

## **Des banques de données biogéographiques vers les banques de données biotopographiques. De nouvelles techniques pour de nouveaux enjeux**

Yvan Barbier et Pierre Rasmont

*Reçu 2002-08-15; accepté 2003-02-19*

**PHYTOPROTECTION 84 : 51-57**

---

Le développement conjoint de la technologie GPS, des Systèmes d'Information Géographique et de la micro-informatique permet à l'heure actuelle d'atteindre un niveau de précision extraordinaire pour les localisations géographiques. Une précision de l'ordre de trois mètres est actuellement disponible pour tout un chacun pouvant disposer d'un récepteur GPS moderne. La précision accrue de ces coordonnées géographiques laisse entrevoir de nouvelles possibilités en termes d'études sur la biogéographie, l'écologie des espèces, les populations, etc. Cependant, les banques de données biogéographiques doivent évoluer dans le même sens. En effet, l'amélioration de plusieurs ordres de grandeur dans la précision des coordonnées géographiques demande une gestion beaucoup plus poussée de celles-ci. Il est à présent impératif de tenir compte des paramètres géodésiques liés aux coordonnées qui y sont notées et conservées. Ces paramètres sont : 1) la nature du système de coordonnées utilisé, 2) le datum associé et 3) la précision des coordonnées. Si ces trois paramètres sont conservés dans les banques de données, il devient alors possible de tirer pleinement profit des nouveaux outils informatiques qui sont maintenant largement disponibles.

### **[From biogeographical databases towards biotopographical databases. New technology for new challenges]**

Recent developments in the GPS technology, Geographical Information Systems and microcomputers allow to achieve an extraordinary level of precision for the geographical localisations. A precision of the order of three meters is now available for everybody who can use a modern GPS receiver. The increased precision of these geographic coordinates allows new possibilities in terms of biogeographical studies, species ecology, populations, etc. However, biogeographical databases have to evolve in the same way. The improvement of several orders of magnitude in the precision of geographic coordinates requires a better management of them. It is now urgent to take into account geodesic parameters linked to the coordinates: 1) the nature of the coordinate system used, 2) the associated datum and 3) the precision of the coordinates. If these three parameters are saved in databases, it is then possible to use successfully the new computers tools that are now widely available.

---

Université de Mons-Hainaut, Laboratoire de Zoologie, Avenue du Champs de Mars, 6, B-7000 Mons. Pour correspondance : yvan.barbier@umh.ac.be

## INTRODUCTION

La biogéographie est la science qui étudie la répartition des êtres vivants sur la Terre. Elle a pour but l'analyse des raisons de la distribution d'un taxon à un moment et à un endroit donnés.

Depuis maintenant plus de 10 ans, la biogéographie profite pleinement de l'avènement de la micro-informatique. Ce phénomène se manifeste par la création et l'exploitation de banques de données biogéographiques. Une telle banque de données est un système informatique qui permet d'organiser, de sauvegarder et d'exploiter les données de nature biogéographique. Elle intègre les trois dimensions fondamentales de la biogéographie : les taxons, le temps et l'espace. Les objectifs qui poussent à la confection de banques de données biogéographiques sont divers : la connaissance de la répartition des espèces, la surveillance (*survey*) des espèces ou des populations, le suivi de métapopulations, les études d'incidence, etc.

L'une des principales difficultés rencontrées lors de la confection des banques de données biogéographiques est le géoréférencement des observations. La dimension spatiale des données biogéographiques est en effet primordiale puisque c'est elle qui permet de dresser les cartes de répartition. Les méthodes de géoréférencement sont diverses. La technique la plus courante est le recours à des «Gazetteers», sorte de dictionnaires qui comportent les coordonnées géographiques des localités d'un pays ou d'une région. Dans certains cas, les observateurs ont accès à des cartes géoréférencées et peuvent fournir des coordonnées plus ou moins précises de leurs observations. Le choix du système utilisé pour noter les coordonnées a fait l'objet de nombreuses discussions et controverses (Cartan 1978). Le système le plus communément adopté est l'UTM (Universal Transverse Mercator) et plus particulièrement sa notation militaire : le MGRS (Military Grid Reference System). Ce système de notation fut notamment utilisé pour le projet «European Invertebrate Survey» (EIS-CIE-EEW). Il s'agit d'un bon

choix, car l'UTM est un système standardisé et international (OTAN-NATO 1983). Il n'est pas lié à un pays en particulier et permet de traverser aisément les frontières lors des opérations de cartographie internationales.

Récemment, le Global Positioning System (GPS) est venu bouleverser très profondément les habitudes des biogéographes. À partir d'un simple récepteur satellite, il est désormais possible d'obtenir très facilement des coordonnées géographiques dont la précision est de l'ordre de 3 m. Cette précision est possible depuis la levée de la *Selective Availability* (SA) par le Congrès américain, le 1<sup>er</sup> mai 2000. Auparavant, la précision n'était que de l'ordre de 100 m (Hofmann-Wellenhof *et al.* 2001). Le GPS permet donc maintenant de localiser des observations biogéographiques de façon extrêmement précise. Cette précision était impossible à atteindre auparavant, même avec de bonnes cartes topographiques, quand celles-ci étaient disponibles.

On pourrait émettre l'objection qu'une telle précision n'est pas requise pour des organismes qui se déplacent sur de grandes distances. Si cela est vrai pour certains taxons, d'autres se déplacent très peu. On peut citer la faune du sol, les espèces sapro-xylophages, les espèces aquatiques, etc. Même les espèces ailées peuvent être très casanières. Ainsi, Rasmont *et al.* (2002) ont estimé les déplacements de trois espèces de bourdons dans les Pyrénées. Plus de 90 % des déplacements n'excèdent pas 50 m.

Les sauts récents de précision modifient considérablement les possibilités de couplage et d'analyse entre les informations fauniques et les informations topographiques existantes. Ces analyses peuvent être menées grâce aux développements récents qui ont été faits dans le domaine des Systèmes d'information géographiques (SIG). Ceux-ci permettent de croiser des informations géographiques qui sont levées à des niveaux de précision semblables. À titre d'exemple, avec une résolution de 10 km, il est impossible d'associer la distribution des observations avec celle

des courbes de niveau en région montagneuse. Même la résolution de 1 km ne permet que difficilement cette mise en relation. Il faut arriver à une précision de 100 m et même 10 m pour pouvoir établir un lien fiable entre l'altitude et la distribution observée.

L'évolution des besoins et des techniques a fait que l'échelle de travail a considérablement changé depuis les premières banques de données biogéographiques. Ce changement est important puisqu'il représente au moins deux ordres de grandeur. On peut dès lors parler de banques de données biotopographiques, car la précision des coordonnées géographiques relève maintenant de la topographie plus que de la géographie. Ce changement d'échelle a des conséquences importantes au niveau de l'organisation des banques de données. Une de ces conséquences est la gestion des coordonnées géographiques. Celle-ci ne peut plus se faire comme il y a 10 ans. L'objectif poursuivi ici est de montrer comment on peut intégrer ces nouveautés dans les banques de données modernes de manière à tirer parti de ces améliorations techniques fondamentales.

## REPRÉSENTATION CARTOGRAPHIQUE

Comme chacun le sait, la Terre n'est pas plate. Elle n'est pas non plus parfaitement sphérique puisqu'elle est légèrement aplatie aux deux pôles. Or, pratiquement toutes nos représentations cartographiques se font sur du papier ou des écrans d'ordinateur en deux dimensions. La difficulté réside donc dans le fait de transformer des coordonnées angulaires sphériques en coordonnées planes. Cette transformation est assurée par des fonctions mathématiques appelées projections. Les projections inverses permettent le passage des coordonnées planes vers les coordonnées sphériques.

Pour les représentations cartographiques à très petite échelle telles que les mappemondes, on peut considérer que la Terre est une sphère. Par contre, pour les représentations à plus grande

échelle, on ne peut plus assimiler la Terre à une sphère. Il est impossible de tenir compte de la forme réelle de la Terre en tout point. Aussi, est-on obligé de faire quelques approximations. La première de celles-ci est d'utiliser le géoïde comme modèle de la forme terrestre. Le géoïde correspond à la surface qui, en tous ses points, est orthogonale à la direction de la verticale. Elle correspond à une approximation de la forme terrestre par extension de la surface des océans. Cette forme est une surface présentant des ondulations qui varient de l'ordre de 100 m (Snyder 1987). Comme il est impossible de tenir compte de la forme réelle du géoïde pour effectuer les calculs de projection, on utilise pour cela un ellipsoïde. Un ellipsoïde est une sphère très légèrement aplatie caractérisée par un petit et un grand axe. Il existe environ 38 ellipsoïdes de par le monde (EPSG 2000). Il est cependant impossible de définir un ellipsoïde qui épouse parfaitement toute la surface du géoïde. Aussi, est-on obligé, localement, de choisir un ellipsoïde qui épouse au mieux la surface du géoïde et, si besoin, de décentrer cet ellipsoïde par rapport au centre de gravité de la terre si ce décentrage améliore encore l'adéquation entre ellipsoïde et géoïde (Fig. 1). La combinaison d'un ellipsoïde, de paramètres de translation et éventuellement de rotation par rapport au centre de gravité de la terre et d'un point d'origine (sur lequel l'ellipsoïde correspond parfaitement à la surface du géoïde) constitue ce qu'on appelle un **datum**. Il existe actuellement plus de 230 datums différents (EPSG 2000).

La combinaison d'une projection (et de ses paramètres), d'un datum et d'un système d'unités forme ce qu'on appelle un **système de coordonnées**. Un tel système nous est très familier, car c'est ce qu'on retrouve sur les cartes de nos Instituts géographiques nationaux. Le système Lambert France I est un exemple de système de coordonnées. Il est défini par la projection conique conforme de Lambert (cinq paramètres non détaillés ici), le datum NTF (Nouvelle Triangulation Française) et des unités exprimées en mètres.

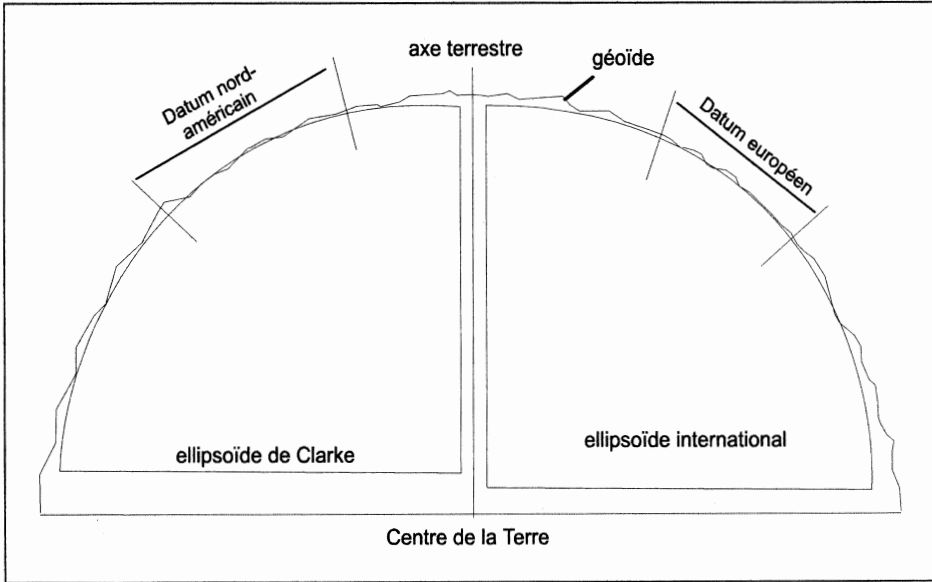


Figure 1. Les datums européen et nord-américain sont adaptés aux formes locales du géoïde.

À ce stade, il est fondamental de bien comprendre que des coordonnées, qu'elles soient sphériques ou projetées, d'un point situé à la surface terrestre seront différentes selon qu'elles sont exprimées dans un datum ou dans un autre. Les écarts observés sont généralement de l'ordre de plusieurs dizaines de mètres, mais peuvent atteindre près de 2 km. À titre d'exemple, la figure 2 illustre une même coordonnée MGRS, exprimée dans trois datums différents. On peut voir que le passage du datum ED50 à WGS84 a relativement peu d'incidence sur la précision du point. Par contre, le passage de ED50 à Amersfoort (le datum utilisé aux Pays-Bas) entraîne une perte de précision plus nette.

Elle est telle que le point initialement situé dans la grille MGRS en FS0000 est maintenant passé en FR0099. Dans une représentation avec des points de 50 ou 100 km, pour une carte européenne par exemple, le point changerait de carré!

## EXEMPLES ET RECOMMANDATIONS

L'expression des coordonnées géographiques n'est donc pas aussi triviale qu'on pourrait l'imaginer. Autrefois, quand on dressait des cartes à l'échelle d'un continent, d'un pays ou d'une ré-

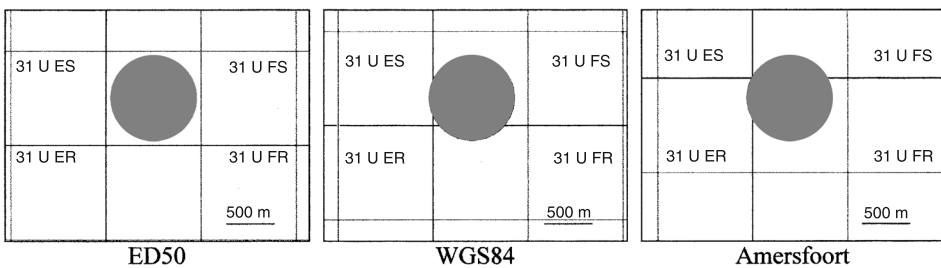


Figure 2. La même coordonnée MGRS (31UFS0000) exprimée dans 3 datums différents.

gion, on pouvait facilement faire abstraction de ces notions de géodésie. Avec la généralisation de l'usage du GPS et des SIG, il est devenu primordial de bien comprendre ces notions si l'on veut tirer pleinement parti de ces nouvelles possibilités et donc de répondre à de nouvelles questions scientifiques.

Il est donc important de faire des **recommandations** quant à la gestion des coordonnées géographiques dans les banques de données.

La première recommandation consiste à toujours noter le système de coordonnées qui est utilisé pour les observations biologiques. Si possible, celui-ci doit figurer sur les étiquettes de provenance. L'utilisateur d'un récepteur GPS doit veiller scrupuleusement à la configuration de celui-ci et plus particulièrement au datum. Par défaut, la plupart des récepteurs sont pré-réglés sur le datum WGS84. Or, en Europe par exemple, la plupart des cartes topographiques qui renseignent les coordonnées sphériques le font en European Data 1950. La différence entre les deux datums est d'environ 100 m. Cette différence n'est plus négligeable lorsqu'on réalise des études de biotopographie.

La seconde recommandation concerne les banques de données. Celles-ci devraient toujours conserver la coordonnée géographique telle qu'elle est donnée par l'observateur. C'est ce qu'on

appelle la **coordonnée originale**. Les informations à conserver en permanence sont :

1. le système de coordonnées;
2. la coordonnée elle-même;
3. le datum associé à la coordonnée originale;
4. l'ordre de grandeur de la précision (par exemple 10 m).

Les auteurs ont appliqué cette recommandation dans deux de leurs logiciels : Data Fauna-Flora et Carto Fauna-Flora (Barbier *et al.* 2000; Barbier et Rasmont 1995, 1996, 2000a, 2000b). Dans les banques de données gérées par ces logiciels, les coordonnées originales sont conservées dans trois champs qui sont en principe intouchables. Le système de coordonnées et la coordonnée sont conservés dans un champ de type texte; les trois premiers caractères servent à identifier le système de coordonnées utilisé et le reste de la chaîne sert à conserver la coordonnée proprement dite. Un autre champ sert à identifier le datum associé et un troisième champ à indiquer le niveau de précision de la coordonnée. Le tableau 1 illustre quelques exemples de coordonnées originales.

Parallèlement à ces champs intouchables, on utilise des champs de calcul qui sont générés à partir des coordonnées originales. Ces champs de calculs peuvent contenir les coordon-

**Tableau 1. Exemples de gestion des coordonnées originales.** Les 3 premiers champs sont intouchables dans la base de données, alors que les 2 derniers champs sont calculés à la demande

Coordonnée originale <sup>a</sup>	Datum	Précision (m)	Latitude/Longitude calculées en degrés décimaux dans le datum WGS84	MGRS calculé dans le datum ED50
deg 50,1666 4,566	WGS84	10	50,1666 4,566	31UFR11845832
deg 50,4333 4,73445	ED50	3	50,43246 4,73318	31UFR2309788264
mgr 31 U FR16326821	ED50	10	50,25275 4,63041	31UFR16326821
lf2 547743 546332	NTF	2	49,91115 1,60995	31UDR0019729679
nma 897223 345223	MER	1000	33,59888 -1,11604	30SXC7419

<sup>a</sup> : les codes qui servent à définir le type de système de coordonnées sont arbitraires. deg = degrés décimaux Greenwich ; mgr = military grid system ; lf2 = Lambert France zone II ; nma = Nord Maroc. D'autres codes sont définis pour les autres systèmes de coordonnées.

nées en latitude/longitude ou en X/Y d'un système quelconque. Ce système présente les avantages suivants :

1. les coordonnées originales et leur précision sont conservées en permanence. Il n'y a donc pas de risque de mauvaise interprétation ou de dégradation de la précision;
2. on peut, au sein d'une même base de données, mélanger des coordonnées de systèmes différents;
3. les coordonnées de travail peuvent être recalculées à la demande dans des systèmes variés;
4. les coordonnées de travail peuvent éventuellement être ajustées à une grille quelconque comme par exemple la grille UTM de 10 km.

Tout ceci nécessite l'utilisation d'algorithmes de conversion de coordonnées géographiques au sein des bases de données. Ceux-ci ne sont pas toujours simples à mettre au point. Il existe cependant une littérature abondante à ce sujet et des ressources Internet nombreuses. Parmi les ouvrages pertinents, citons ceux de Rasmont *et al.* 1986; Rasmont et André 1989; Snyder 1987; EPSG 2000; Iliffe 2000. Parmi les ressources internet les plus intéressantes, citons : [www.epsg.org](http://www.epsg.org), [www.nima.mil](http://www.nima.mil), [www.helsinki.fi/~rlampine/cartogr.html](http://www.helsinki.fi/~rlampine/cartogr.html).

Afin de faciliter l'utilisation de ces algorithmes dans les banques de données, le premier auteur développe actuellement une librairie de programmation dynamique (DLL) qui permet d'assurer facilement des conversions de coordonnées géographiques dans les programmes de la famille Microsoft. Une version d'évaluation est disponible sur demande.

## REMERCIEMENTS

Les développements présentés ici ont été réalisés en partie dans le cadre d'une convention financée par la Région Wallonne (Belgique). Une partie des frais a également été financée par le Conseil de recherche de l'Université de Mons-Hainaut.

## RÉFÉRENCES

- Barbier, Y. et P. Rasmont. 1995.** Carto Fauna-Flora. Cartographie des données biologiques. Cartography of biological data. Université de Mons-Hainaut, Mons, Belgique. IX+93+36 pp.
- Barbier, Y. et P. Rasmont. 1996.** Carto Fauna-Flora. Cartographie des données biologiques. Cartography of biological data. Version 1.2 et utilitaires. Université de Mons-Hainaut, Mons, Belgique. 62 pp.
- Barbier, Y. et P. Rasmont. 2000a.** Carto Fauna-Flora 2.0. Guide d'utilisation. Université de Mons-Hainaut, Mons, Belgique. 59 pp.
- Barbier, Y. et P. Rasmont. 2000b.** CFFedit 2.0. Guide d'utilisation. Université de Mons-Hainaut, Mons, Belgique. 72 pp.
- Barbier, Y., P. Rasmont, M. Dufrière et J.-M. Sibert. 2000.** Data Fauna-Flora 1.0. Guide d'utilisation. Université de Mons-Hainaut, Mons, Belgique. 106 pp.
- Cartan, M. 1978.** Inventaires et cartographies de répartitions d'espèces faune et flore. Éditions du Centre National de la Recherche Scientifique, Paris, France. 127+XIX pp.
- EPSG (European Petroleum Survey Group). 2000.** Guidance Note Number 7. POSC literature pertaining to Geographic and Projected Coordinate System Transformations. 58 pp. Voir aussi <http://www.epsg.org>
- Hofmann-Wellenhof, B., H. Lichtenegger et J. Collins. 2001.** GPS. Theory and practice. Springer-Verlag Wien, New York. 382 pp.
- Iliffe, J.C. 2000.** Datums and map projections for remote sensing, GIS and surveying. Whittles Publishing, New York. 150 pp.
- OTAN – NATO. 1983.** Accord de standardisation. Objet : Systèmes géodésiques, ellipsoïdes, quadrillages et systèmes de coordonnées rectangulaires. STANAG 2211, édition 4, Bureau militaire de standardisation (sans classification).
- Rasmont, P. et J. André. 1989.** Applications d'un logiciel de projection U.T.M. à la surveillance des invertébrés, *in* J. de Beaufort et H. Maurin (réds.), L'utilisation des inventaires informatisés d'invertébrés pour l'identification et la surveillance d'espaces de grand intérêt faunistique. Secrétariat de la Faune et de la Flore, Paris. 25 pp.
- Rasmont, P., M. Dufrière, C. Gaspar, J. Saintenoy-Simon et N. Magis (réds). 1990.** Présentation et code de déontologie de la Fédération des banques de données biogéographiques. Louvain-la-Neuve, France. 17 pp.

**Rasmont, P., S. Iserbyt et Y. Barbier. 2002.**

Les apports de la cartographie moderne.  
Journées d'étude sur la systématique en  
entomologie. Société entomologique de  
France, Paris. Mars 2002.

**Rasmont, P., M.C.D. Speight et B.E.**

**Picton. 1986.** A computer programme for  
conversion of the Irish transverse merca-  
tor projection to the Universal transverse  
mercator projection. *Ir. nat. J.* 22 : 45-50.

**Snyder, J.P. 1987.** Map projections. A

working manual. U.S. geological survey  
professional paper 1395, Washington.  
383 pp.