ici s'appuient sur des éléments cartographiques ponctuels ou surfaciques (polygonaux). Même les cartes en oursins qui, d'une certaine manière, donnent une image du réseau urbain, ne sont pas basées sur des éléments cartographiques linéaires puisque l'information représentée correspond à des relations entre des points figurant des lieux centraux.

Les réseaux dont il est question ici sont constitués de lignes généralement non fermées. Ce sont des lignes brisées ou des segments de droite s'accordant aux inflexions du tracé. On trouve ce genre de réseau dans la nature avec, par exemple, les fleuves et rivières sur lesquels on souhaite représenter un débit horaire ; il peut s'agir aussi de réseaux ferroviaires ou routiers pour lesquels on souhaite montrer le statut des routes (vicinales, départementales, nationales...), la capacité de charge ou la fréquentation.

La Préfecture de la Région Île-de-France a mis en place un système d'observation de la circulation routière sur les principaux axes de la région (fig. 1.22). Dénommé « Service d'Informations pour un Réseau Intelligible aux Usagers », SIRIUS, ce système fournit de nombreuses informations en temps réel, en particulier sur les « bouchons ». Il permet aussi d'élaborer des statistiques destinées à la recherche de solutions pour éviter l'engorgement.

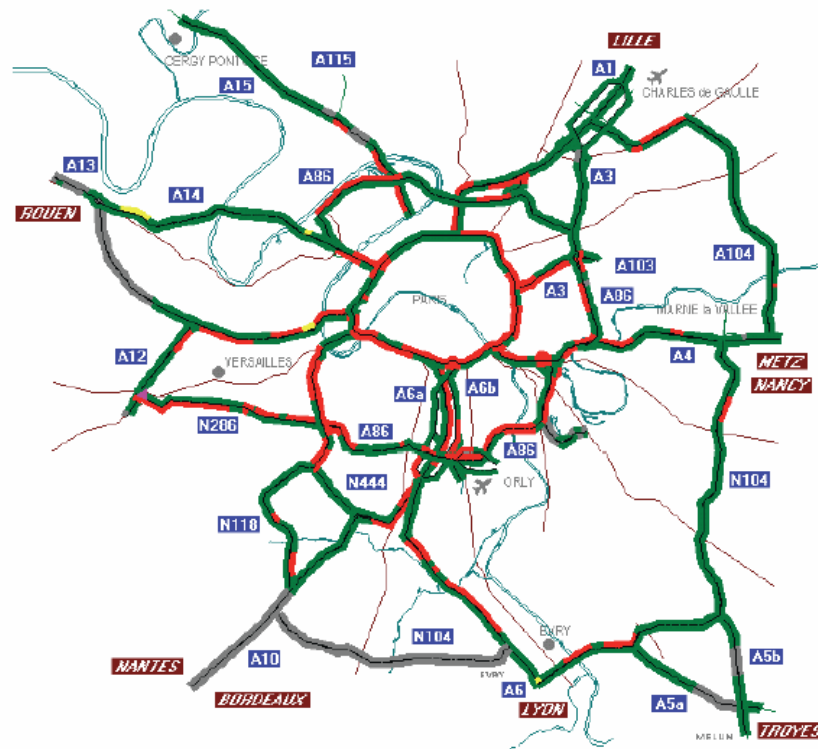


Figure 1.22 : Les axes routiers de la Région Île-de-France observés par le système SIRIUS.
(source : www.ile-de-france.equipement.gouv.fr)

Les statistiques produites par SIRIUS font référence à chaque axe du réseau, par exemple l'autoroute A13 ou bien la route nationale N104. Pour chaque axe, on dispose d'informations brutes comme sa longueur, le nombre d'heures d'encombrements, la longueur d'encombrement. Des indicateurs peuvent être dérivés tel que le HKM (Heures x Kilomètres d'encombrements) : c'est la somme des

produits des longueurs des bouchons (ramenées à une file) par la durée de ces bouchons. Un bouchon est un ralentissement de la circulation débutant lorsque la vitesse descend en dessous de 30 km/h et se terminant lorsque la vitesse remonte au-delà de 60 km/h.

Les statistiques sur les tronçons d'un réseau peuvent être cartographiées de deux façons. La variation des effectifs est traduite par une variation d'épaisseur des lignes figurant chaque tronçon. L'effectif est généralement rapporté à la longueur du tronçon de manière à rendre les effectifs comparables entre eux. Le nombre d'heures de bouchon par kilomètre est un exemple de ce genre de carte (fig. 1.23).

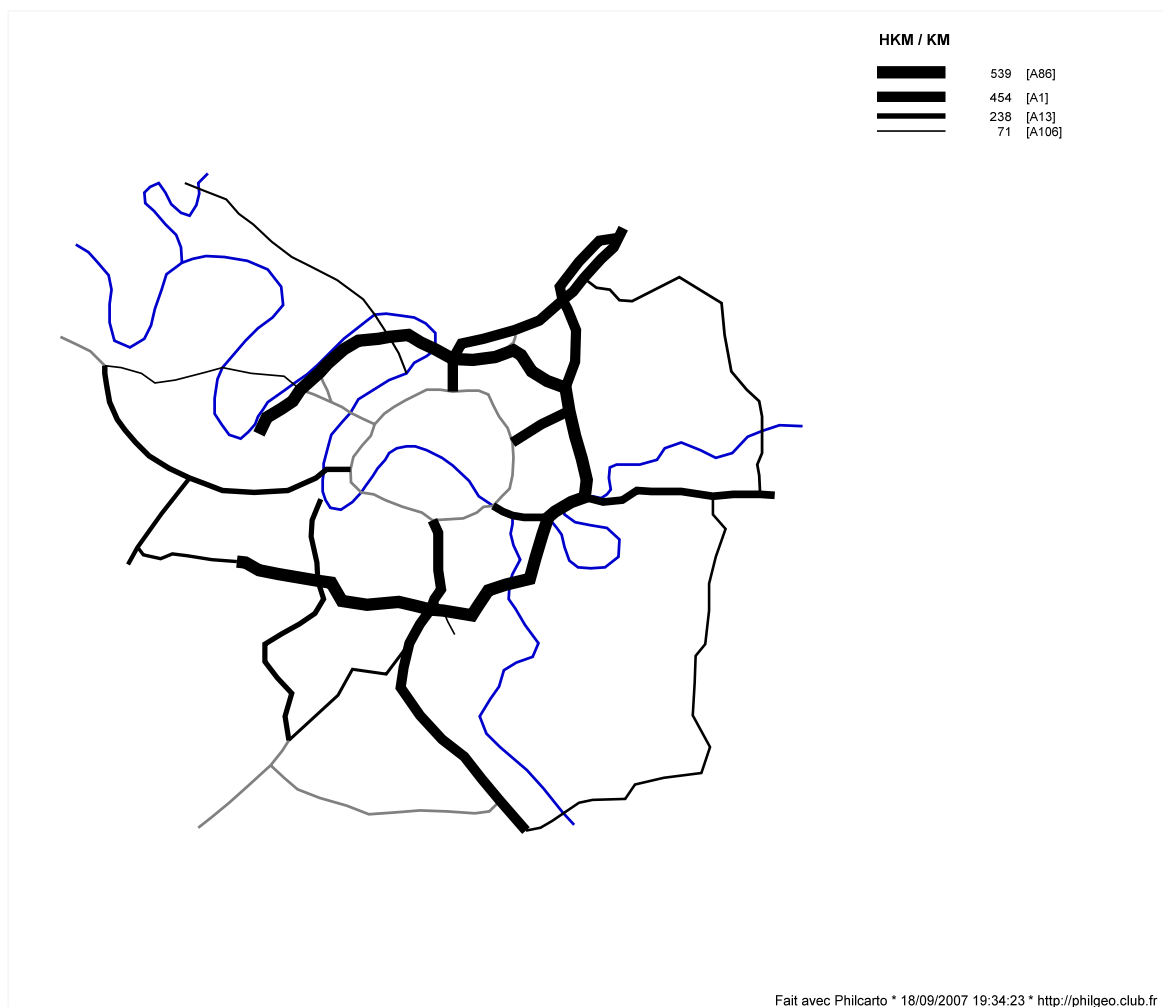


Figure 1.23 : France, Île-de-France.
 Nombre d'heures de bouchons par kilomètre et par an sur les axes du réseau routier principal.
 Fichiers : \Philcarto exemples\IDF SIRIUS\

Lorsque les données enregistrées sur chaque tronçon expriment non pas une quantité, mais un taux ou une modalité discrète, chaque tronçon peut être colorié, après discrétisation et affectation d'une couleur venant d'une gamme ordonnée dans le cas d'un taux, ou bien en affectant une couleur à chaque modalité d'une variable nominale. Par exemple, SIRIUS propose une évaluation de la sévérité des encombrements à l'aide d'un indice dont les valeurs varient de 0 à 2, préalablement discrétisé en 4 catégories discrètes : très sévère (1,5 à 2), sévère (1 à 1,5), fort (0,5 à 1) et réduit (0 à 0,5). Les couleurs traduisant de manière qualitative le degré d'encombrement sont appliquées sur chaque segment du réseau ; la carte obtenue

montre que le grand périphérique est déjà très sévèrement encombré (fig. 1.24)... Notons que chaque tronçon doit être représenté par un trait assez épais pour que les couleurs soient lisibles.

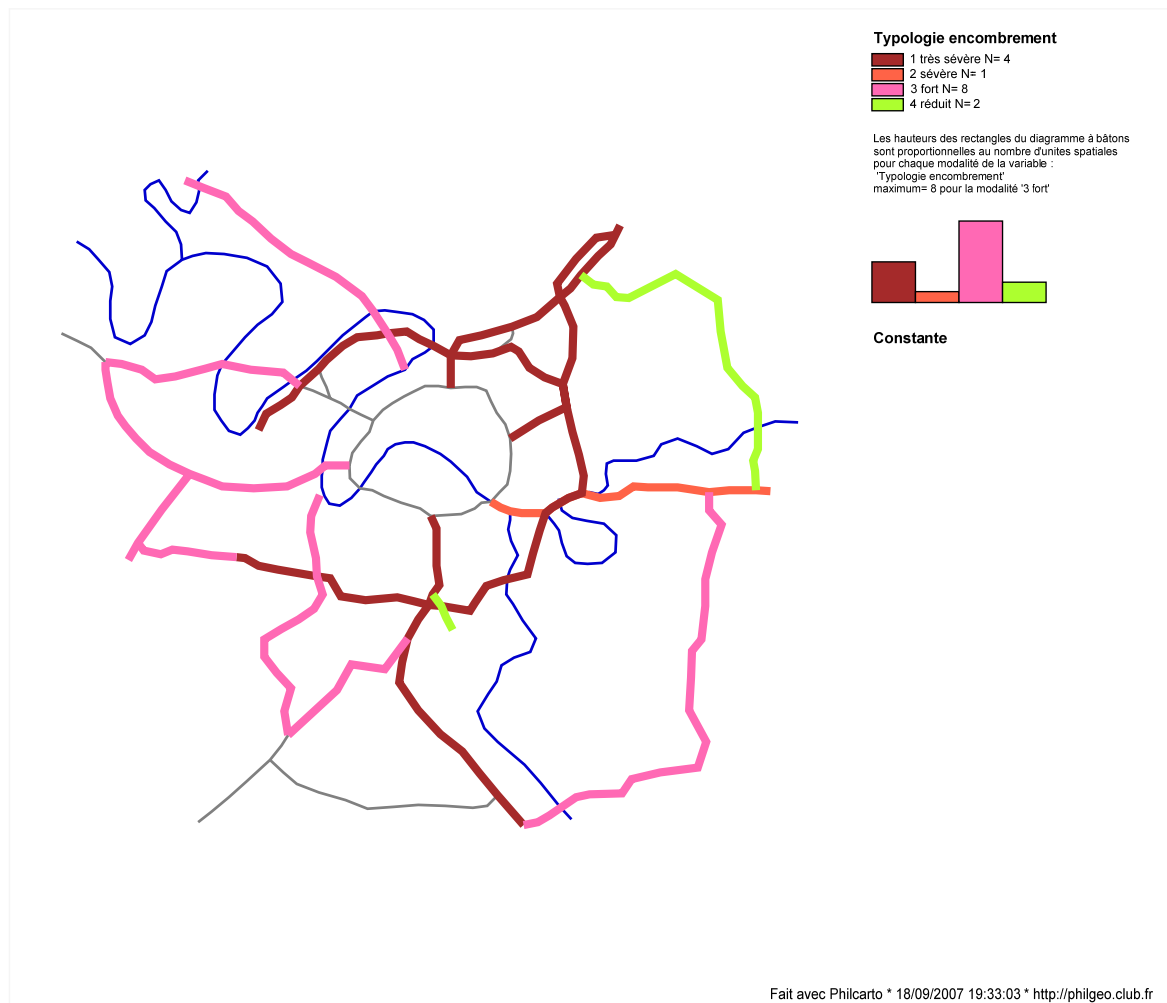


Figure 1.24 : France, Île-de-France. Sévérité des encombrements par axe du réseau routier principal.

On peut aussi combiner sur la même carte la variation de l'épaisseur des lignes fonction de quantités ou d'effectifs enregistrés sur les tronçons avec la variation de couleur exprimant des modalités ou des taux différents.

Les cartes de réseaux

Les cartes de réseaux sont basées sur des lignes représentant les tronçons du réseau.

La représentation des effectifs caractérisant chaque tronçon (flux s'écoulant dans le réseau ou autre mesure de quantité) fait appel à la variation des épaisseurs des lignes proportionnelle à ces effectifs. Pour figurer une modalité discrète issue d'une variable nominale, ou résultant d'un taux préalablement discrétisé, on fait appel à une gamme de couleur ordonnée (pour les taux) ou non (pour les modalités nominales).

Il est possible de combiner variation d'épaisseur et variation de couleur pour représenter simultanément sur chaque tronçon un effectif et une modalité discrète.

Ce mode de représentation est efficace si le maillage résultant du réseau n'est pas trop serré pour permettre la variation d'épaisseur qui, dans tous les cas, ne peut se faire que dans des limites relativement étroites : il apparaît difficile dans un format A4, et pour un réseau assez lâche, de tracer des lignes dont l'épaisseur dépasse un centimètre.

Enfin, l'identification des couleurs des lignes par la légende suppose que ces couleurs soient appliquées sur des lignes assez épaisses, ce qui limite aussi le domaine d'application de ce genre de carte.

1.13. Les cartes de gravitation issues du modèle de Huff

Les cartes en oursins permettent de visualiser les aires d'influences des lieux centraux en fonction des liens qu'ils entretiennent avec les niveaux de la hiérarchie urbaine qui leur sont immédiatement inférieurs et supérieurs. Ceci suppose qu'on possède une information relativement difficile à collecter lorsque les organismes en charge des enquêtes statistiques (comme l'INSEE en France avec l'inventaire communal) ne réalisent pas ce genre d'étude. Lorsqu'on ne dispose pas de ce genre d'information, ou bien lorsqu'elle est disponible et qu'on cherche à la confronter avec le modèle de gravitation, le modèle probabiliste d'influence urbaine de Huff [Huff ; 1973 et 1979] permet d'imaginer ce que pourraient être ces aires d'influences sous certaines contraintes reflétant deux les paramètres du modèle : la masse et la distance.

Appliqué au nombre des admissions dans les hôpitaux des capitales des États du Brésil (fig. 1.25), le modèle de Huff montre que les limites d'égale attraction calculées par le modèle (lignes bleu-ciel) ne correspondent presque jamais aux limites des États. Pour être validées, ces limites d'aires d'influence « théoriques » calculées en fonction des paramètres du modèle doivent naturellement faire l'objet d'une confrontation avec d'autres sources d'information.

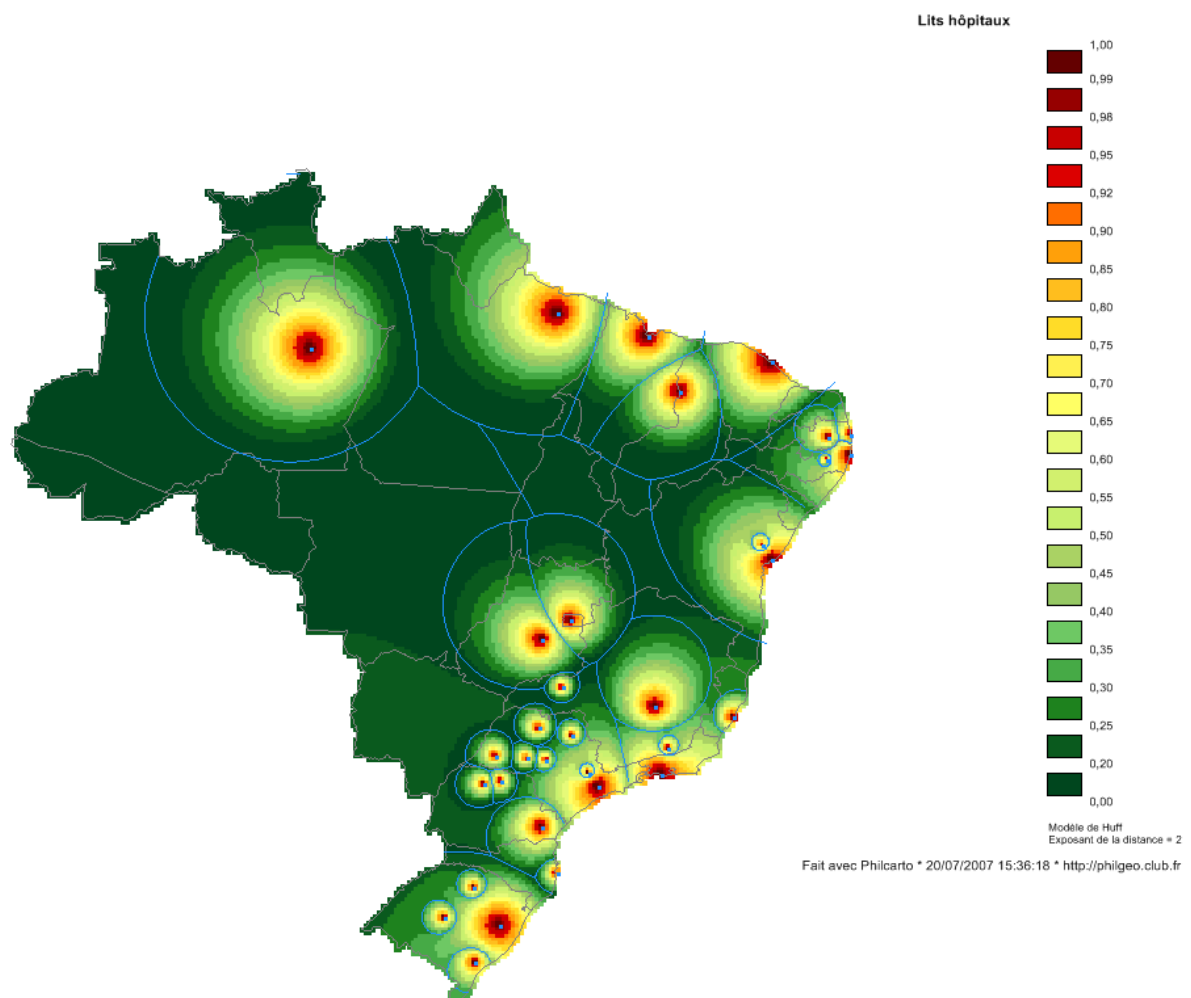


Figure 1.25 : Brésil (États). Limites d'égale influence hospitalière et gradients de probabilités d'influence calculés par le modèle de Huff.

Fichiers : \Philcarto exemples\Brésil gravitation\

Les limites d'égale influence peuvent parfaitement correspondre à une quasi-absence d'influence lorsque les distances sont telles que l'influence tend vers zéro. C'est l'un des intérêts du modèle de Huff que de permettre une évaluation de l'influence en tout lieu de l'espace étudié sous la forme d'une probabilité permettant d'estimer le nombre de personnes susceptibles de se rendre dans chaque centre à partir d'un lieu donné. Ainsi, on peut dresser une carte représentant des gradients de probabilité d'influence.

Appliqué aux admissions dans les hôpitaux des capitales des États du Brésil, le modèle de Huff montre que les limites d'égale attraction calculées par le modèle passent par des probabilités d'influence différentes.

Enfin, on peut tracer la carte des probabilités d'influence sous la forme d'un bloc diagramme vu en perspective. Les angles de vue sont réglables à l'aide de boutons, soit pas à pas, soit en continu (fig. 1.26).

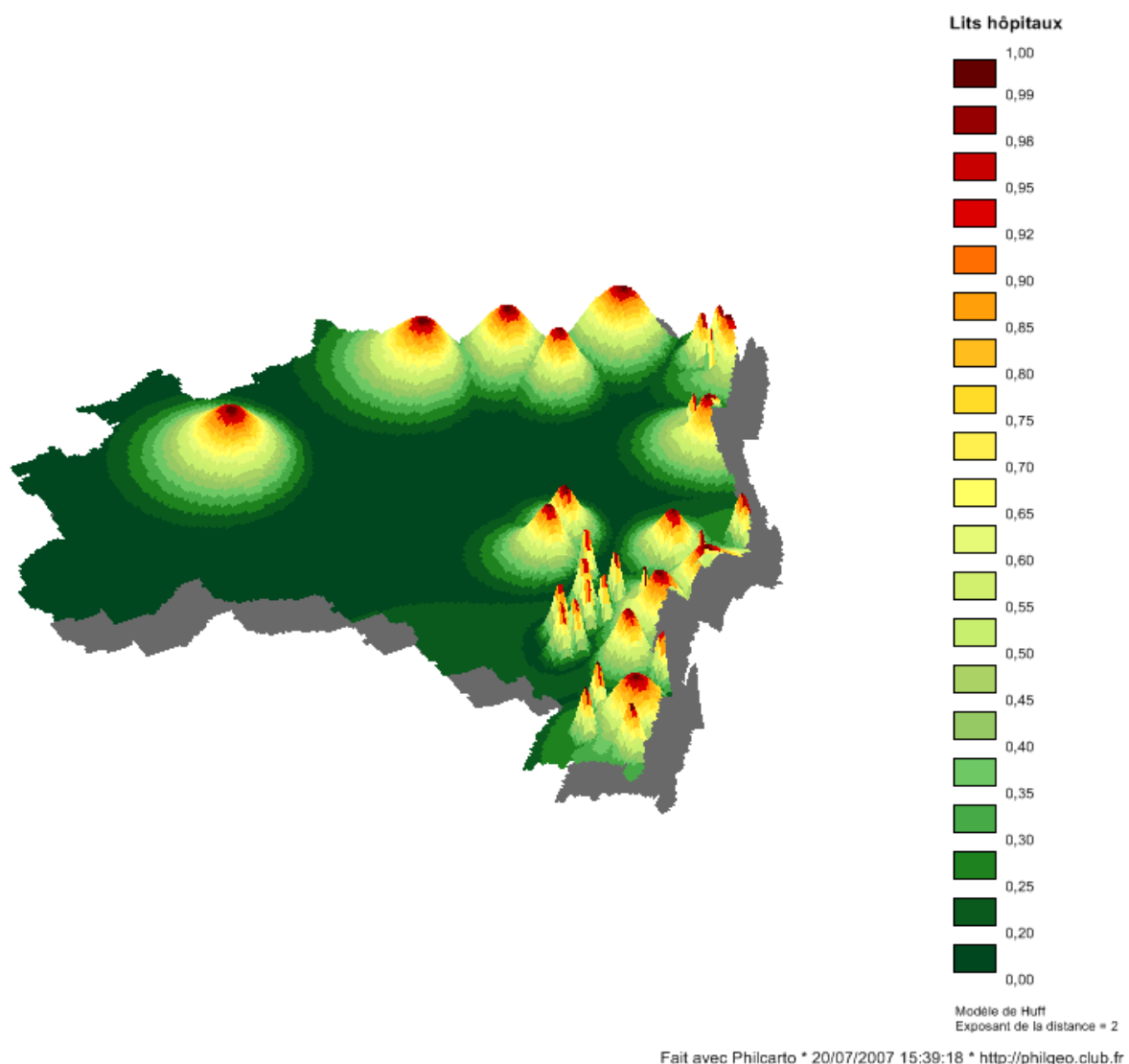


Figure 1.26 : Brésil (capitales des États).
Bloc diagramme des probabilités d'influence vue en direction du Nord-Est.

Les cartes de gravitation issues du modèle de Huff

Les modèles de gravitation donnent une représentation simplifiée des aires d'attraction, c'est-à-dire du partage de l'espace environnant provoqué par l'accumulation des facteurs de centralité en un lieu, dans un espace isotrope.

Par analogie avec la loi de gravitation universelle de Newton, les modèles élémentaires d'attraction urbaine reposent sur trois paramètres : la masse, la distance et l'exposant de la distance.

La masse est un effectif mesuré en chaque lieu central ; il peut s'agir de la population active dans les services à haute valeur ajoutée (conseil financier, ingénierie des réseaux, etc.), du nombre de grands magasins, etc.

La distance est mesurée à vol d'oiseau.

Enfin, la valeur de l'exposant souligne que certains types d'interaction spatiale sont plus sensibles à la distance que d'autres, une valeur élevée traduisant une plus forte résistance au déplacement.

Les points localisant les lieux centraux sont positionnés à l'aide de coordonnées X,Y formant un maillage ponctuel. Un second maillage surfacique ou linéaire peut être superposé au maillage précédent pour former un habillage qui complète la carte et en facilite la lecture.

Les résultats de l'application du modèle de gravitation se traduisent par le tracé des lignes d'iso-attraction entre les centres pris deux à deux : à partir d'un centre donné, le franchissement de cette ligne signifie qu'un autre centre prend le relais et exerce sur l'espace environnant une attraction plus forte que le centre initial. La fermeture des lignes d'iso-attraction délimite les aires d'influence de chaque centre.

Comme les lignes d'iso-attraction ne renseignent pas sur l'intensité des attractions, le modèle de Huff constitue un complément intéressant car il donne en chaque lieu une probabilité d'attraction à l'intérieur de chaque aire d'influence.

Les probabilités d'attraction sont représentées par des cartes en plages de niveaux ou par des blocs-diagrammes.

L'inconvénient fondamental des modèles de gravitation est l'isotropie de l'espace posée comme hypothèse ; le temps de calcul des lignes d'iso-attraction sont relativement long ; les fichiers graphiques des cartes en plages de niveaux sont volumineux.

2. Installer et utiliser *Philcarto*

2.1. Téléchargement et installation

Philcarto peut être téléchargé sur Internet à l'adresse suivante :

<http://philgeo.club.fr>

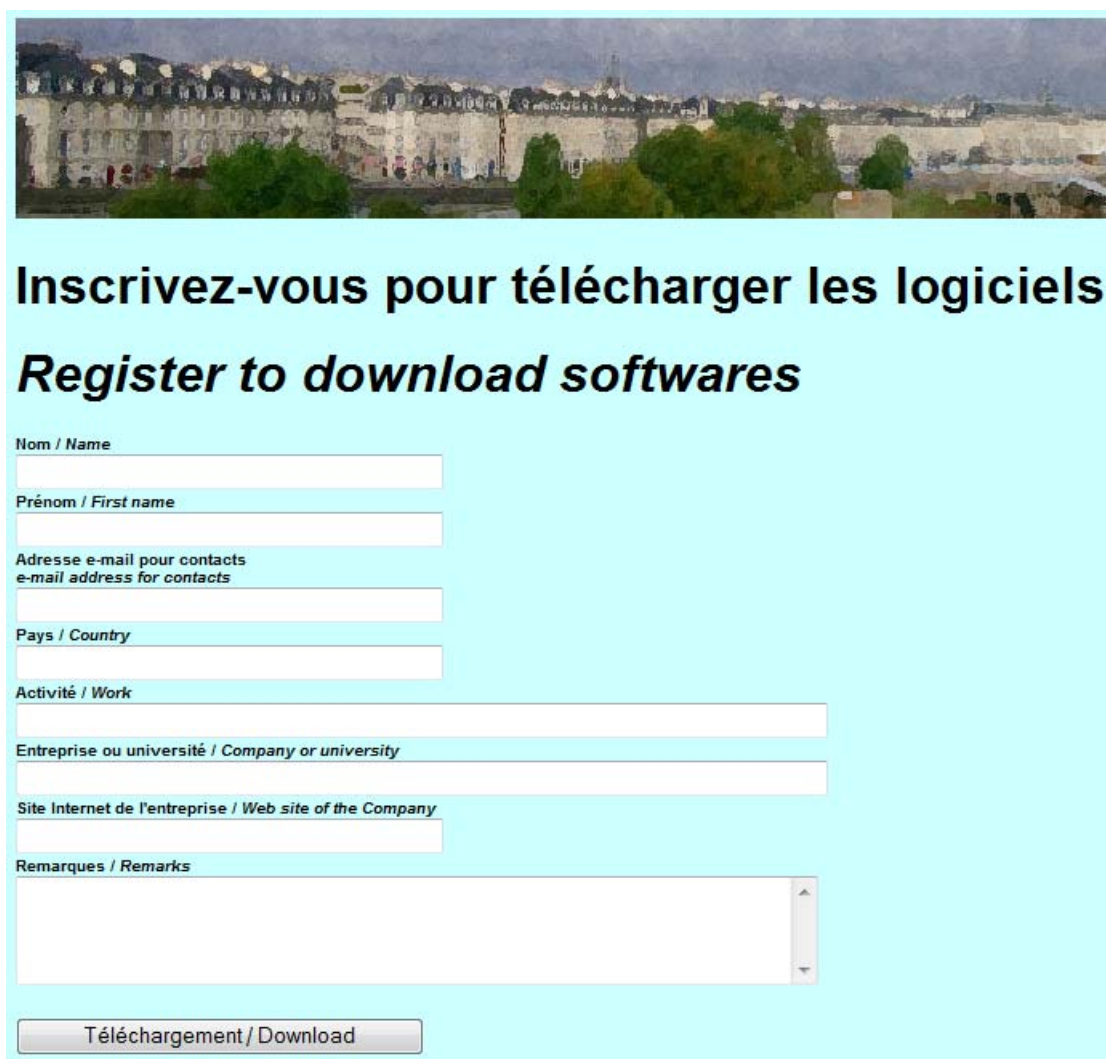
En plus du logiciel, on y trouve également de nombreux fonds de cartes réalisés par des utilisateurs de *Philcarto*, ainsi qu'un programme de digitalisation de fonds de cartes *Phildigit* (cf. chapitre 21) dont l'installation est semblable à celle de *Philcarto* exposée ci-après. Contrairement aux versions précédentes de *Philcarto*, la durée d'utilisation n'est plus limitée. Il est cependant conseillé de visiter de temps à autre le site internet de téléchargement car des corrections sont faites fréquemment.

Cliquer sur le lien :

Logiciels *Philcarto*, *Phildigit*... / Softwares *Philcarto*, *Phildigit*...

Remplir le formulaire d'inscription (fig. 2.1) puis cliquer sur le bouton :

Téléchargement / Download



Inscrivez-vous pour télécharger les logiciels

Register to download softwares

Nom / Name

Prénom / First name

Adresse e-mail pour contacts
e-mail address for contacts

Pays / Country

Activité / Work

Entreprise ou université / Company or university

Site Internet de l'entreprise / Web site of the Company

Remarques / Remarks

Téléchargement / Download

Figure 2.1 : Le formulaire d'inscription pour le téléchargement de *Philcarto*, *Phildigit*...

La page de téléchargement proprement dite s'affiche à l'écran (fig. 2.2). Il est conseillé de visiter cette page de temps à autre car des corrections sont fréquemment apportées aux logiciels (généralement, dès qu'un « bug » est détecté).



Logiciels à télécharger (Windows Vista, XP SP2 et 2000 SP3)
Softwares to download (Windows Vista, XP SP2 et 2000 SP3)

	.Net Framework3 (pour Windows XP SP2 et Windows 2000 SP3 seulement; à installer en premier si ce n'est pas déjà fait) .Net Framework 3 (for windows XP SP2 & Windows 2000 SP3 only: install first if it is not already done)
	Philcarto version 5.xx : cartographie thématique / thematic cartography
	Phildigit version 2.xx : Digitalisation de fonds de cartes / Base maps digitalization
	ShapeSelect version 1.xx : sélection d'éléments dans un fichier .shp / selection of elements in a .shp file
	Anciennes versions des logiciels à télécharger / Old versions of the softwares to download

Lire attentivement la licence d'utilisation de ces logiciels gratuits sans pour autant être "libres".

Figure 2.2 : La page de téléchargement des logiciels

Chacun des liens figurant sur cette page (dont le contenu évoluera à l'avenir) donne accès à une procédure d'installation (sauf pour le lien pointant vers les anciennes versions qui permet de télécharger un fichier **.zip** contenant la procédure d'installation qui doit ensuite être exécutée localement sur l'ordinateur).

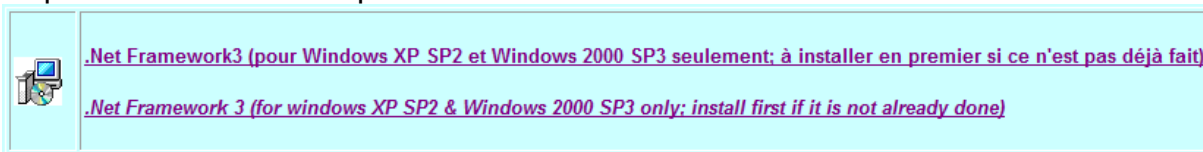
2.1.1. Installation du **.NET Framework**

Philcarto ne peut pas fonctionner si le **.NET Framework** n'est pas préalablement installé sur l'ordinateur. Il s'agit d'un environnement d'exécution des logiciels conçus avec *Visual Studio .NET*. Si vous travaillez sous *Windows Vista*, le **.NET Framework** est déjà installé et vous pouvez passer au §2.1.2.

Si votre système d'exploitation n'est pas *Windows 2000 Service Pack 3* ou *Windows XP Service Pack 2*, vous ne pouvez pas utiliser les dernières versions de logiciels ; dans ce cas, cliquez sur le lien permettant de télécharger les anciennes versions :

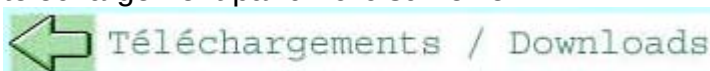
	Anciennes versions des logiciels à télécharger / Old versions of the softwares to download
---	--

Si votre système d'exploitation est *Windows 2000 Service Pack 3* ou *Windows XP Service Pack 2*, vous devez obligatoirement installer le *.NET Framework* en cliquant sur le lien correspondant.



Lire attentivement la page qui décrit l'installation, puis cliquer sur le lien **[.NetFramework](#)** situé en fin de page pour lancer l'installation proprement dite. Suivre ensuite la procédure indiquée qui prend environ une dizaine minutes pour aboutir.

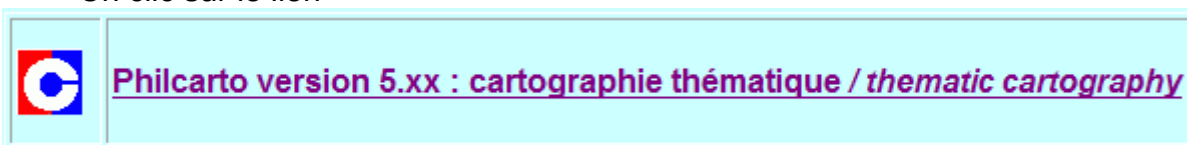
Lorsque le *.NET Framework* est correctement installé revenir à la page de téléchargement par un clic sur le lien



2.1.2. Installation *Philcarto* version 5

Avant d'installer la version 5 de *Philcarto* pour la première fois, il est préférable d'effacer toute éventuelle installation d'une version précédente. Ceci se fait simplement, et de manière habituelle, à l'aide du Panneau de configuration Ajout/Suppression de programme.

Un clic sur le lien



affiche une nouvelle page dans laquelle est présentée la procédure d'installation de *Philcarto*.

Après avoir lu attentivement cette page, cliquer sur le lien **[Philcarto version 5.xx](#)** situé en fin de page pour lancer l'installation proprement dite. Suivre ensuite la procédure indiquée qui prend quelques minutes pour aboutir. Ne rien changer dans les paramètres d'installation.

Lorsque *Philcarto* est correctement installé, vous pouvez télécharger le manuel d'utilisation et les exemples présentés dans ce manuel (les noms des fichiers utilisés pour ces exemples sont indiqués dans des cadres de fond bleu) par un clic sur chacun des deux liens correspondants.

 Les utilisateurs de PC en réseau, les personnes utilisant un ordinateur multi-utilisateurs sont informés que *Philcarto* doit être installé sur chaque ordinateur ; *Philcarto* ne peut pas être installé sur un serveur. Pour simplifier l'installation sur chaque ordinateur d'un réseau, il est préférable de télécharger une seule fois l'installation, de la copier sur chaque ordinateur, et de l'installer ensuite. L'installation doit être faite avec un accès « administrateur ».

2.2. Exécution de *Philcarto*

Pour lancer l'exécution de *Philcarto*, il suffit de double-cliquer sur l'icône du

raccourcis  placé sur le bureau de *Windows* par le programme d'installation. Une autre façon de lancer l'exécution de *Philcarto* consiste à placer un raccourci dans la barre des tâches de *Windows* ; cela se fait très simplement en déplaçant, à l'aide de la souris, l'icône du fichier ***Philcarto.exe*** du répertoire ***C:\Program Files\Philcarto\bin*** vers la barre des tâches. Le raccourci est créé automatiquement ; un seul clic son icône dans la barre des tâches permet ensuite de lancer l'exécution de *Philcarto* (fig. 2.3).



Figure 2.3 : Le raccourci de *Philcarto* est placé dans la barre des tâches.

La fenêtre d'accueil de *Philcarto* s'affiche à chaque lancement du programme (fig. 2.4). Dans le coin supérieur gauche de cette fenêtre, vous pouvez lire le numéro de version (ici 5.01) et la date de compilation qu'il est souhaitable d'indiquer dans toute demande d'aide.



Figure 2.4 : La fenêtre d'accueil de *Philcarto*.

Une fenêtre équipée d'un ascenseur permet de lire la licence d'utilisation. Après avoir pris connaissance des termes de cette licence, et si vous n'êtes pas d'accord

avec ceux-ci, cliquez sur le bouton **Je refuse** ; le programme s'arrêtera. Si vous vous engagez à respecter les termes de la licence, cliquez sur le bouton **J'accepte** et le programme passera à l'étape suivante, celle de l'ouverture des fichiers.

Philcarto fonctionne en Français, Anglais, Espagnol et Portugais. Le Français est sélectionné au lancement du programme ; pour modifier la langue de fonctionnement, il suffit de sélectionner le drapeau correspondant par un clic sur le bouton radio qui le précède.

2.2.1. Ouverture des fichiers fond de carte et données statistiques

Le clic sur **J'accepte** provoque l'affichage du dialogue **Fichier fond de carte ?** (fig. 2.5) permettant de sélectionner un fond de carte dans un répertoire. Le répertoire de base s'appelle **Philcarto alias** ; il est automatiquement créé par *Philcarto* qui le place dans le répertoire **Mes documents** (propre à chaque utilisateur).

Dans ce répertoire **Philcarto alias** sont enregistrés les raccourcis des fichiers « fond de carte » et « données statistiques » ouverts lors des utilisations précédentes de *Philcarto*.

Lors du premier lancement de *Philcarto* ce répertoire est naturellement vide (fig. 2.5). Lors des lancements suivants, et si des fichiers ont été ouverts, les raccourcis sont visibles et peuvent être sélectionnés.

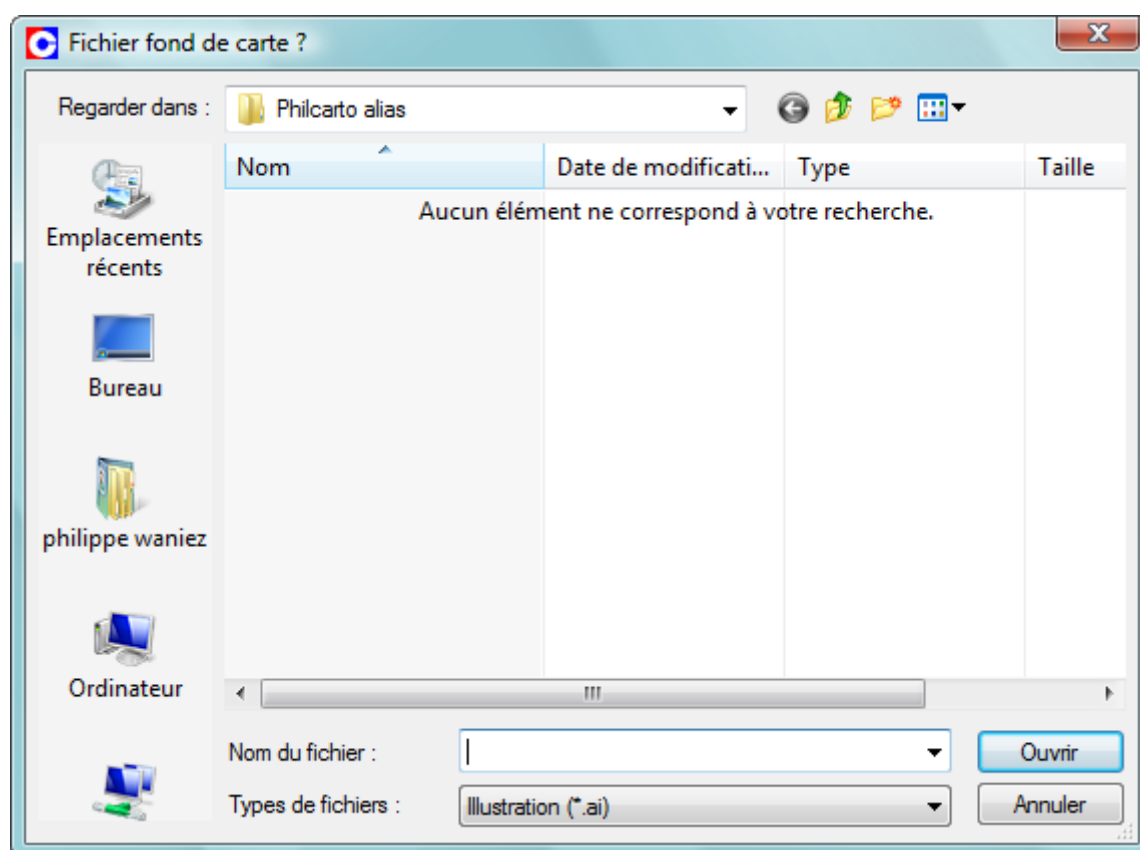


Figure 2.5 : Le dialogue **Fichier fond de carte ?** et le répertoire **Philcarto alias** .

! A tout lancement de *Philcarto*, celui-ci vérifie que chaque raccourci correspond bien au chemin d'un fichier existant. Si un fichier a été déplacé ou

supprimé après la création du raccourci par *Philcarto*, ce raccourci est automatiquement effacé du répertoire  *Philcarto alias*. Autrement dit, on trouve dans le répertoire  *Philcarto alias* des raccourcis qui pointent sur des fichiers fond de carte ou données statistiques existant effectivement dans le chemin enregistré dans le raccourci.

 L'utilisateur reste libre de supprimer, dans  *Philcarto alias*, les raccourcis qu'il juge inutiles, avant de lancer, ou après avoir quitté *Philcarto* et même lors du dialogue d'ouverture des fichiers par un clic droit sur l'icône ou le nom du fichier suivi de la sélection de l'article **Supprimer** du menu local (fig. 2.6).

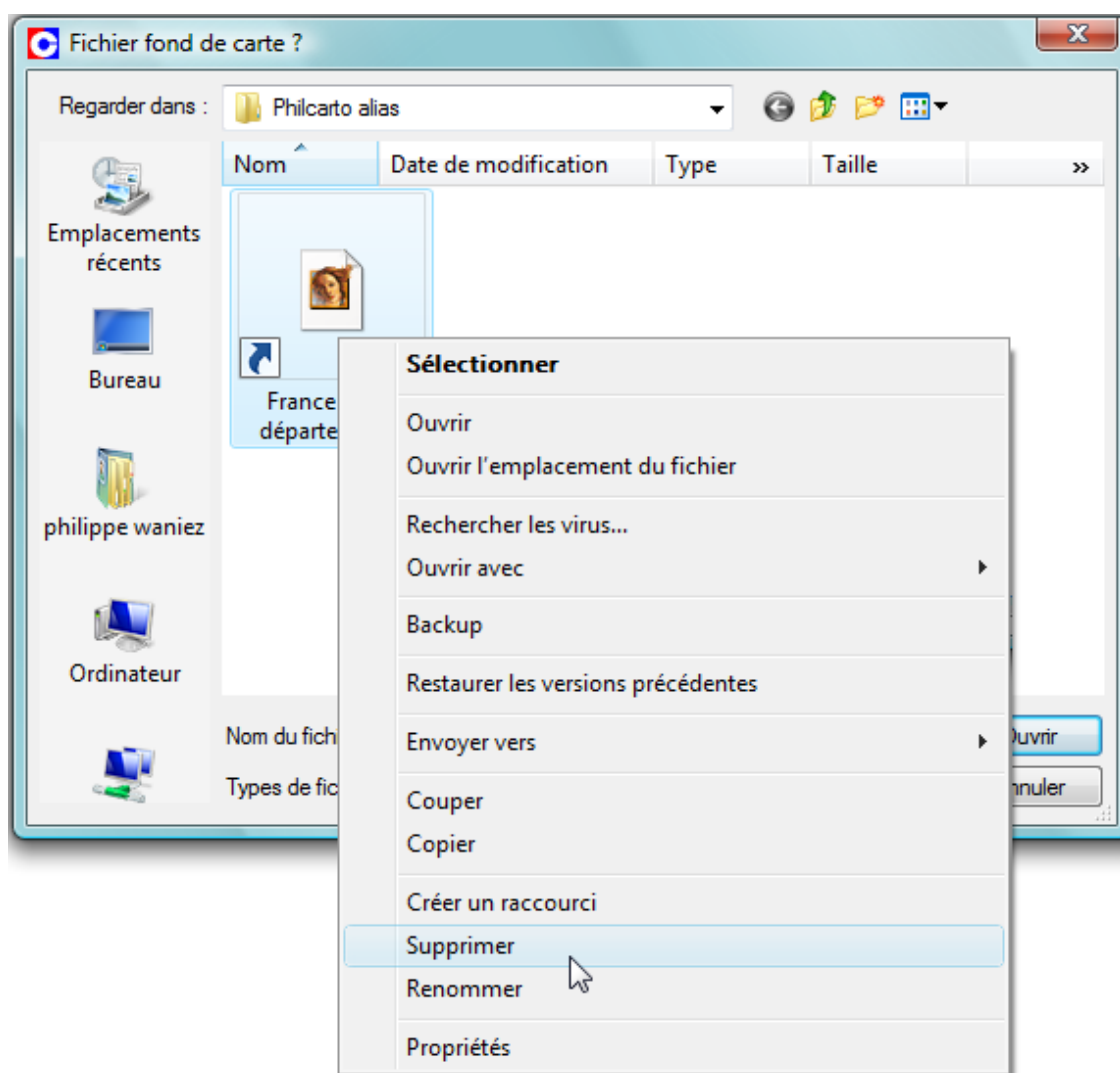


Figure 2.6 : La suppression d'un raccourci dans le dialogue  *Fichier fond de carte ?*.

 Le type de fichier fond de carte est obligatoirement « Illustration » ; il possède le suffixe **.ai** (signifiant *Adobe Illustrator*), cf. chapitre 21.

Si le répertoire  *Philcarto alias* est vide, ou si les raccourcis qui s'y trouvent ne correspondent pas au fichier fond de carte désiré, il faut naviguer dans les

répertoires des disques accessibles pour y rechercher le fichier fond de carte souhaité. Cela se fait à l'aide du menu déroulant situé sur la ligne Regarder dans : situé dans la partie supérieure gauche du dialogue  **Fichier fond de carte ?**. Par exemple, si les fichiers sont situés dans le dossier **\\Philcarto exemples\France départements**, il faut explorer le disque jusqu'à ce répertoire pour y trouver ce fond de carte (fig. 2.7).

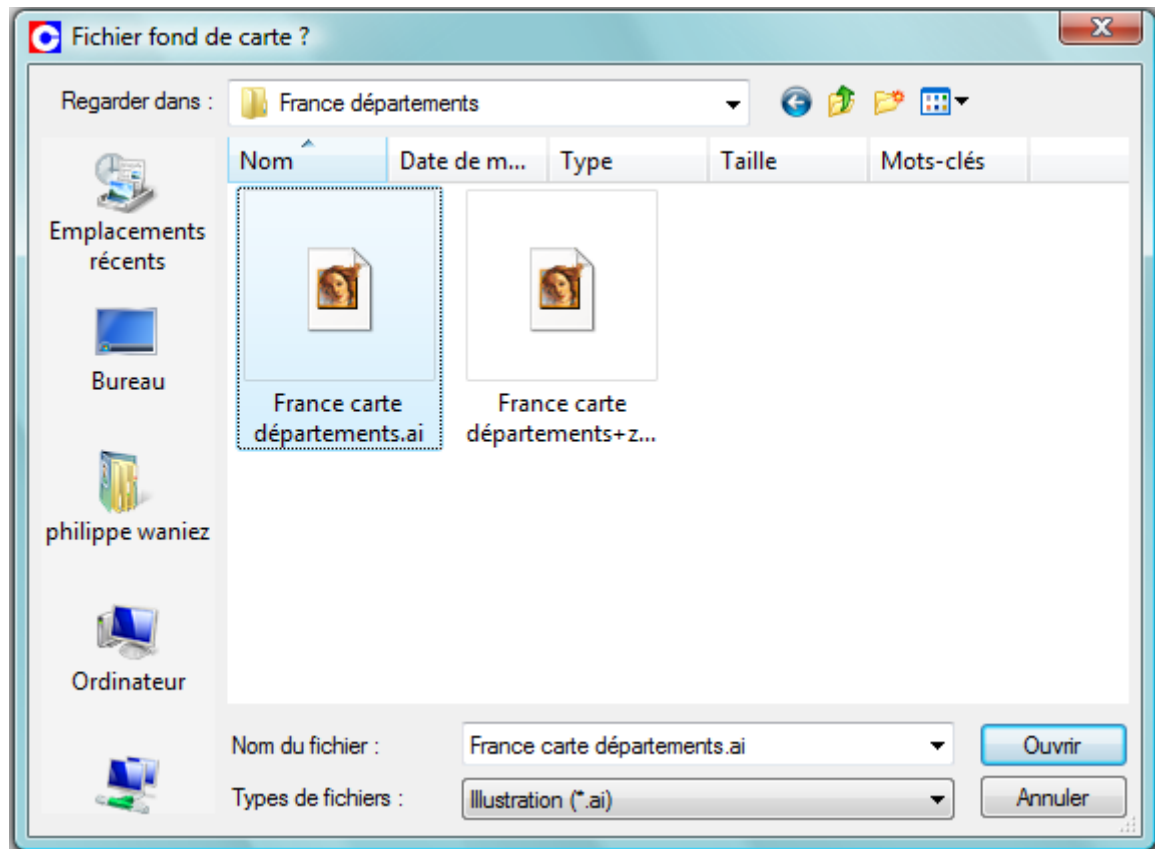


Figure 2.7 : Le dialogue  **Fichier fond de carte ?** et le répertoire **\\Philcarto exemples\France départements**.

Dans le répertoire **\\Philcarto exemples\France départements**, on trouve deux fichiers fond de carte au format **.ai** contenant les départements de France métropolitaine. Ils diffèrent par la présence ou l'absence de deux cartouches contenant Paris et la petite couronne ainsi que les départements de l'Ouest de la région parisienne (fig. 2.8).

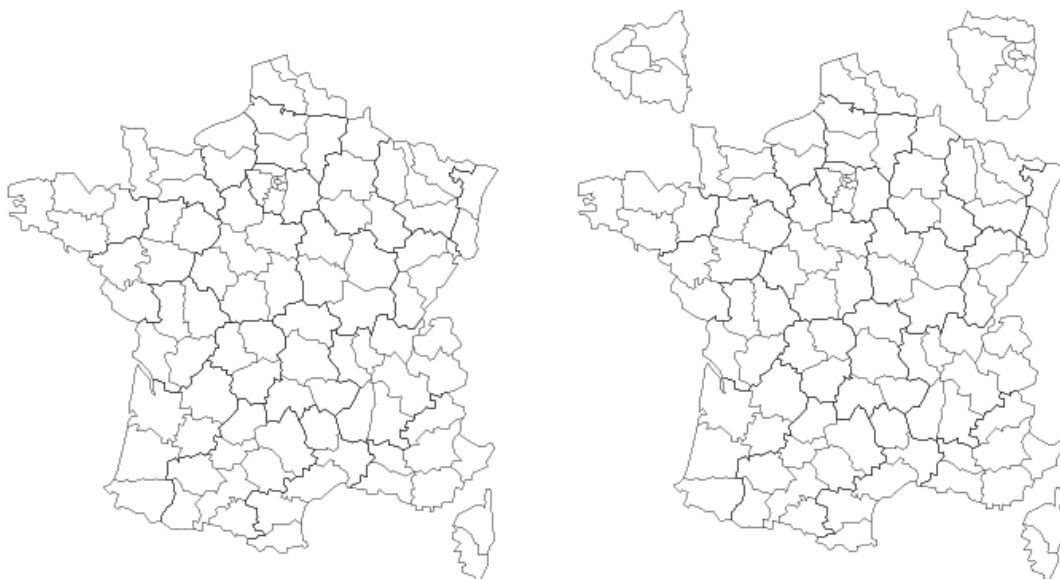


Figure 2.8 : Tracé des deux fonds de cartes : à gauche **France carte départements.ai** et à droite **France carte départements+zoom IDF.ai**.

On sélectionne le fichier **France carte départements.ai** (le fond de carte sans cartouche) par un double clic sur son icône ou un simple clic sur son icône suivie d'un clic sur **Ouvrir** (fig. 2.9).

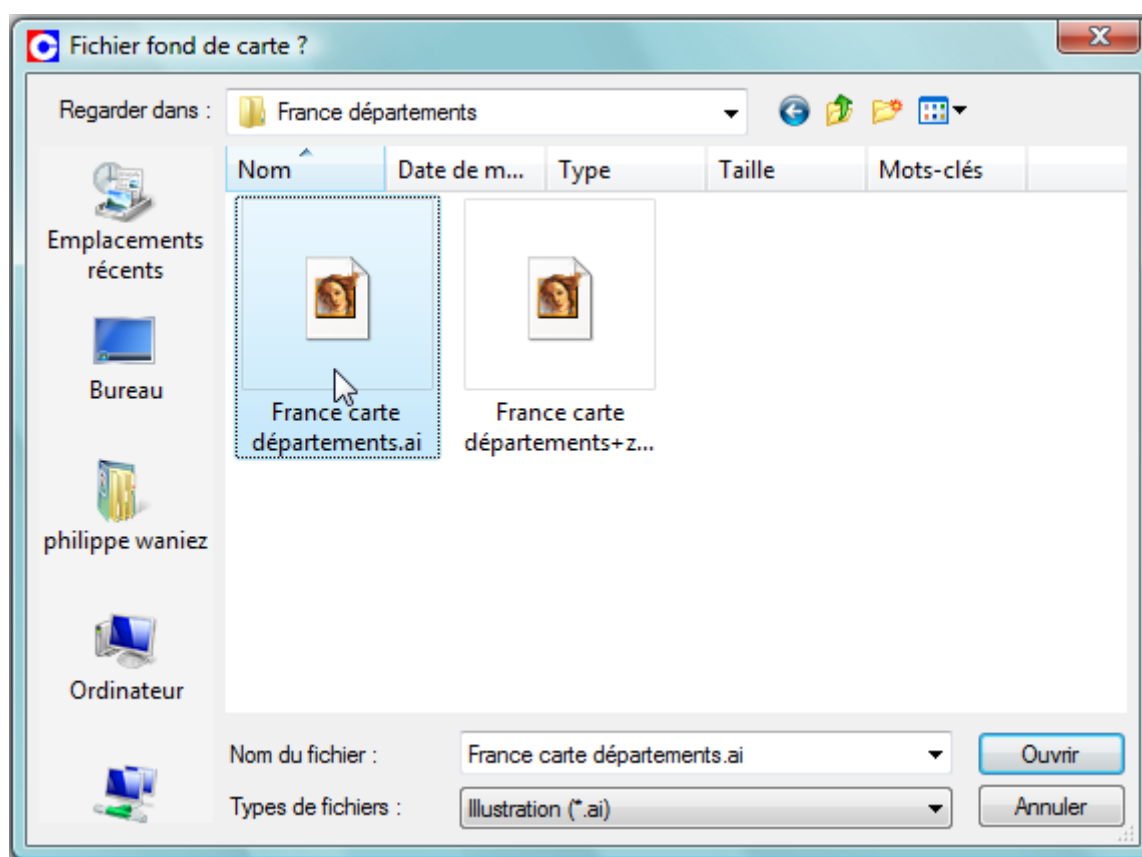
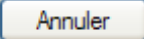


Figure 2.9 : La sélection du fichier fond de carte **France départements .ai**.

Dès que le fichier **France départements.ai** est sélectionné, *Philcarto* place un raccourci de ce fichier dans le répertoire  *Philcarto alias* comme indiqué précédemment. Au prochain lancement de *Philcarto* (c'est-à-dire à la session suivante), on pourra accéder directement à ce fichier s'il n'a pas été effacé ou déplacé entre temps.

Un clic sur  dans le dialogue  *Fichier fond de carte ?* ferme la session de *Philcarto* et provoque le retour à *Windows* (fig. 2.10).

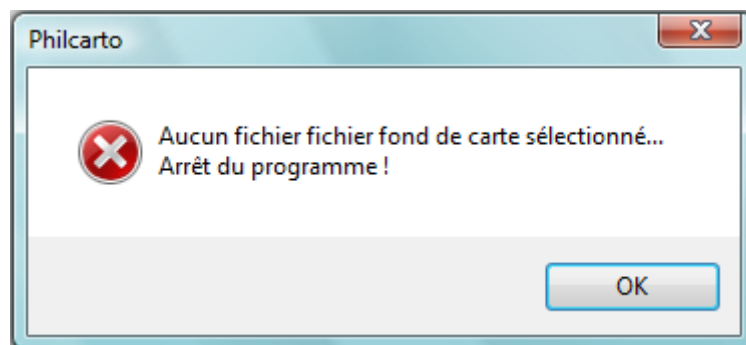


Figure 2.10 : L'alerte de fermeture du programme si aucun fichier fond de carte n'est ouvert.

Si un fichier fond de carte a été sélectionné, *Philcarto* affiche un nouveau dialogue intitulé  *Fichier données statistiques ?* (fig. 2.11) permettant de sélectionner un fichier de données statistiques.

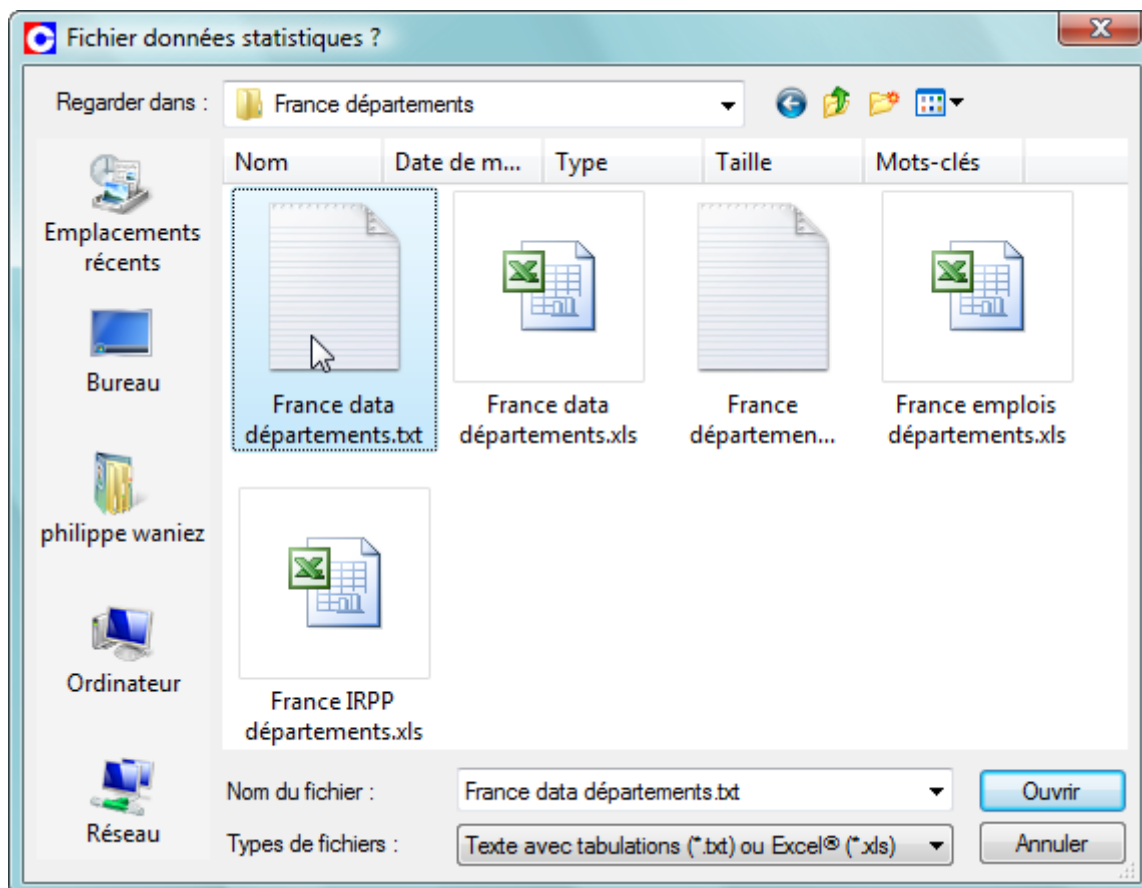
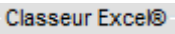
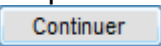


Figure 2.11 : La sélection du fichier de données statistiques **France départements logements.txt**.

Dans le dialogue  Fichier données statistiques ?, deux formats de fichiers sont disponibles : **.txt** qui signifie fichier texte, et **.xls** qui signifie fichier Excel. Ces deux formats de fichier (et seulement ces deux là) peuvent être utilisés avec *Philcarto* .

 Les utilisateurs de la suite *OpenOffice* doivent donc enregistrer les documents créés avec *Calc* au format **.xls**.

Les fichiers de format **.xls** sont des classeurs pouvant contenir plusieurs feuilles, chaque feuille correspondant à un tableau de données statistiques. Si le classeur contient effectivement plusieurs feuilles, le dialogue  **Classeur Excel**® liste les noms des différentes feuilles et attend de l'utilisateur qu'il en choisisse une en cliquant sur son nom, puis valide ce choix à l'aide avec  (fig. 2.12).

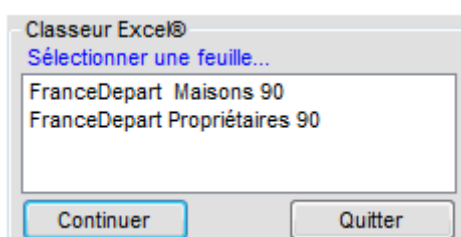
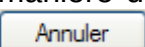


Figure 2.12 : Liste des feuilles présentes dans un classeur Excel.

Les caractéristiques techniques des tableaux de données statistiques et la manière de créer des fichier **.txt** ou **.xls** sont exposées au chapitre 20. Un clic sur  ferme la session et provoque le retour à *Windows* (fig. 2.13).

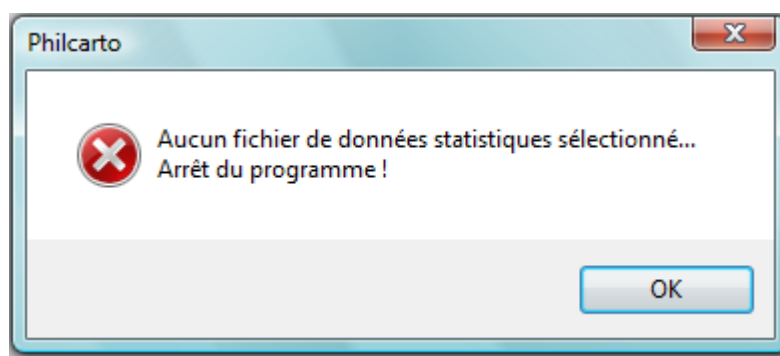


Figure 2.13 : L'alerte de fermeture du programme si aucun fichier de données statistiques n'est ouvert.

Sélectionnons le fichier **France data départements.txt** (le fichier au format texte). Lorsque le fichier fond de carte et le fichier de données statistiques sont sélectionnés dans les deux dialogues successifs, *Philcarto* réalise trois traitements quasiment invisibles pour l'utilisateur (sauf lorsque ces fichiers sont très volumineux. Il lit successivement les deux fichiers et réalise ce que les spécialistes appellent une jointure, c'est-à-dire, la jonction des données contenues dans ces deux fichiers en fonction des valeurs des identifiants des éléments cartographiques qui doivent figurer obligatoirement dans ces deux fichiers. Si *Philcarto* échoue dans la jonction de ces données, un message s'affiche, la session est fermée et l'on retourne à *Windows* (fig. 2.14).

Si ces trois opérations se terminent avec succès (ce qui est le plus souvent le cas...), *Philcarto* affiche un nouveau dialogue intitulé **Types des variables** (fig. 2.15). Cette fenêtre est nouvelle dans la version 5 de *Philcarto*. Elle contient, pour chacune des variables du fichier de données statistiques, son type et son nom. Les variables peuvent être de 3 types différents : effectif (ou quantité...) listée sur fond vert, rapport (ou taux, pourcentage...) listé sur fond bleu, et nominal, listé sur fond rose.

Si le type de variable est donné explicitement dans le suffixe du nom de variable figurant dans le fichier de données statistiques (voir chapitre 19), *Philcarto* lit ce type. En l'absence d'un tel suffixe, *Philcarto* détecte automatiquement le type de variable à partir des valeurs que renferme effectivement chaque variable. Nous reviendrons sur cet aspect des choses au chapitre 20. Notons pour l'instant que le dialogue **Types des variables** est affiché seulement en cas de détection automatique des types des variables, c'est-à-dire en l'absence de suffixe ; il permet de modifier le type, au cas où la détection aurait donné un résultat erroné : il suffit de cliquer sur une ligne de nom de variable, dans le champ **Type**, pour passer au type suivant.

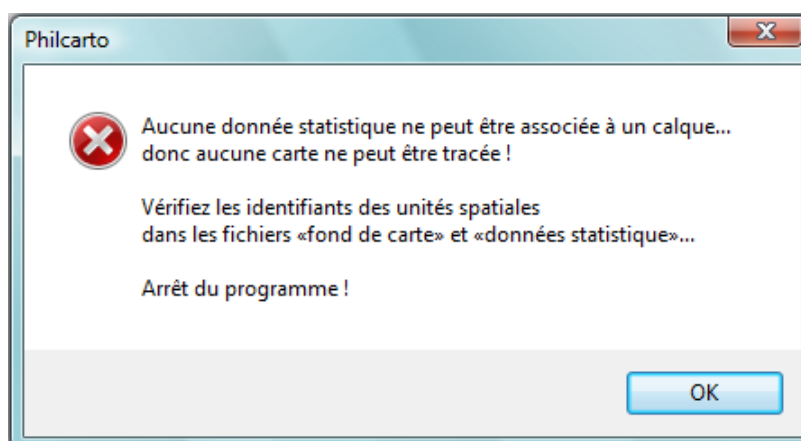


Figure 2.14 : L'alerte de fermeture du programme si la jonction du fond de carte et des données statistiques a échoué.

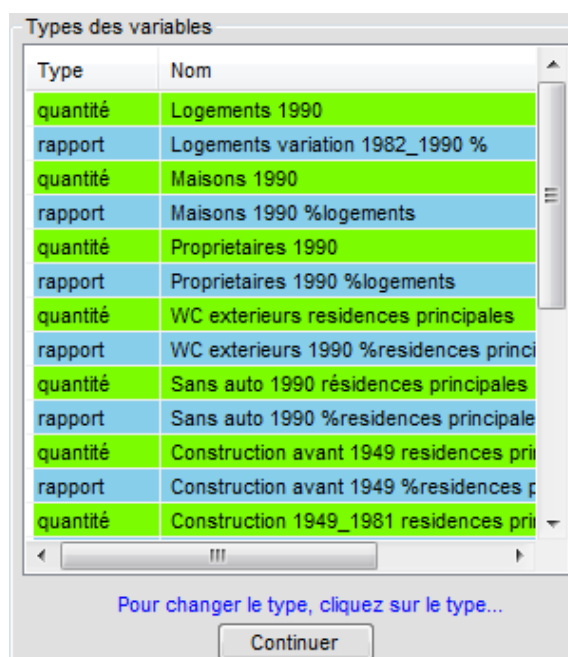
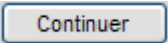


Figure 2.15 : Le dialogue **Types des variables**.

Un clic sur  provoque l'affichage du bureau de *Philcarto*.

2.2.2. Le bureau de *Philcarto*

Le bureau de *Philcarto* occupe la totalité de l'écran (fig. 2.16). Il se compose de 3 éléments principaux : 1° l'espace d'affichage des cartes représentant 4/5 de la largeur de l'écran à partir de la gauche, 2° le cadre **Outils** occupant le quart restant à droite, et 3° la barre de menus classiquement située dans la partie supérieure gauche de l'écran. Enfin, les boutons habituels  (réduire) et  (fermer) situés dans le bandeau supérieur du bureau de *Philcarto*, à droite, complètent le dispositif.

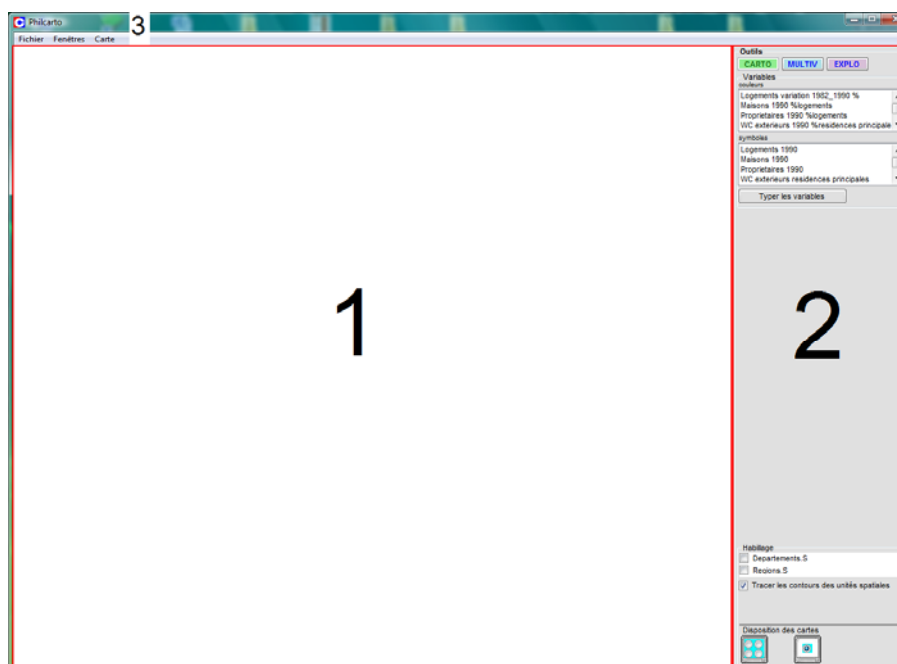


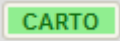
Figure 2.16 : Le bureau de *Philcarto*.

C'est généralement par des actions dans le cadre **Outils** (2) que les cartes sont conçues. Le contenu du cadre **Outils** change en fonction du contexte et permet à l'utilisateur de commander une carte ; celle-ci s'affiche alors dans l'espace d'affichage (1). L'utilisateur peut ensuite modifier la carte affichée à l'aide d'options adaptées au type de carte qu'il vient de réaliser (par exemple, changer la gamme de couleurs d'une carte choroplèthe, changer le calibrage des cercles d'une carte en cercles proportionnels...), options qui sont, elles aussi, choisies dans le cadre **Outils** dont le contenu s'adapte au contexte.

Les menus (3) servent à réaliser des traitements particuliers : ouvrir d'autres fichiers, calculer des indicateurs statistiques nouveaux, enregistrer la carte dans un fichier pour une utilisation ultérieure (importation dans *Word*, par exemple).

2.2.3. La réalisation d'une carte

Dans la partie supérieure du cadre **Outils** se trouvent trois boutons de couleurs différentes :

 donne accès au module de cartographie proprement dit ; c'est ce module là qui est activé lors du lancement de *Philcarto* ;

 conduit au module d'analyse multivariée : analyse factorielle et classification automatique ; ce module produit des facteurs ou des classes sous la

forme de nouvelles variables, mais ne trace pas de carte ; la cartographie des facteurs ou des classes implique un retour au module **CARTO** .

EXPLO propose un environnement d'analyse exploratoire des données sous forme de graphique bivarié ou de diagramme triangulaire interactifs permettant des sélections sur graphique immédiatement répercutées sur la carte.

Lorsque le module **CARTO** est actif, le dialogue **Variables** , situé sous les trois boutons d'accès aux modules, renferme deux listes déroulantes (fig. 2.17). La présence d'une variable dans l'une de ces deux listes dépend de son type. Autrement dit, une même variable ne peut pas figurer simultanément dans les deux listes. Dans la liste **couleurs** , on trouve les variables de type rapport ou nominal. Dans la liste **symboles** apparaissent seulement les variables de type quantité.

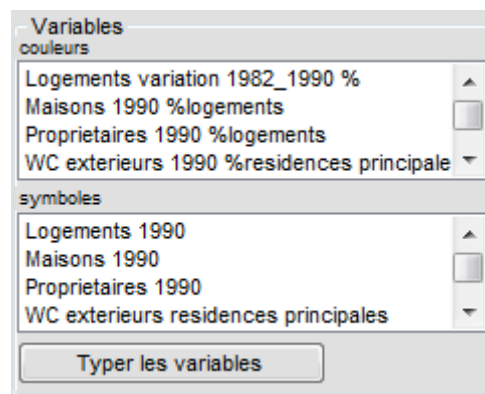


Figure 2.17 : Le dialogue **Variables** .

La sélection d'une variable soit dans la liste **couleurs** , soit dans la liste **symboles** , ou bien d'une variable dans chacune de ces deux listes provoque l'affichage, dans le dialogue **Types de cartes** , des icônes de types de cartes effectivement réalisables compte tenu des éléments cartographiques enregistrés dans le fond de carte (dans ce qui suit il existe dans le fond de cartes des polygones permettant de réaliser des cartes choroplèthes, et des points pouvant servir de centres pour les cartes en cercles proportionnels (fig. 2.18, 2.19 et 2.20).

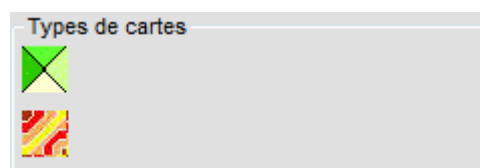


Figure 2.18 : Le dialogue **Types de cartes** lorsqu'une variable est sélectionnée dans la liste **couleurs** .

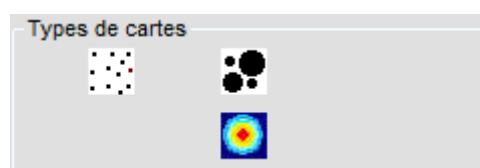


Figure 2.19 : Le dialogue **Types de cartes** lorsqu'une variable est sélectionnée dans la liste **symboles** .

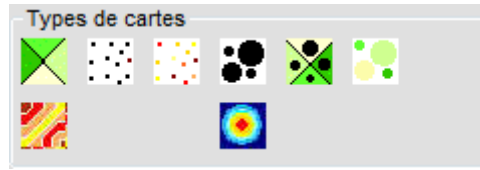


Figure 2.20 : Le dialogue **Types de cartes** lorsqu'une variable est sélectionnée dans la liste **couleurs**, et une seconde variable dans la liste **symboles**.

Le type de carte est indiqué clairement lorsque la flèche de la souris passe sur l'icône, par exemple, pour une carte choroplèthe (fig. 2.21) :

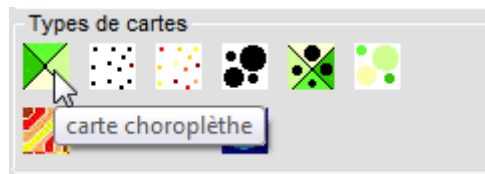


Figure 2.21 : Lorsque la flèche de la souris passe sur l'icône d'un type de carte, ce type est clairement indiqué.

Il suffit alors d'un clic sur l'icône du type de carte souhaité pour qu'une carte soit tracée, et pour qu'un dialogue proposant des outils spécifiques à la réalisation de ce type de carte là s'affiche dans le cadre **Outils**.

3. Les cartes choroplèthes

Dossier : \Philcarto exemples\France départements\
Fond de carte : France carte départements.ai
Statistiques : France data départements.txt

3.1. Sélectionner une variable de couleur

La sélection d'une variable se fait simplement par un clic sur le nom d'une des variables figurant dans la liste **couleurs** du dialogue **Variables**. Ici la variable sélectionnée, **Maisons 1990 %logements**, contient le pourcentage de maisons individuelles dans l'ensemble des logements de chaque département de France métropolitaine au recensement de 1990 (fig. 3.1).

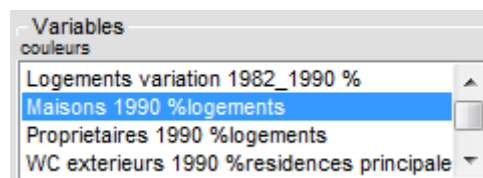


Figure 3.1 : Sélection d'une variable de couleur.

3.2. Afficher une carte choroplèthe

La sélection de la variable de couleur provoque l'affichage du dialogue **Types de cartes**. Dans ce dialogue, cliquer sur l'icône du type carte choroplèthe  (fig. 3.2). La carte s'affiche alors à l'écran (fig. 3.3). Si la méthode de discrétisation n'a pas été changée en cours de session, la discrétisation de la variable choisie est du type  (cf. §3.4).

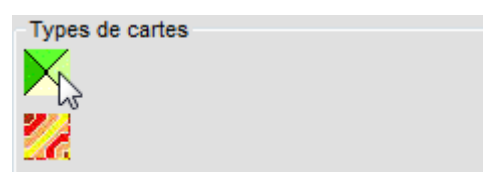


Figure 3.2 : Clic sur l'icône « carte choroplèthe ».

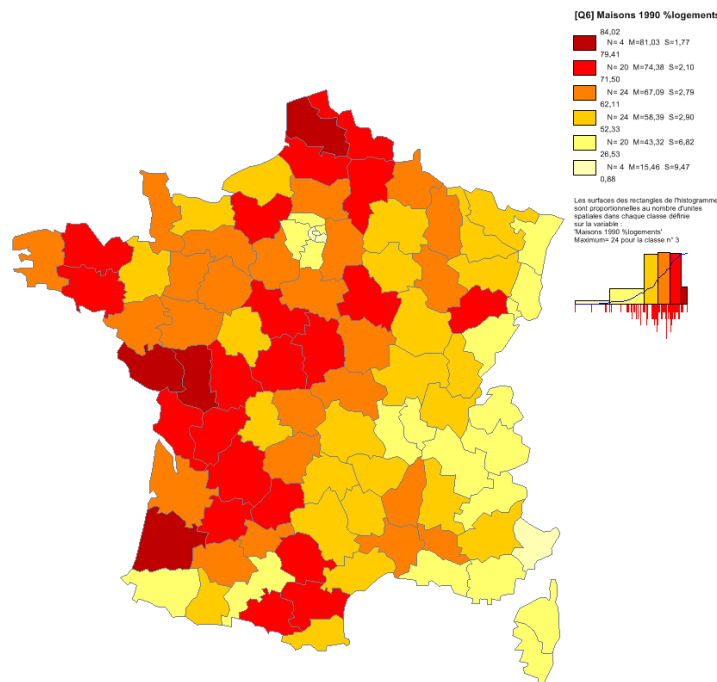


Figure 3.3 : Tracé de la carte choroplèthe avec discrétisation Q6.

3.3. Le dialogue Couleurs

Simultanément au tracé de la carte, les dialogues **couleurs** et **Autocorrélation spatiale** remplacent les dialogues **Variables** et **Types de cartes** dans le cadre **Outils**. (fig. 3.4). Dans le dialogue **couleurs**, des boutons permettent de changer la méthode de discrétisation et la gamme de couleurs, ainsi que de choisir d'autres options qui seront présentées plus loin.

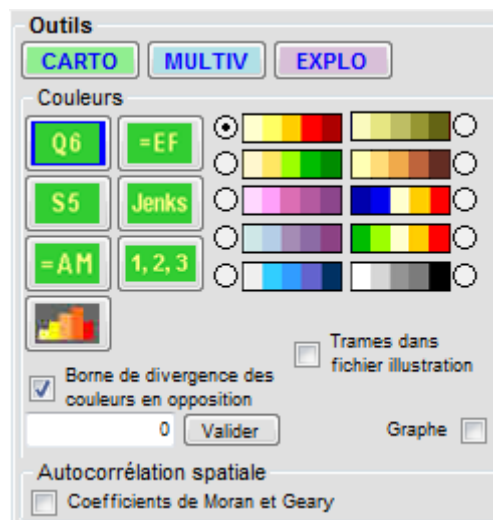


Figure 3.4 : Affichage du dialogue **couleurs**.

3.4. Les différentes méthodes de discrétisation

Lors du lancement de *Philcarto*, la méthode de discrétisation **Q6** est sélectionnée (un cadre bleu entoure son icône). Un clic sur l'un des autres boutons

de discrétisation (à fond vert) permet de changer de méthode, par exemple  pour appliquer la discrétisation de Jenks. Le bouton  est alors sélectionné (une bande bleue apparaît de chaque côté de son icône) pendant que le bouton  est désélectionné (il perd ses bandes bleues). La discrétisation  s'applique alors ensuite à toutes les cartes choroplèthes ou en symboles colorés, tant qu'une autre discrétisation n'est pas sélectionnée par un clic sur le bouton correspondant. Après avoir quitté *Philcarto* et relancé le programme, c'est à nouveau  qui est sélectionnée par défaut. Tout changement de discrétisation se traduit par un nouveau tracé de carte pour laquelle la discrétisation nouvelle s'applique (fig. 3.5). Compte tenu de l'impact de la méthode de discrétisation, cette nouvelle carte peut apparaître assez différente de la précédente...

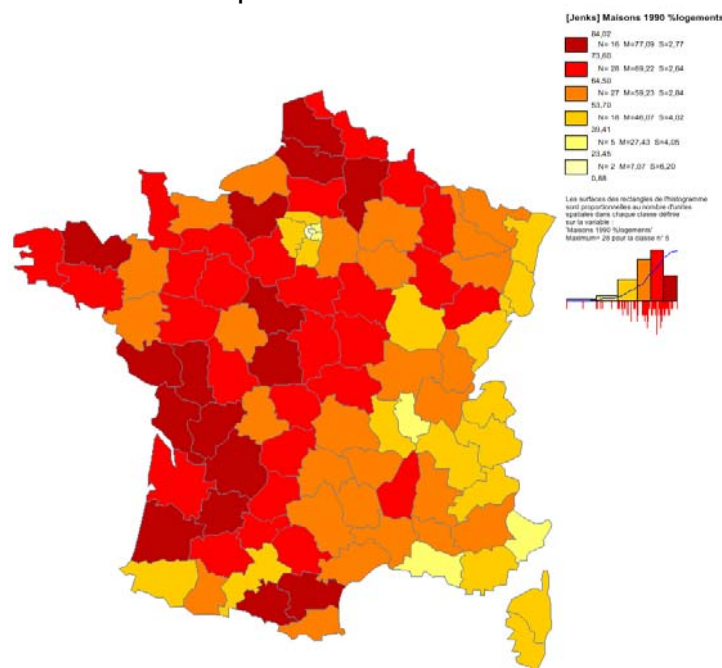


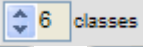
Figure 3.5 : Tracé de la carte choroplèthe avec discrétisation de Jenks.

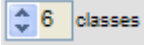
Philcarto propose six méthodes de discrétisation. Voici la signification des icônes sur lesquelles il suffit de cliquer pour changer de mode de discrétisation.

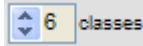
 Discrétisation selon les quartiles avec isolement des queues de la distribution par définition de deux classes extrêmes. Les 6 classes sont définies avec les bornes suivantes : minimum, percentile 5, 1er quartile, médiane, 3e quartile, percentile 95, maximum. Cette discrétisation permet de séparer 5% des unités spatiales ayant les valeurs les plus petites, et 5% les plus grandes.

 Discrétisation sur variable standardisée en 5 classes. Les bornes des classes sont calculées de la manière suivante : minimum, moyenne -1 écart-type, moyenne -0.5 écart type, moyenne $+0.5$ écart-type, moyenne $+1$ écart-type, maximum. En toute rigueur, cette technique ne devrait être employée que pour des

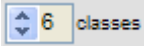
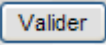
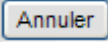
variables présentant une distribution unimodale symétrique dont l'axe de symétrie correspond à la moyenne arithmétique.

 Discrétisation en N classes d'égale amplitude. Le nombre de classes, N, peut varier de 2 à 10 avec le curseur  qui apparaît quand on sélectionne les modes de discrétisation ,  ou .

 Discrétisation en N classes d'égale effectif. Le nombre de classes, N peut varier de 2 à 10 avec le curseur . Si le nombre d'unités spatiales est un multiple de N, et s'il n'y a pas de donnée manquante ou d'ex-æquo, les classes contiennent le même nombre d'unités spatiales. Si N est égal à 4, la discrétisation correspond aux quartiles (chaque classe contient alors 25% du nombre total d'unités spatiales, sauf s'il y a des ex-aequos).

 Discrétisation proposée par Jenks (1977) en N classes. Le nombre de classes, N, peut varier de 2 à 10 avec le curseur . Par une procédure itérative (qui peut s'avérer assez longue lorsque le nombre d'unités spatiales est grand), la discrétisation de Jenks vise à minimiser la variance intra-classes et à maximiser la variance inter-classes, ce qui a pour effet de produire des classes homogènes.

 Par définition, les classes issues de la discrétisation de Jenks sont discontinues. Pour des raisons d'homogénéité des légendes, *Philcarto* indique seulement les bornes inférieures des classes successives (sauf pour la borne correspondant au maximum, bien entendu).

 Discrétisation selon des bornes fixes définies à l'avance par l'utilisateur. Cette solution a été suggérée par un utilisateur qui avait à réaliser une série de cartes avec les mêmes bornes de classes. Le dialogue  permet de saisir les bornes qui deviennent des bornes par défaut tant qu'un autre mode de discrétisation n'a pas été choisi (fig. 3.6). Le nombre de classes peut varier de 2 à 10 avec le curseur . Un clic sur  valide la saisie et efface ce dialogue. Toutes les cartes suivantes utiliseront les mêmes bornes. Le bouton  réinitialise l'ensemble du dialogue alors que le bouton  annule ce mode de discrétisation (c'est la méthode de discrétisation active avant qu'on ait cliqué sur  qui redevient active).

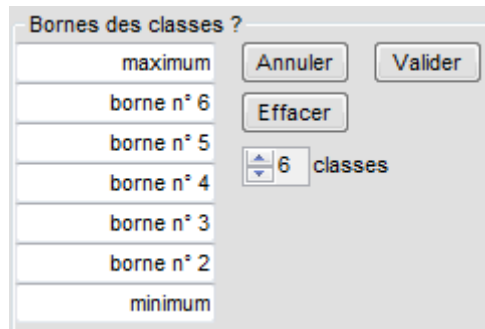


Figure 3.6 : Le choix des bornes des classes
avec le mode de discrétisation **1, 2, 3**.

3.5. Modifier les bornes des classes sur l'histogramme

Quelle que soit la méthode de discrétisation retenue, il est possible de modifier les bornes des classes en agissant directement sur l'histogramme de la variable cartographiée. Les cartographe appellent cette pratique un seuillage à vue. Son principe consiste à ajuster les bornes des classes calculées aux éventuelles ruptures dans la distribution des valeurs visible sur un histogramme de fréquences très détaillé (d'une centaine de classes).

Un clic sur  affiche, dans le cadre **Outils**, le dialogue **Histogramme** contenant un histogramme modifiable de manière interactive (fig. 3.7).

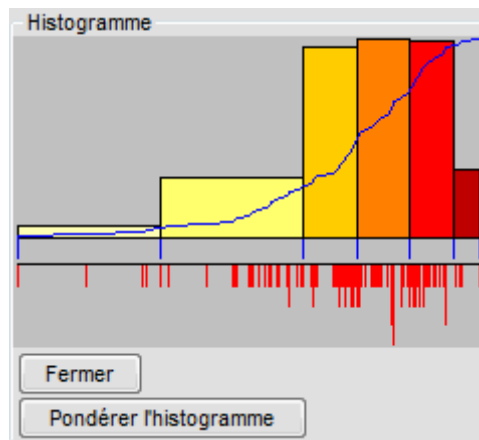


Figure 3.7 : Le dialogue **Histogramme**.

Le dialogue **Histogramme** comprend deux parties :

- les deux tiers supérieurs sont occupés par le graphique formé de l'histogramme proprement dit sur lequel est dessinée la courbe cumulative ; au dessous de l'histogramme figure l'histogramme de fréquences détaillé qui permet de repérer des ruptures dans le continuum des valeurs ;
- la partie inférieure du dialogue comprend des éléments d'interface (boutons, zone d'édition) qui s'affichent en fonction du contexte.

Les bornes des classes sont figurées par des lignes bleues verticales correspondant aux minimums et maximums des rectangles correspondant à chaque classe. Pour changer l'emplacement d'une borne, deux façons de faire peuvent être employées :

- la première consiste à cliquer sur une ligne bleue figurant une borne ; cette ligne prend la couleur magenta qui signifie qu'elle peut être déplacée ; tout en maintenant le bouton de la souris, cette ligne magenta peut être déplacée jusqu'à l'endroit désiré, c'est-à-dire au lieu d'une rupture sur l'histogramme de fréquences (fig. 3.8). Noter qu'une zone d'édition contenant la valeur correspondant à la position de la borne en cours de déplacement apparaît dans la partie inférieure du dialogue ;

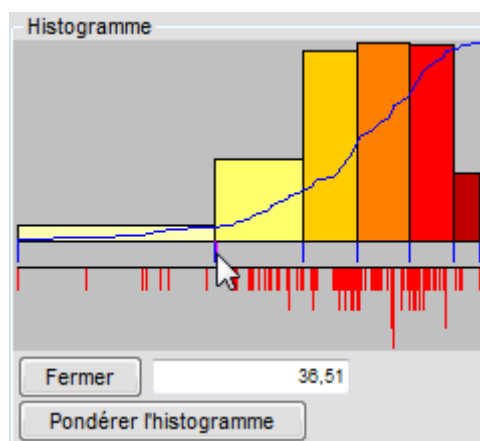


Figure 3.8 : L'histogramme interactif après déplacement des bornes.

- la seconde manière pour changer une borne de classe revient à modifier directement sa valeur dans la zone d'édition qui s'affiche quand on clique sur l'une des lignes bleues figurant cette borne. On entre alors la nouvelle valeur de la borne (ici 36) et on la valide par un clic sur **Appliquer** (fig. 3.9). Naturellement, on peut utiliser les deux méthodes l'une à la suite de l'autre, la première pour déplacer la borne à l'endroit désiré, puis la seconde pour arrondir la valeur de manière à éviter les décimales inutiles.

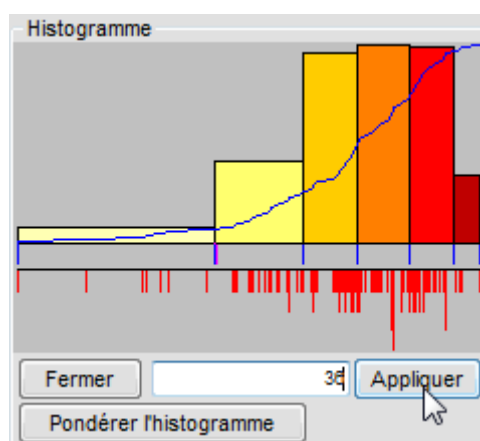


Figure 3.9 : Saisir directement la valeur d'une des bornes de l'histogramme.

L'histogramme qui est tracé par défaut représente le nombre de valeurs dans chaque classe (ici, le nombre de départements). Ceci peut apparaître peu pertinent aux yeux de certains utilisateurs qui préfèrent représenter un effectif en rapport avec la variable cartographiée, comme par exemple, le nombre de maisons individuelles quand c'est leur pourcentage dans l'ensemble des logements qui fait l'objet de la carte choroplèthe. Si ce nombre est enregistré dans une des variables de quantité du

fichier de données statistiques utilisé, un clic sur **Pondérer l'histogramme** permet de sélectionner cette variable dans le dialogue **Variable de pondération ?** (fig. 3.10). Cette liste contient seulement les variables de quantité.

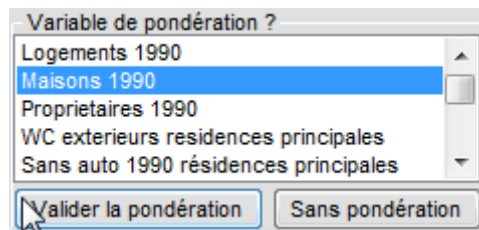


Figure 3.10 : Sélectionner une variable de pondération de l'histogramme.

Un clic sur **Valider la pondération** conduit au tracé de l'histogramme pondéré (fig. 3.11). Le bouton **Sans pondération** permet d'annuler la pondération en cours ou précédemment effectuée. La pondération de l'histogramme n'affecte en rien la carte.

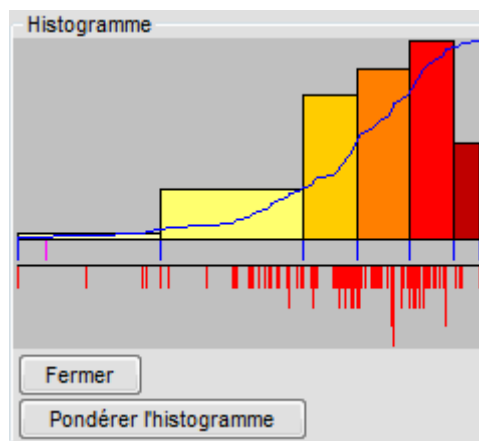


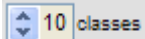
Figure 3.11 : L'histogramme pondéré
A comparer à celui de la figure 3.7.

Enfin, dans le dialogue **Histogramme**, le bouton **Fermer** permet de retourner au dialogue **couleurs**. L'histogramme modifié est tracé dans la légende de la carte. Notons que la pondération de l'histogramme est purement illustrative et ne modifie aucunement la carte, contrairement au déplacement des bornes des classes qui lui peut modifier profondément la carte.

Pour afficher l'histogramme sans pondération, il faut annuler cette pondération en affichant à nouveau le dialogue **Histogramme** par un clic sur , puis choisir **Pondérer l'histogramme**, puis dans le dialogue **Variable de pondération ?** activer **Sans pondération**, ce qui annule la pondération, et enfin, à nouveau dans le dialogue **Histogramme**, choisir **Fermer** pour retourner au dialogue **couleurs**.

3.6. Changer le nombre de classes

Lorsque l'une des méthodes de discrétisation suivantes **Jenks**, **=EF**, **=AM**, est sélectionnée, le dialogue **couleurs** affiche un curseur  **6 classes** permettant

de faire varier le nombre de classes de 2 à 10. En cliquant sur la flèche dirigée vers le haut jusqu'à  , on obtient une carte très contrastée (fig. 3.12).

Le choix du nombre de classes est plus délicat qu'il n'y paraît de prime abord : trop petit, il peut empêcher d'appréhender la complexité de la distribution statistique, trop grand, il peut introduire des ruptures sans signification. Quand si le nombre d'éléments cartographiques surfaciques est de l'ordre de quelques centaines, on choisit généralement entre 4 et 6 classes.

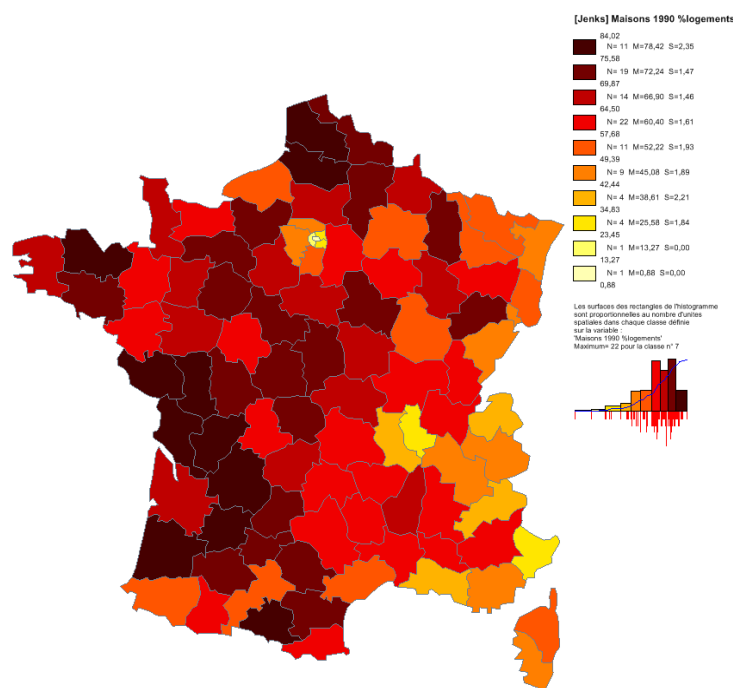


Figure 3.12 : Tracé de la carte choroplèthe avec discrétisation de Jenks en 10 classes.

3.7. Choisir une gamme de couleurs

Lors du lancement de *Philcarto*, la gamme de couleurs  est automatiquement appliquée. *Philcarto* ne donne pas à l'utilisateur le droit de modifier des couleurs dans une de ses gammes ; ceci garantit contre tout choix malvenu comme on en rencontre trop souvent. En revanche, *Philcarto* propose une dizaine de gammes de couleurs, dont deux en opposition et une en niveaux de gris. L'ordre des couleurs dans une gamme peut être inversé en cochant la case ☒ **Inverser couleurs** selon qu'on veut mettre l'accent sur les fortes valeurs (gamme normale) ou sur les faibles (gamme inversée). Par exemple, l'emploi de la gamme inversée  permet de souligner les départements où la proportion de maisons individuelles est faible au regard de l'ensemble des logements, c'est-à-dire dans les départements les plus urbains, notamment Paris et sa petite couronne, Lyon... (fig. 3.13).

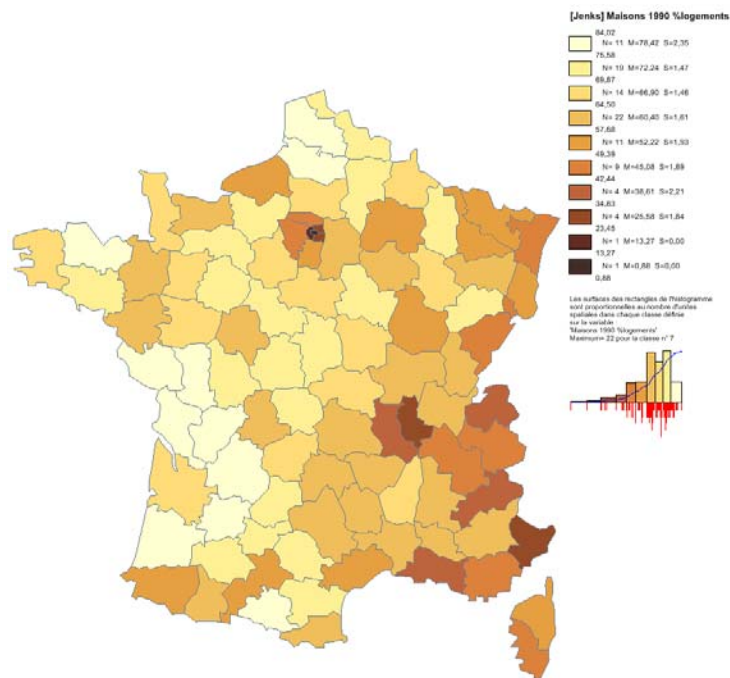


Figure 3.13 : Tracé de la carte choroplèthe avec discrétisation de Jenks en 10 classes et gamme de couleurs inversée.

3.8. Utiliser une gamme de couleurs en opposition.

Choisir l'une des deux gammes de couleurs en opposition  ou  suppose qu'il existe une valeur à partir de laquelle le phénomène mesuré change de nature. Par exemple, dans le cas d'une variation, la partie « chaude » de la gamme de couleurs (du jaune au rouge) exprime un accroissement positif, alors que la partie « froide » (du jaune au bleu ou au vert) traduit une décroissance. Dans ce cas, il existe une valeur « naturelle » de divergence qui est le 0 ou le 100 si les variations sont exprimées en indice de base 100.

Pour montrer l'intérêt de ces gammes de couleurs en opposition, changeons de variable de la manière suivante : cliquer sur **CARTO** qui affiche à nouveau les listes de variables. Dans la liste **couleurs** sélectionnons la variable **Indice Population 1999 Base 100=1990** qui renferme des valeurs supérieures à 100 si la population s'est accrue entre 1990 et 1999, et des valeurs inférieures à 100 si la population a diminué. Avec de telles valeurs, l'adoption d'une gamme de couleurs en opposition est légitime. Un clic sur l'icône  trace la carte choroplèthe avec la méthode de discrétisation et la gamme de couleurs continue actives à ce moment là ; décochons **Inverser couleurs** pour obtenir une gamme de couleurs continue croissante de la valeur minimum à la valeur maximum (fig.3.14).

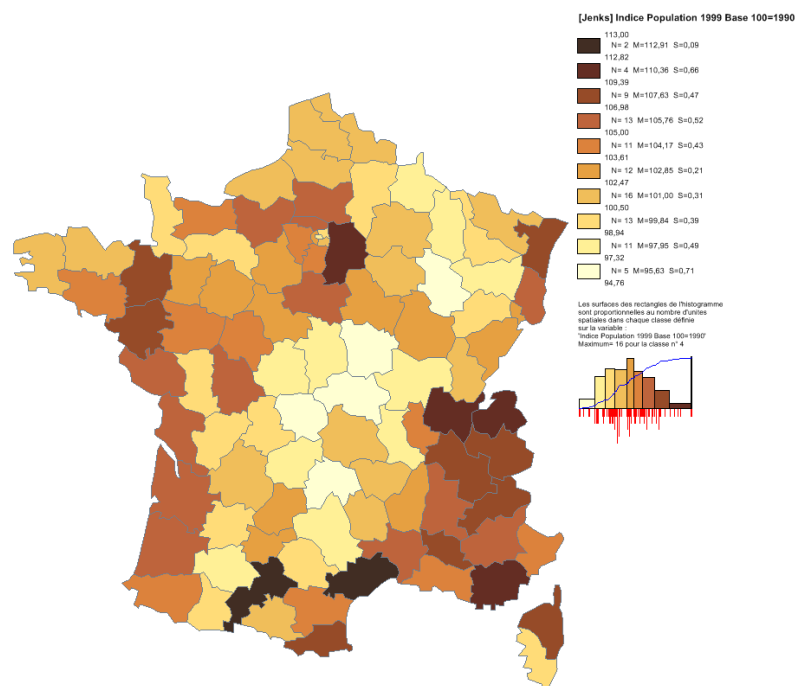


Figure 3.14 : Tracé de la carte choroplèthe avec discrétisation de Jenks en 10 classes et gamme de couleurs

Cette carte présente deux défauts : en légende, la valeur 100 n'apparaît pas (ici la méthode de Jenks a fait que la valeur 100 est comprise dans la classe dont la borne inférieure est 98,94 (qui exprime donc une diminution) et la borne supérieure est 100,5 (donc une augmentation). Le second défaut est l'adoption d'une gamme de couleurs croissante qui ne souligne pas particulièrement les départements ayant perdu des habitants. Il apparaît donc utile d'affecter la valeur 100 comme borne de divergence des couleurs en opposition. Cela se fait simplement à l'aide d'un dialogue permettant de saisir une valeur de divergence (fig. 3.15)

Figure 3.15 : Le dialogue de saisie d'une valeur de divergence pour les gammes de couleurs en opposition.

Par défaut, cette case est cochée et la valeur de divergence est fixée à 0. Ceci signifie que si *Philcarto* trouve des valeurs négatives dans la variable à cartographier, il adopte automatiquement la gamme et 0 comme valeur de divergence. Cette case peut être décochée : ; dans ce cas, la variable est considérée comme continue du minimum au maximum, et elle est traitée comme toute autre variable de couleur. Si la case est cochée, la valeur de divergence peut être modifiée en saisissant une autre valeur que 0, par exemple 100 si la variable est exprimée en indice de base 100 ; un clic sur provoque un nouveau tracé de la carte avec application du nouveau paramètre (fig. 3.16).

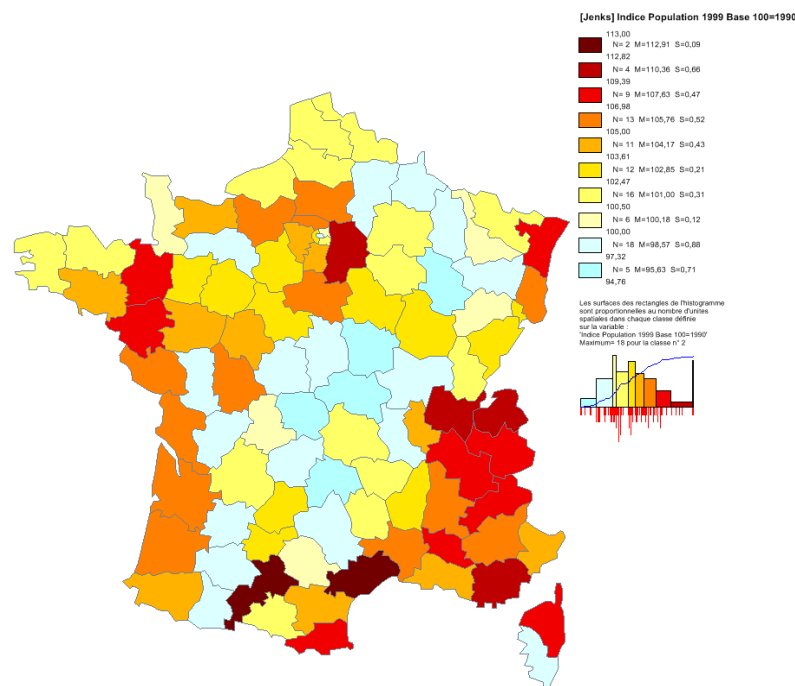


Figure 3.16 : L'utilisation de la gamme pour traduire des diminutions et des augmentations de population.

Cette carte permet de répondre sans hésitation à la question « quels sont les départements ayant perdu des habitants entre 1990 et 1999 ».

Naturellement, rien n'interdit de choisir la gamme au lieu de (fig. 3.17). Pour ne pas déroger à la tradition qui veut que les accroissements sont exprimés par des couleurs chaudes, l'inversion de la gamme de couleurs n'est pas autorisée dans ce cas là.

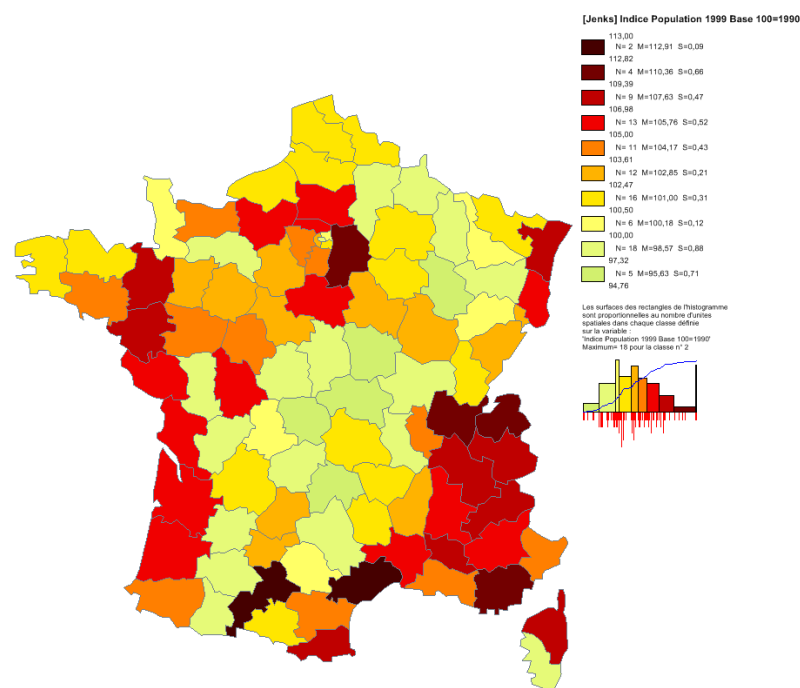


Figure 3.17 : L'utilisation de la gamme pour traduire des diminutions et des augmentations de population.

3.9. Ajouter un habillage

Un habillage est la superposition à la carte représentant des données statistiques d'éléments cartographiques linéaires (routes, fleuves) ou surfaciques (circonscriptions de niveau supérieur comme les régions). Les contours de ces éléments surfaciques sont tracés dans une couleur choisie par l'utilisateur, et il peut y en avoir plusieurs pour une même carte.

Pour commander le tracé d'un habillage linéaire ou surfacique, il faut que le fond de carte comprenne des calques de type **.L** ou **.S** (voir le manuel d'utilisation de *Phildigit*). Ces calques sont listés dans le dialogue **Habillage** situé dans la partie inférieure du cadre **Outils** (fig. 3.18).

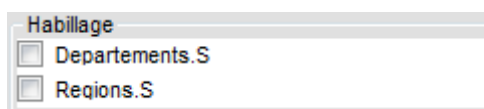


Figure 3.18 : La liste des calques d'habillage.

Dans cette liste, le calque **Departements.S** est celui qui porte les données statistiques. C'est donc le calque **Regions.S** qui peut être utilisé. Un clic dans la case correspondant à ce calque le sélectionne et fait apparaître une série de boutons colorés (fig. 3.19). Il faut impérativement choisir l'une de ces couleurs pour que l'habillage puisse être tracé dans cette couleur-là. On peut choisir une couleur passe-partout comme le noir ou le gris, ou une couleur plus expressive à la condition qu'elle soit en harmonie avec la gamme de couleurs retenue pour tracer la carte.

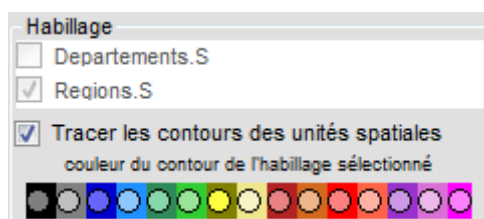


Figure 3.19 : La liste des couleurs d'habillage.

Lorsque le bouton de la couleur choisie est sélectionné (ici le noir), la série de boutons s'efface, le nom du calque est écrit avec un fond coloré par la couleur choisie (fig. 3.20), et la carte est à nouveau tracée, faisant apparaître l'habillage retenu (fig. 3.21). On observe que le trait d'habillage est plus épais que celui des contours des départements qui, de plus, est tracé en gris. La case ☒ **Tracer les contours des unités spatiales** peut être décochée pour laisser subsister seulement les contours d'habillage (fig. 3.22). Ceci est intéressant lorsque la carte choroplèthe se compose de nombreux éléments surfaciques qui parasitent la lecture des ensembles géographiques de même couleur. Noter que les contours d'habillage sont alors tracés en traits fins (fig. 3.23).

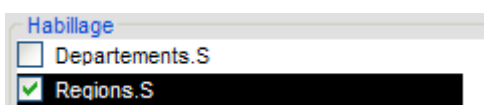


Figure 3.20 : Un calque sélectionné dans la liste des calques d'habillage.

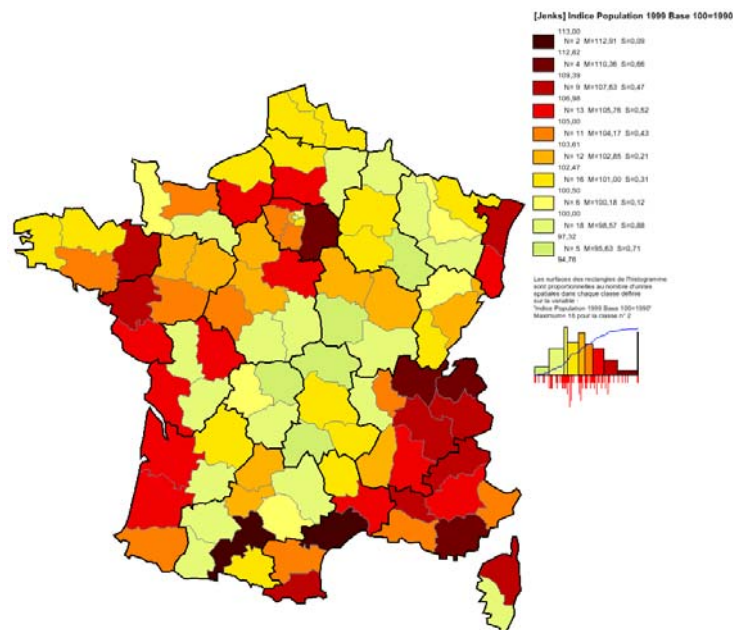


Figure 3.21 : Tracé de la carte avec l'habillage des régions.

Habillage

☐ Departements.S

☒ Regions.S

☐ Tracer les contours des unités spatiales

Figure 3.22 : Décocher la case ☒ Tracer les contours des unités spatiales .

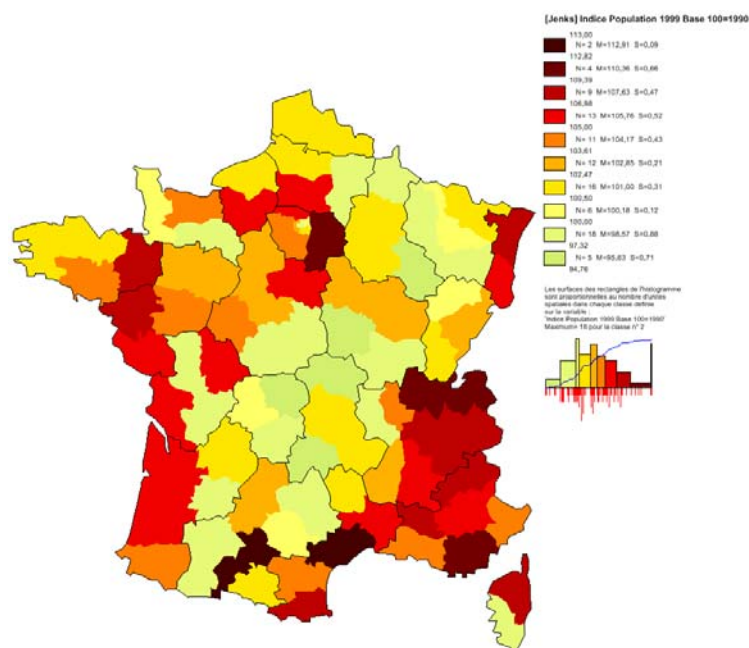


Figure 3.23 : Tracé de la carte avec l'habillage des régions, mais sans les contours des départements.

Dans le cas présent, on préfère laisser cochée la case ☒ Tracer les contours des unités spatiales .

3.10. Identifier les éléments cartographiques

Lorsque la flèche de la souris passe sur un élément cartographique, point ou surface, son identifiant, son nom et la valeur de la variable cartographiée sont affichés dans le bandeau supérieur du bureau de *Philcarto* (fig. 3.24).

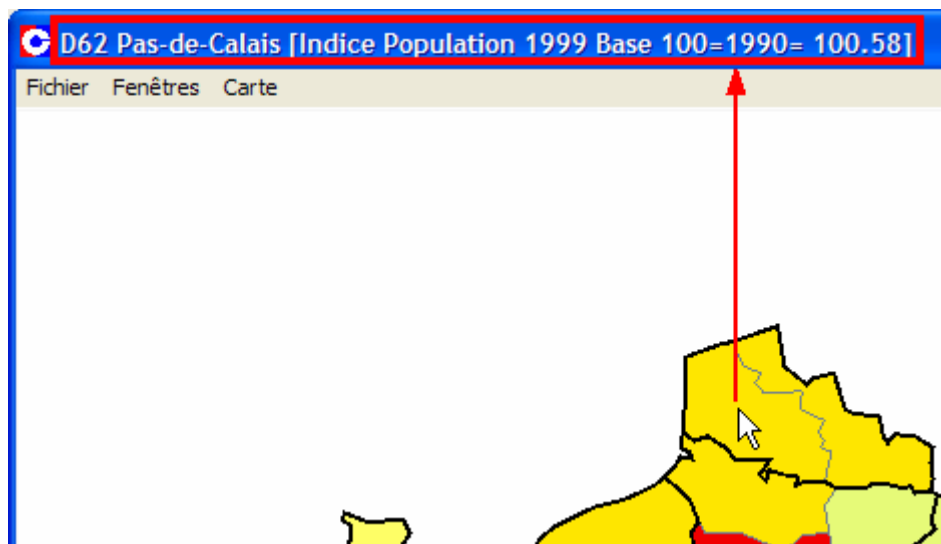


Figure 3.24 : L'identification d'un élément cartographique.

3.11. Enregistrer la carte dans un fichier

Il est possible d'enregistrer dans un fichier une carte affichée à l'écran pour l'utiliser ultérieurement (dans un article, par exemple). Trois formats d'enregistrement graphique sont proposés : le format **.emf** (*Enhanced Meta File*), le format **.ai** (*Adobe Illustrator postscript*) et le format **.svg** (*Scalable Vector Graphics*). Ces trois formats sont vectoriels, ce qui signifie qu'un agrandissement des dimensions de la carte ne se traduit pas par une pixelisation de l'image, les contours restant aussi nets avant et après agrandissement. Un autre intérêt de ces formats vectoriels est la possibilité qu'ils offrent à l'utilisateur de modifier tout élément graphique composant l'image, et ceci élément par élément : la couleur des contours ou de la surface, la position des titres, légende... peuvent ainsi faire l'objet d'un traitement particulier en vue d'une publication par exemple. Dans ce cas, les règles de l'art cartographique doivent aussi être respectées (voir à ce sujet l'ouvrage de Jacques Bertin, *Sémiologie Graphique* récemment réédité par l'Ecole des Hautes Etudes en Sciences Sociales, EHESS, Collection *Les réimpressions*, 444 p., ISBN : 2713212774).

Choisir entre EMF, AI et SVG n'est pas évident. On peut considérer que si la carte doit faire l'objet d'un traitement dans un logiciel de mise en page en vue de son impression dans un livre ou une revue, il est préférable de choisir le format **.ai** ; il est en effet compatible avec les logiciels produits par la société Adobe, notamment *Adobe Illustrator*. Le format **.ai** présente de nombreux avantages techniques dans le domaine de l'infographie. L'utilisateur qui dispose de *Adobe Illustrator* (même une version ancienne puisque le format d'enregistrement correspond à la version 6 du logiciel) aura tout intérêt à enregistrer ses cartes à ce format là. En effet, en plus de sa sophistication, ce format facilite la lecture de la carte car il conserve les valeurs correspondant à chaque élément cartographique dans ce qu'on appelle les remarques d'options d'objet (fig. 3.25). Ceci simplifie grandement le commentaire

des cartes car il n'est pas nécessaire de retourner au tableau de données original pour indiquer la valeur de tel ou tel autre élément.

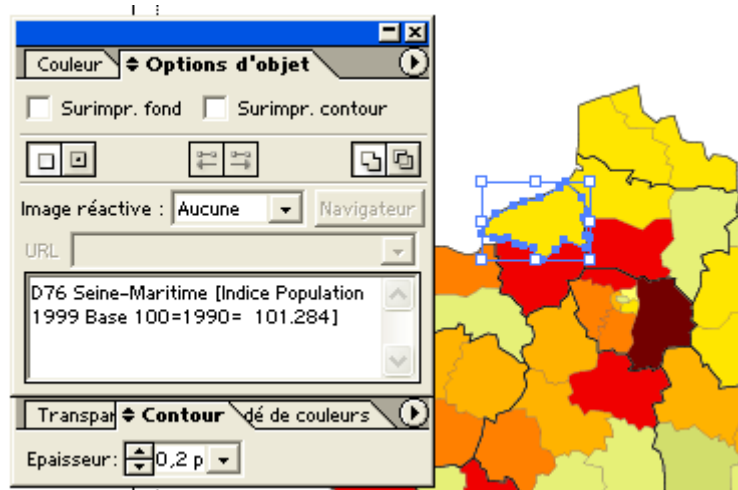


Figure 3.25 : Dans Adobe Illustrator, le département de la Seine-Maritime a été sélectionné. Son identifiant, son nom, et sa valeur sur la variable cartographiée sont lisibles dans les remarques de la fenêtre Options d'objet.

Les utilisateurs de *Adobe Illustrator* seront également intéressés par la possibilité qui leur est offerte de remplacer les couleurs par des trames (que seuls les logiciels Adobe savent restituer). Pour activer cette fonction, il faut cocher la case ☒ Trames dans fichier illustration. Un clic sur , visible quand cette case est cochée, affiche le dialogue **Trames** pour choisir un type de trame (fig. 3.26).

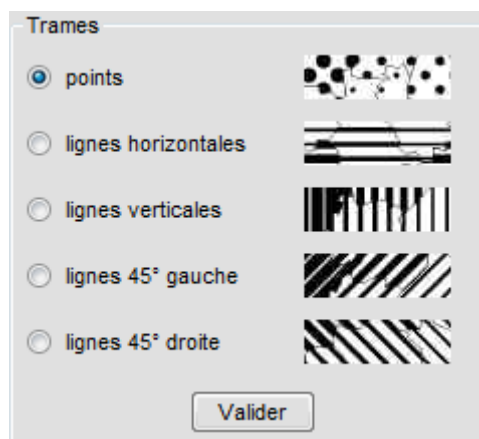


Figure 3.26 : Le dialogue **Trames**.

Par défaut, c'est la trame en points qui est validée quand on choisit . On peut aussi choisir des trames en hachures inclinées ou non. Ces trames, qui sont d'une gestion informatique particulière n'apparaissent pas à l'écran, mais elles remplacent les couleurs quand on enregistre la carte au format **.ai**. L'effet obtenu est plus ou moins réussi selon les configurations spatiales (fig. 3.27 et 3.28).

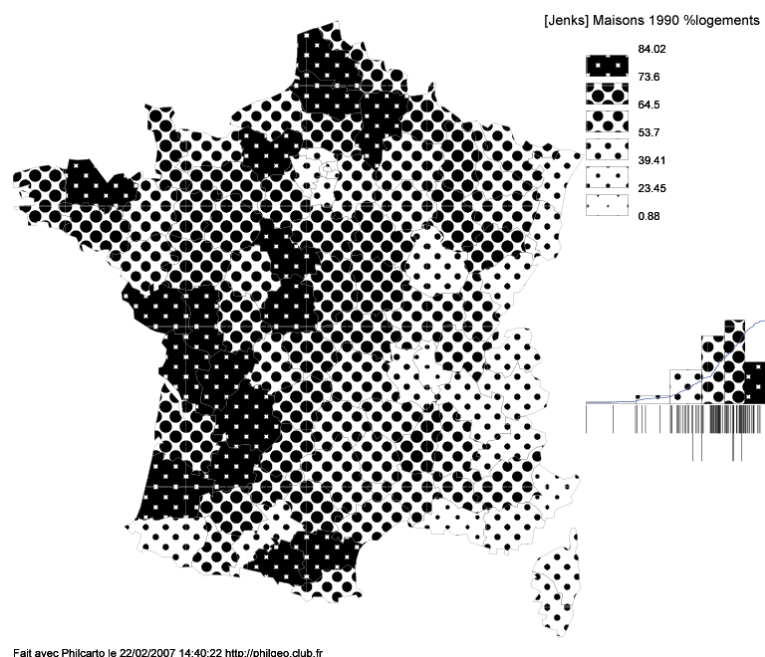


Figure 3.27 : Trames en points.

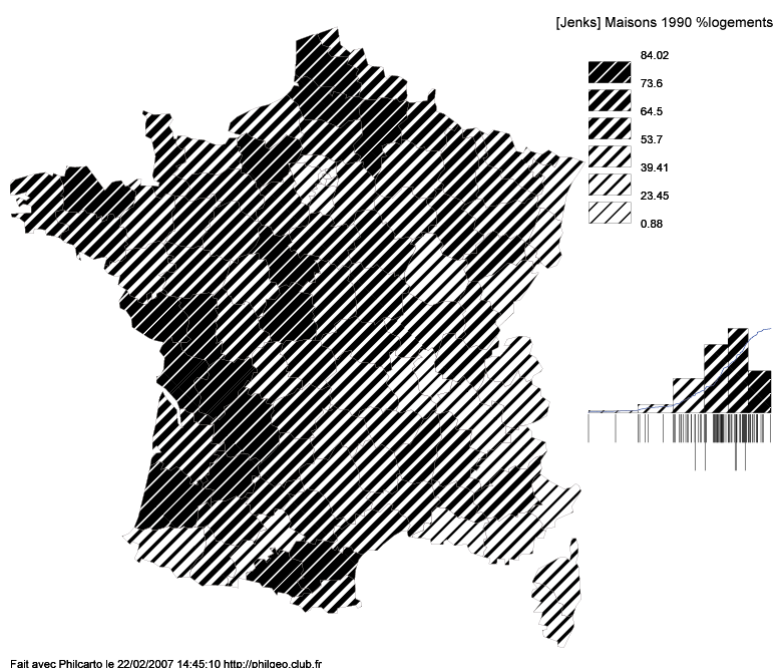
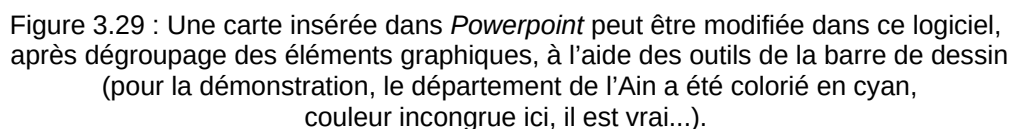


Figure 3.28 : Trames en lignes inclinées à 45°.

Notons qu'il existe un logiciel en *shareware* (vous l'essayez puis vous payez 39\$ US si vous l'adoptez) capable d'éditer des cartes au format **.ai**, sans la possibilité cependant de lire les options d'objet et qui ne reconnaît pas les trames : *Mayura Draw* téléchargeable à l'adresse www.mayura.com.

Si vous ne disposez pas d'un logiciel de dessin capable d'importation au format **.ai**, le format **.emf** peut être choisi ; il est compatible avec la plupart des logiciels plus ou moins onéreux du monde *Windows*. Par exemple, une carte enregistrée au format **.emf** peut être ouverte dans *Powerpoint* à l'aide de l'article **Insertion image à partir du fichier...** et sélectionner le fichier, puis être modifiée dans *Powerpoint* (fig. 3.29) à

Le même genre de chose peut être fait avec plus ou moins de facilité avec *Draw* ou *Impress* de la suite *Open Office*, et naturellement avec les autres logiciel de dessin vectoriel du monde Windows tels *Correl Draw*, *Microsoft Expression*...



87

C'est aussi de manière native que le format **.svg** est accepté par le logiciel de dessin *Inkscape*. Ce logiciel, téléchargeable gratuitement sur le site : <http://www.inkscape.org/>, est un genre d'*Adobe Illustrator* gratuit ! Il est en cours de développement, mais déjà vraiment utilisable. Tous ceux qui n'ont pas les moyens de s'offrir AI peuvent opter pour *Inkscape* afin de réaliser l'habillage final de leurs cartes.

Comme AI, *Inkscape* facilite la lecture de la carte car le format **.svg** conserve les valeurs correspondant à chaque élément cartographique dans ce qu'on appelle les propriétés de l'objet sélectionné (fig. 3.30). Ceci simplifie grandement le commentaire des cartes car il n'est pas nécessaire de retourner au tableau de données original pour indiquer la valeur de tel ou tel autre élément.

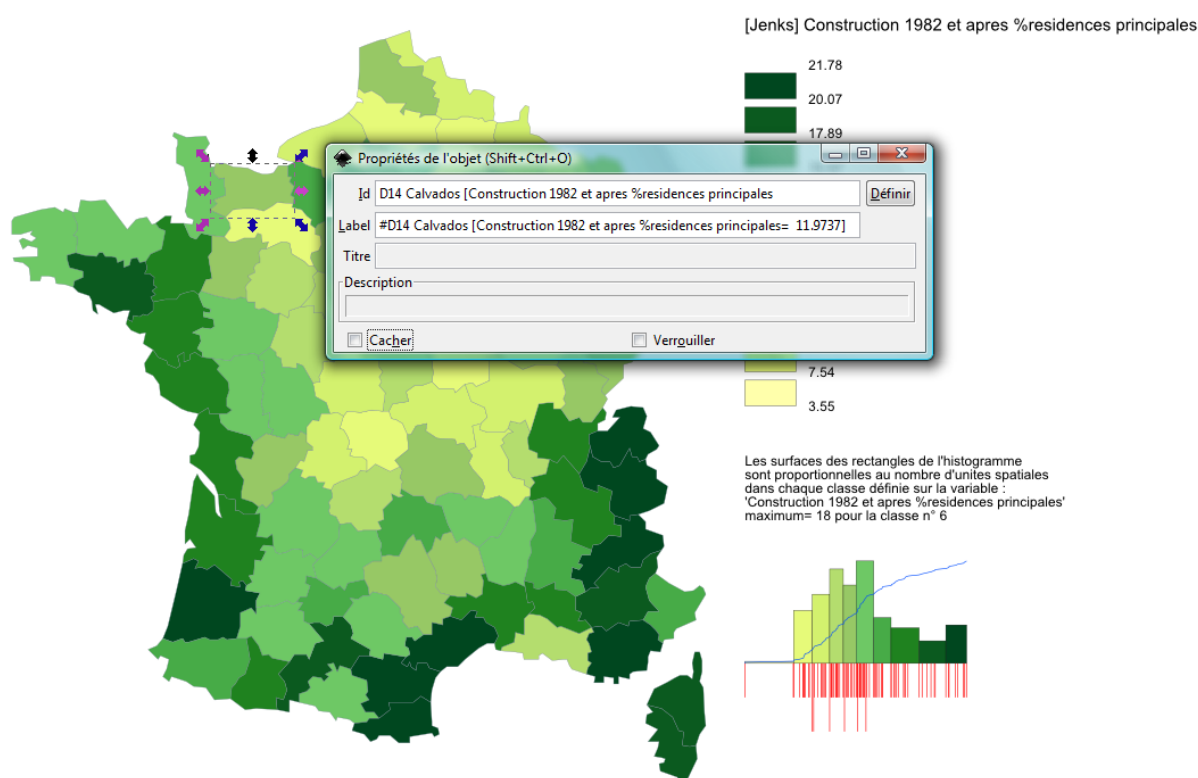


Figure 3.30 : Dans Inkscape, le département du Calvados a été sélectionné. Son identifiant, son nom, et sa valeur sur la variable cartographiée sont lisibles dans les remarques de la fenêtre Propriétés de l'objet.

Pour enregistrer une carte au format **.emf**, **.ai** ou **.svg**, il suffit d'activer l'article correspondant dans le menu **Carte** (fig. 3.31). Le dialogue **Nom du fichier ?** permet dans les trois cas de donner le nom du fichier, **Carte** par défaut, et de préciser le répertoire dans lequel il doit être enregistré (fig. 3.32).

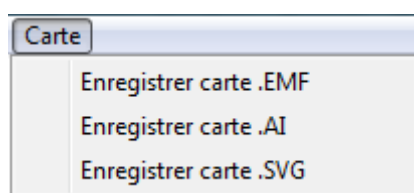


Figure 3.31 : Le menu **Carte** permet d'enregistrer la carte affichée à l'écran soit au format **.emf** (metafile), soit au format **.ai** (Adobe Illustrator), soit au format **.svg** (Scalable Vector Graphics).

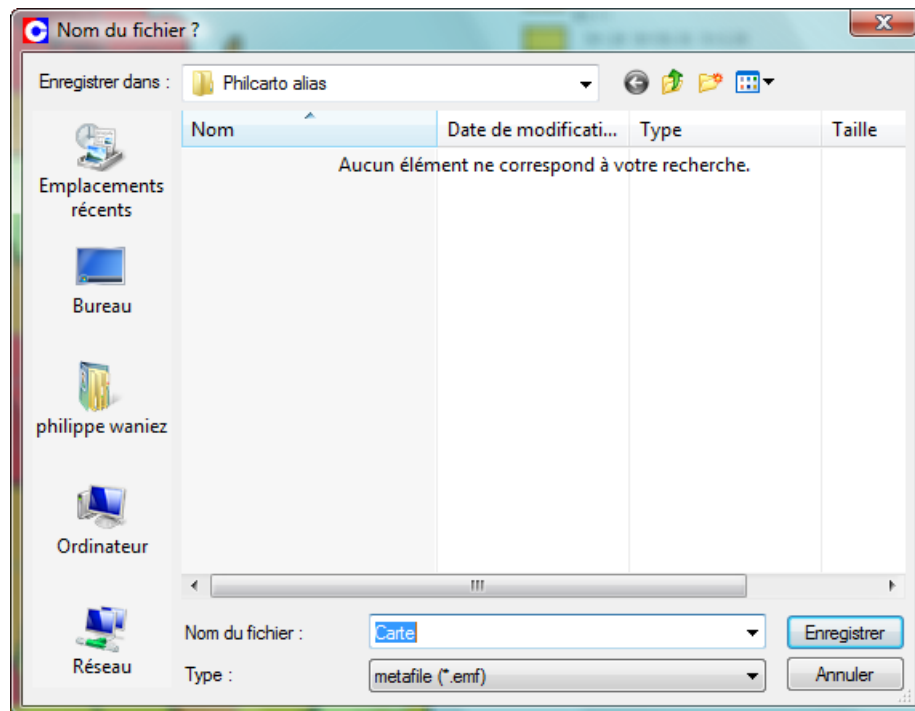


Figure 3.32 : Le dialogue  **Nom du fichier ?** pour l'enregistrement au format **.emf** .

3.12. Evaluer l'autocorrélation spatiale

Les méthodes statistiques supposent que les valeurs sont indépendantes les unes des autres. Or cela n'est pas toujours le cas, ni avec les séries chronologiques (la valeur mesurée au temps $t+1$ n'est pas indépendante de la valeur mesurée au temps t), ni avec les données spatiales (la valeur mesurée au lieu x n'est pas indépendante des valeurs mesurées sur les lieux voisins, y_1 , y_2 ...). Lorsque les données spatialisées ne sont pas indépendantes, on a affaire à une autocorrélation spatiale.

L'autocorrélation spatiale peut être considérée de plusieurs points de vue :

- un facteur qui nuit à la validité des résultats des analyses statistiques ;
- un genre de débordement des informations relevées en un lieu sur les lieux voisins ;
- le signe de l'existence de variables non prises en compte...

Deux mesures de l'autocorrélation spatiale sont souvent proposées dans la littérature statistique : les coefficients d'autocorrélation spatiale de Moran et Geary.

 **Philcarto** permet de calculer ces deux coefficients pour un fond de carte surfacique seulement, à la condition que le fond de carte ne comprenne pas de cartouche mettant en exergue une « sous-régions » (par exemple un zoom sur l'Île-de-France sur la carte des départements français).

Pour montrer comment calculer ces coefficients, réalisons une nouvelle carte par un clic sur  qui affiche à nouveau les listes de variables. Dans la liste  , choisissons la variable  ; un clic sur l'icône  trace la

carte choroplèthe, et la sélection de la gamme de couleurs  produit une carte correcte, en 10 classes avec discrétisation  (fig. 3.33).

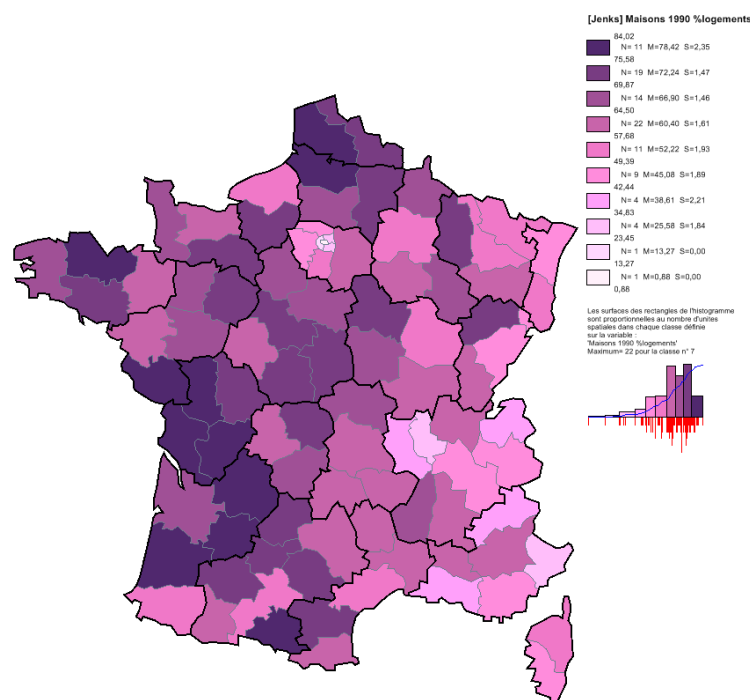


Figure 3.33 : Tracé de la carte choroplèthe avec discrétisation  en 10 classes, gamme de couleurs  et habillage régional.

Le calcul des coefficients d'autocorrélation spatiale de Moran et de Geary est effectué en cochant la case  Coefficients de Moran et Geary du dialogue **Autocorrélation spatiale**. *Philcarto* procède alors au repérage des voisins de chaque élément cartographique, sur une profondeur de voisinage variant de 1 à 10. Ce traitement prend d'autant plus de temps que le nombre d'éléments cartographiques est grand ; cependant, il n'est réalisé qu'une seule fois dans la même session de *Philcarto*. Notons aussi que l'utilisateur n'a pas à définir lui-même une matrice de contiguïté puisque *Philcarto* le fait lui-même. Ceci impose cependant que le fond de carte soit correctement digitalisé, notamment les nœuds de contours.

Sont considérés comme voisin de premier ordre (ou voisin immédiat), les polygones qui ont au moins un point en commun. Une fois les voisins d'ordre 1 à 10 repérés, *Philcarto* calcule les coefficient d'autocorrélation spatiale et les affiche dans la légende de la carte. De manière à rendre les deux coefficients comparables, c'est la valeur 1-Geary qui est indiquée (fig. 3.34). On observe dans ce tableau qu'à l'ordre de voisinage 1, l'autocorrélation est positive mais très modérée ; elle diminue très rapidement ensuite, pour ne jamais prendre une valeur significative ensuite, même si les valeurs deviennent négatives. Ces valeurs confirment néanmoins la réalité de quelques groupements significatifs tant pour les pourcentages élevés (Ouest, Nord) que pour les valeurs faibles (Région Île de France et Sud-Est).

Coefficients d'autocorrélation spatiale		
Ordre	Moran	1-Geary
1	0,47	0,60
2	0,16	0,38
3	0,00	0,13
4	-0,04	0,01
5	-0,05	-0,05
6	-0,09	-0,17
7	-0,11	-0,15
8	-0,06	-0,19
9	-0,09	-0,39
10	-0,16	-0,47

Figure 3.34 : Les coefficients d'autocorrélation spatiale d'ordre 1 à 10 pour la variable **Maisons 1990 %logements**

3.13. Comparer des cartes choroplèthes « à vue »

Loin d'être dépassée, la comparaison de plusieurs cartes permet souvent de faire des observations intéressantes sur la similarité des regroupements spatiaux, et de suggérer des corrélations, parfois sous-tendues par des relations de cause à effet, qui demandent à être confirmées ensuite par des méthodes plus précises. Au moins cette approche permet-elle d'éviter d'appliquer certaines techniques statistiques de manière aveugle comme c'est parfois encore le cas... Pour rendre possible une telle approche, qu'on peut qualifier d'exploratoire, *Philcarto* permet d'afficher simultanément jusqu'à quatre cartes et d'enregistrer la planche ainsi constituée dans un fichier EMF. Naturellement, plus l'écran de l'ordinateur est grand, plus cette visualisation est efficace.

Le dialogue **Disposition des cartes** situé dans la partie inférieure du cadre **Outils**, contient deux icônes représentant un écran d'ordinateur (fig. 3.35). L'écran de droite possède un seul bouton radio actif par défaut : cela signifie qu'une seule carte est tracée à la fois par *Philcarto*.



Figure 3.35 : Le dialogue **Disposition des cartes**.

L'écran situé à gauche comprend quatre boutons radio permettant de désigner l'un des quadrants de l'écran : l'espace d'affichage de la carte est ainsi divisé en quatre rectangles égaux où vient s'afficher toute nouvelle carte. Lorsqu'une carte est affichée en grand format, un clic sur l'un de ces quatre boutons radio transfère le tracé vers le quadrant ainsi désigné, les dimensions de la carte étant alors divisées

par deux. Un clic sur le bouton radio du quadrant supérieur gauche  transfère la carte en 10 classes avec discrétisation de **Jenks** dans ce quadrant là (fig. 3.36).

Cette carte étant affichée, désignons maintenant le quadrant supérieur droit , et sélectionnons le mode de discrétisation **=Ef** . La carte est alors redessinée avec 10 classes d'effectifs égaux et se place à droite de la précédente rendant ainsi possible l'appréciation des différences entre ces deux discrétisations (fig. 3.37).

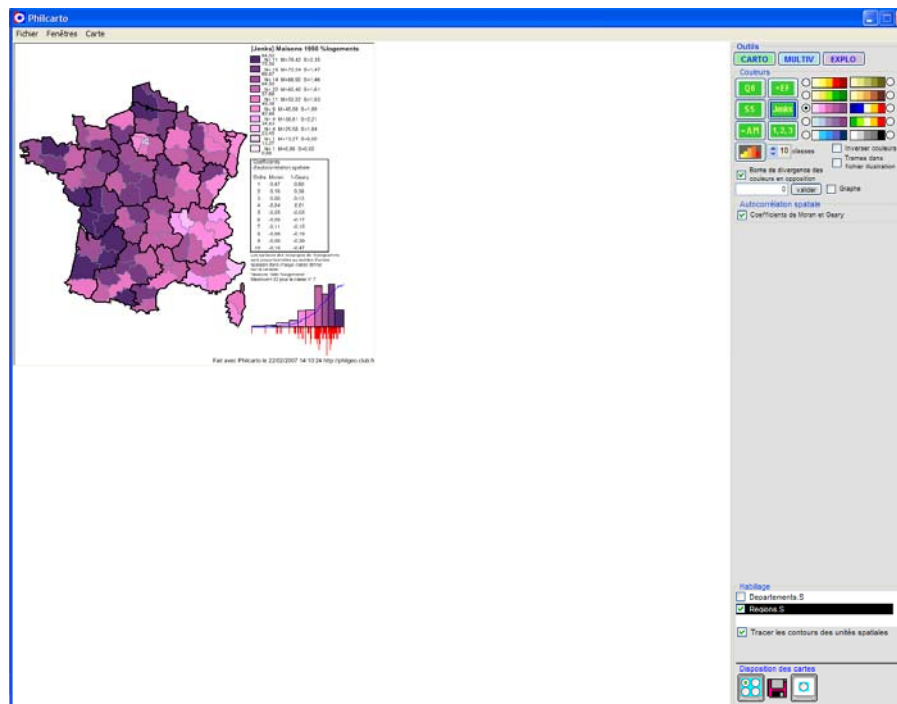


Figure 3.36 : Le transfert de la carte dans le quadrant supérieur gauche (discrétisation de Jenks en 10 classes).

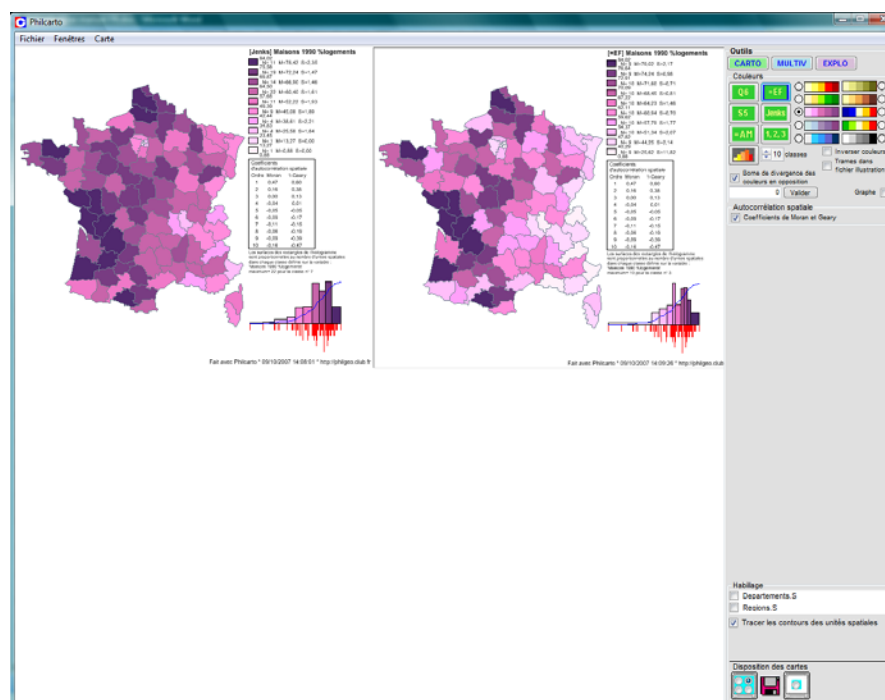


Figure 3.37 : Le tracé de la carte dans le quadrant supérieur droit (à gauche : discrétisation de Jenks en 10 classes ; à droite : discrétisation =EF en 10 classes).

On observe que cette seconde carte est pourvue de l'esquisse d'un cadre à gauche et au-dessus ; ceci signifie que cette carte est active. Autrement dit, elle est encore modifiable à l'aide des outils du dialogue **couleurs** (fig. 3.38). Par exemple, on peut changer la méthode de discrétisation, la gamme de couleurs, le nombre de classes... par exemple 5 classes, et gamme de couleur  (fig. 3.39).

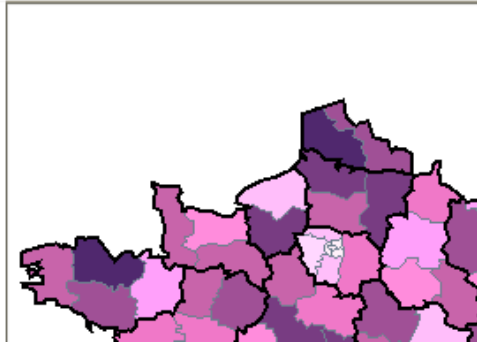


Figure 3.38 : Le cadre situé à gauche et au-dessus de la carte indique qu'elle est active et peut être modifiée.

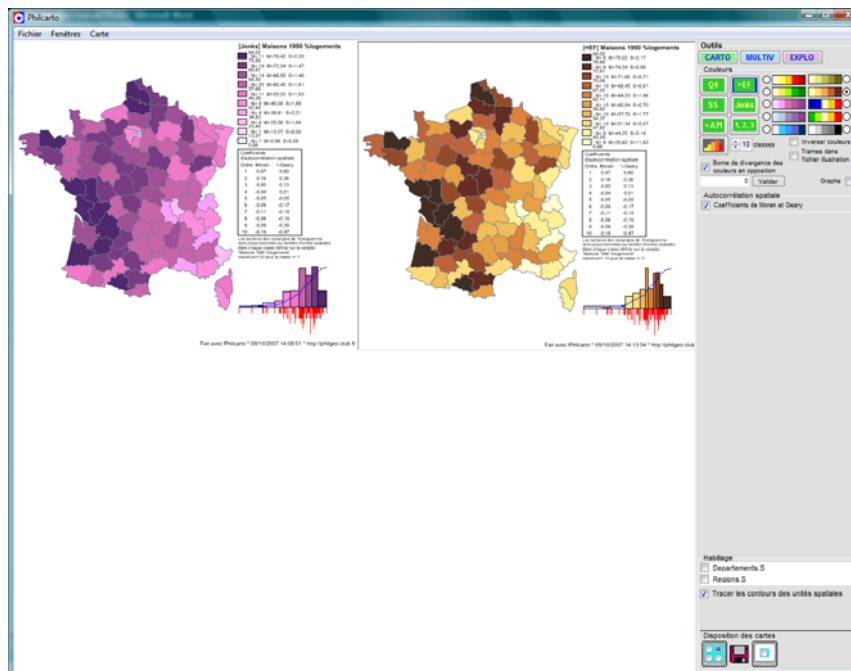


Figure 3.39 : Changement de gamme de couleurs sur la carte active (celle de droite).

On peut réitérer ce genre d'expérience avec d'autres discrétisations, , , et examiner ainsi les différences entre les cartes introduites par chacune de ces méthodes (fig. 3.40) pour apprécier la robustesse des structures spatiales mises en évidence. Un clic sur le bouton radio de l'écran doté d'un seul bouton  affiche en plein écran la carte active (fig. 3.41). L'enregistrement de la planche est possible par un clic sur .

 Pour éviter tout déboire dans la conception d'une planche de cartes, l'emploi des boutons radio  doit suivre la séquence suivante : sélectionner d'abord l'emplacement de la carte avec le bouton radio correspondant, puis tracer la carte.

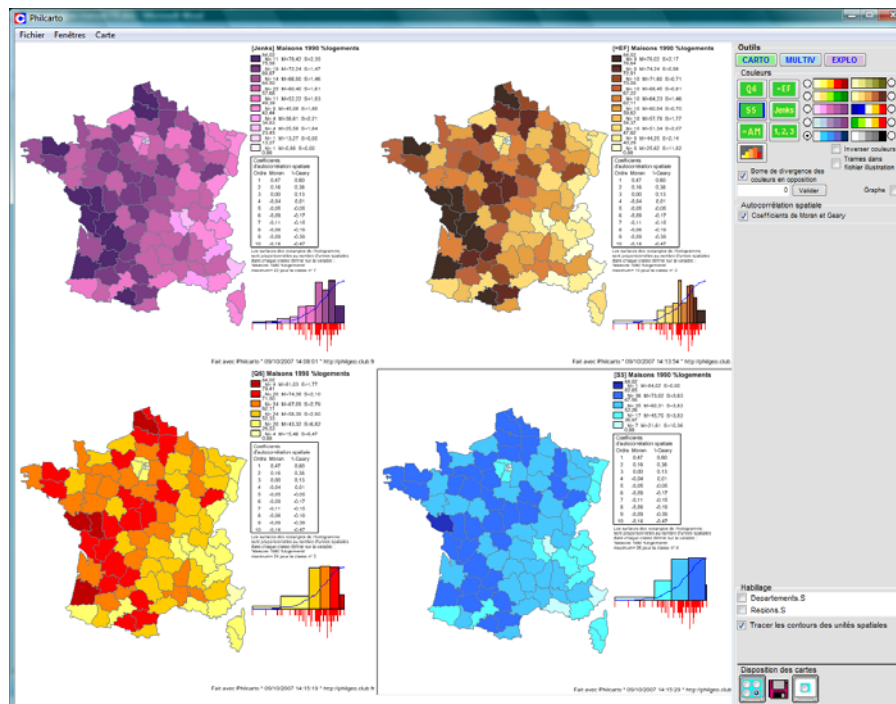


Figure 3.40 : La même variable avec 4 discrétisations différentes : en haut à gauche Jenks 10 classes, en haut à droite =EF 5 classes, en bas à gauche Q6 et en bas à droite S5.

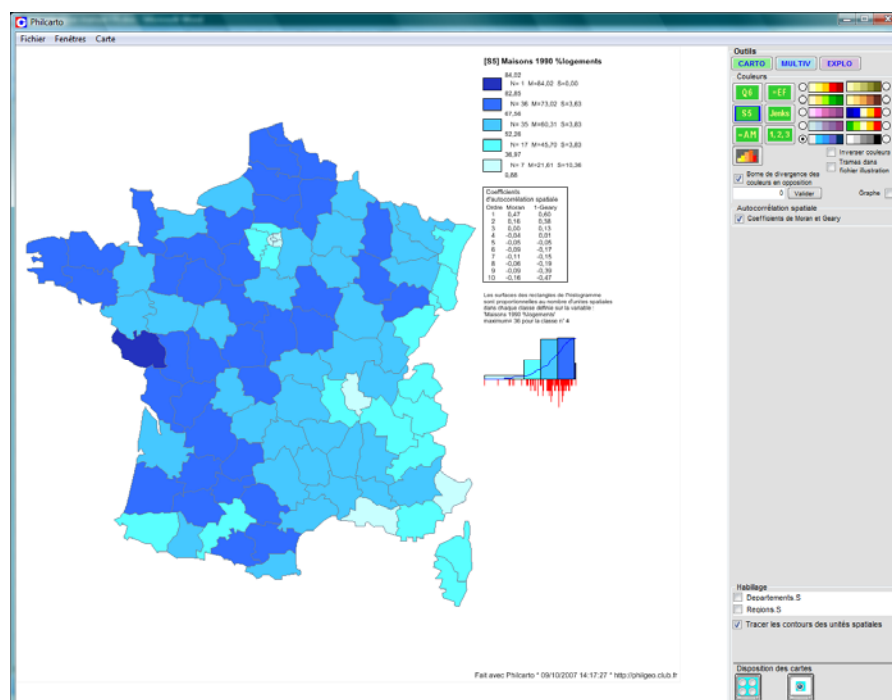


Figure 3.41 : Affichage en plein écran de la carte active avec discrétisation S5.

3.14. La coadjacence des cartes choroplèthes

Lorsque des ressemblances ont été identifiées par comparaison à vue de différentes cartes, il est intéressant d'en préciser leur portée par une mesure de proximité. La statistique propose de nombreuses méthodes de calcul de corrélation, d'association ou de régression... Ces techniques, utilisées avec succès par certains géographes, sont cependant mal adaptées aux problèmes géographiques car elles présentent le défaut de ne pas prendre en considération la position relative des éléments cartographiques, ni l'autocorrélation spatiale, alors même que leur fondement mathématique repose sur l'indépendance supposée des valeurs des variables.

Dans un article publié en 1977, Jean-Claude Muller a proposé une méthode fort intéressante de comparaison de cartes qui repose sur la notion de coadjacence (L'espace Géographique, Tome VI, n°1, janvier-mars 1977, pp.59-72). La coadjacence choropléthique est mesurée de la manière suivante : chaque carte est transformée en graphe en associant un sommet (centroïde ou chef-lieu) à chaque élément cartographique surfacique. Deux sommets sont reliés par un arc si les éléments cartographiques surfaciques correspondants sont contigus et appartiennent à la même classe ; dans le cas contraire, les deux sommets ne sont pas reliés. On obtient ainsi un graphe des éléments cartographiques adjacents, c'est-à-dire voisins et appartenant à la même classe (fig. 3.42).

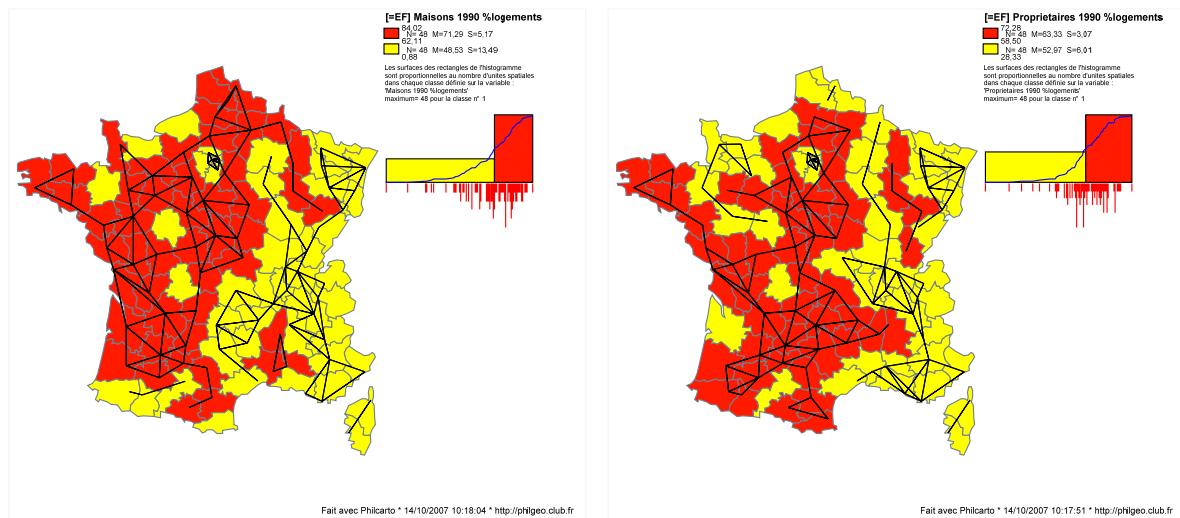


Figure 3.42 : Les cartes choroplèthes en deux classes du % de maisons et du % de propriétaires dans l'ensemble des logements. Le graphe des coadjacences est superposé

Pour obtenir le graphe des unités spatiales adjacentes, il suffit de cocher la case **Graphe** ☒ du dialogue **couleurs**. Si cela n'a pas déjà été fait (par exemple pour un calcul précédent de coefficients d'autocorrélation), *Philcarto* procède au repérage des voisins de chaque élément cartographique, puis le graphe est tracé.

La comparaison de deux cartes choroplèthes consiste à compter le nombre d'arcs communs, distincts et total pour les deux cartes à comparer. Si le pourcentage d'arcs communs approche 100% du nombre total d'arcs, alors les deux graphes sont isomorphiques et les deux cartes choroplèthes sont identiques ; si le pourcentage est proche de 0, les deux cartes choroplèthes sont différentes ; à partir de 50% d'arcs

communs, les valeurs représentées sont liées par un certain type de relation fonctionnelle d'autant plus forte que le pourcentage tend vers 100%.

L'article [Coadjacences choropléthiques](#) du menu [Fenêtres](#) permet d'afficher ces pourcentages dans le dialogue [Coadjacences choropléthiques](#) (fig. 3.43).

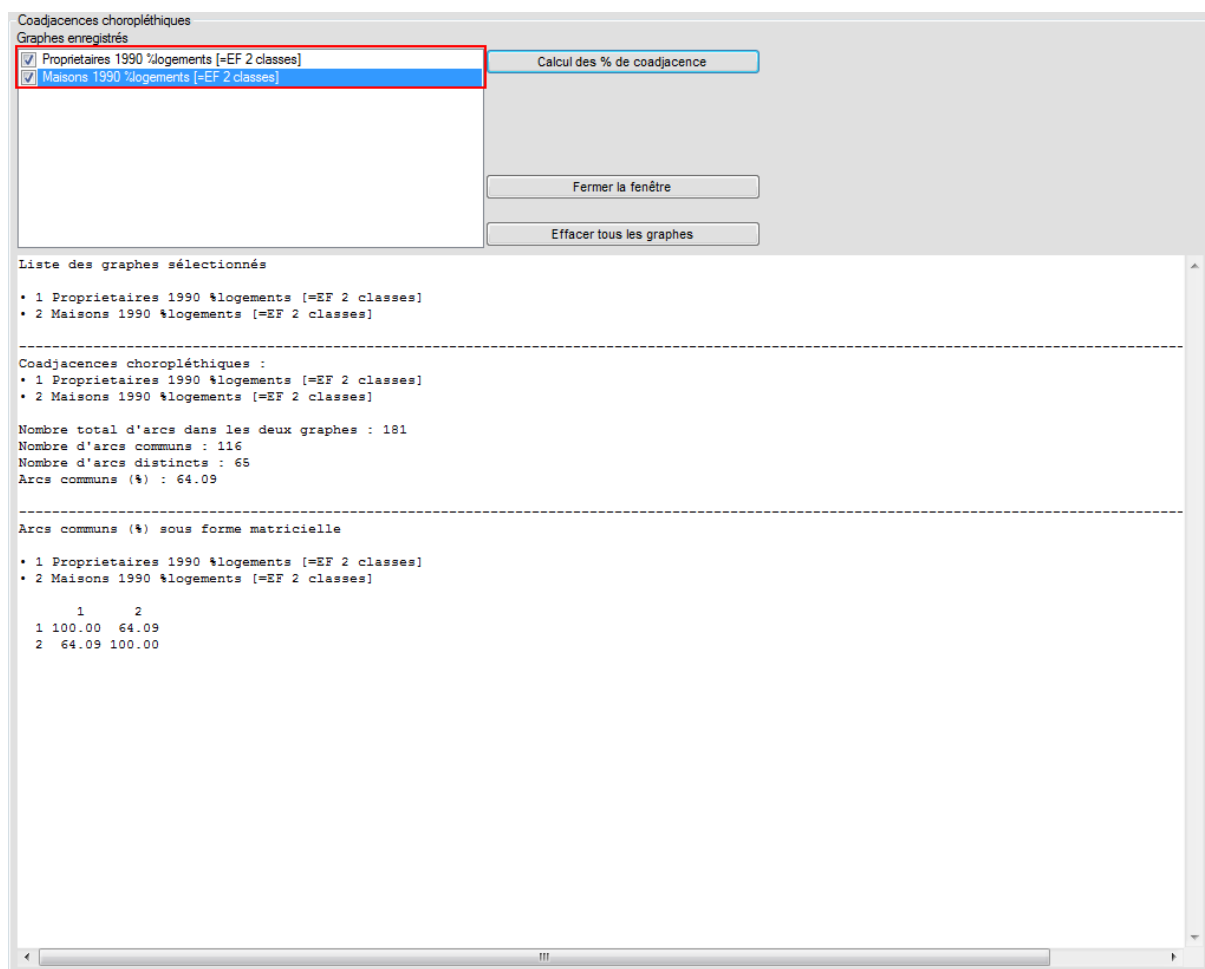


Figure 3.43 : Le dialogue [Coadjacences choropléthiques](#).

Les différents graphes précédemment tracés sont listés dans le cadre Graphes enregistrés ; naturellement, on peut y trouver plus de deux graphes en fonction du nombre de cartes à étudier. Ces graphes portent le nom de la variable cartographiée auquel a été ajouté le mode de discrétisation, par exemple : [Propriétaires 1990 %logements \[=EF 2 classes\]](#) qui correspond à une carte choroplèthe en deux classes d'effectifs égaux. Les graphes dont on souhaite étudier la coadjacence doivent être cochés : ☒ [Maisons 1990 %logements \[=EF 2 classes\]](#). Un clic sur [Calcul des % de coadjacence](#) provoque le calcul puis l'affichage des coadjacences, d'abord sous forme de liste, puis sous forme d'une matrice. Ici, le nombre total d'arcs pour les deux graphes s'élève à 181, dont 116 arcs communs, soit 64% du nombre total d'arcs.

La méthode proposée par Jean-Claude Muller est sensible au nombre de classes : il est préférable que celui-ci soit petit et identique pour tous les graphes à comparer. Elle ne donne pas une mesure absolue de la ressemblance des cartes choroplèthes, mais plutôt une mesure relative à un groupe de cartes.

3.15. Les cartes choroplèthes de variables nominales

Une des nouveautés importantes de la version 5 de *Philcarto* réside dans sa capacité à traiter les données nominales. Comme ce qualificatif l'indique, les valeurs nominales sont tout simplement des chaînes de caractères (au lieu d'être des nombres représentant des effectifs ou des rapports). Naturellement, un nombre peut aussi être considéré comme une valeur nominale s'il n'y a pas de sens à lui appliquer des calculs arithmétiques. Le type de variable nominal est automatiquement détecté si les valeurs sont des textes ; il doit être explicitement défini dans le fichier de données statistiques s'il comporte des valeurs numériques (voir à ce sujet le chapitre 19).

La sélection d'une variable nominale se fait simplement par un clic sur le nom d'une des variables figurant dans la liste **couleurs** du dialogue **Variables**. Ici la variable sélectionnée (fig. 3.44), **Region**, donne pour chaque département son nom en clair (pas sous la forme d'un code). Ainsi, les valeurs nominales sont, par exemple, Alsace ou Languedoc-Roussillon...

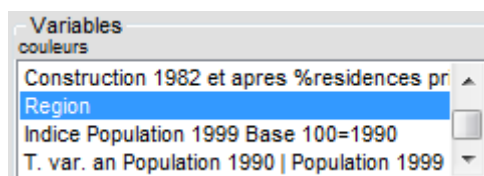


Figure 3.44 : Sélection d'une variable de couleur.

La sélection de la variable de couleur provoque l'affichage du dialogue **Types de cartes**. Cliquer sur l'icône du type carte choroplèthe  (fig. 3.45). La carte s'affiche alors à l'écran (fig. 3.46).

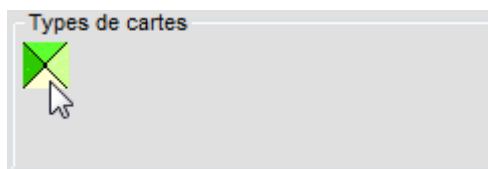


Figure 3.45 : Clic sur l'icône « carte choroplèthe ».

 **Au contraire des variables de rapports, les variables nominales n'ont pas à être discrétisées.**

Pour tracer cette carte, *Philcarto* a d'abord trié les modalités selon l'ordre alphabétique (depuis Alsace, jusqu'à Rhône-Alpes), puis a affecté une couleur à chacune d'elles selon l'ordre de la gamme dite « perroquet » dans laquelle les couleurs ne sont pas ordonnées selon une logique déterminée (la première modalité est représentée en rouge, la seconde en bleu, etc.).

Si les couleurs choisies par *Philcarto* dans la gamme « perroquet » ne sont pas satisfaisantes, il est possible d'en choisir d'autres dans cette gamme qui renferme 30 couleurs différentes. Il suffit pour cela de se rendre dans le dialogue **Couleurs** qui s'est affiché dans le cadre **Outils** juste après le tracé de la carte. On retrouve dans le dialogue **Couleurs** la liste des modalités de variable cartographiée (fig. 3.47).

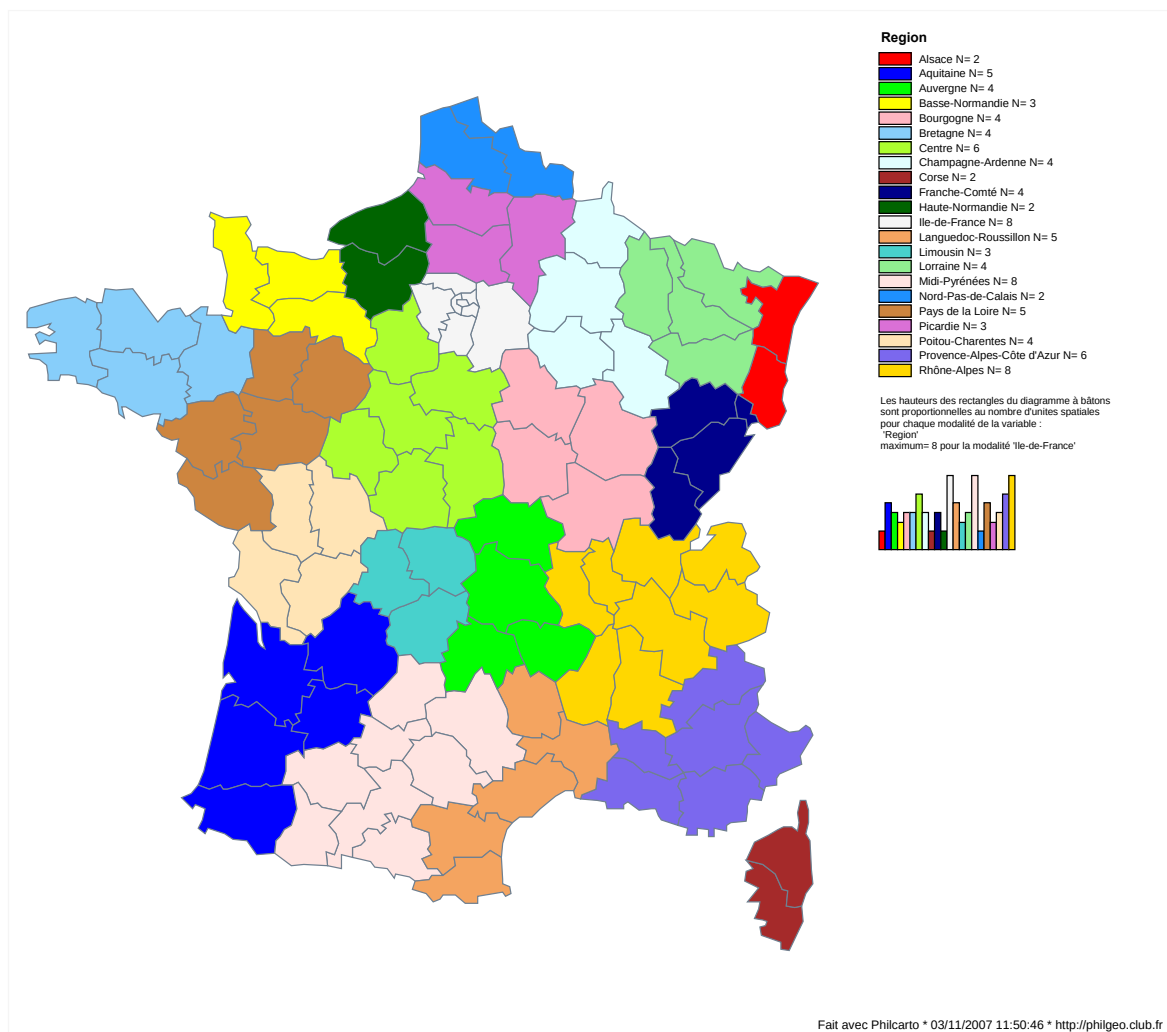


Figure 3.46 : Tracé de la carte choroplèthe du nom de la région de chaque département.



Figure 3.47 : Le cadre Couleurs dans le cas d'une variable nominale.

En regard de chaque modalité figure un bouton radio coloré. Un clic sur l'un de ces boutons radio permet de sélectionner une autre couleur dans la gamme proposée (fig. 3.48). Par exemple, on peut souhaiter modifier le vert de la région Auvergne car elle est adjacente à la région Centre, également figurée en vert, et à la région Limousin qui, elle, est figurée en bleu clair. En cliquant sur la modalité « Auvergne », puis en choisissant la couleur « Tomate », on modifie la couleur de l'Auvergne sur la carte, dans la légende et dans le diagramme à bâtons de la distribution des départements par région (fig. 3.49).



Figure 3.48 : Choisir une couleur différente pour une modalité.

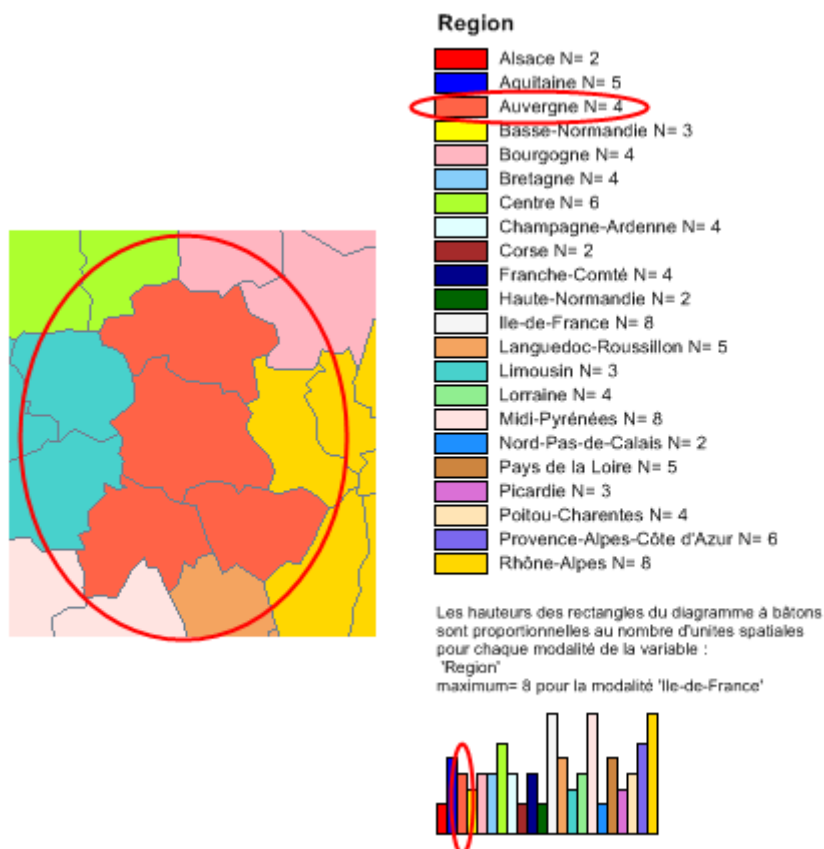


Figure 3.49 : Le changement de couleur se répercute sur la carte, dans la légende, et dans le diagramme à bâtons.

Notons qu'il est possible de pondérer le diagramme à bâtons comme cela est le cas pour les histogrammes des variables de rapport par un clic sur **Pondérer le diagramme à bâtons**. Le dialogue **Variable de pondération ?** permet alors de sélectionner la variable de pondération des bâtons du diagramme, ici **Logements 1990** (fig. 3.50). La validation, par un clic sur **Valider la pondération**, provoque un nouvel affichage de la carte et du diagramme à bâtons. Sans pondération, ce diagramme donne le nombre de départements par région ; en pondérant par le nombre de logements, il indique la répartition des logements par région. La pondération du diagramme à bâtons n'affecte en rien la carte.

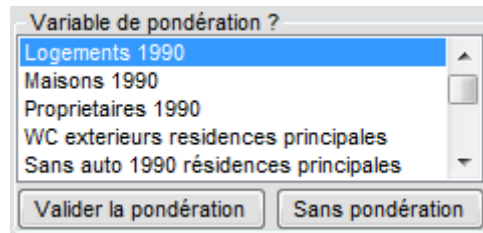


Figure 3.50 : La liste des variables de pondération du diagramme à bâtons.



Figure 3.51 : La pondération du diagramme à bâtons donne une idée juste de la répartition régionale des résidences principales : prééminence de l'Île de France (au centre du diagramme) suivie de Rhône-Alpes (en orangé à droite), puis de Provence-Alpes-Côte d'Azur (en violet à droite) et du Nord-Pas de Calais (en bleu). Les couleurs sont celle de la légende en figure 3.49.

4. Les cartes en cercles proportionnels

Dossier : \Philcarto exemples\France départements\
Fond de carte : France carte départements.ai
Statistiques : France data départements.txt

4.1. Sélectionner une variable de symbole

La sélection d'une variable se fait simplement par un clic sur le nom d'une des variables figurant dans la liste **symboles** du dialogue **Variables**. Ici la variable sélectionnée, **Maisons 1990**, contient le nombre de maisons individuelles dans chaque département de France métropolitaine au recensement de 1990 (fig. 4.1).

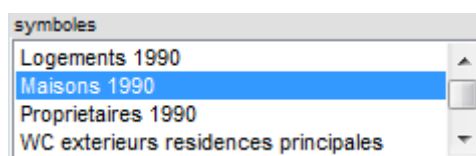


Figure 4.1 : Sélection d'une variable de symbole.

4.2. Afficher une carte en cercles proportionnels

La sélection de la variable de symbole provoque l'affichage du dialogue **Types de cartes**. Cliquer sur l'icône du type carte en cercles proportionnels  (fig. 4.2). La carte s'affiche alors à l'écran (fig. 4.3). Si le calibrage des cercles n'a pas été changé en cours de session, le rayon du plus grand cercle, celui qui correspond à la valeur maximum, est de 30 pixels de l'écran d'affichage.

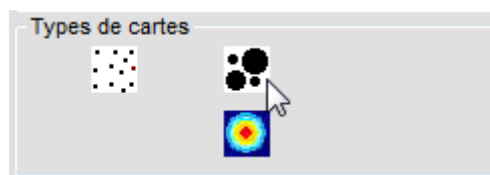


Figure 4.2 : Clic sur l'icône « carte en cercles proportionnels ».

4.3. Le dialogue Cercles

Simultanément au tracé de la carte, le dialogue **Cercles** remplace les dialogues **Variables** et **Types de cartes** dans le cadre **Outils**. (fig. 4.4). Dans le dialogue **Cercles**, des boutons permettent de calibrer les cercles, de changer la couleur des disques et des circonférences, de modifier la légende, d'afficher une carte d'identifiants ou de noms...

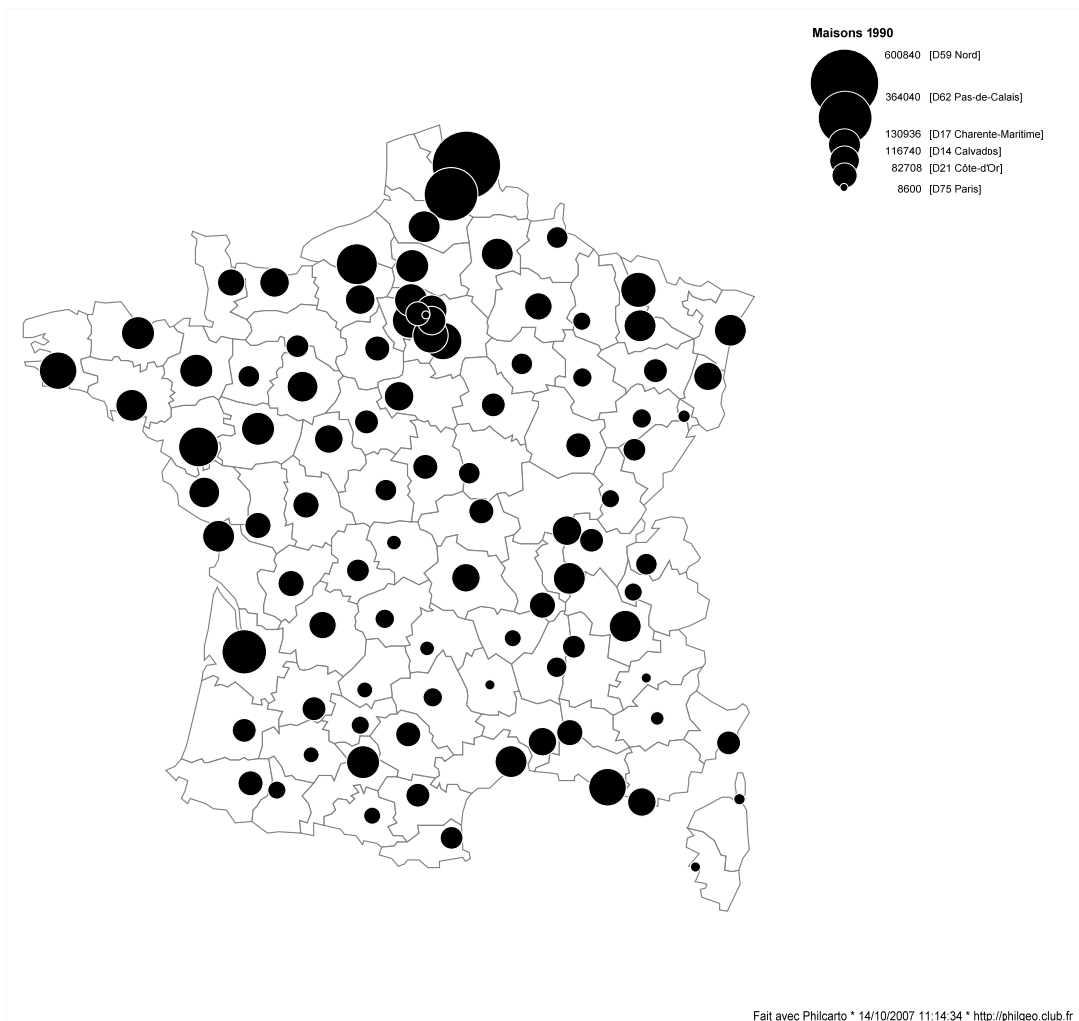


Figure 4.3 : Tracé de la carte en cercles proportionnels.

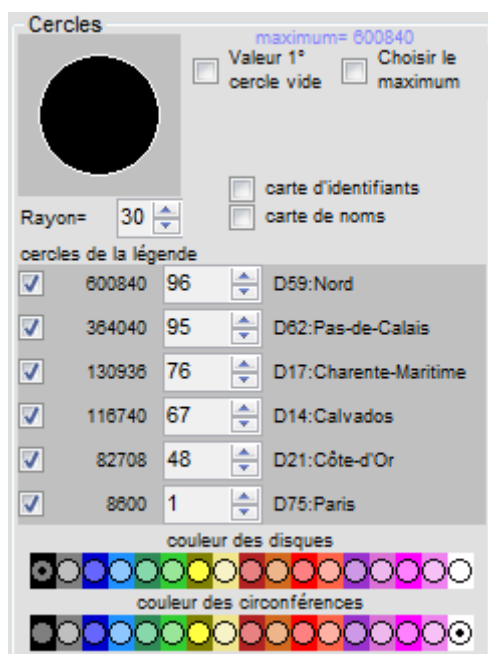


Figure 4.4 : Affichage du dialogue **Cercles**.

4.4. Le calibrage des cercles

Calibrer les cercles consiste à fixer le nombre de pixels d'écran du rayon du plus grand cercle, celui qui correspond à la valeur maximum. En début de session, cette valeur est de 30 pixels ; elle peut varier entre 1 et 40 pixels. Le choix de cette valeur est affaire de perception : la carte ne doit être ni trop claire, ni trop foncée. Pour changer cette valeur, il suffit d'utiliser le curseur **Rayon= 30**. Un clic sur la flèche dirigée vers le haut fait augmenter le nombre de pixels du rayon alors qu'un clic sur la flèche dirigée vers le bas le fait diminuer. Avec 40 pixels, la carte devient plus foncée (fig. 4.5).

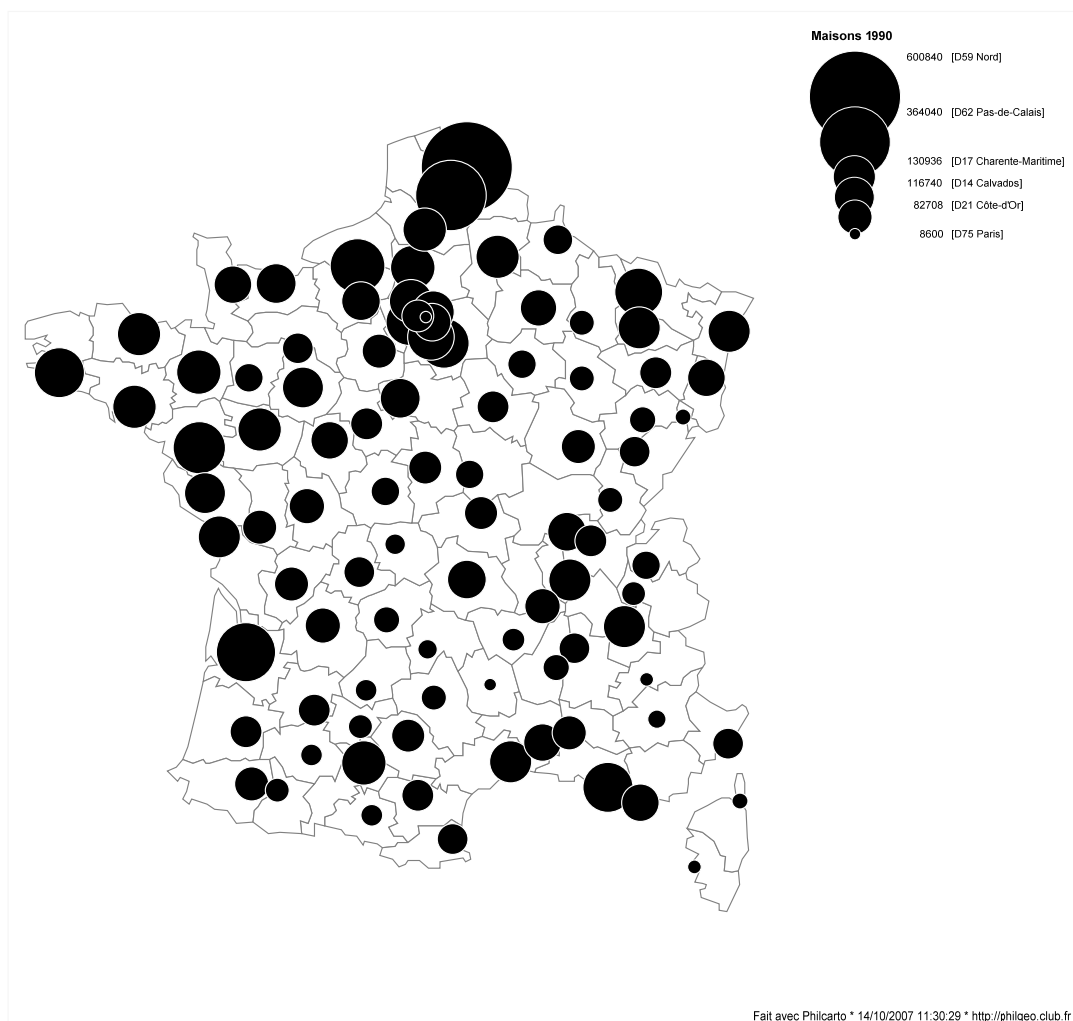


Figure 4.5 : Tracé de la carte en cercles proportionnels avec le rayon maximum calibré à 40.

4.5. Modifier l'apparence des cercles

Par défaut, les disques sont noirs et les circonférences blanches : on dit que les cercles sont détournés. Cela permet de distinguer un petit cercle situé à l'intérieur d'un plus grand, donc de gérer un peu les superpositions. La modification de la couleur de tous les disques et de toutes les circonférences se fait à l'aide de deux séries de boutons radio colorés (fig. 4.6). La première série correspond à la couleur des disques, et la seconde à la couleur des circonférences.



Figure 4.6 : Choisir la couleur des disques et des circonférences.

Pour la carte ci-après (fig. 4.7), on a choisi des disques vert clair et des circonférences bleu foncé.

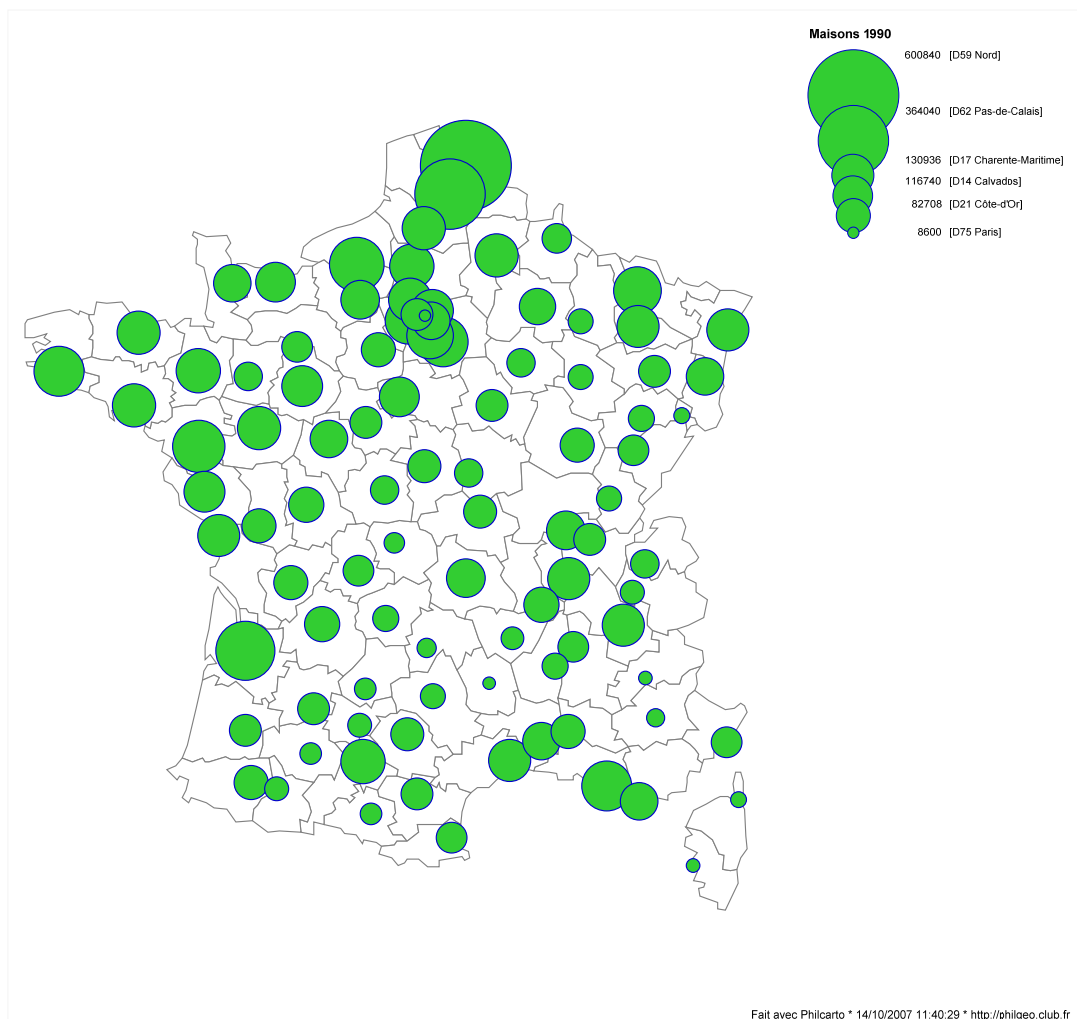


Figure 4.7 : Modification de la couleur des disques et des circonférences.

4.6. Modifier la couleur du fond de la carte

Par défaut, la carte s'affiche sur un fond blanc : il s'agit soit de la couleur de la fenêtre d'affichage, soit de la couleur d'un maillage surfacique associé au maillage ponctuel dont chaque point constitue le centre de chaque cercle. Pour modifier la couleur du fond, il faut choisir un habillage surfacique listé dans le dialogue **Habillage** situé dans la partie inférieure du cadre **Outils**), par exemple, le calque **Regions.S**. Un clic dans la case correspondant à ce calque le sélectionne et fait apparaître une série de boutons radio colorés intitulée **couleur du contour de l'habillage sélectionné** (fig. 4.8). Il faut impérativement choisir l'une de ces couleurs pour que l'habillage puisse être tracé dans cette couleur-là. Par un clic sur le bouton radio correspondant, on peut choisir une couleur passe-partout comme le noir ou le gris, ou une couleur plus expressive à la condition qu'elle soit en harmonie avec la couleur des disques.



Figure 4.8 : La liste des couleurs du contour d'habillage.

Lorsque le bouton de la couleur choisie est sélectionné (ici le gris), la série de boutons radio colorés s'efface, le nom du calque est écrit avec un fond coloré par la couleur choisie, et la carte est à nouveau tracée, faisant apparaître l'habillage retenu. Une nouvelle série de boutons radio colorés s'affiche alors ; intitulée **couleur du fond des habillages surfaciques**, elle permet de choisir la couleur du fond de l'habillage par un clic sur le bouton radio correspondant (fig. 4.9). Notons que ces couleurs sont pâles car le fond ne doit pas parasiter la lecture des cercles ; la couleur du fond doit aussi être choisie en harmonie avec la couleur des disques (fig. 4.10).

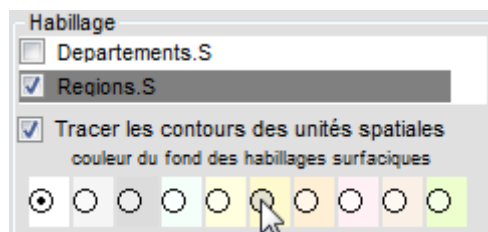


Figure 4.9 : La liste des couleurs du fond d'habillage.

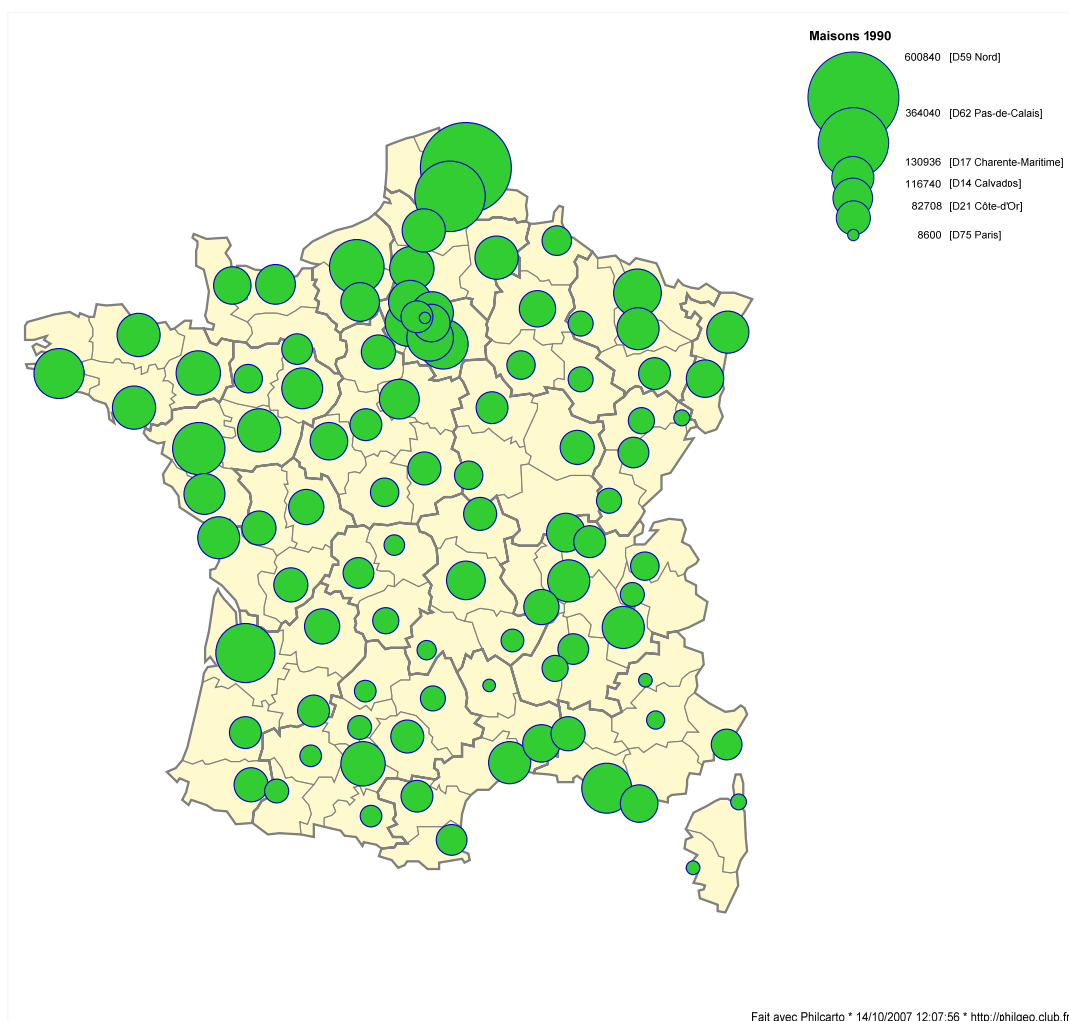


Figure 4.10 : La modification de la couleur du fond de l'habillage

4.7. Choisir les cercles de la légende

La légende d'une carte en cercles est souvent inutilisable. Le traditionnel abaque figurant une courbe sur laquelle peuvent être lues, en théorie, toutes les valeurs comprises entre l'effectif minimum et l'effectif maximum, s'avère bien peu pratique. Une autre solution consiste à faire figurer une série de cercles compris entre le plus petit et le plus grand avec un pas d'accroissement régulier (de 1000 en 1000 par exemple); malheureusement, ces cercles de légende ne correspondent pratiquement jamais à des cercles effectivement présents sur la carte, et leur utilité pour la lecture est toute relative. *Philcarto* préfère figurer en légende un échantillonnage de cercles effectivement présents sur la carte, et indiquer au regard de chaque cercle, la valeur à laquelle il correspond, ainsi que l'identifiant et le nom de l'élément cartographique correspondant (fig. 4.11). Cette présentation, moins classique, présente l'avantage de mieux coller à la réalité des données.

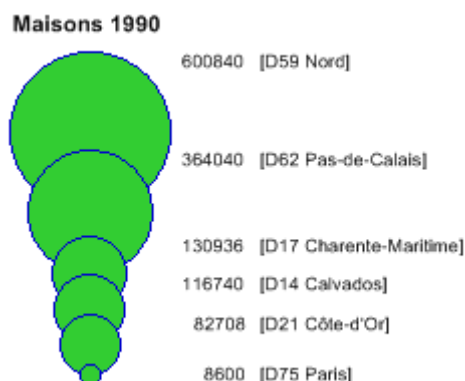


Figure 4.11 : Les cercles de la légende selon l'échantillonnage standard des valeurs.

L'échantillonnage standard des tailles de cercles ne donne pas toujours le meilleur résultat car les distributions de valeurs peuvent s'avérer inadéquates comme c'est le cas sur la figure ci-dessus où trois cercles sont à peu près de même taille. Pour pallier cet inconvénient, *Philcarto* permet de choisir le nombre de cercles et la valeur correspondant à chacun d'eux. Ceci se fait simplement à l'aide d'un ensemble de contrôles rassemblés sous le nom **cercles de la légende** (fig. 4.12).

cercles de la légende

☒	600840	96	D59:Nord
☒	364040	95	D62:Pas-de-Calais
☒	130936	76	D17:Charente-Maritime
☒	116740	67	D14:Calvados
☒	82708	48	D21:Côte-d'Or
☒	8600	1	D75:Paris

Figure 4.12 : Les contrôles permettant de choisir les cercles de la légende.

Ces contrôles sont agencés sous la forme d'un tableau comprenant six lignes correspondant chacune à un cercle de légende ; la légende peut donc se composer de six cercles au plus ; la ligne supérieure correspond au plus grand cercle et la ligne inférieure au plus petit. Chacune des colonnes de ce tableau a une fonction précise. De gauche à droite, on trouve :

- des cases à cocher qui, si elles sont cochées, commandent chacune l'affichage d'un cercle ;
- des valeurs numériques donnant, pour chaque cercle, la valeur statistique qui le caractérise ;
- des curseurs dont les valeurs correspondent, pour chaque cercle, à son rang dans la série des cercles triés du plus petit (rang = 1) au plus grand (rang = nombre total de valeurs) ;
- des textes donnant, pour chaque cercle, l'identifiant et le nom de l'élément cartographique correspondant ; par exemple, le cercle correspondant à la valeur 600840 résidences secondaires est celui du département du Nord, identifié par le code D59.

Grâce à ce tableau de contrôle, il est possible d'élaborer une légende explicite composée de 1 à 6 cercles relatifs à des cercles effectivement présents sur la carte (fig. 4.13 et 4.14).

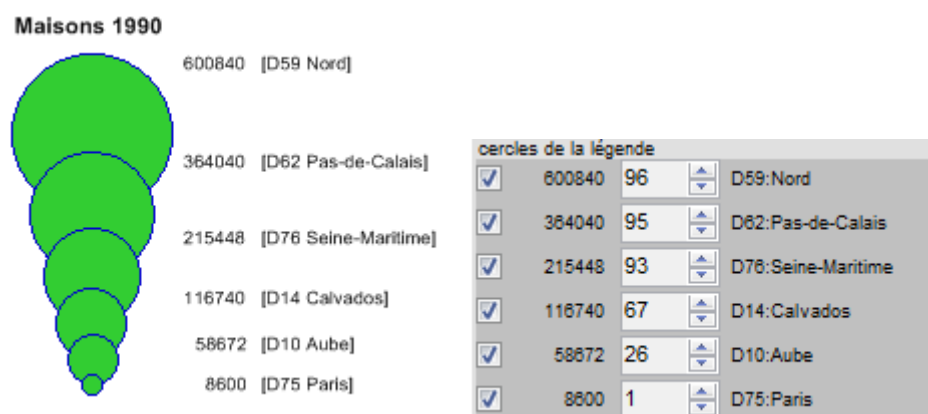


Figure 4.13 : Une légende en 6 cercles et le tableau de contrôles correspondant.

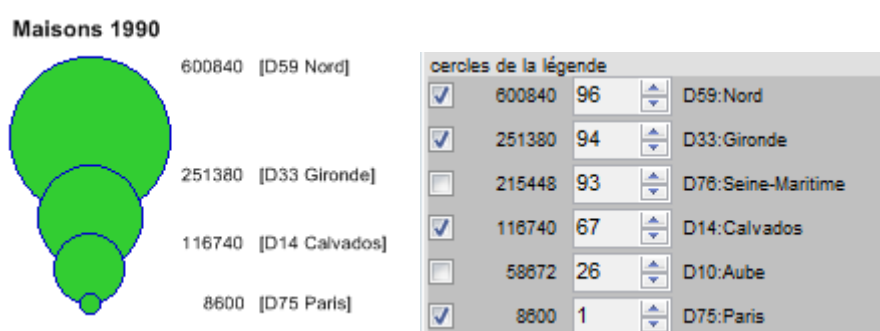


Figure 4.14 : Une légende en 4 cercles et le tableau de contrôles correspondant.

4.8. Disque ou circonférence ?

Dans certains cas, notamment lorsque l'accumulation des cercles dans un secteur de la carte rend la lecture difficile, il peut apparaître utile de ne pas remplir les cercles par une couleur ; en d'autres termes, il s'agit de tracer la circonférence et de laisser le disque sans couleur et cela à partir d'une valeur donnée. La case à cocher ☒ Valeur 1^{re} cercle vide, lorsqu'elle est cochée provoque l'affichage d'une zone d'édition et d'un bouton (4.15).

☒ Valeur 1^{re}
cercle vide

Appliquer

Figure 4.15 : Le dialogue pour le tracé des circonférences seules.

La zone d'édition permet d'entrer la valeur au-delà de laquelle les disques restent vides ; le bouton **Appliquer** valide cette valeur et provoque un nouvel affichage de la carte traduisant la valeur entrée dans la zone d'édition.

Par exemple, sur la carte du nombre de résidences avec WC extérieurs, on souhaite que le disque correspondant à Paris ne soit pas coloré. La valeur de Paris s'élevant à 139612, on entre dans la zone d'édition 135000, nombre inférieur à celui de Paris et supérieur à 127688 compté dans le Nord, second département pour les résidences avec WC extérieurs. Ainsi, la région Île-de-France apparaît moins chargée (fig. 4.16).

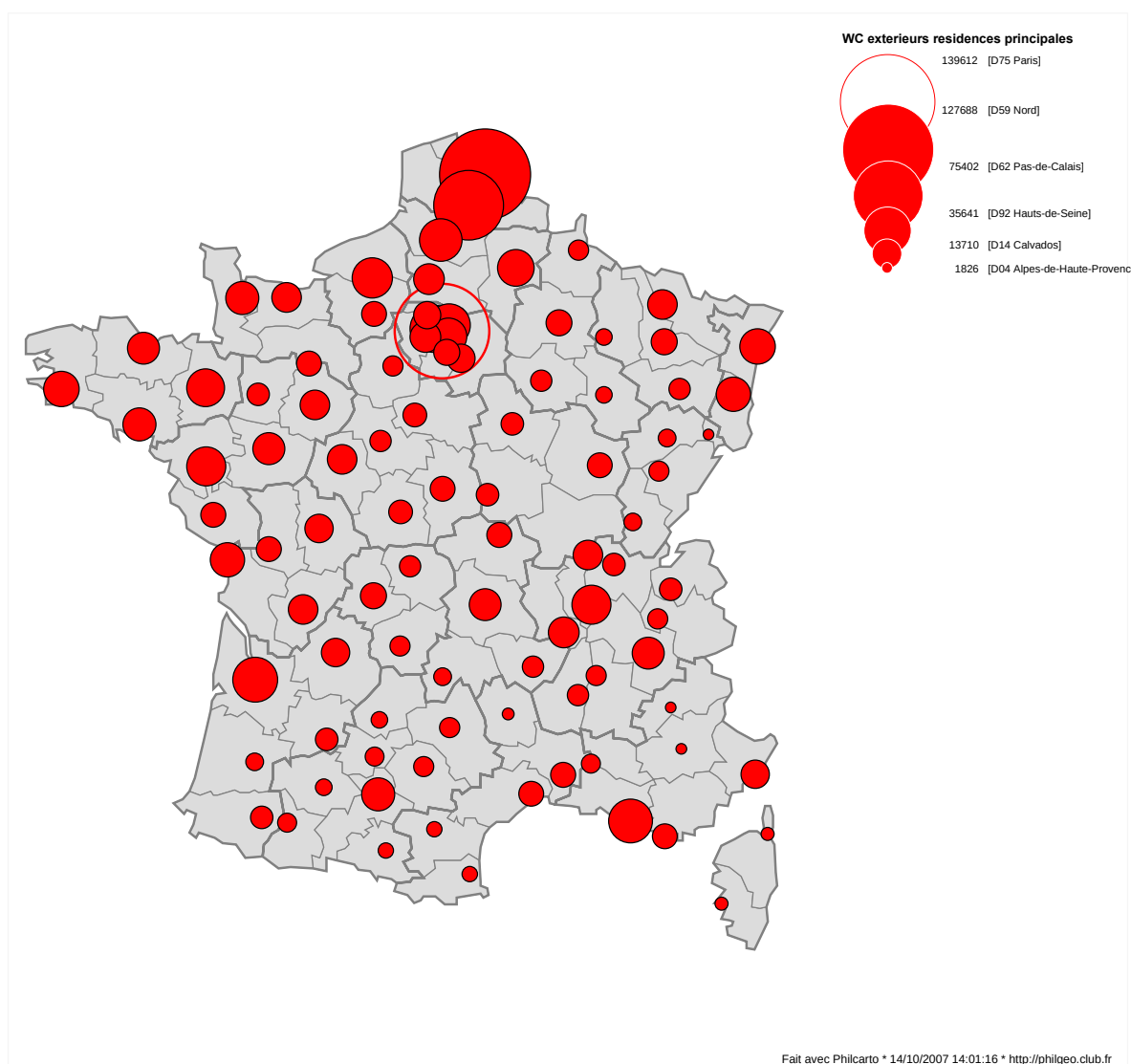


Figure 4.16 : Sur cette carte, le cercle correspondant à Paris est tracé avec sa seule circonférence.

Notons que la circonférence de Paris est tracée de la couleur du disque des autres cercles et que ces autres cercles ont une circonférence noire choisie par l'utilisateur. Signalons aussi que la valeur 0 entrée dans la zone d'édition signifie que pour tous les cercles de la carte, c'est seulement la circonférence qui doit être tracée (fig. 4.17).

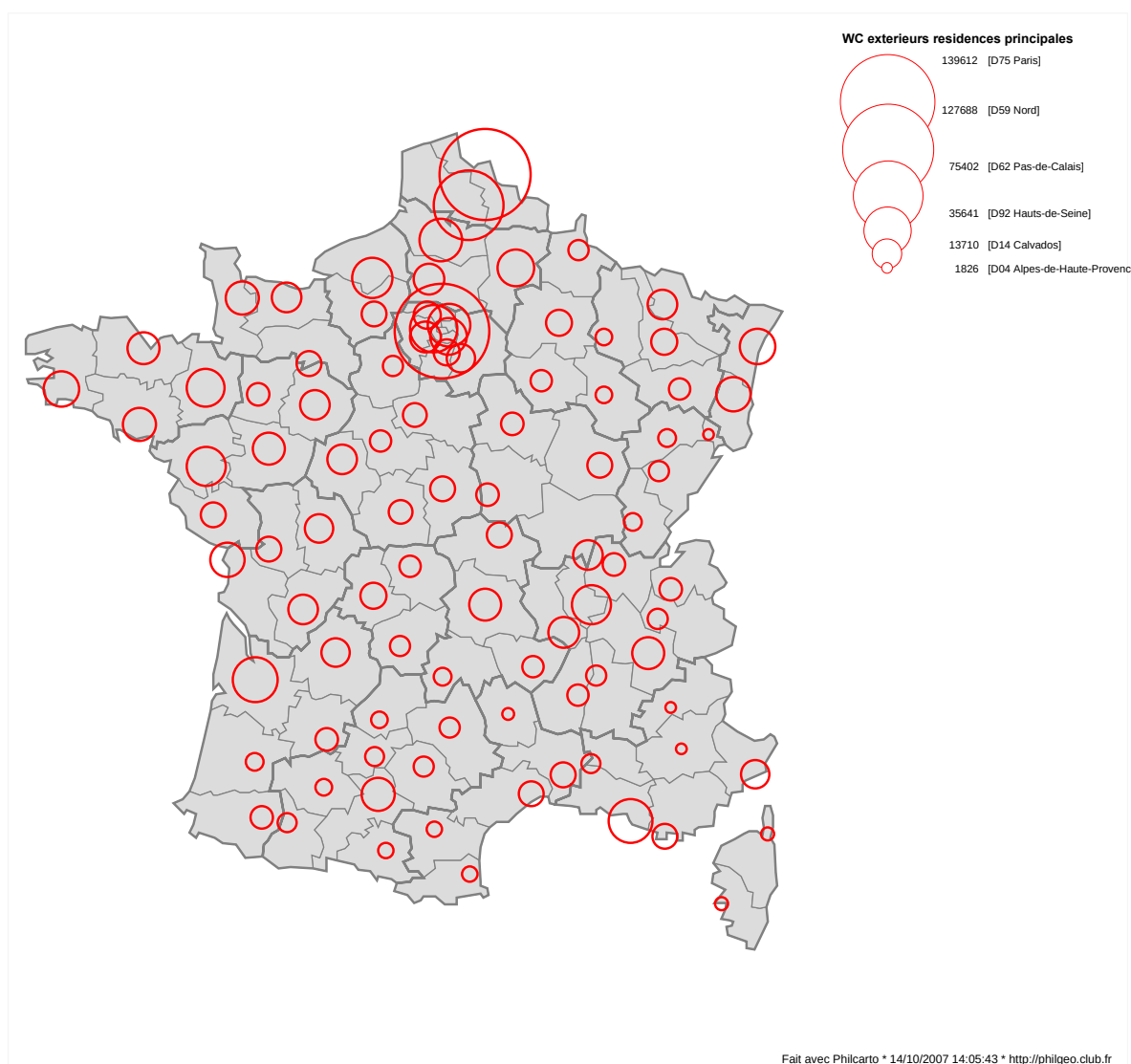


Figure 4.17 : Sur cette carte, seules les circonférences des cercles sont tracées.

4.9. Tracer plusieurs cartes en cercles comparables entre elles

Par défaut, le calibrage des cercles est différent d'une carte à l'autre. De ce fait, les cartes qui représentent des sous-populations extraites d'une même population (par exemple les résidences principales en fonction de la période de construction) ne sont pas comparables entre elles, ce qui peut induire de sérieux problèmes d'interprétation.

Pour résoudre ce problème, *Philcarto* offre la possibilité d'entrer une valeur maximum différente de la valeur maximum de chaque sous-population ; lorsqu'elle existe, cette valeur maximum se substitue à celle de chacune de variable. Ainsi, le calibrage des cercles est identique pour chaque carte et les cartes peuvent être comparées.

Dans le cas du nombre de résidences principales par période de construction, les maximums des trois périodes sont les suivants : avant 1949 : 919 414 ; de 1949 à 1981 : 437 092 ; après 1981 : 108 216. Si l'on n'entre pas de valeur maximum, chaque carte est calibrée en fonction de son propre maximum, et l'on obtient la planche de cartes suivante (fig. 4.18) :

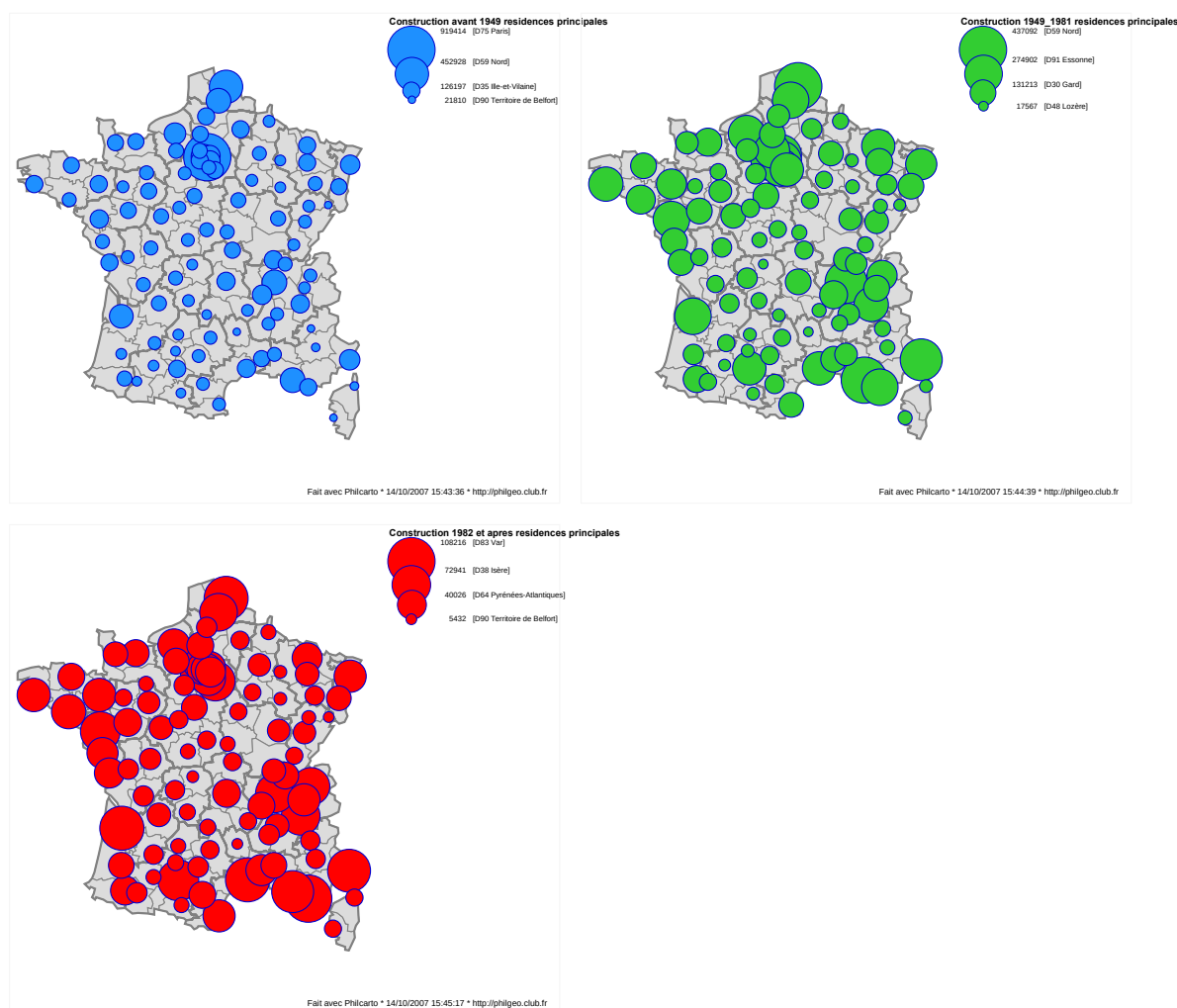


Figure 4.18 : Trois cartes en cercles représentant trois sous-populations extraites de la même population ; le calibrage des cercles est basé sur le maximum de chaque variable.

La planche de cartes obtenue conduit à la conclusion suivante : le nombre de résidences principales construites en 1949 et après est toujours plus élevé que celui des résidences principales construites avant 1949...

Au lieu de laisser *Philcarto* faire le calibrage en fonction du maximum de chaque variable, contraignons-le à utiliser, pour les trois cartes, le maximum des trois maximums, c'est-à-dire 919 414. Pour simplifier, entrons la valeur 920 000 qui correspond au nombre de résidences construites avant 1949, dans la zone d'édition qui s'affiche lorsque la case ☒ Choisir le maximum est cochée (fig. 4.19). Validons cette valeur avec le bouton **Appliquer**, puis traçons à nouveau les trois cartes (fig. 4.20).

☒ Choisir le maximum
 920000
 Appliquer

Figure 4.19 : Entrée de la valeur maximum des trois variables

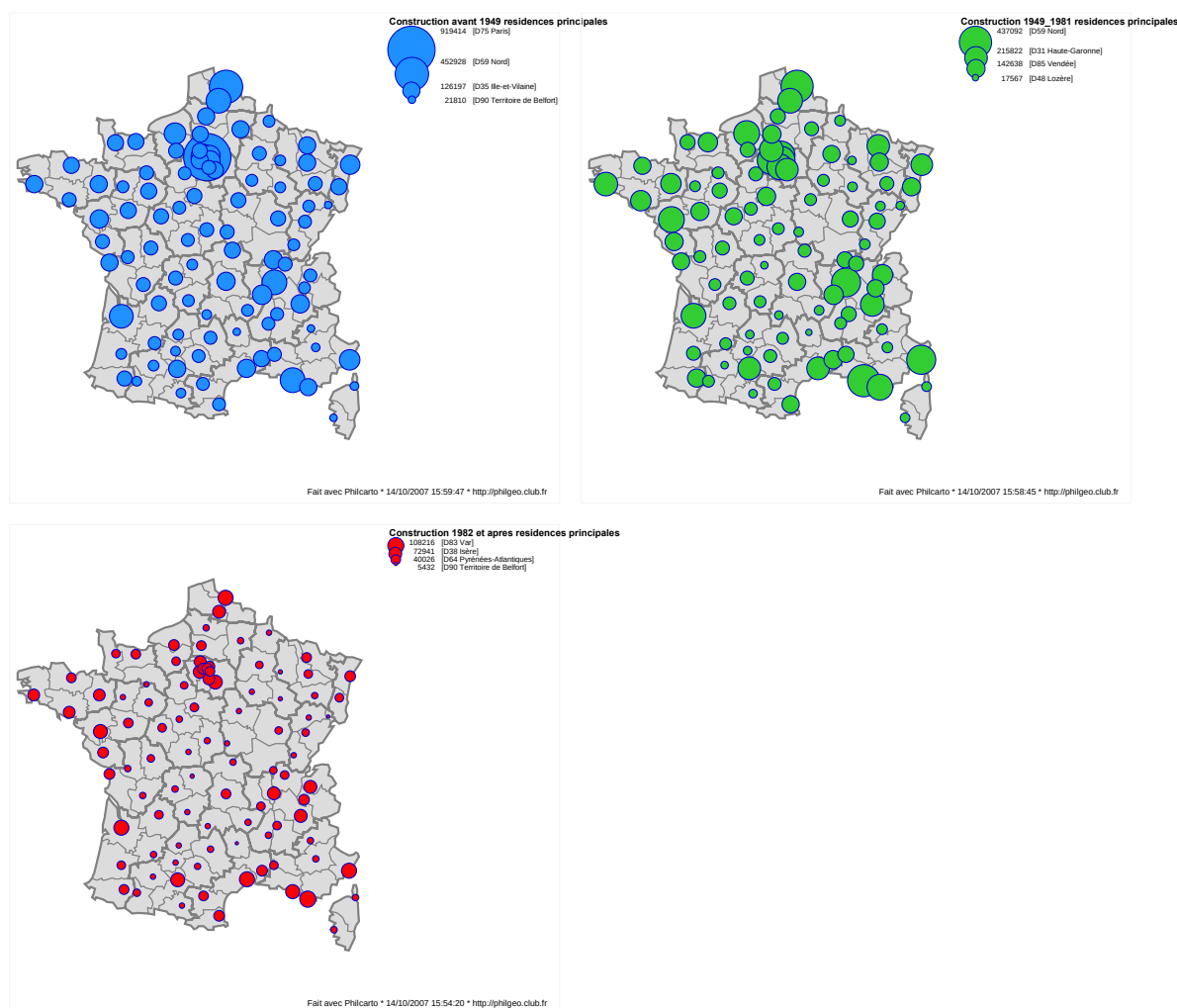


Figure 4.20 : Trois cartes en cercles représentant trois sous-populations extraites de la même population ; le calibrage des cercles est basé sur le maximum des maximums.

Si la carte aux cercles bleus (avant 1949) reste inchangée, les deux autres cartes apparaissent extrêmement différentes de celles de la planche précédente : la période postérieure à 1982 affiche des cercles beaucoup plus petits. On observe aussi l'intense construction de logements dans les départements notamment possédant une façade maritime, alors que dans le centre du pays la construction apparaît plus ralentie.

Notons enfin que, pour revenir au calibrage individuel de chaque carte, il suffit de décocher la case ☐ Choisir le maximum.

4.10. Cercles pour représenter des effectifs négatifs et positifs

Dans les versions précédentes de *Philcarto*, le tracé des cercles n'était possible que sur des effectifs strictement positifs. Les effectifs inférieurs ou égaux à zéro n'étaient pas cartographiés. Pour contourner le problème, il fallait créer une nouvelle variable contenant la valeur absolue de l'effectif, une seconde variable binaire indiquant si oui (=1) ou non (=0) la valeur était positive. Il n'est plus nécessaire de procéder ainsi car *Philcarto* détecte s'il existe des valeurs négatives et positives dans la variable à cartographier. Si cela est le cas, il calibre les cercles en fonction de la

valeur absolue de l'effectif et figure les cercles positifs en rouge, et les cercles négatifs en bleu.

La carte du solde migratoire des départements français entre 1990 et 1999 montre l'intérêt de genre de tracé : l'héliotropisme est évident (fig. 4.21).

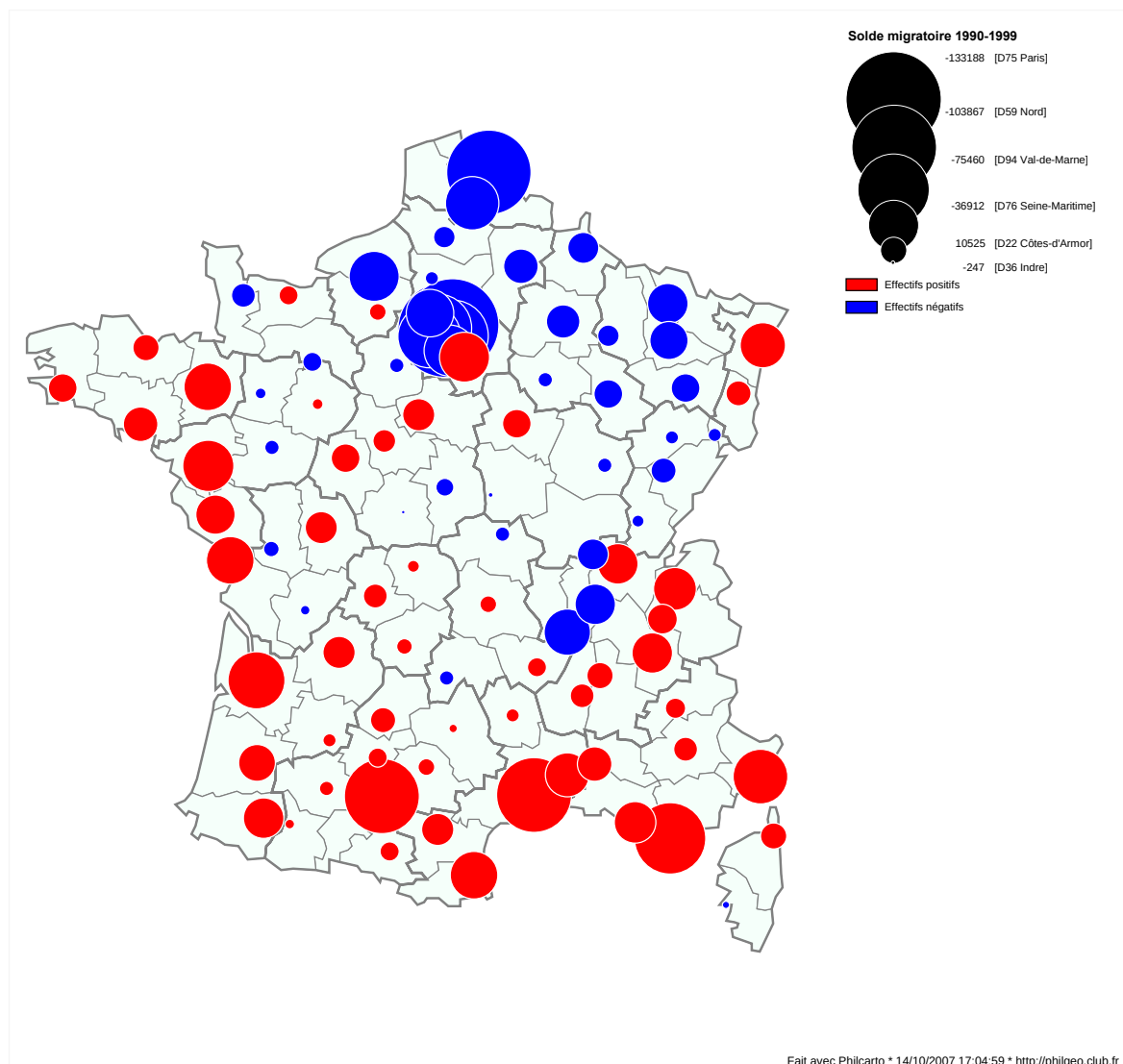


Figure 4.21 : La carte du solde migratoire des départements français entre 1990 et 1999.

Notons que dans le cas de valeurs négatives et positives, on perd la possibilité de choisir la couleur des disques qui sont obligatoirement rouges ou bleus.

4.11. Cartes d'identifiants et/ou de noms.

Le dialogue **Cercles** permet en outre de tracer des cartes sur lesquelles figurent soit les identifiants des éléments cartographiques, soit les noms, soit les deux. Ils sont écrits sur chaque élément cartographique ponctuel. Il suffit de cocher les cases

☐ carte d'identifiants
☐ carte de noms

. Si ces deux cases sont cochées, une case supplémentaire apparaît

pour écrire ces deux éléments sur deux lignes

☒ carte d'identifiants
☒ carte de noms
☒ sur 2 lignes

. Le résultat est

rarement parfait, mais les collisions peuvent être corrigées dans un logiciel de dessin (fig. 4.22).



5. Les cartes en cercles proportionnels sur cartes choroplèthes

Dossier : \Philcarto exemples\France départements\
Fond de carte : France carte départements.ai
Statistiques : France data départements.txt

5.1. Sélectionner une variable de couleur et une variable de symbole

La sélection des deux variables se fait par un clic sur le nom d'une des variables figurant dans la liste **couleurs** et sur celui d'une seconde variable figurant dans la liste **symboles** du dialogue **Variables**. Ici les variables sélectionnées sont respectivement **Maisons 1990 %logements** et **Maisons 1990** (fig. 5.1).

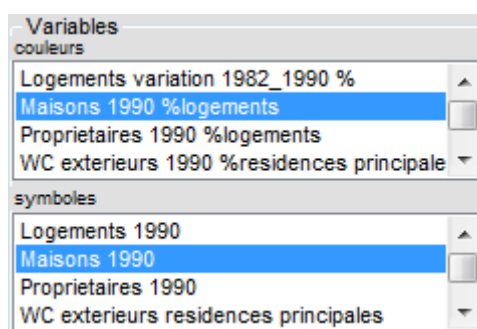


Figure 5.1 : Sélection d'une variable de couleur et de symbole.

5.2. Afficher une carte en cercles sur une carte choroplèthe

La sélection de la variable de symbole provoque l'affichage du dialogue **Types de cartes**. Cliquer sur l'icône du type carte en cercles proportionnels  (fig. 5.2). La carte s'affiche alors à l'écran (fig. 5.3). Toutes les options en cours sont appliquées, aussi bien pour la carte choroplèthe (discrétisation, nombre de classes, gamme de couleurs...) que pour la carte en cercles (calibrage, couleur des cercles, couleur des circonférences).

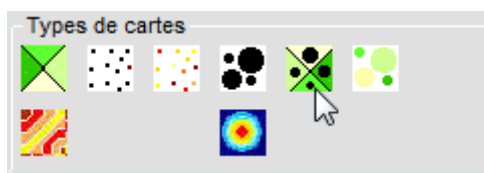


Figure 5.2 : Clic sur l'icône « carte en cercles proportionnels sur carte choroplèthe ».

Naturellement, la surface du cercle gêne la lecture des couleurs dans le cas où ces cercles occultent les surfaces colorées. Dans ce cas, il suffit de ne pas tracer le disque, mais uniquement la circonférence. Ceci se fait en cochant la case ☒ Valeur 1° cercle vide, et en entrant la valeur 0 dans la zone d'édition, ce qui signifie que tous les cercles doivent être évidés (fig. 5.4).

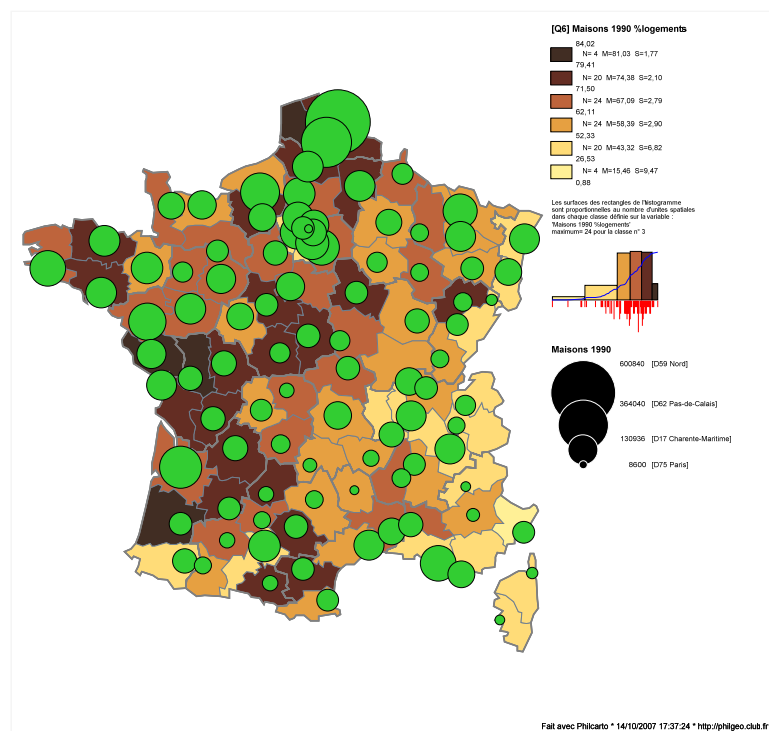


Figure 5.3 : La carte choroplèthe du % de maison individuelles dans l'ensemble des logements sur laquelle est tracée la carte en cercles proportionnels du nombre de maisons individuelles.

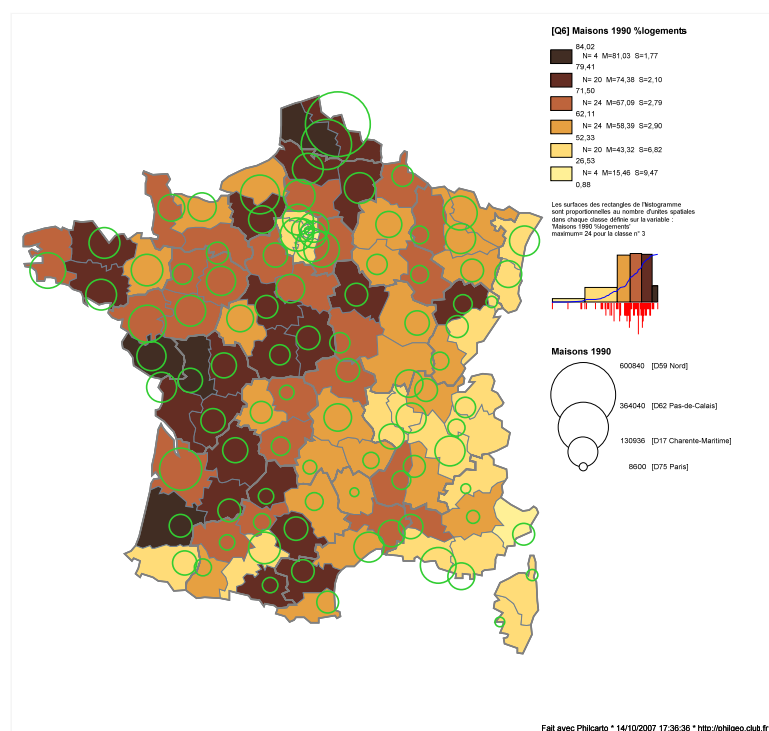


Figure 5.4 : La même carte mais avec le tracé des circonférences seulement.

6. Les cartes en cercles proportionnels colorés

Dossier : \Philcarto exemples\France départements\
Fond de carte : France carte départements.ai
Statistiques : France data départements.txt

6.1. Sélectionner une variable de couleur et une variable de symbole

La sélection des deux variables se fait par un clic sur le nom d'une des variables figurant dans la liste **couleurs** et sur celui d'une seconde variable figurant dans la liste **symboles** du dialogue **Variables**. Ici les variables sélectionnées sont respectivement **Maisons 1990 %logements** et **Maisons 1990** (fig. 6.1).

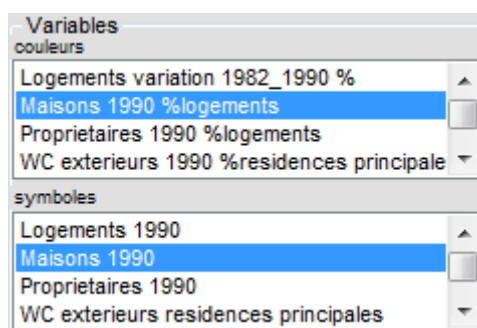


Figure 6.1 : Sélection d'une variable de couleur et de symbole.

6.2. Afficher une carte en cercles proportionnels colorés

La sélection de la variable de symbole provoque l'affichage du dialogue **Types de cartes**. Cliquer sur l'icône du type carte en cercles proportionnels colorés (fig. 6.2). La carte s'affiche alors à l'écran (fig. 6.3). Toutes les options en cours sont appliquées, aussi bien pour l'affectation des couleurs (discrétisation, nombre de classes, gamme de couleurs...) que pour les cercles (calibrage, couleur des circonférences). Dans le cas où la variable de couleurs est une variable nominale, les couleurs sont affectées en fonction des modalités triées et de l'ordre des couleurs dans la gamme de couleurs « perroquet » (la première modalité est représentée en rouge, la seconde en bleu, etc.).

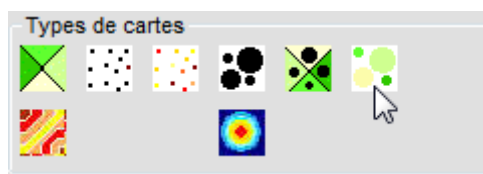


Figure 6.2 : Clic sur l'icône « carte en cercles proportionnels colorés ».

Un habillage, ici les régions, complète la carte et permet d'afficher les cercles sur un fond coloré qui fait ressortir les couleurs. Attention ! penser à choisir une couleur de fond qui n'existe pas dans la gamme de couleurs utilisée pour colorier les cercles (fig. 6.4).

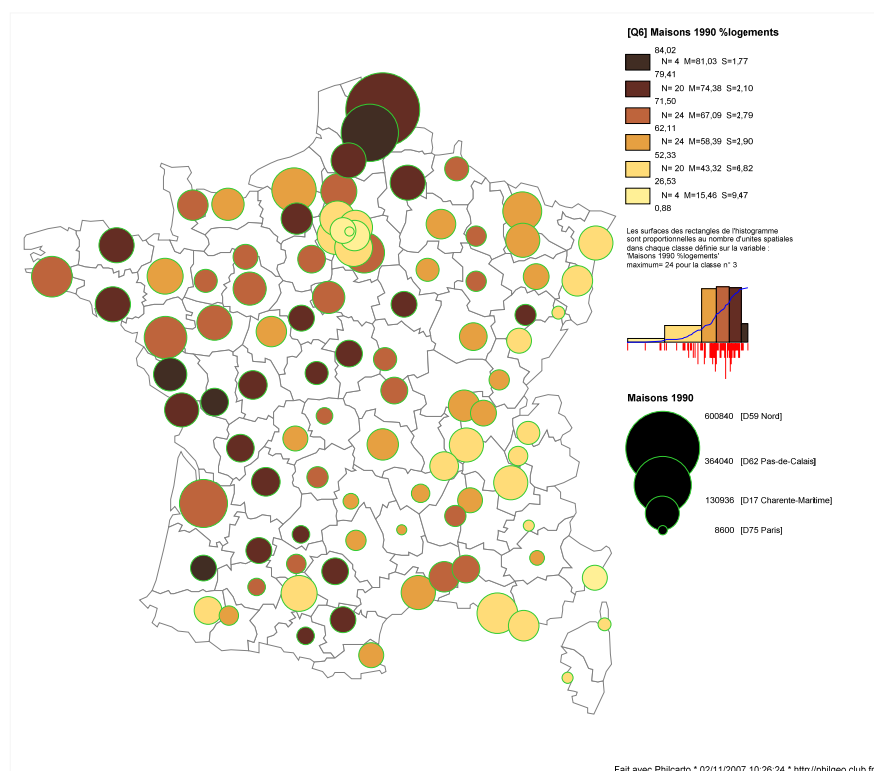


Figure 6.3 : La carte en cercles proportionnels du nombre de maisons individuelles colorés par le % de maison individuelles dans l'ensemble des logements.

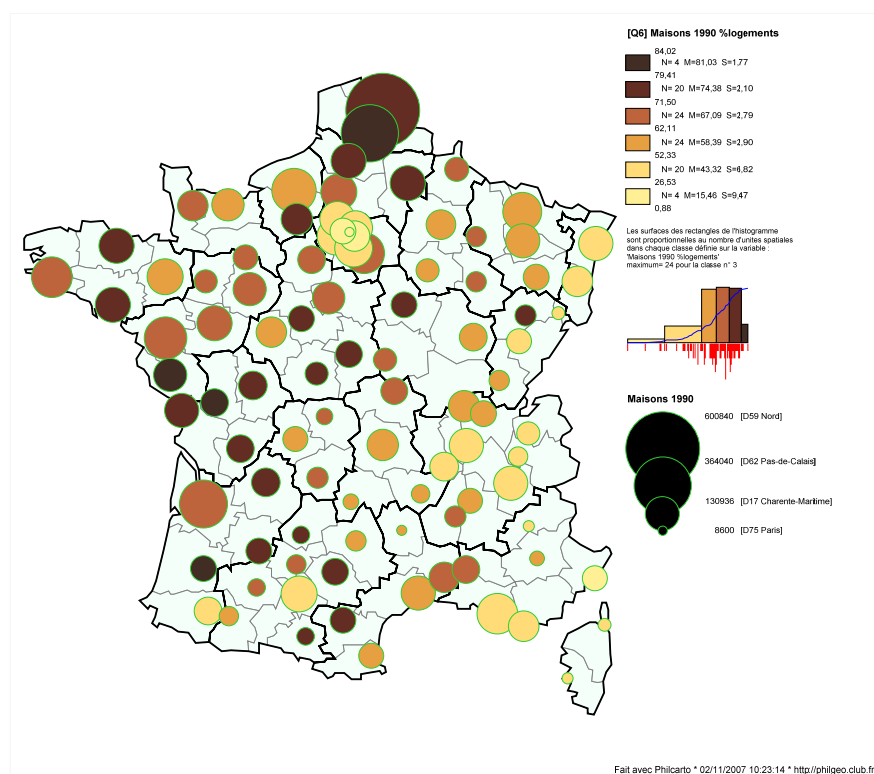


Figure 6.4 : La même carte avec un habillage régional avec un fond de couleur cyan clair.

Dans le cas où la variable d'effectifs comprend des valeurs négatives et positives, c'est la couleur de la circonférence des cercles qui l'indique : rouge clair pour les effectifs positifs et bleu pour les effectifs négatifs. Dans ce cas, on ne peut plus choisir la couleur des circonférences (fig. 6.5).

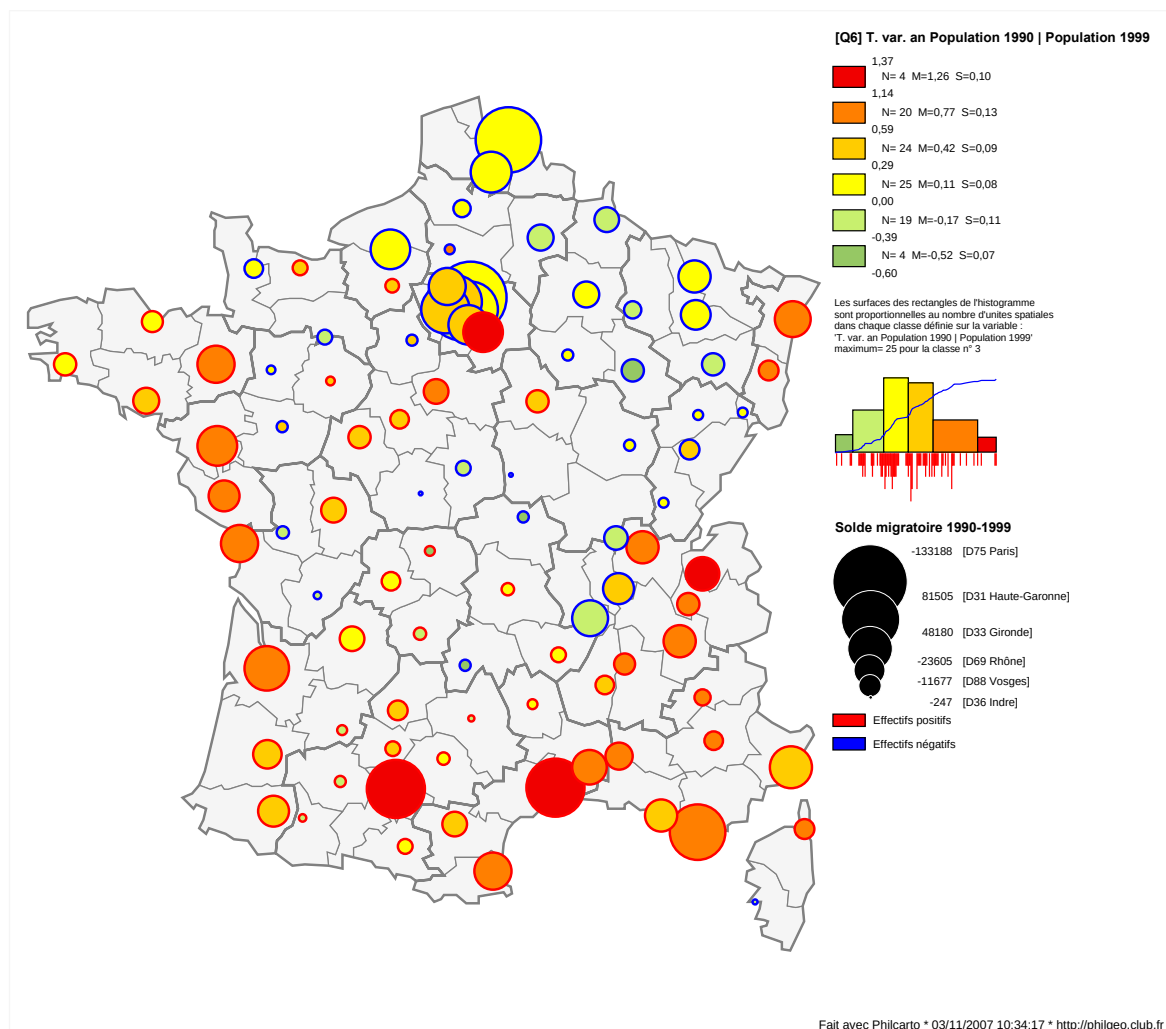


Figure 6.5 : La carte en cercles proportionnels colorés des soldes migratoires colorés par le taux moyen annuel de variation de la population. La couleur de la circonférence des cercles indique si les effectifs du solde migratoire sont négatifs (bleu) ou positifs (rouge).

7. Les cartes en cercles proportionnels affrontés ou en diagrammes à secteurs

Dossier : \Philcarto exemples\France départements\
Fond de carte : France carte départements.ai
Statistiques : France data départements.txt

7.1. Sélectionner deux variables de symbole

La sélection des deux variables se fait par deux clics, l'un pour la première variable, l'autre pour la seconde, sur un nom figurant dans la liste **symboles** du dialogue **Variables**. Ici les variables sélectionnées sont respectivement **Construction avant 1949** et **Construction apres 1949** (fig. 7.1).

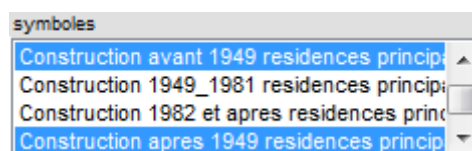


Figure 7.1 : Sélection d'une variable de couleur et de symbole.

Le choix de ces deux variables est guidé par la question du poids respectif des logements anciens et récents dans l'ensemble du parc de résidences principales.

7.2. Afficher la carte en cercles proportionnels affrontés

La sélection de la variable de symbole provoque l'affichage du dialogue **Types de cartes**. Cliquer sur l'icône du type carte en cercles proportionnels affrontés ou en diagrammes à secteurs  (fig. 7.2). La carte s'affiche alors à l'écran (fig. 7.3). Toutes les options en cours sont appliquées comme le calibrage des cercles ou l'habillage.

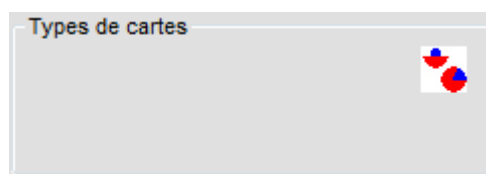


Figure 7.2 : Clic sur l'icône « carte en cercles affrontés ou en diagrammes à secteurs ».

Sauf si la valeur maximum a été préalablement choisie, le calibrage se fait par rapport à la variable qui présente l'effectif le plus élevé ; c'est aussi cette variable là qui figure en légende. L'affectation des couleurs se fait dans l'ordre des variables dans la liste **symboles** du dialogue **Variables**, c'est-à-dire rouge pour la première variable dont le demi-cercle inférieur est tracé, et bleu pour la seconde variable représentée par son demi cercle supérieur.

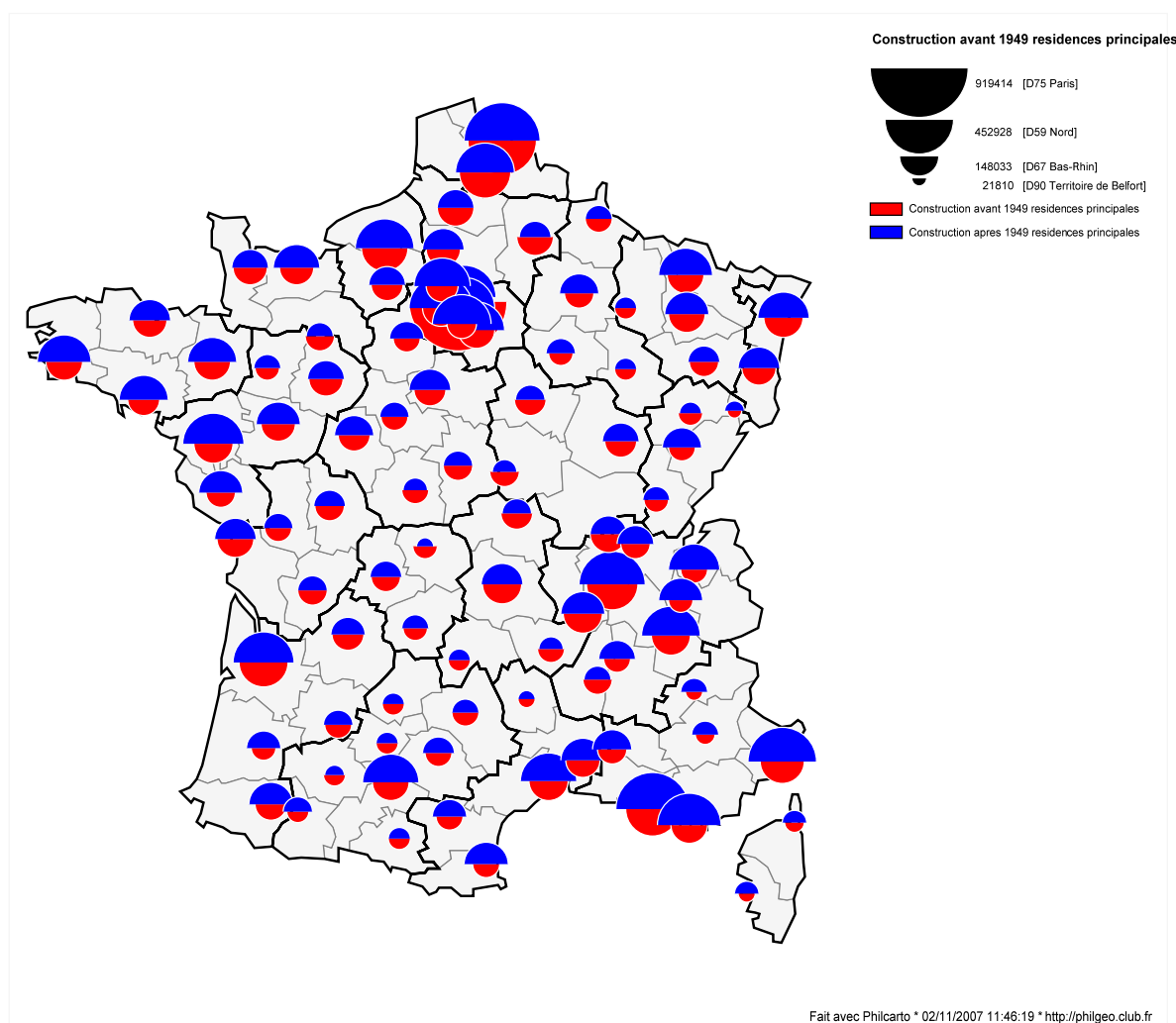


Figure 7.3 : La carte en cercles proportionnels affrontés du nombre de résidences principales construites avant 1949 et après 1949.

7.3. Afficher la carte en diagrammes à secteurs proportionnels

Pour obtenir une carte en diagrammes à secteurs, il suffit de décocher la case  . Sur la carte obtenue (fig. 7.4), la surface des cercles représente maintenant la sommes des effectifs des deux variables sélectionnées. Chaque secteur est proportionnel au pourcentage de la variable qu'il représente dans la somme des effectifs des deux variables.

Ce genre de représentation est, en général, à éviter. Dans certains cas, assez rares, elle peut s'avérer efficace, notamment lorsque la variation de la taille des secteurs suit une direction particulière.

Par ailleurs, rien n'interdit la sélection de plus de deux variables, mais on obtient alors une carte difficile à lire dont la réalisation ne se justifie que lorsque le nombre de cercles est petit (fig. 7.5).

 Les cartes en diagrammes à secteurs proportionnels peuvent être enregistrées seulement au format **.emf**.

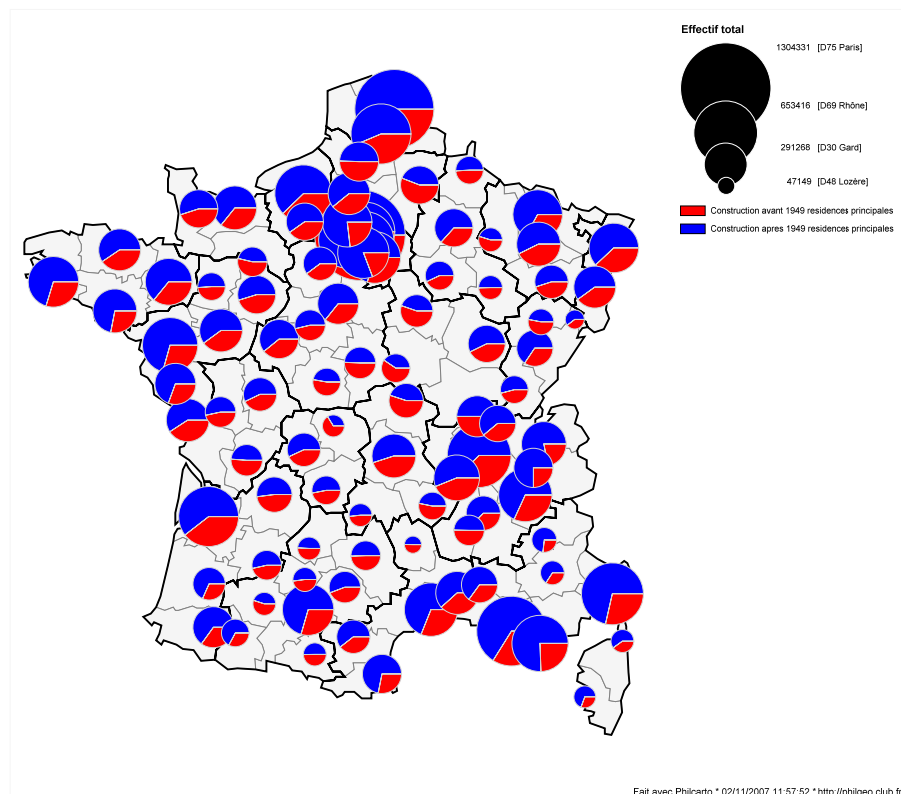


Figure 7.4 : La carte en diagrammes à secteurs proportionnels du nombre de résidences principales construites avant 1949 et après 1949.

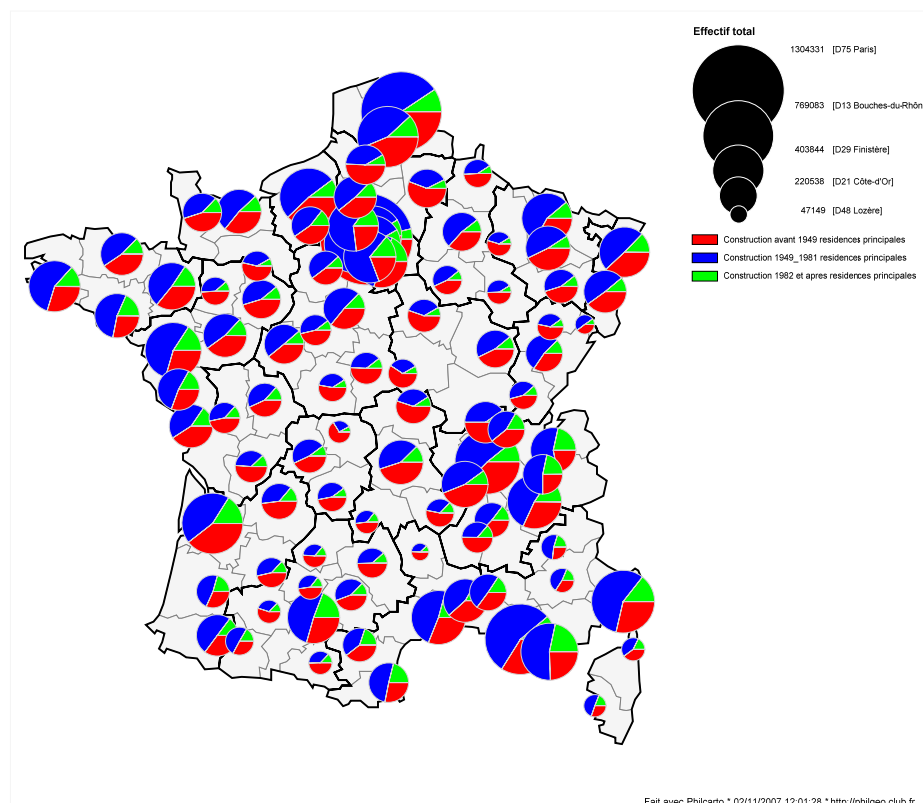


Figure 7.5 : La carte en diagrammes à secteurs proportionnels du nombre de résidences principales construites avant 1949, entre 1949 et 1981, et après 1981.

8. Les cartes en semis de points

Dossier : \Philcarto exemples\France départements\
Fond de carte : France carte départements.ai
Statistiques : France data départements.txt

8.1. Sélectionner une variable de symbole

La sélection d'une variable se fait simplement par un clic sur le nom d'une des variables figurant dans la liste **symboles** du dialogue **Variables**. Ici la variable sélectionnée, **Maisons 1990**, contient le nombre de maisons individuelles dans chaque département de France métropolitaine au recensement de 1990 (fig. 8.1).

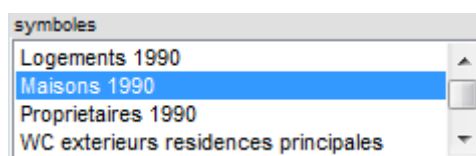


Figure 8.1 : Sélection d'une variable de symbole.

8.2. Afficher une carte en semis de points

La sélection de la variable de symbole provoque l'affichage du dialogue **Types de cartes**. Dans ce dialogue, cliquer sur l'icône du type carte en semis de points (fig. 8.2). La carte s'affiche alors à l'écran (fig. 8.3). Toutes les options en cours sont appliquées comme le calibrage des points ou l'habillage.

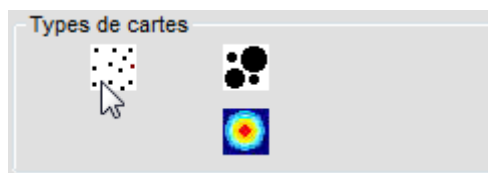


Figure 8.2 : Clic sur l'icône « carte en semis de points ».

8.3. Le dialogue Semis de points

Simultanément au tracé de la carte, le dialogue **Semis de points** remplace les dialogues **Variables** et **Types de cartes** dans le cadre **Outils**. (fig. 8.4). Dans le dialogue **Semis de points**, des contrôles permettent de calibrer les points, de changer leur taille et leur couleur, de modifier la couleur du fond de la carte.

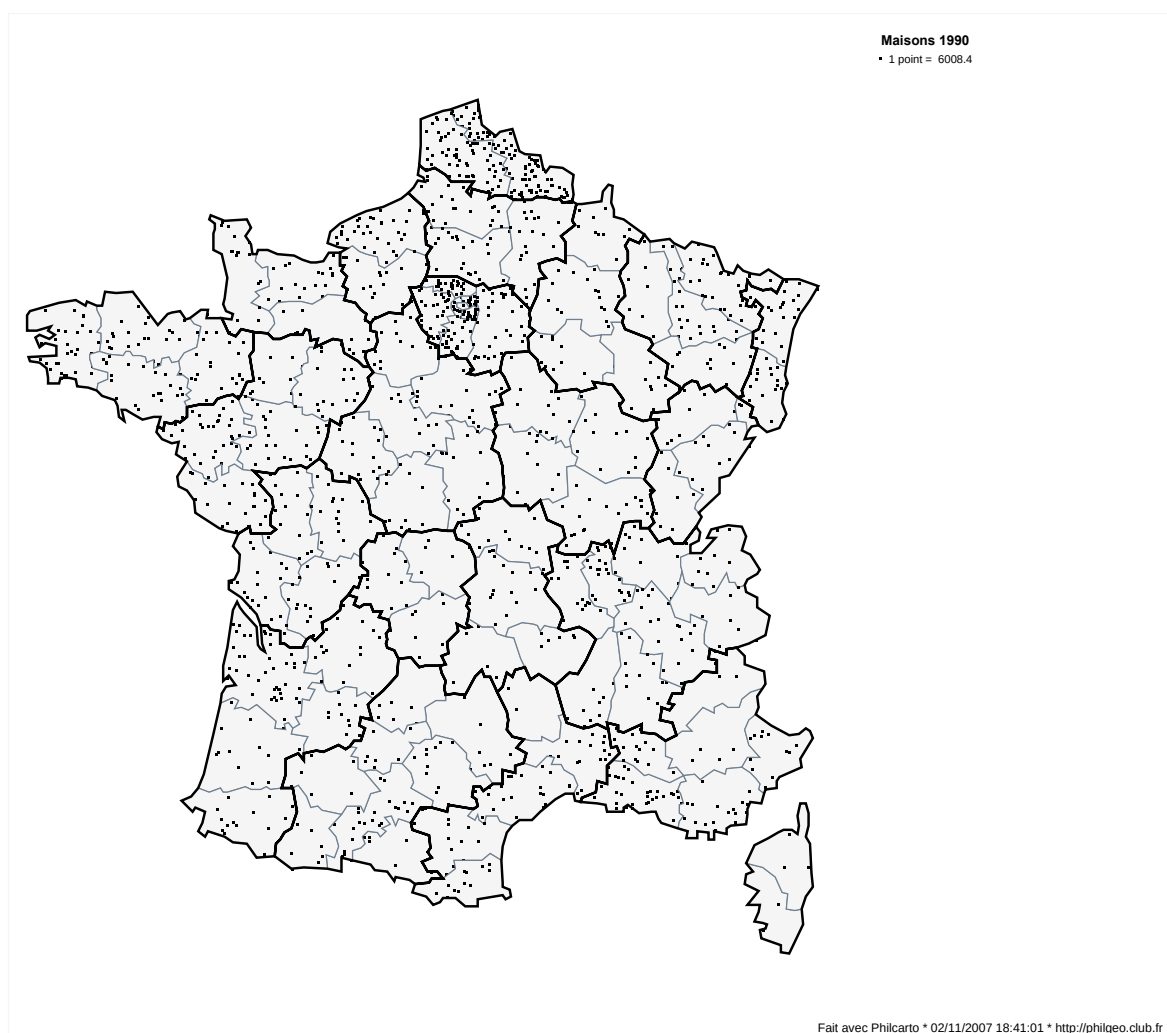


Figure 8.3 : Tracé de la carte en semis de points du nombre de maisons individuelles.

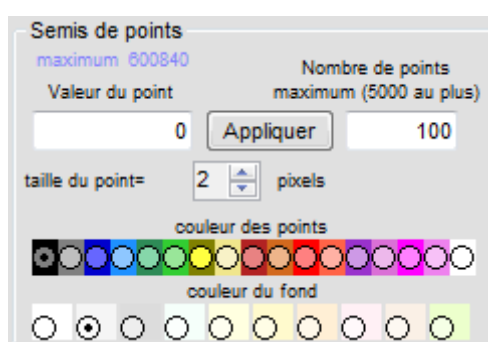


Figure 8.4 : Affichage du dialogue **Semis de points**.

8.4. Le calibrage du point

Il y a deux façons de calibrer la valeur de chaque point du semis. La première, choisie par défaut, consiste à indiquer le nombre de points (fixé arbitrairement à 100 en début de session) correspondant à la valeur maximum de la variable cartographiée ; dans ce cas, la valeur du point est le résultat de la division de cette valeur maximum par le nombre de points. Le nombre maximum de points autorisé

par *Philcarto* s'élève à 5000 ; il est préférable de choisir un nombre plus petit pour ne pas charger, de manière déraisonnable, les fichiers d'enregistrement de ce genre de carte.

Pour modifier le nombre maximum de points, il suffit d'entrer le nombre choisi dans la zone d'édition prévue à cet effet et de valider avec le bouton **Appliquer** (fig. 8.5), par exemple, 300 points correspondent à 600 840 maisons (maximum enregistré dans le département du Nord) soit 2003 maisons pour un point; ce maximum est affiché sous le nom du dialogue **Semis de points**.

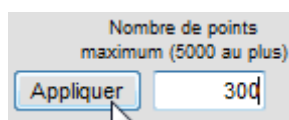


Figure 8.5 : Le calibrage de la valeur du point par le choix du nombre maximum de points correspondant à la valeur maximum de la variable cartographiée.

La seconde façon de calibrer la valeur du point consiste à choisir explicitement sa valeur, par exemple, un point représente 1000 maisons. Cette manière permet d'avoir une légende plus facilement assimilable par le lecteur de la carte. Elle permet aussi de produire une série de carte portant sur des catégories d'une même population comme, par exemple, le nombre de logement par année de construction ; dans ce cas seulement on peut rechercher des comparaisons entre cartes car le calibrage est identique pour toutes les cartes.

Pour fixer la valeur du point, il suffit d'entrer la valeur retenue dans la zone d'édition prévue à cet effet et de valider avec le bouton **Appliquer** (fig. 8.6).

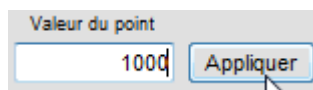


Figure 8.6 : Le calibrage de la valeur du point par le choix explicite de cette valeur.

Qu'on choisisse la première ou la seconde façon de calibrer le point, la validation avec le bouton **Appliquer** se traduit par l'affichage de la carte correspondant au nouveau calibrage (fig. 8.7).

Par défaut, un point correspond à un deux pixels d'écran. Une autre taille peut être choisie entre 1 et 5 pixels (fig. 8.8). Si un seul pixel par point donne souvent une image peu contrastée, en revanche, une taille supérieure à tendance à « graisser » le semis de points. Adopter une taille élevée permet cependant de limiter le nombre de points de la carte et d'éviter ainsi des fichiers trop lourds. La taille du point et le nombre de points doivent donc être choisis de concert ; plusieurs tentatives sont souvent nécessaires...

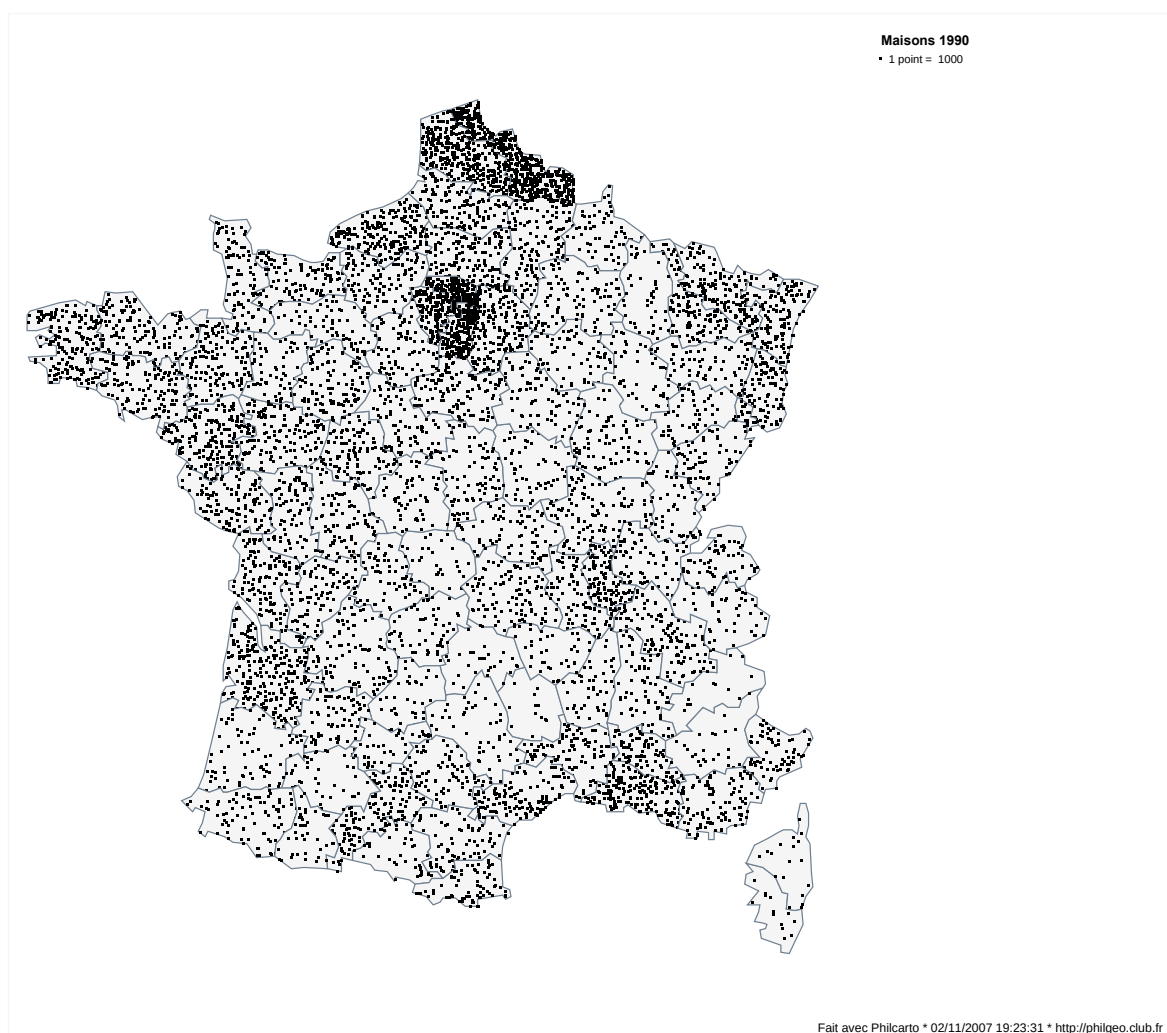


Figure 8.7 : Tracé de la carte en semis de points du nombre de maisons individuelles.
Sur cette carte, un point vaut 1000 maisons, nombre choisi explicitement.

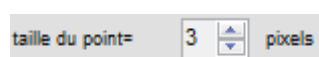


Figure 8.8 : Le choix de la taille du point.

8.5. Modifier la couleur des points et la couleur du fond

Par défaut, les points sont noirs sur fond gris clair. La modification de la couleur de tous les points et du fond de carte se fait à l'aide de deux séries de boutons radio colorés (fig. 8.9). La première série correspond à la couleur des points, et la seconde à la couleur du fond.

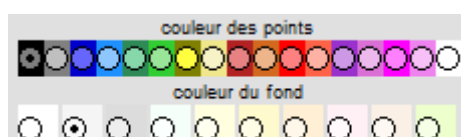


Figure 8.9 : Choisir la couleur des disques et des circonférences.

Pour la carte ci-après (fig. 8.10), on a choisi les paramètres suivants : un point pour 1000 maisons ; les points sont vert printemps, et le fond est jaune clair.

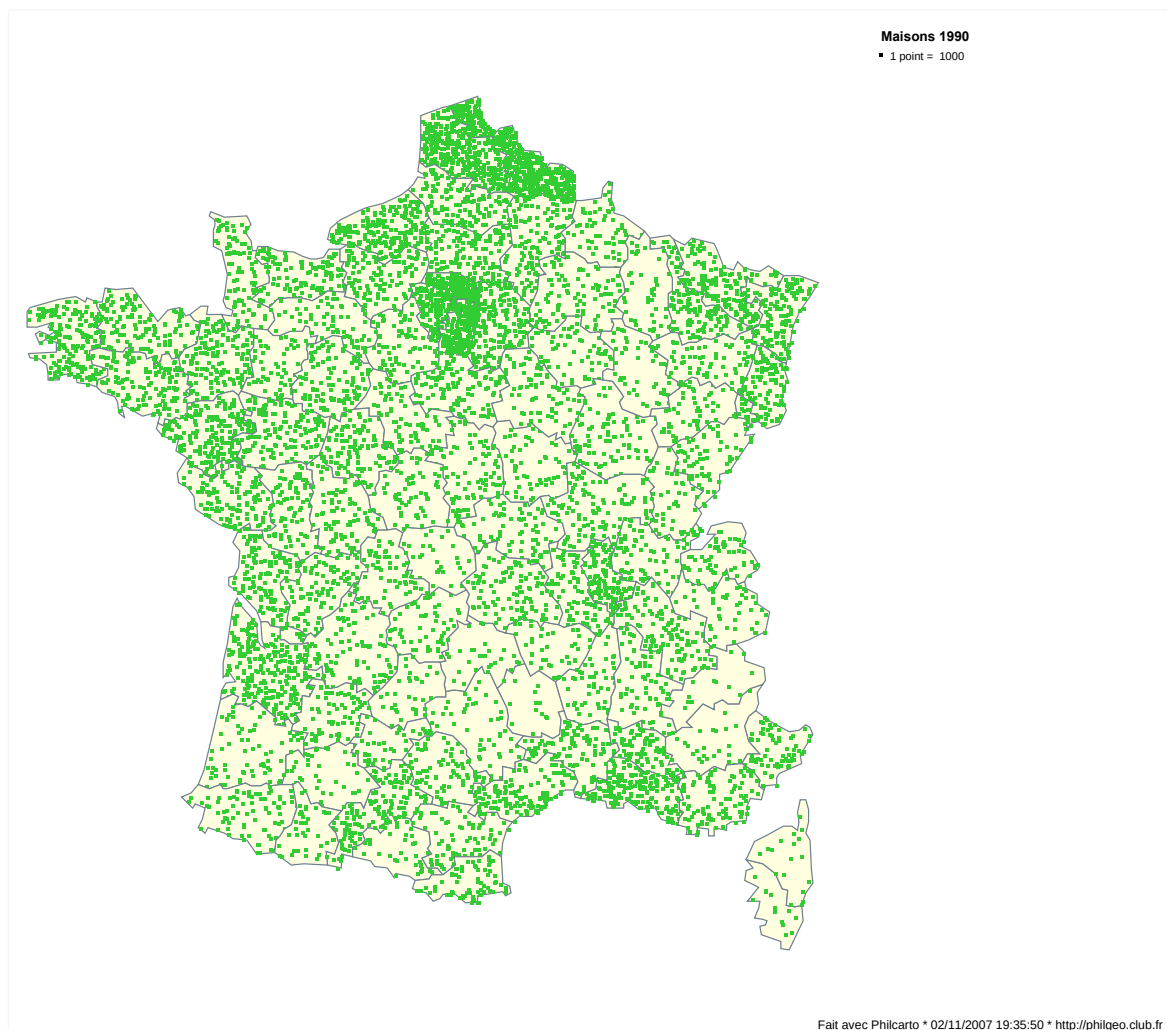


Figure 8.10 : Tracé de la carte en semis de points du nombre de maisons individuelles.
Le calibrage, la couleur des points et la couleur du fond ont été modifiés.

8.6. Points pour des effectifs négatifs et positifs

Comme pour les cartes en cercles, *Philcarto* détecte s'il existe des valeurs négatives et positives dans la variable à cartographier. Si cela est le cas, il figure les points correspondant à des effectifs positifs en rouge, et les points correspondant à des effectifs négatifs en bleu.

La carte du solde migratoire des départements français entre 1990 et 1999 montre l'intérêt de genre de tracé : l'héliotropisme est évident (fig. 8.11). Ici, les contours des départements n'ont pas été tracés en décochant la case ☐ Tracer les contours des unités spatiales située dans le cadre ☐ Habillage ; cela permet de mieux percevoir l'allure générale du semis de points. Un habillage régional au fond bleu clair a été ajouté pour pouvoir néanmoins localiser les concentrations de points.

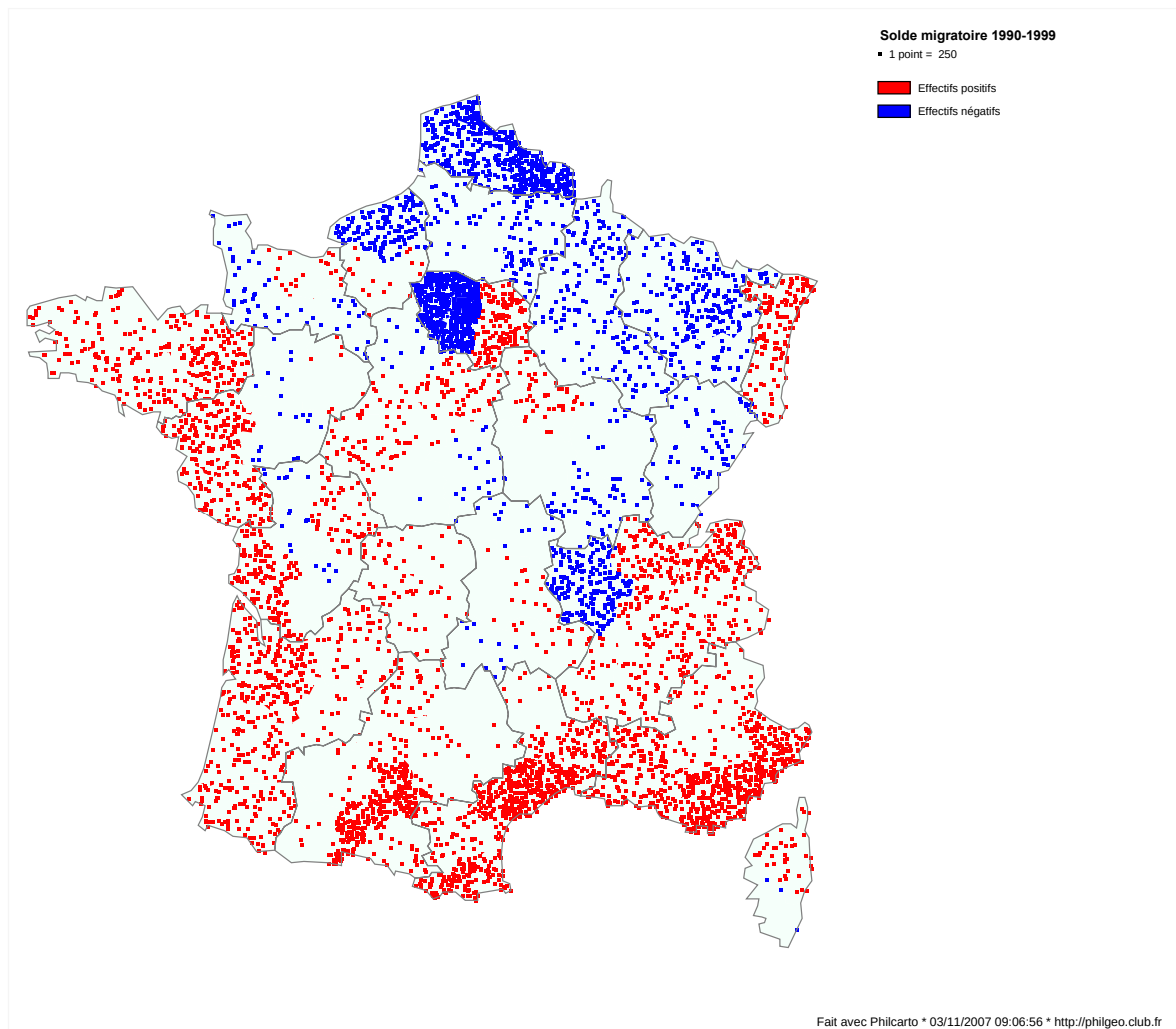


Figure 8.11 : La carte en semis de points du solde migratoire des départements français entre 1990 et 1999.

! Dans le cas de valeurs négatives et positives, on perd la possibilité de choisir la couleur des points qui sont obligatoirement bleus pour les effectifs négatifs ou rouges pour les effectifs positifs.

9. Les cartes en semis de points colorés

Dossier : \Philcarto exemples\France départements\
Fond de carte : France carte départements.ai
Statistiques : France data départements.txt

9.1. Sélectionner une variable de couleur et une variable de symbole

La sélection des deux variables se fait par un clic sur le nom d'une des variables figurant dans la liste **couleurs** et sur celui d'une seconde variable figurant dans la liste **symboles** du dialogue **Variables**. Ici les variables sélectionnées sont respectivement **Maisons 1990 %logements** et **Maisons 1990** (fig. 9.1).

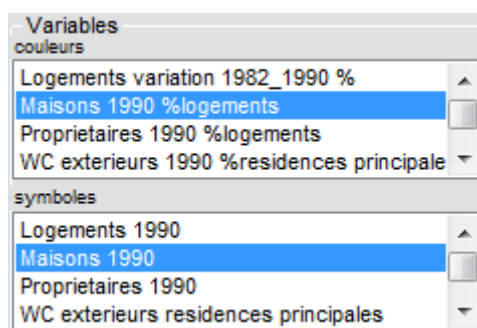


Figure 9.1 : Sélection d'une variable de couleur et de symbole.

9.2. Afficher une carte en semis de points colorés

La sélection de la variable de symbole provoque l'affichage du dialogue **Types de cartes**. Cliquer sur l'icône du type carte en cercles proportionnels colorés (fig. 9.2). La carte s'affiche alors à l'écran (fig. 9.3). Toutes les options en cours sont appliquées, aussi bien pour l'affectation des couleurs (discrétisation, nombre de classes, gamme de couleurs) que pour les points (calibrage). Dans le cas où la variable de couleurs est une variable nominale, les couleurs sont affectées en fonction des modalités triées et de l'ordre des couleurs dans la gamme de couleurs « perroquet » (la première modalité est représentée en rouge, la seconde en bleu, etc.).

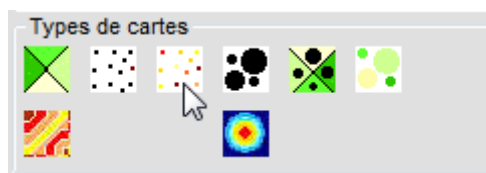


Figure 9.2 : Clic sur l'icône « carte en semis de points colorés ».

Un habillage, ici les régions, complète la carte et permet d'afficher les cercles sur un fond coloré qui fait ressortir les couleurs. Dans le cas où la variable d'effectifs comprend des valeurs négatives et positives, c'est la couleur du fond qui est

modifiée : rose clair pour les effectifs positifs et bleu clair pour les effectifs négatifs. Dans ce cas, on ne peut plus choisir la couleur du fond (fig. 9.4).

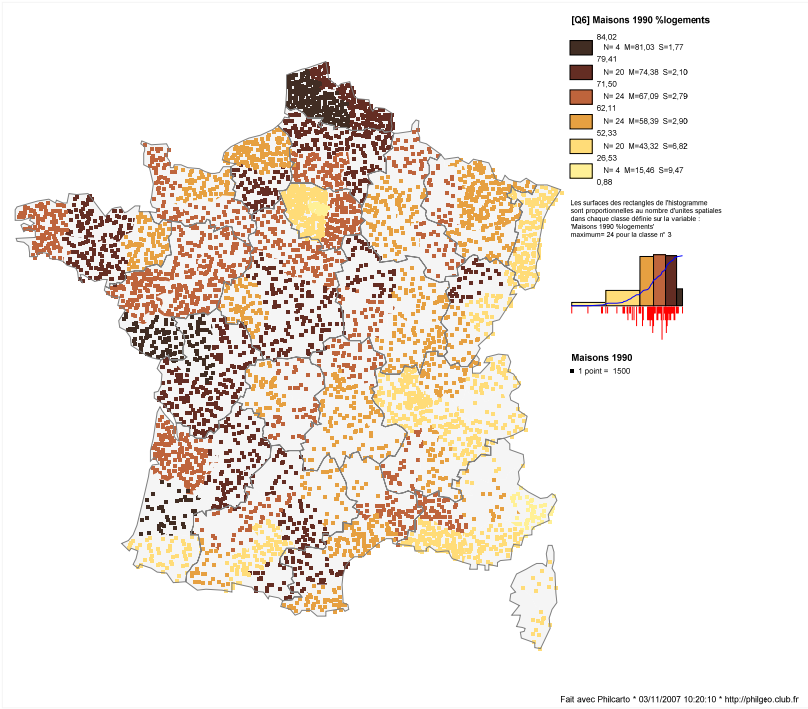


Figure 9.3 : La carte en semis de points colorés du nombre de maisons individuelles colorés par le % de maison individuelles dans l'ensemble des résidences principales.

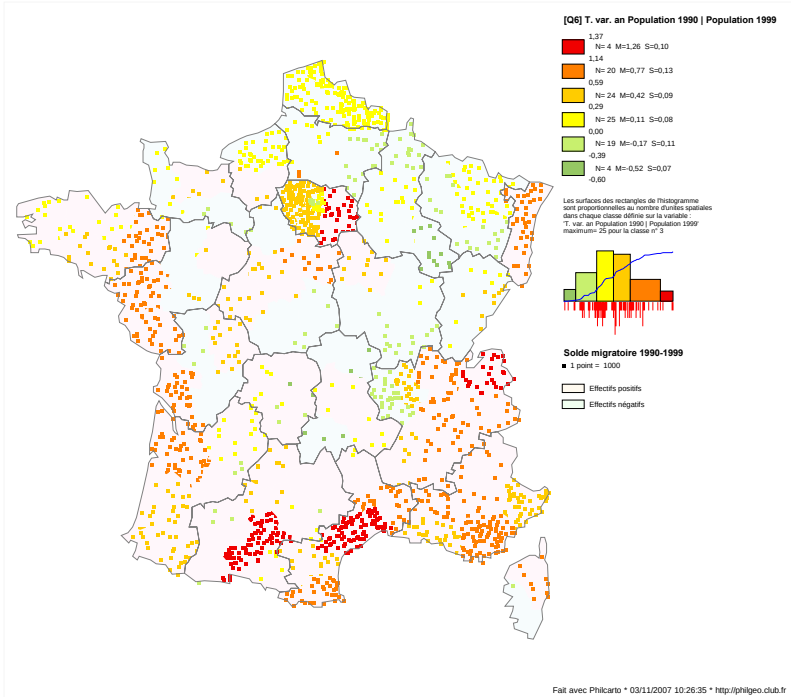


Figure 9.4 : La carte en semis de points colorés des soldes migratoires colorés par le taux moyen annuel de variation de la population. La couleur du fond de carte indique si les effectifs du solde migratoire sont négatifs (bleu clair) ou positifs (rose clair).

10. Les cartes de liens dites « en oursins »

Dossier : \Philcarto exemples\Parana\
Fond de carte : Parana carte.ai
Statistiques : Parana data liens.xls

10.1. Sélectionner une variable de lien

La sélection d'une variable de lien se fait simplement par un clic sur le nom d'une des variables figurant dans la liste **couleurs** du dialogue **Variables**. Ici la variable sélectionnée, **Lien**, définit une relation entre les lieux centraux de l'Etat du Paraná au Brésil (fig. 10.1). Cette variable doit obligatoirement s'appeler « Lien » ; elle est unique dans le fichier de données dont le format est lui-même spécifique aux données de liens ; voir le chapitre sur l'organisation des fichiers de données.

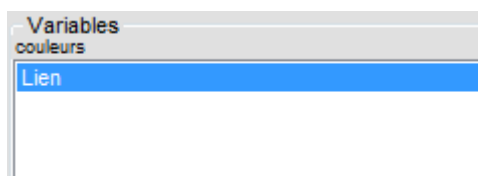


Figure 10.1 : Sélection d'une variable de symbole.

10.2. Afficher une carte de liens

La sélection de la variable de lien provoque l'affichage du dialogue **Types de cartes**. Cliquer sur l'icône du type carte de liens  (fig. 10.2). La carte s'affiche alors à l'écran (fig. 10.3). Toutes les options en cours sont appliquées comme le calibrage des points ou l'habillage.

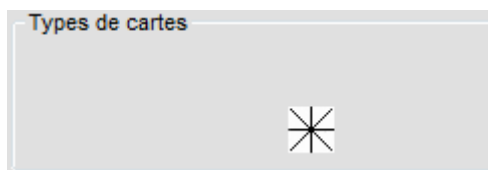


Figure 10.2 : Clic sur l'icône « carte de liens ».

 Par rapport aux versions précédentes de *Philcarto*, la variation d'épaisseur des liens n'est plus possible : cette option produisait une carte peu intelligible. Elle a été remplacée par les cartes de liens colorés selon les modalités d'une variable nominale.

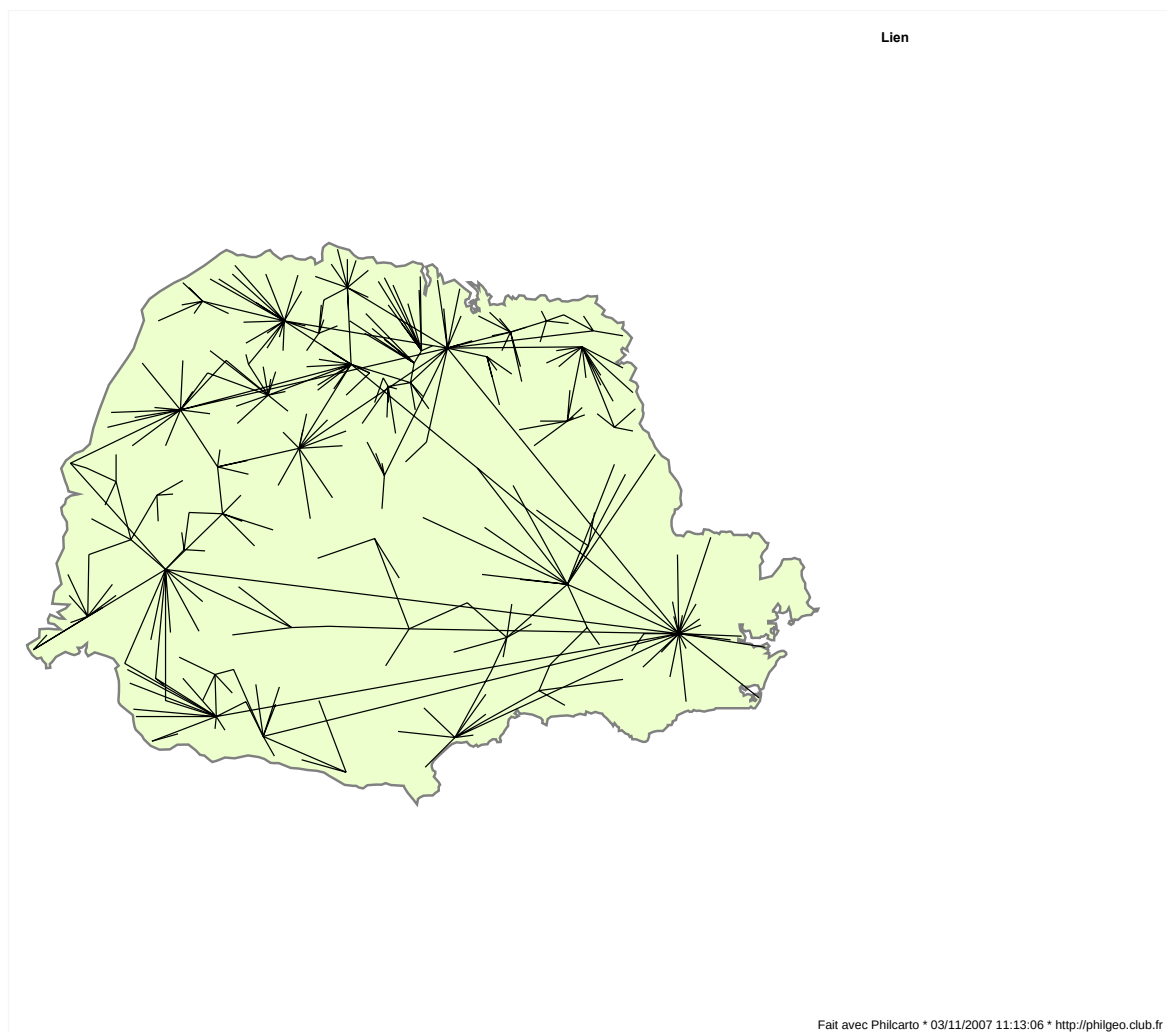


Figure 10.3 : La carte de liens entre lieux centraux de l'Etat du Paraná au Brésil.

11. Les cartes de liens colorés

Dossier : \Philcarto exemples\Parana\
Fond de carte : Parana carte.ai
Statistiques : Parana data liens.xls

11.1. Sélectionner une variable de liens colorés

La sélection d'une variable de liens colorés se fait simplement par un clic sur le nom d'une des variables figurant dans la liste **couleurs** du dialogue **Variables**. Ici la variable sélectionnée, **Lien modal**, définit une relation entre les lieux centraux de l'Etat du Paraná au Brésil (fig. 11.1). Cette variable doit obligatoirement s'appeler « Lien modal » ; elle est unique dans le fichier de données dont le format est lui-même spécifique aux données de liens (voir le chapitre 19).

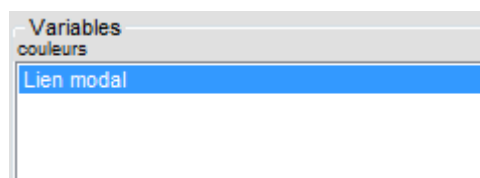


Figure 11.1 : Sélection d'une variable de symbole.

11.2. Afficher une carte de liens

La sélection de la variable de lien provoque l'affichage du dialogue **Types de cartes**. Cliquer sur l'icône du type carte de liens colorés  (fig. 11.2). La carte s'affiche alors à l'écran (fig. 11.3). Toutes les options en cours sont appliquées comme le calibrage des points ou l'habillage.

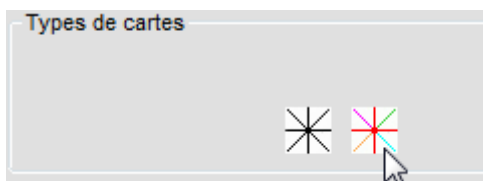


Figure 11.2 : Clic sur l'icône « carte de liens colorés ».

 Par rapport aux versions précédentes de *Philcarto*, la représentation de variables discrétisées qui donnait des résultats peu intelligibles n'est plus possible. Les valeurs doivent être nominales et ne pas excéder 30 modalités.

Les couleurs des liens peuvent être modifiées. Il suffit pour cela de se rendre dans le dialogue **Couleurs** qui s'est affiché dans le cadre **Outils** juste après le tracé de la carte. On retrouve dans le cadre **Couleurs** la liste des modalités de variable cartographiée (fig. 11.4).

En regard de chaque modalité figure un bouton radio coloré. Un clic sur l'un de ces boutons radio permet de sélectionner une autre couleur dans la gamme proposée. Par exemple, si l'on veut que les liens concernant les centres de plus bas

niveau soient tracés en Blanc cassé, il suffit de choisir cette couleur-là à la place du rouge affecté automatiquement par *Philcarto* (fig. 11.5 ; fig. 11.6). En procédant de cette manière avec toutes les couleurs attribuées automatiquement, on obtient une carte qui traduit mieux l'idée de niveau hiérarchique (fig. 11.7).

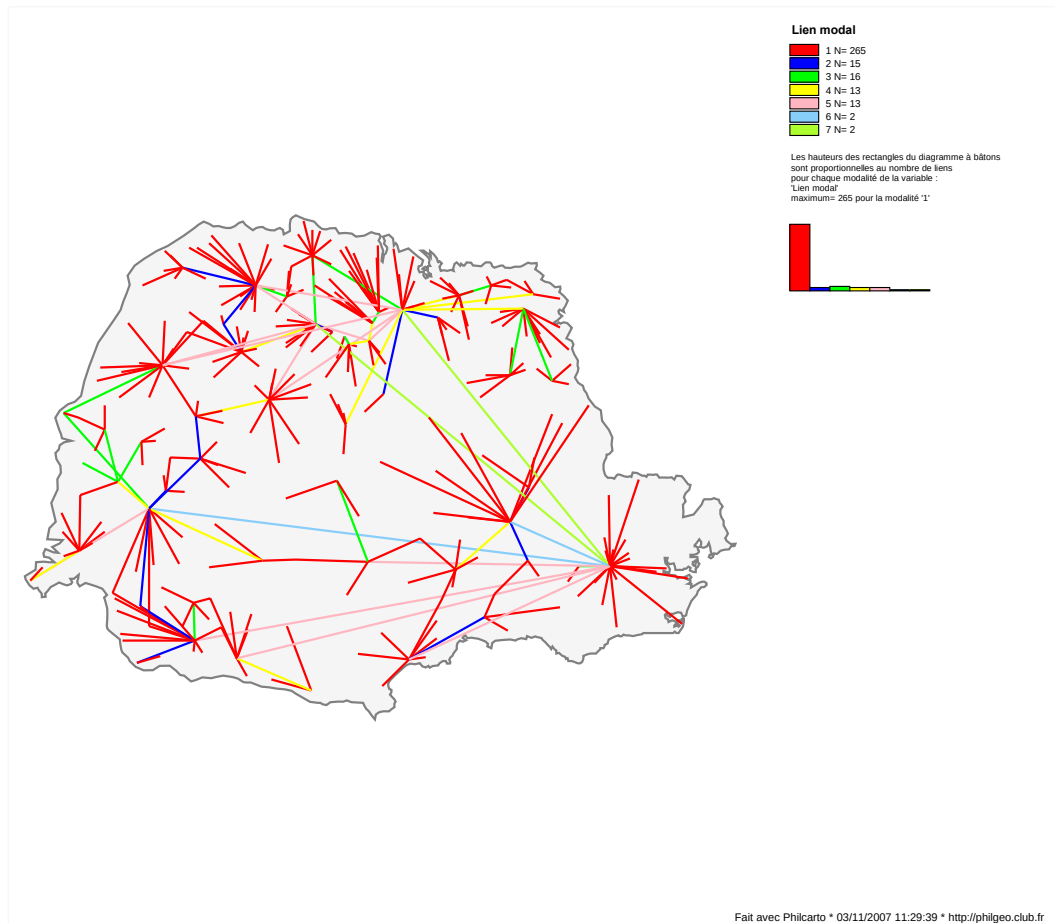


Figure 11.3 : La carte de liens entre lieux centraux de l'Etat du Paraná au Brésil.

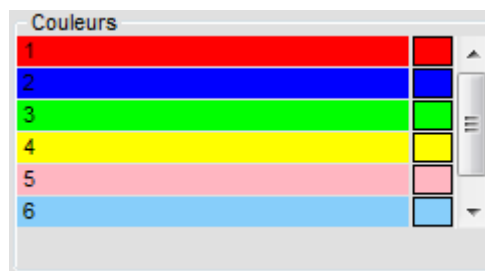


Figure 11.4 : Les couleurs des niveaux de la hiérarchie urbaine.
Le niveau 1 est le plus bas, le niveau 6 est le niveau le plus élevé dans la hiérarchie.

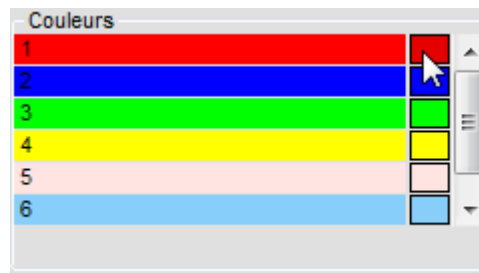


Figure 11.5 : Clic sur le bouton radio rouge pour modifier sa couleur.

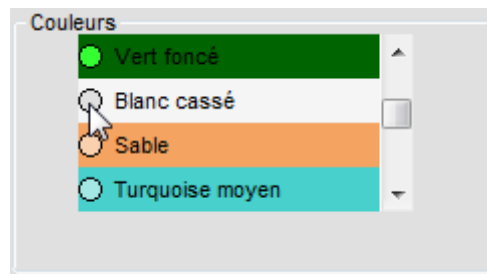


Figure 11.6 : Choix de la couleur « Blanc cassé » pour remplacer le rouge.

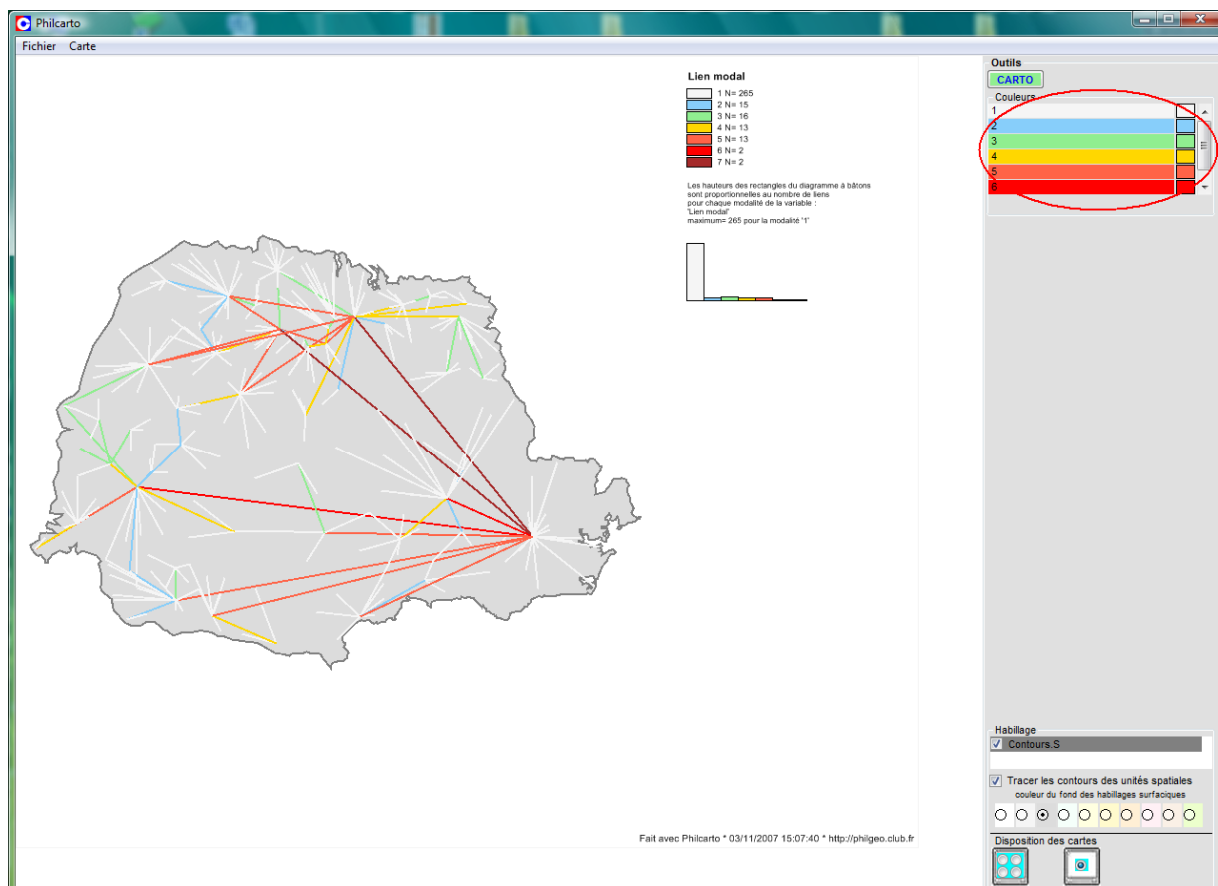


Figure 11.7 : Toutes les couleurs des liens ont été redéfinies pour mieux rendre compte des niveaux hiérarchiques.

12. Les cartes en lignes proportionnelles

Dossier : \Philcarto exemples\IDF SIRIUS\
Fond de carte : IDF carte SIRIUS.ai
Statistiques : IDF data SIRIUS.xls

12.1. Sélectionner une variable de symbole

La sélection d'une variable se fait simplement par un clic sur le nom d'une des variables figurant dans la liste **symboles** du dialogue **Variables**. Ici la variable sélectionnée, **HKM Total**, contient le nombre d'HKM (Heures x Kilomètres d'encombrements), c'est-à-dire la somme des produits des longueurs des bouchons (ramenées à une file) par la durée de ces bouchons, et ceci pour chaque axe majeur de la Région Île de France du système SIRIUS (fig. 12.1).

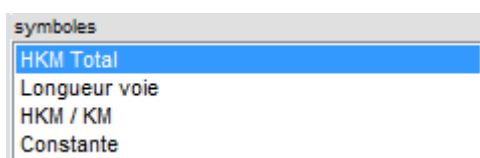


Figure 12.1 : Sélection d'une variable de symbole.

12.2. Afficher une carte en lignes proportionnelles

La sélection de la variable de symbole provoque l'affichage du dialogue **Types de cartes**. Cliquer sur l'icône du type carte en lignes proportionnelles  (fig. 12.2). La carte s'affiche alors à l'écran (fig. 12.3). Si le calibrage des lignes n'a pas été changé en cours de session, l'épaisseur de la ligne la plus épaisse, celle qui correspond à la valeur maximum, est de 10 pixels de l'écran d'affichage.



Figure 12.2 : Clic sur l'icône « carte en cercles proportionnels ».

12.3. Le dialogue Lignes

Simultanément au tracé de la carte, le dialogue **Lignes** remplace les dialogues **Variables** et **Types de cartes** dans le cadre **Outils**. (fig. 12.4). Dans le dialogue **Lignes**, des boutons permettent de calibrer les lignes, de changer leur couleur, de modifier la légende ...

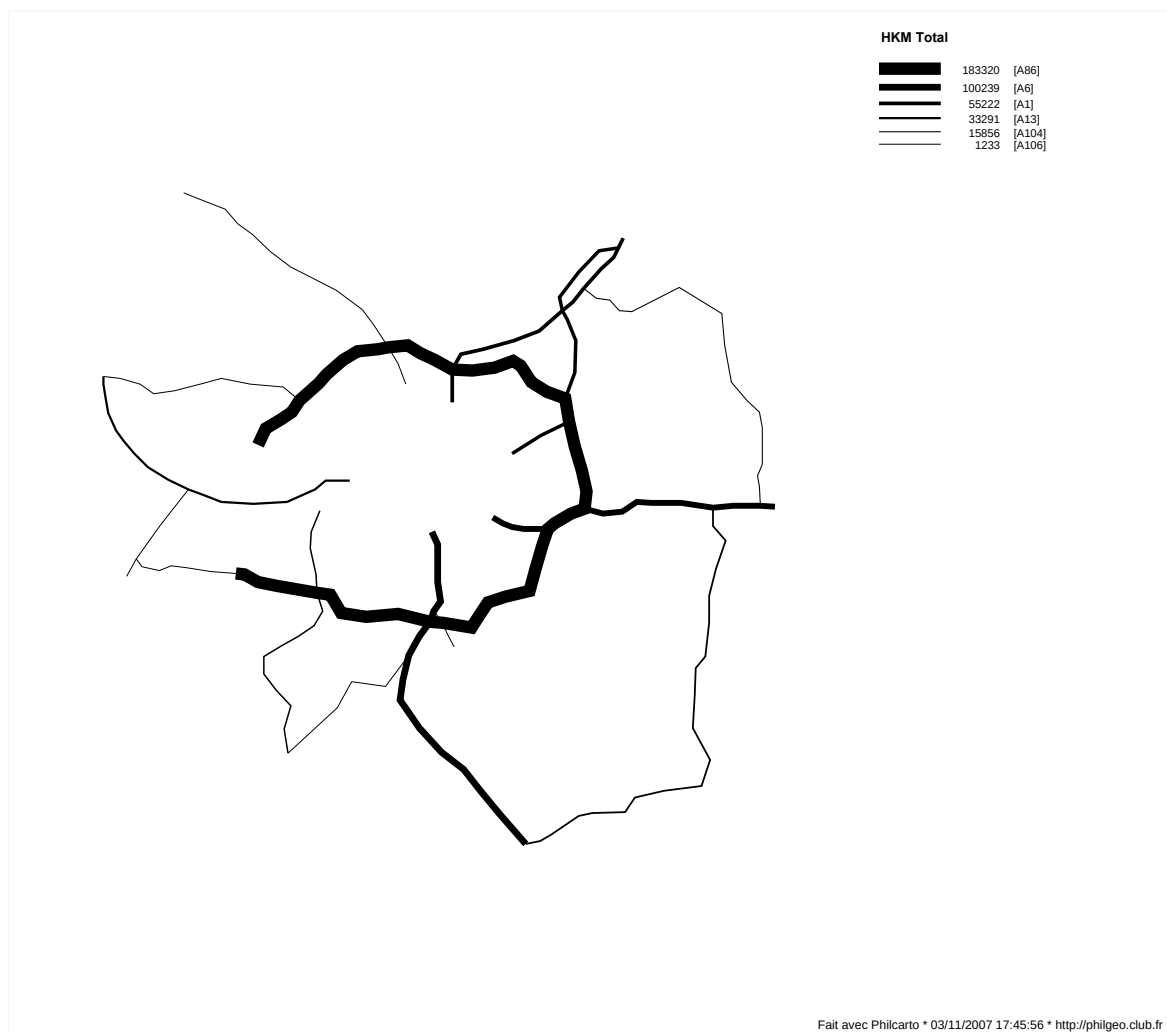


Figure 4.3 : Tracé de la carte en lignes proportionnelles.

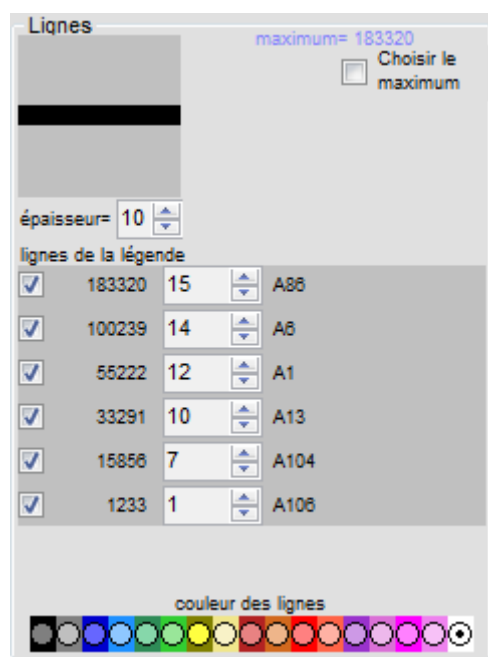


Figure 12.4 : Affichage du dialogue **Lignes**.

12.4. Le calibrage des lignes

Calibrer les lignes consiste à fixer le nombre de pixels d'écran de la ligne la plus épaisse, celle qui correspond à la valeur maximum. En début de session, cette valeur est de 10 pixels ; elle peut varier entre 1 et 40 pixels. Le choix de cette valeur est affaire de perception : la carte ne doit être ni trop claire, ni trop foncée. Pour changer cette valeur, il suffit d'utiliser le curseur `épaisseur= 15` . Un clic sur la flèche dirigée vers le haut fait augmenter le nombre de pixels de l'épaisseur alors qu'un clic sur la flèche dirigée vers le bas le fait diminuer. Avec 20 pixels la carte apparaît déjà trop chargée (fig. 12.5) ; il est préférable de revenir à la valeur par défaut, 10 pixels.

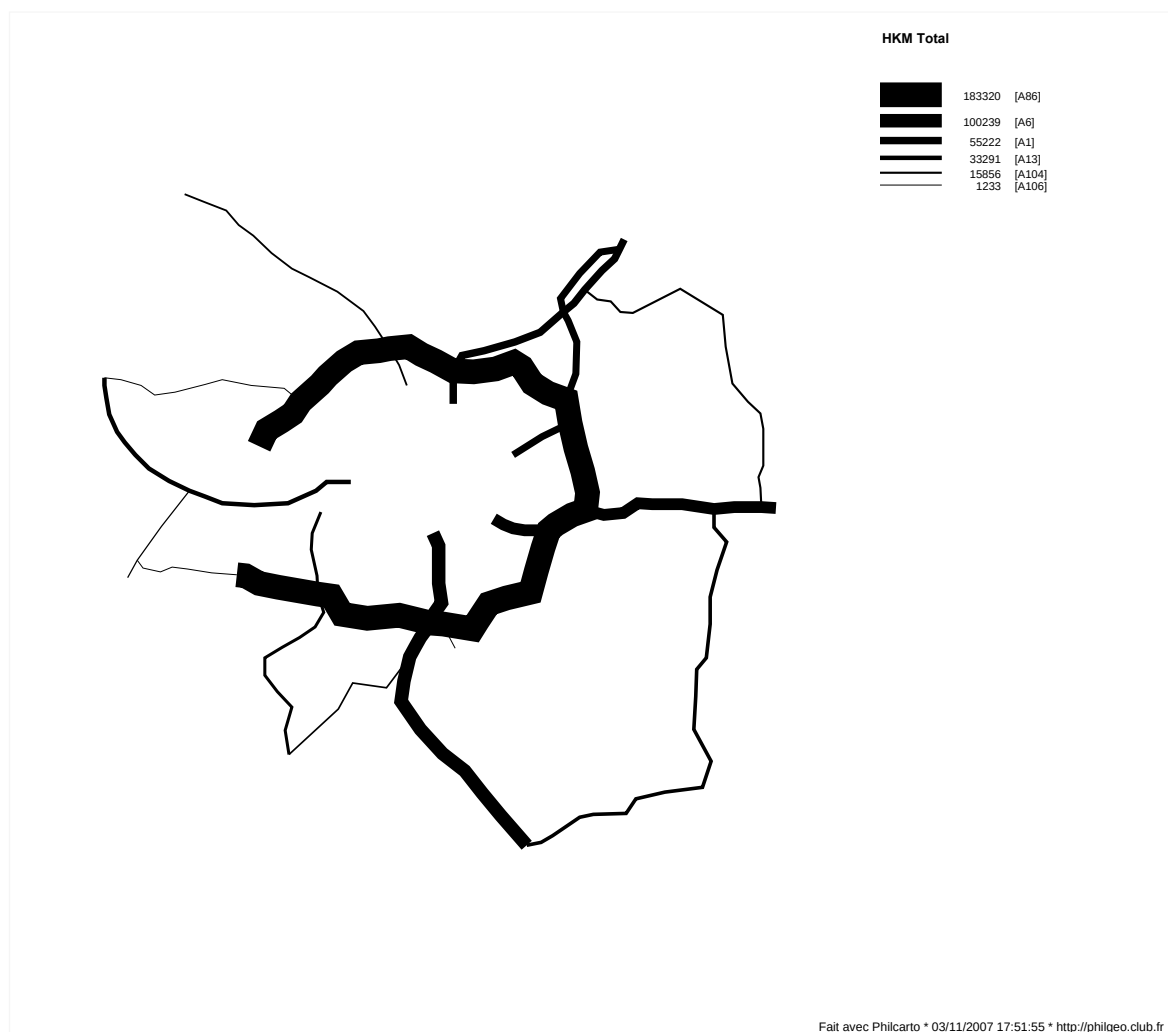


Figure 12.5 : Tracé de la carte en lignes proportionnelles avec l'épaisseur maximum calibrée à 20.

12.5. Modifier l'apparence des lignes

Par défaut, les lignes sont noires. La modification de la couleur de toutes les lignes se fait à l'aide d'une série de boutons radio colorés (fig. 12.6).



Figure 12.6 : Choisir la couleur des lignes.

Pour la carte ci-après (fig. 12.7), on a choisi des lignes rouges.

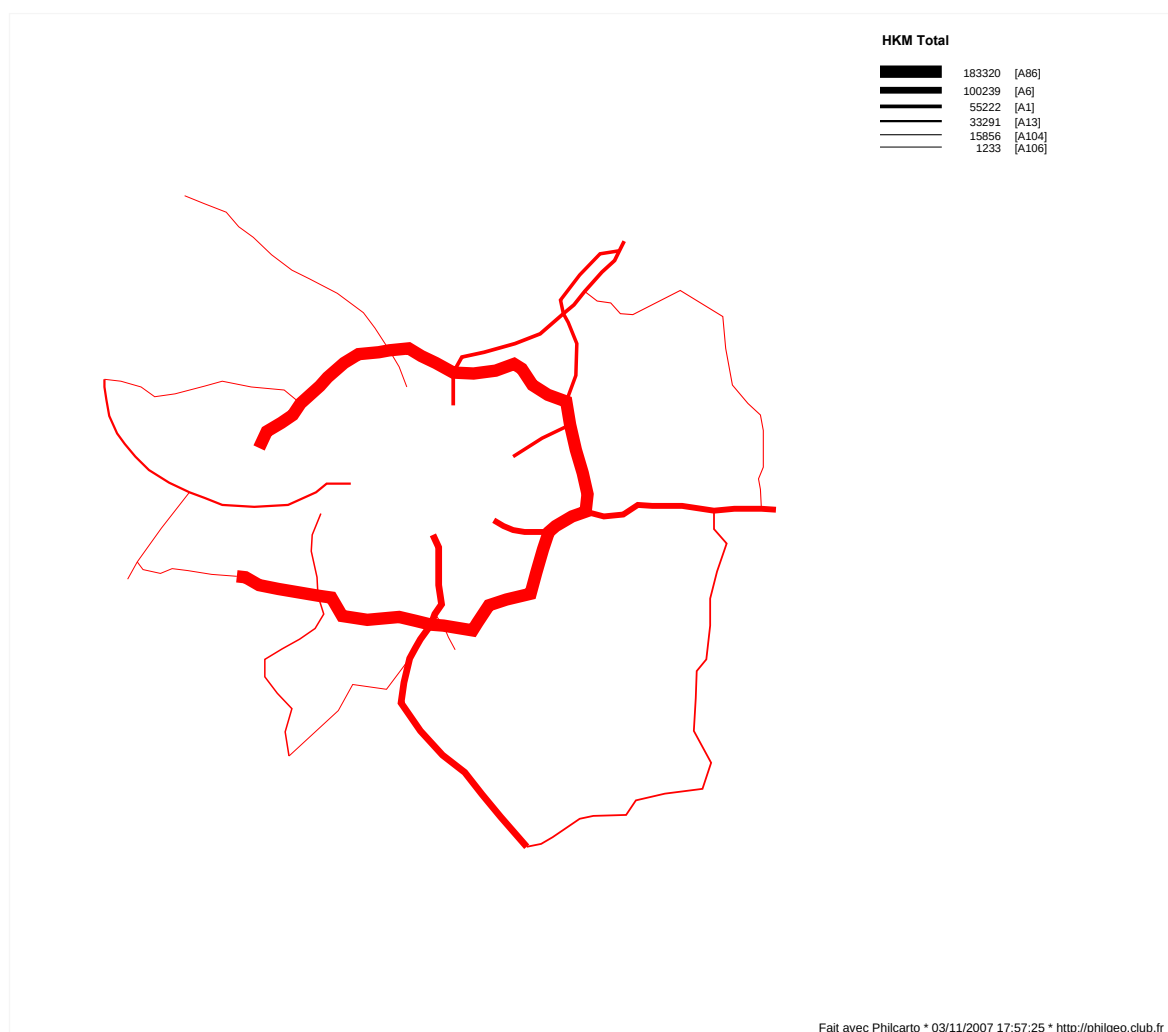


Figure 12.7 : Modification de la couleur des lignes.

12.6. Modifier la couleur de fond de la page

Par défaut, la carte s'affiche sur un fond blanc : il s'agit de la couleur de la fenêtre d'affichage. Pour modifier la couleur du fond, il faut choisir un habillage surfacique listé dans le dialogue **Habillage** situé dans la partie inférieure du cadre **Outils**, par exemple, le calque **Cadre.S**. Un clic dans la case correspondant à ce calque le sélectionne et fait apparaître une série de boutons radio colorés intitulée **couleur du contour de l'habillage sélectionné** (fig. 12.8). Il faut impérativement choisir l'une de ces couleurs pour que l'habillage puisse être tracé dans cette couleur-là. On peut choisir une couleur passe-partout comme le noir ou le gris, ou une couleur plus expressive à la condition qu'elle soit en harmonie avec la couleur des lignes.



Figure 12.8 : La liste des couleurs du contour d'habillage.

Lorsque le bouton radio de la couleur choisie est sélectionné (ici le gris), la série de boutons radio colorés s'efface, le nom du calque est écrit avec un fond coloré par la couleur choisie, et la carte est à nouveau tracée, faisant apparaître l'habillage retenu. Une nouvelle série de boutons radio colorés s'affiche alors ; intitulée **couleur du fond des habillages surfaciques**, elle permet de choisir la couleur du fond de l'habillage par un clic sur le bouton radio correspondant (fig. 12.9). Notons que ces couleurs sont pâles car le fond ne doit pas parasiter la lecture des lignes (fig. 12.10).

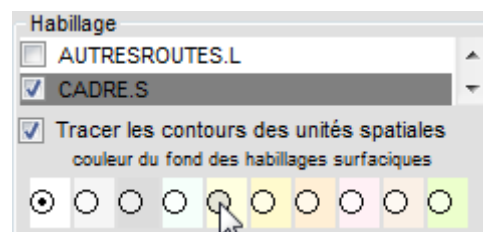


Figure 12.9 : La liste des couleurs du contour d'habillage.

Notons qu'il est possible de tracer plusieurs habillages, comme par exemple les rivières ou les autres routes, en leur affectant une couleur différente. Dans ce but, il faut sélectionner successivement les calques correspondants, puis définir à nouveau la couleur du fond. On obtient ainsi une carte plus parlante (fig. 12.11).

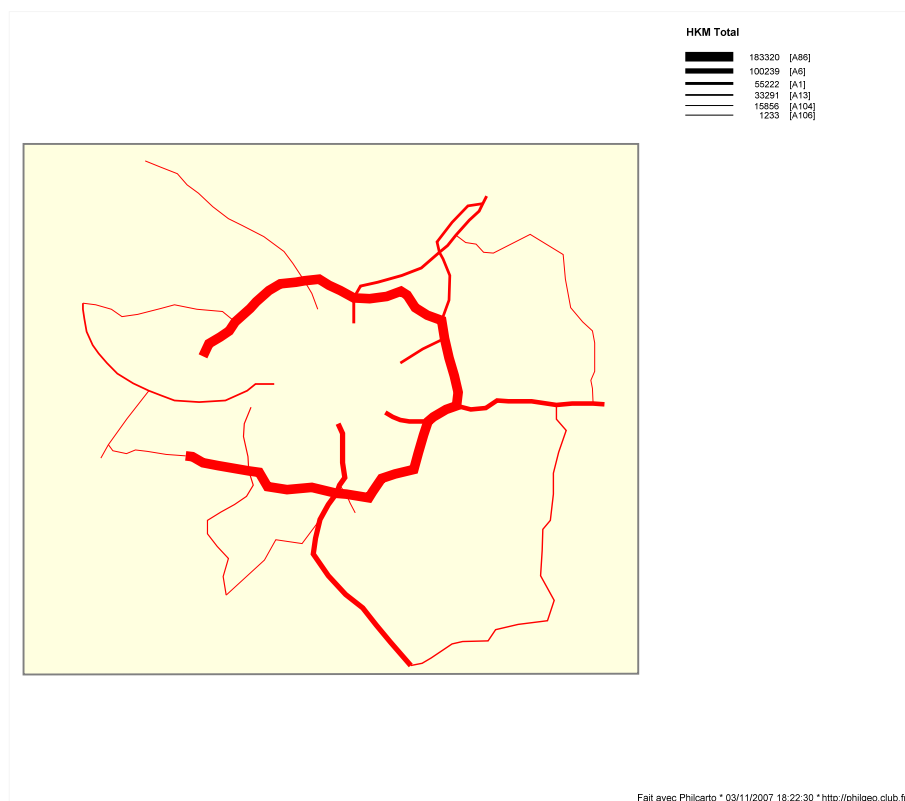


Figure 12.10 : La modification de la couleur du fond de l'habillage.

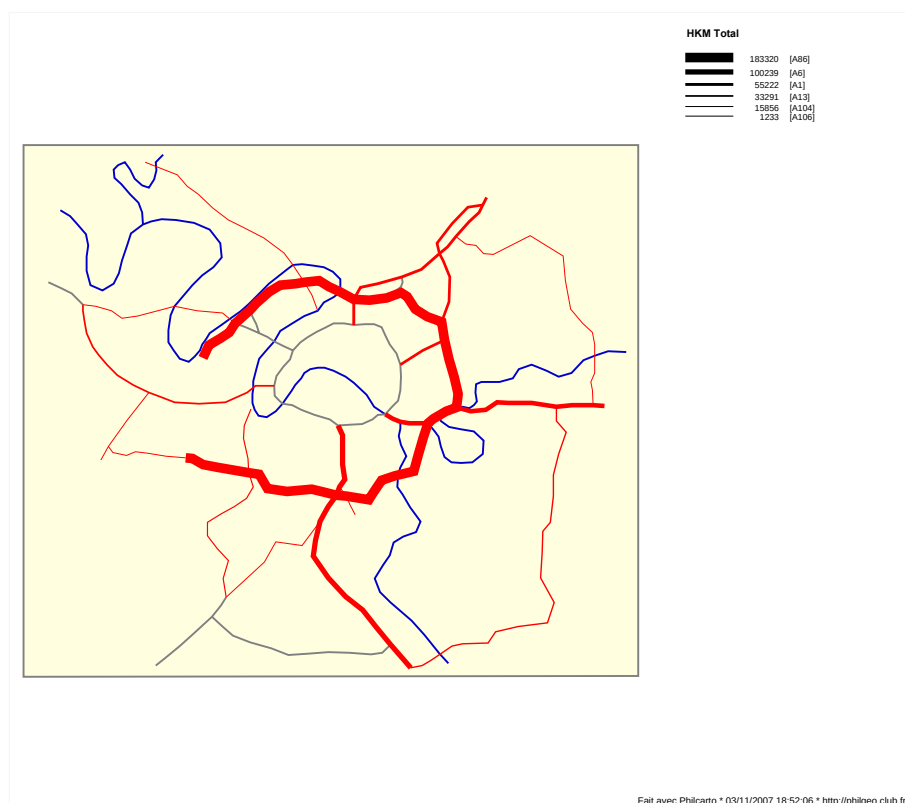


Figure 12.11 : La même carte que sur la figure 12.10 mais avec deux éléments d'habillage supplémentaires : les autres routes principales et les fleuves.

12.7. Choisir les lignes de la légende

La solution retenue consiste à faire figurer la ligne la plus épaisse, la ligne la moins épaisse, et entre elles, quatre autres lignes qui forment un échantillonnage de lignes effectivement présentes sur la carte ; en regard de chaque ligne, on fait figurer, la valeur à laquelle elle correspond, ainsi que l'identifiant et le nom de l'élément cartographique correspondant (fig. 12.12). Cette présentation a l'avantage de coller à la réalité des données.

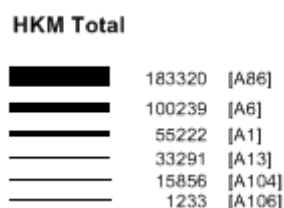


Figure 12.12 : Les lignes de la légende selon l'échantillonnage standard des valeurs.

L'échantillonnage standard des épaisseurs de lignes ne donne pas toujours le meilleur résultat car les distributions de valeurs peuvent s'avérer inadéquates comme c'est le cas sur la figure ci-dessus où trois lignes sont à peu près de même épaisseur. Pour pallier cet inconvénient, *Philcarto* permet de choisir le nombre de lignes et la valeur correspondant à chacun d'eux. Ceci se fait simplement à l'aide d'un ensemble de contrôles rassemblés sous le nom **lignes de la légende** (fig. 12.13).

Checkbox	Value	Spinner	Text
☒	183320	15	A86
☒	100239	14	A6
☒	55222	12	A1
☒	33291	10	A13
☒	15856	7	A104
☒	1233	1	A106

Figure 12.13 : Les contrôles permettant de choisir les cercles de la légende.

Ces contrôles sont agencés sous la forme d'un tableau comprenant six lignes correspondant chacune à une ligne de légende ; la légende peut donc se composer de six lignes au plus ; la ligne supérieure correspond à la ligne la plus épaisse et la ligne inférieure à la ligne la moins épaisse. Chacune des colonnes de ce tableau a une fonction précise. De gauche à droite, on trouve :

- des cases à cocher qui, si elles sont cochées, commandent chacune l'affichage d'une ligne;
- des valeurs numériques donnant, pour chaque ligne, la valeur statistique qui la caractérise ;
- des curseurs dont les valeurs correspondent, pour chaque ligne, à son rang dans la série des lignes triées de la plus fine (rang = 1) à la plus épaisse (rang = nombre total de valeurs) ;
- des textes donnant, pour chaque ligne, l'identifiant et le nom de l'élément cartographique correspondant.

Grâce à ce tableau de contrôle, il est possible d'élaborer une légende explicite composée de 1 à 6 cercles relatifs à des cercles effectivement présents sur la carte (fig. 12.14).



Figure 12.14 : Une légende en 4 lignes et le tableau de contrôles correspondant.

12.8. Tracer plusieurs cartes en lignes comparables entre elles

Par défaut, le calibrage des lignes est différent d'une carte à l'autre. De ce fait, les cartes qui représentent des sous-populations extraites d'une même population (par exemple des trafics routiers décomposés par catégories de véhicules du genre poids lourds, voitures, cycles...) ne sont pas comparables entre elles, ce qui peut induire de sérieux problèmes d'interprétation.

Pour résoudre ce problème, *Philcarto* offre la possibilité d'entrer une valeur maximum différente de la valeur maximum de chaque sous-population ; lorsqu'elle existe, cette valeur maximum se substitue à celle de chacune de variable. Ainsi, le calibrage des cercles est identique pour chaque carte et les cartes peuvent être comparées. Pour faire fonctionner ce dispositif, il suffit de cocher la case ☒ Choisir le maximum, d'entrer une valeur maximum dans zone d'édition et de valider par un clic sur (fig. 12.15).

Figure 12.15 : Choisir une valeur maximum pour rendre les cartes comparables.

13. Les cartes en lignes proportionnelles colorées

Dossier : \Philcarto exemples\IDF SIRIUS\
Fond de carte : IDF carte SIRIUS.ai
Statistiques : IDF data SIRIUS.xls

13.1. Sélectionner une variable de couleur et une variable de symbole

La sélection des deux variables se fait par un clic sur le nom d'une des variables figurant dans la liste **couleurs** et sur celui d'une seconde variable figurant dans la liste **symboles** du dialogue **Variables**. Ici les variables sélectionnées sont respectivement **Sévérité encombrement** et **HKM Total** (fig. 13.1). La sévérité des encombrements est un indice qui varie de 0 à 2, 2 étant la plus grande sévérité.

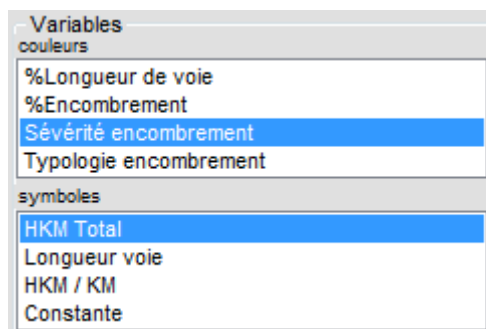


Figure 13.1 : Sélection d'une variable de couleur et de symbole.

13.2. Afficher une carte en lignes proportionnelles colorées

La sélection de la variable de symbole provoque l'affichage du dialogue **Types de cartes**. Cliquer sur l'icône du type carte en lignes proportionnelles colorées  (fig. 13.2). La carte s'affiche alors à l'écran (fig. 13.3). Toutes les options en cours sont appliquées, aussi bien pour l'affectation des couleurs (discrétisation, nombre de classes, gamme de couleurs) que pour l'épaisseur des lignes (calibrage). Dans le cas où la variable de couleurs est une variable nominale, les couleurs sont affectées en fonction des modalités triées et de l'ordre des couleurs dans la gamme de couleurs « perroquet » : la première modalité est représentée en rouge, la seconde en bleu, etc. (fig. 13.4).



Figure 13.2 : Clic sur l'icône « carte en semis de points colorés ».

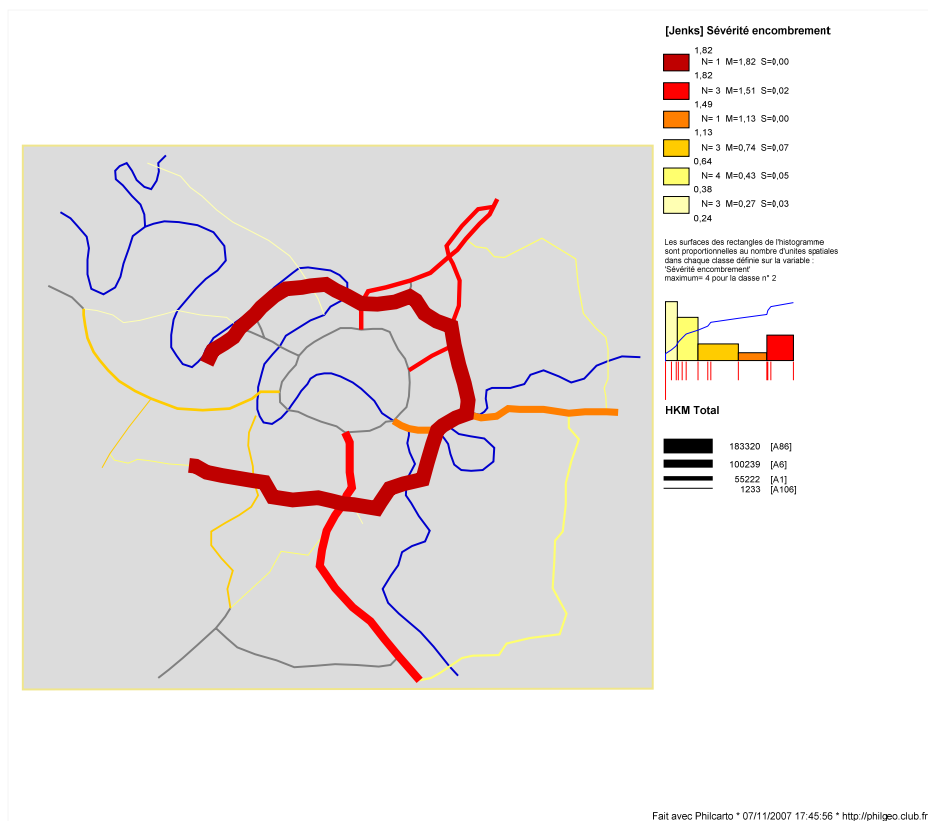


Figure 13.3 : La carte de l'indice de sévérité des encombrements (discrétisation de Jenks).

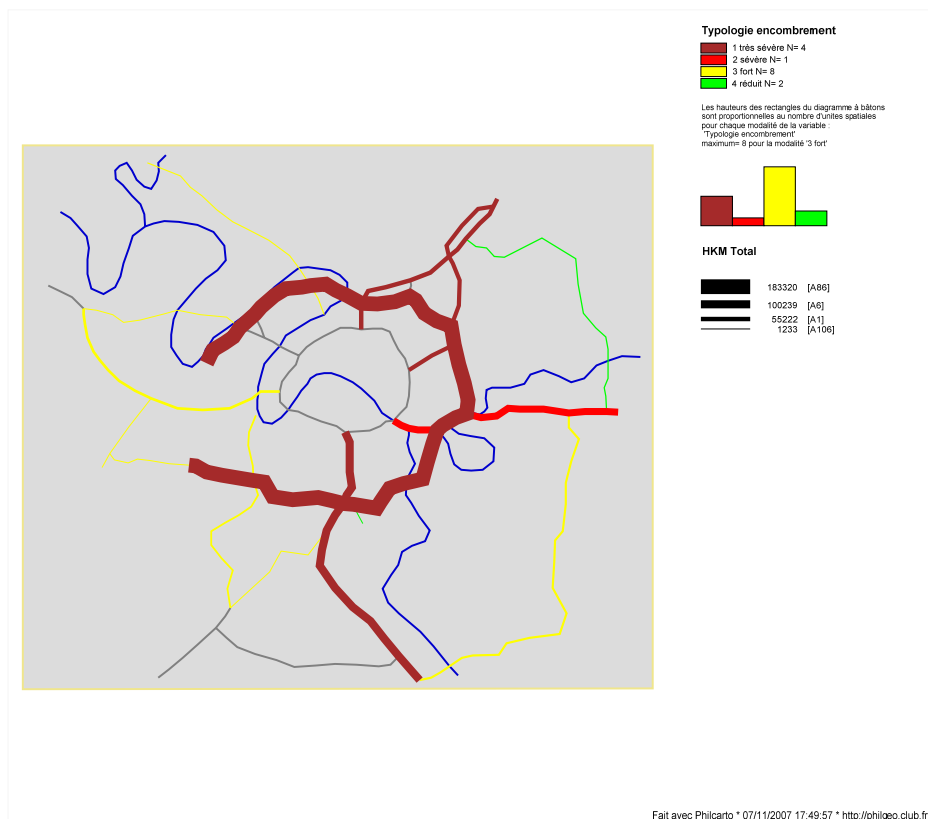


Figure 13.4 : La carte de la typologie des encombrements (4 modalités nominales).

14. Les cartes lissées

Dossier : \Philcarto exemples\France départements\
Fond de carte : France carte départements.ai
Statistiques : France data départements.txt

14.1. Sélectionner une variable de couleur

La sélection d'une variable se fait simplement par un clic sur le nom d'une des variables figurant dans la liste **couleurs** du dialogue **Variables**. Ici la variable sélectionnée, **Variation migratoire %1990-1999 %**, contient le solde migratoire 1990-1999 rapporté à la population en 1990 (%) (fig. 14.1).

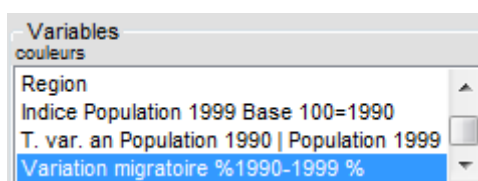


Figure 14.1 : Sélection d'une variable de couleur.

14.2. Afficher une carte choroplèthe non lissée

La sélection de la variable de couleur provoque l'affichage du dialogue **Types de cartes**. Cliquer sur l'icône du type carte choroplèthe . La carte s'affiche alors à l'écran. Choisir la discrétisation de **Jenks**. A l'écran figure maintenant la carte choroplèthe de la variation de la population des départements français due au solde migratoire (fig. 14.2).

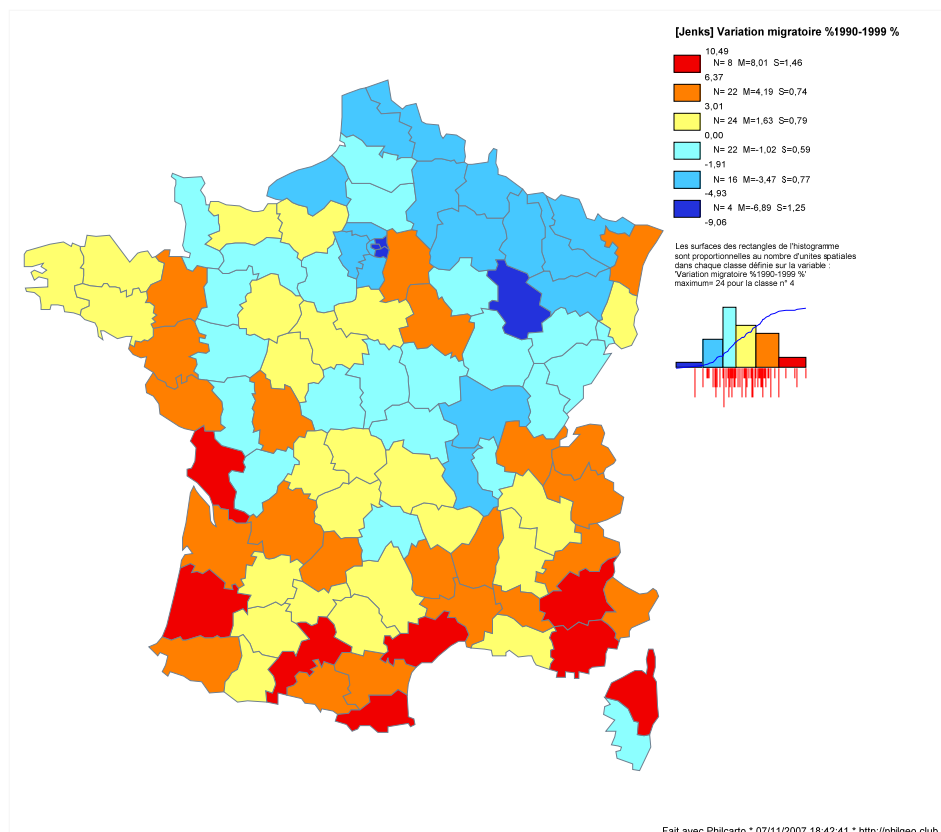


Figure 14.2 : La carte choroplèthe de la variation de la population des départements français 1990-1999 due au solde migratoire.

Cette carte suggère l'existence d'un tropisme Nord-Sud auquel s'ajoute parfois un effet urbain (Toulouse, Montpellier...) qui vient renforcer l'attrait du midi. On peut aussi lire la carte comme traduisant un phénomène de croissance périphérique par la migration, sauf pour le Nord et le Nord-Est qui, décidément, repoussent plus qu'ils n'attirent... Ces tendances de grande ampleur géographiques demeurent cependant difficiles à lire car elles sont en quelque sorte parasitées par le maillage départemental. D'où l'idée de recourir à des techniques de lissages pour traduire de manière plus radicale (et donc simplificatrice) ces observations générales.

14.3. Afficher une carte lissée par triangulation

Un clic sur **CARTO** permet d'afficher à nouveau les listes de variables. Dans la liste **couleurs**, sélectionnons à nouveau la variable **Variation migratoire %1990-1999 %**. La sélection de cette variable provoque l'affichage du dialogue **Types de cartes**. Cliquer sur l'icône du type carte lissée (fig. 14.3).



Figure 14.3 : Clic sur l'icône « carte lissée ».

Si le type de lissage n'a pas été modifié dans le cours de la session, la carte qui s'affiche à l'écran est une carte dite « isoplèthe » obtenue par la méthode de la triangulation de Delaunay (fig. 14.4) car dans le dialogue **Lissage**, le bouton radio **triangulation** est sélectionné.

En même temps que la carte, le dialogue **Lissage** apparaît sur la droite de l'écran (fig. 14.5).

Le calcul des isoplèthes est effectué dans le rectangle encadrant le fond de carte. On peut redécouper ce rectangle de manière à visualiser seulement la partie qui contient la carte de France. Cela se fait en choisissant un habillage, ici **Regions.S**, et en décochant la case **Tracer les contours des unités spatiales** pour que l'habillage soit tracé en trait fin (fig. 14.6).

La carte obtenue apparaît d'emblée plus familière car dans les **Options de tracé**, la case **limiter le tracé à l'habillage** est cochée, ce qui a bien pour effet d'éliminer tous les éléments situés en dehors de la surface couverte par l'habillage (fig. 14.7).

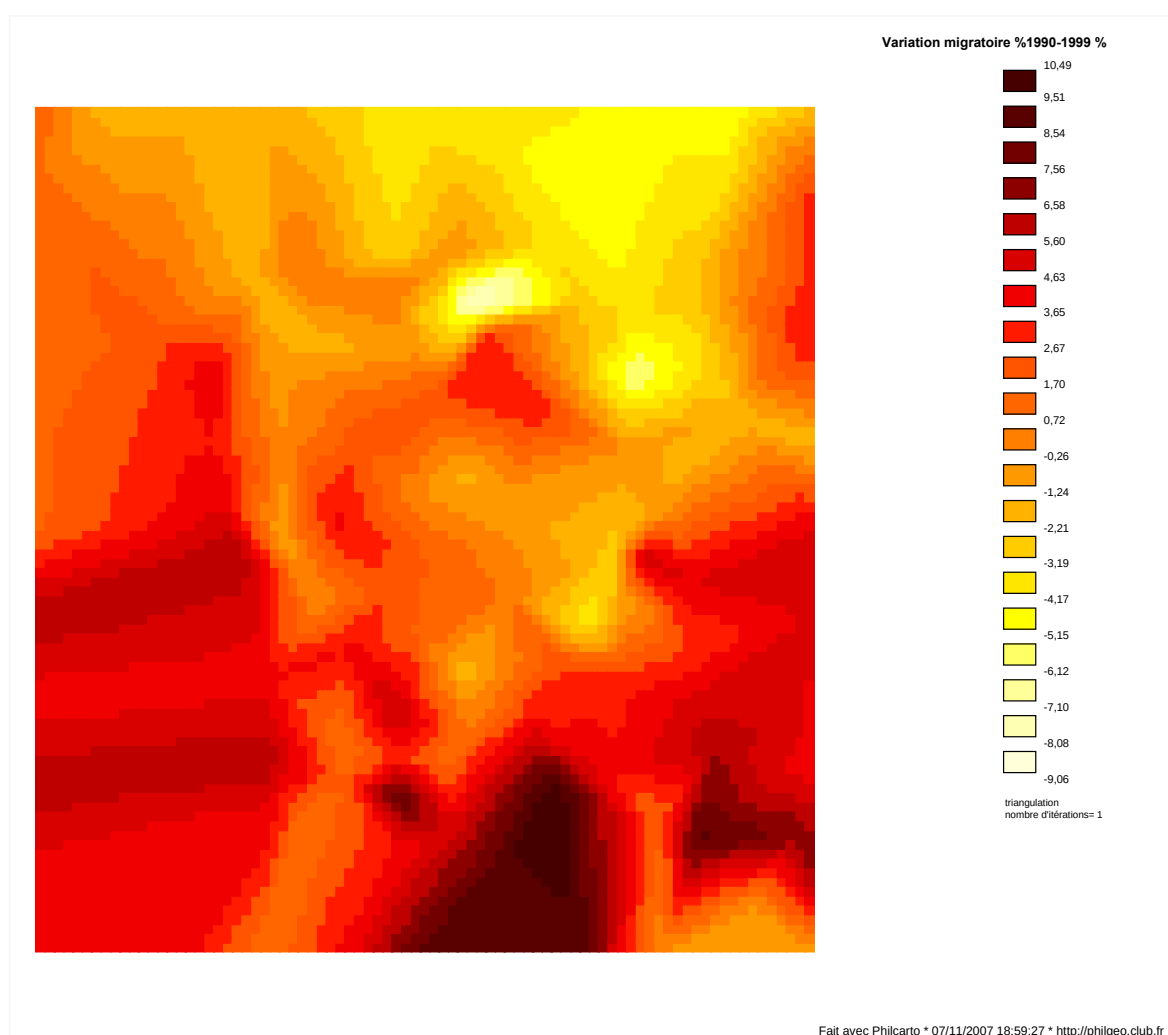


Figure 14.4 : La carte isoplèthe initiale de la variation de la population
Des départements français 1990-1999 due au solde migratoire.

Lissage

☒ triangulation
 ☐ krigeage
☐ surfaces de tendances

Paramètres de la triangulation

nombre d'itérations=

Appliquer

Options de tracé

☒ limiter le tracé à l'habillage
☐ afficher les points de mesure
☐ bloc diagramme

Couleurs

☐ inverser
 ☐ 
☐ 
☒ 
☐ 
☐ 
☐ 
☐ 
☐ 
☐ 
☐ 

Maillage

résolution du maillage=

Appliquer

Niveaux

minimum -9.05518908 maximum=

valeur du niveau du sommet=
 valeur du niveau de la base=
 nombre de niveaux=
 intervalle entre niveaux=

Appliquer

Figure 14.5 : Le dialogue **Lissage**.

Habillage

☐ Departements.S
☒ Regions.S
☐ Tracer les contours des unités spatiales

Figure 14.6 : Placer un habillage sur la carte isoplèthe initiale.

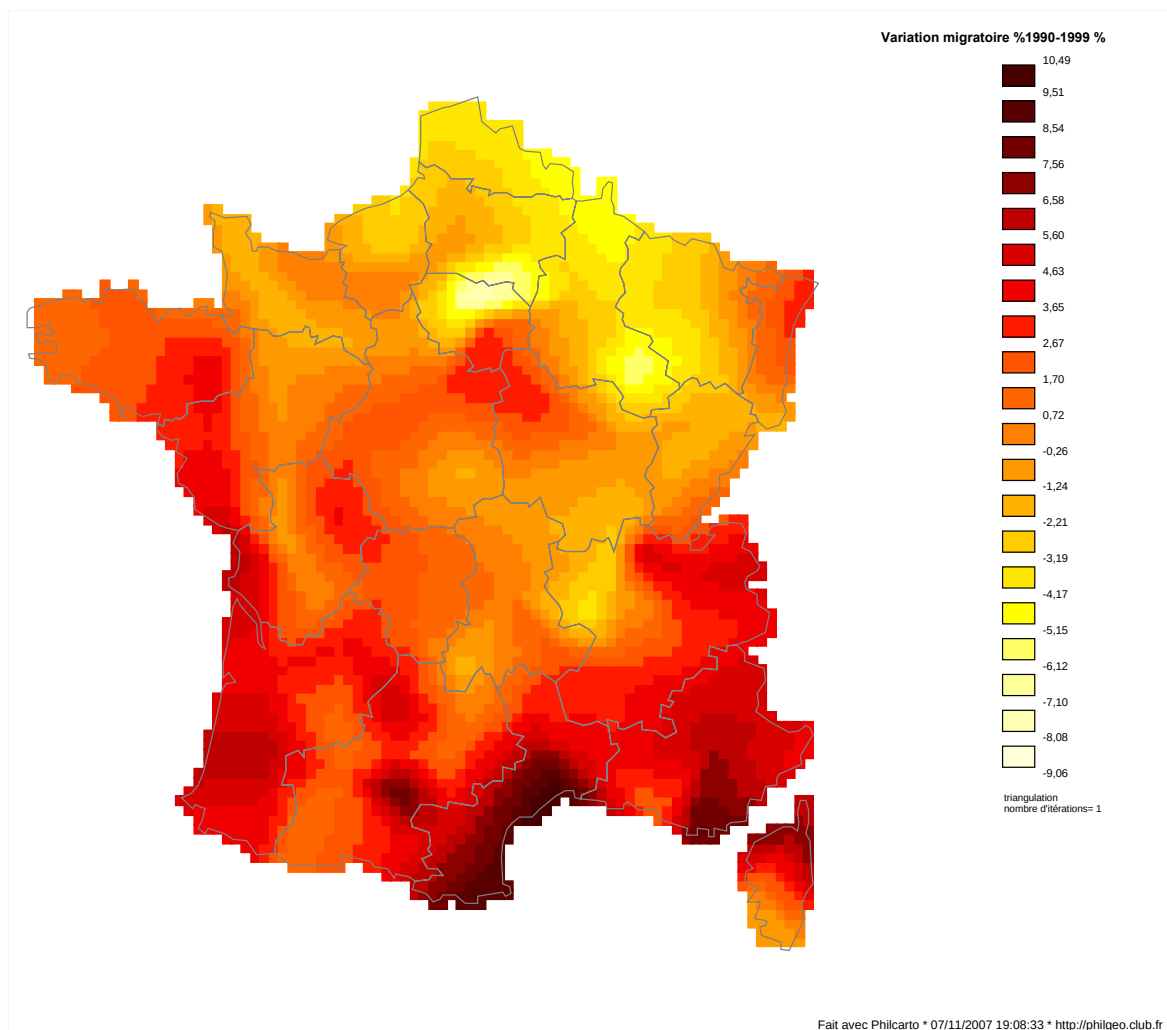


Figure 14.7 : La carte isoplèthe de la variation de la population des départements français 1990-1999 due au solde migratoire, limitée aux contours des régions périphériques.

Quelques améliorations peuvent être apportées à cette carte. La première de ces améliorations concerne la discrétisation des valeurs calculées. Par défaut, l'étendue des valeurs (maximum-minimum) est divisée en 20 classes (c'est une discrétisation en 20 classes d'étendue égale). Les valeurs calculées présentent des décimales, ce qui n'est pas grave en soi, mais qui peut être évité en modifiant le minimum et le maximum dans la rubrique **Niveaux** du dialogue **Lissage**. Par exemple en choisissant comme valeur maximum 10, et comme minimum -10, on obtient, pour 20 classes un pas de discrétisation régulier de 1 (fig. 14.8). La validation de ces valeurs par un clic sur **Appliquer** a pour effet un nouvel affichage de la carte, qui n'est pas vraiment différente de la précédente puisque les valeurs retenues sont très proches du minimum et du maximum de la variable cartographiée.

Niveaux

minimum -9.05518908 maximum=

valeur du niveau du sommet= 10

valeur du niveau de la base= -10

nombre de niveaux= 20

intervalle entre niveaux= 1

Appliquer

Figure 14.8 : Choisir le pas de discrétisation et le nombre de niveaux.

La seconde amélioration a trait aux gammes de couleurs. En effet, la variation migratoire prend des valeurs négatives lorsque l'immigration ne compense pas l'émigration, et des valeurs positives quand l'immigration est plus importante que l'émigration. Dans ce cas, il existe une valeur pivot, la valeur 0, en-dessous de laquelle on observe des pertes et au-dessus de laquelle on a affaire à des gains dus aux mouvement migratoire. C'est dans ce cas qu'on peut appliquer une gamme bicolore en opposition, soit la gamme bleu-rouge , soit la gamme vert-rouge . Sélectionnons la première dans le dialogue **Couleurs** : la carte qui s'affiche avec ses nouvelles couleurs (fig. 14.9) apparaît plus expressive que la précédente.

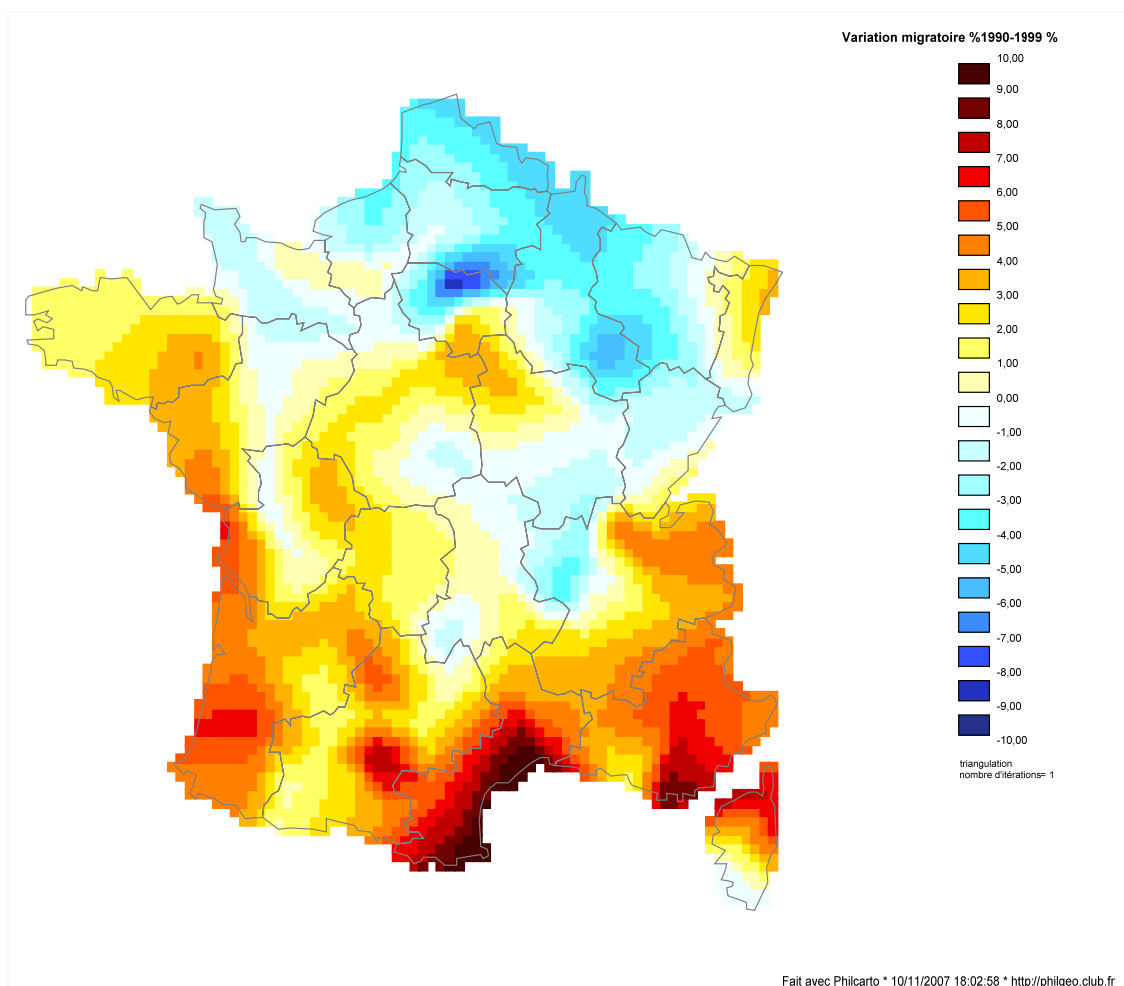


Figure 14.9 : Version bicolore de la carte isopèthe de la variation migratoire 1990-1999.

Par défaut, *Philcarto* attribue autant de couleurs bleues que de couleurs jaunes-oranges-rouges, ce qui est naturellement le cas avec 20 classes. Si la valeur pivot (0 ou 100 s'il s'agit d'un pourcentage...) entre deux caissons bleus ou jaune-orange, il faudrait ajuster le nombre de classes bleues à l'aide du curseur  classes bleues.

Cette carte peut faire l'objet d'une représentation par un bloc diagramme, une perspective qui permet parfois de mieux apprécier la surface calculée. Il suffit de cocher la case  bloc diagramme (fig. 14.10) pour la rendre active. Le bloc diagramme est affiché dès que cette case est cochée (fig. 14.11).

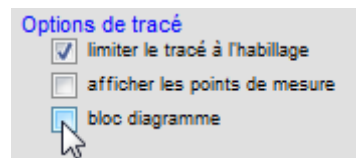


Figure 14.10 : Cocher la case pour visualiser la surface sous forme d'un bloc-diagramme.

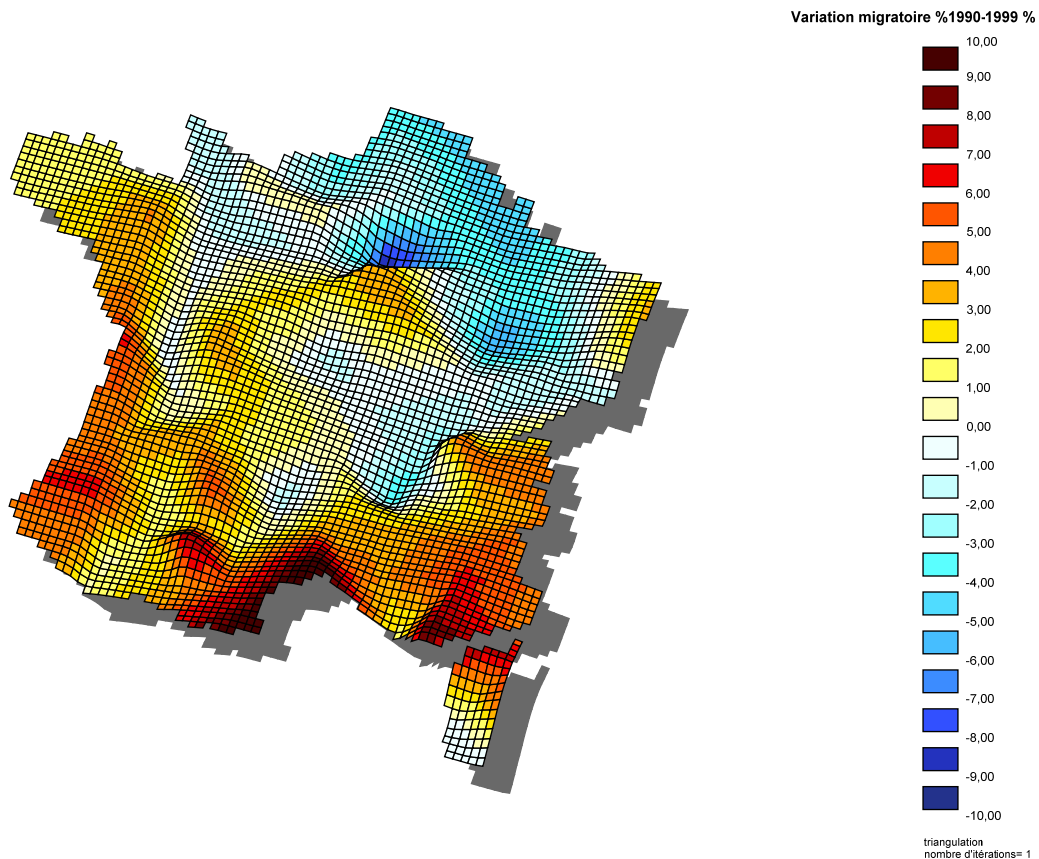


Figure 14.11 : Bloc diagramme de la version bicolore de la carte isoplèthe de la variation migratoire 1990-1999. Vue vers le Nord-Ouest.

Fait avec Philcarto * 10/11/2007 18:20:43 * <http://philgeo.club.fr>

Au moment où s'affiche le bloc diagramme, le dialogue **bloc diagramme** permettant d'en contrôler certaines de ses caractéristiques apparaît dans le cadre **Outils** (fig. 14.12).

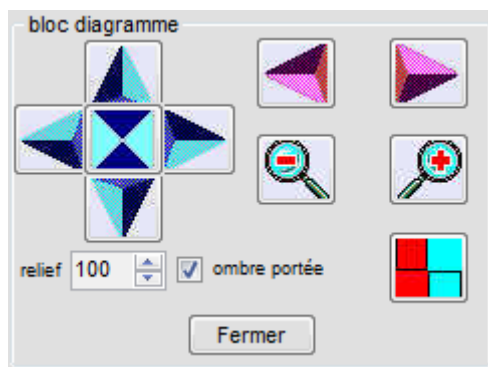


Figure 14.12 : Le dialogue **bloc diagramme**.

Philcarto a choisi une position initiale (celle de la figure 14.11) qui peut s'avérer inadaptée à la distribution géographique étudiée. Par exemple, une « chaîne de montagnes » située au premier plan pourrait masquer un « piémont » localisé à l'arrière. Les boutons de rotation permettent de faire tourner la figure horizontalement et verticalement, de manière à rechercher le meilleur angle de vue possible. Cela est réalisé à l'aide des flèches bleues (fig. 14.13).

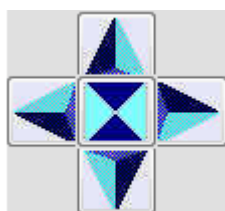


Figure 14.13 : Les boutons de rotation pas à pas du bloc-diagramme.

Ces flèches permettent une rotation pas à pas dans la direction qu'elles indiquent (le pas de rotation n'est pas modifiable par l'utilisateur). Le bouton central remplace la figure en position initiale.

À la droite des boutons de rotation pas à pas, figurent deux flèches de couleur rose qui permettent d'activer la rotation horizontale continue, vers la gauche



ou vers la droite



. Un clic sur l'une ou l'autre de ces flèches provoque la rotation de la figure à une vitesse qui dépend en grande partie de la puissance de l'ordinateur utilisé. Le bloc-diagramme s'anime, et on peut l'arrêter à tout moment par un clic sur



. En cours de rotation continue, il est possible de modifier l'angle de vue vertical par un clic sur l'une ou l'autre des flèches



ou . La case à cocher ☒ **ombre portée** permet de tracer une ombre sous le bloc diagramme permettant ainsi d'améliorer la perception du relief.

Dans le cadre 3D, d'autres éléments complètent le contrôle de la vue du bloc diagramme. Le curseur **relief** 100 permet d'augmenter ou de diminuer (vers la gauche) l'impression de relief. Les loupes assurent les fonctions de zoom sur le bloc

diagramme (au centre du bloc), par rapprochement  ou par éloignement  de l'observateur. Le bouton  permet alternativement d'effacer ou d'afficher le maillage. L'intérêt de l'effacement du maillage est variable selon la configuration géographique et n'a d'utilité que lorsque le nombre de niveaux est grand puisque ce sont les seules couleurs qui, dans ce cas, traduisent le relief. Voici un nouvel affichage du bloc diagramme de la variation migratoire, vue en direction du Sud-Ouest, avec une accentuation de relief de 85% et un zoom 2x (fig. 14.14).

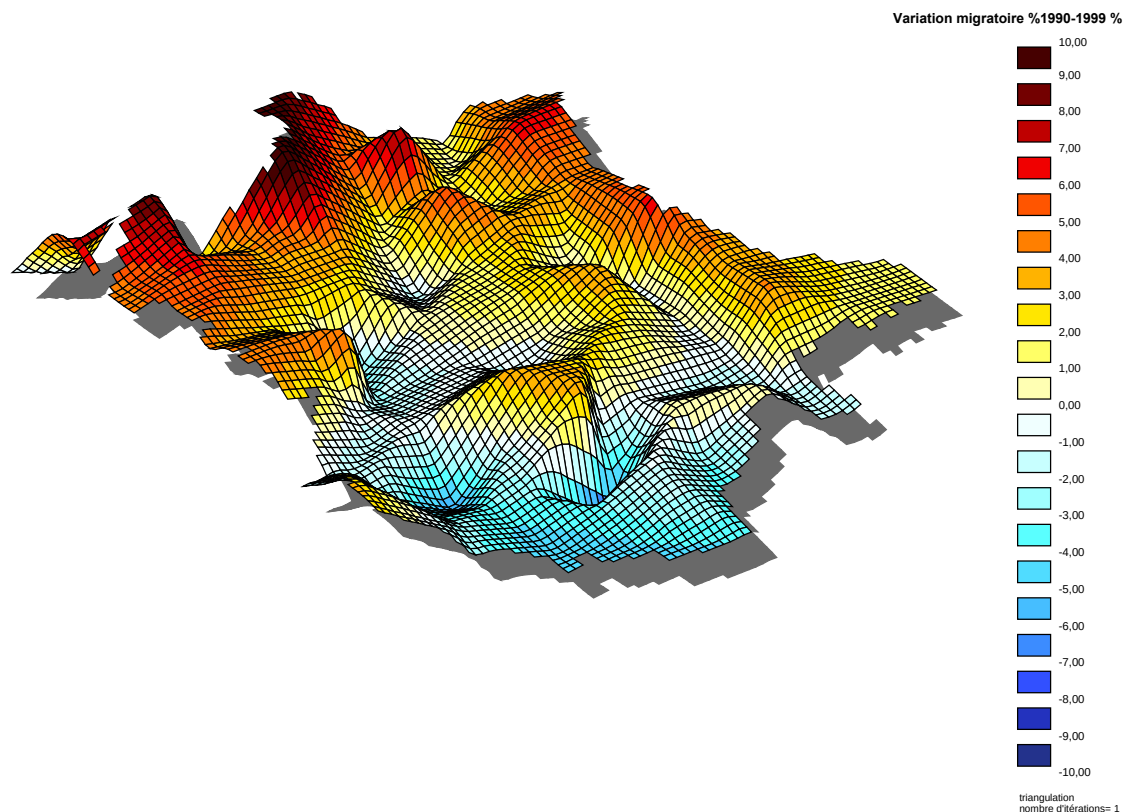
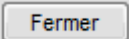


Figure 14.14 : Bloc diagramme de la version bicolore de la carte isoplèthe de la variation migratoire 1990-1999. La vue est orientée vers le Sud-Ouest ce qui a pour effet de placer au premier plan les altitudes les plus basses.

Un clic sur  ferme le bloc-diagramme et affiche à nouveau la carte.

La troisième amélioration possible concerne la résolution du maillage utilisé pour interpoler les valeurs sur l'ensemble de la surface. Par défaut, ce maillage régulier résulte de la division par 100 des longitudes (X) et des latitudes (Y), soit 10 000 rectangles pour l'ensemble de la surface avant l'élimination de tous les éléments situés en dehors de la surface couverte par l'habillage. Ce maillage est satisfaisant pour élaborer la carte et c'est celui qui doit être utilisé dans le cas des bloc diagrammes, surtout s'il s'agit de les faire pivoter.

Pour l'affichage de la carte, une résolution plus élevée permet d'obtenir une image mieux lissée. Pour obtenir ce résultat, il suffit d'augmenter la résolution du maillage avec le curseur prévu à cet effet dans le dialogue **Maillage** (fig. 14.15). Attention cependant à utiliser cette possibilité avec parcimonie : un maillage très fin demande un temps de calcul beaucoup plus long car le nombre de rectangles augmente en raison du carré de la résolution. La nouvelle résolution entre en vigueur dès qu'on clique sur **Appliquer**.

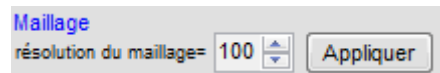


Figure 14.15 : Le dialogue **Maillage**.

Enfin, il est parfois utile de visualiser la position des points de mesure, ce qui se fait simplement avec la case à cocher ☒ **afficher les points de mesure** du dialogue **Options de tracé**.

14.4. Afficher une carte lissée par krigeage

La seconde méthode de lissage proposée par *Philcarto* s'appelle « krigeage ». Ce mot provient du nom de son précurseur, D.G. Krige, un ingénieur des mines sud-africain. C'est le français Matheron qui a formalisé cette méthode et l'a baptisée krigeage. Il s'agit d'une technique d'interpolation spatiale, qui tient compte de la configuration géométrique des points observés et de la structure spatiale propre à la variable estimée. Le krigeage permet d'estimer des valeurs à différents points d'une surface, par exemple, à chaque intersection d'une grille régulière dérivée à partir de points échantillonnés de façon irrégulière.

Le krigeage est bien plus qu'une technique de lissage car il s'agit d'une méthode de géostatistique permettant aussi de calculer l'erreur d'estimation. Un utilisateur sérieux doit bien connaître les conditions d'utilisation de la méthode, en particulier pour choisir la bonne fonction d'estimation selon l'allure du variogramme expérimental observé sur les données.

Pour réaliser un krigeage, il faut choisir le bouton radio ☒ **krigeage** dans le dialogue **Lissage**. Le calcul est immédiatement effectué et la carte affichée (fig. 14.16). On ne peut pas dire que le résultat obtenu soit très différent du lissage par triangulation ! Les plages de couleurs apparaissent cependant moins « tranchées au couteau », ce qui donne un résultat plus lissé encore lorsque la résolution du maillage est élevée (fig. 4.17).

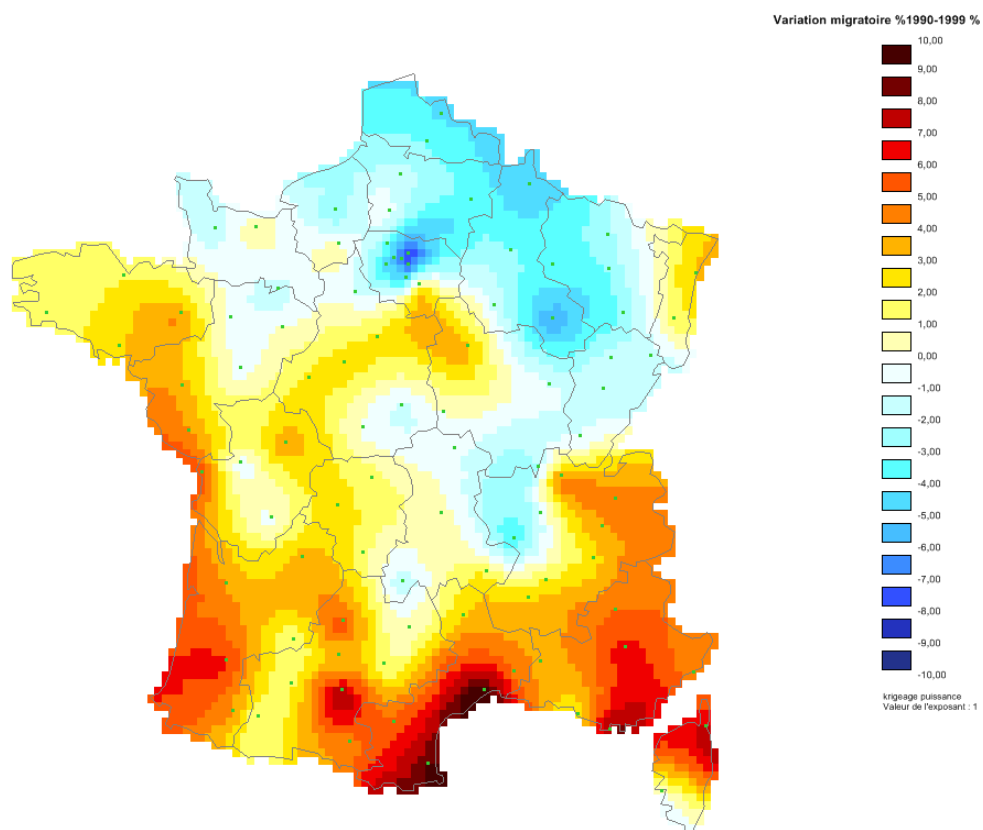


Figure 14.16 : La carte isoplèthe de la variation migratoire 1990-1999 obtenue par krigeage puissance (maillage 100x100). Les points verts sont les points de mesures (préfectures).

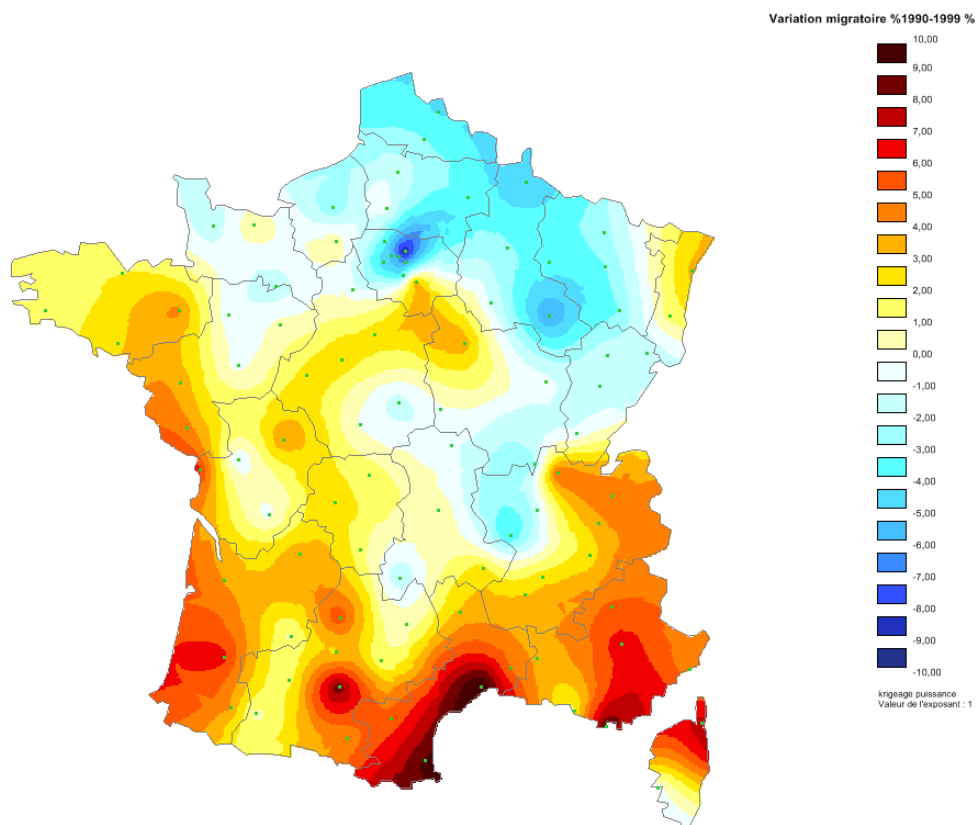


Figure 14.17 : La carte isoplèthe de la variation migratoire 1990-1999 obtenue par krigeage puissance (maillage 500x500). L'habillage est celui des régions.

Par défaut, le krigeage est de type puissance, avec comme exposant 1 (linéaire donc). Il est possible de modifier ses caractéristiques à l'aide du dialogue **Paramètres du krigeage** du dialogue **Lissage** (fig. 14.18).

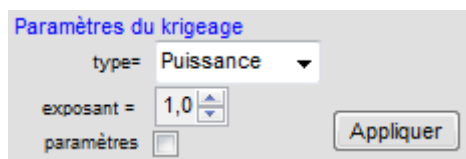


Figure 14.18 : Le dialogue **Paramètres du krigeage**.

Dans le cas du krigeage puissance, l'exposant peut varier de 1 à 5 par pas de 0,1 à l'aide du curseur **exposant = 1,2**.

Le type de krigeage peut être choisi, parmi quatre types différents en déroulant le menu local type (fig. 14.19)

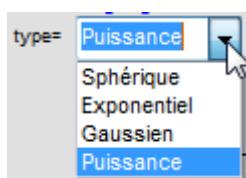


Figure 14.19 : Menu local pour choisir le type de krigeage.

Un clic sur **Appliquer** valide les options de krigeage. Si la case **paramètres** ☒ est cochée, le tracé du nouveau lissage est précédé par l'affichage d'une fenêtre **Variograms** (fig. 14.20) contenant l'ensemble des paramètres modifiables (en anglais car *Philcarto* utilise une bibliothèque externe d'origine italienne...).

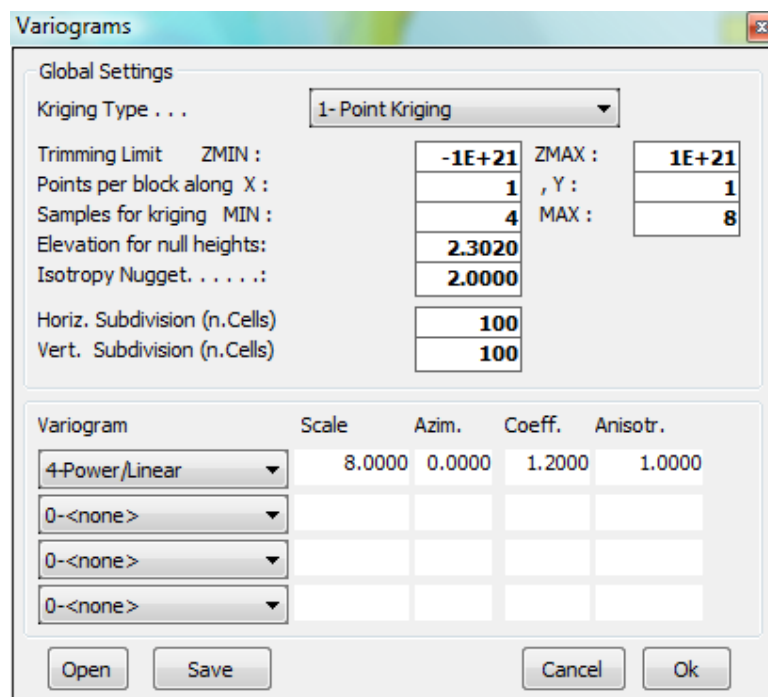


Figure 14.20 : La fenêtre **Variograms**.

14.5. Afficher une carte de surface de tendance

De même que le krigeage, l'analyse en surfaces de tendances (AST) est plus qu'une méthode de lissage même si chaque surface de tendance, prend la forme d'une carte lissée. Il s'agit d'une technique d'ajustement de surfaces caractérisées par des équations polynomiales qui peut être vue de deux façons. La première considère que l'AST est une technique de lissage spatial parmi d'autres, qui a pour objectif d'éliminer des variations locales pour retenir seulement une surface qui rende compte au mieux de la variabilité spatiale globale de la variable analysée, avec une perte contrôlée de variance. La seconde approche voit en l'AST une méthode d'analyse en composantes d'échelles qui cherche à extraire une organisation de l'espace résultant de jeux d'influences qui s'exercent sur des étendues de plus en plus réduites. On fait ainsi l'hypothèse qu'une organisation de l'espace donnée est la résultante de processus :

- généraux affectant non seulement la totalité de l'espace étudié, mais également ses voisins ;
- régionaux couvrant toute l'aire d'étude ;
- locaux, ne concernant qu'une partie seulement de l'espace analysé.

A chacune de ces échelles correspond une composante qui, une fois cartographiée, prend la forme d'une surface de tendance. Dans cette seconde perspective, on ne cherche pas directement le meilleur ajustement, mais on examine la signification éventuelle des surfaces successives [Le Gauffey & Waniez ; 1987] [Riddel ; 1970].

Il existe plusieurs variantes de l'AST. Celle que propose *Philcarto* recourt à la méthode des polynômes bidimensionnels : les coordonnées géographiques (X,Y) sont les variables explicatives d'une variable Z à représenter, soit $z_i = f(x_i, y_i)^P$ où P représente l'ordre du polynôme. On obtient ainsi une série de surfaces de plus en plus complexes quand s'accroît l'ordre du polynôme. Pour réaliser ce module, je me suis inspiré du programme FORTRAN diffusé par Albert Goodman sur le remarquable site internet :

<http://www.deakin.edu.au/~agoodman/masters/index.html>

<http://www.deakin.edu.au/~agoodman/masters/chapter2.html>

Pour tracer une surface de tendance avec *Philcarto*, il faut choisir le bouton radio  **surfaces de tendances** dans le dialogue **Lissage**. Le calcul est immédiatement effectué et la carte affichée (fig. 14.21). Au bas de la légende de cette carte est indiqué qu'il s'agit d'une surface d'ordre 1 qui explique 45,3% de la variance.

Avec *Philcarto*, il est possible de tracer des surfaces d'ordre 1 à 4. A l'ordre 1 l'équation de la surface a pour expression :

$$Z = aX + bY + E$$

Et à l'ordre 2 :

$$Z = aX + bY + cXY + dX^2 + eY^2 + E$$

Où X et Y sont les vecteurs contenant les coordonnées des points de mesure, et Z le vecteur des estimations de la variable à cartographier.

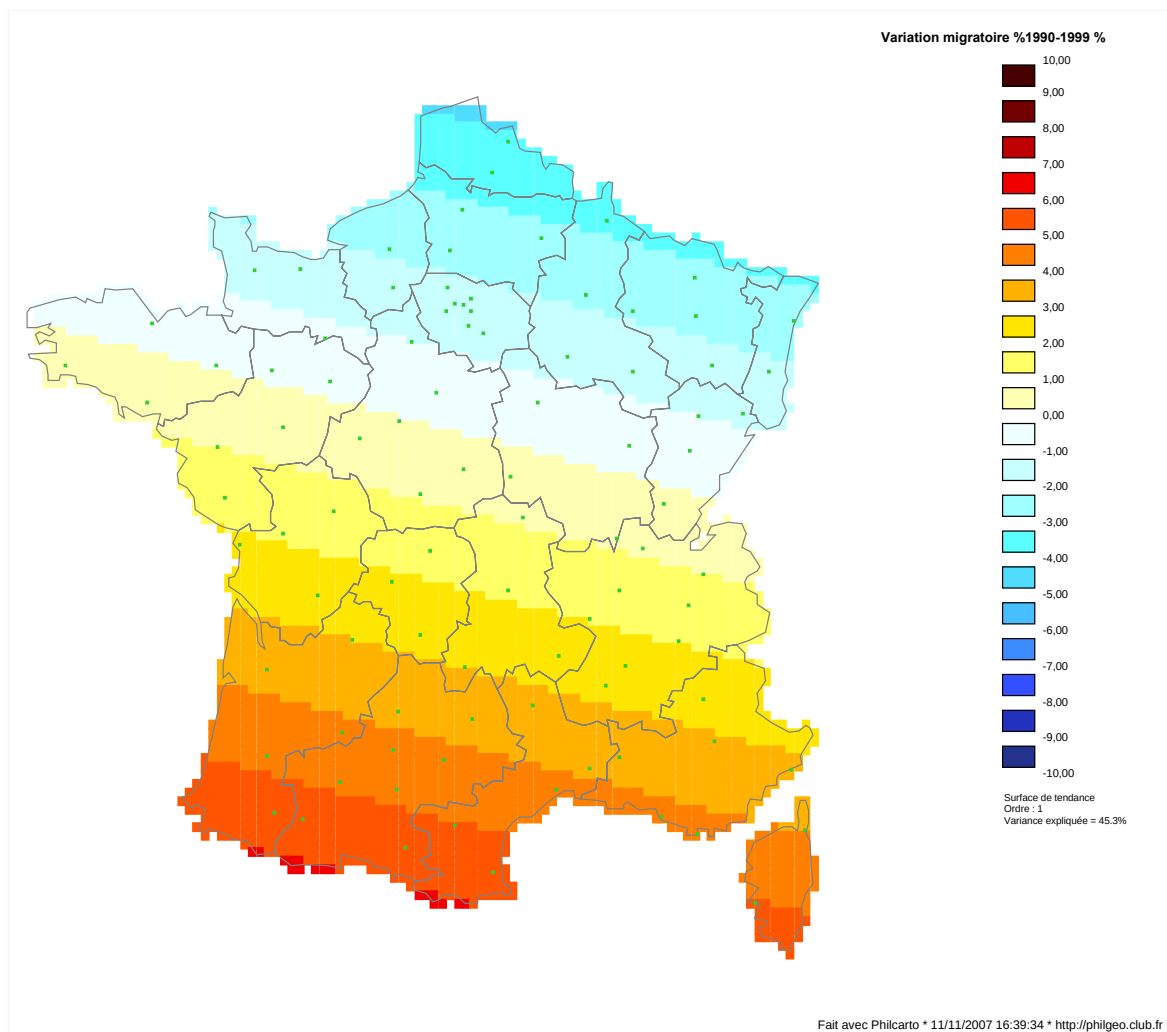


Figure 14.21 : La surface de tendance d'ordre 1 variation migratoire 1990-1999.

L'ordre de la surface à cartographier est contrôlé avec le dialogue **Paramètres de l'analyse en surfaces de tendances** du dialogue **Lissage** (fig. 14.22).

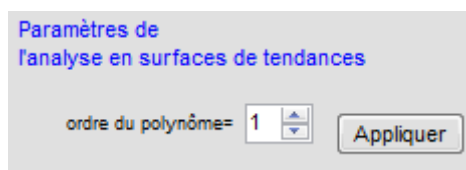


Figure 14.22 : Le dialogue de contrôle de l'ordre du polynôme des surfaces de tendances.

Il suffit d'utiliser le curseur indiquant l'ordre du polynôme (1, 2, 3 ou 4) et de valider ce choix par un clic sur **Appliquer** pour que la nouvelle surface soit tracée. Avec une surface d'ordre 4, on atteint 58,8% de variance expliquée, pour une forme en demi-cuvette dont le fond est au nord du pays, et dont les bords passent par l'arc atlantique, les Pyrénées et la côte méditerranéenne (fig. 14.23). Cette surface peut être visualisée en volume par un bloc diagramme (fig. 14.24).

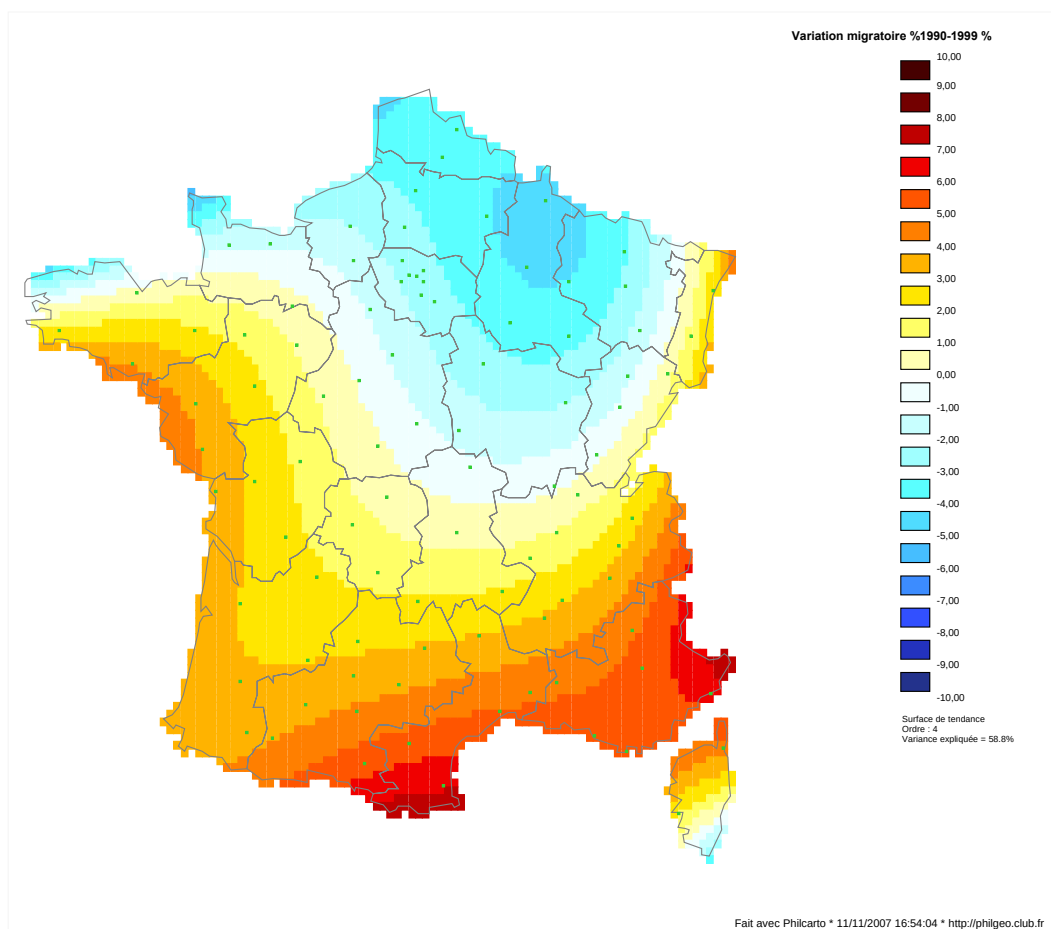


Figure 14.23 : La surface de tendance d'ordre 4 de la variation migratoire 1990-1999.

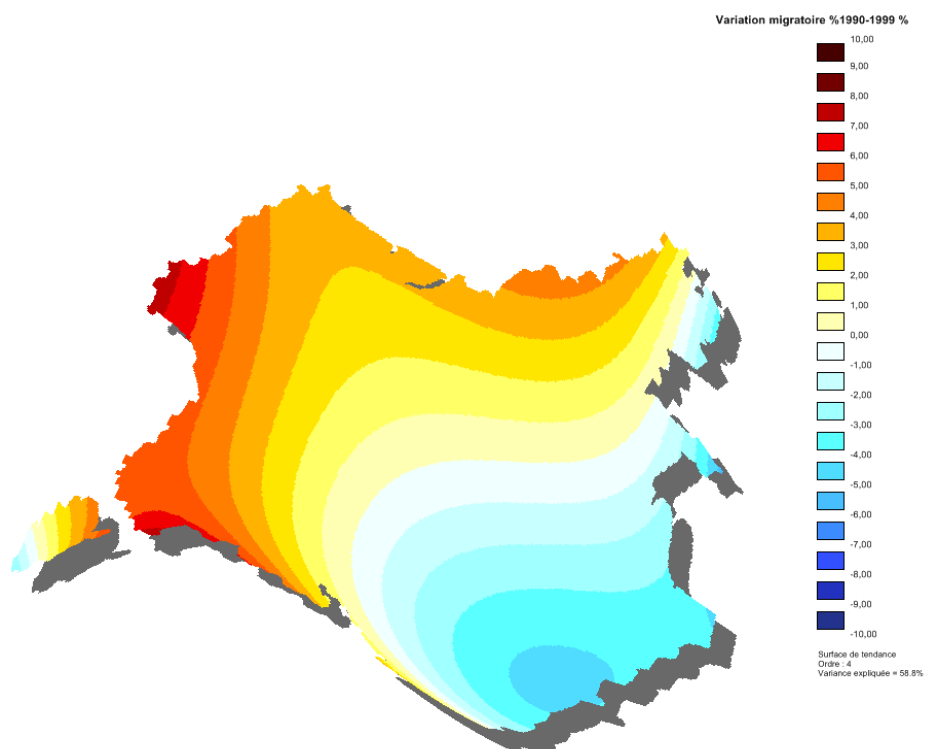


Figure 14.24 : Vue vers le Sud-Ouest de la surface de tendance d'ordre 4 de la variation migratoire 1990-1999 (maillage 500x500).

Comme avec toute méthode d'ajustement, il peut s'avérer utile de calculer les résidus, c'est-à-dire les différences entre les valeurs observées en chaque point et l'estimation de ces valeurs par la surface de tendance. Si la case **calculer résidus** ☒ est cochée, *Philcarto* ajoute, dans la liste **couleurs** du dialogue **Variables** une nouvelle variable contenant ces résidus de la régression polynomiale. La nouvelle variable s'appelle Résidus AST, suivi du nom de la variable, suivi de l'ordre du polynôme, suivi du pourcentage de variance expliquée.

Ces variables de résidus peuvent être cartographiées par un clic sur **CARTO** suivi de la sélection, dans la liste **couleurs** du dialogue **Variables**, de la variable correspondante, ici **Résidus AST [Variation migratoire %1990-1999]**, suivi du tracé d'une carte, ou en cercles proportionnels colorés ; la méthode de discrétisation **S5** est bien adaptée dans ce cas (fig. 14.25).

Cette carte présente des résidus positifs élevés (les valeurs estimées sont plus faibles que les valeurs observées) dans quelques départements du midi (Haute-Garonne, Hérault, Var), et dans nombre de départements du Bassin Parisien (Seine-et-Marne, Yonne, Loiret) ; ces résidus traduisent une attractivité supérieure à ce que la position géographique des département en question laisserait espérer. D'autres départements, au contraire, semblent répulsifs : Seine-Saint-Denis, et Val-de-Marne à proximité de Paris, et plus largement l'ensemble de la Région Île-de-France, à l'exclusion de la Seine-et-Marne, le Rhône (Lyon), les Bouches-du-Rhône (Marseille), mais aussi le Cantal et les Deux-Sèvres.

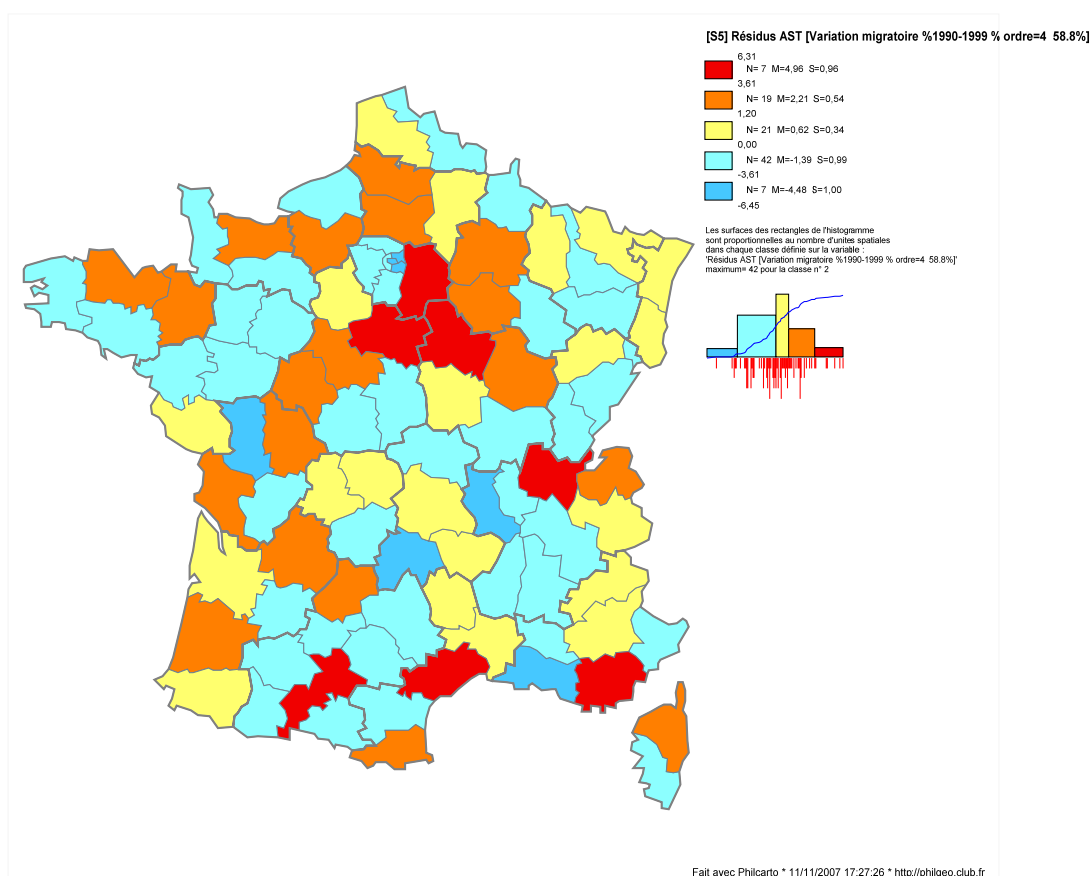


Figure 14.25 : La carte choroplèthe des résidus de la surface de tendance d'ordre 4 de la variation migratoire 1990-1999.

15. Les aires d'influences urbaines : modèles de Reilly et Huff

Dossier : \Philcarto exemples\France gravitation\
Fond de carte : France carte villes.ai
Statistiques : France data villes santé.xls

De nombreux ouvrages exposent la théorie des lieux centraux [King ; 1984]. Présentée par W. Christaller dans les années 1930, cette théorie repose sur l'hypothèse que des lois déterminent la taille, le nombre et la distribution des lieux centraux. Pour valider cette hypothèse, différents modèles ont été testés, mettant en évidence l'importance des champs de forces spatiaux. Leur existence est liée aux déplacements que les consommateurs sont amenés à faire pour satisfaire leurs besoins. Dans le choix du lieu où ils se rendent, ils cherchent à minimiser leurs dépenses, notamment celles de transport. Ainsi, plus les biens et services proposés sont exceptionnels, plus grandes sont les distances franchies. Ce comportement rationnel a pour résultante la polarisation de l'espace en fonction de l'attraction exercée par les villes sur leur espace environnant. Un modèle dérivé de la Gravitation Universelle permet d'évaluer l'influence théorique des lieux centraux et de tracer les limites des aires d'influence. Pour R. Brunet, le modèle de gravitation «a une logique sociale qui le légitime pleinement : plus on est près, plus c'est gros, plus cela offre d'occasions et plus on est attiré» [Brunet ;1997]. Mais comme le rappelle A. Dauphiné, il y a lieu «de multiplier les études pour voir si, malgré ses lacunes axiomatiques, la théorie des places centrales ne rend pas compte de l'organisation des réseaux urbains» [Dauphiné ; 1979].

Les modèles de gravitation donnent une représentation simplifiée des aires d'attraction, c'est-à-dire du partage de l'espace environnant provoqué par l'accumulation des facteurs de centralité en un lieu dans un espace isotrope [Waniez ; 1990]. Par analogie avec la loi de gravitation universelle de Newton, les modèles élémentaires d'attraction urbaine reposent sur trois paramètres : la masse, la distance, et l'exposant de la distance. La masse est un effectif mesuré en chaque lieu central dont la valeur exprime l'accumulation des facteurs de centralité ; il peut s'agir de la population active dans les services à haute valeur ajoutée (conseil financier, ingénierie des réseaux, etc.), du nombre de grands magasins ou de leur superficie, ou bien encore d'une caractéristique particulière à un service dont on cherche à évaluer l'influence, comme par exemple un hôpital. La distance est en général mesurée à vol d'oiseau, l'espace étant, par hypothèse, isotrope. Cette hypothèse est très forte, et explique dans bien des cas les distorsions observées entre les résultats exhibés par le modèle et la réalité empirique. Enfin, la valeur de l'exposant souligne que certains types d'interaction spatiale sont plus sensibles à la distance que d'autres, une valeur élevée traduisant ainsi une forte résistance au déplacement. Une version modifiée de ce modèle a été proposée par D.L. Huff [Huff ; 1973] ; ce nouveau modèle exprime l'attraction en un point de l'espace, non pas de manière absolue, mais en la rapportant à la somme des attractions mesurées sur tout l'espace. Les attractions sont ainsi comparables d'un niveau de la hiérarchie urbaine à l'autre.

Dans ce module, le tracé des limites des aires d'influences est l'adaptation d'un programme FORTRAN réalisé par R. Gambini [Gambini ; 1966]. Le programme de tracé des cartes de probabilités d'influence est le résultat d'une recherche personnelle.

15.1. Sélectionner une variable de masse (d'effectif)

La sélection d'une variable se fait simplement par un clic sur le nom d'une des variables figurant dans la liste **symboles** du dialogue **Variables**. Ici la variable sélectionnée, **Nombre de médecins spécialistes**, contient nombre de médecins spécialistes dans 35 grandes villes françaises (fig. 15.1). On considère que ce nombre de médecins forme une masse attractive sur les espaces environnant ces villes.

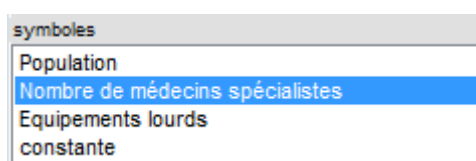


Figure 15.1 : Sélection d'une variable de couleur.

15.2. Afficher la carte des probabilités d'influences

La sélection de la variable de symbole provoque l'affichage du dialogue **Types de cartes**. Cliquer sur l'icône du type carte en cercles proportionnels colorés  (fig. 15.2). La carte s'affiche alors à l'écran (fig. 15.3).

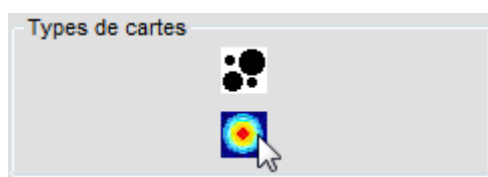


Figure 15.2 : Clic sur l'icône « carte en cercles proportionnels colorés ».

Notons que dans la légende, les probabilités comprises entre 0% et 20% sont rassemblées dans une seule classe de couleur bleu foncé ; la valeur pivot de la gamme de couleurs en opposition est 50% (une chance sur 2).

L'application sur cette carte d'un calque d'habillage départemental la rend plus lisible (fig. 15.4). Chaque lieu central est indiqué par un point vert (application de la case à cocher ☒ **afficher les lieux centraux** du dialogue **Options de tracé**) (fig. 15.5). Sur cette carte, les probabilités d'influence de chaque centre diminuent depuis le centre lui-même, où la probabilité est toujours de 100%, vers la périphérie, de manière plus ou moins abrupte pour atteindre généralement des probabilités inférieures à 50%.

L'aire de forte influence de Paris ($p > 50\%$) apparaît très grande au regard de la plupart des autres villes. Celles-ci voient leur aire de forte influence s'accroître en fonction de la distance à la capitale : aire réduite pour les villes du Bassin Parisien (Orléans, Rouen, Amiens, Reims), plus étendue pour les villes de l'Ouest (Brest, Rennes, Nantes, Poitiers) ainsi que Limoges et Clermont-Ferrand. Lyon présente une configuration semblable à Paris dans la mesure où elle écrase Saint-Etienne de son influence, et, dans une moindre mesure, Grenoble. Le chapelet de grandes villes des deux régions méditerranéennes, depuis Perpignan jusqu'à Nice, fait qu'aucune d'elles, même Marseille, ne peut prétendre à une aire d'influence considérable. La

situation est tout autre pour Bordeaux, et surtout Toulouse dont les aires de forte influence sont très étendues.

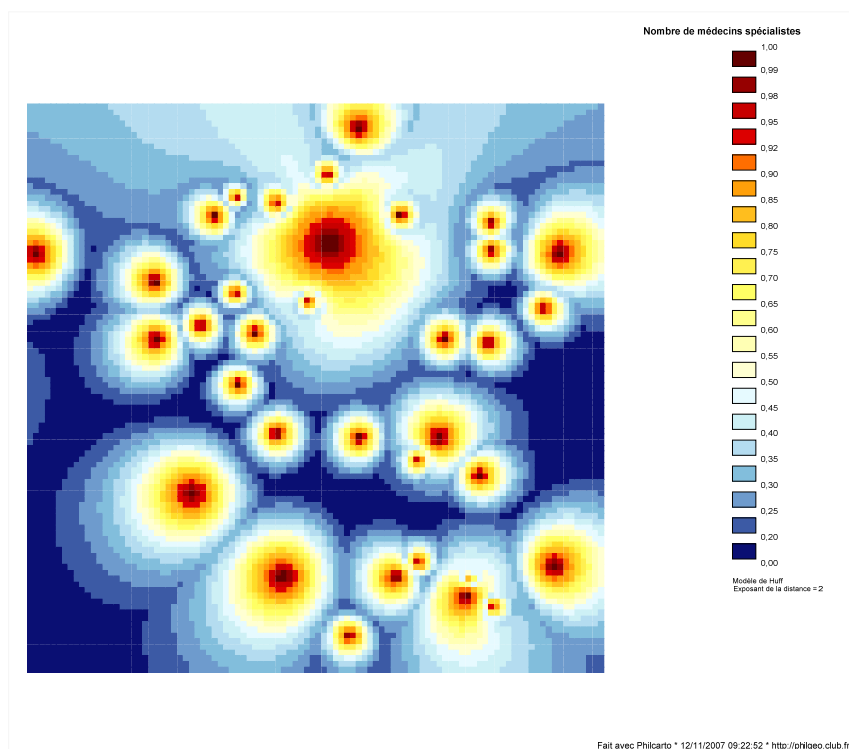


Figure 15.3 : La carte des probabilités d'influences selon le nombre de médecins spécialistes (valeur de l'exposant du modèle de gravitation : 2).

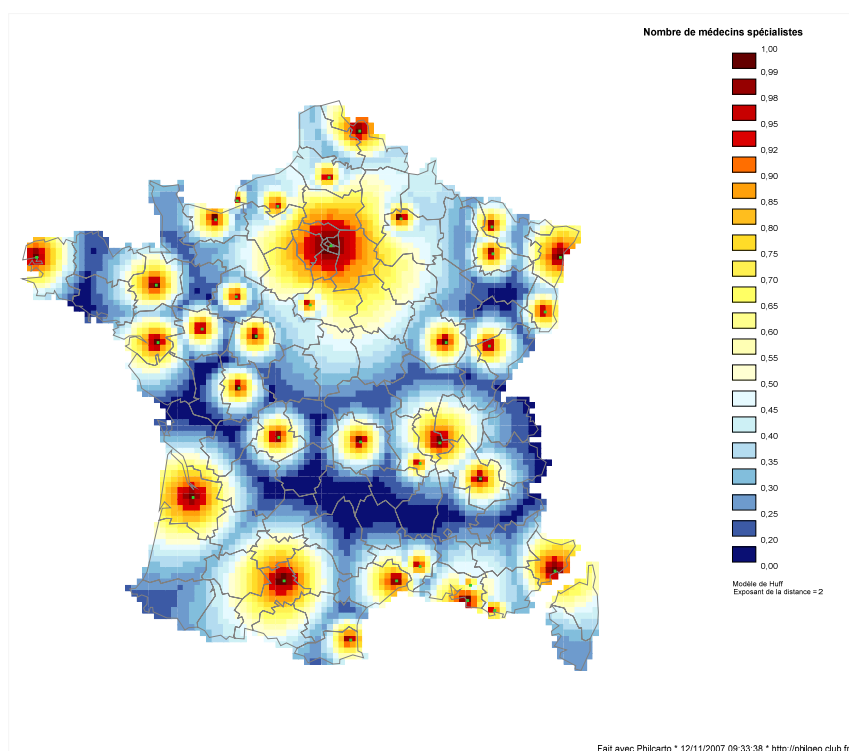


Figure 15.4 : La carte des probabilités d'influences selon le nombre de médecins spécialistes limitées à l'habillage départemental (valeur de l'exposant du modèle de gravitation : 2).

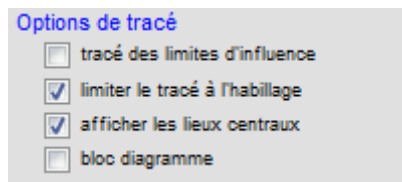


Figure 15.5 : Le dialogue **Options de tracé**.

Le dialogue **Options de tracé** permet aussi le tracé des limites d'influences en cochant la case ☒ **tracé des limites d'influence**. Les limites sont immédiatement calculées et dessinées. Le temps de calcul peut s'avérer long si le nombre de centres est grand. Le dialogue **Couleur des limites d'influence** permet de choisir une couleur qui n'est pas présente dans la gamme utilisée pour les probabilités (fig. 15.6).



Figure 15.6 : Le dialogue **Couleur des limites d'influence**.

Philcarto superpose la carte des limites d'influences à celle des probabilités (fig. 15.7). Dans la plupart des cas, les limites sont tracées dans des espaces de très faibles probabilités ($p < 25\%$). Mais ce n'est pas toujours le cas lorsque les villes sont très proches, telles Marseille et Toulon ou Lyon et Saint-Etienne : dans ces deux cas, les limites sont tracées dans des espaces de probabilités supérieures à 40%. Notons aussi que les aires d'influences d'Orléans, Caen, Rouen, Amiens et Reims sont incluses dans l'aire d'influence de Paris.

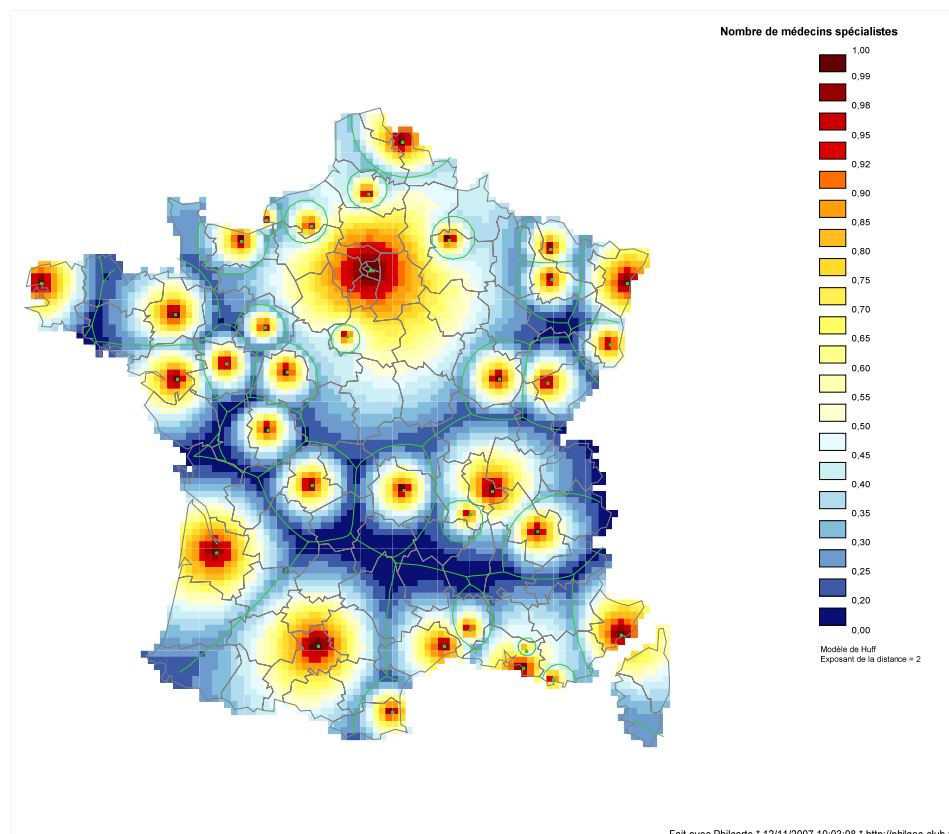


Figure 15.7 : Le tracé des limites d'influences (en vert).

15.3. Le paramétrage du modèle de gravitation

A la suite du tracé de la première carte d'influences, le dialogue **Modèle de gravitation de Huff** s'inscrit dans le cadre **Outils**. On y trouve quelques possibilités de paramétrage du modèle (fig. 15.8).

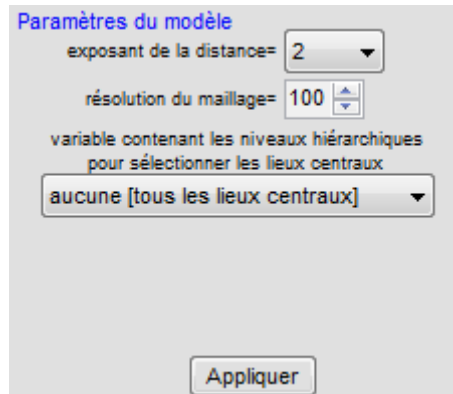


Figure 15.8 : Le dialogue **Paramètres du modèle** de Reilly-Huff.

Le menu local **exposant de la distance= 2** permet de choisir un exposant différent de 2 compris entre 0,5 et 4, par pas de 0,1. Augmenter la valeur de cet exposant revient à donner à la distance un degré de résistance plus élevé et, en conséquence, à limiter les possibilités de choix entre plusieurs lieux centraux. Il suffit de faire dérouler ce menu local pour choisir une nouvelle valeur.

Par défaut, la résolution du maillage est de 100 divisions sur X et sur Y (comme pour les cartes lissées). Cette résolution peut être modifiée jusqu'à 500 pour obtenir des plages de couleurs plus lisses.

Un second menu local **variable contenant les niveaux hiérarchiques pour sélectionner les lieux centraux** permet de choisir une variable contenant une indication de niveau hiérarchique caractérisant chaque centre. Les centres les plus importants doivent être codés 1 ; les niveaux suivants, à partir de la valeur 2, correspondent au rang des autres centres dans la hiérarchie urbaine. Si une variable de niveau est sélectionnée (fig. 15.9), un dialogue apparaît pour sélectionner les niveaux souhaités (fig. 15.10). Dans ce dialogue, il ne faut pas sélectionner les niveaux de façon discontinue, ni exclure un niveau supérieur quand on sélectionne un niveau inférieur (si on choisit le niveau 3, il faut obligatoirement choisir les niveaux 1 et 2 pour rendre compte de l'influence du niveau 3 compte tenu des deux niveaux supérieurs).

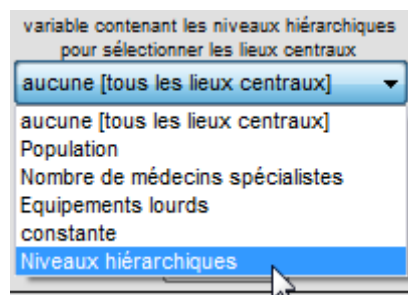


Figure 15.9 : Choisir une variable contenant les niveaux hiérarchiques.

variable contenant les niveaux hiérarchiques
pour sélectionner les lieux centraux

Niveaux hiérarchiques ▼

niveaux sélectionnés

☒ 1

☐ 2

Figure 15.10 : Pour ne retenir dans le modèle que les capitales régionales, sélectionner le niveau 1 seulement.

Un clic sur **Appliquer** permet de valider les paramètres retenus (fig. 15.11) et de tracer la nouvelle carte (fig. 15.12).

Paramètres du modèle

exposant de la distance= 3 ▼

résolution du maillage= 500 ▲▼

variable contenant les niveaux hiérarchiques
pour sélectionner les lieux centraux

Niveaux hiérarchiques ▼

niveaux sélectionnés

☒ 1

☐ 2

Appliquer

Figure 15.11 : Les nouveaux paramètres du modèle...

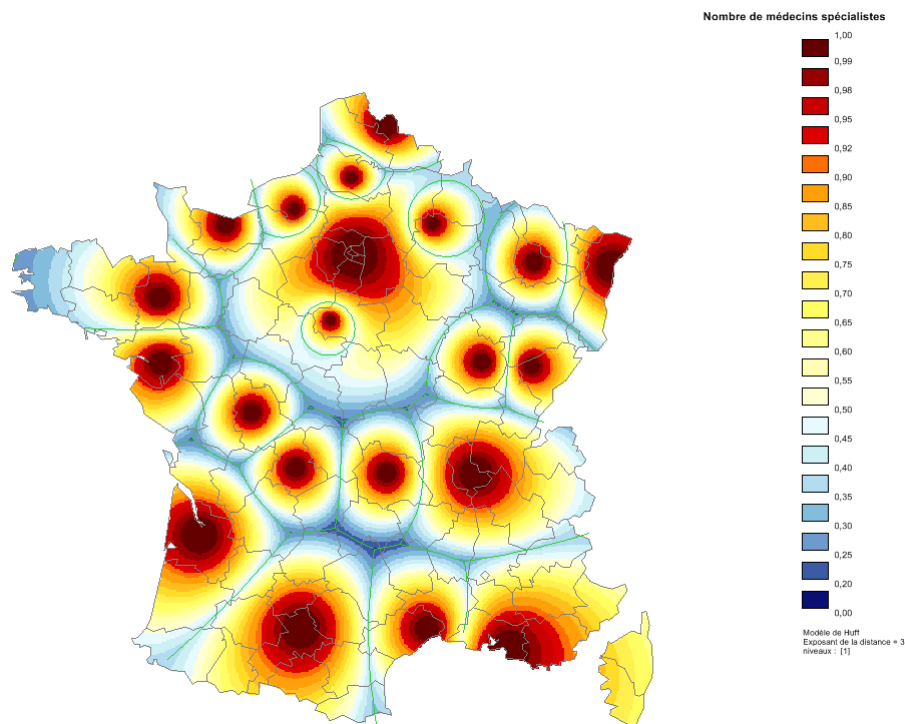


Figure 15.12 : ... et la nouvelle carte (exposant = 3).

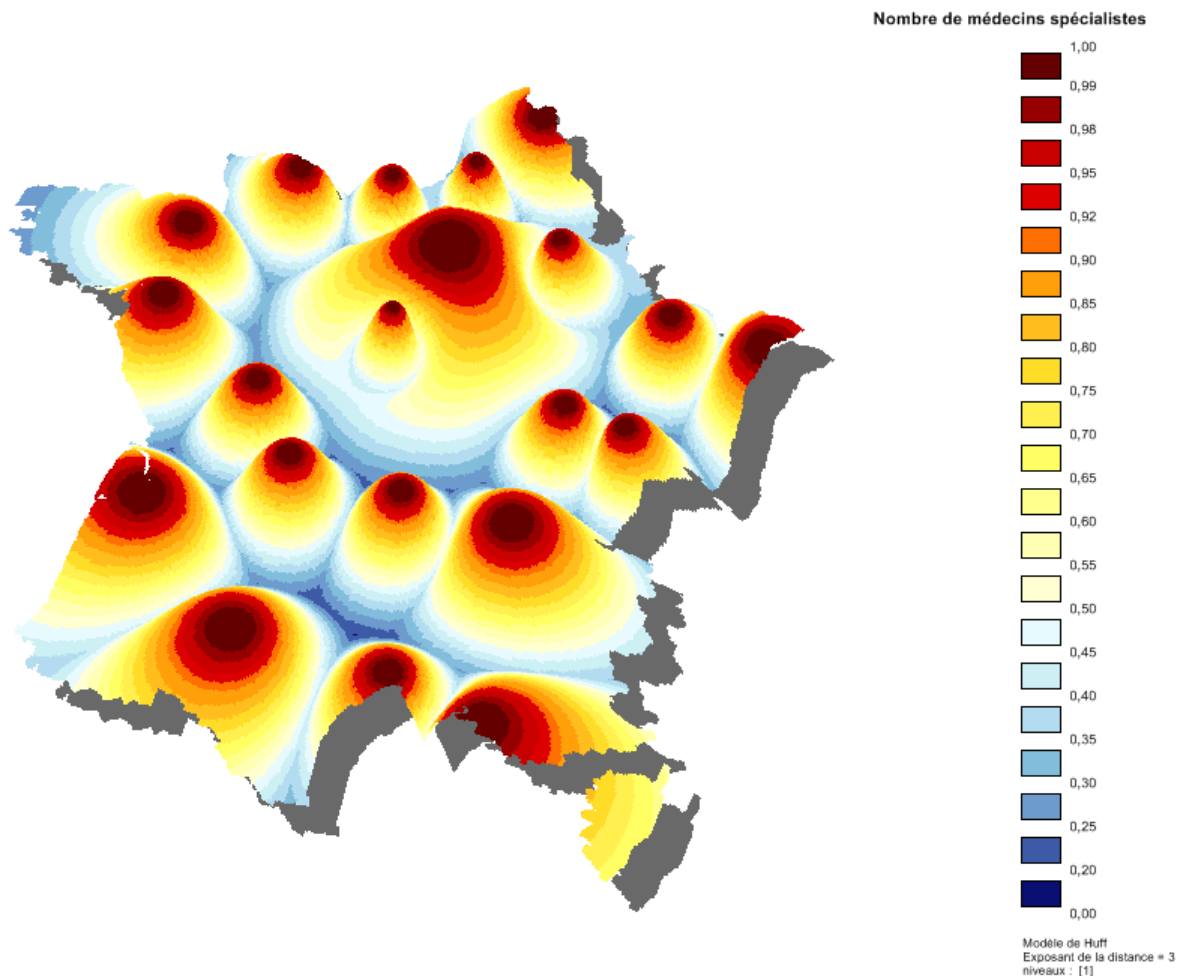


Figure 15.13 : Le bloc diagramme des influences des capitales régionales (exposant = 3) vu vers le Nord-Ouest.

15.4. Le tracé des polygones de Thiessen

Dans le modèle de Reilly, lorsque la masse est une constante, c'est-à-dire, lorsque tous les centres ont la même masse, égale à 1 par exemple, le tracé des limites d'influences prend la forme bien connue des polygones de Thiessen (dits aussi de Voronoï ou de Dirichlet). Ces polygones découpent l'espace de manière telle que chaque point situé à l'intérieur d'un polygone est plus proche du centre du polygone que de l'un quelconque des autres centres de polygones. Les polygones de Thiessen produisent ainsi une partition optimale de l'espace et sont utilisés dans de nombreux domaines tels l'hydrologie, l'écologie... et le géomarketing.

Selon qu'on applique ou non un habillage, les polygones tracés par *Philcarto* sont limités par le cadre de la carte ou par la bordure extérieure de l'habillage (fig. 15.14).

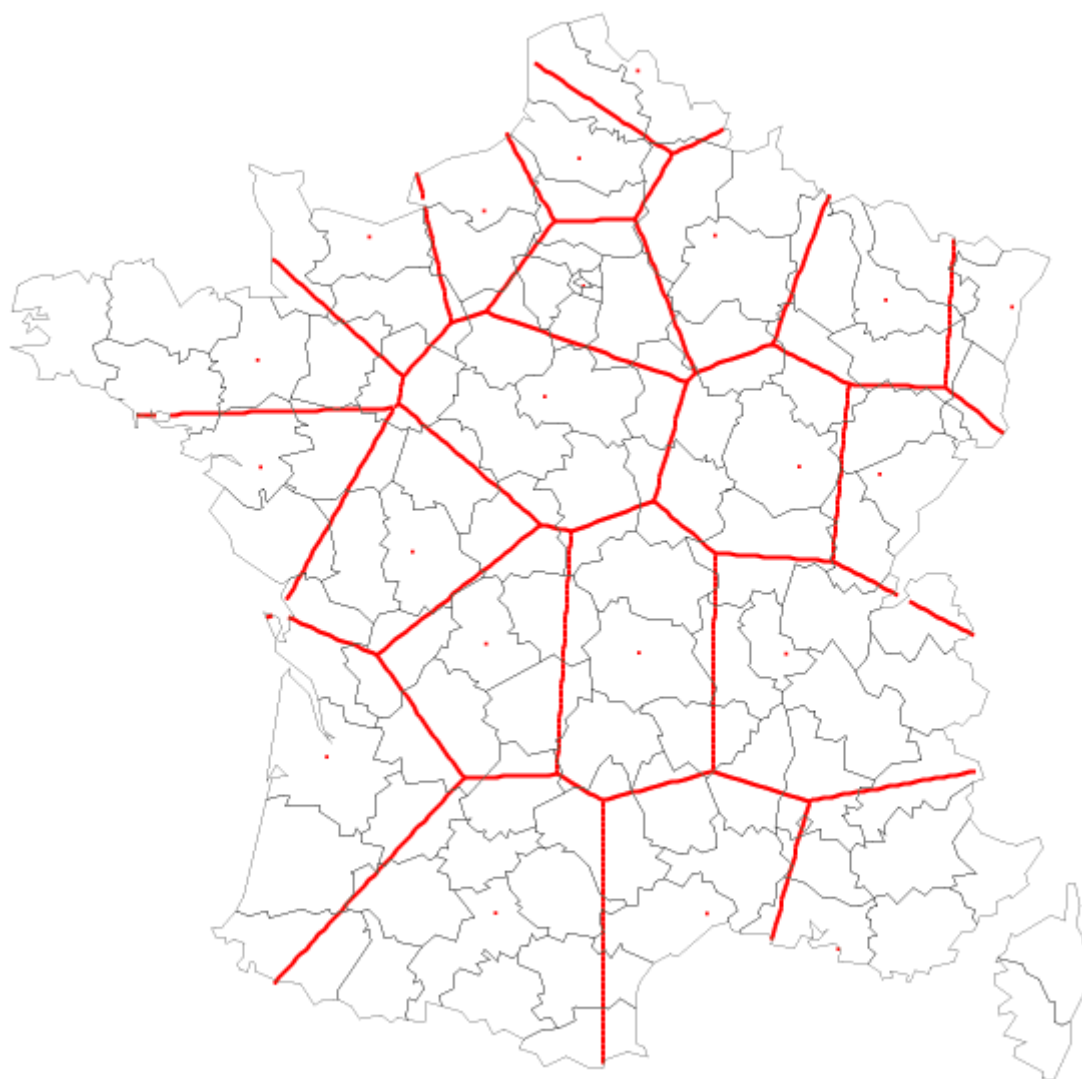


Figure 15.14 : Les polygones de Thiessen des capitales régionales.

16. L'analyse factorielle et la classification automatique

Dossier : \Philcarto exemples\Allemagne\
Fond de carte : Allemagne carte Kreise.ai
Statistiques : Allemagne data Kreise.xls

Les techniques d'analyse multivariée des données proposées par *Philcarto* sont essentiellement descriptives ; elles sont particulièrement efficaces sur des tableaux de données de grande dimension, et n'imposent aucune hypothèse sur les distributions statistiques des variables étudiées, ni le recours à un modèle probabiliste contraignant. Les méthodes rassemblées ici font partie de la boîte à outil du géographe depuis que les ordinateurs ont rendu possibles les calculs lourds et répétitifs qu'implique leur mise en œuvre ; certains géographes en ont même fait des outils privilégiés de l'étude de la différenciation spatiale qui, par essence, résulte d'un complexe de variables. La compréhension des mécanismes en action passe par l'identification des combinaisons de variables qui s'établissent dans l'espace. On cherche ainsi à réduire le « bruit », c'est-à-dire de petites variations considérées comme peu significatives pour mieux souligner les interactions et les associations.

Dans le cadre **Outils**, le bouton **MULTIV** donne accès à des méthodes largement utilisées par les géographes depuis la « révolution quantitative » des années 1960-1970, qui, partie des États-Unis (voir les travaux de Berry, Marble, etc.), a progressé en France dans le courant des années 1970-1980 (voir les travaux de D. Pumain, T. Saint-Julien, V. Rey, à Paris, F. Auriac, M. Vigouroux, J. Charre, M. Le Berre et des autres membres du Groupe Dupont à Avignon).

Le premier groupe de techniques comprend les méthodes dites « factorielles » qui, « derrière le foisonnement des données [dégagent] les quelques facteurs qu'elles représentent et qui expliquent, au sens statistique du terme, l'essentiel de la variation analysée » [Claval ; 1977]. Les méthodes factorielles se décomposent en plusieurs techniques qui diffèrent en fonction de la nature des données à analyser, et des objectifs assignés à l'analyse. Les deux techniques les plus couramment utilisées en géographie et qui figurent dans *Philcarto* sont :

- l'Analyse en Composantes Principales (ACP) dont les bases ont été jetées par Hotelling en 1933, et qui a commencé à être utilisée en sciences humaines dans les années 1960, en psychologie notamment. Cette technique s'applique essentiellement aux tableaux de mesures ;

- l'Analyse Factorielle des Correspondances (AFC) dont les principes théoriques remontent aux travaux de Hartley et Fisher au cours des années 1940, et qui a été exposée pour la première fois sous ce nom en 1963 par J.-P. Benzécri. Elle s'applique principalement aux tableaux de correspondances, mais permet également, à la condition de subtils recodages, d'analyser des tableaux de données nominales.

Le second ensemble de techniques produit des classifications automatiques. Dans *Philcarto*, c'est la Classification Ascendante Hiérarchique (CAH) qui permet de construire rigoureusement des typologies emboîtées souvent considérées, par les géographes, comme des synthèses régionales. La CAH s'applique aussi bien aux tableaux de mesures qu'aux tableaux de correspondances.

L'exposé de ces différentes méthodes sort du cadre du présent manuel. Mais il est impossible de les mettre en pratique sans en comprendre les fondements mathématiques, finalement assez simples pour peu qu'on ne cherche pas à réaliser les calculs sans l'aide d'un ordinateur...

Pour aborder les principes de l'Analyse des Données, on ne saurait trop conseiller la lecture du remarquable et magnifique (et plus encore !) ouvrage de J.-P. Fénelon « Qu'est-ce que l'Analyse des Données » [Fénelon ; 1981]. Pour un exposé plus orienté vers la géographie, l'ouvrage de L. Sanders « L'analyse statistique des données en géographie » constitue une très bonne référence [Sanders ; 1990].

Pour ce qui concerne plus spécifiquement la classification en géographie, la lecture de l'ouvrage de P. Dumolard « L'espace différencié » est très enrichissante [Dumolard ; 1981] .

Comme on l'a compris, la mise au point des méthodes d'analyse multivariées a fait l'objet de très nombreuses recherches tant chez les mathématiciens et statisticiens que chez les utilisateurs de ces techniques. Ainsi, on trouve sur Internet une grande variété de sites se rapportant à ce domaine. C'est là que j'ai puisé les codes sources de l'ACP et de l'AFC que j'ai adaptés à *Philcarto* à partir des programmes FORTRAN que F. Murtagh, professeur associé à l'Observatoire Astronomique de Strasbourg, met à la disposition des chercheurs.

Pour la CAH, j'ai utilisé le code source FORTRAN du sous-programme AGREGA publié par M. Jambu et M.-O. Lebeaux dans leur ouvrage « Classification automatique pour l'analyse des données, Tome 2 : logiciels ». Ce sous-programme met en œuvre un algorithme original en deux étapes qui réduit considérablement le temps nécessaire à l'élaboration de la hiérarchie, ce qui rend possible l'utilisation de la CAH sur quelques milliers d'individus.

16.1. Choisir une méthode d'analyse multivariée

Après avoir cliqué sur **MULTIV** du cadre **Outils**, le nouveau dialogue **Analyse des données** s'affiche dans ce cadre. Il contient un ensemble de boutons permettant de sélectionner un type d'analyse multivariée (fig. 16.1).

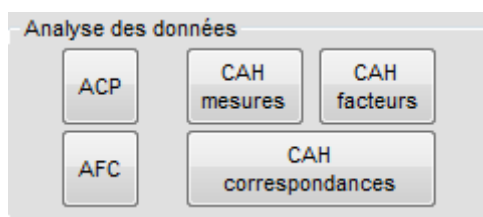


Figure 16.1 : Le dialogue **Analyse des données** .

Les deux boutons situés dans la partie gauche du dialogue **Analyse des données** permettent de choisir entre l'Analyse en Composantes Principales **ACP** qui s'applique à un tableau de mesures et l'Analyse Factorielle des Correspondances **AFC** qui s'applique à un tableau de correspondances.

Les trois autres boutons, situés à droite, donnent accès aux différentes variantes de la Classification Ascendante Hiérarchique **CAH** sur un tableau de mesures, sur un tableau de facteurs et sur un tableau de correspondances.

Dans le cas d'un tableau de mesures, l'ACP ou la CAH procèdent préalablement à un centrage et à une réduction des variables (moyenne de chaque variable = 0, écart-type de chaque variable = 1).

Dans le cas d'un tableau de correspondances, l'AFC ou la CAH mesurent les distances entre profils selon la métrique du Khi^2 .

Dans d'une CAH sur un tableau de facteurs, ceux-ci sont laissés tels quels : ils sont centrés par définition et conservent leur propre variance (moyenne de chaque facteur = 0, écart-type = racine carré de la valeur propre associée à chaque facteur).

16.2. Réaliser une ACP

Un clic sur **ACP** provoque l'affichage d'une liste de variables dans le dialogue **Analyse des données** (fig. 16.2). On trouve dans cette liste les variables de couleurs, c'est-à-dire des rapports, pourcentages, taux, indices... à l'exclusion de quantités ou d'effectifs. Si l'on souhaite introduire des effectifs dans l'ACP (ce qui n'est pas recommandé), il faut changer le type des variables concernées.

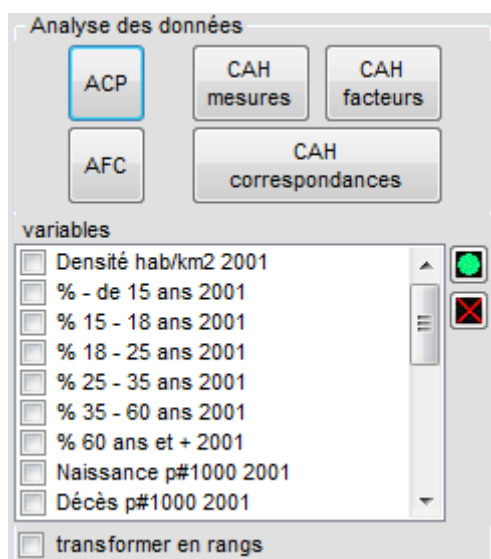


Figure 16.2 : Le dialogue **Analyse des données** après le choix d'une ACP

La sélection des variables de l'analyse se fait en cochant les cases figurant en regard des noms des variables dans la liste **variables**. Un clic sur  sélectionne toutes les variables (à utiliser avec précaution, une vérification de la liste des variables sélectionnées s'imposant toujours), alors que  sert à désélectionner toutes les variables. L'ascenseur situé à droite de la liste permet de faire défiler la liste si elle est trop longue pour être contenue dans le cadre.

Si la case ☐ **transformer en rangs** est décochée (option par défaut), l'ACP est réalisée par diagonalisation d'une matrice d'inertie composée de coefficients de corrélation de Pearson. Si elle est cochée, les variables sont préalablement transformées en rangs et la diagonalisation s'applique sur une matrice d'inertie composée de coefficients de corrélation de rangs de Spearman ; le rang le plus élevé sur chaque variable de l'analyse correspond à la valeur maximum de chaque variable.

Sélectionnons les variables correspondant à la répartition par classes d'âges de la population, et lançons l'ACP par un clic sur **Calculer** (fig. 16.3).

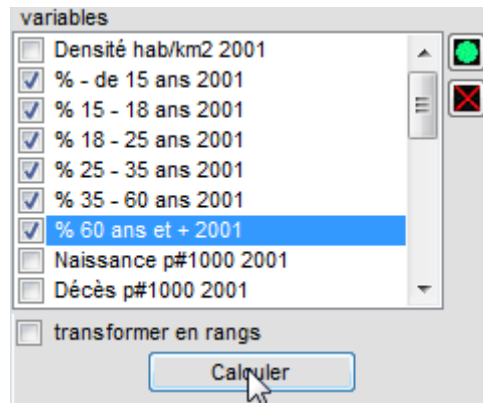


Figure 16.3 : Le choix des variables de l'ACP et le lancement du calcul.

Après quelques secondes d'attente la fenêtre **Analyse multivariée** s'affiche à l'écran (fig. 16.4).

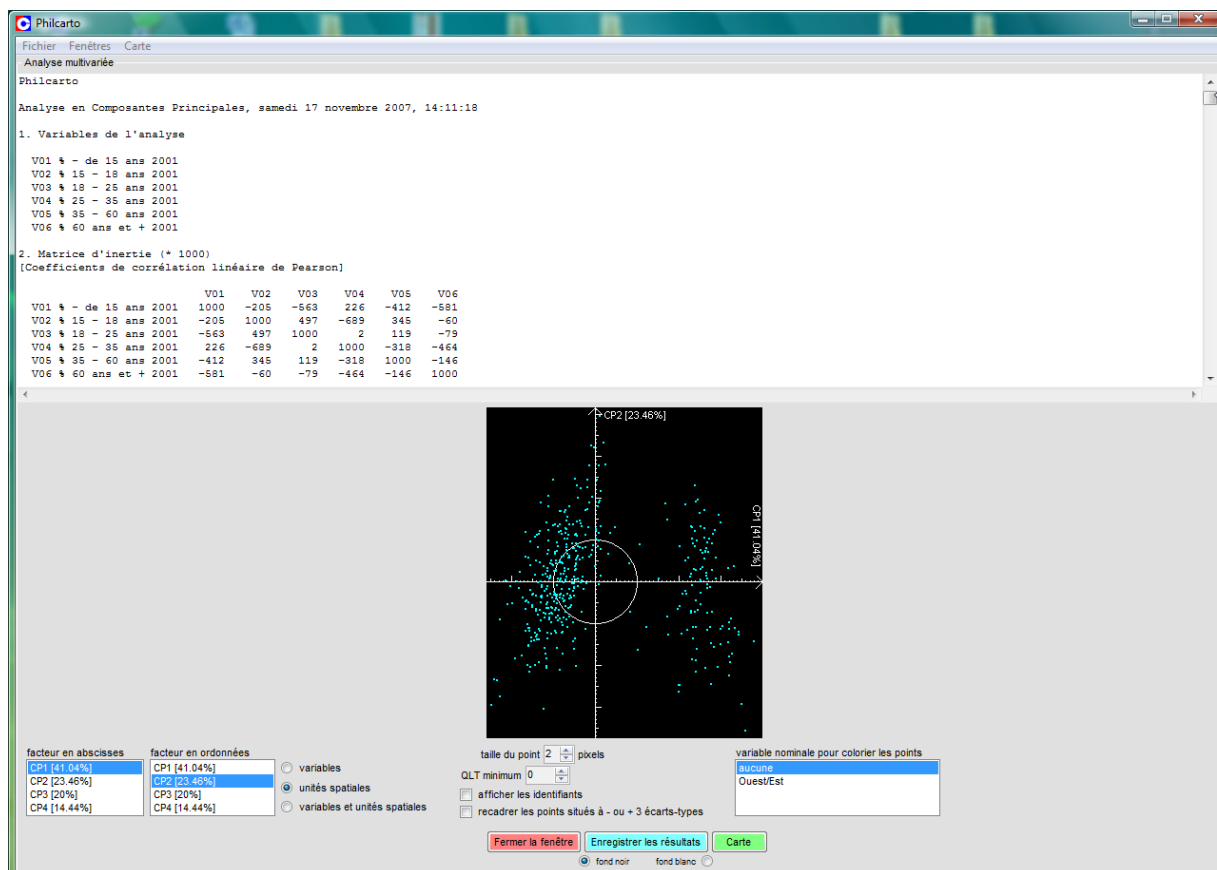


Figure 16.4 : La fenêtre **Analyse multivariée** dans le cas d'une ACP.

Cette fenêtre comprend deux parties principales :

- dans la moitié supérieure, on trouve le *listing* de l'ACP avec son contenu classique : liste des variables, matrice d'inertie, résultat de la diagonalisation (histogramme des valeurs propres), corrélations des variables avec les composantes principales, communautés (CO^2), contributions relatives des variables (CTR), coordonnées des unités spatiales sur les composantes principales, contributions relatives des

unités spatiales aux composantes (CTR), qualités de représentation des unités spatiales sur les composantes (CO²).

- dans la partie inférieure figure un graphique de composantes principales prises deux à deux dont les caractéristiques changent en fonction des choix de l'utilisateur : choix des facteurs en abscisses et en ordonnées (fig. 16.5), éléments à représenter (fig. 16.6), caractéristiques des points figurant sur le graphique (fig. 16.17) : taille du point et qualité de représentation minimale pour l'affichage réglables à l'aide d'un curseur, affichage des identifiants de chaque point (déconseillé lorsque le nombre de points est grand), recadrage dans un intervalle compris entre -3 et +3 écarts-types pour les points qui sont extérieurs à ces limites.

facteur en abscisses	facteur en ordonnées
CP1 [41.04%]	CP1 [41.04%]
CP2 [23.46%]	CP2 [23.46%]
CP3 [20%]	CP3 [20%]
CP4 [14.44%]	CP4 [14.44%]

Figure 16.5 : Le choix des facteurs du graphique en abscisses et en ordonnées.

☐ variables
☒ unités spatiales
☐ variables et unités spatiales

Figure 16.6 : Le choix des éléments à représenter sur le graphique.

taille du point pixels
 QLT minimum
☐ afficher les identifiants
☐ recadrer les points situés à - ou + 3 écarts-types

Figure 16.7 : Le choix des caractéristiques des points du graphique.

En déplaçant le pointeur de la souris sur un point, on peut lire son identifiant, son nom, sa qualité de représentation cumulée sur les deux composantes constituant le graphique, et les valeurs du couple des coordonnées (fig. 16.8).

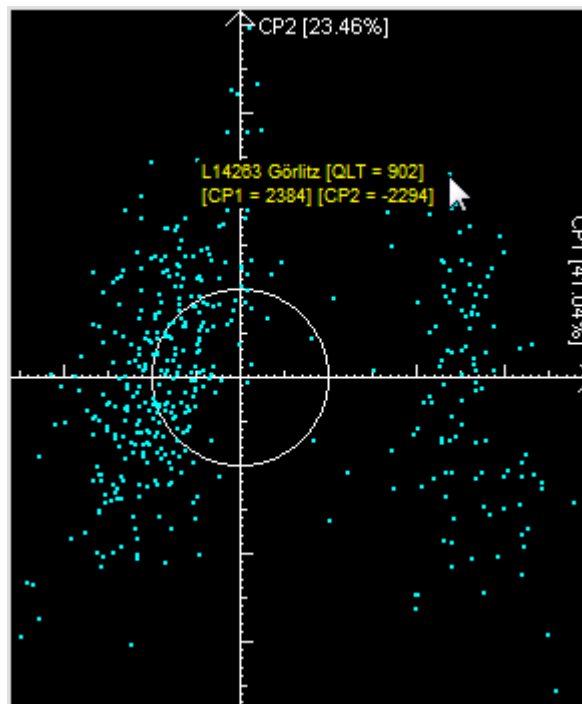


Figure 16.8 : Identifier un point sur le plan factoriel.

Lorsque les résultats de l'ACP ont fait l'objet d'une première interprétation, il est possible de visualiser les cartes des coordonnées sur chaque composante principale par un clic sur **Carte**.

L'utilisateur se retrouve alors tout simplement face au dialogue **Variables** du cadre **Outils** (fig. 16.9). Dans la liste **couleurs** du dialogue **Variables**, les noms des composantes principales calculées ont été ajoutées à la liste des variables initiales ; le nom de chacune de ces nouvelles variables indique son rang dans l'analyse (CP1, CP2...), le taux d'inertie qu'elle a absorbé suivi du mot **Coordonnées** qui indique qu'il s'agit bien des coordonnées des unités spatiales sur la composante ainsi dénommée. Dans la liste **symboles**, on trouve aussi de nouvelles variables, dénommées de la même façon à une différence près : le mot **Coordonnées** a été remplacé par le mot **Qualités** qui indique que dans ces nouvelles variables sont enregistrées les qualités de représentation.

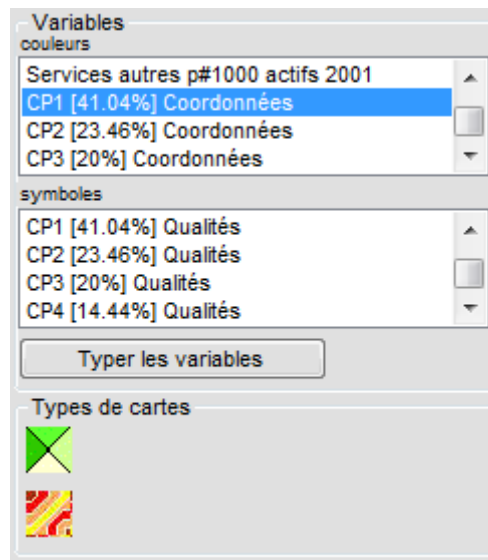


Figure 16.9 : Les listes de variables **couleurs** et **symboles** mises à jour après l'ACP.

Lorsque l'affichage des listes de variables est consécutif à une ACP, la variable **CP1...Coordonnées** est automatiquement sélectionnée. En revanche, la variable **CP1...Qualités** n'est pas sélectionnée et, si nécessaire, la liste doit faire l'objet d'un défilement pour atteindre cette variable-là. Si l'on souhaite réaliser une carte choroplèthe des coordonnées sur CP1, il suffit de cliquer sur l'icône . La carte s'affiche immédiatement dans une configuration bien adaptée à l'ACP : discrétisation  et gamme de couleurs en opposition (bleu pour les valeurs négatives, rouge pour les valeurs positives, et gris pour les valeurs proches de 0). Après habillage, le résultat obtenu est très structuré dans l'espace (fig. 16.10) :

à l'Est, l'ancienne RDA ne comporte que des couleurs jaune-orange (sauf Berlin qui est en gris ;

à l'Ouest l'ancienne RFA est majoritairement en bleu clair au Nord et au Sud, alors que sa partie centrale est grise.

Au lieu d'une carte choroplèthe, on peut aussi bien réaliser une carte en surface de tendance en sélectionnant l'icône  (fig. 16.11) ; la surface de tendance d'ordre 4 traduit efficacement l'organisation spatiale de cette première composante dont elle représente d'ailleurs 78% de la variance (mais l'estimation faite pour Berlin est mauvaise, ce qu'une carte des résidus permettrait de mettre en évidence).

Une autre possibilité de cartographie consiste à combiner qualité de représentation représentée par des cercles proportionnels et coordonnées sur les composantes représentées, une fois discrétisées, par des couleurs (fig. 16.12). Cela permet d'éviter d'interpréter la position des unités spatiales mal représentées, notamment sur les composantes de rang supérieur à 1 ou 2.

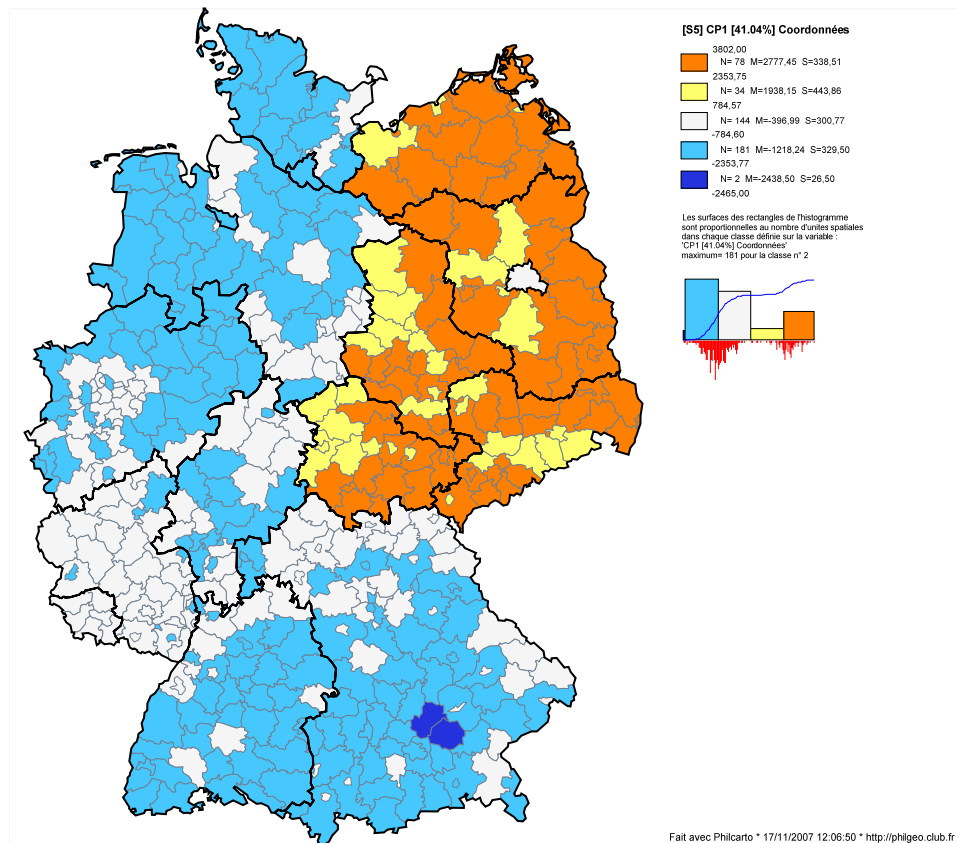


Figure 16.10: La carte des coordonnées sur la première composante principale.

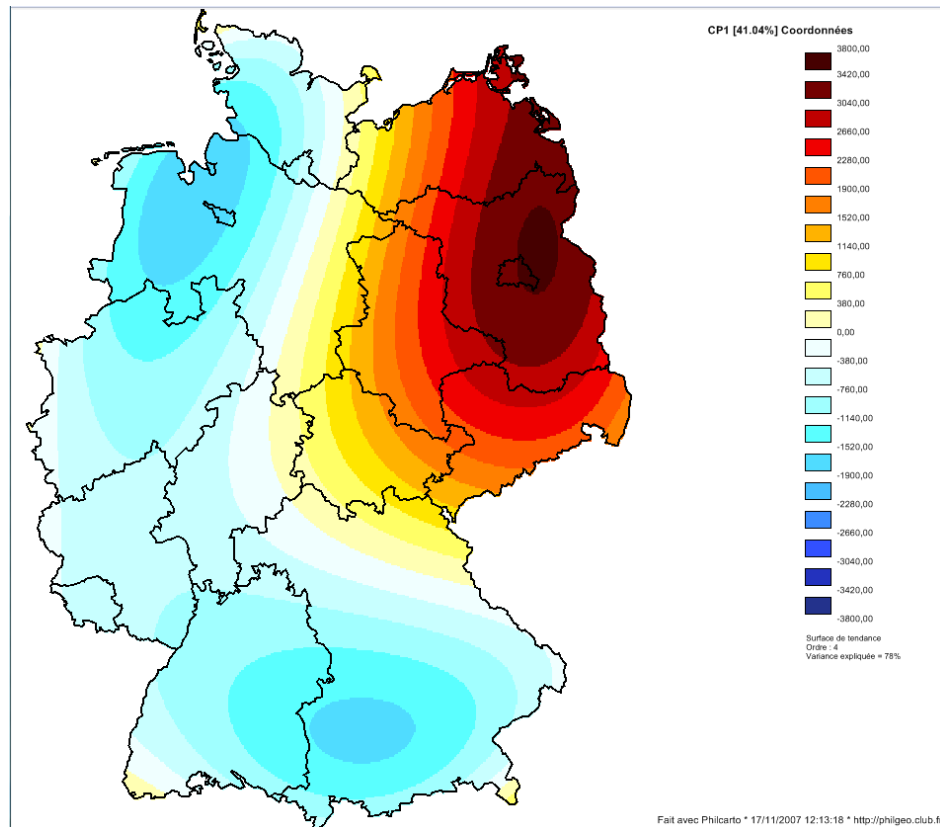


Figure 16.11 : La carte en surface de tendance d'ordre 4 de la première composante principale.

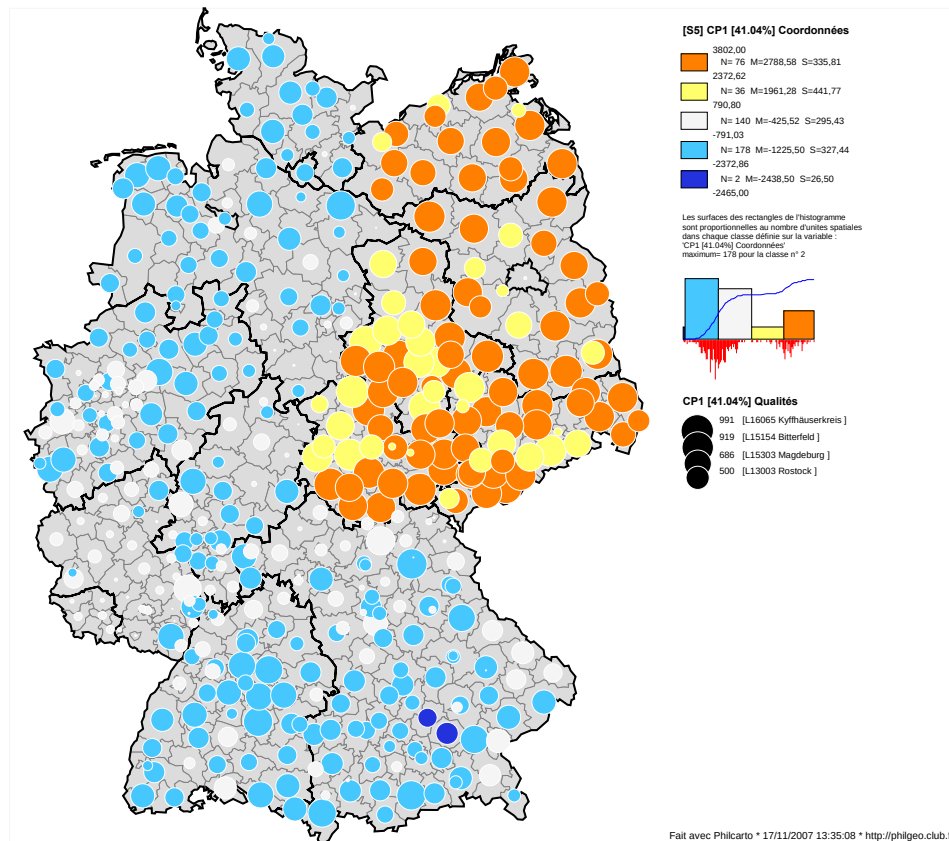


Figure 16.12 : La carte combinant qualité de représentation et coordonnées sur la composante principale n°1.

Lorsque la carte d'une composante est tracée, il est souvent nécessaire de revenir à la fenêtre contenant les résultats de l'analyse. Cela se fait de manière simple en activant l'article **Résultats ACP** du menu **Fenêtres** (fig. 6.13), ou en tapant la combinaison de touches du clavier Ctrl 1.

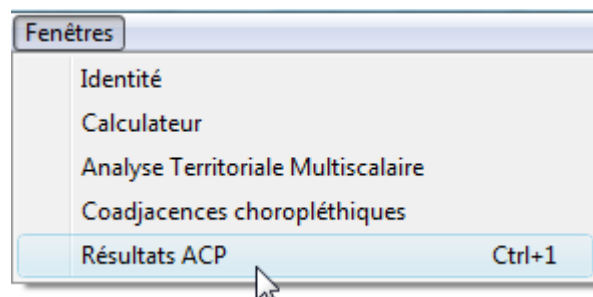


Figure 16.13 : Afficher à nouveau les résultats de l'ACP.

La fenêtre **Analyse multivariée** masque alors la carte. Dans la moitié inférieure de cette fenêtre, on trouve des fonctions qui n'ont pas encore été décrites.

Le bouton **Fermer la fenêtre** permet d'afficher à nouveau la carte en cours.

Le bouton **Enregistrer les résultats** permet d'enregistrer la totalité des résultats de l'analyse dans un dossier unique ; de manière optionnelle, les graphiques enregistrés peuvent être dessinés en blanc sur fond noir ou en noir sur fond blanc selon que l'un ou l'autre des deux boutons radio **fond noir** **fond blanc** est activé. Un clic sur

Enregistrer les résultats ouvre le dialogue **Rechercher un dossier** qui permet de sélectionner ou de créer un dossier destiné à contenir les différents fichiers de résultats. Soit le dossier existe et dans ce cas, il suffit de le sélectionner là où il se trouve et de valider par **OK** ; soit il n'existe pas encore et clic sur **Créer un nouveau dossier** ouvre une zone d'édition dans laquelle on peut écrire le nom du dossier créé (fig. 16.14).

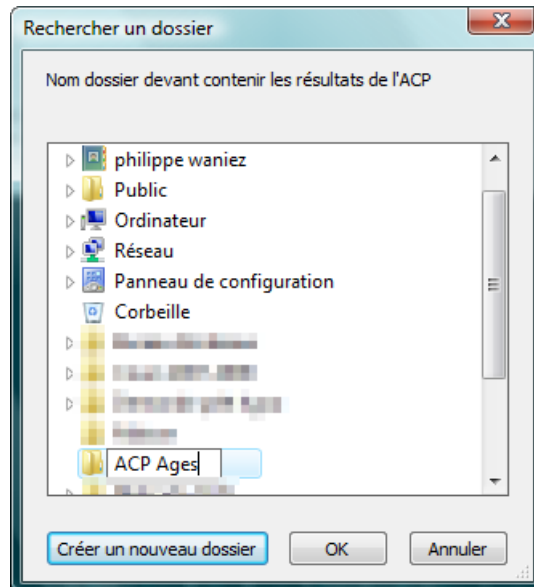


Figure 16.14 : Créer un nouveau dossier de résultats d'ACP.

La validation à l'aide de **OK** ; provoque l'enregistrement de tous les résultats de l'ACP, c'est-à-dire, pour chaque couple de composantes principales, le graphique des observations (unités spatiales), celui de variables, et celui des observations et des variables, soit 3 graphiques par couple de composantes. Ces graphiques sont au format **.emf**. A ces fichiers s'ajoute un fichier texte contenant le *listing* de l'ACP (fig. 16.15).

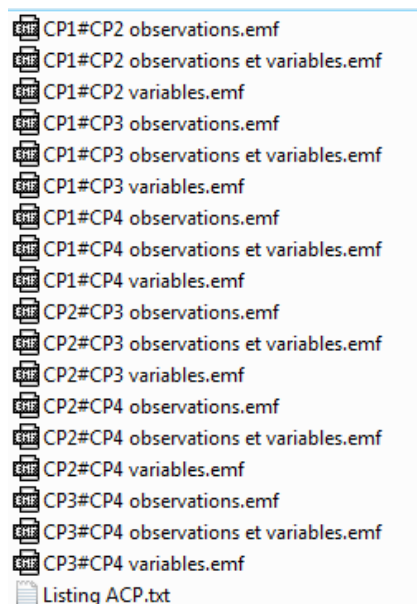


Figure 16.15 : Un exemple d'enregistrement de résultats d'ACP dans un dossier.

Si le fichier de données contient des variables nominales, leurs noms apparaissent dans la liste `variable nominale pour colorier les points`. Dans notre fichier, par exemple, une variable nominale dénommée Ouest/Est prend les valeurs RFA pour les Kreise situés dans l'ancienne Allemagne de l'Ouest, et RDA s'il s'agit de Kreise des nouveaux Länder de l'Est (fig. 16.16). Le graphique colorié en fonction de ces deux modalités apporte une clé de compréhension de la composante n°1 située en abscisse (fig. 16.17) : les Kreise de l'Ouest apparaissant en bleu et majoritairement du côté des coordonnées négatives, alors que les Kreise de l'Est, en rouge sur le graphique, figurent presque tous du côté positif de la composante n°1.

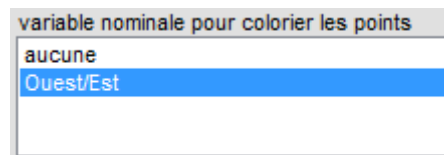


Figure 16.16 : La sélection de la variable nominale Ouest/Est pour colorier les points.

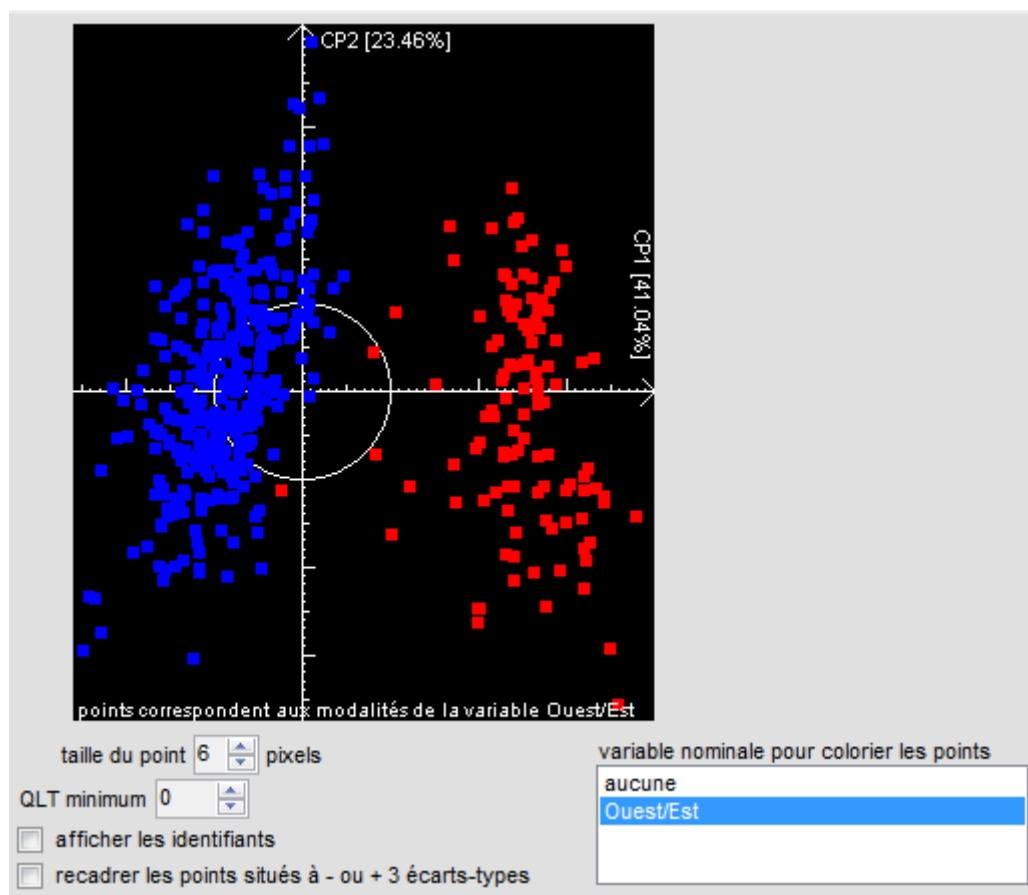


Figure 16.17 : Les Kreise de l'Ouest et les Kreise de l'Est se séparent en deux groupes distincts sur la première composante principale.

Dans ce genre de situation, c'est-à-dire lorsqu'il existe une césure nette des observations en deux sous-ensembles, il peut être intéressant de réaliser l'ACP sur l'un de ces sous-ensembles, puis de plonger les observations de l'autre sous-ensemble dans la structure révélée par l'ACP sur le premier. Cette manière de faire implique que les observations du premier sous-ensemble, les Kreise de la RFA, par

exemple, soient considérées comme des éléments principaux et codés 0 ; à l'inverse, les observations du second sous-ensemble, les Kreise de la RDA sont considérés comme des éléments supplémentaires et sont codés 1. Dans notre fichier de données, c'est la variable Rôle qui contient ces codes. Ce genre de variable binaire, si elle existe dans le tableau de données, est automatiquement détectée lors de la lecture du fichier de données. Lorsqu'une ACP est demandée (par un clic sur **ACP**), le nom de la ou des variables binaires s'inscrit dans le dialogue **variable binaire** (fig. 16.18) où il est possible de sélectionner l'une ou l'autre de ces variables (ici une seule variable dénommée **ObservationsSupplémentaires**).

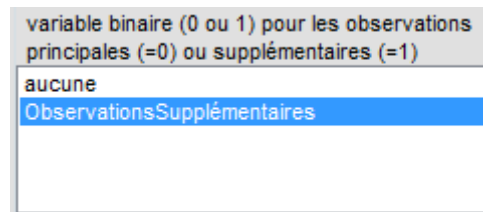


Figure 16.18 : Le dialogue **variable binaire** pour sélectionner les observations principales et supplémentaires.

Les résultats de l'ACP sont naturellement différents comme en témoigne la forme du nuage de point formé des observations principales et supplémentaires sur les deux premières composantes (fig. 16.19).

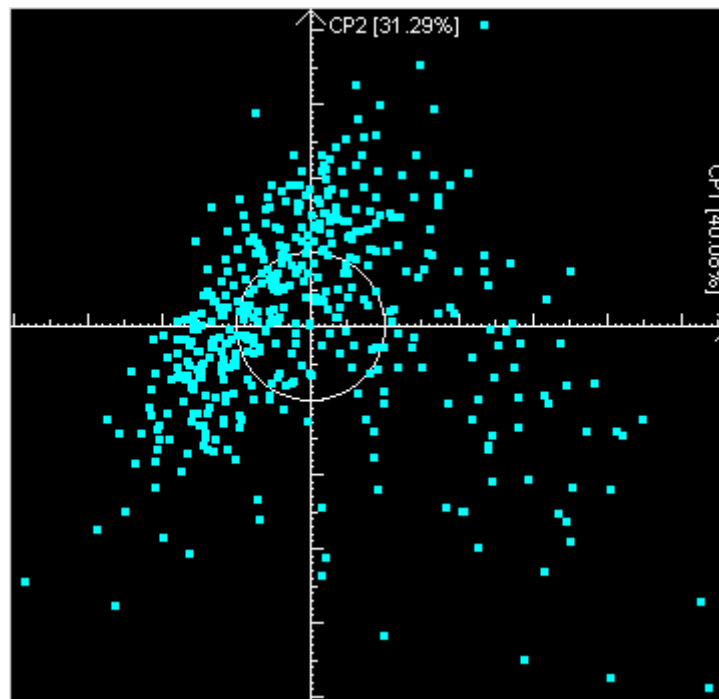


Figure 16.19 : Le nouveau plan factoriel basé sur les données des seuls Kreise de RFA (mais la totalité des Kreise, RFA et RDA sont représentés).

En sélectionnant la variable Ouest/Est dans la liste **variable nominale pour colorier les points**, on observe, en bleu les Kreise principaux de RFA et en rouge les Kreise supplémentaires de RDA (fig. 16.20). Notons qu'il y aurait lieu de s'interroger sur la forme en boomerang du nuage de points.

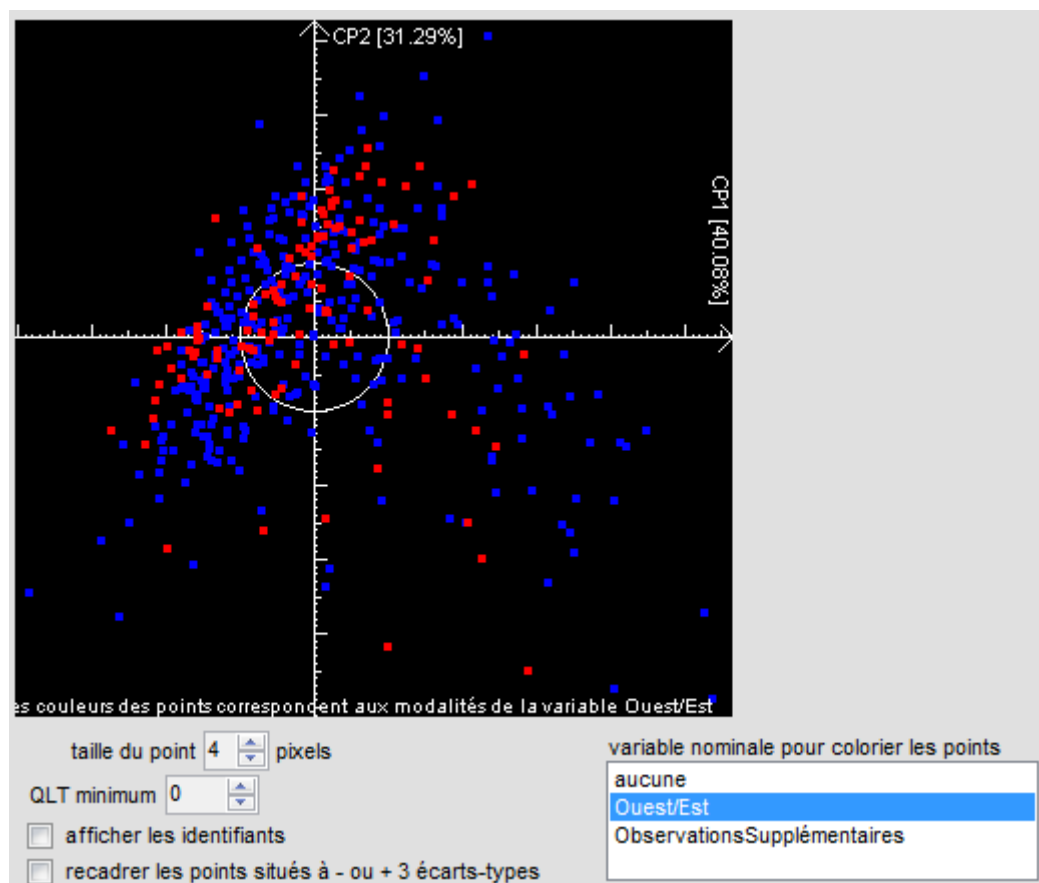


Figure 16.20 : Le nouveau plan factoriel basé sur les données des seuls Kreise de RFA (en bleu les Kreise principaux de RFA et en rouge les Kreise supplémentaires de RDA).

Dans le listing de l'ACP, aux chapitres des coordonnées et des contributions des observations, les identifiants des observations supplémentaires sont suivis du caractère * (fig. 16.21) . Notons aussi que la contribution des observations supplémentaires est toujours nulle (fig. 16.22), ce qui est parfaitement logique puisque la mise en observation supplémentaire consiste précisément à ne pas les laisser participer au calcul des composantes principales.

L10044	Saarlouis	136	1368	519
L10045	Saarpfalz-Kreis	467	1784	283
L10046	St, Wendel	-99	1338	173
L11000*	Berlin	2195	-1281	1782
L12051*	Brandenburg an der Havel	1096	1524	-959
L12052*	Cottbus	864	-1788	-776

Figure 16.21 : Les coordonnées des observations principales et supplémentaires ; noter la présence du caractère * à la suite des identifiants des observations supplémentaires.

L10044	Saarlouis	0	3	1
L10045	Saarpfalz-Kreis	0	5	0
L10046	St, Wendel	0	3	0
L11000*	Berlin	0	0	0
L12051*	Brandenburg an der Havel	0	0	0
L12052*	Cottbus	0	0	0

Figure 16.22 : Les contributions des observations principales et supplémentaires ; noter les valeurs nulles des contributions des observations supplémentaires.

16.3. Réaliser une CAH

Après avoir cliqué sur **MULTIV** du cadre **Outils**, un clic sur **CAH mesures** du dialogue **Analyse des données** provoque l'affichage d'une liste de variables. On trouve dans cette liste les variables de couleurs, c'est-à-dire des rapports, pourcentages, taux, indices... à l'exclusion de quantités ou d'effectifs. Si l'on souhaite introduire des effectifs dans l'ACP (ce qui n'est pas recommandé), il faut changer le type des variables concernées. Se reporter à l'ACP pour la présentation du fonctionnement de ce dialogue (§16.2).

Sélectionnons les variables correspondant à la répartition par classes d'âges de la population, et lançons la CAH par un clic sur **Calculer** (fig. 16.23).

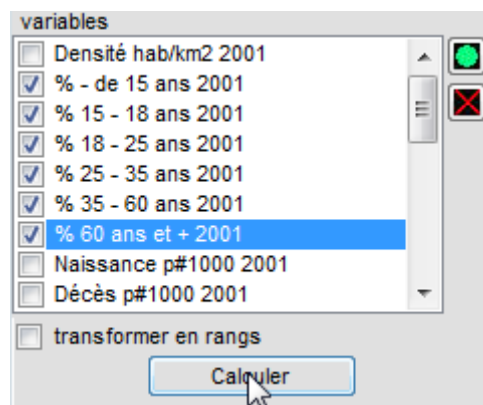


Figure 16.23 : Le choix des variables de l'ACP et le lancement du calcul.

Après quelques secondes d'attente la fenêtre **Analyse multivariée** s'affiche à l'écran (fig. 16.24). Cette fenêtre comprend deux parties principales :

- dans la moitié supérieure, on trouve le *listing* de la CAH avec son contenu classique : liste des variables, histogramme des indices de niveaux de la hiérarchie, paramètres des classes de chaque partition, c'est-à-dire, pour chaque variable : Nombre d'unités spatiales, Minimums et Maximums, Moyennes, Ecart-types, Coefficients de variation, Distances aux moyennes (dans l'unité de chaque variable), Distances aux moyennes (en écart-type de chaque variable), Interprétation des distances aux moyennes.
- dans la partie inférieure l'arbre de la hiérarchie, dessiné horizontalement pour plus de commodité. La hiérarchie étant indicée par des indices de niveaux, l'arbre est gradué de 0 à la valeur de l'indice pour deux classes. Un trait vertical rouge coupe l'arbre à droite, juste au-dessous du dernier nœud, ce qui définit deux classes, celle qui s'est agrégée à l'ensemble à ce dernier nœud (en bleu à l'extrême droite de l'arbre), et celle qui contenait toutes les autres observations (en rouge).

Au-dessous de l'arbre, un curseur **nombre de classes** 2 permet de faire varier le nombre de classes (de 2 à 12 classes au plus) ; une modification du nombre de classes se solde par un déplacement du trait rouge vertical pour indiquer à quel niveau se situe la partition correspondante. Ici, avec 4 classes, ce trait rouge se situe à l'indice de niveau 0,49 soit 55,35% de l'inertie totale expliquée par ces 4 classes (fig. 16.25).

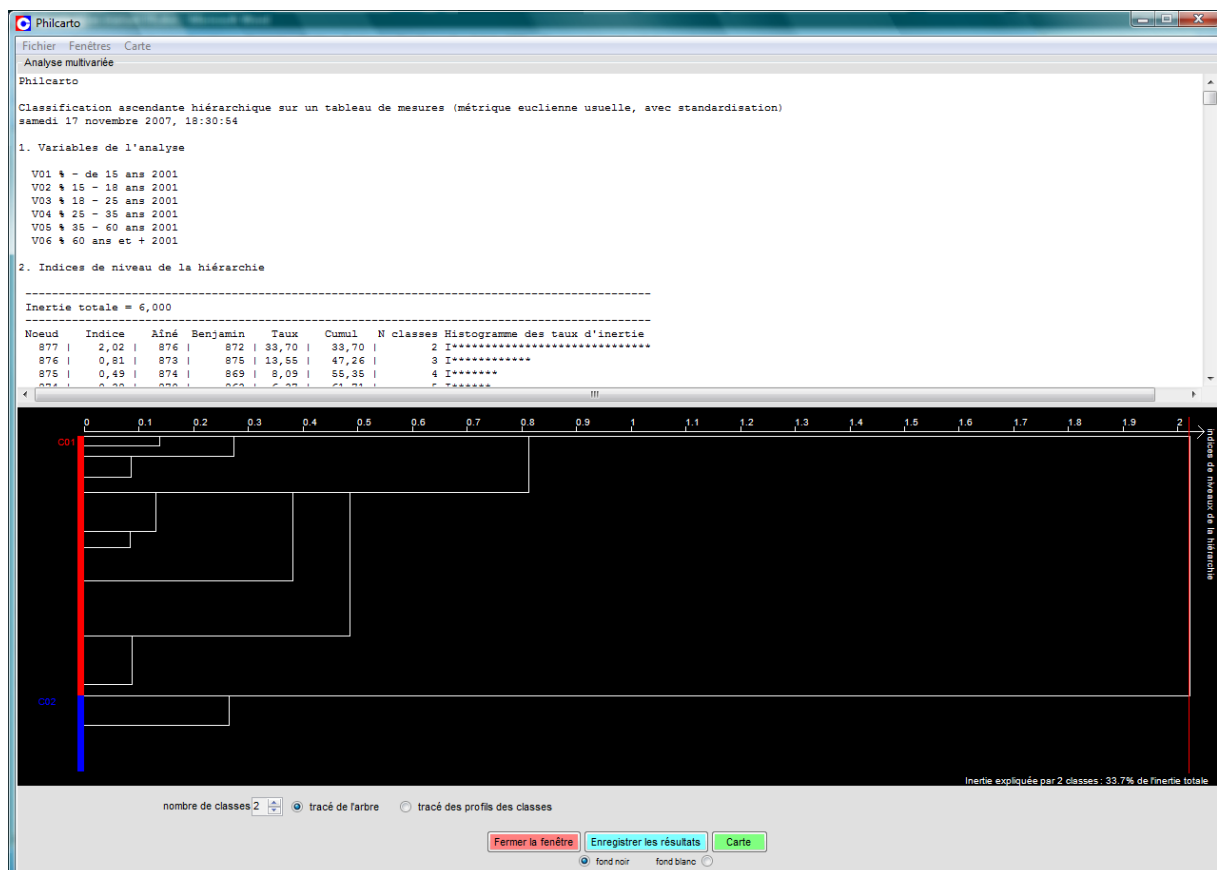


Figure 16.24 : La fenêtre **Analyse multivariée** dans le cas d'une CAH.

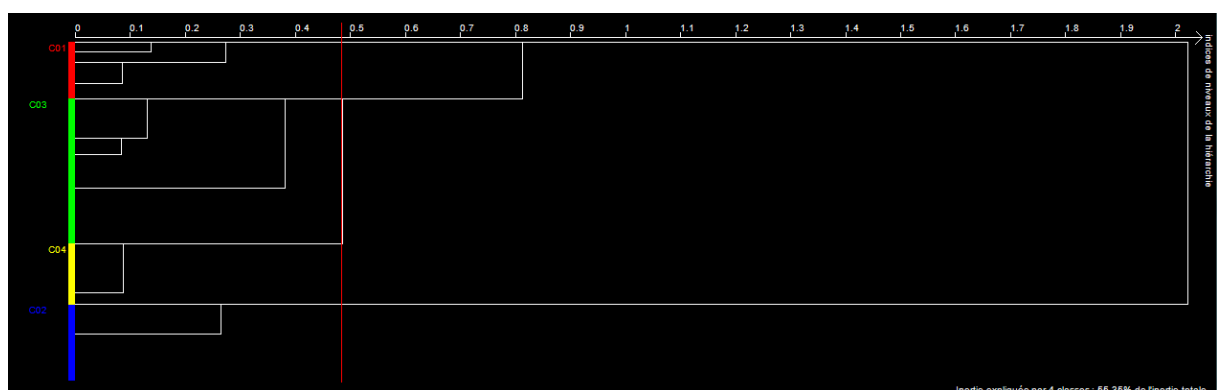


Figure 16.25 : La partition de l'arbre de classification en quatre classes.

Les boutons ☒ **tracé de l'arbre** ☐ **tracé des profils des classes** permettent d'afficher alternativement soit l'arbre, soit les profils des classes. Lorsque le profil des classes est sélectionné, un diagramme à bâtons par classe (fig. 16.26) est dessiné dans la couleur de la classe ; ce diagramme traduit les déviations, sur chaque variable, exprimées en écarts-types, des moyennes des classes par rapport à la moyenne générale.

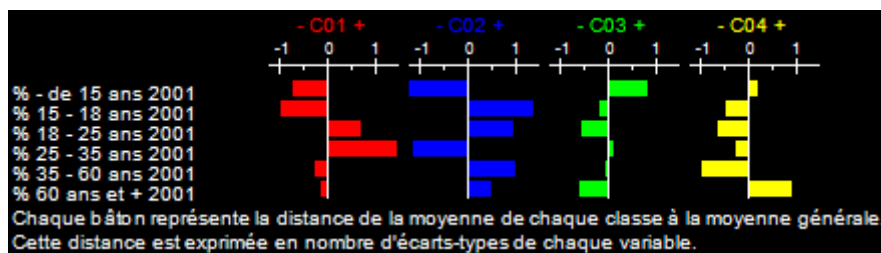


Figure 16.26 : Les diagrammes à bâtons des moyennes des classes par rapport aux moyennes générales.

Ces diagrammes à bâtons permettent une interprétation facile et rapide de chaque classe. Par exemple, la classe n°4 est caractérisée par une proportion anormalement élevée de personnes âgées, et corrélativement un déficit des autres classes d'âges (sauf pour les – de 15 ans situés très légèrement au-dessus de la moyenne). On aura donc affaire ici aux Kreise les plus « vieillis ». Les autres classes s'interprètent de la même façon. La classe n°1 est la classe des jeunes adultes (18-35 ans) ; la classe n°2 est celle du déficit des jeunes actifs de 25-35 ans et de la surreprésentation des adultes plus âgés et de leurs enfants adolescents (15-18 ans) ou étudiants (18-25 ans). Enfin la classe n°3 est caractérisée par la surreprésentation des enfants.

Lorsque les résultats de la CAH ont fait l'objet d'une première interprétation, il est possible de visualiser les cartes des classes par un clic sur [Carte](#).

L'utilisateur se retrouve alors tout simplement face au dialogue [Variables](#) du cadre [Outils](#) (fig. 16.27). Dans la liste [couleurs](#) du dialogue [Variables](#), les noms des partitions calculées ont été ajoutées à la liste des variables initiales ; le nom de chacune de ces nouvelles variables indique son rang dans l'analyse (Partition n°1, Partition n°2...) et cela pour les 11 partitions successives (12 classes) ; on trouve aussi dans le nom de chaque nouvelle variable, le taux d'inertie absorbé par la partition. La liste [symboles](#) reste inchangée.

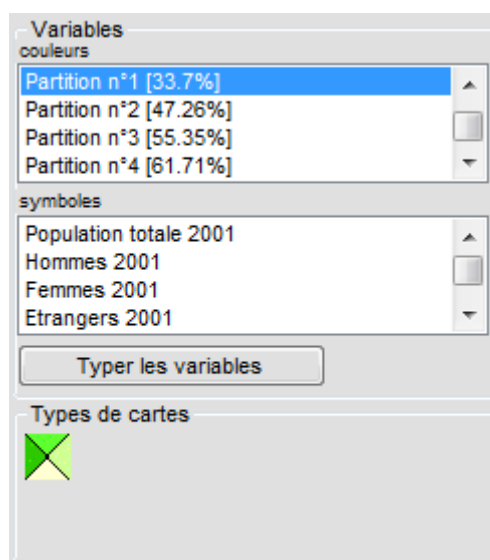


Figure 16.27 : Les listes de variables couleurs et [symboles](#) affichées après la CAH.

Lorsque l'affichage des listes de variables est consécutif à une CAH, la variable **Partition n°1 [...]** est automatiquement sélectionnée. Si l'on souhaite réaliser une

carte choroplèthe de la partition n°3 (4 classes), il suffit de sélectionner cette variable et de cliquer sur l'icône . La carte s'affiche immédiatement dans une configuration bien adaptée à la CAH : la gamme de couleurs perroquet traduit chaque classe dans la couleur qui est la sienne sur l'arbre de classification et sur les diagrammes à bâtons des déviations (fig. 16.28)

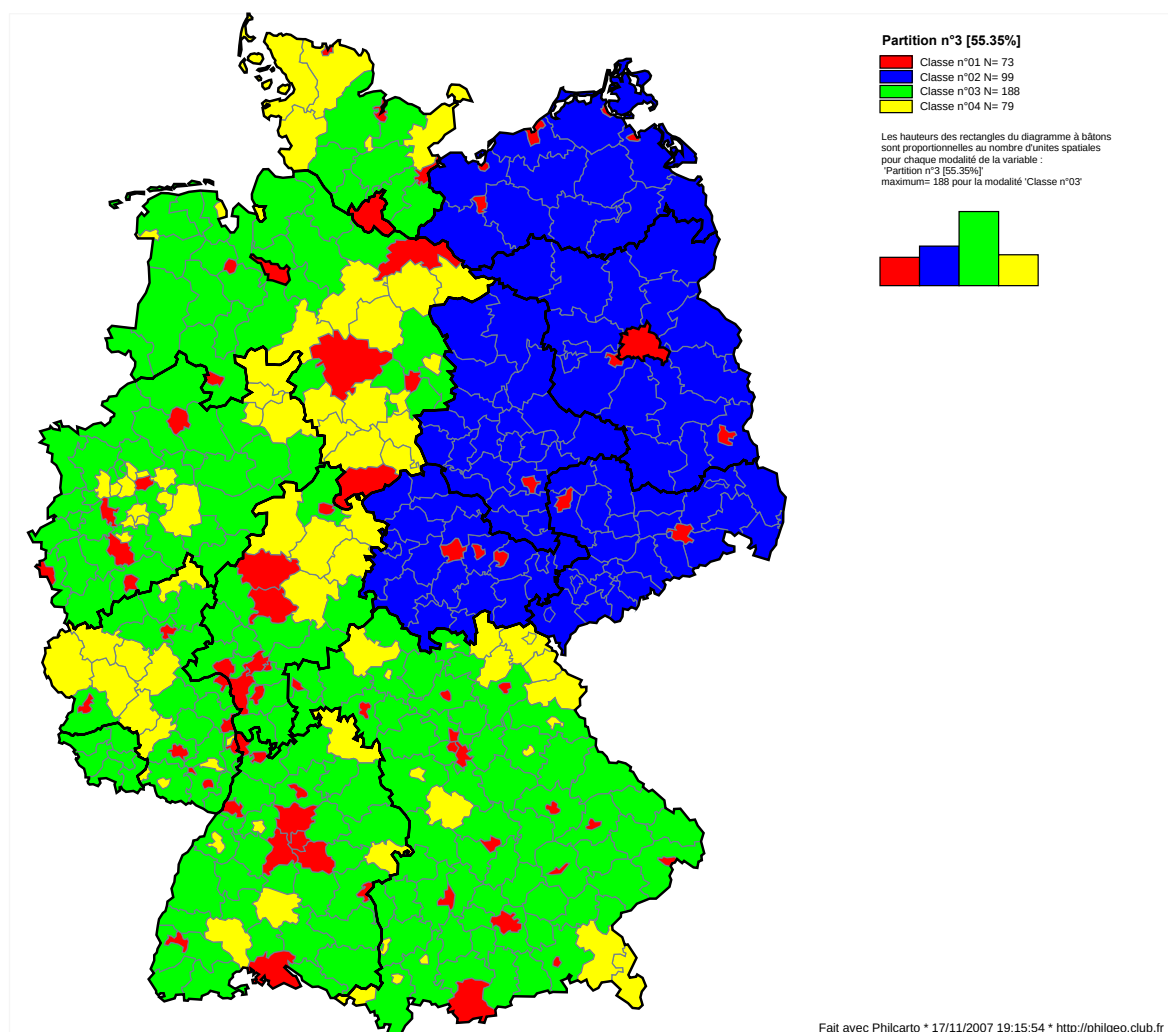


Figure 16.28 : La carte de la partition n°3 issue de la CAH sur les classes d'âges.

La carte obtenue est très suggestive : d'une part, elle met en avant l'ancienne RDA (classe n°2 en bleu) dont on a dit qu'elle présentait un déficit dans la tranche 25-35 ans, celle des jeunes actifs ; d'autre part, elle souligne l'attraction sur ces jeunes actifs des grandes villes (classe n°1 en rouge). Enfin, dans les Länder de l'Ouest, on observe l'existence de régions vieilles (classes n°4 en jaune). Les Kreise de l'Ouest de l'Allemagne forment la classe n°3 (en vert), avec un profil inverse de celui de la classe jaune ; compte tenu du grand nombre de Kreise composant cette classe, 188, il y aurait sans doute intérêt à prolonger la partition de l'arbre au-delà de 4 classes, de manière à examiner l'éventuelle dislocation en plusieurs classes de ce bloc trop important que constitue la classe n°3.

Lorsque la carte d'une partition est tracée, il est souvent nécessaire de revenir à la fenêtre contenant les résultats de l'analyse. Cela se fait de manière simple en

activant l'article **Résultats CAH** du menu **Fenêtres** (fig. 16.29), ou en tapant la combinaison de touches du clavier Ctrl 1.

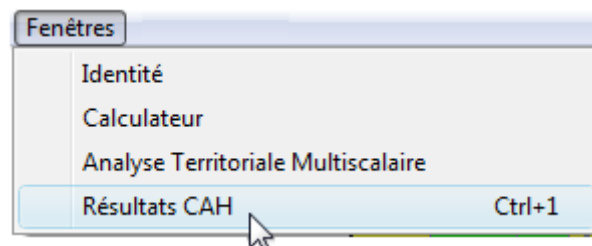


Figure 16.29 : Afficher à nouveau les résultats de l'ACP.

La fenêtre **Analyse multivariée** masque alors la carte. Dans la moitié inférieure de cette fenêtre, on trouve des fonctions qui n'ont pas encore été décrites.

Le bouton **Fermer la fenêtre** permet d'afficher à nouveau la carte en cours.

Le bouton **Enregistrer les résultats** permet d'enregistrer la totalité des résultats de l'analyse dans un dossier unique ; de manière optionnelle, les graphiques enregistrés peuvent être dessinés en blanc sur fond noir ou en noir sur fond blanc selon que l'un ou l'autre des deux boutons ☒ **fond noir** ☐ **fond blanc** est activé. Un clic sur **Enregistrer les résultats** ouvre le dialogue **Rechercher un dossier** permettant de choisir ou de créer un dossier destiné à contenir les différents fichiers de résultats. Soit le dossier existe et dans ce cas, il suffit de le sélectionner là où il se trouve et de valider par **OK** ; soit il n'existe pas encore et un clic sur **Créer un nouveau dossier** ouvre une zone d'édition dans laquelle on peut écrire le nom du dossier créé (fig. 16.30).

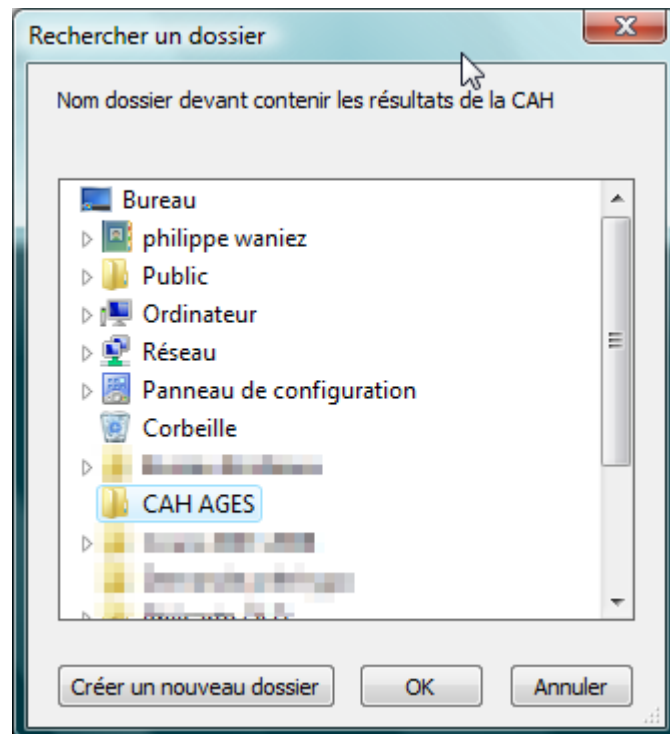
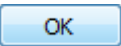


Figure 16.30 : Créer un nouveau dossier de résultats de CAH.

La validation par  provoque l'enregistrement de tous les résultats de la CAH, c'est-à-dire les arbres partitionnés de 2 à 12 classes et les profils correspondants, sous forme de diagrammes à bâtons. Ces graphiques sont au format **.emf**. A ces fichiers s'ajoute un fichier texte contenant le *listing* de la CAH au format **.txt** (fig. 16.31).

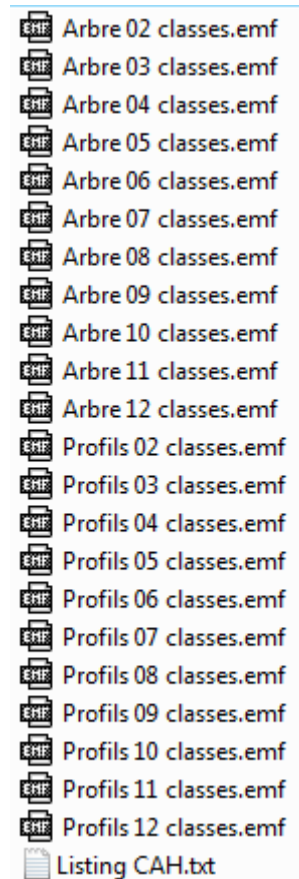


Figure 16.31 : Un exemple d'enregistrement de résultats de CAH dans un dossier.

16.4. Combiner CAH et ACP

Souvent, la CAH est précédée d'une ACP qui a pour but de réduire l'univers des variables à quelques composantes majeures. Cela a pour effet d'accélérer la CAH puisque le nombre de variables entrant dans les calculs est réduit. Contrairement à une idée reçue, cela n'est pas une obligation car on peut parfaitement opérer une CAH directement sur un tableau de données sans l'avoir soumis à l'analyse factorielle. Ainsi, il est possible de faire suivre la CAH par une ACP dont l'objectif est de permettre de placer les classes dans la structure des données issue de l'ACP. Dans ce second cas de figure, il s'agit de visualiser les coordonnées des unités spatiales sur un plan des composantes 1 et 2, en coloriant chacun des points avec la couleur de la classe auxquels ils appartiennent.

Voici un premier exemple de procédure composée des trois étapes suivantes :

- 1° ACP sur les 6 variables de classes d'âges ;
- 2° CAH sur les 2 premières composantes principales (soit 65,4% d'inertie expliquée par le plan factoriel formé par les composantes 1 et 2) ;
- 3° visualisation des classes issues de la CAH sur le plan factoriel.

La première étape, celle de l'ACP, correspond en tous points à la présentation déjà faite (§16.2) ; puis, en fin d'ACP, les coordonnées sur les composantes principales sont enregistrées et affichées dans la liste **couleurs** du dialogue **Variables**. Un nouveau clic sur **MULTIV** dans le cadre **Outils**, permet de choisir le mode de classification **CAH** du dialogue **Analyse des données**, ce qui provoque l'affichage d'une liste de variables ; on trouve les coordonnées sur les composantes principales enregistrées en fin d'ACP. Sélectionnons les deux premières composantes principales, et lançons la CAH par un clic sur **Calculer** (fig. 16.32).

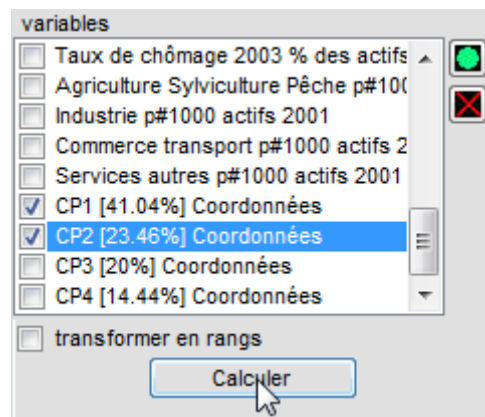


Figure 16.32 : La sélection des coordonnées sur les deux premières composantes principales.

Après quelques secondes d'attente la fenêtre **Analyse multivariée** s'affiche à l'écran et montre les résultats de la CAH. Un clic sur **Carte** permet d'observer que les 11 partitions ont bien été enregistrées dans le tableau de données : elles sont listées dans la liste **couleurs** du dialogue **Variables** (fig. 16.33)

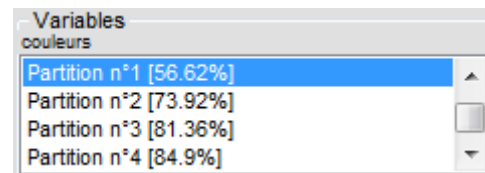


Figure 16.33 : Les partitions issues de la CAH réalisée sur les deux premières composantes principales.

Pour visualiser les classes de ces partitions dans le plan factoriel formé par les composantes 1 et 2, il faut tout d'abord refaire l'ACP, avec les mêmes 6 variables de la première étape, ce qui ne prend que quelques secondes.

Lorsque le plan factoriel est affiché, on trouve les 11 partitions dans la liste **variable nominale pour colorier les points**. Il suffit alors de sélectionner l'une ou l'autre des partitions pour que les points formant chaque classe soient coloriés de la couleur représentant chaque classe sur l'arbre et sur les diagrammes des profils des classes (fig. 16.34).

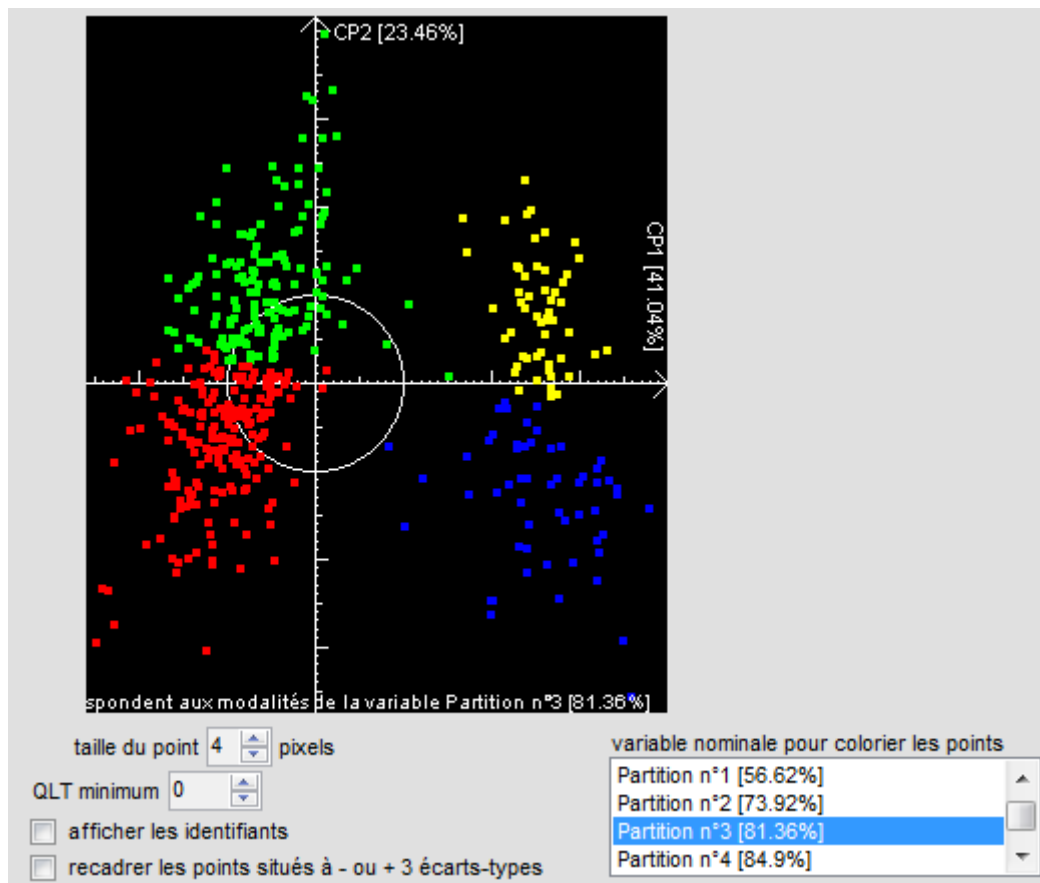


Figure 16.34 : La visualisation, dans le plan factoriel des composantes 1 et 2, des 4 classes de la partition n°3 issues de la CAH réalisée sur ces deux premières composantes principales.

Voici un second exemple de procédure composée de deux étapes seulement :

- 1° CAH sur les 6 variables de classes d'âges ;
- 2° ACP sur les mêmes 6 variables suivie de la visualisation des classes issues de la CAH sur le plan factoriel des composantes 1 et 2.

La première étape, celle de la CAH, correspond en tous points à la présentation déjà faite (§16.3) ; puis, en fin de CAH, les 11 partitions sont enregistrées dans le tableau de données.

La seconde étape, celle de l'ACP, correspond en tous points à la présentation déjà faite (§16.2) ; lorsque le plan factoriel est affiché, on trouve les 11 partitions dans la liste **variable nominale pour colorier les points**. Il suffit alors de sélectionner l'une ou l'autre des partitions pour que les points formant chaque classe soient coloriés de la couleur représentant chaque classe sur l'arbre et sur les diagrammes des profils des classes (fig. 16.35). Sur ce graphique, on observe un certain chevauchement des classes. Ceci est normal puisque c'est la totalité de l'inertie qui est prise en compte lors du calcul des classes, et non pas seulement l'inertie du plan factoriel des composantes 1 et 2. Cela peut paraître gênant pour la classe n°1, en rouge qui concerne à la fois les Kreise de l'Est (en bleu) et ceux de l'Ouest. Mais comme cette classe est celle qui est principalement formée par les grandes villes, ce chevauchement est au contraire une information supplémentaire puisqu'il traduit une structure par âges caractéristique des grandes villes, tant à l'Ouest qu'à l'Est.

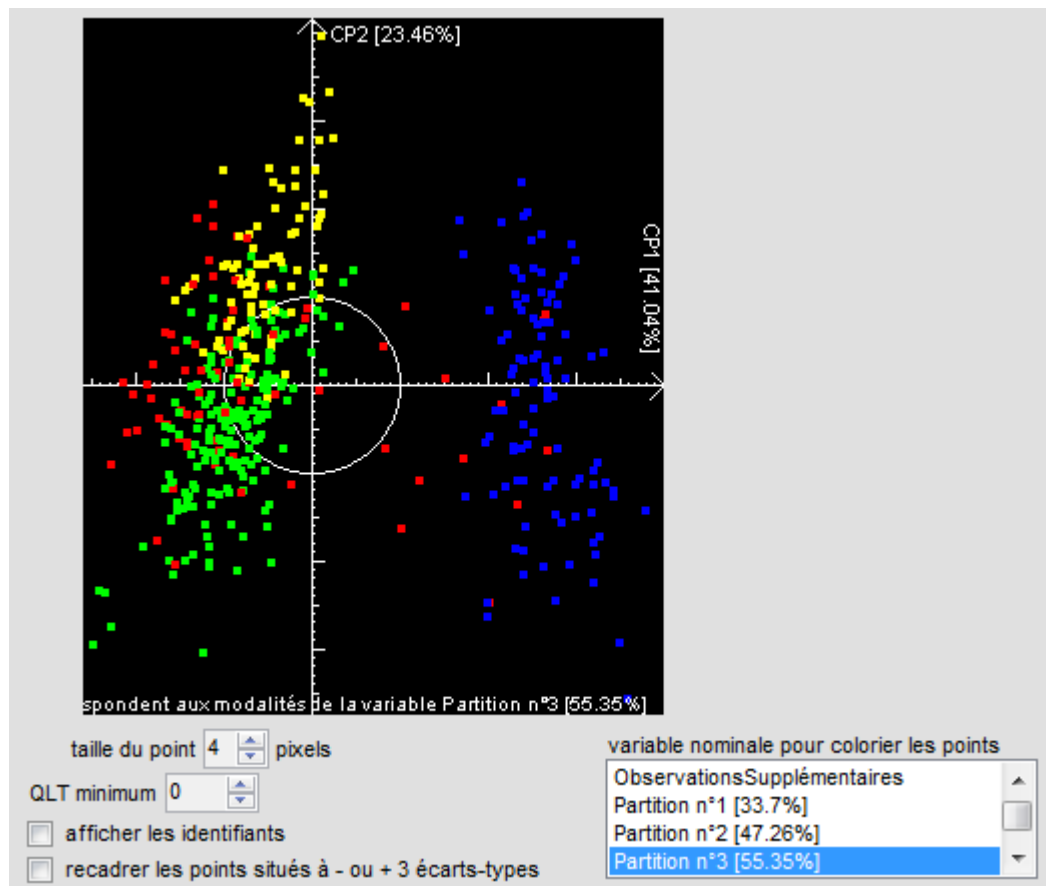


Figure 16.35 : La visualisation, dans le plan factoriel des composantes 1 et 2, des 4 classes de la partition n°3 issues de la CAH réalisée sur les mêmes variables que l'ACP.

Pour conclure ce chapitre, soulignons que tous ces traitements demandaient, il n'y a pas si longtemps, une programmation laborieuse avec des logiciels coûteux. Avec *Philcarto* les méthodes d'analyse multivariée des données appliquées à la cartographie thématique sont désormais à la portée de tous. Le temps gagné sur les problèmes techniques peut être utilisé pour mieux comprendre ces techniques afin, d'une part, d'éviter les erreurs d'interprétation des résultats et d'autre part de tirer la « substantifique moelle » des données produite souvent à grands frais.

17. L'analyse exploratoire des données

Dossier : \Philcarto exemples\Allemagne\
Fond de carte : Allemagne carte Kreise.ai
Statistiques : Allemagne data Kreise.xls

La cartographie exploratoire des données cherche à pallier les difficultés rencontrées lors de l'analyse de certaines données à l'aide de méthodes usuelles, difficultés non spécifiques aux données géographiques puisqu'on les retrouve dans toutes les sciences humaines et même dans certains domaines des sciences naturelles. « Il est difficile d'analyser les données géographiques à partir de modèles statistiques traditionnels pour des raisons tenant aux interdépendances spatiales, [et] à la complexité des structures spatiales (...) » [Unwin ; 1992]. Les tentatives faites par certains statisticiens pour adapter les méthodes statistiques classiques (régression, analyse factorielle) aux données spatiales conduisent souvent à des solutions complexes et difficiles à mettre en pratique dans le cas d'une étude en vraie grandeur.

Nombre de chercheurs, de la biologie jusqu'aux sciences sociales, doivent recourir à des méthodes moins « dures », moins contraignantes, plus intuitives aussi. Il y a déjà une trentaine d'années, J.W. Tukey, le père de l'Analyse Exploratoire des Données (AED) proposait un ensemble de techniques d'analyse, dont l'avantage principal était d'inciter le chercheur à s'interroger directement sur les données elles-mêmes, et pas seulement sur les paramètres statistiques qui les résument (moyenne, variance, etc.). Il insistait sur la nécessité de réagir à l'apparition de valeurs exceptionnelles ou manquantes, aux ruptures dans les distributions, aux regroupements, etc. [Tukey ; 1977]. La plupart de ces techniques reposaient sur des représentations graphiques permettant d'observer les données sous divers angles, de manière à « piéger » l'information significative, celle qui permet d'avancer dans la connaissance de l'objet étudié. De ces préceptes est issue une véritable « école » d'analyse des données statistiques basée sur l'observation des représentations graphiques. L'ouvrage « Visualizing Data » [Cleveland ; 1993] représente la somme pédagogique indispensable à qui veut s'immerger pleinement dans les méthodes exploratoires d'analyse des données.

À l'époque de Tukey, l'AED demeurait une approche presque aussi lourde, en termes de temps passé, que celle plus courante de la statistique descriptive. Grâce aux ordinateurs dotés d'interfaces graphiques, l'AED a pris une nouvelle dimension : l'interactivité. Aujourd'hui, le chercheur peut non seulement observer les diverses configurations des données sur les représentations graphiques, mais il peut aussi agir sur elles, en les sélectionnant, les coloriant, les masquant, etc. Et ce n'est pas le hasard si les premiers logiciels d'AED sont apparus sur l'ordinateur grand public le plus en avance à l'époque en ce domaine : le *Macintosh* de la société Apple Computer. Dans un ouvrage publié en 1991, j'ai tenté de dresser un état comparatif des différents logiciels qui se partageaient le marché de l'AED sur *Macintosh* [Waniez ; 1991]. C'était aussi l'occasion de présenter les principes nécessaires à une exploration fructueuse. Mais on déplorait déjà à cette occasion qu'aucun de ces logiciels n'ait une fonction de cartographie, absence bien regrettable pour qui veut analyser de la sorte des données spatialisées.

Pour trouver de bons exemples de ce que devrait être un vrai logiciel de Cartographie Exploratoire des Données, il faut se tourner vers les prototypes développés par des chercheurs comme *Polygon Explorer*, *Regard*, *Manet* sur *Macintosh* ou *Exploremap* sur PC.

En plus d'une certaine variété de graphiques statistiques associés à diverses fonctions de sélection, ces logiciels sont dotés d'une fenêtre Carte qui répond, comme les autres fenêtres graphiques, aux actions de l'utilisateur. De ce fait, la Cartographie Exploratoire des Données permet d'appréhender de manière subtile les configurations spatiales. Ainsi s'établit entre le chercheur et les données une grande proximité favorisée par la carte et les graphiques qui se transforment au gré des questions qui leur sont posées.

Il y a une dizaine d'années, j'ai diffusé *Philexplo* pour *Macintosh* qui s'inspirait de *Regard*, réalisé sous la direction de A. Unwin au Trinity College de Dublin. Il en différait cependant sous plusieurs aspects : des types de graphiques différents avec, par exemple, la remise à l'ordre du jour des diagrammes triangulaires, mais en les plongeant dans un environnement interactif qui augmentait encore leur efficacité. La seconde différence importante résidait dans le caractère pleinement opérationnel de *Philexplo*, sur quelques milliers d'unités spatiales, faisant de l'exploration cartographique un instrument mobilisable rapidement par un chercheur non-spécialiste des techniques d'information géographique.

Les fonctions d'exploration statistique disponibles dans la version 5 de *Philcarto* pour *Windows* bénéficient de l'expérience de la programmation de ce genre de logiciel interactif acquise avec *Philexplo*. Sur ces premières bases, j'ai ajouté des fonctions nouvelles comme l'ajustement Lowess par exemple. *Philexplo* traitait seulement des cartes en plages de couleurs ; la méthode est maintenant étendue aux cercles colorés et aux lignes.

17.1. Réaliser une analyse exploratoire

Après avoir cliqué sur **EXPLO** du cadre **Outils**, une nouvelle variable, nommée Exploration n°1, apparaît dans la liste **couleurs** du dialogue **Variables** (fig. 17.1). Simultanément, l'icône  des cartes choroplèthes s'affiche dans le dialogue **Types de cartes** du cadre **Outils**. L'analyse exploratoire commence par un clic sur cette icône.

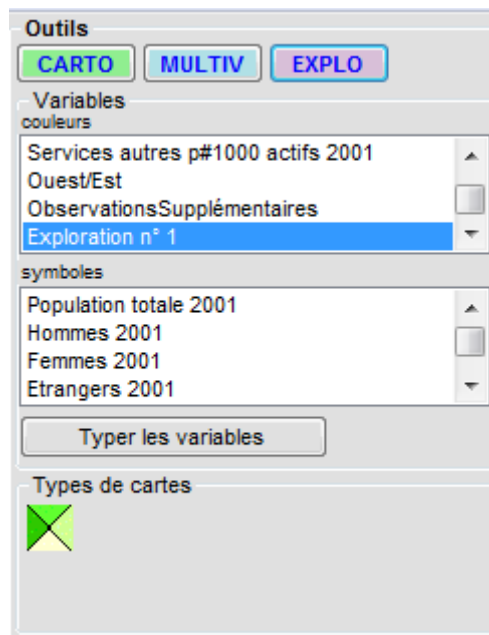


Figure 17.1 : L'ajout d'une variable Exploration n°1 et l'affichage de l'icône du type de carte choroplèthe.

Le bureau de *Philcarto* est alors modifié de la façon suivante (fig. 17.2) :

- une carte entièrement grise est affichée ; elle traduit l'état de la variable Exploration n°1 qui est une variable nominale prenant la seule modalité « 000 » ;
- dans le cadre **Outils**, un dialogue **Couleurs** apparaît ; il renferme une ligne 000 qui se termine par un bouton du même gris que celui de la carte ;
- un second dialogue nommé **Exploration** (fig. 17.3) comprend différents éléments nécessaires à l'exploration proprement dite :
 - ❖ l'icône du graphique bivarié  ;
 - ❖ l'icône du diagramme triangulaire  ;
 - ❖ une série de 30 boutons radio colorés permettant de choisir la couleur de sélection ; le bouton radio rouge est sélectionné en début d'exploration.
 - ❖ deux boutons de gestion des afficheurs de couleurs **tous** **aucun** ;
 - ❖ un rectangle gris qui contiendra par la suite les afficheurs des couleurs effectivement présentes sur la carte.

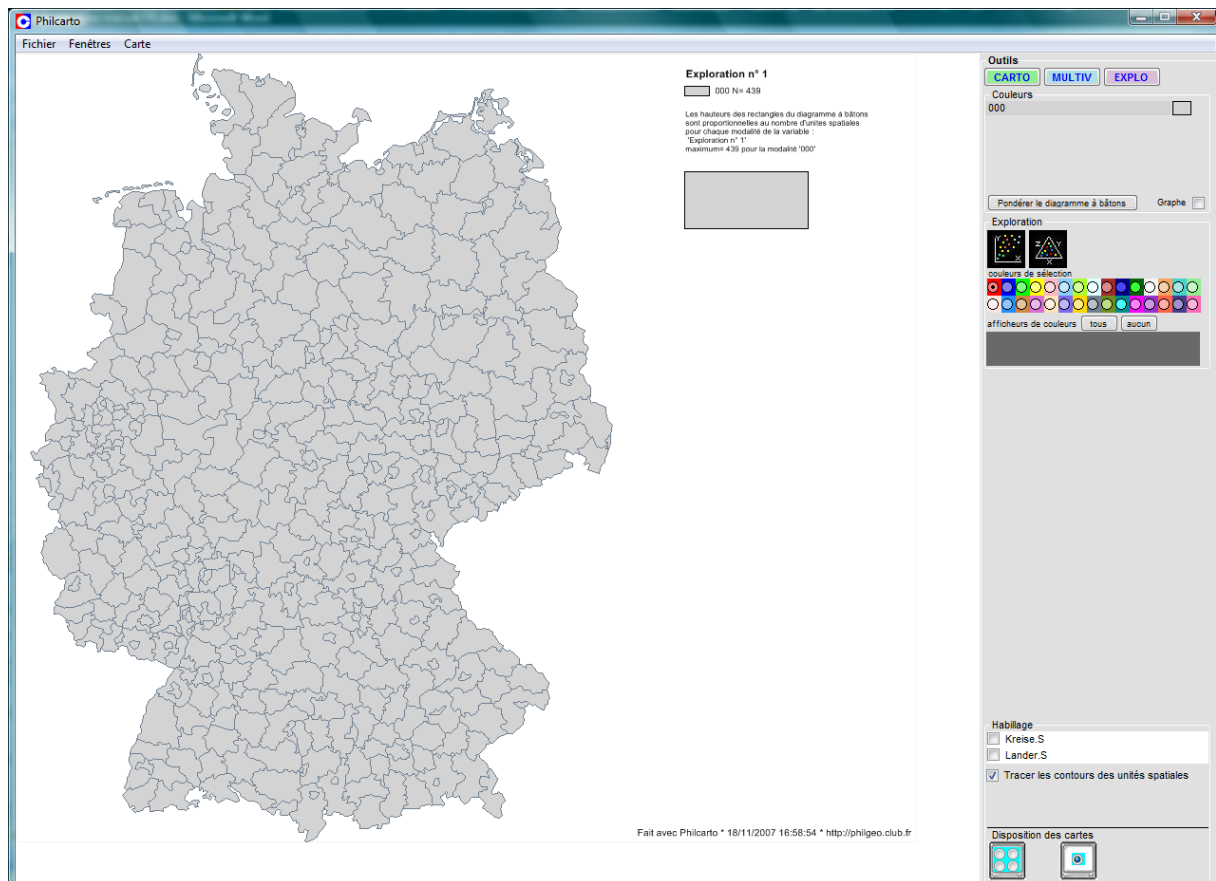


Figure 17.2 : Le bureau de *Philcarto* en début d'analyse exploratoire.

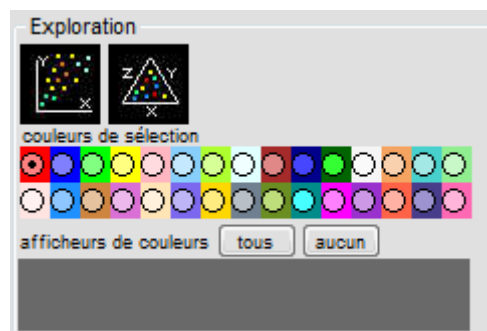


Figure 17.3 : Le dialogue **Exploration** en début d'analyse exploratoire.

17.2. L'analyse exploratoire bivariable

L'exploration bivariable permet d'étudier les relations qu'entretiennent deux variables. Elle commence par un clic sur l'icône du graphique bivié  ; une fenêtre intitulée  **Graphique bivié** apparaît à la gauche du cadre **Outils** (fig. 17.4). Elle est équipée de deux listes de variables permettant de sélectionner la **variable exogène (X)** et la **variable endogène (Y)** ; ce vocabulaire est celui de la régression et peut être remplacé, si l'on préfère, par variable en abscisse et variable en ordonnée.

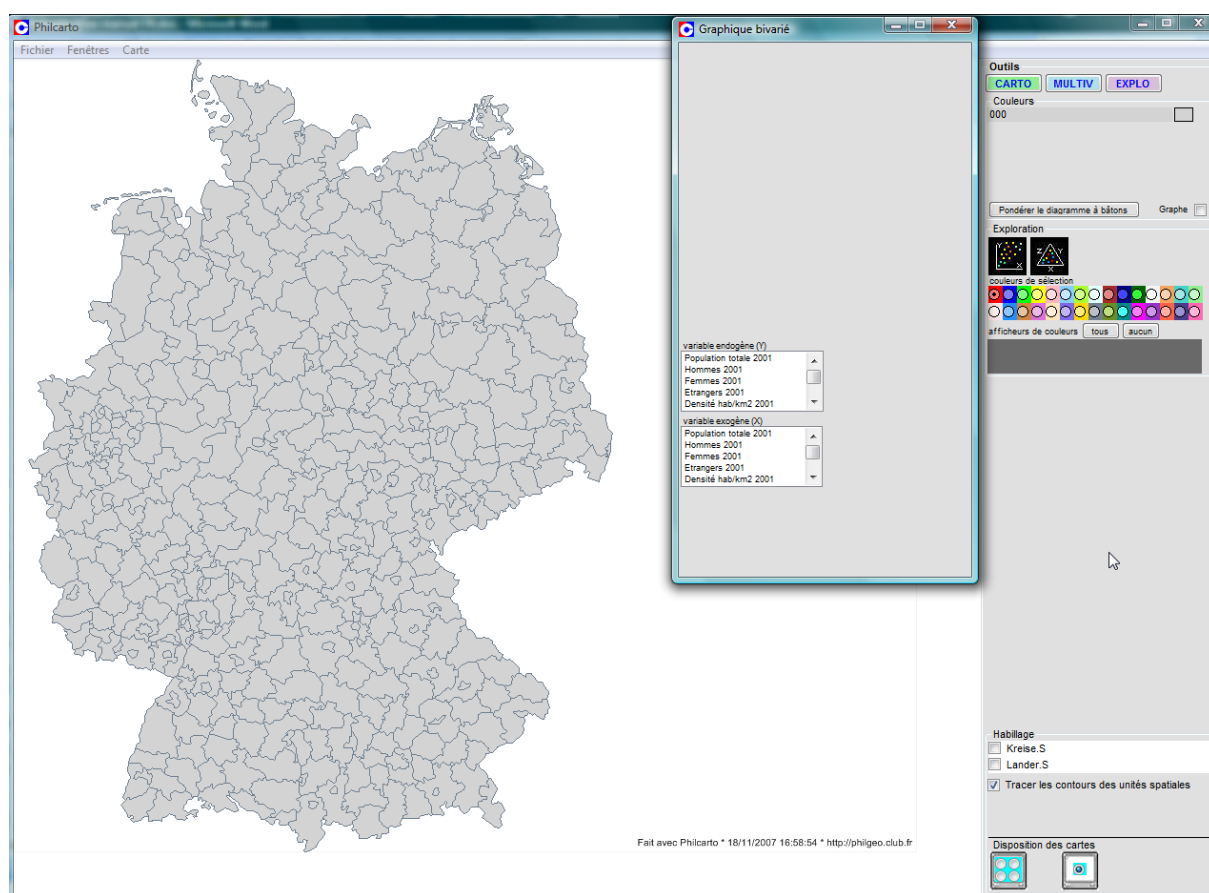


Figure 17.4 : Le bureau de *Philcarto* en début d'analyse exploratoire bivariable.

Par exemple, si l'on souhaite montrer qu'il existe une relation entre la présence des étrangers et le niveau d'urbanisation approché par la densité de population, on désigne en Y la variable **Etrangers 2001 %** et en X la variable **Densité hab/km2 2001**. Dès que deux variables sont sélectionnées, le graphique bivié est immédiatement tracé (fig. 17.5). Simultanément, de nombreux contrôles sont ajoutés dans la fenêtre

Graphique bivié.

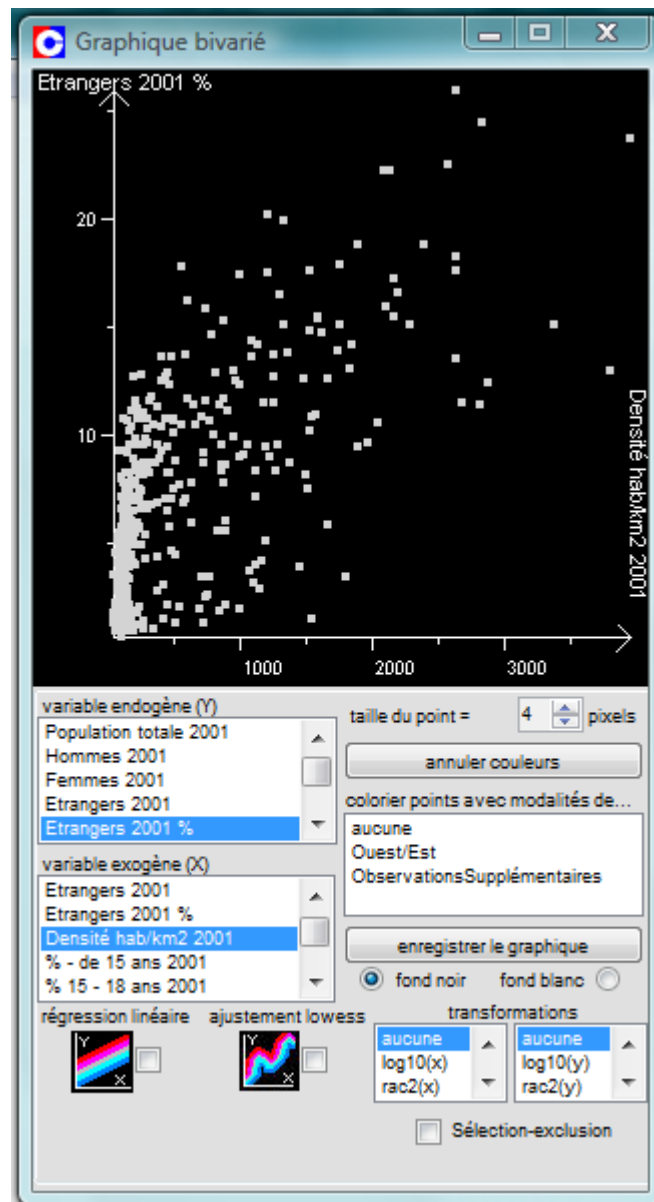


Figure 17.5 : Le graphique bivarié de la densité de population (X) et du % d'étrangers (Y).

On observe sur le graphique une forte dissymétrie de la distribution des densités : beaucoup de faibles valeurs et peu de valeurs élevées. Si cela est bien conforme à la réalité, la concentration des points proches de la valeur 0 en abscisse empêche d'observer quoi que ce soit dans ce voisinage. Une transformation logarithmique sur X à l'aide du dialogue **transformations** permet d'étaler la distribution des densités et d'obtenir un graphique plus lisible, sur lequel il est possible d'identifier chaque point en plaçant dessus le pointeur de la souris (fig. 17.6).

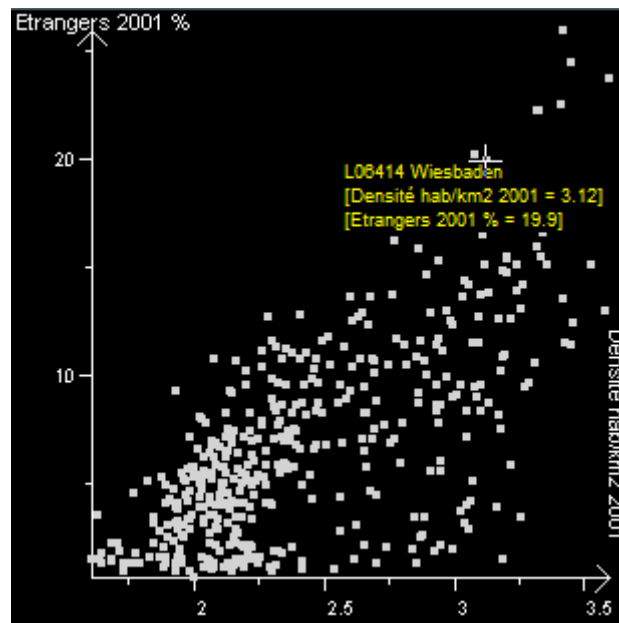


Figure 17.6 : Le graphique bivariable de la densité de population (X) et du % d'étrangers (Y) après transformation de la densité en logarithme décimal ; la valeur de la densité indiquée pour Wiesbaden est le logarithme décimal de la densité.

Sur ce graphique, on remarque qu'il existe effectivement une tendance qui fait que les étrangers sont en plus forte proportion dans les Kreise densément peuplés (donc dans les grandes villes). Pour confirmer cette impression, sélectionner les points situés dans la partie supérieure droite du graphique en utilisant le lasso (fig. 17.7) (pour faire apparaître le lasso, cliquer au point de départ de la sélection avec le bouton gauche de la souris et déplacer la souris en maintenant ce bouton enfoncé). Ce secteur-là du nuage de points comprend les Kreise les plus denses et présentant la plus forte proportion d'étrangers dans leur population.

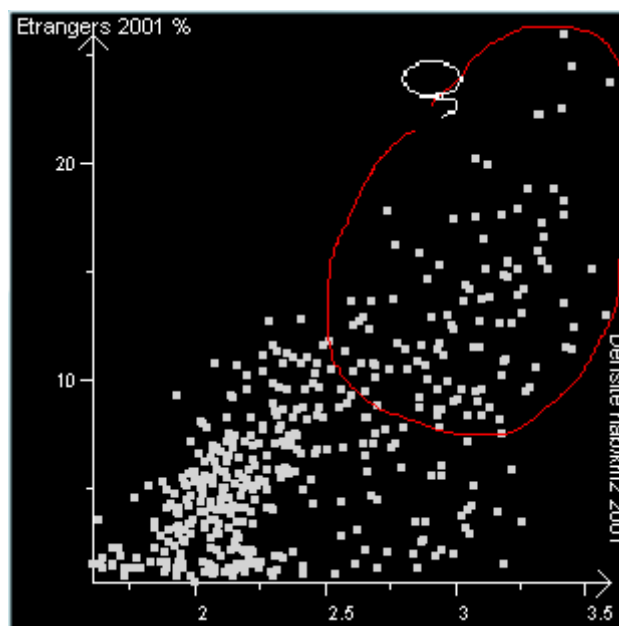


Figure 17.7 : La sélection de la partie supérieure droite du nuage de points.

Lorsque le bouton gauche de la souris est relâché, les points situés à l'intérieur du lasso sont affichés en rouge et les Kreise concernés sont coloriés en rouge sur la carte (fig. 17.8). Simultanément, un nouvel élément de couleur rouge et prenant la valeur « 001 » est affiché dans le dialogue **Couleurs** du cadre **Outils**. Une case à cocher rouge s'inscrit dans les **afficheurs de couleurs**.

On observe sur cette carte que dans la partie Est de l'Allemagne, seul le Kreis de Berlin est sélectionné. Dans la partie Ouest du pays, la sélection concerne soit des villes importantes isolées (Kassel, Münster...) soit de grands bassins industriels-urbains tels que les bassins de la Ruhr (Dortmund), du Rhin-Main (Francfort), du Rhin-Neckar (Mannheim), les régions de Stuttgart et de Munich.

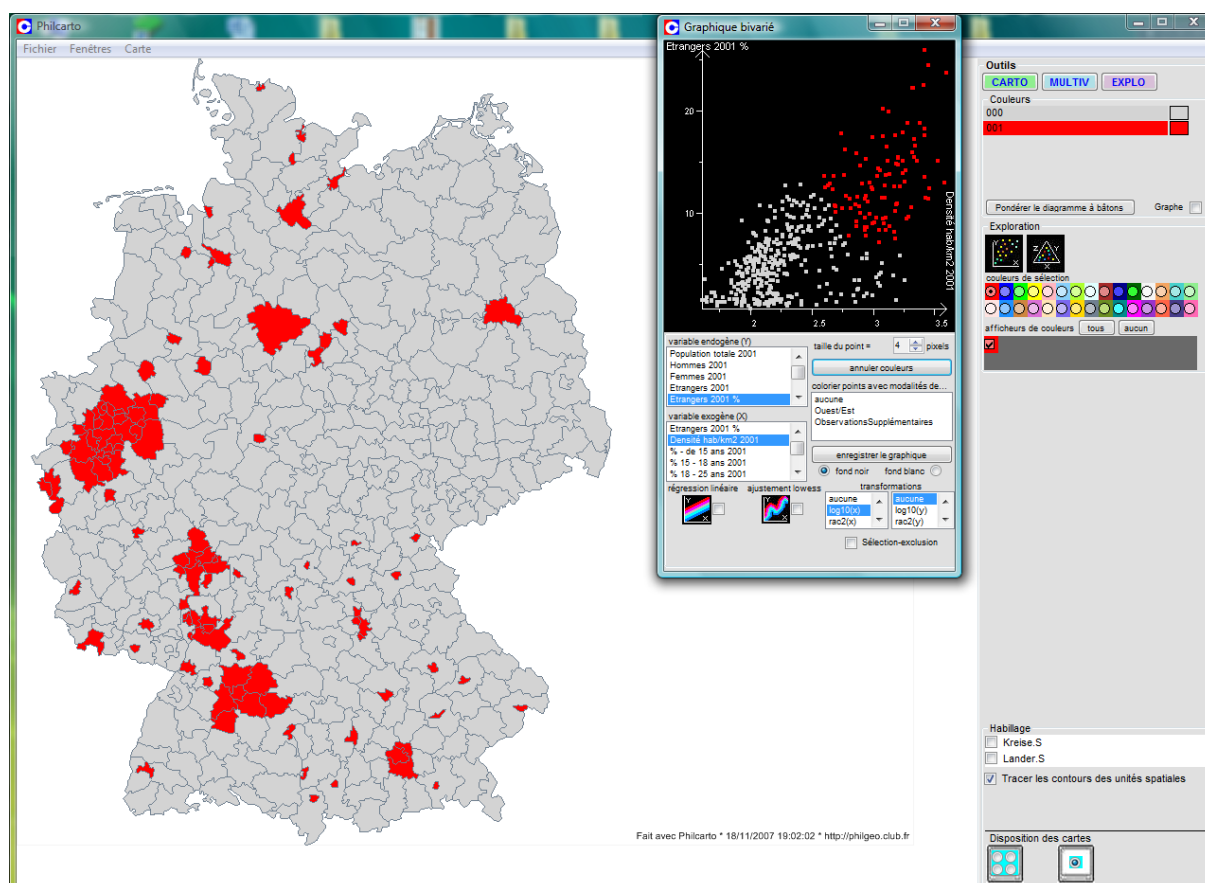


Figure 17.8 : La sélection de la partie supérieure droite du nuage de points.

Pourquoi la sélection des points apparaît-elle en rouge ? Tout simplement parce que c'est le bouton radio de couleur rouge qui est sélectionné dans le dialogue **couleurs de sélection**.

A quoi sert la case à cocher rouge ? Lorsque cette case est cochée, les éléments cartographiques de même couleur sont visibles sur la carte ; s'il est décoché, ces éléments sont occultés : ils sont tracés en blanc sur la carte et disparaissent du graphique (fig. 17.9). Dans le vocabulaire de l'analyse exploratoire, ce procédé s'appelle le masquage ; il permet d'afficher certaines sélections et pas les autres. Pour faire réapparaître la sélection rouge, il suffit de cocher la couleur rouge dans le dialogue **couleurs de sélection**.

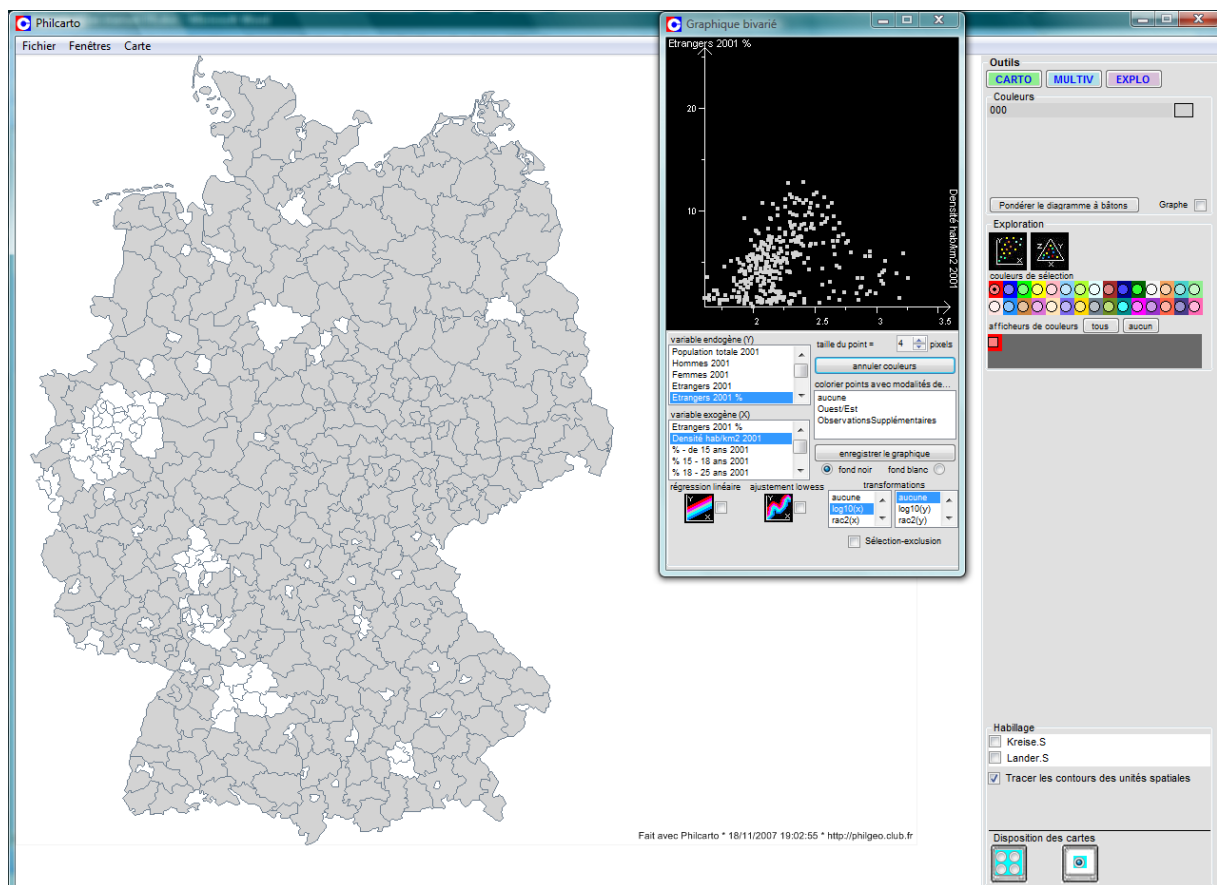


Figure 17.9 : L'afficheur de couleur rouge étant décoché, les éléments rouges sont tracés en blanc sur la carte et disparaissent du graphique bivarié.

L'absence des villes de l'ancienne RDA dans la sélection rouge est intrigante. Cela signifie-t-il que les villes de l'Est n'attirent aucun étranger ? Pour tenter de répondre à cette question, sélectionnons en vert clair les points situés dans la partie inférieure droite du graphique. Pour cela, choisissons cette couleur dans le dialogue **couleurs de sélection**, et sélectionnons au lasso les points concernés sur le graphique.

La carte obtenue concerne majoritairement des kreise de l'Est (fig. 17.10) que l'on peut identifier en plaçant le pointeur de la souris sur les éléments cartographiques verts, du Nord au Sud : Rostock, Schwerin, Brandenburg, Leipzig, Dresde, Erfurt, Jena, et le bassin industriel de Chemnitz (qu'on ne résiste pas à appeler de son ancien nom : Karl Marx Stadt !). Autrement dit, on trouve dans la sélection verte, la plupart des grandes régions urbaines et industrielles (à l'exclusion notable de Berlin), ce qui traduit une attraction des grandes villes sur la population d'origine étrangère, mais à un niveau moindre à l'Est qu'à l'Ouest. Un tel résultat invite donc à poursuivre l'étude de la relation en séparant les Kreise des nouveaux Länder de ceux de l'ancienne RFA.

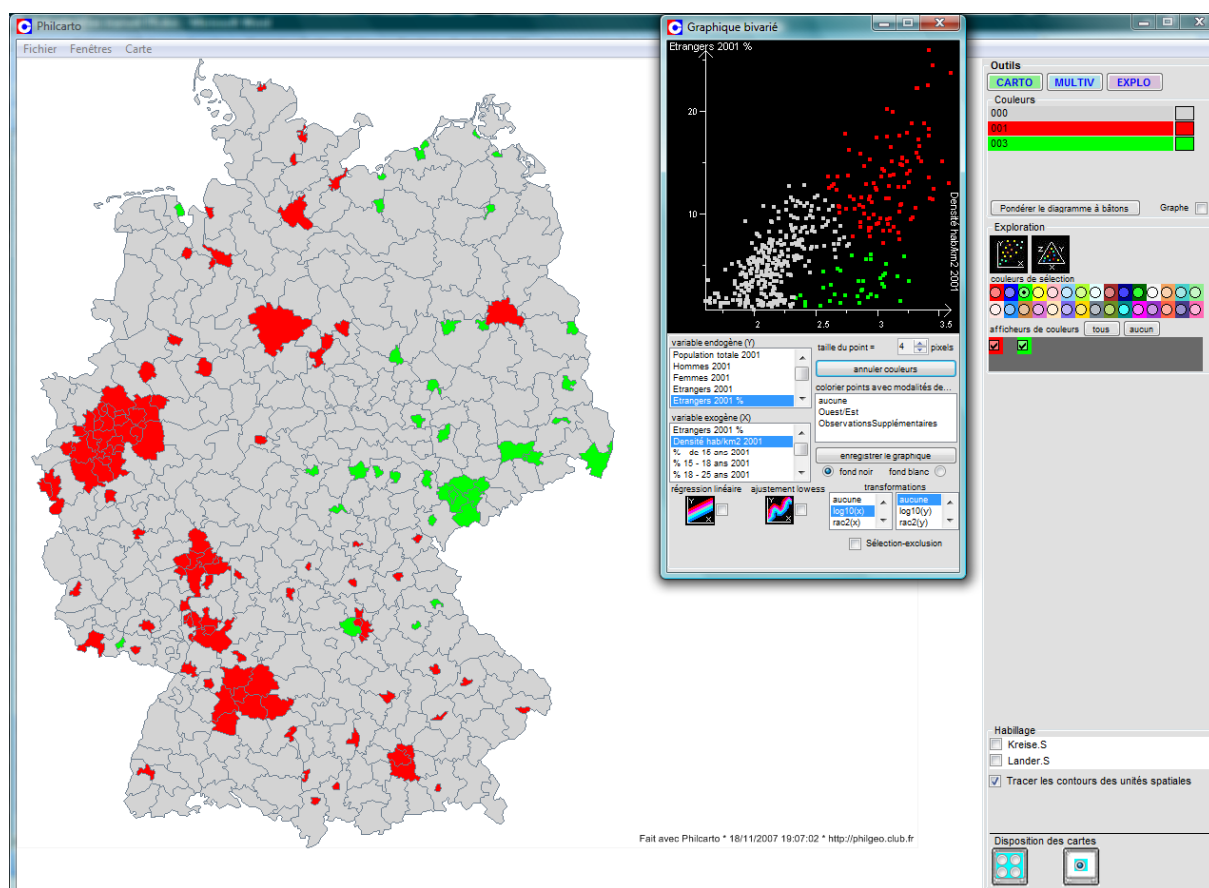


Figure 17.10 : Le bureau de *Philcarto* avec les sélections rouges et vertes sur le graphique bivarié.

Poursuivre l'étude de la relation entre le pourcentage d'étrangers dans la population totale et la densité de population interprétée comme indicateur d'urbanisation revient à chercher à modéliser cette relation. *Philcarto* propose deux solutions pour y parvenir : la régression linéaire et l'ajustement Lowess.

L'analyse graphique précédente ayant montré qu'il est juste d'étudier séparément les deux anciennes Allemagnes, voici comment sélectionner l'ex Allemagne de l'Ouest. Tout d'abord, annulons toutes les couleurs par un clic sur **annuler couleurs** dans fenêtre **Graphique bivarié**. Au-dessous de ce bouton figure la liste de variables nominales **colorier points avec modalités de...** (fig. 17.11); dans cette liste, la variable Ouest/Est indique le rattachement des Kreise avant la réunification, soit à l'Ouest (ancienne RFA), soit à l'Est (ancienne RDA). La sélection de cette variable se traduit par le coloriage des points en fonction de ces rattachements, et l'inscription de deux cases à cocher dans le dialogue **afficheurs de couleurs** (fig. 17.12).

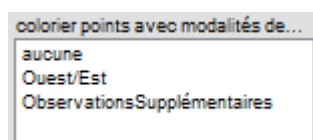


Figure 17.11 : La liste des variables nominales pour colorier les points.

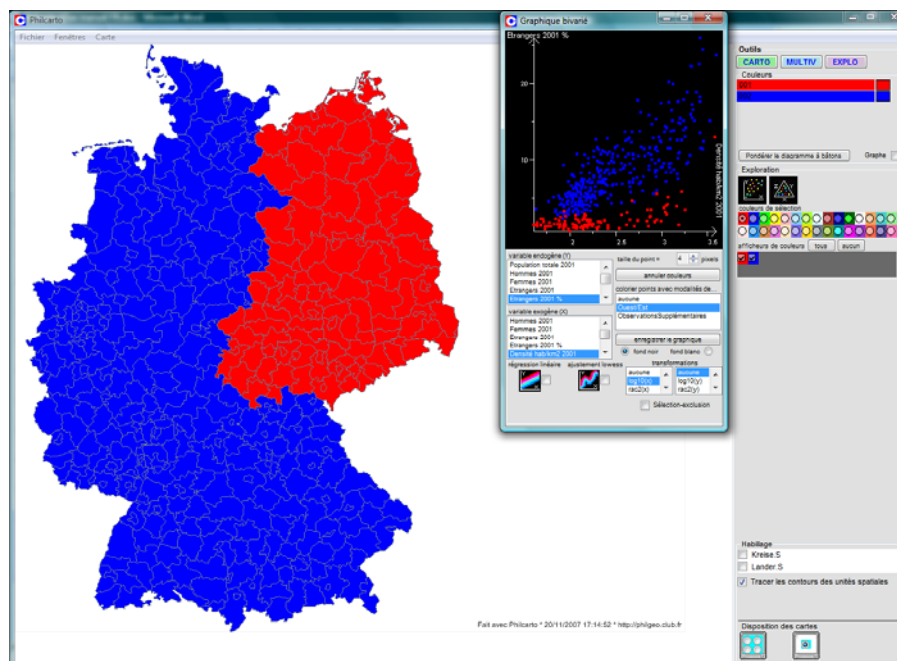


Figure 17.12 : Le bureau de *Philcarto* après la sélection de la variable Ouest/Est.

Pour que les analyses à venir portent seulement sur l'Allemagne de l'Ouest, il suffit de décocher l'afficheur de couleur rouge (fig. 17.13). L'ancienne RDA devient de ce fait inactive pour la suite des opérations.

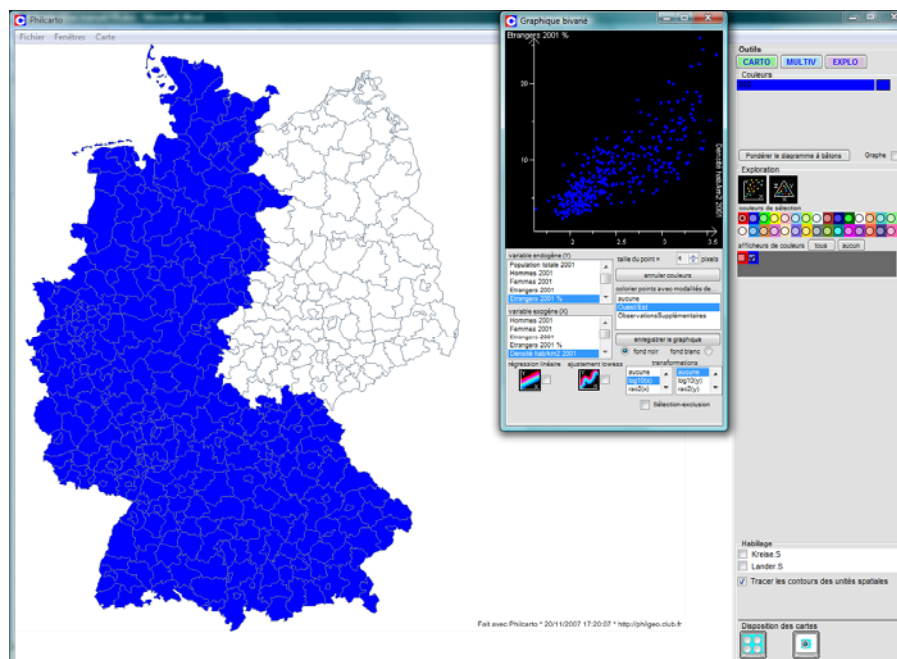


Figure 17.13 : Le bureau de *Philcarto* après inactivation de l'ancienne RDA.

Le tracé de la droite de régression sur l'ensemble des points bleus est provoqué par un clic sur . Les paramètres de la droite et le coefficient de détermination ($R^2=0,63$) sont affichés (fig. 17.14). Puis les écarts à la droite de

régression sont calculés ; ces résidus sont alors divisés par leur écart-type et recodés en classe de la manière suivante :

- en-dessous de -1,5 écart-type : très fort résidu négatif, bleu foncé ;
- entre -1,5 et -1,0 écart-type : fort résidu négatif, bleu ;
- entre -1,0 et -0,5 écart-type : résidu négatif modéré, bleu clair ;
- entre -0,5 et +0,5 écart-type : résidu non significatif, gris ;
- entre +0,5 et 1,0 écart-type : résidu positif modéré, jaune clair ;
- entre +0,5 et +1,5 écart-type : fort résidu positif, jaune ;
- au-dessus de +1,5 écart-type : très fort résidu positif, rouge.

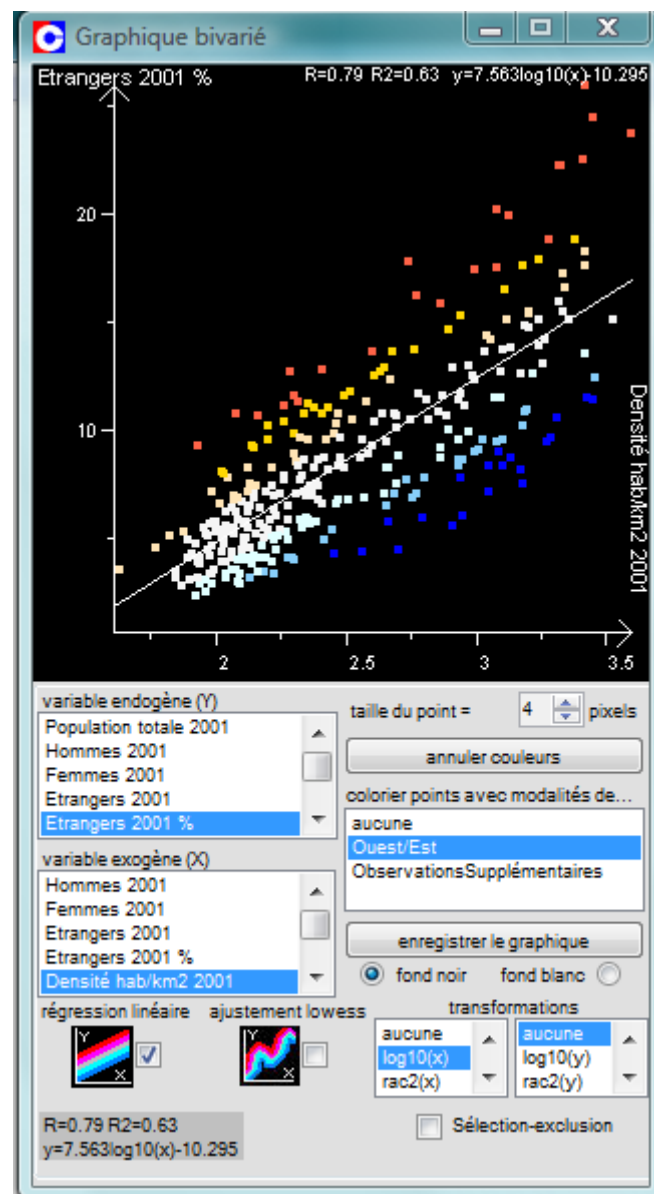


Figure 17.14 : Le tracé de la droite de régression et l'étude des résidus.

La carte des résidus est affichée à la suite du recodage des résidus, dans les mêmes couleurs que sur le graphique bivarié (fig. 17.15). On observe une certaine concentration des résidus positifs dans le Baden-Württemberg (Stuttgart) et la Bavière (Münich) ce qui traduit une attractivité de ces régions sur les étrangers supérieure au reste de la RFA. Au contraire, dans la moitié nord, les résidus sont souvent négatifs, comme dans la Ruhr, dans la Sarre, dans la région de Brème... ce

qui exprime une moindre attractivité qui s'explique peut-être par un différentiel de dynamisme des régions industrielles et de demande de travailleurs étrangers plus réduite.

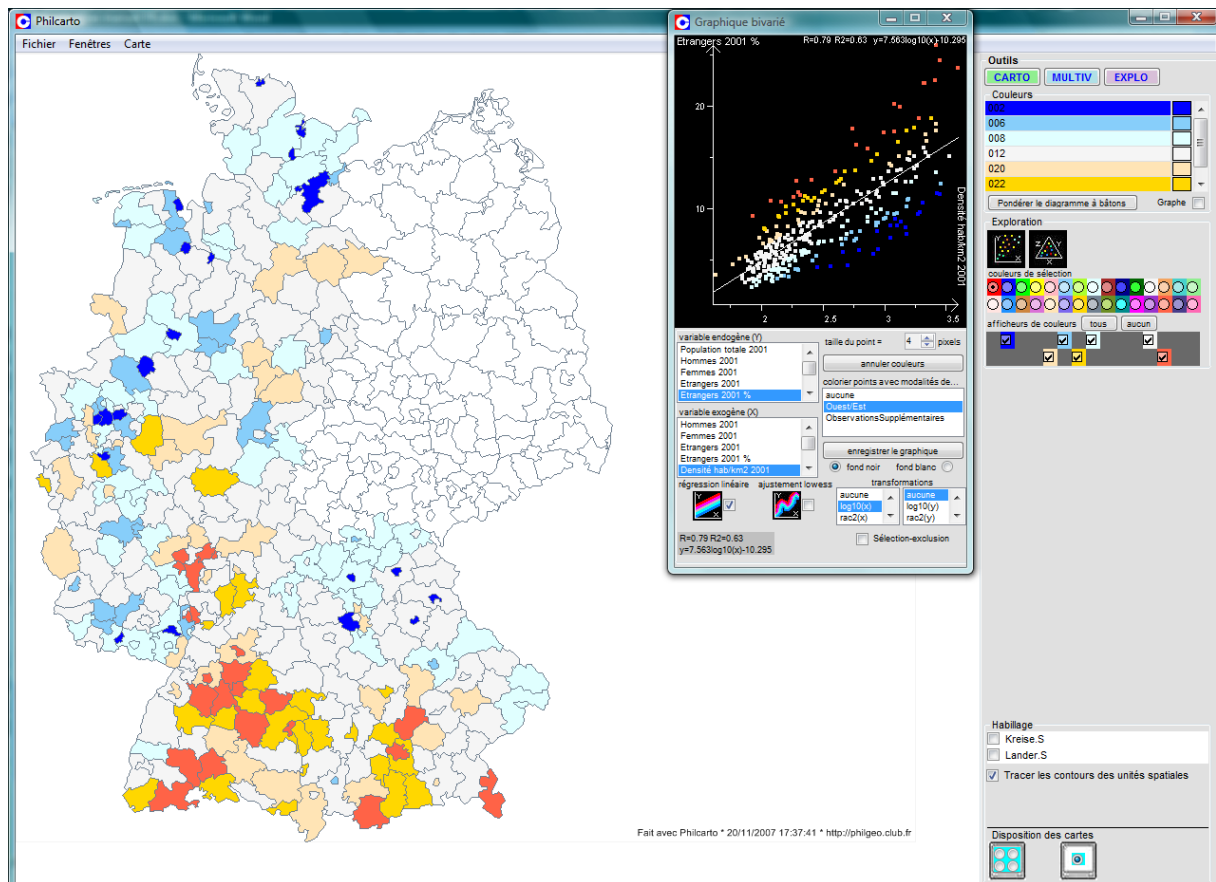


Figure 17.15 : Le bureau de *Philcarto* après le calcul des résidus de régression et l'affichage de la carte correspondante.

Philcarto propose une seconde méthode d'ajustement dite Lowess (pour LOcally WEighted RegreSSion, ou régression pondérée localement, également connue sous le nom de Loess). Elle peut être utilisée lorsque le nuage de points du graphique bivarie ne s'approche pas de la forme d'un ovale très allongé et lorsque les transformations usuelles ont échoué pour linéariser la relation.

Pour réaliser cette fonction, j'ai utilisé un programme FORTRAN diffusé sur Internet par W.S. Cleveland des *Bell Laboratories* à Murray Hill NJ 07974, USA, programme que j'ai adapté à *Philcarto* (<http://www.netlib.org/go/lowess.f>). Pour une présentation complète de l'ajustement Lowess, il est souhaitable de se reporter à l'ouvrage de cet auteur [Cleveland ; 1994].

Pour réaliser un tel ajustement avec *Philcarto*, il faut cliquer sur  ; le reste du traitement est identique à celui de la régression, à une différence importante près : l'utilisation du tenseur exprimé en proportion du nombre d'observations à analyser. Fixé par défaut à 0,5 (soit la moitié du nombre des observations), ce paramètre peut varier entre 0,1 et 1. Plus il se rapproche de 1, moins les aspérités de la relation sont prises en compte ; un tenseur de 1 donne un résultat semblable à celui de la régression. Le réglage du tenseur se mesure donc en % du nombre de

tenseur = 20 %
R2= 0.65

de la fenêtre

 Graphique bivarié. Un genre de coefficient de détermination est également calculé.

Les **afficheurs de couleurs** permettent d'afficher seulement les couleurs souhaitées à la fois sur le graphique et sur la carte. Dans le cas des ajustements, cela permet de ne retenir que les résidus forts ou très forts, simplement en décochant les cases à cocher ,  et  : . On obtient alors un graphique et une carte qui montre mieux l'attractivité des Kreise de l'ex-Allemagne de l'Ouest (fig. 17.16).

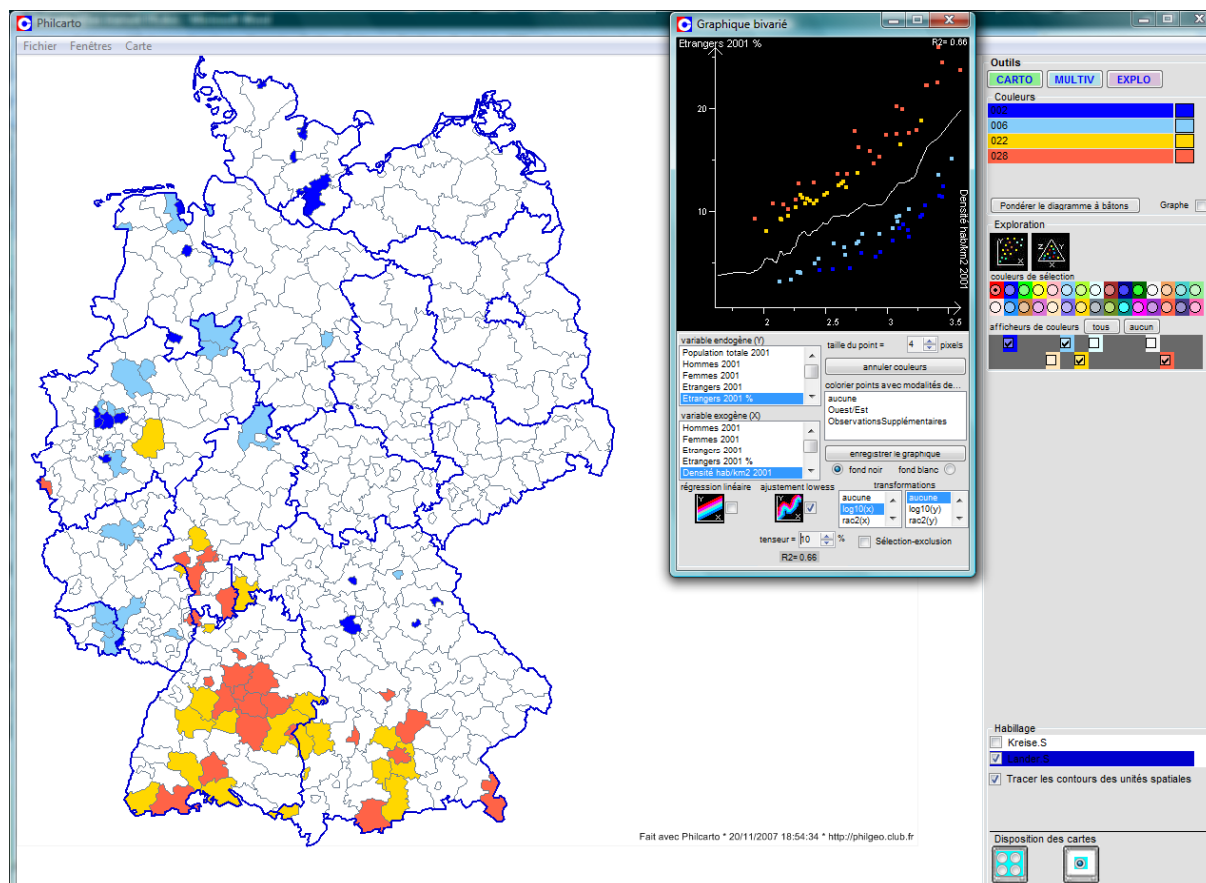


Figure 17.16 : Le bureau de *Philcarto* avec un ajustement Lowess et la carte des résidus les plus forts.

Pour conclure à propos des graphiques bivariés, soulignons que le bouton  permet de créer ou de sélectionner un dossier et d'y enregistrer non seulement le graphique, mais aussi les caractéristiques statistiques des sélections (moyenne, écart-type...) sur toutes les variables du tableau de données, et pour chacune des couleurs présentes sur le graphique.

17.3. Le diagramme triangulaire et l'exploration trivariée

Dossier : \Philcarto exemples\Brésil utilisation sol\
Fond de carte : Brésil carte micro-régions.ai
Statistiques : Brésil data micro-régions.xls

Dans un diagramme triangulaire, la somme des valeurs pour chaque unité spatiale doit toujours être égale à 100. Si l'addition des trois variables choisies ne correspond pas à 100, *Philcarto* calcule le pourcentage de chacune d'elles. On peut donc fournir en entrée du diagramme triangulaire :

- soit trois variables de pourcentages dont la somme est égale à 100 ;
- soit trois variables d'effectifs dont la somme a une signification (ne pas additionner des choux et des navets) et constitue de ce fait une population pour laquelle chaque variable définit une sous-population dont on peut calculer le pourcentage dans cette population totale.

L'exploration trivariée commence par un clic sur l'icône du graphique diagramme triangulaire  ; une fenêtre intitulée  **Diagramme triangulaire** apparaît à la gauche du cadre **Outils** (fig. 17.17). Elle est équipée de trois listes de variables permettant de sélectionner les trois variables, X, Y et Z ; X est l'axe horizontal gradué de gauche à droite, Y l'axe situé sur la droite du triangle et gradué de bas en haut, et Z l'axe situé sur la gauche du triangle et gradué de haut en bas.

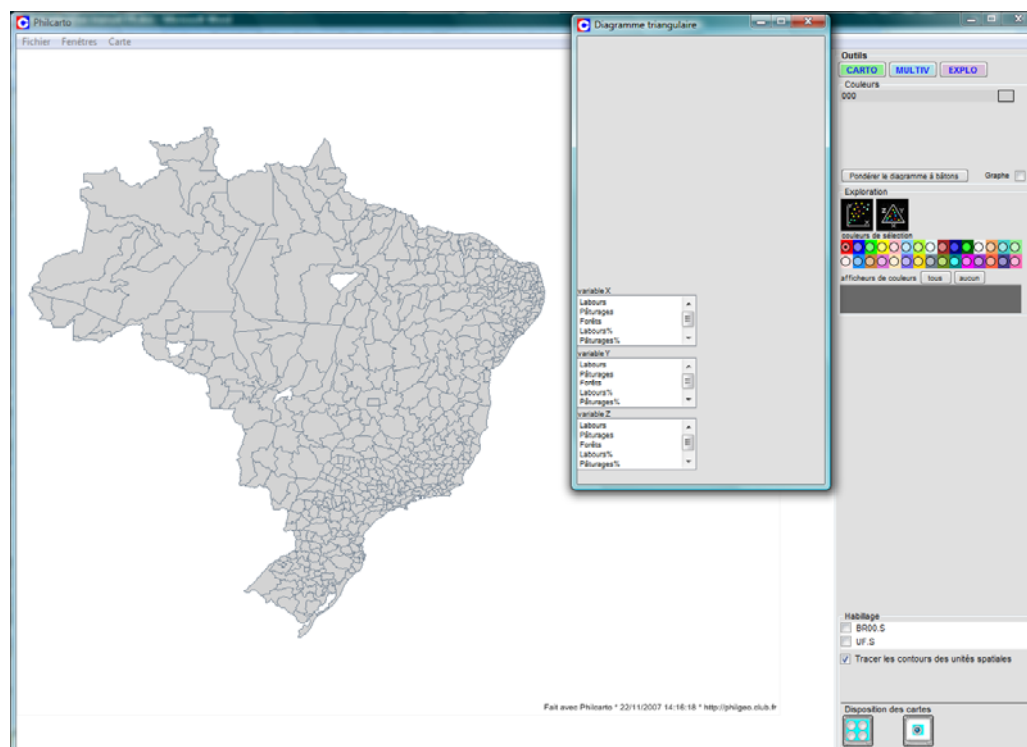


Figure 17.17 : Le bureau de *Philcarto* en début d'analyse exploratoire trivariée.

Par exemple, si l'on souhaite étudier les différents types d'utilisation du sol dans les micro-régions brésiliennes, il suffit de sélectionner les variables **Labours%**, **Pâturages%** et **Forêts %**. Dès que trois variables sont sélectionnées, le diagramme triangulaire est immédiatement tracé. Simultanément de nombreux contrôles sont ajoutés dans la fenêtre **Diagramme triangulaire**. Le dialogue **taille du point= 2 pixels** permet de régler la visibilité des points, ici 4 pixels (fig. 17.18).

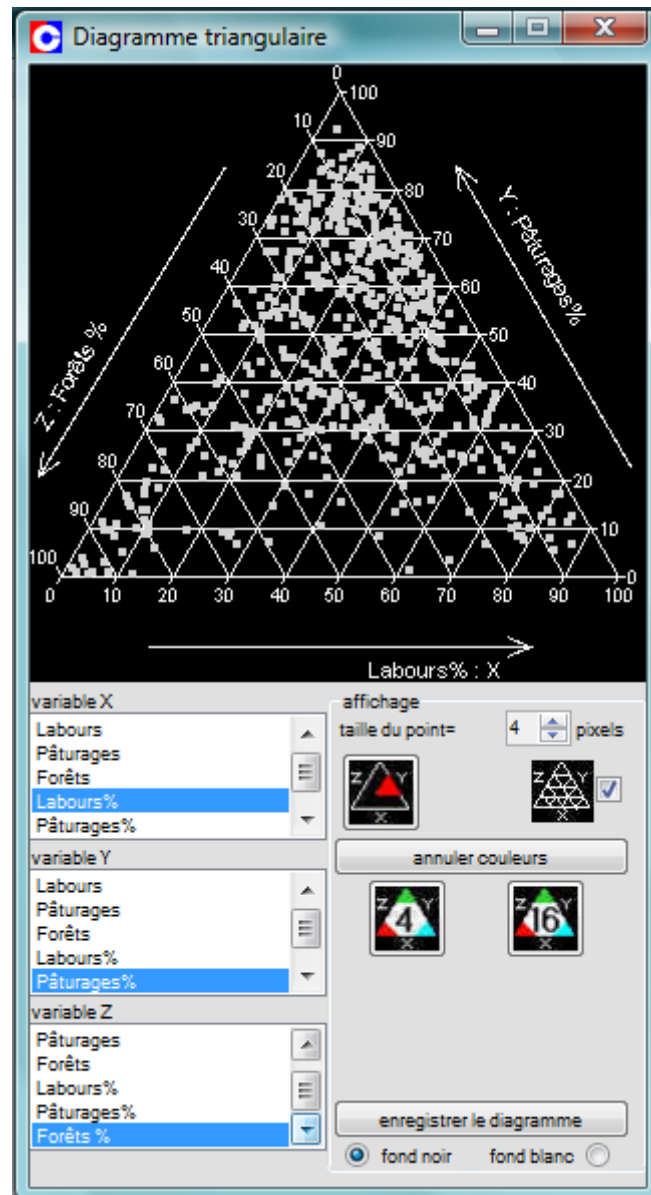


Figure 17.18 : Le diagramme triangulaire des trois types d'utilisation du sol.

La sélection au lasso sur le diagramme triangulaire fonctionne de la même façon que sur le graphique bivarié (§17.2). Si l'on souhaite sélectionner, par exemple, les micro-régions où prédominent les labours, il suffit d'entourer les points situés dans l'angle inférieur droit du triangle pour qu'elles soient sélectionnées à la fois dans le diagramme triangulaire et sur la carte (ici en marron car cette couleur a été préalablement choisie dans le dialogue couleur de sélection). On peut faire de même pour la forêt (en vert dans l'angle inférieur gauche) et les pâturages (en jaune dans l'angle supérieur). La carte obtenue montre l'importance de la forêt en

Amazonie, des pâturages dans le Centre-Ouest (Cerrados) et le sud du Rio Grande do Sul, et des labours à la fois sur le littoral nordestin, dans l'intérieur de São Paulo, et à l'Ouest de Santa Catarina (fig. 17.19).

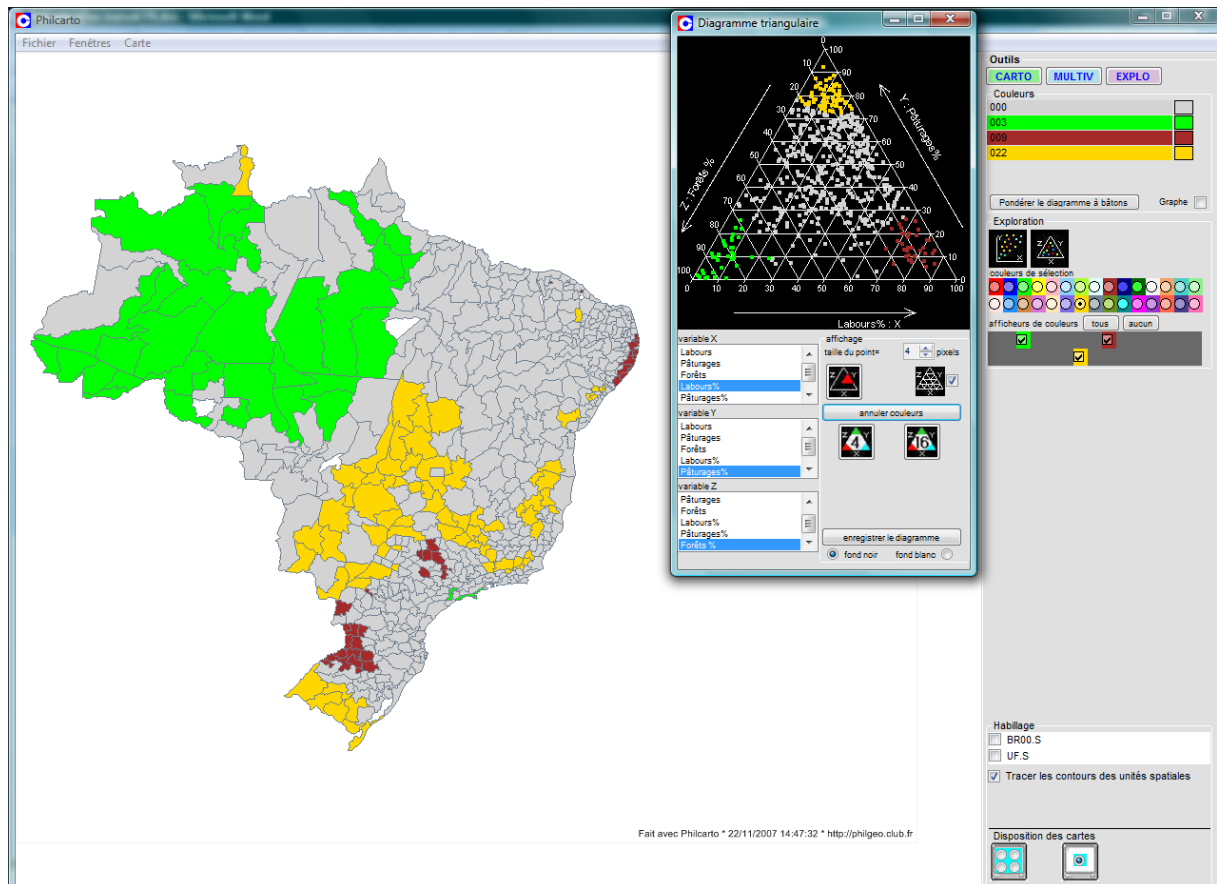


Figure 17.19 : Le bureau de *Philcarto* après la sélection des points situés dans les trois angles du diagramme triangulaire.

Comme une façon d'explorer le diagramme triangulaire consiste à sélectionner les points situés dans les angles, deux boutons permettent de réaliser automatiquement une partition géométrique du triangle soit en 4 classes , soit en 16 classes . Ces partitions sont basées sur des sous-triangles équilatéraux délimités à l'intérieur du triangle principal pour 4 classes, ou des 4 premiers sous-triangles pour 16 classes (fig. 17.20)

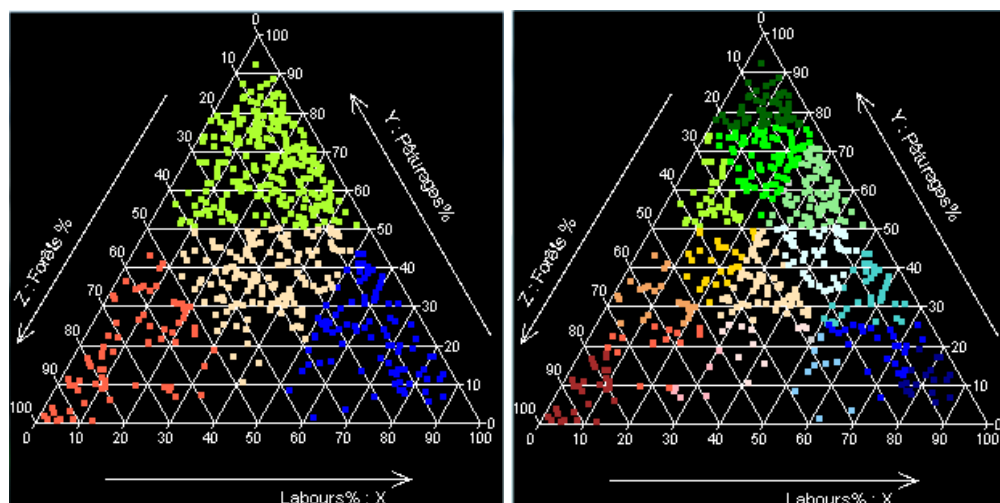


Figure 17.20 : Le diagramme triangulaire partitionné géométriquement en 4 classes et en 16 classes.

Avec une partition en 4 classes, la carte obtenue présente une structure simple avec une occupation massive du territoire par la forêt dans la Région Nord, des pâturages dans toute la partie centrale du pays, et des labours dans l'intérieur de São Paulo, et à l'Ouest de Santa Catarina (fig. 17.21). De son côté, la carte de la partition en 16 classes conserve les principales césures observables sur la carte en 4 classes, mais montre de nombreuses nuances, notamment dans les espaces les plus monolithiques.

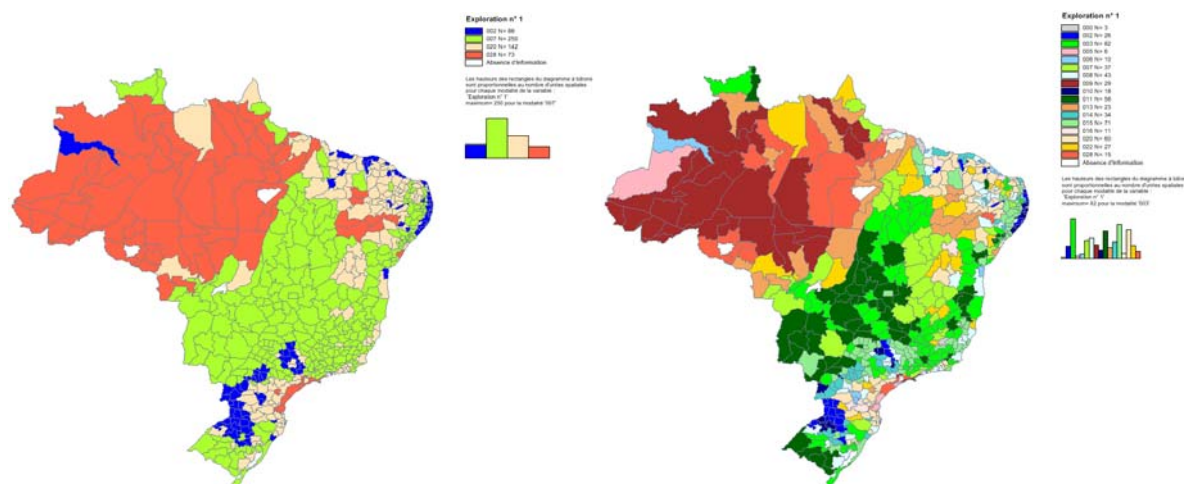


Figure 17.21 : Le diagramme triangulaire partitionné géométriquement en 4 classes et en 16 classes.

Le bouton [enregistrer le diagramme](#) permet de créer ou de sélectionner un dossier et d'y enregistrer non seulement le diagramme triangulaire (sur fond noir ou blanc, en fonction du bouton sélectionné ☒ fond noir ☐ fond blanc), mais aussi les caractéristiques statistiques des sélections (moyenne, écart-type...) sur toutes les variables du tableau de données, et pour chacune des couleurs présentes dans le diagramme.

Dans certains cas (mais pas ici), les points sont concentrés dans un secteur du diagramme. Il faut alors zoomer dans le diagramme pour mieux analyser la répartition des points. Le bouton  permet de réaliser ce zoom : après avoir cliqué sur ce bouton, il faut définir la base d'un sous-triangle par deux clics

successifs. Le sous triangle est alors tracé puis le zoom s'effectue. Les partitions en 4 ou 16 classes réalisées ensuite le sont dans les limites de sous triangle de zoom.

Le bouton  permet de réafficher à nouveau le graphique sans zoom.

Lorsque le nuage de points est très dense, sa forme est plus lisible si l'on supprime le treillis des sous-triangles de repérage en décochant la case . Ainsi, il est plus facile d'identifier chaque point en plaçant le curseur dessus (fig. 17.22). Le réaffichage du treillis se fait en cochant à nouveau cette case.

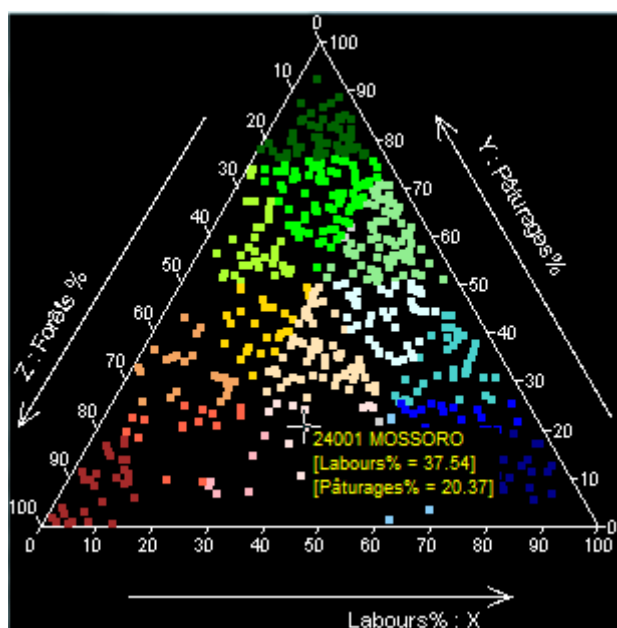


Figure 17.22 : Le diagramme triangulaire sans le treillis de repérage et avec un point identifié.

Si des variables nominales existent dans le tableau de données, leurs noms sont listés dans une liste intitulée **colorier points avec modalités de...** Lorsqu'une CAH a été préalablement réalisée sur les mêmes variables que celles du diagramme triangulaire, on trouve dans cette liste les variables correspondant aux 11 partitions (fig. 17.23).

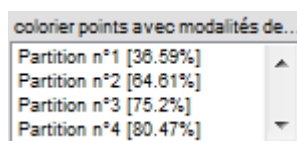


Figure 17.23 : La liste des partitions issues de la CAH pour colorier les points du diagramme.

En cliquant successivement sur le nom de chaque partition, il est possible de les visualiser l'une après l'autre et d'observer ainsi les emboîtements successifs des classes dans le graphique et sur la carte (fig. 17.24). L'interprétation de cette carte ne pose pas de problème particulier :

- **rouge** : domination des labours ;
- **rose** : mixte labours et pâturages ;
- **bleu** : domination des pâturages ;
- **jaune** : mixte pâturages et forêt ;
- **vert** : domination de la forêt.

Ces couleurs sont les mêmes que celles utilisées pour tracer les diagrammes à bâtons des profils des classes de la CAH (fig. 17.25).

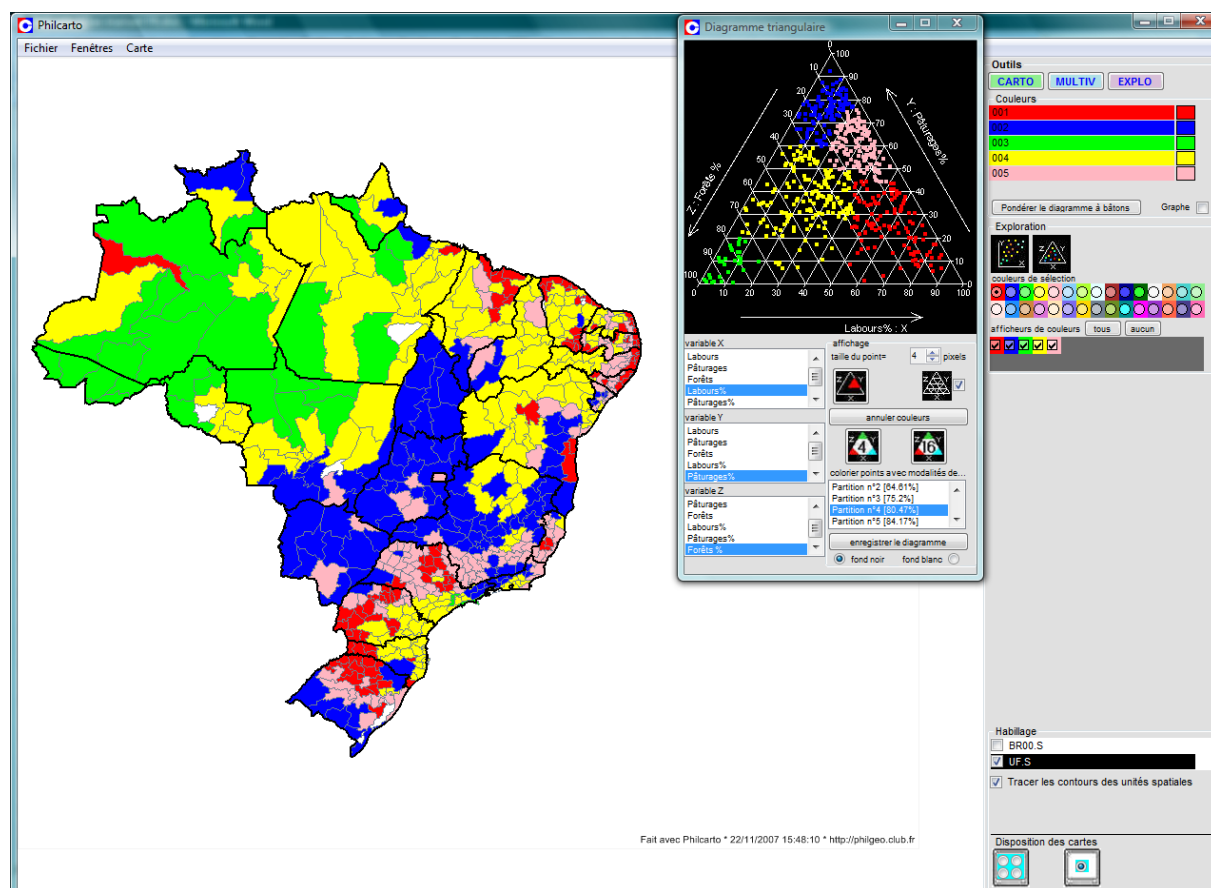


Figure 17.24 : Le bureau de *Philcarto* après la sélection des points en fonction des 5 classes de la partition n°4 issue de la CAH sur les mêmes variables que celles du diagramme triangulaire.

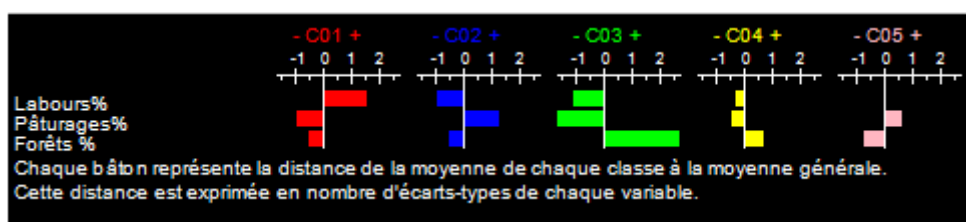


Figure 17.25 : Les profils des classes de la CAH sur les 3 types d'utilisation du sol.

18. L'Analyse Territoriale Multiscale

Dossier : \Philcarto exemples\Union Européenne NUTS 2\
Fond de carte : Union Européenne carte NUTS 2.ai
Statistiques : Union Européenne data NUTS 2.xls

L'analyse Territoriale Multiscale, ATM, est une nouvelle méthode de recherche proposée par Claude Grasland et Liliane Lizzi (UMS 2414 RIATE, CNRS-DATAR-Université Paris 7). Développée dans le cadre du projet européen *Hypercarte*, l'ATM repose sur l'idée qu'il n'est pas possible d'apprécier la situation d'un territoire particulier sans prendre en considération à la fois les caractéristiques des territoires voisins et les différentes entités régionales de plus ou moins grandes dimensions auquel il appartient. L'ATM prête ainsi une attention particulière à l'agrégation et à la désagrégation des éléments constitutifs des territoires aux niveaux local, régional et national (quelle que soit la terminologie utilisée, l'idée reste la même, celle qu'une entité territoriale appartient en même temps à plusieurs territoires de dimensions successivement plus étendues).

Un rapport issu du projet *Hypercarte* disponible sur Internet à l'adresse :

http://www.ums-riate.com/documents/MTA_4th%20version-livret2.pdf

montre que l'ATM consiste à prendre en compte, pour chaque entité de base de l'Union Européenne (par exemple chacune des 280 unités spatiales du maillage NUTS2), trois types de déviations (ou distance, ou écart) :

- une déviation par rapport à l'ensemble de l'UE (dite « macro » dans *Philcarto*) ;
- une déviation par rapport au pays dont l'entité fait partie (dite « meso » dans *Philcarto*) ;
- une déviation par rapport à ses voisins immédiats (dite « micro » dans *Philcarto* »).

Ces déviations sont calculées pour un indicateur statistique qui prend la forme d'un rapport, par exemple, le produit intérieur brut par habitant en parités de pouvoir d'achat (PIB en PPA). Pour pouvoir calculer cet indicateur, il faut disposer du PIB en PPA et de la population de chaque unité spatiale : l'ATM exige donc des quantités, des effectifs, des valeurs... pour être mise en œuvre.

Au niveau macro, le calcul de la déviation revient à calculer la différence entre le PIB en PPA par habitant de chaque région NUTS2 et le PIB en PPA par habitant de l'Union Européenne. Une différence positive exprime une situation meilleure que celle de l'UE, en moyenne ; une différence négative traduit une position en retrait par rapport à la moyenne de l'UE.

Au niveau meso, le calcul de la déviation revient à calculer la différence entre le PIB en PPA par habitant de chaque région NUTS2 et le PIB en PPA par habitant du pays dans lequel est située la région NUTS2 en question. Une différence positive exprime une situation meilleure que celle du pays d'appartenance, en moyenne ; une différence négative traduit une position en retrait par rapport à la moyenne de l'UE.

Au niveau micro, le calcul de la déviation revient à calculer la différence entre le PIB en PPA par habitant de chaque région NUTS2 et le PIB en PPA par habitant de l'ensemble formé des régions NUTS2 ayant une limite commune avec la région en question. Une différence positive exprime une situation meilleure que celle des

régions voisines, en moyenne ; une différence négative traduit une position en retrait par rapport à la moyenne de l'UE.

Ces déviations sont calculées sous la forme d'un indice dans lequel la base 100 correspond à la valeur de l'espace de référence (ici, l'UE au niveau macro, le pays au niveau meso, et l'ensemble des voisins au niveau micro). Il est recommandé d'utiliser une gamme de couleurs en opposition avec 100 comme valeur pivot.

Il n'y a donc dans l'ATM rien de bien compliqué sur le plan mathématique puisqu'il s'agit seulement d'arithmétique. Seule la recherche des unités spatiales voisines fait appel à un algorithme sophistiqué qui permet d'éviter de définir au préalable une matrice de contiguïté fastidieuse à obtenir et à saisir.

L'originalité de l'ATM est donc plus conceptuelle que technique : obtenir trois vues différentes sur les mêmes données en fonction d'un espace de référence variable constitue une réelle avancée scientifique car cela permet enfin de prendre en considération l'idée selon laquelle la « réalité » apparaît différente en fonction de l'échelle d'analyse.

Une fois les trois déviations calculées, il est intéressant de les combiner pour obtenir une typologie. Les auteurs proposent d'étudier les déviations deux à deux. Avec *Philcarto*, il est possible de réaliser une classification des unités spatiales en faisant une classification ascendante hiérarchique (CAH) sur le tableau de mesures formé par les trois variables contenant les déviations.

L'ouverture du fond de carte et du fichier de données statistiques se traduit par l'apparition des noms des variables dans les listes **couleurs** et **symboles**. On trouve dans la liste **symboles** les variables qui vont être mises en rapport pour calculer, par exemple, le PIB en PPA par habitant, c'est-à-dire les variables **Population Totale 1999** et **PIB PPA 1999** ; le fait que ces variables se trouvent dans la liste **symboles** indique qu'il s'agit de données brutes, d'effectifs et de valeurs (fig. 18.1).

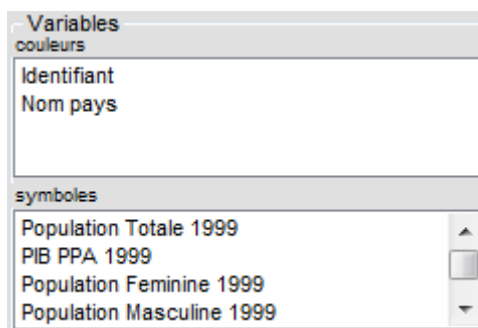


Figure 18.1 : les listes de variables pour l'ATM.

Dans la liste **couleurs** figurent deux variables : **Identifiant** contient les identifiants des unités spatiales (les mêmes que ceux utilisés pour coder le fond de carte). Ces identifiants étant hiérarchiques, les deux premiers caractères donnent les codes des pays (AT pour Autriche, FR pour France...) ; on peut donc en extraire les codes de pays et les utiliser pour définir le niveau meso. **Nom pays** donne, pour chaque unité spatiale NUTS 2, le nom en clair de son pays d'appartenance. Ce nom permet de réaliser directement la carte choroplèthe des pays, chaque nom étant considéré comme une modalité de la variable nominale **Nom pays** (fig. 18.2).

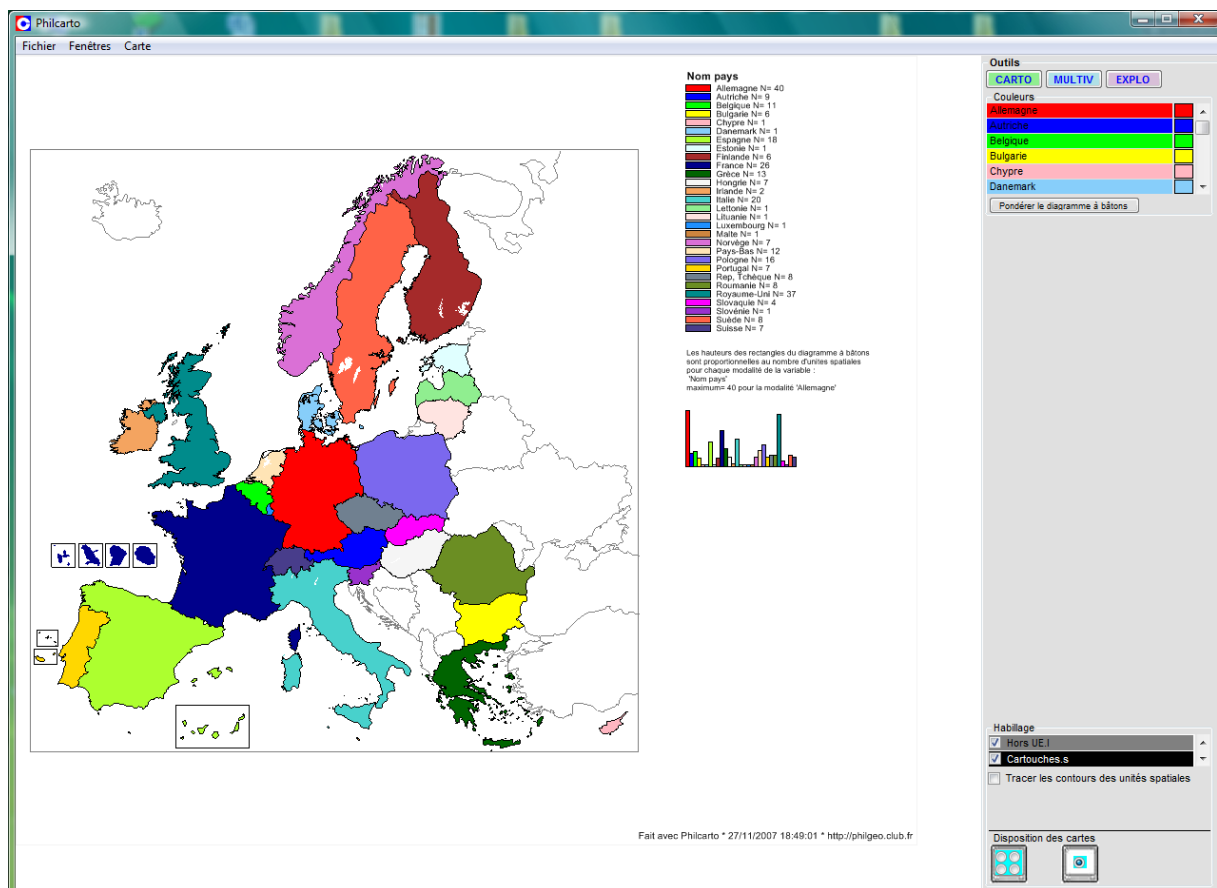


Figure 18.2 : Le bureau de *Philcarto* avec la carte choroplèthe des noms de pays ; les contours des unités NUTS 2 ne sont pas tracés ; trois habillages complètent la carte : contours des pays de l'UE, contours des pays hors UE, et cartouches des îles.

Philcarto propose une interface simple pour le calcul des déviations aux trois niveaux macro, meso et micro. On accède à cette fonction par l'article **Analyse Territoriale Multiscale** du menu **Fenêtres**. (fig. 18.3)

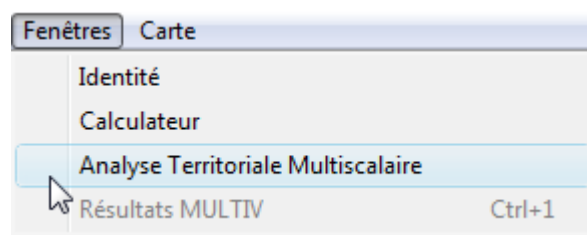


Figure 18.3 : L'accès à la fonction d'ATM de *Philcarto*.

Un dialogue intitulé **Analyse Territoriale Multiscale** s'affiche à l'écran (fig. 18.4). Il contient sur la gauche deux listes de variables de quantités ou d'effectifs. Dans la liste supérieure, on sélectionne la variable du numérateur, par exemple **PIB PPA 1999**. Dans la seconde liste, située au-dessous de la précédente, on choisit la variable du dénominateur, par exemple **Population Totale 1999**. Le résultat du calcul revient donc à calculer le PIB en PPA par habitant en 1999.

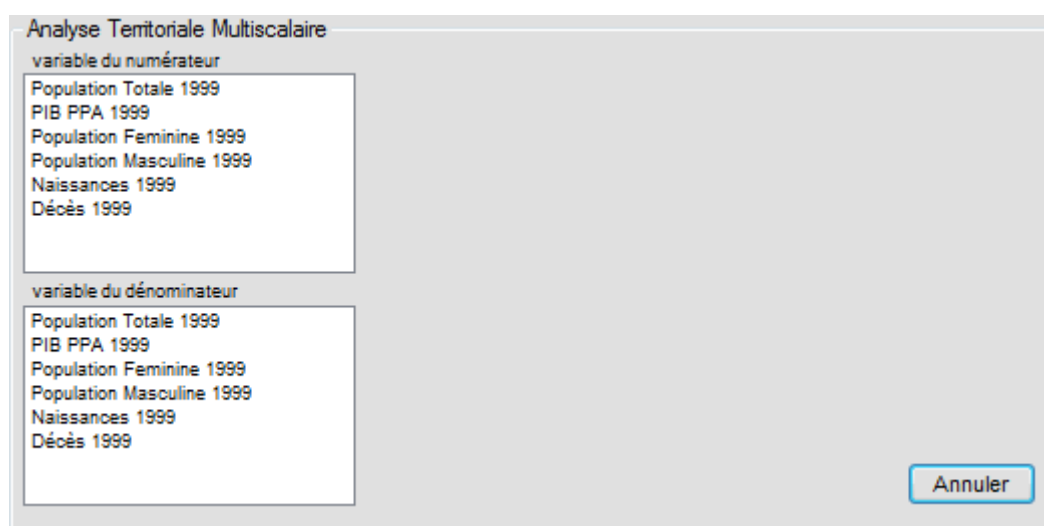


Figure 18.4 : Le dialogue **Analyse Territoriale Multiscale**.

Une fois choisies les variables du numérateur et du dénominateur, des éléments complémentaires de dialogue font leur apparition (fig. 18.5). Ils permettent : de choisir le nombre de niveaux de l'ATM (trois en principe, deux dans certains cas) et d'indiquer comment est codée l'échelle meso dans les données. Deux possibilités de codage sont proposées : soit les codes du niveau meso sont inclus dans l'identifiant, et dans ce cas il faut indiquer comment l'extraire, soit ce code constitue les modalités d'une variable discrète et dans ce cas, il suffit d'en indiquer le nom dans la liste figurant sous le bouton radio ☒ contenu par la variable...

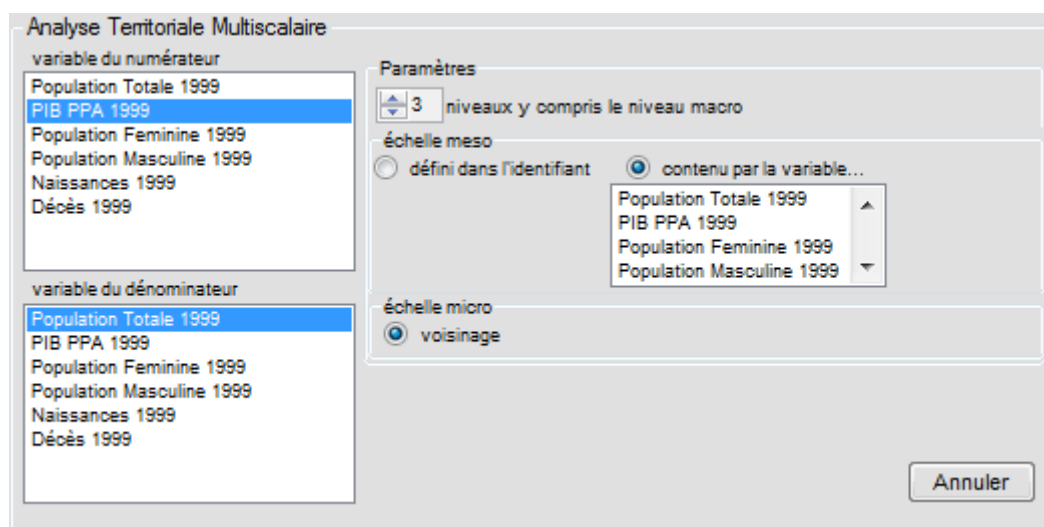


Figure 18.5 : Le choix des paramètres de l'ATM.

Sélectionner une variable nominale contenant les codes du niveau meso est la manière la plus simple de définir le niveau meso. Ici, il suffit de sélectionner la variable Nom pays (fig. 18.6). Cela suppose qu'on ait prévu une telle variable dans le tableau de données.

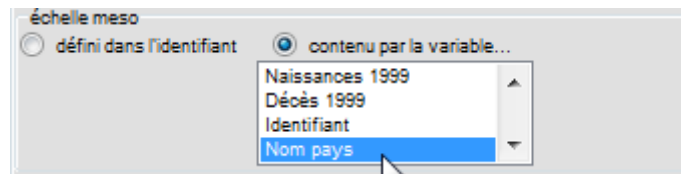


Figure 18.6 : La variable Nom pays contient le nom du pays d'appartenance de chaque unité spatiale.

S'il n'existe pas de variable définissant le niveau meso dans le tableau de données, il est possible de l'extraire des identificateurs des unités spatiales si celui-ci a été conçu sur une base hiérarchique incluant ce niveau. Dans le cas des pays de l'UE, les deux premières lettres de l'identifiant donnent le code du pays. Dans ce cas, on choisit le bouton radio ☒ défini dans l'identifiant. Un dialogue complémentaire permet d'indiquer à quel caractère commence le codage du niveau meso et combien de caractères il faut pour le définir complètement (fig. 18.7). Un exemple est donné correspondant à la première unité spatiale du tableau de données : si l'identifiant est AT11, le code du pays est AT ; il commence au premier caractère et comprend deux caractères. Il faut donc régler en conséquence les curseurs **caractère initial** et **longueur** (fig. 18.7). Naturellement, le format de l'identifiant doit être toujours le même, avec le code du niveau meso toujours à la même place et toujours de même longueur.

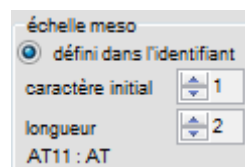


Figure 18.7 : Le dialogue d'extraction du niveau meso à partir de l'identifiant de l'unité spatiale.

Lorsque le niveau meso est défini, un clic sur **Calculer** déclenche le calcul des écarts. Dans ce calcul, l'étape la plus longue concerne le niveau micro puisqu'il est nécessaire de rechercher tous les voisins de chaque unité spatiale.

 **Au niveau micro, les résultats produits par *Philcarto* sur l'exemple de l'Union Européenne diffèrent légèrement de ceux produits par *Hypercarte* en raison du fond de carte utilisé ici, moins détaillé que celui du projet *Hypercarte*, ce qui produit quelques distorsions par rapport à ce dernier. Autrement dit, les résultats sont légèrement sensibles au degré de généralisation du fond de carte à partir duquel sont établis les voisinages.**

A l'issue de tous les calculs, le bureau de *Philcarto* s'affiche à nouveau. Dans la liste **couleurs** du dialogue **Variables**, trois nouvelles variables ont pris place, correspondant aux trois niveaux de l'ATM (fig. 18.8). Le nom de ces variables indique qu'il s'agit de variables provenant de l'ATM, aux niveaux macro, meso et

micro ; la suite du nom indique quelles variables ont permis le calcul des écarts, ici le PIB en PPA 1999 divisé par la population totale en 1999.

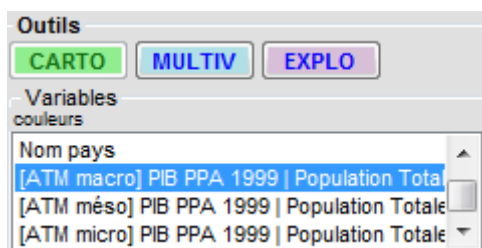
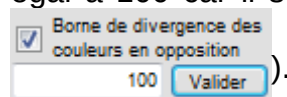


Figure 18.8 : Les trois nouvelles variables créées par l'ATM.

Ces trois nouvelles variables peuvent être cartographiées de différentes façons ; nous présentons ci-après les trois cartes choroplèthes. La discrétisation est celle de Jenks en 8 classes, avec une gamme de couleurs en opposition et un pivot égal à 100 car il s'agit d'un indice (penser à donner cette valeur dans le dialogue



La carte du niveau macro (fig. 18.9) est des plus connues. Elle montre clairement l'existence d'une césure qui correspond aux deux anciens blocs. Du côté de l'Est, les anciennes RDA et Tchécoslovaquie apparaissent en moins mauvaise posture que la Pologne, la Hongrie, la Roumanie et la Bulgarie. Du côté de l'Ouest, la « banane » joignant la moitié Nord de l'Italie au Sud de la Grande Bretagne est clairement visible avec des écarts le plus souvent positifs ; les capitales occupent généralement une position favorable (Londres, Paris, Oslo, Stockholm, Helsinki...).

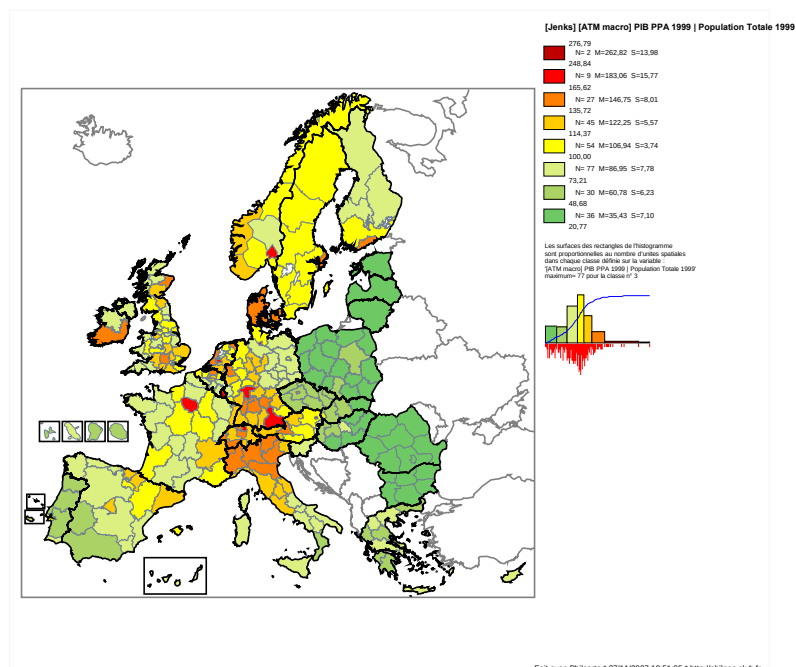


Figure 18.9 : Carte des écarts au PIB en PPA par habitant de l'Union Européenne (niveau macro).

Sur la carte du niveau meso (fig. 18.10), le clivage Est/Ouest n'existe plus. En revanche, d'importantes différences de certains pays de l'UE font leur apparition, comme par exemple entre l'Est et l'Ouest de la Pologne. D'autres pays, au contraire, apparaissent plus homogènes, comme la France ou la Norvège.

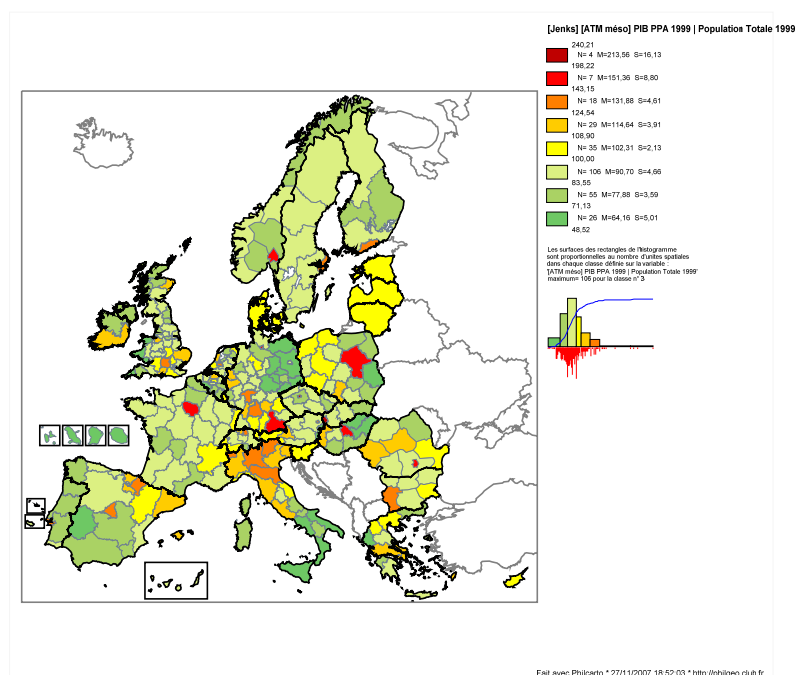


Figure 18.10 : Carte des écarts au PIB en PPA par habitant de chaque pays (niveau meso).

La carte du niveau micro (fig. 18.11) montre l'importance des centres urbains. Bien sûr, les capitales des Etats de l'UE apparaissent déjà en meilleure posture que leur espace environnant sur les deux précédentes cartes.

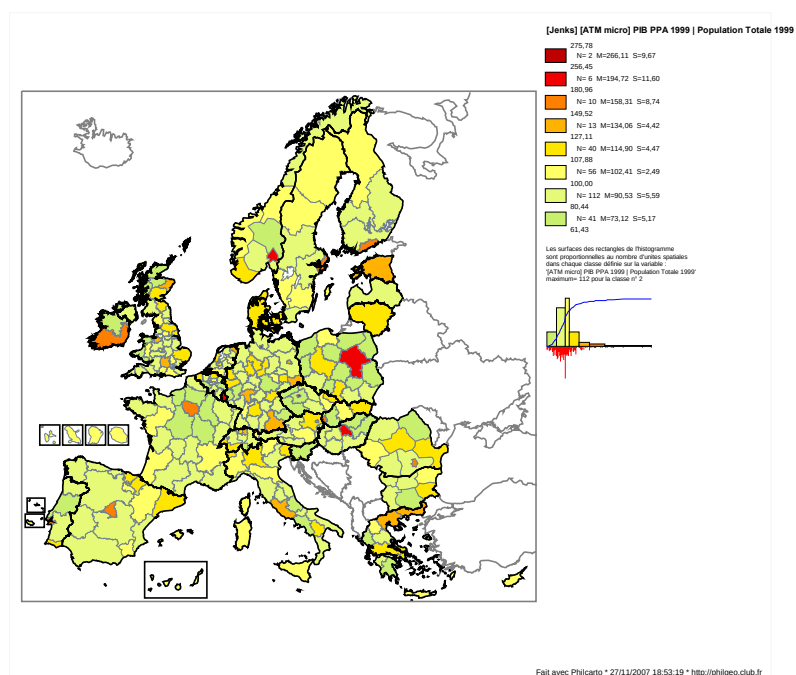


Figure 18.11 : Carte des écarts au PIB en PPA par habitant de chaque voisinage (niveau micro).

Ce qui est nouveau sur la carte du niveau micro, c'est la place plus confortable occupée par certaines régions abritant des villes de second niveau : Lyon, Bordeaux, Nantes en France, Edimbourg ou Liverpool au Royaume Uni, Stavanger, Bergen, Trondheim en Suède, Thessaloniki en Grèce. Notons cependant que les différences entre les cartes sont bien plus importantes entre le niveau macro et le niveau meso, qu'entre le niveau meso et le niveau micro.

Lorsque les écarts sont calculés, le second objectif de l'ATM peut être atteint : il s'agit de combiner les trois niveaux pour mettre en évidence leurs relations pour en dresser une typologie pouvant être cartographiée, elle aussi. *Philcarto* propose toute une gamme de méthodes pour réaliser une telle étude : analyse exploratoire des écarts considérés par paires sur un graphique bivarié ; analyse en composantes principales et classification automatique sur les trois niveaux simultanément... On se limite ici à l'exposé de l'ACP et de la CAH.

Comme les écarts sont enregistrés dans trois variables de couleurs, on peut leur appliquer une ACP (fig. 18.12).

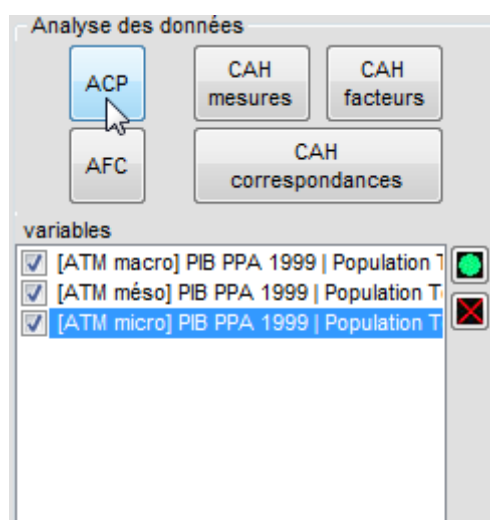


Figure 18.12 : La définition d'une ACP sur les trois niveaux d'écarts de l'ATM.

L'examen du cercle des corrélations (fig. 18.13) montre d'une part l'existence d'un premier facteur unipolaire absorbant 75,9% de l'inertie totale. Les trois niveaux sont localisés du côté gauche de l'axe des abscisses, ce qui signifie que les écarts sont en faveur des régions dont les coordonnées sur la première composante sont négatives. Cependant, le niveau macro se démarque de deux autres pour former la seconde composante qui n'absorbe que 16,4% de l'inertie totale ; ce niveau se situe également dans la partie négative de la seconde composante et suit donc le même schéma de lecture que pour la première : les écarts sont en faveur des régions dont les coordonnées sur la seconde composante sont négatives.

Autrement dit, les écart mesurés sur les trois niveaux sont convergents, mais si les niveaux meso et micro se différencient peu l'un de l'autre, en revanche, le niveau micro, lui, apporte une information supplémentaire utile à la compréhension de la différenciation régionale de l'Union Européenne.

La carte des unités NUTS 2 sur la première composante principale (fig. 18.14) souligne la position très favorable de la plupart des capitales des Etats de l'UE (en bleu foncé). En dehors de celles-ci, certaines régions comme le Pays Basque, la Catalogne, le Nord de l'Italie apparaissent en bonne position.

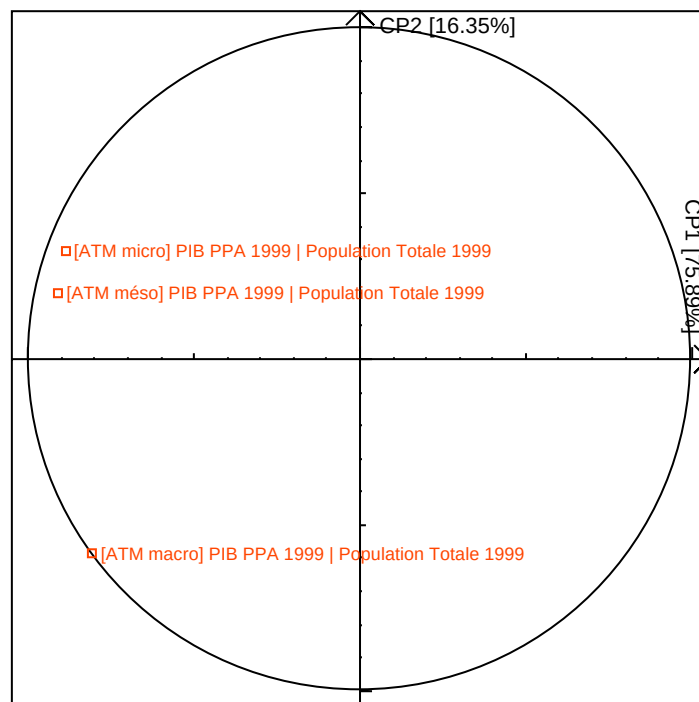


Figure 18.13 : La localisation des trois niveaux de l'ATM sur le cercle des corrélations de l'ACP.

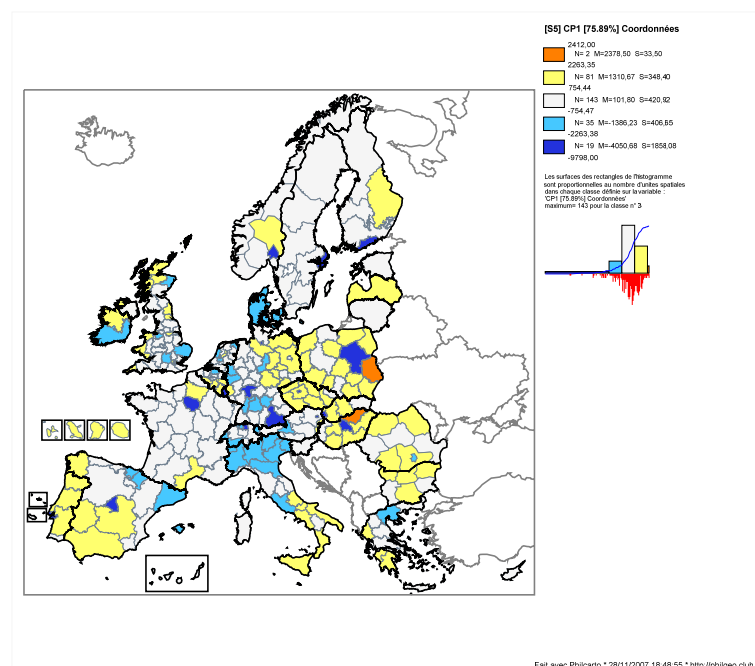


Figure 18.14 : La carte des unités NUTS 2 sur la première composante principale de l'ACP des écarts de l'ATM.

En revanche, le Sud de l'Espagne, le Portugal, le Sud de l'Italie et plus généralement les anciens pays de l'Est, demeurent en mauvaise posture.

La carte des unités NUTS 2 sur la seconde composante principale (fig. 18.15) transcrit de façon radicale l'opposition entre l'Europe Centrale, « banane bleue » élargie qui bénéficie d'un niveau de PIB très supérieur à la moyenne européenne et l'Europe Orientale qui demeure en situation précaire. L'Ouest et le Sud de l'Europe, de la Normandie à l'Andalousie participent peu à ce système d'opposition et occupent une position moyenne sur la seconde composante principale.

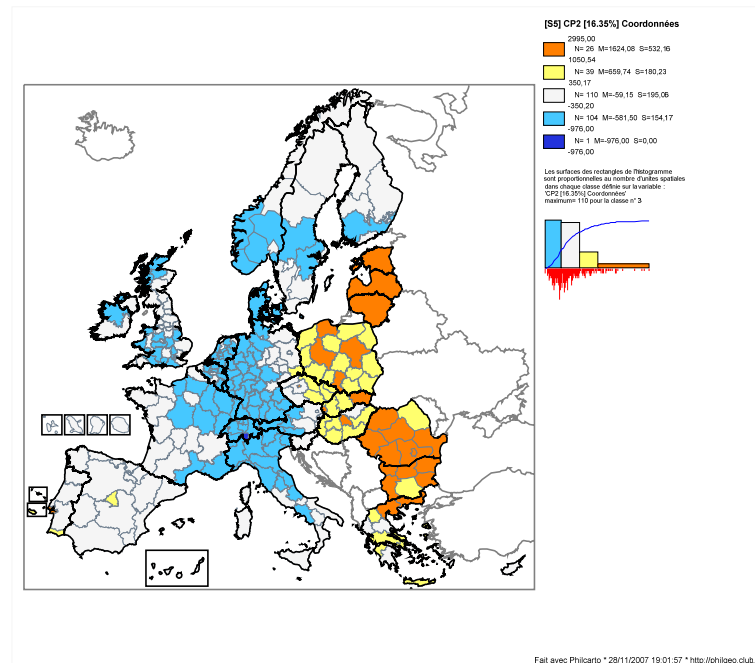


Figure 18.15 : La carte des unités NUTS 2 sur la première composante principale de l'ACP des écarts de l'ATM.

Une autre façon de combiner les trois niveaux d'écarts consiste à leur appliquer une CAH sur tableau de mesure. Avec 5 classes, l'inertie absorbée s'élève à 72,3% de l'inertie totale. Les profils des classes sont particulièrement bien discriminés (fig. 18.16) :

- la classe n°1 rassemble les régions qui présentent des écarts toujours négatifs avec une accentuation particulière pour le niveau macro. La carte (fig. 18.17) montre qu'il s'agit de la majeure partie de la Péninsule Ibérique, du Sud de l'Italie, de la majeure partie de la Grèce, et des pays de l'ancien bloc de l'Est. Toutes ces régions présentent un déficit important par rapport à la moyenne européenne, par rapport à leur moyenne nationale, et par rapport aux régions environnantes.
- dans la classe n°2, tous les écarts sont positifs. On y trouve donc les régions en bonne santé économique, souvent des capitales d'Etats, mais pas toutes ces capitales.
- la classe n°3 est aussi caractérisée par des écarts positifs sur les trois niveaux, mais de façon modérée pour le niveau macro, et très modérée pour les niveaux meso et micro. Les régions concernées sont parfois celles qui abritent des villes dynamiques.
- la classe n°4 caractérise les régions très dynamiques tirées vers le haut par de très grandes villes telles Londres, Paris, Hambourg, Oslo. Ce qui les caractérise au premier chef, c'est leur capacité à régner non seulement

sur le PIB européen, mais surtout sur le PIB de leur pays et de leur région. Ces villes concentrent les richesses produites à tous les niveaux.

- enfin, la classe n°5 concerne des régions qui ne sont pas en mauvaise posture, puisqu'elles se situent à un niveau moyen sur le plan européen, mais elles apparaissent peu productives au niveaux meso et surtout micro : une grande partie des régions françaises entrent dans cette catégorie, mais aussi un grand nombre de régions de l'ex Europe de l'Ouest.

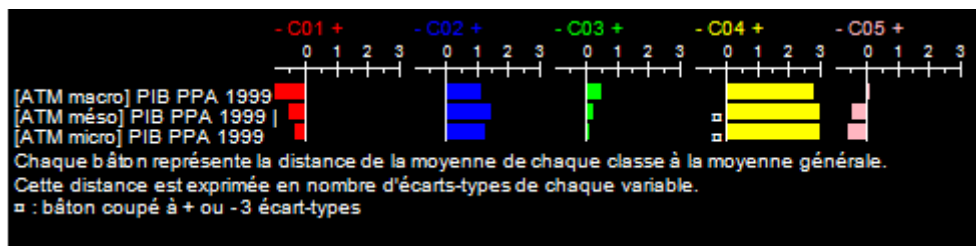


Figure 18.16 : Les profils des 5 classes de la CAH sur les trois niveaux d'écarts de l'ATM.

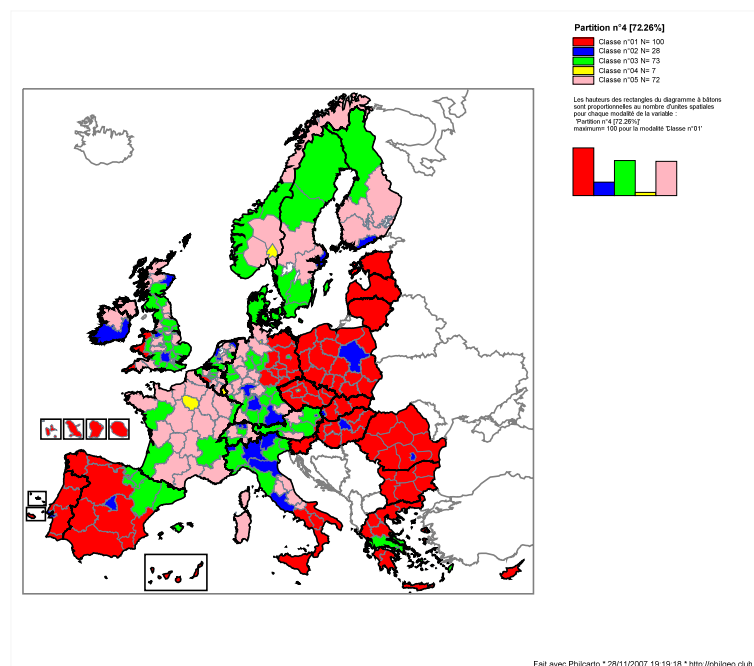


Figure 18.17 : La carte des 5 classes de la CAH sur les trois niveaux d'écarts de l'ATM.

19. Les données statistiques

19.1. Les matrices d'information spatiale

Les données dont il s'agit ici correspondent principalement à celles qui sont publiées, dans le monde entier, par les instituts nationaux de statistique en charge des recensements de toute sorte (population et logements, agriculture, etc.) ou d'enquêtes plus spécialisées (éducation, santé, etc.). Dans ces annuaires (ou leur équivalent informatique), l'espace national est divisé en régions, départements et communes (avec un nombre de niveaux et des dénominations qui peuvent différer d'un pays à l'autre) ; chaque élément de cette division de l'espace est désigné par le terme générique « unité spatiale ». Pour chaque thème recensé ou enquêté, on dispose de matrices (**M**) dont les lignes sont les unités spatiales qui correspondent, dans le fond de carte, aux éléments cartographiques ponctuels, linéaires ou surfaciques ; dans les colonnes figurent les variables, terme équivalent à « descripteurs », « attributs », « caractéristiques », « caractères », etc. Il s'agit de qualités, de propriétés ou de paramètres mesurables, au moyen desquels on peut décrire les unités spatiales et les comparer aux autres. Cela peut être un groupe de descripteurs, ou bien la partition d'un ensemble en catégories pertinentes vis-à-vis du thème étudié. Les qualités observées, les variables de l'analyse, peuvent être de nature qualitative ou quantitative, si bien qu'il convient de distinguer différents types de tableaux de données. Dans tous les cas, les données sont agrégées, ce qui signifie que l'on n'a jamais accès aux informations individuelles (dans le cas des recensements et autres enquêtes officielles, le cadre réglementaire garantit la confidentialité des réponses).

19.1.1. Les tableaux de correspondances

Ils réalisent le croisement de l'ensemble des unités spatiales (en lignes) avec celui des catégories d'objets ayant fait l'objet d'une mesure (en colonnes) ; par exemple, les communes d'un département et les catégories de superficie des exploitations agricoles qui s'y trouvent. La cellule formée par le croisement d'une ligne et d'une colonne contient le nombre d'exploitations (n_{ij}) appartenant à la catégorie **j**, relevées dans l'unité spatiale **i**. Dans ce genre de tableau, il est possible de calculer l'effectif total d'une ligne en sommant les valeurs des cases d'une même ligne, pour toutes les colonnes, noté ($n_{i.}$). De même, on peut calculer l'effectif total d'une colonne en sommant les valeurs des cases d'une même colonne, pour toutes les lignes, noté ($n_{.j}$). L'effectif total du tableau est noté ($n_{..}$).

$$M = \begin{bmatrix} n_{1,1} & n_{1,2} & \cdots & n_{1,J} \\ n_{2,1} & n_{2,2} & \cdots & n_{2,J} \\ \vdots & \vdots & \ddots & \vdots \\ n_{I,1} & n_{I,2} & \cdots & n_{I,J} \end{bmatrix}$$

Placés sur des éléments cartographiques ponctuels, les effectifs enregistrés dans les tableaux de correspondances peuvent être représentés en utilisant des

symboles dont la surface est proportionnelle à ces effectifs ; c'est le cas des cartes en cercles réalisées par *Philcarto*.

19.1.2. Les tableaux de mesures quantitatives de rapports ou d'intervalles

Comme dans le cas des tableaux de correspondances, les lignes des tableaux de mesures quantitatives représentent les unités spatiales. Mais ils en diffèrent car les colonnes contiennent des mesures enregistrées sur une échelle d'intervalle ou de rapport.

Sur une échelle d'intervalle, chaque unité spatiale possède une valeur qui lui est propre au regard du phénomène mesuré, et les différences entre observations sont exprimées dans la même unité de mesure, à partir d'une valeur zéro arbitraire. Les températures relevées dans des stations météorologiques, sur l'échelle Celsius ou Fahrenheit, sont des exemples évidents de ce genre de mesure. L'échelle de rapport possède les mêmes propriétés que la précédente, mais un zéro naturel au lieu d'un zéro arbitraire ; les distances entre lieux ou les superficies correspondent à cette échelle. Les tableaux de pourcentages sont également des tableaux de mesures quantitatives résultant du rapport entre deux effectifs et sont en général dérivés d'un tableau de correspondances.

Placées sur des éléments cartographiques surfaciques, les valeurs enregistrées dans les tableaux de mesures, de rapports ou d'intervalles peuvent être représentées à l'aide de plages de couleurs. On peut aussi colorier les surfaces des symboles proportionnels ; dans ce cas, on a affaire à un mode de cartographie bivarié : des cercles pour traduire les effectifs ou les quantités, des couleurs pour exprimer les mesures. *Philcarto* réalise l'un et l'autre type de carte. Dans ces deux cas, une étape préalable de discrétisation de la variable de mesure quantitative s'avère indispensable car seul un petit nombre de couleurs peut être distingué sur une carte (4 à 9 en général) ; pour partitionner l'étendue de la variable en classes, on recourt à l'une des méthodes de discrétisation plus ou moins automatiques proposées par *Philcarto*.

19.1.3. Les tableaux de valeurs nominales

Avec l'échelle nominale, les différentes valeurs n'ont aucun ordonnancement particulier et servent seulement à indiquer quelle modalité a été relevée sur une unité spatiale comme par exemple la couleur politique du maire de la commune. Lorsque les modalités ne prennent que deux valeurs pour exprimer, par exemple, la présence ou l'absence d'une espèce végétale, il s'agit d'une échelle binaire. Avec *Philcarto*, le nombre de modalités différentes ne doit pas dépasser 30.

Placées sur des éléments cartographiques surfaciques, les valeurs enregistrées dans les tableaux de mesures nominales ou ordinales peuvent être représentées à l'aide de plages de couleurs. On peut aussi colorier les surfaces des symboles proportionnels ; dans ce cas, on a affaire à un mode de cartographie bivarié : des cercles pour traduire les effectifs ou les quantités, des couleurs pour exprimer les mesures. Enfin, dans le cas d'éléments cartographiques linéaires, ce sont ces lignes qui sont directement coloriées. La gamme de couleurs ne présente volontairement aucun ordre ; elle est souvent désignée par le nom « gamme perroquet » car elle couvre l'étendue du bleu, au vert, au jaune et au rouge. Contrairement au cas des mesures de rapport ou d'intervalle, aucune discrétisation n'est réalisée : les modalités ordonnées selon l'alphabet sont cartographiées directement.

19.1.4. Les données manquantes

Il arrive assez fréquemment qu'on ne dispose pas de toutes les valeurs pour toutes les unités spatiales. Plusieurs cas de figure peuvent apparaître :

- les données manquent de manière sporadique dans quelques cases du tableau. Dans ce cas, les valeurs manquantes doivent être codées à l'aide du caractère **X** (majuscule) ;
- les données manquent dans toutes les cases d'une même ligne : la ligne peut être supprimée ;
- les données manquent dans toutes les cases d'une même colonne : la colonne peut être supprimée.

 Au moment de la lecture du fichier de données statistiques, *Philcarto* supprime les variables dont les valeurs, pour toutes les unités spatiales, sont manquantes ou égales à zéro et affiche un dialogue d'information.

19.1.5. Les fichiers des matrices d'information spatiale

Dossier : \Philcarto exemples\France départements\
Fond de carte : France carte départements.ai
Statistiques : France data départements.xls

La création, la saisie, la correction, et l'enregistrement des matrices d'information spatiale nécessite l'utilisation d'un logiciel extérieur à *Philcarto*. Le plus simple est d'employer un tableur du genre *Excel* pour éditer les cases du tableau de données.

Philcarto reconnaît deux types de format de fichier, le choix du format se faisant dans le dialogue  Fichier données statistiques ? :

- le format texte (**.txt**) avec des tabulations pour séparer les valeurs les unes des autres ;
- le format feuille de calcul *Excel* (**.xls**) Dans le cas d'un fichier *Excel* contenant plusieurs feuilles, *Philcarto* affiche le dialogue  Classeur Excel® qui permet de choisir la feuille sur laquelle il doit travailler (fig. 19.1).

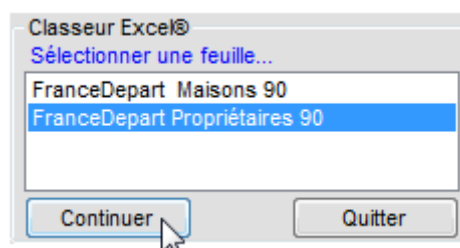


Figure 19.1 : La sélection d'une feuille de calcul dans un classeur *Excel* multifeuille.

La première colonne du fichier contient les identifiants des unités spatiales composés de chiffres et de lettres majuscules à l'exclusion de tout autre caractère (fig. 19.2). Naturellement, le même identifiant doit désigner la même unité spatiale dans le fichier fond de carte et dans le fichier de données statistiques.

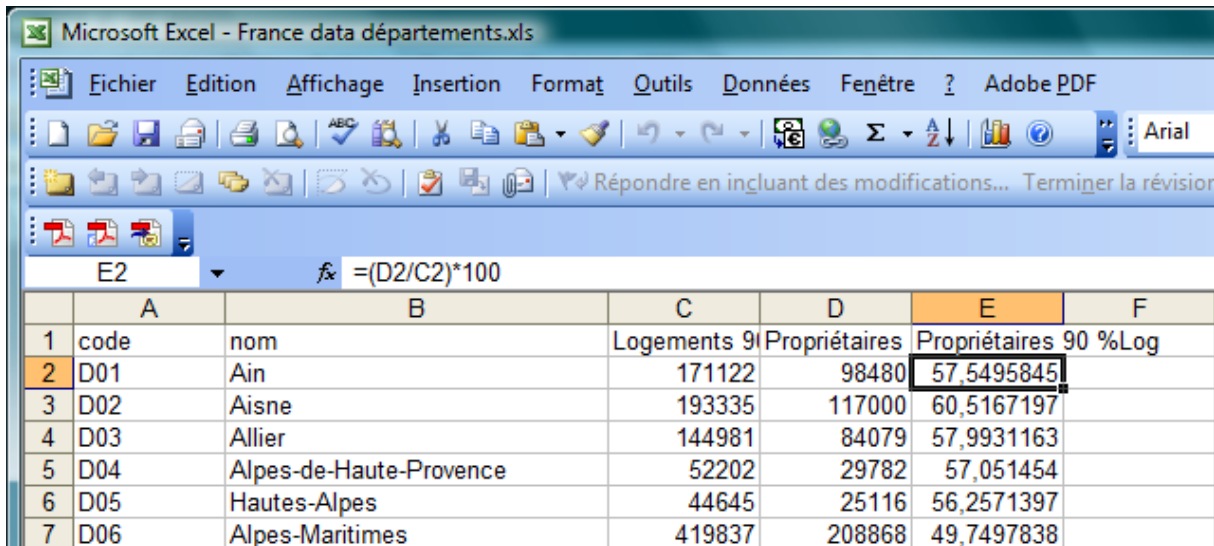
La seconde colonne peut renfermer, de manière optionnelle, le nom en clair de chaque unité spatiale. C'est ce nom qui apparaît dans la barre supérieure de la fenêtre *Philcarto* lorsqu'on déplace le pointeur de la souris sur une des unités spatiales de la carte affichée dans cette fenêtre.

Les colonnes suivantes contiennent les données statistiques à proprement parler. Pour simplifier l'exposé ci-dessus, on a distingué les matrices d'information spatiale en fonction de la nature des données qui s'y trouvent. Cependant, à l'intérieur d'un même fichier peuvent coexister des matrices de natures différentes sans que cela occasionne la moindre gêne à *Philcarto*.

code	nom	Logements 9	Propriétaires	Propriétaires 90 %	Log
D01	Ain	171122	98480	57,5495845	
D02	Aisne	193335	117000	60,5167197	
D03	Allier	144981	84079	57,9931163	
D04	Alpes-de-Haute-Provence	52202	29782	57,051454	
D05	Hautes-Alpes	44645	25116	56,2571397	
D06	Alpes-Maritimes	419837	208868	49,7497838	
D07	Ardèche	104854	66517	63,4377325	

Figure 19.2 : Un extrait de matrice d'information spatiale.

Lorsqu'on ouvre ce fichier avec *Excel* (fig. 19.3), on observe que la première ligne de la feuille est occupée par les noms des variables : la première cellule de cette première ligne contient le nom de la variable des identifiants des unités spatiales, au choix de l'utilisateur, la seconde celle de leurs noms en clair. Le nom de cette seconde variable est obligatoire **nom**, et cela dans toutes les langues.



	A	B	C	D	E	F
1	code	nom	Logements 9	Propriétaires	Propriétaires 90 %	Log
2	D01	Ain	171122	98480	57,5495845	
3	D02	Aisne	193335	117000	60,5167197	
4	D03	Allier	144981	84079	57,9931163	
5	D04	Alpes-de-Haute-Provence	52202	29782	57,051454	
6	D05	Hautes-Alpes	44645	25116	56,2571397	
7	D06	Alpes-Maritimes	419837	208868	49,7497838	

Figure 19.3 : Un extrait de matrice d'information spatiale ouvert dans *Excel*.

Les noms des variables suivantes sont aussi au choix de l'utilisateur. Sachant que ces noms vont apparaître dans *Philcarto* dans des listes déroulantes permettant de choisir les variables à cartographier, il est préférable de définir des noms courts (une dizaine de caractères) et si possible significatifs ; éviter cependant des noms du genre **Var01**, **Var02**, etc.

Les noms des variables peuvent être typés sous la forme d'un suffixe qui permet à *Philcarto* de distribuer les variables dans les listes **couleurs** et **symboles** de

 Dans ce fichier les valeurs décimales sont indiquées par des virgules. Ceci est la notation officielle française. *Philcarto* interprète comme marqueur des décimales la virgule ou le point qui, lui, est la notation américaine. Il est important de souligner qu'il ne faut en aucun cas utiliser les séparateurs de milliers dans les fichiers textes car *Philcarto* ne saurait pas les interpréter correctement.

 Il est souhaitable que les noms de variables figurant dans la première ligne du tableau soient courts, une vingtaine de caractères tout au plus.

 Dans les feuilles de calcul *Excel*, les cellules peuvent contenir des formules. *Philcarto* évalue le résultat de ces formules au moment de la lecture de la feuille.

A côté d'un tableur qui permet de saisir et corriger les données, on conseille aussi d'avoir à portée de main un éditeur de texte permettant de regarder de plus près le texte enregistré par *Excel*. *WordPad* fourni en standard avec *Windows* est trop limité. *UltraEdit* est recommandé car il permet de repérer très rapidement les éventuels problèmes rencontrés au cours de la lecture d'un fichier texte. Par exemple, il arrive que des lignes vides soient enregistrées par *Excel* en fin de fichier. *UltraEdit* permet de repérer ces lignes et de les supprimer très facilement. Ce *shareware* peut être téléchargé à l'adresse : <http://www.UltraEdit.com>

19.2. Les matrices de liens

La réalisation de cartes de liens nécessite l'enregistrement d'un type de tableau de données particulier, la matrice de liens spatiaux. Ces tableaux permettent de définir des liens entre deux points, nommés point d'origine et point de destination.

 Comme *Philcarto* trace un segment non orienté entre deux points, la notion de point d'origine et de point de destination n'a pas de signification particulière, origine et destination étant interchangeables lors du tracé du segment. Cependant, sachant qu'on peut ajouter une flèche avec un logiciel de dessin soit sur l'origine soit sur la destination, il faut veiller, si l'on souhaite pouvoir utiliser cette possibilité, à indiquer correctement quel est le point d'origine et quel est le point de destination. Cette précaution dans la constitution du fichier de données permet d'ajouter une flèche à tous les segments d'une même carte, en une seule opération.

Chaque enregistrement du fichier correspond à un point d'origine pour lequel l'enregistrement décrit au moins un point de destination (plusieurs points de destination si un même point d'origine est relié à plusieurs points de destination). Autrement dit, un même point de départ peut être relié à plusieurs points de destination, et cela s'écrit dans un seul enregistrement pour un point de départ donné.

 Dans les versions précédentes de *Philcarto* le format des matrices de liens était différent : chaque enregistrement décrivait un lien et un seul. Les fichiers utilisés

précédemment pour réaliser des cartes de liens ne sont donc plus compatibles avec la version 5 de *Philcarto* et doivent donc être reformatés selon le format décrit ci-après.

19.2.1. Le format de la matrice pour les cartes de liens non colorés

Dossier : \Philcarto exemples\Parana\
Fond de carte : Parana carte.ai
Statistiques : Parana data liens.xls, feuille : Réseau

Dans le cas des cartes de liens non colorés, le format d'enregistrement est le suivant (fig. 19.5) :

- colonne n°1 dénommée obligatoirement **Origine** : identifiant du point d'origine du lien ;
- colonne n°2, optionnelle et dénommée obligatoirement **Nom** si elle est présente : nom du point d'origine du lien
- colonne n°3 dénommée obligatoirement **Lien** : identifiant du point de destination du premier lien ;
- colonnes suivantes sans nom et en nombre variable, en fonction du nombre de destination : identifiant du point de destination des liens suivants.

Origine	Nom	Lien		
410010	Abatiá	412410		
410020	Adrianópolis	410690		
410040	Almirante Tamandaré	410690		
410045	Altamira do Paraná	412800		
410050	Altônia	412810		
410060	Alto Paraná	411840		
410070	Alto Piquiri	412810		
410080	Alvorada do Sul	411370		
410090	Amaporã	411840		
410100	Ampére	410480	410840	
410110	Andirá	411180	410240	
410120	Antonina	410690		
410130	Antônio Olinto	412560		
410140	Apucarana	411370	411520	
410150	Arapongas	410140		

Figure 19.5 : Un exemple de matrice de liens non colorés.

19.2.2. Le format de la matrice pour les cartes de liens colorés

Dossier : \Philcarto exemples\Parana\
Fond de carte : Parana carte.ai
Statistiques : Parana data liens.xls, feuille : Réseau + hiérarchie

Dans le cas des cartes de liens colorés, le format d'enregistrement est le suivant (fig. 19.6) :

- colonne n°1 dénommée obligatoirement **Origine** : identifiant du point d'origine du lien ;
- colonne n°2, optionnelle et dénommée obligatoirement **Nom** si elle est présente : nom du point d'origine du lien
- colonne n°3 dénommée obligatoirement **Lien modal** : identifiant du point de destination du premier lien ;
- colonne n°4 sans nom : modalité nominale caractérisant le premier lien (30 modalités différentes, chaîne de caractères ou nombre, au maximum sont permises) ;
- colonnes suivantes sans nom, en nombre variable en fonction du nombre de destination et par couples de deux colonnes pour chaque lien : identifiant du point de destination du lien suivi par la modalité qui le caractérise.

Origine	Nom	Lien modal				
410010	Abatiá	412410	1			
410020	Adrianópolis	410690	1			
410040	Almirante Tamandaré	410690	1			
410045	Altamira do Paraná	412800	1			
410050	Altônia	412810	1			
410060	Alto Paraná	411840	1			
410070	Alto Piquiri	412810	1			
410080	Alvorada do Sul	411370	1			
410090	Amaporã	411840	1			
410100	Ampére	410480	1	410840	1	
410110	Andirá	411180	1	410240	1	
410120	Antonina	410690	1			
410130	Antônio Olinto	412560	1			
410140	Apucarana	411370	5	411520	5	
410150	Arapongas	410140	4			

Figure 19.6 : Un exemple de matrice de liens non colorés.

19.3. L'enregistrement des variables créées par *Philcarto*

Plusieurs des traitements statistiques réalisés par *Philcarto* produisent de nouvelles variables : sélections de l'analyse exploratoire des données, classes de la classification ascendante hiérarchique, coordonnées sur les facteurs de l'analyse en composantes principales ou de l'analyse des correspondances, écarts de l'analyse territoriale multiscalaire...

En cours de session, *Philcarto* stocke ces nouvelles variables en mémoire à la suite des données lues dans le fichier des données statistiques ouvert en début de session, ou à l'aide de l'article **Autres statistiques** du menu **Fichier**. Ce stockage temporaire est perdu en fin de session. Pour conserver les variables ajoutées par *Philcarto*, il faut utiliser l'article **Enregistrer statistiques** du menu **Fichier**.

En sélectionnant **Oui**, *Philcarto* ouvre le dialogue **Enregistrer statistiques** ; le nom par défaut de ce fichier est **Statistiques** (fig. 19.7). Le type de fichier par défaut est **Texte (séparateur: tabulation) (*.txt)** ; mais *Philcarto* peut aussi enregistrer les données statistiques au format *Excel* en changeant le type de fichier pour **Classeur Microsoft Excel® (*.xls)** en utilisant la flèche de ce menu local.

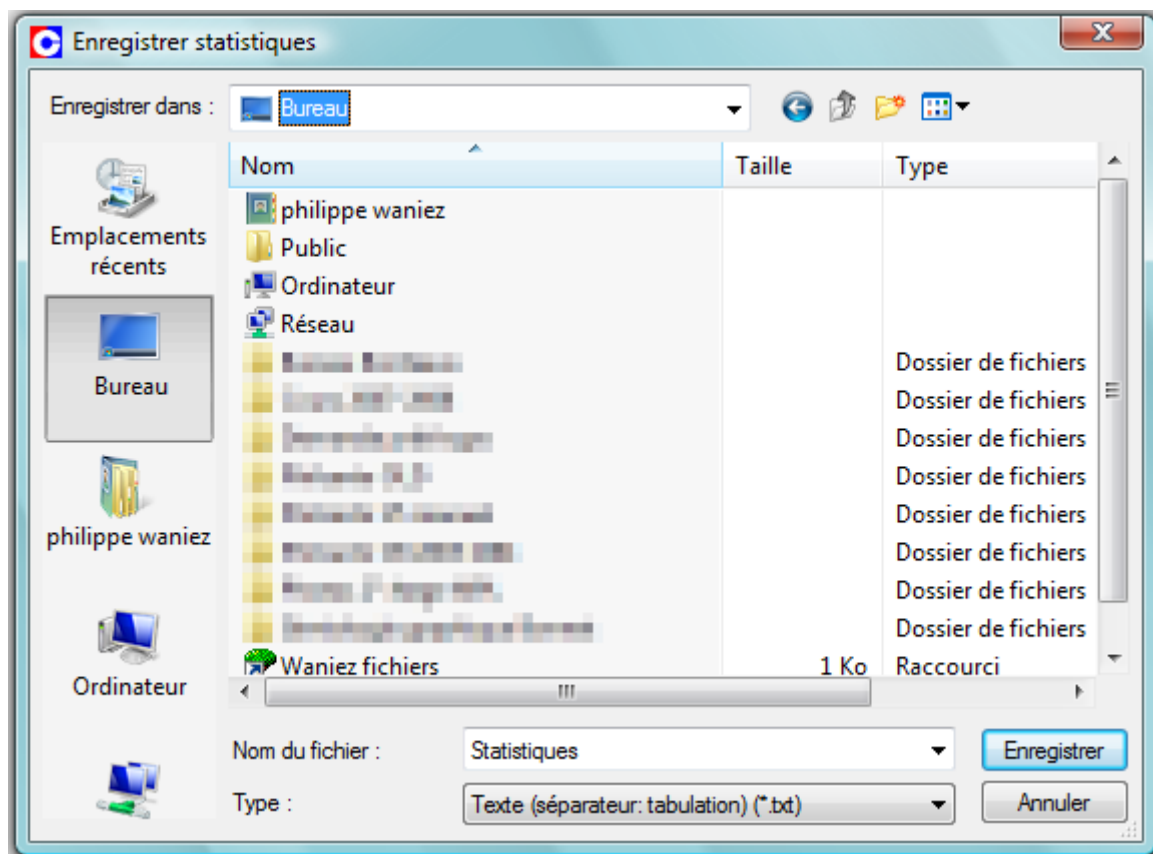


Figure 19.7 : Le dialogue **Enregistrer statistiques**.

Un clic sur **Enregistrer** ouvre le dialogue **Enregistrer les données statistiques** destiné à sélectionner, si une telle sélection est utile, les variables à conserver dans le nouveau fichier (fig. 19.8).

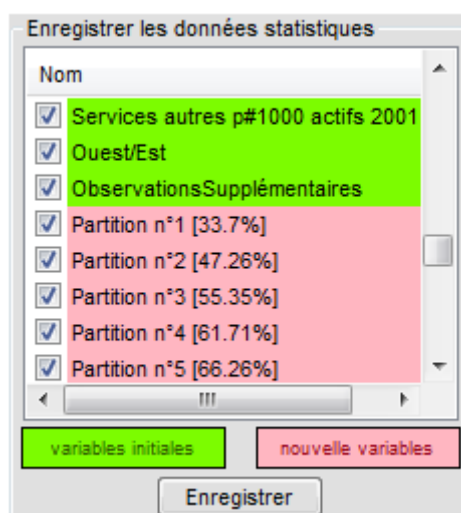


Figure 19.8 : Le dialogue **Enregistrer les données statistiques**.

Cette liste se compose de deux parties : en vert, les variables du fichier d'origine et en rose les variables créées en cours de session. Pour les premières comme pour les secondes, il est possible de ne pas retenir telle ou telle autre

variable en décochant la case qui figure en regard de son nom. Un clic sur **Enregistrer** enregistre effectivement les variables restées cochées dans le fichier nommé précédemment.

Si les nouvelles variables n'ont pas été enregistrées en cours de session, *Philcarto* affiche une alerte pour demander s'il faut conserver les nouvelles variables créées durant la session qui se ferme (fig. 19.9). En sélectionnant **Oui**, *Philcarto* ouvre le même dialogue que précédemment.

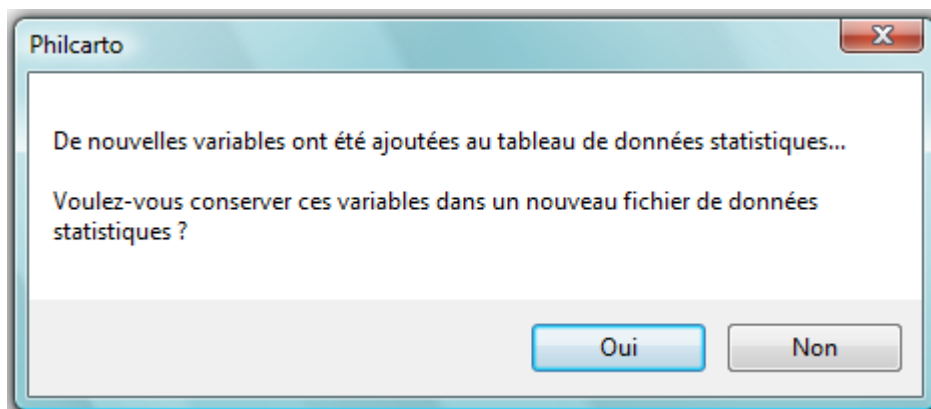


Figure 19.9 : Le message s'affichant en fin de session et permettant l'enregistrement des variables créées par *Philcarto*.

 Lorsque *Philcarto* enregistre un fichier avec des valeurs décimales, la marque de la décimale est le point «.». Ouvert dans *Excel* avec un système français, ces valeurs décimales sont considérées par *Excel* comme du texte. Pour réutiliser dans *Excel* des fichiers enregistrés par *Philcarto*, il faut donc préalablement remplacer les points (.) par des virgules (,) à l'aide de *UltraEdit*, par exemple, ou bien faire de même directement dans *Excel*.

19.4. Le calcul d'indices

Il est parfois nécessaire de calculer un nouvel indicateur à partir des variables du tableau de données. Dans ce but, on peut naturellement utiliser *Excel* ou *OpenOffice* ; le calculateur de *Philcarto* permet de réaliser des calculs impliquant simultanément plusieurs variables dans le but de leur appliquer des opérations courantes et répétitives. On accède au calculateur en activant l'article **Calculateur** du menu **Fenêtres** ; le dialogue **Calculateur** affiche, dans sa partie gauche, la liste des variables pouvant faire l'objet d'un calcul avec le calculateur : il s'agit de variable de quantités ou d'effectifs (fig. 19.10). La partie droite du calculateur dialogue **Calculer...** liste les différents calculs qu'il est possible de réaliser ; cette liste se compose de trois parties : des calculs courants du genre pourcentage ; le calcul de l'indice de concentration de Gini ; un ensemble d'indices de spécialisation ou de diversification.

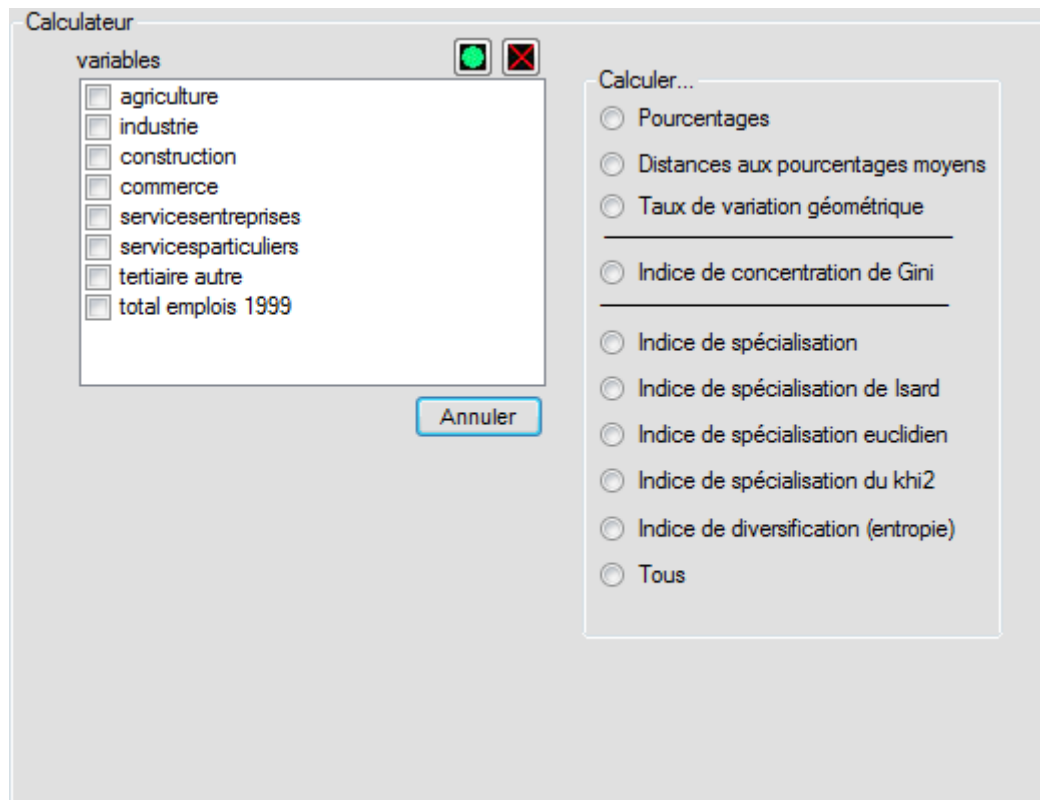


Figure 19.10 : Le dialogue du calculateur.

19.4.1. Calculer des pourcentages

Dossier : \Philcarto exemples\France départements\
Fond de carte : France carte départements.ai
Statistiques : France emplois départements.xls

Il est facile, mais fastidieux, de réaliser ce genre de calcul avec un tableur. Le calculateur réalise ce calcul sur une série de variables de façon simple : dans la liste de variables on sélectionne celle pour lesquelles on veut obtenir un pourcentage ; si elle est présente, ne pas sélectionner la variable contenant le total des autres variables. Le bouton  permet de sélectionner toutes les variables de la liste `variables` ; le bouton  désélectionne toutes les variables. Dans le dialogue `Calculer...`, sélectionner le bouton radio `Pourcentages`. Une seconde liste de variable apparaît alors dans la partie gauche du dialogue `Calculateur` ; dénommée `variable de base du pourcentage (optionnelle)`, elle permet de choisir la variable de base du pourcentage, celle qui contient l'effectif total de chaque unité spatiale. Si la variable contenant le total est présente dans le tableau de données, on peut la sélectionner dans cette liste. On peut aussi laisser le calculateur faire la somme des variables sélectionnées dans la liste `variables`, en ne sélectionnant aucune variable de base. Un clic sur `Calculer` lance le calcul (fig. 19.11). Le dialogue `Calculateur` disparaît, et laisse place à la liste des variables de `couleurs` qui, maintenant, contient les pourcentages (fig. 19.12) qui, de ce fait, peuvent être cartographiés (fig. 19.13).

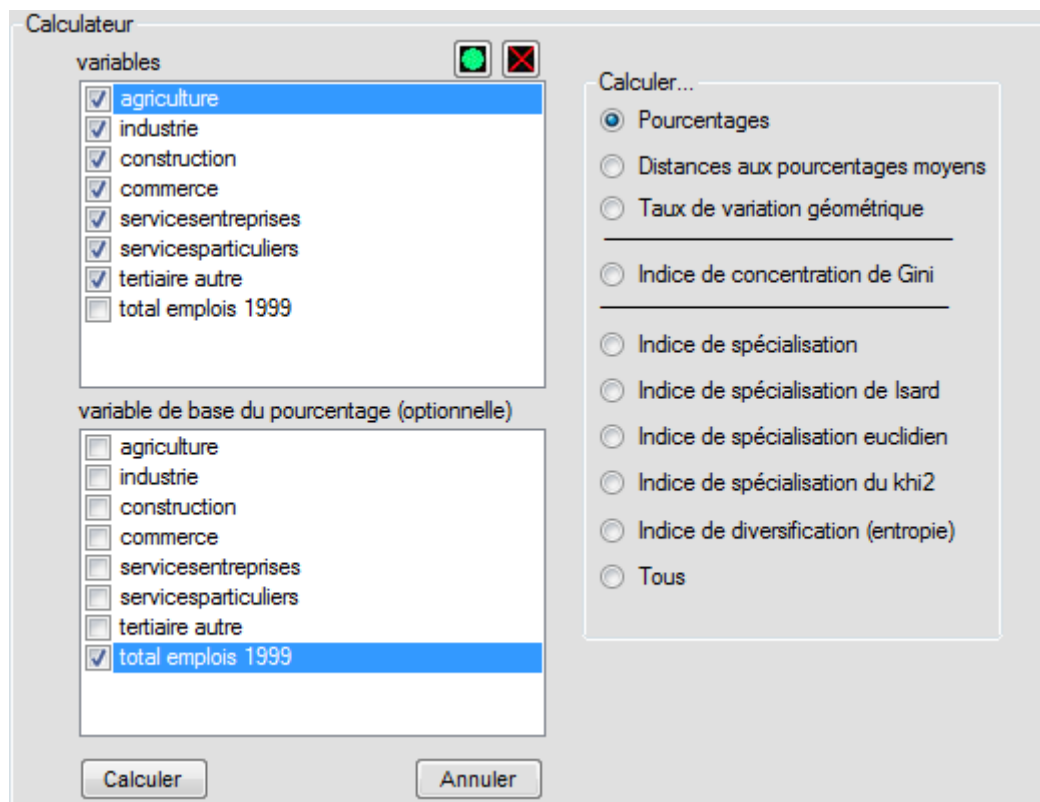


Figure 19.11 : Le **Calculateur** du pourcentage de chaque secteur d'activité dans l'ensemble des emplois.

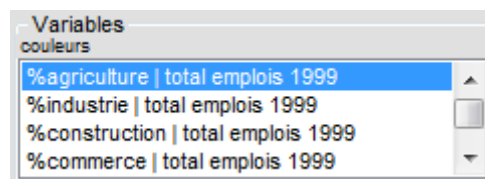


Figure 19.12 : La liste des variables de pourcentages.

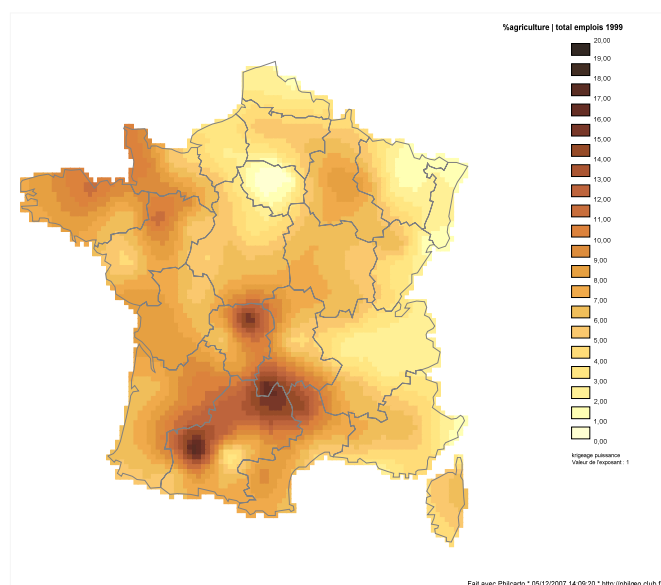


Figure 19.13 : Une carte lissée du pourcentage de l'emploi agricole.

19.4.2. Calculer des distances au pourcentage moyen

Dossier : \Philcarto exemples\France départements\
Fond de carte : France carte départements.ai
Statistiques : France emplois départements.xls

Si les effectifs qui ont permis de calculer des pourcentages sont différents d'une unité spatiale à l'autre, la moyenne arithmétique des pourcentages n'est pas égale au pourcentage moyen. Dans ce cas, et si l'on dispose des effectifs (et non pas seulement des pourcentages), il est préférable de calculer les **Distances aux pourcentages moyens** ; il s'agit du même calcul que celui qui est fait pour l'Analyse Territoriale Multiscalaire pour le niveau macro ; la différence avec cette méthode réside dans le fait qu'on peut calculer ces distances pour un ensemble de variables (au lieu d'une seule variable pour l'ATM).

Le choix de variables se fait de la même façon que pour le calcul des pourcentages (fig. 19.14). Un clic sur **Calculer** lance le calcul.

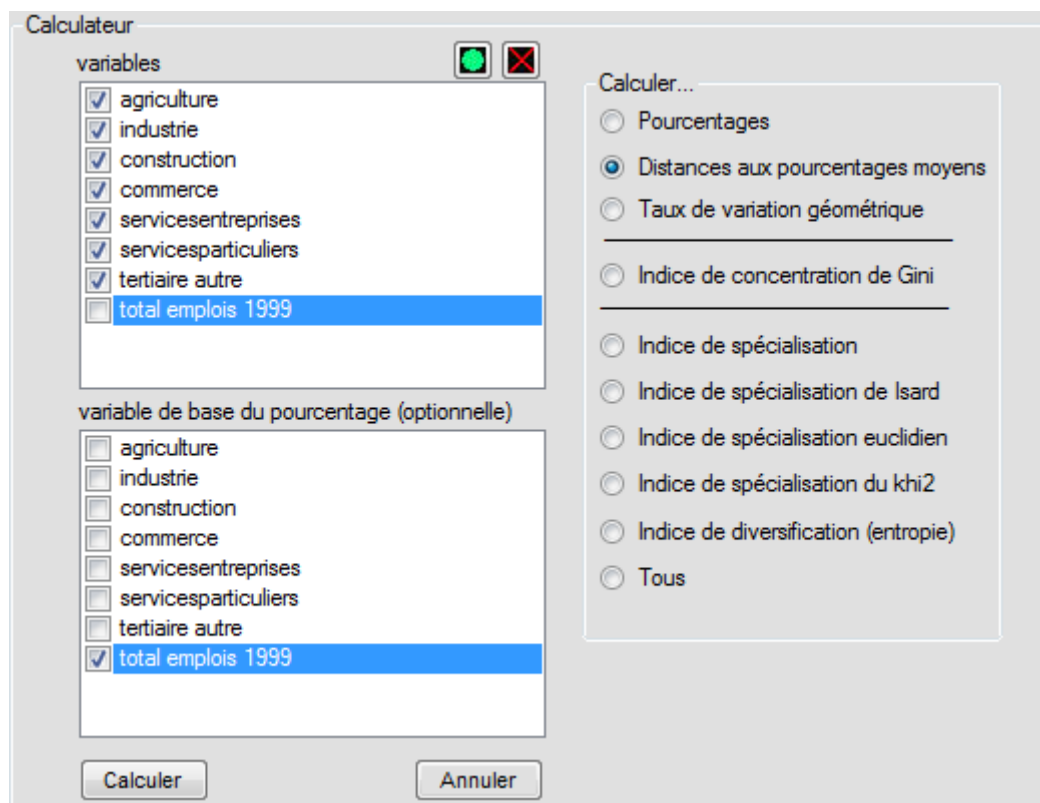


Figure 19.14 : Le **Calculateur** des distances aux pourcentages moyens de chaque secteur d'activité dans l'ensemble des emplois.

Le calcul est réalisé en quatre étapes :

- calcul du total des effectifs de chaque variable ;
- calcul du pourcentage de chaque variable dans ce total ;
- calcul du pourcentage de chaque variable dans chaque unité spatiale ;
- calcul de la différence entre ces deux pourcentages.

La cartographie des distance fait avantageusement appel à une gamme de couleurs en opposition (fig. 19.15)

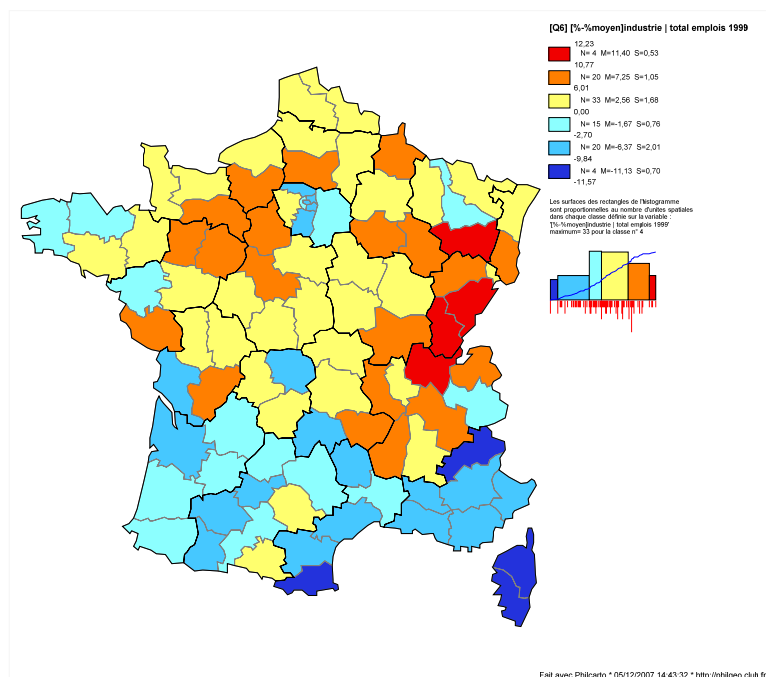


Figure 19.15 : La carte choroplèthe des distances pourcentage moyen de l'emploi industriel.

19.4.3. Calculer un taux de variation géométrique

Dossier : \Philcarto exemples\France départements\
Fond de carte : France carte départements.ai
Statistiques : France data départements.xls

Pour calculer la variation d'une population sur une période donnée, un simple calcul de variation, basé sur la différence entre les deux populations divisée par la population en début de période, ne suffit pas, surtout si la variation est forte. Il est préférable de calculer un taux de variation géométrique basé sur le principe des intérêts composés. On obtient ainsi un taux de variation moyen pour chaque unité de temps (le plus souvent l'année). En démographie, c'est ainsi qu'on calcule le taux de variation moyen annuel d'une population entre deux recensements.

L'obtention des taux de variation géométriques exige deux variables : la population en début de période et la population en fin de période ; de plus il faut connaître la durée de la période. Dans le calculateur, on sélectionne le bouton radio ☒ **Taux de variation géométrique**, puis la **variable de début** de période et la **variable de fin** de période ; enfin, on indique la **longueur de la période** (fig. 19.16). Un clic sur **Calculer** lance le calcul. Lorsque les taux calculés, les nouvelles variables sont affichées dans la liste **couleurs** du dialogue **Variables** du dialogue **Outils**. Si les taux présentent des valeurs négatives et positives, leur cartographie fait avantageusement appel à une gamme de couleurs en opposition (fig. 19.17)

Calculateur

variable de début

- ☐ Propriétaires 1990
- ☐ WC extérieurs résidences principales
- ☐ Sans auto 1990 résidences principales
- ☐ Construction avant 1949 résidences principales
- ☐ Construction 1949_1981 résidences principales
- ☐ Construction 1982 et après résidences principales
- ☐ Construction après 1949 résidences principales
- ☒ Population 1990
- ☐ Population 1999
- ☐ Solde migratoire 1990-1999

variable de fin

- ☐ Propriétaires 1990
- ☐ WC extérieurs résidences principales
- ☐ Sans auto 1990 résidences principales
- ☐ Construction avant 1949 résidences principales
- ☐ Construction 1949_1981 résidences principales
- ☐ Construction 1982 et après résidences principales
- ☐ Construction après 1949 résidences principales
- ☐ Population 1990
- ☒ Population 1999
- ☐ Solde migratoire 1990-1999

Calculer...

- ☐ Pourcentages
- ☐ Distances aux pourcentages moyens
- ☒ Taux de variation géométrique
- ☐ Indice de concentration de Gini
- ☐ Indice de spécialisation
- ☐ Indice de spécialisation de Isard
- ☐ Indice de spécialisation euclidien
- ☐ Indice de spécialisation du khi2
- ☐ Indice de diversification (entropie)
- ☐ Tous

longueur de la période

Calculer Annuler

Figure 19.16 : Le Calculateur d'un taux de variation géométrique.

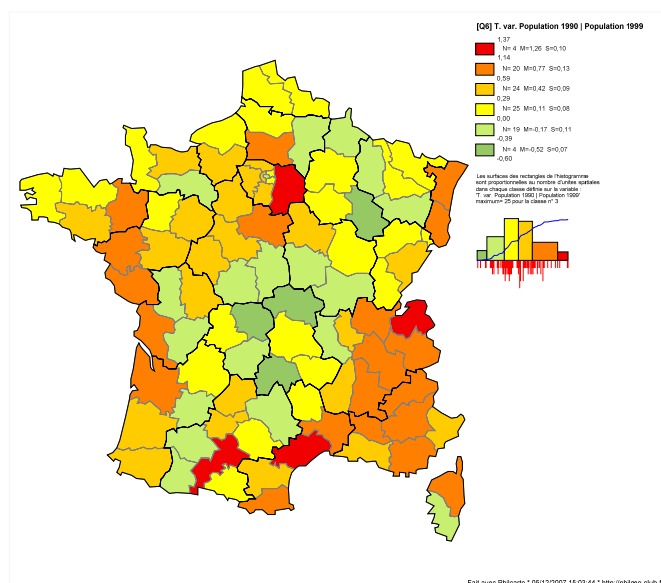


Figure 19.17 : La carte choroplèthe du taux de variation moyen annuel de la population 1990-1999.

19.4.4. Calculer un indice de concentration de Gini

Dossier : \Philcarto exemples\France départements\
Fond de carte : France carte départements.ai
Statistiques : France IRPP départements.xls

Le calcul de l'indice de concentration de **Gini** exige que chaque variable corresponde à une catégorie pour laquelle on doit indiquer la valeur du centre de catégorie. A l'intersection d'une ligne et d'une colonne, on trouve un effectif qui correspond au nombre d'individus caractérisés par la catégorie de valeurs représentée par cette colonne dans l'unité spatiale correspondant à la ligne.

Par exemple, pour chaque tranche de Revenu Fiscal de Référence (RFR) de l'Impôt sur le Revenu des Personnes Physiques (IRPP), les noms des variables doivent contenir la valeur du centre de catégorie. Ainsi, au lieu de donner le nom : « RFR 0 à 7500€ », on l'écrit de la manière suivante « RFR 0 à 7500€ C=3750 », « C= » indiquant que la valeur qui suit est bien un centre de catégorie ; aucun caractère ne doit figurer après la valeur du centre. Si la dernière catégorie est ouverte, on choisit un centre supérieur ou égal à la borne inférieure de la catégorie.

Pour calculer l'indice de Gini, on sélectionne le bouton radio ☒ Indice de concentration de Gini, puis les variables devant entrer dans le calcul (fig. 19.18). Un clic sur affiche les centres des catégories (fig. 19.19) ; on peut alors les modifier si nécessaire ; un nouveau clic sur lance le calcul proprement dit.

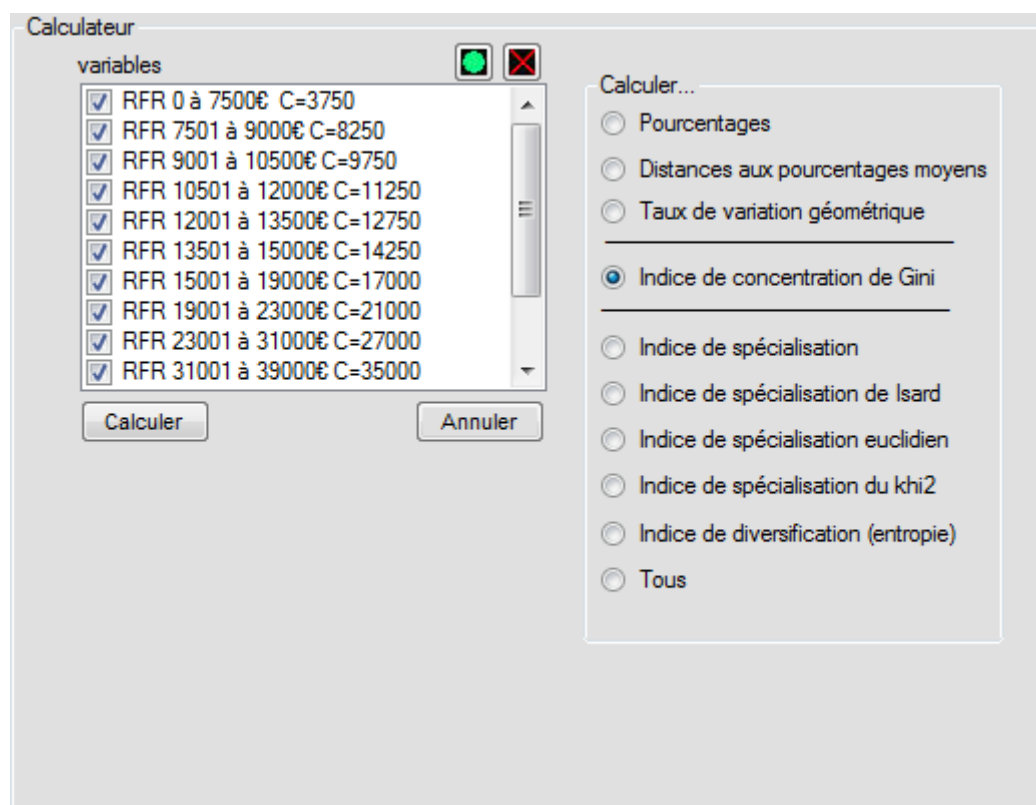


Figure 19.18 : Le **Calculateur** d'un indice de concentration de Gini.

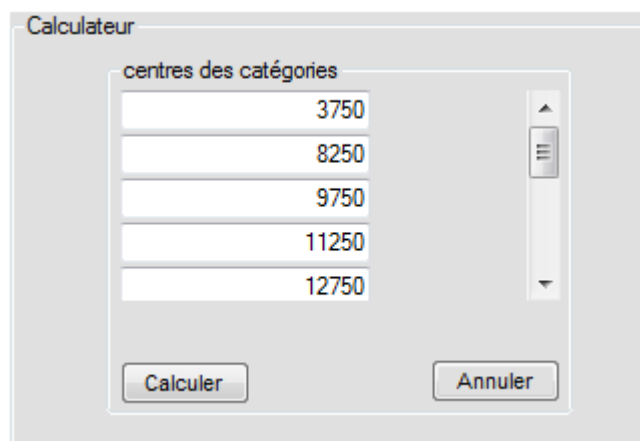


Figure 19.19 : L'affichage des centres de catégories pour le calcul d'un indice de concentration de Gini.

Lorsque l'indice de Gini est calculé, la nouvelle variable est affichée dans la liste **couleurs** du dialogue **Variables** du dialogue **Outils**. Notons que l'indice de Gini mesure la concentration, mais ne donne aucune information sur les variables qui sont à l'origine de la concentration. Pour interpréter les valeurs de cet indice, il faut donc le mettre en relation avec les variables d'origine (fig. 19.20).

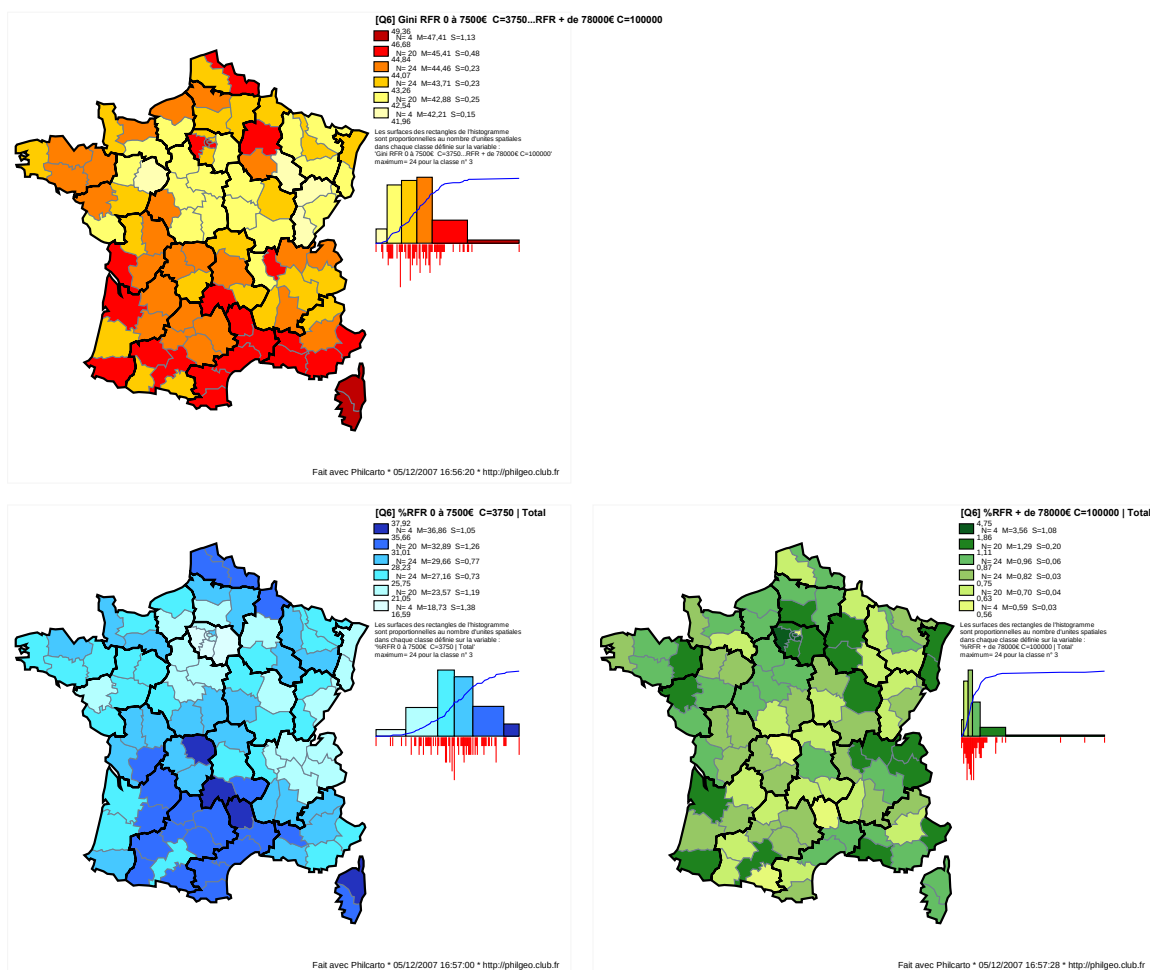


Figure 19.20 : L'interprétation de l'indice de Gini (carte rouge) ; % de foyers fiscaux de la catégorie de revenus la plus basse (carte bleu) et de la catégorie la plus élevée (carte verte).

Par exemple, la carte de l'indice de concentration de Gini du revenu fiscal de référence des contribuables de l'IRPP présente de fortes valeurs à la fois dans la moitié Sud de la France, et dans certains départements de la Région Île-de-France. Compte tenu de ce qu'on sait de la situation économique des régions concernées, la concentration est sans doute d'origine différente. Dans le Sud-Ouest, départements littoraux exclus, la concentration correspond à celle des contribuables à faible revenu fiscal de référence. A l'inverse, à Paris, dans les Hauts-de-Seine, les Yvelines, dans la Marne (champagne...) et sur la Côte d'Azur, ce sont les plus hauts revenus qui sont concentrés.

L'indice de Gini n'est donc pas aisé à interpréter dans la mesure où les indications qu'il donne peuvent correspondre à des situations assez différentes. Pour un exposé facile à comprendre de l'indice de Gini, se reporter à : Hubler J., 1996. *Statistiques pour l'économie*. Paris, Bréal, Col. Lexifac, 219 p.

19.4.5. Calculer des indices de spécialisation ou de diversification

Dossier : \Philcarto exemples\France départements\
Fond de carte : France carte départements.ai
Statistiques : France emplois départements.xls

La spécialisation d'un sous-ensemble géographique résulte soit d'une fonction spécifique assurée par ce sous-ensemble, soit de son appartenance à un domaine particulier. C'est cette spécialisation qui fait qu'un sous-ensemble apparaît original par rapport à l'ensemble de référence ; c'est aussi elle qui traduit la complémentarité fonctionnelle assurant la cohésion de l'ensemble.

Pour mesurer la nature et l'intensité d'une spécialisation dans un domaine particulier tel que l'activité, l'occupation du sol ou bien encore la stratification sociale, on fait appel à des indices synthétiques. Le calcul de ces indices fait appel à un tableau de correspondances (cf. §20.1.1). Il existe un grand nombre d'indices ; *Philcarto* en propose quatre (fig. 19.21). Les valeurs de ces indices sont comprises entre 0 et 1 ; pour faciliter la lecture des cartes, *Philcarto* les multiplie par 100.

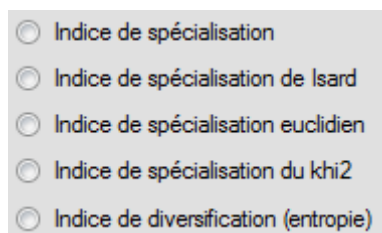


Figure 19.21 : Les indices de diversification et de spécialisation calculés par *Philcarto*.

La signification des indices de spécialisation est expliquée dans l'excellent ouvrage : Pumain D., Saint-Julien T., 2001. *Les interactions spatiales*. Paris, Armand Colin, Col. Cursus, 192 p.

Philcarto propose aussi un indice de diversification basé sur la mesure de l'entropie. La méthode de calcul de cet indice là est exposée dans : Novi M., 1998. *Pourcentages et tableaux statistiques*. Paris, PUF, Col. Que sais-je, 125 p.

Le calculateur peut produire soit l'un des cinq indices proposés, soit tous ces indices en une seule fois. Il suffit pour cela de choisir le bouton radio **Tous** et de sélectionner les variables devant entrer dans le calcul (fig. 19.22). Un clic sur **Calculer** lance le calcul proprement dit.

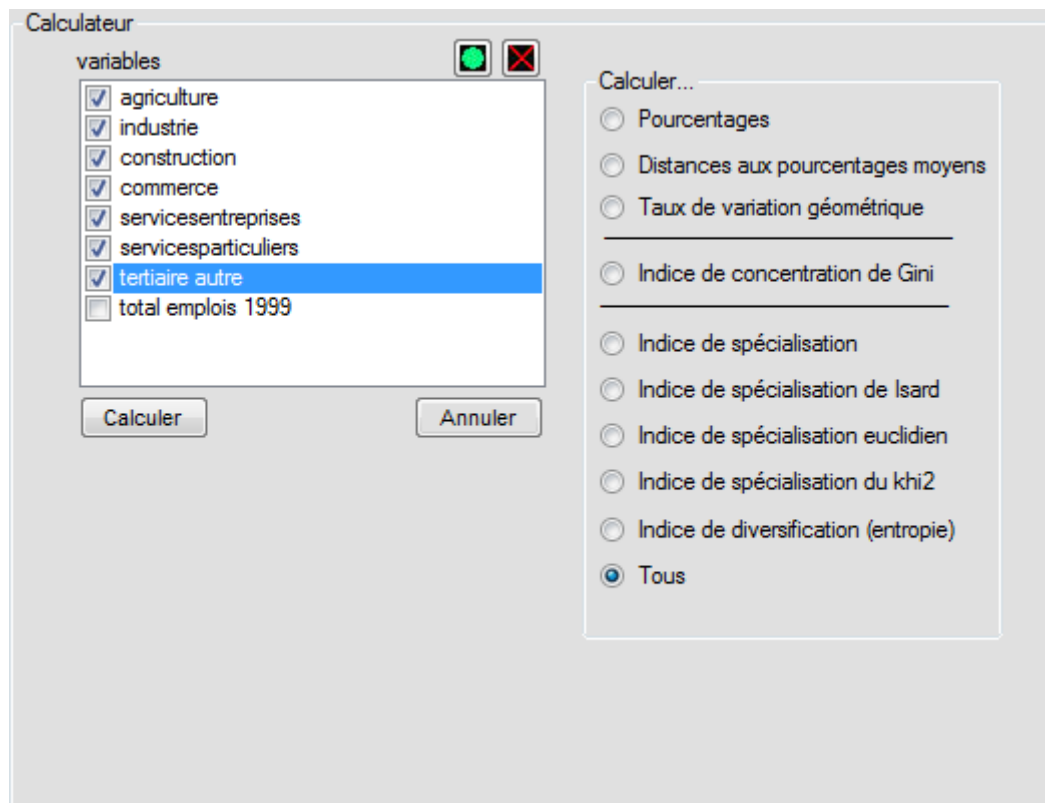


Figure 19.22 : Le **Calculateur** de tous les indices de spécialisation et de diversification.

Lorsque les indices sont calculés, les nouvelles variables sont affichées dans la liste **couleurs** du dialogue **Variables** du dialogue **Outils** (fig. 19.23). Les cartes des spécialisation de Isard, d'Euclide et du Khi2 n'apparaissent pas très différentes (coadjacence de 50% entre Isard et Euclide). En revanche, celle de la diversification est assez différente (fig. 19.24).

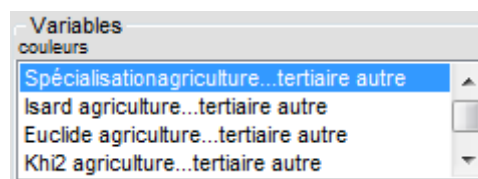


Figure 19.23 : La liste des indices de spécialisation calculés.

Si ces indices mesurent la spécialisation ou la diversification, ils ne donnent cependant aucune information sur les modalités ou groupes de modalités qui engendrent ces différences. Comme avec l'indice de Gini, l'analyse doit se poursuivre par l'examen de cartes de pourcentages de chaque modalité afin d'identifier celles d'entre-elles qui sont particulièrement discriminantes dans tel ou tel autre regroupement visible sur la carte. Dans ce but, on peut aussi recourir aux techniques de l'Analyse des Données.

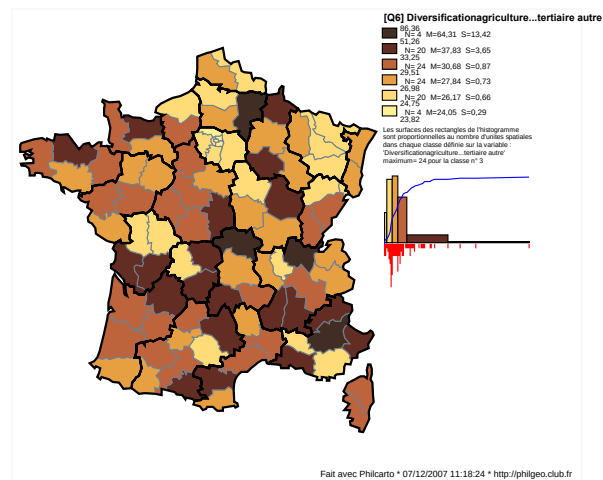
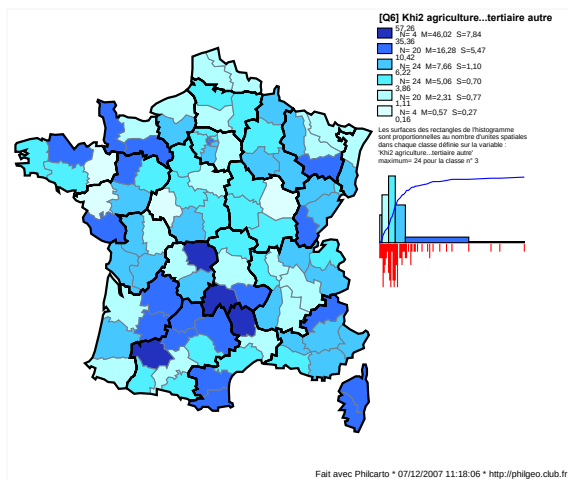
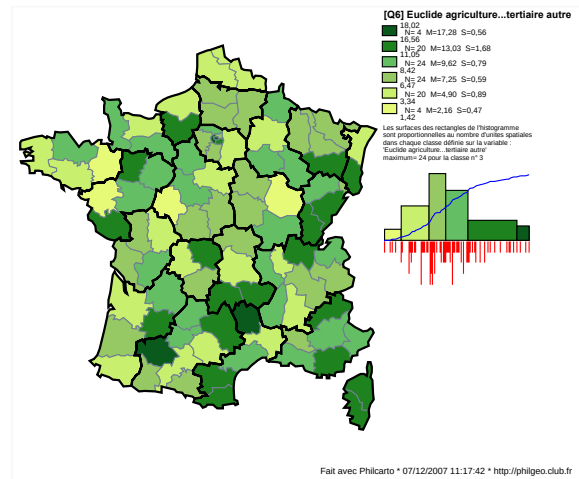
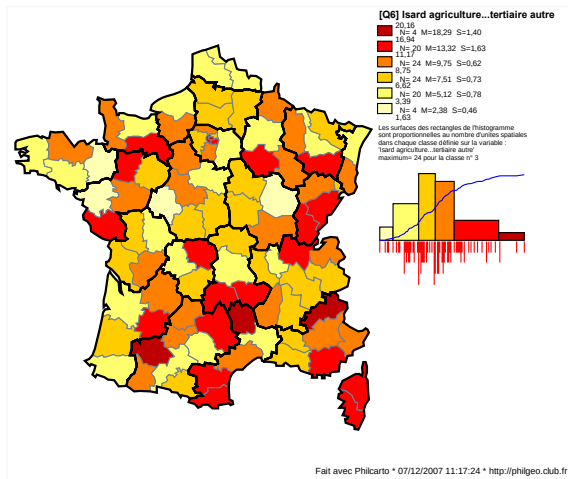


Figure 19.24 : Les cartes de indices de spécialisation (Isard en rouge, Euclide en vert, Khi2 en bleu) et de diversification (en marron).

20. Les fonds de cartes

Les unités spatiales constituant le fond de carte sont figurées soit par des surfaces, soit par des points, soit par des lignes. Selon l'échelle de cartographie, le figuré peut changer ; par exemple, une ville peut être représentée par une surface à l'échelle 1:50 000 alors que ses limites ne pouvant être visibles à l'échelle 1:10 000 000, on la localise sur le fond de carte par un simple point. Comme logiciel mono-échelle, *Philcarto* ne prend pas en compte de telles subtilités : le fond de carte doit avoir été préalablement dessiné à une échelle donnée et enregistré en format *Adobe Illustrator (.ai)*.

Si ce format graphique, c'est-à-dire la manière dont sont codés sur le plan informatique les éléments composant le dessin, a été choisi parmi les centaines de formats connus, c'est parce que de nombreux graphistes savent utiliser *Adobe Illustrator*, l'un des leaders du marché, et sont capables de réaliser des fonds de cartes de bonne qualité à partir de documents scannés. Un autre avantage du format *.ai* est de permettre l'enregistrement des cartes réalisées par *Philcarto* en conservant les caractéristiques géométriques du fond de carte : si le fond de carte fourni en entrée est déjà au format d'édition, les cartes en sortie de *Philcarto* le seront aussi ; elles pourront également être habillées par des éléments cartographiques complémentaires préparés à l'avance (fleuves, routes, échelle, etc.) par de simples « copier/coller » réalisés avec *Adobe Illustrator*.

Malheureusement, *Adobe Illustrator* ne fait pas partie des logiciels gratuits, beaucoup s'en faut : il faut compter environ 850€ TTC pour l'acheter sur le site Internet de la société Adobe. Un tel prix le rend peu accessible aux particuliers. C'est la raison pour laquelle j'ai conçu et réalisé *Phildigit*, téléchargeable à la même adresse que *Philcarto*. *Phildigit* se situe au niveau élémentaire de la digitalisation : il sera utilisé avec profit par des utilisateurs occasionnels, ne disposant pas d'un logiciel de dessin coûteux, ne souhaitant pas investir trop de temps ou d'argent dans une activité dont ils ne deviendront jamais spécialistes ; quelques heures et un peu d'attention leur suffiront pour réaliser un fond de carte utilisable. Ceci ne signifie pas qu'on ne puisse réaliser un travail de bonne qualité avec *Phildigit*. Autrement dit, *Phildigit* permet à un utilisateur débutant ou pressé, de réaliser simplement un fond de carte vectoriel, sur la base d'un document scanné à l'aide d'un simple scanner de bureau. Le document d'origine peut être un croquis réalisé à partir d'un relevé de terrain, mais plus généralement provenant d'une publication comme en produisent les organismes de collecte et de diffusion de données concernant la population, la santé, la société, l'environnement, etc. dans le cadre d'un maillage de l'espace usuel (communes, départements...) ou spécifique (stations météorologiques, parcellaire...). La version 2 de *Phildigit* permet d'importer des fonds de carte enregistrés dans des formats spécifiques aux principaux SIG du marché. Ceci permet d'exploiter les très nombreux fichiers disponibles sur Internet notamment. Une recherche sur la toile s'impose donc avant d'entreprendre tout travail de dessin.

Pour apprendre à digitaliser un fond de carte avec *Phildigit*, ou bien pour rendre utilisables avec *Philcarto* des fonds de cartes enregistrés aux formats des SIG (*.shp/.shx/.dbf* pour *ArcView*, *.mid/.mif* pour *Mapinfo*), il faut télécharger le manuel de *Phildigit* sur le site internet : <http://philgeo.club.fr> .

Bibliographie

Bertin J., 1977, *La graphique et le traitement graphique de l'information*. Paris, Flammarion, 273 p.

Brunet R., 1987. *La carte Mode d'Emploi*. Paris, Fayard/Reclus, 270 p.

Brunet R., 1997. *Champs et contrechamps, Raisons de géographe*. Paris, Belin, Col. Mappemonde, 317 p.

Cauvin C., Reymond H., Serradj A., 1987, *Discrétisation et représentation cartographique*. Montpellier, GIP RECLUS, Col. Reclus Modes d'Emploi, 116 p.

Claval P., 1977, *La nouvelle géographie*. Paris, PUF, Col. Que sais-je, n°1693, 126 p.

Cleveland W.S., 1993. *Visualizing data*. Summit, NJ, Hobart Press, 360 p.

Cleveland W.S., 1994. *The elements of graphing data*. Summit, NJ, Hobart Press, 297 p.

Dangermond J., 1990, A classification of software components commonly used in geographic information systems. *Introductory readings in Geographic Information Systems*. Donna J. Peuquet and Duane F. Marble editors. London, Taylor & Francis, pp. 30-51.

Dauphiné A., 1979. *Espace, région et système*. Paris, Economica, Col. Geographia, n°1, 167 p.

Dumolard P., 1981. *L'espace différencié*. Paris, Economica, Col. Geographia, n°2, 208 p.

Fénelon J.-P., 1981. *Qu'est-ce que l'Analyse des Données*. Paris, Editions Lefonen, 311 p.

Gambini R., 1966. *A Computer Program for Calculating Lines of Equilibrium Between Multiple Centers of Attraction*. Center for Regional Studies, University of Kansas, Lawrence, Kansas.

Gratton Y. (2002). Le krigeage: la méthode optimale d'interpolation spatiale. *Les articles de l'Institut d'analyse géographique*, juillet 2002 (www.iag.asso.fr).

Griffith D.A., 1987. *Spatial Autocorrelation. A primer*. Association of American Geographers, Col. Ressource Publications in Geography, 86 p.

Hartwig F., Dearing B., 1982. *Exploratory Data Analysis*. Sage University Paper, Series : Quantitative Applications in the Social Sciences, n°16. Beverly Hills, CA, Sage Publications, 83p.

Huff D. L., 1973. The delineation of a National System of Planning Regions on the Basis of Urban Spheres of Influence. *Regional Studies*, Vol. 7, pp. 323-329, Pergamon Press.

Huff D.L., 1979. *Ireland's Urban System*. Economic Geography, Vol. 55, n°3, pp. 196-212, 1979.

Haslett J., 1992. Spatial data analysis – challenges. *The statistician* 41, pp. 271-284.

Hubler J., 1996. *Statistiques pour l'économie*. Paris, Bréal, Col. Lexifac, 219 p.

Jambu M., Lebeaux M.-O., 1978. *Classification automatique pour l'analyse des données*. T. 2 : logiciels. Paris, Dunod, 399 p.

King L., 1984. *Central Place Theory*. Beverly Hills, Sage Publications, Col. Scientific Geography series, Vol. 1, 96 p.

Ruppert J., 1995. A Delaunay Refinement Algorithm for Quality 2-Dimensional Mesh Generation. *Journal of Algorithms* 18(3):548-585, May 1995

Le Gauffey Y., Waniez P., 1987. *Initiation à l'analyse en surfaces de tendances*. Paris, Ed. de l'ORSTOM, 35 p.

Mac Dougall E.B., 1992. *Exploratory analysis, dynamic statistical visualization and geographic information systems*. Cartography and geographic information systems, Vol. 19, n°4, pp. 237-246.

Minvielle E., Souiah S.-A., 2003, *L'analyse statistique et spatiale*. Nantes, Editions du Temps, 284 p.

Novi M., 1998. *Pourcentages et tableaux statistiques*. Paris, PUF, Col. Que sais-je, 125 p.

Pumain D., Saint-Julien T., 2001. *Les interactions spatiales*. Paris, Armand Colin, Col. Cursus, 192 p.

Riddel J.B., 1970. *The spatial dynamics of modernization in Sierra-Leone*. Northwestern University Press, 142 p.

Sanders L., 1990. *L'analyse statistique des données en géographie*. Montpellier, GIP Reclus, Col. Alidade, 267 p.

Tukey J.W., 1977. *Exploratory Data Analysis*. Reading, MA., Addison-Wesley Publishing Company, 688 p.

Unwin A., Sloan B, Wills G., 1992. Interactive graphical methods for trade flows. *Proceedings of the conference on new techniques and technologies*. Bonn, RFA.

Velleman P.F., 1981. *Applications, Basics, and Computing of Exploratory Data Analysis*. Boston, MA, Duxbury Press.

Waniez P., 1991. *Analyse exploratoire des données*. Montpellier, GIP RECLUS, Col. Reclus Modes d'Emploi, n°17, 159 p.

Waniez P., 1990. Cartographie des aires d'influences urbaines: méthode ancienne, techniques nouvelles. *Mappemonde* 1, pp. 27-32.

Waniez P., 2000. Fonction hospitalière et système urbain au Brésil. *Cahiers Santé*, Vol. 10, n°2, pp. 145-153.

Zanin C., Trémélo M.-L., 2003, *Savoir faire une carte*. Paris, Belin, Col. Sup., 200 p.