Data Request Form

To obtain data from the European Vegetation Archive (EVA), please first make an enquiry to the EVA database administrator Ilona Knollová (ikuzel@sci.muni.cz) whether the data meeting 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.

- Applicants' name:

Remigiusz Pielech (main author)
Idoia Biurrun, Iwona Dembicz & Jürgen Dengler (GrassPlot Governing Board)
Idoia Biurrun, Florian Jansen, Borja Jimenez-Alfaro & Anna Kuzemko (EVA Governing Board)

- Applicant's institutional address:

University of Argiculture in Kraków, al. 29 Listopada 46, 31-425 Kraków, Poland

- Applicant's e-mail:

remekpielech@gmail.com

- Project title:

Biases in species richness data in large phytosociological databases

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

A just submitted manuscript from the GrassPlot database (Biurrun et al. subm.: Benchmarking plant diversity of Palaearctic grasslands and other open habitats) found unexpectedly strong differences in vascular plant species richness between mean values from GrassPlot and the few published mean values from large EVA member databases, where for nearly all vegetation class x region x grain size combinations the GrassPlot data had much higher richness. Given the wide and increasing usage of EVA and GrassPlot in many research projects, we consider it important for users to be aware of potential biases (incomplete richness records, preferential recording of particular species-rich or species-poor stands) in both databases. This would allow users to take counter-measures, e.g. excluding some data or regions or applying correction factors. With this study we thus aim to determine how often there are significant differences in mean richness estimates derived from both databases, and if so in which direction and how strong. We would quantify these for all combinations of countries and vegetation classes that are represented in both databases, but possibly aggregate at a higher level (country group or group of related vegetation classes). While GrassPlot directly contains richness data for the standard grain sizes 1 m², 10 m² and 100 m², in the case of EVA generalised additive models (GAMs) would be modelled to get values for 10 m², 100 m² and, where appropriate, also for 1 m². To account for potential reasons of incomplete records, we possibly will carry out the comparisons also for subsets of both databases that are expected to have higher data quality (EVA: the more recent plots; GrassPlot: only the plots of the EDGG nested plot sampling). Moreover, to assess other potential factors that could lead to biased richness estimates at the country or country group level, we will compare spatial aggregation patterns in both databases and possibly apply

resampling to avoid undue impacts of oversampled regions. Specifically for GrassPlot data, where for each dataset the sampling strategy (random, systematic, preferential) is already available (but will be further refined during the project), we will test how strong the biasing effects are.

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

The applicants' listed above will be the core authors of the paper, but potentially also opt-in authors from EVA or GrassPlot could get involved into data analysis.

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

End of 2021 (but if unforeseen complications occur potentially early 2022)

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

All

- 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?

Plots need to be assigned to a country

- Vegetation types needed (syntaxa):

All vegetation types that are also in GrassPlot, i.e. not only grasslands, but anything except forests, tall shrublands, true aquatic and segetal communities. Please note that e.g. dwarf shrub heathlands, garrigues, mires, springs, salt marshes, rocky communities, any type of arctic-alpine communities are included. We will provide a list of the target vegetation classes when approved.

- Other data selection criteria:

Plot size and vegetation class are available

- Envisaged publications:

One paper, with J. Veg. Sci. as preferred publication outlet

- 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 for 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 from 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. 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: http://iavs.org/Governance/Code-of-Professional-Ethics.aspx). 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.

We will offer co-authorship to one nominated person per each EVA database that contributes at least 1% of the final dataset on EVA side, provided this person makes an intellectual contribution to the paper.

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

Remigiusz Pielech is Custodian of EU-PL-003; all other applicants are also custodians or members of EVA databases

- We agree with the terms of EVA Data Property and Governance Rules as approved on 26 May 2012 (http://euroveg.org/download/eva-rules.pdf).
- In any result obtained based on this data, we will cite the EVA report paper (Chytrý et al. 2016; https://doi.org/10.1111/avsc.12191). In addition, we 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).

Krákow & Wädenswil, 25 December 2020

Remigiusz Pielech & Jürgen Dengler