

Database of Scale-Dependent Phytodiversity Patterns in Palaeartic Grasslands (GrassPlot)

(GIVD code: EU-00-003)

Data Property and Governance Rules (Bylaws)

Discussed and approved by the founding GrassPlot Consortium members, namely Idoia Biurrun, Steffen Boch, Alessandro Chiarucci, Timo Conradi, Francesco de Bello, Iwona Dembicz, Jürgen Dengler, Goffredo Filibeck, Itziar García-Mijangos, Riccardo Guarino, Monika Janišová, Swantje Löbel, Alireza Naqinezhad, Santiago Soliveres, Manuel J. Steinbauer, David Storch, Werner Ulrich and Viktoria Wagner, on 19 March 2017. Modified version approved by unanimous vote of the members on 31 December 2018.

1. Status and purpose of the GrassPlot Database

The Database of Scale-Dependent Phytodiversity Patterns in Palaeartic Grasslands (short: GrassPlot) is a collaborative initiative within the framework of the *Eurasian Dry Grassland Group* (EDGG) of the *International Association for Vegetation Science* (IAVS). The purpose of GrassPlot is to establish and maintain a common data repository of high-quality vegetation-plot observations (i.e. records of plant taxon co-occurrence at particular sites) of grasslands and related vegetation types from the whole Palaeartic biogeographic realm, and to facilitate the use of these data for non-commercial purposes, mainly academic research and applications in nature conservation and ecological restoration. GrassPlot focuses specifically on multi-scale and multi-taxon sampling in precisely delimited plots with extensive environmental data.

Obligatory requirements for data contributions to GrassPlot are: (a) origin in the Palaeartic biogeographic realm; (b) grassland vegetation in the wide sense, i.e. terrestrial and semi-terrestrial vegetation types dominated by hemicryptophytes, therophytes, geophytes, and occasionally bryophytes, lichens and chamaephytes (forests, shrublands, aquatic, ruderal and segetal vegetation are not considered); (c) careful sampling of precisely delimited plots with the aim of complete species lists, (d) providing details of sampling methodology (in particular, whether rooted or shoot presence was recorded and which plot shape was used), and (e) meeting one of the following criteria (or a combination of these): (i) data for one or several of the *eight standard grain sizes* (0.0001; 0.001 or 0.0009; 0.01; 0.1 or 0.09; 1; 10 or 9; 100; 1000 or 900 or 1024 m²) or (ii) *nested-plot series with at least four different grain sizes*. Plots meeting the following additional criteria are particularly welcome: (f) precise GPS coordinates; (g) complete sampling of one or several macroscopic non-vascular taxa of the terricolous vegetation (bryophytes, lichens, “algae”) in addition to vascular plants; (h) multi-scale sampling; (i) direct cover estimates of species in percent for at least one grain size; and (j) environmental variables measured or determined at the plot scale (vegetation structure, topography, soil, land use).

2. Data contributions to the GrassPlot Database

Persons who are willing to contribute their own published or unpublished plot records or plot records of other authors which they digitised from the literature can apply to become a member of the GrassPlot Consortium. Applications should be sent to the Executive Committee of GrassPlot. If

the offered data are not present in the database yet and meet the technical requirements of sampling and data storage, such offers will normally be accepted. Generally, data must be provided in an electronic format, but exceptionally unpublished data in paper format will be accepted if they fill important gaps. The complete data of the EDGG Field Workshops (formerly: *EDGG Research Expeditions*) automatically become part of the GrassPlot Database once they are prepared.

In GrassPlot, the data of a specific project (e.g. a dissertation or an EDGG Field Workshop) are treated as an *individual dataset*. GrassPlot does not support the artificial splitting of coherent datasets into several independent datasets. A dataset in GrassPlot has a *Dataset Custodian* and one or several *data owners* (one of them being the Dataset Custodian), whose maximum number is defined by the GrassPlot Score, and who through data contribution become part of the GrassPlot Consortium (see Article 3). As dataset owners can be listed those persons who carried out the field work, determined the plants, analysed the soils or entered and harmonized the data. It is the responsibility of the Dataset Custodian to forward communication received from the GrassPlot Governing Board to the individual data owners and to inform the Executive Committee about any changes of their contact data.

The *GrassPlot Score* rates the overall size and value of datasets in the GrassPlot context: First the numbers of all plots in a contributed dataset are summed up, using the following weights: 2 for plots $\geq 100 \text{ m}^2$, 1 for plots $\geq 1 \text{ m}^2$; 0.5 for plots $\geq 0.01 \text{ m}^2$ and 0.25 for plots $< 0.01 \text{ m}^2$. To this value, five times the number of nested-plot series (spanning at least three orders of magnitude) and the square root of the diameter of the study area in km (distance of the two plots furthest away from each other) are added. To derive the GrassPlot Score, the resulting raw score is multiplied by 2 if compositional data are also provided or if in addition to vascular plants also bryophytes and lichen have been sampled, or multiplied by 4 if both criteria are fulfilled. The fourth root of the GrassPlot Score rounded to the next integer defines the maximum number of data owners that can be assigned to one dataset.

The individual GrassPlot Score of a data owner is the summation of the fractional GrassPlot Scores of all datasets of which this person is co-owner of. For the fractional scores each owner of a dataset is counted equal, except the Dataset Custodian who is counted twice. The individual GrassPlot Scores of data owners are used to weight votes in elections of the GrassPlot Governing Board (see Article 3).

Having contributed plot records to GrassPlot does not limit other usages of these data by their owners. Unpublished data, i.e. data that do not exist in form of individual plot records in any publically accessible print or online sources, can be withdrawn from the GrassPlot Database by their Dataset Custodians at any time. Such withdrawal does not affect usages and publications that have already been approved at that point of time.

3. GrassPlot Consortium

The founding *Consortium of the Database of Scale-Dependent Phytodiversity Patterns in Palaeartic Grasslands* (short: *GrassPlot Consortium*) consists of the 17 researchers who participated physically or remotely in the initial GrassPlot Workshop, 6–10 March 2017, in Bayreuth, Germany. Individuals or groups of individuals who agree with these Bylaws and contribute data to GrassPlot that meet the specifications defined by the GrassPlot Governing Board become members of the Consortium once their data are included in the common database, whereby the maximum number of data owners and thus Consortium members per dataset is defined by the GrassPlot Score (see above). Further, the

GrassPlot Governing Board can appoint researchers without plot data provision as Consortium members who substantially contributed to the development of the GrassPlot initiative, be it via managing the database, preparing ancillary datasets or leading GrassPlot publications.

The GrassPlot Consortium has the following functions: It (i) elects the GrassPlot Governing Board, and (ii) approves and potentially modifies these Bylaws. Moreover, the whole GrassPlot Consortium has to be informed about any intended usage of data retrieved from the GrassPlot directly or indirectly. Only members of the GrassPlot Consortium are entitled to propose paper projects using GrassPlot data, but non-members can submit a proposal jointly with a GrassPlot Consortium member. Communication within the GrassPlot Consortium is via e-mail; it thus is the responsibility of each GrassPlot Consortium member to provide the Executive Committee with a functioning e-mail address.

Decisions by the GrassPlot Consortium (including elections) are made via electronic ballot (i.e. e-mail) within a 14-day period by simple majority among the returned votes. In these ballots, each GrassPlot Consortium member has one, two or three votes, based on the amount of data they contributed to the database. With an individual GrassPlot Score above 10,000 one has three votes, with an individual GrassPlot Score above 2,000 two votes and else one vote. Those who have been appointed as Consortium member without contribution of plot data always have one vote.

4. Governing Board, Executive Committee, Custodian and Deputy Custodian

GrassPlot is represented and governed by its *Governing Board*. The Governing Board is elected by the GrassPlot Consortium for two-year renewable terms. Eligible are all GrassPlot Consortium members who have been nominated by a GrassPlot Consortium member (including themselves) during a nomination period of at least 14 days. In the election each member can vote for up to seven candidates, with the votes being weighted with the factors indicated in Article 3. The seven candidates with the highest number of votes are elected; in case of a tie for the seventh position, all candidates with the same number of votes are also elected. If there are as many or fewer nominees than positions to be filled, the candidates are appointed without election. If more than half of the regular positions of the Governing Board have become vacant, they are re-filled by an election for the remaining time of the current term.

The Governing Board is responsible for managing the GrassPlot Database and any ancillary data, processing incoming data requests as described in Article 6, representing GrassPlot on conferences and similar occasions, attracting additional data or paper projects, preparing grant proposals as well as for maintaining a GrassPlot website and updating the GrassPlot data on the GIVD website. The Governing Board can appoint additional persons from inside and outside the GrassPlot Consortium to assist them with the data management. The Governing Board elects from its members the *GrassPlot Custodian and Deputy Custodian*, who together form the *GrassPlot Executive Committee*. The Executive Committee is responsible for running the everyday business of GrassPlot and serving as main contact for requests from inside and outside the Consortium.

Unlike the procedure described above, the first Governing Board for the period March 2017 – March 2019 will be assembled without prior election in order to facilitate a quick set-up of the first database version. This Governing Board will consist of Idoia Biurrun, Timo Conradi, Iwona Dembicz, Jürgen Dengler, Riccardo Guarino, Alireza Naqinezhad and Viktoria Wagner, with Jürgen Dengler being Custodian and Idoia Biurrun Deputy Custodian, and was approved by the founding GrassPlot Consortium members.

5. Data availability regimes

At the time of data submission or update, the Dataset Custodian assigns one of the following data availability regimes to the dataset contributed:

Regime 1: *Restricted-access data* require an explicit consent from the Dataset Custodian each time their use is requested. If the Dataset Custodian does not give such an approval (possibly after negotiation of details) within 14 days of distributing the paper proposal, the data will not be made available for the particular project.

Regime 2: *Semi-restricted-access data* are generally available for all projects that are in agreement with these Bylaws. The Dataset Custodian will be informed by the GrassPlot Executive Committee at least 14 days before the release of data. If no objection is raised during this period, this will be considered as implicit permission to use these data for the particular purpose.

Regime 3: *Free-access data* (typically those digitised from published sources) are available to a wider community of users. Their release does not require explicit or implicit consent of their originator for release.

Contributors are encouraged to assign Regime 2 to their data. Data of EDGG Research Expeditions, EDGG Field Workshops or other official EDGG activities automatically are assigned to Regime 2, but without the possibility to deny the use of the data on a case-by-case basis.

6. Data requests and terms of data use

Individuals or groups of individuals who would like to use the GrassPlot data for research or application purposes must submit a proposal to the Custodian describing the aims, basic methods and an approximate delineation of the data set they require. Only proposals by GrassPlot Consortium members will be considered, but non-members can submit a proposal jointly with a GrassPlot Consortium member. Only supra-regional proposals that are based on multiple datasets are considered; regional projects should direct their requests directly to the respective Dataset Custodian(s).

The proposal should contain (1) project title, (2) names and addresses of the applicant(s), indicating who of them is a Consortium member, (3) brief description of aims and methods of the study, (4) specification of the data needed (geographic area, plot sizes, compositional data, cryptogam treatment), (5) estimated time of delivery of results, e.g. manuscript submission, (6) envisaged publication(s), and (7) explicit statement that the applicant agrees with these Bylaws.

For incoming proposals, the Governing Board will check (1) whether the applicant is eligible to obtain data according to these Bylaws, and if so, (2) whether these Bylaws are respected, whether (3) there is a reasonable link between the aims, expected outputs and data requested, and (4) there are any conflicts with already approved projects. If these initial assessments have been positive, the Executive Committee will send the request to all Dataset Custodians and all data owners whose plots are concerned.

Within 14 days after such a request has been sent out, contributors of Regime-1 data can agree to provide their data and those of Regime-2 data disagree with their usage in this particular case. After

the deadline and based on the received answers of the Dataset Custodians, the Executive Committee will release the available data to the authors of the proposal.

The Governing Board will make efforts to deliver the data as fast as possible; however, the speed of this service may be affected by the available labour force and funding. Upon data delivery the applicants must provide the result (e.g. a manuscript ready for submission) to the Executive Committee within two years. After this time they lose their mandate to use the data unless they ask for extension and get it granted. The Executive Committee is entitled to ask for the progress of projects using GrassPlot data at any point of time. Each publication and conference contribution using GrassPlot data must be sent to the Executive Committee before submission for a check whether the data were used in accordance with these Bylaws. This check will not deal with the scientific quality of the manuscript and the Executive Committee will keep any information contained in the manuscript as confidential until the paper is published.

The applicants are not allowed to pass any data obtained from GrassPlot (including those under Regime 3) to third parties or use them for other purposes than stated in the original proposal. Data obtained from GrassPlot can be used for non-commercial purposes only and the final product (publication, report, software application) should contain a proper citation of the GrassPlot Database and of the original data sources (or at least main sources if there are many sources, possibly in an electronic appendix). Recommended references will be provided by the Executive Committee.

The titles of projects that have received data from GrassPlot, together with the names of applicants, can be published on the GrassPlot website. The applicants will send the papers based on GrassPlot data to the Custodian upon their publication; links to these papers can be added to the GrassPlot website. The applicants are encouraged to report any errors found in the data to the custodians of the original databases.

7. Authorship

For the first report-style paper about the philosophy, content and functionality of GrassPlot as well as for the first descriptive paper on alpha diversity patterns ("*data papers*", "*opt-out papers*"), all data owners who have contributed their data by the deadline specified by the lead will be offered co-authorship, regardless the extent of their contribution (in terms of data or paper writing). If updates on the content and functionality of GrassPlot are published, it is upon the discretion of the Governing Board to decide whether co-authorship is again offered to all data owners or only to those who newly joined or contributed new content since the last such database report.

Authorship of other papers or reports completely or partly based on data received from GrassPlot ("*research papers*", "*opt-in papers*") requires that, in addition to data provision, one makes an intellectual contribution to at least one of the following aspects: (i) design of the study; (ii) statistical analyses; (iii) interpretation of the results; (iv) writing up or (v) revising the text.

Potential *opt-in authors* (those authors who have not already been listed by the applicant in the application), have to declare their interest to join as co-authors during the same 14-day period during which also the Dataset Custodians have to respond to data requests. Their declaration should contain a specification why they are interested in the particular paper project and what they could contribute, and should be sent to the Executive Committee. For the combined EDGG Field Workshop datasets, the Governing Board will nominate at maximum as many opt-in authors as such datasets

are included in the request, based on expressed interests of the Workshop participants, taking into account their amount of contribution and how often they had been nominated in the past. Additionally, the Governing Board can nominate up to two opt-in authors from the Governing Board members or additional persons involved in the database management. For each other dataset that contributes at least 2% of the data of the particular study, the respective Dataset Custodian can nominate one of the data owners of the dataset as opt-in author. If someone is Dataset Custodian in several smaller datasets that together reach the 2% threshold, this is also accepted. The Executive Committee collects the information on proposed opt-in authors and will send them upon approval to the lead author of the publication for inclusion into the preliminary author list. Any other member of the GrassPlot Consortium not covered by the previous criteria (persons without data contribution or those whose data are not used in the particular study, data owners of datasets smaller than 2% or additional candidates from bigger datasets) can also apply to become opt-in author, but it is upon the discretion of the lead author(s) of the paper project whether they accept such an offer or not.

Generally, the lead author(s) of a paper project decide when and how to involve the approved opt-in authors. However, they are requested to provide all potential co-authors adequate opportunity and time to contribute. This means that outcomes should be shared with the full author team at least at three points of time, i.e. (i) when there is a concrete outline of the methodological approach, (ii) when there are the first results and (iii) at least two weeks before submission of a planned paper or presentation of the results on a conference. If a person preliminarily listed as opt-in author, despite requests from the lead author, has not made any intellectual contribution to the paper by the time of submission, the lead author should remove the respective name after consultation with the Governing Board. In case of dispute over authorship, both lead authors and opt-in authors can appeal to the Governing Board.

8. Relationship of GrassPlot to GIVD, EVA and sPlot

GrassPlot is registered in the *Global Index of Vegetation-Plot Databases* (GIVD) under ID EU-00-003. The GrassPlot Governing Board regularly updates the information in GIVD and intends to prepare a GIVD *Long Database Report* for the journal *Phytocoenologia* shortly after foundation. Whenever GrassPlot data are used, as a minimum this Long Database Report has to be cited.

GrassPlot is collaborating both with the *European Vegetation Archive* (EVA) and the global plot database "*sPlot*". For those data that are suitable for EVA and sPlot, i.e. typically those with plot sizes from 1–100 m² and with cover estimates, GrassPlot aims at facilitating the data provision to these continental and global plot databases. For this purpose, GrassPlot keeps record which of its contained data are already in EVA or sPlot and which additionally would be suitable. For the non-contained but suitable data from Europe (in the sense of EVA), GrassPlot will encourage and support data provision to a suitable regional or national EVA member database. For suitable data from outside Europe that are not yet contained in sPlot, GrassPlot will ask the Dataset Custodians for their agreement to contribute their data via GrassPlot to the next release of sPlot. Note that in this case the rules of data use follow the *Governance and Data Property Rules of the sPlot Working Group* (http://www.idiv-biodiversity.de/sdiv/workshops/workshops-2013/splot/join/content_815683/sPlot-Rules_approved.pdf), which can deviate from the rules applicable for data directly retrieved from GrassPlot.