

Université de Montréal

**Évaluation de l'unicité écologique sur des échelles spatiales
étendues et continues**

par

Gabriel Dansereau

Département de sciences biologiques

Faculté des arts et des sciences

Mémoire présenté en vue de l'obtention du grade de

Maître ès sciences (M.Sc.)

en Sciences biologiques

22 avril 2021

Université de Montréal

Faculté des arts et des sciences

Ce mémoire intitulé

Évaluation de l'unicité écologique sur des échelles spatiales étendues et continues

présenté par

Gabriel Dansereau

a été évalué par un jury composé des personnes suivantes :

Anne-Lise Routhier

(président-rapporteur)

Timothée Poisot

(directeur de recherche)

Pierre Legendre

(codirecteur)

Élise Filotas

(membre du jury)

Résumé

...sommaire et mots clés en français...

Abstract

...summary and keywords in english...

Table des matières

Résumé	5
Abstract	7
Liste des tableaux	11
Table des figures.....	13
Liste des sigles et des abréviations	15
Remerciements.....	17
Introduction.....	19
0.1. Avant-propos.....	19
0.2. Mise en contexte	20
0.3. Biodiversité.....	22
0.4. Données	23
0.5. Modèles	23
First Article. Titre de l'article	25
1. Introduction	26
Bibliography	31

Liste des tableaux

Table des figures

Liste des sigles et des abréviations

KQ-Methode	Méthode des moindres carrés, de l'allemand <i>Methode der kleinsten Quadrate</i>
MCMC	Monte Carlo par chaînes de Markov, de l'anglais <i>Markov Chain Monte Carlo</i>
MSE	Erreur quadratique moyenne, de l'anglais <i>Mean Square Error</i>
NDR	Retract d'un voisinage, de l'anglais <i>Neighbourhood Deformation Retract</i>
OLS	Moindres carrés ordinaires, de l'anglais <i>Ordinary Least Square</i>
ZFC	Théorie des ensembles de Zermelo-Fraenkel avec l'axiome du choix

Remerciements

...remerciements...

Introduction

0.1. Avant-propos

Ce mémoire est parti d'une idée assez simple. Celle de combiner les grandes bases de données en ligne sur biodiversité et les méthodes d'intelligence artificielle pour prédire la distribution géographique des sites de biodiversité exceptionnelle.

Il existe de nombreuses façons de mesurer la biodiversité et d'identifier les endroits où celle-ci est la plus exceptionnelle. Ces sites sont notamment importants à des fins de conservation et de gestion des aires protégées. Par contre, plusieurs de ces mesures ont été développées pour un petit nombre de sites ou d'espèces et nécessitent de connaître la composition en espèces exacte à un endroit donné. Or, la composition exacte est difficile, voire impossible, à connaître pour certains endroits, notamment en région éloignée, ce qui empêche l'utilisation de certaines mesures sur de très grandes échelles spatiales.

Des sources de données nouvelles se sont cependant développées récemment. Il s'agit notamment de bases de données en ligne massives comme GBIF, eBird et iNaturalist. Celles-ci sont imparfaites, mais contiennent tout de même beaucoup d'informations pouvant être mises à profit, notamment pour une évaluation de la biodiversité à grande étendue spatiale. Ces bases de données, largement basées sur des contributions citoyennes, sont cependant le reflet de biais déjà présents dans les données écologiques, notamment en faveur des villes, et ne règlent pas le manque de données en région éloignées.

D'un autre côté, nous assistons également au développement de méthodes prédictives en intelligence artificielle de plus en plus poussées et de plus en plus performantes. Celles-ci offrent

la possibilité d'obtenir des prédictions assez justes pour la distribution des espèces, ainsi que des inférences compensant entre autres pour le manque de données.

Par contre, regrouper ces différents éléments, les bases de données massives et les méthodes d'intelligence artificielle, pour la prédiction de la biodiversité soulève des questions d'applicabilité, puisque les mesures de biodiversité et le cadre théorique derrière celles-ci n'ont justement pas été développées pour une échelle aussi étendue et un si grand volume de données.

Mon projet est donc à la jonction de ces éléments. J'ai exploré les bases de données et les méthodes prédictives, puis j'ai cherché à vérifier comment elles peuvent s'intégrer avec les mesures de biodiversité actuelles, en particulier pour comprendre ce qu'elles font ressortir lorsque utilisées de cette façon. Je souhaite que ces travaux amènent une nouvelle vue des possibilités actuellement ouvertes, ainsi qu'une nouvelle compréhension des concepts soulevés. J'espère que ceux-ci pourront être utiles pour la conservation et la gestion des aires protégées.

- Idée simple
- Mesures pour biodiversité → identification sites exceptionnels
 - MAIS besoin composition en espèces exacte
 - Donc impossible pour certains endroits
 - Particulièrement plus difficile pour grandes étendues spatiales
- Grandes bases de données sur biodiversité
 - Reflet des biais échantillonnages → surtout villes, peu données au Nord
- Méthodes d'intelligence artificielle + en + performantes
 - Possibilité prédictions justes, inférences compensant manque données
- Or, soulève questions sur l'applicabilité des mesures, sur ce qu'elles montrent réellement.

0.2. Mise en contexte

L'identification des zones clés de biodiversité est l'une des priorités pour la conservation et la gestion des aires protégées. En particulier, il y a actuellement un besoin pour développer des méthodes permettant d'identifier les sites les plus importants pour la biodiversité de façon efficace

sur de grandes étendues spatiales. Or, identifier de tels endroits implique plusieurs questions complexes. En premier, il est nécessaire de définir ce que constituent des zones clés de biodiversité. Plusieurs définitions et plusieurs mesures ont été suggérées à ce sujet, mais celles-ci varient généralement quand à l'étendue spatiale ou aux régions ciblées. Ensuite, au-delà de la définition de la biodiversité, il est nécessaire de trouver des données qui permettent d'évaluer avec justesse le caractère unique ou exceptionnel des sites. La récolte de données en écologie est parfois difficile à certains endroits, par exemple pour certaines régions éloignées. Lorsque nécessaire, les observations directes peuvent parfois être remplacées par des prédictions réalisées à partir de données plus générales. Par contre, une panoplie de méthodes prédictives existent et la plupart d'entre elles n'ont pas été évaluées spécifiquement avec certaines mesures de biodiversité. Finalement, il est également nécessaire d'adapter à la fois les mesures de biodiversité et les méthodes prédictives aux grandes étendues spatiales. La biodiversité varie parfois différemment en fonction des échelles. Il en est de même quand à la performance des mesures. Intégrer le tout peut donc s'avérer complexe et implique d'avoir une compréhension développée des définitions de la biodiversité, des données et des méthodes disponibles, ainsi que des facteurs pouvant influencer la biodiversité en fonction des échelles spatiales.

Dans mon mémoire, je me suis intéressé à cette question en cherchant à vérifier l'applicabilité d'une mesure donnée, celles des contributions locales à la diversité bêta, pour identifier les zones de biodiversité exceptionnelle à grande étendue spatiale. De plus, j'ai cherché à vérifier si cette méthode pouvait être appliquée à des prédictions de distribution d'espèces produites à partir de données provenant de grandes bases de données citoyennes. Mon mémoire est donc divisé en trois sections. La première comporte une mise en contexte, ainsi qu'une revue de littérature présentant les concepts pertinents. La seconde partie consiste en un article scientifique présentant les résultats de mes travaux et analyses. La dernière partie consiste en un retour sur les résultats, en lien avec la mise en contexte présentée dans la première section.

0.3. Biodiversité

La biodiversité comporte une composante spatiale très importante, dont la conception a évolué avec le temps.

Notre conception de la biodiversité dans l'espace a évolué avec le temps, allant d'une conception large de la structure spatiale de la biodiversité d'une région

(Whittaker1960VegSis; Whittaker1972EvoMea) a été le premier à décrire les concepts de diversité alpha, bêta et gamma.

La diversité bêta représente la variation dans la composition en espèces au sein des sites d'une région géographique d'intérêt (Legendre et al., 2005).

La diversité bêta totale au sein d'une communauté peut, entre autres, être décomposée en contributions locales à la diversité bêta (*local contributions to beta diversity*, LCBD), ce qui permet d'identifier les sites possédant une composition en espèces exceptionnelle, donc une biodiversité unique (Legendre et De Cáceres, 2013).

La plupart des études ayant utilisé la mesure des LCBD l'ont utilisé à échelle locale, donc sur des étendues spatiales restreintes, et sur un petit nombre de sites (da Silva et Hernández, 2014; Heino et al., 2017; Heino et Grönroos, 2017; Legendre et De Cáceres, 2013). Par exemple, l'étude ayant présenté la mesure l'a utilisé sur des communautés de poissons échantillonnées à intervalles le long d'une rivière (Legendre et De Cáceres, 2013).

Quelques études ont utilisé la mesure des LCBD sur de plus grandes étendues spatiales, donc comportant potentiellement une plus forte hétérogénéité spatiale, mais ces études comportaient un nombre de sites assez faible (Poisot et al., 2017; Taranu et al., 2020; Yang et al., 2015).

Quelques études récentes l'ont utilisée sur des données arrangées en grille, donc spatialement continues (D'Antraccoli et al., 2020; Legendre et Condit, 2019; Tan et al., 2017; Tan et al., 2019). Cependant, celles-ci portaient également sur des échelles spatiales restreintes.

Une étude récente l'a utilisée sur un grand nombre de sites (Niskanen et al., 2017). Cette dernière a également cherché à prédire la mesure directement en fonction de conditions environnementales.

Une avenue intéressante a été explorée par **Vasconcelos2018ExpImp** en utilisant la mesure sur des prédictions réalisés par des modèles de distribution d'espèces en fonction des conditions climatiques actuelles et de scénarios de changements climatiques.

0.4. Données

De plus en plus de données sont disponibles en ligne, et ce, massivement. Il faut donc passer vers une approche dirigée par les données disponibles dans une optique de synthèse, ce qui permet généralement de révéler de nouvelles informations écologiques à partir de données existantes (Poisot et al., 2019).

De grandes bases de données sur la biodiversité fournissent des informations écologiques à exploiter, notamment eBird, GBIF et iNaturalist.

En même temps, nous disposons désormais de données de plus en plus précises sur les conditions environnementales partout sur le globe. Par exemple, WorldClim et CHELSA fournissent des données climatiques, alors que Copernicus et EarthEnv fournissent des informations sur l'utilisation du territoire. Dans les deux cas, ces informations sont parfois disponibles à des échelles spatiales très fines.

0.5. Modèles

L'étude de la biodiversité dans l'espace peut également profiter de l'essor des méthodes prédictives, entre autres celles issues de l'intelligence artificielle. Les modèles de distribution d'espèces (*species distribution models*, SDM) (Guisan et Thuiller, 2005) servent notamment à prédire la distribution des espèces en fonction des conditions environnementales à partir d'observations déjà réalisées.

Il existe une grande variété de SDM, allant des premiers modèles comme BIOCLIM (**Nix1986BioAna**), à des modèles plus récents comme MAXENT (**Phillips2006MaxEnt**; **Phillips2008ModSpe**) et à différentes méthodes apparentées à l'intelligence artificielle (et non spécifique à la distribution d'espèces), notamment les *Randoms Forests* (**Breiman2001RanFor**) ou encore les *Boosted Regression Trees* (**Elith2008WorGui**).

Carlson (2020) ont récemment suggéré d'utiliser les *Bayesian Additive Regression Trees* (BART)(Chipman et al., 2010) pour les SDM, ce qui permet d'obtenir de meilleurs résultats en réduisant le sur-ajustement, tout en permettant d'évaluer l'incertitude sous une formulation bayésienne.

Plusieurs de ces méthodes représentent toutefois des méthodes d'apprentissage supervisé, de sorte qu'elles ont besoin d'être entraînées sur des données déjà étiquetées. La principale conséquence au niveau des SDM est donc le besoin de disposer de données d'absence, en plus de données de présence, afin de pouvoir entraîner les algorithmes. Or, les données d'absence sont plus difficiles à obtenir, notamment en raison du problème du double-zéro (**Legendre2012NumEco**).

La base de données *eBird* comporte toutefois un avantage à ce sujet, puisqu'il s'agit d'une base de données semi-structurée (Johnston et al., 2020). Les données (et donc les observations) y sont structurées par listes d'observations. En rapportant leurs observations, les utilisateurs doivent déclarer si celles-ci constituent une liste complète des espèces détectées lors de leur échantillonnage. Ainsi, cela permet un peu plus justement d'inférer la non-détection d'autres espèces

First Article.

Titre de l'article

by

Hima Ginère¹

(¹) 1252i rue complexe
Université du plan complexe

This article was submitted in Une revue.

The main contributions of Gabriel Dansereau for this articles are presented.

- Calcul de telle chose;
- Vérification de telle équation;
- Idée pour telle définition;
- Démonstration de tel théorème.

Le coauteur1 a suggéré telle chose.

Le coauteur2 a fait telle calcul.

RÉSUMÉ. Le résumé en français.

Mots clés : Mots clés

ABSTRACT. The english abstract.

Keywords: Key words

1. Introduction

Beta diversity, defined as the variation in species composition among sites in a geographic region of interest (Legendre et al., 2005), is an essential measure to describe the organization of biodiversity through space. Total beta diversity within a community can be partitioned into local contributions to beta diversity (LCBD) (Legendre & De Cáceres, 2013), which allows the identification of sites with exceptional species composition, hence unique biodiversity. Such a method, focusing on specific sites, is useful for both community ecology and conservation biology, as it highlights areas that are most important for their research or conservation values. However, the use of LCBD indices is currently limited in two ways. First, LCBD indices are typically used on data collected over local or regional scales with relatively few sites, for example on fish communities at intervals along a river or stream (Legendre & De Cáceres, 2013). Second, LCBD calculation methods require complete information on community composition, such as a community composition matrix Y ; thus, they are inappropriate for partially sampled sites (e.g. where data for some species is missing), let alone for unsampled ones. Accordingly, the method is of limited use to identify areas with exceptional biodiversity in regions with sparse sampling. However, predictive approaches are increasingly common given the recent development of computational methods, which often uncover novel ecological insights from existing data (Poisot et al., 2019), including in unsampled or lesser-known locations, as well as larger spatial scales. Here, we examine whether the LCBD method can assess ecological uniqueness over broad and continuous scales based on predictions of species distributions and evaluate whether this reveals novel ecological insights regarding the identification of exceptional biodiversity areas.

Species distribution models (SDMs) (Guisan & Thuiller, 2005) can bring a new perspective to LCBD studies by filling in gaps and performing analyses on much broader scales. In a community

matrix Y , such as required for LCBD calculation, ecological communities are abstracted as assemblages of species present at different sites. Viewing communities as such opens the perspective of predicting community composition from predictions of individual species, which is precisely the aim of SDMs. Community-level modelling from SDMs is not an especially novel idea (Ferrier et al., 2002; Ferrier & Guisan, 2006), but it is increasingly relevant with the advent of large-scale, massive, and open data sources on species occurrences, often contributed by citizens, such as eBird and GBIF. At their core, SDMs aim at predicting the distribution of a species based on information about where the species was previously reported, matched with environmental data at those locations, and then make predictions at other (unsampled) locations based on their respective environmental conditions. However, going from single-species SDMs to a whole community is not a trivial task, and many solutions have been suggested, such as stacked species distribution models (S-SDMs) (Ferrier & Guisan, 2006), spatially explicit species assemblage modelling (SESAM) (Guisan & Rahbek, 2011), joint species distribution models (JSDMs) (Pollock et al., 2014), and hierarchical modelling of species communities (HMSC) (Ovaskainen et al., 2017). These alternative methods all have different strengths, but even S-SDM, in a sense the most simple and less community-specific method, has been shown to provide reliable community predictions (Norberg et al., 2019; Zurell et al., 2020). This is important, as in the context of large-scale studies with a high number of sites and species, reducing the model complexity with a simpler yet efficient model such as an S-SDM can reduce the number of computations in an important way. Regardless of the method used, community-level analyses can be applied to the resulting community prediction, but this has been lacking for community measures other than species richness (Ferrier & Guisan, 2006). Notably, the LCBD framework has, to our knowledge, never been applied to SDM results. The computation of local contributions to beta diversity (LCBD) on SDM predictions, however, raises the issue of calculating the uniqueness scores on much larger community matrices than on the typical scales on which it has been used.

The total number of sites will increase (1) because of the continuous scale of the predictions, as there will be more sites in the region of interest than the number of sampled sites, and (2) because of the larger spatial extent allowed for the SDM predictions. A high number of SDM-predicted

sites with a large extent opens up the possibility of capturing a lot of variability of habitats and community composition, but also many very similar ones, which could change the way that exceptional sites contribute to the overall variance in the large-scale community. LCBD scores have typically been used at local or regional scales with relatively few sites (da Silva & Hernández, 2014; Heino et al., 2017; Heino & Grönroos, 2017; Legendre & De Cáceres, 2013). Some studies did use the measure over broader, near-continental extents (Poisot et al., 2017; Taranu et al., 2020; Yang et al., 2015), but the total number of sites in these studies was relatively small. Recent studies also investigated LCBD and beta diversity on sites distributed in grids or as pixels of environmental raster layers, hence continuous scales, but these did not cover large extents and a high number of sites (D’Antraccoli et al., 2020; Legendre & Condit, 2019; Tan et al., 2017; Tan et al., 2019). Niskanen et al. (2017) predicted LCBD values of plant communities (and three other diversity measures) on a continuous scale and a high number of sites (> 25 000) using Boosted Regression Trees (BRTs). However, they modelled the diversity measures directly instead of modelling species distributions first, as we are suggesting here. They obtained lower predictive accuracy for LCBD than for their other diversity measures, mentioning that it highlighted the challenge of predicting LCBD specifically. They also computed LCBD indices at a regional scale, not a continental one, while using a fine spatial resolution (1 km x 1 km). Therefore, the distribution of LCBD values at broad, continuous scales with a high number of sites and predicted species assemblages remains to be investigated.

Measuring ecological uniqueness from LCBD indices on extended continuous scales also raises the question of which sites will be identified as exceptional and for what reason. The method intends that sites should stand out and receive a high LCBD score whenever they display an exceptional community composition, be it a unique assemblage of species that may have a high conservation value or a richer or poorer community than most in the region (Legendre & De Cáceres, 2013). Both the original study and many of the later empirical ones have shown a negative relationship between LCBD scores and species richness (da Silva & Hernández, 2014; Heino et al., 2017; Heino & Grönroos, 2017; Legendre & De Cáceres, 2013), although other studies observed both negative and positive relationships at different sites (Kong et al., 2017) or quadrats (Yao et

al., 2021). Therefore, this relationship should still be investigated, especially at broad continuous scales, where LCBD indices have not yet been used. Total beta diversity increases with spatial extent (Barton et al., 2013) and is strongly dependent on scale, notably because of higher environmental heterogeneity and sampling of different local species pool (Heino et al., 2015), which could potentially add some variation to the relationship. Neither the previous studies at broad spatial extents (Poisot et al., 2017; Taranu et al., 2020; Yang et al., 2015), on spatially continuous data (D’Antraccoli et al., 2020; Tan et al., 2019), or on a high number of sites (Niskanen et al., 2017) have specifically measured the variations of the richness-LCBD relationship according to different regions and spatial extents. These studies brought forward relevant elements which now need to be combined.

This study shows that species distribution modelling offers relevant LCBD and community-level predictions on broad spatial scales, similar to those obtained from occurrence data and providing uniqueness assessments in poorly sampled regions. Our results further highlight a changing relationship between site richness and LCBD values depending on (i) the region on which it is used, as species-poor and species-rich regions display different uniqueness profiles; and on (ii) the scale at which it is applied, as increasing the spatial extent can merge the uniqueness profiles of contrasting subregions to create a new, distinct one at a broader scale. Hence, our method could prove useful to identify beta diversity hotspots in unsampled locations on large spatial scales, which could be important targets for conservation purposes.

Bibliographie

- Barton, P. S., Cunningham, S. A., Manning, A. D., Gibb, H., Lindenmayer, D. B. et Didham, R. K. (2013). The Spatial Scaling of Beta Diversity. *Global Ecology and Biogeography*, 22(6), 639-647. <https://doi.org/10.1111/geb.12031>
- Carlson, C. J. (2020). Embarcadero : Species Distribution Modelling with Bayesian Additive Regression Trees in R. *Methods in Ecology and Evolution*, 11(7), 850-858. <https://doi.org/10.1111/2041-210X.13389>
- Chipman, H. A., George, E. I. et McCulloch, R. E. (2010). BART : Bayesian Additive Regression Trees. *Annals of Applied Statistics*, 4(1), 266-298. <https://doi.org/10.1214/09-AOAS285>
- da Silva, P. G. et Hernández, M. I. M. (2014). Local and Regional Effects on Community Structure of Dung Beetles in a Mainland-Island Scenario. *PLOS ONE*, 9(10), e111883. <https://doi.org/10.1371/journal.pone.0111883>
- D'Antraccoli, M., Bacaro, G., Tordoni, E., Bedini, G. et Peruzzi, L. (2020). More Species, Less Effort : Designing and Comparing Sampling Strategies to Draft Optimised Floristic Inventories. *Perspectives in Plant Ecology, Evolution and Systematics*, 45, 125547. <https://doi.org/10.1016/j.ppees.2020.125547>
- Ferrier, S., Drielsma, M., Manion, G. et Watson, G. (2002). Extended Statistical Approaches to Modelling Spatial Pattern in Biodiversity in Northeast New South Wales. II. Community-Level Modelling. *Biodiversity & Conservation*, 11(12), 2309-2338. <https://doi.org/10.1023/A:1021374009951>
- Ferrier, S. et Guisan, A. (2006). Spatial Modelling of Biodiversity at the Community Level. *Journal of Applied Ecology*, 43(3), 393-404. <https://doi.org/10.1111/j.1365-2664.2006.01149.x>

- Guisan, A. et Rahbek, C. (2011). SESAM – a New Framework Integrating Macroecological and Species Distribution Models for Predicting Spatio-Temporal Patterns of Species Assemblages. *Journal of Biogeography*, 38(8), 1433-1444. <https://doi.org/10.1111/j.1365-2699.2011.02550.x>
- Guisan, A. et Thuiller, W. (2005). Predicting Species Distribution : Offering More than Simple Habitat Models. *Ecology Letters*, 8(9), 993-1009. <https://doi.org/10.1111/j.1461-0248.2005.00792.x>
- Heino, J., Bini, L. M., Andersson, J., Bergsten, J., Bjelke, U. et Johansson, F. (2017). Unravelling the Correlates of Species Richness and Ecological Uniqueness in a Metacommunity of Urban Pond Insects. *Ecological Indicators*, 73, 422-431. <https://doi.org/10.1016/j.ecolind.2016.10.006>
- Heino, J. et Grönroos, M. (2017). Exploring Species and Site Contributions to Beta Diversity in Stream Insect Assemblages. *Oecologia*, 183(1), 151-160. <https://doi.org/10.1007/s00442-016-3754-7>
- Heino, J., Melo, A. S., Bini, L. M., Altermatt, F., Al-Shami, S. A., Angeler, D. G., Bonada, N., Brand, C., Callisto, M., Cottenie, K., Dangles, O., Dudgeon, D., Encalada, A., Göthe, E., Grönroos, M., Hamada, N., Jacobsen, D., Landeiro, V. L., Ligeiro, R., ... Et Townsend, C. R. (2015). A Comparative Analysis Reveals Weak Relationships between Ecological Factors and Beta Diversity of Stream Insect Metacommunities at Two Spatial Levels. *Ecology and Evolution*, 5(6), 1235-1248. <https://doi.org/10.1002/ece3.1439>
- Johnston, A., Hochachka, W. M., Strimas-Mackey, M. E., Gutierrez, V. R., Robinson, O. J., Miller, E. T., Auer, T., Kelling, S. T. et Fink, D. (2020). Analytical Guidelines to Increase the Value of Citizen Science Data : Using eBird Data to Estimate Species Occurrence. *bioRxiv*, 574392. <https://doi.org/10.1101/574392>
- Kong, H., Chevalier, M., Laffaille, P. et Lek, S. (2017). Spatio-Temporal Variation of Fish Taxonomic Composition in a South-East Asian Flood-Pulse System. *PLOS ONE*, 12(3), e0174582. <https://doi.org/10.1371/journal.pone.0174582>

- Legendre, P., Borcard, D. et Peres-Neto, P. R. (2005). Analyzing Beta Diversity : Partitioning the Spatial Variation of Community Composition Data. *Ecological Monographs*, 75(4), 435-450. <https://doi.org/10.1890/05-0549>
- Legendre, P. et Condit, R. (2019). Spatial and Temporal Analysis of Beta Diversity in the Barro Colorado Island Forest Dynamics Plot, Panama. *Forest Ecosystems*, 6(1), 7. <https://doi.org/10.1186/s40663-019-0164-4>
- Legendre, P. et De Cáceres, M. (2013). Beta Diversity as the Variance of Community Data : Dissimilarity Coefficients and Partitioning. *Ecology Letters*, 16(8), 951-963. <https://doi.org/10.1111/ele.12141>
- Niskanen, A. K. J., Heikkinen, R. K., Väre, H. et Luoto, M. (2017). Drivers of High-Latitude Plant Diversity Hotspots and Their Congruence. *Biological Conservation*, 212, 288-299. <https://doi.org/10.1016/j.biocon.2017.06.019>
- Norberg, A., Abrego, N., Blanchet, F. G., Adler, F. R., Anderson, B. J., Anttila, J., Araújo, M. B., Dallas, T., Dunson, D., Elith, J., Foster, S. D., Fox, R., Franklin, J., Godsoe, W., Guisan, A., O'Hara, B., Hill, N. A., Holt, R. D., Hui, F. K. C., ... Et Ovaskainen, O. (2019). A Comprehensive Evaluation of Predictive Performance of 33 Species Distribution Models at Species and Community Levels. *Ecological Monographs*, 89(3), e01370. <https://doi.org/10.1002/ecm.1370>
- Ovaskainen, O., Tikhonov, G., Norberg, A., Blanchet, F. G., Duan, L., Dunson, D., Roslin, T. et Abrego, N. (2017). How to Make More out of Community Data ? A Conceptual Framework and Its Implementation as Models and Software. *Ecology Letters*, 20(5), 561-576. <https://doi.org/10.1111/ele.12757>
- Poisot, T., Guéveneux-Julien, C., Fortin, M.-J., Gravel, D. et Legendre, P. (2017). Hosts, Parasites and Their Interactions Respond to Different Climatic Variables. *Global Ecology and Biogeography*, 26(8), 942-951. <https://doi.org/10.1111/geb.12602>
- Poisot, T., LaBrie, R., Larson, E., Rahlin, A. et Simmons, B. I. (2019). Data-Based, Synthesis-Driven : Setting the Agenda for Computational Ecology. *Ideas in Ecology and Evolution*, 12. <https://doi.org/10.24908/iee.2019.12.2.e>

- Pollock, L. J., Tingley, R., Morris, W. K., Golding, N., O'Hara, R. B., Parris, K. M., Vesk, P. A. et McCarthy, M. A. (2014). Understanding Co-Occurrence by Modelling Species Simultaneously with a Joint Species Distribution Model (JSDM). *Methods in Ecology and Evolution*, 5(5), 397-406. <https://doi.org/10.1111/2041-210X.12180>
- Tan, L., Fan, C., Zhang, C., von Gadow, K. et Fan, X. (2017). How Beta Diversity and the Underlying Causes Vary with Sampling Scales in the Changbai Mountain Forests. *Ecology and Evolution*, 7(23), 10116-10123. <https://doi.org/10.1002/ece3.3493>
- Tan, L., Fan, C., Zhang, C. et Zhao, X. (2019). Understanding and Protecting Forest Biodiversity in Relation to Species and Local Contributions to Beta Diversity. *European Journal of Forest Research*, 138(6), 1005-1013. <https://doi.org/10.1007/s10342-019-01220-3>
- Taranu, Z. E., Pinel-Alloul, B. et Legendre, P. (2020). Large-Scale Multi-Trophic Co-Response Models and Environmental Control of Pelagic Food Webs in Québec Lakes. *Oikos*, n/a(n/a). <https://doi.org/10.1111/oik.07685>
- Yang, J., Sorte, F. A. L., Pyšek, P., Yan, P., Nowak, D. et McBride, J. (2015). The Compositional Similarity of Urban Forests among the World's Cities Is Scale Dependent. *Global Ecology and Biogeography*, 24(12), 1413-1423. <https://doi.org/10.1111/geb.12376>
- Yao, J., Huang, J., Ding, Y., Xu, Y., Xu, H. et Zang, R. (2021). Ecological Uniqueness of Species Assemblages and Their Determinants in Forest Communities. *Diversity and Distributions*, 27(3), 454-462. <https://doi.org/10.1111/ddi.13205>
- Zurell, D., Zimmermann, N. E., Gross, H., Baltensweiler, A., Sattler, T. et Wüest, R. O. (2020). Testing Species Assemblage Predictions from Stacked and Joint Species Distribution Models. *Journal of Biogeography*, 47(1), 101-113. <https://doi.org/10.1111/jbi.13608>