

## Table of contents

The dataset consists of the following tables:

Observation_standalone.csv	A table that can be used alone, since all necessary information of species and phases are decoded. Location is only specified by coordinates, altitude and distance to water.
Observation.csv	Observation data with coding that can be decoded by the tables below
Site.csv	List of observation sites, including coordinates
Parish.csv	List of the old geographical administration unit parish ("socken" in Swedish)
County.csv	List of the geographical administration unit county ("län" in Swedish)
Observer.csv	List of observers
Species.csv	List of defined objects for the observations, mainly plant and bird species, but also a set of agricultural activities
Phenophase.csv	List of defined phenological phases

## Resources

### Observation\_standalone.csv

The table lists the actual observations, together with the most important parameters needed for analyses of the dataset. More detailed information, e.g. site names and observers are excluded. The table columns are defined as described below.

Obs_id	The unique ID of each observation
Latitude	Latitude in decimal degrees North (WGS84)
Longitude	Longitude in decimal degrees East (WGS84)
Altitude	Altitude above sea level (m)
Distance_to_water	Distance to the sea or any of the tree biggest lakes in Sweden; Vänern, Vättern and Mälaren (m)
Species_code	Object code, as used in the table Species.csv
Species_SciName	Scientific name (if applicable)
Phase_code	Phenophase code, as used in the table Phenophase.csv
Phase_EngName	English description of the phase
Date	Observation date in the ISO format (YYYY-MM-DD)
Year	Year (extracted from Date)
Day_of_Year	Day number of the year from January 1 <sup>st</sup> (extracted from Date)
Quality	Quality, as defined by the observer or the digitizer (1=high, 0=low). Observers noted low quality when the date for the phase to occur was not well defined. Digitizers noted low quality when digits was hard to interpret in the form, or if a period was defined instead of a distinguished date.
Outlier	Outliers were identified by testing location/date/species/phase combinations to algorithms that can distinguish unrealistic dates for each combination. Only a subset could be tested. 1=outlier, 0=not outlier and NULL=not tested.

### Observation.csv

The table lists the actual observations, as they have been interpreted when digitizing the original forms. The table columns are defined as described below.

Obs_id	The unique ID of each observation
Site_code	The site code, as used in the Site.csv table
Observer_id	The identity code for the observer, as used in the Observer.csv table
Species_code	Code for the object that has been observed, as used in the table Species.csv
Phase_code	Code for the phenophase that has been observed for the object, as used in the table Phenophase.csv
Date	Observation date in the ISO format (YYYY-MM-DD)
Year	Year (extracted from Date)
Day_of_Year	Day number of the year from January 1 <sup>st</sup> (extracted from Date)
Quality	Quality, as defined by the observer or the digitizer (1=high, 0=low). Observers noted low quality when the date for the phase to occur was not well defined. Digitizers noted low quality when digits was hard to interpret in the form, or if a period was defined instead of a distinguished date.
Outlier	Outliers were identified by testing location/date/species/phase combinations to algorithms that can distinguish unrealistic dates for each combination. Only a subset could be tested. 1=outlier, 0=not outlier and NULL=not tested.

### Site.csv

The table lists the observation sites. Most sites are well defined places, like an estate, school or church, but in some places they are just defined as the town or village where the observer lived. The table columns are defined as described below.

Site_code	The site code, as used in the Observation.csv table
Site_name	The Swedish name of the site
Parish_code	The parish code, as used in the Parish.csv table
Latitude	Latitude in decimal degrees North (WGS84)
Longitude	Longitude in decimal degrees East (WGS84)
SWEREF99_N	North coordinate in the Swedish national grid SWEREF99 TM (EPSG 3006)
SWEREF99_E	East coordinate in the Swedish national grid SWEREF99 TM (EPSG 3006)
RT90_X	X (north) coordinate in the former Swedish national grid RT90 2,5 gon W (EPSG 3021)
RT90_Y	Y (east) coordinate in the former Swedish national grid RT90 2,5 gon W (EPSG 3021)
Altitude	Altitude above sea level (m)
Distance_to_water	Distance to the sea or any of the tree biggest lakes in Sweden; Vänern, Vättern and Mälaren (m)
pep_s_id	Site ID used in the PEP725 database (see <a href="http://www.pep725.eu">www.pep725.eu</a> )

### Parish.csv

The table lists the old geographical administration unit parish, called “socken” in Swedish. This administrative unit was in use at the time when the data was collected. Observation sites were sorted by county and parish. The table columns are defined as described below.

Parish_code	The parish code, as used in the Site.csv table
Parish_name	The Swedish name of the parish
County_code	The county code, as used in the County.csv table
CountryPart_code	Sweden is divided into three major parts: 1. Götaland 2. Svealand 3. Norrland

### County.csv

The table lists the Swedish counties, the geographical administration unit, called “län” in Swedish. The names are updated to the currently used names. Parishes were located by the county it currently belongs to, which in some cases can divert from the original, noted by the observer. The table columns are defined as described below.

County_code	The county code, as used in the Parish.csv table
County_name	The Swedish name of the county

### Observer.csv

The table lists the observers of the phenological observations, where it has been possible to read it out in the old documents. The registered observer is most probably the actual observer, but in some cases it could be expected that the observer is another, unknown person under the registered observer’s jurisdiction. The table columns are defined as described below. Note: For names that have not been able to distinguish, the name is set to “NULL”, “NN”, “?” or similar.

Observer_id	The identity code for the observer, as used in the Observation.csv table
Observer_name	The name of the observer
Site_code	The identity code of the site where the observer made his observations, as used in the Observation.csv and Site.csv tables

### Species.csv

The table defines the objects of the phenological observations, mainly species of plants and birds, but also of a few other organisms and agricultural activities. The table columns are defined as described below. Note: For objects where certain definitions are not applicable, the value is set to “NULL” (e.g. the scientific name of “hay”).

ObjectGroup_code	Subdivision code for objects 1. Plants and trees 2. Birds 3. Other animals than birds 4. Agricultural crops/diseases/activities
Species_code	Object code, as used in the table Observations.csv
Species_SwName	Common Swedish name of the object
Species_SciName	Scientific name (if applicable)

gss_id	Species code used in the PEP725 database (see <a href="http://www.pep725.eu">www.pep725.eu</a> ), if applicable
--------	---

### Phenophase.csv

The table defines the phases by which the phenological observations are made for. The phases are separately designed for each group of object (e.g. plants, birds or agricultural activities). The table columns are defined as described below. Note: For objects where certain definitions are not applicable, the value is set to “NULL”.

Phase_code	Phenophase code, as used in the table Observations.csv
Phase_SwName	Swedish description of the phase
Phase_EngName	English description of the phase
SWENPN_code	Corresponding phase code, used by the Swedish National Phenology Network (SWE-NPN) in the “Nature’s calendar” observation database (see <a href="http://www.naturenskalender.se">www.naturenskalender.se</a> ), if applicable
BBCH_code	Phase code, as used in the PEP725 database (see <a href="http://www.pep725.eu">www.pep725.eu</a> ), if applicable. The code is an extension of the BBCH protocol.
BBCH_description	Description of the phase, according to the PEP725 database