



HAL
open science

Apports de l'analyse factorielle des correspondances sur variables instrumentales en typologie des stations : illustration sur la plaine de la Lanterne (Haute-Saône)

Jean-Claude Gégout, François Houllier

► To cite this version:

Jean-Claude Gégout, François Houllier. Apports de l'analyse factorielle des correspondances sur variables instrumentales en typologie des stations : illustration sur la plaine de la Lanterne (Haute-Saône). *Revue forestière française*, 1993, 45 (5), pp.539-547. 10.4267/2042/26453 . hal-03443798

HAL Id: hal-03443798

<https://hal.science/hal-03443798>

Submitted on 23 Nov 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

APPORTS DE L'ANALYSE FACTORIELLE DES CORRESPONDANCES SUR VARIABLES INSTRUMENTALES EN TYPOLOGIE DES STATIONS : ILLUSTRATION SUR LA PLAINE DE LA LANTERNE (HAUTE-SAÔNE) *

J.-Cl. GÉGOUT - F. HOULLIER

L'étude des relations entre la végétation et les facteurs de milieu est à la base de la plupart des études de typologie des stations forestières. La stratégie usuellement utilisée pour déterminer les principaux gradients responsables de la structure de la végétation, ainsi que pour établir le caractère indicateur des espèces repose principalement sur l'analyse factorielle des correspondances (AFC) et sur l'utilisation des profils écologiques (Brêthes, 1989).

Cependant ces méthodes d'analyse des données laissent subsister plusieurs problèmes :

— Les cas où les gradients exprimés par les premiers axes de l'AFC sont dus à la superposition de facteurs du milieu et de facteurs liés à la dynamique et/ou à la structure de la végétation, sont difficilement solubles (Becker et Le Goff, 1988 ; Mercier, 1988).

— La détermination du degré de pertinence biologique des variables de milieu relevées sur le terrain reste, tant pour le chargé d'étude que pour les utilisateurs finaux du catalogue, un casse-tête éludé dans la plupart des études actuelles.

— L'étude de l'autécologie des espèces par l'utilisation des profils écologiques se heurte à la multiplicité des variables initiales qui sont très souvent partiellement redondantes.

Une méthode récente d'analyse des données, nommée analyse canonique des correspondances (ACC) ou analyse factorielle des correspondances sur variables instrumentales (AFCVI) selon les auteurs (Ter Braak, 1986, 1987 ; Lebreton *et al.*, 1988a, 1988b ; Mercier, 1988) permet de résoudre de façon satisfaisante ces différents problèmes.

Notre objectif est d'illustrer sur un exemple simple, mais réel, le principe et l'intérêt de l'AFCVI pour la typologie des stations.

* Le catalogue des stations dont sont issues les données utilisées dans cet article a été financé par la Région Franche-Comté et le Fonds forestier national. Ce financement a été acquis par l'intermédiaire de la Direction régionale de l'Agriculture et de la Forêt (DRAF) de Franche-Comté (Service régional de la Forêt et du Bois).

MATÉRIEL ET MÉTHODES

La région

La Plaine de la Lanterne est située en Franche-Comté dans le département de la Haute-Saône entre les villes de Luxeuil et de Lure. Elle a fait l'objet d'une étude phytocéologique en vue de l'élaboration d'un catalogue des stations (Gégout, 1992).

Les conditions climatiques de la région sont homogènes avec une température moyenne annuelle de 9,3°C et une pluviométrie égale à 960 mm. Les formations géologiques sont constituées d'alluvions fluviales et fluvioglaciales de nature siliceuse. Des limons de faible épaisseur (30 à 70 cm) recouvrent fréquemment ces alluvions. Ce contexte géologique est à l'origine des formes très douces du paysage ; les pentes, en effet, ne dépassent guère 5 %.

Les données

Cent-six relevés phytocéologiques ont été réalisés en milieu forestier.

À partir des relevés floristiques en présence/absence, nous avons constitué un tableau phytosociologique, F (figure 1, ci-dessous), formé de n lignes, les relevés (n = 106), et de p colonnes, les espèces. Nous nous sommes limités aux 85 espèces présentes dans plus de deux relevés (p = 85).

Parmi les variables écologiques observées, nous avons retenu :

- le pH ;
- les formes d'humus (6 modalités) : horizon tourbeux, moder dysmoder, mull moder, mull acide, mull mésotrophe, mull eutrophe ;

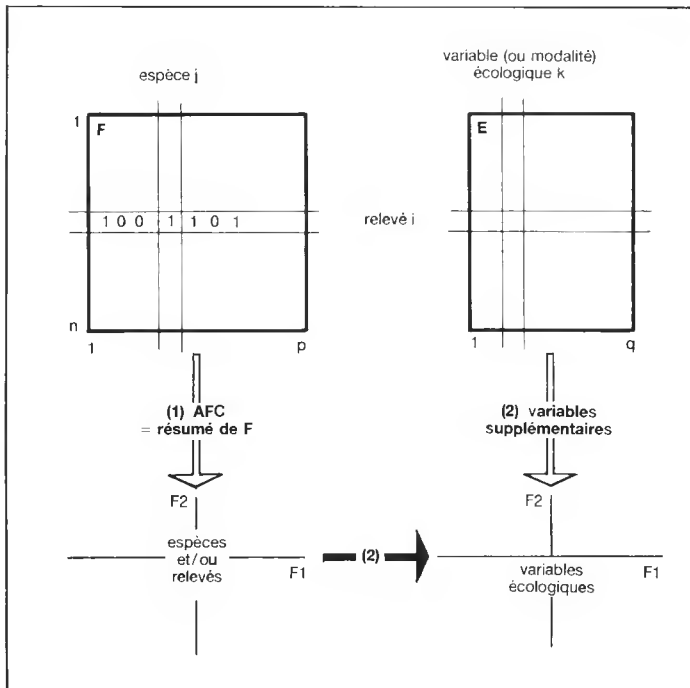


Figure 1
PRINCIPE DE L'ÉTUDE DES RELATIONS MILIEU-VÉGÉTATION lorsqu'on effectue une AFC du tableau floristique F avec les variables écologiques (E) en variables supplémentaires.

Les variables écologiques sont introduites *a posteriori* en vue d'une interprétation indirecte des résultats.

— l'hydromorphie (5 modalités) : hydromorphie absente, traces d'hydromorphie à plus de 50 cm de la surface, traces d'hydromorphie à moins de 50 cm sans décoloration forte à 20 cm (chroma > 2), traces d'hydromorphie à moins de 50 cm avec une décoloration forte (chroma < 3) à 20 cm, sol tourbeux.

Ces variables, quantitative et qualitatives, ont permis de générer un tableau de données écologiques E à n lignes et q colonnes correspondant aux modalités des variables écologiques ($q = 1 + 6 + 5 = 12$; pour respectivement le pH, les formes d'humus, et l'hydromorphie).

Calculs effectués

Le tableau F a d'abord été analysé par une AFC. Nous avons ensuite procédé à l'AFCVI des tableaux F et E. Ces calculs ont été effectués sur un microordinateur PC compatible avec le logiciel BIOMEKO v4.0 (développé au Centre d'Étude fonctionnelle et évolutive Louis Emberger à Montpellier). D'autres logiciels existent pour réaliser ces analyses, notamment ADE (développé sur MacIntosh à l'Université Claude Bernard - Lyon I) et CANOCO (développé par Ter Braak à Wageningen).

AFC DU TABLEAU PHYTOSOCIOLOGIQUE : PRINCIPE ET RÉSULTATS

L'AFC fournit un moyen de résumer le tableau F, en extrayant des axes factoriels synthétiques qui rendent compte des principaux gradients écologiques sous-jacents à la répartition de la végétation et le long desquels il est possible de positionner les espèces ou, symétriquement, les relevés [pour un exposé mathématique, se reporter, par exemple, à Saporta (1990)]. Ces axes factoriels, qui constituent le meilleur résumé possible du tableau F, sont notamment utilisés pour générer des graphiques appelés plans factoriels (figure 2, ci-dessous).

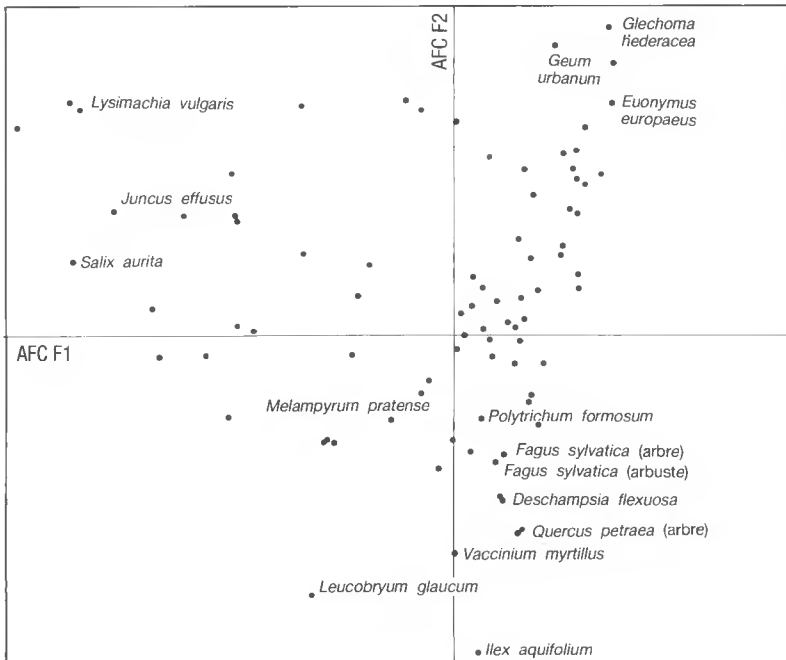


Figure 2
**PLAN FACTORIEL F1-F2
 DE L'AFC DU TABLEAU
 FLORISTIQUE F :
 REPRÉSENTATION DES
 ESPÈCES**

Les espèces représentées sont celles qui font l'objet d'un commentaire dans le texte.

L'interprétation de ces axes et plans factoriels repose en particulier sur le taux d'inertie associé à chaque axe, sur la représentation des variables écologiques comme variables supplémentaires et sur la connaissance préalable de l'autécologie des espèces.

● *Taux d'inertie*

L'inertie du tableau F mesure l'hétérogénéité des relevés par rapport à un relevé moyen. (Mathématiquement, elle est définie à partir d'une distance entre relevés, la distance du Chi-2, et du poids affecté à chaque relevé, en l'occurrence le nombre d'espèces qu'il contient).

Le taux d'inertie associé à un axe factoriel mesure la part prise par cet axe dans l'inertie totale de F : c'est une mesure de son efficacité à rendre compte de la structure du tableau F. Les axes factoriels sont ordonnés selon leur taux d'inertie décroissant dont la somme vaut 100 %.

Dans notre exemple, l'inertie totale du nuage est égale à 4,98. Les inerties absorbées par les quatre premiers axes sont égales respectivement à 0,45, 0,37, 0,26 et 0,22. Il en résulte des taux d'inertie qui sont respectivement égaux à 8,9, 7,5, 5,2 et 4,4 %. Seul le plan déterminé par les deux premiers axes semble interprétable.

● *Examen des variables écologiques supplémentaires* (figure 3, p. 543)

Les variables écologiques n'ont pas directement servi à l'élaboration du résumé du tableau F par l'AFC. Elles sont utilisées, *a posteriori*, comme variables supplémentaires : elles sont projetées sur les plans factoriels pour fournir des indications sur la signification écologique des axes factoriels.

Pour l'axe 1, l'examen des modalités des variables écologiques dans le plan factoriel des relevés montre l'importance de l'excès d'eau des sols. En effet, cet axe semble bien corrélé avec la profondeur d'apparition des premières traces d'hydromorphie, au lessivage de l'horizon A₂ et à l'apparition de la tourbe.

Pour l'axe 2, les mulls eutrophes et les pH supérieurs à 5 sont bien individualisés et bien corrélés au pôle positif, les pH inférieurs à 4 semblent plutôt liés au pôle négatif, les autres valeurs de pH et formes d'humus sont très dispersées.

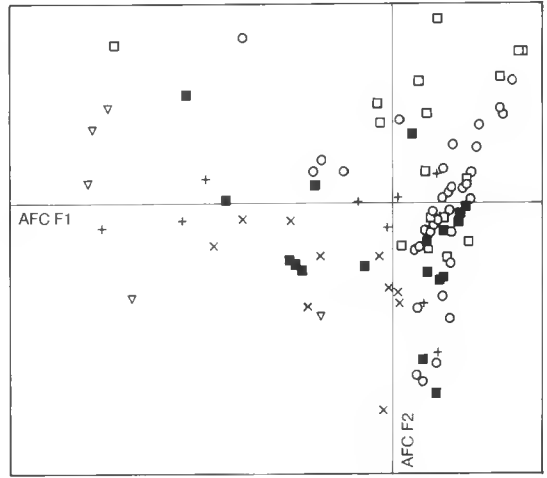
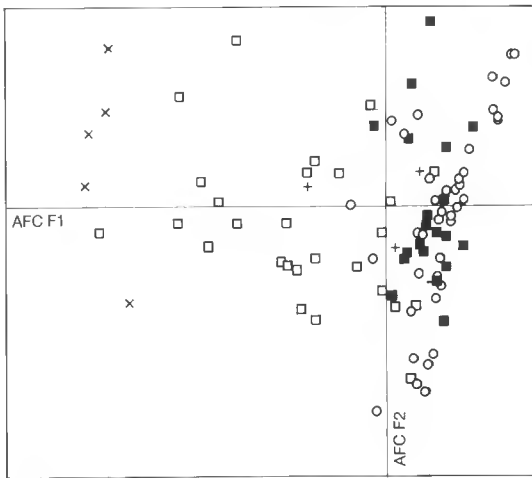
● *Examen des espèces* (figure 2, p. 541)

La représentation des espèces dans le plan factoriel 1-2 montre pour l'axe 1 un gradient très net allant des espèces de milieu drainé vers les espèces hygrophiles (par exemple, *Salix aurita*, *Juncus effusus*, *Lysimachia vulgaris*).

Le pôle positif de l'axe 2 est manifestement occupé par des espèces aux exigences trophiques élevées (par exemple, *Glechoma hederacea*, *Geum urbanum*, *Euonymus europaeus*,...). Par contre, le pôle négatif est occupé, d'une part, par des espèces acidiphiles (*Vaccinium myrtillus*, *Leucobryum glaucum*, *Deschampsia flexuosa*), d'autre part, par des espèces plus ou moins indifférentes vis-à-vis du niveau trophique (*Ilex aquifolium*, *Quercus petraea*).

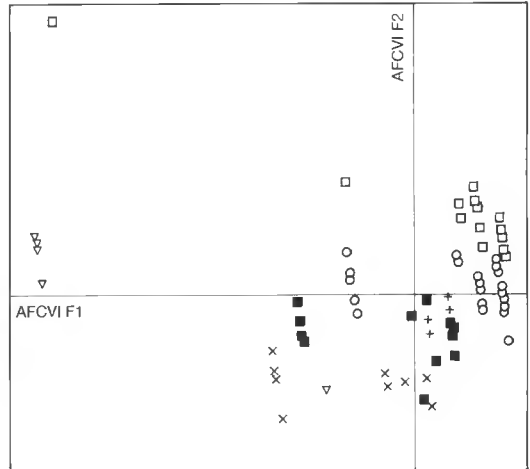
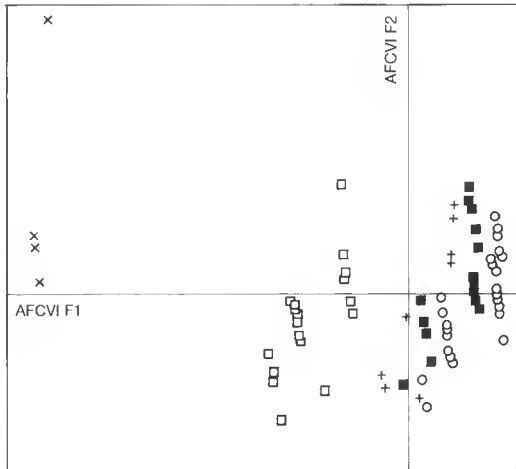
Une observation plus détaillée des résultats montre que six espèces contribuent fortement au pôle négatif de l'axe 2 : il s'agit de deux espèces acidiphiles (*Vaccinium myrtillus*, *Polytrichum formosum*) et de quatre ligneux (*Ilex aquifolium*, *Fagus sylvatica* aux stades adulte et arbustif, et *Quercus petraea* au stade adulte). Cette importance des espèces ligneuses, qui jouent un grand rôle dans la dynamique des peuplements, conduit à un diagnostic double pour la signification de cet axe : acidité et maturation du peuplement.

Figure 3 COMPARAISON DE L' AFC (EN HAUT) ET DE L' AFCVI (EN BAS) :
 représentation des modalités des variables écologiques dans le plan F1-F2 des relevés :
 classes d'hydromorphie (à gauche) et types d'humus (à droite).



- Hydromorphie absente
- Hydromorphie > 50 cm
- + Hydromorphie < 50 cm sans A₂ décoloré
- Hydromorphie < 50 cm avec A₂ décoloré
- × Sol organique humide (sol tourbeux, gley à anmoor)

- ▽ Horizon tourbeux
- × Moder, dysmoder
- + Mull moder
- Mull acide
- Mull mésotrophe
- Mull eutrophe



● Conclusion partielle

Dans la Plaine de la Lanterne, le gradient hydrique exprimé par l'axe 1 est tout à fait intéressant pour la typologie. En revanche, l'axe 2 exprime un facteur invariant intéressant, le niveau trophique du sol, et un facteur non pérenne, le degré de maturité du peuplement. Pour élaborer la typologie, il convient d'éliminer l'effet de ce dernier facteur.

PRINCIPE ET APPORTS DE L'AFCVI

Là aussi, nous nous contentons d'esquisser le principe de la méthode en renvoyant le lecteur à des références plus techniques pour une présentation détaillée (Ter Braak, 1986, 1987 ; Lebreton *et al.*, 1988a, 1988b ; Chessel et Mercier, 1992).

Comme précédemment en AFC, il s'agit d'obtenir le meilleur résumé possible du tableau F, mais ce résumé est contraint à être directement dépendant des variables écologiques du tableau E (on parle d'ailleurs parfois « d'analyse factorielle sous contraintes »). Le principe de l'AFCVI consiste donc (figure 4, ci-dessous) :

- à traiter simultanément les deux tableaux couplés F et E (le couplage est assuré par les relevés) ;
- à définir, comme en AFC, l'inertie du tableau F des relevés phytosociologiques (on utilise la même distance et la même pondération) ;
- à contraindre les axes factoriels à être des combinaisons linéaires des variables écologiques exprimées dans le tableau E : ces variables sont dites « instrumentales ».

Le résumé obtenu par une AFCVI est à la fois plus directement interprétable et moins bon en terme d'inertie. Remarquons au passage que la symétrie de l'AFC entre relevés et espèces est brisée et que le nombre maximal d'axes factoriels obtenus par une AFCVI est déterminé par le nombre des modalités des variables écologiques du tableau E.

Les aides à l'interprétation usuelles en AFC restent disponibles. Il est notamment possible de construire des plans factoriels et d'y représenter simultanément, avec certaines précautions, les relevés, les espèces et les variables du milieu.

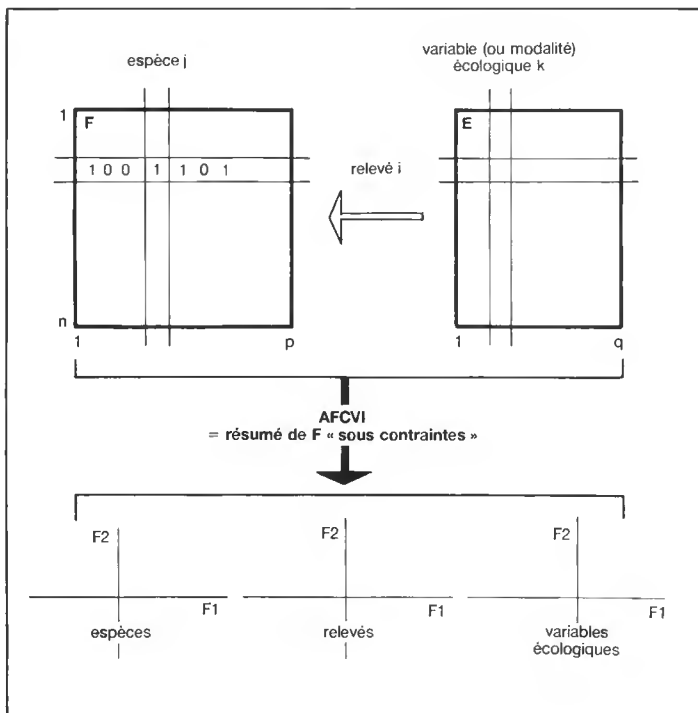


Figure 4
PRINCIPE DE L'ÉTUDE DES RELATIONS MILIEU-VÉGÉTATION lorsqu'on effectue une AFCVI du tableau floristique F avec les variables écologiques (E) en variables instrumentales.

Les variables écologiques sont spécifiées et introduites *a priori* : elles participent directement à l'élaboration des axes factoriels.

● *Taux d'inertie*

L'inertie totale du nuage est égale à 0,96. Les inerties absorbées par les deux premiers axes sont égales à 0,36 et 0,20. Les taux d'inertie sont donc respectivement égaux à 37,4 et 20,4 %. On observe que les inerties sont inférieures à celles de l'AFC ; cela est normal puisque les axes de l'AFCVI sont contraints par les variables du milieu.

La comparaison des inerties de l'AFCVI par rapport à celles de l'AFC présente, pour la typologie, un intérêt certain. Elles permettent de déterminer la capacité des variables du milieu à rendre compte des gradients écologiques exprimés par la végétation.

Mathématiquement, l'efficacité des variables de E pour prédire la structure de la végétation (tableau F) peut être calculée en faisant le rapport de la somme de l'inertie absorbée par les r premiers axes de l'AFCVI sur celle de l'inertie absorbée par les r premiers axes de l'AFC. Plus ce rapport, qui varie de 0 à 1, est proche de l'unité, meilleure est la pertinence des variables pour expliquer la répartition de la végétation. Dans notre exemple, nous obtenons une valeur de 0,81 pour le premier axe. Cette valeur élevée prouve que la classification des sols de la Plaine de la Lanterne vis-à-vis de l'hydromorphie est biologiquement pertinente.

● *Examen des relevés et des variables écologiques*

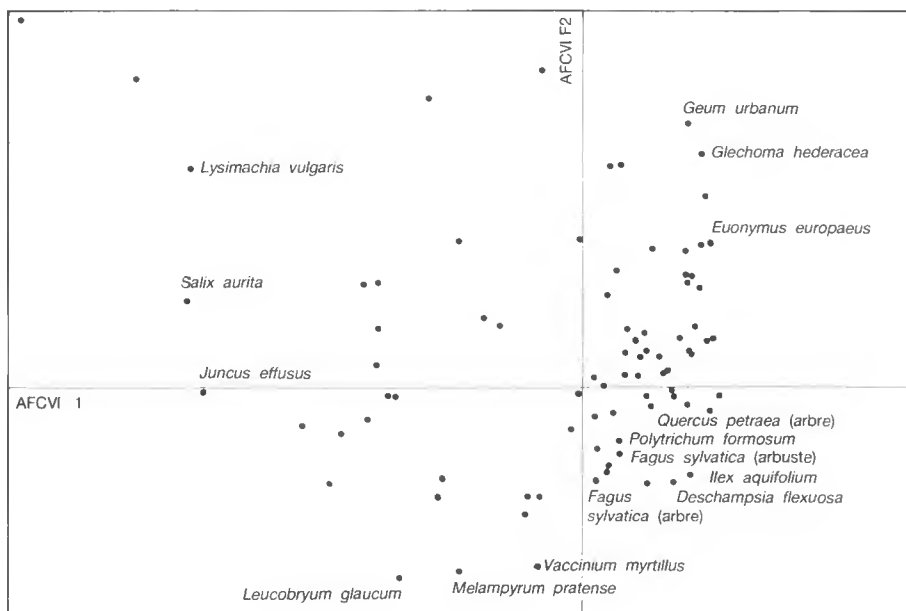
La position comparée des variables de milieu dans le plan factoriel 1-2 des relevés de l'AFC et de l'AFCVI est indiquée à la figure 3 (p. 543).

L'ordination proposée par l'AFCVI est beaucoup plus satisfaisante :

— Pour l'AFC comme pour l'AFCVI, trois groupes de relevés semblent s'individualiser vis-à-vis de l'hydromorphie. Il s'agit des sols tourbeux, des sols hydromorphes avec décoloration forte

Figure 5 **PLAN FACTORIEL F1-F2 DE L'AFCVI DES TABLEUX F ET E (VARIABLES INSTRUMENTALES) : REPRÉSENTATION DES ESPÈCES**

Les espèces mentionnées sont celles qui font l'objet d'un commentaire dans le texte.



à 20 cm, et de tous les autres sols. Dans le dernier groupe, on observe en outre avec l'AFCVI, que les classes d'hydromorphie les plus faibles induisent de légères modifications de la végétation, modifications qui n'étaient pas évidentes dans les résultats de l'AFC.

— Pour ce qui est du niveau trophique, là où l'AFC isolait quatre grandes formes d'humus (figure 3, p. 543) : tourbe, mull eutrophe, moder (avec difficulté) et autres formes d'humus, l'AFCVI en isole cinq : en plus des classes de l'AFC, elle distingue en effet les mulls mésotrophes du groupe des mulls acides et des mulls moders.

Ainsi, alors que l'AFC indiquait un lien assez lâche entre végétation et formes d'humus, l'AFCVI montre que chaque forme d'humus induit une végétation spécifique (mis à part les mulls acides et les mulls moders qui ont une végétation semblable). En outre, la dispersion des points autour des centres de gravités de chaque forme d'humus indique que les variations de pH ont une influence non négligeable sur la composition spécifique des relevés.

● *Examen des espèces*

Les résultats obtenus sur l'axe trophique (figure 5, p. 545) sont beaucoup plus conformes à l'expérience de terrain qu'avec les résultats de l'AFC. La Myrtille (*Vaccinium myrtillus*), le Leucobryum glauque (*Leucobryum glaucum*) et le Mélampyre des prés (*Melampyrum pratense*), d'une part, le Chêne sessile (*Quercus petraea*) et le Houx (*Ilex aquifolium*), d'autre part, sont positionnés de manière plus correcte. La position des espèces à l'autre extrémité de l'axe trophique et sur l'axe 1 reste, par rapport à l'AFC, globalement inchangée.

CONCLUSIONS

L'utilisation de l'AFCVI sur les données de la Plaine de la Lanterne illustre l'intérêt de cette méthode.

D'une part, les gradients exprimés par les axes de l'AFCVI sont uniquement ceux retenus comme intéressants pour la typologie (hydromorphie et niveau trophique). Les relevés peuvent donc être classés et regroupés de façon optimale par rapport à ces facteurs.

D'autre part, l'optimum écologique des espèces correspond à une réalité abiotique — puisque l'effet de maturation a été éliminé — et synthétique (contrairement à ce qui se passe avec la méthode des profils écologiques).

Enfin, l'AFCVI confirme ou infirme l'importance des facteurs du milieu retenus. Elle permet, en outre, un regroupement des classes qui n'ont qu'une faible influence sur la végétation. Ainsi, pour notre exemple, mulls acides et mulls moders d'une part et les trois classes de plus faible hydromorphie d'autre part peuvent être regroupés ; ce qui n'était pas évident avant analyse.

Pratiquement, par leur simplicité d'utilisation sur le terrain, les classes d'hydromorphie validées par l'AFCVI (tourbe, sol grisâtre à 20 cm, sol drainé) peuvent même se substituer avantageusement à l'utilisation de la végétation lors d'un diagnostic stationnel.

L'exemple de la Plaine de la Lanterne montre donc que l'AFCVI peut compléter avantageusement l'AFC ; tout particulièrement lorsque les relations entre les premiers axes de l'AFC et les descripteurs écologiques utilisés sont perturbées par des phénomènes liés à la physionomie et à la dynamique des peuplements.

Cependant plusieurs problèmes restent posés. Certains sont inhérents à ce type d'études observationnelles : la pertinence de l'AFCVI est conditionnée par le choix des variables de milieu

mesurées et par le plan d'échantillonnage. D'autres sont liés à la méthode même : l'AFCVI est une méthode linéaire (cas des variables quantitatives), le codage des données est pour une large part arbitraire (cas des variables qualitatives), il n'existe pas de test statistique de la pertinence biologique des variables écologiques (l'utilisation des pourcentages d'inertie reste empirique).

<p>J.-Cl. GÉGOUT Unité Écosystèmes forestiers et Dynamique des Paysages</p>	<p>F. HOULLIER Unité Dynamique des Systèmes forestiers</p>
<p>Laboratoire INRA/ENGREF de Recherches en Sciences forestières ÉCOLE NATIONALE DU GÉNIE RURAL, DES EAUX ET DES FORÊTS 14, rue Girardet 54042 NANCY CÉDEX</p>	

Remerciements

Les auteurs remercient vivement Michel Buffet, Sonia Darracq, Jean-Luc Dupouey, Jean-Claude Pierrat et Jean-Claude Rameau pour leurs remarques et suggestions sur les premières versions du manuscrit.

BIBLIOGRAPHIE

- BECKER (M.), LE GOFF (N.). — Diagnostic stationnel et potentiel de production. — *Revue forestière française*, vol. XL, n° spécial « Diagnostics en forêt », 1988, pp. 29-43.
- BRÉTHES (A.). — La Typologie des stations forestières : recommandations méthodologiques. — *Revue forestière française*, vol. XLI, n° 1, 1989, pp. 7-26.
- CHEssel (D.), MERCIER (P.). — Couplage de triplets statistiques et liaisons espèces-environnement. In : Biométrie et environnement, rassemblés par J.-D. Lebreton. — Paris : Masson, 1992.
- GÉGOUT (J.-C.). — Typologie des stations forestières de la Plaine de la Lanterne (Dépression périvosgienne). — 1992. — 117 p.
- LEBRETON (J.-D.), CHEssel (D.), PRODON (R.), YOCCOZ (N.). — L'Analyse des relations espèces - milieu par l'analyse canonique des correspondances. I - Variables de milieu quantitatives. — *Acta Oecologia, Oecologica generalis*, vol. 9, n° 1, 1988, pp. 53-67.
- LEBRETON (J.-D.), CHEssel (D.), RICHARDO-COULET (M.), YOCCOZ (N.). — L'Analyse des relations espèces - milieu par l'analyse canonique des correspondances. II - Variables de milieu qualitatives. — *Acta Oecologia, Oecologica generalis*, vol. 9, n° 2, 1988, pp. 137-151.
- MERCIER (P.). — Approche méthodologique de l'étude des relations entre la structure spatiale du peuplement ligneux et la végétation du sous-bois. — *Annales des Sciences forestières*, vol. 45, n° 3, 1988, pp. 275-290.
- SAPORTA (G.). — Probabilités, analyse des données et statistiques. — Éditions Technip, 1990. — 493 p.
- TER BRAAK (C.J.F.). — Canonical correspondence analysis : a new eigenvector technique for multivariate direct gradient analysis. — *Ecology*, vol. 67, n° 5, 1986, pp. 1167-1179.
- TER BRAAK (C.J.F.). — The analysis of vegetation-environment relationships by canonical correspondence analysis. — *Vegetatio*, 69, 1987, pp. 69-77.