

Introduction to the ‘chronosphere’ R package

Adam T. Kocsis, Nussaibah B. Raja

2019-11-28

1. Introduction

1.1. Installation

To install this alpha version of the package, you must download it either from the CRAN servers or its dedicated GitHub repository (<http://www.github.com/adamkocsis/chronosphere/>). All minor updates will be posted on GitHub as soon as they are finished, so please check this regularly. The version on CRAN will be lagging for some time, as it takes the servers many days to process everything. All questions and bugs can be reported at the github issues board (<https://github.com/adamkocsis/chronosphere/issues>). Instead of spending it on actual research, a tremendous amount of time was invested in making this piece of software streamlined and user-friendly. If you use a dataset of the package in a publication, please cite both its reference(s) and the chronosphere’ package itself.

After installing, from the CRAN, from a source or with `devtools::install_github()`, you can attach the package with:

```
library(chronosphere)
```

1.2 General features

The purpose of the ‘chronosphere’ project is to facilitate, streamline and fasten the finding, acquisition and importing of Earth science data in R. Although the package currently focuses on deep time datasets, the scope of the included datasets will be much larger, spanning from a single variable published as supplementary material in a journal article, to GIS data or the entire output of GCM models. ‘chronosphere’ intends to decrease the gap between research hypotheses and the finding, download and importing of datasets.

2. RasterArray

Faster data importing and better organization represents a considerable part of this process. Spatially explicit data are excellent candidates to demonstrate how more efficient data organization can speed up research. Although R has an excellent infrastructure for handling raster data (Shijmans), the arrangement of individual layers are rather limited, which can be a problem, when a large number of layers have to be considered. RasterStacks and RasterBricks are very efficient for organizing RasterLayers according to a single dimension (e.g. depth for 3D variables, or time), multidimensional structures are preferred.

To offer a more effective solution, the ‘chronosphere’ package includes the definition of the RasterArray S4 class. RasterArrays represent hybrids between RasterStacks and regular R arrays. In short, they are virtual arrays of RasterLayers, and can be thought of as regular arrays, that include entire rasters as elements rather than single numeric, logical or character values. As regular R users are familiar with subsetting, combinations and structures of regular arrays (including formal vectors and matrices), the finding, extraction and storage of spatially explicit data is much easier in such containers.

2.1. Structure

RasterArrays do not directly inherit from Raster* objects of the raster package, as a considerable number of main functions differ, but they rather represent a wrapper object around regular RasterStacks - stacks of

individual RasterLayers. This ensures that whenever users are unfamiliar with the methods of RasterArray class objects, they can always reduce their data to stacks or individual layers.

The formal class RasterArray has only two slots: a stack and an index. The stack includes the Raster data in an unfolded manner, similarly to how matrices and arrays are essentially just vectors with additional attributes. The stack slot incorporates a single RasterStack object, which represents the data content of the object. The index slot, on the other hand, describes the structure of the RasterArray. It is a vector/matrix/array of integers each representing an index of the layers in the stack. The configuration (dimensions) of the index represents the entire array.

The ‘chronosphere’ package includes two demo datasets: a set of ancient topographies (PaleoDEMs, SCotese..) and time series of bioclimatic variables (annual mean temperature and precipitation) from the CHELSA project. (Sref) These datasets can be attached with the data command.

```
data(dems)
data(clim)
```

The structure of RasterArrays can be inspected if the object’s name is typed into the console:

```
dems

## class          : RasterArray
## RasterLayer properties:
## - dimensions   : 181, 361 (nrow, ncol)
## - resolution   : 1, 1 (x, y)
## - extent       : -180.5, 180.5, -90.5, 90.5 (xmin, xmax, ymin, ymax)
## - coord. ref.  : +proj=longlat +datum=WGS84 +ellps=WGS84 +towgs84=0,0,0
## Array properties:
## - dimensions   : 10 (vector)
## - num. layers   : 10
## - proxy:
##      0         5         10        15        20        25        30        35
## "dem_0" "dem_5" "dem_10" "dem_15" "dem_20" "dem_25" "dem_30" "dem_35"
##      40        45
## "dem_40" "dem_45"
```

The first part of the console output includes properties of the individual RasterLayers stored in the stack. These layers have to share essential attributes that allow them to be stored in a single stack (extent, resolution, CRS).

The second part of the output is a visualization of the structure of the RasterArray itself. In the case of the DEMs, 10 layers are stored in the stack, each layer having its individual name (e.g. ‘dem_0’). It is a single dimensional array (vector), and each element has its name in the array (“0”). The differentiating between the names of layers and the names of elements allows different subsetting and replacement rules for the two, which both can be handy - depending on the needs of the user.

The structure of the RasterArray can be visualized, analyzed or processed using the proxy object. Proxies are essentially the same as the index slots of the RasterArray, but instead of including the indices of the layers they represent, proxies include the names of the layers. These can be accessed using the proxy() function.

```
proxy(dems)

##      0         5         10        15        20        25        30        35
## "dem_0" "dem_5" "dem_10" "dem_15" "dem_20" "dem_25" "dem_30" "dem_35"
##      40        45
## "dem_40" "dem_45"
```

Proxies are displayed as the second parts of the console output when the name of the object is typed into the

console (show method).

```
clim

## class          : RasterArray
## RasterLayer properties:
## - dimensions   : 181, 361 (nrow, ncol)
## - resolution   : 1, 1 (x, y)
## - extent       : -180.5, 180.5, -90.5, 90.5 (xmin, xmax, ymin, ymax)
## - coord. ref.  : +proj=longlat +datum=WGS84 +ellps=WGS84 +towgs84=0,0,0
## Array properties:
## - dimensions   : 10, 2 (nrow, ncol)
## - num. layers   : 20
## - proxy:
##      bio1      bio12
## 2001 "dem_0.1" "dem_0.2"
## 2002 "dem_5.1" "dem_5.2"
## 2003 "dem_10.1" "dem_10.2"
## 2004 "dem_15.1" "dem_15.2"
## 2005 "dem_20.1" "dem_20.2"
## 2006 "dem_25.1" "dem_25.2"
## 2007 "dem_30.1" "dem_30.2"
## 2008 "dem_35.1" "dem_35.2"
## 2009 "dem_40.1" "dem_40.2"
## 2010 "dem_45.1" "dem_45.2"
```

This RasterArray has 10 rows (annual means) and two variables/columns: temperature (bio1) and precipitation (bio12). With the proxy() function it is easy to interact with this object, or to query or analyze it.

```
proxy(clim)

##      bio1      bio12
## 2001 "dem_0.1" "dem_0.2"
## 2002 "dem_5.1" "dem_5.2"
## 2003 "dem_10.1" "dem_10.2"
## 2004 "dem_15.1" "dem_15.2"
## 2005 "dem_20.1" "dem_20.2"
## 2006 "dem_25.1" "dem_25.2"
## 2007 "dem_30.1" "dem_30.2"
## 2008 "dem_35.1" "dem_35.2"
## 2009 "dem_40.1" "dem_40.2"
## 2010 "dem_45.1" "dem_45.2"
```

RasterArrays are fairly easy to construct: one only needs a stack of the data and an regular vector/matrix/array including integers. For instance, the dems object can be recreated from scratch without problems.

```
# a stack of rasters
stackOfLayers <- dems@stack
# an index object
ind <- 1:10
names(ind) <- letters[1:10]
# a RasterArray
nra <- RasterArray(index=ind, stack=stackOfLayers)
nra

## class          : RasterArray
## RasterLayer properties:
## - dimensions   : 181, 361 (nrow, ncol)
```

```
## - resolution : 1, 1 (x, y)
## - extent : -180.5, 180.5, -90.5, 90.5 (xmin, xmax, ymin, ymax)
## - coord. ref. : +proj=longlat +datum=WGS84 +ellps=WGS84 +towgs84=0,0,0
## Array properties:
## - dimensions : 10 (vector)
## - num. layers : 10
## - proxy:
## a b c d e f g h
## "dem_0" "dem_5" "dem_10" "dem_15" "dem_20" "dem_25" "dem_30" "dem_35"
## i j
## "dem_40" "dem_45"
```

The attributes of the index object are defining the structure of the RasterArray. RasterArrays can be created with the combination of individual RasterLayers (or RasterArrays) using the `combine()` function.

```
# one raster
r1 <- raster()
values(r1) <- 1
# same structure, different value
r2 <- raster()
values(r2) <- 2
comb <- combine(r1, r2)
comb
```

```
## class : RasterArray
## RasterLayer properties:
## - dimensions : 180, 360 (nrow, ncol)
## - resolution : 1, 1 (x, y)
## - extent : -180, 180, -90, 90 (xmin, xmax, ymin, ymax)
## - coord. ref. : +proj=longlat +datum=WGS84 +ellps=WGS84 +towgs84=0,0,0
## Array properties:
## - dimensions : 2 (vector)
## - num. layers : 2
## - proxy:
## r1 r2
## "layer.1" "layer.2"
```

Matrix-like RasterArrays can also be created easily with the, `cbind()`, and `rbind()` functions.

```
# bind dems to itself
cbind(dems, dems)

## class : RasterArray
## RasterLayer properties:
## - dimensions : 181, 361 (nrow, ncol)
## - resolution : 1, 1 (x, y)
## - extent : -180.5, 180.5, -90.5, 90.5 (xmin, xmax, ymin, ymax)
## - coord. ref. : +proj=longlat +datum=WGS84 +ellps=WGS84 +towgs84=0,0,0
## Array properties:
## - dimensions : 10, 2 (nrow, ncol)
## - num. layers : 20
## - proxy:
## [,1] [,2]
## 0 "dem_0.1" "dem_0.2"
## 5 "dem_5.1" "dem_5.2"
## 10 "dem_10.1" "dem_10.2"
## 15 "dem_15.1" "dem_15.2"
```

```
## 20 "dem_20.1" "dem_20.2"
## 25 "dem_25.1" "dem_25.2"
## 30 "dem_30.1" "dem_30.2"
## 35 "dem_35.1" "dem_35.2"
## 40 "dem_40.1" "dem_40.2"
## 45 "dem_45.1" "dem_45.2"
```

2.2. RasterArray attributes and function to query

Functions that query and change attributes of the RasterArray resemble general arrays more than Raster* objects. They are connected to the index slot of the RasterArray and return values accordingly.

The number of elements represented in the RasterArray can be queried with the `length()` function:

```
length(dems)
```

```
## [1] 10
```

This RasterArray has 10 elements. The number of column and row names can be queried in a similar way:

```
nrow(clim)
```

```
## [1] 10
```

```
ncol(clim)
```

```
## [1] 2
```

These functions are summarized in the `dim()` function. This, however, unlike the regular `dim()` method of vectors, return the length of the RasterArray-vector, rather than just NULL.

```
dim(dems)
```

```
## [1] 10
```

```
dim(clim)
```

```
## [1] 10 2
```

The organization of RasterLayers can be greatly facilitated with names. The `names()`, `colnames()`, `rownames()` and `dimnames()` functions work the same way on RasterArrays as if they were arrays of simple numeric, logical or character values. The `names()` function returns the names of individual elements of a vector-like RasterArray.

```
names(dems)
```

```
## [1] "0" "5" "10" "15" "20" "25" "30" "35" "40" "45"
```

The `colnames()` and `rownames()` functions are more relevant for matrix-like RasterArrays, such as `clim`.

```
colnames(clim)
```

```
## [1] "bio1" "bio12"
```

```
rownames(clim)
```

```
## [1] "2001" "2002" "2003" "2004" "2005" "2006" "2007" "2008" "2009" "2010"
```

All name-related methods can be used for replacement as well. For instance, you can quickly rename the names of the columns of the `clim` object this way:

```
clim2 <- clim
colnames(clim2) <- c("temp", "prec")
clim2
```

```
## class          : RasterArray
## RasterLayer properties:
## - dimensions   : 181, 361 (nrow, ncol)
## - resolution   : 1, 1 (x, y)
## - extent       : -180.5, 180.5, -90.5, 90.5 (xmin, xmax, ymin, ymax)
## - coord. ref.  : +proj=longlat +datum=WGS84 +ellps=WGS84 +towgs84=0,0,0
## Array properties:
## - dimensions   : 10, 2 (nrow, ncol)
## - num. layers   : 20
## - proxy:
##      temp      prec
## 2001 "dem_0.1" "dem_0.2"
## 2002 "dem_5.1" "dem_5.2"
## 2003 "dem_10.1" "dem_10.2"
## 2004 "dem_15.1" "dem_15.2"
## 2005 "dem_20.1" "dem_20.2"
## 2006 "dem_25.1" "dem_25.2"
## 2007 "dem_30.1" "dem_30.2"
## 2008 "dem_35.1" "dem_35.2"
## 2009 "dem_40.1" "dem_40.2"
## 2010 "dem_45.1" "dem_45.2"
```

Just as you would do it with normal arrays, the you can query/rewrite all names with the `dimnames()` function, that uses a list to store the names in every dimension.

```
dimnames(clim2)[[1]] <- 1:10
clim2
```

```
## class          : RasterArray
## RasterLayer properties:
## - dimensions   : 181, 361 (nrow, ncol)
## - resolution   : 1, 1 (x, y)
## - extent       : -180.5, 180.5, -90.5, 90.5 (xmin, xmax, ymin, ymax)
## - coord. ref.  : +proj=longlat +datum=WGS84 +ellps=WGS84 +towgs84=0,0,0
## Array properties:
## - dimensions   : 10, 2 (nrow, ncol)
## - num. layers   : 20
## - proxy:
##      temp      prec
## 1 "dem_0.1" "dem_0.2"
## 2 "dem_5.1" "dem_5.2"
## 3 "dem_10.1" "dem_10.2"
## 4 "dem_15.1" "dem_15.2"
## 5 "dem_20.1" "dem_20.2"
## 6 "dem_25.1" "dem_25.2"
## 7 "dem_30.1" "dem_30.2"
## 8 "dem_35.1" "dem_35.2"
## 9 "dem_40.1" "dem_40.2"
## 10 "dem_45.1" "dem_45.2"
```

Besides the names of the elements in the `RasterArray`, every layer has its own name in the stack. These can be accessed with `layers()` function:

```
layers(clim)
```

```
## [1] "dem_0.1" "dem_5.1" "dem_10.1" "dem_15.1" "dem_20.1" "dem_25.1"
## [7] "dem_30.1" "dem_35.1" "dem_40.1" "dem_45.1" "dem_0.2" "dem_5.2"
```

```
## [13] "dem_10.2" "dem_15.2" "dem_20.2" "dem_25.2" "dem_30.2" "dem_35.2"  
## [19] "dem_40.2" "dem_45.2"
```

The total number of cells in the RasterLayer or the entire stack can be accessed with the `ncell()` and `nvalues()` functions, respectively.

```
ncell(dems)
```

```
## [1] 65341
```

```
nvalues(dems)
```

```
## [1] 653410
```

2.3. Subsetting and replacement

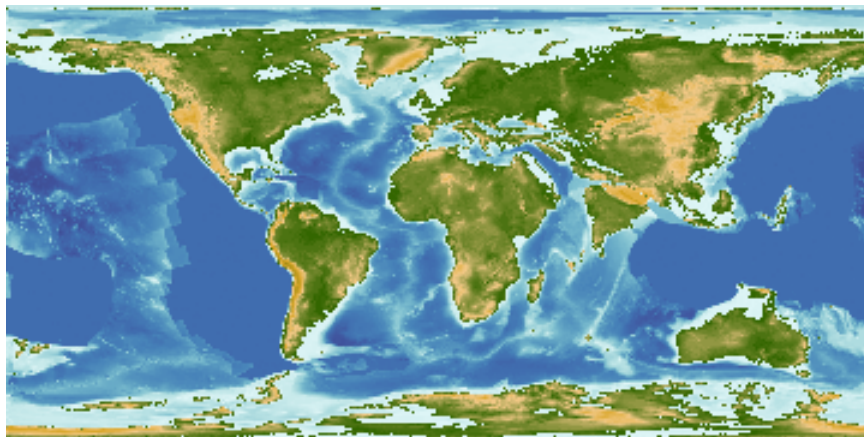
Facilitating the accession of items is the primary purpose of RasterArrays. These either focus on the layers (stack items, double bracket operator “[[”) or the elements of the RasterArray (single bracket operator “[”).

2.3.1 Layer selection - Double bracket [[

This form of subsetting and replacement are inherited from the RasterStack class. Individual layers can be accessed directly from the stack using either the position index, the name of the layer or the logical value pointing to the position. Whichever is used, the RasterArray wrapper is omitted and output will be a RasterLayer or RasterStack class object.

A single layer can be accessed using its name, regardless of its position in the RasterArray. This can be visualized either with the default `plot()` or the more general `mapplot()` function.

```
one <- dems[["dem_45"]]  
mapplot(one, col="earth")
```



Returning a single RasterArray. Multiple elements will be the format of a stack:

```
dems[[c(1,2)]]
```

```
## class      : RasterStack
## dimensions : 181, 361, 65341, 2  (nrow, ncol, ncell, nlayers)
## resolution : 1, 1  (x, y)
## extent     : -180.5, 180.5, -90.5, 90.5  (xmin, xmax, ymin, ymax)
## crs        : +proj=longlat +datum=WGS84 +ellps=WGS84 +towgs84=0,0,0
## names      : dem_0, dem_5
## min values : -9000, -7000
## max values : 6000, 6300
```

Which are the first two RasterLayers in the stack of the RasterArray.

Double brackets can also be used for replacements, but as this has no effect on the structure of the array, changes implemented with this method are more difficult to trace. For instance,

```
# copy
dem2 <- dems
dem2[["dem_0"]] <- dem2[["dem_5"]]
```

will rewrite the values in the first element of dem2, but that will not be evident in the RasterArray's structure.


```
# but these two are now the same
```

```
dem2[[1]]
```

```
## class      : RasterLayer
## dimensions  : 181, 361, 65341  (nrow, ncol, ncell)
## resolution  : 1, 1  (x, y)
## extent     : -180.5, 180.5, -90.5, 90.5  (xmin, xmax, ymin, ymax)
## crs        : +proj=longlat +datum=WGS84 +ellps=WGS84 +towgs84=0,0,0
## source     : memory
## names      : dem_0
## values     : -7000, 6300  (min, max)
## zvar       : z
```

```
dem2[[2]]
```

```
## class      : RasterLayer
## dimensions  : 181, 361, 65341  (nrow, ncol, ncell)
## resolution  : 1, 1  (x, y)
## extent     : -180.5, 180.5, -90.5, 90.5  (xmin, xmax, ymin, ymax)
## crs        : +proj=longlat +datum=WGS84 +ellps=WGS84 +towgs84=0,0,0
## source     : memory
## names      : dem_5
## values     : -7000, 6300  (min, max)
## zvar       : z
```

2.3.2 Single bracket

Features offered by the double bracket (“[[” operator are virtually identical with those of RasterStacks. The true utility of RasterArrays become evident with simple array-type subsetting.

Unlike Raster* objects of the raster package, single brackets will get and replace items from the RasterArray as if they were simple arrays. For example, single elements of the DEMs can be selected with the age of the DEM, passed as a character subscript.

```
dems["30"]
```

```
## class      : RasterLayer
## dimensions  : 181, 361, 65341  (nrow, ncol, ncell)
## resolution  : 1, 1  (x, y)
## extent     : -180.5, 180.5, -90.5, 90.5  (xmin, xmax, ymin, ymax)
## crs        : +proj=longlat +datum=WGS84 +ellps=WGS84 +towgs84=0,0,0
## source     : memory
## names      : dem_30
## values     : -8000, 10200  (min, max)
## zvar       : z
```

returning the 30Ma RasterLayer. By default the RasterArray container is dropped, but it can be conserved, if the drop argument is set to FALSE.

```
dem30 <- dems["30", drop=FALSE]
class(dem30)
```

```
## [1] "RasterArray"
## attr(,"package")
## [1] "chronosphere"
```

Beyond bounds accessing is valid for single dimensional RasterArrays (vector-like ones):

```
dems[4:12]
```

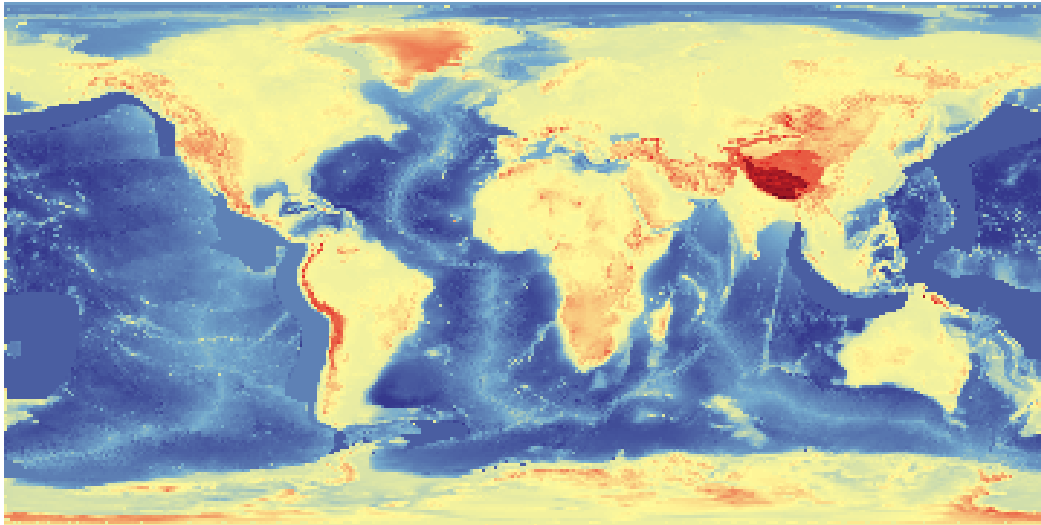
```
## class          : RasterArray
## RasterLayer properties:
## - dimensions   : 181, 361 (nrow, ncol)
## - resolution   : 1, 1 (x, y)
## - extent       : -180.5, 180.5, -90.5, 90.5 (xmin, xmax, ymin, ymax)
## - coord. ref.  : +proj=longlat +datum=WGS84 +ellps=WGS84 +towgs84=0,0,0
## Array properties:
## - dimensions   : 9 (vector)
## - num. layers   : 7
## - proxy:
##      15      20      25      30      35      40      45      <NA>
## "dem_15" "dem_20" "dem_25" "dem_30" "dem_35" "dem_40" "dem_45"      NA
##      <NA>
##      NA
```

Missing values are legitimate parts of RasterArrays. These gaps in the data are not represented in the stacks, but only in the index slots of the RasterArrays. They can be inserted or added into the layers.

```
demna <- dems
demna[3] <- NA
```

Multidimensional subscripts work in a similar fashion. If a single layer is desired from the RasterArray, that can be accessed using the names of the margins.

```
# character is necessary, as the row named "2003" is necessary
one <- clim["2003", "bio1"]
mapplot(one)
```



similarly to entire rows,

```
clim["2005", ]
```

```
## class          : RasterArray
## RasterLayer properties:
## - dimensions   : 181, 361 (nrow, ncol)
## - resolution   : 1, 1 (x, y)
## - extent       : -180.5, 180.5, -90.5, 90.5 (xmin, xmax, ymin, ymax)
## - coord. ref.  : +proj=longlat +datum=WGS84 +ellps=WGS84 +towgs84=0,0,0
## Array properties:
## - dimensions   : 2 (vector)
## - num. layers   : 2
## - proxy:
##      bio1      bio12
## "dem_20.1" "dem_20.2"
```

or columns

```
clim[, "bio12"]
```

```
## class          : RasterArray
```

```
## RasterLayer properties:
## - dimensions : 181, 361 (nrow, ncol)
## - resolution : 1, 1 (x, y)
## - extent : -180.5, 180.5, -90.5, 90.5 (xmin, xmax, ymin, ymax)
## - coord. ref. : +proj=longlat +datum=WGS84 +ellps=WGS84 +towgs84=0,0,0
## Array properties:
## - dimensions : 10 (vector)
## - num. layers : 10
## - proxy:
##      2001      2002      2003      2004      2005      2006
## "dem_0.2" "dem_5.2" "dem_10.2" "dem_15.2" "dem_20.2" "dem_25.2"
##      2007      2008      2009      2010
## "dem_30.2" "dem_35.2" "dem_40.2" "dem_45.2"
```

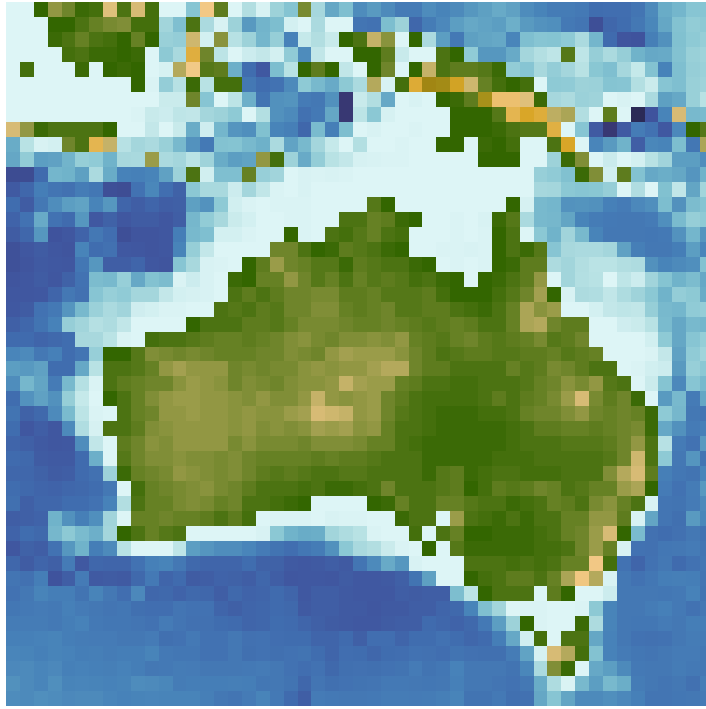
2.4 Inherited from Raster*

As the spatial information is contained entirely in the RasterStacks, a number methods are practically inherited from RasterStack class. For instance, all RasterLayer of the RasterArray can be cropped in a single line of code.

```
# crop to Australia
ext <- extent(c(
  xmin = 106.58,
  xmax = 157.82,
  ymin = -45.23,
  ymax = 1.14
))

# cropping all DEMS (Australia drifted in)
au<- crop(dems, ext)

# select the first element
mapplot(au[1], col="earth")
```

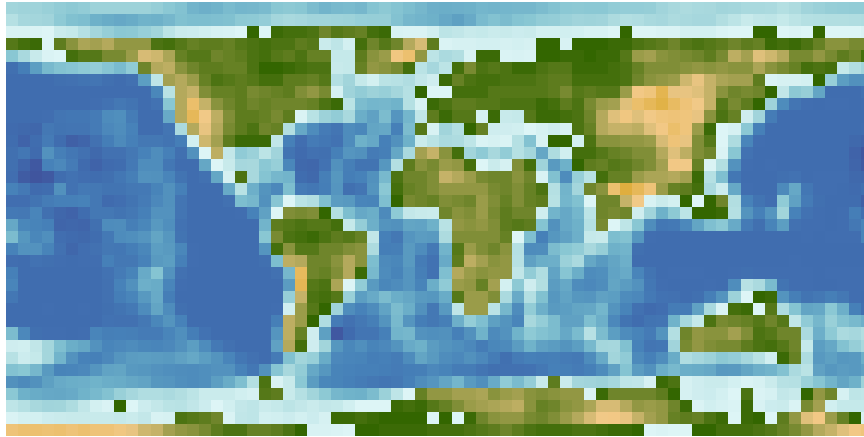


Other functions such as aggregation or resampling works just the same.

```
template <- raster(res=5)

# resample all DEMS
coarse <- resample(dems, template)

# plot an element
mapplot(coarse["45"], col="earth")
```



3. Plotting

`mapplot()` - `RasterLayer` - `RasterArray`[1 dimension] - `RasterArray`[2 dimension]

each chunk has to have a unique name (after `r`) otherwise the `vignettebuilder` will give an error.

The `echo` field sets whether the code is visible. `eval` triggers whether it is evaluated.