

# Présentation de la version 7 de Phytobase S

## Généralités

*Phytobase S 7* est une adaptation de la base de données relationnelle *Phytobase 7*, application 4<sup>ème</sup> Dimension (version 2003) développée au Laboratoire d'écologie végétale de l'Université de Neuchâtel en collaboration avec le Service Conseil Zones Alluviales. Elle s'applique à la gestion des données phytosociologiques des zones alluviales de Suisse selon la méthode synusiale intégrée, dans laquelle les relevés floristiques sont effectués séparément pour chaque strate de végétation.

*Phytobase S 7* renferme l'ensemble des relevés des zones alluviales de Suisse de la thèse de Roulier (1998), dont la typologie sert de référentiel pour le diagnostic des nouveaux relevés. L'homogénéité des syntaxons élémentaires a été améliorée en se basant sur deux mesures de similarité entre les relevés synusiaux (Jaccard en présence-absence et Steinhaus en tenant compte des recouvrements).

Les catalogues floristiques et syntaxonomiques ont été mis à jour et complétés. La nomenclature de base des espèces reste celle de la première édition de *Flora Europaea*, mais l'utilisateur a la possibilité d'afficher et d'exporter les données floristiques selon la nomenclature de *Flora Helvetica* ou de la *Banque de Données Nomenclaturale de la Flore de France* (BDNFF 2003). Le catalogue floristique intègre les indications caractéristiques et différentielles syntaxonomiques selon la dernière version du référentiel de Julve (Baseflor 2003), ainsi que les stratégies C-S-R selon la classification de Grime.

Les relevés synusiaux et phytocénotiques des différentes campagnes de suivi des zones alluviales sont saisis dans cette base de données où ils peuvent faire l'objet de différents diagnostics (diagnostic syntaxonomique, diagnostic de l'homécie, spectre biologique, spectre dynamique des essences, valeurs écologiques indicatrices de Landolt et d'Ellenberg, signature CSR, indices de biodiversité, etc.).

En particulier, les données des transects permanents sont organisées en sessions et bandes et peuvent être mises en valeur sous forme de graphiques et de tableaux.

## Diagnostic syntaxonomique des relevés

Le diagnostic automatique du groupement végétal (syntaxon élémentaire SyE pour les relevés synusiaux ou cœnotaxon élémentaire CoeE pour les relevés phytocénologiques) peut être effectué pour chacun de ces relevés, en le comparant aux relevés centroïdes des syntaxons ou des cœnotaxons élémentaires répertoriés dans la base.

La procédure de calcul de l'indice de similarité  $S_{rc}$  entre un relevé synusial ou phytocénologique  $r$  et le relevé centroïde  $c$  d'un syntaxon ou d'un cœnotaxon élémentaire utilise les données suivantes :

$r$	relevé testé (synusie ou phytocénose)
$c$	relevé centroïde (syntaxon ou coenotaxon élémentaire)
$i$	élément constitutif du relevé ou/et du relevé centroïde (espèce pour le diagnostic d'un relevé synusial ou syntaxon élémentaire pour celui d'un relevé phytocénologique)
$R_{ir}$	recouvrement de l'élément $i$ dans le relevé $r$
$R_{ic}$	recouvrement moyen de l'élément $i$ dans le relevé centroïde $c$
$F_{ic}$	fréquence relative de l'élément $i$ dans le relevé centroïde $c$
$A_{ic}$	fréquence absolue de l'élément $i$ dans le relevé centroïde $c$
$n_r$	nombre d'éléments du relevé $r$
$n_c$	nombre d'éléments du relevé centroïde $c$
$n_{rc}$	nombre d'éléments communs au relevé $r$ et au relevé centroïde $c$

## Calcul des recouvrements

Le recouvrement ( $R_{ir}$ ) d'une espèce (ou respectivement d'un SyE)  $i$  dans un relevé synusial (ou respectivement d'un relevé phytocénologique)  $r$  est déduit du code de dominance de Braun-Blanquet selon le tableau suivant :

Code de dominance	Recouvrement	Recouvrement minimum	Recouvrement maximum
r	0.1%	0.01%	0.1%
+	0.5%	0.1%	1%
1	3%	1%	5%
2	14%	5%	25%
3	32%	25%	50%
4	57%	50%	75%
5	90%	75%	100%

Chaque relevé synusial est transformé de manière à ce que la somme des recouvrements de toutes les espèces soit égal à 100% (dominance et recouvrement relatifs).

Dans le relevé centroïde  $c$  du syntaxon élémentaire, le *recouvrement moyen*  $R_{ic}$  de l'espèce  $i$  est calculé sur le sous-ensemble des relevés où elle est présente ; cette valeur est donc indépendante de la fréquence.

## Les mesures de fidélité

L'*indice de caractérisation*  $IC_{ij}$  mesure la fidélité relative d'une espèce  $i$  à un syntaxon élémentaire  $j$  (SyE) par rapport à l'ensemble des SyE contenus dans la base de données (entre 0 et 1) :

$$IC_{ij} = \frac{1}{2} \left( \frac{F_{ij}}{\sum_{k=1}^N F_{ik}} + \frac{A_{ij}}{\sum_{k=1}^N A_{ik}} \right)$$

où  $N$  est le nombre total de SyE dans lesquels l'espèce  $i$  est présente. Cette formule permet d'éviter de donner trop de poids aux SyE particulièrement riches (fréquences absolues élevées) ou pauvres (fréquences relatives élevées) en relevés. En outre, les indices de caractérisation de chaque espèce sont automatiquement ajustés de manière à ce que leur valeur maximale soit égale à 1, ceci pour éviter qu'ils prennent des valeurs trop faibles quand le nombre de SyE est important.

L'*indice de fidélité*  $U$  est une autre mesure de la fidélité d'une espèce à un groupe de relevés (SyE mais aussi alliance, ordre, classe ou homécie), calculée selon la méthode décrite par Bruelheide (2000), dérivée de l'approximation d'une distribution binomiale ou hypergéométrique par la distribution normale. Cette valeur  $U$  reflète les différences de fréquence relative de l'espèce entre le groupe et l'ensemble des relevés utilisés pour la définition des groupes. Elle tient compte aussi de la différence de taille des groupes et augmente avec le volume de la base. La fidélité  $U$  est nulle quand l'espèce est aussi fréquente dans le groupe que dans l'ensemble des groupes, positive quand cette fréquence est supérieure et négative quand elle est inférieure.

Sachant qu'un groupe de relevés  $p$  est formé de  $Z_p$  relevés dans un total de  $Z$  relevés de référence et que la fréquence absolue d'une espèce  $i$  dans l'ensemble des groupes vaut  $A_i$ , la fréquence absolue théorique de cette espèce dans le groupe  $p$  est  $\mu = (Z_p / Z) A_i$ . Pour calculer l'indice de fidélité, on

compare cette valeur théorique avec la fréquence absolue observée  $A_{ip}$ .

On a donc :

$Z$  = nombre total de relevés de référence dans la base de données  
 $Z_p$  = nombre de relevés de référence dans le groupe  $p$   
 $A_i$  = fréquence absolue de l'espèce  $i$  dans les  $Z$  relevés de la base  
 $A_{ip}$  = fréquence absolue de l'espèce  $i$  dans les  $Z_p$  relevés du groupe  $p$   
 $P = Z_p / Z$  = proportion de relevés dans le groupe  $p$   
 $\mu = A_i P$  = fréquence absolue théorique de l'espèce  $i$  dans le groupe  $p$

La fidélité  $U_{ip}$  de l'espèce  $i$  au groupe  $p$  se calcule ainsi :

$$U_{ip} = s \frac{|A_{ip} - \mu| - 0.5}{\sqrt{A_i P (1 - P)}} = s \frac{\left| A_{ip} - A_i \frac{Z_p}{Z} \right| - 0.5}{\sqrt{A_i \frac{Z_p}{Z} \left( 1 - \frac{Z_p}{Z} \right)}}$$

$s = 1$  si  $A_{ip} > \mu + 0.5$

$s = 0$  si  $\mu - 0.5 \leq A_{ip} \leq \mu + 0.5$

$s = -1$  si  $A_{ip} < \mu - 0.5$

### Diagnostic syntaxonomique d'un relevé synusial

L'attribution ou le rattachement d'un relevé synusial à un SyE se base sur un **diagnostic syntaxonomique automatique**, qui se fait normalement en deux phases :

1. En cliquant d'abord sur le bouton *CL OR AL*, on déclenche la méthode qui aboutit à l'affichage du spectre syntaxonomique du relevé (contributions des espèces caractéristiques des différentes classes phytosociologiques, pondérées ou non par le recouvrement, représentés dans le tableau de gauche) et au diagnostic de la classe, de l'ordre et de l'alliance. Les informations utilisées dans cette méthode concernant l'appartenance phytosociologique des espèces sont inspirées de la classification synusiale de JULVE *et al.* (1997), en tenant compte partiellement de la fidélité des espèces aux syntaxons de *Phytobase* et sont contenues dans la table [IDIOTAXON]. Le bouton *CATMINAT* déclenche aussi un diagnostic de la classe, de l'ordre et de l'alliance (ainsi que de leurs subdivisions éventuelles : sous-classe, sous-ordre et sous-alliance), mais basé ici sur le nouveau référentiel phytosociologique CATMINAT de JULVE (Baseveg 2003), en tenant compte des espèces

caractéristiques transgressives et des espèces différentielles. Le bouton *Diagnostic fidélité* permet de faire un diagnostic de la classe, de l'ordre et de l'alliance basé uniquement sur la fidélité des espèces aux syntaxons représentés dans la base, sans tenir compte des indications contenues dans la table [IDIOTAXON]. Le spectre des classes et des alliances apparaît dans la partie inférieure du cadre gauche. Pour chaque diagnostic, le spectre des classes peut être visualisé sous forme d'un histogramme.

2. En cliquant ensuite sur le bouton *Diagnostic SyE*, dans le cadre droit du formulaire, on déclenche une méthode qui compare le relevé synusial aux relevés centroïdes des syntaxons élémentaires de la même catégorie synusiale (A, B, H ou M) stockés dans la base. Un dialogue permet de limiter la comparaison aux syntaxons appartenant aux classes les mieux représentées, en fonction du résultat du diagnostic précédent. Plusieurs indices de similarité sont calculés en fonction de la proportion d'espèces partagées avec le relevé centroïde du SyE, selon différentes options de pondération. Dans tous les cas, l'indice de similarité varie entre 0 (aucun élément commun) et 1 (ressemblance maximum). Le nom du syntaxon élémentaire le plus proche (en tête de liste) apparaît alors à l'écran.

Dans une première série de diagnostics, les espèces du relevé sont pondérées par leur recouvrement ; les espèces du relevé centroïde sont pondérées par le produit de leur recouvrement moyen et de leur fréquence relative (diagnostic **R-F**), par le produit de leur recouvrement moyen et de leur indice de caractérisation (diagnostic **R-C**) ou par le produit de leur recouvrement moyen et de leur indice de fidélité (diagnostic **R-U**). Le détail du calcul pour le diagnostic **R-F** est le suivant :

$$S_{rc} = \frac{2 \sum_{i=1}^{N_{rc}} \min(R_{ir}, R_{ic} \cdot F_{ic})}{\sum_{i=1}^{N_r} R_{ir} + \sum_{i=1}^{N_c} R_{ic} \cdot F_{ic}}$$

Pour le diagnostic **R-U**, l'indice de fidélité est divisé par la fidélité maximale observée pour la catégorie de syntaxon considérée (A, B, H ou M), de manière à ce que la fidélité relative obtenue soit comprise entre -1 et 1, et donc que l'indice de similarité calculé ne soit pas supérieur à 1 :

$$S_{rc} = \frac{2 \sum_{i=1}^{N_{rc}} \min \left( R_{ic}, R_{ic} \frac{U_{ic}}{U_{\max}} \right)}{\sum_{i=1}^{N_r} R_{ir} + \sum_{i=1}^{N_c} R_{ic} \frac{U_{ic}}{U_{\max}}}$$

Dans une deuxième série de diagnostics, les espèces du relevé sont traitées en présence-absence ; les espèces du relevé centroïde sont pondérées par leur fréquence relative (diagnostic **F**), par leur indice de caractérisation (diagnostic **C**) ou par leur indice de fidélité (diagnostic **U**). Le détail du calcul pour le diagnostic **F** est le suivant :

$$S_{rc} = \frac{2 \sum_{i=1}^{N_{rc}} F_{ic}}{N_r + \sum_{i=1}^{N_c} F_{ic}}$$

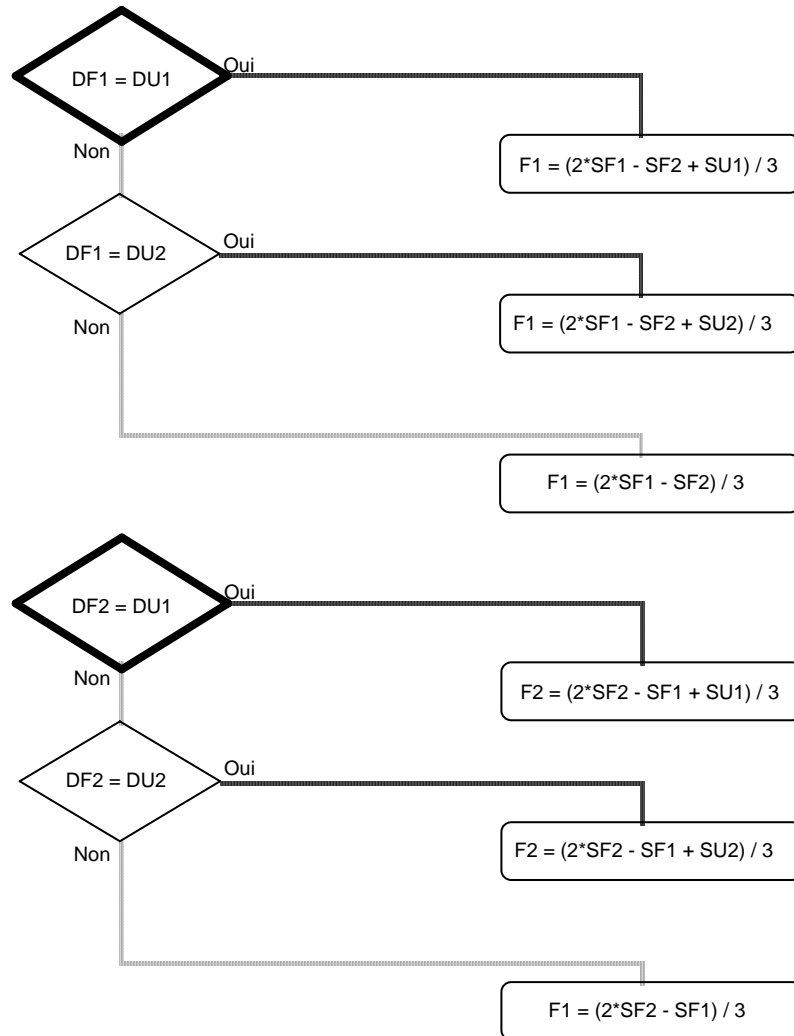
Pour le diagnostic **U**, l'indice de fidélité est divisé par la fidélité maximale de manière à ce que l'indice de similarité calculé ne soit pas supérieur à 1 :

$$S_{rc} = \frac{2 \sum_{i=1}^{N_{rc}} \frac{U_{ic}}{U_{\max}}}{N_r + \sum_{i=1}^{N_c} \frac{U_{ic}}{U_{\max}}}$$

Un septième indice probabiliste de similarité (diagnostic **P**) correspond à la moyenne des rapports entre la fréquence absolue de chaque espèce dans le SyE et sa fréquence absolue dans tous les SyE de la même catégorie (A, B, H ou M).

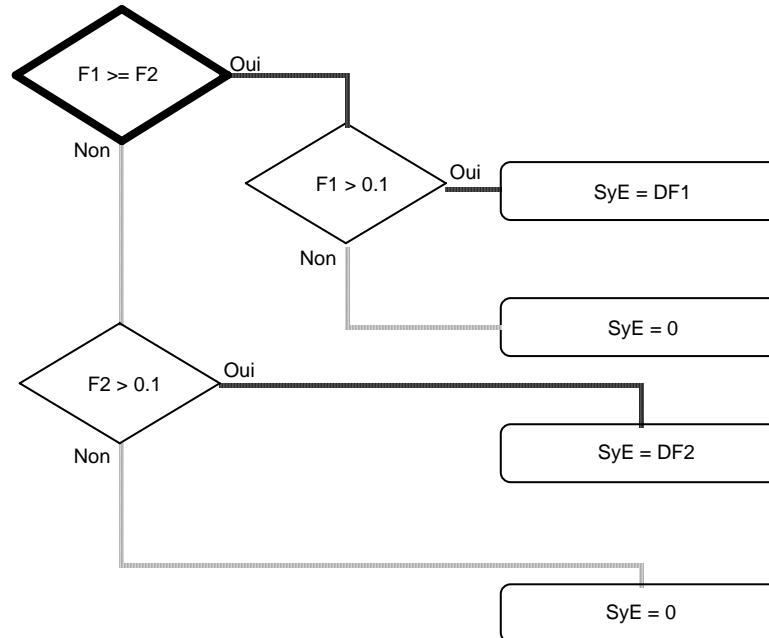
La moyenne des six premières similarités est utilisée pour effectuer le **diagnostic global**. Le résultat du diagnostic s'affiche sous forme d'une liste des groupements végétaux classés dans l'ordre décroissant de leur moyenne des similarités avec le relevé testé. Le premier groupement de la liste est retenu comme étant le plus proche du relevé testé et son nom s'affiche dans la fenêtre du diagnostic. La *fiabilité* du diagnostic global dépend de la différence entre les deux similarités moyennes les plus élevées ainsi que de la valeur de la première. Si la différence entre les deux similarités moyennes les plus élevées dépasse le quart de la valeur de la première et si celle-ci est supérieure à 0.2, le relevé peut être *rattaché* au premier SyE ; sinon le diagnostic global est jugé insuffisamment fiable. Si de surcroît la première similarité est supérieure à 0.5, on considère que le relevé pourrait être *attribué* au premier SyE.

Outre le diagnostic global du SyE qui est utilisé pour le rattachement automatique des nouveaux relevés, **Phytobase S** propose une autre aide à la décision : le **diagnostic combiné**. Celui-ci privilégie d'abord le diagnostic **R-F** – qui prend en compte le recouvrement et la fréquence relative – et secondairement le diagnostic **R-U**, basé sur le recouvrement et la fidélité. Des coefficients de fiabilité F1 et F2 sont calculés pour les deux premiers SyE obtenus à partir du diagnostic **R-F**. La procédure est représentée dans les organigrammes suivants :



- DF1 Premier SyE selon le diagnostic R-F
- DF2 Deuxième SyE selon le diagnostic R-F
- DU1 Premier SyE selon le diagnostic R-U
- DU2 Deuxième SyE selon le diagnostic R-U
- SF1 Indice de similarité du premier SyE selon le diagnostic R-F
- SF2 Indice de similarité du deuxième SyE selon le diagnostic R-F
- SU1 Indice de similarité du premier SyE selon le diagnostic R-U
- SU2 Indice de similarité du deuxième SyE selon le diagnostic R-U
- F1 Coefficient de fiabilité du premier SyE selon le diagnostic R-F
- F2 Coefficient de fiabilité du deuxième SyE selon le diagnostic R-F

Parmi les deux premiers SyE diagnostiqués par le diagnostic **R-F**, on retient celui qui obtient le coefficient de fiabilité le plus élevé, sauf si celui-ci ne dépasse pas la valeur limite de 0.1. Dans ce cas, on considère qu'il n'existe pas dans la base de données un SyE auquel on peut rattacher de façon fiable le relevé. La procédure est la suivante :



La mise à jour des diagnostics peut se faire globalement pour une sélection de relevés synusiaux, avec ou sans exportation d'un tableau de résultats détaillés.

## Diagnostic phyto-écologique des synusies

Les **valeurs écologiques indicatrices** de Landolt (L lumière, T température, K continentalité, F humidité, R acidité, N richesse en nutriments, D aération du sol et H humus, entre 1 et 5, valables pour la Suisse) et d'Ellenberg (L, T, K, F, R et N, entre 1 et 9, valables pour l'Europe centrale) sont calculées pour chaque relevé synusial en pondérant les espèces par leur recouvrement. Les espèces dont la valeur écologique indicatrice est 0 (inconnue ou à large amplitude) ne sont pas prises en compte dans la moyenne pondérée.

Les **indices de biodiversité** sont dérivés du calcul de l'entropie généralisée (Rényi 1961, Hill 1973, Legendre & Legendre 1998) :

$$H_a = \frac{\log \sum_{i=1}^q p_i^a}{1-a}$$



$a$  est un nombre entier positif ou nul (généralement compris entre 0 et 2)

$q$  est le nombre d'espèces

$p_i$  est la fréquence relative ou l'abondance relative de l'espèce  $i$

Pour les relevés synusiaux, le calcul de  $p_i$  se base sur le recouvrement  $R_i$  de chaque espèce, déduit du code de dominance :

$$p_i = \frac{R_i}{\sum_{i=1}^q R_i}$$

La mesure d'entropie  $H$  est convertie en mesure de diversité  $N$  (nombre positif ou nul, sans borne supérieure) :

$$N_a = \exp H_a$$

Les trois premiers ordres correspondent aux principaux indices de diversité utilisés en écologie :

Ordre	Entropie	Diversité	Nom
a = 0	$H_0 = \log q$	$N_0 = q$	Richesse spécifique ( <i>species richness</i> )
a = 1	$H_1 = -\sum_{i=1}^q p_i \cdot \log p_i$	$N_1 = \exp H_1$	Entropie ou diversité de Shannon
a = 2	$H_2 = -\log \sum_{i=1}^q p_i^2$	$N_2 = \frac{1}{\sum_{i=1}^q p_i^2}$	Entropie ou diversité de Simpson

Plus l'ordre  $a$  de l'indice est élevé, plus le poids relatif des espèces peu recouvrantes est faible. Cette composante de la diversité est mesurée par l'indice de *régularité* (*evenness*), calculé comme le rapport de deux mesures de diversité, par exemple :

$$E = \frac{N_2}{N_0}$$

Ainsi calculée, la régularité varie entre 0 et 1 (toutes les espèces présentent le même recouvrement), ne dépend que des différences de recouvrement et n'est pas corrélée avec la composante richesse spécifique.

Les **homécies** sont des catégories dynamiques et écologiques de synusie correspondant aux « groupes comportementaux » définis par Roulier (1997) :

Catégorie	Code	Couleur	Nom
A	Ab	Bleu	Synusies arborescentes pionnières
	Ar	Rouge	Synusies arborescentes climaciques
	Av	Vert	Synusies arborescentes post-pionnières
B	Bb	Bleu	Synusies arbustives pionnières
	Bm	Marron	Synusies arbustives palustres
	Br	Rouge	Synusies arbustives post-pionnières à climaciques, mésophiles
	Bv	Vert	Synusies arbustives post-pionnières, mésohygrophiles
H	Hb	Bleu	Synusies herbacées pionnières
	Hm	Marron	Synusies herbacées palustres
	Ho	Orange	Synusies herbacées xéroclines, hélioclines
	Hr	Rouge	Synusies herbacées climaciques, mésophiles, sciaphiles
	Hv	Vert	Synusies herbacées post-pionnières, hygrophiles

La fidélité  $U$  des espèces aux homécies est calculée d'après la méthode de Bruelheide (voir plus haut) pour chaque catégorie de synusie (A, B ou H) en se basant sur les relevés centroïdes des SyE, chacun étant ayant été rattaché à une homécie.

Le diagnostic de l'homécie est désormais calculé directement pour chaque relevé synusial, en se basant sur le recouvrement et la fidélité des espèces. Il coïncide dans la plupart des cas avec le diagnostic indirect passant par le rattachement un SyE (seule méthode de diagnostic utilisée dans les versions précédentes de Phytobase S).

### Mise en valeur des transects

La nouvelle version de Phytobase S propose une organisation hiérarchique des données issues de l'application du protocole standardisé pour le suivi de la végétation des zones alluviales de Suisse.

Les bandes qui composent un transect sont assimilées à des phytocénoses, composées elles-mêmes de synusies arborescentes (A1, a2), arbustives (B1, b2) et herbacées (H1, h2). A chaque phytocénose est assigné un numéro de transect, un numéro de bande à l'intérieur du transect (de 1 à  $x$ ) et un numéro de session (année des relevés).

Le **profil des homécies** est dessiné pour chaque session du transect en superposant les six strates de végétation. Il représente les homécies diagnostiquées directement à partir des relevés synusiaux, symbolisées par des carrés de couleur alignés selon l'ordre des bandes.

Les changements de composition floristique entre deux sessions de relevés sont calculés dans Phytobase S et représentés par un **graphique du turnover** par catégorie de synusie (A, B et H). Les relevés des sous-strates 1 et 2 sont réunis en un seul relevé par catégorie de synusie, en additionnant les recouvrements des espèces. Deux options sont disponibles pour le calcul du turnover entre deux sessions de relevés des strates de la même bande : distance de Jaccard (en présence-absence) ou distance de Bray-Curtis (tenant compte des différences de recouvrement). Le turnover moyen sur l'ensemble du transect est calculé pour chaque catégorie de synusie.

D'autres graphiques sont disponibles pour les transects : **profil de la richesse spécifique totale** (avec possibilité de superposer l'ensemble des sessions pour le même transect), **profil des recouvrements** des strates, **profil des valeurs écologiques indicatrices** de Landolt ou d'Ellenberg.

Les données servant à la construction de chacun de ces graphiques sont automatiquement exportées en format texte tabulé.

FG, 20.12.2003