Package 'npphen'

June 3, 2022

Type Package

Title Vegetation Phenological Cycle and Anomaly Detection using Remote
Sensing Data

Version 1.5.2

Date 2022-05-05

Author Roberto O. Chávez [cre, aut], Sergio A. Estay [cre, aut], José A. Lastra [cre,ctb] & Carlos G. Riquelme [ctb]

Maintainer José A. Lastra <jose.lastra@pucv.cl>

Description Calculates phenological cycle and anomalies using a non-parametric
approach applied to time series of vegetation indices derived from remote sensing data
or field measurements. The package implements basic and high-level functions for
manipulating vector data (numerical series) and raster data (satellite derived products).
Processing of very large raster files is supported. For more information, please check
the following paper:
Estay, S., Chávez, R.O. (2018) <doi:10.1101/301143>.

License GPL (>= 3)

Depends R (>= 3.6.0), raster, ks

Imports methods, graphics, stats, rgdal, lubridate, snow

Suggests rgeos (>= 0.3-8),tcltk

Encoding UTF-8

RoxygenNote 7.1.2

NeedsCompilation no

LazyData true

Repository CRAN

Date/Publication 2022-06-03 21:50:02 UTC

R topics documented:

Index   **16**

Bdesert_stack   *Blooming Desert Northern Chile*

Description

MODIS NDVI RasterStack of the Atacama Desert, Northern Chile, combining MOD13Q1 & MYD13Q1
between 2000.02.18 and 2021.06.26, cleaned using the detailed QA band. Blooming deserts are rare
greening events to be seen as positive NDVI anomalies in Atacama

Usage

```
Bdesert_stack
```

Format

RasterStack EPSG:32719

Source

https://lpdaac.usgs.gov/

ExtremeAnom   *ExtremeAnom*

Description

Based on the annual reference frequency distribution of a vegetation index time series (e.g. a nu-
meric vector of NDVI), calculates anomalies and how extreme these anomalies are (rfd position
ranging from 0 to 100)

Usage

```
ExtremeAnom(x, dates, h, refp, anop, rge, output = "both", rfd = 0.9)
```

Arguments

x	Numeric vector. A time series of a vegetation index (e.g. LAI, NDVI, EVI) or any other variable with seasonal behavior. The code has been optimized to work with integer values. Please re-scale the input values if necessary (e.g. NDVI ranging from 0.0000 to 1.0000, multiply by 10,000
dates	A date vector. The number of dates must be equal to the number of "x" values (numeric input vector).
h	Numeric. Indicates the geographic hemisphere. This argument defines the starting date of the growing season. h=1 for the Northern Hemisphere (season starting on January 1st), h=2 for the Southern Hemisphere (season starting on July 1st).
refp	Numeric vector with the correlative number of dates to be used as reference period. For example, refp = c(1:388) for MODIS Vegetation Index 16-days composites MOD13Q1 (18/02/2000 – 18/12/2016)
anop	Numeric vector with the correlative number of dates for the period in which the anomalies and rfd position (how extreme the anomalies are) will be calculated. For example anop = c(389:411) for the complete 2017 of MODIS Vegetation Index 16-days composites MOD13Q1 (01/01/2017 – 19/12/2017). anop y refp can be overlapped.
rge	Numeric vector with two values setting the minimum and maximum values of the response variable (e.g. NDVI) used in the analysis. We suggest the use of theoretically based limits. For example in the case of MODIS NDVI or EVI, it ranges from 0 to 10,000, so rge = c(0,10000)
output	Character string. Defines the output values. 'both' (default) returns both anomalies and rfd position together as a single numeric vector, 'anomalies' returns only anomalies, 'rfd' returns only rfd values (how extreme the anomalies are) and 'clean' returns only extreme anomalies, i.e. anomalies at which a given rfd is overpass (e.g. 0.90). This critical threshold is set by the users using the rfd argument.
rfd	Numeric. This argument only applies when the argument output='clean'. It defines the percentile (from 0 to 0.99) of the reference frequency distribution, for which anomalies are not flagged as extreme anomalies. For example, if 'rfd = 0.90' only anomalies falling outside the '0.90 rfd' (default value) will be flagged as extreme anomalies while the rest will be neglected (NA values). Please notice that 'rfd = 0.90' implies that the 5% of the most extreme positive and 5% of the most extreme negative anomalies will be considered.

Details

Calculates anomalies and a probabilistic measure of how extreme the anomalies are using a numeric vector of vegetation canopy greenness, e.g. satellite based Normalized Difference Vegetation Index (NDVI) or Enhanced Vegetation Index (EVI). For this purpose, it divides the time series (numeric vector) of vegetation greenness into 2: the reference period, from which the annual phenological reference is calculated (same as Phen function), and the observation period, for which we want to calculate anomalies. This anomalies can be filtered by the position of the observation within the historical rfd. Users can, for example, set 'rfd = 0.95' to consider only anomalies that outside the 95% rfd of historical records.

Value

numerical vector

See Also

ExtremeAnoMap

Examples

```
library(lubridate)

## Testing with the MegaDrought_stack from Central Chile (NDVI),
#showing extreme negative anomalies (browning)##

# Load data
#RasterStack
data("MegaDrought_stack")
#Dates
data("modis_dates")

# Extracting a time series from a particular pixel
md_pixel <- cellFromXY(MegaDrought_stack,c(313395,6356610))
md_pixelts <- as.numeric(MegaDrought_stack[md_pixel])
plot(modis_dates,md_pixelts, type='l')

# Anomaly detection for the given pixel
anomRFD_MD <- ExtremeAnom(x = md_pixelts,dates = modis_dates,
h = 2,refp = c(1:423), anop = c(884:929),rfd = 0.9,output = 'both',rge = c(0,10000))

#Basic plot

selection <- which(anomRFD_MD[47:92] > 90)

barplot(names=format.Date(modis_dates[884:929],format='%Y-%m'),
anomRFD_MD[1:46],col= ifelse(anomRFD_MD[1:46] < 0,"red","blue"),
main ='Anomalies whit rfd > 0.90')
abline(h=0)

## Testing with the Bdesert_stack from the Atacama Desert, Northern Chile (NDVI),
#showing extreme positive anomalies (greening)##

# Load data
#RasterStack
data("Bdesert_stack")
#Dates
data("modis_dates")

# Extracting a time series from a particular pixel
bd_pixel<-cellFromXY(Bdesert_stack,c(286638,6852107))
```

```
bd_pixelts<-as.numeric(Bdesert_stack[bd_pixel])
plot(modis_dates,bd_pixelts, type = 'l')

# Anomaly detection for the given pixel
anomRFD_BD <- ExtremeAnom(x = bd_pixelts,dates = modis_dates,
h = 2,refp = c(1:423), anop = c(723:768),rfd = 0.9,output = 'both',rge = c(0,10000))

#Basic plot

selection <- which(anomRFD_BD[47:92] > 95)

#basic plot
barplot(names=format.Date(modis_dates[723:768],format='%Y-%m'),
anomRFD_BD[1:46],col= ifelse(anomRFD_BD[1:46] < 0,"red","blue"),
main ='Anomalies whit rfd > 0.95')
abline(h=0)
```

ExtremeAnoMap   *ExtremeAnoMap*

Description

Based on the annual reference frequency distribution of a vegetation index time series (e.g. a raster stack of NDVI), calculates anomalies and how extreme these anomalies are (rfd position ranging from 0 to 100).

Usage

```
ExtremeAnoMap(s, dates, h, refp, anop, rge, output = "both", rfd = 0.9,
  nCluster, outname, format, datatype)
```

Arguments

s	Raster stack. A time series of a vegetation index (e.g. LAI, NDVI, EVI) or any other variable with seasonal behavior.The code has been optimized to work with integer values. Please re-scale the input raster stack if necessary (e.g. NDVI ranging from 0.0000 to 1.0000, multiply by 10,000).
dates	A date vector. The number of dates must be equal to the number of layers of raster stack.
h	Numeric. Indicates the geographic hemisphere to define the starting date of the growing season. h=1 if the vegetation is in the Northern Hemisphere (season starting at January 1st), h=2 if it is in the Southern Hemisphere (season starting at July 1st).
refp	Numeric vector with the correlative number of dates to be used as reference period. For example, refp = c(1:393) for MODIS Vegetation Index 16-days composites (18/02/2000 – 06/06/2017)

anop	Numeric vector with the correlative number of dates for the period in which the anomalies will be calculated. For example refp = c(21:43) for the first complete year for MODIS Vegetation Index 16-days composites (01/01/2001 – 19/12/2001). anop y refp can be overlapped
rge	Numeric vector containing minimum and maximum values of the response variable used in the analysis. We suggest the use of theoretically based limits. For example in the case of MODIS NDVI or EVI, it ranges from 0 to 10,000, so rge = c(0,10000).
output	Character string. Defines the output values. 'both' (default) returns both anomalies and rfd position together as a single numeric vector, 'anomalies' returns only anomalies, 'rfd' returns only rfd values (how extreme the anomalies are) and 'clean' returns only extreme anomalies, i.e. anomalies at which a given rfd is overpass (e.g. 0.90). This critical threshold is set by the users using the rfd argument.
rfd	Numeric. This argument only applies when the argument output='clean'. It defines the percentile (from 0 to 0.99) of the reference frequency distribution, for which anomalies are not flagged as extreme anomalies. For example, if 'rfd = 0.90' only anomalies falling outside the '0.90 rfd' (default value) will be flagged as extreme anomalies while the rest will be neglected (NA values). Please notice that'rfd = 0.90' implies that the 5% of the most extreme positive and 5% of the most extreme negative anomalies will be considered.
nCluster	Numeric. Number of CPU cores to be used for computational calculations
outname	Character vector with the output path and filename with extension or only the filename and extension if work directory was set. For example outname ="output_anom.tif". See [writeRaster](#)
format	Character. Output file type. See [writeFormats](#)
datatype	Character. Output data type. See [dataType](#)

Details

Similar to [ExtremeAnom](#), it calculates phenological anomalies but using a raster stack instead of a numeric vector of vegetation canopy greenness values (e.g. Leaf Area Index, LAI) or satellite based greenness proxies such as the Normalized Difference Vegetation Index (NDVI) or Enhanced Vegetation Index (EVI). For this purpose, it divides the time series (raster stack) of vegetation greenness into 2: the reference period, from which the annual phenological cycle is calculated (same as [PhenMap](#) function), and the observation period, for which we want to calculate anomalies with respect to the annual phenological cycle. Negative anomalies correspond to observed values lower than the reference and positive anomalies to values higher than the reference. his anomalies can be filtered by the position of the observation within the historical rfd. Users can, for example, set 'rfd = 0.95' to consider only anomalies that outside the 95% rfd of historical records.

Value

RasterStack

See Also

[ExtremeAnom](#)

Examples

```
## Not run:
##DEPENDING ON HARDWARE, THIS PROCESS CAN BE HIGHLY TIME CONSUMING##

## Testing with the MegaDrought_stack from Central Chile (NDVI),
#showing extreme negative anomalies (browning)##

# Load data
#RasterStack
data("MegaDrought_stack")
#Dates
data("modis_dates")

library(snow)

# Define the number of cores to be use. In this example we use 1
nc1 <- 1
ExtremeAnoMap(s=MegaDrought_stack,dates=modis_dates,h=2,refp = c(1:423),
anop = c(884:929),rfd = 0.9,output = 'both',nCluster=nc1,outname="anomRFD_MD.tif",
format="GTiff", datatype="INT2S", rge=c(0,10000))
#map_an1 <- stack("anomRFD_MD.tif")#run only for load anomaly brick
#plot(map_an1)


## Testing with the Bdesert_stack from the Atacama Desert, Northern Chile (NDVI),
#showing extreme positive anomalies (greening)##

# Load data
RasterStack
data("Bdesert_stack")
#Dates
data("modis_dates")

# Define the number of cores to be use. In this example we use 1
nc1 <- 1

ExtremeAnoMap(s=Bdesert_stack,dates=modis_dates,h=2,refp = c(1:423),
anop = c(723:768),rfd = 0.9,output = 'both',nCluster=nc1,outname="anomRFD_BD.tif",
format="GTiff", datatype="INT2S", rge=c(0,10000))
#map_an1 <- stack("anomRFD_BD.tif")#run only for load anomaly brick
#plot(map_an1)


## End(Not run)
```

MegaDrought_stack   *Mega drought in Central Chile*

Description

MODIS NDVI RasterStack of Central Chile combining MOD13Q1 & MYD13Q1 between 2000.02.18
and 2021.06.26, cleaned using the detailed QA band. Megadrought affected Central Chile between
2010 and 2015 and again in 2019-2022

Usage

```
MegaDrought_stack
```

Format

RasterStack EPSG:32719

Source

<https://lpdaac.usgs.gov/>

---

modis_dates	*A vector with dates for both RasterStack datasets: Mega drought and Blooming Desert, with 929 dates*

Description

A vector with dates for both RasterStack datasets: Mega drought and Blooming Desert, with 929
dates

Usage

```
modis_dates
```

Format

Vector of dates

Source

<https://lpdaac.usgs.gov/>

---

npphen	*npphen*

Description

The functions in this package estimate the expected annual phenological cycle from time series or raster stack of vegetation (greenness) indexes. The algorithm to estimate the annual phenological cycle (used by the functions in npphen) uses a bivariate kernel density estimator in the index-time space. In this space, the x-axis corresponds to days of the growing season (1-365) and the y-axis to the vegetation index values, which range's values are set using the rge argument (see each function's vignette for details). The expected value of the index for each day of the growing season (the expected phenology) is approximated by the maximum value of the kernel at that day. Anomalies are calculated as deviations from the expected values at given days. The package implements basic and high-level functions for manipulating vector data (numerical series) and raster data (satellite derived products). Processing of very large raster files is supported. For methodological details of kernel density estimation see Wand & Jones (1994).

Author(s)

Roberto O. Chávez <roberto.chavez.o@pucv.cl>
Sergio A. Estay <sergio.estay@uach.cl>
José A. Lastra <jose.lastra@pucv.cl>
Carlos G. Riquelme <carlosriquelmemv@gmail.com>

References

Wand, M.P. & Jones, M.C. (1994) Kernel smoothing. Crc Press.

See Also

Phen, PhenMap, ExtremeAnom, ExtremeAnoMap, PhenKplot

Phen	*Phen*

Description

Estimates the annual phenological cycle from a time series of vegetation greenness.

Usage

```
Phen(x, dates, h, frequency = "16-days", rge)
```

Arguments

x	Numeric vector. A time series of a vegetation index (e.g. LAI, NDVI, EVI) or any other variable with seasonal behavior. The code has been optimized to work with integer values. Please re-scale the input values if necessary (e.g. NDVI ranging from 0.0000 to 1.0000, multiply by 10,000
dates	A date vector. The number of dates must be equal to the number of "x" values (numeric input vector).
h	Numeric. Indicates the geographic hemisphere to define the starting date of the growing season. h=1 if the vegetation is in the Northern Hemisphere (season starting at January 1st), h=2 if it is in the Southern Hemisphere (season starting at July 1st)
frequency	Character string. Defines the number of samples for the output phenology and must be one of the this: 'daily' giving output vector of length 365, '8-days' giving output vector of length 46 (i.e MOD13Q1 and MYD13Q1), 'monthly' giving output vector of length 12,'bi-weekly' giving output vector of length 24 (i.e. GIMMS) or '16-days' (default) giving output vector of length 23 (i.e MOD13Q1 or MYD13Q1).
rge	Numeric vector with two values setting the minimum and maximum values of the response variable (e.g. NDVI) used in the analysis. We suggest the use of theoretically based limits. For example in the case of MODIS NDVI or EVI, it ranges from 0 to 10,000, so rge = c(0,10000)

Details

Derives the annual phenological cycle for a standard growing season using a numeric vector of vegetation canopy greenness values (e.g. Leaf Area Index, LAI) or satellite based greenness proxies such as the Normalized Difference Vegetation Index (NDVI) or Enhanced Vegetation Index (EVI). A vector with dates for the greenness values is also required.

Value

A numeric vector, where each value represents the expected greeness at that date

See Also

[PhenMap,PhenKplot](#)

Examples

```
library(lubridate)

## Testing raster data from Central Chile (NDVI), h=2##

# Load data
#RasterStack
data("MegaDrought_stack")
```

```
#Dates
data("modis_dates")

#Creates Raster time series using a raster stack and a date database from Central Chile
# Obtain data from a particular pixel generating a time series
md_pixel <- cellFromXY(MegaDrought_stack,c(313395,6356610))
md_pixelts <- as.numeric(MegaDrought_stack[md_pixel])
plot(modis_dates,md_pixelts, type='l')

# Phenology for the given pixel
Phen(x=md_pixelts,dates=modis_dates,h=2,frequency='16-days',rge=c(0,10000))

## Testing with the Bdesert_stack from the Atacama Desert, Northern Chile (NDVI), h=2 ##

# Load data
#RasterStack
data("Bdesert_stack")
#Dates
data("modis_dates")

#Creates Raster time series using a raster stack and a date database from Northern Chile
# Obtain data from a particular pixel generating a time series
bd_pixel<-cellFromXY(Bdesert_stack,c(286638,6852107))
bd_pixelts<-as.numeric(Bdesert_stack[bd_pixel])
plot(modis_dates,bd_pixelts, type = 'l')

# Phenology for the given pixel
Phen(x=bd_pixelts,dates=modis_dates,h=2,frequency='16-days',rge=c(0,10000))
```

PhenKplot   *PhenKplot*

Description

Plot the most probable vegetation greenness values.

Usage

```
PhenKplot(x, dates, h, xlab, ylab, rge)
```

Arguments

x   Numeric vector. A time series of a vegetation index (e.g. LAI, NDVI, EVI) or
                   any other variable with seasonal behavior. The code has been optimized to work
                   with integer values. Please re-scale the input values if necessary (e.g. NDVI
                   ranging from 0.0000 to 1.0000, multiply by 10,000

dates   A date vector. The number of dates must be equal to the number of "x" values
                   (numeric input vector).

h	Numeric. Indicates the geographic hemisphere to define the starting date of the growing season. h=1 if the vegetation is in the Northern Hemisphere (season starting at January 1st), h=2 if it is in the Southern Hemisphere (season starting at July 1st)
xlab	Character vector (or expression) giving plot title in x axis label
ylab	Character vector (or expression) giving plot title in y axis label
rge	Numeric vector with two values setting the minimum and maximum values of the response variable (e.g. NDVI) used in the analysis. We suggest the use of theoretically based limits. For example in the case of MODIS NDVI or EVI, it ranges from 0 to 10,000, so rge =c(0,10000)

Details

It is the graphical version of the [Phen](#) function. It calculates and plot a likelihood map of the vegetation-greenness – time space using a numeric vector of vegetation canopy greenness values (e.g. Leaf Area Index (LAI) or greenness proxies values such as the Normalized Difference Vegetation Index (NDVI) or Enhanced Vegetation Index (EVI). Also a vector with dates for the greenness values is required. This function calculates the confidence areas on a per year basis. Functions for confidence intervals on per day basis are under development. This function is partially based on the ci2d function on package gplots.

See Also

[Phen](#), [PhenMap](#)

Examples

```
library(lubridate)

## Testing raster data from Central Chile (NDVI), h=2##

# Load data
#RasterStack
data("MegaDrought_stack")
#Dates
data("modis_dates")

# Generate a Raster time series using a raster stack and a date database from Central Chile
# Obtain data from a particular pixel generating a time series
md_pixel <- cellFromXY(MegaDrought_stack,c(313395,6356610))
md_pixelts <- as.numeric(MegaDrought_stack[md_pixel])
plot(modis_dates,md_pixelts, type='l')

# Phenology for the given pixel
PhenKplot(x=md_pixelts,dates=modis_dates,h=2,xlab="DGS", ylab="NDVI", rge=c(0,10000))


## Testing with the Bdesert_stack from the Atacama Desert, Northern Chile (NDVI), h=2 ##
```

```
# Load data
#RasterStack
data("Bdesert_stack")
#Dates
data("modis_dates")

# Generate a Raster time series using a raster stack and a date database from Northern Chile
# Obtain data from a particular pixel generating a time series
bd_pixel<-cellFromXY(Bdesert_stack,c(286638,6852107))
bd_pixelts<-as.numeric(Bdesert_stack[bd_pixel])
plot(modis_dates,bd_pixelts, type = 'l')

# Phenology for the given pixel
PhenKplot(x=bd_pixelts,dates=modis_dates,h=2,xlab = 'DGS',ylab = 'NDVI',rge=c(0,10000))
```

PhenMap   *PhenMap*

Description

Estimates annual Land Surface Phenology (LSP) using time series of a vegetation greenness raster stack.

Usage

```
PhenMap(s, dates, h, frequency = "16-days", nCluster, outname, format,
    datatype, rge)
```

Arguments

s	Raster stack of a vegetation index (e.g. LAI, NDVI, EVI) or any other variable with seasonal behavior. The code has been optimized to work with integer values. Please re-scale the input raster stack if necessary (e.g. NDVI ranging from 0.0000 to 1.0000, multiply by 10,000
dates	A date vector. The number of dates must be equal to the number of "x" values (numeric input vector).
h	Numeric. Indicates the geographic hemisphere to define the starting date of the growing season. h=1 if the vegetation is in the Northern Hemisphere (season starting at January 1st), h=2 if it is in the Southern Hemisphere (season starting at July 1st)
frequency	Character string. Defines the number of samples for the output phenology and must be one of the this: 'daily' giving output vector of length 365, '8-days' giving output vector of length 46 (i.e MOD13Q1 and MYD13Q1), 'monthly' giving output vector of length 12,'bi-weekly' giving output vector of length 24 (i.e. GIMMS) or '16-days' (default) giving output vector of length 23 (i.e MOD13Q1 or MYD13Q1).

nCluster	Numeric. Number of CPU cores to be used for computational calculations
outname	Character vector with the output path and filename with extension or only the filename and extension if work directory was set. For example outname = "output_phen.tif". See `writeRaster`
format	Character. Output file type. See `writeFormats`
datatype	Character. Output data type. See `dataType`
rge	A vector containing minimum and maximum values of the response variable used in the analysis. We suggest the use of theoretically based limits. For example in the case of MODIS NDVI or EVI, it ranges from 0 to 10,000, so rge =c(0,10000)

Details

Derives the annual Land Surface Phenological (LSP) cycle for a standard growing season using a raster stack of satellite based greenness values such as the Normalized Difference Vegetation Index (NDVI) or Enhanced Vegetation Index (EVI). The LSP cycle is calculated for all pixels of the input raster stack in the same way as for the `Phen` function. The output is a multiband raster where every band is the expected greenness value at a given time step of the standard growing season. For example, for MODIS Vegetation Index 16-days composites the number of time steps of the growing season is 23 (frequency = "16-days"), and therefore, the output raster will have 23 bands. A vector with dates for the greenness values is also required.

Value

RasterStack

See Also

`Phen`

Examples

```
## Not run:
##DEPENDING ON HARDWARE, THIS PROCESS CAN BE HIGHLY TIME CONSUMING##

## Testing raster data from Central Chile (NDVI), h=2##

# Load data
#RasterStack
data("MegaDrought_stack")
#Dates
data("modis_dates")

# Making the LSP raster, n bands = 23

library(snow)

# Define the number of cores to be use. In this example we use 1
nc1<-1
```

```
PhenMap(s = MegaDrought_stack,dates = modis_dates,h = 2,
frequency = '16-days',nCluster = nc1,outname = 'phen_MD.tif',
format = 'GTiff',datatype = 'INT2S',rge = c(0,10000))
#map1 <- stack("phen_MD.tif")#run only for load phenology brick
#plot(map1)

## Testing with the Bdesert_stack from the Atacama Desert, Northern Chile (NDVI), h=2 ##

# Load data
RasterStack
data("Bdesert_stack")
#Dates
data("modis_dates")

# Making the LSP raster, n bands = 23
# Define the number of cores to be use. In this example we use 1
nc1<-1

PhenMap(s= Bdesert_stack,dates=modis_dates,h=2,
frequency='16-days', nCluster=nc1,outname="phen_BD.tif",
format="GTiff", datatype="INT2S",rge=c(0,10000))
#map2 <- stack("phen_BD.tif") #run only for load phenology brick
#plot(map2)

## End(Not run)
```

phents	*A dataframe with 8-days MODIS NDVI of a pixel with deciduous Nothofagus macrocarpa forest, Central Chile, using MOD13Q1/MYD13Q1 combined*

Description

A dataframe with 8-days MODIS NDVI of a pixel with deciduous Nothofagus macrocarpa forest, Central Chile, using MOD13Q1/MYD13Q1 combined

Usage

```
phents
```

Format

A data frame with 929 rows and 2 variables:

dates dates for each record

NDVI NDVI value for a pixel of deciduous N. macrocarpa forest

Source

https://lpdaac.usgs.gov/

Index