

Supplementary information for the paper "Environmental conditions for alternative tree cover states in high latitudes" by B. Abis and V. Brovkin, 2016.

All the datasets used are included in the Data folder.

All the scripts for performing the analysis and reproducing the plots are in the Analysis folder.

All data and functions to load for python are in the Analysis/files_to_load_and_import folder.

In general, the scripts are meant for Ipython notebook, but python versions are also provided.

CDO 1.7.1 and NCO 4.4.8 were used for most of the preprocessing of the original datasets.

Command history is included in the meta-data of each file in Data. Additional processing, as in `canadian_fire_manipulation.R`, performed with RStudio Version 0.99.441. The rest of the analysis uses Python and IPython through notebooks. Python 2.7.10, IPython 4.0.1.

Tables showing results of the classification with detailed information about classes and bin division, and the correlation between variables, are reported in `Supplementary_Tables.pdf`.

Python scripts contain paths to datasets and files that should be updated according to instruction (see notes below). The scripts are numbered and contain the following:

- 0 filtering of the original datasets according to materials section (uses datasets in Data)
- 1 plots of the original datasets on the boreal area (uses datasets in Data)
- 2 Generalised Additive Models applied in all four regions with the entire datasets and with random 1000 gridcells samples (repetitively) (uses data and modules from `files_to_load_and_import`)
- 3--6 plots for phase-space analysis for every region (can be adapted to plot every combination of variables) (use data and modules from `files_to_load_and_import`)
- 7 exemplification of the bin division for the classification (uses data and modules from `files_to_load_and_import`)
- 8--11 classification algorithm in every region with creation of thresholds and masks to count only significant states (including plots per region) (use data and modules from `files_to_load_and_import`)
- 12 plot of the possible alternative states of the entire boreal area (projection is adaptable) (uses data and modules from `files_to_load_and_import`)
- 13 tests for multimodality and standard variation comparison

Original data consists of:

- Percentage tree cover fraction (TCF) from 0.05 degrees MODIS MOD44B V1 C5 2001--2010 product [1]
- Mean annual rainfall (MAR) from CRU TS3.22 Precipitation dataset 1998--2010 [2]
- Mean seasonal soil moisture (MSSM) from CPC Soil Moisture dataset 1998--2010 [3]
- Mean minimum 2m temperature (MTmin) from NCEP/NCAR Reanalysis 1998--2010 [4]
- Permafrost zonation index (PZI) from Global Permafrost Zonation Index Map [5]
- Fire frequency (FF) from GFED4 burned area dataset 1996--2012 and Canadian National Fire Database 1980--2014 [6,7]
- Growing degree days above 0 degrees C (GDD0) from NCEP Reanalysis (NMC initialised) 1998--2010 [4]
- Soil texture type (ST) from improved FAO soil type dataset [8]
- Mean thaw depth (MTD) from Arctic EASE-Grid Mean Thaw Depths [9]
- Surface elevation from Global 30-Arc-Second Elevation Dataset [10]
- Land cover type from Global Land Cover 2000 product (GLC2000) [11]

R packages installed are as follows:

- `silvermantest`-package (from the Philipps-Universität Marburg https://www.uni-marburg.de/fb12/stoch/forschung/rpackages/silvermantest_manual.pdf)
- `mgcv`-package (<https://cran.r-project.org/web/packages/mgcv/mgcv.pdf>)
- `data.table`-package (<https://cran.r-project.org/web/packages/data.table/vignettes/datatable-intro.pdf>)

Python packages installed are as follows:

- python: stable 2.7.12
- basemap (1.0.7)
- brewer2mpl (1.4.1)
- cdo (1.2.6)
- ipykernel (4.2.0)
- ipython (4.0.1)
- ipython-genutils (0.1.0)
- ipywidgets (4.1.1)
- jupyter (1.0.0)
- jupyter-client (4.1.1)
- jupyter-console (4.0.3)
- jupyter-core (4.0.6)
- matplotlib (1.5.0)
- nco (0.0.2)
- netCDF4 (1.2.1)
- nose (1.3.7)
- notebook (4.0.6)
- numpy (1.10.1)
- palettable (2.1.1)
- pandas (0.17.1)
- path.py (8.1.2)
- pickleshare (0.5)
- Pillow (2.7.0)
- py (1.4.31)
- rpy2 (2.7.4)
- scikit-learn (0.17)
- scipy (0.16.1)
- seaborn (0.6.0)
- setuptools (18.7.1)
- sklearn (0.0)
- snakeviz (0.4.0)
- sympy (0.7.6.1)

Data references:

1. John R.G. Townshend, Mark Carroll, Charlene DiMiceli, Robert Sohlberg, Matthew Hansen, and Ruth DeFries. Vegetation continuous fields mod44b, 2010 percent tree cover, collection 5, version 1, 2010. URL <http://icdc.zmaw.de>. University of Maryland, College Park, Maryland, downloaded 08/02/2013, provided on 0.05 degree Climate Modelling Grid in NetCDF by the Integrated Climate Data Center (ICDC) University of Hamburg, Hamburg, Germany.
2. Ian Harris, Philip D. Jones, Timothy J. Osborn, and David H. Lister. Cru ts3.22: Climatic research unit (cru) time-series (ts) version 3.22 of high resolution gridded data of month-by-month variation in climate (jan. 1901- dec. 2013), 2014. URL <http://dx.doi.org/10.5285/18BE23F8-D252-482D-8AF9-5D6A2D40990C>. University of East Anglia Climatic Research Unit. NCAS British Atmospheric Data Centre, 24 September 2014.
3. Huug van den Dool, Jin Huang, and Yun Fan. Performance and analysis of the constructed analogue method applied to U.S. soil moisture over 1981-2001. J. Geophys. Res., 108(D16): 8617, 2003. ISSN 0148-0227. doi: 10.1029/2002jd003114. URL <http://dx.doi.org/10.1029/2002JD003114>. CPC Soil Moisture data provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at <http://www.esrl.noaa.gov/psd/>.
4. Kalnay et al. The ncep/ncar 40-year reanalysis project, 1996. NCEP Reanalysis data provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at <http://www.esrl.noaa.gov/psd/>.
5. Stephan Gruber. Derivation and analysis of a high-resolution estimate of global permafrost zonation. The Cryosphere, 6(1):221–233, 2012. ISSN 1994-0424. doi: 10.5194/tc-6-221-2012. URL <http://www.the-cryosphere.net/6/221/2012/>.
6. Louis Giglio, James T. Randerson, and Guido R. van der Werf. Analysis of daily, monthly, and annual burned area using the fourth-generation global fire emissions database (gfed4). Journal of Geophysical Research: Biogeosciences, 118(1):317–328, 2013. ISSN 2169-8961. doi: 10.1002/jgrg.20042. URL <http://dx.doi.org/10.1002/jgrg.20042>.
7. Canadian Forest Service. Canadian national fire database - agency fire data, 2014. Natural Resources Canada, Canadian Forest Service, Northern Forestry Centre, Edmonton, Alberta <http://cwfis.cfs.nrcan.gc.ca/ha/nfdb>.
8. Stefan Hagemann and Tobias Stacke. Impact of the soil hydrology scheme on simulated soil moisture memory. Climate Dynamics, 44(7):1731–1750, 2014.
9. Tingjun Zhang, James McCreight, and Roger G. Barry. Arctic ease-grid freeze and thaw depths, 1901 - 2002, version 1, 2006. Boulder, Colorado USA. NSIDC: National Snow and Ice Data Center.
10. U.S. Geological Survey. Global 30 arc-second elevation (gtopo30), 1996. U.S. Geological Survey, Sioux Falls, South Dakota.
11. GLC2000 database. The global land cover map for the year 2000, 2003. URL <http://www-gem.jrc.it/glc2000>. European Commission Joint Research Centre.