

Rapport BIO500

Éloïse Paquette, Danaé Vaillancourt, Élodie Michel et Camille Breton

This manuscript was compiled on April 22, 2025

Résumé

À l'aide de R Studio et du langage SQL, des figures ont été créées à partir d'une base de données sur les lépidoptères afin d'analyser l'effet des variations temporelles et spatiales sur leurs communautés. Après plusieurs analyses, on remarque une augmentation en richesse spécifique au fil des générations, soit très forte après les années 2000 mais très variable en deçà, probablement par une fluctuation d'effort d'échantillonnage ou par plusieurs pressions écologiques. Il y a toutefois une forte concentration de la richesse spécifique dans le sud de la province du Québec et plus faible au nord, une hétérogénéité expliquée par les gradients d'altitude et de climat. Enfin, la diversité des espèces dans les dates d'arrivée et de départ suggère une niche temporelle propre à chacune, aidant à la réduction de la compétition interspécifique pour les ressources.

Introduction

Au sein d'une communauté et d'un territoire, les espèces s'adaptent aux nouvelles conditions et aux changements environnementaux sur une échelle de temps variable. Les interactions entre les espèces peuvent alors modifier la composition de la communauté et sa diversité (Nieto-Sánchez et al., 2015). Cependant au sein de la famille des lépidoptères, comment les variations spatiales et temporelles influencent-elles la structure de leurs communautés ? Est-ce que certaines espèces seront impactées plus que d'autres ?

Méthode

La base de données des lépidoptères est une liste d'observations de différentes espèces de papillons, comprenant leur nom scientifique, la date, le lieu où l'échantillon a été prélevé à l'aide de coordonnées géographiques (latitude et longitude), l'heure, le jour et l'année d'observation ainsi que tous les droits d'auteur, comprenant le créateur, le titre, les licences, la source originale et le éditeur. Cette dernière comprenant environ 441 891 observations et ce, sur un intervalle de données allant de 1859 à 2023.

Afin de pouvoir utiliser cette base de données, nous avons procédé à une revalidation en éliminant et triant les lacunes humaines potentielles dans l'enregistrement des données à l'aide du logiciel R studio. Une variété de tests a été exécuté pour parvenir à l'obtention de données valides et d'une reproductibilité dans nos résultats. Des erreurs non négligeables, comme des NA non-expliquer, des échantillons répétés, des valeurs irréels dans les dates, les années, les heures et même les coordonnées, devant être filtrer pour être des points seulement au Québec, ont été corrigés et l'identification des bons types par colonnes ont été modifiés.

Ensuite, une fois la base de données considérée comme valide, nous avons pu procéder à la création de tables, avec

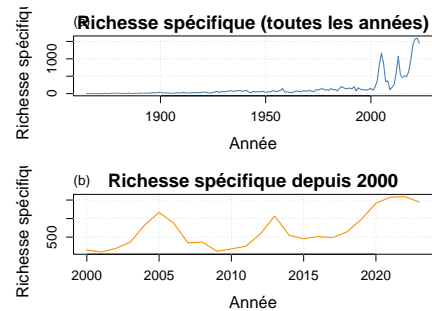


Fig. 1. Richesse spécifique au fil des années.

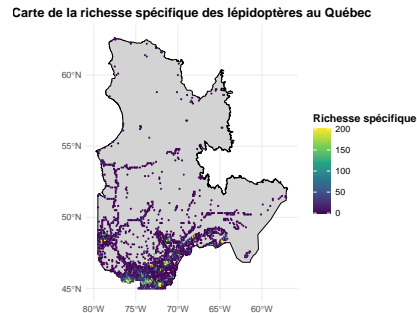


Fig. 2. Richesse spatiale

l'aide du langage informatique SQL, en regroupant les colonnes pertinentes ensembles, créer les liens entre ces dernières et ainsi optimiser l'analyse de nos données. En faisant des requêtes dans nos différentes tables, nous avons alors été en mesure de créer des figures nous aidant à répondre à notre question de départ. Finalement, une fois tous nos manipulations insérées dans des scripts différents dans le logiciel R, nous avons procédé à l'automatisation de notre code à l'aide du package Target et ainsi créer une chaîne de traitement permettant à nos résultats d'être reproductibles.

Résultats & Discussion

Discussion figure 1

```
library(dplyr)
library(ggplot2)
library(lubridate)
```

```
n_especes <- 10
```

Significance Statement

```
pheno <- tar_read(fig_phenologie)
```

```
pheno <- pheno %>%
  mutate(date_obs = as.Date(date_obs),
         jour_annee = yday(date_obs),
         annee = year(date_obs))
```

```
summary_data <- pheno %>%
  group_by(taxa_name, annee) %>%
  summarise(start_day = min(jour_annee), end_day = max(jour_annee),
            group_by(taxa_name) %>%
  summarise(start_jour_moyen = round(mean(start_day)),
            end_jour_moyen = round(mean(end_day)),
            presence_days = end_jour_moyen - start_jour_moyen,
            .groups = "drop") %>%
  mutate(start_date = as.Date(start_jour_moyen, origin = "1993-01-01"),
         end_date = as.Date(end_jour_moyen, origin = "2023-01-01")) %>%
  arrange(desc(presence_days)) %>%
  slice_head(n = n_especes)
```

```
figure_pheno <- ggplot(summary_data, aes(y = reorder(taxa_name, presence_days), x = start_date, xend = end_date)) +
  geom_segment(aes(yend = taxa_name), color = "#8888aa", size = 1) +
  geom_point(aes(x = start_date), color = "#1f78b4", size = 4) +
  geom_point(aes(x = end_date), color = "#e66101", size = 4) +
  geom_text(aes(x = start_date, label = format(start_date, "%d %b")),
            vjust = 2, color = "#1f78b4", fontface = "bold", size = 3.5) +
  geom_text(aes(x = end_date, label = format(end_date, "%d %b")),
            vjust = 2, color = "#e66101", fontface = "bold", size = 3.5) +
  annotate("rect", xmin = max(summary_data$end_date) - 10, xmax = max(summary_data$end_date) + 10,
            ymin = 0.5, ymax = nrow(summary_data) + 0.5, fill = "gray50", alpha = 0.8) +
  annotate("text", x = max(summary_data$end_date) + 25, y = nrow(summary_data) + 1,
            label = "Jours de présence", color = "white", fontface = "bold", size = 5) +
  geom_text(aes(x = max(summary_data$end_date) + 25, y = 0.5, label = presence_days), color = "white",
            fontface = "bold", size = 5) +
  annotate("text", x = max(summary_data$end_date) + 10, y = nrow(summary_data) + 1,
            label = "Jours de présence", color = "gray50", fontface = "italic", size = 5) +
  annotate("point", x = min(summary_data$start_date), color = "#1f78b4", size = 4) +
  annotate("text", x = min(summary_data$start_date), color = "#1f78b4", fontface = "bold", size = 4) +
  annotate("point", x = min(summary_data$start_date), color = "#e66101", size = 4) +
  annotate("text", x = min(summary_data$start_date), color = "#e66101", fontface = "bold", size = 4) +
  annotate("text", x = min(summary_data$start_date), label = "Départ moyen", hjust = 0, color = "white",
            fontface = "bold", size = 5) +
  expand_limits(y = nrow(summary_data) + 1.5) +
  scale_x_date(date_labels = "%B", date_breaks = "1 month") +
  labs(y = NULL, x = NULL, title = paste0("Phénologie moyenne des ", n_especes, " espèces les plus observées")) +
  theme_minimal(base_size = 16) +
  theme(axis.text.x = element_text(angle = 45, hjust = 1),
        axis.text.y = element_text(margin = margin(r = 10)),
        plot.title = element_text(size = 18, face = "bold", hjust = 0.5) +
        coord_cartesian(clip = "off") +
  theme_minimal(base_size = 16) +
  theme(
    axis.text.x = element_text(angle = 30, hjust = 1),
    plot.title = element_text(size = 18, face = "bold", hjust = 0.5),
    plot.margin = margin(1, 2, 1, 2, "cm")
  ))
```

```
# Afficher la figure
figure_pheno
```

La richesse spécifique observée à travers le temps révèle une dynamique marquée par une accélération récente. Pendant plus d'un siècle, la richesse observée est demeurée faible et relativement constante, ce qui pourrait s'expliquer par un effort d'échantillonnage limité ou des lacunes dans l'enregistrement des données (Bowler et al., 2025). Ce n'est qu'à partir des années 2000 que l'on observe une augmentation brutale du nombre d'espèces recensées, comme le montrent la figure 1. Cette tendance peut être liée à l'émergence de nouvelles technologies de collecte de données (ex. plateformes participatives comme iNaturalist ou eButterfly), à une mobilisation accrue des citoyens scientifiques, ainsi qu'à une meilleure accessibilité des données d'identification. Ces facteurs peuvent conduire à une augmentation apparente de la richesse spécifique, sans nécessairement refléter une transformation écologique réelle des communautés (Bowler et al., 2025).

Même si cette figure ne montre pas directement les variations spatiales, la montée récente de la richesse spécifique pourrait aussi être le reflet d'un élargissement spatial des efforts de surveillance. L'intégration de nouveaux sites ou régions dans les bases de données pourrait contribuer à accroître la richesse observée à l'échelle régionale. La figure (b), centrée sur les données postérieures à l'an 2000, révèle une forte variabilité interannuelle de la richesse spécifique avec des pics marqués autour de 2005, 2013 et 2021. Cette variabilité pourrait être liée à des fluctuations dans l'effort d'échantillonnage, mais elle pourrait aussi refléter des changements écologiques réels, notamment en lien avec les effets du changement climatique sur la phénologie et la distribution des espèces (Menéndez et al., 2006). L'augmentation progressive de la richesse observée vers 2020-2022 pourrait également coïncider avec une hausse de la participation citoyenne à la collecte de données, amplifiée par des plateformes comme iNaturalist. Ces résultats soulignent l'importance de considérer les facteurs à la fois écologiques et méthodologiques lorsqu'on interprète des tendances temporelles. Enfin, l'analyse temporelle longue permet de constater que la structure des communautés n'est pas stable dans le temps, et qu'elle est soumise à de multiples pressions, à la fois écologiques et méthodologiques. Ces observations rappellent l'importance d'interpréter avec prudence les tendances en biodiversité, surtout lorsque les données proviennent de sources hétérogènes ou de périodes très contrastées en termes d'effort d'observation (Bowler et al., 2025).

Discussion figure 2

Mettre la figure 2
La carte de la richesse spécifique des lépidoptères au Québec illustre de façon marquée les variations spatiales dans la composition des communautés. On remarque une très forte concentration de la richesse spécifique dans le sud de la province. Ces zones coïncident largement avec les régions les plus densément peuplées, les plus facilement accessibles et les plus transformées par l'activité humaine (Québec, 2021). En revanche, la richesse spécifique diminue de manière marquée en remontant vers le nord.

Plusieurs facteurs peuvent expliquer cette hétérogénéité. Sur le plan écologique, les gradients de latitude, d'altitude

et de climat influencent la diversité des espèces : les milieux plus chauds et plus diversifiés du sud offrent des conditions favorables à une plus grande variété de lépidoptères (White et Kerr, 2006). De plus, les régions du sud du Québec abritent une mosaïque d'habitats (zones agricoles, forêts mixtes, milieux humides) propices à l'établissement d'un plus grand nombre d'espèces.

Cette variation spatiale entraîne des structures de communautés très contrastées entre les régions : dans le sud, la compétition entre espèces peut être plus intense, avec des dynamiques de partage de niche et de spécialisation plus complexes. À l'inverse, dans le nord, les communautés sont probablement dominées par un nombre restreint d'espèces généralistes, capables de tolérer des conditions plus rudes et de coloniser des habitats plus homogènes.

Il faut cependant souligner que cette carte peut aussi refléter des biais d'échantillonnage. Les zones plus accessibles ont historiquement été davantage échantillonnées, ce qui peut amplifier artificiellement la richesse apparente dans le sud. Il est donc important d'interpréter ces résultats avec prudence et de compléter les analyses avec des données d'effort d'échantillonnage si disponibles. En somme, la variation spatiale influence profondément la structure des communautés de lépidoptères au Québec, en déterminant à la fois la richesse spécifique, la composition des espèces et les interactions écologiques entre elles.

Discussion figure 3

Mettre la figure 3

La figure 3 met en évidence la variabilité temporelle de la présence des espèces de lépidoptères les plus observées au cours de l'année. Cette diversité dans les dates d'arrivée et de départ suggère une niche temporelle propre à chaque espèce, réduisant potentiellement la compétition interspécifique pour les ressources (Ziv et Smallwood, 2000). Certaines espèces, comme *Plodia interpunctella*, sont actives dès le mois de janvier, alors que d'autres, telles que *Acleris maccana*, n'apparaissent qu'en juin.

On remarque également une forte variabilité dans la durée de présence, allant de 127 jours pour certaines espèces à 243 jours pour *Coleotechnites thujaella*. Cette disparité reflète probablement des différences dans les stratégies de vie, les tolérances climatiques et les cycles biologiques. En effet, la température est un facteur crucial pour les ectothermes. Chez les insectes, comme les lépidoptères, qui, par leur durée de développement généralement courte, vont avoir plusieurs générations au cours de la saison, chacune soumise à des conditions climatiques différentes (Gibert, 2012). Avec une hausse de la température, les espèces moins plastiques, vont plutôt changer leur aire de répartition en colonisant ou en abandonnant des habitats qui sont devenus favorables ou défavorables. Tandis que les espèces à longue durée de présence vont plutôt répondre à ces changements en modifiant la phénologie de leur cycle de vie ou bien par une adaptation au milieu par plasticité phénotypique. Ces différentes techniques dépendront de l'importance du changement, l'échelle de temps considéré et les traits d'histoire de vie des organismes (Gibert, 2012).

Les espèces présentant sur une longue durée pourraient alors jouer un rôle structurant plus constant dans la communauté, alors que les espèces plus éphémères pourraient avoir un impact plus ponctuel mais important durant leur pic

d'abondance. Les variations temporelles illustrées influencent donc directement la structure des communautés, en affectant les interactions écologiques (comme la compétition, la prédation ou la pollinisation) et en modulant la diversité observée au fil des saisons.

Conclusion

À la lumière des résultats obtenus, il est clair que les variations spatiales et temporelles au fil des générations influencent la structure des communautés de lépidoptères, en modifiant la dynamique de concentration en richesse spécifique dans les différentes régions du Québec ainsi que le temps de présence des espèces durant une année. Ces variations s'expliquent majoritairement par les pressions écologiques et environnementales, mais aussi par les fluctuations ou le manque d'effort d'échantillonnage lors de la collecte de données par les chercheurs. Il demeure que la réalité actuelle ne peut être ignorée : les hausses de température, l'urbanisation croissante et la destruction des habitats exigent une adaptation rapide des espèces (Gibert, 2012). Mais, seront-elles capables de s'adapter à temps?

Bibliographie

#Référence

- Bowler, D. E., Boyd, R. J., Callaghan, C. T., Robinson, R. A., Isaac, N. J. B. et Pocock, M. J. O. (2025). Treating gaps and biases in biodiversity data as a missing data problem. *Biological Reviews*, 100(1), 50-67. <https://doi.org/10.1111/brv.13127>
- Gibert, P. (2012). Plasticité phénotypique et réponses adaptatives aux changements environnementaux chez les insectes. *Université Claude Bernard Lyon 1*,. <https://cnrs.hal.science/tel-02309704/document>
- Menéndez, R., Megías, A. G., Hill, J. K., Braschler, B., Willis, S. G., Collingham, Y., Fox, R., Roy, D. B. et Thomas, C. D. (2006). Species richness changes lag behind climate change. *Proceedings of the Royal Society B: Biological Sciences*, 273(1593), 1465-1470. <https://doi.org/10.1098/rspb.2006.3484>
- Nieto-Sánchez, S., Gutiérrez, D. et Wilson, R. J. (2015). Long-term change and spatial variation in butterfly communities over an elevational gradient: driven by climate, buffered by habitat. *Diversity and Distributions*, 21(8), 950-961. <https://doi.org/10.1111/ddi.12316>
- Québec, I. de la statistique du. (2021). Bulletin sociodémographique, 25(2).
- White, P. et Kerr, J. T. (2006). Contrasting spatial and temporal global change impacts on butterfly species richness during the 20th century. *Ecography*, 29(6), 908-918. <https://doi.org/10.1111/j.2006.0906-7590.04685.x>
- Ziv, Y. et Smallwood, J. A. (2000). Gerbils and Heteromyids — Interspecific Competition and the Spatio-Temporal Niche. In S. Halle et N. C. Stenseth (dir.), *Activity Patterns in Small Mammals: An Ecological Approach* (p. 159-176). Springer. https://doi.org/10.1007/978-3-642-18264-8_11