

ReadMe - V_heath_plant_species_composition_experimental

Dataset responsible: Kari Anne Bråthen and Hanna Böhner
(kari.brathen@uit.no, hanna.bohner@uit.no)

17.02.2021

1 Protocol

Plant species composition in all herbivore and large herbivore exclosures as well as control plots has been assessed following the COAT protocol ‘protocol_plant_measurements_heath_and_meadow_varanger’.

Large herbivore exclosures were set up and sampled for the first time in 2018. All herbivore exclosures were set up and sampled for the first time in 2019.

2 Description of the dataset

The dataset includes three different types of files and all files are saved as ;-separated txt-files:

- One data file per year (`_YEAR.txt`)
- One coordinate file with coordinates of all sites (`_coordinates.txt`)
- One auxiliary file with information about which sites are included in the study design (`_aux.txt`)

2.1 V_heath_plant_species_composition_experimental_YEAR.txt

These files contain presence/absence data of all vascular plant species. In addition, abundance data is included if available. Difficult species were registered on genus level. A species list and more information about how species data was recorded can be found in the protocol.

Example of the first rows of the data files:

```
sn_region    sn_locality    sn_section    sc_type_of_sites_ecological
1 varanger    vestre_jakobselv    bearylveaijohka    heath_near
2 varanger    vestre_jakobselv    bearylveaijohka    heath_near
3 varanger    vestre_jakobselv    bearylveaijohka    heath_near
      sn_site      sc_plot_treatment    sn_plot    t_year      t_date    v_observer
1 vj_be_hn_a    all_herbivore_exclosure    1    2018      <NA>      <NA>
2 vj_be_hn_a    all_herbivore_exclosure    1    2018      <NA>      <NA>
3 vj_be_hn_a      control    1    2018    2018-07-25      jr
      v_number_of_pins    v_species    v_plant_part    v_abundance    v_presence
1      NA    agr_mer    reproductive      NA      NA
2      NA    agr_mer    vegetative      NA      NA
3      3    agr_mer    reproductive      NA      0
      v_comment
1 all_herbivore_exclosures started in 2019
2 all_herbivore_exclosures started in 2019
3      <NA>
```

Description of the columns included in the data files:

Column name	Description	Possible values
sn_region	Study region	varanger
sn_locality	Locality (within region)	vestre_jakobselv
sn_section	Section (within locality)	bearalveaijohka, torvhaugdalen
sc_type_of_sites_ecological	Habitat type	heath_near
sn_site	Unique Site ID	e.g. vj_be_hn_a, vj_be_hn_e2, NA, NA
sc_plot_treatment	Plot treatment	all_herbivore_exclosure, control, large_herbivore_exclosure
sn_plot	Point frequency plot	1-16
t_year	Sampling year	e.g. 2019
t_date	Sampling date	YYYY-MM-DD
v_observer	Initials of observer	e.g. kab (Kari Anne Bråthen)
v_number_of_pins	Number of pins per PF plot	NA, 3, 24
v_species	Species or genus (Abbreviations)	e.g. bis_viv, car_sp
v_plant_part	Plant part	reproductive, vegetative, wood
v_abundance	Number of point intercepts	[numeric]
v_presence	Presence (1) or absence (0)	0, 1
v_comment	Comments	[character]

2.2 V_heath_plant_species_composition_experimental_coordinates.txt

This file contains the coordinates of all sites included in the study design. Coordinates are given in decimal degrees and UTM 33 (WGS 84).

Example of the first rows of coordinate files:

```
sn_site    e_dd    n_dd e_utm33 n_utm33
1 vj_to_hn_c 29.08530 70.30827 1025566 7861453
2 vj_to_hn_d 29.11366 70.30286 1026747 7861110
3 vj_be_hn_a 28.95137 70.28681 1021193 7857958
```

2.3 V_heath_plant_species_composition_experimental_aux.txt

This file contains further information about the dataset such as the years when sites were first included in the study design and when sites were excluded from the study design.

Example of the first rows of auxiliary-files:

	sn_region	sn_locality	sn_section	sn_site
1	varanger	vestre_jakobselv	bearalveaijohka	vj_be_hn_a
2	varanger	vestre_jakobselv	bearalveaijohka	vj_be_hn_a
3	varanger	vestre_jakobselv	bearalveaijohka	vj_be_hn_a
	sc_plot_treatment	year_first	year_last	v_comment
1	control	2018	NA	NA
2	large_herbivore_exclosure	2018	NA	NA
3	all_herbivore_exclosure	2019	NA	NA

* year_last is NA if the site is still included in the study design

3 Data cleaning and formatting

Rawdata is cleaned and formatted in three steps:

1. Data cleaning: All rawdata files entered in excel-templates are cleaned and saved as txt-files using the scripts `01_check_and_reformat_point_frequency_fieldsheets_all_herbivore_exclosures.R`, `01_check_and_reformat_point_frequency_fieldsheets_large_herbivore_exclosures.R` and `01_check_and_reformat_point_frequency_fieldsheets_observational.R`. The scripts check for correct spelling, correct format, outliers and missing observations, adds other necessary columns (e.g. region, locality and habitat) and saves the data as txt-files. Each file is processed separately and possible mistakes are corrected in the scripts.

In particular the scripts check for:

- **sn_site:** Correct spelling of all site names and missing observation. Missing observations will be included with NA for abundance.
- **t_date:** Correct format (yyyy-mm-dd), dates in other formats are reformatted.
- **v_observer:** Correct format (initials and lowercase letters), observer is reformatted e.g. if full names were used.
- **Species and functional group names:** Correct spelling, all species names will be converted to abbreviations (e.g. vac_myr).
- **v_abundance:** Outliers in abundance, 'x' will be replaced with 0.1, empty cells will be filled with 0 and weird entries (e.g. if there was a problem with the keyboard) will be corrected.
- **v_comment:** Correct spelling and format (lowercase letters and english). Comments are edited or translated if necessary.
- The columns **sn_region**, **sn_locality**, **sn_section**, **t_year** and **t_season** are added.

A comment is added if corrections go beyond simple typing mistakes and lead to differences between rawdata and cleaned data.

2. Data formatting: All cleaned files are compiled and formatted using the script `02_make_datafiles_from_point_frequency_experimental.R`. The script formats the data according to the requirements of the COAT dataprotal and produces one file for each dataset derived from observational plant measurements in heath and meadow sites:

- `V_meadow_vascular_plant_abundance_experimental_YEAR.txt`
(point intercepts on functional group level in meadow sites)
- `V_meadow_plant_species_composition_experimental_YEAR.txt`
(presence and abundance data of all vascular plant species in meadow sites)
- `V_meadow_ground_cover_experimental_YEAR.txt`
(abundance of for example Litter, Mosses, Lichen and small rodent activity in meadow sites)
- `V_meadow_shrub_height_experimental_YEAR.txt`
(height of thicket forming *Salix sp.* and *Betula nana* in meadow sites)
- `V_heath_vascular_plant_abundance_experimental_YEAR.txt`
(point intercepts on functional group level in heath sites)
- `V_heath_plant_species_composition_experimental_YEAR.txt`
(presence and abundance data of all vascular plant species in heath sites)
- `V_heath_ground_cover_experimental_YEAR.txt`
(abundance of for example Litter, Mosses, Lichen and small rodent activity in heath sites)
- `V_heath_shrub_height_experimental_YEAR.txt`
(height of thicket forming *Salix sp.* and *Betula nana* in heath sites)

3. Quality check: A final quality check is performed on each dataset. All variables are checked and observations are plotted together with the years before.