AssortedFunciton R Package 697 Report

For my 697 Project I developed an R package for the analysis of ecological data. This package has three main focuses. The first focus is on the analysis of time series state data and network analysis. The second focus of this data set is on identifying and analyzing discrete events of varying lengths in time series data. With the third focus being on the efficient analysis of spatial data in a raster format.

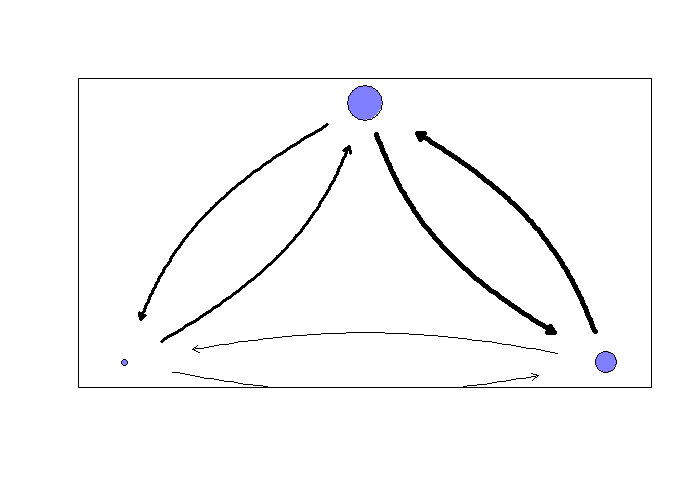
**Section 1**

Understanding where animals move is highly important, however in the marine world can be difficult. One strategy for determining the movement of marine animals is using passive acoustic telemetry. This involves tagging an animal with an acoustic transmitter and releasing it into the wild. The researcher then deploys underwater listening stations to record if the individual is within 200m of the receiver. While informative this data can be very large and difficult to work with, analyze, and generalize to the species as a whole to characterize their movements. Several functions in this package were built to address this problem. One analysis technique is based on network theory and on transition matrix models, with the general idea being you can summarize all the location or “states” of the individual by examining what states they an individual is in as well as what the subsequent state that the individual is in. This is represented as an n x n matrix in which the number inside represent the number of times an individual transistioned from one state (column) to another state (row). Below is an example of a list of states Table 1. And an example of what the matrix would look like (Table 2).

|  |  |
| --- | --- |
| Time | State |
| 1 | State 1 |
| 2 | State 1 |
| 3 | State 3 |
| 4 | State 3 |
| 5 | State 2 |
| 6 | State 3 |
| 7 | State 1 |
| 8 | State 1 |

|  |  |  |  |
| --- | --- | --- | --- |
|  | State 1 | State 2 | State 3 |
| State 1 | 2 |  | 1 |
| State 2 |  |  | 1 |
| State 3 | 1 | 1 | 1 |

I created a function transMat that will generate these transition matrixes from the raw list of states. Furthermore I created a function transBub that will display these transition matrixes visually (Figure 1). Where the size of the bubble represents the number of times the individual was in each state. While the thickness of the line represents the number of times each transition occurred.



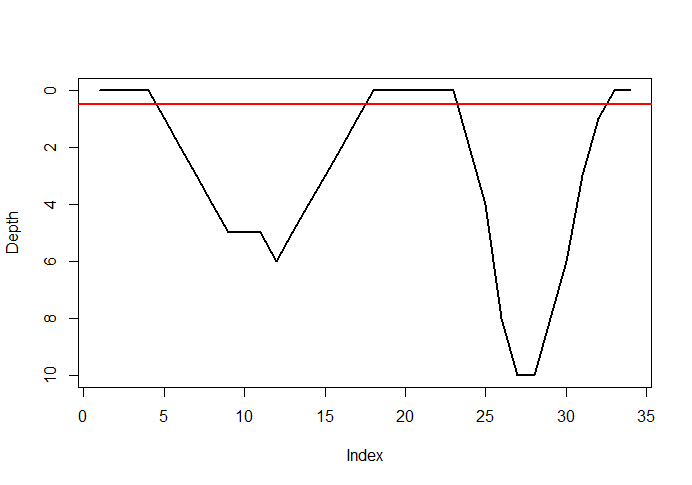
This functionality has been used in multiple conference presentations as well as in thesis and soon publications.

Additionally there are several other functions in AssFunc that are useful for the analysis of acoustic telemetry data. There is logGen, which generates a log file. This is a file that summarizes the number of detections on each receiver, across any unit of time you give it. For instance it will generate the number of detections on each receiver every 5 minutes, 1hr or 1 day, between a supplied start and end time. This can be useful for calculating daily presence or in generating the input for particle filters.

**Section 2**

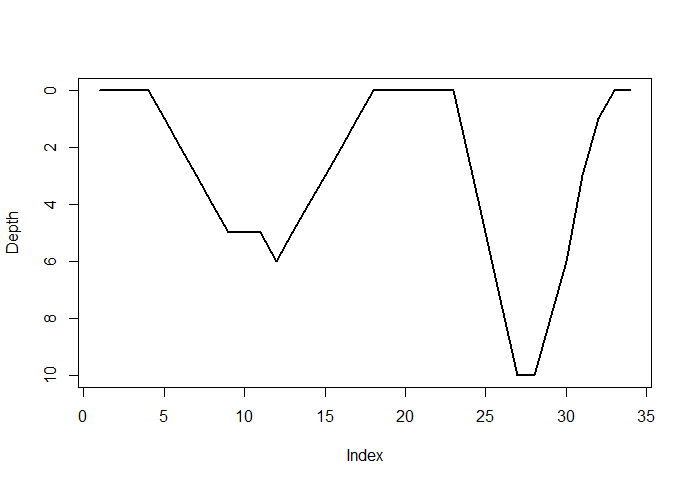
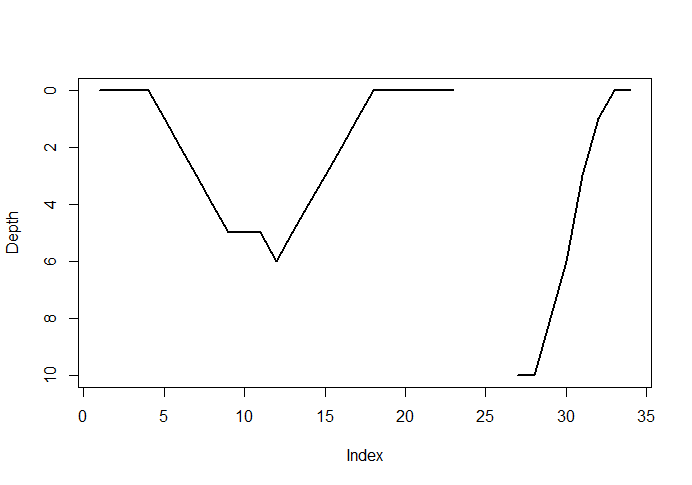
With the increased ability to collect data, there is an increase in the demand for automated analysis. Many datasets are thousands or millions of datapoints long where it is not feasible to by hand identify events. This is commonly true for datasets that are time series in nature that might be collected at fine scale resolutions (>1hz). This package has a events functions which purpose was to help analyze this data.

For air breathing diving animals such as turtles, identifying the characteristics of individual’s dives is important and often used, with each dive being an independent dataset. However identifying each dive individually is problematic, and hence why I created the events function. A typical data set for a turtle might look somewhat like figure 2.



The red like represents the classification of a dive. So there are two discrete dives, one to 6 meters and one to 10 meters. Using the event function it will tell you the start and end of each dive based on the threshold value. So this command event(dive>.5) will provide and output of 5,17 and 23,32. Which corresponds to the start point (5 and 23) and the end point(17, 32) of each dive respectively.

I additionally realized that this was much more useful than merely dive data and could be used for any time series event, with one of the most useful being interpolating during NAs. Using the eventInterp function, it will fill in any event (usually where there is a string of NA) with linearly interpolated values based on the values on either side of the event. For instance if a segment of data was missing from the previous depth record (Figure 3A) it could be interpolated based on the values on each side of the string of NAs (Figure 3b)



The last function in the event series is eventFunc. This function applies a function to every event individually and returns a values for each one. For instance for the turtle dive data, you couple apply a function to get the maximum value of each dive, or the duration of each dive separately.

**Section 3**

The last series of functions is focused on the analysis of spatial data. There is a function to calculate the distance in meters between two points that are in latitude and longitude. As well as a function to convert a series of latitude and longitudes into a Cartesian coordinate frame, such as meters from some local origin.

Finally the analysis of large rasters can be very slow using the focal function from the raster package. In order to speed of this focal command, I made my own called focalfast. This will apply any function you supply over the entire raster over whatever focal window you supply. This function performs the same as focal but approximately 10 times faster. However, it comes at the cost of flexibility, with reductions in how data such as boarders are handled.

**Section 4**

Finally there is a few other assorted functions like recLoc, which generates the a matrix with the name and location of each state from a list of states. Or sunCalc which generates the sunrise and sunset time for any day of the year, for any supplied latitude and longitude.