

crawl: An R Package for Modeling Animal Movement from Satellite Telemetry Data

Devin S. Johnson

*Alaska Fisheries Science Center,
National Marine Fisheries Service,
Seattle, Washington
Email: devin.johnson@noaa.gov*

2015 NMFS PST Mini-Symposium
November 18, 2015



R Packages: Low-Level Software Tools for Fisheries Management

Devin S. Johnson

*Alaska Fisheries Science Center,
National Marine Fisheries Service,
Seattle, Washington
Email: devin.johnson@noaa.gov*

2015 NMFS PST Mini-Symposium
November 18, 2015



The Problems

- NMFS science centers collect a large amount of animal telemetry data every year
- Used to determine where animals spend time in various behaviors
- Predictions have to be made for times with no observed location
- Must be done for many animals
- Must often be repeated for many projects

The Problems

- NMFS science centers collect a large amount of animal telemetry data every year
- Used to determine where animals spend time in various behaviors
- Predictions have to be made for times with no observed location
- Must be done for many animals
- Must often be repeated for many projects

The Problems

- NMFS science centers collect a large amount of animal telemetry data every year
- Used to determine where animals spend time in various behaviors
- Predictions have to be made for times with no observed location
- Must be done for many animals
- Must often be repeated for many projects

The Problems

- NMFS science centers collect a large amount of animal telemetry data every year
- Used to determine where animals spend time in various behaviors
- Predictions have to be made for times with no observed location
- **Must be done for many animals**
- **Must often be repeated for many projects**

The Problems

- NMFS science centers collect a large amount of animal telemetry data every year
- Used to determine where animals spend time in various behaviors
- Predictions have to be made for times with no observed location
- Must be done for many animals
- Must often be repeated for many projects

correlated random walk library of R Functions

- `crawl` is an add-on package for the R statistical environment
- Contains a collection of functions and objects that allow a user to fit continuous-time correlated random walk models to telemetry data
- As part of the PST project:
 - updated to use compiled C++ code for computationally efficient model fitting
 - Detailed example documentation added (vignette)
 - Web-based app included for data pre-processing (Shiny app)
- Obtained from the Comprehensive R Archive Network (CRAN) the “official” repository for R packages

correlated random walk library of R Functions

- `crawl` is an add-on package for the R statistical environment
- Contains a collection of functions and objects that allow a user to fit continuous-time correlated random walk models to telemetry data
- As part of the PST project:
 - updated to use compiled C++ code for computationally efficient model fitting
 - Detailed example documentation added (vignette)
 - Web-based app included for data pre-processing (Shiny app)
- Obtained from the Comprehensive R Archive Network (CRAN) the “official” repository for R packages

correlated random walk library of R Functions

- `crawl` is an add-on package for the R statistical environment
- Contains a collection of functions and objects that allow a user to fit continuous-time correlated random walk models to telemetry data
- As part of the PST project:
 - updated to use compiled C++ code for computationally efficient model fitting
 - Detailed example documentation added (vignette)
 - Web-based app included for data pre-processing (Shiny app)
- Obtained from the Comprehensive R Archive Network (CRAN) the “official” repository for R packages

correlated random walk library of R Functions

- `crawl` is an add-on package for the R statistical environment
- Contains a collection of functions and objects that allow a user to fit continuous-time correlated random walk models to telemetry data
- As part of the PST project:
 - updated to use compiled C++ code for computationally efficient model fitting
 - Detailed example documentation added (vignette)
 - Web-based app included for data pre-processing (Shiny app)
- Obtained from the Comprehensive R Archive Network (CRAN) the “official” repository for R packages

correlated random walk library of R Functions

- `crawl` is an add-on package for the R statistical environment
- Contains a collection of functions and objects that allow a user to fit continuous-time correlated random walk models to telemetry data
- As part of the PST project:
 - updated to use compiled C++ code for computationally efficient model fitting
 - Detailed example documentation added (vignette)
 - Web-based app included for data pre-processing (Shiny app)
- Obtained from the Comprehensive R Archive Network (CRAN) the “official” repository for R packages

correlated random walk library of R Functions

- `crawl` is an add-on package for the R statistical environment
- Contains a collection of functions and objects that allow a user to fit continuous-time correlated random walk models to telemetry data
- As part of the PST project:
 - updated to use compiled C++ code for computationally efficient model fitting
 - Detailed example documentation added (vignette)
 - Web-based app included for data pre-processing (Shiny app)
- Obtained from the Comprehensive R Archive Network (CRAN) the “official” repository for R packages

The R Environment

- The R environment is a language and program for data analysis
- One of the most popular computing languages in the world
- R is an interactive high-level computing language
(No compiling source code)
- Users can customize R with packages
 - Contain functions and objects for custom analysis
 - Can post these packages for others to use
- Best of all it's open source and free!

The R Environment

- The R environment is a language and program for data analysis
- One of the most popular computing languages in the world
- R is an interactive high-level computing language
(No compiling source code)
- Users can customize R with packages
 - Contain functions and objects for custom analysis
 - Can post these packages for others to use
- Best of all it's open source and free!

The R Environment

- The R environment is a language and program for data analysis
- One of the most popular computing languages in the world
- R is an interactive high-level computing language
(No compiling source code)
- Users can customize R with packages
 - Contain functions and objects for custom analysis
 - Can post these packages for others to use
- Best of all it's open source and free!

The R Environment

- The R environment is a language and program for data analysis
- One of the most popular computing languages in the world
- R is an interactive high-level computing language
(No compiling source code)
- Users can customize R with packages
 - Contain functions and objects for custom analysis
 - Can post these packages for others to use
- Best of all it's open source and free!

The R Environment

- The R environment is a language and program for data analysis
- One of the most popular computing languages in the world
- R is an interactive high-level computing language
(No compiling source code)
- Users can customize R with packages
 - Contain functions and objects for custom analysis
 - Can post these packages for others to use
- Best of all it's open source and free!

The R Environment

- The R environment is a language and program for data analysis
- One of the most popular computing languages in the world
- R is an interactive high-level computing language
(No compiling source code)
- Users can customize R with packages
 - Contain functions and objects for custom analysis
 - Can post these packages for others to use
- Best of all it's open source and free!

The R Environment

- The R environment is a language and program for data analysis
- One of the most popular computing languages in the world
- R is an interactive high-level computing language
(No compiling source code)
- Users can customize R with packages
 - Contain functions and objects for custom analysis
 - Can post these packages for others to use
- **Best of all it's open source and free!**

R Packages

- Collection of custom objects and functions for analysis
- Allows an complex analysis to be repeated / altered on demand
- Scientists can easily share newly developed statistical methodology
- Provides avenues for documentation and reproducibility
- Problem:
 - Produces low-level tools
 - Substantial R knowledge necessary

R Packages

- **Collection of custom objects and functions for analysis**
- Allows an complex analysis to be repeated / altered on demand
- Scientists can easily share newly developed statistical methodology
- Provides avenues for documentation and reproducibility
- Problem:
 - Produces low-level tools
 - Substantial R knowledge necessary

R Packages

- Collection of custom objects and functions for analysis
- **Allows an complex analysis to be repeated / altered on demand**
- **Scientists can easily share newly developed statistical methodology**
- **Provides avenues for documentation and reproducibility**
- **Problem:**
 - **Produces low-level tools**
 - **Substantial R knowledge necessary**

R Packages

- Collection of custom objects and functions for analysis
- Allows an complex analysis to be repeated / altered on demand
- Scientists can easily share newly developed statistical methodology
- Provides avenues for documentation and reproducibility
- Problem:
 - Produces low-level tools
 - Substantial R knowledge necessary

R Packages

- Collection of custom objects and functions for analysis
- Allows an complex analysis to be repeated / altered on demand
- Scientists can easily share newly developed statistical methodology
- Provides avenues for documentation and reproducibility
- Problem:
 - Produces low-level tools
 - Substantial R knowledge necessary

R Packages

- Collection of custom objects and functions for analysis
- Allows an complex analysis to be repeated / altered on demand
- Scientists can easily share newly developed statistical methodology
- Provides avenues for documentation and reproducibility
- **Problem:**
 - Produces low-level tools
 - Substantial R knowledge necessary

R Packages

- Collection of custom objects and functions for analysis
- Allows an complex analysis to be repeated / altered on demand
- Scientists can easily share newly developed statistical methodology
- Provides avenues for documentation and reproducibility
- Problem:
 - Produces low-level tools
 - Substantial R knowledge necessary

Mid-level Packages?

- R packages can be made to require a minimum of R skill
- Long-form documentation and examples can be included
- Shiny apps can be added for some GUI interaction
- Users can avoid a large amount of scripting and programming

Mid-level Packages?

- R packages can be made to require a minimum of R skill
- Long-form documentation and examples can be included
- Shiny apps can be added for some GUI interaction
- Users can avoid a large amount of scripting and programming

Mid-level Packages?

- R packages can be made to require a minimum of R skill
- Long-form documentation and examples can be included
- Shiny apps can be added for some GUI interaction
- Users can avoid a large amount of scripting and programming

Mid-level Packages?

- R packages can be made to require a minimum of R skill
- Long-form documentation and examples can be included
- Shiny apps can be added for some GUI interaction
- Users can avoid a large amount of scripting and programming

Mid-level Packages?

- R packages can be made to require a minimum of R skill
- Long-form documentation and examples can be included
- Shiny apps can be added for some GUI interaction
- Users can avoid a large amount of scripting and programming

A brief intermission

A crawl demo

agTrend: another R package from NMML

- Estimate abundance trends for aggregations of survey sites
- E.g, abundance trend of Western Aleutian Steller sea lions
- Problem:
 - Not all sites are surveyed in a given year
 - How can abundance be aggregated over all sites each year?
- Solution:
 - Bayesian nonparametric regression analysis
 - A relatively complex analysis!
- Package constructed directly to satisfy management needs
- Analysis repeatable with on-demand aggregation groups
e.g., trends for all rookeries that start with the letter S!

agTrend: another R package from NMML

- Estimate abundance trends for aggregations of survey sites
- E.g, abundance trend of Western Aleutian Steller sea lions
- Problem:
 - Not all sites are surveyed in a given year
 - How can abundance be aggregated over all sites each year?
- Solution:
 - Bayesian nonparametric regression analysis
 - A relatively complex analysis!
- Package constructed directly to satisfy management needs
- Analysis repeatable with on-demand aggregation groups
e.g., trends for all rookeries that start with the letter S!

agTrend: another R package from NMML

- Estimate abundance trends for aggregations of survey sites
- E.g, abundance trend of Western Aleutian Steller sea lions
- Problem:
 - Not all sites are surveyed in a given year
 - How can abundance be aggregated over all sites each year?
- Solution:
 - Bayesian nonparametric regression analysis
 - A relatively complex analysis!
- Package constructed directly to satisfy management needs
- Analysis repeatable with on-demand aggregation groups e.g., trends for all rookeries that start with the letter S!

agTrend: another R package from NMML

- Estimate abundance trends for aggregations of survey sites
- E.g, abundance trend of Western Aleutian Steller sea lions
- **Problem:**
 - Not all sites are surveyed in a given year
 - How can abundance be aggregated over all sites each year?
- **Solution:**
 - Bayesian nonparametric regression analysis
 - A relatively complex analysis!
- Package constructed directly to satisfy management needs
- Analysis repeatable with on-demand aggregation groups
e.g., trends for all rookeries that start with the letter S!

agTrend: another R package from NMML

- Estimate abundance trends for aggregations of survey sites
- E.g, abundance trend of Western Aleutian Steller sea lions
- Problem:
 - Not all sites are surveyed in a given year
 - How can abundance be aggregated over all sites each year?
- Solution:
 - Bayesian nonparametric regression analysis
 - A relatively complex analysis!
- Package constructed directly to satisfy management needs
- Analysis repeatable with on-demand aggregation groups
e.g., trends for all rookeries that start with the letter S!

agTrend: another R package from NMML

- Estimate abundance trends for aggregations of survey sites
- E.g, abundance trend of Western Aleutian Steller sea lions
- Problem:
 - Not all sites are surveyed in a given year
 - How can abundance be aggregated over all sites each year?
- Solution:
 - Bayesian nonparametric regression analysis
 - A relatively complex analysis!
- Package constructed directly to satisfy management needs
- Analysis repeatable with on-demand aggregation groups
e.g., trends for all rookeries that start with the letter S!

agTrend: another R package from NMML

- Estimate abundance trends for aggregations of survey sites
- E.g, abundance trend of Western Aleutian Steller sea lions
- Problem:
 - Not all sites are surveyed in a given year
 - How can abundance be aggregated over all sites each year?
- Solution:
 - Bayesian nonparametric regression analysis
 - **A relatively complex analysis!**
- Package constructed directly to satisfy management needs
- Analysis repeatable with on-demand aggregation groups
e.g., trends for all rookeries that start with the letter S!

agTrend: another R package from NMML

- Estimate abundance trends for aggregations of survey sites
- E.g, abundance trend of Western Aleutian Steller sea lions
- Problem:
 - Not all sites are surveyed in a given year
 - How can abundance be aggregated over all sites each year?
- Solution:
 - Bayesian nonparametric regression analysis
 - A relatively complex analysis!
- Package constructed directly to satisfy management needs
- Analysis repeatable with on-demand aggregation groups
e.g., trends for all rookeries that start with the letter S!

agTrend: another R package from NMML

- Estimate abundance trends for aggregations of survey sites
- E.g, abundance trend of Western Aleutian Steller sea lions
- Problem:
 - Not all sites are surveyed in a given year
 - How can abundance be aggregated over all sites each year?
- Solution:
 - Bayesian nonparametric regression analysis
 - A relatively complex analysis!
- Package constructed directly to satisfy management needs
- Analysis repeatable with on-demand aggregation groups
e.g., trends for all rookeries that start with the letter S!

R Packages by NMML/AFSC

CRAN Mid-level production packages

Package	Analysis of ...
crawl	Telemetry data
agTrend	Population trends
stocc	Spatial occupancy data
marked	Mark-recapture data
RMark	Mark-recapture data
DSpat	Line transect data
hierachicalDS	Line transect data
mrds	Line transect data

github Low-level or development packages

<https://github.com/NMML>

Conclusion

- R provides a way to bundle complex analysis into tidy packages
- Can create demos, apps, long-form documentation
- R is platform independent
- Multiple locations to host packages: CRAN (official) or GitHub
- Still requires use of R console
- Next: Shiny allows one to use the power of R (packages) without the R console

Conclusion

- R provides a way to bundle complex analysis into tidy packages
- Can create demos, apps, long-form documentation
- R is platform independent
- Multiple locations to host packages: CRAN (official) or GitHub
- Still requires use of R console
- Next: Shiny allows one to use the power of R (packages) without the R console

Conclusion

- R provides a way to bundle complex analysis into tidy packages
- Can create demos, apps, long-form documentation
- **R is platform independent**
- Multiple locations to host packages: CRAN (official) or GitHub
- Still requires use of R console
- Next: Shiny allows one to use the power of R (packages) without the R console

Conclusion

- R provides a way to bundle complex analysis into tidy packages
- Can create demos, apps, long-form documentation
- R is platform independent
- Multiple locations to host packages: CRAN (official) or GitHub
- Still requires use of R console
- Next: Shiny allows one to use the power of R (packages) without the R console

Conclusion

- R provides a way to bundle complex analysis into tidy packages
- Can create demos, apps, long-form documentation
- R is platform independent
- Multiple locations to host packages: CRAN (official) or GitHub
- Still requires use of R console
- Next: Shiny allows one to use the power of R (packages) without the R console

Conclusion

- R provides a way to bundle complex analysis into tidy packages
- Can create demos, apps, long-form documentation
- R is platform independent
- Multiple locations to host packages: CRAN (official) or GitHub
- Still requires use of R console
- Next: Shiny allows one to use the power of R (packages) without the R console

Questions?

This PDF was later amended to make the document 508 compliant.

