

Projet Hydrogen

Abdoulaye Diabakhaté

26 juin 2018

Table des matières

1	Best Subset Of Environmental Variables With Maximum (Rank) Correlation With Community Dissimilarities	3
1.1	Introduction	3
2	Fraction de taille 0_0.2	4
3	Fraction de taille 0.22_3	4
4	Fraction de taille 20_180	4
5	Fraction de taille 5_20	5
6	Fraction de taille 180_2000	5
7	Fraction de taille 0.8_5	5
8	Conclusion	6

1 Best Subset Of Environmental Variables With Maximum (Rank) Correlation With Community Dissimilarities

1.1 Introduction

La fonction [bioenv](#) calcule une matrice de dissimilarité de communauté en utilisant [vegdist](#).

Ensuite, elle sélectionne tous les sous-ensembles possibles de variables environnementales, met à l'échelle les variables et calcule les distances euclidiennes pour ce sous-ensemble en utilisant [dist](#).

Ensuite, elle trouve la corrélation entre les dissimilarités communautaires et les distances environnementales, et pour chaque taille de sous-ensembles, enregistre le meilleur résultat.

Il y'a $2^p - 1$ sous-ensembles de p-variables.

En effet, un coefficient de corrélation (typiquement le coefficient de corrélation de rang de Spearman) est calculé entre les deux matrices et le meilleur sous-ensemble de variables environnementales peut alors être identifié et soumis ensuite à un test de permutation pour déterminer la signification.

La méthode est également largement acceptée par la communauté scientifique en raison de sa flexibilité à travers une grande variété de données et est complètement non paramétrique ; l'article [Clarke et Ainsworth \(1993\)](#) décrivant la méthode compte 674 citations sur Google Scholar au moment de cette publication.

Pour toutes les fractions qui suivront, nous avons 15 variables en compétition :

Lat, Long, T, Sal, chl_a, O₂, NO₃⁻, NO₃⁻, NO₂⁻, NH₄⁺, SSD, Phos, Si, depth, Fe

2 Fraction de taille 0__0.2

Pour cette fraction de taille composée de [11 matrices de distances](#), 10 distances ont retenu 8 paramètres qui sont :[Lat-Long-T-Chla-O2_l-NO3-NH4_l-Fe](#).

L'unique distance restant, correspondant à la matrice [whittaker_prevalence](#) a retenu en plus des 8 paramètres précédemment cités, [Sal-NO3m_l-SSD](#).

3 Fraction de taille 0.22__3

Pour cette fraction de taille composée de [11 matrices de distances](#), 7 distances ont retenu 2 paramètres qui sont : [T-Si](#).

2 distances ont retenu [T-Phos](#) et enfin 2 autres en ont retenu [T-Phos-O2_l](#)

4 Fraction de taille 20__180

Pour cette fraction de taille composée de [11 matrices de distances](#), l'ensemble de ces matrices ont retenu [T-Chla](#).

5 Fraction de taille 5__20

Pour cette fraction de taille composée de 13 matrices de distances, 7 distances ont retenu 8 paramètres qui sont : [Lat-T-Sal-Chla-O2_l-NO3m_l-NH4_l-SSD](#).

4 distances ont retenu la même liste sauf [SSD](#) et enfin 2 distances ont retenu la même liste sauf [NO3m_l](#).

6 Fraction de taille 180__2000

Pour cette fraction de taille composée de 13 matrices de distances, 2 distances ont retenu 4 paramètres qui sont : [T-Chla-NH4_l-SSD](#).

6 distances ont retenu la même liste sauf [Chla-SSD](#) ; 3 distances ont retenu la même liste sauf [NH4_l-SSD](#), et enfin 1 distance a retenu la même liste, avec [Lat](#) en plus.

7 Fraction de taille 0.8__5

Pour cette fraction de taille composée de 13 matrices de distances, 2 distances ont retenu 8 paramètres qui sont : [Lat-T-chla-NO3m_l-NH4_l-Fe-SSD-Si](#) ; 7 distances ont retenu [T-Sal-Chla-NO3m_l-Fe-SSD](#) et enfin 4 distances ont retenu [T-Chla-NH4_l-Fe-SSD](#).

8 Conclusion

On peut noter que pour une fraction donnée, on a quasiment le même comportement des variables pour toutes les distances alors que pour 2 fractions données on a un comportement différent.