**This manuscript is contextually identical with the following published paper:**
Chytry M, Hennekens SM, Jimenez-Alfaro B, etal. (2016) European Vegetation Archive (EVA): an integrated database of European vegetation plots. APPLIED VEGETATION SCIENCE 19: 173-180. DOI: 10.1111/avsc.12191

**The original published PDF available in this website:**
http://onlinelibrary.wiley.com/doi/10.1111/avsc.12191/abstract

# European Vegetation Archive (EVA): an integrated database of European vegetation plots

Milan Chytry, Stephan M. Hennekens, Borja Jiménez-Alfaro, Ilona Knollová, Jürgen Dengler, Florian Jansen, Flavia Landucci, Joop H.J. Schaminée, Svetlana Aćić, Emiliano Agrillo, Didem Ambarlı, Pierangela Angelini, Iva Apostolova, Fabio Attorre, Christian Berg, Erwin Bergmeier, Idoia Biurrun, Zoltán Botta-Dukát, Henry Brisse, Juan Antonio Campos, Luis Carlón, Andraž Čarni, Laura Casella, János Csiky, Renata Ćušterevska, Zora Dajić Stevanović, Jiřı Danihelka, Els De Bie, Patrice de Ruffray, Michele De Sanctis, W. Bernhard Dickoré, Panayotis Dimopoulos, Dmytro Dubyna, Tetiana Dziuba, Rasmus Ejrnæs, Nikolai Ermakov, Jörg Ewald, Giuliano Fanelli, Federico Fernández-González, Úna FitzPatrick, Xavier Font, Itziar García-Mijangos, Rosario G. Gavilán, Valentin Golub, Riccardo Guarino, Rense Haveman, Adrian Indreica, Deniz Işık Gürsoy, Ute Jandt, John A.M. Janssen, Martin Jiroušek, Zygmunt Kącki, Ali Kavgacı, Martin Kleikamp, Vitaliy Kolomiychuk, Mirjana Krstivojević Ćuk, Daniel Krstonošić, Anna Kuzemko, Jonathan Lenoir, Tatiana Lysenko, Corrado Marcenó, Vassiliy Martynenko, Dana Michalcová, Jesper Erenskjold Moeslund, Viktor Onyshchenko, Hristo Pedashenko, Aaron Pérez-Haase, Tomáš Peterka, Vadim Prokhorov, Valerijus Rašomavičius, Maria Pilar Rodríguez-Rojo, John S. Rodwell, Tatiana Rogova, Eszter Ruprecht, Solvita Rūsina, Gunnar Seidler, Jozef Šibík, Urban Šilc, Željko Škvorc, Desislava Sopotlieva, Zvjezdana Stančić, Jens-Christian Svenning, Grzegorz Swacha, Ioannis Tsiripidis, Pavel DanTurtureanu, Emin Uğurlu, Domas Uogintas, Milan Valachovič, Yulia Vashenyak, Kiril Vassilev, Roberto Venanzoni, Risto Virtanen, Lynda Weekes, Wolfgang Willner, Thomas Wohlgemuth & Sergey Yamalov

**Keywords**

Biodiversity informatics; Database; Ecoinformatics; European Vegetation Survey; International Association for Vegetation Science; Phytosociological data; Relevé; Vegetation database; Vegetation plot

**Abbreviations**

EVA = European Vegetation Archive; EVS = European Vegetation Survey; GIVD = Global Index of Vegetation-Plot Databases; IAVS = International Association for Vegetation Science.

Abstract

The European Vegetation Archive (EVA) is a centralized database of European vegetation plots developed by the IAVS Working Group European Vegetation Survey. It has been in development

since 2012 and first made available for use in research projects in 2014. It stores copies of national and regional vegetationplot databases on a single software platform. Data storage in EVA does not affect on-going independent development of the contributing databases, which remain the property of the data contributors. EVA uses a prototype of the database management software TURBOVEG 3 developed for joint management of multiple databases that use different species lists. This is facilitated by the SynBioSys Taxon Database, a system of taxon names and concepts used in the individual European databases and their corresponding names on a unified list of European flora. TURBOVEG 3 also includes procedures for handling data requests, selections and provisions according to the approved EVA Data Property and Governance Rules. By 30 June 2015, 61 databases from all European regions have joined EVA, contributing in total 1 027 376 vegetation plots, 82% of them with geographic coordinates, from 57 countries. EVA provides a unique data source for large-scale analyses of European vegetation diversity both for fundamental research and nature conservation applications. Updated information on EVA is available online at http://euroveg.org/eva-database.

**Introduction**

Records of the occurrence and abundance of plant species found in vegetation plots represent a specific type of biodiversity data. Unlike records of individual species occurrence, these fine-resolution data that often cover large geographic ranges are suitable for analysing species co-occurrence patterns in local communities, classifying vegetation, defining vegetation types, exploring vegetation–environment relationships, bioindication, $\alpha$- and $\beta$-diversity pattern assessment and for other purposes. In addition to their use in fundamental ecological, macroecological, biogeographical and biodiversity research, vegetation-plot data are also an invaluable source of information for nature conservation, monitoring vegetation change over time and other practical applications (Dengler et al. 2012b). For these reasons, vegetation sampling in plots has been used extensively by vegetation scientists since the late 19th century (e.g. Schroter & Kirchner 1886 –1902; Warming 1895), and lately data from both historical and recent plots have been assembled in numerous national or regional databases (Dengler et al. 2011, 2012b).

The tradition and intensity of vegetation-plot sampling is much stronger in Europe than on any other continent. Of the 2.8 million vegetation plots contained in 182 databases registered in the Global Index of Vegetation-Plot Databases (GIVD) by May 2012, 66.5% were from European databases (Jansen et al. 2012b). However, the existence of multiple European databases with different formats, incompatible species lists and various access limitations has been a significant obstacle to the full use of this enormous resource for research and applications at the international scale. Therefore, the Working Group European Vegetation Survey (EVS) of the International Association for Vegetation Science (IAVS) has worked towards developing a centralized database of European vegetation plots and mechanisms for the use of these data in international analyses of European vegetation diversity and in more general biodiversity studies.

Here we present the new, centralized database called the European Vegetation Archive (EVA), give a brief history of the underlying initiative, describe the technical procedures and data property rules involved, and provide an overview of its current content.

**Brief history of the EVA initiative**

The first national projects of vegetation-plot databases were started in the 1980s in France (Brisse et al. 1995), the Netherlands (Schamin ee et al. 1989) and Switzerland (Wohlgemuth 1992). The establishment of several new databases in the 1990s was stimulated by the release of the database management program TURBOVEG (Hennekens 1996). This program was accepted in 1994 by the EVS as an international standard for storing vegetation-plot data and subsequently installed in several countries (Schaminee & Hennekens 1995).

By the early 2000s, many databases using TURBOVEG or other management software existed in different countries (Ewald 2001), but the major obstacle to their integration was their use of different species lists, usually following the taxonomy and nomenclature of the national floras or checklists. A related issue was the absence of a modern and complete taxonomic checklist of the European flora (Dengler et al. 2012a). Therefore, in 2002 the SynBioSys Europe team was established, involving vegetation scientists from several European countries who aimed at the development of an information system on European vegetation, integrating vegetation-plot data from different European databases (Schaminee et al. 2007). For this purpose, a working checklist of European vascular plants, bryophytes, lichens and macro-algae ('SynBioSys Taxon Database') was established, and the species lists used in several European vegetation-plot databases were linked to this checklist by regional experts (E. Bergmeier, J. Danihelka, W.B. Dickore, N. Ermakov and R. Haveman). However, the SynBioSys Europe project failed to obtain funding, and its original plans were not realized. Nevertheless, the amount of vegetation-plot observations stored in European databases kept growing, reaching at least 1.8 million by 2009 (Schaminee et al. 2009). The visibility of information about vegetation-plot databases was significantly enhanced by the launch of the web-based database Global Index of Vegetation-Plot Databases (GIVD; Dengler et al. 2011), which contains metadata about the content of individual databases (although not actual vegetationplot data) and is continually updated by the managers of these databases.

As a natural continuation of these developments, the European Vegetation Archive (EVA) was started in February 2012 by agreement of the owners or managers of a few key European databases, including the national databases of Austria (Willner et al. 2012), the Czech Republic (Chytry & Rafajová 2003), Germany (Ewald et al. 2012), Italy (Landucci et al. 2012), Slovakia (Šibík 2012), the Netherlands (Schaminee et al. 2012), the UK (Rodwell 2012), the database of the German federal state Mecklenburg-Vorpommern (Jansen et al. 2012a) and the Database of Dry Grasslands in the Nordic and Baltic Region (Dengler & Rusin a 2012). At the annual meeting of the European Vegetation Survey in Vienna in May 2012, the EVA Data Property and Governance Rules (http://euroveg.org/download/eva-rules.pdf) were approved, and the EVA Coordinating Board elected. Subsequently, new databases joined the initiative. In autumn 2012, the EVA website was launched (http://euroveg.org/eva-database). Managing a large database consisting of multiple databases that use different species lists and follow different data property rules required new database management software. Therefore, Stephan Hennekens developed a prototype of TURBOVEG 3, which provided the necessary tools. This was an entirely new software product rather than an update of the previous version of TURBOVEG. It was first used for the Braun-Blanquet project led by Borja Jiménez-Alfaro at Masaryk University (Brno), in which vegetation-plot data from many European databases were collected for the purpose of characterization of European vegetation alliances (Jiménez-Alfaro et al. 2014). In autumn 2013, Stephan Hennekens at Alterra (Wageningen) and Borja Jiménez-Alfaro and Ilona Knollov a at Masaryk University were appointed as EVA database managers, and contributing databases were uploaded to a single platform. Continuous intensive collaboration among these three data managers led to considerable improvement of data quality and

provided necessary feedback for further development and testing individual data management functions in TURBOVEG 3. In parallel, the update of the SynBioSys Taxon Database continued with the help of the EVA Taxonomic Advisory Board (Erwin Bergmeier, Luis Carlon, Jiřı Danihelka, Jürgen Dengler and Florian Jansen). In spring 2014 the first version of EVA was released for use in research and applied projects.

**Management of EVA data**

The EVA database stores copies of individual national or regional databases or parts of these following the EVA Data Property and Governance Rules. These contributing databases continue their activities of data acquisition and quality control, and send updated versions to EVA from time to time. EVA collaborates preferentially with comprehensive national databases or, where these do not exist, large regional or thematic databases. Smaller regional databases are expected to be primarily integrated into national databases, with their data contributed to EVA through these. The data stored in EVA remain the property of the owners of the contributing databases. For each contributing database there is a designated custodian (optionally also a deputy custodian), who is the owner, manager or other designated representative that acts on the particular database's behalf.

Data can be stored in EVA under three access regimes selected by the custodian: (1) restricted access – data are available only to EVA data contributors, and with each use requiring the custodian's explicit consent; (2) semirestricted access – data are available to EVA data contributors unless the custodian explicitly objects to their use for specific projects; and (3) free access – data are available to a wider community of users. Under regimes (1) and (2), the right to use the data also includes the other persons, besides the custodians, involved in the establishment and maintenance of the contributing databases. Additionally, other researchers can obtain restricted or semi-restricted data for research projects via cooperation with EVA data contributors. It is an aim of EVA to increase the amount of free access vegetation plots, but the specific arrangements regarding data access regimes depend entirely on the decisions of the custodians of each contributing database.

The EVA is managed using a functioning prototype of TURBOVEG 3. This prototype still does not include several functions needed for the full management of primary databases (these functions are available in TURBOVEG 2; Hennekens & Schaminee 2001); therefore, TURBOVEG 3 has not yet been provided to the managers of the contributing databases. Many of the contributing databases continue to be managed using TURBOVEG 2 (Hennekens & Schaminee 2001), in which these functions are available.

TURBOVEG 3 is linked to the SynBioSys Taxon Database, which provides the connection of each name used in the contributing databases to a unified taxonomic concept and nomenclature. The output data format can contain both the unified and the original taxon names to give the users an opportunity for checking whether the unified names correctly reflect the concepts behind the original names.

Metadata on the EVA databases are managed in cooperation with GIVD (Dengler et al. 2011). The EVA data can be requested via standard forms available at the EVA website. Using these forms, the applicant should provide the project description and the specification of the data required. Upon receipt of a data request, an EVA database manager checks which contributing databases contain the required data and sends a request for approval to the custodians of the restricted access databases or a notification to the custodians of the semi-restricted or free access databases. If the required permissions are given (restricted access data) or no objections are raised (semi-restricted access

data) within three weeks, the data are released to the applicant. Descriptions of all projects that use the EVA data are published on the EVA website.

**Current content of EVA**

By 30 June 2015, EVA comprised 61 databases, including comprehensive national databases, large regional databases and thematic databases focused on certain broad vegetation types across the whole of Europe or a large part of the continent (Appendices S1 and S2). In addition to Europe in the physico-geographic sense, EVA also includes data from Cyprus, the Anatolian part of Turkey and the Macaronesian archipelagos. If a contributing database focusing primarily on Europe also includes data from adjacent regions, especially northern Africa and the Near East, these plots are also included.

As of 30 June 2015, EVA contained a total of 1 027 376 vegetation plots from 57 countries (Appendix S3). This is nearly half of the 2 131 753 presumably nonduplicated plots contained in the GIVD-registered European databases (unpubl. data based on www.givd.info, accessed 30 June 2015). The difference between the GIVD-registered and EVA data amounts is mainly because only a stratified subset of 102 327 plots from the Dutch National Vegetation Database (which contains about 600 000 plots in total) was included in EVA in order to limit the disproportion in sampling intensity between the Netherlands and the other countries. Moreover, the GIVD figure can include some duplicated plots that are not indicated as duplicates. The geographic distribution of plots across Europe is unequal, with the highest concentration in Central and Northwest Europe and major gaps in the Nordic countries, Russia (except for Tatarstan, Bashkortostan and the Lower Volga Valley) and Belarus (Fig. 1, Appendix S3). Of the vegetation plots included in EVA, 30% provide data in restricted access regime, 66% in semi-restricted access regime and 4% in free access regime. Most plots (82%) are georeferenced with latitude/longitude coordinates, but in some regions most coordinates do not represent precise locations but either central points of grid cells (e.g. about 10 km x 10 km in size) or even larger geographic entities. The sampling year is recorded for 86% of plots, ranging from 1885 to 2014, with most plots (83% of those with recorded dates) sampled between 1971 and 2014 (Fig. 2). Plot size is indicated for 64% of the plots, and some kind of assignment to vegetation types (phytosociological syntaxa or informally defined habitat types) is available for 69% of the plots.

**Current EVA uses and outlook**

The EVA is a new data resource with huge potential to support fundamental research and applied projects at the international scale across Europe. By 30 June 2015, EVA data had been provided to 17 projects (http://euroveg.org/eva-database-eva-projects), most of them focusing on international vegetation surveys and classification of selected vegetation types. Macroecological projects focusing on plant invasions across vegetation types or patterns of fine-scale species richness of selected broad vegetation types across Europe have also made use of EVA data, as have projects focusing on species distributions. Additionally, EVA has provided species data for assessment of plant indicator values. An important development in the field of European nature conservation policy was the request in 2013 from the European Environment Agency (EEA) to determine the floristic compositions of the EUNIS habitat types.

Describing the compositions of these is an on-going programme of EEA, for which two major groups of European habitats have already been reviewed (forests in 2013, heathlands and scrub in 2014) by

using the available vegetation-plot databases and published sources (Schaminee et al. 2013, 2014), which were simultaneously integrated into EVA. We expect that many new projects will follow.

In 2014, EVA also became the exclusive European partner of a broader international initiative, sPlot (Dengler et al. 2014), hosted by the German Centre of Integrative Biodiversity Research (iDiv). The aim of sPlot is to collect representative data sets of vegetation plots from global

The EVA still has many gaps in geographic coverage as well as in representation of certain vegetation types. It also inevitably contains various biases inherent to sets of data assembled from multiple sources and originally collected for various purposes (Michalcov a et al. 2011; Chytry et al. 2014). However, the overall effect of these biases is relatively small in comparison with the wealth of information provided, and, moreover, the biases can partly be accounted for when selecting, analysing and interpreting the data. The great willingness shown by many database owners and managers to share their data in just the first three years of EVA's existence inspires optimism that it will continue to grow and that data gaps will shrink.

**Acknowledgements**

Our major thanks go to thousands of European vegetation scientists of several generations who collected the original vegetation-plot data in the field, published them or made their unpublished data available to others, and to those who spent myriad hours digitizing data and managing the contributing databases. EVA data management has been partly funded by the Czech Science Foundation (Centre of Excellence PLADIAS, 14-36079G).

**References**

Brisse, H., de Ruffray, P., Grandjouan, G. & Hoff, M. 1995. The phytosociological database "SOPHY" Part 1: calibration of indicator plants, Part II: socio-ecological classification of the relev es. Annali di Botanica 53: 177–223.

Chytry, M. & Rafajová, M. 2003. Czech National Phytosociological Database: basic statistics of the available vegetation-plot data. Preslia 75: 1–15.

Chytry, M., Tichy, L., Hennekens, S.M. & Schaminee, J.H.J. 2014. Assessing vegetation change using vegetation-plot databases: a risky business. Applied Vegetation Science 17: 32–41.

Dengler, J. & Rusin a, S. 2012. Database dry grasslands in the Nordic and Baltic Region. Biodiversity & Ecology 4: 319.

Dengler, J., Jansen, F., Glockler, F., Peet, R.K., De C aceres, M., Chytry, M., Ewald, J., Oldeland, J., Lopez-Gonzalez, G., ( . . .) & Spencer, N. 2011. The Global Index of Vegetation-Plot Databases (GIVD): a new resource for vegetation science. Journal of Vegetation Science 22: 582–597.

Dengler, J., Berendsohn, W., Bergmeier, E., Chytry, M., Danihelka, J., Jansen, F., Kusber, W.-H., Landucci, F., Muller, A., ( € . . .) & von Raab-Straube, E. 2012a. The need for and the requirements of EuroSL, an electronic taxonomic reference list of all European plants. Biodiversity & Ecology 4: 15–24.

Dengler, J., Oldeland, J., Jansen, F., Chytry, M., Ewald, J., Finckh, M., Glockler, F., Lopez-Gonzalez, G., Peet, R.K. & Schaminee, J.H.J. (eds.) 2012b. Vegetation databases for the 21st century. Biodiversity & Ecology 4: 15–24.

Dengler, J., Bruelheide, H., Purschke, O., Chytry, M., Jansen, F., Hennekens, S.M., Jandt, U., Jim enez-Alfaro, B., Kattge, J., (. . .) & sPlot Consortium 2014. sPlot – the new global vegetation-plot database for addressing trait–environment relationships across the world's biomes. In: Mucina, L., Price, J.N. & Kalwij, J.M. (eds.) Biodiversity and vegetation: patterns, processes, conservation. p. 90. Kwongan Foundation, Perth, AU.

Ewald, J. 2001. Der Beitrag pflanzensoziologischer Datenbanken zur vegetationsokologischen Forschung. € Berichte der Reinhold-Tuxen-Gesellschaft € 13: 53–69.

Ewald, J., May, R. & Kleikamp, M. 2012. VegetWeb – the national online repository of vegetation plots from Germany. Biodiversity & Ecology 4: 173–175.

Hennekens, S.M. 1996. TURBOVEG: Software package for input, processing and presentation of phytosociological data. User's guide. University of Lancaster, Lancaster, UK.

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

Jansen, F., Dengler, J. & Berg, C. 2012a. VegMV – the vegetation database of Mecklenburg-Vorpommern. Biodiversity & Ecology 4: 149–160.

Jansen, F., Glockler, F., Chytr € y, M., De Cáceres, M., Ewald, J., Lopez-Gonzalez, G., Oldeland, J., Peet, R.K., Schaminée, J.H.J. & Dengler, J. 2012b. News from the Global Index of Vegetation-Plot Databases (GIVD): the metadata platform, available data, and their properties. Biodiversity & Ecology 4: 77–82.

Jimenez-Alfaro, B., Chytry, M., Hennekens, S., Knollová, I., Schaminee, J. & Braun-Blanquet project consortium 2014. The Braun-Blanquet project: evaluating and characterizing European vegetation alliances. In: Carni, A., Juvan, N. & Ribeiro, D. (eds.), 23rd International Workshop of the European Vegetation Survey, Ljubljana 8–12 May 2014, Book of Abstracts. p. 33. ZRC SAZU, Ljubljana, SI.

Kattge, J., D ıaz, S., Lavorel, S., Prentice, I.C., Leadley, P., Bonisch, G., Garnier, E., Westoby, M., Reich, P.B., ( € . . .) & Wirth, C. 2011. TRY – a global database of plant traits. Global Change Biology 17: 2905–2935.

Landucci, F., Acosta, A.T.R., Agrillo, E., Attorre, F., Biondi, E., Cambria, V.E., Chiarucci, A., Del Vico, E., De Sanctis, M., (. . .) & Venanzoni, R. 2012. VegItaly: the Italian collaborative project for a National Vegetation Database. Plant Biosystems 146: 756–763.

Michalcov a, D., Lvončık, S., Chytry, M. & Hájek, O. 2011. Bias in vegetation databases? A comparison of stratified–random and preferential sampling. Journal of Vegetation Science 22: 281–291.

Rodwell, J. 2012. UK National vegetation classification database. Biodiversity & Ecology 4: 381.

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

Schamin ee, J.H.J., van Wirdum, G. & Westhoff, V. 1989. Naar een nieuw overzicht van de plantengemeenschappen in Nederland. De Levende Natuur 90: 204–209.

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

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

Schaminee, J.H.J., Hennekens, S.M. & Ozinga, W.A. 2012. The Dutch national vegetation database. Biodiversity & Ecology 4: 201–209.

Schaminee, J.H.J., Chytry, M., Hennekens, S.M., Jim enez-Alfaro, B., Mucina, L., Rodwell, J.S. & Tichy, L. 2013. Review of EUNIS forest habitat classification. European Environment Agency [report no. EEA/NSV/13/005], Copenhagen, DK.

Schaminee, J.H.J., Chytry, M., Hennekens, S.M., Janssen, J.A.M., Jiménez-Alfaro, B., Knollová, I., Mucina, L., Rodwell, J.S. & Tichy, L. 2014. Vegetation analysis and distribution maps for EUNIS habitats. European Environment Agency [report no. EEA/NSV/13/005], Copenhagen, DK.

Schröter, C. & Kirchner, O. 1886 –1902. Die Vegetation des Bodensees. Stettner, Lindau, DE.

Šibík, J. 2012. Slovak vegetation database. Biodiversity & Ecology 4: 429.

Warming, E. 1895. Plantesamfund – Grundtræk af den økologiske plantegeografi. P.G. Philipsens, København, DK.

Willner, W., Berg, C. & Heiselmayer, P. 2012. Austrian vegetation database. Biodiversity & Ecology 4: 333.

Wohlgemuth, T. 1992. Die vegetationskundliche Datenbank. Schweizerische Zeitschrift fur Forstwesen € 143: 22–36.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1. An overview of the vegetation-plot databases included in EVA with their GIVD codes, custodians and numbers of all plots and georeferenced plots.

Appendix S2. Published references to EVA databases.

Appendix S3. Numbers and densities of vegetation plots included in EVA by countries.

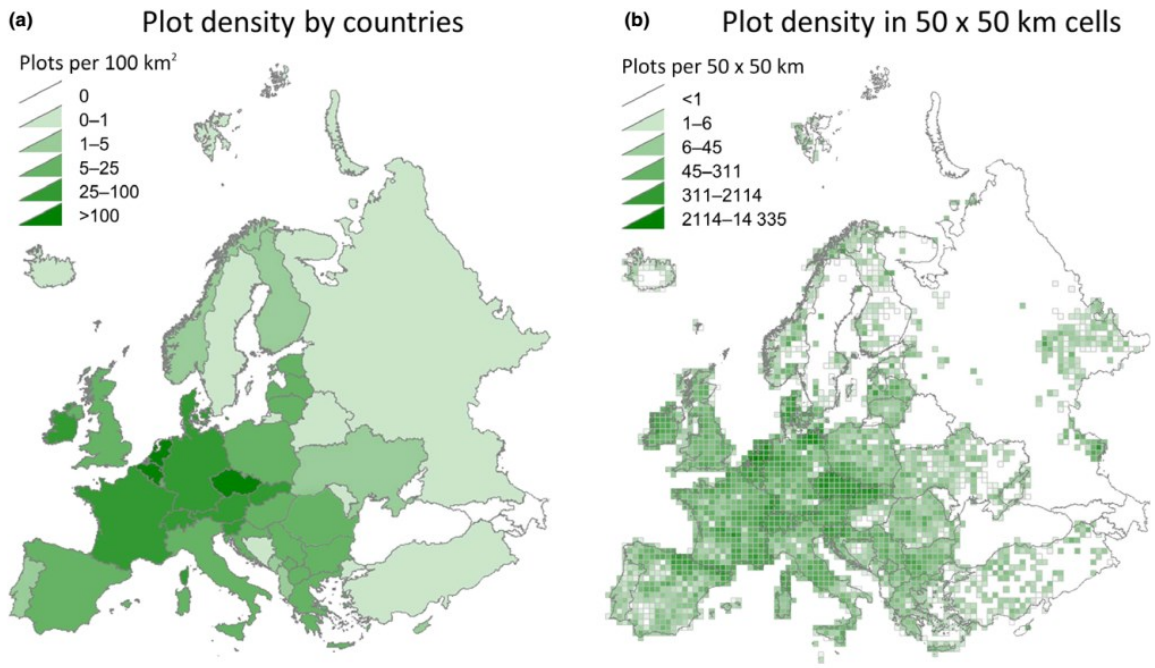


Fig. 1. Distribution of the vegetation plots included in EVA per country (a) and density of georeferenced plots in 50 km 9 50 km grid cells (b). Plot density by countries was recalculated to 100-km2 units. Some differences in the content of these two figures are due to the fact that in some countries many plots are not georeferenced. Data accessed on 30 June 2015.

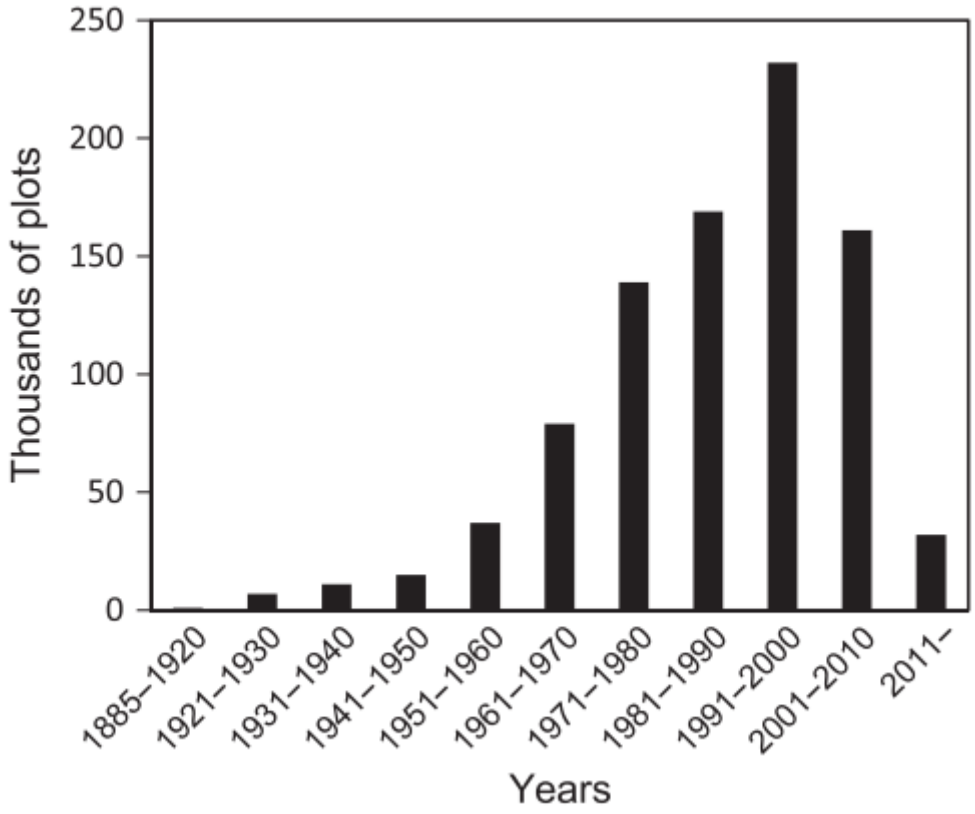


Fig. 2. Sampling dates of the vegetation plots included in EVA.