

Plotting Tools for a Molecular Dynamics Trajectory Database

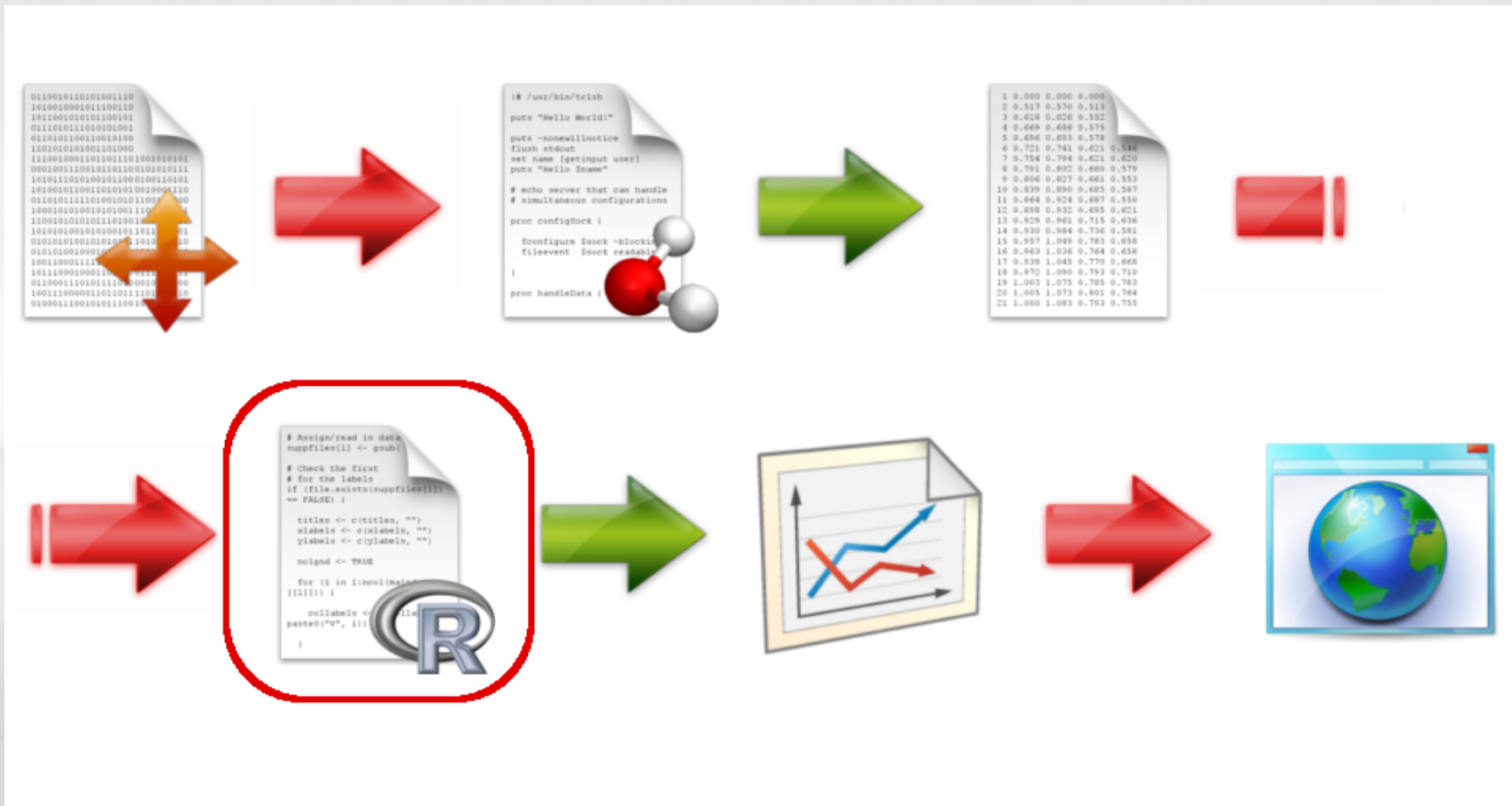
Jeremiah Boswell
Western Michigan University

Dr. Thomas C. Bishop
Louisiana Tech University

Topics of Discussion

- | Big Picture
- | Overview of R
- | Development Experience
- | Outcome

The Workflow



What is R?

Programming Language

- Developed by Bell Laboratories
- Based on S programming language
- Feels like a scripting language
 - Source code is interpreted, not compiled
 - “Maximum processing per statement” approach

Software environment

- Can run interactively
 - GUI on Windows systems



Early Stages

!Previous experience

- !Knew a little syntax, some functions
- !No file I/O, standalone source file experience
- !General programming knowledge

!Linux!

!Documentation

!Proof-of-concept programs

- !One per plot type (line, box, violin)
- !Read special test files

Gaining Momentum

┆rmsd2plot.R

- ┆Could be run non-interactively

- ┆One analysis type, but plot types and file structure could vary

- ┆Much time spent on special checks and error handling

┆First specification of objective (First dat2plot.R)

- ┆Multiple plots per image

 - ┆Many bugs

 - ┆rand2dat.tcl for test data

The Real Challenge

Second specification of objective

- Much more data available for testing

- Much more functionality needed

 - Range, interactivity, number of plots, labels

Development questions raised

- Program per analysis, number of plots, or only one?

- How to intuitively get information from user?

Time constraints

- Two late evenings later...

Results

- | dat2plot.R, iteration 3

- | Single program for all operations

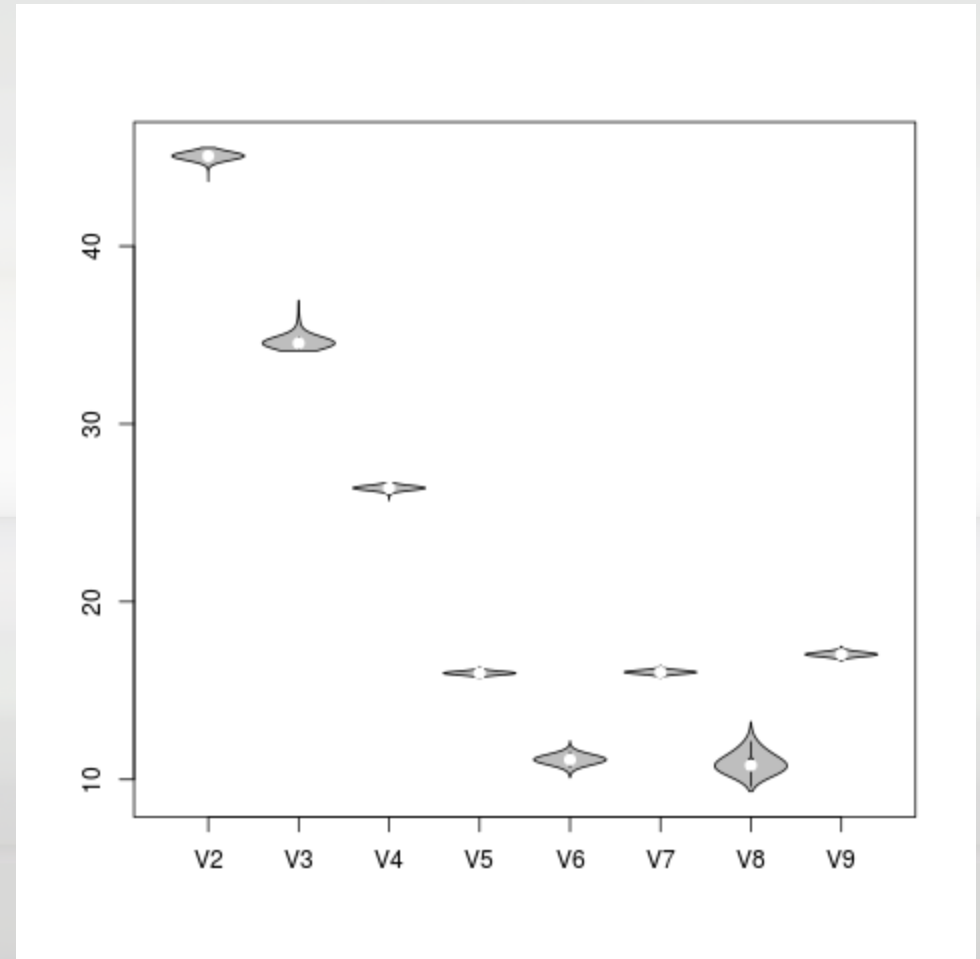
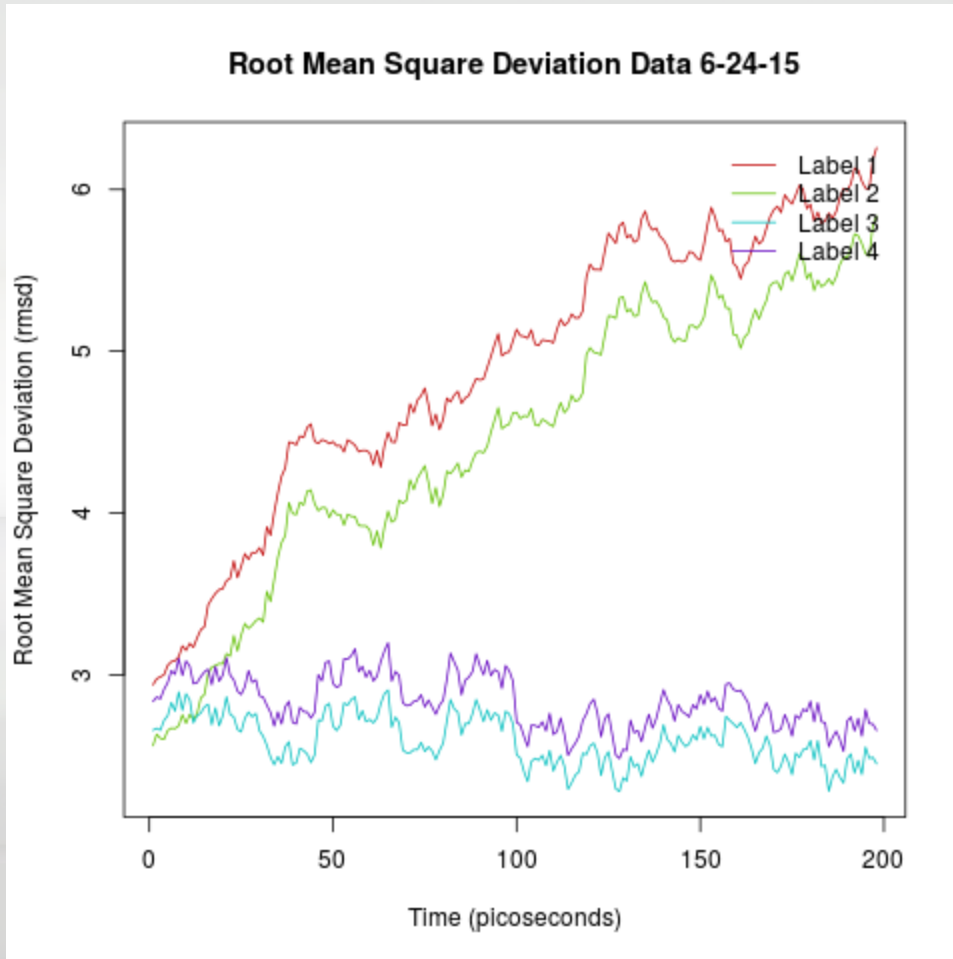
- | Code length increased by five times!

- | Plot types and data range user arguments

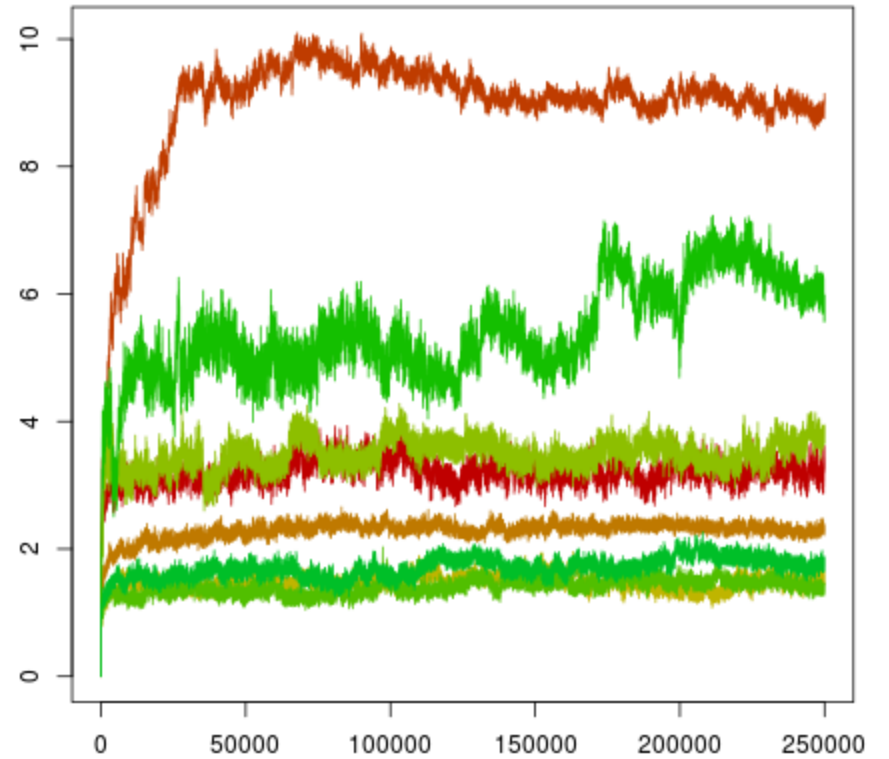
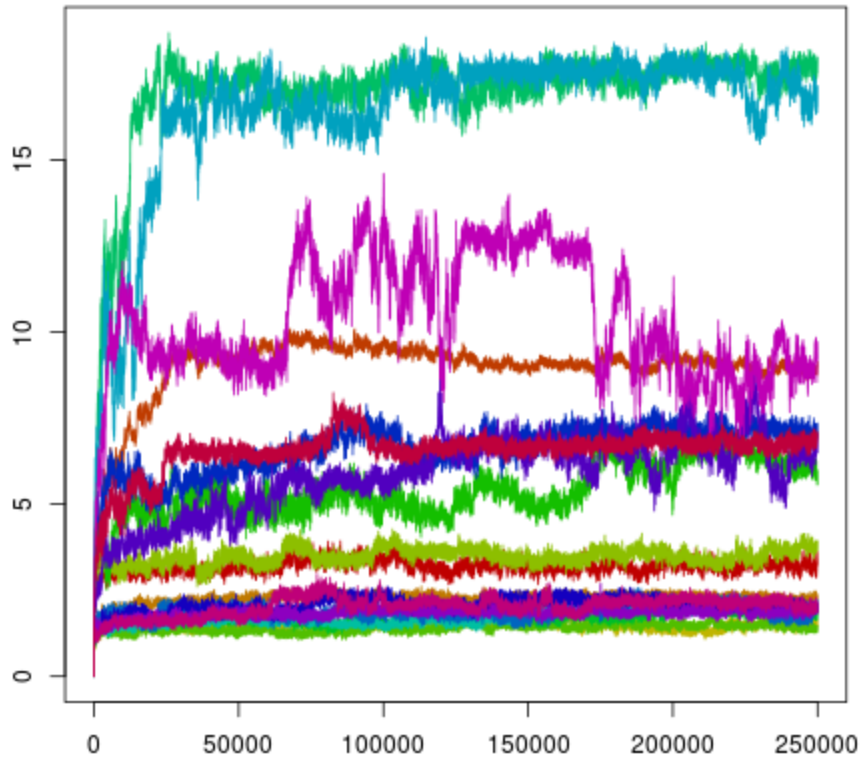
- | Interactive mode

- | **Pics or it didn't happen!**

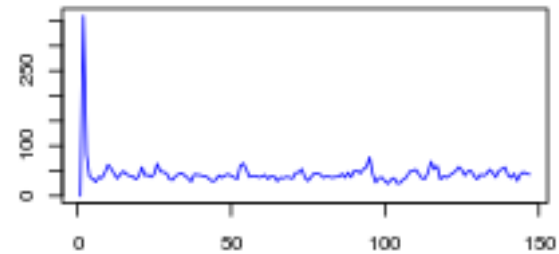
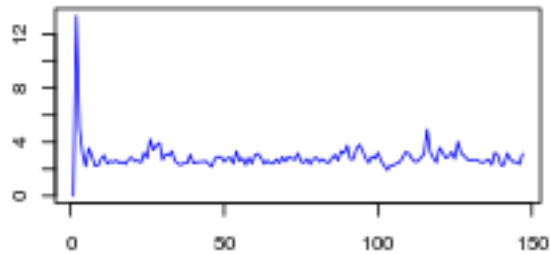
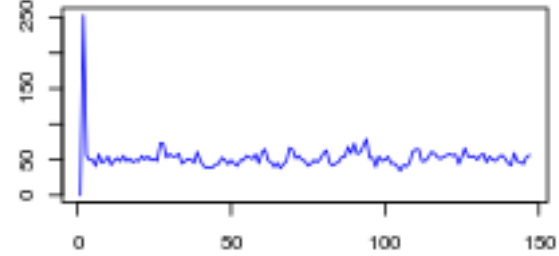
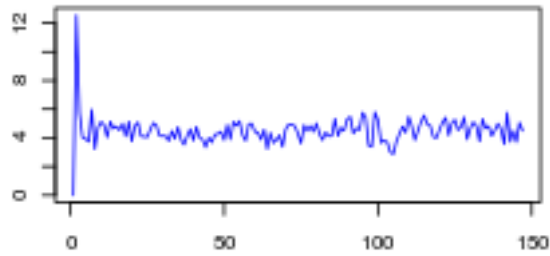
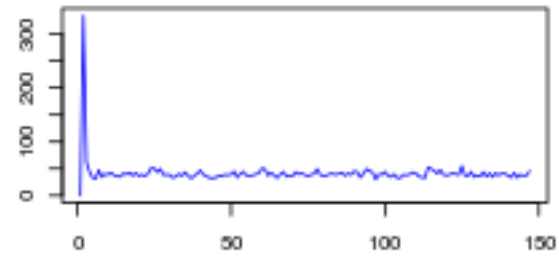
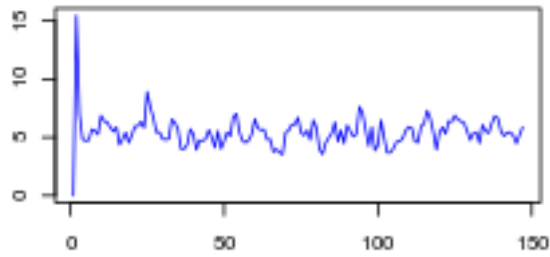
Legend Display and Plot Types



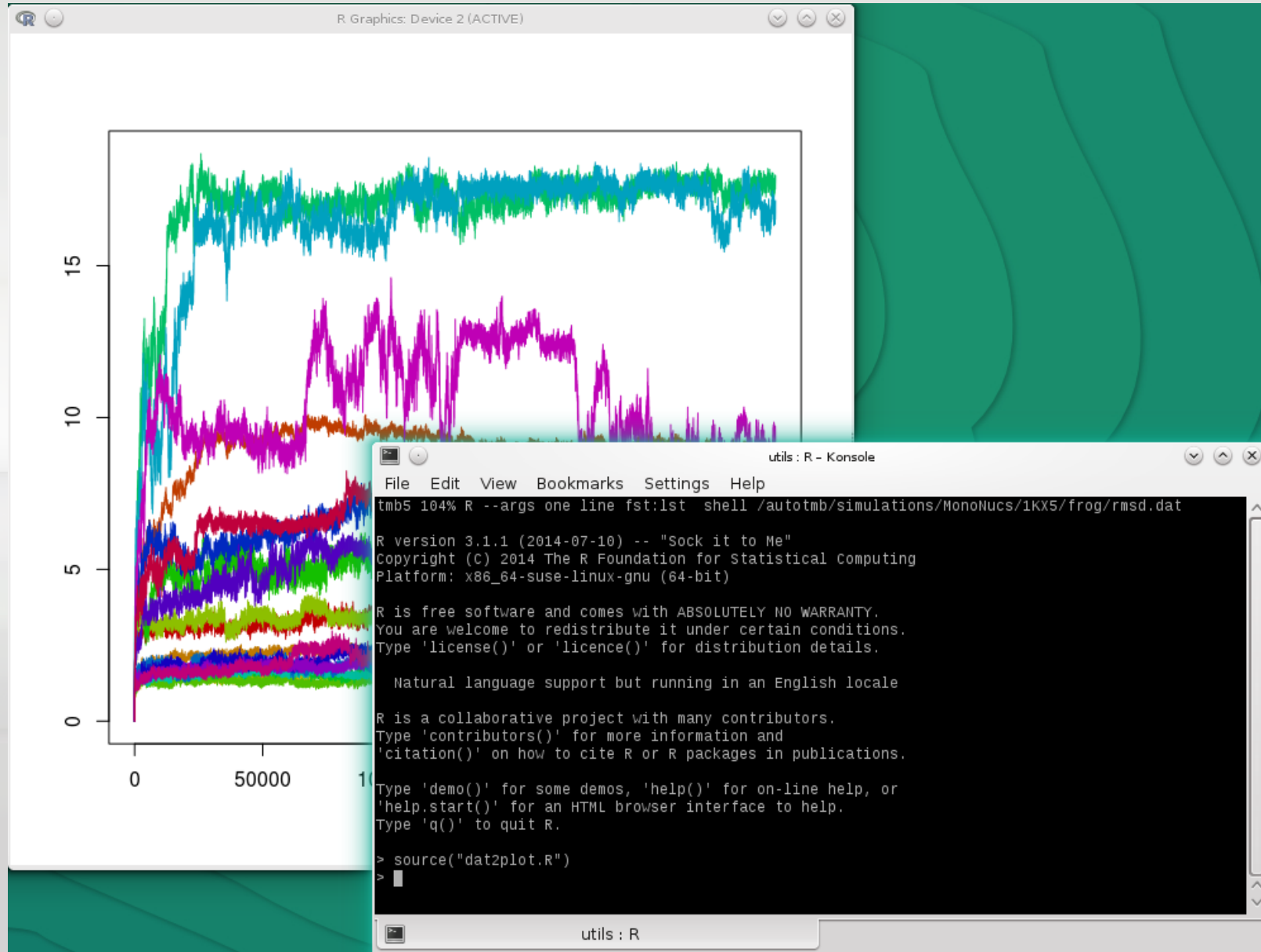
Range Specification



Array Plot Functionality



Interactive Shell Mode



The image displays two overlapping windows from an R environment. The background window, titled "R Graphics: Device 2 (ACTIVE)", shows a time-series plot with multiple colored lines (green, cyan, magenta, brown, red, blue, purple, yellow, orange) plotted against a white background. The x-axis ranges from 0 to 1, with a major tick at 50000. The y-axis ranges from 0 to 15, with major ticks at 0, 5, 10, and 15. The lines show an initial sharp increase followed by a period of relative stability with some noise.

The foreground window, titled "utils : R - Konsole", is a terminal window showing the R startup sequence. The text in the terminal is as follows:

```
File Edit View Bookmarks Settings Help
tmb5 104% R --args one line fst:lst shell /autotmb/simulations/MonoNucs/1KX5/frog/rmsd.dat
R version 3.1.1 (2014-07-10) -- "Sock it to Me"
Copyright (C) 2014 The R Foundation for Statistical Computing
Platform: x86_64-suse-linux-gnu (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> source("dat2plot.R")
>
```

Acknowledgements

Dr. Thomas C. Bishop

Seneca Joseph, Anik Karan, and Anthony Agee
(and Sandeep Perla)

Lindsay Gouedy and Jessica Wasserman

Louisiana Tech Faculty @ morning lectures

Fellow REU students

Support by the National Science Foundation
under the NSF EPSCoR Cooperative Agreement
No. EPS-1003897 with additional support from the
Louisiana Board of Regents.

Questions?