

Ukrainian Grasslands Database (GIVD ID: EU-UA-001)

Data Property and Governance Rules

Discussed and approved by electronic vote of all data contributors at that point of time on 20 October 2020

1. Status and purpose of UKRAINIAN GRASSLANDS DATABASE

The UKRAINIAN GRASSLANDS DATABASE (UGD) is a collaborative initiative within the framework of the Working Groups European Vegetation Survey (EVS), European Dry Grassland Group (EDGG) and Ecoinformatics of the International Association for Vegetation Science (IAVS). The purpose of the UGD is to establish and maintain a common data repository of vegetation-plot observations (i.e. records of plant taxon co-occurrence at particular sites, typically sized between 1 and 100 m², further called relevés) of grasslands and related vegetation types from Ukraine and to facilitate the use of these data for non-commercial purposes, mainly academic research and applications in nature conservation and ecological restoration. Further the UGD aims at contributing its data to such analyses at subcontinental to global scales and thus is a member of the European Vegetation Archive (EVA; <http://euroveg.org/eva-database>), the global vegetation-plot database sPlot (<http://www.idiv-biodiversity.de/sdiv/workshops/workshops-2013/splot>) and the Global Index of Vegetation-Plot Databases (GIVD; <http://www.givd.info>).

2. Governance of the Ukrainian Grassland Database

Through contribution and inclusion of relevés, contributing persons (further: *Contributors*) become members of the *UGD Consortium* (further: *Consortium*), unless they wish to contribute data without Consortium membership (only possible for free access data). In addition, the Consortium can appoint persons involved in the database management but without relevé contribution as Consortium members per majority vote.

The Consortium has the following functions: It (i) elects the Custodian and Deputy Custodian for two-year renewable terms, (ii) approves and potentially modifies these Rules, and (iii) has the right to appoint database managers as Consortium members. Moreover, the whole Consortium has to be informed about any intended usage of data retrieved from the UGD directly or indirectly. Communication with the Consortium is via e-mail; it thus is the responsibility of each Consortium member to provide the Custodian with a functioning e-mail address.

Decisions by the 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 votes, each Consortium member has one, two or three votes, based on the amount of data and work he or she contributed to the database. For that purpose, the number of relevés contributed by a certain Contributor is counted and multiplied with a weighing factor (1 for relevés of other authors digitised from the literature, 2 for own published relevés, 3 for own unpublished relevés). If the resulting sum is 1–999 the member has one vote, for 1000–2999 two votes and for ≥ 3000 three votes. Up to two (additional) votes can be given to database managers by a decision of the Consortium.

The UGD is represented and governed by its *Custodian and Deputy Custodian*. These two persons, jointly or alone, are coordinating the database management and serve as contact for Consortium members and potential future data contributors. Further, they ensure that the data of the UGD are contributed to and regularly updated in the EVA and sPlot databases as well as the GIVD metadata base. Finally, any data use requests from EVA, sPlot, UGD Consortium members or external persons are to be sent to the Custodian, who then handles them according to these Rules.

3. Data contributions to the UGD

Persons who are willing to contribute their own published or unpublished relevés or relevés of other authors which they digitised from the literature can apply to become a member of the UGD Consortium. Applications should be sent to the Custodian of the UGD. 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.

Within the UGD each individual relevé will be assigned to one natural person as the Contributor. This could either be the person who made the relevé or the person who digitised a published relevé of someone else. If a person or institution provides a dataset to the UGD to which different persons contributed relevés, it is the obligation of the provider to assign to each relevé one single Contributor. In case of free-access data, providers can decide that these data are to be stored without Contributor information. It is the obligation of the data provider to provide and update the names and contact data of the Contributors to the Custodian.

Upon inclusion of provided relevés, their Contributor becomes member of the UGD Consortium. Having contributed relevés to the UGD does not limit other usages of these relevés by their Contributor. Contributors are encouraged to provide data corrections and updates to the Custodian. Unpublished data, i.e. data that do not exist in form of individual relevés in any publically accessible print or online sources, can be withdrawn from the UGD by their Contributor at any time. Such a withdrawal does only affect usages and publications that are not yet started at that point of time. If a contributor withdraws all of his/her contributed data, his/her membership in the UGD Consortium is terminated.

4. Data availability regimes

At the time of data submission or update, each Contributor assigns one of the following data availability regimes to the data contributed by him/her, either for the whole dataset or its individual relevés:

Regime 1: *Restricted-access data* are available for members of the Consortium only, and their use depends on the explicit permission of the respective Contributor. The Contributor must be contacted any time when there is a request to use these data. The terms of data use must be negotiated in each specific case.

Regime 2: *Semi-restricted-access data* are available for members of the Consortium only, but without the need for a specific permission in each single case. The Contributor will be informed by the Custodian or Deputy-Custodian at least 10 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* are available to a wider community of users. These data can be released based on the proposal to the Custodian, with no need for approval.

Unpublished relevés made by the Contributor himself/herself can be assigned to any of the three regimes, while relevés that have been published before can only be assigned to Regimes 2 or 3. The UGD initiative encourages a gradual transfer of data in the UGD from Regime 1 to Regime 3, however, the decision on such transfers is entirely with the Contributors.

5. Direct data requests and terms of data use

Individuals or groups of individuals who would like to use the UGD 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. With respect to data under Regimes 1 and 2 only proposals submitted by Consortium members will be considered. Other persons may ask for these data provided their request is approved by one of the Consortium members.

The proposal should contain (1) applicant's name and address, (2) project title, (3) brief description of aims and methods of the study, (4) estimated time of delivery of results, e.g. manuscript submission, (5) specification of the data needed (geographic area described as countries, counties or range of geographic coordinates, vegetation type or other selection criteria), (6) envisaged publications, (7) proposal how to handle the authorship in publications based on the data from the UGD, (8) explicit statement that the applicant agrees with these Rules.

For incoming proposals, the Custodian and Deputy Custodian will jointly check (1) whether the applicant is eligible to obtain data according to these Rules, and if so, (2) whether there is a reasonable link between the aims, expected outputs and data requested, and (3) whether these Rules are respected. If these initial assessments were positive, the Custodian or Deputy Custodian will send the request to all Contributors whose relevés are concerned, listing the requested relevés, together with information to which data availability regime they belong. Within three weeks 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 Contributors, UGD will release the available data to the authors of the proposal.

The Custodian and Deputy Custodian (possibly together with other Consortium members involved in the database management) 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 should deliver the result (e.g. a manuscript ready for submission) to the Custodian and Deputy-Custodian within three years. After this time they lose their mandate to use the data. Each paper using UGD data must be sent to the Custodian and Deputy-Custodian before submission for a check whether the data were used in accordance with these Rules. This check will not deal with the scientific quality of the manuscript and the Custodian and Deputy-Custodian 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 the UGD (including those under Regime 3) to third parties or use them for other purposes than stated in the original proposal. Data

obtained from the UGD can be used for non-commercial purposes only and the final product (publication, report, software application) should contain a proper citation of the original data sources (or at least main sources if there are many sources, possibly in an electronic appendix) and of the UGD. Recommended references will be provided by the Custodian and Deputy-Custodian.

The titles of projects that have received data from the UGD, together with the names of applicants, can be published on a UGD website (<https://edgg.org/databases/Regional-databases>). The applicants will send the papers based on the UGD data to the Custodian upon their publication; links to these papers can be added to a UGD website. The applicants are encouraged to report any errors found in the data to the custodians of the original databases.

The applicants are encouraged to report any errors found in the data to the custodians of the original databases.

6. Authorship

For report-style papers (“data papers”) about the philosophy, content and functionality of the UGD, including, but not restricted to, Long and Short Database Reports in *Vegetation Classification and Survey* or another journal, all members of the Consortium will be offered co-authorship, regardless the extent of their contribution (in terms of data or paper writing).

Authorship of papers or reports (“research papers”) based on data received directly from the UGD will be based on the authorship arrangement proposed by the lead author (applicant) that has been, possibly after modification following an individual negotiation, accepted by Contributors of Regime-1 data by explicit agreement and by Regime-2 data implicitly by not objecting within the deadline.

The recommended good practice for the applicants is that they offer co-authorship of research papers at least to those Contributors whose data (1) account for more than 10% of the final dataset used in the particular project or (2) are particularly important for the project result (e.g. they represent a unique vegetation type or geographic area which is not documented in other sources as well as Custodian, Deputy-Custodian or database managers when they were significantly involved in data preparation. Becoming co-author implies the possibility to contribute to the general concept of the product, data analysis and/or interpretation of the results. When in doubt, an inclusive, rather than a restrictive criterion should be applied.

7. Data contributions to and data usages by EVA and sPlot

All data contained in the UGD will be regularly uploaded into the *European Vegetation Archive (EVA)* with the same data availability regimes as within the UGD itself. To the *sPlot database*, which has only one uniform data availability regime similar to semi-restricted access, data of the Regimes 2 and 3 will be uploaded plus those of Regime 1 where the Contributors explicitly agreed with that.

Details of data usage and availability of data from the UGD obtained via EVA or sPlot are regulated by the *EVA Data Property and Governance Rules* (<http://euroveg.org/download/eva-rules.pdf>) and 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), respectively, and can deviate from the Rules applicable for data directly retrieved from the UGD.

When EVA or sPlot are preparing *data papers* about their content, the acting Custodian (and if possible also the Deputy-Custodian) will become co-authors of these on behalf of the Consortium. Announcements of intended data usage for *research papers* from EVA or sPlot will be received by the Custodian and forwarded to the whole Consortium. These information mails will contain an outline which relevés are concerned (normally with a list identifying relevé IDs, their data availability regimes and their Contributors) and specification of potential co-authorship offers. In the case of EVA, Contributors of Regime-1 data have the possibility to agree with use of their data in that specific case and Contributors of Regime-2 data can decline the use of their data in that case, both by response to the Custodian within the set deadline. No such possibility exists in the case of sPlot. Also within the deadline, Consortium members can announce their wish to become active co-authors of paper projects announced by EVA or sPlot, preferably supported with a statement of possible contributions (beyond the raw data). The Custodian or Deputy-Custodian will forward such applications to the lead author of the respective paper. In case there are more applicants for co-authorship in a specific paper project than the lead author is willing to accept (based on the EVA or sPlot Rules, respectively), the Consortium will assign the co-authorship option by majority vote. Criteria to be considered in this vote are (a) fractional contribution of data in that particular case; (b) specific expertise or interests, and (c) balance among Consortium members, considering also previous authorship assignments.

The data in EVA and sPlot are available for supra-national to continental and for continental to global analyses, respectively, proposed by representatives of those databases that compose the EVA and sPlot Consortia. Member of the UGD Consortium who wishes to conduct such analyses with data from EVA or sPlot can apply for this through the Custodian, following the Rules of EVA and sPlot, respectively. Application is restricted to Contributors who contributed Regime-2 or Regime-3 data (in the case of EVA) or who agreed with the transfer of their contributed data to sPlot (in the case of sPlot).

8. Data management

Custodian and Deputy-Custodian are responsible for the data management of the UGD. They can appoint additional database managers when needed. Data management comprises four major fields: (i) maintaining a file with names, contact data and voting weights for all Consortium members; (ii) setting-up, maintaining and updating the UGD in Turboveg format, including for each relevé information on its Contributor, type of data (own unpublished, own published, digitised from literature), data availability regime and, in case of Regime 1, whether the data can be contributed to sPlot; (iii) registering the UGD in GIVD and keeping the metadata there up-to-date; and (iv) providing the data to EVA and sPlot, including updates when appropriate.

9. Relationship with GIVD

After approval of these Rules, the Custodian and Deputy-Custodian will, together with all Consortium members, prepare and submit a database paper describing the content of the UGD to be published in an international journal, e.g. in cooperation with GIVD as Long (or Short) Database Report in *Vegetation Classification and Survey*. Whenever data from the UGD are used, whether retrieved directly or via EVA or sPlot, as a minimum requirement this paper shall be cited as reference for the data source together with the unique GIVD ID of the UGD.