ReadMe - V_heath_vascular_plant_abundance_observational

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

16.11.2020

1 Protocol

Plant abundance has been measured using point intercept frequency following the COAT protocol 'protocol_plant_measurements_heath_and_meaodw_varanger'.

1.1 Timing of sampling and changes in that

From 2005 to 2008 and from 2018 onwards, plant measurements were conducted once a year during peak growing season (from late July until beginning of August).

From 2009 to 2016, plant measurements were conducted twice a year, early in the growing season (early July) and late in the growing season (end of August and beginning of September).

In 2017, plant measurements were conducted only late in the growing season (end of August and beginning of September) due to a very late snow melt.

1.2 Spatial layout of sampling and changes in that

From 2005 onwards, plant measurements were conducted in 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 2009.

A third locality, Ifjordfjellet, has been included from 2009 to 2016.

A new type of heath sites which are located further away from the productive river valleys were included 2020.The new sites are located in the already existing sections (5 sites in komagdale_nedre, 5 sites in komagdalen_ovre, 5 sites in sandfjorddalen, 5 sites in bearalveaijohka and 4 sites in torvhaugdalen).

Some heath plots have been moved in 2018. More detailed information about which sites were included in the study design can be found in the auxialiary file 'V_heath_vascular_plant_abundance_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 frequeny 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 estaimates for the major functional groups (Soininen et al. (2018)).

1.3 Other changes in the protcol

Plant abundance has been registrerd on species level from 2005 to 2008 and from 2018 onwards but is grouped into the functional groups for plant abundance estimates. From 2009 to 2018, plant abundacne has been registred on functional group level. However, some species have always been registred to species level and *Vaccinium myrtillus* has been added to these species in 2020. A list with all functional groups and species can be found in section 3.1.

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_vascular_plant_abundance_observational_YEAR.txt

These files contain point intercept data of the following functional groups and species which can be converted to biomass estimates:

```
 [1] "Betula nana"                    "Broad leaved grasses"
 [3] "Calamagrostis phragmitoides"    "Deciduous dwarf shrubs"
 [5] "Deciduous vascular cryptogams" "Deschampsia cespitosa"
 [7] "Empetrum nigrum"                "Equisetum sp"
 [9] "Evergreen dwarf shrubs"         "Evergreen non woody"
[11] "Green willows"                  "Grey willows"
[13] "Hemiparasites"                  "Nardus stricta"
[15] "Narrow leaved grasses"          "Prostrate willows"
[17] "Rumex acetosa"                  "Salix herbacea"
[19] "Sedges and rushes"              "Small forbs"
[21] "Tall forbs"                     "Vaccinium myrtillus"
```

Example of the first rows of the data files:

```
  sn_region      sn_locality sn_section sc_type_of_sites_ecological    sn_site
1  varanger vestre_jakobselv  jakobselv                  heath_near vj_vj_hn_d
2  varanger vestre_jakobselv  jakobselv                  heath_near vj_vj_hn_d
3  varanger vestre_jakobselv  jakobselv                  heath_near vj_vj_hn_d
  sn_plot t_year     t_date t_season v_observer v_number_of_pins
1       1   2005 2005-07-27   summer         td               20
2       1   2005 2005-07-27   summer         td               20
3       1   2005 2005-07-27   summer         td               20
  v_functional_group v_plant_part v_abundance v_comment
1            bet_nan reproductive           0      <NA>
2            bet_nan   vegetative          39      <NA>
3            bet_nan         wood           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)	vestre_jakobselv, komagdalen, ifjordfjellet
sn_section	Section (within locality)	jakobselv, komagdalen_ovre, komagdalen_midtre, komagdalen_nedre, bearalveaijohka, torvhaugdalen, eastordalen, iesjohka, storelva, sandfjorddalen, giksjohka
sc_type_of_sites_ecological	Habitat type	heath_near, heath_far
sn_site	Unique Site ID	e.g. vj_vj_hn_d, ko_ko_hn_a, ko_ko_hn_e, ko_kn_hn_b, vj_be_hn_d, vj_to_hn_d, if_ie_hn_b, if_st_hn_c, ko_sa_hn_b, if_gi_hn_a
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, spring, fall
v_observer	Initials of observer	e.g. kab (Kari Anne Bråthen)
v_number_of_pins	Number of pins per PF plot	20, 3
v_functional_group	Functional group or species	bet_nan, broad_leaved_grasses, cal_phr, deciduous_dwarf_shrubs, deciduous_vascular_cryptogams, des_ces, emp_nig, equ_sp, evergreen_dwarf_shrubs, evergreen_non_woody, green_willows, grey_willows, hemiparasites, nar_str, narrow_leaved_grasses, prostrate_willows, rum_ace, sal_her, sedges_rushes, small_forbs, tall_forbs, vac_myr
v_plant_part	Plant part	reproductive, vegetative, wood
v_abundance	Number of point intercepts	[numeric]
v_comment	Comments	[character]

2.2 V_heath_vascular_plant_abundance_observational_coordinates.txt

This file contains the coordinates of all sites included in the study desgin. 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 if_ea_hn_a 27.36533 70.40762 959977.6 7858316
2 if_ea_hn_b 27.34243 70.42363 958774.6 7859895
3 if_ea_hn_c 27.34975 70.42428 959028.7 7860022
```

2.3 V_heath_vascular_plant_abundance_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 ifjordfjellet eastordalen if_ea_hn_a        ae1h       2009
2  varanger ifjordfjellet eastordalen if_ea_hn_b        ae2h       2009
3  varanger ifjordfjellet eastordalen if_ea_hn_c        ae3h       2009
  year_last v_comment
1      2016        NA
2      2016        NA
3      2016        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 Soininen 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 keybaord) 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.