Journal of Vegetation Science **22** (2011) 582–597

SPECIAL FEATURE: ECOINFORMATICS

The Global Index of Vegetation-Plot Databases (GIVD): a new resource for vegetation science

Jürgen Dengler, Florian Jansen, Falko Glöckler, Robert K. Peet, Miquel De Cáceres, Milan Chytrý, Jörg Ewald, Jens Oldeland, Gabriela Lopez-Gonzalez, Manfred Finckh, Ladislav Mucina, John S. Rodwell, Joop H. J. Schaminée & Nick Spencer

Keywords
Biodiversity; Data sharing; Ecoinformatics; GBIF; Global change; Macroecology; Metadata; Phytosociology; Relevé; Scientific reward

Abbreviations
EML, Ecological Metadata Language; GBIF, Global Biodiversity Information Facility; GEOSS, Global Earth Observation System of Systems; GIVD, Global Index of Vegetation-Plot Databases; TDWG, Biodiversity Information Standards (formerly: Taxonomic Database Working Group); XML, Extensible Markup Language

Received 29 September 2010
Accepted 24 January 2011
Co-ordinating Editor: Ingolf Kühn

Dengler, J. (corresponding author, dengler@botanik.uni-hamburg.de); **Oldeland, J.** (oldeland@botanik.uni-hamburg.de) & **Finckh, M.** (mfinckh@botanik.uni-hamburg.de): Biodiversity, Evolution and Ecology of Plants, Biocentre Klein Flottbek and Botanical Garden, University of Hamburg, Ohnhorststr. 18, 22609 Hamburg, Germany
Jansen, F. (jansen@uni-greifswald.de) & **Glöckler, F.** (falko.gloeckler@gmx.de): Landscape Ecology & Nature Conservation, Institute of Botany and Landscape Ecology, University of Greifswald, Grimmer Str. 88, 17487 Greifswald, Germany
Peet, R.K. (peet@unc.edu): Department of Biology CB#3280, University of North Carolina, Chapel Hill, NC 27599-3280, USA
De Cáceres, M. (miquelcaceres@gmail.com): Centre Tecnològic Forestal de Catalunya, Ctra. St. Llorenç de Morunys km 2, 25280 Solsona, Spain
Chytrý, M. (chytry@sci.muni.cz): Department of Botany and Zoology, Masaryk University, Kotlářská 2, 611 37 Brno, Czech Republic
Ewald, J. (joerg.ewald@hswt.de): Botany & Vegetation Science, Faculty of Forestry, University of Applied Sciences Weihenstephan-Triesdorf, Hans-Carl-von-Carlowitz-Platz 3, 85354 Freising, Germany
Lopez-Gonzalez, G. (g.lopez-gonzalez@leeds.ac.uk): School of Geography, University of Leeds, Leeds, LS2 9JT, UK

Abstract

Question: How many vegetation plot observations (relevés) are available in electronic databases, how are they geographically distributed, what are their properties and how might they be discovered and located for research and application?

Location: Global.

Methods: We compiled the Global Index of Vegetation-Plot Databases (GIVD; http://www.givd.info), an Internet resource aimed at registering metadata on existing vegetation databases. For inclusion, databases need to (i) contain temporally and spatially explicit species co-occurrence data and (ii) be accessible to the scientific public. This paper summarizes structure and data quality of databases registered in GIVD as of 30 December 2010.

Results: On the given date, 132 databases containing more than 2.4 million non-overlapping plots had been registered in GIVD. The majority of these data were in European databases (83 databases, 1.6 million plots), whereas other continents were represented by substantially less (North America 15, Asia 13, Africa nine, South America seven, Australasia two, multi-continental three). The oldest plot observation was 1864, but most plots were recorded after 1970. Most plots reported vegetation on areas of 1 to 1000 m^2; some also stored time-series and nested-plot data. Apart from geographic reference (required for inclusion), most frequent information was on altitude (71%), slope aspect and inclination (58%) and land use (38%), but rarely soil properties ($< 7\%$).

Conclusions: The vegetation plot data in GIVD constitute a major resource for biodiversity research, both through the large number of species occurrence records and storage of species co-occurrence information at a small scale, combined with structural and plot-based environmental data. We identify shortcomings in available data that need to be addressed through sampling under-represented geographic regions, providing better incentives for data collection and sharing, developing user-friendly database exchange standards, as well as tools to analyse and remove confounding effects of sampling biases. The increased availability of data sets conferred by registration in GIVD offers significant opportunities for large-scale studies in community ecology, macroecology and global change research.

Mucina, L. (l.mucina@curtin.edu.au): Department of Environment & Agriculture, Curtin Institute for Biodiversity & Climate, Curtin University, GPO Box U1987, Perth, WA 6845, Australia
Rodwell, J.S. (johnrodwell@tiscali.co.uk): Lincoln Institute, University of Manchester, Samuel Alexander Building, Oxford Road, Manchester M13 9PL, UK
Schaminée, J.H.J. (joop.schaminee@wur.nl): Radboud University Nijmegen and Wageningen UR, PO Box 47, 6700 AA Wageningen, The Netherlands
Spencer, N. (spencern@landcareresearch.co.nz): Landcare Research New Zealand, PO Box 40, Lincoln 7460, New Zealand

Doi: 10.1111/j.1654-1103.2011.01265.x

Introduction

Vegetation plot observations or relevés, which are broadly defined as records of plant taxon co-occurrence at particular sites, constitute the primary descriptive data on which much of vegetation science is based, and serve as the most important data resource available to vegetation scientists. These data contribute in many ways to a better understanding of vegetation as an object, as well as to the daily practice of nature management. Vegetation classification provides an invaluable framework for describing the context within which ecological and biodiversity research is conducted or applied. At the core of any credible classification of plant communities, there should be a data system that openly archives and disseminates underlying vegetation plot information. Vegetation science seeks general patterns in composition and dynamics of plant communities, yet these context-dependent phenomena are often obscured by the idiosyncrasies of the local environment and its history. This local context can only be recognized and corrected if the local study is placed in a much broader context, which is made possible by access to large databases of vegetation plots. Clearly, the creation and management of vegetation plot databases are activities essential for the advancement of vegetation science (Ewald 2003; Bekker et al. 2007; Le Duc et al. 2007; Schaminée et al. 2007, 2009).

Although millions of vegetation plots have been recorded and partly digitized for local and regional purposes, access to this massive and widely scattered resource was extremely limited prior to the advent of electronic database technologies and digital communication means. One of the first large-scale, modern vegetation databases was the Dutch National Vegetation Database founded in 1988 (Schaminée et al. 2006). The introduction of the database program TURBOVEG at the same time, followed by its widespread adoption (Schaminée & Hennekens 1995; Hennekens & Schaminée 2001), stimulated the development of regional and national vegetation databases in many countries. Rodwell (1995) gave the first major overview on existing vegetation-plot data in European countries without specifically differentiating computerized data. He concluded that 'well over one million relevés' were available on this continent alone. The first global survey of vegetation-plot databases was done by Ewald (2001), who suggested that there were over one million computerized plots worldwide. Schaminée et al. (2009) estimated the number of plots from Europe at 4.3 million, with more than 1.8 million stored electronically. Progress has been made outside Europe as well, with mature national vegetation plot databases having been developed for many countries, such as New Zealand, South Africa and the United States.

The original purpose for most of the current large vegetation plot databases was the preparation of detailed national or regional vegetation classification systems or vegetation maps (Rodwell 1991et seq.; Schaminée et al. 1995 et seq.; Valachovič et al. 1995 et seq.; Berg et al. 2001 et seq.; Mucina et al. 2006; Chytrý 2007 et seq.; Willner & Grabherr 2007). In many cases, the databases have been used for related studies of large-scale vegetation patterns (Wohlgemuth et al. 1999; Duckworth et al. 2000; van Kley & Schaminée 2003; Lososová et al. 2004). However, the current use of vegetation-plot databases goes far beyond traditional descriptive vegetation science, and these databases now serve in multiple applications in basic ecological and biogeographical research, as well as in applied studies addressing pressing environmental issues. For example, vegetation-plot data have been used to identify species responses to environmental gradients (Coudun & Gégout 2005), species niche width (Fridley et al. 2007), and to model spatial distributions of species (Coudun & Gégout 2007), plant communities (Brzeziecki et al. 1993; Marage & Gégout 2009), site conditions (Holtland et al. 2010) or biogeographical analyses (Loidi et al. 2010). Vegetation-plot databases also assist in studies of distribution patterns of species traits across landscapes or habitats (Baker et al. 2004; Malhado et al. 2009) and in testing specific hypotheses on plant adaptations (Ozinga et al. 2005, 2007, 2009; Lososová et al. 2006; Lososová & Láníková 2010). Numerous other fields of ecological research can benefit from vegetation plot databases, including the estimation of species pools (Ewald 2002), quantification of habitat suitability for metapopulation studies (Münzbergová & Herben 2004, or estimation of the values of environmental variables at sites where measured data are absent (Tichý et al. 2010). The large numbers of historical vegetation plot observations make these databases obvious candidates for studies of regional and global change. Examples of the use of vegetation plot databases for assessing processes associated with global change include analyses of species response to climate change (Lenoir et al. 2008), nitrogen deposition (Dupré et al. 2010; Maskell et al. 2010), estimation of carbon stocks across landscapes (Hall et al. 2001) and quantification of levels of alien plant invasion across habitats or landscapes (Brown & Peet 2003; Chytrý et al. 2008a, b; Vilà et al. 2010). Long-term ecological research (often using permanent vegetation plots for monitoring purposes) is a vibrant field profiting much from the versatility of computerized vegetation plot databases (e.g. Bakker et al. 2002; Smits et al. 2002). Climate change research profits from using long-term plot data, as studies on carbon storage in tropical forests (Lewis et al. 2009; Phillips et al. 2009, 2010) clearly suggest.

Doi: 10.1111/j.1654-1103.2011.01265.x

Some of the current impediments to vegetation-plot data distribution and sharing are of a technical nature (Wiser et al. 2011). Even though most vegetation-plot data worldwide have been gathered with moderately compatible field methods for the past 100 years, vegetation data are stored using widely differing database systems, with different data models and supporting different data formats. Yet, the disparate data formats are not the main obstacle in data exchange. Rather, the use of data follows availability, which typically has been local or regional, or at best national. As a consequence, broad-scale data sharing to investigate international and especially global-scale questions has been limited.

A promising avenue to stimulate integrative large-scale analyses is to increase the visibility of vegetation data through the use of global biodiversity websites. For example, some vegetation databases have contributed their floristic content to organisations such as the Global Biodiversity Information Facility (GBIF; Wheeler 2004). Unfortunately, publishing vegetation data through GBIF is not sufficient for vegetation plots as the GBIF data set largely follows the DarwinCore species occurrence standard, which does not incorporate the essential feature of vegetation data: the joint occurrence of taxa (e.g. Wiser et al. 2011; Finckh, M., Muche, M., Schmiedel, U. & Jürgens, N., unpublished data). One step towards a solution is adoption of an international exchange standard that goes beyond DarwinCore to fully support vegetation-plot data, which is the direction advocated and advanced by Wiser et al. (2011).

Numerous groups have discussed methods of facilitating data sharing across the ecological and biodiversity sciences. As in other areas of ecology, many important vegetation-plot data sets are maintained and controlled by groups with vested interests and who are unlikely to relinquish control, although they are often willing to allow use of their data. In the short term, the best available option for increased sharing appears to be one of providing a central registry where data sets can be discovered and mechanisms of access determined. This is the model behind Metacat, the on-line system for registering ecological data established in 2000 by the US National Center for Ecological Analysis and Synthesis, and the engine behind the data registry used by journals of the Ecological Society of America. As of 2006, Metacat provided metadata for more than 12 000 ecological data sets (Jones et al. 2006). In order to maximize documentation and discovery of vegetation-plot data, vegetation science needs a model similar to that of Metacat, but specifically focused on vegetation-plot data.

In this paper, we describe the design and initial content of the Global Index of Vegetation-Plot Databases (GIVD), an on-line portal designed specifically for vegetation-plot data, devoted to fostering data exchange by providing metadata on existing databases. In particular, we report the history and organisation of GIVD, provide statistics of its current content and point out avenues for collaborative research and application.

About GIVD

Historical development of GIVD

The first steps toward GIVD were made by F. Jansen and F. Glöckler who, on behalf of a German working group (Vegetation Databases Section of NetPhyD), established an on-line archive of metadata on vegetation databases from Germany and surrounding countries, under which basic information on 18 databases with nearly one million vegetation plots had been registered by 2008 (see http://geobot.botanik.uni-greifswald.de/portal/vegbank).

Creating a global archive of metadata was the logical next step. Fifteen scientists from eight countries, including the authors of the present article, serve as the Steering Committee of this international initiative. The metadata facility was launched under the name Global Index of Vegetation-Plot Databases (GIVD) in August 2010 (http://www.givd.info). GIVD is currently hosted at the Institute of Botany and Landscape Ecology at the University of Greifswald and is endorsed by the German working group on vegetation databases (http://netphyd.floraweb.de/?q=node/42), the European Vegetation Survey (EVS; http://www.evsitalia.eu/default.asp), and the EcoInformatics Working Group (http://www.bio.unc.edu/faculty/peet/vegdata/), the latter two being working groups of the International Association for Vegetation Science (IAVS). Apart from direct access to all uploaded metadata, the GIVD homepage provides search functions for databases meeting specific requirement and analytical function such as summary statistics.

In order to provide GIVD with a critical mass of metadata, members of the Steering Committee compiled an initial global list of vegetation plot databases ('search list') and contacted the database managers directly. The invitation to upload metadata was distributed to other vegetation scientists via several learned societies around the world as well as through personal contacts of the Steering Committee members. The search list is continuously updated based on information made available to us.

Admissible databases and required metadata

GIVD's definition of a 'vegetation-plot database' combines criteria of unambiguousness in space and time, completeness of sampling and mode of data storage. We defined as eligible for GIVD '*an electronic database that contains plot-based vegetation observations of species co-occurrence data of all vascular plants, bryophytes, lichens, and macroscopic algae (or*

Doi: 10.1111/j.1654-1103.2011.01265.x

of an explicit subset of these) sampled at a certain point in time at a given geographical location.' While any measure for species or individual performance (e.g. cover, abundance, biomass or dimensional measurements) and even presence/absence data are acceptable, it is essential that the aim of the recording was a complete species list from the defined set of plants visible in the plot at the sampling time. In order to ensure that recorded absences are true absences and to distinguish GIVD databases from the numerous databases of floristic occurrence records, we defined an upper plot size limit of 1 ha (10 000 m^2), which, however, may be exceeded in case of nested plot series that also contain subplots smaller than the threshold. There are no restrictions regarding the size of the database, its ownership (private or institutional) or its status (from planned via ongoing capture to finished), except that the database owners are willing to share their data with others, notwithstanding specific agreements or restrictions (e.g. embargo period, reduced spatial resolution or appropriate attribution).

The GIVD web portal allows registered users to upload and modify metadata for proprietary databases. After uploading the data and performing a basic consistency check, each database is assigned a unique identifier, including continent and country (e.g. EU-NL-001), which allows easy (cross-)referencing to the databases. We envisage that this identifier might become a similarly powerful tool as the permanent acronyms assigned to the world's herbaria by the *Index Herbariorum* (Thiers 2010). Information required for registration of a database includes: contact details, general description of the database and scope, ranges of recording years and plot sizes, number of non-overlapping vegetation plots, number of plots with plot size not recorded in the database, fractions of plots from different countries and number of records from different guilds. Most of the database managers also entered metadata into the various non-required fields (see Appendix S1 for overview of all fields and their definitions).

Analysis of the Content of GIVD

Existing vegetation-plot data and their coverage by GIVD

As of 30 December 2010, GIVD contained metadata for 132 vegetation-plot databases (Appendix S2). Together, these databases contained 2 444 701 non-overlapping vegetation plots (i.e. each time series and nested plot series is counted only once), and 4 540 486 distinct vegetation plot observations (which includes subplots in nested samples and repeated observations in time series). In comparison, our search list at the same date included 232 known databases (including the 132 in GIVD) with an estimated total of 3.5 million non-overlapping plots. This figure is the sum for GIVD registered databases and the estimated 1.1 million non-overlapping plots for the 100 non-registered databases (11 000 was the average number of plots in those non-registered databases for which such information was available). This means that we have presently achieved coverage of 57% in terms of databases and approximately 70% in terms of plot numbers in relation to the databases known to us. At this stage, only a few of the big databases known to the Steering Committee have not been registered in GIVD. Notwithstanding missing databases and taking into account a potential 10% overlap in records, the number of non-overlapping plots in GIVD for Europe (1 630 233) is close to the total number estimated by Schaminée et al. (2009; 1 852 000) and above their value for plots in 'central databases' of countries or federal states (1 451 000).

Geographic distribution of vegetation-plot data

Of the included databases, 83 were from Europe, 15 from North America, 13 from Asia, nine from Africa, seven from South America, two from Australasia, while three covered data from more than one continent: ForestPlots.net (00-00-001; Lopez-Gonzalez et al. 2011) with tropical forest plots and two Russian databases stretching from Europe to Asia. The ranking of continents was similar when based on plot numbers. When related to the respective land surfaces, the coverage by available non-overlapping plots was best in Europe (158 plots per 1000 km^2), moderate in North America (23), low in Australasia (8.6) and Africa (4.0), and very low in Asia (0.80) and South America (0.32), while no plots were available from Antarctica. The mean plot density calculated for all terrestrial surfaces was 16.4 plots per 1000 km^2.

Plots from 104 of the 234 countries or equivalent units (according to ISO 3166) were included in the registered databases (Appendix S3). While half of the countries were represented by a single database, only two had ten or more databases registered, notably Germany ($n = 27$) and the United States ($n = 12$). Eighteen countries (Austria, Bulgaria, Czech Republic, France, Hungary, Ireland, Japan, Namibia, The Netherlands, New Zealand, Poland, Portugal, Slovakia, Slovenia, Spain, South Africa, Taiwan, United Kingdom) had comprehensive national databases, whereas for all other countries only regional, habitat-related or project-related databases were registered. On a per-country basis, the density of plots varied widely (Figs. 1 and 2, Appendix S3), with a maximum in The Netherlands (18 031 plots/km^2).

Characteristics of the registered databases

The size of individual databases ranged from 12 plots to 600 000 plots, the latter in the Dutch National Vegetation

Doi: 10.1111/j.1654-1103.2011.01265.x

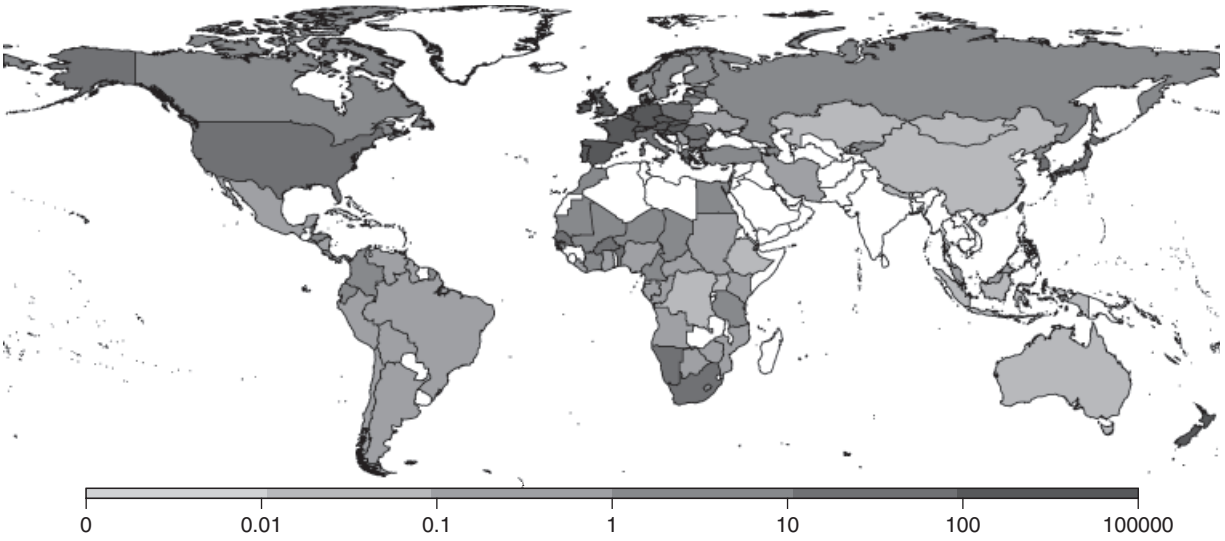

Fig. 1. Density of non-overlapping vegetation plots (per 1000 km^2) available for countries and equivalent geographical units of the world (based on GIVD as of 30 December 2010). None of the registered databases contained data for the white countries. Note the non-linear density scaling.

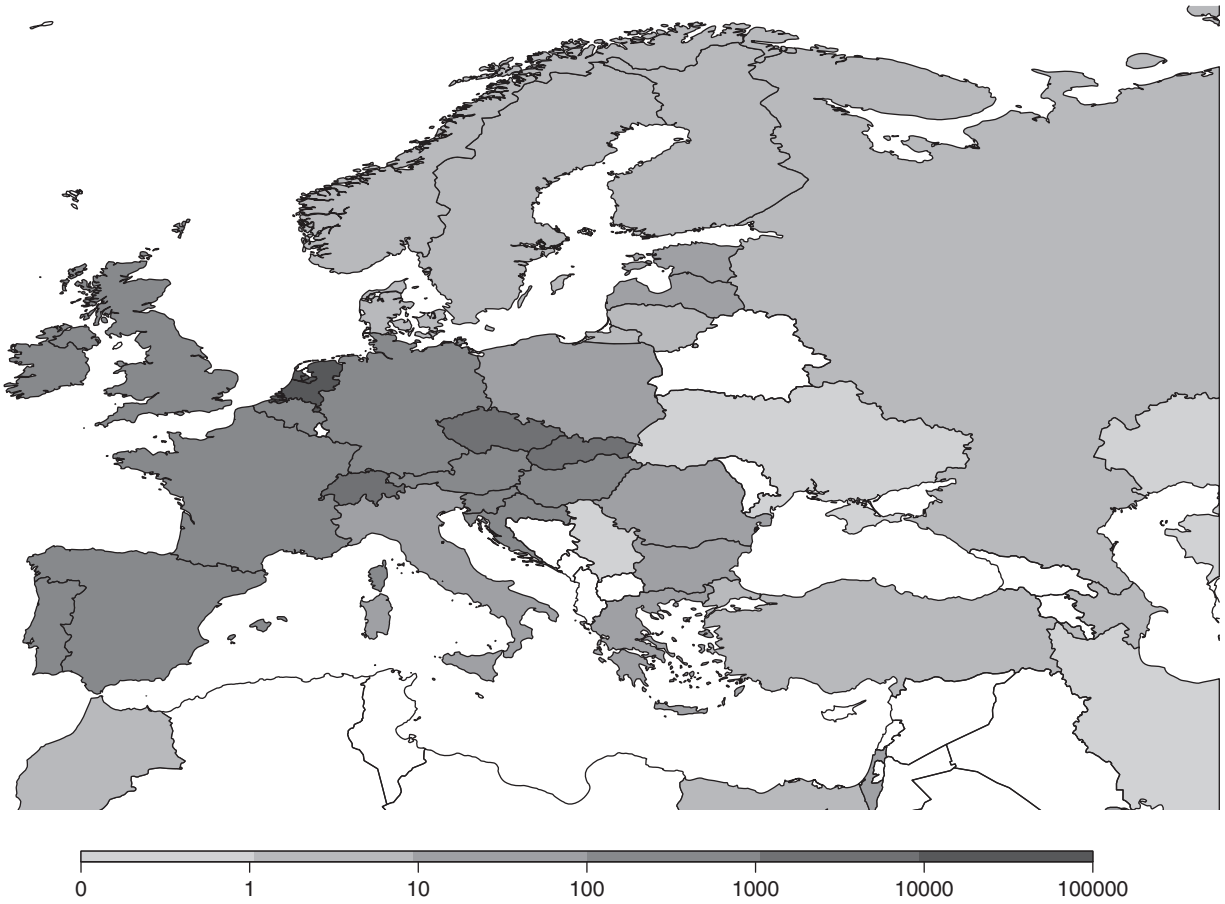

Fig. 2. Density of non-overlapping vegetation plots (per 1000 km^2) available for countries and equivalent geographical units in Europe (based on GIVD as of 30 December 2010). None of the registered databases contained data for the white countries. Note that the density scaling is non-linear and differs from Fig. 1 in order to differentiate among the European territories.

Doi: 10.1111/j.1654-1103.2011.01265.x

Database (EU-NL-001; Schaminée et al. 2006) (Fig. 3). For those databases that provided an estimate of how many non-overlapping vegetation plots existed within their scope of sampling ($n = 69$), the average 'completeness' was 65%, i.e. according to their assessment, nearly twice as many plots existed as compared to those included. Of the databases with status information, two were planned, 15 were emerging, 40 were in the process of databasing historical data, 55 had added all historical data and were continuing to accumulate new observations and 14 were finished. As for database formats, both general standard software such as Microsoft Access, Microsoft Excel or MySQL, and special programs for vegetation databases such as TURBOVEG (Hennekens & Schaminée 2001) and BIOTABase (Finckh, M., Muche, M., Schmiedel, U. & Jürgens, N., unpublished data) were used. The content of one database is freely available on-line, 28 are free upon request, 74 according to specific agreements, three of the registered databases employ 'blocking periods', while none provides data with reduced geographic resolution to external users. On-line upload functions were available for 14 databases, while 24 offered on-line search tools. The most frequent other content in the registered databases apart from plot location and species composition was vegetation classifications, synoptic tables and vegetation maps.

Characteristics of the available vegetation-plot observations

While most databases contained discrete plot records that were only recorded once, at least 33 also contained time-series records (allowing the study of temporal changes) and 13 contained nested plots (i.e. plots of different size from the same location, allowing the study of scale effects). The plot sizes in the databases ranged from 0.01 to 10 000 m^2 (250 000 m^2 for nested plot series with smaller sizes contained). The majority of plots had sizes between 10 m^2 and 1000 m^2, but for nearly one-fifth of them, the plot size was not recorded at all (Fig. 4). The geometric mean of minimum and maximum plot sizes found in individual databases ranged from 1 to 10 000 m^2, with a median across all databases of 32 m^2. The plot size variability (i.e. ratio of biggest to smallest size) within individual databases ranged from 1 to 1 000 000, with a median of 72.

The oldest plots in the databases dated from 1864 (in The Netherlands) and the numbers of plots showed a continuously increasing trend since then (Fig. 5). Woody plants (if present) were recorded in nearly all plots, while herbaceous vascular plants were recorded slightly less often (Fig. 6). The proportion of plots where terricolous bryophytes (17%) and lichens (10%) were considered was much lower, whereas records of algae and non-terricolous (i.e. epiphytic, lignicolous or saxicolous) non-vascular plants were overall negligible (Fig. 6). However, there was one large database that specialized on epiphytic lichens, with 20 000 plots (NA-US-012).

The most frequently used performance measure in the GIVD databases was cover (71% of plots with available information), followed by number of individuals (24%), dimensional measurements of individuals (22%) and biomass (21%), while the amount of pure

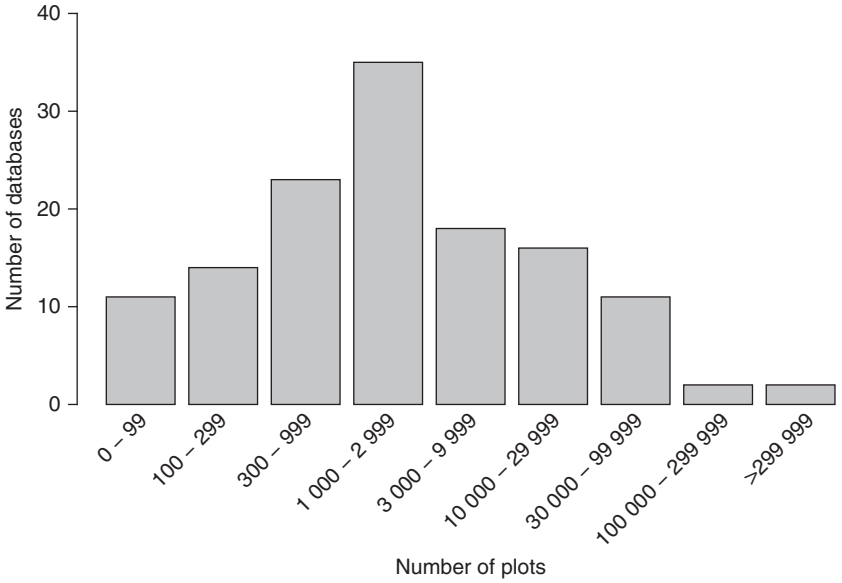

Fig. 3. Size distribution (in terms of the number of non-overlapping plots) of the 132 databases with available size data registered in GIVD on 30 December 2010.

Doi: 10.1111/j.1654-1103.2011.01265.x

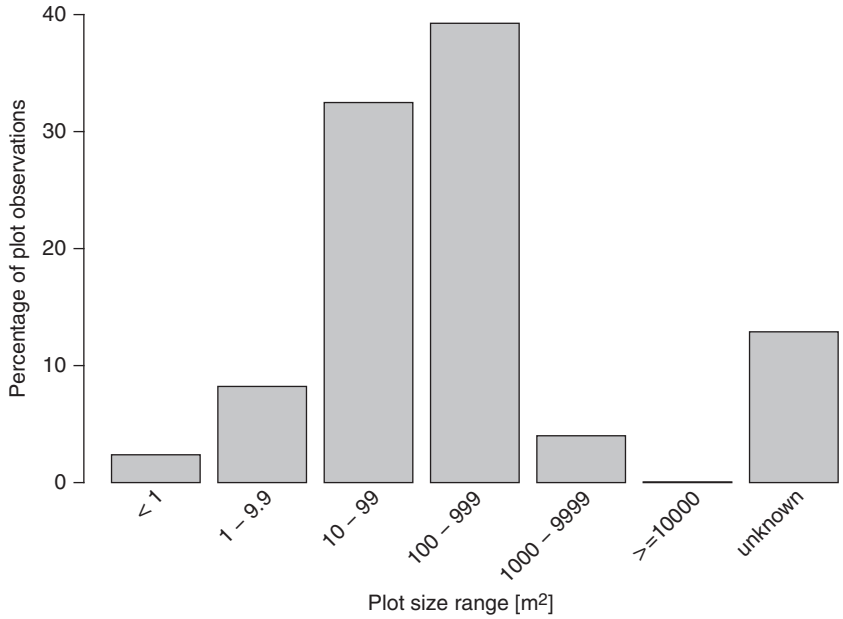

Fig. 4. Plot size distribution of the 4 172 217 vegetation plot observations (including individual observations from time-series and nested plots) with information on plot size contained in the 132 databases registered in GIVD on 30 December 2010.

presence–absence data and data with other performance measures was negligible. While 20% of the plots were accompanied by precise location data (GPS co-ordinates), 39% had been recorded with a precision of at least 1 km, 21% with a precision of 1 to 10 km and 25% were only accompanied by coarser geographic information (e.g. political or natural geographical units). The most frequently recorded plot-based environmental variables were altitude (46%), slope aspect and inclination (38%) and land-use categories (25%). By contrast, soil variables were included rarely, with coverage of certain visible fractions such as litter or stones (8%) and soil pH (4.5%) being the most frequent. Many vegetation-plot observations in the databases were assigned to a vegetation classification system, either to that of their primary source (32%) or to a uniform classification within the respective database (58%).

Discussion

Expected roles of GIVD: increased vegetation data availability

The primary goal of GIVD is to facilitate the use of vegetation data through increasing its visibility and availability worldwide. With more than 2.4 million records of non-overlapping plots whose owners are willing to share them, and using an estimate of roughly 20 species per plot, we calculate that the GIVD registered databases contain some 48 million geo-referenced species records. Compared to the approximately 217 million species records for all groups of organisms contained in the Global Biodiversity Information Facility (GBIF; http://data.gbif.org/; September 2010), this is a significant achievement when the scope (all taxa from bacteria to vertebrates), level of financial support (from institutions and countries around the world) and longevity (since the year 2000) of GBIF is considered. Only in the case of a few GIVD registered databases are the contained species records presently available via GBIF.

In addition to geo-referenced single-species records (as in floristic mapping data or occurrence records from herbaria), vegetation-plot data contain detailed and standardized information on community composition and, in many cases, structural and environmental data directly measured on the plots (Mucina et al. 2000b; Ewald 2003; Dengler et al. 2008). If we further consider the small-scale spatial resolution of species composition and environmental heterogeneity, the potential impact of the data that can be retrieved via GIVD is enormous, because they allow analyses that potentially go far beyond what traditional macroecological and biogeographical approaches have achieved. The ability to assess the conditions under which various species co-occur at moderately small scales is unique to plot data. Together with the availability of new software, adequate techniques and powerful computers, the species co-occurrence data offer novel and unique opportunities for ecological research within the rapidly expanding field of eco-informatics (e.g. Ozinga & Schaminée 2004).

Doi: 10.1111/j.1654-1103.2011.01265.x

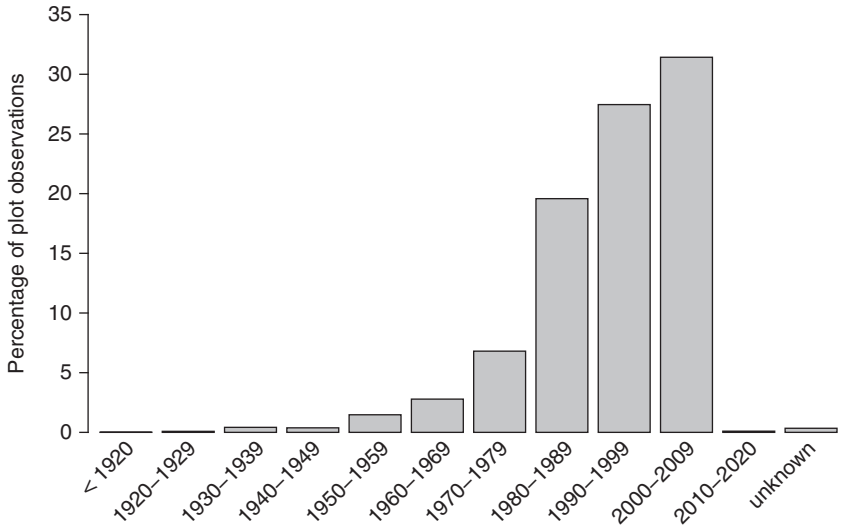

Fig. 5. Temporal distribution of the 3 526 565 vegetation plot observations (including time-series and nested plots) with information on recording year contained in the 132 databases registered in GIVD on 30 December 2010.

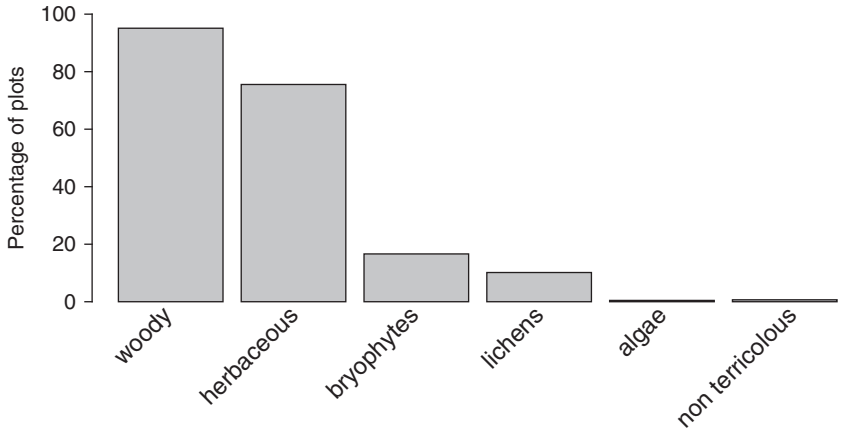

Fig. 6. Proportion of guilds recorded in the 2 444 701 non-overlapping plots with available data contained in the 132 databases registered in GIVD on 30 December 2010.

Expected roles of GIVD: incentivize data sharing

The wealth of data registered in GIVD indicates that vegetation scientists are generally willing to share their data with colleagues, but any initiative of this scale has to be sensitive to questions of data origin, authorship and ownership. Wiser et al. (2011) suggest that the major impediment to data sharing and aggregation is not a reluctance on the part of the owners, while other work indicates that the reasons for such hesitancy might be deeply embedded and complex (e.g. Parr & Cummings 2005; Nelson 2009). The motivation for not sharing data might be the assumption that there will always be someone else who 'has a smarter idea about what to do' with data than the database owner (J. Boyle cited in Nelson 2009). But even if there is not the fear of being 'scooped' by someone else producing papers from the data (Costello 2009), many researchers might feel that sharing their data will not result in due recognition of their efforts in having created or compiled data (Cassey & Blackburn 2006; Parr 2007; Costello 2009; Nelson 2009). In fact, initiatives like GIVD might sometimes be seen as imperialistic, shifting the balance of interest in data from their many original dispersed sources to a single global sink, and emphasizing metadata management rather than original data collection. Even very modest and small-scale surveys can become part of the self-respect of local, regional or national practitioners, and the identity of their

Doi: 10.1111/j.1654-1103.2011.01265.x

organisations or institutions as well as the impression of losing control or credit for work accomplished can be very threatening (Waterton & Ellis 2004). Such considerations are especially important because the present 'reward system' is practically restricted to citation of publications, whereas data publication is not usually rewarded. Many authors do not feel obliged to cite the data sets used for their analyses at all, or do it only by citing a web address but not the originator (see also Costello 2009).

Generally, we consider it is best to formally include as co-authors the originators of such databases whose data are of key importance and extensively used for analyses in a certain paper. This would give adequate credit to the work of these individuals, without which the presented analyses would not exist, as well as ensuring that the peculiarities of the data are adequately accounted for during the analyses. If such a direct involvement of the database originators is not feasible, for whatever reason, the minimum requirement should be the adequate 'citation' of their database. However, even if the third-party authors are willing, it is not always easy for them to find an appropriate citation for a specific database because only a few of these have ever been fully documented in a printed publication. One solution to this dilemma would be to establish data or databases themselves as entities that are cited and included in bibliometric analyses, equally to printed publications (Costello 2009), but this probably could only be achieved in the medium to the long term. Thus, it is probably easier and faster to overcome the observed reluctance of journal editors to publish database reports that contain a brief description of the database structure and content. Such a report would provide the database orginator(s) with a reference in a peer-reviewed journal that could be cited whenever data from that database are used. Most such published reports on vegetation-plot databases have been reasonably well cited (e.g. Mucina et al. 2000a; Wiser et al. 2001; Chytrý & Rafajová 2003; Gégout et al. 2005). The 'Data papers' in *Ecology*, which are a combination of a published abstract and a data set deposited in *Ecological Archives*, are another good example. For the vegetation databases contained in GIVD at a certain date, the Steering Committee ensured that all of them (unless the authors wished otherwise) are published as citable 'Database Reports' of one to several pages, including a standardized fact sheet in a special volume of the peer-reviewed serial *Biodiversity & Ecology* (Dengler et al. in press).

Editors of scientific journals play a key role in stimulating the establishment and easy access to vegetation databases. Not only should they change their presently restrictive attitude towards publishing database reports towards a more encouraging one, but they are also in a position to enforce appropriate reference whenever a study uses data from databases. Finally, ecological journals could follow the practice in other disciplines and require formal archiving of the underlying data in a public archive when an article is published (Costello 2009; Nelson 2009). The Central European journal *Tuexenia* adopted such a practice for plot data in 2005, and the data gathered are included in the database VegetWeb (EU-DE-013; Ewald et al. 2007). Similarly, the US Federal standard for vegetation classification mandates that plot data be publically archived with a unique web accession code, such as that used by VegBank (NA-US-002; US FGDC 2008).

An interesting observation could be made with regard to the Dutch National Vegetation Database (EU-NL-001). At the beginning, in 1988, people were reluctant to share their data and add them to the national database, but nowadays the offer of data from all kinds of source (private, institutional, governmental) is enormous and almost difficult to handle. The experience with the Czech National Phytosociological Database (EU-CZ-001; Chytrý & Rafajová 2003) is very similar. The national databases guarantee the visibility of data and their use within a wide spectrum of applications. Researchers involved in the development of the 'The West African Vegetation Database' (AF-00-001) identified as major incentives for online data contribution: (a) simple user interface; (b) access to data controlled by data contributors, (c) using the database as to showcase data and promote collaborations and (d) data security, including regular backups of the data (Janßen et al. 2011). We hope that GIVD will stimulate similar developments in other countries.

Expected roles of GIVD: stimulation of vegetation surveys in under-sampled regions

Data density varies considerably among continents and among countries or regions within continents. For example, while countries with a strong tradition of phytosociology (Braun-Blanquet 1964; Westhoff & van der Maarel 1973; Ewald 2003; Dengler et al. 2008) usually have good data coverage, there are some exceptions, such as Italy, Poland and Romania, where a negligible proportion of their plot legacy is digitized (Fig. 2 and Appendix S3). On the global scale, data availability appears to be mostly good in the temperate zones, moderate in mediterranean and boreal zones, but poor in tropical and arctic zones; for many countries in the subtropical zone, not a single plot is registered (Fig. 1). It is strange that relatively more plot data are available from tropical than from subtropical countries. Hence, one of the benefits of GIVD is that it can direct future effort regarding the collection of new field data and electronic capture of existing vegetation plot observations to those regions on Earth that are presently not well covered.

Doi: 10.1111/j.1654-1103.2011.01265.x

Problems and solutions beyond GIVD

Despite the potential of GIVD to help solve the problems of vegetation data availability and sharing, many issues still prevent the successful integration of data for analyses from multiple vegetation-plot databases. Here we briefly summarize the major impediments and consider possible solutions, when available:

(1) *Technical problems*: Important impediments to vegetation data exchange and combination come from the fact that vegetation data are stored in different database structures and use different vocabularies (Wiser et al. 2011). The solution is to adopt and conform to international standards for data exchange (Berendsohn et al. 1999; Jones et al. 2006; Williams et al. 2006). In the broader field of ecology, the availability and adoption of EML has facilitated automated data exchange (Jones et al. 2006). In the realm of vegetation-plot data, the wide adoption of TURBOVEG means that many plot records have basically the same conformation and could be relatively easily merged into a single data set. The first broad-based international standard for vegetation plot data was the VegBank XML standard (http://www.vegbank.org), which was designed to be flexible enough to accommodate all significant forms of vegetation plot data. The Veg-X data exchange standard (Wiser et al. 2011) was built on the foundation laid by VegBank XML and its derivative VegetWeb, and has been adopted by TDWG as a formal component of the standard for taxon occurrence data.

(2) *Taxonomic problems*: It is far from trivial to combine data from different sources correctly and without unnecessarily losing taxonomic information and precision. Even if the data to be combined are based on the same database system (such as TURBOVEG), the species lists (and other authority data) often vary widely. For example, combining data based on different checklists is not an easy task as species names and concepts often have a many-to-many relationship, where one name can refer to several different circumscriptions and one circumscription can refer to many different names. A conceptual basis for the solution was first elaborated by Berendsohn (1995; see also Berendsohn et al. 2003), and the first vegetation-plot database to fully adopt a taxon concept model was VegBank (http://www.vegbank.org). The VegBank solution, which expands on the 'Berlin model' of Berendsohn through recognition of the transient nature of on-line data sources and the need to recognize continuous versioning, was fully embraced in the US standard for vegetation-plot data (US FGDC 2008; Jennings et al. 2009). Further suggestions have been provided in Kennedy et al. (2005), Franz et al. (2008), Franz & Peet (2009) and Jansen & Dengler (2010). All of these works emphasize the need to identify taxon occurrences not just to name but also to concept (a 'name-reference couplet').

(3) *Variability of plot size*: Plot sizes vary widely, not only among but also within vegetation types and regions, and even in time (Fig. 4; see also Chytrý & Otýpková 2003). This imposes serious problems for joint analyses, as nearly all aspects of biodiversity are scale-dependent. Investigations into species richness (Crawley & Harral 2001; Dengler 2009a), species fidelity and classification (Dengler et al. 2009), vegetation–environment relationships (Reed et al. 1993) and ordination patterns (Otýpková & Chytrý 2006) can be seriously affected by plot size. For the future, the consequences from these findings should be to use standardized plot sizes (Chytrý & Otýpková 2003; Dengler 2003). Application of multi-scale sampling can also be considered (Whittaker et al. 1979; Stohlgren et al. 1995; Peet et al. 1998; Dengler 2009b). To overcome the potential confounding effects of varying plot sizes within historical data, additional studies such as those already conducted by Otýpková & Chytrý (2006) and Dengler et al. (2009) are needed to quantify these effects and to develop appropriate analytical solutions. Alternatively, only data with identical or at least similar plot sizes can be chosen from databases for a specific analysis (Michl et al. 2010), an approach that is facilitated through GIVD by making the retrieval of the respective data easier.

(4) *Other aspects of sampling methodology:* Although vegetation plot data collection protocols are moderately compatible, the differences in their implementation can bias the value of the 'performance measure' sampled. For instance, in forest ecology, the diameter measurement and post-field data management of trees with buttresses can potentially bias basal area and biomass calculations (Clark 2002; Phillips et al. 2002). Similarly, species presence and cover observations can be affected by the methodology used, number of observers and observation time (Kercher et al. 2003; Archaux et al. 2006; Vittoz & Guisan 2007). A further complication arises when vegetation plot data are used for analysing vegetation change, as turnover calculations of a given performance measure can be affected by erroneous measurements or observations (Vittoz & Guisan 2007). Additionally, researchers need to consider the differences in sampling intervals among plots and standardize census length when comparing turnover rates from different sources (Lewis et al. 2004). A more detailed documentation of the sampling methodologies used in the databases registered in GIVD in the future would facilitate the selection of compatible data sets for inclusion in large-scale analyses.

(5) *Geographical and ecological biases*: In the history of phytosociology, preferential location of plots in places where assumed diagnostic species occurred together was common practice, leading to biased species co-occurrence patterns in the data sets (Glavač 1996; Botta-Dukát et al.

Doi: 10.1111/j.1654-1103.2011.01265.x

2007; Diekmann et al. 2007). Even when the sampling within each single project meets the statistical requirements, the combination of data from many projects in one database inevitably leads to unequal and non-random distribution of vegetation-plot observations in geographical and ecological space as well as in time. However, the sheer volume of data, especially as more and more databases become available, might offer a partial solution to this problem. For example, biases caused by subjective plot placement of individual researchers might be 'averaged out' when data from numerous researchers are combined. Large databases (or combinations of databases) further offer the possibility to stratify data and analyse their more balanced subsamples (Knollová et al. 2005; Haveman & Janssen 2008; Lengyel et al. in press).

Future perspectives

While this article presents only a snapshot on the data presently available in GIVD, we anticipate further growth. Our longer-term vision is to develop GIVD in ways similar to Metacat (Jones et al. 2006), so that, ultimately, users who query GIVD will not only receive information on which databases contain data suitable for the intended analyses, but they will also discover other data from distributed databases, with GIVD acting as the central node. The next logical step would be integration of all plot data into a central or distributed database so that plot data can be discovered and retrieved as needed. The VegBank archive is the closest implementation to this model currently available in that it accepts and distributes all forms of plot data, can be queried for data that meet a given need and data can be cited by accession code or data set. The BIEN project at the US National Center for Ecological Analysis and Synthesis (NCEAS) represents a significant step forward in that it attempts to provide access in one public archive to all western hemisphere plant occurrence data and includes both single observations and co-occurrences in plots (http://www.nceas.ucsb.edu/projects/12290). The BIEN data are expected to be publicly available in late 2011 and could well serve as a model for distribution of co-occurrence data worldwide. GIVD has a practical advantage over VegBank and BIEN in that data are controlled by the providers, which results in much greater willingness of data owners to participate. The GIVD Steering Committee will work toward integrated access to data registered in GIVD via a standard query system coupled with data distribution based on the Veg-X data exchange protocol (Wiser et al. 2011).

At this stage, we can only estimate the number of potential 'customers' for GIVD. However, as an indication of likely importance, initiatives such as Satellite-driven Earth Observation Systems like GEOSS (Lautenbacher 2006) badly need high-quality baseline data on the occurrence of plant species or plant assemblages with geo-referenced plots of a known reference size and inventory. This in-situ information will help to calibrate and validate products of Earth Observation Systems on land cover, tree cover and vegetation height (Muchoney & Williams 2010), and could significantly contribute to the success of GEOSS in the future. Development of second-generation dynamic global vegetation models (second DGVM; Fisher et al. 2010) will also require much more information on the spatial distribution of plant species, their ecology and associated plant functional data than was incorporated in past efforts. By coupling species-specific trait characteristics (e.g. mean plant height, specific leaf area, growth form) found in trait databases, such as LEDA (Kleyer et al. 2008) or TRY (http://www.try-db.org), with plot-based distribution information on those species, GIVD could support further refinement of DGVMs.

Conclusions

The large amount of data registered in GIVD so far demonstrates that numerous vegetation plots are available in digital format and many researchers and institutions are willing to share these data. Beyond quantity, vegetation-plot databases provide types of data that are otherwise rare within GBIF. Instead of just single-species presence–absence records, they contain species co-occurrence and importance data, mostly combined with structural and environmental data, and sometimes arranged in connected series of records over time (time series) or across multiple spatial scales (nested plots). Further, they combine small spatial grain with large spatial extent, a property particularly suitable for studying the causations underlying macroecological patterns and for monitoring global change, while the majority of ecological studies conducted to date are either small grain/small extent or large grain/large extent (Dengler 2009b).

We consider the establishment of GIVD a major step towards more intensive and more effective use of vegetation-plot data. Other problems connected with vegetation-plot databases still need to be solved, such as sampling under-represented geographic regions, providing better incentives for data collection and sharing, developing user-friendly exchange standards between databases, as well as tools to analyse and reduce confounding effects of sampling biases. Suggestions for further improvement of content and functionality of GIVD are welcome.

Acknowledgements

The idea of GIVD was conceived by J.D., J.E., M.F. and F.J. and then developed further by the GIVD Steering

Doi: 10.1111/j.1654-1103.2011.01265.x

Committee, including all persons listed as authors. F.G. and F.J. are in charge for the technical realization of the on-line platform, while mainly J.D., F.G. and F.J. were responsible for the collation of data and their quality control. Numerical analyses were carried out by F.J. and J.D., and the figures were prepared by F.J. The text was first drafted by J.D. with substantial contributions by R.K.P., M.D.C., M.C., J.E., J.O. and G.L.-G., while all authors contributed to its revision. We acknowledge the German Federal Agency of Nature Conservation (BfN), the project BIOTA Southern Africa and the Floristisch-Soziologische Arbeitsgemeinschaft e. V. (FlorSoz) for financial support of the meeting in Hamburg where the GIVD idea was born. Finally, we thank Ingolf Kühn and two anonymous referees for their helpful comments on the original version of this manuscript.

References

Archaux, F., Gosselin, F., Bergès, L. & Chevalier, R. 2006. Effects of sampling time, species richness and observer on the exhaustiveness of plant censuses. *Journal of Vegetation Science* 17: 299–306.

Baker, T.R., Phillips, O.L., Malhi, Y., Almeida, S., Arroyo, L., Di Fiore, A., Killeen, T.J., Laurance, S.G., Laurance, W.F., Lewis, S.L., Lloyd, J., Monteagudo, A., Neill, D.A., Patiño, S., Pitman, N.C., Silva, J.N. & Vásquez Martínez, R. 2004. Variation in wood density determines spatial patterns in Amazonian forest biomass. *Global Change Biology* 10: 545–562.

Bakker, J.P., Elzinga, J.A. & de Vries, Y. 2002. Effects of long-term cutting in a grassland system: possibilities for restoration of plant communities on nutrient-poor soils. *Applied Vegetation Science* 5: 107–120.

Bekker, R.M., van der Maarel, E., Bruelheide, H. & Woods, K. 2007. Long-term datasets: from descriptive to predictive data using ecoinformatics. *Journal of Vegetation Science* 18: 457–462.

Berendsohn, W.G. 1995. The concept of 'potential taxa' in databases. *Taxon* 44: 207–212.

Berendsohn, W.G., Anagnostopolous, A., Hagedorn, G., Jakupovic, J., Nimis, P.L., Valdés, B., Güntsch, A., Pankhurst, R.J. & White, R.J. 1999. A comprehensive reference model for biological collections and surveys. *Taxon* 48: 511–562.

Berendsohn, W.G., Döring, M., Geoffroy, M., Glück, K., Güntsch, A., Hahn, A., Kusber, W.-H., Li, J., Röpert, D. & Specht, F. 2003. The Berlin model: a concept-based taxonomic information model. In: Berendsohn, W.G. (ed.) *MoReTax – Handling factual information linked to taxonomic concepts in biology*. pp. 15–42. Federal Agency for Nature Conservation, Bonn, DE.

Berg, C., Dengler, J. & Abdank, A. (eds.) 2001. *Die Pflanzengesellschaften Mecklenburg-Vorpommerns und ihre Gefährdung – Tabellenband*. Weissdorn, Jena, DE.

Botta-Dukát, Z., Kovács-Láng, E., Rédei, T., Kertész, M. & Garadnai, J. 2007. Statistical and biological consequences of preferential sampling in phytosociology: theoretical considerations and a case study. *Folia Geobotanica* 42: 141–152.

Braun-Blanquet, J. 1964. *Pflanzensoziologie – Grundzüge der Vegetationskunde*. 3rd ed. Springer, Wien, AT.

Brown, R.L. & Peet, R.K. 2003. Diversity and invasibility of southern Appalachian plant communities. *Ecology* 84: 32–39.

Brzeziecki, B., Kienast, F. & Wildi, O. 1993. A simulated map of the potential natural forest vegetation of Switzerland. *Journal of Vegetation Science* 4: 499–508.

Cassey, P. & Blackburn, T.M. 2006. Reproducibility and repeatability in ecology. *BioScience* 56: 958–959.

Chytrý, M. (ed.) 2007. *Vegetace České republiky 1. Travinná a keřičková vegetace – Vegetation of the Czech Republic 1: Grassland and heathland vegetation*. Academia, Praha, CZ.

Chytrý, M. & Otýpková, Z. 2003. Plot sizes used for phytosociological sampling of European vegetation. *Journal of Vegetation Science* 14: 563–570.

Chytrý, M. & Rafajová, M. 2003. Czech national phytosociological database: basic statistics of the available vegetation-plot data. *Preslia* 75: 1–15.

Chytrý, M., Jarošík, V., Pyšek, P., Hájek, O., Knollová, I., Tichý, L. & Danihelka, J. 2008a. Separating habitat invasibility by alien plants from the actual level of invasion. *Ecology* 89: 1541–1553.

Chytrý, M., Maskell, L.C., Pino, J., Pyšek, P., Vila, M., Font, X. & Smart, S.M. 2008b. Habitat invasions by alien plants: a quantitative comparison among Mediterranean, subcontinental and oceanic regions of Europe. *Journal of Applied Ecology* 45: 448–458.

Clark, D.A. 2002. Are tropical forests an important global carbon sink? Revisiting the evidence from long-term inventory plots. *Ecological Applications* 12: 3–7.

Costello, M.J. 2009. Motivating online publication of data. *BioScience* 59: 418–427.

Coudun, C. & Gégout, J.-C. 2005. Ecological behaviour of herbaceous forest species along a pH gradient: a comparison between oceanic and semicontinental regions in northern France. *Global Ecology and Biogeography* 14: 263–270.

Coudun, C. & Gégout, J.-C. 2007. Quantitative prediction of the distribution and abundance of *Vaccinium myrtillus* with climatic and edaphic factors. *Journal of Vegetation Science* 18: 517–524.

Crawley, M.J. & Harral, J.E. 2001. Scale dependence in plant biodiversity. *Science* 291: 864–868.

Dengler, J. 2003. *Entwicklung und Bewertung neuer Ansätze in der Pflanzensoziologie unter besonderer Berücksichtigung der Vegetationsklassifikation*. Galunder, Nümbrecht, DE.

Dengler, J. 2009a. Which function describes the species–area relationship best? – A review and empirical evaluation. *Journal of Biogeography* 36: 728–744.

Dengler, J. 2009b. A flexible multi-scale approach for standardised recording of plant species richness patterns. *Ecological Indicators* 9: 1169–1178.

Dengler, J., Chytrý, M. & Ewald, J. 2008. Phytosociology. In: Jørgensen, S.E. & Fath, B.D. (eds.) *Encyclopedia of ecology*. pp. 2767–2779. Elsevier, Oxford, UK.

Dengler, J., Löbel, S. & Dolnik, C. 2009. Species constancy depends on plot size – a problem for vegetation classification and how it can be solved. *Journal of Vegetation Science* 20: 754–766.

Dengler, J., Chytrý, M., Ewald, J., Finckh, M., Jansen, F., Lopez-Gonzalez, G., Oldeland, J., Peet, R.K. & Schaminée, J.H.J. (eds.) (in press) *Vegetation databases for the 21st century*. Biocentre Klein Flottbek and Botanical Garden, Hamburg, DE.

Diekmann, M., Kühne, A. & Isermann, M. 2007. Random versus non-random sampling: effects on patterns of species abundance, species richness and vegetation–environment relationships. *Folia Geobotanica* 42: 179–190.

Duckworth, J.C., Bunce, R.G.H. & Malloch, A.J.C. 2000. Vegetation–environment relationships in Atlantic European calcareous grasslands. *Journal of Vegetation Science* 11: 15–22.

Dupré, C., Stevens, C.J., Ranke, T., Bleeker, A., Peppler-Lisbach, C., Gowing, D.J.G., Dise, N.B., Dorland, E., Bobbink, R. & Diekmann, M. 2010. Changes in species richness and composition in European acidic grasslands over the past 70 years: the contribution of cumulative atmospheric nitrogen deposition. *Global Change Biology* 16: 344–357.

Ewald, J. 2001. Der Beitrag pflanzensoziologischer Datenbanken zur vegetationsökologischen Forschung. *Berichte der Reinhold-Tüxen-Gesellschaft* 13: 53–69.

Ewald, J. 2002. A probabilistic approach to estimating species pools from large compositional matrices. *Journal of Vegetation Science* 13: 191–198.

Ewald, J. 2003. A critique for phytosociology. *Journal of Vegetation Science* 14: 291–296.

Ewald, J., Conrad, S. & Kleikamp, M. 2007. Vegetationsaufnahmen aus Tuexenia Band 25 und Band 26 gehen online. *Tuexenia* 27: 417–420.

Fisher, R., McDowell, N., Purves, D., Moorcroft, P., Sitch, S., Cox, P., Huntingford, C., Meir, P. & Woodward, F.I. 2010. Assessing uncertainties in a second-generation dynamic vegetation model caused by ecological scale limitations. *New Phytologist* 187: 666–681.

Franz, N.M. & Peet, R.K. 2009. Towards a language for mapping relationships among taxonomic concepts. *Systematics and Biodiversity* 7: 5–20.

Franz, N.M., Peet, R.K. & Weakley, A.S. 2008. On the use of taxonomic concepts in support of biodiversity research and taxonomy. In: Wheeler, Q.D. (ed.) *The new taxonomy*. pp. 63–86. Taylor & Francis, Boca Raton, FL, US.

Fridley, J.D., Vandermast, D.B., Kuppinger, D.M., Manthey, M. & Peet, R.K. 2007. Co-occurrence-based assessment of habitat generalists and specialists: a new approach for the measurement of niche width. *Journal of Ecology* 95: 707–722.

Gégout, J.-C., Coudun, C., Bailly, G. & Jabiol, B. 2005. EcoPlant: a forest site database linking floristic data with soil and climate variables. *Journal of Vegetation Science* 16: 257–260.

Glavač, V. 1996. *Vegetationsökologie – Grundfragen, Aufgaben, Methoden*. Fischer, Jena, DE.

Hall, G.M.J., Wiser, S.K., Allen, R.B., Beets, P.N. & Goulding, C.J. 2001. Strategies to estimate national forest carbon stocks from inventory data: the 1990 New Zealand baseline. *Global Change Biology* 7: 389–403.

Haveman, R. & Janssen, J.A.M. 2008. The analysis of long-term changes in plant communities using large databases: the effect of stratified resampling. *Journal of Vegetation Science* 19: 355–362.

Hennekens, S.M. & Schaminée, J.H.J. 2001. TURBOVEG, a comprehensive data base management system for vegetation data. *Journal of Vegetation Science* 12: 589–591.

Holtland, W.J., ter Braak, C.J.F. & Schouten, M.G.C. 2010. Iteratio: calculating environmental indicator values for species and relevés. *Applied Vegetation Science* 13: 369–377.

Janßen, T., Schmidt, M., Dressler, S., Hahn-Hadjali, K., Hien, M., Konaté, S., Lykke, A.M., Mahamane, A., Sambou, B., Sinsin, B., Thiombiano, A., Wittig, R. & Zizka, G. 2011. Addressing data property rights concerns and providing incentives for collaborative data pooling: the West African Vegetation Database. *Journal of Vegetation Science* 22, doi: 10.1111/j.1654-1103.2011.01271.x.

Jansen, F. & Dengler, J. 2010. Plant names in vegetation databases – a neglected source of bias. *Journal of Vegetation Science* 21: 1179–1186.

Jennings, M.D., Faber-Langendoen, D., Loucks, O.L., Peet, R.K. & Roberts, D. 2009. Standards for associations and alliances of the U.S. National Vegetation Classification. *Ecological Monographs* 79: 173–199.

Jones, M.B., Schildhauer, M.P., Reichman, O.J. & Bowers, S. 2006. The new bioinformatics: integrating ecological data from the gene to the biosphere. *Annual Review of Ecology, Evolution and Systematics* 37: 519–544.

Kennedy, J., Kukla, R. & Paterson, T. 2005. Scientific names are ambiguous as identifiers for biological taxa: their context and definition are required for accurate data integration. In: Ludäscher, B. & Raschid, L. (eds.) *Data integration in the life sciences. Proceedings of the second international workshop, San Diego, CA, July 20–22, DILS 2005, LNBI 3615, 2005*. pp. 80–95. Springer, Berlin, DE.

Kercher, S.M., Frieswyk, C.B. & Zedler, J.B. 2003. Effects of sampling teams and estimation methods on the assessment of plant cover. *Journal of Vegetation Science* 14: 899–906.

Kleyer, M., Bekker, R.M., Knevel, I.C., Bakker, J.P., Thompson, K., Sonnenschein, M., Poschlod, P., van Groenendael, J.M., Klimeš, L., Klimešová, J., Klotz, S., Rusch, G.M., Hermy, M., Adriaens, D., Boedeltje, G., Bossuyt, B., Dannemann,

A., Endels, P., Götzenberger, L., Hodgson, J.G., Jackel, A.-K., Kühn, I., Kunzmann, D., Ozinga, W.A., Römermann, C., Stadler, M., Schlegelmilch, J., Steendam, H.J., Tackenberg, O., Wilmann, B., Cornelissen, J.H.C., Eriksson, O., Garnier, E. & Peco, B. 2008. The LEDA Traitbase: a database of life-history traits of the Northwest European flora. *Journal of Ecology* 96: 1266–1274.

Knollová, I., Chytrý, M., Tichý, L. & Hájek, O. 2005. Stratified resampling of phytosociological databases: some strategies for obtaining more representative data sets for classification studies. *Journal of Vegetation Science* 16: 479–486.

Lautenbacher, C.C. 2006. The global earth observation system of systems: science serving society. *Space Policy* 22: 8–11.

Le Duc, M., Yang, L. & Marrs, R. 2007. A database application for long-term ecological field experiments. *Journal of Vegetation Science* 18: 509–516.

Lengyel, A., Chytrý, M. & Tichý, L. in press. Heterogeneity-constrained random resampling of phytosociological databases. *Journal of Vegetation Science* 22: 175–183.

Lenoir, J., Gégout, J.-C., Marquet, P.A., de Ruffray, P. & Brisse, H. 2008. A significant upward shift in plant species optimum elevation during the 20th century. *Science* 320: 1768–1771.

Lewis, S.L., Phillips, O.L., Sheil, D., Vinceti, B., Baker, T.R., Brown, S., Graham, A.W., Higuchi, N., Hilbert, D.W., Laurance, W.F., Lejoly, J., Malhi, Y., Monteagudo, A., Nunez, P., Sonke, B., Nur Supardi, M.N., Terborgh, J.W. & Vasquez, M.R. 2004. Tropical forest tree mortality, recruitment and turnover rates: calculation, interpretation, and comparison when census intervals vary. *Journal of Ecology* 92: 929–944.

Lewis, S.L., Lopez-Gonzalez, G., Sonke, B., Affum-Baffoe, K., Baker, T.R., Ojo, L.O., Phillips, O.L., Reitsma, J.M., White, L., Comiskey, J.A., Djuikouo, M.N., Ewango, C.E.N., Feldpausch, T.R., Hamilton, A.C., Gloor, M., Hart, T., Hladik, A., Lloyd, J., Lovett, J.C., Makana, J.R., Malhi, Y., Mbago, F.M., Ndangalasi, H.J., Peacock, J., Peh, K.S.H., Sheil, D., Sunderland, T., Swaine, M.D., Taplin, J., Taylor, D., Thomas, S.C., Votere, R. & Woll, H. 2009. Increasing carbon storage in intact African tropical forests. *Nature* 457: 1003–1006.

Loidi, J., Biurrun, I., Campos, J.A., García-Mijangos, I. & Herrera, M. 2010. A biogeographical analysis of the European Atlantic lowland heathlands. *Journal of Vegetation Science* 21: 832–842.

Lopez-Gonzalez, G., Lewis, S.L., Burkitt, M. & Phillips, O.L. 2011. ForestPlots.net: a web application and research tool to manage and analyse tropical forest plot data.. *Journal of Vegetation Science* 22, doi: 10.1111/j.1654-1103.2011.01312.x.

Lososová, Z. & Láníková, D. 2010. Differences in trait compositions between rocky natural and artificial habitats. *Journal of Vegetation Science* 21: 520–530.

Lososová, Z., Chytrý, M., Cimalová, Š., Kropáč, Z., Otýpková, Z., Pyšek, P. & Tichý, L. 2004. Weed vegetation of arable land in Central Europe: gradients in diversity and species composition. *Journal of Vegetation Science* 15: 415–422.

Lososová, Z., Chytrý, M., Kühn, I., Hájek, O., Horáková, V., Pyšek, P. & Tichý, L. 2006. Patterns of plant traits in annual vegetation of man-made habitats in central Europe. *Perspectives in Plant Ecology, Evolution and Systematics* 8: 69–81.

Malhado, A.C.M., Malhi, Y., Whittaker, R.J., Ladle, R.J., ter Steege, H., Aragão, L.E.O.C., Quesada, C.A., Araújo, A.M., Phillips, O.L., Peacock, J., Lopez-Gonzalez, G., Baker, T.R., Butt, N., Anderson, L.O., Arroyo, L., Almeida, S., Higuchi, N., Killeen, T., Monteagudo, A., Neill, D., Pitman, N., Prieto, A., Salomão, R., Silva, N., Vásquez-Martínez, R. & Laurance, W.F. 2009. Spatial trends in leaf size of Amazonian rainforest trees. *Biogeosciences* 6: 1563–1576.

Marage, D. & Gégout, J.-C. 2009. Importance of soil nutrients in the distribution of forest communities on a large geographical scale. *Global Ecology and Biogeography* 18: 88–97.

Maskell, L.C., Smart, S.M., Bullock, J.M., Thompson, K. & Stevens, C.J. 2010. Nitrogen deposition causes widespread loss of species richness in British habitats. *Global Change Biology* 16: 671–679.

Michl, T., Dengler, J. & Huck, S. 2010. Montane-subalpine tall-herb vegetation (*Mulgedio-Aconitetea*) in central Europe: large-scale synthesis and comparison with northern Europe. *Phytocoenologia* 40: 117–154.

Muchoney, D. & Williams, M. 2010. Building a 2010 biodiversity conservation data baseline: contributions of the Group on Earth Observations. *Ecological Research* 25: 937–946.

Mucina, L., Bredenkamp, G.J., Hoare, D.B. & McDonald, D.J. 2000a. A national vegetation database for South Africa. *South African Journal of Science* 96: 497–498.

Mucina, L., Schaminée, J.H.J. & Rodwell, J.S. 2000b. Common data standards for recording relevés in field survey for vegetation classification. *Journal of Vegetation Science* 11: 769–772.

Mucina, L., Rutherford, M.C. & Powrie, L.W. 2006. Logic of the map: approaches and procedures. In: Mucina, L. & Rutherford, M.C. (eds.) *The vegetation of South Africa, Lesotho and Swaziland*. pp. 12–29. SANBI, Pretoria, ZA.

Münzbergová, Z. & Herben, T. 2004. Identification of suitable unoccupied habitats in metapopulation studies using co-occurrence of species. *Oikos* 105: 408–414.

Nelson, B. 2009. Empty archives. *Nature* 461: 160–163.

Otýpková, Z. & Chytrý, M. 2006. Effects of plot size on the ordination of vegetation samples. *Journal of Vegetation Science* 17: 465–472.

Ozinga, W.A. & Schaminée, J.H.J. 2004. The development of ecological information systems, a new tool in biological research. In: *Recent Research Developments in Environmental Biology*, 1: 531–551.

Ozinga, W.A., Hennekens, S.M., Schaminée, J.H.J., Bekker, R.M., Prinzing, A., Bonn, S., Poschlod, P., Tackenberg, O.,

Thompson, K., Bakker, J.P. & van Groenendael, J.M. 2005. Assessing the relative importance of dispersal in plant communities using an ecoinformatics approach. *Folia Geobotanica* 40: 53–67.

Ozinga, W.A., Hennekens, S.M., Schaminée, J.H.J., Smits, N.A.C., Bekker, R.M., Römermann, C., Klimeš, L., Bakker, J.P. & van Groenendael, J.M. 2007. Local above-ground persistence of vascular plants: life-history trade-offs and environmental constraints. *Journal of Vegetation Science* 18: 489–497.

Ozinga, W.A., Römermann, C., Bekker, R.M., Prinzing, A., Tamis, W.L.M., Schaminée, J.H.J., Hennekens, S.M., Thompson, K., Poschlod, P., Kleyer, M., Bakker, J.P. & van Groenendael, J.M. 2009. Dispersal failure contributes to plant losses in NW Europe. *Ecology Letters* 12: 66–74.

Parr, C.S. 2007. Open sourcing ecological data. *BioScience* 57: 309–310.

Parr, C.S. & Cummings, M.P. 2005. Data sharing in ecology and evolution. *Trends in Ecology and Evolution* 20: 362–363.

Peet, R.K., Wentworth, T.R. & White, P.S. 1998. A flexible, multipurpose method for recording vegetation composition and structure. *Castanea* 63: 262–274.

Reed, R.A., Peet, R.K., Palmer, M.W. & White, P.S. 1993. Scale dependence of vegetation–environment correlations: a case study of a North Carolina piedmont woodland. *Journal of Vegetation Science* 4: 329–340.

Phillips, O.L., Malhi, Y., Vinceti, B., Baker, T., Lewis, S.L., Higuchi, N., Laurance, W.F., Núñez, V.P., Vásquez, M.R., Laurance, S.G., Ferreira, L.V., Stern, M., Brown, S. & Grace, J. 2002. Changes in the biomass of tropical forests: evaluating potential biases. *Ecological Applications* 12: 576–587.

Phillips, O.L., Aragao, L., Lewis, S.L., Fisher, J.B., Lloyd, J., Lopez-Gonzalez, G., Malhi, Y., Monteagudo, A., Peacock, J., Quesada, C.A., van der Heijden, G., Almeida, S., Amaral, I., Arroyo, L., Aymard, G., Baker, T.R., Banki, O., Blanc, L., Bonal, D., Brando, P., Chave, J., de Oliveira, A.C.A., Cardozo, N.D., Czimczik, C.I., Feldpausch, T.R., Freitas, M.A., Gloor, E., Higuchi, N., Jimenez, E., Lloyd, G., Meir, P., Mendoza, C., Morel, A., Neill, D.A., Nepstad, D., Patino, S., Penuela, M.C., Prieto, A., Ramirez, F., Schwarz, M., Silva, J., Silveira, M., Thomas, A.S., ter Steege, H., Stropp, J., Vasquez, R., Zelazowski, P., Davila, E.A., Andelman, S., Andrade, A., Chao, K.J., Erwin, T., Di Fiore, A., Honorio, E., Keeling, H., Killeen, T.J., Laurance, W.F., Cruz, A.P., Pitman, N.C.A., Vargas, P.N., Ramirez-Angulo, H., Rudas, A., Salamao, R., Silva, N., Terborgh, J. & Torres-Lezama, A. 2009. Drought sensitivity of the Amazon rainforest. *Science* 323: 1344–1347.

Phillips, O.L., van der Heijden, G., López-González, G., Aragão, L.E.O.C., Lewis, S.L., Lloyd, J., Malhi, Y., Monteagudo, A., Almeida, S., Alvarez, D.E., Amaral, I., Andelman, S., Andrade, A., Arroyo, L., Aymard, G., Baker, T.R., Blanc, L., Bonal, D., Alves de Oliveira, A.C., Chao, K.-J., Dávila, C.N., da Costa, L., Feldpausch, T.R., Fisher, J.B., Fyllas, N.M., Freitas, M.A., Galbraith, D., Gloor, E., Higuchi, N., Honorio, E., Jiménez, E., Keeling, H., Killeen, T.J., Lovett, J.C., Meir, P., Mendoza, C., Morel, A., Núñez, V.P., Patiño, S., Peh, K., Peña, K., Cruz, A., Prieto, A., Quesada, C.A., Ramírez, F., Ramírez, H., Rudas, A., Salamão, R., Schwarz, M., Silva, J., Silveira, M., Sonké, B., Sota Thomas, A., Stropp, J., Vásquez, R., Taplin, J. & Vilanova, E. 2010. Drought mortality relationships for tropical forests. *New Phytologist* 187: 631–646.

Rodwell, J.S. (ed.) 1991. *British plant communities. Volume 1 – Woodland and scrub*. Cambridge University Press, Cambridge, UK.

Rodwell, J.S. 1995. The European Vegetation Survey questionnaire: an overview of phytosociological data, vegetation survey programmes and databases in Europe. *Annali di Botanica* 53: 87–98.

Schaminée, J.H.J. & Hennekens, S.M. 1995. Update of the installation of Turboveg in Europe. *Annali di Botanica* 53: 159–161.

Schaminée, J.H.J., Stortelder, A.H.F. & Weesthoff, V. (eds.) 1995. *De Vegetatie von Nederland – Deel 1. Inleiding tot de plantensociologie – grondslagen, methoden en toepassingen*. Opulus Press, Uppsala, SE.

Schaminée, J.H.J., Janssen, J.A.M., Haveman, R., Hennekens, S.M., Heuvelink, G.B.M., Huiskes, H.P.J. & Weeda, E.J. 2006. *Schatten voor de natuur – Achtergronden, inventaris en toepassingen van de Landelijke Vegetatie Databank*. KNNV Uitgeverij, Utrecht, NL.

Schaminée, J.H.J., Hennekens, S.M. & Ozinga, W.A. 2007. Use of the ecological information system SynBioSys for the analysis of large datasets. *Journal of Vegetation Science* 18: 463–470.

Schaminée, J.H.J., Hennekens, S.M., Chytrý, M. & Rodwell, J.S. 2009. Vegetation-plot data and databases in Europe: an overview. *Preslia* 81: 173–185.

Smits, N.A.C., Schaminée, J.H.J. & van Duuren, L. 2002. 70 years of permanent plot research in The Netherlands. *Applied Vegetation Science* 5: 121–126.

Stohlgren, T.J., Falkner, M.B. & Schell, L.D. 1995. A Modified-Whittaker nested vegetation sampling method. *Vegetatio* 117: 113–121.

Thiers, B. 2010. *Index Herbariorum: A global directory of public herbaria and associated staff*. New York Botanical Gardens Virtual Herbarium, New York, NY, US. Available at: http://sweetgum.nybg.org/ih/(accessed 22 September 2010).

Tichý, L., Hájek, M. & Zelený, D. 2010. Imputation of environmental variables for vegetation plots based on compositional similarity. *Journal of Vegetation Science* 21: 88–95.

US FDGC [U.S. Federal Geographic Data Committee] 2008. *National Vegetation Classification Standard, Version 2* [FGDC Document number FGDC-STD-005-2008] available at: http://www.fgdc.gov/standards/projects/FGDC-standards-projects/vegetation/NVCS_V2_FINAL_2008-02.pdf [U.S. Federal Geographic Data Committee Secretariat c/o U.S. Geological Survey] Accessed 22 September 2010.

Doi: 10.1111/j.1654-1103.2011.01265.x

Valachovič, M., Ot'ahel'ová, H., Stanová, V. & Ìàglocký, Š. 1995. *Rastlinné spoločenstvá Slovenska 1. Pionierska vegetácia (Plant communities of Slovakia 1. Pioneer vegetation)*. Veda, Bratislava, SK.

van Kley, J.E. & Schaminée, J.H.J. 2003. Large-scale ordination and gradient analysis of salt-marsh communities in the Netherlands in the light of the Dutch National Vegetation Classification. *Phytocoenologia* 33: 335–347.

Vilà, M., Pino, J., Montero, A. & Font, X. 2010. Are island plant communities more invaded than their mainland counterparts? *Journal of Vegetation Science* 21: 438–446.

Vittoz, P. & Guisan, A. 2007. How reliable is the monitoring of permanent vegetation plots? A test with multiple observers. *Journal of Vegetation Science* 18: 413–422.

Waterton, C. & Ellis, R. 2004. Environmental citizenship in the making: the participation of volunteer naturalists in UK biological recording. *Science & Public Policy* 31: 95–105.

Westhoff, V. & van der Maarel, E. 1973. The Braun-Blanquet approach. In: Whittaker, R.H. (ed.) *Ordination and classification of communities*. pp. 617–726. Junk, The Hague, NL.

Wheeler, Q.D. 2004. What if GBIF? *BioScience* 54: 717–717.

Whittaker, R.H., Niering, W.A. & Crisp, M.D. 1979. Structure, pattern, and diversity of a mallee community in New South Wales. *Vegetatio* 39: 65–76.

Williams, R.J., Martinez, N.D. & Golbeck, J. 2006. Ontologies for ecoinformatics. *Web Semantics: Science, Services and Agents on the World-Wide Web* 4: 237–242.

Willner, W. & Grabherr, G. (eds.) 2007. *Die Wälder und Gebüsche Österreichs*. Spektrum Akademischer Verlag, M ü nchen, DE.

Wiser, S.K., Bellingham, P.J. & Burrows, L.E. 2001. Managing biodiversity information: development of New Zealand's National Vegetation Survey databank. *New Zealand Journal of Ecology* 25: 1–17.

Wiser, S.K., Spencer, N., De Cáceres, M., Kleikamp, M., Boyle, B. & Peet, R.K. 2011. Veg-X – An exchange standard for plot-based vegetation databases. *Journal of Vegetation Science* 22, Doi: 10.1111/j.1654-1103.2010.01245.x.

Wohlgemuth, T., Schutz, M., Keller, W. & Wildi, O. 1999. Computed ecograms of Swiss forests. *Botanica Helvetica* 109: 169–191.

Supporting Information

Additional supporting information may be found in the online version of this article:

Appendix S1. Overview and definition of all required and non-required fields contained in GIVD on 30 December 2010.

Appendix S2. Overview of the 132 vegetation-plot databases registered in GIVD on 30 December 2010.

Appendix S3. Overview of the data availability via GIVD on a per-country basis on 30 December 2010.

Please note: Wiley-Blackwell is not responsible for the content and functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Doi: 10.1111/j.1654-1103.2011.01265.x

Appendix S1. Overview and definition of all required (in bold face) and non-required fields contained in GIVD on 30 December 2010.

The fields are arranged and grouped into major categories as in the online forms. Entries starting with "-" indicate subfields of the preceding field. An asterisk (*) in the column Type indicates that this field can be "repeated" if required. When the Type is "Select list" or "Proportions", the addition of "(multiple)" indicates that several options can be chosen for one database or one vegetation-plot record can fall into more than one category. In cases of select lists, the available options are presented in italics, possibly followed by definitions in brackets.

Field name	Type	Definition or available options	Remark
(1) Fields provided by the system or the GIVD Steering Committee			
Database ID	Identifier	Unique identifier for each database, consisting of two groups of two letters and one group of three numerals, separated by hyphens	The first group of letters denominates the continent (AF, AN, AS, AU, EU, NA, SA, or 00 for multi-continental), the second group of letters the country (ISO 3166 country code or 00 for multi-national), while the three numerals at the end are a consecutive number within each of the countries
Registered since	Date and time		
Last update	Date and time		
(2) Description of the database			
Name of the database	Text	Name of the database, potentially accompanied with an acronym	The name should be meaningful in a global context
Contact data of the responsible researcher(s): - **First name** - **Surname** - Department - Institution - Street - ZIP-code - Town - Country - **E-mail**	Various*		To ensure continuous availability also in the future, preferably more than one contact person should be named
Scope	Text	Concise verbal outline which kind of data are collected (i.e. criteria for inclusion)	

Field name	Type	Definition or available options	Remark
Status	Select list	*planned* (concretely planned but no data entered yet) *emerging* (first data entered but database not fully functional or not yet in the final format) *ongoing capture* (in the process of databasing historic data) *completed and continuing* (all historic data entered and continuing to accumulate new records) *finished* (data entry discontinued)	
Storage format/program	Select list (multiple)	*Turboveg* *MS Access* *MySQL* *PostgreSQL* *Oracle* *Excel* *other*	In the case of *other*, a text field is available to provide details
Export format	Select list (multiple)	*Turboveg* *MS Access* *SQL* *Excel* *Open Document* *PDF* *CSV file* *plain text file* *other*	In the case of *other*, a text field is available to provide details
Owner - private?	Text	Name and address of the institution or individual that owns the database	The field private is clicked when the database is owned by an individual researcher
URL	URL	Weblink to homepage of database	

Field name	Type	Definition or available options	Remark
Availability	Select list (multiple)	*free online* *free upon request* *according to specific agreement* *after blocking period* *available with reduced precision* *no yet available* (only for *planned* and *emerging* databases)	A database is only then eligible for GIVD when the host is basically willing to share the contained data with other researchers; the conditions for data sharing are defined here
Online search	Yes/no		
Online upload	Yes/no		
Other contents	Select list (multiple)	*classification* *synoptic tables* *vegetation maps* *other*	In the case of *other*, a text field is available to provide details
(3) Overall statistics of the database			
Number of sources	Numeric	Number of different sources included in the database	Each published source and each unpublished project is counted as one
List of primary sources	Text	Bibliographic references of the sources	Instead of filling in the field, a text file can be uploaded
Overlap with other databases	Select list (multiple)*	<All other databases that are already in GIVD> *Other overlapping databases*	For each of the partially overlapping databases, the number of joint independent plots can be specified
Number of vegetation-plot observations	Numeric		Each sub-plot in nested plots and each recording date in time series is counted separately
Number of non-overlapping vegetation plots	Numeric		Only the non-overlapping plots are counted (i.e. all records in one time series or in one nested-plot series together count as one)
Estimate of existing non-overlapping vegetation-plots	Numeric	Estimate of the existing vegetation-plot data that fall within the scope of the database (including those in the database, those in digital format, but not yet included, and those in paper format)	
Nested plots: - Number of series - Maximum plot-size range (min size – max size) - Maximum number of plot sizes in one series - Mean number of plot sizes in one series	Various		Only available when *nested plots* has been chosen

Field name	Type	Definition or available options	Remark
Time series: - Number of locations - Maximum plot-size range (start year – end year) - Repetition - Maximum number of plot sizes in one series - Mean number of plot sizes in one series	Various	The subfield "Repetition" has a select list (multiple) with the following options: *more than once a year* *every year* *every 2 years* *every 3-10 years* *less than every 10 years* *other*	Only available when *time series* has been chosen
Comments on nested plots / time series	Text	Free-text field for explanation of the sampling design	Only available when *nested plots* and/or *time series* have been chosen
(4) Statistics referring to the total number of vegetation-plot records			
Plot size minimum [m²]	Numeric	Size of the smallest (sub-) plot contained in the database	
Plot size maximum [m²]	Numeric	Size of the biggest plot contained in the database	
Plot size categories	Proportions	*< 1m²* *1 - <10 m²* *10 - < 100 m²* *100 - < 1000 m²* *1000 - < 10000 m²* *≥ 10000 m²* ***unknown***	Instead of filling in the categorical fields, a file with precise numbers per plot size can be uploaded
Year (oldest)	Numeric	Year of the oldest vegetation-plot record contained in the database	
Year (youngest)	Numeric	Year of the youngest vegetation-plot record contained in the database	

Field name	Type	Definition or available options	Remark
Decades	Proportions	*1910–1919* *1920–1929* *1930–1939* *1940–1949* *1950–1959* *1960–1969* *1970–1979* *1980–1989* *1990–1999* *2000–2009* *2010–2019* *unknown*	Instead of filling in the categorical fields, a file with precise numbers per year can be uploaded
(5) Statistics referring to the number of independent plots			
Covered countries	Proportions*		The countries are selected from a pull-down menu and then the individual proportion is entered
Guilds	Proportions (multiple)	*woody plants* *herbaceous plants* *bryophytes* (terricolous or aquatic bryophytes) *lichens* (terricolous or aquatic lichens) *algae* (terricolous or aquatic algae) *non-terricolous taxa* (epiphytic, saxicolous, or lignicolous non-vascular plants)	Each plot is assigned to one or several of the named six "guilds" depending on which of these would have been recorded when present (even if the guild is not present in individual plots)
Performance measure	Proportions (multiple)	*presence-absence only* *cover* *number of individuals* *dimensional measurements of individuals* *biomass* *other*	
Vegetation classification	Proportions (multiple)	*from primary source* *with uniform classification*	Proportions of vegetation-plot records that have been assigned to a vegetation classification

Field name	Type	Definition or available options	Remark
Geographic localisation	Proportions	*GPS coordinates (precision 25 m or less)* *point coordinates less precise than GPS (up to 1 km)* *small grid (≤ 10 km)* *geographical units or coarser grid coarser scale (> 10 km)*	Each plot is assigned to the highest-precision category available
Structural data	Proportions (multiple)	*cover of vegetation layers* *height of vegetation layers*	
Environmental data	Proportions (multiple)	*altitude* *slope aspect* *slope inclination* *microrelief* *surface cover other than plants* (open soil, litter, bare rock,…) *soil pH* *soil depth* *other soil attributes* *land use categories*	In the case of *other soil attributes*, a text field is available to provide details
(6) Additional data			
Number of taxon names	Numeric		
Number of valid taxa	Numeric		
Plant nomenclature	Text	Description which nomenclatural reference is applied for which plant group	
Publication(s) describing the data	Text	Bibliographic references of the publication(s)	
Publications(s) analysing the data	Text	Bibliographic references of the publication(s)	Instead of filling in the field, a text file can be uploaded
Classification(s) using the data	Text	Bibliographic references of the publication(s)	
Comments	Text	Free-text field for any further comments not covered by the standard fields	

Appendix S2. Overview of the 132 vegetation-plot databases registered in GIVD on 30 December 2010. A selection of relevant parameters are presented; for definitions, see Appendix S1.

ID of the database	National database	Name of the database	URL	Status	Availability	Storage format/program	Online search	Online upload	Normal plots	Nested plots	Time series	Number of sources	Number of vegetation-plot observations	Number of non-overlapping vegetation plots	Estimate of existing non-overlapping vegtation plots	Year (oldest)	Year (youngest)	Plot size minimum [m²]	Plot size maximum [m²]	Woody vascular plants [%]	Herbaceous vascular plants [%]	Bryophytes [%]	Lichens [%]	Algae [%]	Non-terricolous taxa [%]
00-00-001		ForestPlots.net	http://www.forestplots.net	ongoing capture	according to a specific agreement	other	yes	yes	x	x	x	6	2653	2653		1939	2010	800	20000	100	0	0	0	0	0
00-RU-001		Forest of Southern Ural		completed and continuing	according to a specific agreement	Turboveg	no	no	x				2270	2270		1977	2009	100	1000	100	100	100	90	0	0
00-RU-002		Database of Masaryk University`s vegetation research in Siberia		ongoing capture	according to a specific agreement	Turboveg	no	no	x				1041	1041		2003	2010	100	100	100	100	100	100	0	0
AF-00-001		West African Vegetation Database	http://www.westafricanvegetation.org/	emerging	free online	MySQL	yes	yes					16074	12000	40000	1980	2009	25	10000	90	70	0	0	0	3
AF-00-003		BIOTA Southern Africa Biodiversity Observatory Vegetation Database	http://www.biota-africa.org	ongoing capture	according to a specific agreement	MS Access, other	no	no	x	x	x	1	12808	4083	4083	2001	2010	0.01	10000	100	100	0	0	0	0
AF-00-004		East Africa		ongoing capture	according to a specific agreement	other	no	no	x			1	1786	1786		1991	2010	5	1000	100	100	0	0	0	0
AF-00-005		FLOTROP		emerging	free online	other			x				31800	31800		1929	2007			100	100	0	0	0	0
AF-00-006		SWEA-Dataveg	http://www.wetlands.uni-bonn.de	ongoing capture	according to a specific agreement	MS Access	no	no	x				206	206	1000	2009	2009	10	10	2	98	0	0	0	0
AF-00-007		Namib Desert Region Vegetation Database	http://www.biota-africa.org	ongoing capture	according to a specific agreement	MS Access, other	no	no	x		x	1	10952	10902		1980	2010	1	10000	100	100	0	0	0	0
AF-MA-001		BIOTA Marocco	http://www.biota-africa.org	completed and continuing	according to a specific agreement	other	no	no	x	x	x	4	44984	1790	44984	1952	2009	0.25	10000	100	100	0	0	0	0
AF-NA-001	x	National Phytosociological Database of Namibia		ongoing capture	according to a specific agreement	Turboveg	no	no	x			20	10526	10526	10526	1956	2010	100	1000	100	100	0	1	0	0
AF-ZA-001	x	National Vegetation Database of South Africa	http://www.sanbi.org.za	ongoing capture	according to a specific agreement	Turboveg			x			306	47414	47414	47414	1968	2010	1	400	100	88	3.7	2.9	0	0
AS-00-001		Korean Forest Database	http://www.butbn.cas.cz/skorea	completed and continuing	according to a specific agreement	Turboveg	no	no	x	x		41	1994	1844	10000	1978	2007	25	900	98	100	11	0	0	5
AS-CN-001		Database of grassland in China	establish in 2 months	completed and continuing		Excel	no	no	x			1	356	356		1995	2010	4	4	100	100	0	0	0	0
AS-EG-001		Vegetation database of Sinai in Egypt		emerging	according to a specific agreement	Turboveg	yes	no	x				1622	1622		1983	2010	4	5000						
AS-IL-001		INPA biological databases	http://ww2.bgbm.org/natureinfo	ongoing capture	free upon request	other			x		x		5000	1000		1979	2010	1	1000	100	100	0	0	0	0
AS-IR-001		Vegetation database of Iran		planned	after blocking period	Excel	no	no	x				0	0		1930	2010	1	400	5	95	0	0	0	0
AS-IR-002		Mountain wetland vegetation		emerging	after blocking period	Excel	no	no					1000	1000		2007	2008	1	100	3	92	5	0	0	0
AS-IR-003		Vegetation database of the Hyrcanian area, N Iran		emerging		Excel							100	0		1964	2009	100	400	90	10	0	0	0	0
AS-JP-001	x	PRDB	http://www.ffpri.affrc.go.jp/labs/prdb/EnglishVer/index-e.html	emerging	according to a specific agreement	other	no	no	x				22000	22000		1950	2008	1	2500	100	100	1	1	0	5
AS-KG-001		Vegetation database of the interdisciplinary project: The Impact of the Transformation Process on Human-Environmental Interactions in Southern Kyrgyzstan	http://www.kirgistan.uni-hamburg.de		according to a specific agreement	Excel	no	no				9	698	698	698	2001	2009	16	500	100	100	12	0	0	0
AS-KW-001		Kuwait Vegetation Database		ongoing capture	according to a specific agreement	Turboveg	no	no	x				202	202	400	1970	1999	0.5	110	100	100	0	0	0	0
AS-TR-001		Vegetation plot data and database of the grassland communities in Anatolia		ongoing capture	according to a specific agreement	Turboveg	no	no	x			46	1878	1878	20000	1974	2009	4	500	0	100	0	0	0	0
AS-TW-001	x	National Vegetation Database of Taiwan		completed and continuing	not yet available	Turboveg	no	no	x				8035	8035		1979	2007	4	1000	93.7	50	0	0	0	0
AS-TW-002		Taiwan forest dynamics plot	http://fdp.tfri.gov.tw/	completed and continuing	according to a specific agreement	Excel	yes	yes					3	3		1995	2008			100	0	0	0	0	0
AU-AU-001		PPBio Australasia LTER sites	http://www.griffith.edu.au/ppbio	completed and continuing		other			x				60	60		2007	2010	10000	10000	100	0	0	0	0	0
AU-NZ-001	x	New Zealand National Vegetation Databank	http://nvs.landcareresearch.co.nz	completed and continuing		other	yes	no	x	x	x		468000	77000		1960	2010	100	400	70	20	4	1	0	5
EU-00-002		Dry Grasslands in the Nordic and Baltic Region	http://www.biologie.uni-hamburg.de/bzf/syst/wg_dry_grasslands_nordic/wg_dgnb1_eng.htm	ongoing capture	according to a specific agreement	other	no	no	x			125	7675	7675	20000	1922	2009	1	100	100	100	70.7	70.7	71	4.2
EU-00-003		South East European Dry Grassland Database	http://www.edgg.org/subgroups.htm#Southeastern	emerging	according to a specific agreement	Excel	no	no	x			2	204	204	50000	2009	2010	1	100	100	100	100	100	100	0
EU-00-004	x	SIVIM - Iberian and Macaronesian Vegetation Information System	http://www.sivim.info	ongoing capture	free online	other	yes	yes	x			1350	110000	110000	150000	1926	2010	1	400	100	100	50	30	1	0
EU-00-005		Mountain tall herbs: Mulgedio-Aconitetea and related vegetation types in Europe		ongoing capture	according to a specific agreement	Turboveg	no	no	x			100	2852	2852	10000	1911	2008	0.25	3000	100	100	69.7	69.7	0	0
EU-00-006		Ecological Conditions database (EC)	http://abiotic.wur.nl	ongoing capture	free upon request	MS Access, Excel	no	no	x		x	91	8229	8130	25000	1936	2010	1	900	50	100	30	5	1	0
EU-00-007		VIOLETEA - Heavy Metal Grasslands		ongoing capture	free upon request	Excel	no	no	x			17	656	656	1000	1970	2007	1	100	0	100	0	0	0	0
EU-00-008		Deciduous forests on acidic soils (NW Europe)		finished	free upon request	other	no	no	x			75	4430	4430		1936	2007	100	900	100	100	99	99	0	0
EU-00-009		Vegetationsdaten der Oberrheinaue	http://www.fva-fr.de	finished	free upon request	MS Access, MySQL	yes	yes	x				1100	1100		1994	2004	25	250	70	30	0	0	0	0
EU-AT-001	x	Austrian Vegetation Database	http://vegedat.vinca.at/	ongoing capture	free upon request	Turboveg	no	no	x			530	40000	40000	100000	1926	2010	0.1	4000	100	100	20	5	0	0
EU-AZ-001		Shahdag region Greater Caucasus Azerbaijan		completed and continuing	not yet available	Turboveg	no	no		x	x		1094	222	222	2007	2008	0.01	100	100	100	0	0	0	0
EU-BE-001		VLAVEDAT the vegetation database of Flanders		ongoing capture	according to a specific agreement	Turboveg	no	no	x			295	26180	26180	55000	1927	2003	0.1	10000	45.7	98.98	31.3	5.28	2.1	3.8
EU-BG-001	x	Bulgarian Vegetation Database			according to a specific agreement	Turboveg			x			20	5901	5901		1949	2010	4	100	30	100	20	2	0	0
EU-CH-001		Permanent.Plot.ch	http://www.unil.ch/ecospat/page48113_en.html	completed and continuing	according to a specific agreement	MS Access	no	no			x		3633	861	17000	1883	2009	1	10000	16	100	12	0	0	0
EU-CH-002		Swiss Biodiversity Monitoring BDM (Z9 Plants)	http://www.biodiversitymonitoring.ch	completed and continuing	according to a specific agreement	other	no	no			x		3000	1600		2001	2010	10	10	100	0	0	0	0	0
EU-CH-004		National Inventory of Swiss Bryophytes NISM	http://www.nism.uzh.ch/	completed and continuing	according to a specific agreement	Oracle	yes	no	x			1	1221	1221		1984	2006	100	100	0	0	100	0	0	0
EU-CH-005		Swiss Forest Vegetation			free upon request	Oracle			x				14500	14500						100	0	0	0	0	0
EU-CH-006		Dry meadows and pastures of Switzerland	http://www.bafu.admin.ch/schutzgebiete-inventare/07849/index.html?lang=de	completed and continuing	according to a specific agreement	MS Access	no	no	x			1	13724	13724	13724	1995	2009	28	28	0	100	0	0	0	0
EU-CH-007		Swiss mire monitoring / Wirkungskontrolle Moorschutz Schweiz	http://www.wsl.ch/fe/oekologie/biotop/index_DE	completed and continuing	according to a specific agreement	Excel, other	no	no			x	1	32446	24015	25000	1996	2010	10	1200	100	100	100	0	0	0
EU-CH-008		Phytobase C2005	http://www.bafu.admin.ch/schutzgebiete-inventare/07839/index.html?lang=de	completed and continuing	according to a specific agreement	MS Access	no	no	x			1	1241	1241	1241	1987	2008	10	900	100	100	0	0	0	0
EU-CH-009		Phytobase S2008	http://www.bafu.admin.ch/schutzgebiete-inventare/07839/index.html?lang=de	completed and continuing	according to a specific agreement	MS Access	no	no	x			1	1314	1314	1314			20	900	75	75	25	0	0	25
EU-CZ-001	x	Czech National Phytosociological Database	http://www.sci.muni.cz/botany/vegsci/dbase.php?lang=cz	ongoing capture	free upon request	Turboveg	no	no	x			2088	95932	95932	105000	1922	2009	0.09	10000	100	100	60	50	0.3	2
EU-CZ-002		Database of Czech Forest Classification system		completed and continuing	according to a specific agreement	Turboveg	no	no	x		x	1	48439	32387		1926	2005	5	8100	100	100	0	0	0	0
EU-DE-001		Vegbank MV	http://geobot.botanik.uni-greifswald.de/portal/index.php?option=com_content&task=view&id=111&Itemid=346	completed and continuing	free upon request	Turboveg	no	no	x			467	53842	53842	150000	1928	2010	0.1	2000	100	100	25	10	1	1
EU-DE-002		BERGWALD		completed and continuing	free upon request	MS Access	no	no	x			37	4934	4934	7000	1938	1997	5	2600	100	100	0	0	0	0
EU-DE-003		WINALPecobase	www.winalp.info	finished	free upon request	MS Access	no	no	x				1505	1505	7000	2009	2010	196	196	100	100	89	4	0	0

ID of the database	National database	Name of the database	URL	Status	Availability	Storage format/program	Online search	Online upload	Normal plots	Nested plots	Time series	Number of sources	Number of vegetation-plot observations	Number of non-overlapping vegetation plots	Estimate of existing non-overlapping vegtation plots	Year (oldest)	Year (youngest)	Plot size minimum [m²]	Plot size maximum [m²]	Woody vascular plants [%]	Herbaceous vascular plants [%]	Bryophytes [%]	Lichens [%]	Algae [%]	Non-terricolous taxa [%]
EU-DE-004		Polygono-Poetea annuae of Germany		emerging	according to a specific agreement	Excel	no	no	x				582	582	1500	1936	2010	0.01	50	100	100	20	0	0	0
EU-DE-005		Temperate deciduous forests of the Elbe-Weser region (Lower Saxony, Germany)	http://www.zalf.de/home_zalf/institute/lse/lse/mitarbeiter/wulf/general.htm	finished	according to a specific agreement	Excel	no	no	x			2	415	415	415	1986	1989	100	400	100	100	100	0	0	0
EU-DE-006		Temperate deciduous forests of the Prignitz region (NW Brandenburg, Germany)	http://www.zalf.de/home_zalf/institute/lse/lse/mitarbeiter/wulf/general.htm	finished	according to a specific agreement	Excel	no	no	x			2	232	232	232	1996	1999	100	400	100	100	100	0	0	0
EU-DE-007		Temperate deciduous forests of the Uckermark region (NE Brandenburg, Germany)	http://www.zalf.de/home_zalf/institute/lse/lse/mitarbeiter/wulf/general.htm	ongoing capture	according to a specific agreement	Excel	no	no	x			1	500	500	500	1992	2007	100	400	100	100	100	0	0	0
EU-DE-008		Cytisus_SFB299	http://www.uni-giessen.de/landscape	finished	according to a specific agreement	Excel	yes	no	x			1	220	220		1997	1997	16	25	100	100	0	0	0	0
EU-DE-009		BioChangeMeadows		completed and continuing	free upon request	Turboveg	no	no	x				1092	1092		1939	2008	15	25	0	100	0	0	0	0
EU-DE-010		Forests of the Oldenburg Region (NW Germany)		ongoing capture	free upon request	Turboveg	no	no	x				1522	1522		1990	2008	16	400	100	100	70	0	0	0
EU-DE-011		Forests and Grasslands of the Lower Werra Region		ongoing capture	free upon request	Turboveg	no	no	x				565	565		2002	2009	25	100	100	100	0	0	0	0
EU-DE-012		Nardus swards of Germany		finished	free upon request	other	no	no	x				419	419		1986	1989	5	25	100	100	100	100	0	0
EU-DE-013		VegetWeb	http://www.floraweb.de/vegetation/aufnahmen.html	ongoing capture	free online	MySQL	yes	yes	x			211	26692	26692	2000000	1934	2007			100	100	7	6	0	0
EU-DE-014		GVRD Vegetation Database Halle	http://www.biologie.uni-halle.de/bot/vegetation_db//index.php?Lang=E	ongoing capture	free upon request	Turboveg, MS Access, MySQL	yes	no				618								100	100	50	8	1	0
EU-DE-015		Successional permanent plot database (Lower Saxony, Germany)		completed and continuing	free upon request	MS Access	no	no	x		x	1	760	23		1968	2006	125	200	100	100	10	0	0	0
EU-DE-016		Database of strict forest reserves (NW-Germany)		completed and continuing	free upon request	MS Access	no	no	x	x	x	1	3600	3600		1988	2010	100	400	100	100	80	50	0	20
EU-DE-017		Temperate deciduous and coniferous forests of the Sollig Hills (Lower Saxony, Germany)		completed and continuing	free upon request	MS Access	no	no	x		x	1	2901	2057		1966	2008	100	400	100	100	100	0	0	0
EU-DE-018		Temperate deciduous forests of the Göttinger Wald (S-Lower Saxony, Germany)		completed and continuing	free upon request	MS Access	no	no	x		x	1	608	322		1955	2010	1	400	100	100	50	0	0	0
EU-DE-019		Pine forests on acidic soils (Germany)		finished	free upon request	other	no	no	x			73	3198	3198	4500			50	900	100	100	100	100	0	0
EU-DE-020		Dry Grasslands of Germany	http://www.edgg.org/subgroups.htm#Arbeitsgruppe	planned	according to a specific agreement	Turboveg	no	no	x				0	0	70000	1993	1997	10	10	100	100	0	0	0	0
EU-DE-021		Main-Kinzig + Bergland: Vegetation of the central German highland region		ongoing capture	according to a specific agreement	Turboveg	no	no	x			10	1562	1562	20000	1972	2008	0.12	450	100	100	60	60	0	0
EU-DE-022		Frankenalb		ongoing capture		other			x		x	1	2540	2500		1985	2009	1	1000	100	100	100	100	0	0
EU-DE-023		Post-mining vegetation database (Eastern Germany)		completed and continuing	according to a specific agreement	MS Access	no	no	x		x	3	5194	3247		1994	2009	1	10000	100	100	0	0	0	0
EU-DE-024		Fichtelgebirge		completed and continuing		MS Access						1	532	532		2005	2006	10000	10000	100	100	0	0	0	0
EU-DE-025		Grafenwoehr Training Area		completed and continuing	according to a specific agreement	MS Access		no	x			1	595	595		2008	2008	10000	10000	100	100	0	0	0	0
EU-DE-026		Vaccinio-Pinetea		emerging	not yet available	Excel							2000	2000		1950	2000	30	200	100	100	100	0	0	0
EU-DE-027		BioChangeFields		completed and continuing	free upon request	Turboveg							1216	1216		1951	2009	10	100	0	100	0	0	0	0
EU-EE-001		Estonian vegetation	http://www.botany.ut.ee/	ongoing capture	according to a specific agreement	Excel, other	no	no					1430	1430		1980	2010	1000	2000	100	100	100	5	40	5
EU-FI-001		NW Finnish Lapland		completed and continuing	according to a specific agreement	Excel							200	200	500	1986	1990	0.25	4	100	100	100	100	0	0
EU-FR-001		Vigie-flore	http://www.vigie-flore.fr	ongoing capture	according to a specific agreement	MySQL	no	no		x	x		981	117	117	2009	2010	1	1	0	100	0	0	0	0
EU-FR-002		FLOREM		ongoing capture	according to a specific agreement		no	no	x				2000	2000		1970	2010	0.25	100	0	100	0	0	0	0
EU-FR-003	x	SOPHY	http://sophy.univ-cezanne.fr/sophy.htm	ongoing capture	according to a specific agreement	other			x				203278	203278		1915	2010	1	400	100	100	1	1	0	0
EU-GB-001	x	UK National Vegetation Classification Database		finished	not yet available	MS Access			x			40	35000	35000		1959	1979	1	2500	100	100	100	100	1	0
EU-GB-002		Environmental Change Network	http://www.ecn.ac.uk	completed and continuing	according to a specific agreement	Oracle	yes	no			x		290	12		1993		2	10	100	100	100	100	0	0
EU-GB-003		Countryside Survey	http://www.countrysidesurvey.org.uk/	completed and continuing	free online	Oracle	yes	yes	x	x	x	1	58698	49165	49165	1978	2007	2	200	100	100	70	20	0	0
EU-GR-001		KRITI - Vegetation of Crete		completed and continuing	according to a specific agreement	Turboveg	no	no	x				6500	6500	8000	1945	2010	1	400	100	100	10	1	0.1	0
EU-GR-002		Greek Woodland Vegetation Database		completed and continuing	according to a specific agreement	Turboveg	no	no	x				3500	3500	6000	1963	2010	50	800	100	100	60	10	0	0
EU-GR-003		Isoeto-Nano-Juncetea in Greece and the Aegean region		completed and continuing	according to a specific agreement	Turboveg	no	no	x			1	300	300	350	1985	2010	0.5	4	100	100	10	0	0	0
EU-GR-004		Segetal weed vegetation of Greece		ongoing capture	according to a specific agreement	Turboveg	no	no	x			3	200	200	250	1983	2010	16	50	100	100	0	0	0	0
EU-HR-001		Phytosociological database of non-forest vegetation in Croatia		ongoing capture	free upon request	Turboveg	no	no	x			195	5728	5728		1927	2009	0.4	60	0	100	0	0	0	0
EU-HU-001		Regional Vegetation Database of Kiskunsag		finished	according to a specific agreement	Excel	no	no	x			1	605	605	3000	2006	2008	400	400	100	100	0	0	0	0
EU-HU-002		Long-term database of sandy grassland of Fulophaza		completed and continuing	according to a specific agreement	Excel	no	no			x	1	1200	200	3000	2000	2010	16	16	100	100	0	0	0	0
EU-HU-003	x	CoenoDat Hungarian Phytosociological Database		completed and continuing	according to a specific agreement	Turboveg	no	no	x			184	11000	11000	45000	1929	2007	0.25	2500	100	100	11	1	0	0
EU-IE-001	x	Irish Vegetation Database	http://nationalvegetationdatabase.biodiversityireland.ie	ongoing capture	according to a specific agreement	Turboveg	no	no	x			103	21995	21995	26000	1949	2008	0.1	20	20	100	35	10	4	0
EU-IT-001		www.anarchive.it	http://www.anarchive.it	finished	free online	PostgreSQL	yes	yes					3871	3871		1970	2010	0.04	1000	40	60	0.1	0.1	0.1	0.1
EU-LV-001		Semi-natural grasslands of Latvia		completed and continuing	according to a specific agreement	Turboveg	no	no	x		x	1	2500	1965	3000	1997	2009	1	125	100	100	80	80	0	0
EU-LV-002		Mires of Latvia		emerging	according to a specific agreement	Turboveg	no	no	x				2027	2027	2027	1995	2009	1	4	100	100	100	0	0	0
EU-NL-001	x	Dutch National Vegetation Database	http://www.synbiosys.alterra.nl/lvd	completed and continuing	according to a specific agreement	Turboveg, PostgreSQL	yes	no	x		x	80	627000	600000	750000	1864	2010	0.1	1000	100	100	0	0	0	0
EU-NL-002		Vegetation of Dutch Road verges			according to a specific agreement	Turboveg	no	no	x				2552	2552		1993	1088	2	9	100	100	80	80	0	0
EU-PL-001		Polish Vegetation Database - SynBiotSilesiae		emerging	free upon request	Turboveg	no	no	x			159	25107	25107	180000	1952	2009	1	5000	46	100	29	1	0	0
EU-RO-001		Vegetation database of Dry Grasslands in the Southeast Romania		emerging	free online	Excel			x			313	2500	2500		1941	2005	1	100	100	100	0	0	0	0
EU-RO-002		Vegetation database of the dry grasslands from the Transylvanian Basin - Romania		ongoing capture	according to a specific agreement	Excel	no	no	x			21	515	515	4000	1940	2008	1	200	100	100	2	0	0	0
EU-RO-003		Oak-hornbeam forest database of the Transylvanian Basin - Romania		completed and continuing	according to a specific agreement	Turboveg	no	no				16	524	524		1941	2005	10	400	100	100	0	0	0	0
EU-RU-002		Lower Volga Valley Phytosociological Database		completed and continuing	according to a specific agreement	Turboveg	no	no	x		x	11000	11000	11000	11500	1928	2010	3	2500	5	100	5	2	3	1
EU-RU-003		Vegetation Database of the Volga and the Ural Rivers Basins		ongoing capture	according to a specific agreement	Turboveg, Excel	no	no	x			3000	3000	3000		1968	2010	0.1	100	10	90	10	5	1	0
EU-RU-004		Coastal Vegetation Database of north-western Seas of Russia		completed and continuing	according to a specific agreement	Turboveg	no	no	x			1529	1529	1529	2000	1995	2003	0.04	600	2.75	100	34.3	13.9	6.5	0
EU-RU-005		Coastal Vegetation Database of southern Seas of Russia		completed and continuing	according to a specific agreement	Turboveg	no	no	x			1735	1735	1735	2000	2004	2006	0.5	600	8.65	100	5.71	1.84	1	0

ID of the database	National database	Name of the database	URL	Status	Availability	Storage format/program	Online search	Online upload	Normal plots	Nested plots	Time series	Number of sources	Number of vegetation-plot observations	Number of non-overlapping vegetation plots	Estimate of existing non-overlapping vegtation plots	Year (oldest)	Year (youngest)	Plot size minimum [m²]	Plot size maximum [m²]	Woody vascular plants [%]	Herbaceous vascular plants [%]	Bryophytes [%]	Lichens [%]	Algae [%]	Non-terricolous taxa [%]
EU-RU-006		Moscow region forests		completed and continuing	according to a specific agreement	Turboveg	no	no	x		x		650	400		1987	2010	100	400	95	100	80	0	0	0
EU-RU-007		Biodiversity of Murmansk region		completed and continuing	according to a specific agreement	MS Access	no	no	x				350	350		2005	2009	625	625	90	100	100	100	5	0
EU-SI-001	x	Vegetation database of Slovenia	http://hacquetia.zrc-sazu.si/%C5%A0ilc%20Urban/myweb3/links.htm	ongoing capture	according to a specific agreement	Turboveg	no	no	x		x	390	15938	15773	20000	1932	2009	0.2	4000	100	100	31.7	0.05	0	0
EU-SK-001	x	Slovak Vegetation Database	http://ibot.sav.sk/cdf/index.html	ongoing capture	free upon request	Turboveg	no	no	x			1325	51581	51581	60000	1919	2010	0.01	5000	100	100	60	60	0.2	2
NA-00-002		BIOTREE-NET	www.biotreenet.com	emerging	not yet available	PostgreSQL	yes		x			53	2019	2019		1969	2009	100	54000	100	0	0	0	0	0
NA-CA-001		Understory vegetation in old-growth and clearcut forest, Vancouver Island, Canada	http://esapubs.org/Archive/ecol/E084/047/suppl-1.htm	finished	free online	other	no	no	x			2	72	72	72	1996	1996	100	100	100	100	100	0	0	0
NA-CA-002		Vegetation database of Québec (MRNF)		finished	according to a specific agreement		no	no	x				28425	28425	28425	1986	2000	400	400	100	100	100	100	0	0
NA-US-001		Forest Inventory National Database (FIADB)	http://fia.fs.fed.us/tools-data/default.asp	completed and continuing	free online	MS Access, other	yes	no			x	3	1929131	495284	495284	1966	2008	14	1013	100	1	0	1	0	0
NA-US-002		VegBank	http://vegbank.org	completed and continuing	free online	PostgreSQL	yes	yes	x			58	22629	22629		1971	2007	12	1000	100	100	0	0	0	0
NA-US-003		Jasper Ridge Woody Plant Communities			free upon request	other	no	no	x			1	44	44	44	2002	2002	400	400	100	0	0	0	0	0
NA-US-004		Thirty years of permanent vegetation plots, Mount St. Helens, Washington	http://faculty.washington.edu/moral	completed and continuing	free upon request	Excel					x		1743	92		1978	22	250	250	100	100	100	0	0	0
NA-US-005		Santa Rita Experimental Range long term	http://ag.arizona.edu/SRER/data.html	completed and continuing	free online	Excel	no	yes			x	1	2084	132	132	1954	37	9	9	100	100	0	0	0	0
NA-US-006		Carolina Vegetation Survey (North Carolina, United States)	http://cvs.bio.unc.edu	completed and continuing	free upon request	MS Access	no	no	x	x	x	10	117195	8153		1976	2009	0.01	1000	100	100	0	0	0	0
NA-US-007		FIA database (FIADB) - Vegetation diversity and structure indicator (VEG)	http://fia.fs.fed.us/tools-data/default.asp	completed and continuing	free online	MS Access	no	no		x	x	1	26607	1568	2000	2001	2008	1	672	100	0	0	0	0	0
NA-US-008		University of Wisconsin Plant Ecology Laboratory	http://www.botany.wisc.edu/PEL/	completed and continuing	according to a specific agreement	MS Access	yes	no	x			14	700	700		1945	2008	1	1	100	100	0	0	0	0
NA-US-009		Observatory Woods Gridded Vegetation, Wisconsin USA	http://uwarboretum.org/	completed and continuing	according to a specific agreement	MS Access	no	no		x		1	96	38		2007	2007	1	6	100	50	0	0	0	0
NA-US-011		Desert grassland net-primary productivity in central New Mexico	http://sev.lternet.edu/project_details.php?id=SEV129	completed and continuing	free online	other	yes	yes						220		1999	2010	1	1	10	90	0	0	0	0
NA-US-012		Epiphytic lichen synusiae in forested areas of the US with 20,000 relevés	http://fia.fs.fed.us/lichen/data/	ongoing capture	free online	MS Access, Oracle	yes	no	x			1	4941	4941	4941	1994	2010	4000	4000	0	0	0	0	0	100
NA-US-013		Natural Heritage Vegetation Database for West Virginia		completed and continuing	according to a specific agreement	MS Access	no	no	x		x		3722	3712	4000	1963	2010	25	400	100	100	22	22	1	0
SA-00-001		Ephemeral wetland vegetation in extra- and oro-tropical South America	http://www.biologie.uni-freiburg.de/data/bio2/geobotanik/index.html	completed and continuing	according to a specific agreement	Turboveg	no	no	x			26	514	514		1960	2008	1	100	2	97	1	0	0	0
SA-BO-001		Project Database of Bolivian Ecoregions		finished	according to a specific agreement	MySQL	yes	yes	x	x			126	118	104	2000	2003	25	800	22	26	0	0	0	52
SA-BR-001		Plants in the Brazilian PPBio Data Repository	http://ppbio.inpa.gov.br		after blocking period	other	yes	yes	x				1638	1638		2001	2010	1	10000	100	100	0	0	0	0
SA-CL-001		CL-Dataveg		ongoing capture	according to a specific agreement	Turboveg, MS Access	no	no	x			18	650	650	1500	1975	2008	1	100	1	99	0	0	0	0
SA-CO-001		Páramo Region (High Andean Mountain)		emerging	not yet available	Excel	no	no	x			45	800	800	1500	1973	2006	1	100	90	90	50	50	0	0
SA-CO-002		Colombian Caribbean Region		emerging	not yet available	Excel	no	no	x			12	320	320	500	1976	2009	20	1200	90	70	40	40	0	0
SA-EC-001		Ecuador forest plots		ongoing capture	according to a specific agreement	other	no	no	x				230	230		2000	2010	400	400	100	10	0	0	0	0

Appendix S3. Overview of the data availability via GIVD on a per-country basis on 30 December 2010. Number of databases as well as total number of vegetation-plot observations, number of non-overlapping plots, and density of non-overlapping plots per 1000 km² of land surface are presented (note that a small proportion of the plots in GIVD-registered databases have not been assigned to countries).

Rank (according to number of non-overlapping plots)	Country code (according to ISO 3166)	Country	Continent(s)	Area [km²]	Number of databases containing plots from the country	Number of vegetation-plot observations	Number of non-overlapping plots	Density of non-overlapping plots [plots per 1000 km²]
1	NL	Netherlands	EU	33880	5	638001	610902	18031.35
2	US	United States	NA	9158960	13	2108439	537060	58.64
3	FR	France	EU	550100	7	195481	194617	353.78
4	DE	Germany	EU	348950	32	182560	178706	512.12
5	CZ	Czech Republic	EU	77270	5	159363	143311	1854.68
6	ES	Spain	EU+AF	499040	3	99193	99193	198.77
7	GB	United Kingdom (incl. Isle of Man and Channel Islands)	EU	241930	3	93988	84177	347.94
8	NZ	New Zealand	AU	267990	1	468000	77000	287.32
9	ZA	South Africa	AF	1214470	3	55887	52609	43.32
10	SK	Slovakia	EU	48080	4	50873	50873	1058.09
11	CH	Switzerland	EU	40000	9	58802	46199	1154.98
12	AT	Austria	EU	82450	7	40537	40537	491.66
13	CA	Canada	NA	9093510	3	28950	28950	3.18
14	BE	Belgium	EU	30280	3	26975	26975	890.85
15	PL	Poland	EU	306290	6	25899	25899	84.56
16	JP	Japan	AS	364500	1	22000	22000	60.36
17	IE	Ireland	EU	68890	1	21995	21995	319.28
18	RU	Russia	AS+EU	16380940	11	21604	21352	1.30
19	NA	Namibia	AF	823290	4	24317	18821	22.86
20	PT	Portugal	EU+AF	91500	1	18700	18700	204.37
21	SI	Slovenia	EU	20140	2	15939	15774	783.22
22	BF	Burkina Faso	AF	273600	3	14878	12841	46.93
23	HU	Hungary	EU	92100	4	11918	10918	118.55
24	GR	Greece	EU	128900	4	10497	10497	81.44
25	TW	Taiwan	AS	35800	2	8038	8038	224.53
26	SN	Senegal	AF	192530	2	7435	6824	35.44
27	IT	Italy	EU	294110	3	5922	5922	20.14
28	BG	Bulgaria	EU	110630	2	5901	5901	53.34
29	HR	Croatia	EU	55920	1	5728	5728	102.43
30	NE	Niger	AF	1266700	2	5782	5171	4.08
31	LV	Latvia	EU	62050	3	5141	4606	74.23
32	RO	Romania	EU	229870	7	4442	4442	19.32
33	TD	Chad	AF	1259200	1	3943	3943	3.13
34	BJ	Benin	AF	110620	2	4201	3386	30.61
35	EE	Estonia	EU	42390	2	2644	2644	62.37
36	MR	Mauritania	AF	1025220	1	2385	2385	2.33
37	BR	Brazil	SA	8459420	2	1932	1932	0.23
38	TR	Turkey	AS+EU	769630	2	1881	1881	2.44
39	ML	Mali	AF	1220190	2	1846	1846	1.51
40	MA	Morocco	AF	446300	2	44984	1790	4.01
41	TZ	Tanzania	AF	883590	3	1778	1778	2.01
42	CI	Cote d'Ivoire	AF	318000	1	1717	1717	5.40
43	EG	Egypt	AF+AS	995450	2	1622	1622	1.63
44	KR	Korea, Republic	AS	98730	1	1635	1512	15.31
45	CM	Cameroon	AF	465400	2	1364	1364	2.93
46	MX	Mexico	NA	1908690	1	1330	1330	0.70
47	CO	Colombia	SA	1038700	4	1153	1153	1.11
48	SE	Sweden	EU	410330	3	1143	1143	2.79
49	IL	Israel (incl. West Bank and Gaza Strip)	AS	21710	1	5000	1000	46.06
50	IR	Iran	AS	1636200	3	1100	1000	0.61
51	CL	Chile	SA	748800	2	733	733	0.98
52	VE	Venezuela	SA	882050	2	719	719	0.82
53	KG	Kyrgyzstan	AS	191800	1	698	698	3.64
54	NG	Nigeria	AF	910770	2	668	668	0.73
55	MZ	Mozambique	AF	784090	1	616	616	0.79
56	SD	Sudan	AF	2376000	1	572	572	0.24
57	NO	Norway	EU	306250	2	562	562	1.84

Rank (according to number of non-overlapping plots)	Country code (according to ISO 3166)	Country	Continent(s)	Area [km²]	Number of databases containing plots from the country	Number of vegetation-plot observations	Number of non-overlapping plots	Density of non-overlapping plots [plots per 1000 km²]
58	CV	Cape Verde	AF	4030	1	477	477	118.36
59	FI	Finland	EU	304590	3	411	411	1.35
60	DJ	Djibouti	AF	23180	1	382	382	16.48
61	CN	China	AS	9327430	2	370	367	0.04
62	UA	Ukraine	EU	579350	4	318	318	0.55
63	EC	Ecuador	SA	276840	2	289	289	1.04
64	BO	Bolivia	SA	1084380	3	295	287	0.26
65	LS	Lesotho	AF	30350	1	284	284	9.36
66	AU	Australia	AU	7682300	2	283	283	0.04
67	AR	Argentina	SA	2736690	1	279	279	0.10
68	KP	Korea, Democratic People's Republic	AS	120410	1	285	264	2.19
69	AO	Angola	AF	1246700	1	263	262	0.21
70	CF	Central African Republic	AF	622980	2	228	228	0.37
71	PE	Peru	SA	1280000	2	226	226	0.18
72	AZ	Azerbaijan	AS+EU	82600	2	1094	222	2.69
73	CR	Costa Rica	NA	51060	1	206	206	4.03
74	KW	Kuwait	AS	17820	1	202	202	11.34
75	PA	Panama	NA	74430	1	200	200	2.69
76	KE	Kenya	AF	569140	2	192	192	0.34
77	DK	Denmark	EU	42430	1	177	177	4.17
78	MY	Malaysia	AS	328550	1	175	175	0.53
79	GA	Gabon	AF	257670	1	172	172	0.67
80	BW	Botswana	AF	566730	2	153	153	0.27
81	ZW	Zimbabwe	AF	386850	1	142	142	0.37
82	HN	Honduras	NA	111890	1	140	140	1.25
83	LT	Lithuania	EU	62680	1	123	123	1.96
84	AD	Andorra	EU	470	1	110	110	234.04
85	KZ	Kazakhstan	AS+EU	2699700	2	110	110	0.04
86	GF	French Guiana	SA	88150	1	91	91	1.03
87	LI	Liechtenstein	EU	160	1	89	89	556.25
88	SV	El Salvador	NA	20720	1	68	68	3.28
89	RS	Serbia	EU	88360	1	66	66	0.75
90	GH	Ghana	AF	227540	2	66	66	0.29
91	ID	Indonesia	AS+AU	1811570	1	66	66	0.04
92	CD	Congo, Republic	AF	2267050	1	65	65	0.03
93	CG	Congo, Democratic Republic	AF	341500	1	63	63	0.18
94	LR	Liberia	AF	96320	2	44	44	0.46
95	NI	Nicaragua	NA	121400	1	42	42	0.35
96	GY	Guyana	SA	196850	1	42	42	0.21
97	ET	Ethiopia	AF	1000000	2	36	36	0.04
98	BZ	Belize	NA	22810	1	32	32	1.40
99	MN	Mongolia	AS	1566500	1	28	28	0.02
100	BN	Brunei Darussalam	AS	5270	1	10	10	1.90
101	UG	Uganda	AF	197100	1	8	8	0.04
102	NP	Nepal	AS	143000	1	4	4	0.03
103	GQ	Equatorial Guinea	AF	28050	1	3	3	0.11
104	GW	Guinea-Bissau	AF	28120	1	3	3	0.11
		Total		**108615920**		**4537823**	**2442039**	**22.48**