



# European Vegetation Archive Data Request Form



To obtain data from the European Vegetation Archive (EVA), including the ReSurveyEurope Database, please first enquire the EVA database administrator Ilona Knollová (ikuzel@sci.muni.cz) whether the data that meet your needs are available. If they are, please fill in the form below and submit it to Ilona or another member of the EVA Coordinating Board (or ReSurveyEurope Board if you ask for data from the ReSurveyEurope Database).

- Applicant's name:

Zeke Marshall

- Applicant's institutional address:

UK Centre for Ecology & Hydrology, Library Avenue, Bailrigg, Lancaster, LA1 4AP, UK

- Applicant's e-mail:

[zekmar@ceh.ac.uk](mailto:zekmar@ceh.ac.uk)

- Project title:

The development and testing of methods for the analysis of large matrices in vegetation science.

- Are you asking for core EVA data (non-repeated vegetation surveys) or for ReSurveyEurope data (repeated vegetation surveys)?

Core EVA at present.

- Brief description of the aims and methods of the study:

### **Outline**

The fundamental unit of community ecology is the association or co-occurrence matrix and computational algorithms for key analyses in numerical ecology such as ordination and clustering are based on such matrices (Legendre and Legendre, 2012).

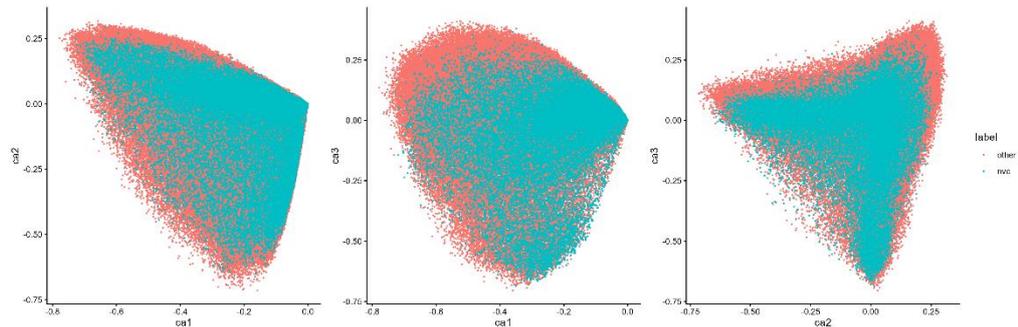
However, mathematical operations using regular (dense) matrices are limited by computational resource, with the memory requirements for storing matrices scaling quadratically relative to matrix size, preventing application to large datasets such as the European Vegetation Archive (EVA).

Encouragingly, this problem has already been circumvented in numerous other fields of research through the use of sparse matrices, matrix free methods, and Arnoldi and Krylov methods (Lehoucq et al., 1998).

This project aims to explore the use of the Julia programming language (Bezanson et al., 2017), which has excellent support for the aforementioned methods, for performing common analyses in numerical ecology, with a focus on vegetation science.

Prototyping has already begun using the GBNVPD (~4000 species x ~270,000 plots) and initial results are promising, as demonstrated in Figure 1 below.

Should routines for basic analyses such as correspondence analysis be successfully created, the code will be made available in a Julia package.



**Figure 1:** The first three axes of a correspondence analysis for the GBNVPD, with the plot coordinates grouped into the NVC dataset (blue) and all other datasets (red).

### References

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>

Legendre, P., & Legendre, L. (2012). *Numerical Ecology*. Elsevier.

Lehoucq, R. B., Sorensen, D. C., & Yang, C. (1998). *ARPACK users' guide: Solution of large-scale eigenvalue problems with implicitly restarted Arnoldi methods*. SIAM. ISBN 978-0-89871-407-4

- Will someone else be involved in data editing or analysis in addition to the applicant?

Not at present.

- Estimated time of delivery of results (e.g., manuscript submission):

October 2026.

- Geographic area needed (e.g., countries or range of geographic coordinates):

All countries in the EVA target area excluding Greenland, Kazakhstan (European part), Russia (European part), Svalbard and Jan Mayen, North Africa, and the Middle East.

- Do you need plots to be georeferenced? If so, what is the minimum accuracy of plot location (in metres or kilometres) needed for your project?

Yes, 10km.

- Vegetation types needed (syntaxa):

All.

- Other data selection criteria:

Would it be possible to provide the EUNIS assignments for each plot?

- Envisaged publications:

A software article in the *Journal of Vegetation Science* or other international journal.



# European Vegetation Archive Data Request Form



- **Data deposition.** Some journals require data used for the analysis to be stored in a public repository to ensure the repeatability of the analysis. According to EVA Rules, you are not allowed to store the original vegetation-plot data obtained from EVA. However, if you plan to publish in such a journal, you may deposit a reduced EVA-derived dataset that (1) would make it possible to repeat the analysis published in the paper and (2) does not contain any information not used in the analysis. For example, such a dataset can contain only a subset of species (e.g., only angiosperms or only neophytes), or replace species names with codes, or replace species cover values with presences/absences, or remove all the header data, or replace the exact plot coordinates by coarse grid-cell coordinates etc. If you plan to deposit reduced information from vegetation plots, please describe here what might be deposited. If the project developed so that you needed to deposit more information than specified here, you would need to ask specific permission from the Custodians of the EVA databases used in your analysis before the dataset is deposited.

N/A

- **Plant trait data from the TRY consortium.** If you plan to combine your analysis of vegetation-plot data with plant trait data, you can also request a dataset of 18 gap-filled traits for a large number of plant taxa prepared by the TRY consortium. These traits include Leaf area, Specific leaf area, Leaf fresh mass, Leaf dry matter content, Leaf C, Leaf N, Leaf P, Leaf N per area, Leaf N:P ratio, Leaf delta15N, Seed mass, Seed length, Seed number per reproductive unit, Dispersal unit length, Plant height, Stem specific density, Stem conduit density, and Conduit element length. This dataset can be provided to you by the EVA manager together with the vegetation-plot data. If you use this dataset, you must inform about your project the TRY data contributors who might be potentially interested and invite them as potential co-authors, assuming they will make an intellectual contribution to your paper. The list of the TRY data contributors will be sent to you together with the gap-filled trait dataset.

No

- **Specification of the co-authorship arrangements in publications based on the requested data.** Note that the EVA Rules recommend that co-authorship is offered to a representative of each database providing data that are particularly important for the project (e.g., a relatively large proportion of the final dataset used in the analyses or data from unique vegetation types or under-represented geographic areas). This database representative should be an expert in the topic of the project (not necessarily the custodian or deputy custodian), and this person should contribute to the project more than just by providing the existing data, e.g. by intellectual contribution to the concept of the paper, preparation of new data, or helping with data analysis, interpretation of the results or writing parts of the paper (see the IAVS Code of Professional Ethics: [https://www.iavs.org/page/governance\\_code-of-professional-ethics](https://www.iavs.org/page/governance_code-of-professional-ethics)). The project leader should enable active participation by regularly informing potential co-authors about the progress of the project from its early stage. The project leader should also make final co-authorship arrangements based on the real input of the individual contributors.

I am happy to welcome any contributors as collaborators and co-authors. The data contributors who are interested in this collaboration should register in the EVA online form. Following the EVA Rules, co-authorship should be connected with an intellectual contribution.

- **Eligibility of the applicant to receive EVA or ReSurveyEurope data.** Specify to which EVA or ReSurveyEurope database the applicant has contributed; if the applicant is not the custodian or deputy custodian of an EVA or ReSurveyEurope database, give a name of a custodian or deputy custodian who supports this data request.

I am the custodian of the GBNVPD (GIVD-code EU-GB-007).



# European Vegetation Archive Data Request Form



- I agree with the terms of EVA Data Property and Governance Rules as approved on 26 May 2012 (<http://euroveg.org/download/eva-rules.pdf>).
- If I ask for ReSurveyEurope data, I agree with the terms of ReSurveyEurope Data Property and Governance Rules as approved on 6 April 2022 (<http://euroveg.org/download/resurveyeurope-rules.pdf>).
- In any result obtained based on EVA core data (non-repeated vegetation surveys), I will cite the EVA report article (Chytrý et al. 2016; <https://doi.org/10.1111/avsc.12191>). In any result obtained based on the ReSurveyEurope data (repeated vegetation surveys), I will cite the ReSurveyEurope report article as soon as it is published. In addition, I will cite individual source databases used in my project (if possible, in the list of References; if not possible, at least as a list of databases in the electronic supplementary material).
- If I ask for the plant trait data from TRY, I agree to invite to my project the TRY data contributors following the list received from the EVA database manager.

Newton Aycliffe, 07/02/2026

Zeke Marshall