- Under Unix, shlib modules for add-on packages are now linked against R as a shared library ('libR') if this exists. (This allows for improved embedding of R into other applications.)

- New mechanism for explicitly registering native routines in a DLL/shared library accessible via .C(), .Call(), .Fortran() and .External(). This is potentially more robust than the existing dynamic lookup, since it checks the number of arguments, type of the routine.

- New mechanism allowing registration of C routines for converting R objects to C pointers in .C() calls. Useful for references to data in other languages and libraries (e.g. C and hdf5).

- The internal ftp/http access code maintains the event loop, so you can download whilst running tcltk or Rggobi, say. It can be hooked into package **XML** too.

New features in version 1.2.3

- Support for configuration and building the Unix version of R under Mac OS X. (The 'classic' Macintosh port is 'Carbonized' and also runs under that OS.)

- dotchart() and stripchart() become the preferred names for dotplot() and stripplot(), respectively. The old names are now deprecated.

- Functions in package **ctest** now consistently use +/-Inf rather than NA for one-sided confidence intervals.

New features in version 1.2.2

- The Macintosh port becomes a full member of the R family and its sources are incorporated as from this release. See 'src/macintosh/INSTALL' for how that port is built.

- The API header files and export files 'R.exp' are released under LGPL rather than GPL to allow dynamically loaded code to be distributed under licences other than GPL.

- postscript() and xfig() devices now make use of genuine Adobe afm files, and warn if characters are used in string width or height calculations that are not in the afm files.

- Configure now uses a much expanded search list for finding a FORTRAN 77 compiler, and no longer disallows wrapper scripts for this compiler.

- New Rd markup \method{GENERIC}{CLASS} for indicating the usage of methods.

- print.ftable() and write.ftable() now have a 'digits' argument.

- undoc() has a new 'lib.loc' argument, and its first argument is now called 'package'.

Changes on CRAN

by Kurt Hornik and Friedrich Leisch

CRAN packages

The following extension packages from 'src/contrib' were added since the last newsletter.

AnalyzeIO Functions for I/O of ANALYZE image format files. By Jonathan L Marchini.

CoCoAn Two functions to compute correspondence analysis and constrained correspondence analysis and to make the associated graphical representation. By Stephane Dray.

GLMMGibbs Generalised Linear Mixed Models are an extension of Generalised Linear Models to include non-independent responses. This package allows them to be fitted by including

functions for declaring factors to be random effects, for fitting models and generic functions for examining the fits. By Jonathan Myles and David Clayton.

GeneSOM Clustering Genes using Self-Organizing Map. By Jun Yan.

Oarray Generalise the starting point of the array index, e.g. to allow x[0, 0, 0] to be the first element of x. By Jonathan Rougier.

PTAk A multiway method to decompose a tensor (array) of any order, as a generalisation of SVD also supporting non-identity metrics and penalisations. 2-way SVD with these extensions is also available. The package includes also some other multiway methods: PCAn (Tucker-n) and PARAFAC/CANDECOMP with these extensions. By Didier Leibovici.

RArcInfo This package uses the functions written by Daniel Morissette (`danmo@videotron.ca`) to read geographical information in Arc/Info V 7.x format to import the coverages into R variables. By Virgilio Gómez-Rubio.

RMySQL DataBase Interface and MySQL driver for R. By David A. James and Saikat DebRoy.

RandomFields Simulating and analysing random fields using various methods. By Martin Schlather.

SuppDists Ten distributions supplementing those built into R. Inverse Gauss, Kruskal-Wallis, Kendall's Tau, Friedman's chi squared, Spearman's rho, maximum F ratio, the Pearson product moment correlation coefficient, Johnson distributions, normal scores and generalized hypergeometric distributions. In addition two random number generators of George Marsaglia are included. By Bob Wheeler.

adapt Adaptive quadrature in up to 20 dimensions. By Thomas Lumley.

blighty Function for drawing the coastline of the United Kingdom. By David Lucy.

bqtl QTL mapping toolkit for inbred crosses and recombinant inbred lines. Includes maximum likelihood and Bayesian tools. By Charles C. Berry.

cramer Provides R routine for the multivariate non-parametric Cramer-Test. By Carsten Franz.

event.chart The function `event.chart` creates an event chart on the current graphics device. S original by J. Lee, K. Hess and J. Dubin. R port by Laura Pontiggia.

geoR Functions to perform geostatistical analysis including model-based methods. By Paulo J. Ribeiro Jr and Peter J. Diggle.

gregmisc Misc Functions written/maintained by Gregory R. Warnes. By Gregory R. Warnes. Includes code provided by William Venables and Ben Bolker.

lgtdl A very simple implementation of a class for longitudinal data. See the corresponding paper in the DSC-2001 proceedings. By Robert Gentleman.

lokern Kernel regression smoothing with adaptive local or global plug-in bandwidth selection., By Eva Herrmann (FORTRAN 77 & S original); Packaged for R by Martin Maechler.

lpridge `lpridge` and `lpepa` provide local polynomial regression algorithms including local bandwidth use. By Burkhardt Seifert (S original); packaged for R by Martin Maechler.

maxstat Maximally selected rank and Gauss statistics with several p-value approximations. By Torsten Hothorn and Berthold Lausen.

meanscore The Mean Score method for missing and auxiliary covariate data is described in the paper by Reilly & Pepe in Biometrika 1995. This likelihood-based method allows an analysis using all available cases and hence will give more efficient estimates. The method is applicable to cohort or case-control designs. By Marie Reilly PhD and Agus Salim.

netCDF Read data from netCDF files. By Thomas Lumley, based on code by Steve Oncley and Gordon Maclean.

nlrq Nonlinear quantile regression routines. By Roger Koenker and Philippe Grosjean.

odesolve This package provides an interface for the ODE solver lsoda. ODEs are expressed as R functions. By R. Woodrow Setzer with fixups by Martin Maechler.

panel Functions and datasets for fitting models to Panel data. By Robert Gentleman.

permax Functions intended to facilitate certain basic analyses of DNA array data, especially with regard to comparing expression levels between two types of tissue. By Robert J. Gray.

pinktoe Converts S tree objects into HTML/perl files. These web files can be used to interactively traverse the tree on a web browser. This web based approach to traversing trees is especially useful for large trees or for trees where the text for each variable is verbose. By Guy Nason.

sma The package contains some simple functions for exploratory microarray analysis. By Sandrine Dudoit, Yee Hwa (Jean) Yang and Benjamin Milo Bolstad, with contributions from Natalie Thorne, Ingrid Lönnstedt, and Jessica Mar.

sna A range of tools for social network analysis, including node and graph-level indices, structural distance and covariance methods, structural equivalence detection, $p*$ modeling, and network visualization. By Carter T. Butts.

strucchange Testing on structural change in linear regression relationships. It features tests/methods from the generalized fluctuation test framework as well as from the F test (Chow test) framework. This includes methods to fit, plot and test fluctuation processes

(e.g., CUSUM, MOSUM, recursive/moving estimates) and F statistics, respectively. Furthermore it is possible to monitor incoming data online. By Achim Zeileis, Friedrich Leisch, Bruce Hansen, Kurt Hornik, Christian Kleiber, and Andrea Peters.

twostage Functions for optimal design of two-stage-studies using the Mean Score method. By Marie Reilly PhD and Agus Salim.

CRAN mirrors the R packages from the Omegahat project in directory 'src/contrib/Omegahat'. The following are recent additions:

OOP OOP style classes and methods for R and S-Plus. Object references and class-based method definition are supported in the style of languages such as Java and C++. By John Chambers and Duncan Temple Lang.

RGnumeric A plugin for the Gnumeric spreadsheet that allows R functions to be called from cells within the sheet, automatic recalculation, etc. By Duncan Temple Lang.

RJavaDevice A graphics device for R that uses Java components and graphics APIs. By Duncan Temple Lang.

RSPython Allows Python programs to invoke S functions, methods, etc. and S code to call Python functionality. This uses a general foreign reference mechanism to avoid copying and translating object contents and definitions. By Duncan Temple Lang.

SLanguage Functions and C support utilities to support S language programming that can work in both R and S-Plus. By John Chambers.

SNetscape R running within Netscape for dynamic, statistical documents, with an interface to and from JavaScript and other plugins. By Duncan Temple Lang.

SXalan Process XML documents using XSL functions implemented in R and dynamically substituting output from R. By Duncan Temple Lang.

Slcc Parses C source code, allowing one to analyze and automatically generate interfaces from S to that code, including the table of S-accessible native symbols, parameter count and type information, S constructors from C objects, call graphs, etc. By Duncan Temple Lang.

Recommended packages

The number of packages on CRAN has grown over 100 recently, ranging from big toolboxes with a long history in the S community (like **nlme** or **survival**) to more specialized smaller packages (like the geo-statistical packages described later in this volume of *R News*). To make orientation for users, package developers and providers of binary distributions easier we have decided to start categorizing the contributed packages on CRAN.

Up to now there have been 2 categories of packages: Packages with priority *base* that come directly with the sources of R (like **ctest**, **mva** or **ts**) and all the rest. As a start for finer granularity we have introduced the new priority *recommended* for the following packages: **KernSmooth**, **VR** (bundle of **MASS**, **class**, **nnet**, **spatial**), **boot**, **cluster**, **foreign**, **mgcv**, **nlme**, **rpart** and **survival**.

Criteria for priority *recommended* are:

- Actively maintained and good quality code that can be installed on all (major) platforms.

- Useful for a wider audience in the sense that users "would expect this functionality" from a general purpose statistics environment like R.

- Depend only on packages with priority *base* or *recommended*.

All binary distributions of R (that are available from CRAN) will at least include the recommended packages in the future. Hence, if you write a package that depends on a recommended package it should be no problem for users of your package to meet this requirement (as most likely the recommended packages are installed anyway). In some sense the "minimum typical R installation" is the base system plus the recommended packages.

A final note: This is of course **no judgement at all** about the quality of all the other packages on CRAN. It is just the (subjective) opinion of the R Core Team what a minimum R installation should include. The criterion we used most for compiling the list of recommended packages was whether we thought a package was useful for a wide audience.

MacOS and MacOS X

Starting from release 1.3.0 of R two different versions of R for Macintosh are available at CRAN in two separate folders: 'bin/macos' and 'bin/macosx'.

The 'bin/macos' folder contains a version intended to run on Macintosh machines running System 8.6 to 9.1 and which will run on MacOS X as a carbon application, thus natively. The folder has the same structure as it was for release 1.2.x. Now the base distribution is available as two versions:

'rm130.sit' contains the base R distribution without the recommended packages;

'rm130_FULL.sit' contains the base distribution along with all recommended packages.

It is possible at a later time to download the archive 'recommended.sit' that contains the additional recommended packages not included in the base-only package. As usual the other contributed packages can be found in a separate folder.

The 'bin/macosx' folder contains a MacOS X specific build that will run on a X11 server and it is based on the Darwin kernel, i.e., it is a Unix build that runs on MacOS X. This is provided by Jan de Leeuw (`deleeuw@stat.ucla.edu`). It comes in three versions:

'R-1.3.0-OSX-base.tar.gz' has the R base distribution. It has Tcl/Tk support, but no support for GNOME.

'R-1.3.0-OSX-recommended.tar.gz' has the R base distribution plus the recommended packages. It has Tcl/Tk support, but no support for GNOME.

'R-1.3.0-OSX-full.tar.gz' has the R base distribution plus 134 compiled packages. It is compiled with both GNOME and Tcl/Tk support.

The 'bin/macosx' folder contains two folders, one containing some additional dynamic libraries upon which this port is based upon, and another giving replacements parts complied with the ATLAS optimized BLAS.

'ReadMe.txt' files are provided for both versions.

Other changes

GNU a2ps is a fairly versatile any-text-to-postscript processor, useful for typesetting source code from a wide variety of programming languages. 's.ssh', 'rd.ssh' and 'st.ssh' are a2ps style sheets for S code, Rd documentation format, and S transscripts, respectively. These will be included in the next a2ps release and are currently available from the "Other Software" page on CRAN.

Kurt Hornik
Wirtschaftsuniversität Wien, Austria
Technische Universität Wien, Austria
`Kurt.Hornik@R-project.org`

Friedrich Leisch
Technische Universität Wien, Austria
`Friedrich.Leisch@ci.tuwien.ac.at`

Date-Time Classes

by Brian D. Ripley and Kurt Hornik

Data in the form of date and/or times are common in some fields, for example times of diagnosis and death in survival analysis, trading days and times in financial time series, and dates of files. We had been considering for some time how best to handle such data in R, and it was the last of these examples that forced us to the decision to include classes for dates and times in R version 1.2.0, as part of the **base** package.

We were adding the function `file.info`. Finding information about files looks easy: Unix users take for granted listings like the following (abbreviated to fit the column width):

```
auk% ls -l
total 1189
...       948 Mar 20 14:12 AUTHORS
...      9737 Apr 24 06:44 BUGS
...     17992 Oct  7  1999 COPYING
...     26532 Feb  2 18:38 COPYING.LIB
...      4092 Feb  4 16:00 COPYRIGHTS
```

but there are a number of subtle issues that hopefully the operating system has taken care of. (The example was generated in the UK in April 2001 on a machine set to the C locale.)

- The format. Two formats are used in the extract above, one for files less than 6 months' old and one for older files. Date formats have an international standard (ISO 8601), and this is not it! In the ISO standard the first date is 2001-03-20 14.12. However, the format is not even that commonly used in the UK, which would be 20 Mar 2001 14:12. The month names indicate that this output was designed for an anglophone reader. In short, the format should depend on the *locale*.

- Time zones. Hopefully the times are in the time zone of the computer reading the files, and take daylight saving time into account, so the first time is in GMT and the second in BST. Somewhat more hopefully, this will be the case even if the files have been mounted from a machine on another continent.

  Note that this can be an issue even if one is only interested in survival times in days. Suppose a patient is diagnosed in New Zealand and dies during surgery in California?

We looked at existing solutions, the R packages **chron** and **date**. These seem designed for dates of the accuracy of a day (although **chron** allows partial days), are US-centric and do not take account of time zones. It was clear we had to look elsewhere for a comprehensive solution.