

LUXNAT: banque de données biogéographiques

par Marc MEYER

Introduction

Le développement de projets de cartographie biogéographique en Europe (p. ex. Atlas Florae Europaea, Cartographie des Invertébrés Européens) montrait clairement que des banques de données fiables et performantes doivent être à la base de toute représentation cartographique.

Depuis quelques années le Musée national d'histoire naturelle de Luxembourg a ainsi initié la création d'une banque de données biogéographiques sur ordinateur central avec le soutien du "Centre Informatique de l'Etat". La première version fût une variante rigide ne permettant pratiquement aucune utilisation de fichiers dictionnaires. Aujourd'hui la version centrale de LUXNAT se trouve dans l'environnement VM/CMS et correspond intégralement à une banque de données relationnelle gérée par le standard SQL (IBM SQL/DS, REXX-SQL, ISPF, etc). A côté de cette version centrale, qui contient toujours la dernière mise à jour des tables, nous avons installé un réseau local avec plusieurs stations de travail sur lesquelles on peut transférer des parties de LUXNAT identiques à la structure centrale. Ceci permet une utilisation simultanée des informations par différents utilisateurs sans perte de temps.

La structure relationnelle de LUXNAT

Dès le début des développements informatiques, le concept de LUXNAT correspondait à une banque de données biogéographiques universelle sans limitation ni dans l'espace, ni dans les groupes taxonomiques. Toutes les informations concernant les observations d'organismes vivants à n'importe quelle partie du globe terrestre peuvent être saisies. Ceci est une condition essentielle si l'on veut intégrer des collections scientifiques dans un tel programme. Il est évident que les données saisies dans LUXNAT sont validées par des spécialistes et que les fichiers actuels ne contiennent donc que des informations provenant des domaines et des régions traités par les collaborateurs scientifiques du Musée.

Une banque de données centralisant des observations biogéographiques de tout genre doit obligatoirement suivre une certaine standardisation des informations. Ceci implique l'identification préalable des taxons en cause (selon la nomenclature utilisée par LUXNAT!) et des lieux d'observation (coordonnées Gauss-Krueger, resp. UTM; localisation administrative de la localité la plus proche). Avant de pouvoir faire une saisie dans LUXNAT, ces informations doivent donc être validées par l'auteur lui-même. Actuellement des aides synonymiques ne sont pas prévues, mais elles peuvent être intégrées facilement en cas de besoin.

La flexibilité indispensable dans le cas d'une banque de données universelle est garantie par la structure relationnelle qui permet une gestion facile des fichiers et une recherche performante de toutes les variables contenues dans les tables.

Exemple:

Système traditionnel (1 ligne dans un fichier)

esp. loc. coord. écosyst. date auteur origine indiv.
 Papilio machaon L. / Luxembourg / 32UKV9987 / mesobrometum / 03081991 / MEYER Marc / obs. de terrain / 2 femelles

Système relationnel (la même ligne dans 4 tables; codes numériques remplaçant des variables en texte clair en *italiques*):

loc. coord. écosyst.
 Table 1 (= OBSLOC): *113 / 32U KV 99 87 / E110* -> lien vers table 2

date aut. orig.
 Table 2 (= OBSERV): *03 08 1991 / 81 / 1* -> lien vers table 3

esp.
 Table 3 (= OBSESP): *1256* -> lien vers table 4

indiv.
 Table 4 (= INDIVIDU): 2.

Les redondances dans les données sont évitées par la séparation de chaque information unitaire (l'observation minimale se composant du taxon, du lieu, de la date et de la personne) aux niveaux suivants:

1. le lieu d'observation (OBSLOC)
2. l'observation (OBSERV)
3. le taxon observé (OBSESP)
4. les individus/abondances observés (INDIVIDU, PHYTOSOC 1)].

Cette séparation des tables offre la possibilité à des spécialistes de certains groupes taxonomiques d'ajouter des variables spécifiques à leur groupe selon leurs besoins.

La structure générale de la banque de données LUXNAT est montrée dans le schéma relationnel de la figure 1. Les 3[4] tables mentionnées ci-dessus constituent la partie centrale de la structure arborescente. Les variables décrivant les observations standardisées sont indiquées dans le tabl. 1.

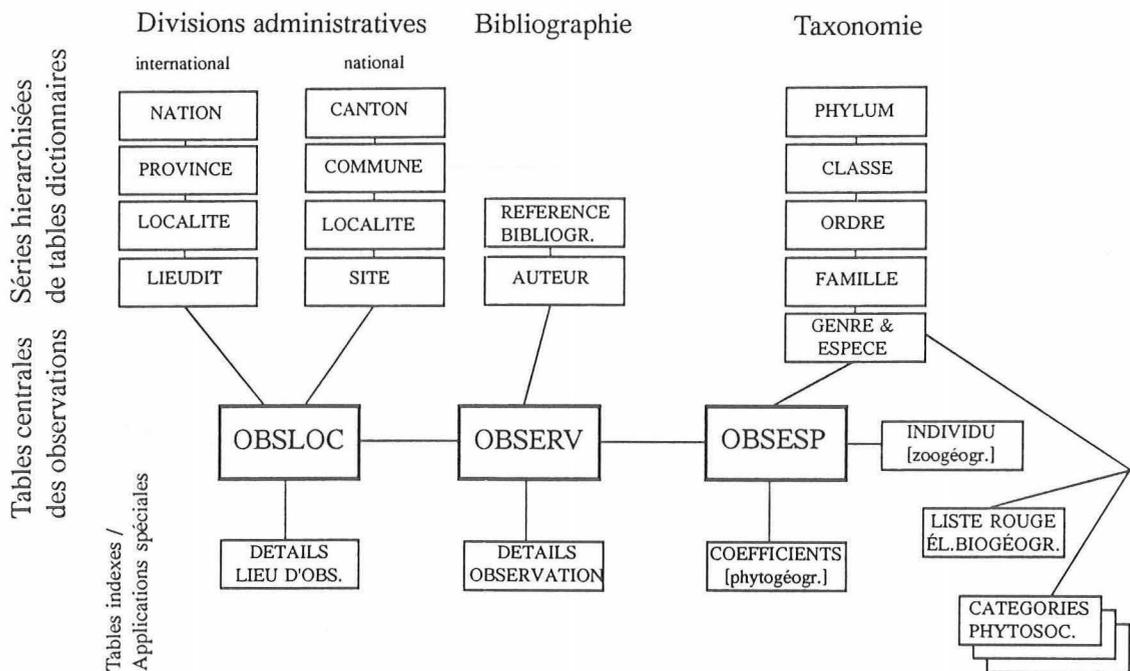


Fig. 1. Structure de LUXNAT.

Tableau 1. Tables centrales (fichiers d'observations) de LUXNAT. (Les tables et données obligatoires sont mentionnées en caractères gras)

<i>Table</i>	<i>Variables</i>	<i>Remarques</i>
OBSLOC	nation coordonnées UTM coord. Gauss-Krüger altitude spécifications conc. l'écosystème auteur réf. bibliographique date variables relationnelles	(données étrangères) (données nationales) vers tables FLOUER, OBSERV.
OBSERV	confidentialité de l'observ. date détails conc. le lieu d'observation origine des données auteur réf. bibliographique variables relationnelles	(accessible au public ou non!) (ou période) vers tables OBSLOC, OBSESP [et PHYTOSOC 1]
OBSESP	code de l'espèce variables relationnelles	vers tables OBSERV et INDIVIDU / PHYTOSOC 2.
INDIVIDU	statut et nombre d'individus variable relationnelle	(réservée aux animaux!) vers table OBSESP.
PHYTOSOC 1	surface de la parcelle inventoriée % des niveaux de stratification végétale variable relationnelle	(réservée aux plantes!) vers table OBSERV.
PHYTOSOC 2	coefficient d'abondance coefficient de dominance variable relationnelle	(réservée aux plantes!) vers table OBSESP.

Deux séries de tables hiérarchisées ("dictionnaires") sont utilisées pour permettre la production de résultats plus compréhensifs:

* *Tables taxonomiques*

Classification et nomenclature basées sur l'ouvrage de PARKER (1982), mais seulement les catégories "phylum", "classe", "ordre" et "famille" ont été retenues, la nomenclature des espèces comme unités taxonomiques fondamentales reposant cas par cas sur les propositions des spécialistes en cause:

(PHYL -> CLAS -> ORD -> FAM -> ANIMAL / PLANTE)

* *Tables des provinces/localités*

(données étrangères: NATION -> PROVINCE -> LOCALITÉ -> LIEUDESCR; données nationales: CANTON -> COMMUNE -> LOCALITÉ -> LIEUDIT).

La redondance dans ces tables "dictionnaires" est évitée par l'utilisation de codes numériques dans les tables centrales (décrivant l'observation) qui relient celles-ci aux dictionnaires dans lesquels se trouvent les définitions exactes des données en texte clair, p. ex. l'auteur (AUTEUR), la référence bibliographique (REFERENCE), le type d'écosystème (ECOSYST), le type d'habitat (HABITAT) et les espèces (ANIMAL/PLANTE).

Utilisation de standards

Le manque de standards internationaux en matière de biogéographie a provoqué des incompatibilités importantes entre les banques de données existantes.

En premier lieu il faut regretter l'absence de classification des organismes vivants généralement acceptée par les biologistes, mais une banque de données universelle ne pourra pas prendre en considération tous les différents points de vue des spécialistes. La classification universelle proposée par PARKER (1982) a donc été adoptée pour les besoins de LUXNAT, puisqu'elle présente l'avantage d'être valable pour tous les groupes taxonomiques du monde. Si une révision générale a été publiée ultérieurement pour un certain groupe d'organismes, les codes sont facilement adaptables. Pour la production de listes de sortie, seulement les catégories "phylum", "classe", "ordre" et "famille" ont été retenues.

Des changements de nomenclature ou de systématique causent un problème majeur aux gérants de banques de données, aussi certains ont prévu des places réservées à des extensions de codes numériques. Dans LUXNAT l'utilisation d'un code double évite ces problèmes.

1. Le premier code est un nombre courant interne généré par l'ordinateur et représentant la variable relationnelle assurant le lien du taxon vers les autres tables. Ce code est **inchangeable** et lié de façon permanente au taxon en question.

2. Le second code est un code taxonomique "classique" définissant numériquement la position systématique du taxon. Ce code **pourra être adapté** à tous les changements taxonomiques éventuels.

Exemple d'un code d'espèce:

<i>code famille</i>	<i>nom de l'espèce</i>	<i>code systém.</i>	<i>numéro crt. esp.</i>
4358	<i>Papilio machaon</i> LINNAEUS, 1758	10010	1660

Le Conseil de l'Europe a publié une liste des écosystèmes européens et cette liste a été adoptée par le projet des Communautés Européennes "CORINE-Biotopes" (CORINE, 1986). La même liste est utilisée par LUXNAT après quelques adaptations compatibles, notamment en ce qui concerne les écosystèmes anthropogènes. Le tableau 2 montre un extrait de cette liste.

Tableau 2. Extrait de la liste standard des écosystèmes européens
(version française modifiée du standard CORINE)

A400	Forêts des sites humides
A410	Forêts galeries alluviales (<i>Salicetalia purpureae</i>)
A411	Salicion triandro-viminalis
A412	Salicion albae
A420	Forêts alluviales (<i>Fraxino-Alnion glutinosae</i>)
A421	En zone de source (<i>Carici-remotae-Fraxinetum</i>)
A422	En zone de ruisseaux (<i>Stellario-Alnetum glutinosae</i>)
A423	En zone de rivières (<i>Pruno-Fraxinetum</i>)
A430	Forêts de marais et tourbières
A431	Aulnaies (<i>Alnetalia glutinosae</i>)
A432	Sauricaies (<i>Salicetalia auritae</i>)
A433	Boulaies (<i>Betula</i> sp.)

CORINE-Biotopes utilise une classification standardisée des districts administratifs des membres de la CE (code "NUTS"). Ce code a été adopté pour les besoins de LUXNAT, après avoir été élargi aux nations européennes non membres de la CE.

Extrait du code NUTS:

France 2000	
Est	2400
Lorraine	2410
Meurthe-et-Moselle	2411
Meuse	2412
Moselle	2413
Vosges	2414
Alsace	2420
Bas-Rhin	2421
Haut-Rhin	2412

Projets en cours d'élaboration

Le développement récent des micro-ordinateurs a permis de réaliser récemment un réseau local permettant l'installation d'extraits conformes de LUXNAT sur PC avec la possibilité de gestion on-line (stations de travail), resp. off-line (laptops utilisés sur le terrain ou ordinateurs privés des collaborateurs scientifiques du Musée).

Un lien direct entre LUXNAT et un système d'information géographique (GIS) sera développé dans les mois prochains.

Les tables dictionnaires de LUXNAT serviront à installer un système de documentation général pour l'ensemble des sections du Musée national d'histoire naturelle (Fig. 2). Ceci englobe le catalogage des collections scientifiques, mais aussi de la bibliothèque, des collections de tirés à part, de la diathèque, etc.

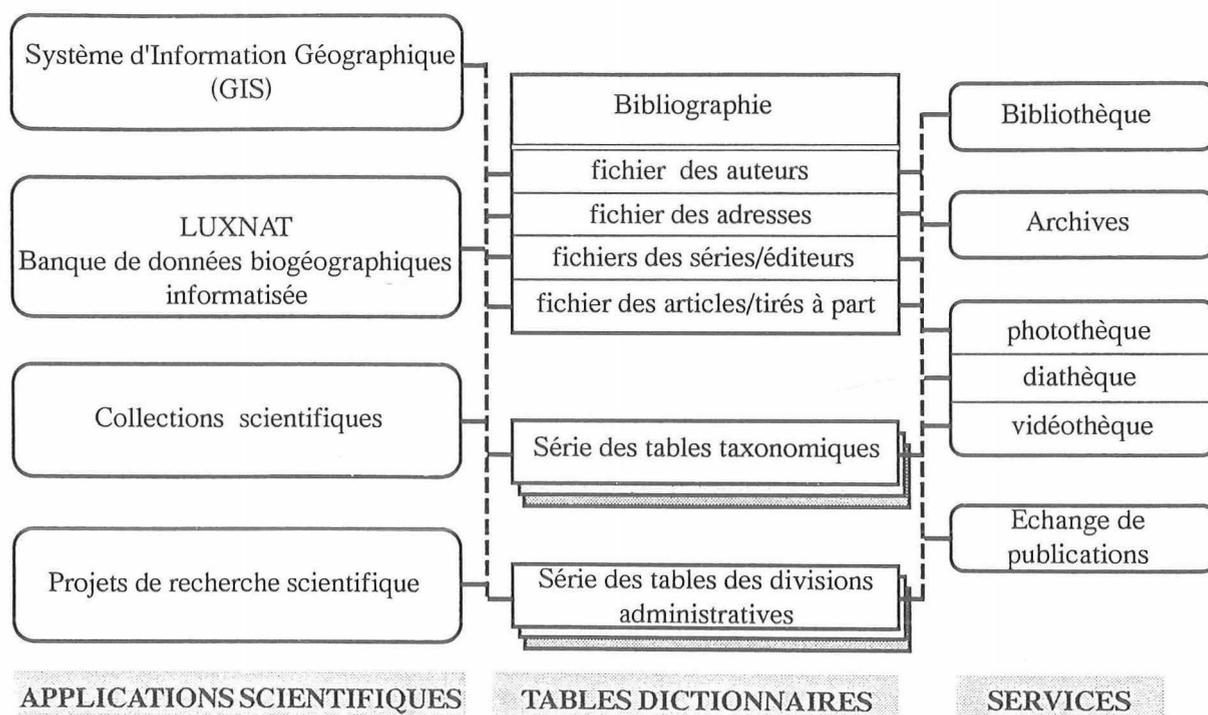


Fig. 2. Système de documentation relationnel pour le Musée National d'Histoire Naturelle de Luxembourg

Bibliographie

CORINE (WYATT, B. K. ed.), 1986. Register of sites of importance for nature conservation in Europe, Specifications for data content and format of the standard site record. CORINE / BIOTOPE / 86-2.2., Brussels.

DUFOUR, C. & GONSETH, Y., 1986. Contribution à la méthode informatique en faunistique. *Documenta Faunistica Helvetica* 1, 1-30.

MEYER, M., 1989. Actual situation of the butterfly populations in Luxembourg. Proceedings of the 1987 colloquy of EIS, Paris: Muséum National d'Histoire Naturelle.

PARKER, S.P., 1982. Synopsis and classification of living organisms. 2 vols., New York etc.: McGraw-Hill Book Company.

RASMONT, P., DUFRÊNE, M. & MEYER, M., 1989. Pour un réseau européen de banques de données fauniques. *Notes fauniques de Gembloux*, 18: 65-68.

Marc MEYER
Musée national d'histoire naturelle
Section Zoologie
24, rue Münster
L-2160 LUXEMBOURG