
Analyse spatiale de la communauté végétale de la portion nord du désert du Chihuahua

Guillaume Blanchet, Pierre Legendre

*Département de sciences biologiques,
Université de Montréal,
Case postale 6128, succursale Centre-ville,
Montréal (Québec) Canada, H3C 3J8*

RÉSUMÉ. Cet article présente les résultats d'analyses statistiques réalisées sur la communauté végétale de la portion nord du désert du Chihuahua afin d'élucider sa structure spatiale. Ces analyses avaient pour but d'approfondir les résultats de Muldavin [MUL 02] en répondant aux questions suivantes : 1) à quelle échelle spatiale la communauté est-elle significativement structurée ? 2) Jusqu'à quel point peut-on expliquer la diversité bêta par les variables environnementales ? 3) Quelles sont les associations significatives d'espèces ?

MOTS-CLÉS : Concordance de Kendall, coordonnées principales d'une matrice de voisinage (CPMV), désert du Chihuahua, diversité bêta, espèces indicatrices, partition de la variation, krigeage, partitionnement par la méthode des K centroïdes (K-means).

1 Introduction

Le désert du Chihuahua couvre quelque 630000 km² et s'étend du Nouveau-Mexique (USA) jusqu'à l'état de San Luis Potos (Mexique). La flore de ce désert est d'une grande richesse. Nos analyses porteront sur un échantillon de 1510 sites situés dans la portion septentrionale de ce désert au Nouveau-Mexique et au Texas. 529 espèces de plantes furent répertoriées par Muldavin. Ce spécialiste de la végétation du désert chihuahuaien a développé un indice d'affinité floristique, fondé sur l'abondance de 3 espèces indicatrices, qui lui a permis d'établir la limite nord du désert du Chihuahua au Nouveau-Mexique [MUL 02].

Dans les écosystèmes, un très grand nombre de variables influencent la distribution des espèces vivant dans un milieu donné. Pour structurer l'échantillonnage, il importe de se poser des questions précises qui permettront de déterminer quelles variables doivent être observées ou mesurées. Les analyses présentées dans cet article ont pour but de répondre à trois questions à propos du désert du Chihuahua : 1) à quelle échelle spatiale la communauté végétale est-elle significativement structurée ? 2) Jusqu'à quel point peut-on expliquer la variation de composition entre les sites (diversité bêta) par la variation en altitude des sites d'échantillonnage ? 3) Quelles sont les associations significatives d'espèces ?

Différentes méthodes seront utilisées afin de répondre à ces questions. Pour la première question, les coordonnées principales d'une matrice de voisinages (CPMV) seront utilisées [BOR 02, LEG 05b], alors que la deuxième question trouvera réponse avec le partitionnement de la variation [BOR 92]. Une analyse de concordance de Kendall [LEG 05a] sera utilisée pour répondre à la troisième question. Avant ces dernières analyses, le tableau des abondances d'espèces a subi une transformation de Hellinger qui consiste à calculer la racine carrée des données préalablement transformées en abondances relatives par site [LEG 01].

2 Structure spatiale de la communauté végétale

L'analyse en coordonnées principales d'une matrice de voisinage (CPMV) réalise une décomposition spectrale des relations spatiales entre les sites. Cette décomposition permet de modéliser la structure spatiale multi-échelle du tableau sites \times espèces en utilisant uniquement les variables CPMV dérivées des positions géographiques des sites. Les CPMV peuvent décrire des structures spatiales contrôlées par des processus à échelle très large, comme la présence de chaînes de montagnes de part et d'autre de la région d'échantillonnage, mais aussi des phénomènes locaux comme la présence d'une espèce dans une zone très spécifique de la région échantillonnée. 71 CPMV significatives, sur 722 au départ, ont été retenues par sélection progressive pas à pas pour décrire la structure spatiale de la végétation du désert du Chihuahua. Elles représentent des influences spatiales à échelle large (taches d'environ 100 km : Fig. 1, gauche) à moyenne (taches d'environ 10 km). Une CPMV se distingue : elle explique une portion significative de variation spatiale dans une zone particulière ; la tache mesure environ 10 km de diamètre (Fig. 1, droite).

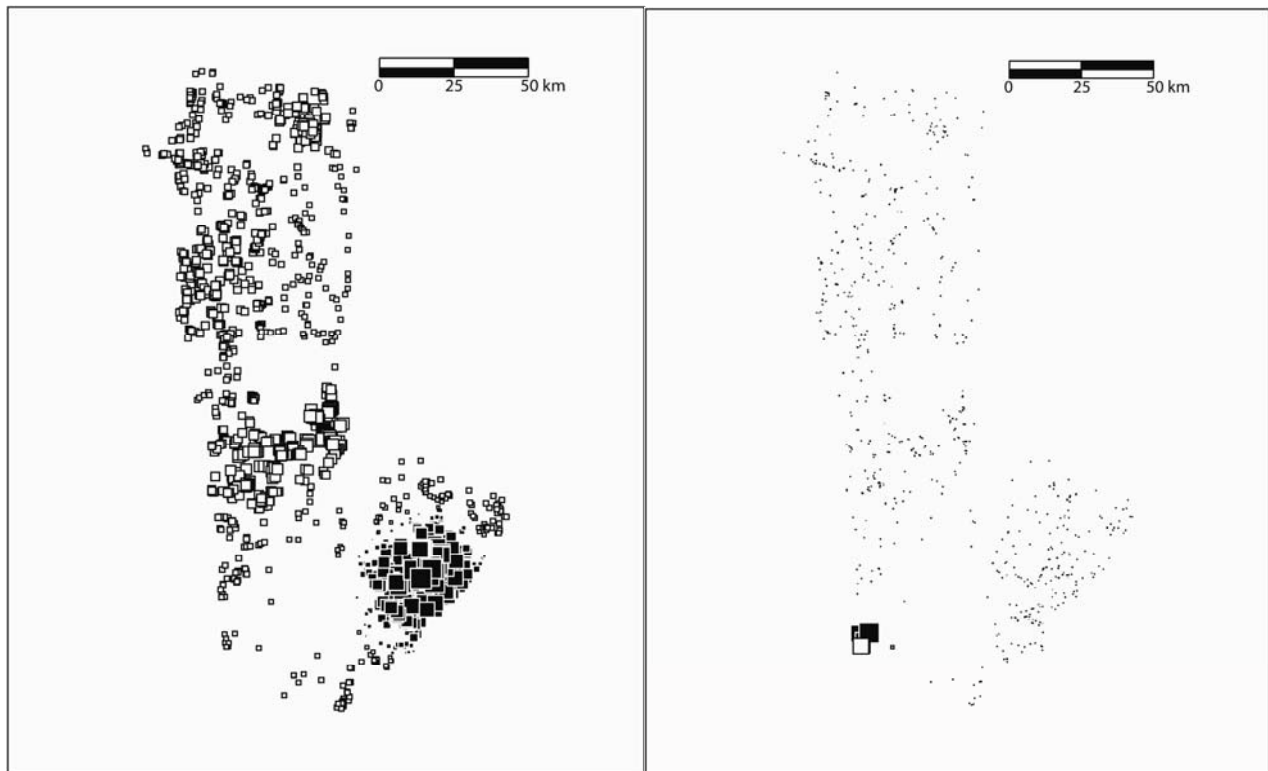


Figure 1. Carte de deux des variables CPMV expliquant une portion significative de la variation du tableau sites \times espèces. La taille des carrés est proportionnelle à la valeur de la variable CPMV à chaque site ; carrés blancs : signe négatif ; carrés noirs : signe positif. La carte de gauche représente la première CPMV qui explique un phénomène à grande échelle ; celle de droite correspond à un phénomène local situé dans la portion sud-ouest de la carte.

3 Partition de la diversité bêta : influences environnementale et spatiale

Nous ajouterons maintenant une variable environnementale importante, l'altitude, à l'analyse et tenterons de déterminer jusqu'à quel point nous pouvons expliquer la variation de composition entre les sites (diversité bêta) par la variation en altitude des sites d'échantillonnage.

Pour vérifier si l'altitude avait de l'importance pour expliquer la variation en composition végétale entre les sites, nous avons partitionné les sites par la méthode des K centroïdes (K -means), puis testé l'hypothèse que les sites se regroupent en fonction de l'altitude. La méthode des K centroïdes regroupe les

sites selon leurs proximités ; l'algorithme a préservé la distance de Hellinger entre les sites puisque les données avaient subi une transformation de Hellinger avant le partitionnement [LEG 01]. Nous avons retenu la solution minimisant la variation intragroupe, pour $K = 2$ à $K = 10$ groupes, après 100 attributions aléatoires des objets aux groupes initiaux. La statistique de Calinski-Harabasz fut utilisée pour déterminer le nombre optimal de groupes au sens des moindres carrés. Cette statistique indiqua que la division des sites en 5 groupes était optimale. Une analyse de variance portant sur la variable altitude supporta l'hypothèse d'une forte relation entre ces 5 groupes et l'altitude des sites ($F = 185.80$, $P < 0.0001$). La cartographie de l'altitude par krigeage avait été nécessaire pour obtenir des estimations de l'altitude aux 1510 sites d'échantillonnage alors que des données d'altitude n'étaient disponibles que pour 1471 sites.

L'examen visuel de la distribution des 5 groupes de sites superposés à la topographie sur la carte permet d'apprécier leur relation avec l'altitude. Les deux groupes de sites les plus importants sont situés à basse altitude, dans une vallée. Un groupe se trouve au flanc des deux chaînes de montagne alors que les deux autres groupes recouvrent les régions de plus hautes altitudes.

La partition de la variation spatiale du tableau sites \times espèces (diversité bêta), par analyse canonique de redondance (ACR ou RDA), a permis de déterminer quelle fraction de la variation des données est influencée par quelles variables. La communauté végétale du désert du Chihuahua est structurée à 19.4 % (R^2 ajusté, $P = 0.001$; $R^2 = 57.9\%$) par les 722 variables CPMV et à 5.5 % (R^2 ajusté, $P = 0.001$; $R^2 = 5.6\%$) par l'altitude. 71 des 722 CPMV ont été retenues par sélection progressive des variables explicatives en ACR. Cette méthode permet aussi de calculer, par soustraction, la fraction [b] qui correspond à la variation expliquée conjointement par les deux types de variables (CPMV et altitude); cette portion compte pour 4.4 % (R^2 ajusté) de la variation du tableau sites \times espèces (Fig. 2).

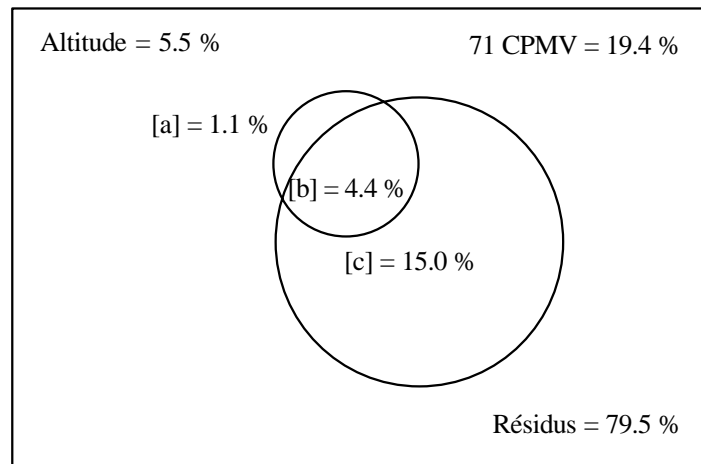


Figure 2. Diagramme de Venn partitionnant la variation spatiale (R^2 ajustés) de la communauté végétale du désert du Chihuahua. [a + b] représente la portion de la variation expliquée par l'altitude alors que [b + c] est la portion représentée par les variables CPMV. Les pourcentages sont relatifs à la variation totale du tableau sites \times espèces après transformation de Hellinger ($n = 1510$, $p = 529$). Les portions de variation expliquées par les cercles sont approximatives.

Nous avons cherché à savoir si certaines espèces étaient particulièrement associées à chacun des 5 groupes de sites. Pour ce faire, nous avons utilisé une méthode de recherche des espèces indicatrices des groupes d'une partition [DUF 97]. Au niveau de signification $\alpha = 0.01$, nous avons identifié 170 espèces indicatrices de l'un ou l'autre des cinq groupes de sites.

4 Les associations significatives d'espèces

Les associations d'espèces sont intéressantes car elles permettent de synthétiser certaines caractéristiques environnementales au niveau des associations plutôt qu'à celui des espèces individuelles. Par ailleurs, les

associations d'espèces sont moins sujettes à l'erreur d'échantillonnage que les espèces individuelles ; on peut donc les utiliser pour prédire certaines caractéristiques du milieu physique [LEG 05a]. Les espèces furent partitionnées en groupes en appliquant la méthode des K centroïdes appliquée aux vecteurs propres d'une analyse en composante principale (ACP) de la matrice de corrélations entre les espèces. Le coefficient de Calinski-Harabasz présentait un maximum pour la division des espèces en 9 groupes.

Le coefficient de concordance de Kendall (W) a permis de tester la concordance des espèces réunies en groupes préliminaires et d'identifier celles qui sont significativement concordantes avec d'autres espèces de leur groupe [LEG 05a]. Le nombre d'espèces significativement concordantes par groupe va de 47 pour chacun des deux plus grands groupes à 28, 19, 18, 16, 15, 15, et enfin 9 pour le plus petit groupe.

5 Discussion

La plupart des 71 variables CPMV significatives représentent la variation spatiale de la communauté végétale à échelle large dans le nord du désert du Chihuahua. L'influence que pourraient avoir les deux chaînes de montagnes sur la répartition spatiale des espèces en est un bon exemple. Il faut cependant noter la présence possible d'une ou de quelques espèces endémiques à une zone particulière du désert, la dernière variable CPMV significative ayant détecté une structure spatiale très locale.

La partition de la variation a montré que 22.4 % (R^2 ajusté) de la diversité bêta était explicable par des variables CPMV alors que l'altitude n'expliquait que 21% de cette variation spatiale. Puisque les variables CPMV significatives représentent surtout de la variation à échelle large, cela suggère que la variation qu'elles modélisent a pour origine des variables physiques du milieu. Il nous faudra ajouter des variables physiques ou pédologiques au tableau des variables environnementales. Considérant le fait qu'il y a une vallée dans la région d'échantillonnages, il est cependant possible que la dispersion de certaines espèces se soit arrêtée parce qu'elles ne pouvaient pas passer la barrière naturelle créée par la chaîne de montagnes. Il se peut également qu'un microclimat dans la vallée favorise la reproduction de certaines espèces.

Muldavin [MUL 02], a utilisé trois espèces de plantes, qu'il a considérées comme indicatrices, pour établir la limite nord du désert du Chihuahua au Nouveau-Mexique. Les résultats de notre recherche d'espèces indicatrices fournissent de nouvelles espèces qui pourraient être utilisées pour construire de nouveaux indices pour délimiter le désert du Chihuahua.

6 Bibliographie

- [BOR 92] BORCARD D., LEGENDRE P., DRAPEAU P., "Partialling out the spatial component of ecological variation", *Ecology*, vol. 73, n° 3, 1992, p. 1045-1055.
- [BOR 02] BORCARD D., LEGENDRE P., "All-scale analysis of ecological data by means of principal coordinates of neighbour matrices", *Ecological Modelling*, vol. 153, 2002, p. 51-68.
- [DUF 97] DUFRÊNE M., LEGENDRE P., "Species assemblages and indicator species: the need for a flexible asymmetrical approach", *Ecological Monographs*, vol. 67, n°3, 1997, p. 345-366.
- [LEG 05a] LEGENDRE P., "Species association: the Kendall coefficient of concordance revisited", *Journal of Agricultural, Biological and Environmental Statistics*, 2005 (*sous presse*).
- [LEG 05b] LEGENDRE P., BORCARD D., "Quelles sont les échelles spatiales importantes dans un écosystème ?", in: *Analyse statistique de données spatiales*, Driesbeke J.-J., Lejeune M., Saporta G., éditeurs, Éditions TECHNIP, Paris, 2005 (*sous presse*).
- [LEG 01] LEGENDRE P., GALLAGHER E.D., "Ecologically meaningful transformations for ordination of species data", *Oecologia*, vol. 129, 2001, p. 271-280.
- [MUL 02] MULDAVIN E.H., "Some floristic characteristics of the northern Chihuahuan Desert: a search for its northern boundary", *Taxon*, vol. 51, 2002, p. 453-462.