



Forecast verification using R

Matthew Pocerlich
NCAR – Research Application
Laboratory
pocerlich@ucar.edu

Why R?

- R is Open Source!
- R is free
- Runs on all operating systems
- All code is visible
- Over 1,000 packages contributed.
- Very large, active user base.

The R Community

- Developers
 - R Core Group (17 members), only 2 have left since 1997
 - Major update in April/October (freeze dates, beta versions, bug tracking, ...)
- Mailing lists
 - Help list ~ 100 messages/day, archived, searchable.
- Packages – more than 1000 packages (9/2006)
- R News - ~quarterly news letter

www.r-project.org

- Contains everything
 - Source code
 - Documentation
 - Newsletter
 - Mailing list
 - Packages

Introduction to our environment.

- The Xemacs environment
 - Buffer
 - R prompt
- Quick demo

Benefits of Using Scripts

- Scripts are text files, are readable
- Document methods
- Allows results to be reproduced
- Allows code to be reused

Emacs Speaks Statistics

- Provides syntax-based
- Object name completion
- Key stroke short cuts
- Command history
- See summary of shortcuts
- Alt-x R to invoke R with Xemacs.

A word about R in Windows

- Added GUI features
- Build in script
- <control> R sends a line or highlighted section into R.
- Install package with GUIs
- Save graphics by point and click.

Packages in R

- Contributed by people world wide.
- Apply and extend R capabilities to meet the needs of specific communities.
- Accompany many statistical textbooks
- Accompany applied articles (Adrian Raftery, Doug Nychka, Tilman Gneiting, Barbara Casati, Matt Briggs)

Verification, climate study and meteorology related packages

- verification
- fields
- radiosondes
- extRemes
- BMA(Bayesian Model Averaging)
- BMAensemble
- circular
- ROCR

Packages

- Packages must be installed to call.
- Packages must be called to use.
- Base packages are installed by default.

Everything is an object – everything has a class

- numeric; binary
- character
- factor
- matrix – only contains numbers
- Dates – POSIXct, POSIXlt
- dataframe – numbers, text and dates – same lengths
- List – different types of data, different lengths

Verification package

- Organized base on forecast/observation types
- Functions include
 - Skill Scores
 - Heidke, Peirce, ...
 - Graphics
 - ROC
 - Attribute diagrams
 - Conditional quantile plots
 - Other contributions
 - Scale intensity (Casati)
 - Measurement error(Briggs)
 - Circular rps (Gneiting)