

# ReadMe - V\_heath\_plant\_species\_composition\_observational

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

22.01.2021

## 1 Protocol

Plant species composition has been assessed using point intercept frequency following the COAT protocol ‘protocol\_plant\_measurements\_heath\_and\_meadow\_varanger’.

### 1.1 Timing of sampling and changes in that

Plant species composition has been conducted at the peak of the growing season. Data has been collected from 2005 to 2008 and from 2019 onwards.

### 1.2 Spatial layout of sampling and changes in that

From 2005 onwards, plant measurements were conducted in two localities, Vestre Jakobselv and Komagdalen.

The sections ‘jakobselv’ and ‘komagdalen\_nedre’ have been included from 2005 to 2008 and the section ‘sandfjorddalen’ has been included since 2019.

From 2020 onwards, a new type of heath sites, which are located further away from the productive river valleys, have been included in the study design.

More detailed information about which sites were included in the study design can be found in the auxiliary file ‘V\_heath\_plant\_species\_composition\_observational\_aux.txt’.

From 2005 to 2008, plant measurements were conducted in 13 50x42 cm plots placed along 4 lines within a 15x15 m sampling quadrat at each site. Point intercept frequency was conducted using a 50x42 cm point frequency frame with 20 pins.

From 2009 onwards, plant measurements were conducted in 24 triangle plots (with a sidelength of 40 cm) placed along 3 lines within the 15x15 m sampling quadrats. Point intercept frequency was conducted using a triangle point frequency frame with 3 pins.

The changes in the study design (3 pins vs 20 pins and sampling early and late vs sampling in the middle of the growing season) have been compared and both design yielded comparable plant abundance estimates for the major functional groups (Soininen et al. (2018)).

### 1.3 Other changes in the protocol

Plant abundance has been registered on species level from 2005 to 2008. From 2019 onwards, presence of all vascular plant species is registered but abundance of difficult species (mainly grasses) is registered on functional group level.

## 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\_observational\_YEAR.txt

These files contain presence/absence data of all vascular plant species. In addition, abundance data is included if available. 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  komagdalen komagdalen_midtre          heath_near
2  varanger  komagdalen komagdalen_midtre          heath_near
3  varanger  komagdalen komagdalen_midtre          heath_near
      sn_site sn_plot t_year      t_date t_season v_observer v_number_of_pins
1 ko_km_hn_a      1   2005 2005-08-03  summer      ke           20
2 ko_km_hn_a      1   2005 2005-08-03  summer      ke           20
3 ko_km_hn_a      1   2005 2005-08-03  summer      ke           20
v_species v_plant_part v_abundance v_presence v_comment
1  agr_mer reproductive      0          0      <NA>
2  agr_mer  vegetative      0          0      <NA>
3  agr_sp  reproductive      0          0      <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)	komagdalen, vestre_jakobselv
sn_section	Section (within locality)	komagdalen_midtre, komagdalen_nedre, komagdalen_ovre, bearalveaijohka, torvhaugdalen, jakobselv, sandfjorddalen
sc_type_of_sites_ecological	Habitat type (heath near or far from productive river valleys)	heath_near, heath_far
sn_site	Unique Site ID	e.g. ko_km_hn_a, ko_km_hn_e, ko_ko_hn_c, vj_be_hn_c, vj_to_hn_c, vj_vj_hn_d, ko_sa_hn_f, vj_be_hn_d2, ko_km_hf_b, ko_ko_hf_b
sn_plot	PF plot (24 triangles in each quadrat)	1-24
t_year	Sampling year	e.g. 2019
t_date	Sampling date	YYYY-MM-DD
t_season	Sampling season	summer
v_observer	Initials of observer	e.g. kab (Kari Anne Bråthen)
v_number_of_pins	Number of pins per PF plot	20, 3
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\_observational\_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 ko_ko_hn_a 29.94898 70.32916 1056664 7871442
2 ko_ko_hn_b 29.96922 70.32893 1057411 7871603
3 ko_ko_hn_c 29.97894 70.33058 1057721 7871871
```

## 2.3 V\_heath\_plant\_species\_composition\_observational\_aux.txt

This file contains further information about the dataset such as old site names (for example used in raw data files before 2019) and 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 sn_site_old year_first
1 varanger komagdalen komagdalen_mitdre ko_km_hn_a      koh2a      2005
2 varanger komagdalen komagdalen_mitdre ko_km_hn_b      koh2b      2005
3 varanger komagdalen komagdalen_mitdre ko_km_hn_c      koh2c      2005
year_last v_comment
1      NA      NA
2      NA      NA
3      NA      NA
```

\* year\_last is NA if the site is still included in the study design

## 3 Data cleaning and formatting

From 2005 to 2008, rawdata has been cleaned by Virve Ravolainen. From 2009 to 2013, rawdata has been cleaned by Virve Ravolainen and Eeva Soinen and from 2014 to 2018 by Kari Anne Bråthen. All pre-cleaned files were formatted meeting the requirements of the COAT dataportal by Hanna Boehner.

From 2019 onwards, 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 script `01_check_and_reformat_point_frequency_fieldheets_observational.R`. The script checks 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 script.

In particular the script checks 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_rawdata_observational.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_observational_YEAR.txt`  
(point intercepts on functional group level in meadow sites)
- `V_heath_vascular_plant_abundance_observational_YEAR.txt`  
(point intercepts on functional group level in heath sites)
- `V_meadow_plant_species_composition_observational_YEAR.txt`  
(presence and abundance data of all vascular plant species in meadow sites)
- `V_heath_plant_species_composition_observational_YEAR.txt`  
(presence and abundance data of all vascular plant species) in heath sites)
- `V_meadow_ground_cover_observational_YEAR.txt`  
(abundance of for example Litter, Mosses, Lichen and small rodent activity in meadow sites)
- `V_heath_ground_cover_observational_YEAR.txt`  
(abundance of for example Litter, Mosses, Lichen and small rodent activity in heath sites)
- `V_tall_shrub_shrub_height_observational_YEAR.txt`  
(height of *Betula nana* and thicket forming *Salix sp.* in meadow sites)
- `V_heath_shrub_height_observational_YEAR.txt`  
(height of *Betula nana* and thicket forming *Salix sp.* in heath sites)
- `V_tall_shrub_thicket_edge_observational_YEAR.txt`  
(canopy height and extend of willo thickets in meadow plots in meadow 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.

## 4 Calculation of state variables

Point intercepts can be converted to biomass esatimates. Contact Kari Anne Bråthen for more information about the conversion factors.

## References

Soininen, Eeva M., John-Andre Henden, Virve T. Ravolainen, Nigel G. Yoccoz, Kari Anne Bråthen, Siw T. Killengreen, and Rolf A. Ims. 2018. “Transferability of Biotic Interactions: Temporal Consistency of Arctic Plant–rodent Relationships Is Poor.” *Ecology and Evolution* 8 (19): 9697–9711. doi:[10.1002/ece3.4399](https://doi.org/10.1002/ece3.4399).