

# Diversity of Collembola under various types of anthropogenic load on ecosystems of European part of Russia

Nataliya Kuznetsova<sup>‡</sup>, Natalya Ivanova<sup>§</sup>

<sup>‡</sup> Moscow State Pedagogical University, Moscow, Russia

<sup>§</sup> Institute of Mathematical Problems of Biology RAS – the Branch of Keldysh Institute of Applied Mathematics of Russian Academy of Sciences, Pushchino, Russia

Corresponding author: Nataliya Kuznetsova ([mpnk@yandex.ru](mailto:mpnk@yandex.ru))

Academic editor: Dmitry Schigel

## Abstract

### Background

Despite the key role played by soil organisms in the functioning of terrestrial ecosystems and provisioning of ecosystem services (Barrios 2007, Bardgett and Putten 2014), available open data on soil biodiversity are incongruously scarce (Eisenhauer 2017, Cameron 2018). This is especially true for Russia, but contrasts long traditions of soil zoological research and large volumes of data that were collected during the second half of the 20<sup>th</sup> century for the territory of the former USSR. Last year, 41,928 georeferenced occurrences of soil-dwelling arthropods Collembola were digitised and [published through GBIF.org](#). This work continues these activities. The article combines descriptions of three new sampling-event datasets about the various types of anthropogenic load on the diversity and the abundance of Collembola, small arthropods involved in the destruction of organic residues in the soil:

1. **Collembola of winter wheat fields in the Kaluga Region: conservation treatment versus conventional one** (Kuznetsova et al. 2020). The following variants were studied: 1) treatment with organic fertilisers and tillage, without mineral fertilisers and pesticides, 2) the same, but without tillage, only discing; 3) with mineral fertilisers, pesticides and tillage. Special multi-scale sampling design was used. The material was collected on 24-26 July 2019 in Kaluga Region, European part of Russia. Data on 2226 records on 7302 specimens of 32 species in six fields in 486 soil cores are presented.
2. **Collembola of broadleaved forests along gradient of urbanisation in Moscow** (Kuznetsova and Ageeva 2020). Sampling plots were placed in oak and lime forests located at different distances from the centre of Moscow. The material was

collected in different seasons of 1990–1991. Data on 1737 records on 6873 specimens of 64 species (17 series of sampling, 720 soil cores) are presented.

3. **Collembola in clear cutting areas of Arkhangelsk Region: spatial and temporal series of the data** (Kuznetsova and Klyueva 2020). Sampling plots were in birch forests of different ages with spruce underbrush and in old spruce forest. The study was carried out in July of 1970–1971 and 1984 in Arkhangelsk Region, European part of Russia. In 1970, cores were taken at sites where the forest was restored 15, 30 and 80 years after clear cuttings, as well as in a 180-year-old spruce forest. In 1984, sampling was repeated in two plots. Data on 1468 records on 18788 specimens of 47 species (seven series of sampling, 720 soil cores) are presented.

## New information

These datasets contribute to filling gaps in the global biodiversity distribution of the Collembola. All datasets present new information about effects of agricultural treatments, urbanisation and clear cutting on springtail diversity and abundance in ecosystems of the European part of Russia.

## Keywords

sampling event, springtails, winter wheat fields, conservation treatment in agriculture, urbanisation, clear cutting, multi-scale sampling design, chronosequence method, reforestation, secondary succession

## Introduction

Collembola, or springtails, is one of the most diverse and abundant groups of small arthropods in soil (Petersen and Luxton 1982, Hopkin 1997). Sampling-event datasets on Collembola in anthropogenic sites have not been included in GBIF until now. There is only poor data on the effects of the considered types of anthropogenic impact on the diversity of Collembola on the territory of the European part of Russia in literature, as well. The information about effects of conservation treatment in organic agriculture on springtail diversity is absent. Only few papers about effects of urbanisation in Moscow are known (Kuznetsova 1994, Sterzynska and Kuznetsova 1997), although without information about springtails of urban forests. Data on the changes of Collembola diversity during reforestation, i.e. under secondary succession after clear cutting, was partly provided in publications (Kuznetsova 2005).

## General description

**Purpose:** The purpose of the data paper is to present information on Collembola for ecosystems under very common types of anthropogenic load (agriculture, urbanisation, clear cutting) in the European part of Russia. We pursued this to show the diversity and the abundance in the most detailed sample-event form.

## Sampling methods

**Description:** The data paper based on three datasets:

**The “agricultural” dataset** (Kuznetsova et al. 2020) provides information on the number of individuals of springtail species in soil cores collected in winter wheat fields under different agricultural treatments in the Kaluga Region. The following variants were studied:

1. treatment with organic fertilisers and tillage, without mineral fertilisers and pesticides (Fig. 1);
2. the same, but without tillage, only discing;
3. with mineral fertilisers, pesticides and tillage.

Two fields were considered for each variant. Data on 2226 records on 7302 specimens of 32 species in six fields in 486 soil cores are presented.

**The “urban” dataset** (Kuznetsova and Ageeva 2020) provides information on the number of individuals of springtail species in soil cores collected in oak and lime forests located at different distances from Moscow centre. Sampling plots in Neskuchny Sad (lime, *Tilia cordata*) (Fig. 2) and in the Central Botanical Garden (oak, *Quercus robur*) were closer to the centre than others. Sampling plots near Uzkoe Hall (oak) and in the Bitsa Park (lime) were close to the boundaries of Moscow. Sampling plots in the area of Troitsk (lime) and Shishkin Les village (oak) were in the surroundings of Moscow (now belonging to the territory of New Moscow). The material was collected from 1990-1991. In lime forests, three series of cores were taken in June 1990 and one in October 1990. In oak forests, 12 series were taken in different seasons of 1990–1991. One series was taken under the larch trees on the territory of the Central Botanical Garden. Data on 1737 records on 6873 specimens of 64 species (17 series of sampling, 720 soil cores) are presented.

**The “clear cutting” dataset** (Kuznetsova and Klyueva 2020) provides information on the number of individuals of springtail species in soil cores collected in birch forests of different ages with spruce underbrush and in old spruce forest in Arkhangelsk Region, European part of Russia. In 1970, cores were taken at sites where the forest was restored 15, 30 and 80 years after clear cuttings (Fig. 3), as well as in a 180-year-old spruce forest.

In 1984, sampling was repeated in two plots. Data on 1468 records on 18788 specimens of 47 species (seven series of sampling, 211 soil cores) are presented.

**Sampling description: The “agricultural” dataset.** The material was received using the multi-scale sampling design. The method is appropriate to study the structure of biodiversity (Lande 1996), communities and populations at different spatial scales (Azovsky et al. 2000). Fractal arrangement of cores saves sample effort because the same core is used for analysis at different scales (Marsh and Ewers 2013). The approach is common in hydrobiology, entomology etc., but it continues to be rare in soil zoology (Kuznetsova and Saraeva 2018). We used a small size of the corer (8 cm<sup>2</sup> in section) due to the necessary special attention of diversity and spatial structure of population at the micro level. Soil was investigated down to 20 cm. A total of 81 cores were taken in each field when sampling. Cores were placed in the corners of different-scale equilateral triangles inscribed in squares with sides 10 cm, 25 cm, 1 m and 10 m. The different-scale triangles were designed following the principles of fractal geometry. The sample design is described in detail by Saraeva et al. (2015).

**The “urban” dataset.** Sites of oak and lime forests were studied at different distances from the megapolis centre. A similar approach was applied by different authors (Weigmann and Kratz 1987 in Berlin; Sterzynska 1990 in Warsaw; etc.). Design of sampling was based on regular arrangements which cover different forest microsites. The regular approach to sampling is as common in soil zoology as random sampling (Petersen and Luxton 1982). A quadrangular frame of 5×5 cm was used for sampling. Each core was subdivided into three layers: 1) litter, 2) soil 0–5 cm and 3) soil 5–10 cm. Cores were arranged along lines between trees: near tree trunks (cores with numbers 1, 5, 6, 10, 11, 15), under middle of tree crowns (numbers 2, 4, 7, 9, 12, 14) and in a gap between trees (numbers 3, 8, 13). A total of 15 cores in three lines were taken when sampling. One line included five cores. The distance between lines of cores was about 10 m. The sample design is described in detail by Potapov and Kuznetsova (2011).

**The “clear cutting” dataset.** Design of sampling was based on the regular arrangement which covers different forest microsites. Forests of different ages after clear cutting were studied according to the chronosequence method (Johnson and Miyanishi 2008). It was supplemented with material taken at the same sites 14 years after the first sampling. Thus, the real changes in Collembola diversity can be estimated. A quadrangular frame of 5×5 cm was used for sampling in all sites. Each core was subdivided into three layers: 1) litter (L-layer) with ground cover, 2) fermentative layer (F) and 3) humus layer (H) with 2 cm of mineral soil. Cores were placed along lines between trees: near tree trunks (cores with numbers 1, 5, 6, 10), under middle of tree crowns (numbers 2, 4, 7, 9) and in a gap between trees (numbers 3, 8). Ten cores were taken when sampling. The distance between two lines of cores was about 10 m.

### **Extraction of Collembola from cores**

Tullgren funnels were used for Collembola extraction from soil cores into 70% alcohol (Gilyarov 1975, Petersen and Luxton 1982).

## Laboratory processing

All individuals of springtails were mounted on slides in Phoera liquid according to a standard procedure (Potapov and Kuznetsova 2011). Springtails were identified to species level using a microscope.

**Quality control:** Keys by Fjellberg (Fjellberg 1980, Fjellberg 1998, Fjellberg 2007), Babenko et al. (1994), Potapov (2002)) and particular taxonomic articles were used for species identification. Experts on different families of springtails were consulted.

**Step description:** Data on species were digitised, standardised, the quality of data was checked and errors were corrected and then published.

## Geographic coverage

**Description:** All material was collected in the European part of Russia (Fig. 4): Moscow City, Arkhangelsk, Kaluga Regions.

**Coordinates:** 54°35'6" and 63°57'43.2" Latitude; 34°52'48" and 40°37'55.2" Longitude.

## Taxonomic coverage

**Taxa included:**

Rank	Scientific Name
phylum	Arthropoda
order	Collembola

## Temporal coverage

**Formation period:** July 24, 2019 - July 26, 2019 for DOI 10.15468/rv6g98; 1990-1991 for DOI: 10.15468/e25d3s; 1970-1971, 1984 for DOI: 10.15468/z38wxq.

## Usage licence

**Usage licence:** Open Data Commons Attribution License

## Data resources

**Data package title:** Diversity of Collembola under various types of anthropogenic load on ecosystems of the European part of Russia.

**Number of data sets: 3**

**Data set name:** Collembola of winter wheat fields in the Kaluga Region: conservation treatment versus conventional one.

**Character set:** UTF-8

**Download URL:** <https://www.gbif.org/dataset/575c5097-521d-47ef-908b-cc659ff249f4>

**Data format:** Darwin Core Archive format.

**Description:** The dataset includes two Darwin Core (Wieczorek 2012) tables related by the eventID field – Events and Associated occurrences. The occurrence table includes only occurrence-present records.

Column label	Column description
eventID (Event Core)	An identifier for the event (core) <a href="https://dwc.tdwg.org/terms/#dwc:eventID">https://dwc.tdwg.org/terms/#dwc:eventID</a>
locationID (Event Core)	An identifier for the place of field data collection <a href="https://dwc.tdwg.org/terms/#dwc:locationID">https://dwc.tdwg.org/terms/#dwc:locationID</a>
dynamicProperties (Event Core)	Description of the event <a href="https://dwc.tdwg.org/terms/#dwc:dynamicProperties">https://dwc.tdwg.org/terms/#dwc:dynamicProperties</a>
eventDate (Event Core)	Field data collection date (YYYY-MM-DD) <a href="https://dwc.tdwg.org/terms/#dwc:eventDate">https://dwc.tdwg.org/terms/#dwc:eventDate</a>
samplingProtocol (Event Core)	Sampling protocol <a href="https://dwc.tdwg.org/terms/#dwc:samplingProtocol">https://dwc.tdwg.org/terms/#dwc:samplingProtocol</a>
country (Event Core)	Country name (Russian Federation) <a href="https://dwc.tdwg.org/terms/#dwc:country">https://dwc.tdwg.org/terms/#dwc:country</a>
countryCode (Event Core)	The standard code for the Russian Federation according to ISO 3166-1-alpha-2 (RU) <a href="https://dwc.tdwg.org/terms/#dwc:countryCode">https://dwc.tdwg.org/terms/#dwc:countryCode</a>
stateProvince (Event Core)	Region ('oblast') name. The first-level administrative division <a href="https://dwc.tdwg.org/terms/#dwc:stateProvince">https://dwc.tdwg.org/terms/#dwc:stateProvince</a>
institutionCode (Event Core)	Short name of the institution (MPGU) <a href="https://dwc.tdwg.org/terms/#dwc:institutionCode">https://dwc.tdwg.org/terms/#dwc:institutionCode</a>
decimalLatitude (Event Core)	The geographic latitude in decimal degrees of the geographic centre of the data sampling place <a href="https://dwc.tdwg.org/terms/#dwc:decimalLatitude">https://dwc.tdwg.org/terms/#dwc:decimalLatitude</a>
decimalLongitude (Event Core)	The geographic longitude in decimal degrees of the geographic centre of the data sampling place <a href="https://dwc.tdwg.org/terms/#dwc:decimalLongitude">https://dwc.tdwg.org/terms/#dwc:decimalLongitude</a>
geodeticDatum (Event Core)	Spatial reference system (SRS) upon which the geographic coordinates given in decimalLatitude and decimalLongitude are based <a href="https://dwc.tdwg.org/terms/#dwc:geodeticDatum">https://dwc.tdwg.org/terms/#dwc:geodeticDatum</a>
coordinatePrecision (Event Core)	The fraction of a degree corresponding to the number of significant digits in the source coordinates <a href="https://dwc.tdwg.org/terms/#dwc:coordinatePrecision">https://dwc.tdwg.org/terms/#dwc:coordinatePrecision</a>

coordinateUncertaintyInMetres (Event Core)	The maximum uncertainty distance in metres <a href="https://dwc.tdwg.org/terms/#dwc:coordinateUncertaintyInMeters">https://dwc.tdwg.org/terms/#dwc:coordinateUncertaintyInMeters</a>
sampleSizeValue (Event Core)	Size of the sampling core <a href="https://dwc.tdwg.org/terms/#dwc:sampleSizeValue">https://dwc.tdwg.org/terms/#dwc:sampleSizeValue</a>
sampleSizeUnit (Event Core)	The unit of measurement of the size sampling core <a href="https://dwc.tdwg.org/terms/#dwc:sampleSizeUnit">https://dwc.tdwg.org/terms/#dwc:sampleSizeUnit</a>
eventID (Occurrence Extension)	An identifier for the event (core) <a href="https://dwc.tdwg.org/terms/#dwc:eventID">https://dwc.tdwg.org/terms/#dwc:eventID</a>
occurrenceID (Occurrence Extension)	An identifier for the record <a href="https://dwc.tdwg.org/terms/#dwc:occurrenceID">https://dwc.tdwg.org/terms/#dwc:occurrenceID</a>
basisOfRecord (Occurrence Extension)	Basis of the record (PreservedSpecimen) <a href="https://dwc.tdwg.org/terms/#dwc:basisOfRecord">https://dwc.tdwg.org/terms/#dwc:basisOfRecord</a>
scientificName (Occurrence Extension)	Scientific name <a href="https://dwc.tdwg.org/terms/#dwc:scientificName">https://dwc.tdwg.org/terms/#dwc:scientificName</a>
kingdom (Occurrence Extension)	The full scientific name of the kingdom (Animalia) <a href="https://dwc.tdwg.org/terms/#dwc:kingdom">https://dwc.tdwg.org/terms/#dwc:kingdom</a>
order (Occurrence Extension)	The full scientific name of the order (Collembola) <a href="https://dwc.tdwg.org/terms/#dwc:order">https://dwc.tdwg.org/terms/#dwc:order</a>
taxonRank (Occurrence Extension)	The taxonomic rank <a href="https://dwc.tdwg.org/terms/#dwc:taxonRank">https://dwc.tdwg.org/terms/#dwc:taxonRank</a>
lifeStage (Occurrence Extension)	The life stage of individuals. Here it is used for juvenile individuals indicated. <a href="https://dwc.tdwg.org/terms/#dwc:lifeStage">https://dwc.tdwg.org/terms/#dwc:lifeStage</a>
individualCount (Occurrence Extension)	The number of individuals represented in the event <a href="https://dwc.tdwg.org/terms/#dwc:individualCount">https://dwc.tdwg.org/terms/#dwc:individualCount</a>
identifiedBy (Occurrence Extension)	List of persons, who identified collected Collembola <a href="https://dwc.tdwg.org/terms/#dwc:identifiedBy">https://dwc.tdwg.org/terms/#dwc:identifiedBy</a>
recordedBy (Occurrence Extension)	List of persons who collected field data <a href="https://dwc.tdwg.org/terms/#dwc:recordedBy">https://dwc.tdwg.org/terms/#dwc:recordedBy</a>

**Data set name:** Collembola of broadleaved forests along gradient of urbanisation in Moscow.

**Character set:** UTF-8

**Download** [336e3eb6-0ed7-46a7-8a13-1faec0d3f8f2](https://www.gbif.org/dataset/336e3eb6-0ed7-46a7-8a13-1faec0d3f8f2)

**URL:** <https://www.gbif.org/dataset/336e3eb6-0ed7-46a7-8a13-1faec0d3f8f2>

**Data format:** Darwin Core Archive format

**Description:** The dataset includes two Darwin Core tables related by the eventID field – Events and Associated occurrences. The occurrence table includes only occurrence-present records.

Column label	Column description
parentEventID (Event Core)	An identifier for core <a href="https://dwc.tdwg.org/terms/#dwc:parentEventID">https://dwc.tdwg.org/terms/#dwc:parentEventID</a>
eventID (Event Core)	An identifier for the event (layer) <a href="https://dwc.tdwg.org/terms/#dwc:eventID">https://dwc.tdwg.org/terms/#dwc:eventID</a>
locationID (Event Core)	An identifier for the place of field data collection <a href="https://dwc.tdwg.org/terms/#dwc:locationID">https://dwc.tdwg.org/terms/#dwc:locationID</a>
dynamicProperties (Event Core)	Description of the event <a href="https://dwc.tdwg.org/terms/#dwc:dynamicProperties">https://dwc.tdwg.org/terms/#dwc:dynamicProperties</a>
minimumDepthInMetres (Event Core)	The lesser depth of a range of depth below the local surface, in metres <a href="https://dwc.tdwg.org/terms/#dwc:minimumDepthInMeters">https://dwc.tdwg.org/terms/#dwc:minimumDepthInMeters</a>
maximumDepthInMetres (Event Core)	The greater depth of a range of depth below the local surface, in metres <a href="https://dwc.tdwg.org/terms/#dwc:maximumDepthInMeters">https://dwc.tdwg.org/terms/#dwc:maximumDepthInMeters</a>
eventDate (Event Core)	Field data collection date (YYYY-MM-DD) <a href="https://dwc.tdwg.org/terms/#dwc:eventDate">https://dwc.tdwg.org/terms/#dwc:eventDate</a>
samplingProtocol (Event Core)	Sampling protocol <a href="https://dwc.tdwg.org/terms/#dwc:samplingProtocol">https://dwc.tdwg.org/terms/#dwc:samplingProtocol</a>
country (Event Core)	Country name (Russian Federation) <a href="https://dwc.tdwg.org/terms/#dwc:country">https://dwc.tdwg.org/terms/#dwc:country</a>
countryCode (Event Core)	The standard code for the Russian Federation according ISO 3166-1-alpha-2 (RU) <a href="https://dwc.tdwg.org/terms/#dwc:countryCode">https://dwc.tdwg.org/terms/#dwc:countryCode</a>
institutionCode (Event Core)	Short name of the institution (MPGU) <a href="https://dwc.tdwg.org/terms/#dwc:institutionCode">https://dwc.tdwg.org/terms/#dwc:institutionCode</a>
decimalLatitude (Event Core)	The geographic latitude in decimal degrees of the geographic centre of the data sampling place <a href="https://dwc.tdwg.org/terms/#dwc:decimalLatitude">https://dwc.tdwg.org/terms/#dwc:decimalLatitude</a>
decimalLongitude (Event Core)	The geographic longitude in decimal degrees of the geographic centre of the data sampling place <a href="https://dwc.tdwg.org/terms/#dwc:decimalLongitude">https://dwc.tdwg.org/terms/#dwc:decimalLongitude</a>
geodeticDatum (Event Core)	Spatial reference system (SRS) upon which the geographic coordinates given in decimalLatitude and decimalLongitude are based <a href="https://dwc.tdwg.org/terms/#dwc:geodeticDatum">https://dwc.tdwg.org/terms/#dwc:geodeticDatum</a>
coordinateUncertaintyInMetres (Event Core)	The maximum uncertainty distance in metres <a href="https://dwc.tdwg.org/terms/#dwc:coordinateUncertaintyInMeters">https://dwc.tdwg.org/terms/#dwc:coordinateUncertaintyInMeters</a>
coordinatePrecision (Event Core)	The fraction of a degree corresponding to the number of significant digits in the source coordinates <a href="https://dwc.tdwg.org/terms/#dwc:coordinatePrecision">https://dwc.tdwg.org/terms/#dwc:coordinatePrecision</a>
sampleSizeValue (Event Core)	Size of the sampling core <a href="https://dwc.tdwg.org/terms/#dwc:sampleSizeValue">https://dwc.tdwg.org/terms/#dwc:sampleSizeValue</a>
sampleSizeUnit (Event Core)	The unit of measurement of the size sampling core <a href="https://dwc.tdwg.org/terms/#dwc:sampleSizeUnit">https://dwc.tdwg.org/terms/#dwc:sampleSizeUnit</a>
rightsHolder (Event Core)	A person owning rights over datasets <a href="https://dwc.tdwg.org/terms/#dcterms:rightsHolder">https://dwc.tdwg.org/terms/#dcterms:rightsHolder</a>
references (Event Core)	Bibliographic references pointing to the data <a href="https://dwc.tdwg.org/terms/#dcterms:references">https://dwc.tdwg.org/terms/#dcterms:references</a>



habitat (Event Core)	A category or description of the habitat <a href="https://dwc.tdwg.org/terms/#dwc:habitat">https://dwc.tdwg.org/terms/#dwc:habitat</a>
eventID (Occurrence Extension)	An identifier for the event (layer) <a href="https://dwc.tdwg.org/terms/#dwc:eventID">https://dwc.tdwg.org/terms/#dwc:eventID</a>
occurrenceID (Occurrence Extension)	An identifier for the record <a href="https://dwc.tdwg.org/terms/#dwc:occurrenceID">https://dwc.tdwg.org/terms/#dwc:occurrenceID</a>
basisOfRecord (Occurrence Extension)	Basis of the record (PreservedSpecimen) <a href="https://dwc.tdwg.org/terms/#dwc:basisOfRecord">https://dwc.tdwg.org/terms/#dwc:basisOfRecord</a>
scientificName (Occurrence Extension)	Scientific name <a href="https://dwc.tdwg.org/terms/#dwc:scientificName">https://dwc.tdwg.org/terms/#dwc:scientificName</a>
kingdom (Occurrence Extension)	The full scientific name of the kingdom (Animalia) <a href="https://dwc.tdwg.org/terms/#dwc:kingdom">https://dwc.tdwg.org/terms/#dwc:kingdom</a>
order (Occurrence Extension)	The full scientific name of the order (Collembola) <a href="https://dwc.tdwg.org/terms/#dwc:order">https://dwc.tdwg.org/terms/#dwc:order</a>
taxonRank (Occurrence Extension)	The taxonomic rank <a href="https://dwc.tdwg.org/terms/#dwc:taxonRank">https://dwc.tdwg.org/terms/#dwc:taxonRank</a>
lifeStage (Occurrence Extension)	The life stage of individuals. Here it is used for juvenile individuals indicated. <a href="https://dwc.tdwg.org/terms/#dwc:lifeStage">https://dwc.tdwg.org/terms/#dwc:lifeStage</a>
individualCount (Occurrence Extension)	The number of individuals represented in the event <a href="https://dwc.tdwg.org/terms/#dwc:individualCount">https://dwc.tdwg.org/terms/#dwc:individualCount</a>
recordedBy (Occurrence Extension)	List of persons who collected field data <a href="https://dwc.tdwg.org/terms/#dwc:recordedBy">https://dwc.tdwg.org/terms/#dwc:recordedBy</a>
identifiedBy (Occurrence Extension)	List of persons who identified collected Collembola <a href="https://dwc.tdwg.org/terms/#dwc:identifiedBy">https://dwc.tdwg.org/terms/#dwc:identifiedBy</a>

**Data set name:** Collembola in clear cutting areas of Arkhangelsk Region: spatial and temporal series of the data.

**Character set:** UTF-8

**Download** <https://www.gbif.org/dataset/36ceb840-3011-4411-8739-1a0f02272489>

**URL:** <https://www.gbif.org/dataset/36ceb840-3011-4411-8739-1a0f02272489>

**Data format:** Darwin Core Archive format

**Description:** The dataset includes two Darwin Core tables related by the eventID field – Events and Associated occurrences. The occurrence table includes only occurrence-present records.

Column label	Column description
parentEventID (Event Core)	An identifier for core <a href="https://dwc.tdwg.org/terms/#dwc:parentEventID">https://dwc.tdwg.org/terms/#dwc:parentEventID</a>

eventID (Event Core)	An identifier for the event (layer) <a href="https://dwc.tdwg.org/terms/#dwc:eventID">https://dwc.tdwg.org/terms/#dwc:eventID</a>
samplingProtocol (Event Core)	Sampling protocol <a href="https://dwc.tdwg.org/terms/#dwc:samplingProtocol">https://dwc.tdwg.org/terms/#dwc:samplingProtocol</a>
sampleSizeValue (Event Core)	Size of the sampling core <a href="https://dwc.tdwg.org/terms/#dwc:sampleSizeValue">https://dwc.tdwg.org/terms/#dwc:sampleSizeValue</a>
sampleSizeUnit (Event Core)	The unit of measurement of the size sampling core <a href="https://dwc.tdwg.org/terms/#dwc:sampleSizeUnit">https://dwc.tdwg.org/terms/#dwc:sampleSizeUnit</a>
decimalLatitude (Event Core)	The geographic latitude in decimal degrees of the geographic centre of the data sampling place <a href="https://dwc.tdwg.org/terms/#dwc:decimalLatitude">https://dwc.tdwg.org/terms/#dwc:decimalLatitude</a>
decimalLongitude (Event Core)	The geographic longitude in decimal degrees of the geographic centre of the data sampling place <a href="https://dwc.tdwg.org/terms/#dwc:decimalLongitude">https://dwc.tdwg.org/terms/#dwc:decimalLongitude</a>
geodeticDatum (Event Core)	Spatial reference system (SRS) upon which the geographic coordinates given in decimalLatitude and decimalLongitude are based <a href="https://dwc.tdwg.org/terms/#dwc:geodeticDatum">https://dwc.tdwg.org/terms/#dwc:geodeticDatum</a>
coordinatePrecision (Event Core)	The fraction of a degree corresponding to the number of significant digits in the source coordinates <a href="https://dwc.tdwg.org/terms/#dwc:coordinatePrecision">https://dwc.tdwg.org/terms/#dwc:coordinatePrecision</a>
coordinateUncertaintyInMetres (Event Core)	The maximum uncertainty distance in metres <a href="https://dwc.tdwg.org/terms/#dwc:coordinateUncertaintyInMeters">https://dwc.tdwg.org/terms/#dwc:coordinateUncertaintyInMeters</a>
countryCode (Event Core)	The standard code for the Russian Federation according to ISO 3166-1-alpha-2 <a href="https://dwc.tdwg.org/terms/#dwc:countryCode">https://dwc.tdwg.org/terms/#dwc:countryCode</a>
country (Event Core)	Country name (Russian Federation) <a href="https://dwc.tdwg.org/terms/#dwc:country">https://dwc.tdwg.org/terms/#dwc:country</a>
stateProvince (Event Core)	Region ('oblast') name. The first-level administrative division <a href="https://dwc.tdwg.org/terms/#dwc:stateProvince">https://dwc.tdwg.org/terms/#dwc:stateProvince</a>
locationID (Event Core)	An identifier for the place of field data collection <a href="https://dwc.tdwg.org/terms/#dwc:locationID">https://dwc.tdwg.org/terms/#dwc:locationID</a>
habitat (Event Core)	A category or description of the habitat <a href="https://dwc.tdwg.org/terms/#dwc:habitat">https://dwc.tdwg.org/terms/#dwc:habitat</a>
verbatimEventDate (Event Core)	Field data collection date in original data <a href="https://dwc.tdwg.org/terms/#dwc:verbatimEventDate">https://dwc.tdwg.org/terms/#dwc:verbatimEventDate</a>
eventDate (Event Core)	Field data collection date (YYYY-MM-DD/DD) <a href="https://dwc.tdwg.org/terms/#dwc:eventDate">https://dwc.tdwg.org/terms/#dwc:eventDate</a>
institutionCode (Event Core)	Short name of the institution (MPGU) <a href="https://dwc.tdwg.org/terms/#dwc:institutionCode">https://dwc.tdwg.org/terms/#dwc:institutionCode</a>
language (Event Core)	Language of datasets (EN) <a href="https://dwc.tdwg.org/terms/#dc:language">https://dwc.tdwg.org/terms/#dc:language</a>
dynamicProperties (Event Core)	Description of the event <a href="https://dwc.tdwg.org/terms/#dwc:dynamicProperties">https://dwc.tdwg.org/terms/#dwc:dynamicProperties</a>
eventID (Occurrence Extension)	An identifier for the event (layer) <a href="https://dwc.tdwg.org/terms/#dwc:eventID">https://dwc.tdwg.org/terms/#dwc:eventID</a>
occurrenceID (Occurrence Extension)	An identifier for the record <a href="https://dwc.tdwg.org/terms/#dwc:occurrenceID">https://dwc.tdwg.org/terms/#dwc:occurrenceID</a>

basisOfRecord (Occurrence Extension)	Basis of the record (PreservedSpecimen) <a href="https://dwc.tdwg.org/terms/#dwc:basisOfRecord">https://dwc.tdwg.org/terms/#dwc:basisOfRecord</a>
scientificName (Occurrence Extension)	Scientific name <a href="https://dwc.tdwg.org/terms/#dwc:scientificName">https://dwc.tdwg.org/terms/#dwc:scientificName</a>
kingdom (Occurrence Extension)	The full scientific name of the kingdom (Animalia) <a href="https://dwc.tdwg.org/terms/#dwc:kingdom">https://dwc.tdwg.org/terms/#dwc:kingdom</a>
order (Occurrence Extension)	The full scientific name of the order (Collembola) <a href="https://dwc.tdwg.org/terms/#dwc:order">https://dwc.tdwg.org/terms/#dwc:order</a>
taxonRank (Occurrence Extension)	The taxonomic rank <a href="https://dwc.tdwg.org/terms/#dwc:taxonRank">https://dwc.tdwg.org/terms/#dwc:taxonRank</a>
lifeStage (Occurrence Extension)	The life stage of individuals. Here it is used for juvenile individuals indicated. <a href="https://dwc.tdwg.org/terms/#dwc:lifeStage">https://dwc.tdwg.org/terms/#dwc:lifeStage</a>
individualCount (Occurrence Extension)	The number of individuals represented in the event <a href="https://dwc.tdwg.org/terms/#dwc:individualCount">https://dwc.tdwg.org/terms/#dwc:individualCount</a>
recordedBy (Occurrence Extension)	List of persons who collected field data <a href="https://dwc.tdwg.org/terms/#dwc:recordedBy">https://dwc.tdwg.org/terms/#dwc:recordedBy</a>
identifiedBy (Occurrence Extension)	List of persons who identified collected Collembola <a href="https://dwc.tdwg.org/terms/#dwc:identifiedBy">https://dwc.tdwg.org/terms/#dwc:identifiedBy</a>

## Acknowledgements

The material for the “agricultural” dataset was funded by the Russian Fund for Basic Research, project 19-516-60002. The authors are grateful to Tatyana Ageeva, Nellie Klyueva and Mikhail Potapov who participated in the sampling of the material of different datasets. We are also thankful to Mikhail Potapov, Ksenia Panina, Yulia Shveenkova and Anna Bokova who helped in processing and identification of the material. The work of Natalya Ivanova was done within the framework of the Basic Scientific Research theme of the Institute of Mathematical Problems of Biology RAS – the Branch of Keldysh Institute of Applied Mathematics of the Russian Academy of Sciences.

## References

- Azovsky A, Chertoprood M, Kucheruk N, Rybnikov P, Sapozhnikov F (2000) Fractal properties of spatial distribution of intertidal benthic communities. *Marine Biology* 136: 581-590. <https://doi.org/10.1007/s002270050718>
- Babenko A, Potapov M, Stebaeva S, Chernova N (1994) Keys of springtails of the fauna of Russia and adjacent countries: Family Hypogastruridae. [Определитель коллембол фауны России и сопредельных стран: Семейство Hypogastruridae]. Nauka, Moskva, 336 pp. [In Russian].

- Bardgett R, Putten Wvd (2014) Belowground biodiversity and ecosystem functioning. *Nature* 515: 505-511. <https://doi.org/10.1038/nature13855>
- Barrios E (2007) Soil biota, ecosystem services and land productivity. *Ecological Economics* 64: 269-285. <https://doi.org/10.1016/j.ecolecon.2007.03.004>
- Cameron E, et al. (2018) Global gaps in soil biodiversity data. *Nature Ecology & Evolution* 2 (7): 1042-1043. <https://doi.org/10.1038/s41559-018-0573-8>
- Eisenhauer N, et al. (2017) Priorities for research in soil ecology. *Pedobiologia* 63: 1-7. <https://doi.org/10.1016/j.pedobi.2017.05.003>
- Fjellberg A (1980) Identification keys to Norwegian Collembola. *Geography. Norsk Entomologisk Forening, As, Norway*, 152 pp.
- Fjellberg A (1998) The Collembola of Fennoscandia and Denmark, Part I: Poduromorpha. *Fauna Entomologica Scandinavica*. 35. Brill, 184 pp.
- Fjellberg A (2007) The Collembola of Fennoscandia and Denmark, Part II: Entomobryomorpha and Symphyleona. *Fauna Entomologica Scandinavica*. Vol. 42. Brill, 266 pp. <https://doi.org/10.1163/ej.9789004157705.i-265>
- Gilyarov M (1975) Sampling of small arthropods (microfauna) and nematodes. In: Gilyarov M (Ed.) *Metody pochvenno-zoologicheskikh issledovaniy*. Nauka, Moskva. [In Russian].
- Hopkin S (1997) *Biology of the springtails (Insecta: Collembola)*. Oxford University Press, Oxford, 330 pp. <https://doi.org/10.1017/S1742758400019263>
- Johnson E, Miyanishi K (2008) Testing the assumptions of chronosequences in succession. *Ecology Letters* 11 (5): 419-431. <https://doi.org/10.1111/j.1461-0248.2008.01173.x>
- Kuznetsova N (1994) Collembolan guild structure as an indicator of tree plantation conditions in urban areas. *Memorabilia Zoologica (Warszawa)* 49: 197-205.
- Kuznetsova N (2005) Organization of communities of soil-dwelling Collembola. [Организация сообществ почвообитающих коллембол]. *GNO Prometey MPGU, Moskva*, 244 pp. [In Russian].
- Kuznetsova N, Saraeva A (2018) Beta-diversity partitioning approach in soil zoology: A case of Collembola in pine forests. *Geoderma* 332: 142-152. <https://doi.org/10.1016/j.geoderma.2017.09.030>
- Kuznetsova NA, Ageeva TZ (2020) Collembola of broadleaved forests along gradient of urbanization in Moscow. Sampling event dataset. 1.4. Moscow Pedagogical State University (MPGU).. Release date: 2020-8-25. URL: <https://doi.org/10.15468/e25d3s>
- Kuznetsova NA, Klyueva NM (2020) Collembola in clear cutting areas of Arkhangelsk Region: spatial and temporal series of the data. Sampling event dataset. 1.7. Moscow Pedagogical State University (MPGU). Release date: 2020-8-22. URL: <https://doi.org/10.15468/z38wxq>
- Kuznetsova NA, Potapov MB, Panina KS, Antipova MD, Bokova AI (2020) Collembola of winter wheat fields in the Kaluga Region: conservation treatment versus conventional one. Sampling event dataset. 1.4. Moscow Pedagogical State University (MPGU). Release date: 2020-8-25. URL: <https://doi.org/10.15468/rv6g98>
- Lande R (1996) Statistics and partitioning of species diversity, and similarity among multiple communities. *Oikos* 76: 5-13. <https://doi.org/10.2307/3545743>
- Marsh J, Ewers R (2013) A fractal-based sampling design for ecological surveys quantifying  $\beta$ -diversity. *Methods in Ecology and Evolution* 4: 63-72. <https://doi.org/10.1111/j.2041-210x.2012.00256.x>

- Petersen H, Luxton M (1982) A comparative analysis of soil fauna populations and their role in decomposition processes. *Oikos* 39: 288-388. <https://doi.org/10.2307/3544689>
- Potapov M (2002) Synopses on Palaearctic Collembola. Isotomidae. Vol. 3. Senckenberg Museum of Natural History Görlitz, 603 pp.
- Potapov M, Kuznetsova N (2011) Methods of research of microarthropod communities: a manual for students and postgraduates. [Методы исследования сообществ микроартропод: пособие для студентов и аспирантов]. *Tovarishchestvo Nauchnykh Izdaniy KMK, Moskva*, 77 pp. [In Russian].
- Saraeva A, Potapov M, Kuznetsova N (2015) Different-scale distribution of Collembola in uniform ground cover: Sphagnum moss. *Entomological Review* 95 (5): 557-577. <https://doi.org/10.1134/S0013873815050012>
- Sterzynska M (1990) Communities of Collembola in natural and transformed soils of the linden-oak-hornbeam sites of the Mazovian Lowland. *Fragmenta Faunistica* 34 (11): 165-262. <https://doi.org/10.3161/00159301FF1990.34.11>
- Sterzynska M, Kuznetsova N (1997) Comparative analysis of dominant species in springtail communities (Hexapoda: Collembola) of urban greens in Moscow and Warsaw. *Fragmenta Faunistica (Warszawa)* 40 (2): 15-26. <https://doi.org/10.3161/00159301FF1997.40.2.015>
- Weigmann G, Kratz W (1987) Oribatid mites in urban zones of West Berlin. *Biology and Fertility of Soils* 3: 81-84. <https://doi.org/10.1007/BF00260583>
- Wiczorek J, et al. (2012) Darwin Core: An evolving community-developed biodiversity data standard. *PLOS One* 7 (1): e29715. <https://doi.org/10.1371/journal.pone.0029715>



Figure 1.  
Studied field without mineral fertilisers and pesticides. Photo by Nataliya Kuznetsova.



Figure 2.

Lime forest in Neskuchny Sad, Moscow. Photo by Anna Bokova.



Figure 3.  
Clear cutting area in surroundings of Lomovoe at present. Photo by Irina Amosova.





Figure 4.

Geographic coverage. **A.** study areas in the European part of Russia; **B.** sampling sites (red dots) near Lomovoe (Arkhangelsk Region); **C.** sampling sites (red dots) in Kaluga and Moscow Regions. Vector layers from [VMap0](#), administrative borders from OpenStreetMap via [gis-lab.info](#).