

¹ MetaCommunityMetrics.jl: A Julia Package for Spatiotemporal Metacommunity Analysis

³ **Yan Yin Jenny Cheung**  ¹ and **Laura Melissa Guzman**  ¹

⁴ ¹ Department of Entomology, Cornell University, United States ¶ Corresponding author

DOI: [10.xxxxxx/draft](https://doi.org/10.xxxxxx/draft)

Software

- [Review](#) 
- [Repository](#) 
- [Archive](#) 

Editor: 

Submitted: 22 October 2025

Published: unpublished

License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC BY 4.0](#))

⁵ Summary

⁶ Community ecologists are often interested in inferring ecological processes from observational data. This involves calculating different metrics such as beta diversity. Calculating these ⁷ metrics is easy and fast due to the availability of R packages. However, community ecology today is growing beyond analysing single datasets. Large biodiversity databases have expanded ⁹ the sets of spatio-temporal community data, and there is a growing interest in using simulations ¹⁰ to study ecological processes. For these applications, R packages become slow to use.

¹² MetaCommunityMetrics.jl is a Julia package that helps ecologists to analyze large datasets ¹³ of species observations to understand how community composition varies across space and time. By implementing computationally optimized versions of established software and new ¹⁴ complementary functions that are designed for handling large spatiotemporal datasets, this ¹⁵ package enables ecologists to efficiently analyze large-scale spatiotemporal community datasets ¹⁶ that would be computationally prohibitive with existing tools. We designed intuitive interfaces ¹⁷ that would feel familiar to ecologists transitioning from R. Our function naming conventions ¹⁸ closely match their R equivalents, and we maintained parameter naming and ordering consistency ¹⁹ where possible. Some of the metrics we have implemented include beta diversity decompositions ²⁰ in space and time, dispersal-niche continuum index, and variability metric, but a full list of ²¹ metrics is listed in Table 1. Full documentation and examples, including installation instructions ²² and help files, are available at <https://cralibe.github.io/MetaCommunityMetrics.jl/>.

Table 1: A summary of the metrics and functions available in MetaCommunityMetrics.jl.

Metrics	Functions	Details
Beta diversity decompositions in space and time	<code>beta_diversity()</code> <code>spatial_beta_div()</code> <code>temporal_beta_div()</code>	Re-implementation of R package adespatial(Dray et al., 2023)
Dispersal-niche continuum index (DNCI)	<code>create_groups()</code> <code>plot_groups()</code> <code>DNCI_multigroup()</code>	Re-implementation of R script from Guzman et al. (2022)
Niche overlap index Occupied patches proportion	<code>niche_overlap()</code> <code>prop_patch()</code>	A novel grouping function for DNCI analysis A novel visualization function for DNCI analysis Re-implementation of R script from R package DNCIper (Gibert, 2024) Novel implementation Re-implementation of R script from Guzman et al. (2022)

Metrics	Functions	Details
Variability Metric	CV_meta()	Re-implementation of R script from Wang et al. (2019)
Niche hypervolume measurements	MVNH_det() MVNH_dissimilarity() average_MVNH_det() average_MVNH_dissimilarity()	Re-implementation from R package MVNH(Lu et al., 2021) Re-implementation from R package MVNH(Lu et al., 2021) Novel extension of MVNH_det() Novel extension of MVNH_dissimilarity()

24 Statement of need

25 Many valuable R ([R Core Team, 2024](#)) packages have been developed to aid in community
 26 analyses by providing ecological metrics that can summarize processes from biodiversity patterns
 27 and have been widely adopted in the field, such as adespatial ([Dray et al., 2023](#)), codyn
 28 ([Hallett et al., 2020](#)), and vegan ([Oksanen et al., 2024](#)). However, conducting community
 29 analyses in R is often computationally expensive, especially when working with large datasets,
 30 a limitation that becomes critical in workflows requiring repeated metric calculations. These
 31 include simulation studies for method validation, bootstrap resampling for uncertainty estimation
 32 ([Efron, 1992](#)), null model approaches for statistical testing ([Gotelli & Graves, 1996](#)), and the
 33 analysis of large biodiversity databases, such as [BioTime](#), all of which can create significant
 34 computational bottlenecks when processing multiple large datasets iteratively. To improve
 35 the efficiency of these computationally intensive analyses, Julia becomes a logical choice
 36 for re-implementations. Julia is known for its just-in-time compilation that optimizes code
 37 execution to levels comparable with lower-level languages like C or Fortran ([Bezanson et al.,](#)
 38 [2017](#)), yet Julia remains relatively uncommon in ecological workflows.

39 State of the field

40 To position MetaCommunityMetrics.jl within the Julia ecosystem, we conducted a
 41 comprehensive review of existing Julia packages under the [EcoJulia organization](#), a
 42 community-driven effort providing tools for ecological and environmental analysis, and the
 43 [PoisotLab](#), which develops many Julia packages focusing on quantitative and computational
 44 ecology. We ensured that our package builds beyond the tools currently existing in Julia even
 45 when sharing similar ecological concepts. For example, our beta diversity metrics share the
 46 concept of beta diversity with Diversity.jl ([Robert Reeve & Harris, 2024](#)). However, our
 47 implementation decomposes beta diversity into species replacement and richness differences
 48 (or abundance difference when abundance data is used instead of occurrence data) ([Baselga,](#)
 49 [2010, 2012, 2013; Legendre, 2014](#)), which differentiates our approach from Diversity.jl,
 50 which emphasizes alpha, beta, and gamma diversity measures at the metacommunity and
 51 community levels ([Richard Reeve et al., 2014](#)). To our knowledge, no existing Julia packages
 52 implement the specific metrics listed in [Table 1](#). MetaCommunityMetrics.jl addresses this
 53 gap by providing the first Julia implementations of these metrics, which are essential for
 54 research in community ecology but were previously only available through computationally
 55 expensive R implementations.

56 Software Design

57 We built MetaCommunityMetrics.jl on three design principles: computational efficiency,
 58 usability for spatiotemporal data, and methodological validation.

59 First, we extended single-snapshot metrics (single site and/or time point) to operate on
60 multidimensional datasets (multiple sites and/or time points) via aggregation functions
61 (`spatial_beta_div()`, `temporal_beta_div()`, `average_MVNH_det()`, `average_MVNH_dissimilarity()`),
62 trading complexity for usability and eliminating manual looping.

63 Second, we prioritized performance over comprehensive coverage (compared to the R
64 implementations). Our beta diversity focuses on the Podani family approach rather than all
65 five coefficient families, `MVNH_dissimilarity` implements only Bhattacharyya distance, and
66 `DNCI` lacks built-in visualizations. We plan to expand these in future releases.

67 Third, we use vectorized operations where possible. We also employ Cholesky decomposition
68 for covariance calculations (a computational optimization over LU decomposition in the original
69 R implementation), providing substantial performance gains for positive-definite matrices
70 ([Haddad, 2024](#)).

71 Fourth, our `DNCI` implementation accepts empty sites and singletons using Zero-adjusted
72 Bray-Curtis dissimilarity, filters uninformative species (completely absent or ubiquitous),
73 and implements status flags for five edge cases rather than throwing errors or returning
74 NAs, supporting simulation workflows where sparse matrices are meaningful. We developed
75 `create_groups()` and `plot_groups()` to automate site grouping (required by Vilmi et al.
76 ([2021](#)) but manual in R) and visualize results. See [documentation](#) for details.

77 Finally, all implementations were validated against R equivalents to ensure methodological
78 consistency (see [documentation](#)).

79 Research Impact Statement

80 `MetaCommunityMetrics.jl` delivers substantial computational advantages over existing tools.
81 For example, we re-implemented the `beta.div.comp` function from the R package `adespatial`
82 ([Dray et al., 2023](#)), achieving up to $\sim 58\times$ faster execution times while maintaining
83 methodological consistency (see [Figure 1](#)). As shown in [Figure 1](#), performance gains range
84 from $\sim 3\times$ to $58\times$ across all re-implemented functions using [three testing datasets](#) (large:
85 53,352 observations; medium: 26,676 observations; small: 5,325 observations), positioning
86 `MetaCommunityMetrics.jl` as an efficient solution for diverse ecological research needs, from
87 standard biodiversity analyses to computationally intensive simulation studies and large-scale
88 data processing.

89 Further, the first empirical application of Approximate Bayesian Computation methods to
90 connect metacommunity theory with observed data (a long-term phytoplankton community
91 monitoring data) is using this package to calculate summary statistics. The manuscript of this
92 application is under preparation.

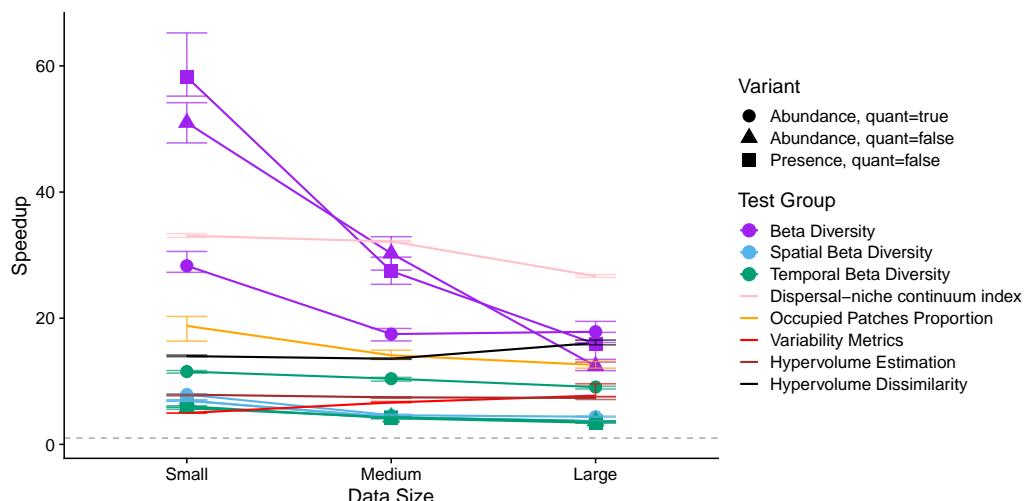


Figure 1: A plot showing the speedup of all benchmarked functions across three datasets: small (5,335 observations), medium (26,676 observations), and large (53,352 observations). Speedup is calculated as the R median execution time divided by the Julia median execution time. Median speedup and its confidence interval of each comparison is shown. The grey dashed line shows where speedup equals to 1, that is both R and Julia's implementations require the same amount of time. The calculation of beta diversity has two variants: quant = true (abundance data is used) and quant = false (occurrence data is used). Details about this benchmarking can be found in the documentation under the [benchmark results section](#).

AI usage disclosure

We used Claude, a generative AI tool, for proofreading and identifying grammatical errors in the manuscript and software documentation, and for providing optimization suggestions during software development. All writing suggestions were manually reviewed by the authors and selectively accepted. For optimization suggestions during software development, all suggestions were manually reviewed by the authors and selectively accepted, verified to produce the same results as before optimization, and benchmarked to confirm improvements in execution time and/or memory usage.

Acknowledgement

We are grateful to the authors of the pre-existing R packages/implementations that we have implemented in `MetaCommunityMetrics.jl`. We express our gratitude to everyone who has tried out this package and provided feedback on how to improve it. LMG acknowledges funding from USC and Cornell University, YYJC acknowledges funding from the Wrigley Institute.

References

- Baselga, A. (2010). Partitioning the turnover and nestedness components of beta diversity. *Global Ecology and Biogeography*, 19(1), 134–143. <https://doi.org/10.1111/j.1466-8238.2009.00490.x>
- Baselga, A. (2012). The relationship between species replacement, dissimilarity derived from nestedness, and nestedness. *Global Ecology and Biogeography*, 21(12), 1223–1232. <https://doi.org/10.1111/j.1466-8238.2011.00756.x>
- Baselga, A. (2013). Separating the two components of abundance-based dissimilarity: Balanced

- 114 changes in abundance vs. Abundance gradient. *Methods in Ecology and Evolution*, 4(6),
 115 552–557. <https://doi.org/10.1111/2041-210x.12029>
- 116 Bezanson, J., Edelman, A., Karpinski, S., & Shah, V. B. (2017). Julia: A fresh approach to
 117 numerical computing. *SIAM Review*, 59(1), 65–98. <https://doi.org/10.1137/141000671>
- 118 Dray, S., Bauman, D., Blanchet, G., Borcard, D., Clapé, S., Guénard, G., Jombart, T.,
 119 Larocque, G., Legendre, P., Madi, N., & Wagner, H. H. (2023). *Adespatial: Multivariate
 120 multiscale spatial analysis*. <https://doi.org/10.32614/cran.package.adespatial>
- 121 Efron, B. (1992). Bootstrap methods: Another look at the jackknife. In S. Kotz & N. L.
 122 Johnson (Eds.), *Breakthroughs in statistics: Methodology and distribution* (pp. 569–593).
 123 Springer New York. https://doi.org/10.1007/978-1-4612-4380-9_41
- 124 Gibert, C. (2024). *DNCImper: Assembly process identification based on SIMPER analysis*.
 125 <https://github.com/gibertcorentin/DNCImper>
- 126 Gotelli, N. J., & Graves, G. R. (1996). *Null models in ecology*. Smithsonian Institution Press.
- 127 Guzman, L. M., Thompson, P. L., Viana, D. S., Vanschoenwinkel, B., Horvath, Z., Ptacnik, R.,
 128 Jeliazkov, A., Gascón, S., Lemmens, P., Anton-Pardo, M., & others. (2022). Accounting for
 129 temporal change in multiple biodiversity patterns improves the inference of metacommunity
 130 processes. *Ecology*, 103(6), e3683. <https://doi.org/10.1002/ecy.3683>
- 131 Haddad, C. N. (2024). Cholesky factorization. In *Encyclopedia of optimization* (pp. 1–3).
 132 Springer. https://doi.org/10.1007/978-3-030-54621-2_67-1
- 133 Hallett, L., Avolio, M. L., Carroll, I. T., Jones, S. K., MacDonald, A. A. M., Flynn, D. F.,
 134 B., Slaughter, P., Ripplinger, J., Collins, S. L., Gries, C., & Jones, M. B. (2020). *Codyn:
 135 Community dynamics metrics*. <https://doi.org/10.5063/F1N877Z6>
- 136 Legendre, P. (2014). Interpreting the replacement and richness difference components of beta
 137 diversity. *Global Ecology and Biogeography*, 23(11), 1324–1334. <https://doi.org/10.1111/geb.12207>
- 138 Lu, M., Winner, K., & Jetz, W. (2021). A unifying framework for quantifying and comparing
 139 n-dimensional hypervolumes. *Methods in Ecology and Evolution*, 12(10), 1953–1968.
 140 <https://doi.org/10.1111/2041-210X.13665>
- 141 Oksanen, J., Simpson, G. L., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara,
 142 R. B., Solymos, P., Stevens, M. H. H., Szoecs, E., Wagner, H., Barbour, M., Bedward, M.,
 143 Bolker, B., Borcard, D., Carvalho, G., Chirico, M., De Caceres, M., Durand, S., ... Weedon,
 144 J. (2024). *Vegan: Community ecology package*. <https://doi.org/10.32614/CRAN.package.vegan>
- 145 R Core Team. (2024). *R: A language and environment for statistical computing*. R Foundation
 146 for Statistical Computing. <https://doi.org/10.32614/r.manuals>
- 147 Reeve, Robert, & Harris, C. (2024). *Diversity.jl (v0.5.15)*. <https://doi.org/10.5281/zenodo.12790658>
- 148 Reeve, Richard, Leinster, T., Cobbold, C. A., Thompson, J., Brummitt, N., Mitchell, S.
 149 N., & Matthews, L. (2014). How to partition diversity. *arXiv Preprint arXiv:1404.6520*.
 150 <https://doi.org/10.48550/arXiv.1404.6520>
- 151 Vilmi, A., Gibert, C., Escarguel, G., Happonen, K., Heino, J., Jamoneau, A., Passy, S. I.,
 152 Picazo, F., Soininen, J., Tison-Rosebery, J., & others. (2021). Dispersal-niche continuum
 153 index: A new quantitative metric for assessing the relative importance of dispersal versus
 154 niche processes in community assembly. *Ecography*, 44(3), 370–379. <https://doi.org/10.1111/ecog.05356>
- 155 Wang, S., Lamy, T., Hallett, L. M., & Loreau, M. (2019). Stability and synchrony across
 156 ecological hierarchies in heterogeneous metacommunities: Linking theory to data.

DRAFT