

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

Yan Yin Jenny Cheung <sup>1</sup> and Laura Melissa Guzman <sup>1</sup>✉

<sup>1</sup> Department of Entomology, Cornell University, United States ✉ Corresponding author

DOI: [10.21105/joss.10167](https://doi.org/10.21105/joss.10167)

## Software

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

Editor: [Evan Spotte-Smith](#)  

## Reviewers:

- [@FrancisBanville](#)
- [@julienmaldar](#)

Submitted: 22 October 2025

Published: 29 July 2026

## License

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

## Summary

Community ecologists are often interested in inferring ecological processes from observational data. This involves calculating different metrics such as beta diversity, which measures the dissimilarity between communities in terms of species composition or abundance. Calculating these metrics is easy and fast due to the availability of R packages, particularly for datasets with a limited number of sites and sampling events. However, community ecology today is growing beyond small-scale, single-dataset analyses. Large biodiversity databases have expanded the sets of spatiotemporal community data, and there is a growing interest in using simulations to study ecological processes. While R packages can achieve high performance through tools such as Rcpp, this requires re-implementation in C++, which presents a barrier for most ecologists. Julia offers native high performance while remaining as accessible as R, making it a more practical path to efficient computation for large-scale ecological analyses. However, broader adoption of Julia in ecology depends on having a rich ecosystem of packages, which this package aims to contribute to.

MetaCommunityMetrics.jl is a Julia package that helps ecologists analyze datasets of species observations to understand how community composition varies across space and time. By implementing computationally optimized versions of established software and new complementary functions designed for large spatiotemporal datasets, this package enables ecologists to efficiently perform analyses that would be computationally prohibitive with existing tools. We designed intuitive interfaces that would feel familiar to ecologists transitioning from R. Our function naming conventions closely match their R equivalents, and we maintained parameter naming and ordering consistency where possible. Some of the metrics we have implemented include beta diversity decompositions in space and time, dispersal-niche continuum index, and variability metrics measuring temporal variability in abundance across spatial scales and organizational levels of a metacommunity. A full list is provided in [Table 1](#). Full documentation and examples, including installation instructions 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                                                                                      | Ecological Meaning                                                                                                |
|-------------------------------------------------|---------------------------------|----------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------|
| Beta diversity decompositions in space and time | <code>beta_diversity()</code>   | Re-implementation of R package <code>adespatial</code> ( <a href="#">Dray et al., 2023</a> ) | Community dissimilarity decomposed into species replacement and richness/abundance differences                    |
|                                                 | <code>spatial_beta_div()</code> | Re-implementation of R script from Guzman et al. ( <a href="#">2022</a> )                    | Spatial variation in community composition decomposed into species replacement and richness/abundance differences |

| Metrics                                | Functions                                 | Details                                                             | Ecological Meaning                                                                                                 |
|----------------------------------------|-------------------------------------------|---------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------|
| Dispersal-niche continuum index (DNCI) | <code>temporal_beta_div()</code>          | Re-implementation of R script from Guzman et al. (2022)             | Temporal variation in community composition decomposed into species replacement and richness/abundance differences |
|                                        | <code>DNCI_create_groups()</code>         | A novel grouping function for DNCI analysis                         | Groups sites geographically using hierarchical clustering for DNCI analysis                                        |
|                                        | <code>DNCI_plot_groups()</code>           | A novel visualization function for DNCI analysis                    | Visualizes geographic site groupings for DNCI analysis                                                             |
|                                        | <code>DNCI_multigroup()</code>            | Re-implementation of R script from R package DNCIper (Gibert, 2024) | Identifies whether dispersal or niche processes drive community assembly                                           |
| Niche overlap index                    | <code>niche_overlap()</code>              | Novel implementation                                                | Mean niche overlap between all species pairs based on relative abundance distributions                             |
| Occupied patches proportion            | <code>prop_patches()</code>               | Re-implementation of R script from Guzman et al. (2022)             | Proportion of sites occupied by each species                                                                       |
| Variability Metric                     | <code>CV_meta()</code>                    | Re-implementation of R script from Wang et al. (2019)               | Temporal variability in abundance across spatial scales and organizational levels                                  |
| Niche hypervolume measurements         | <code>MVNH_det()</code>                   | Re-implementation from R package MVNH (Lu et al., 2021)             | Volume of a species' niche hypervolume                                                                             |
|                                        | <code>MVNH_dissimilarity()</code>         | Re-implementation from R package MVNH (Lu et al., 2021)             | Niche dissimilarity between two species using Bhattacharyya distance                                               |
|                                        | <code>average_MVNH_det()</code>           | Novel extension of <code>MVNH_det()</code>                          | Mean niche hypervolume across all species in a community                                                           |
|                                        | <code>average_MVNH_dissimilarity()</code> | Novel extension of <code>MVNH_dissimilarity()</code>                | Mean niche dissimilarity across all species pairs in a community                                                   |

## Statement of need

Many valuable R (R Core Team, 2024) packages have been developed to aid in community analyses by providing ecological metrics that can summarize processes from biodiversity patterns and have been widely adopted in the field, such as `adespatial` (Dray et al., 2023), `codyn` (Hallett et al., 2020), and `vegan` (Oksanen et al., 2024). However, conducting community analyses in R is often computationally expensive, especially when working with large datasets, a limitation that becomes critical in workflows requiring repeated metric calculations. These include simulation studies for method validation, bootstrap resampling for uncertainty estimation

(Efron, 1992), null model approaches for statistical testing (Gotelli & Graves, 1996), and the analysis of large biodiversity databases, such as BioTime (Dornelas et al., 2025), all of which can create significant computational bottlenecks when processing multiple large datasets iteratively. To improve the efficiency of these computationally intensive analyses, Julia becomes a logical choice for re-implementations. Julia is known for its just-in-time compilation that optimizes code execution to levels comparable with lower-level languages like C or Fortran (Bezanson et al., 2017), yet Julia remains relatively uncommon in ecological workflows.

## State of the field

To position `MetaCommunityMetrics.jl` within the Julia ecosystem, we conducted a comprehensive review of existing Julia packages under the [EcoJulia](#), a community-driven effort providing tools for ecological and environmental analysis, and the [PoisotLab](#), which develops many Julia packages focusing on quantitative and computational ecology. The packages under [EcoJulia](#) cover a range of ecological analyses, including spatial ecology, diversity measurement, phylogenetics, and ecosystem simulations, while the [PoisotLab](#) packages focus primarily on species distributions and species interaction networks. While some of these packages implement diversity measures relevant to community ecology, none implement the specific metacommunity-scale metrics listed in [Table 1](#). We ensured that our package builds beyond the tools currently existing in Julia even when sharing similar ecological concepts. For example, our beta diversity metrics share the concept of beta diversity with `Diversity.jl` (Robert Reeve & Harris, 2024). However, our implementation decomposes beta diversity into species replacement and richness differences (or abundance difference when abundance data is used instead of occurrence data) (Baselga, 2010, 2012, 2013; Legendre, 2014), which differentiates our approach from `Diversity.jl`, which emphasizes alpha, beta, and gamma diversity measures at the metacommunity and community levels (Richard Reeve et al., 2014). `MetaCommunityMetrics.jl` is, to our knowledge, the first Julia package to implement the metrics listed in [Table 1](#), which are essential for research in community ecology but were previously only available through computationally expensive R implementations. However, some of the [EcoJulia](#) and [PoisotLab](#) packages are compatible with `MetaCommunityMetrics.jl` as they can be used upstream to retrieve and process data. For instance, `SpeciesDistributionToolkit.jl` provides a wrapper for the GBIF API (GBIF: The Global Biodiversity Information Facility, 2024) that allows users to query and retrieve georeferenced species occurrence records, which can then be formatted into a site-by-species matrix and analyzed using `MetaCommunityMetrics.jl` to investigate metacommunity processes, such as the interplay between dispersal and niche-based processes in shaping community composition across space and time.

## Software Design

We built `MetaCommunityMetrics.jl` on three design principles: usability, computational efficiency, and accuracy.

First, we prioritized usability for spatiotemporal data and simulation workflows. We extended single-snapshot metrics (single site and/or time point) to operate on multidimensional datasets (multiple sites and/or time points) via aggregation functions (`spatial_beta_div()`, `temporal_beta_div()`, `average_MVNH_det()`, `average_MVNH_dissimilarity()`), eliminating manual looping. Our DNCI implementation also accepts empty sites and singletons using Zero-adjusted Bray-Curtis dissimilarity, filters uninformative species (completely absent or ubiquitous), and implements status flags for five edge cases rather than throwing errors or returning NAs, supporting simulation workflows where sparse matrices are meaningful. We developed `DNCI_create_groups()` and `DNCI_plot_groups()` to automate site grouping (required by Vilmi et al. (2021) but manual in R) and visualize grouping results.

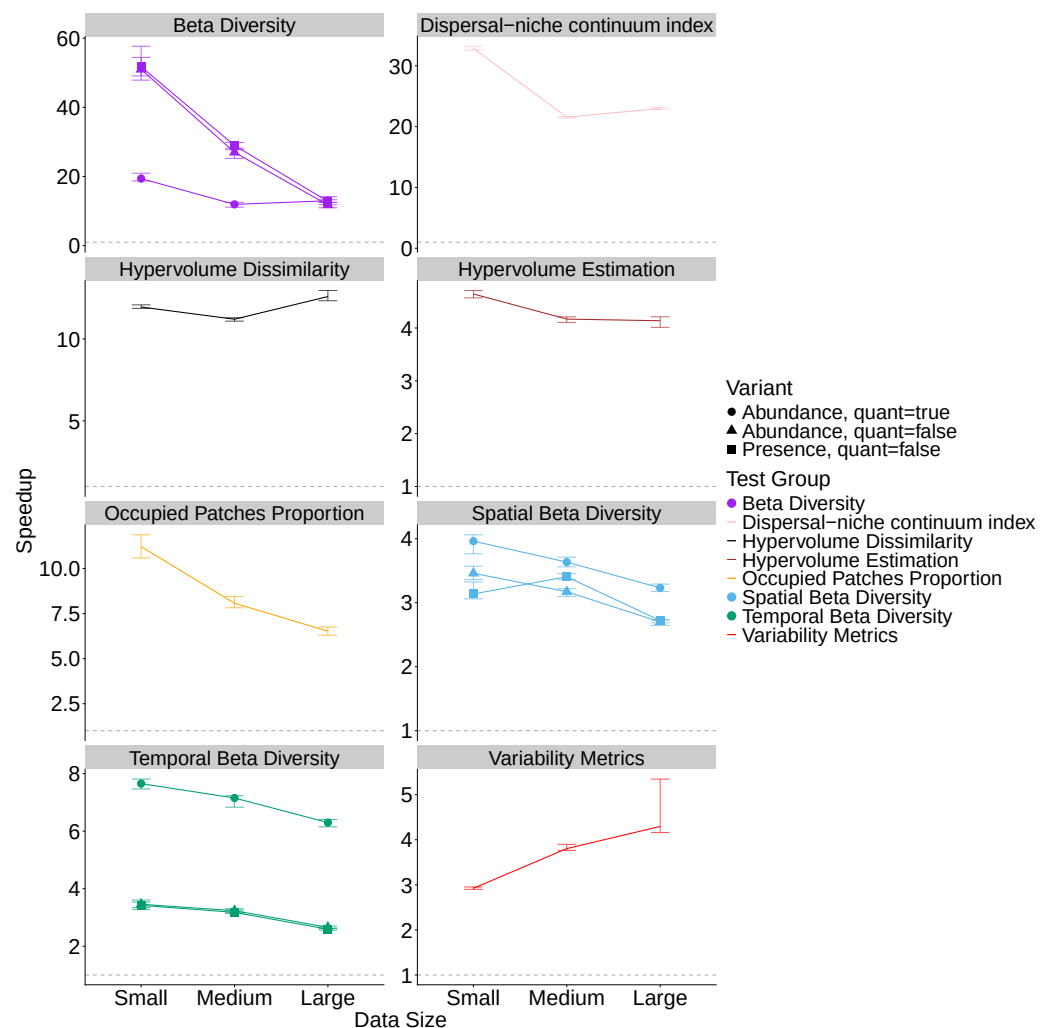
Second, we prioritized computational efficiency over comprehensive coverage (compared to the R implementations). Our beta diversity focuses on the Podani family approach (a mathematical framework that decomposes pairwise dissimilarity between sites into two distinct ecological components: species replacement and richness differences) rather than all five coefficient families, `MVNH_dissimilarity()` implements only the Bhattacharyya distance (a measure of dissimilarity between two probability distributions used here to quantify niche dissimilarity between two multivariate normal hypervolumes), and `DNCl` lacks built-in visualizations. We plan to expand these in future releases. To maximize performance, we use vectorized operations where possible and employ Cholesky decomposition in `MVNH_dissimilarity()` for the Mahalanobis component of the Bhattacharyya distance calculation, replacing the LU decomposition (a matrix factorization method that expresses a matrix as the product of a lower and upper triangular matrix) used in `solve()` in the original R implementation, providing substantial performance gains for positive-definite matrices (Haddad, 2024).

Finally, to ensure accuracy, all implementations with equivalent R functions were validated using the package's built-in sample data by running both Julia and R implementations on identical inputs within the same script via `RCall`, and computing the absolute differences between outputs. For all deterministic functions, differences were smaller than 0.0001, which we considered negligible and attributable to floating-point differences between implementations. For the `DNCl`, which is permutation-based, validation was conducted by running both implementations 100 times with 1000 permutations each and comparing the distributions of results, which showed agreement in ecological interpretation across all group pairs.

## Research Impact Statement

`MetaCommunityMetrics.jl` delivers substantial computational advantages over existing tools. For example, we re-implemented the `beta.div.comp` function from the R package `adespatial` (Dray et al., 2023), achieving up to  $\sim 52\times$  faster execution times while maintaining methodological consistency (see Figure 1). As shown in Figure 1, performance gains range from  $\sim 3\times$  to  $52\times$  across all re-implemented functions using three testing datasets (large: 53,352 observations; medium: 26,676 observations; small: 5,325 observations), positioning `MetaCommunityMetrics.jl` as an efficient solution for diverse ecological research needs, from standard biodiversity analyses to computationally intensive simulation studies and large-scale data processing. Memory usage is generally comparable between Julia and R implementations, with notable reductions in Julia for the variability metrics function (e.g., 12.61 MiB vs 60.23 MiB for the large dataset). However, the `DNCl` implementation requires substantially more memory in Julia than in R (e.g., 127.52 MiB vs 77.79 MiB for the large dataset). This difference arises because the R implementation generates the column-sum-preserving permuted matrices one at a time within the permutation loop, immediately computing the result and discarding the matrix, so only a single permuted matrix is held in memory at any given time. The use of parallelization in the R benchmark further reduces the peak memory usage. In contrast, the Julia implementation generates all permuted matrices upfront and stores them simultaneously in memory before processing. Direct comparisons should also be interpreted with caution given the differences in how each language measures memory allocation.

Beyond computational performance, the development of `MetaCommunityMetrics.jl` was driven in part by the needs of an ongoing empirical application of Approximate Bayesian Computation methods to connect metacommunity theory with observed data from a long-term phytoplankton community monitoring dataset. Specifically, the requirement to efficiently calculate summary statistics for this application directly informed the design and implementation of several functions in the package. The manuscript of this application is under preparation.



**Figure 1:** A plot showing the speedup of all benchmarked functions across three datasets: small (5,335 observations, 10% of the full dataset), medium (26,676 observations, 50% of the full dataset), and large (53,352 observations), all derived from the sample dataset available via `load_sample_data()`. The sample dataset is a subset of rodent abundance data from the Portal Project (Ernest et al., 2018), a long-term study of a Chihuahuan desert ecosystem, covering 24 sites, 21 species, and 117 sampling events from 2010 to 2023. The small and medium datasets were generated by random subsampling of the sample dataset with a fixed random seed and saved for consistent reuse across benchmark runs. Each function was benchmarked using 100 samples in both Julia and R. Speedup is calculated as the R median execution time divided by the Julia median execution time. Median speedup and its 95% bootstrap confidence interval (based on 2000 bootstrap iterations) are shown for each comparison. The grey dashed line indicates where speedup equals 1, meaning both implementations require the same execution time. Beta diversity has two variants: `quant = true` (abundance data) and `quant = false` (occurrence data). When `quant = false`, abundance and presence inputs produce equivalent results, but the abundance input requires an additional internal conversion to binary, which accounts for the difference in speedup between the two inputs.

## 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 changes in abundance vs. abundance gradient. *Methods in Ecology and Evolution*, 4(6), 552–557. <https://doi.org/10.1111/2041-210x.12029>
- Bezanson, J., Edelman, A., Karpinski, S., & Shah, V. B. (2017). Julia: A fresh approach to numerical computing. *SIAM Review*, 59(1), 65–98. <https://doi.org/10.1137/141000671>
- Dornelas, M., Antão, L. H., Bates, A. E., Brambilla, V., Chase, J. M., Chow, C. F. Y., Fontrodona-Eslava, A., Magurran, A. E., Martins, I. S., Moyes, F., Sagouis, A., Adu-Acheampong, S., Acquah-Lamptey, D., Adam, D., Ajani, P. A., Albaina, A., Almaraz, P., An, J., Anderson, R. S., ... Zvuloni, A. (2025). BioTIME 2.0: Expanding and improving a database of biodiversity time series. *Global Ecology and Biogeography*, 34(5), e70003. <https://doi.org/10.1111/geb.70003>
- Dray, S., Bauman, D., Blanchet, G., Borcard, D., Clappe, S., Guénard, G., Jombart, T., Larocque, G., Legendre, P., Madi, N., & Wagner, H. H. (2023). *Adespatial: Multivariate multiscale spatial analysis*. <https://doi.org/10.32614/cran.package.adespatial>
- Efron, B. (1992). Bootstrap methods: Another look at the jackknife. In S. Kotz & N. L. Johnson (Eds.), *Breakthroughs in statistics: Methodology and distribution* (pp. 569–593). Springer New York. [https://doi.org/10.1007/978-1-4612-4380-9\\_41](https://doi.org/10.1007/978-1-4612-4380-9_41)
- Ernest, S. M., Yenni, G. M., Allington, G., Bledsoe, E. K., Christensen, E. M., Diaz, R. M., Geluso, K., Goheen, J. R., Guo, Q., Heske, E., & others. (2018). The portal project: A long-term study of a chihuahuan desert ecosystem. *BioRxiv*, 332783. <https://doi.org/10.1101/332783>
- GBIF: The Global Biodiversity Information Facility. (2024). *What is GBIF?* <https://www.gbif.org/what-is-gbif>
- Gibert, C. (2024). *DNCImper: Assembly process identification based on SIMPER analysis*. <https://github.com/gibertcorentin/DNCImper>
- Gotelli, N. J., & Graves, G. R. (1996). *Null models in ecology*. Smithsonian Institution Press.
- Guzman, L. M., Thompson, P. L., Viana, D. S., Vanschoenwinkel, B., Horvath, Z., Ptacnik, R., Jeliazkov, A., Gascón, S., Lemmens, P., Anton-Pardo, M., & others. (2022). Accounting for temporal change in multiple biodiversity patterns improves the inference of metacommunity



- processes. *Ecology*, 103(6), e3683. <https://doi.org/10.1002/ecy.3683>
- Haddad, C. N. (2024). Cholesky factorization. In *Encyclopedia of optimization* (pp. 1–3). Springer. [https://doi.org/10.1007/978-3-030-54621-2\\_67-1](https://doi.org/10.1007/978-3-030-54621-2_67-1)
- Hallett, L., Avolio, M. L., Carroll, I. T., Jones, S. K., MacDonald, A. A. M., Flynn, D. F. B., Slaughter, P., Ripplinger, J., Collins, S. L., Gries, C., & Jones, M. B. (2020). *Codyn: Community dynamics metrics*. <https://doi.org/10.5063/F1N877Z6>
- Legendre, P. (2014). Interpreting the replacement and richness difference components of beta diversity. *Global Ecology and Biogeography*, 23(11), 1324–1334. <https://doi.org/10.1111/geb.12207>
- Lu, M., Winner, K., & Jetz, W. (2021). A unifying framework for quantifying and comparing n-dimensional hypervolumes. *Methods in Ecology and Evolution*, 12(10), 1953–1968. <https://doi.org/10.1111/2041-210X.13665>
- Oksanen, J., Simpson, G. L., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., Solymos, P., Stevens, M. H. H., Szoecs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Borcard, D., Carvalho, G., Chirico, M., De Caceres, M., Durand, S., ... Weedon, J. (2024). *Vegan: Community ecology package*. <https://doi.org/10.32614/CRAN.package.vegan>
- R Core Team. (2024). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://doi.org/10.32614/r.manuals>
- Reeve, Robert, & Harris, C. (2024). *Diversity.jl* (v0.5.15). <https://doi.org/10.5281/zenodo.12790658>
- Reeve, Richard, Leinster, T., Cobbold, C. A., Thompson, J., Brummitt, N., Mitchell, S. N., & Matthews, L. (2014). How to partition diversity. *arXiv Preprint arXiv:1404.6520*. <https://doi.org/10.48550/arXiv.1404.6520>
- Vilmi, A., Gibert, C., Escarguel, G., Happonen, K., Heino, J., Jamoneau, A., Passy, S. I., Picazo, F., Soininen, J., Tison-Rosebery, J., & others. (2021). Dispersal-niche continuum index: A new quantitative metric for assessing the relative importance of dispersal versus niche processes in community assembly. *Ecography*, 44(3), 370–379. <https://doi.org/10.1111/ecog.05356>
- Wang, S., Lamy, T., Hallett, L. M., & Loreau, M. (2019). Stability and synchrony across ecological hierarchies in heterogeneous metacommunities: Linking theory to data. *Ecography*, 42(6), 1200–1211. <https://doi.org/10.1111/ecog.04290>