

R discussion, Michael Sumner

February 21, 2013

Outline

Gridded data toolbox

Spatial tools in R

Discussion of R examples

The wider suite of R tools

Gridded data toolbox

- ▶ data repository managed by Ben with a suite of perl and Matlab tools, to support Matlab and R users
- ▶ bathymetry, ocean colour, ocean currents, fronts, sea ice, sea level pressure, sea surface height/anomaly, sea surface temperature, winds, WOCE data and derived layers
- ▶ extract functions with spatial (longlat) and temporal range, options for spatial and temporal resolution and product variants

```
## load up the extractor functions
source("//aad.gov.au/files/AADC/Scientific_Data/Data/gridded/tools/R/be
data.dir = "//aad.gov.au/files/AADC/Scientific_Data/Data/gridded/data/"

## extract bathymetry over kerguelen plateau
bb = extract.bathymetry(roi.longitude = c(60, 90), roi.latitude = c(-70
    data.source = "etopo2", data.directory = data.dir)

## extract full grid for two consecutive dates
ic = extract.seaice(dates = c("1/10/2001", "2/10/2001"), data.directory
    data.source = "smmr_ssmi")
```

Toolbox - general issues

- ▶ improvements in efficiency and understanding of how to work with the resulting data objects, more options for outputs
- ▶ toolbox includes functions for display and data manipulation, and these could be updated in a number of ways (lots of options)
- ▶ not all Matlab-support is mirrored in R, due to limitations in tools and resources
- ▶ some features are "workarounds" for R, for e.g. since the source data (in HDF) cannot be read by existing R tools
- ▶ advanced use of GIS "vector" data? i.e. points, lines, polygons, good support in R, also map images, aerial photos, etc.

Given the complexity of coverage and resolution for each product, there's no easy answer on what *should* be provided, but there's room for improvement on what we have.

Toolbox - specific changes

Some improvements that I consider:

- ▶ more error checking, systematic tests, regular automated inventory of existing data
- ▶ integrate with Spatial and raster tools in R, map projections
- ▶ more support for DateTimeClasses
- ▶ standard map presentation tools, e.g. colour palettes, layouts
- ▶ output options, as Spatial, raster, NetCDF, GIS, etc.
- ▶ tools for data stored in projected grids, to avoid destructive resampling
- ▶ integrated tools for other GIS data: lines, polygons
- ▶ more examples!
- ▶ organize code into R packages (though questions about package repository and external access)

R tools for handling spatial data, gridded and vector

- ▶ traditionally a loose set of tools for spatial data
- ▶ prior to 2005 no serious organization outside of individual packages
- ▶ new classes (sp, raster, others) and tools for transforming between the variety of classes
- ▶ much more powerful but in some ways more complicated, since it all relies on extension packages

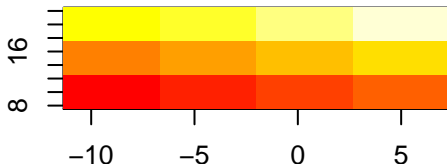
R data model of vectors, arrays, lists, and indexing

- ▶ vectors
- ▶ "atomic" vectors, character, complex, numeric, integer, logical, numeric, complex, raw, NULL
- ▶ "recursive" vectors, lists, expressions
- ▶ matrices and arrays, atomic vectors with dimension
- ▶ indexing with [and [[

R's image

Traditionally an "image" in R is just a matrix with coordinate vectors.

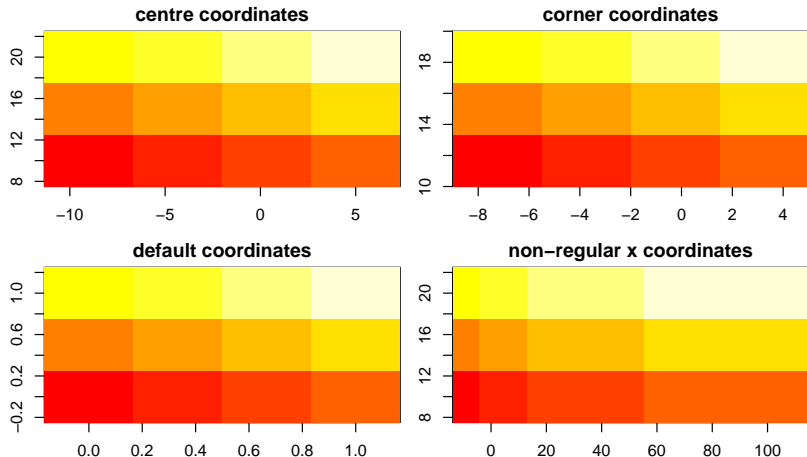
```
mat <- matrix(1:12, ncol = 3, nrow = 4)
xcentre <- seq(-9, 5, length = nrow(mat))
ycentre <- seq(10, 20, length = ncol(mat))
im1 <- list(x = xcentre, y = ycentre, z = mat)
par(cex = 0.8)
image(im1)
```



R's image, more details

```
xcorner <- seq(min(xcentre), max(xcentre), length = length(xcentre) + 1)
ycorner <- seq(min(ycentre), max(ycentre), length = length(ycentre) + 1)
op <- par(mfrow = c(1, 2))
image(xcentre, ycentre, mat, main = "centre coordinates")
image(xcorner, ycorner, mat, main = "corner coordinates")
## same as image(list(x = xcentre, y = ycentre, z = mat))
image(mat, main = "default coordinates")
image(xcentre + c(0, 5, 25, 80), ycentre, mat, main = "non-regular x co
par(op)
```

R's image, centres, corners, irregular



- ▶ `image()` will take a list, or just the matrix, or the x, y and z components separately
- ▶ coordinates can be on the centres or corners, and they may be non-regular

Raster graphics in R

Some of the current tools reflect that a true *raster image* engine is relatively new to R.

- ▶ support for palettes and a variety of colour models is very good
- ▶ *rasterImage()* with raster graphics primitives, takes colour value inputs and provides auto-scaling and interpolation

Mapping in R - traditional

- ▶ matrices and coordinate vectors, 3D arrays
- ▶ **spatial** and **fields** packages
- ▶ **spatstat** for point patterns, huge package
- ▶ traditional points, lines and polygons stored as lists of vectors (e.g. **maps** package)

Mapping in R - recent

- ▶ **sp** and **raster** provide *multi-layer* grids, stored like database tables
- ▶ **sp** stores points, lines, and polygons with attribute data much like shapefiles/databases
- ▶ raster sources and formats, map projections, visualization
- ▶ general "overlay" concepts, important options/decisions and the variety of tools and approaches
- ▶ **raster** single and multi-layer grids, many simplifications to processing with gridded and vector data and memory-handling
- ▶ out-of-memory tools in **rgdal**, **raster**, **data.table**, **ff**, **bigmemory** and various database packages
- ▶ **adehabitat***, **trip**, **diveMove**, **argosfilter** for tracking data.
- ▶ **rgeos**, **rasterVis**, **animation**, **geosphere**, **spacetime**, **dismo**, etc.

Improvements for transect analysis

- ▶ R code to determine "optimal" transects
- ▶ techniques to sample from time series grids

R Examples

Accessing monthly ocean colour data

Merging vector and raster data

Communication

- ▶ Shared data repositories, software tools, documentation, source control
- ▶ "Literate programming" tools for reporting (documents that are compiled reproducibly with embedded code)
- ▶ Animations with animation package

Finding help with R

- ▶ there's an **R for Dummies** book
- ▶ The R Inferno, gotchas for R users <http://www.burns-stat.com/documents/books/the-r-inferno/>
- ▶ R on <http://stackoverflow.com>
- ▶ do read and use the mailings lists
<http://www.r-project.org/mail.html>
- ▶ Task Views on CRAN, for domain-specific materials
<http://cran.csiro.au/web/views/>
- ▶ Contributed docs
<http://cran.csiro.au/other-docs.html>
- ▶ re-read the FAQs <http://cran.csiro.au/faqs.html>
- ▶ **sos** package