

POUR L'EXPÉRIMENTATION DES SIG EN GÉOGRAPHIE, LE SIG PARANÁ

Philippe WANIEZ*

Mirian VIZINTIM**

Violette BRUSTLEIN***

RÉSUMÉ Avec la présentation du SIG Paraná, sont exposées les étapes d'une expérience, conduite par deux géographes et un cartographe, destinée à évaluer le degré d'innovation du logiciel GéoConcept par rapport à d'autres instruments plus courants d'analyse géographique.

ABSTRACT The presentation of SIG Paraná provides the opportunity to detail the evaluation by two geographers and a cartographer of the GéoConcept software compared to other, more commonly used, tools for geographical analysis.

RESUMEN Con la presentación del SIG Paraná se exponen las etapas de una experiencia llevada por dos geógrafos y un cartógrafo y destinada a evaluar el grado de innovación del programa GéoConcept en comparación con otros instrumentos más corrientes de análisis geográfico.

• AGRICULTURE • CARTOMATIQUE • LOGICIEL • PARANÁ • SIG

• AGRICULTURE • AUTOMATIC MAPPING • GIS • PARANÁ • SOFTWARE

• AGRICULTURA • CARTOMÁTICA • PARANÁ • PROGRAMA • SIG

Nombreux sont aujourd'hui les géographes qui, peu ou prou, pratiquent la cartomatique ou cartographie des données statistiques assistée par ordinateur. Au minimum, il s'agit de la «mise à plat» d'indicateurs considérés comme pertinents pour exprimer la répartition spatiale de divers phénomènes. En permettant de faire varier à l'infini, ou presque, le nombre de classes, les seuils statistiques et les gammes de couleurs, tout micro-ordinateur de bas de gamme peut tracer ces cartes d'analyse. Mais l'utilisateur court néanmoins le risque bien connu de se perdre face à la multiplication inévitable des documents produits par la machine. Au cours des années 1980, les techniques désormais «classiques» de l'analyse des données (analyses factorielles, classifications automatiques, modèles de régression, etc.) se sont largement développées, afin d'élaborer les synthèses exigées par le volume croissant des corpus statistiques; avec l'avènement des micro-ordinateurs et des logiciels presse-boutons et bon marché, l'us s'est souvent transformé en abus, faute d'une bonne connaissance des conditions d'application de ces méthodes (1).

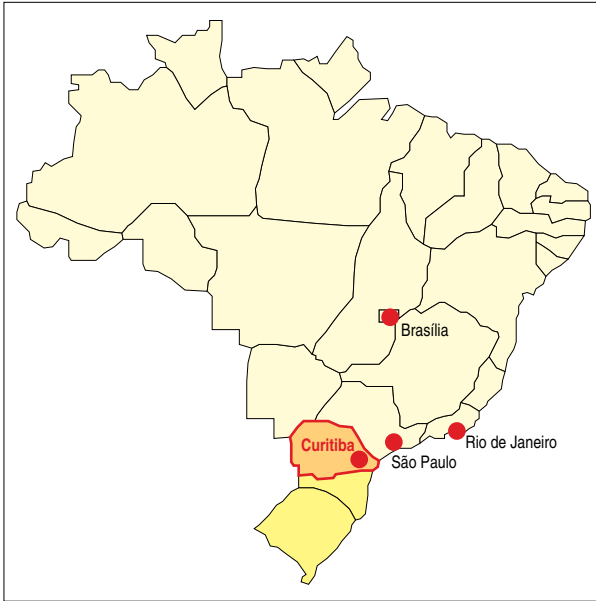
À l'heure où l'on parle beaucoup des systèmes d'information géographique — plus que l'on ne les pratique, il est vrai —,

l'arrivée sur le marché du logiciel de SIG «performants et conviviaux» a de quoi inquiéter les spécialistes de l'analyse spatiale: plus encore qu'avec la cartomatique ou l'analyse des données, ne risque-t-on pas de voir l'émergence d'un nouvel engouement où le «tout SIG» serait la panacée? Pour tenter d'éviter un tel écueil, et afin que l'outil de recherche demeure à la place qui lui revient, celle d'instrument au service d'une problématique scientifique, il apparaît indispensable qu'à côté des opérations de formation des géographes — telle l'École Nationale d'été sur le SIG qui s'est déroulée à Montpellier au mois de septembre 1992 —, se multiplient les tests de ces nouvelles techniques de stockage, d'analyse et de représentation de l'information géographique, et cela en relation avec les programmes de recherche en cours. C'est dans cette perspective qu'est présenté ici le SIG Paraná réalisé dans le cadre du programme *Fronteiras, des frontières agricoles brésiliennes au marchés mondiaux* réalisé conjointement par le GIP Reclus et l'ORSTOM. L'objectif de cet article n'est pas de présenter les résultats de fond sur les transformations de l'agriculture du Paraná, mais d'exprimer un point de vue illustré par des exemples extraits de cette recherche.

* GIP Reclus, Maison de la Géographie, Montpellier; ORSTOM, UR 5C: Territoires.

** Université de Londrina, Paraná, Brésil.

*** CNRS, GIP Reclus, Maison de la Géographie, Montpellier.

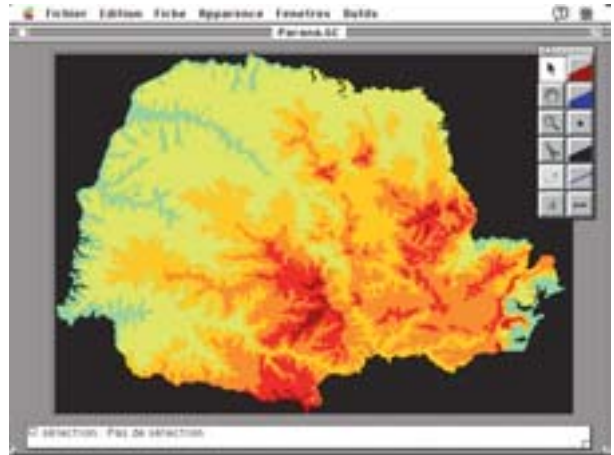


1. L'État du Paraná dans la région Sud du Brésil

Les sources d'information

L'État du Paraná est le plus septentrional de la région Sud du Brésil (fig. 1). Son secrétariat d'État à l'Agriculture a publié, en 1989, un atlas de facture classique, composé d'une quarantaine de planches cartographiques se rapportant au milieu naturel, à la gestion du territoire, aux voies de communication et aux activités économiques. Ces cartes, imprimées à l'échelle du 1/1 400 000, forment un ensemble cohérent, pour rendre compte des principaux ensembles régionaux, et homogène sur le plan cartographique: même échelle et même projection. Dans cet atlas, on a sélectionné un certain nombre de cartes pouvant faire partie d'un SIG: les niveaux d'altitude (fig. 2), les ensembles géomorphologiques, les types de sols (fig. 3), les températures moyennes annuelles, les précipitations moyennes annuelles, les cours d'eau et les bassins hydrographiques (fig. 4) et les routes, les limites municipales, et les chefs-lieux de *municípios* (fig. 5) (2).

Parallèlement à ces cartes, on dispose du Système d'Analyse des *Municípios* Brésiliens Agrégés (SAMBA) réalisé en coopération avec l'Institut Brésilien de Géographie et de Statistique (IBGE). Il s'agit d'un vaste ensemble de données censitaires sur la population et l'agriculture, dans lequel l'unité d'observation est le *município* (4 500 pour l'ensemble du Brésil, 310 pour le seul Paraná). Après diverses analyses, réalisées précédemment à l'aide des logiciels SAS et UNISAS sur gros système IBM, une soixantaine d'indicateurs se rapportant à la superficie des exploitations, à la structure foncière, à l'utilisation du sol, aux techniques culturales et aux diverses productions, ont été sélectionnés.



2. Les niveaux d'altitude



3. Les types de sols

À l'issue de la phase d'acquisition des données, les informations à analyser se composent d'un jeu de cartes thématiques sur papier et d'une matrice d'information spatiale sur support magnétique, dont les colonnes renferment des indicateurs statistiques agricoles et dont les lignes se rapportent aux *municípios* présents sur la carte du même nom.

La structuration des données dans le SIG

Le logiciel *GéoConcept*, sur micro-ordinateur Macintosh, a été retenu pour le SIG Paraná. Il s'agit d'un système de conception récente, conçu et diffusé par la société française Alsoft, dont l'un des atouts est de fonctionner à la fois sur Macintosh et sur PC/Windows.

Avec *GéoConcept*, la notion de plan ou couche d'information (*layer*) n'existe pas. Les données sont structurées par objet, sans pour autant que ce vocable recouvre les concepts qu'il désigne



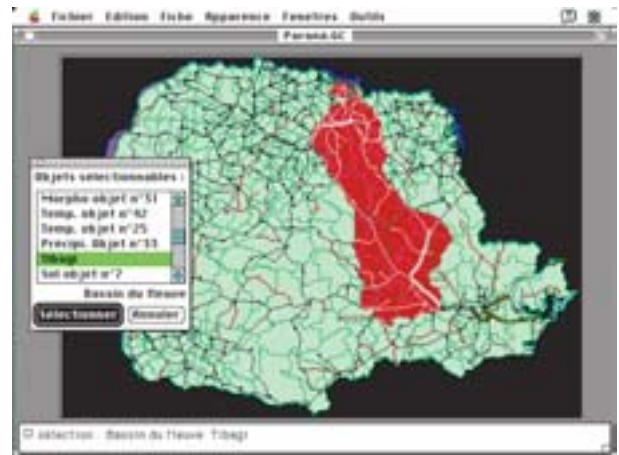
4. Les cours d'eau et les bassins hydrographiques



5. Les routes, les limites municipales et les chefs-lieux de *municípios*

habituellement en informatique. Plus simplement, un objet est un élément d'une carte tel un niveau d'altitude, une unité de sol, un bassin hydrographique ou un *município*. Un objet est défini par son genre (ponctuel, linéaire, aréal ou textuel), son type (type de sol), son sous-type (latosol rouge), son nom et les coordonnées de ses contours.

Tous les objets de tous les types sont empilés sur le bureau du SIG dans un ordre défini lors de la configuration du système. Ainsi, un point de l'espace appartient à la fois à un niveau d'altitude, à une unité de sol, à un bassin hydrographique et à un *município*, etc. Cette structuration des données peut être vérifiée en désignant un point de l'espace, le logiciel proposant alors une liste d'objets sélectionnables auquel ce point appartient (fig. 6). De même, plusieurs objets de types différents peuvent être inclus, former une intersection, être situés à une certaine distance. Ainsi, un *município* peut être localisé dans un bassin hydrographique: il est alors inclus dans ce bassin; et être desservi par



6. L'empilement d'objets géographiques

Après avoir désigné un point de l'espace, une fenêtre s'affiche, proposant la sélection de tous les objets auxquels appartient ce point. Ici, le bassin du fleuve Tibagi est sélectionné.

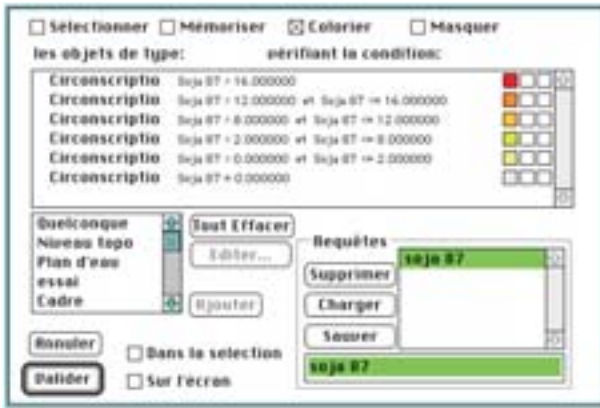


7. La fiche d'identification du *município* de Londrina

Le *município* est sélectionné en rouge. La fiche se compose d'un ensemble de champs obligatoires, type (circonscription), sous-type (*município*), nom (Londrina), premières coordonnées géographiques de cet objet (x et y), et de champs ajoutés après la création du système et dont les valeurs sont importées d'un fichier de statistiques agricoles (répartition des exploitations par classes de superficie, force de travail, etc.).

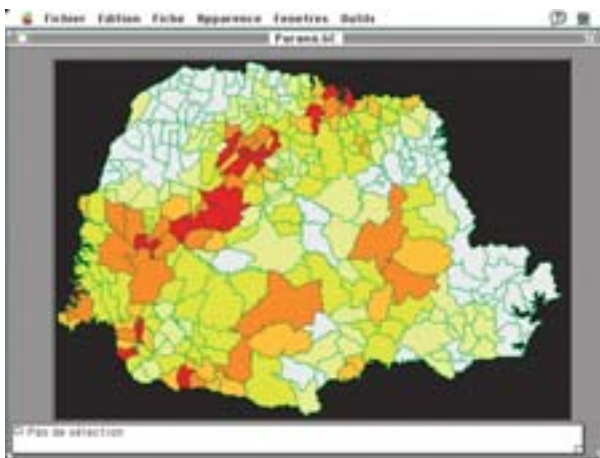
une route: il est en intersection avec cette route. Ces propriétés seront utilisées plus loin par les requêtes topologiques.

L'entrée des données dans le système ne pose pas de problème particulier puisque *GéoConcept* importe des fichiers textuels (ASCII) selon un format décrit par le fabricant. Les cartes sur papier ont tout d'abord été scannées, puis redessinées avec le logiciel *Adobe Illustrator* sur micro-ordinateur Macintosh, et enregistrées dans un fichier contenant la description de la carte en langage Postscript (comprenant notamment l'identification



8. Requête statistique simple

Elle est mise en jeu pour la réalisation de la carte de la figure 9.

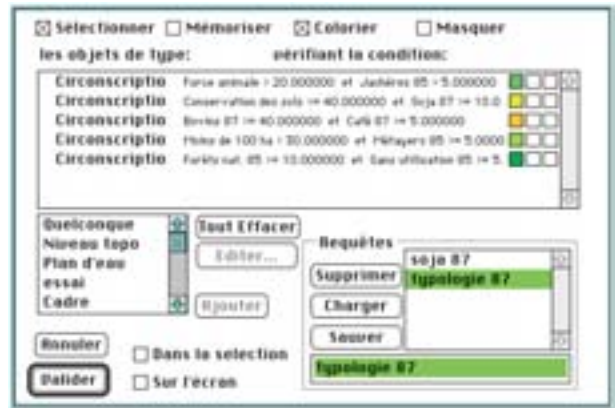


9. Part du soja en 1987

et les coordonnées de chaque objet). Ce fichier Postscript est ensuite mis en conformité au format reconnu par *GéoConcept*, puis importé dans le SIG après que celui-ci ait été configuré pour le recevoir, c'est-à-dire après définition de chaque type et sous-type. Lors de cette importation, une fiche d'identification est automatiquement créée pour chaque objet et contient son type, son sous-type et ses coordonnées. Cette fiche peut être étendue afin d'accueillir d'autres attributs propres à ces objets. Si, par exemple, la fiche-type des *municipios* doit être complétée pour recevoir les statistiques agricoles, après configuration de cette fiche, les données en format texte seront simplement importées (fig. 7).

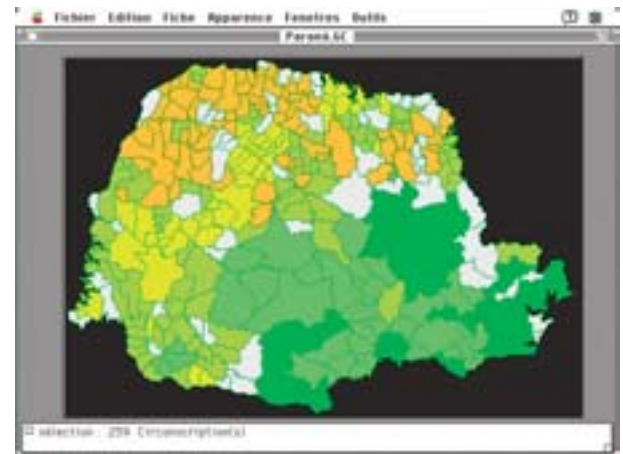
L'utilisation du SIG

Pour que l'effort demandé par la structuration des données soit justifié, il faut que le SIG soit en mesure de réaliser des analyses, difficiles à pratiquer autrement. Les traitements se font



10. Requête statistique complexe

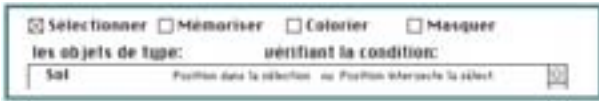
Chaque ligne donne les caractéristiques d'un type particulier de mise en valeur: 1. agriculture de subsistance, avec jachère et traction animale; 2. production intensive de soja; 3. élevage bovin et production de café; 4. petit métayage; 5. terrains forestiers ou inutilisés. Ces types ont été construits après analyse en composantes principales et classification ascendante hiérarchique réalisées hors de *GéoConcept* qui n'intervient ici que pour le tracé de la carte.



11. Une typologie multivariée des *municipios*

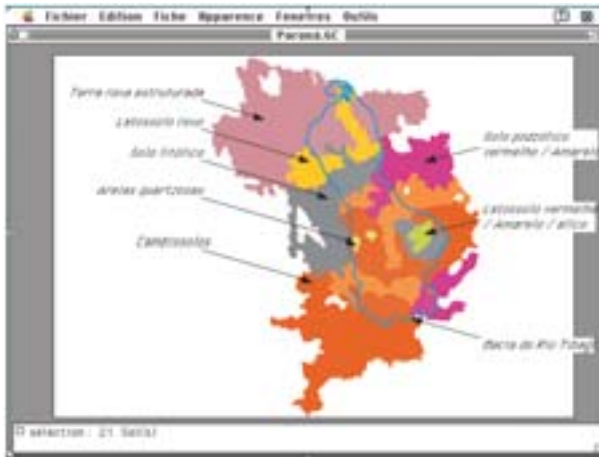
sous la forme de requêtes. Il s'agit de questions adressées au système, dont la réponse est une sélection d'objets conformes à la demande. Ces requêtes sont de courts textes, énonçant les conditions à réunir et, éventuellement l'action à exercer sur les objets sélectionnés (coloriage, masquage, etc.). Elles peuvent concerner les attributs composant les fiches d'identification (requêtes statistiques), faire appel aux positions des objets dans l'espace (requêtes topologiques) ou combiner les deux (requêtes mixtes).

L'archétype d'une requête statistique revient à colorier un fond de carte comme en cartomatique: qui peut le plus peut le moins... La requête (fig. 8) désigne le type d'objet (circonscrip-



12. Requête topologique

Après sélection directe sur la carte du bassin du fleuve Tibagi (fig. 6), requête nécessaire à l'obtention des unités de sol incluses ou en intersection avec ce bassin.



13. Les unités de sol incluses ou en intersection avec le bassin du Tibagi

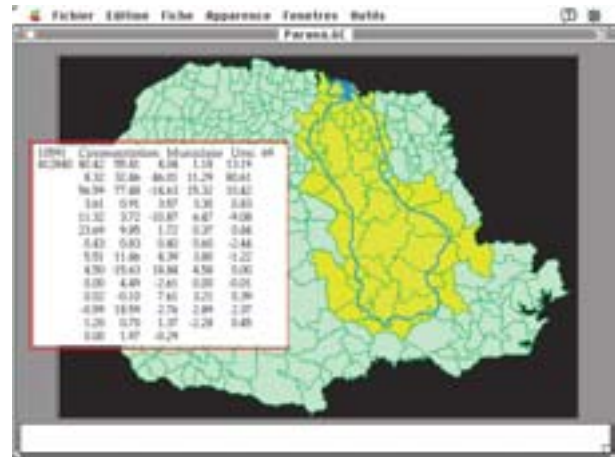
tion municipale) et la manière dont ils doivent être coloriés (fig. 9) en fonction des valeurs prises par la variable «soja 87», qui correspond à un pourcentage de soja dans la valeur totale de la production. Jusque-là, rien de bien nouveau! Plus complexe apparaît la construction d'une typologie basée sur un ensemble de 10 variables (fig. 10); si le SIG est capable de tracer une carte (fig. 11) difficile à réaliser à l'aide d'un logiciel de cartomatique, aucun outil d'analyse des données n'est proposé, et il faut faire appel à un logiciel extérieur pour identifier les indicateurs discriminants. Même si elles autorisent la cartographie interactive des données, les requêtes statistiques n'apparaissent donc pas comme une innovation bien significative.

Avec les requêtes topologiques, on aborde l'aspect le plus nouveau du concept SIG. De telles requêtes permettent d'examiner les positions respectives d'objets de types différents et de répondre à des questions telles que: «Quels sont les sols présents dans tel bassin hydrographique?». Deux requêtes sont nécessaires, la première pour sélectionner le bassin hydrographique à analyser, ici le bassin du fleuve Tibagi; la seconde (fig. 12) pour sélectionner les unités de sol dont la position est dans la sélection, à savoir celles strictement incluses dans le bassin, et celles dont la position intersecte la sélection et qui chevauchent plusieurs bassins hydrographiques. La carte obtenue (fig. 13) montre l'existence d'un fort contraste de qualité agronomique entre le sud du bassin, doté de sols relativement



14. Autre requête topologique

Après sélection directe sur la carte du bassin du fleuve Tibagi (fig. 6), requête nécessaire à l'obtention des *municípios* inclus ou en intersection avec ce bassin.

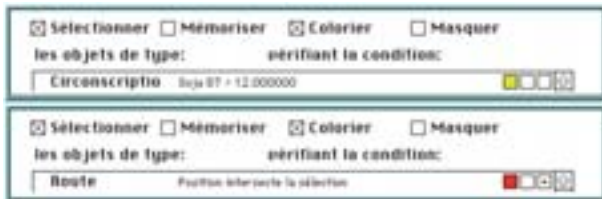


15. Les *municípios* inclus ou en intersection avec le bassin du Tibagi

Dans l'encadré, les indicateurs agricoles de l'un de ces *municípios* (Uraí) tels qu'ils sont exportés vers un logiciel d'analyse statistique ou autre.

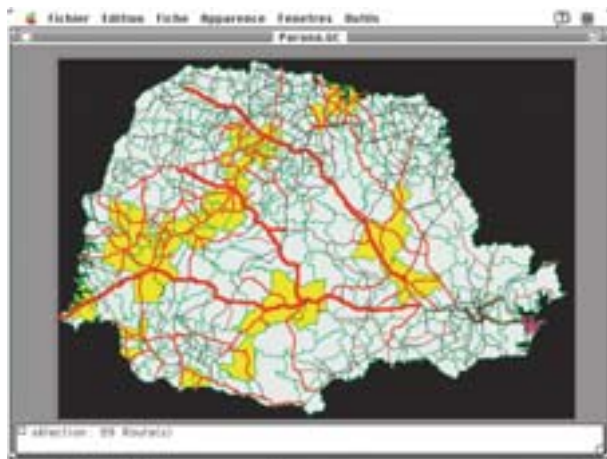
mauvais comme les *cambissolos*, et le nord où dominent la *terra roxa* et les *latossolos roxos*, bien plus fertiles et aptes à la mécanisation. De la même manière, on peut sélectionner les *municípios* appartenant entièrement ou en partie à ce bassin hydrographique (fig. 14) afin d'exporter leurs données agricoles ((fig. 15) vers un logiciel d'analyse statistique pour, par exemple, réaliser une analyse de variance, permettant d'apprécier quelles sont les variables pour lesquelles les *municípios* du bassin sont significativement différents de ceux qui n'y sont pas. En généralisant à tout type d'objet géographique la pratique ancienne de la superposition afin de sélectionner ceux qui occupent une position particulière vis-à-vis des autres, les requêtes topologiques améliorent une voie qu'il était laborieux et fastidieux de parcourir auparavant.

Enfin, rien n'interdit de combiner requêtes statistiques et requêtes topologiques dans des requêtes mixtes. Ceci permet de mettre en évidence la combinaison de la localisation géographique avec la position statistique. Ainsi, on sait que le soja, culture d'exportation par excellence, nécessite une infrastructure routière reliant zones de production et port d'exportation. Pour examiner la position des aires productrices de soja par rapport au réseau routier, deux requêtes doivent se succéder (fig. 16). La première sélectionne les *municípios* où le soja,



16. Requête mixte

Elle combine deux requêtes s'exécutant consécutivement et destinées à sélectionner et colorier en rouge les routes qui sont en intersection avec les *municípios* où le soja, en 1987, représente plus de 12% de la valeur totale de la production agro-pastorale (en jaune).



17. Les routes en intersection avec les *municípios* fortement producteurs de soja, en 1987

Les couloirs d'exportation reliant ces aires de production au port de Paranaguá (magenta) sont figurés en trait rouge plus épais.

avec plus de 12% de la valeur totale de la production agricole, est économiquement très important; la seconde retient les routes qui traversent ces *municípios*, c'est-à-dire sont en intersection avec eux (fig. 17). On note que deux aires de production secondaires se sont développées sur les axes reliant la principale aire de production, au nord-ouest de l'État, au port de Paranaguá (en magenta): c'est une preuve supplémentaire de l'importance des couloirs d'exportation (trait rouge plus épais) sur la dynamique spatiale de ce genre de produit agricole. En permettant d'adopter simultanément un point de vue spatial et un point de vue statistique, les requêtes mixtes permettent de tester des hypothèses ou de confronter aisément les données aux connaissances acquises par ailleurs, qu'elles soient établies scientifiquement ou qu'elles correspondent aussi à des idées reçues ou «de bon sens».

Une innovation «incontournable»?

Cet article relate une expérience et expose un point de vue basé sur cette expérience. Ses conclusions n'ont pas vocation à être

généralisées mais, après quelques mois de mise en pratique de *GéoConcept*, voici quelques réflexions à l'usage de ceux qui veulent se «lancer» dans une opération de même nature.

En premier lieu, on peut «ensiler» à loisir des volumes considérables de données sans pour autant produire *ipso facto* une information utile, faisant progresser la connaissance de l'espace géographique, objet de la recherche.

En second lieu, il faut insister sur la lourdeur relative des opérations d'acquisition et de structuration des données. Même avec un logiciel aussi facile à utiliser que *GéoConcept*, la présence d'un cartographe dans l'équipe du projet est indispensable. De même, en l'absence de bases de données déjà constituées, il est préférable de choisir avec circonspection les cartes et les données statistiques à introduire dans le SIG. Dans la mesure du possible, il est même souhaitable de préciser quelles opérations seront réalisées, sur quels objets et dans quels buts. Cela rejoint la préoccupation précédente, même si, dans la pratique, on ne peut exclure une part de découverte lors de l'exploration des données.

Enfin, en admettant qu'un tel outil conduise à systématiser l'approche hypothético-déductive et invite à ne plus considérer les données cartographiques ou statistiques seulement comme des illustrations, il n'exclut, en aucun cas, le recours à d'autres instruments d'analyse, statistique notamment. La standardisation des opérations d'importation-exportation des données sur Macintosh incite d'ailleurs à cet usage complémentaire des logiciels. Le SIG apparaît alors comme une sorte de «tour de contrôle» qui organise l'arrivée des données de natures diverses destinées à être mises en relation, et d'où partent, soit ces mêmes données restructurées et sélectionnées pour une analyse ultérieure, soit des résultats de requêtes prouvant le bien fondé d'une hypothèse.

Incontournables les SIG? Oui, sans aucun doute! Mais gardons-nous des pratiques abusives comme celles qui ont été observées à propos de la cartomatique et de l'analyse des données, et, à l'heure actuelle, celles de la mode «tout SIG», dont l'effet immédiat est de... brouiller les cartes.

(1) Comment ne pas sourire de ces cartes de résidus de régression linéaire qui montrent la variation systématique des résidus avec la variable à expliquer, démontrant ainsi, et seulement, que ce modèle est inadapté aux données?

(2) Les cartes de l'atlas présentées ici apparaissent, sur les figures 2 à 5, telles qu'elles s'affichent sur le bureau de *Géoconcept*.

Références bibliographiques

Atlas do Estado do Paraná, 1989, Curitiba, 73 p., 38 planches.
WANIEZ Ph., 1990, *Systèmes d'information géographique sur Macintosh*, Paris, Eyrolles, coll. Informatique, 120 p.