4-24-2020

Animaltracker: Streamlining Spatio-Temporal Analysis and Visualization of High Sampling Rate Animal Data

Thea Sukianto
Boise State University

Joe Champion
Boise State University

Sergio Arispe
Oregon State University

Dylan Mikesell
Boise State University

Animaltracker: Streamlining Spatio-Temporal Analysis and Visualization of High Sampling Rate Animal Data

Abstract

As GPS technology continues to advance, spatio-temporal data is collected by such devices at an increasingly higher sampling rate. This data can provide valuable insights on animal behavior patterns. However, processing hundreds of thousands of observations manually harbors potential for inefficiency. Furthermore, in-depth statistical analysis and visualization of the data would often require the use of additional tools. For the task of streamlining the processing, analysis, and visualization pipeline for cattle GPS collar data, we develop the animaltracker package in the statistical programming language R. With R Shiny, we construct a three-panel web application as the core feature of the package. The first panel allows for user-driven customization of data processing through filtering, elevation augmentation, and exporting. On the second panel, we visualize elevation time series and sampling rate by animal among others. On the final panel, we calculate statistical summaries for user-selected variables. More advanced visualization and analysis is possible through R functions included in the package, while data validation and outlier detection is also possible through a second Shiny application. By compiling these utilities in a framework designed for high-volume data such as the R environment, we provide a convenient platform that maximizes the efficiency of the pipeline.

ANIMALTRACKER

STREAMLINING SPATIO-TEMPORAL ANALYSIS AND VISUALIZATION OF HIGH SAMPLING RATE ANIMAL DATA

[1]THEA SUKIANTO, [1]JOE CHAMPION, [2]SERGIO ARISPE, [1]DYLAN MIKESELL

BOISE STATE UNIVERSITY[1], OREGON STATE UNIVERSITY[2]

BOISE STATE UNIVERSITY

Oregon State University

OVERVIEW

Scientists studying animal behavior have begun using low-cost geolocation (GPS) trackers, but struggle to migrate large raw data files (100,000+ observations) using spreadsheet software. *Animaltracker* is a new *R* package to automate data cleaning and visualization of GPS logs. The package bundles *R Shiny* web applications with data processing functions to provide customizable data processing, interactive plots of animal locations over time, and statistical tools for augmenting, analyzing, and exporting the cleaned data.

DATA PROCESSING

The data workflow is summarized in a simple 4-step process:

1. Data Upload

- Upload archive of raw GPS data (.csv format)

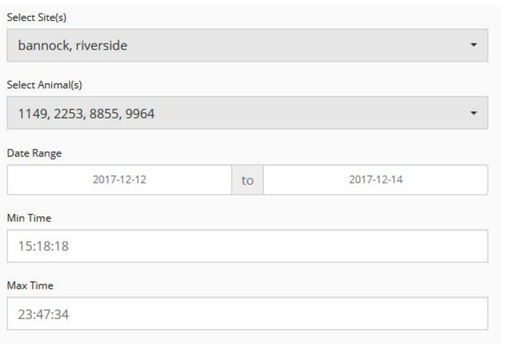
2. Select Data

- 4 filtering options:
  - **Site** (data source), **Animal** (ID number), **Date** (min/max), **Time** (24h min/max)

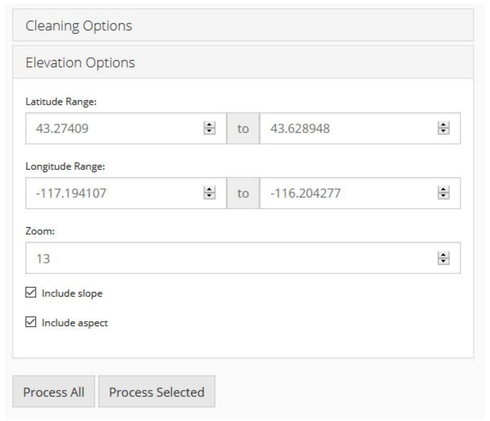
3. Data Processing

- **Cleaning Options:**
  - **Discard erroneous data** (yes/no)
    - e.g. (0,0) coordinates
- **Elevation Append Options:**
  - **Latitude** (min/max), **Longitude** (min/max), **Zoom** (geodesic zoom 1-18), **Slope** (yes/no), **Aspect** (yes/no)
- Data selected in **(2)** with options in **(3)** is then processed automatically

4. Data Download



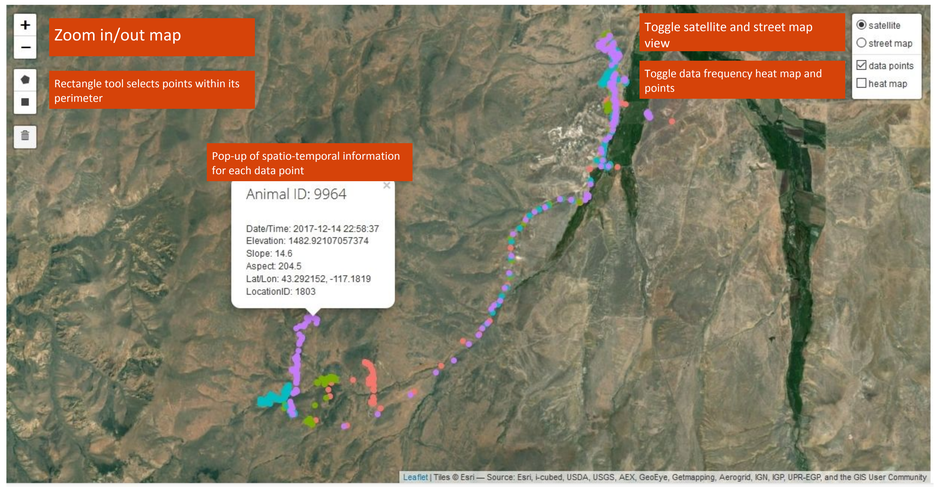
Panel of options for Step 2.



Panel of options for Step 3.

MAPPING

The data selected in **(2)** of processing is displayed on an interactive map.



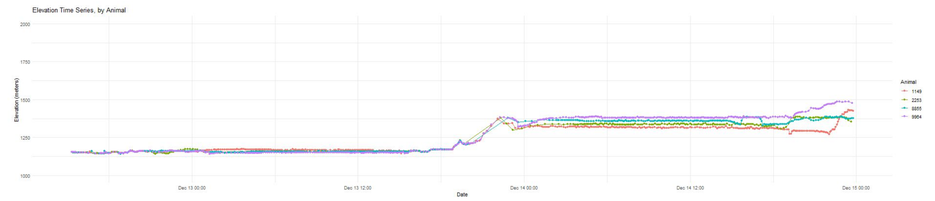
Map Info

Current zoom level = 13

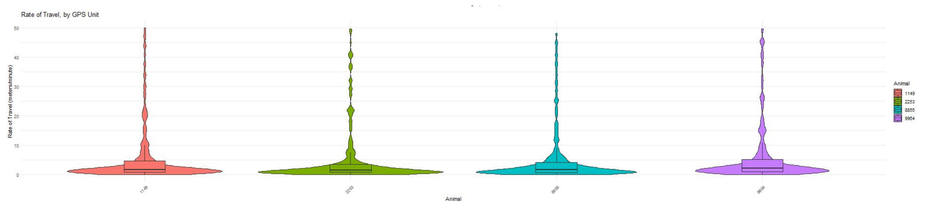
Map zoom level dynamically changes default zoom in data processing options

DATA ANALYSIS

After the raw data is uploaded, plots and statistical summaries are generated by the app. As the data is processed, the plots and summaries change to reflect the new data.



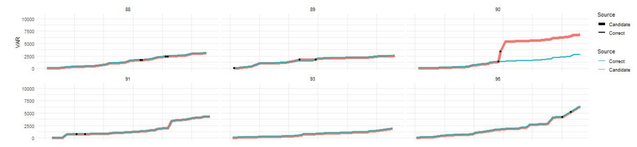
Time series plot of elevation (m) for animals 1149, 2253, 8855, and 9964.



Violin plot of rate of travel (m/min) for animals 1149, 2253, 8855, and 9964.

DATA VALIDATION

- Compare animal GPS data processed by different methods
  - Flag system
    - Flag distance, rate, and course when above a user-defined threshold
  - Extreme value detection
    - Modified z-score method
  - Time series plots of variables for each animal



Time series plots of cumulative distance (m) for animals 88, 89, 90, 91, 93, and 96, processed by different methods.

ADDITIONAL FEATURES

R functions bundled in the animaltracker R package

- Data Visualization
  - Measurement intervals
    - Quantile-quantile (Q-Q) plots
    - Histograms
  - Boxplots for altitude distribution by animal ID
  - Comparison
    - Side-by-side violin plots
    - Faceted time series plots
- App functionality from R CMD

FUTURE DEVELOPMENT

- Optimize RAM consumption of elevation lookup
- .kmz, .kml, and .shp file support
  - Visualize fencing
  - Visualize water sources
    - Calculate distance to closest water source from fixed point
- More advanced time series analysis

View animaltracker v0.1.0 docs on CRAN