



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:

Mingwei Li

- Applicant's institutional address:

PhD Student in Political Economy
School of Economics, Zhejiang University
West Quad, Zijingang Campus
866 Yuhangtang Road
Hangzhou, Zhejiang 310058
P. R. China

- Applicant's e-mail:

12601001@zju.edu.cn

- Project title:

Prospective validation of co-occurrence-based dark-diversity estimates across European vegetation resurveys

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

Both core EVA data (dated non-repeated vegetation surveys) and ReSurveyEurope data (repeated vegetation surveys). ReSurveyEurope will provide the focal temporal outcomes, while core EVA will provide the dated historical reference observations needed to construct regional species pools and co-occurrence predictors using information available no later than each focal baseline.

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

Aim. The study will test the prospective predictive validity of the co-occurrence quantity used in hypergeometric dark-diversity estimation. For each vascular plant species not recorded in a focal plot at baseline but represented in a baseline-defined regional reference pool, we will ask whether its baseline co-occurrence probability predicts later recorded occurrence in the same plot beyond the simpler benchmark of baseline regional prevalence.

Baseline-only predictor construction. Species pools and both predictors will be frozen at each focal baseline. Reference matrices will use only observations sampled by that date. The focal observation will always be excluded; where coverage permits, the entire focal ReSurvey project/database will be held out for independent validation. We will compare prevalence-only, co-occurrence-only and combined models using the same candidate pool.

Temporal outcome and modelling. The analysis unit will be a plot-species candidate. General later recorded occurrence will be analysed together with a higher-confidence persistent endpoint, defined for suitable time series as absence at baseline followed by repeat-confirmed later presence (for example, presence in both the penultimate and final surveys). Survey duration, number and timing of survey occasions, plot area, relocation uncertainty and source project will be modelled explicitly. Predictive value will be evaluated out of sample using calibration, discrimination and proper predictive scores, with hierarchical or equivalent resampling procedures for dependence among species, plots and projects.

Ecological-context domain. The confirmatory analysis will focus on permanent or accurately relocated plots with at least three comparable survey occasions, comparable plot area and sampling methods, and evidence that the broad ecological context remained comparable. Independent habitat, land-use, management and environmental information will define this subset where available; composition-derived trajectories will serve only as a sensitivity analysis. A linked secondary analysis will examine how predictive transferability changes under succession, disturbance or other evidence of species-pool transition.

Heterogeneity and transferability. Prespecified moderators - vegetation type, resurvey interval, ecological-context stability, disturbance or succession, plot comparability, relocation precision and regional-reference coverage - will be used to test when incremental predictive performance remains transferable.

Interpretation. Non-arrival will not be interpreted as proof of ecological unsuitability. The estimand is the incremental temporal information in a baseline co-occurrence estimate beyond prevalence, and the ecological contexts and time horizons over which that information remains transferable.

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

Professor Deming Luo, Business School, Ningbo University, will contribute to study design, statistical modelling, robustness assessment and interpretation. Raw data will be accessible only to authorised project members. Custodian-nominated representatives may subsequently join under the EVA/ReSurveyEurope authorship procedures.

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

Target manuscript submission within 18 months of data delivery, and no later than 24 months after data delivery.

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

All European countries represented in ReSurveyEurope. Dated core EVA reference plots are requested from the same countries and, where they fall within prespecified spatial neighbourhoods around focal plots, adjacent cross-border areas. Baseline-specific temporal and spatial filters will be applied after delivery so that reference data remain matched to each focal plot and baseline year.

- 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. Please provide all available plot coordinates, including corner coordinates, together with precision/uncertainty and the plot (re-)location method. No single accuracy threshold should be imposed during extraction: analysis thresholds will be prespecified by spatial scale and varied in sensitivity tests. Coordinates are required for regional pool construction, spatial stratification and, where permitted, linkage to coarse contextual covariates used solely to assess ecological-context comparability. Exact coordinates will not be published, shared or deposited.



European Vegetation Archive

Data Request Form



- Vegetation types needed (syntaxa):

All terrestrial vegetation types represented in ReSurveyEurope, stratified by available syntaxonomic, EUNIS or other habitat classifications. The requested biological data are the complete vascular-plant component. Aquatic vegetation and observations without a complete vascular-plant list will not be used in the primary analysis.

- Other data selection criteria:

General extraction principle. Please provide a broad candidate selection and retain records when optional metadata are missing; missing optional fields should be documented rather than used automatically as extraction exclusions. All eligible survey occasions and all standard and available extended header fields described below are requested because they are needed for temporal ordering, taxonomic harmonisation, comparability screening, context-transition classification, spatial pool construction, project holdout, quality control or sensitivity analysis.

A. ReSurveyEurope focal series. Request series with at least two survey dates and complete vascular-plant lists, retaining all occasions without date limits. Two occasions support general later occurrence; three or more comparable occasions support confirmatory persistent appearance. The delivery should include species occurrence and all available cover/abundance information by vegetation layer, the cover scale, harmonised taxon names and identifiers, and original taxon names/codes or taxonomic-resolution flags where retained and available. Source database/GIVD identifiers, project identifiers and stable plot-observation identifiers are required for linkage, provenance and leave-one-project/database-out validation.

ReSurveyEurope standard fields. Please include RS_PROJECT, RS_SITE, RS_PLOT, RS_OBSERV, DATE, SURF_AREA, PLOT_SHAPE, LONGITUDE, LATITUDE, LONGITUDE2-LONGITUDE4, LATITUDE2-LATITUDE4, PRECISION, LOC_METHOD, MANIPULATE, MANIPTYPE, OWNER and EVA_ACCESS wherever available. The extraction should not discard a series solely because MANIPULATE = Y, because this field can include both experiments and natural disturbance/succession. MANIPTYPE and project metadata will be used to retain controls, exclude experimental treatments from the confirmatory analysis, and identify natural-disturbance or succession trajectories for the prespecified transferability analysis.

ReSurveyEurope extended metadata and documentation. Please include, where available: country/locality; sampling author and bibliographic reference; sampling design and methods; total and layer-specific vegetation cover; elevation, slope and aspect; bedrock and soil; land use, management, recent land-use change and disturbance; vegetation, habitat, EUNIS and syntaxonomic classifications; plot permanence, relocation and comparability notes; remarks and quality flags. Please also provide the ReSurveyEurope Project Metadata Form for each delivered project, or equivalent project-level documentation, including the project reference, sampling design/methods, measured environmental variables and treatment/disturbance description.

B. Core EVA historical reference observations. Please provide dated, georeferenced core EVA plots from the requested geography and vegetation scope with complete vascular-plant lists. No global date cutoff is requested because each record will be filtered relative to the baseline year of a focal ReSurveyEurope plot. Please include species occurrence, all available cover/abundance and layer information, cover scale, harmonised taxon names/identifiers and available original names/codes; and all available standard or extended header information relevant to date, plot area, coordinates and uncertainty, source database/GIVD/project and observation identifiers, country/locality, author/reference, sampling methods, vegetation/habitat/EUNIS/syntaxon, elevation, slope, aspect, total/layer cover, substrate/soil, land use/management/disturbance, remarks and quality flags.

C. Delivery metadata. For both components, please provide the data dictionary, missing-value conventions, available taxonomic harmonisation/crosswalk, source-database/access-regime mapping, and a field-availability record by source project/database. Species, header and project-level files should use stable link keys. These items are needed to prevent apparent colonisations caused by taxonomic/identifier changes and to distinguish missing values from unavailable fields.



European Vegetation Archive

Data Request Form



- Envisaged publications:

One main peer-reviewed scientific article on the prospective predictive validity and temporal transferability limits of co-occurrence-based dark-diversity estimation across European vegetation resurveys.

- 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.

Analysis code, model specifications, workflow documentation and non-EVA-derived or fully synthetic materials will be archived publicly where appropriate. We do not plan to deposit original EVA/ReSurveyEurope plot data or a reconstructable reduced version. If a journal requires deposition of any EVA-derived analytical data, its exact content and spatial/taxonomic resolution will first be agreed in writing with the relevant custodians and Boards. Exact coordinates, source plot identifiers, raw source metadata and reconstructable plot-by-species records will not be published or deposited.

- 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). The project leader should enable active participation by regularly informing potential



European Vegetation Archive Data Request Form



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.

Representatives of databases that make a particularly important contribution to the final analysis - through a large share of retained observations or unique vegetation types or under-represented regions - will be invited early to contribute intellectually to study design, data validation, analysis, interpretation and writing. Custodian-nominated representatives who express interest will receive regular progress updates and a genuine opportunity to participate. Final authorship will reflect actual contributions and follow the EVA and ReSurveyEurope Rules and the IAVS Code of Professional Ethics; other data contributors will be acknowledged and all required source databases will be cited.

- 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.

The applicant is not a custodian or deputy custodian of an EVA/ReSurveyEurope database. Professor Milan Chytrý, an EVA/ReSurveyEurope database custodian and member of the ReSurveyEurope Board, has confirmed in writing that he supports this data request.

- 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.

Hangzhou, 26 August 2026

Mingwei Li