David Sweeney

Professor Stacy DeRuiter, Mathematics & Statistics

A Toolkit for Analysis of Data from Animal-borne Movement Loggers

The goal of our research project was to develop a freely available, fully-documented software tool kit to enable high-quality, reproducible, sophisticated analyses of data from high-resolution animal-borne tags, while also facilitating comparison of results between studies, tag-types, and computational software. This project was conducted in collaboration with a team of researchers from the University of St. Andrews. Before this project began, the most comprehensive and authoritative set of tools for tag data analysis was a set of Matlab scripts, which were developed by Mark Johnson and others. Our project expanded upon and enhanced this already accessible toolkit, enabling its tools to be used on data from a variety of different tag types. We developed more tools used for the visualization and statistical analysis of multi-sensor data. We also improved documentation of previously designed tools by creating a user guide to all the tools. Finally, we created additional versions of all Matlab tools in both R and Octave, allowing for the free access to this toolkit. All of this was done thanks to a generous grant from the US Office of Naval Research.

Within the grand scope of this project, the focus of my work involved translating the Matlab tools' scripts into R's distinct software language. This was quite a task given that I had never even heard of Matlab before joining this research group. Another part of my work involved thoroughly testing the developed tools to ensure that they were all working as intended. Thankfully, Octave has become increasingly flexible in recent years, allowing functions written in Matlab code to be run without much trouble. This made the process of debugging our tools easier, but with over 60 different tools, debugging could be quite challenging at times. I also spent a large portion of my time writing and formatting helpful documentation of each tool in each software program. In order to ensure that future users of these tools understand how each tool was designed to be used, proper documentation was required. This job involved not only writing helpful documentation within each program, but also writing help pages within our project website (http://www.animaltags.org./doku.php?id=start).

Around the middle of the summer, I was tasked with composing a tool that would allow for the automated detection of animal behavior events from signal data. The tool I developed is titled detect_peaks (http://www.animaltags.org./doku.php?id=tagwiki:tools:processing:detect_peaks) and it works by taking a signal vector and determining the times at which peaks in the signal spike above a predetermined threshold level. The peaks that exceed the threshold are thus determined to be significant animal behavior events. One of the many possible uses of this function is determining the time at which a marine organism performs a prey capture attempt. Initial testing of this function on data from a lunge feeding whale shows my tool to have a 96.67% true positive detection rate and a 11.16% false positive detection rate. Throughout the process of writing this tool, I realized just how far I have progressed as a software developer over the course of this summer. In addition, my knowledge of the field of biologging has increased tremendously by reading a lot of scientific literature related to behavioral event detection methods.

This project will culminate in a trip by our team to St. Andrews, Scotland to help lead a three-day workshop (August 7 - 9) on how to use our toolkit. There will be about 30 participants at the workshop, representing at least 6 countries and 20 home institutions. The workshop will consist of a series of lectures and practical/working sessions, all designed to provide the participants with a solid foundation for the use of these tools.