METADATA

CLASS I. DATA SET DESCRIPTORS

A. Data set identity

Farmland biodiversity and agricultural management on 237 farms in 13 European and two African regions

B. Data set identification code

- Bee_data.txt
- Dictionary.txt
- Earthworm_data.txt
- Habitat_data.txt
- HabitatType_Explanation.txt
- Management_data.txt
- Region_Description.txt
- Spider_data.txt
- Surrounded_landscape_data.txt
- Vascular_plant_data.txt

C. Data set description

Farmland is a major land cover type in Europe and Africa and provides habitat for numerous species. The severe decline in farmland biodiversity of the last decades has been attributed to changes in farming practices, and organic and low-input farming are assumed to mitigate detrimental effects of agricultural intensification on biodiversity. Since the farm enterprise is the primary unit of agricultural decision making, management-related effects at the field scale need to be assessed at the farm level. Therefore, in this study, data were collected on habitat
characteristics, vascular plant, earthworm, spider, and bee communities and on the corresponding agricultural management in 237 farms in 13 European and two African regions. In 15 environmental and agricultural homogeneous regions, 6–20 farms with the same farm type (e.g., arable crops, grassland, or specific permanent crops) were selected. If available, an equal number of organic and non-organic farms were randomly selected. Alternatively, farms were sampled along a gradient of management intensity. For all selected farms, the entire farmed area was mapped, which resulted in total in the mapping of 11 338 units attributed to 194 standardized habitat types, provided together with additional descriptors. On each farm, one site per available habitat type was randomly selected for species diversity investigations. Species were sampled on 2115 sites and identified to the species level by expert taxonomists. Species lists and abundance estimates are provided for each site and sampling date (one date for plants and earthworms, three dates for spiders and bees). In addition, farmers provided information about their management practices in face-to-face interviews following a standardized questionnaire. Farm management indicators for each farm are available (e.g., nitrogen input, pesticide applications, or energy input). Analyses revealed a positive effect of unproductive areas and a negative effect of intensive management on biodiversity. Communities of the four taxonomic groups strongly differed in their response to habitat characteristics, agricultural management, and regional circumstances. The data has potential for further insights into interactions of farmland biodiversity and agricultural management at site, farm, and regional scale.

D. Key words

agricultural management; arable crop; bee; BioBio; earthworm; grassland; habitat diversity; permanent crop; spider; Tunisia; Uganda; vascular plant
CLASS II. RESEARCH ORIGIN DESCRIPTION

A. Overall project description

A.1. Identity

The sampling of habitat, species and agricultural management data was part of the EU research project BioBio (http://www.biobio-indicator.org/).

A.2. Originators

The project was coordinated by Felix Herzog at Agroscope, Institute for Sustainability Sciences ISS, in Zurich, Switzerland, felix.herzog@agroscope.admin.ch. A consortium of 16 partners from 15 countries developed and conducted the project.

A.3. Period of study

The published data was collected during the growing season either in 2010 or in 2011 depending on the region.

A.4. Objectives

The main objective of the EU project BioBio (Biodiversity indicators for organic and low-input farming systems) was to identify a set of indicators which are scientifically sound, generic at the European scale, and relevant and useful for stakeholders to assess biodiversity in farmland. The partners from Tunisia, Uganda, and Ukraine acted as representatives for International Corporation Partner Countries to test the usability of the indicator set beyond the borders of the EU, Norway, and Switzerland.
A.5. Research motivation

Intensification of farming has caused a critical decline of biodiversity in farmland (Robinson and Sutherland 2002). Efforts are being made to halt this loss of biodiversity at national and EU-scale, such as organic or low-input farming or set-aside areas (Kleijn et al. 2006, Whittingham et al. 2007, Kampmann et al. 2012, Tuck et al. 2014). In order to evaluate the effectiveness of such efforts and to understand the underlying mechanisms, any assessment method needs to adequately capture the complexity of farmland and its biodiversity (Büchs 2003). The unit at which assessments take place is crucial for the relevance of its results. Individual farm enterprises are the entity for which landmark decisions are made and should therefore be addressed primarily, and this was taken as the unit of reference within the BioBio project (Herzog et al. 2013, Siebert et al. 2006, Schneider et al. 2014).

Based on a literature review, we evaluated relevant levels and aspects of biodiversity on farms. We recorded the habitat characteristics of the farmland on 237 farms in Europe and Africa. For the assessment of species diversity, we selected a set of target taxonomic groups which are affected by agricultural management, are not redundant, cover a wide range of trophic levels, and fulfil distinct ecological functions. These were vascular plants, earthworms, spiders, and bees (Marc et al. 1999, Matzdorf et al. 2008, Paoletti 1999, Sauberer et al. 2004, Tscharntke et al. 1998), which were sampled on representative subsamples of the agricultural habitats on each farm. In order to investigate the sensitivity of habitats and species to agricultural management, we also collected information on agricultural management. This information is provided at the farm scale. To investigate effects of the surrounding landscape, land cover of ten habitat classes in a 250 m buffer around the sampled sites was estimated for eight regions. The dataset is a one-year biodiversity sampling, and this may be not sufficient to capture effects and processes acting on species abundances, in particular for arthropods.
Recent analyses of these data revealed significant positive effect of organic farming on species richness and abundance, but to a much lower extent at farm scale than at site scale (Nascimbene et al. 2012, Schneider et al. 2014). Further, the effects were taxa specific, being more distinct for plants and bees than for earthworms and spiders (Schneider et al. 2014). Major influences on all four taxa had regional characteristics such as e.g., the duration of the vegetation period, the soil conditions, or the general management intensity (Lüscher et al. 2015). Nevertheless, negative effects on species richness and abundance of high mineral fertilization and pesticide applications and positive effects of various unproductive habitats, were common (Lüscher et al. 2014a, Moreno et al. 2015). Moreover, the taxonomic groups responded individually to surrounding landscape, agricultural management, and the scale (site, farm) of analysis (Kovács-Hostyánszki et al. 2013). Hence, the four taxonomic groups were applicable and meaningful in different biogeographic regions and farm types (Herzog et al. 2012). A cost estimation and up-scaling of this work contribute to a successful implementation of biodiversity monitoring at the European scale (Geijzendorffer et al. 2016, Targetti et al. 2014, Targetti et al. 2015).

We hope that further analyses of this data set will enhance knowledge on community patterns in these four contrasting taxonomic groups and that it will usefully complement other available data sets. In our perspective, this data provides an example for a broad-scale biodiversity survey with the potential to spread out in space and time.

A.6. Sources of funding

Part of this work was funded by the European Union (project BioBio: KBBE-227161, 2009 – 2012), with further funding provided by the Austrian Ministry for Science and Research and by the Lendület program of the Hungarian Academy of Sciences.
B. Specific subproject description

B.1. Site description

B.1.a. Farm selection

In BioBio, 12 regions were selected to reflect the variability of biogeographic regions and the major farm types across Europe. In Ukraine and two additional regions in Africa, the BioBio indicators were tested for their applicability beyond the borders of the EU, Norway, and Switzerland (Fig. 1). The regions chosen had relatively homogeneous environmental conditions such as topography, soil characteristics, temperature, and precipitation. In each region, we randomly selected an equal number of organic and non-organic farms where possible. Due to limited availability of certified farms or willingness of farmers to participate, the number of farms was unbalanced in the regions CH (10 organic, 9 non-organic), HU (7 organic, 11 non-organic) and NL (11 organic, 3 non-organic). In two regions (GB and NO), farms were selected in pairs because of their location along a geographical and intensity gradient. Farms that were classified as organic had to have been continuously managed according to the standards of organic farming for at least five years. For non-organic farms, no additional constraints were required. If no farms existed that were certified for organic farming, up to 20 farms were picked along a management intensity gradient (BG and ED) or in two groups of low and high management intensity (UA). All selected farms within a region belonged to the same farm type, e.g., arable crops, grassland, mixed farming, or permanent crops. They were representative for a specific combination of region and farm type but not implicitly for all farms in the region (Schneider et al. 2014). For the region IT, a vineyard region, farms had to be selected from three production areas that are separated from each other, because the number of organic farms in one single area was too small.
Fig. 1: Location of study regions. IDs for the regions are explained in Region_Description.txt. The region UG is located in Uganda and not shown on the map. Black circles indicate 12 regions with organic and non-organic farms; three grey circles indicate regions where farms along a management intensity gradient or in two groups of low and high management intensity were investigated.

B.1.b. Habitat mapping

Once a farm was selected and the farmer agreed to participate in the project, the entire farmed area was mapped based on field observation using a standard habitat mapping procedure (Bunce et al. 2008). We recorded all areal (at least 5 m wide and covering 400 m$^2$) and linear
habitats (at least 0.5 m wide and 30 m long). First, the habitats were delineated based on dominant Raunkiær plant life forms in combination with a standardized set of environmental and structural descriptors (humidity, acidity, nutrient supply, tree cover, and understory composition). Cultivated herbaceous crops were further separated into four types: summer-sown non-entomophilic annual, winter-sown non-entomophilic annuals, entomophilic/and or bee attracting annuals, and perennials. A complete list of the 194 resulting habitat types and their descriptions is provided in HabitatType_Explanation.txt. If available, additional information to the mapped habitats is provided: a code for humidity, a code for soil conditions, the occurrence of trees, the occurrence of bushes, and the main-covering species. As the habitat mapping took place before the species sampling, the species with highest cover might deviate from the species with the maximal cover in the vegetation sample. In the meantime, the dominance of species may have changed due to vascular plant phenology. In addition, the species with highest covers was assessed for the whole habitat, whereas the vegetation sample covered just a part of the area, where the relative occurrence of species may be different.

B.2. Sampling design

We applied a stratified sampling design. Per farm, one site per habitat type was selected randomly (Fig. 2). Therefore, the number of selected sites differed among farms, depending on the number of existing habitat types.
Fig. 2: Schematic representation of the sampling design for three exemplary farms. Farms are indicated by different shading. Each farm consists of different areal and linear habitat types, indicated by different fill patterns. Species symbols indicate the randomly selected sites where plant, earthworm, spider, and bee species were sampled (reprinted from Lüscher et al. 2014b, with permission from Elsevier).

B.2.a. Exception: Region UA

For the region UA, the Kiev region in Ukraine, the sampling was adapted to the large size of fields and farms (mean farm size in the region UA was 4,600 ha, mean farm size in all other regions was 68 ha). Six farms were selected that cultivated *Glycine max*, *Hordeum vulgare*, *Triticum aestivum*, *Zea mays*, and *Fagopyrum esculentum* or *Medicago sativa*. At least one grassland habitat and four types of linear habitats had to be present. In each selected farm of
the region UA, 10 habitat types were considered for the species sampling. To account for the large size of the habitats (median arable crop field size was 174 ha), species were sampled on three sites per habitat. Habitat mapping was completed for the entire farmed area, based on aerial photographs.

B.2.b. Spatial location

As part of the description of the study regions, we provide the central coordinates for each region. Due to an agreement with the farmers about privacy issues, data that will allow recognizing individual farms cannot be published.

B.3. Research methods

The field recording handbook was published by Dennis et al. (2012), it also comprises a critical ex-post evaluation of the methods used.

B.3.a. Species sampling

On all selected sites, vascular plant, earthworm, spider, and bee species were sampled using standardized protocols during one growing season, generally in 2010 (Fig. 3, Dennis et al. 2012). As an exception, in the region IT, 12 of the 90 investigated fields were sampled in 2011. In all regions in the International Corporation Partner Countries (TN, UA and UG), field work was conducted in 2011.
Vascular plant species

Vascular plant species were identified and their respective ground cover was estimated by experienced field botanists in one survey of 10 m × 10 m in the central part of the selected site. If the site was a cultivated herbaceous crop or hayfield, the survey was moved to the edge.
with at least 3 m distance from the border. In linear habitats, the plant survey measured 1 m × 10 m.

Earthworm species

Earthworms were sampled at three locations per selected site, which were at least 20 m apart from the border and 10 m apart from each other. In narrow areal and in linear habitats, the locations were close to the midline and 10 m apart along the line feature. A metal frame of 0.3 m × 0.3 m was placed on the humid soil. First, we poured a solution of allyl isothiocyanate (0.1 g/l) into the metal frame (two applications at a 5 min interval). This encouraged earthworms to move to the surface where they were collected. After 10 min, we excavated a 20 cm deep soil core of the whole frame, and hand sorted the soil for 20 min per core. The collected earthworms were stored in ethanol (80%) or formaldehyde (4%). In the data file, the earthworms are listed separately for the two methods, i.e. allyl isothiocyanate application and hand sorting. Expert taxonomists identified the specimen to the species level in the lab. Juveniles, i.e. individuals without a clitellum, were excluded from the data set. In the region TN, no earthworm samples were taken, because this taxonomic group was completely absent. In the region UG, seven morphospecies and *Eisenia fetida* were distinguished.

Spider species

Spiders were sampled from soil surface and vegetation at five locations per selected site. We used a modified shredder vac/blower (Stihl SH 86-D) to suck the spiders within a haphazardly placed ring of 0.357 m diameter for 30 s. The locations were at least 20 m apart from the border and at least 10 m apart from each other. In narrow areal and in linear habitats, the locations were close to the midline and 10 m apart along the line feature. In very short linear habitats, the locations were closer to each other. Spider sampling was conducted on three dates. The exact dates depended on the region, the vegetation cover, and the weather.
conditions (dry and warm). Collected spiders were inverted in a zip-seal bag, stored in a cool box, and deep-frozen in the lab. Then, we separated the spiders from additional material such as sand, small pieces of vegetation, or other invertebrates and stored them in ethanol (80%). Expert taxonomists identified each specimen to the species level. Juveniles were excluded from the data set. In the region UG, no spider samples were taken, because the spider taxonomy was not sufficiently stable.

Bee species

Bees were sampled in a transect walk with aerial netting for 15 min. The transect was 2 m × 100 m and crossed the middle of the vascular plant survey. If the habitat was shorter than 100 m, the transect length was divided, e.g., in two 50 m-long transects. Bees were sampled on three dates during good weather conditions, i.e. temperature higher than 15° C, sunny, and not too windy. The sampling dates depended on the region. Where identification was not possible in the field, captured bees were transferred to a killing jar, pinned, and identified by expert taxonomists in the lab. *Apis mellifera* individuals were not captured, but counted in the field. In the regions AT and NL, bees were sampled on only two dates. In the region UG, 132 morphospecies and *Apis sp.* were distinguished.

B.3.b. Details of species taxonomy

In general, species were identified to the species level. For the taxonomy of vascular plants, we referred to Flora Europaea (Tutin 2001). The species names of earthworms were harmonized using the Fauna Europaea (De Jong et al. 2014) and relying on Pérez-Losada et al. (2015). To check the species names of spiders, we consulted the world spider catalog (Platnick 2010). The species names of bees were harmonized using the Fauna Europaea.
B.4. Agricultural management data

Data on agricultural management were collected in face-to-face interviews with the farmers, following a standardized questionnaire, which had been translated in the corresponding local language. In a first part of the questionnaire, the farmers provided information at farm scale. They were asked about the farming system (organic or non-organic) and its duration, about the number and types of livestock units, about fuel and electricity consumption, and about expenditures on fertilizers, crop protection and concentrate feed stuff. Naturally, depending on the farm type (arable crops, grassland, mixed farming, or permanent crops) different aspects were more or less important. In a second part of the questionnaire, the farmers listed the variety of their cultivated crops and their area. Individually, for each crop under the same management, they provided information on the amount and type of nitrogen input, the number of pesticide applications, the mechanical field operations, and the livestock density. Organic farms did not use mineral nitrogen. Natural pesticides were included in the counts, and we distinguished between herbicides, fungicides, and insecticides. Pesticide applications were related to the cultivated crops, e.g., fungicides were mainly applied in the vineyards of the region IT. Mechanical field operations included soil cultivation, fertilization, pesticide application, mowing, turning, bale making, and loading. Here, as well, the type of field operation was related to the crop. The crop specific information was then scaled up to average values for the entire production area per farm with the indicator tool DIALECTE (Doublet 1994 – 2004). The duration of the interviews depended on the complexity of the farm and lasted from one to more than three hours. Management data at farm scale are only partly available for the regions TN and UG and are lacking for the region UA.
B.5. Surrounding landscape data

In eight regions (AT, CH, DE, FR, GB, HU, NL, and NO), based on aerial photographs, surrounding landscape characteristics of the sampled sites were elaborated. The percentage cover of ten habitat classes, which were distinguishable on the photographs, was estimated in a buffer zone of 250 m around the sampled sites. Each habitat class, covering at least 10% of the area in the buffer zone was recorded.

CLASS III. DATA SET STATUS AND ACCESSIBILITY

A. Status

A.1. Latest update

02 May 2016. This is the first version of this data set.

A.2. Latest archive date

02 May 2016

A.3. Metadata status

02 May 2016. This is the first version of the metadata.

B. Accessibility

B.1. Storage location and medium

B.2. Contact person

Felix Herzog at Agroscope, Institute for Sustainability Sciences, Zürich, Switzerland
felix.herzog@agroscope.admin.ch or the corresponding contact person per region, see Region_Description.txt.

B.3. Copyright restrictions

This data set is freely available for non-commercial scientific use, given the appropriate scholarly citation.

B.4. Proprietary restrictions

B.4.a. Release date

B.4.b. Citation

B.4.c. Disclaimers
Tab. 1: Overview of available data.

<table>
<thead>
<tr>
<th>Region</th>
<th>FarmType</th>
<th>Habitat data</th>
<th>Vascular plant data</th>
<th>Earthworm data</th>
<th>Spider data</th>
<th>Bee data</th>
<th>Management data</th>
<th>Surr. landscape data</th>
<th>No of farms (org)</th>
<th>No of mapped elements</th>
<th>No of sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT</td>
<td>Arable crops</td>
<td>O</td>
<td>O</td>
<td>Q</td>
<td>U a)</td>
<td></td>
<td>O</td>
<td>O</td>
<td>16 (8)</td>
<td>604</td>
<td>124</td>
</tr>
<tr>
<td>BG</td>
<td>Grassland</td>
<td>O</td>
<td>O</td>
<td>Q</td>
<td>O</td>
<td>O</td>
<td>X</td>
<td></td>
<td>16</td>
<td>344</td>
<td>142</td>
</tr>
<tr>
<td>CH</td>
<td>Grassland</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>X</td>
<td>19 (10)</td>
<td>490</td>
<td>139</td>
</tr>
<tr>
<td>DE</td>
<td>Mixed</td>
<td>O</td>
<td>O</td>
<td>Q</td>
<td>Q</td>
<td>O</td>
<td>O</td>
<td></td>
<td>16 (8)</td>
<td>1672</td>
<td>129</td>
</tr>
<tr>
<td>ED</td>
<td>Dehesa</td>
<td>O</td>
<td>Q</td>
<td>O</td>
<td>Q</td>
<td>O</td>
<td>X</td>
<td></td>
<td>10</td>
<td>801</td>
<td>145</td>
</tr>
<tr>
<td>EO</td>
<td>Olives</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>X</td>
<td></td>
<td>20 (10)</td>
<td>488</td>
<td>85</td>
</tr>
<tr>
<td>FR</td>
<td>Arable crops</td>
<td>O</td>
<td>O</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>16 (8)</td>
<td>1250</td>
<td>227</td>
</tr>
<tr>
<td>GB</td>
<td>Grassland</td>
<td>O</td>
<td>O</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>20 (10)</td>
<td>2529</td>
<td>236</td>
</tr>
<tr>
<td>HU</td>
<td>Mixed</td>
<td>O</td>
<td>O</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>Q</td>
<td></td>
<td>18 (7)</td>
<td>443</td>
<td>156</td>
</tr>
<tr>
<td>IT</td>
<td>Vineyards</td>
<td>O</td>
<td>O</td>
<td>Q</td>
<td>U</td>
<td>Q</td>
<td>X</td>
<td></td>
<td>18 (9)</td>
<td>587</td>
<td>74</td>
</tr>
<tr>
<td>NL</td>
<td>Horticulture</td>
<td>O</td>
<td>Q</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>14 (11)</td>
<td>425</td>
<td>113</td>
</tr>
<tr>
<td>NO</td>
<td>Grassland</td>
<td>O</td>
<td>Q</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12 (6)</td>
<td>577</td>
<td>119</td>
</tr>
<tr>
<td>TN</td>
<td>Olives</td>
<td>O</td>
<td>O</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>20 (8)</td>
<td>534</td>
<td>142</td>
</tr>
<tr>
<td>UA</td>
<td>Arable crops</td>
<td>O</td>
<td>O</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6</td>
<td>461</td>
<td>180</td>
</tr>
<tr>
<td>UG</td>
<td>Mixed</td>
<td>O</td>
<td>O</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>16</td>
<td>133</td>
<td>104</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>a)</th>
<th>b)</th>
<th>c)</th>
<th>d)</th>
<th>e)</th>
<th>f)</th>
<th>g)</th>
<th>h)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>no third survey</td>
<td>no separation of methods</td>
<td>fauna sampled on two organic and two non-organic farms (some additional gaps, e.g. lacking third survey for bees)</td>
<td>samples pooled per survey, (in the region GB, for one survey separate samples, for two surveys pooled)</td>
<td>no fauna sampled on six organic farms (some additional gaps), i.e. five organic and three non-organic farms are more or less complete</td>
<td>no earthworms found</td>
<td>data only partly or not available</td>
<td>no sufficiently stable spider taxonomy</td>
</tr>
</tbody>
</table>

Supplemental descriptors

Note that for the investigations within the EU project BioBio, certain samples were excluded. Due to region specific conditions, the separation of different habitat types was under development during the project. Therefore, in some cases, several habitats of the same type within one farm were sampled. Such oversampled habitats were excluded for project-related analyses. Here, all sampled data are included.
B.5. Costs

None.

Acknowledgments

We thank all farmers who allowed access to their fields and provided information on the agricultural management. Helpful comments of two anonymous reviewers are acknowledged.

Literature cited


