hyperSpec Plotting functions

Claudia Beleites <cbeleites@units.it> CENMAT and DI3, University of Trieste

May 26, 2011

Vignette under Development

This file is currently undergoing a thorough revision. Changes may happen frequently.

Reproducing the Examples in this Vignette

All spectra used in this manual are installed automatically with *hyperSpec*. Note that some definitions are executed in vignette.defs, and others invisibly at the beginning of the file in order to have the code as similar as possible to interactive sessions.

Contents

1	Predefined functions	2
2	Arguments for plot	4
3	Spectra 3.1 Stacked spectra	6 9
4	Calibration Plots, (Depth) Profiles, and Time Series Plots 4.1 Calibration plots	10 10 12
5	Levelplot	12
6	Spectra Matrix	13
7	False-Colour Maps: plotmap	14
8	3 D (with rgl)	17
9	Using ggplot2 with hyperSpec objects	17
10	Troubleshooting 10.1 No output is produced	19 19

11	Interactive Graphics	19
	11.1 spc.identify: finding out wavelength, intensity and spectrum	19
	11.2 map.identify: finding a spectrum in a map plot	1
	11.3 map.sel.poly: selecting spectra inside a polygon in a map plot	1
	11.4 Related functions provided by base graphics and lattice	2
	11.5 Interactively changing graphics	2

Suggested Packages

latticeExtra: available deldir: available rgl: available ggplot2: available

In addition tripack, playwith, and latticist are mentioned, but not used in this vignette.

Preliminary Calculations

For some plots of the chondro dataset, the pre-processed spectra and their cluster averages \pm one standard deviation are more suitable:

```
> chondro.preproc <- chondro - spc.fit.poly.below (chondro)
Fitting with npts.min = 15
> chondro.preproc <- sweep (chondro.preproc, 1, mean, "/")
> chondro.preproc <- sweep (chondro.preproc, 2, apply (chondro.preproc, 2, quantile, 0.05), "-")
> cluster.cols <- c ("dark blue", "orange", "#C02020")
> cluster.meansd <- aggregate (chondro.preproc, chondro$clusters, mean_pm_sd)
> cluster.means <- aggregate (chondro.preproc, chondro$clusters, mean)</pre>
```

For details about the pre-processing, please refer to the example work flow in vignette ("chondro"), or the help? chondro.

1 Predefined functions

hyperSpec comes with 6 major predefined plotting functions.

plot main switchyard for most plotting tasks

levelplot hyperSpec has a method for lattice[?] function levelplot

plotspc plots spectra

plotc calibration plot, time series, depth profile

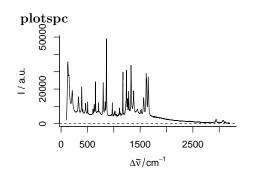
plotc is a *lattice* function

plotmap more specialized version of levelplot for map or image plots.

plotmap is a lattice function

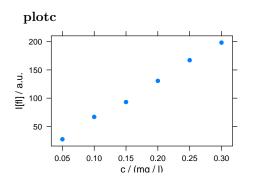
plotvoronoi more specialized version of plotmap that produces Voronoi tesselations. plotvoronoi is a *lattice* function

plotmap, plotvoronoi, and levelplot are *lattice* functions. Therefore, in loops, functions, Sweave chunks, etc. the lattice object needs to be printed explicitly by e.g. print (plotmap (object)) (R FAQ: Why do lattice/trellis graphics not work?).



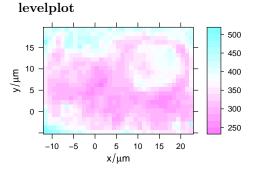
plots the spectra, i.e. the intensities \$spc over the wavelengths @wavelength.

> plotspc (paracetamol)

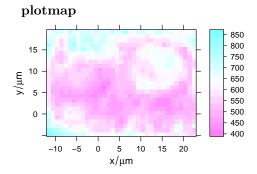


plots an intensity over a single other data column, e.g.

- calibration
- time series
- depth profile
- > plotc (flu)



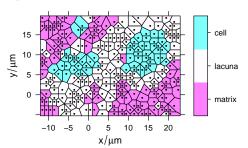
plots a false colour map, defined by a formula.
> levelplot (spc ~ x * y, chondro, aspect = "iso")
Warning: Only first wavelength is used for plotting



plotmap is a specialized version of level plot. It uses a single value (e.g. average intensity or cluster membership) over two data columns (default x and y

> plotmap (chondro)

plotvoronoi



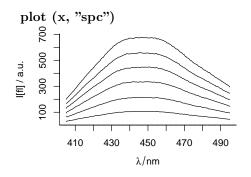
plotmap is a specialized version of level plot. It uses a single value (e. g. average intensity or cluster membership) over two data columns (default $x\$ and $y\$

> plotvoronoi (sample (chondro, 300), clusters ~ x * y) deldir 0.0-13

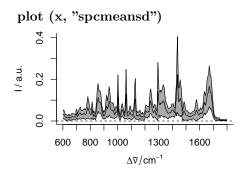
Please note: The process for determining duplicated points has changed from that used in version 0.0-9 (and previously).

2 Arguments for plot

hyperSpec's plot method uses its second argument to determine which of the specialized plots to produce. This allows some handy abbreviations. All further arguments are handed over to the function actually producing the plot.



is equivalent to plotspc (flu)
> plot (flu, "spc")

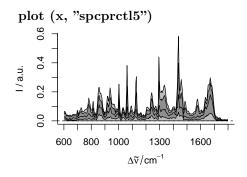


plots mean spectrum \pm 1 standard deviation > plot (chondro.preproc, "spcmeansd")

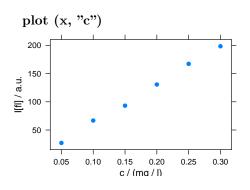


plots median, $16^{\rm th}$ and $84^{\rm th}$ percentile for each wavelength. For Gaussian distributed data, $16^{\rm th}$, $50^{\rm th}$ and $84^{\rm th}$ percentile are equal to mean \pm standard deviation. Spectroscopic data frequently are not Gaussian distributed. The percentiles give a better idea of the true distribution. They are also less sensitive to outliers.

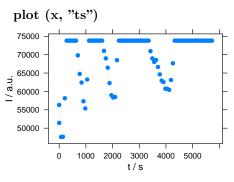
> plot (chondro.preproc, "spcprctile")



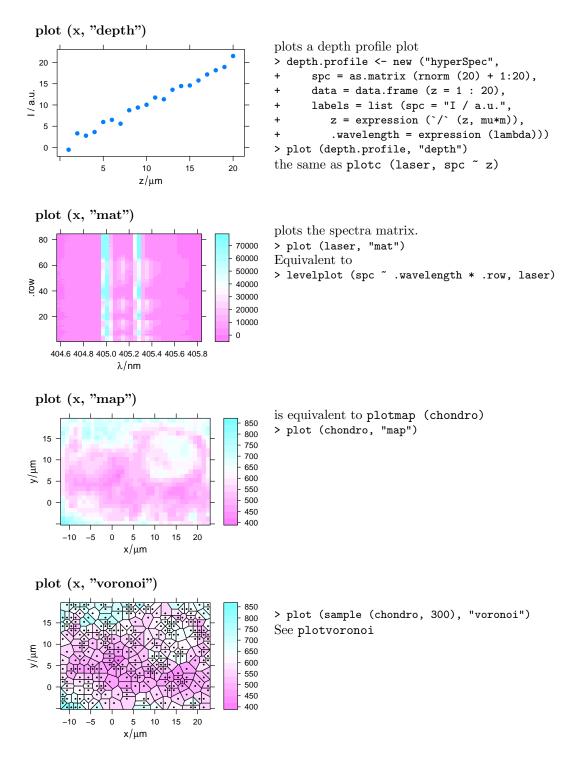
like "spcprct1" plus $5^{\rm th}$ and $95^{\rm th}$ percentile. > plot (chondro.preproc, "spcprct15")



> plot (flu, "c")
is equivalent to plotc (flu)



plots a time series plot
> plot (laser [,, 405], "ts")
equivalent to plotc (laser, spc ~ t)

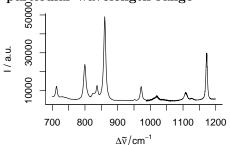


3 Spectra

plotspc

plotspc offers a variety of parameters for customized plots. To plot ...

particular wavelength range



if only one wavelength range is needed, the extract command (see vignette ("introduction")) is handiest:
> plotspc (paracetamol [,, 700 ~ 1200])

wavelengths. If wl.range already contains indices use wl.index = TRUE.

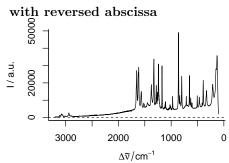


use wl.range = list (600 ~ 1800, 2800 ~ 3100). Cut the wavelength axis appropriately with xoffset = 750> plotspc (paracetamol,

wl.range = c (300 - 1800, 2800 - max),

xoffset = 750)

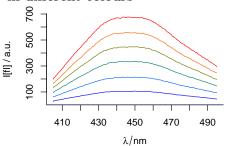
If available, the package plotrix[1] is used to produce the cut



use wl.reverse = TRUE

> plotspc (paracetamol, wl.reverse = TRUE)

in different colours



use col = vector.of.colours > plotspc (flu, col = matlab.dark.palette (6))

dots instead of lines

```
use lines.args = list (pch = 20, type = "p")
> plotspc (paracetamol [,, 2800 ~ 3200],
+ lines.args = list (pch = 20, type = "p"))
```

use lines.args = list (type = "h")
> plot (barbiturates [[1]], lines.args = list (type = "h"))

more spectra into an existing plot



use add = TRUE
> plotspc (chondro [30,,])
> plotspc (chondro [300,,], add = TRUE, col = "blue")

Summary characteristics



func may be used to calculate summary characteristics prior to plotting. To plot e.g. the standard deviation of the spectra, use:

> plotspc (chondro.preproc, func = sd)



zeroline takes a list with parameters to abline, NA suppresses the line.

```
> plotspc (paracetamol,
```

```
+ zeroline = list (col = "red"))
```

adding to a spectra plot



plotspc uses base graphics. After plotting the spectra, more content may be added to the graphic by abline, lines, points, etc.

```
> plot (laser, "spcmeansd")
```

```
> abline (v = c(405.0063, 405.1121, 405.2885, 405.3591),
```

3.1 Stacked spectra

stacked



```
use stacked = TRUE
> plotspc (cluster.means,
+ col = cluster.cols,
```

stacked = TRUE)

Stacking groups of spectra



The spectra to be stacked can be grouped: stacked = factor. Alternatively, the name of the grouping extra data column can be used:

```
> plot (cluster.meansd,
```

```
+ stacked = ".aggregate",
```

+ fill = ".aggregate",
+ col = cluster.cols)

⁺ col = c("black", "blue", "red", "darkgreen"))

Manually giving yoffset



Stacking values can also be given manually as numeric values in yoffset:

```
> plotspc (cluster.meansd,
+ yoffset = rep (0:2, each = 3),
+ col = rep (cluster.cols, each = 3))
```

Dense stacking



```
To obtain a denser stacking:
> yoffsets <- apply (cluster.means [[]], 2, diff)
> yoffsets <- apply (yoffsets, 1, min)
> plot (cluster.means, yoffset = c (0, cumsum (yoffsets)),
+ col = cluster.cols)
```

Elaborate example

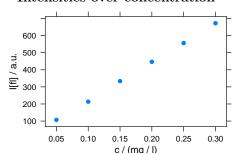


4 Calibration Plots, (Depth) Profiles, and Time Series Plots

plotc

4.1 Calibration plots

Intensities over concentration

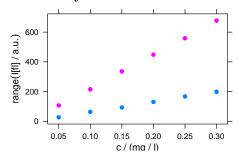


Plotting the Intensities of one wavelength over the concentration for univariate calibration:

> plotc (flu [,, 450])

The default is to use the first intensity only.

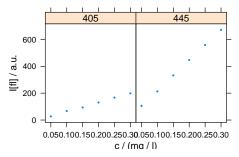
Summary Intensities over concentration



A function to compute a summary of the intensities before drawing can be used:

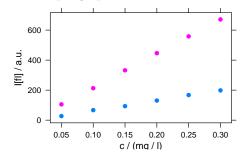
> plotc (flu, func = range, groups = .wavelength)
If func returns more than one value, the different results are accessible by .wavelength.

Conditioning: plotting more traces separately



> plotc (flu [,, c (405, 445)], spc ~ c | .wavelength,
+ cex = .3, scales = list (alternating = c(1, 1)))

Grouping: plot more traces in one panel

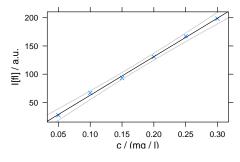


> plotc (flu [,, c (405, 445)], groups = .wavelength)

Changing Axis Labels (and other parameters)

```
Arguments for xyplot can be given to plotc:
                                                plotc (flu [,, 450],
  600
                                                         ylab = expression (I ["450 nm"] / a.u.),
  500
1450 nm/a.u.
                                                         xlim = range (0, flu$c + .01),
  400
                                                         ylim = range (0, flu$spc + 10),
  300
                                                         pch = 4)
  200
  100
           0.05
                0.10
                     0.15
                           0.20
                                 0.25
                                      0.30
                   c / (ma / I)
```

Adding things to the plot: customized panel function

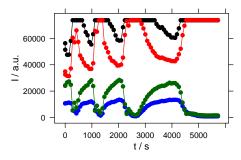


```
As plotc uses the lattice function xyplot, additions to the plot must be made via the panel function:
> panelcalibration <- function (x, y, ..., clim = range (x), level = .
```

```
+ panel.xyplot (x, y, ...)
+ lm <- lm (y ~ x)
+ panel.abline (coef (lm), ...)
+ cx <- seq (clim [1], clim [2], length.out = 50)
+ cy <- predict (lm, data.frame (x = cx),
+ interval = "confidence",
+ level = level)
+ panel.lines (cx, cy [,2], col = "gray")
+ panel.lines (cx, cy [,3], col = "gray")
+ }
> plotc (flu [,,405], panel = panelcalibration,
+ pch = 4, clim = c (0, 0.35), level = .99)
```

4.2 Time series and other Plots of the Type Intensity-over-Something

Abscissae other than c



Other abscissae may be specified by explicitly giving the model formula:

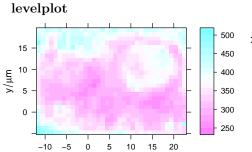
5 Levelplot

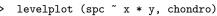
hyperSpec's levelplot can use two special column names:

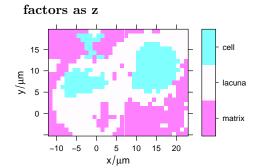
.wavelength for the wavelengths

.row for the row index (i. e. spectrum number) in the data

Besides that, it behaves exactly like level plot. Particularly, the data is given as the second argument:







x/µm

If the colour-coded value is a factor, the display is adjusted to this fact:

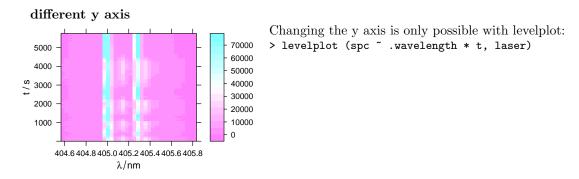
> levelplot (clusters ~ x * y, chondro)

6 Spectra Matrix

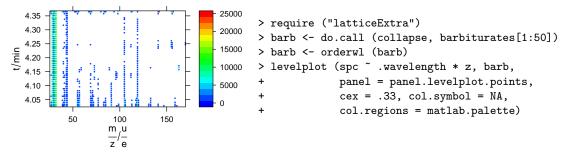
It is often useful to plot the spectra against an additional coordinate, e.g. the time for time series, the depth for depth profiles, etc.

This can be done by plot (object, "mat"). The actual plotting is done by levelplot, so the plots can be grouped or conditioned.





colour-coded points: different panel function

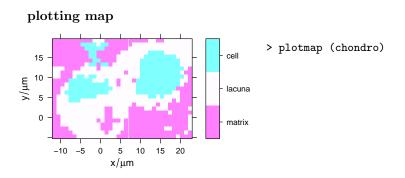


contour lines Contour lines may be added to all levelplot based plots: plot (flu, "mat", 600 contour = TRUE, 400 450 labels = TRUE, 300 250 .row col = "#00000080",at = seq (0, 700, by = 50))

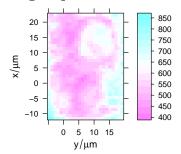
7 False-Colour Maps: plotmap

λ/nm

plotmap is a specialized version of levelplot. The spectral intensities may be summarized by a function before plotting (default: mean). The same scale is used for x and y axes (aspect = "iso").

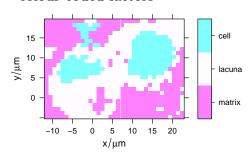


plotting maps with other than x and y



specify the colour-coded variable, abscissa and ordinate as formula: colour.coded ~ abscissa * ordinate > plotmap (chondro, spc ~ y * x)

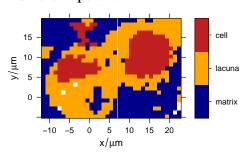
colour-coded factors



> plotmap (chondro, clusters ~ x * y)

If the colour-coded variable is a factor, each level gets its own colour, and the legend is labeled accordingly.

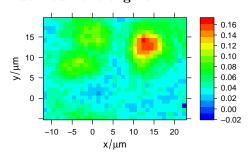
different palette



To plot with a different palette, use trellis.args = list (col.regions = palette).

> print (plotmap (chondro, clusters ~ x * y, col.regions = cluster.cols))

defined wavelengths



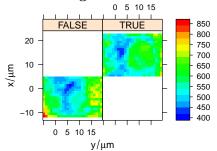
To plot a map of the average intensity at particular wavelengths use extraction:

> plotmap (chondro.preproc [, , c(728, 782, 1098,

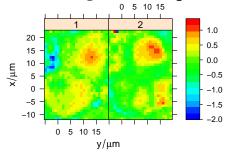
1240, 1482, 1577)],

col.regions = matlab.palette)

Conditioning



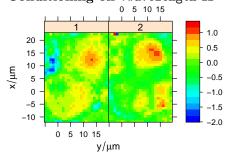
Conditioning on .wavelength



plotmap automatically applies the function in *func* before plotting. This defaults to the mean. In order to suppress this, use func = NULL. This allows conditioning on the wavelengths.

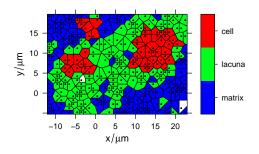
To plot e.g. the first two score maps of a principal component analysis:

Conditioning on .wavelength II



Alternatively, use levelplot directly:

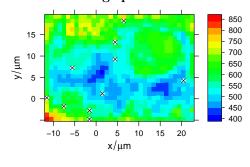
Voronoi plot



> plotvoronoi (sample (chondro, 300), clusters ~ x * y,
+ col.regions = matlab.palette(20))

Voronoi uses panel.voronoi from latticeExtra[2]. The tesselation is calculated by default using deldir[3], but tripack[4] can also be used. tripack seems to faster in general, but may "hang" with certain data sets (particularly regular grids with missing spectra as in this example). Furthermore, it is not FOSS (free and open source software).

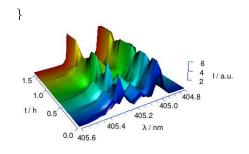
Mark missing spectra



If the spectra come from a rectangular grid, missing positions can be marked with this panel function:

8 3 D (with rgl)

3D plots with rgl



rgl[5] offers fast 3d plotting in R. As rgl's axis annotations are sometimes awkward, they may better be set manually:

9 Using ggplot2 with hyperSpec objects

hyperSpec objects do not yet directly support plotting with ggplot2 [6]. Nevertheless, ggplot2 graphics can easily be obtained.

In general, as.long.df transforms a hyperSpec object into a long-form data.frame that is suitable for qplot, ggplot, etc:

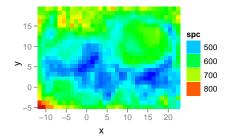
plot spectra with as.long.df

```
> df <- as.long.df (flu, rownames = TRUE)
> ggplot (df, aes (x = .wavelength, y = spc,
+ colour = c, groups = .rownames)) +
+ geom_line () +
+ xlab (labels (flu)$.wavelength) +
+ ylab (labels (flu)$spc)
```

The two special columns .wavelength and .rownames contain the wavelength axis and allow to distinguish the spectra.

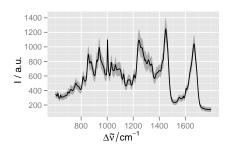
Note however, that long data.frames can be very memory consuming as they are of size $nrow \cdot nwl \times (ncol + 2)$ with respect to the dimensions of the hyperSpec object. Thus, e.g. the chondro data set (2 MB) as hyperSpec object) needs 24 MB as long-format data.frame. It is therefore highly recommended to calculate the particular data to be plotted beforehand. Moreover, depending on the particular plot as.data.frame or as.t.df may be more suitable than as.long.df:

Map with gaplot2



```
> df <- as.long.df (apply (chondro, 1, mean))
> ggplot (df,aes (x = x, y = y, fill = spc)) +
+    geom_tile() +
+    scale_fill_continuous ("spc", matlab.palette ()) +
+    scale_x_continuous (expand = c (0, 0)) +
+    scale_y_continuous (expand = c (0, 0)) +
+    coord_equal ()
```

Mean \pm standard deviation with ggplot2



Cut spectra with qqplot2



```
Cut axes can be simulated by faceting. A factor is needed that indicates the respective ranges:
```

```
> df <- paracetamol [,, c( min ~ 1800, 2800 ~ max)] / 1e4
> df <- as.long.df (df)
> df$range <- factor (df$.wavelength > 2000)
> ggplot (df, aes (x = .wavelength, y = spc)) +
+ geom_line () +
+ facet_grid (. ~ range, labeller = function (...) "",
+ scales = "free", space = "free") +
+ scale_x_continuous (breaks = seq (0, 3200, 400)) +
+ opts (strip.background = theme_blank ())
```

10 Troubleshooting

10.1 No output is produced

plotmap, plotvoronoi, levelplot, and plotc use *lattice* functions. Therefore, in loops, functions, Sweave chunks, etc. the lattice object needs to be printed explicitly by print (plotmap (object)) (R FAQ: Why do lattice/trellis graphics not work?). The same holds for *ggplot2* graphics.

For suggestions how the lattice functions can be redefined so that the result is printed without external print command, see the file vignettes.defs.

11 Interactive Graphics

hyperSpec offers basic interaction, spc.identify for spectra plots, and map.identify and map.sel.poly for maps. The first two identify points in spectra plots and map plots, respectively. map.sel.poly selects the part of a hyperSpec object that lies inside the user defined polygon.

11.1 spc.identify: finding out wavelength, intensity and spectrum

spc.identify allows to measure points in graphics produced by plotspc. It works correctly with reversed and cut wavelength axes.

```
> spc.identify (plotspc (paracetamol, wl.range = c (600 ~ 1800, 2800 ~ 3200), xoffset = 800))
```

The result is a data.frame with the indices of the spectra, the wavelength, and its intensity.

11.2 map.identify: finding a spectrum in a map plot

map.identify returns the spectra indices of the clicked points.

```
> map.identify (chondro)
```

11.3 map.sel.poly: selecting spectra inside a polygon in a map plot

map.sel.poly returns a logical indicating which spectra are inside the polygon drawn by the user:
> map.sel.poly (chondro)

11.4 Related functions provided by base graphics and lattice

For base graphics (as produced by plotspc), locator may be useful as well. It returns the clicked coordinates. Note that these are *not* transformed according to xoffset & Co.

For lattice graphics, grid.locator may be used instead. If it is not called in the panel function, a preceding call to trellis.focus is needed:

```
> plot (laser, "mat")
> trellis.focus ()
> grid.locator ()
```

identify (or panel.identify for lattice graphics) allows to identify points of the plot directly. Note that the returned indices correspond to the plotted object.

11.5 Interactively changing graphics

hyperSpec's lattice functions work with playwith[7] and latticist [8]. These packages allow easy customization of the plots and also identification of points.

References

- [1] Lemon J. Plotrix: a package in the red light district of r. R-News, 6(4):8-12, 2006.
- [2] Deepayan Sarkar and Felix Andrews. latticeExtra: Extra Graphical Utilities Based on Lattice, 2010. URL http://CRAN.R-project.org/package=latticeExtra. R package version 0.6-14.
- [3] Rolf Turner. deldir: Delaunay Triangulation and Dirichlet (Voronoi) Tessellation., 2010. URL http://CRAN.R-project.org/package=deldir. R package version 0.0-13.
- [4] Fortran code by R. J. Renka. R functions by Albrecht Gebhardt. With contributions from Stephen Eglen <stephen@anc.ed.ac.uk>, Sergei Zuyev, and Denis White. tripack: Triangulation of irregularly spaced data, 2009. URL http://CRAN.R-project.org/package=tripack. R package version 1.3-4.
- [5] Daniel Adler and Duncan Murdoch. rgl: 3D visualization device system (OpenGL), 2011. URL http://CRAN.R-project.org/package=rgl. R package version 0.92.798.
- [6] Hadley Wickham. ggplot2: elegant graphics for data analysis. Springer New York, 2009. ISBN 978-0-387-98140-6. URL http://had.co.nz/ggplot2/book.
- [7] Felix Andrews. playwith: A GUI for interactive plots using GTK+, 2010. URL http://CRAN.R-project.org/package=playwith. R package version 0.9-53.
- [8] Felix Andrews. latticist: A Graphical User Interface for Exploratory Visualisation, 2010. URL http://latticist.googlecode.com/. R package version 0.9-43.

Session Info

```
[,1]
sysname "Linux"
release "2.6.38-8-generic"
version "#42-Ubuntu SMP Mon Apr 11 03:31:24 UTC 2011"
nodename "cbdesktop"
machine "x86_64"
login "unknown"
user "cb"
```

R version 2.13.0 (2011-04-13)

Platform: x86_64-pc-linux-gnu (64-bit)

locale:

[1] LC_CTYPE=de_DE.UTF-8 LC_NUMERIC=C LC_TIME=de_DE.UTF-8 LC_MESSAGES=de_DE.UTF-8

[10] LC_TELEPHONE=C

attached base packages:

[1] grid stats graphics grDevices utils datasets methods base

other attached packages:

[1] ggplot2_0.8.9 proto_0.3-9.2 [5] plotrix_3.2 deldir_0.0-13

reshape_0.8.4 plyr_1.5.2 latticeExtra_0.6-14 RColorBrewer_1.0-2

[9] hyperSpec_0.96-20110504 lattice_0.19-26

loaded via a namespace (and not attached):

[1] digest_0.4.2 tools_2.13.0