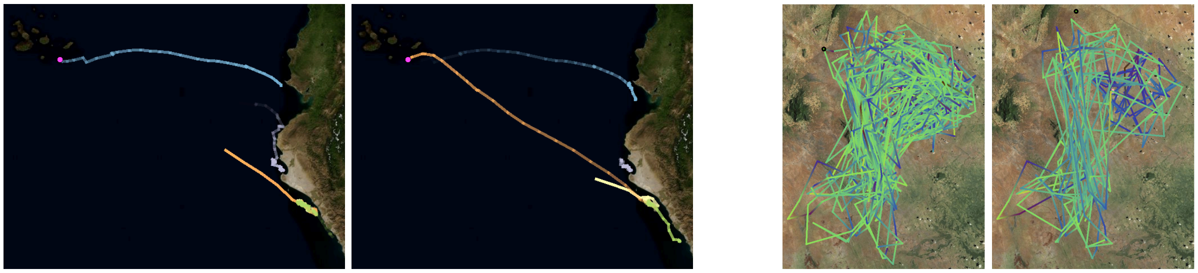
Lessons Learned from Tool Development for Animal Movement Analysis

Florian Mansmann* University of Konstanz

David Spretke† University of Konstanz

Halldor Janetzko‡ University of Konstanz



(a) Track Animation: Interval 1 (b) Track Animation: Interval 2 (c) Normal Tracks (d) Simplified Tracks

Figure 1: Key geographic features of the Animal Ecology Explorer system

Abstract

Animal tracking has made significant technological progress in the last years with lighter, more capable sensors and better battery usage. However, analyzing these movement tracks is still challenging for most biologists since it requires extensive computer science knowledge for data preprocessing, geographical mapping and visualization to interpret the collected data. While our overall goal is to simplify this process for a wide range of animal trackers, this paper specifically focuses on the lessons that we learned in the process of developing tools for biologists.

1 Introduction

In a joint project with biologists from the Max Plank Institute for Ornithology we had the task to develop an interactive tool for animal tracking biologists to be used on the collected datasets in Movebank (see http://www.movebank.org). Like in other cross-discipline projects, communication is key to understand each other's worlds. Learning about current workflows, analysis tasks, systems and each other's languages enabled us to develop a tool to be used in practice. Current plans are to integrate the tool as a java web start application into the Movebank website to allow exploration.

This paper shares three principle lessons that we learned in the process of collaborating with animal trackers involving data preprocessing, data exploration and specific tool requirements.

2 Lesson 1: Data can be messy

We all experience in our daily information visualization work that real-world data are often messy and data preprocessing is the most expensive and time-consuming step. Dealing with data coming from animal traces is often even more complicated: not only that noise and outlier will occur, but the rate of measurements may vary over time. The variation of sampling rates will influence all following analysis and visualization steps as the parts with high sample rates are over-represented in the low-sampled measurements.

*e-mail: Florian.Mansmann@uni-konstanz.de

†e-mail: David.Spretke@uni-konstanz.de

‡e-mail: Halldor.Janetzko@uni-konstanz.de

Furthermore, it is often important to enrich the data with calculated statistical values, like correlations or aggregations. Biologists therefore may use external applications such as MATLAB or R to add further features to the data set. The lack of an integrative platform for movement analysis allowing a straightforward enrichment and an easy-to-use visualization hinders biologists from concentrating on their main task: the generation and validation of hypotheses. As a consequence we have plans to add the possibility to write and execute R-scripts to our tool and integrate the results directly into the currently analyzed trajectories.

3 Lesson 2: Exploring Movement in Context

Animal movement in itself is rarely the key interest of biological studies. It is rather the cross-relation between movement with different additional measurements (e.g. for acceleration or body temperature) or external condition such as weather conditions. To enable comprehensive analyses, a system thus first needs to provide a powerful data linkage between the trajectories and external data such as weather and second needs to enable investigation of movement in the context of additional sensor measurements and external conditions. Since we built on top of the Movebank system, which enables data linkage of studies with weather conditions, we could fully focus on how to explore this movement in context.

Our first approach was to link line charts showing external data or additional sensory measurements with the map representation showing the trajectories. Brushing an area on the map or a part of the chart highlights the marked data in all linked views (see Fig. 2). Movement of the selection can then be used for detailed investigation, such as replaying the animal behavior by moving the selection in a chart along the x-axis representing time as shown in Figure 1 a) and b).

When analyzing several tracked animals at once in one chart we learned from the biologists that not only the individuals' concurrent behavior in time is of interest, but also simultaneous behavior with respect to the distance traveled. In other word, setting cumulative distance for the x-axis of a chart will almost automatically align the migration patterns of two individuals despite the fact that their measurements were taken at different points in time. In addition, more space is given in the chart to this relatively short but interesting time segment.

After the initial approach of encoding values by coloring bars of constant height to compare overall trends, we discovered Hori-

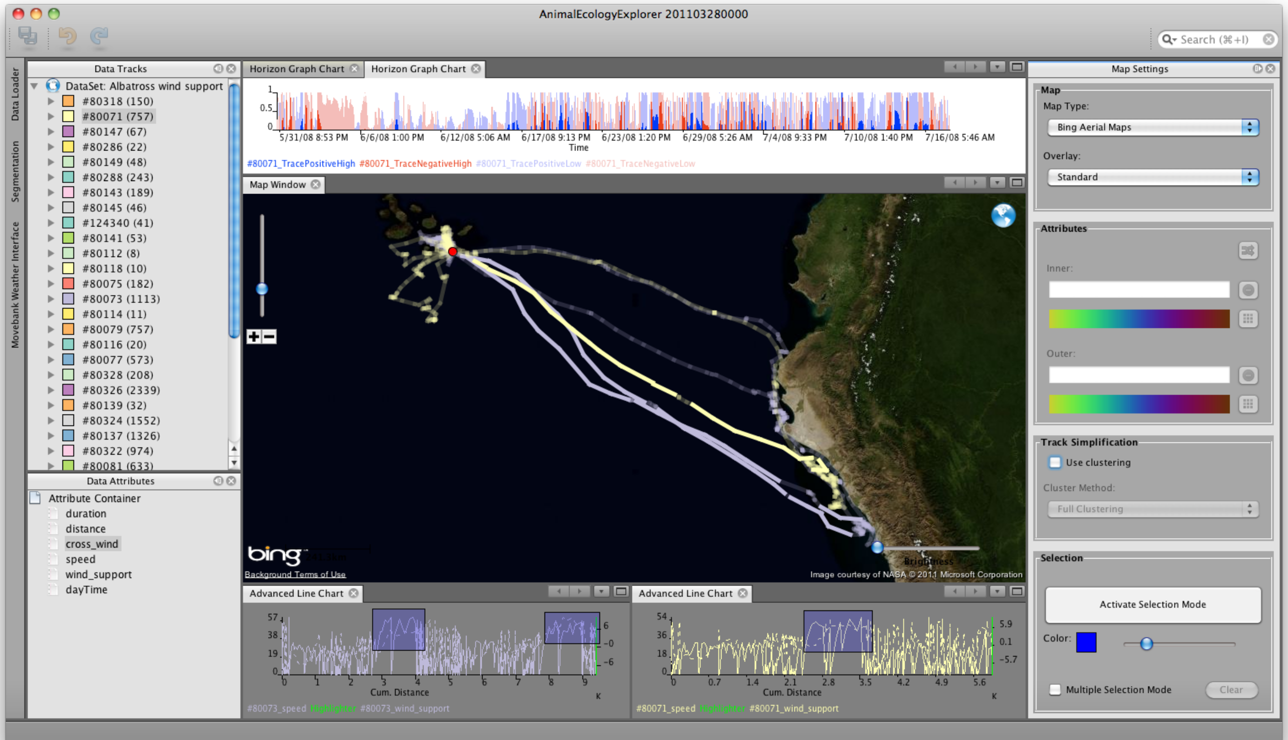


Figure 2: Visual interface of the Animal Ecology Explorer showing albatrosses migrating from the Galapagos islands to the mainland and back. On the left side the study and the tracked animals can be selected. The horizon graph at the top shows disturbing cross-wind over time, whereas the two line charts at the bottom show cummulated distance vs. speed and wind support. Brushing high speed and high wind support regions in the charts highlights the respective segments on the north-west return routes in the linked map.

zon Graphs [1], which are almost as scalable while showing more details of the underlying time series.

Further analysis tasks involved segmentation and annotation of animal movement trajectories to meaningfully group behaviorally similar segments for a joint annotation. A major issue in this case is the setting of segmentation and clustering parameters, which in our tool can be done interactively. Meaningful pre-settings, fast displaying of the results and iterative adjustment of parameters thereby support the biologists in this difficult task.

In general, we learned that visualizing behavior of migrating birds works pretty well on a map, whereas highly overlapping stationary animal movement can result in significant visual clutter. A preliminary solution to this problem is the track simplification option as shown in Figure 1 c) and d). This option removes points from the trajectory that do not contribute to the space and time characteristics of the respective trajectory and can be seen as a compromise between exactness and visibility. Close points in space and time are thereby replaced by only one representative point. In addition to significantly improving visibility, this option can also have positive effects on the tool's performance. However, we are currently still investigating alternative map or abstract representations to deal with this issue.

4 Lesson 3: Tool Requirements

The visual interactive analysis of animal movement behavior is a challenging task and requires a highly flexible tool, which supports the analyst in a suitable manner. In the majority of cases, biologists do not have access to high resolution displays for their analysis tasks. They often have to work on systems with small screens, which get overloaded quite fast providing only little space for the visualization windows and/or settings boxes. To address that problem, our Animal Ecology Explorer system builds upon a space-efficient window system, including functionality such as hide/unhide, maximize/minimize, dock/undock, and drag-and-drop for all windows. This allows the analyst to use our tool, even if the screen size is small.

Furthermore, we have learned that biologists get a lot of great ideas for new features while using our tool. Our system was thus designed in a modular fashion to easily integrate new functionality. A generic drag-and-drop concept for tracks and data attributes enables the biologists to generate suitable visualizations for their analysis task.

5 Conclusions

This paper showed a number of principle lessons learned for data analysis tool development for animal trackers. In summary, there is a huge potential for simplifying the current data analysis process in an integrated data preparation, enrichment and exploration system.

Acknowledgements

We thank the Movebank team M. Wikelski, R. Kays, K. Safi, B. Kranstauber and S. Davidson for the many rounds of feedback and for sharing their tracking data.

References

[1] S. Few. Time on the horizon., Jun/Jul 2008. Visual Business Intelligence Newsletter, Online at http://www.perceptualedge.com/articles/visual_business_intelligence/time_on_the_horizon.pdf.