

ECHELLE ET DONNEES NUMERIQUES

*Par G. Troispoux et L. Coudercy
C.E.R.T.U/France*

L'information géographique sous sa forme numérique s'impose de plus en plus face aux traditionnelles cartes papier. L'échelle, paramètre indispensable à la description de toute représentation cartographique sur un support papier, devient, avec l'avènement du numérique, une notion plus difficile à percevoir, plus subtile, qui se décline de différentes façons. Constante sur un support papier, l'échelle des données numériques s'avère être une "variable" dont la maîtrise est alors plus difficile.

Autour de cette notion d'échelle, nous allons essayer de comparer les deux approches (carte papier et données numériques) en s'appuyant sur les autres paramètres qui définissent et précisent l'information géographique dans son ensemble, et qui sont liés à l'échelle des données.

Carte papier :

Echelle d'une carte papier :

L'échelle est le rapport entre la dimension d'un objet mesurée sur la carte et sa valeur réelle. Dans le cas d'une carte papier, elle est forcément fixe. Les échelles les plus couramment utilisées pour la cartographie française peuvent être réparties dans les gammes d'échelles suivantes, en considérant que cette classification peut varier selon les métiers et les usages :

Les très grandes échelles :

1 :200, 1 :500, 1 :1000,

Les grandes échelles :

1: 2 000, 1: 5 000, 1: 10 000

Les échelles moyennes :

1: 25 000, 1: 50 000, 1: 100 000

Les petites échelles :

1: 250 000, 1: 500 000, 1: 1 000 000...

Contenu :

Le contenu d'une carte papier est lié à l'échelle pour de simples règles de lisibilité et de représentation graphique, mais aussi par les habitudes issues de la longue maturation des cartes existantes. Il est tout naturellement limité par le fait qu'il n'y a pas de place sur la feuille de papier pour dessiner et représenter tous les détails. Dans de nombreux cas, un choix est indispensable parmi les objets pouvant figurer sur une carte. Ce choix n'est pas exclusivement fonction de leurs dimensions, qui doivent rester en rapport avec l'échelle de la carte, mais il peut dépendre de leur importance pour les utilisateurs ou de leur fonction. Il faut également tenir compte de la représentation sémiologique des données qui contribue à limiter ce contenu, soit par la place occupée par les symboles, soit par le nombre limité de classes affichables.

Précision :

La précision représente la justesse d'une mesure. Dans le cas d'une carte, il s'agit donc de la différence entre la position papier d'un objet et sa position terrain. Elle revêt divers aspects.

Précision de saisie :

Elle est dépendante de la technique utilisée. La restitution des levés étant purement graphique, la précision de saisie sera forcément limitée par le pouvoir séparateur de l'œil, c'est à dire sa capacité à dissocier deux détails.

Cette valeur est comprise entre 1 et $2/10^{\text{ème}}$ de mm et correspond donc à la précision de saisie maximale pour une échelle donnée. Il serait inutile de pratiquer des levés avec une précision meilleure.

Précision de restitution :

C'est la précision finale de la carte.

Elle est à priori moins bonne que la précision de saisie, car les phases de cartographie et de reprographie détériorent légèrement la qualité initiale des levés. En particulier, dès lors qu'un levé est cartographié, c'est à dire que l'on a appliqué un certain nombre de règles d'esthétique et de lisibilité, puis imprimé le résultat sur un support papier non stable, cette précision devient variable selon les éléments de la carte et peut atteindre plusieurs dixièmes de mm.

Par exemple, pour une carte au 1:10 000, la précision graphique de 1 à $2/10^{\text{ème}}$ de mm, se traduit à cette échelle par une précision de saisie de 1 à 2 m. La précision de restitution de la carte sera moins bonne et la dégradation introduite, variable.

Déformation due à l'écriture cartographique :

Cartographier un levé topographique, c'est appliquer un signe conventionnel à chaque objet. Il faut que tout élément graphique isolé soit perceptible, que sa forme soit distinguée, que deux éléments voisins soient séparés, enfin que les différentes valeurs soient différenciées.

Pour des règles d'esthétique et de lisibilité, chaque signe doit avoir une épaisseur minimale. Pour cartographier l'ensemble des objets et pouvoir les différencier visuellement, il faut souvent déplacer certains d'entre eux de quelques dixièmes de millimètres.

On constate donc que la phase de cartographie altère la précision finale de la carte.

Par exemple, le seuil de séparation de deux traits est de 0,15 mm. L'épaisseur minimale d'un trait est également de 0,15 mm. Si l'on veut représenter une route à l'échelle de 1:25 000 par un signe conventionnel à deux traits, la largeur minimale du signe sera donc de 0,45 mm, soit 11,25 m sur le terrain à cette échelle. Ainsi, toutes les routes dont la largeur est inférieure à 11,25 m auront forcément une représentation amplifiée sur une telle carte. Aussi, il faudra déplacer d'autant, voire plus, tous les objets qui bordent la route.

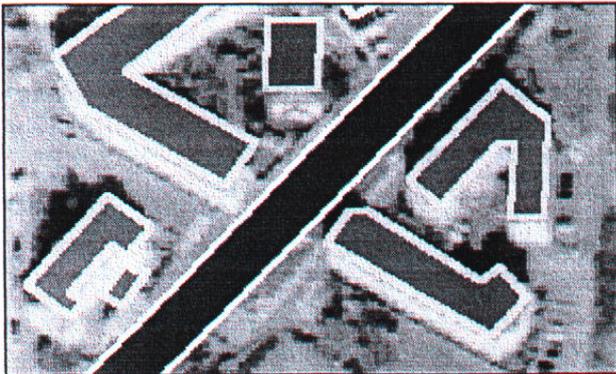


Figure 1 : emprises d'une route et de bâtiments sur une photographie aérienne

La figure 1 montre un extrait de photographie aérienne sur laquelle sont positionnés quelques bâtiments levés par photogrammétrie ainsi que l'emprise surfacique d'une route.

La figure 2 montre la même zone extraite de la carte IGN au 1:25 000 très agrandie. On y a reporté les bâtiments levés sur la photographie qui demeure l'image de la réalité. Cet exemple montre la dégradation subie par les données lors de la phase cartographique que l'on peut résumer ainsi:

- L'emprise cartographiée de la route déborde largement sur les bâtiments de droite qu'il a fallu déplacer.
- Le bâtiment en bas à droite a été grossi.

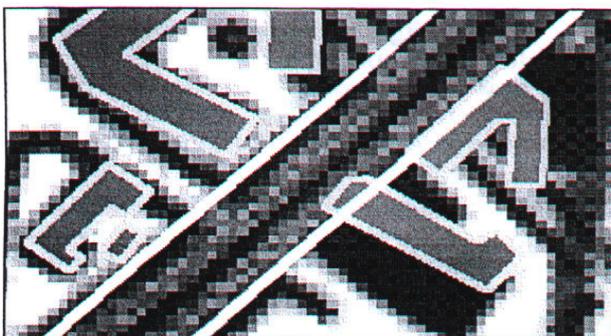


Figure 2 : extrait de la carte IGN au 1:25 000 sur laquelle ont été reportés les bâtiments présents sur la photographie

Précision relative

C'est la précision du positionnement des objets les uns par rapport aux autres.

Cette précision est liée à la précision absolue de la donnée, mais aussi à la rigueur du dessin cartographique. En effet, le respect des règles cartographiques amène à déplacer certains objets, tout en tenant compte autant que possible de la position relative des objets voisins. Dans le cas d'objets déplacés, elle peut être meilleure que la précision absolue.

Généralisation :

La généralisation est le procédé de simplification systématique de la représentation d'un objet pour l'adapter à une échelle donnée.

Ce procédé peut être mis en œuvre dès la saisie ou lors de la représentation cartographique. Toute carte, quelle que soit son échelle, suppose une généralisation des tracés par rapport à la réalité du terrain. A une échelle donnée, une carte peut être plus ou moins généralisée. Dans le cas d'une généralisation trop poussée, on parle alors de schématisation. La généralisation est étroitement liée à la précision mais aussi au contenu et à la densité des données ainsi qu'à la symbolisation qu'on désire appliquer.

La figure 3 ci-dessous montre le réseau de voirie au 1: 1000000 entre Routes500® à gauche, et Route120® à droite (les exemples sont basés sur des bases de données, plus facile à manipuler).

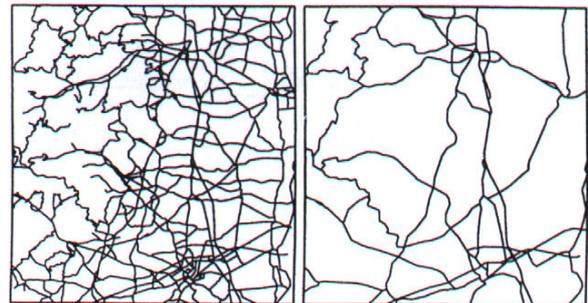


Figure 3 : lien entre généralisation et contenu

Cette illustration souligne surtout l'un des deux aspects indissociables de la généralisation : le contenu et la densité d'information, qui sont différents selon le niveau de généralisation souhaité.

La figure 4 superpose Routes500® (traits pleins) et Route120® (tiretés). Cette illustration souligne le second aspect de la généralisation : la simplification des formes qui influe sur la précision.

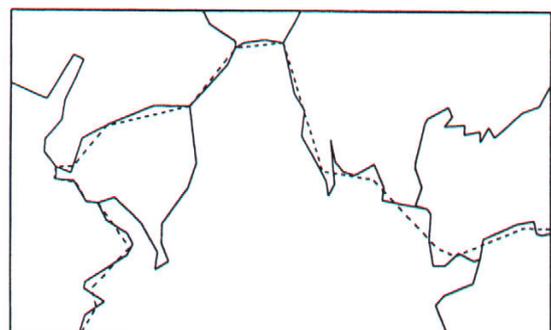


Figure 4 : lien entre généralisation et précision

Grain :

Le grain est défini comme la plus petite entité qu'il est possible de reproduire sur une carte. A une échelle donnée, sur une carte papier, il est directement lié aux dimensions propres des objets à faire figurer.

On peut cependant toujours contourner cet aspect par une représentation à l'aide d'un signe conventionnel. Par exemple, sur les cartes au 1:25 000, les calvaires, objets trop petits pour être visibles à cette échelle, sont pourtant figurés par des symboles.

Exhaustivité :

L'exhaustivité est variable selon la nature des objets. Elle est fonction des spécifications de saisie et du grain. Elle dépend également de l'objectif à atteindre et du type de carte à réaliser (topographique ou thématique).

Données numériques :

Echelle :

Dans le cas d'une base de données géographiques, la notion d'échelle devient plus floue du fait de la nature même des données en lien avec les outils qui permettent de saisir, manipuler, visualiser et restituer cette information. On ne peut plus parler de l'échelle d'une base de données géographiques mais d'une gamme d'échelles optimale pour l'utiliser.

La gamme d'échelles est évidemment liée à la précision de saisie. Mais elle est à adapter en fonction des modes de restitution et d'usage envisagés. Par exemple, en fonction de la résolution et de la taille de l'écran sur lequel on désire visualiser les données, on choisira une échelle adaptée à la fois au travail à réaliser et au confort oculaire.

Cette échelle de travail et de confort n'est pas forcément celle qui est le plus en adéquation avec la précision des données, pour une restitution cartographique papier.

Nous essaierons, dans la suite de ce dossier, de préciser tous ces aspects.

Restitution :

La restitution se fait soit sur un écran vidéo, soit sur un support papier. Ces deux moyens de restituer l'information sont totalement différents. Pour un travail donné, une échelle adaptée à un type de restitution est rament adaptée à l'autre. De nombreux paramètres influent sur l'échelle de restitution comme :

- les caractéristiques physiques des systèmes ;
- le degré de symbolisation des données ;
- la restitution en couleur ou en noir et blanc ;
- le contenu des données...

Précision :

Précision de saisie :

Comme pour une carte papier, la précision de saisie reste liée à la technique utilisée. C'est cette précision qui définit

la gamme d'échelles la mieux adaptée à l'utilisation ultérieure de ces données. Elle peut être variable selon la nature des objets saisis et doit être décrite et quantifiée, par classe d'objets, dans les spécifications du produit. Elle n'est plus limitée systématiquement par l'échelle de restitution et la règle du $1/10^{\text{ème}}$ de mm.

La précision devient très importante pour les données numériques car elle remplace, en partie, la traditionnelle échelle des cartes papier. On parle de précision métrique, décimétrique. De ce paramètre, on déduit la gamme d'échelles optimale pour l'utilisation des données.

A l'inverse d'une carte papier dont l'échelle et la légende permettent de déduire sa précision et son contenu, une base de données numériques est spécifiée par la précision des éléments qui la constituent.

Deux types de saisie sont couramment pratiqués, qui influent directement sur la précision des données. Le premier consiste en des levés topographiques soit directement sur le terrain soit par photogrammétrie. La précision résultante est celle normalement définie dans le cahier des charges.

L'intérêt d'une telle saisie réside dans le fait que les données ne subissent aucune altération par la mise en forme cartographique, la précision définitive restant égale à celle des levés. L'autre type de saisie consiste à numériser des cartes existantes. Dans ce cas, on possède des données numériques dont la précision de saisie est moins bonne que la précision de restitution cartographique des cartes utilisées et qui n'est pas homogène sur le territoire.

C'est le cas de la BDCARTO® qui résulte de la numérisation des cartes au 1:50 000 de l'IGN, ce qui aboutit à une précision non homogène. Si l'on superpose les données de la BDCARTO® sur une orthophotographie ou sur une carte à plus grande échelle (Figures 5 et 6), on constate que la superposition est nettement meilleure en zone rurale qu'en zone urbaine. Ceci s'explique par le nombre important de données à cartographier au 1/ 50.000 dans un noyau urbain qui subit, de fait, des déformations liées à une écriture cartographique plus poussée qu'en zone rurale. Dans ce dernier cas, la densité des objets étant plus faible, leur positionnement reste plus proche de la précision de saisie.

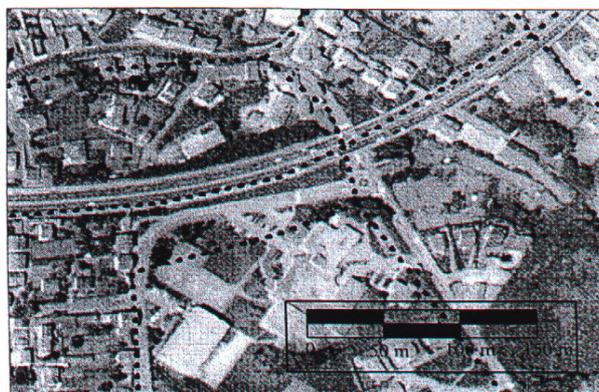


Figure 5 : superposition de la BDCARTO® et d'une carte au 1.25 000 en milieu rural

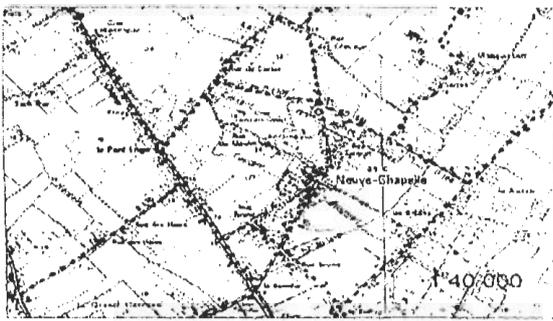


Figure 6 : superposition de la BDCARTO© et d'une carte au 1:25 000 en milieu rural

Précision de restitution :

La plupart des données numériques commercialisées sont vendues sans aucun habillage cartographique. L'avantage réside dans le fait que les utilisateurs disposent alors des données brutes de saisie sans dégradation de précision. La précision de restitution ne dépend plus alors que des moyens mis en œuvre par l'utilisateur : imprimante, sémiologie, généralisation...

Précision relative :

Dans la mesure où les données n'ont reçu aucune symbolisation cartographique, la précision relative reste celle de la saisie.

Grain

Comme pour une carte papier, il définit une limite de saisie des données qui se traduit souvent par une restriction concernant les dimensions minimales des objets.

On peut également limiter le contenu d'une classe d'objets pour des raisons plus sémantiques comme la nature ou la fonction. Par exemple, on peut exclure les tronçons de route dont la longueur est inférieure à 500 m ou seulement exclure ceux qui mènent à une propriété privée, voire une combinaison des deux.

Pour une base d'occupation des sols, le grain s'exprime également par la plus petite zone saisie. Ainsi dans la BDCARTO, le grain est variable selon la classe de la nomenclature concernée : il est plus fin pour l'urbain que pour les bois.

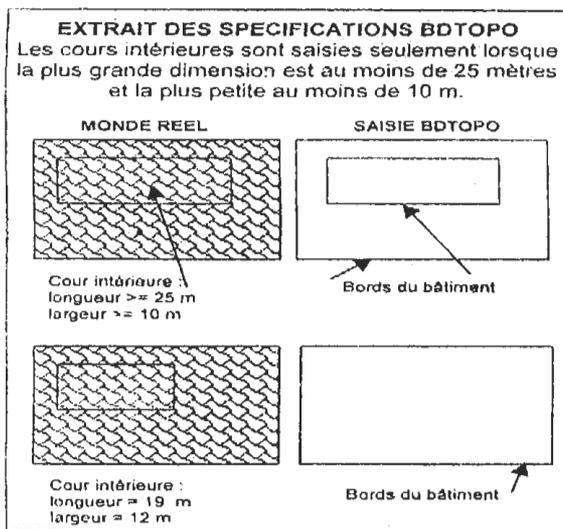


Figure 7 : extrait des spécifications BDTOPO

Enfin, par extension, les plus petits tronçons linéaires saisis pour décrire les détails d'un objet qui découlent du niveau de généralisation choisi représentent une forme de grain, souvent importante à connaître. L'exemple de la figure 7 montre un extrait des spécifications de la BDTOPO dont les limites exprimées en mètres ont un lien direct avec le grain.

Cette notion de grain définit ainsi une sélection pour chaque classe d'objets de la base de données, tout en fixant, par type d'objet, l'unité graphique élémentaire que l'on pourra espérer représenter. Cette notion mériterait d'être affinée, car, si l'on voit assez bien ce qu'elle recouvre, elle est parfois difficilement représentable par une grandeur unique.

La connaissance du grain, ajoutée à la connaissance de la précision, permettent de déduire une gamme d'échelles utilisable pour les données ainsi spécifiées.

Contenu :

Le contenu d'une base de données reste évidemment lié à la précision et aux dimensions des objets à saisir.

Cependant la problématique est ici différente de celle d'une carte papier. Cette dernière est en effet destinée à être imprimée avec la totalité des objets qu'elle contient. En revanche, on peut considérer une base de données comme une collection d'objets qui n'ont pas forcément vocation à être restitués simultanément. Dans ce cas, les objets pouvant être contenus dans une base sont directement fonction de sa précision et de son grain, ais deviennent complètement

indépendants de sa représentation. Ainsi, rien n'empêche de constituer une base de données avec un grand nombre de classes d'objets en sachant qu'elles ne seront jamais représentées simultanément.

Exhaustivité :

Pour une base de données, l'exhaustivité revêt plus d'importance que pour une carte papier car de cette exhaustivité dépend en grande partie, la qualité des analyses. Elle est étroitement liée aux spécifications du produit et notamment au grain, qui impose des règles de sélection, ainsi qu'aux techniques de saisie. En revanche, elle peut être plus facilement atteinte dans la mesure où elle reste indépendante de la représentation cartographique de la base.

Besoins de précision et de grain :

Les données numériques exploitées par des outils logiciel de type SIG sont particulièrement adaptées à l'analyse, qu'elle soit spatiale ou thématique, mais servent aussi à la cartographie. Selon le type d'usage, les besoins en terme de précision et de grain ne sont évidemment pas les mêmes. Dans le cas d'une analyse spatiale, la qualité attendue d'une base n'est pas uniquement fonction de la précision géométrique des données ou de l'échelle comme cela est le cas pour une carte papier.

D'autres paramètres influent alors comme la structuration des données, l'exhaustivité, les spécifications, ou la cohérence. En particulier, les utilisateurs peuvent, pour certains usages, être plus sensibles à la cohérence géométrique entre couches d'information qu'à la précision réelle. Ainsi, quand on veut connaître les rivières bordant une commune, la cohérence géométrique entre la couche rivière et la couche de limite administrative est indispensable.

Une base de données issue d'une saisie topographique sera souvent plus adaptée à l'analyse. A l'inverse, pour des restitutions cartographiques, une base de données issue de la digitalisation d'une carte sera plus simple d'emploi, une part importante du travail d'écriture cartographique étant alors déjà réalisée.

L'échelle ou la précision :

La précision et le grain remplacent l'échelle :

Comme nous venons d'essayer de le montrer, la notion d'échelle n'a plus tout à fait la même signification quand elle fait référence à des données numériques.

Attribut indispensable d'une carte papier, elle permet d'en déduire une précision moyenne et d'avoir une bonne idée sur son contenu.

Pour une base de données numériques, c'est plutôt la précision des données et la connaissance du grain qui permettent d'en déduire une gamme d'échelles utilisable.

C'est bien la précision et le grain qui ont pris le pas sur l'échelle pour qualifier une base de données.

Par exemple, en annonçant une précision "décamétrique" pour la BDCARTO®, l'IGN donne un ordre de grandeur qui se situe autour de 10 à 40 mètres de précision sur chaque point. En tenant compte de la traditionnelle résolution graphique du 1:10^{ème} de mm, cette précision équivaut alors à une précision de restitution de cartes dont la gamme d'échelles la mieux adaptée est entre 1:50 000 et 1:200 000. Ces valeurs correspondent à la gamme d'échelles de travail couramment admise, c'est à dire celle où toute étude, toute analyse garde un sens cohérent avec la représentation et la généralisation des données.

Le grain est défini par les spécifications de contenu : en prenant l'exemple des tronçons de route, les sélections sur ces objets imposent des longueurs minimales de 200 m, c'est à dire 1 mm à l'échelle du 1: 200000 que l'on peut considérer comme une limite extrême pour la représentation graphique d'un objet linéaire.

A ces deux paramètres principaux que sont la précision et le grain, s'ajoute la connaissance de la technique de saisie utilisée. Il est clair que les bases de données issues de la numérisation de cartes papier auront une précision finale moins bonne et moins homogène qu'une base de données issue de levés topographiques ou photogrammétriques.

L'échelle de restitution :

Un autre aspect de la représentation des données est bien sûr l'impression sur papier et les multiples sorties

graphiques qui sont un aboutissement logique à tout travail sur ce type de données.

Contrairement aux habitudes "d'avant l'informatique" où l'échelle est fixe et ne peut prendre que certaines valeurs entérinées par l'usage et l'habitude, on voit apparaître avec l'informatique des échelles quelconques (16l:13500, 1:4 560...). En effet de nombreuses cartes éditées aujourd'hui à partir d'outils logiciel sont destinées à illustrer des rapports d'étude édités dans le format A4. Quelle que soit la taille terrain concernée, l'échelle prend alors des valeurs inhabituelles.

La qualité de ces impressions dépend de plusieurs paramètres, tels que la qualité de l'imprimante utilisée, de son format, du papier, de l'encre mais aussi des compétences de l'utilisateur en sémiologie graphique et de sa capacité à trouver la meilleure adéquation entre l'objectif à atteindre et les solutions logiciel et matériel dont il dispose. Ajoutons à cela que les logiciels actuels sont souvent très limités dans leur possibilité sémiologique.



Figure 8 : réseau routier de la ville de Brest au 1:200 000 issu de la BDCARTO® avec représentation des routes à deux traits

L'exemple figure 8 montre un extrait de la BDCARTO® avec une symbolisation très simple qui s'avère être totalement illisible dans les noyaux urbains à l'échelle du 1:200 000, échelle haute de la gamme d'échelles courantes de cette donnée.

Elle commence seulement à être lisible à l'échelle du 1 :50 000 (Figure 9).

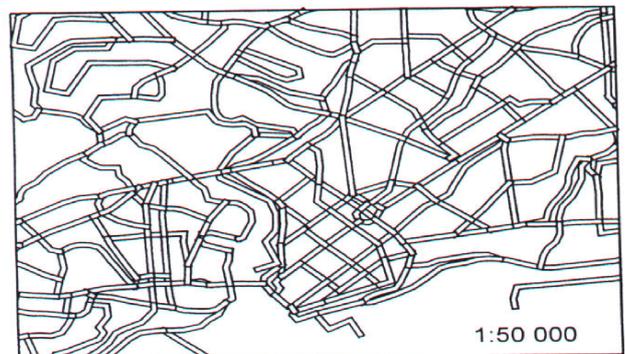


Figure 9 : réseau routier de la ville de Brest au 1:50 000 issu de la BDCARTO®

A propos du zoom

Le "zoom" est un outil particulièrement bien adapté au maniement des données numériques sur écran. Mais, en permettant de grossir indéfiniment l'échelle d'une représentation graphique, il devient dangereux, voire totalement inadapté, tant pour l'utilisateur que pour le producteur. En effet, on ne pourra jamais empêcher certains opérateurs de travailler à une échelle du 1: 1 000 avec des données dont la précision est de 10 mètres, ou de superposer des données du cadastre avec la BDCARTO®, même si cela n'a aucun sens.

Vu sous un angle producteur, cet effet de zoom peut se révéler également fort regrettable. Par exemple, dans le cas d'une saisie photogrammétrique, étant donnée l'excellente qualité des optiques utilisées, un zoom trop important sur les photographies peut amener l'opérateur à saisir une finesse de détails bien au delà des spécifications du produit.

Ceci peut se justifier pour un meilleur confort de travail (calage de fonds raster, duplication de la géométrie...). L'opérateur doit être conscient des limites d'une telle manipulation et les intégrer dans ses méthodes de travail.

Non homogénéité des bases :

Contrairement aux cartes papier, aucune contrainte intrinsèque d'homogénéité de la précision et du grain ne pèse sur les bases de données géographiques. En effet, l'ubiquité de l'informatique permet des mélanges de données de précisions différentes au sein d'un même produit.

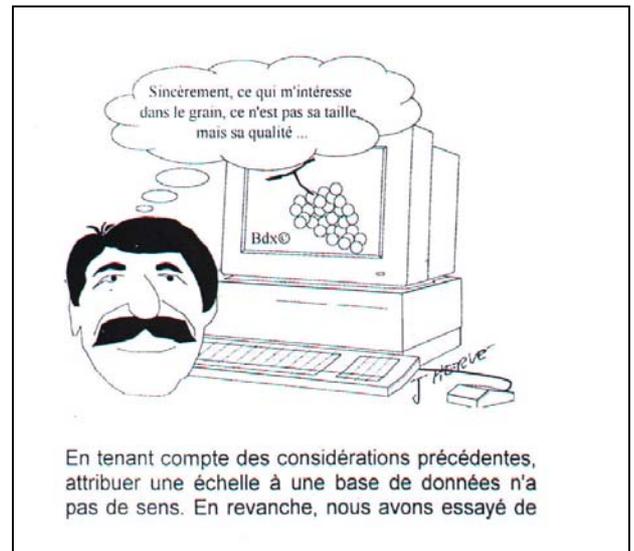
On a déjà vu que le grain, y compris au sein d'une même classe d'objet, pouvait prendre des valeurs variables. On peut citer aussi le cas de la BDTPOPO®, où les bâtiments comportent des détails de l'ordre du mètre, alors que les routes, quelle que soit leur largeur, sont représentées par un simple trait. Ces particularités de spécifications découlent du mode de recueil de l'information qui se fait à partir d'images photoaériennes et qui empêche de voir clairement les limites des chaussées. Ceci aboutit à généraliser des objets de plus de 7 mètres pour la classe route, alors que la base est annoncée comme métrique.

Le fait que dans une même base, on puisse trouver des classes d'objets ayant une précision et un grain fortement différent, n'est pas sans conséquence pour l'utilisateur final, Il lui faut prendre soin de bien analyser les spécifications détaillées d'une base, pour être capable d'en tirer le meilleur partie, et éviter des contre emplois.

Conclusion

Contrairement à la cartographie classique où une carte traditionnelle possède une échelle unique qui fixe à la fois son contenu, sa précision et sa représentation, la notion d'échelle liée à une base de données numériques devient plus complexe mais plus souple.

Tout d'abord, la nature même de ces données permet de les rendre indépendantes de leur représentation et d'offrir, ainsi, plus d'aisance au niveau des échelles utilisables. Cependant, les logiciels et matériels manipulant ces données, tout en assurant une bonne ergonomie, imposent certaines restrictions dues à leurs performances et à leurs limites physiques.



En tenant compte des considérations précédentes, attribuer une échelle à une base de données n'a pas de sens. En revanche, nous avons essayé de montrer que la description de la technique de saisie de cette base **et plus spécifiquement la con-naissance de sa précision et de son grain**, informations parfois propres à chacune des classes d'objets qui la composent, permettent de mieux appréhender la qualité d'usage de ces données.

Enfin, en règle générale, la description d'une base de données géographiques, tant pour son contenu que pour son utilisation potentielle, est beaucoup plus complexe à percevoir et à spécifier qu'une carte papier et laisse l'utilisateur face à des critères de choix à reconstruire. Il y a donc encore beaucoup de travail pour affermir les réflexions actuelles, afin de mieux guider les utilisateurs.

Bibliographie :

- Des données localisées aux systèmes d'information géographique ; conseil aux services, METL, 1996
- Cohérence entre Corine Land Cover et la BDCARTO ; rapport d'étude, CERTU, 1997
- SIG et couches de polygones : le rôle du grain, C ETE Méditerranée, 1998
- Cartographie générale, tome 1, notions générales et principes d'élaboration. R. Cuenin, IGN, 1972
- Topographie, topométrie, R. d'Hollander, IGN, 1976
- La carte mode d'emploi, R. Brunet, Fayard/Reclus 1987.