Received: 26 July 2021 | Revised: 17 November 2021 | Accepted: 4 January 2022

DOI: 10.1111/avsc.12460

Applied Vegetation Science IAVS

RESEARCH ARTICLE

Two sides of one medal: Arable weed vegetation of Europe in phytosociological data compared to agronomical weed surveys

Jana Bürger[1] | Filip Küzmič[2] | Urban Šilc[2] | Florian Jansen[3] | Erwin Bergmeier[4] | Milan Chytrý[5] | Alicia Cirujeda[6] | Silvia Fogliatto[7] | Guillaume Fried[8] | Denise F. Dostatny[9] | Bärbel Gerowitt[1] | Michael Glemnitz[10] | José L. González-Andújar[11] | Eva Hernández Plaza[11] | Jordi Izquierdo[12] | Michaela Kolářová[13] | Zdeňka Lososová[5] | Helen Metcalfe[14] | Jevgenija Ņečajeva[15] | Sandrine Petit[16] | Gyula Pinke[17] | Valerijus Rašomavičius[18] | Christoph von Redwitz[19] | Matthias Schumacher[20] | Lena Ulber[19] | Francesco Vidotto[7]

[1]Crop Health, Faculty of Agricultural and Environmental Sciences, University of Rostock, Rostock, Germany

[2]Research Centre of the Slovenian Academy of Sciences and Arts (ZRC SAZU), Jovan Hadži Institute of Biology, Ljubljana, Slovenia

[3]Landscape Ecology, Faculty of Agricultural and Environmental Sciences, University of Rostock, Rostock, Germany

[4]Department of Vegetation and Phytodiversity Analysis, Georg-August University of Göttingen, Göttingen, Germany

[5]Department of Botany and Zoology, Faculty of Science, Masaryk University, Brno, Czech Republic

[6]Plant Protection, Agrifood Research and Technology Centre of Aragón (CITA), Zaragoza, Spain

[7]Department of Agricultural, Forest and Food Sciences (DISAFA), University of Torino, Grugliasco, Italy

[8]Plant Health Laboratory, Anses, Montferrier-sur-Lez, France

[9]National Centre for Plant Genetic Resources, Plant Breeding and Acclimatization Institute – National Research Institute, Błonie, Poland

[10]Leibniz Centre for Agricultural Landscape Research, Müncheberg, Germany

[11]Instituto de Agricultura Sostenible, Spanish National Research Council (CSIC), Córdoba, Spain

[12]Department of Agri-food Engineering and Biotechnology, Universitat Politècnica de Catalunya, Castelldefels, Spain

[13]Department of Agroecology and Crop Production, Faculty of Agrobiology, Food and Natural Resources, Czech University of Life Sciences Prague, Praha, Czech Republic

[14]Sustainable Agricultural Sciences, Rothamsted Research, Harpenden, United Kingdom

[15]Institute for Plant Protection Research 'Agrihorts', Latvia University of Life Sciences and Technologies, Jelgava, Latvia

[16]Agroécologie, AgroSup Dijon, INRAE, Université de Bourgogne Franche-Comté, Dijon, France

[17]Faculty of Agricultural and Food Sciences, Széchenyi István University, Mosonmagyaróvár, Hungary

[18]Nature Research Centre, Institute of Botany, Vilnius, Lithuania

[19]Julius Kuehn-Institut, Federal Research Centre for Cultivated Plants, Institute for Plant Protection in Field Crops and Grassland, Braunschweig, Germany

[20]Department of Weed Science, Institute of Phytomedicine, Faculty of Agricultural Sciences, University of Hohenheim, Stuttgart, Germany

Correspondence
Jana Bürger, Crop Health, Faculty of Agricultural and Environmental Sciences, University of Rostock, Rostock, Germany.
Email: jana.buerger@uni-rostock.de

Funding information
The work on this project was supported by Deutsche Forschungsgemeinschaft DFG (grant number BU 3097/1-1 for Jana

Abstract

Questions: Two scientific disciplines, vegetation science and weed science, study arable weed vegetation, which has seen a strong diversity decrease in Europe over the last decades. We compared two collections of plot-based vegetation records originating from these two disciplines. The aim was to check the suitability of the collections for joint analysis and for addressing research questions from the opposing domains.

Bürger), Czech Science Foundation (grant number 19-28491X for Milan Chytrý and Zdeňka Lososová), and by Slovenian Research Agency (research core funding No. P1-0236 for Urban Šilc and Young researcher programme for Filip Küzmič)

Co-ordinating Editor: Borja Jiménez-Alfaro

We asked: are these collections complementary? If so, how can they be used for joint analysis?
Location: Europe.
Methods: We compared 13 311 phytosociological relevés and 13 328 records from weed science, concerning both data collection properties and the recorded species richness. To deal with bias in the data, we also analysed different subsets (i.e., crops, geographical regions, organic vs conventional fields, center vs edge plots).
Results: Records from vegetation science have an average species number of 19.0 ± 10.4. Metadata on survey methodology or agronomic practices are rare in this collection. Records from weed science have an average species number of 8.5 ± 6.4. They are accompanied by extensive methodological information. Vegetation science records and the weed science records taken at field edges or from organic fields have similar species numbers. The collections cover different parts of Europe but the results are consistent in six geographical subsets and the overall data set. The difference in species numbers may be caused by differences in methodology between the disciplines, i.e., plot positioning within fields, plot sizes, or survey timing.
Conclusion: This comparison of arable weed data that were originally sampled with a different purpose represents a new effort in connecting research between vegetation scientists and weed scientists. Both collections show different aspects of weed vegetation, which means the joint use of the data is valuable as it can contribute to a more complete picture of weed species diversity in European arable landscapes.

KEYWORDS
agriculture, arable weeds, Arable Weeds and Management in Europe data collection, European Weed Vegetation database, phytosociology, segetal plants, species richness, vegetation survey, vegetation-plot data, weed survey

1 | INTRODUCTION

Surveying weed vegetation, also called *segetal* or *arable vegetation*, has a long tradition, especially in Europe. Phytosociologists were the first to conduct surveys in and around arable fields, aiming to classify plant communities and describe the spatial and temporal variability of those plant communities. Examples from northeastern Germany date back to 1944, with an increasing number of surveys from the early 1960s (Manthey, 2003). These early surveys and classifications already identified main drivers shaping species composition in fields, associated with site conditions (climate, soil type and pH) or crop types (e.g., cereals vs row crops) as well as management (overview in Hüppe & Hofmeister, 1990).

A deeper insight into how field management influences arable weed vegetation was initiated around the 1980s and 1990s through research coming from an agronomic motivation (Hanzlik & Gerowitt, 2016). Weed scientists initiated surveys as a complement to field experiments as surveys allow to cover large spatial scales with a wider range of environmental conditions (Andersson & Milberg, 1998; Salonen, 1993). They aimed to understand weed species occurrence and community composition in relation to certain cropping practices. The agronomic purposes of such surveys were to identify important and emerging weed species or shifts in community composition caused by changes in land-use practices or climate and different tolerances to herbicides. A decade later, there was a growing interest to promote agro-ecological weed management approaches for a better balance of production, ecosystem services provided by weeds and the preservation of rare species (Storkey & Westbury, 2007).

As a result, there are currently two scientific disciplines holding data on arable vegetation in Europe: vegetation science and weed science. They overlap in parts of their scientific scope (Figure 1), but joint research is rare. Sometimes, publications find their way into the other domain. This has become much easier since the dawn of the digital era with large, easily searchable literature databases, open access and scientific networks expanding online (e.g., Dengler et al., 2011). Nevertheless, only a few research groups bridge both disciplines.

Another development facilitated in the digital era is the collection of formerly disparate data sets in large databases and networks like GBIF for species occurrence data (GBIF: The Global Biodiversity Information Facility, 2020), TRY for plant-trait data (Kattge et al., 2020), the CESTES database for combined species, trait and environment data (Jeliazkov et al., 2020), or EVA and sPlot for vegetation-plot data from Europe (Chytrý et al., 2016) and globally (Bruelheide et al., 2019). These collections have enabled research on a much larger geographical scale than before, mainly by making primary research data easier to access and to re-use (Chytrý et al., 2014).

The need for biodiversity monitoring (status quo and trends) has also been recognised for arable landscapes (Hurford et al., 2020), which have seen the strongest decrease in diversity and population sizes of individual species compared to other habitats in Europe over the last decades (Eichenberg et al., 2021; Leuschner & Ellenberg, 2017). This need has spurred the development of efficient sampling schemes (Wietzke & Leuschner, 2020), but also a call for an integrative data culture using data from disparate sources and involving stakeholders (Kühl et al., 2020).

In this paper, we report on two collections of vegetation-plot data that were assembled to make records of weed vegetation visible and available for further analysis. These collections broadly originate from the two scientific disciplines concerned with weed vegetation mentioned earlier. This is most notable from the bodies that helped to mobilise the data: the International Association for Vegetation Science (http://iavs.org), through its working group *European Vegetation Survey* (EVS) and the *European Vegetation Archive* (EVA) database, on the one hand and the European Weed Research Society (https://ewrs.org), through its working group *Weeds and Biodiversity*, on the other.

Records of arable weed vegetation in EVA date back to the 1930s but have very little associated information regarding the agricultural background of the surveyed fields or metadata on sampling methodology. In contrast, the weed science-based collection holds data starting from the 1990s but is accompanied by extensive management information and detailed metadata of survey methods.

In this paper, we aim to check the suitability of the collections for joint analysis and for addressing research questions from the opposing domains. The paper is structured into three parts: (1) description of the collections and comparison on a qualitative level, (2) basic comparative statistical analysis of both collections, and (3) exploration of whether the two collections are complementary and how they can be used for potential joint analyses.

2 | MATERIAL AND METHODS

2.1 | Terms and definitions

Plot-based vegetation data are lists of all (vascular) plant species recorded from a discrete area of variable size with a measure of each species' abundance (Braun-Blanquet, 1964; du Rietz, 1936). However, details of the sampling scheme might differ significantly between studies. Plots differ in size and shape. On arable land, plots originate either from the *field core*, i.e., the inner field with some distance to the field's boundary (Figure 2), from the *field edge*, i.e., next to the boundary of the cultivated area, or from the *field margin*, which is usually outside of the cultivation area, but immediately adjacent to the field boundary (grass margins, field strips etc.). The vegetation data are accompanied at least by a description of the field location and sampling year, but surveyors may also record a number of site characteristics, e.g., soil pH, slope, altitude or crop species.

In this paper, we use the following terminology to distinguish between observation units: a *field* is a distinct parcel of land and usually the smallest unit of a farmer's management (Figure 2). Within a field or field parts, the surveyor places one or several *plots*, i.e., distinct small areas on which the actual vegetation is recorded. Each time the vegetation is recorded, this is called a plot *observation*. When compiling a database, each line or *record* corresponds to one observation, which can be identified by field ID, possibly field part attribute, plot ID and observation date. Observations may be aggregated to a higher-level observation unit, for example, by summing up all the observations from the same field or field part; the aggregated data are also called *records*.

2.2 | Data collection: weed vegetation subset of the European Vegetation Archive (EVA-W)

This data collection was obtained from the EVA (Chytrý et al., 2016), which includes the *European Weed Vegetation Database* (GIVD ID EU-00-028; Küzmič et al., 2020) and several other databases with plot observations of weed vegetation. The plots from EVA were first selected using the expert system based on species lists characterising

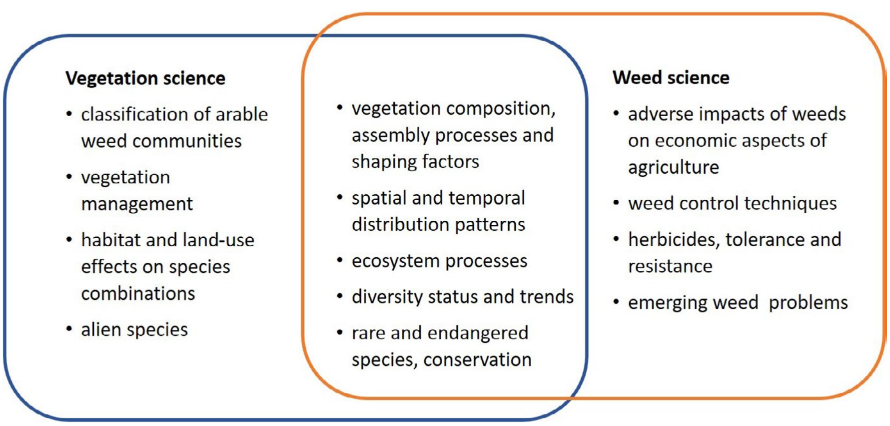

FIGURE 1 Overlap between the objectives of vegetation science and weed science in arable landscapes

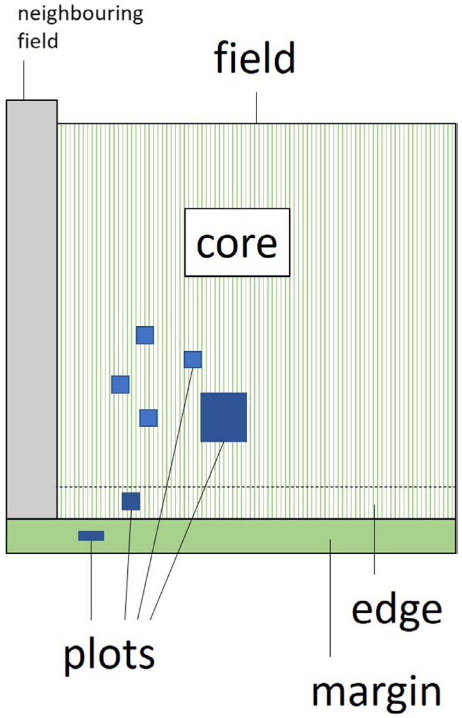

FIGURE 2 Terminology of weed vegetation surveys used in this study

weed vegetation classes from EuroVegChecklist (Mucina et al., 2016). Plots assigned to the classes *Papaveretea rhoeadis*, *Chenopodietea*, *Sisymbrietea* and *Digitario sanguinalis-Eragrostietea minoris* were retained.

The data stored in the EVA come from various sources, e.g., scientific papers, technical reports, monographs, doctoral theses and final projects of students such as master theses. The selected data sets either come from studies of several different vegetation types including weed vegetation, or studies focussing on a selected crop type or area with weed vegetation as the main subject. Since the main aim in many of these studies was the floristic and ecological characterisation of plant communities (Braun-Blanquet, 1964), usually well developed and species-rich communities were recorded. The metadata (e.g., crop type, plot position in the field, georeference or record date) are often missing either because such data were not recorded in the original publication or not digitised.

Because the records with weed vegetation were selected from EVA according to the occurring species and species communities, they could also originate from non-arable sites with ruderal or disturbed vegetation. For the comparison to the Arable Weeds and Management in Europe (AWME) data collection, we wanted to restrict the selection to plots that originated from arable sites. We therefore performed a second selection step and identified the plots from arable land either from specific keywords in the location description (i.e., 'cereals', 'rye field') or when the species list included a crop species with a cover class >2 on the Braun-Blanquet scale. We added a georeference to approximately 400 records which only had a description of their location via geocoding. Afterwards, we limited the data set to georeferenced plots and excluded a small number of plots with longitudes >39°E (Russia) due to the scarcity of data in this region. Finally, we only used records dating from 1980 onwards to make the time periods of both data collections more comparable. Species abundance was transformed to presence/absence data. Within our study, we use the abbreviation EVA-W (*weed subset of EVA*) for this data set. Details on the used data sets can be found in Appendix S1.

2.3 | Data collection: Arable Weeds and Management in Europe (AWME)

Arable Weeds and Management in Europe (GIVD ID EU-00-033; Bürger et al., 2020) is a collection of vegetation-plot data created through a call to members of the Working Group *Weeds and Biodiversity* of the European Weed Research Society in June 2019. The aim of this initiative was to make the data visible and more readily available for further research. We received 36 data sets of weed surveys (for details see Appendix S2) conducted between 1996 and 2018. Some of them have not been previously published. Temporal and spatial scales vary strongly between the data sets, with study periods between 1 and 15 years long and spatial extents of the study areas varying from less than 1000 ha up to national surveys. Nearly all records are georeferenced. However, in two data sets, the coordinates were assigned based on the settlement rather than the exact field location to respect an anonymity agreement with the field owners.

We found a variety of sampling designs. Some studies were done with a narrow focus on one crop or study region, with a number of fields surveyed exactly once, sometimes within a study period of more than one year. Other studies were more complex. Some visited the same field more than once. Other studies took samples in various parts of the same fields, for example, aiming to study differences between field core and field edge or, more rarely, between plots treated with herbicides and control plots.

Most surveys were carried out at a time of highly developed vegetation, mainly before crop harvest. Others aimed at assessing the vegetation before a necessary weed control decision, or at assessing the efficacy of a herbicide treatment. The latter surveys took place in very early crop development stages, i.e., few weeks after crop sowing, or in spring prior to the period of rapid vegetation growth. To quantify the species abundance, studies used various forms of cover estimates, including original and modified scales by Braun-Blanquet (1964), Londo (1976), Barralis (1976), Rasiņš and Tauriņa (1982), or measured density by counting individual plants.

Most studies performed at the field edge excluded species before analysis that were spill-over from the field margins (i.e., woody or grassland plants that were not considered to be arable species). Field observations that recorded no weed plants (species number = 0) appeared in some cases where the study design included a survey after herbicide treatment, but were not included in the collection.

All data sets were formatted to uniform spreadsheets of species data, a catalogue of field metadata and the accompanying management and environmental data prior to assembly in combined text files. Records were classified by field part (core/edge/margin) and given a marker whether weed control had taken place before the survey (treated/untreated).

Observations made on the same date from the same field, field part and weed control status were aggregated into one record. The plot size of the aggregated record was set to the sum of the original plots. Species occurrence was transformed to presence/absence format. The timing of the weed survey was classified in AWME in relation to the cropping season (early, mid- or late season), which may differ in calendar dates depending on the sowing time of the crop.

2.4 | Analysis

The plant taxonomic concepts and nomenclature of both collections were unified according to the Euro+Med PlantBase (Euro+Med, 2021). Only vascular plants were retained since bryophytes, fungi and algae are rarely recorded in arable vegetation. Infraspecific taxa were merged to the species level. Taxa identified only to the genus level were kept in the individual records and used for plot species richness calculation but were removed for the accumulation curves to prevent inflation of species numbers. For individual records, we judged it neglectable if nested taxa (e.g., both *Chenopodium* and *Chenopodium album*) were present (Jansen & Dengler, 2010).

Some taxonomically problematic species (e.g., from the genera *Taraxacum* and *Achillea*) were merged into aggregates following mostly concepts in Euro+Med (2021), Ehrendorfer (1967) and Martinčič et al. (2007). If the cultivated crop was included in the vegetation record, we deleted it. Volunteer crops were retained. We found no records duplicated in both collections.

To compare species richness between data collections, we plotted sample-based *species accumulation curves*, generated by repeated accumulation in random order (Gotelli & Colwell, 2001; Oksanen et al., 2020; Ugland et al., 2003).

Then we calculated species number per record for both data collections and used this measure as our unit of analysis for subsequent comparisons. We tested the effect of various variables on species number per record. We used Kruskal–Wallis and Wilcoxon rank tests to assess whether species numbers differed significantly between groups (regarding data collection, crop, weed control treatment or survey date). The effect of plot size on species number per record was tested by linear log–log regression, followed by an ANOVA. The correlation of species number and latitude was tested using Spearman's rank correlation coefficient.

The continental-scale data collection comprises large gradients in environmental conditions like climate and soils. These correspond to differences in the cultivated crops (species and management systems) and the associated weed species pools. The environmental gradients and the regionally adapted management introduce variability in the data sets, which may impede the analysis of causal relationships, e.g., between crop type and species number. Furthermore, the sampling density in geographical space differs significantly in both data collections.

Although most of these challenges could be met by advanced statistical methods in a study aiming at deeper analysis, we chose here a relatively simple approach. To account for the bias, we use smaller geographical subsets for comparisons additionally to the overall data sets. These subsets within two data collections cover either the same or adjacent geographical areas with similar sampling densities. The subsets have the advantage of more similar cropping systems, environment and agricultural cropping history which makes comparisons easier. For each of six regions, we present species accumulation curves and the comparison of species numbers per record in different crops.

Data preparation, geocoding and analysis were performed using the statistical software environment R 4.0.4 (R Core Team, 2021).

3 | RESULTS

3.1 | Quantitative description

Arable Weeds and Management in Europe contains more than twice as many records as the EVA-W, but the AWME records originate from a lower number of distinct locations. For many fields, the AWME contains more than one record, for example from repeated visits to the same field. The ratio of records to distinct geographical coordinates is 2:1 in EVA-W and 6:1 in AWME. In EVA-W, coordinates are often assigned to records according to the locality description; therefore, different fields of the same settlement can have the same coordinates. In AWME, some studies — as mentioned before — have repeated records from the same fields.

The two data collections cover different parts of Europe (Figure 3a). There are countries for which both collections contribute records, but largely the distribution of plots is largely complementary between the collections. Strongholds of vegetation-plot data collections like the Netherlands and the Czech Republic are very well covered in EVA-W. Countries with large agricultural areas like Germany and France feature highly in the AWME but with few records in EVA-W. Restricting the time span of EVA-W to plots surveyed beginning in 1980 substantially lowered the coverage in Spain, Austria, Switzerland, Poland and Serbia, when compared to the potentially available records in the EVA.

The EVA-W collection contains nearly twice as many species as the AWME collection (Figure 3b). The increase of the species accumulation curve is steeper and higher for EVA-W. This corresponds to the number of rarely recorded species: about 1200 taxa in EVA-W are included in less than 10 records, compared to 550 taxa in AWME (Table 1). The majority of species recorded in AWME are also present in EVA-W (shared species in Figure 3c), but EVA-W contains a majority of species that are not present in AWME.

In AWME, 16% of taxa were identified only at the genus level. This proportion is approximately twice the proportion in EVA-W. In AWME, these genus-level taxa also occur in a larger proportion of fields. The most frequent of these taxa is *Matricaria* (which can indicate several species, including *Matricaria chamomilla* and

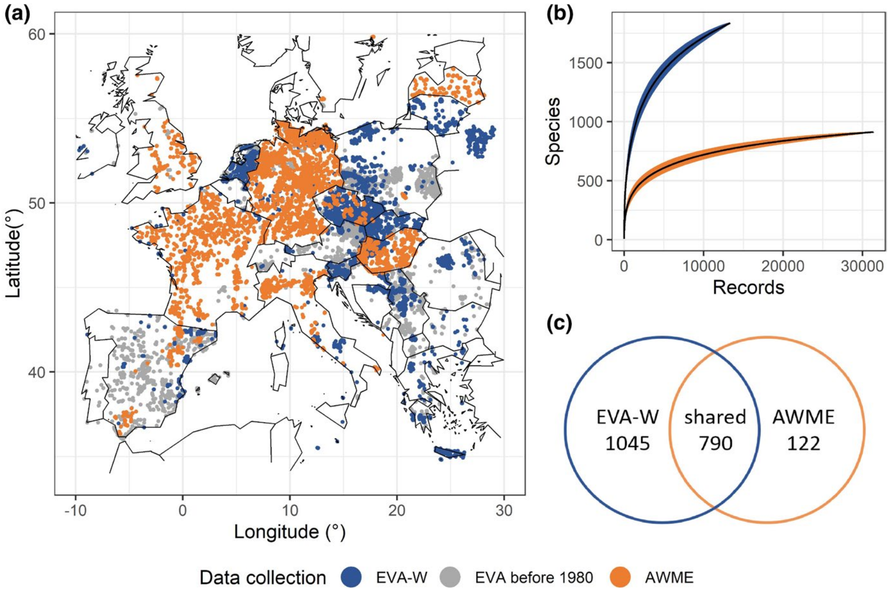

FIGURE 3 Records and species in the collections of weed vegetation data: Arable Weed and Management in Europe (AWME) and European Vegetation Archive subset Weeds (EVA-W). (a) Field locations of vegetation plots in the two data collections. Locations shaded in grey are records stored in the European Vegetation Archive taken before 1980. (b) Species accumulation curves. (c) Venn diagram of shared species

Tripleurospermum inodorum), which is present in 9% of records in AWME.

Summary statistics regarding plot metadata are shown in Table 2. Information on the surveyed field part or management system in the field was available only for a very minor part of the EVA-W records. We therefore excluded this information from the further analysis and set the value of these variables to unknown (NA) for the whole EVA-W data set.

3.2 | Comparisons of species number per record

3.2.1 | Field parts, plot size and latitude

Species numbers per record differ considerably between the two data collections, with a mean of 8.5 species [standard deviation (SD) ± 6.4, median 7] in the AWME collection and 19.0 species in the EVA-W (SD ± 10.4, median 18), ($\chi^2 = 10\,855$, $df = 1$, $p < 0.001$), despite larger plot sizes in AWME than in EVA-W (Table 2). Within the AWME, core plots have significantly lower species numbers than edge plots: the averages are 8.0 (SD ± 5.8, median 7) and 17.0 (SD ± 9.4, median 15) species, respectively, which are both significantly lower than the EVA-W average ($\chi^2 = 12{,}873$, $df = 2$, $p < 0.001$). Approximately half of the species in AWME are found in edge and core plots, 30% only in core plots and 17% only in edge plots (Figure 4b).

Linear log–log models relating species number and plot size (Figure 4a) suggest an increase of species number with plot size for the EVA-W and the AWME edge plots and a decrease for AWME core plots, but had rather poor fits (R^2 between 0.06 and 0.3). We found no association between latitude and species numbers.

3.2.2 | Crops

Species numbers differ little between crops in AWME (lowest average: cereal = 7.4 ± 6.7, highest average: potato = 10.7 ± 9.0). The range is larger in EVA-W, where records from maize and sunflower plots have the lowest average species number and cereal plots the highest (Table 3). A linear mixed-effects model showed a significant overall effect of crop on species number ($F = 71.26$, $p < 0.001$).

Within crops, AWME field-core plots have consistently the smallest species number (Figure 5, case number in Appendix S3). In three out of five crops with sufficient plot numbers (not including potato), species number per plot is on average higher or the same in the AWME edge plots as in EVA-W.

Species numbers are lower in AWME records after weed control ('treated': herbicide or mechanical weeding) compared to records taken before treatment or on field parts excluded from treatment as part of the study design. For the EVA-W plots, no information is available on whether they received a weed control treatment. Records in AWME from treated plots have on average 7.9 species, records from untreated plots 8.7 species, compared to the 19.0 species in the EVA-W plots (all significantly different at $p < 0.001$, Kruskal–Wallis test). The same pattern is found for the comparison of treated vs untreated plots within a crop (Figure 6).

The timing of the weed survey was classified for the AWME data in relation to the cropping season. Early surveys, taken after crop sowing and mid-season surveys, mostly reported similar species numbers. Records taken late in the season, near to harvest, usually show higher species numbers except for beet and untreated maize plots in AWME.

3.3 | Geographical subsets

The findings from the overall data set were confirmed for six geographical subsets (Figure 7). The AWME subsets include between 150 and 300 species, which is much less than the EVA-W subsets, which contain between 250 and over 600 species. In all subsets, the species number is higher in EVA-W records than in AWME records,

TABLE 1 Quantitative properties of two data collections regarding records and taxa

	EVA-W	AWME
Number of records	13 311	31 328
Unique georeferences	6824	4945
Number of taxa	1977	1082
Proportion of taxa occurring in only 1 record	24%	18%
Proportion of taxa occurring in max. 10 records	62%	52%
Taxa recorded only at genus level	141	170
Mean frequency of these taxa (proportion of records)	0.09%	0.5%
Maximum frequency of these taxa (proportion of records)	2%	9%

Abbreviations: AWME, Arable Weed and Management in Europe; EVA-W, European Vegetation Archive subset on Weeds.

TABLE 2 Available metadata in two data collections

		EVA-W	AWME
Plot size (m²)	Min	0.09	0.125
	Median	25	120
	Max	9 999	20 000
	Unknown	15%	0%
Crops (% of records)	Beet	3	3
	Cereal	54	42
	Maize	11	29
	Oilseed rape	5	13
	Potato	8	0.6
	Sunflower	0.6	4
	Other or unknown	18.5	8.6
Sowing seasons (% of records)	Autumn	9	37
	Spring	26	47
	Unknown	65	16
Cropping system (% of records)	Conventional	0	98
	Organic	0	1
	Unknown	100	1
Field part (% of records)	Field core	0	94
	Field edge	0	6
	Margin	0	0.1
	Unknown	100	0

Abbreviations: AWME, Arable Weed and Management in Europe; EVA-W, European Vegetation Archive subset on Weeds.

except for records from organic fields (subsets from Spain and Czechia), which consistently reach the same species richness as the records from EVA-W.

4 | DISCUSSION

In our study, we show that arable vegetation records collected for vegetation science research have consistently higher species numbers than records from weed science surveys. In the unfiltered comparison of species numbers, the EVA-W and AWME collections seem to originate from different vegetation and quite different conclusions on the state of species richness in European arable landscapes would be drawn depending on the data collection used.

However, the difference decreases substantially when focussing on subsets from organic management or records taken at field edges. This is in line with many studies on the effect of organic farming on weed species richness (Hyvönen et al., 2003; Winqvist et al., 2011) and the possible role of field-margin vegetation in preserving diversity in arable landscapes (Fried et al., 2009; Metcalfe et al., 2019). The difference between species numbers in AWME and EVA-W also depends on crop types: it is higher in autumn-sown cereals and oilseed rape but much less pronounced in maize and other crops sown in late spring.

First, we will discuss which methodological factors may have contributed to our results and then give an outlook how the data might be used with respect to the encountered biases.

4.1 | Methodological causes for the differences between data sets

The majority of EVA-W plots have limited associated metadata except for plot location, plot size, crop and survey date. Therefore, we must make assumptions regarding the survey methodology of these data. In contrast, information on sampling methodology is available in detail from the publications associated with the AWME records (Appendix S2).

The EVA-W collection consists mainly of data from a phytosociological background (EVA defines itself as a repository of 'records of plant-taxon co-occurrence at particular sites, also called phytosociological relevés'; http://euroveg.org/eva-database; 10 Nov 2021). In this context, preferential sampling was and maybe still is the most commonly used sampling strategy because for research purposes such as vegetation classification, it aims to record typically developed, species-rich stands of a certain vegetation type (Moore et al., 1970; Mueller-Dombois & Ellenberg, 1974). Although choosing plots at the field edge was once considered wrong (Meisel, 1979), nowadays field edges and extensively managed fields are the only place where arable weed communities can develop well, which is even more true for studies of rare and endangered species (Meyer et al., 2015; Pinke, 2000; van Elsen, 1989). Consequently, it is highly likely that most records from the EVA-W collection (filtered to be from 1980 onwards) were taken near the field boundary or in extensively managed fields.

Due to spill-over from adjacent habitats, vegetation records from edges and margins include many species that are not arable species, or which would not be able to persist in the field interior (Metcalfe et al., 2019; Munoz et al., 2020; Romero et al., 2008). Therefore,

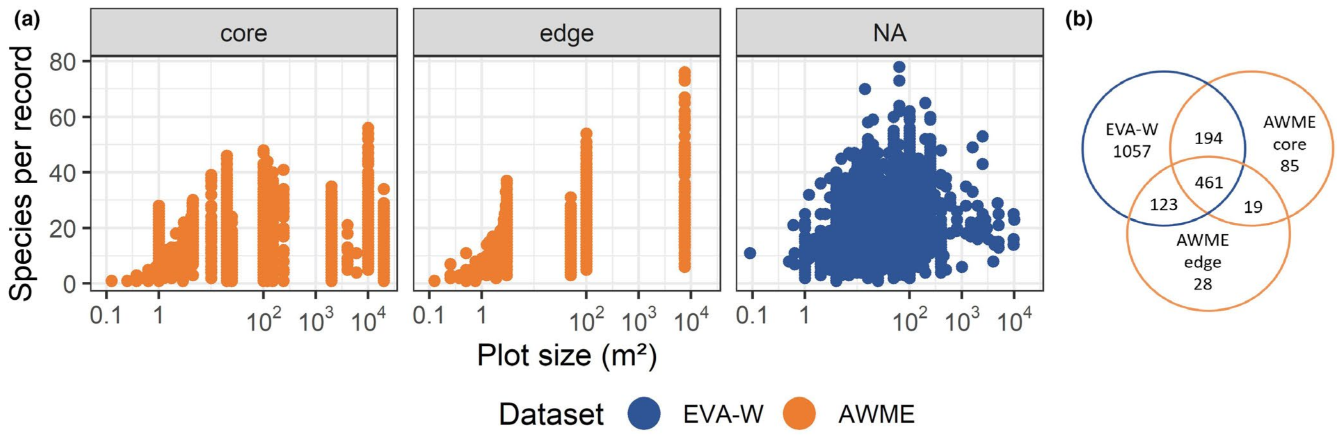

FIGURE 4 Comparison of data sets by surveyed field part. (a) Species numbers per record in relation to plot size. (b) Shared and exclusive species. AWME records taken on grass margins ($n = 30$) were left out of this figure for better readability. AWME, Arable Weed and Management in Europe; EVA-W, European Vegetation Archive subset on Weeds; NA, information not available

TABLE 3 Species number per record in crop classes of two data collections

	EVA-W				AWME			
Crop	Median	Mean	SD	*n*	Median	Mean	SD	*n*
Beet	18	18.9	8.63	356	7	8.38	5.73	906
Cereal	19	20.1	9.92	7225	5	7.41	6.68	13 087
Maize	9	12.5	9.77	1413	8	8.69	5.36	9040
Oilseed rape	14	15.1	6.79	705	9	9.55	5.43	4068
Potato	22	21.9	7.43	1066	8	10.7	8.96	197
Sunflower	12	14.3	7.07	82	9	10.2	6.13	1310

Note: Crop has a significant effect on species number (EVA-W $\chi^2 = 1056.5$, $df = 5$, $p < 0.001$, AWME $\chi^2 = 1479$, $df = 5$, $p < 0.001$).

Abbreviations: AWME, Arable Weed and Management in Europe; EVA-W, European Vegetation Archive subset on Weeds; *n*, number of records; SD, standard deviation.

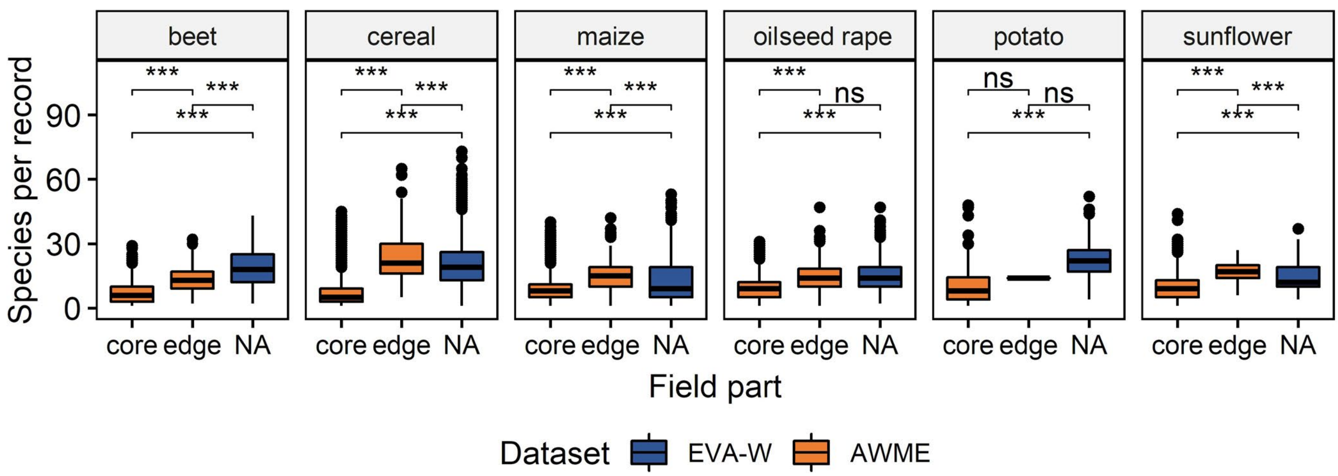

FIGURE 5 Species number per record in two data collections and six crops. Pairwise comparison of means with Wilcoxon test; number of records varies strongly in the subgroups. ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$; ns, not significant. NA, information not available. See Appendix S3 for case numbers. AWME, Arable Weed and Management in Europe; EVA-W, European Vegetation Archive subset on Weeds; NA, information not available

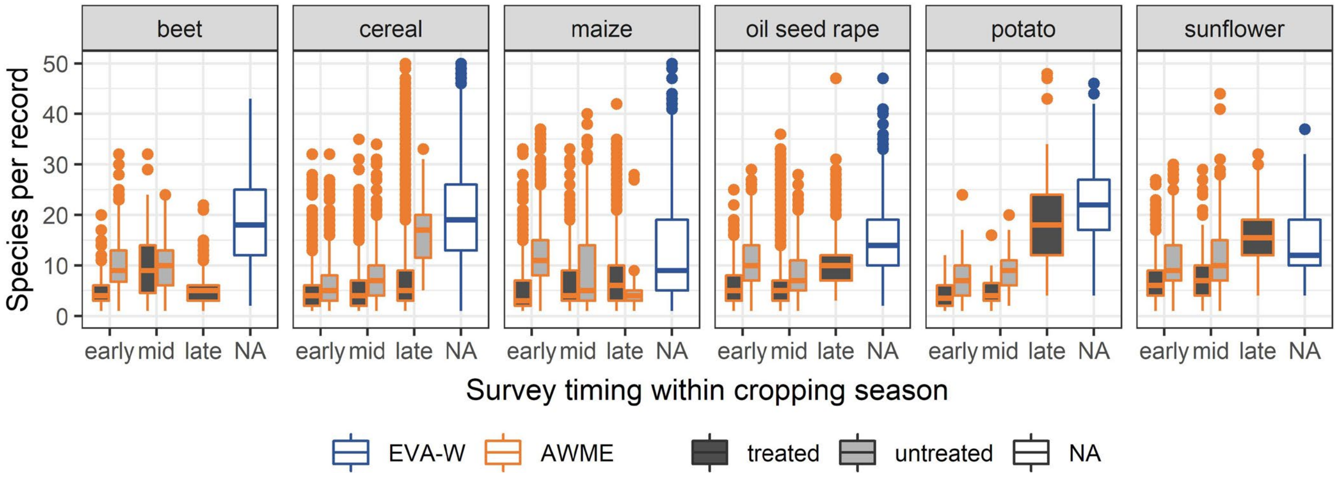

FIGURE 6 Species number per record in relation to survey time within cropping season. Classification relative to cropping season: early: short time after sowing; mid-season: row closure in row crops/heading of cereals; late season: full vegetation/height of summer/near or after harvest. AWME, Arable Weed and Management in Europe; EVA-W, European Vegetation Archive subset on Weeds; NA, information not available

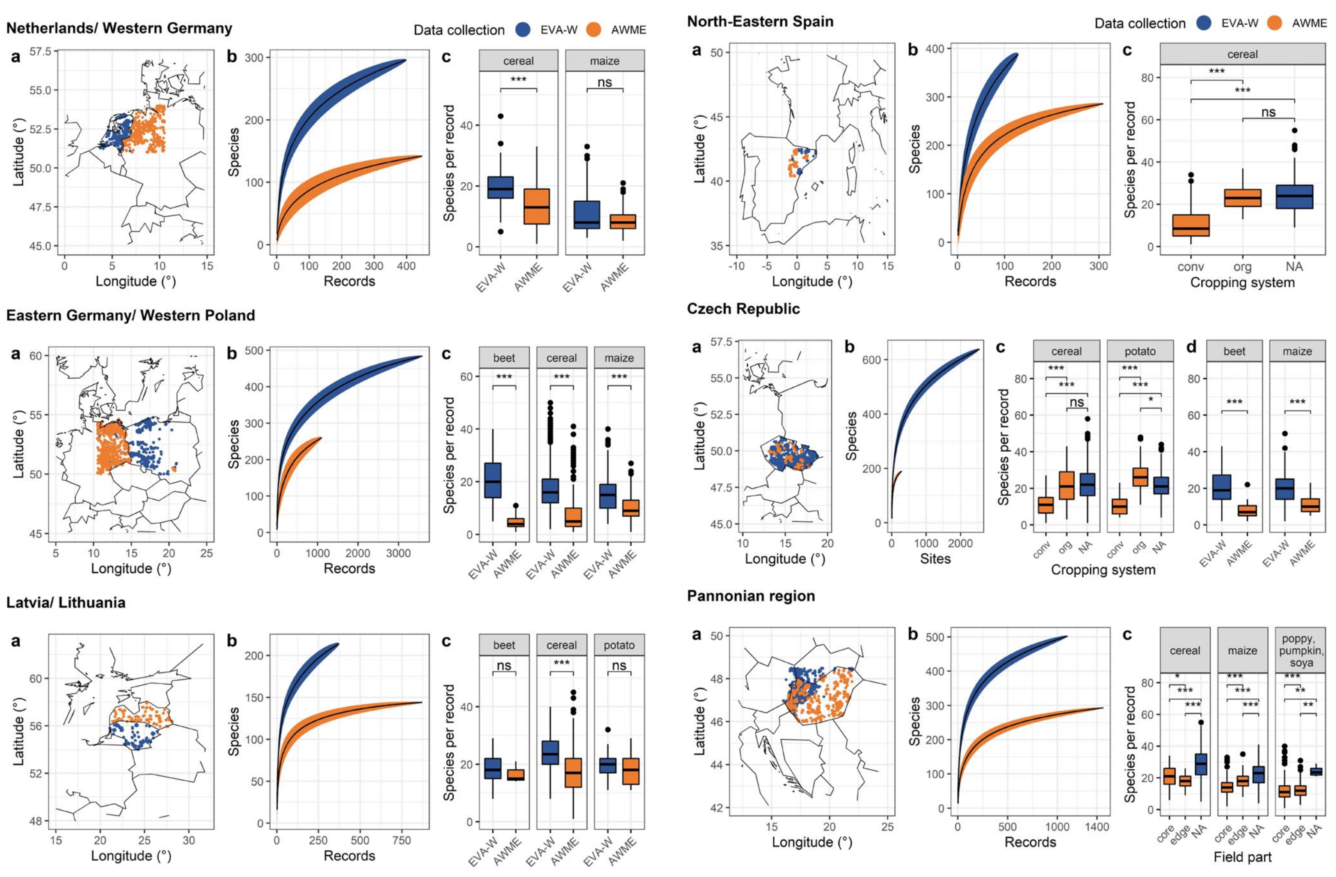

FIGURE 7 Comparing EVA-W and AWME records in six geographical subsets. (a) Plot locations. (b) Species accumulation curves, leaving out taxa recorded at genus level. (c, d) Species number per record for a subset of crops which are present in both data collections, counting taxa recorded at the genus level as a species. conv, conventional management; org, organic management; AWME, Arable Weed and Management in Europe; EVA-W, European Vegetation Archive subset on Weeds Weeds; NA, information not available

many weed science studies that focus on weed occurrence and competition explicitly keep a distance to the field boundary (Hanzlik & Gerowitt, 2016) or eliminate non-arable plants like woody saplings from the record (Cirujeda et al., 2011). Such non-arable species are likely the main reason for the higher species numbers in EVA-W compared to AWME. This probably affects the overall species number even more than the species number per record.

The time periods covered by the two collections may have influenced the proportion of samples from extensively managed fields. EVA-W plots were sampled from 1980 to 2013, AWME plots mainly in the decade between 2000 and 2010 (Appendix S4). There is a 16-year difference, which covers a period of intensive changes in agricultural practice all over Europe, with a strong decline in extensively managed fields. This time period includes the end of political divide between Western and Eastern European countries, followed by the enlargement of the European Union with many countries afterwards influenced by the Common Agricultural Policy. Analyses of how agricultural intensification has influenced arable weed vegetation in general have been published by various authors (Batáry et al., 2017; Richner et al., 2015). However, a detailed and comparative analysis of how different measures (e.g., large fields and mechanisation in collective farming of the East, higher inputs of fertiliser and herbicides in the West, segregation between plant and animal production, giving up marginal land when it is not profitable under CAP) affected arable weed vegetation and the loss of extensively managed arable land throughout the decades after 1990 and in various countries still needs to be undertaken on the continental scale (Sutcliffe et al., 2015).

Munoz et al. (2020) discuss the effect of regional species pools on the number of species found in a survey. Whilst many weed surveys are carried out in the main areas of intensive agriculture in Central Europe, a large part of the arable weed diversity is located in the Mediterranean region or in traditional livestock farming areas. The authors found this effect in their study of weed species in France. It was also discussed for Spain by Cirujeda et al. (2011). It most probably also affects our comparison of the whole data sets given the many records of EVA-W taken in Greece and Crete and a large number of AWME records in France and Northern Germany. However, this does not apply to the comparisons of subsets, except for the regional subset from Spain, where the EVA-W records originate from the Pyrenees, which is a marginal arable production area (Baldock et al., 1994).

Survey season may be another methodological cause for differing average species numbers per record. Usually, vegetation surveys are recommended to be taken at the peak of vegetation, i.e., when plants are well developed, have all organs, are flowering or setting seed and can therefore be more easily identified. Some of the studies included in AWME sampled the weed vegetation inside fields much earlier in the season, soon after crop and weed emergence and/or prior to herbicide treatment (Fried et al., 2008; Heard et al., 2003; Redwitz & Gerowitt, 2018). Under the early-season field conditions with sparse soil cover and low competition, rarer species may be more easily spotted, but species with later germination can be missing and some plants cannot be identified at this early stage and are therefore recorded at the genus level.

The last methodological aspect we want to discuss are exclosures, which are sometimes used in weed science studies to get a better picture of potential weed vegetation with no weed control. Records from such untreated plots generally have higher species numbers than records from treated plots in the same crop. The untreated plots may give useful insights into the effect of herbicide treatment or weed competition on yield. The additional species are often associated with other crops of the rotation, for example typical weeds from oilseed rape occurring in a maize crop (de Mol et al., 2015). Species composition and richness of these plots are similar to those of conservation headlands, a species conservation measure where field edges are cropped, but fertiliser and herbicide input is restricted (Albrecht et al., 2016). These plots could also be used as a bridge for comparing the treated core plots, plots from edge and margins and the phytosociological records.

4.2 | Joint use of the data collections

Despite the notable differences, it is clear that both data collections were recorded from the same landscape element: arable fields, although mostly from different sections. While weed scientists with an agronomic background tend to study weed vegetation in the field interior, contemporary phytosociologists or scientists interested in the status of rare and endangered segetal species survey field edges or even the margins outside of fields.

Since the two data collections shed light on different parts of arable fields, they are like two sides of the same medal. Depending on the specific research question, it will be necessary to combine the information gathered from both collections together to obtain a more accurate picture of this habitat. Additionally, future monitoring activities on arable plant diversity should include the different parts of arable fields. Arable land use has always had a history of spatial differences in disturbance and management (Jansen et al., 2009). Analysis of long-term trends in change of the European arable flora will then be facilitated by integrating all available data sources. Similar to the data sources made available by the efforts for the *European Weed Vegetation Database* (Küzmič et al., 2020), many more studies exist in the weed science that can be included in the AWME collection, including resurvey studies, surveys from before 1996 and those from further European countries, especially from Scandinavia.

The contrast in species richness at the field scale described in this study should be extended further to the landscape scale. It is known that more heterogeneous agricultural landscapes harbour higher species diversity than homogeneous agricultural landscapes with large field units and a low crop diversity (Benton et al., 2003; Dostatny, 2013; Gabriel et al., 2005; Storkey et al., 2012). We can now offer a continental-scale data set to study the interdependence between landscape heterogeneity and vegetation diversity inside and at the edges of arable fields.

The two data collections could also find a joint use on the species level. Filling geographical gaps in one collection with data from the other collection will be useful for research into geographical species distributions either for obtaining more accurate species distribution models of agronomically important weeds (Bürger et al., 2018) or for asking macroecological questions like whether species are more abundant in the centre of their distribution. The main benefit of using the EVA-W and AWME collections in this context is the confirmed provenance of the data from agricultural fields, rather than ruderal and other disturbed habitats.

As for any data-driven analysis, an appropriate selection procedure is needed when using EVA-W and AWME data together. Tailored to the research question, consideration must be given to the factors discussed in subsection 4.1. The analysis must especially deal with issues of the uneven distribution of records from different crops and different sampling density.

From the experience we share in this study, we would like to stress the importance of recording, presenting and digitising all available plot information regardless of the study purpose at that time. There are probably many more records from arable fields and margins stored within the EVA. However, without an explicit marker, it was not possible to extract them. Moreover, records of arable weed vegetation should always include information on the plot location relative to the field boundary, on the crop and whenever possible on the management practise.

AUTHOR CONTRIBUTIONS

Jana Bürger, Filip Küzmič, Urban Šilc and Florian Jansen conceived the ideas, designed the analysis, interpreted the results and led the writing of the manuscript; Jana Bürger and Filip Küzmič collected and curated the data, Jana Bürger performed the analysis. All authors contributed data and critically reviewed the manuscript drafts.

DATA AVAILABILITY STATEMENT

Data of the EVA-W data set are available on request to the European Vegetation Archive. AWME data are available on request from the data custodian (helen.metcalfe@rothamsted.ac.uk).

ORCID

Jana Bürger https://orcid.org/0000-0003-3898-6664
Filip Küzmič https://orcid.org/0000-0002-3894-7115
Urban Šilc https://orcid.org/0000-0002-3052-699X
Florian Jansen https://orcid.org/0000-0002-0331-5185
Erwin Bergmeier https://orcid.org/0000-0002-6118-4611
Milan Chytrý https://orcid.org/0000-0002-8122-3075
Alicia Cirujeda https://orcid.org/0000-0001-9646-8422
Silvia Fogliatto https://orcid.org/0000-0002-7309-8550
Guillaume Fried https://orcid.org/0000-0002-3653-195X
Denise F. Dostatny https://orcid.org/0000-0002-6544-1531
Bärbel Gerowitt https://orcid.org/0000-0001-9435-3453
Michael Glemnitz https://orcid.org/0000-0002-6506-1889
José L. González-Andújar https://orcid.org/0000-0003-2356-4098
Eva Hernández Plaza https://orcid.org/0000-0002-2378-2602
Jordi Izquierdo https://orcid.org/0000-0002-3949-1295
Michaela Kolářová https://orcid.org/0000-0002-9817-1195
Zdeňka Lososová https://orcid.org/0000-0001-9152-7462
Helen Metcalfe https://orcid.org/0000-0002-2862-0266
Jevgenija Ņečajeva https://orcid.org/0000-0002-0828-9721
Sandrine Petit https://orcid.org/0000-0001-8781-8873
Gyula Pinke https://orcid.org/0000-0002-9956-1363
Valerijus Rašomavičius https://orcid.org/0000-0003-1314-4356
Christoph von Redwitz https://orcid.org/0000-0002-1591-5081
Matthias Schumacher https://orcid.org/0000-0002-5256-1673
Lena Ulber https://orcid.org/0000-0003-2829-1527
Francesco Vidotto https://orcid.org/0000-0002-0971-1445

REFERENCES

Albrecht, H., Cambecèdes, J., Lang, M. & Wagner, M. (2016) Management options for the conservation of rare arable plants in Europe. *Botany Letters*, 163(4), 389–415. https://doi.org/10.1080/23818107.2016.1237886

Andersson, T.N. & Milberg, P. (1998) Weed flora and the relative importance of site, crop, crop rotation and nitrogen. *Weed Science*, 46, 30–38. https://doi.org/10.1017/S0043174500090135

Baldock, D., Beaufoy, G. & Clark, J. (1994) *The nature of farming: Low intensity farming systems in nine European countries*. London: Institute for European Environmental Policy.

Barralis, G. (1976). *Méthode d'étude des groupements adventices des cultures annuelles. 5éme Colloque International sur l'Ecologie et la Biologie des Mauvaises Herbes*. Dijon.

Batáry, P., Gallé, R., Riesch, F., Fischer, C., Dormann, C.F., Mußhoff, O. et al. (2017) The former Iron Curtain still drives biodiversity-profit trade-offs in German agriculture. *Nature Ecology & Evolution*, 1(9), 1279–1284. https://doi.org/10.1038/s41559-017-0272-x

Benton, T.G., Vickery, J.A. & Wilson, J.D. (2003) Farmland biodiversity: is habitat heterogeneity the key? *Trends in Ecology & Evolution*, 18(4), 182–188. https://doi.org/10.1016/S0169-5347(03)00011-9

Braun-Blanquet, J. (1964) *Pflanzensoziologie: Grundzüge der Vegetationskunde*. Wien: Springer.

Bruelheide, H., Dengler, J., Jiménez-Alfaro, B., Purschke, O., Hennekens, S.M., Chytrý, M. et al. (2019) sPlot – A new tool for global vegetation analyses. *Journal of Vegetation Science*, 30(2), 161–186. https://doi.org/10.1111/jvs.12710

Bürger, J., Metcalfe, H., von Redwitz, C., Cirujeda, A., Fogliatto, S., Fried, G. et al. (2020) Arable weeds and management in Europe. *Vegetation Classification and Survey*, 1, 169–170. https://doi.org/10.3897/VCS/2020/61419

Bürger, J., de Mol, F., Fried, G., González-Andújar, J.L., Hernández Plaza, E., Ņečajeva, J. et al. (2018) *Merging European arable weed survey datasets - Case study: Increase accuracy of Species Distribution Models*. Book of Abstracts 18th European Weed Research Society Symposium Lubljana, 244.

Chytrý, M., Hennekens, S.M., Jiménez-Alfaro, B., Knollová, I., Dengler, J., Jansen, F. et al. (2016) European Vegetation Archive (EVA): an integrated database of European vegetation plots. *Applied Vegetation Science*, 19(1), 173–180. https://doi.org/10.1111/avsc.12191

Chytrý, M., Tichý, L., Hennekens, S.M. & Schaminée, J.H.J. (2014) Assessing vegetation change using vegetation-plot databases: a risky business. *Applied Vegetation Science*, 17(1), 32–41. https://doi.org/10.1111/avsc.12050

Cirujeda, A., Aibar, J. & Zaragoza, C. (2011) Remarkable changes of weed species in Spanish cereal fields from 1976 to 2007. *Agronomy for Sustainable Development*, 31(4), 675–688. https://doi.org/10.1007/s13593-011-0030-4

De Mol, F., von Redwitz, C. & Gerowitt, B. (2015) Weed species composition of maize fields in Germany is influenced by site and crop sequence. *Weed Research*, 55(6), 574–585. https://doi.org/10.1111/wre.12169

Dengler, J., Jansen, F., Glöckler, F., Peet, R.K., De Cáceres, M., Chytrý, M. et al. (2011) The Global Index of Vegetation-Plot Databases (GIVD): a new resource for vegetation science. *Journal of Vegetation Science*, 22(4), 582–597. https://doi.org/10.1111/j.1654-1103.2011.01265.x

Dostatny, D.F. (2013) The function of small farms in supporting biological diversity of agricultural ecosystems. *Roczniki Naukowe Ekonomii Rolnictwa i Rozwoju Obszarów Wiejskich*, 100(4), 34–42.

Du Rietz, G.E. (1936) Classification and nomenclature of vegetation units 1930–1935. *Svensk Botanisk Tidskrift*, 30(3), 580–589.

Ehrendorfer, F. (1967) *Liste der Gefäßpflanzen Mitteleuropas*, 1st edition. Wien: Notring der wissenschaftlichen Verbände Österreichs.

Eichenberg, D., Bowler, D.E., Bonn, A., Bruelheide, H., Grescho, V., Harter, D. et al. (2021) Widespread decline in Central European plant diversity across six decades. *Global Change Biology*, 27(5), 1097–1110. https://doi.org/10.1111/gcb.15447

Euro+Med (2021) *Euro+Med PlantBase - the information resource for Euro-Mediterranean plant diversity*. Retrieved from http://ww2.bgbm.org/EuroPlusMed/

Fried, G., Norton, L.R. & Reboud, X. (2008) Environmental and management factors determining weed species composition and diversity in France. *Agriculture, Ecosystems & Environment*, 128(1–2), 68–76. https://doi.org/10.1016/j.agee.2008.05.003

Fried, G., Petit, S., Dessaint, F. & Reboud, X. (2009) Arable weed decline in Northern France: Crop edges as refugia for weed conservation? *Biological Conservation*, 142(1), 238–243. https://doi.org/10.1016/j.biocon.2008.09.029

Gabriel, D., Thies, C. & Tscharntke, T. (2005) Local diversity of arable weeds increases with landscape complexity. *Perspectives in Plant Ecology, Evolution and Systematics*, 7(2), 85–93. https://doi.org/10.1016/j.ppees.2005.04.001

GBIF: The Global Biodiversity Information Facility (2020) *What is GBIF?* Retrieved from https://www.gbif.org/what-is-gbif

Gotelli, N.J. & Colwell, R.K. (2001) Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecology Letters*, 4(4), 379–391. https://doi.org/10.1046/j.1461-0248.2001.00230.x

Hanzlik, K. & Gerowitt, B. (2016) Methods to conduct and analyse weed surveys in arable farming: a review. *Agronomy for Sustainable Development*, 36, 11. https://doi.org/10.1007/s13593-015-0345-7

Heard, M.S., Hawes, C., Champion, G.T., Clark, S.J., Firbank, L.G., Haughton, A.J. et al. (2003) Weeds in fields with contrasting conventional and genetically modified herbicide–tolerant crops. I. Effects on abundance and diversity. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 358(1439), 1819–1832. https://doi.org/10.1098/rstb.2003.1402

Hüppe, J. & Hofmeister, H. (1990) Syntaxonomische Fassung und Übersicht über die Ackerunkrautgesellschaften der Bundesrepublik Deutschland. *Berichte der Reinhold-Tüxen-Gesellschaft*, 2, 61–81.

Hurford, C., Wilson, P. & Storkey, J. (Eds) (2020) *The Changing Status of Arable Habitats in Europe: A Nature Conservation Review* (1st edition 2020). Cham: Springer International Publishing.

Hyvönen, T., Ketoja, E., Salonen, J., Jalli, H. & Tiainen, J. (2003) Weed species diversity and community composition in organic and conventional cropping of spring cereals. *Agriculture, Ecosystems & Environment*, 97(1–3), 131–149. https://doi.org/10.1016/S0167-8809(03)00117-8

Jansen, F. & Dengler, J. (2010) Plant names in vegetation databases - a neglected source of bias. *Journal of Vegetation Science*, 21(6), 1179–1186. https://doi.org/10.1111/j.1654-1103.2010.01209.x

Jansen, F., Zerbe, S. & Succow, M. (2009) Changes in landscape naturalness derived from a historical land register - A case study from NE Germany. *Landscape Ecology*, 24(2), 185–198. https://doi.org/10.1007/s10980-008-9297-5

Jeliazkov, A., Mijatovic, D., Chantepie, S., Andrew, N., Arlettaz, R., Barbaro, L. et al. (2020) A global database for metacommunity ecology, integrating species, traits, environment and space. *Scientific Data*, 7(1), 6. https://doi.org/10.1038/s41597-019-0344-7

Kattge, J., Bönisch, G., Díaz, S., Lavorel, S., Prentice, I.C., Leadley, P. et al. (2020) TRY plant trait database - enhanced coverage and open access. *Global Change Biology*, 26(1), 119–188. https://doi.org/10.1111/gcb.14904

Kühl, H.S., Bowler, D.E., Bösch, L., Bruelheide, H., Dauber, J., Eichenberg, D., Fernández, N. et al. (2020) Effective biodiversity monitoring needs a culture of integration. *One Earth*, 3(4), 462–474. https://doi.org/10.1016/j.oneear.2020.09.010

Küzmič, F., Šilc, U., Lososová, Z., Mucina, L., Chytrý, M., Knollová, I. et al. (2020) European Weed Vegetation Database – a gap-focused vegetation-plot database. *Phytocoenologia*, 50(1), 93–100. https://doi.org/10.1127/phyto/2019/0337

Leuschner, C. & Ellenberg, H. (2017). Vegetation of Arable Fields, Gardens and Vineyards. In Leuschner, C. & Ellenberg, H. (Eds.), *Vegetation ecology of Central Europe. Volume 2. Ecology of Central European non-forest vegetation: Coastal to alpine, natural to man-made habitats* (1st edition, pp. 779–839). Cham: Springer International. https://doi.org/10.1007/978-3-319-43048-5_12

Londo, G. (1976) The decimal scale for releves of permanent quadrats. *Vegetation*, 33, 61–64.

Manthey, M. (2003). *Vegetationsökologie der Äcker und Ackerbrachen Mecklenburg-Vorpommerns*. Dissertationes Botanicae: Vol. 373. Berlin, Stuttgart. Gebrüder Bornträger.

Martinčič, A., Wraber, T., Jogan, N., Podobnik, A., Ravnik, V., Turk, B., et al. (2007) *Mala flora Slovenije: ključ za določanje praprotnic in semenk*, 4th edition. Lubljana: Tehniška založba Slovenije.

Meisel, K. (1979) Veränderungen der Segetalvegetation in der Stolzenauer Wesermarsch seit 1945. *Phytocoenologia*, 6, 118–130. https://doi.org/10.1127/phyto/6/1979/118

Metcalfe, H., Hassall, K.L., Boinot, S. & Storkey, J. (2019) The contribution of spatial mass effects to plant diversity in arable fields. *Journal of Applied Ecology*, 56(7), 1560–1574. https://doi.org/10.1111/1365-2664.13414

Meyer, S., Bergmeier, E., Becker, T., Wesche, K., Krause, B. & Leuschner, C. (2015) Detecting long-term losses at the plant community level - arable fields in Germany revisited. *Applied Vegetation Science*, 18(3), 432–442. https://doi.org/10.1111/avsc.12168

Moore, J.J., Fitzsimons, S.J.P., Lambe, E. & White, J. (1970) A comparison and evaluation of some phytosociological techniques. *Vegetatio*, 20(1–4), 1–20. https://doi.org/10.1007/BF01965526

Mucina, L., Bültmann, H., Dierßen, K., Theurillat, J.-P., Raus, T., Čarni, A. et al. (2016) Vegetation of Europe: hierarchical floristic classification system of vascular plant, bryophyte, lichen and algal communities. *Applied Vegetation Science*, 19(Suppl. 1), 3–264. https://doi.org/10.1111/avsc.12257

Mueller-Dombois, D. & Ellenberg, H. (1974) *Aims and methods of vegetation ecology*. New York: Jon Wiley & Sons.

Munoz, F., Fried, G., Armengot, L., Bourgeois, B., Bretagnolle, V., Chadoeuf, J. et al. (2020) Ecological specialization and rarity of arable weeds: insights from a comprehensive survey in France. *Plants*, 9(7), 824. https://doi.org/10.3390/plants9070824

Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P. & McGlinn, D., et al. (2020). *vegan: Community Ecology Package (Version 2.5-7) [Computer software]*. Retrieved from https://CRAN.R-project.org/package=vegan

Pinke, G. (2000) Die Ackerwildkraut-Gesellschaften extensiv bewirtschafteter Felder in der Kleinen Ungarischen Tiefebene. *Tuexenia*, 20, 335–364.

R Core Team (2021) *R: A language and environment for statistical computing (Version 4.0.4)*. Vienna: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Rasiņš, A. & Tauriņa, M. (1982) *Nezāļu kvantitātes uzskaites metodika Latvijas PSR apstākļos*. Riga.

von Redwitz, C. & Gerowitt, B. (2018) Maize-dominated crop sequences in northern Germany: Reaction of the weed species communities. *Applied Vegetation Science*, 21(3), 431–441. https://doi.org/10.1111/avsc.12384

Richner, N., Holderegger, R., Linder, H.P. & Walter, T. (2015) Reviewing change in the arable flora of Europe: a meta-analysis. *Weed Research*, 55(1), 1–13. https://doi.org/10.1111/wre.12123

Romero, A., Chamorro, L., Sans, F.X. (2008) Weed diversity in crop edges and inner fields of organic and conventional dryland winter cereal crops in NE Spain. *Agriculture, Ecosystems & Environment*, 124(1–2), 97–104. https://doi.org/10.1016/j.agee.2007.08.002

Salonen, J. (1993) Weed infestation and factors affecting weed incidence in spring cereals in Finland - a multivariate approach. *Agricultural and Food Science in Finland*, 2(6), 525–536. https://doi.org/10.23986/afsci.72678

Storkey, J., Meyer, S., Still, K.S. & Leuschner, C. (2012) The impact of agricultural intensification and land-use change on the European arable flora. *Proceedings of the Royal Society B: Biological Sciences*, 279(1732), 1421–1429. https://doi.org/10.1098/rspb.2011.1686

Storkey, J. & Westbury, D.B. (2007) Managing arable weeds for biodiversity. *Pest Management Science*, 63(6), 517–523. https://doi.org/10.1002/ps.1375

Sutcliffe, L.M.E., Batáry, P., Kormann, U., Báldi, A., Dicks, L.V., Herzon, I. et al. (2015) Harnessing the biodiversity value of Central and Eastern European farmland. *Diversity and Distributions*, 21(6), 722–730. https://doi.org/10.1111/ddi.12288

Ugland, K.I., Gray, J.S. & Ellingsen, K.E. (2003) The species-accumulation curve and estimation of species richness. *Journal of Animal Ecology*, 72(5), 888–897. https://doi.org/10.1046/j.1365-2656.2003.00748.x

Van Elsen, T. (1989) Ackerwildkraut-Gesellschaften herbizidfreier Ackerränder und des herbizidbehandelten Bestandesinnern im Vergleich. *Tuexenia*, 9, 75–105.

Wietzke, A., Leuschner, C. (2020) Surveying the arable plant diversity of conventionally managed farmland: A comparison of methods. *Environmental Monitoring and Assessment*, 192(2), 98. https://doi.org/10.1007/s10661-019-8042-7

Winqvist, C., Bengtsson, J., Aavik, T., Berendse, F., Clement, L.W., Eggers, S. et al. (2011) Mixed effects of organic farming and landscape complexity on farmland biodiversity and biological control potential across Europe. *Journal of Applied Ecology*, 48(3), 570–579. https://doi.org/10.1111/j.1365-2664.2010.01950.x

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

Appendix S1. Details of data sets included in the EVA-W arable weed vegetation subset of the European Vegetation Archive

Appendix S2. Details of data sets included in Arable Weeds and Management in Europe data collection (as of 2021)

Appendix S3. Record numbers presented in Figure 5 of the main paper.

Appendix S4. Distribution of records over time in two data collections.

How to cite this article: Bürger, J., Küzmič, F., Šilc, U., Jansen, F., Bergmeier, E., Chytrý, M., et al (2022) Two sides of one medal: Arable weed vegetation of Europe in phytosociological data compared to agronomical weed surveys. *Applied Vegetation Science*, 25, e12460. https://doi.org/10.1111/avsc.12460