

Partición de la variación en comunidades de ácaros: influencia ambiental y espacial

Lisanny de la Rosa (Matrícula 100675664)

2026-02-09

Contents

Resumen (Abstract)	1
1. Introducción	2
1.2. Problema y justificación	2
1.3. Recorrido teórico mínimo exigido (checklist)	2
2. Materiales y métodos	3
2.1. Datos	3
2.2. Preparación de datos (decisiones analíticas)	4
2.3. Técnica principal (núcleo del ejercicio)	4
2.4. Verificación mínima de supuestos / diagnósticos	4
2.5. Reproducibilidad	4
3. Resultados	6
3.1. Resultados del AED	6
3.2. Resultados de la técnica principal	6
3.3. Resultados complementarios (opcionales)	6
4. Discusión	6
4.1. Respuesta directa a preguntas/hipótesis	6
4.2. Interpretación conectada con teoría	6
4.3. Implicaciones biogeográficas/ecológicas (y conservación si aplica)	6
5. Aporte del trabajo	6
6. Limitaciones	6
7. Conclusiones	6
Referencias	7
Apéndice A. Bitácora de decisiones (incluye IA)	7
A.1. Decisiones analíticas	7
A.2. Uso de IA (si aplica)	7

Resumen (Abstract)

Escribe 150–250 palabras. Incluye: (i) problema, (ii) datos, (iii) técnica principal, (iv) 1–2 hallazgos, (v) conclusión principal y 1 limitación.

Palabras clave: palabra1; palabra2; palabra3; ...

1. Introducción

El estudio de comunidades biológicas en diferentes ambientes permite comprender cómo factores ambientales y espaciales influyen en la distribución de las especies. En ecología numérica, los análisis de partición de varianza son herramientas útiles para identificar qué proporción de la variación en la composición de especies se debe a condiciones ambientales específicas y cuál responde a la estructura espacial del área de muestreo (Borcard, Legendre & Drapeau, 1992).

En la República Dominicana, los ecosistemas de musgos presentes en zonas húmedas de la Cordillera Central y en áreas montañosas como Valle Nuevo ofrecen un hábitat ideal para pequeños organismos como los ácaros. Estos microartrópodos cumplen funciones importantes en la descomposición y el reciclaje de nutrientes, por lo que entender su distribución resulta clave para valorar la salud de los ecosistemas (Borcard & Legendre, 1994).

Al aplicar técnicas como la transformación de Hellinger y la partición de varianza, se puede evaluar cómo variables como la densidad del sustrato, el contenido de agua y el tipo de sustrato, junto con la ubicación espacial de las muestras, explican la composición de las comunidades de ácaros. Este enfoque permite no solo interpretar patrones ecológicos locales, sino también generar información que puede ser extrapolada a otros ambientes tropicales de la isla (Borcard, Gillet & Legendre, 2011; Borcard & Legendre, 2002).

1.2. Problema y justificación

El análisis de comunidades biológicas enfrenta el reto de distinguir entre los efectos ambientales y los efectos espaciales en la distribución de las especies. En el caso de los ácaros que habitan en musgos, resulta complejo determinar si su presencia y abundancia se deben principalmente a condiciones del sustrato (densidad, humedad, tipo) o a la proximidad espacial entre sitios de muestreo. Este problema es relevante porque la falta de claridad puede llevar a interpretaciones erróneas sobre los factores que realmente controlan la biodiversidad en un ecosistema.

En la República Dominicana, los ecosistemas montañosos como los de Valle Nuevo y la Cordillera Central son espacios de alta diversidad biológica y endemismo. Sin embargo, aún existe poca información sobre la dinámica de comunidades de microartrópodos en estos ambientes. La justificación de este estudio radica en que, al aplicar métodos de partición de varianza, se puede identificar qué proporción de la variación en la composición de especies de ácaros se explica por factores ambientales y cuál por la estructura espacial. Esto permite comprender mejor los procesos ecológicos que sostienen la biodiversidad en ambientes tropicales húmedos y aporta herramientas para la conservación y manejo de estos ecosistemas.

Además, el uso de datasets como *mite*, ampliamente documentados en la literatura científica (Borcard, Legendre & Drapeau, 1992; Borcard & Legendre, 1994; Borcard & Legendre, 2002; Borcard, Gillet & Legendre, 2011), ofrece un marco metodológico sólido que puede ser replicado en estudios locales. De esta manera, se justifica la práctica como un ejercicio que combina teoría ecológica con aplicaciones estadísticas modernas, contribuyendo a la formación en ecología numérica y a la generación de conocimiento aplicable en la República Dominicana.

1.3. Recorrido teórico mínimo exigido (checklist)

AED/EDA

Se realizó una exploración inicial de los datos de abundancia de ácaros en musgos (*mite*), verificando la distribución de especies y posibles valores extremos. Cada fila representa un sitio de muestreo y cada columna una especie, lo que permite identificar patrones de abundancia y rareza.

Transformaciones

Se aplicó la transformación de **Hellinger** a la matriz de abundancias, lo cual reduce el peso de especies muy dominantes y facilita el uso de técnicas basadas en distancias euclidianas.

Distancias y justificación

La transformación Hellinger permite trabajar con la **distancia euclidiana**, que es adecuada para análisis de partición de varianza y métodos lineales. En otros contextos, Bray–Curtis o Jaccard serían alternativas, pero aquí la elección se justifica por la compatibilidad con el método.

Técnica principal

La técnica principal es la **partición de varianza (varpart)**, que descompone la variación en la comunidad en fracciones explicadas por variables ambientales y espaciales.

Limitaciones de la técnica

- Supone que las variables ambientales y espaciales seleccionadas son representativas.
- No capta procesos biológicos no medidos (ej. interacciones entre especies).
- La fracción no explicada puede ser alta, reflejando limitaciones en los datos.

Interpretación

La interpretación se centra en los diagramas de partición de varianza, donde se visualizan las fracciones explicadas por ambiente, espacio y su intersección. No se utilizan biplots en este caso, pero sí diagramas de Venn.

Escala y muestreo

Cada muestra corresponde a un sitio de musgo muestreado, con coordenadas espaciales (x, y) y variables ambientales asociadas.

Implicaciones biogeográficas/ecológicas

El análisis permite inferir si la distribución de ácaros está más determinada por condiciones locales (densidad, humedad, tipo de sustrato) o por la proximidad espacial. Esto aporta información sobre procesos de dispersión y control ambiental en comunidades microartrópodos.

Reproducibilidad

El análisis es reproducible en R mediante el paquete **vegan**, con un pipeline que incluye: carga de datos, transformación Hellinger, construcción de polinomios espaciales y aplicación de **varpart**. Se recomienda fijar semillas (`set.seed()`) para garantizar replicabilidad en pasos aleatorios.

2. Materiales y métodos

2.1. Datos

2.1.1. Dataset y origen

En este ejercicio se utiliza el dataset mite, acompañado de sus tablas asociadas mite.env y mite.xy, disponibles en el paquete vegan de R.

mite: matriz de abundancias de especies de ácaros recolectados en muestras de musgo.

mite.env: variables ambientales medidas en los mismos sitios de muestreo (densidad del sustrato, contenido de agua y tipo de sustrato).

mite.xy: coordenadas espaciales de los sitios de muestreo, que permiten analizar patrones espaciales.

Este conjunto de datos fue recolectado originalmente por Daniel Borcard (1989) y ha sido ampliamente utilizado en estudios de ecología numérica y análisis de comunidades, especialmente en la demostración de técnicas de partición de varianza y análisis espacial.

```
#>
#> Partition of variance in RDA
#>
#> Call: varpart(Y = Y, X = mite.env[, c("SubsDens", "WatrCont",
#> "Substrate")], S)
#>
#> Explanatory tables:
```

```

#> X1:  mite.env[, c("SubsDens", "WatrCont", "Substrate")]
#> X2:  S
#>
#> No. of explanatory tables: 2
#> Total variation (SS): 27.205
#>           Variance: 0.39428
#> No. of observations: 70
#>
#> Partition table:
#>
#>           Df R.squared Adj.R.squared Testable
#> [a+c] = X1      8  0.42839      0.35343     TRUE
#> [b+c] = X2      5  0.35495      0.30455     TRUE
#> [a+b+c] = X1+X2 13  0.54761      0.44259     TRUE
#> Individual fractions
#> [a] = X1|X2      8              0.13804     TRUE
#> [b] = X2|X1      5              0.08916     TRUE
#> [c]              0              0.21539     FALSE
#> [d] = Residuals              0.55741     FALSE
#> ---
#> Use function 'rda' to test significance of fractions of interest

```

2.1.2. Diseño de muestreo y variables

Describe: unidad muestral, variables biológicas (respuesta) y ambientales (explicativas), tipos de variables, rangos, NA, etc.

2.1.3. Limitaciones del dataset

Señala limitaciones **reales** (tamaño muestral, sesgos, variables omitidas, representatividad, escalas, etc.).

2.2. Preparación de datos (decisiones analíticas)

2.2.1. Limpieza y manejo de NA

2.2.2. Transformaciones y estandarización

Explica qué transformación aplicas y **por qué**.

2.2.3. Medida(s) de distancia (si aplica)

Declara la(s) distancia(s) y justifica (tipo de datos, objetivo, supuestos).

2.3. Técnica principal (núcleo del ejercicio)

Describe la técnica, qué responde, qué representa cada salida y sus supuestos.

2.4. Verificación mínima de supuestos / diagnósticos

Aquí no se trata de “cumplir por cumplir”: explica **qué chequeas** y qué implica.

2.5. Reproducibilidad

Declara cómo garantizas reproducibilidad: semilla, sesión, orden de ejecución.

```

#> R version 4.4.0 (2024-04-24)
#> Platform: x86_64-pc-linux-gnu

```

```

#> Running under: Ubuntu 22.04.4 LTS
#>
#> Matrix products: default
#> BLAS: /usr/lib/x86_64-linux-gnu/openblas-pthread/libblas.so.3
#> LAPACK: /usr/lib/x86_64-linux-gnu/openblas-pthread/libopenblas-p-r0.3.20.so; LAPACK version 3.10.0
#>
#> locale:
#> [1] LC_CTYPE=C.UTF-8 LC_NUMERIC=C LC_TIME=C.UTF-8
#> [4] LC_COLLATE=C.UTF-8 LC_MONETARY=C.UTF-8 LC_MESSAGES=C.UTF-8
#> [7] LC_PAPER=C.UTF-8 LC_NAME=C LC_ADDRESS=C
#> [10] LC_TELEPHONE=C LC_MEASUREMENT=C.UTF-8 LC_IDENTIFICATION=C
#>
#> time zone: Etc/UTC
#> tzcode source: system (glibc)
#>
#> attached base packages:
#> [1] stats graphics grDevices utils datasets methods base
#>
#> other attached packages:
#> [1] readr_2.1.5 tidyr_1.3.1 dplyr_1.1.4
#> [4] ggplot2_3.5.2 SpadeR_0.1.1 iNEXT_3.0.1
#> [7] adespatial_0.3-23 indicpecies_1.7.14 cluster_2.1.6
#> [10] vegan_2.6-6.1 lattice_0.22-6 permute_0.9-7
#>
#> loaded via a namespace (and not attached):
#> [1] DBI_1.2.3 deldir_2.0-4 s2_1.1.6
#> [4] rlang_1.1.5 magrittr_2.0.3 ade4_1.7-22
#> [7] e1071_1.7-14 compiler_4.4.0 mgcv_1.9-1
#> [10] png_0.1-8 vctrs_0.6.5 reshape2_1.4.4
#> [13] stringr_1.5.1 pkgconfig_2.0.3 wk_0.9.1
#> [16] crayon_1.5.3 fastmap_1.2.0 utf8_1.2.4
#> [19] promises_1.3.2 rmarkdown_2.27 tzdb_0.4.0
#> [22] ps_1.8.1 purrr_1.0.4 xfun_0.53
#> [25] seqinr_4.2-36 jsonlite_1.8.9 progress_1.2.3
#> [28] later_1.4.1 adegenet_2.1.10 uuid_1.2-0
#> [31] jpeg_0.1-10 parallel_4.4.0 prettyunits_1.2.0
#> [34] R6_2.5.1 stringi_1.8.4 RColorBrewer_1.1-3
#> [37] boot_1.3-30 Rcpp_1.0.14 bookdown_0.39
#> [40] knitr_1.47 adephylo_1.1-16 httpuv_1.6.15
#> [43] Matrix_1.7-0 adegraphics_1.0-21 splines_4.4.0
#> [46] igraph_2.0.3 tidyselect_1.2.1 rstudioapi_0.17.1
#> [49] yaml_2.3.10 phylobase_0.8.12 websocket_1.4.1
#> [52] processx_3.8.5 tibble_3.2.1 plyr_1.8.9
#> [55] withr_3.0.2 shiny_1.10.0 evaluate_0.24.0
#> [58] sf_1.0-16 units_0.8-5 spData_2.3.1
#> [61] proxy_0.4-27 xml2_1.3.6 pillar_1.9.0
#> [64] KernSmooth_2.23-24 generics_0.1.3 sp_2.1-4
#> [67] chromote_0.4.0 hms_1.1.3 munsell_0.5.1
#> [70] scales_1.3.0 xtable_1.8-4 rnc1_0.8.7
#> [73] class_7.3-22 glue_1.8.0 tools_4.4.0
#> [76] interp_1.1-6 XML_3.99-0.16.1 grid_4.4.0
#> [79] spdep_1.3-5 ape_5.8 RNeXML_2.4.11
#> [82] latticeExtra_0.6-30 colorspace_2.1-0 nlme_3.1-165
#> [85] cli_3.6.3 fansi_1.0.6 gtable_0.3.5

```

```
#> [88] digest_0.6.37      classInt_0.4-10     htmltools_0.5.8.1
#> [91] lifecycle_1.0.4     httr_1.4.7          mime_0.12
#> [94] MASS_7.3-61
```

3. Resultados

Presenta resultados en el mismo orden que Métodos. Cada figura/tabla debe tener: (i) qué muestra, (ii) resultado clave, (iii) vínculo con preguntas.

3.1. Resultados del AED

3.2. Resultados de la técnica principal

3.3. Resultados complementarios (opcionales)

4. Discusión

4.1. Respuesta directa a preguntas/hipótesis

Responde P1, P2, etc., citando evidencia concreta (figuras/tablas).

4.2. Interpretación conectada con teoría

Vuelve a los conceptos del checklist y muestra que los entiendes: qué significa el patrón, qué no significa, y bajo qué supuestos.

4.3. Implicaciones biogeográficas/ecológicas (y conservación si aplica)

¿Qué sugiere tu análisis sobre procesos, filtros ambientales, dispersión, estructura espacial, heterogeneidad?

5. Aporte del trabajo

En 3–6 líneas: ¿qué aporta tu trabajo? (Puede ser metodológico: “muestra cómo...”; o conceptual: “sugiere que...”)

6. Limitaciones

Lista 3–6 limitaciones **reales**:

- Dataset (muestra, variables, escala)
- Método (supuestos, sensibilidad a parámetros, distancia, transformación)
- Generalización (qué NO puedes concluir)

7. Conclusiones

3–6 bullets o un párrafo corto. Deben ser defendibles y estar sustentadas en resultados.

Referencias

Apéndice A. Bitácora de decisiones (incluye IA)

Registra decisiones clave y evidencia de verificación. No es “para delatarte”: es para mostrar trazabilidad.

A.1. Decisiones analíticas

- Transformación elegida: ... (por qué)
- Distancia elegida: ... (por qué)
- Parámetros clave: ... (por qué)
- Alternativas probadas: ... (qué cambió)

A.2. Uso de IA (si aplica)

Si usaste IA, documenta **qué preguntaste** y **cómo verificaste**.

- Objetivo de la consulta: ...
 - Resumen de la respuesta recibida: ...
 - Cómo verifiqué (doc, ayuda de R, manual, paper, etc.): ...
 - Qué adopté y qué descarté: ...
-