**Supplemental Material**

# Example application: Vegetation greenness trends for a landscape on Disko Island

Here we provide an example analysis of interannual changes in vegetation greenness from 2000 to 2020 across a ~4 km2 study area on Disko Island off the western coast of Greenland (Figure 2). The study area (approximate center 69.27°N, 53.46°W) is located on the eastern slopes of the Blæsedalen valley just east of Qeqertarsuaq (Godhavn). The close proximity of the valley to the University of Copenhagen’s Arctic Station has made the area subject to much ecological and geological research, including multiple long-term monitoring projects and experiments (https://arktiskstation.ku.dk). Climatically, the site lies within the transition zone between the low and high Arctic, with basaltic soils on discontinuous permafrost (Xu et al. 2021) covered by erect dwarf shrub tundra (Walker et al. 2002). We characterize annual maximum vegetation greenness using the Normalized Difference Vegetation Index (NDVImax) derived from Landsat satellite observations. Landsat NDVImax relates to vegetation productivity and aboveground biomass in tundra ecosystems (Johansen and Tømmervik 2014, Berner et al. 2018, Berner et al. 2020). Here, we focus on the period from 2000 to 2020 because there were limited Landsat data available prior to 2000 in this region, as shown below. We provide the scripts associated with this example as supplemental files and in this section guide the reader through the analysis code with example output figures and tables that are generated by the *LandsatTS* functions (excluding Figure 2).

A picture containing text

Description automatically generated

Figure 2. (a) Relative changes in Landsat annual maximum NDVI (NDVImax) from 2000 to 2020 across the study area on Disko Island. (b) Location of Disko Island off the western coast of Greenland. Figure created using QGIS (v3.20; QGIS.org 2021). Background imagery from Google Satellite © 2022 CNES / Airbus used with fair use permission. Underlying digital elevation model from the U.S. National Snow and Ice Data Center (Howat et al. 2014, Howat et al. 2015).

*Part 1: Export Landsat time series from Google Earth Engine*

First the user needs to export Landsat time series for sample locations in the study area using GEE (Code Box 1). For this they need to prepare the environment, set the boundaries of the study area and then retrieve the Landsat pixel center coordinates using the *lsat\_get\_pixel\_centers()* function. Next, the Landsat records are exported for the pixel center locations using *lsat\_export\_ts()*. Here, we choose to export only Landsat observations between day of year 152 (beginning of June) and 273 (end of September). The user then waits for GEE to finish the exports. Progress can be monitored using the GEE task manager in the web browser (<https://code.earthengine.google.com/tasks>) or on the R console, using the *ee\_monitoring()* function provided by *rgee*. For the example, it took ~2 days to export the 19 files (totaling ~692 MB) associated with this example analysis. The CSV files containing the raw exports then need to be copied from the user’s Google Drive to the local machine that will carry out the subsequent processing using *LandsatTS*. The files can be copied manually or using the *ee\_drive\_to\_local()* function provided by *rgee*. Once the records are available locally, they need to be cleaned and processed into vegetation index time series as detailed in the next section.

Code Box 1: Export Landsat time series from Google Earth Engine

# Load required R packages

require(lsatTS)

require(rgee)

require(sf)

require(ggplot2)

require(*data.table*)

# Initialize Google Earth Engine

ee\_Initialize()

# Create sf polygon of the study area

aoi.poly <- st\_polygon(list(matrix(

c(-332950,-2243300,

-334950,-2243300,

-334950,-2245300,

-332950,-2245300,

-332950,-2243300),

ncol = 2,

byrow = T)))

*# Transform polygon to WGS84 lat long*

aoi.poly <- aoi.poly %>%

st\_sfc(crs = 3413) %>%

st\_transform(crs = 4326) %>%

st\_as\_sf()

# Get the central coordinates for each of the 4557 Landsat pixels in study area

aoi.pts <- lsat\_get\_pixel\_centers(aoi.poly)

# Export summer Landsat surface reflectance measurements for each pixel to a folder

# called “earth\_engine/lsat\_disko” on the user’s Google Drive.

lsat\_export\_ts(

pixel\_coords\_sf = aoi.pts,

startJulian = 152,

endJulian = 273,

prefix = 'disko',

drive\_export\_dir = 'earth\_engine/lsat\_disko')

## Part 2: Derive vegetation greenness time series from the raw Landsat data

To derive the vegetation greenness time series from the raw exports of Landsat time series, the records first need to be imported to R as a *data.table* object, re-formatted using *lsat\_general\_prep()* and cleaned with *lsat\_clean\_data()* to filter out clouds, snow, and water, as well as radiometric and geometric errors (Code Box 2). For the study area on Disko Island, *lsat\_clean\_data()* removed 1,817,683 of 2,452,693 observations (74.11%) in the data cleaning process. The availability of Landsat observations for all point locations (“sample.ids”) in the remaining dataset can then be visualized using *lsat\_summarize\_data\_avail*(). In the case of the pixel centers across the study area on Disko Island, the number of observations is poor before the year 2000, as highlighted by the graph that is automatically generated by the function (Figure 3). Therefore, we later limit the analysis of vegetation greenness to the years between 2000 and 2020. Finally, the NDVI is calculated using the *lsat\_calc\_spec\_index()*. The dataset is then ready for the sensor cross-calibration and phenological modelling.

Code Box 2: Derived vegetation greenness time series from the raw Landsat data

# Import CSV exported with GEE as *data.table*

data.files <- list.files(‘~/earth\_engine/lsat\_disko’, full.names = T)

lsat.dt <- do.call("rbind", lapply(data.files, fread))

# (Re-)format the imported raw data

lsat.dt <- lsat\_general\_prep(lsat.dt)

# Clean data by filtering clouds, snow, and water, as well as radiometric and geometric errors.

lsat.dt <- lsat\_clean\_data(lsat.dt)

# Summarize the availability of Landsat data for each pixel

lsat\_summarize\_data\_avail(lsat.dt)

# Compute the Normalized Difference Vegetation Index (NDVI)

lsat.dt <- lsat\_calc\_spec\_index(lsat.dt, si = 'ndvi')

Chart, scatter chart

Description automatically generated

Figure 3. Availability of quality screened Landsat observations across years for sample locations in the study area on Disko Island as returned by the *lsat\_summarize\_data\_avail()* function. Summaries are based on observations acquired between day of year 152 (beginning of June) and 273 (end of September). Note the limited availability of observations before the year 2000. Lines with points denote median counts while shaded error bands encompass the 2.5th to 97.5th percentiles of counts among sample locations.

*Part 3: Cross-sensor calibration and phenological modelling*

The derived NDVI time series need to be calibrated across the different Landsat sensors, and then NDVImax estimated using the phenological modelling approach (Code Box 3). We start by cross-calibrating the time series using *lsat\_calibrate\_rf()*. As the number of observations in the Disko Island dataset is too small to calibrate the random forest models, we use the pre-processed dataset of high latitude observations included with *LandsatTS*. The function saves the models in a specified output directory and generates a series of graphs (Figure 4) and tabular data (Table 3) that help with evaluating model performance. As desired, the calibration reduced the median bias between the Landsat 7 observations and the Landsat 5 and 8 observations visually (Figure 4) and statistically (Table 3). Next, as a step towards estimating annual NDVImax, we fit phenological models to the calibrated NDVI time series using *lsat\_fit\_phenological\_curves()*. The function automatically returns a figure with Landsat observations and fitted phenological curves for nine random sample locations in the dataset (Figure 5). Each phenological curve characterizes the seasonal progression of NDVI using observations pooled over a multi-year period (here an 11 year moving window) and should be smooth and hump-shaped. Beware of phenological curves with long straight lines that could suggest inadequate seasonal distribution of data used when fitting the curves. Once the models are fitted, the summary statistics (including the estimated NDVImax) are extracted using *lsat\_summarize\_growing\_seasons()*. The *lsat\_evaluate\_phenological\_max()* can be used to output a figure that allows for visually assessing the performance of modelled NDVImax (Figure 6). In the case of this Disko Island dataset, modeled estimates of NDVImax tendto be biased slightly low (~1%) when only one or two observations are available from a growing season (Figure 6), yet there were rarely such few observations during the period from 2000 to 2020 (Figure 3). The final step following the cross-calibration and phenological modelling is the time series analysis.

Code Box 3: Cross-calibration and phenological modelling

# Cross-calibrate NDVI among sensors using random forest models

# Outputs in Figure 4 and Table 3.

Lsat.dt <- lsat\_calibrate\_rf(

lsat.dt,

band.or.si = ‘ndvi’,

train.with.highlat.data = T,

outdir = ‘output/ndvi\_xcal\_smry/’,

overwrite.col = T)

# Fit phenological models (cubic splines) to time series at

# each sample location (Figure 5)

lsat.pheno.dt <- lsat\_fit\_phenological\_curves(lsat.dt, si = ‘ndvi’)

# Summarize spectral characteristics for each growing season

lsat.gs.dt <- lsat\_summarize\_growing\_seasons(lsat.pheno.dt, si = ‘ndvi’)

# Evaluate the estimates of annual maximum NDVI (Figure 6)

lsat.eval.dt <- lsat\_evaluate\_phenological\_max(lsat.pheno.dt, si = ‘ndvi’)

Chart

Description automatically generated

Figure 4. Relationships between Landsat 7 NDVI and both (a) Landsat 5 NDVI and (b) Landsat 8 NDVI using (left panels) original data and (right panels) data that were cross-calibrated with random forest models. Each point is a sample location from the Arctic – Boreal domain where there were temporally overlapping measurements from pairs of Landsat satellites. Orange diagonal lines depict 1:1 relationships. Model performance metrics are provided in Table 3. Note that cross-calibration substantially reduces biases between sensors but does increase scatter.

Table 3. Summary of original biases, performance of random forest models for cross-sensor calibration, and post-calibration biases in NDVI between Landsat 7 ETM+ and either Landsat 5 TM or Landsat 8 OLI. Error metrics were derived internally by the random forest using out-of-bag (OOB, i.e., withheld) data and further assessed using cross-validation, which yielded nearly identical results albeit with further information on post-calibration biases.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **Satellite**  **sensor** | **Original Data** | |  | **OOB Error Metrics** | | |  | **Cross-Validated Error Metrics** | | | | | | **Median**  **bias** | **Median**  **% bias** |  | **r2** | **RMSE** | **N** |  | **r2** | **RMSE** | **N** | **Median**  **bias** | **Median**  **% bias** | | Landsat 5 TM | -0.04 | -6.1 |  | 0.98 | 0.03 | 4315 |  | 0.98 | 0.03 | 1438 | +0.001 | +0.1 | | Landsat 8 OLI | +0.03 | +4.6 |  | 0.97 | 0.03 | 4881 |  | 0.97 | 0.03 | 1627 | -0.001 | -0.1 | |  | |  | | |  | | | | |
|  |  |  |  |  |  |  |  |  |  |

A picture containing calendar

Description automatically generated

Figure 5. Seasonal progression of Landsat NDVI and phenological curves for nine random sample locations from the study area on Disko Island. Each point is an observation sorted by the day of year it was acquired and colored by the year of acquisition. Each phenological curve was fit to observations pooled over an 11-year window centered on each focal year.

Chart, box and whisker chart

Description automatically generated

Figure 6. Raw estimates of annual maximum NDVI (NDVImax) are biased low when few Landsat observations are available from a given growing season, whereas phenologically modeled estimates of NDVImax are minimally impacted by the availability of observations. The figure summarizes how raw and modeled estimates of NDVImax differ from observed NDVImax based on number of observations, as determined using *lsat\_evaluate\_phenological\_max().*

## Part 3: Analyze vegetation greenness time series

Finally, the trend in the NDVImax across years for each sample location (pixel center) in our study area on Disko Island is calculated using the *lsat\_calc\_trend()* function (Code Box 4). Note how we use the “yrs” argument to restrict the time series analysis to the years between 2000-2020 to avoid using the low number of observations in the record prior to the turn of the millennium. Figure 7 shows a histogram of precent change in NDVImax across the study area and a time series of annual mean NDVImax by trend category, both of which are generated by the function. These figures indicate extensive browning across the study area in recent decades*.*

Code Box 4: Analyze vegetation greenness time series

# Compute temporal trend in NDVImax (Figure 7)

lsat.trend.dt <- lsat\_calc\_trend(lsat.gs.dt, si = 'ndvi.max', yrs = 2000:2020)

Chart, histogram

Description automatically generated

Figure 7. (a) Histogram of relative change in Landsat NDVImax from 2000 to 2020 among sample locations in the study area on Disko Island. Landsat NDVImax decreased (browned) across much of the study area over the past two decades. (b) Annual mean Landsat NDVImax from 2000 to 2020 for sample locations grouped by their concomitant temporal trend. Trends were assessed for each sample location by removing temporal autocorrelation and then applying a Mann-Kendall trend test. Error bands depict ±1 standard error.

*Results from the example study*

This example analysis showed that from 2000 to 2020, annual maximum vegetation greenness (i.e., NDVImax) systematically decreased (α = 0.10; browned) across 51% of the study area on Disko Island, whereas vegetation greenness systematically increased (α = 0.10; greened) across 3% of this study area (Figure 2a and 8). There were no systematic changes across the remaining 46% of the study area (Figure 2a and 8). Overall, vegetation greenness decreased by an average of 5.7 ± 8.4% (± 1 SD) during this period. The predominance of browning in this study area contrasts with widespread greening in the Arctic (Myers-Smith et al. 2020, Frost et al. 2021), where Landsat observations indicate that average Arctic vegetation greenness increased 3.9% from 2000 to 2020 (Berner et al. 2020, Mekonnen et al. 2021). Nevertheless, browning in this study area is broadly consistent with findings from recent pan-Arctic analyses using Landsat (Berner et al. 2020) and MODIS (Frost et al. 2021) satellite data that show regional browning in southwestern Greenland. The causes of browning in southwestern Greenland warrant further investigation but are potentially linked to hotter and drier conditions suppressing shrub and other vegetation growth and, in some areas, to defoliation from moths (*Eurois occulta*) or browsing by muskoxen (*Ovibos moschatus*) (Forchhammer 2017, Gamm et al. 2018, Prendin et al. 2020). This analysis demonstrates a general workflow that can be used to not only explore long-term changes in vegetation greenness across focal landscapes, but also to perform sample-based analyses across large geographic domains.

# References

Berner, L. T., P. Jantz, K. D. Tape, and S. J. Goetz. 2018. Tundra plant aboveground biomass and shrub dominance mapped across the North Slope of Alaska. Environmental Research Letters **13**:035002.

Berner, L. T., R. Massey, P. Jantz, B. C. Forbes, M. Macias-Fauria, I. H. Myers-Smith, T. Kumpula, G. Gauthier, L. Andreu-Hayles, B. Gaglioti, P. J. Burns, P. Zetterberg, R. D'Arrigo, and S. J. Goetz. 2020. Summer warming explains widespread but not uniform greening in the Arctic tundra biome. Nature communications **11**:4621.

Forchhammer, M. 2017. Sea-ice induced growth decline in Arctic shrubs. Biology Letters **13**.

Frost, G. V., M. J. Macander, U. S. Bhatt, L. T. Berner, J. W. Bjerke, H. E. Epstein, B. C. Forbes, S. J. Goetz, M. J. Lara, T. Park, G. K. Phoenix, S. P. Serbin, H. Tømmervik, D. A. Walker, and D. Yang. 2021. Tundra greenness [in “State of the Climate in 2020”]. Bulletin of the American Meteorological Society **102**:S290–S292.

Gamm, C. M., P. F. Sullivan, A. Buchwal, R. J. Dial, A. B. Young, D. A. Watts, S. M. Cahoon, J. M. Welker, and E. Post. 2018. Declining growth of deciduous shrubs in the warming climate of continental western Greenland. Journal of Ecology **106**:640-654.

Howat, I., A. Negrete, and B. Smith. 2015. MEaSUREs Greenland Ice Mapping Project (GIMP) Digital Elevation Model, Version 1. NASA National Snow and Ice Data Center Distributed Active Archive Center. doi: <https://doi.org/10.5067/NV34YUIXLP9W>. [2021-11-23], Boulder, Colorado USA.

Howat, I. M., A. Negrete, and B. E. Smith. 2014. The Greenland Ice Mapping Project (GIMP) land classification and surface elevation data sets. The Cryosphere **8**:1509-1518.

Johansen, B., and H. Tømmervik. 2014. The relationship between phytomass, NDVI and vegetation communities on Svalbard. International Journal of Applied Earth Observation and Geoinformation **27**:20-30.

Mekonnen, Z. A., W. J. Riley, L. T. Berner, N. J. Bouskill, M. S. Torn, G. Iwahana, A. L. Breen, I. H. Myers-Smith, M. G. Criado, Y. Liu, E. S. Euskirchen, S. J. Goetz, M. C. Mack, and R. F. Grant. 2021. Arctic tundra shrubification: a review of mechanisms and impacts on ecosystem carbon balance. Environmental Research Letters **16**:053001.

Myers-Smith, I. H., J. T. Kerby, G. K. Phoenix, J. W. Bjerke, H. E. Epstein, J. J. Assmann, C. John, L. Andreu-Hayles, S. Angers-Blondin, P. S. A. Beck, L. T. Berner, U. S. Bhatt, A. D. Bjorkman, D. Blok, A. Bryn, C. T. Christiansen, J. H. C. Cornelissen, A. M. Cunliffe, S. C. Elmendorf, B. C. Forbes, S. J. Goetz, R. D. Hollister, R. de Jong, M. M. Loranty, M. Macias-Fauria, K. Maseyk, S. Normand, J. Olofsson, T. C. Parker, F.-J. W. Parmentier, E. Post, G. Schaepman-Strub, F. Stordal, P. F. Sullivan, H. J. D. Thomas, H. Tømmervik, R. Treharne, C. E. Tweedie, D. A. Walker, M. Wilmking, and S. Wipf. 2020. Complexity revealed in the greening of the Arctic. Nature Climate Change **10**:106-117.

Prendin, A. L., M. Carrer, M. Karami, J. Hollesen, N. Bjerregaard Pedersen, M. Pividori, U. A. Treier, A. Westergaard-Nielsen, B. Elberling, and S. Normand. 2020. Immediate and carry-over effects of insect outbreaks on vegetation growth in West Greenland assessed from cells to satellite. Journal of Biogeography **47**:87-100.

QGIS.org. 2021. QGIS Geographic Information System. QGIS Association. <http://www.qgis.org>.

Walker, D. A., W. A. Gould, H. A. Maier, and M. K. Raynolds. 2002. The Circumpolar Arctic Vegetation Map: AVHRR-derived base maps, environmental controls, and integrated mapping procedures. International Journal of Remote Sensing **23**:4551-4570.

Xu, W., A. Prieme, E. J. Cooper, M. A. Mörsdorf, P. Semenchuk, B. Elberling, P. Grogan, and P. L. Ambus. 2021. Deepened snow enhances gross nitrogen cycling among Pan-Arctic tundra soils during both winter and summer. Soil Biology and Biochemistry **160**:108356.