
ANIMOVE
P1.y=diag(c(0, 0.001, 0.001))
playPar(mov.model=~1, err.model=list(x=~errX, y=~errY),
data=nfsNew, fixPar=c(NA, 1, NA, 1, NA, NA, NA, NA))
<- crwMLE(mov.model=~1, err.model=list(x=~errX, y=~errY),
data=nfsNew, coord=c("longitude", "latitude"),
Time.name="Time", initial.state=initial.drift,
fixPar=c(NA, 1, NA, 1, NA, NA, NA, NA),
control=list(maxit=2000,trace=1, REPORT=10),

August 2023

Movement data in R

Segmentation

Segmentation

Segmentation allows discerning behaviorally or contextually more homogeneous sections of an otherwise heterogeneous trajectory.

It provides better insight in the processes driving different movement modes or just to be used to do separate analysis. Helps answering the following questions:

- What is the animal doing? (Exploratory)
- Why is the animal doing what it's doing? (Explanatory)
- Can we anticipate the movement response of the animal? (Predictive)

Segmentation

Presentation based mainly on this paper:

Journal of Animal Ecology

British Ecological Society

Journal of Animal Ecology 2016, **85**, 69–84 doi: 10.1111/1365-2656.12379

SPECIAL FEATURE: STUCK IN MOTION? RECONNECTING QUESTIONS AND TOOLS IN MOVEMENT ECOLOGY

What is the animal doing? Tools for exploring behavioural structure in animal movements

Eliezer Gurarie[1,2]*, Chloe Bracis[3], Maria Delgado[4,5], Trevor D. Meckley[6], Ilpo Kojola[7] and C. Michael Wagner[6]

[1]*Department of Biology, University of Maryland, College Park, MD 20742, USA;* [2]*School of Environmental and Forest Sciences, University of Washington, Seattle, WA 98195, USA;* [3]*Quantitative Ecology and Resource Management, University of Washington, Seattle, WA 98195, USA;* [4]*Department of Biosciences, University of Helsinki, 00014 Helsinki, Finland;* [5]*Research Unit of Biodiversity (UMIB, UO-CSIC-PA), Oviedo University – Campus Mieres, 33600 Mieres, Spain;* [6]*Department of Fisheries and Wildlife, Michigan State University, East Lansing, MI 48824, USA; and* [7]*Natural Resources Institute, Box 16, FI-96301 Rovaniemi, Finland*

Edelhoff *et al. Movement Ecology* (2016) 4:21
DOI 10.1186/s40462-016-0086-5

Movement Ecology

REVIEW Open Access

CrossMark

Path segmentation for beginners: an overview of current methods for detecting changes in animal movement patterns

Hendrik Edelhoff*, Johannes Signer and Niko Balkenhol

This is also an interesting read

Segmentation Methods: overview

- **Supervised** (e.g. random forest, to a certain extent K-means): some knowledge is expected from the user, who has to provide certain information prior the implementation, such as:
  - Training dataset
  - Number of expected groups/classes
- **Unsupervised** (e.g. EMbC and others presented today): the user selects the algorithm but no previous knowledge is required (user knowledge needed later to interpret the classification)

Segmentation Methods: overview

- **Classification** (e.g. K-means, K-medians, BPMM, EMbC, random forest): apply clustering algorithms and selection criteria to determine the number of partitions a data set can parsimoniously be subdivided into (origins in machine learning)
- **Phenomenological** (e.g. BCPA): identify the structure or periodicity of movement paths (origins in time series and signal processing)
- **Mechanistic** (e.g. MRW): estimate transitions between states in a statistical framework (derived from the generalized mixed modelling tradition). Note: Not designed to be exploratory, but to address explanatory questions

Segmentation Methods: overview

- **Univariate clustering** (e.g. BPMM, BCPA, lavielle, multimode): grouping based on only one variable.
- **Bivariate clustering** (e.g. EMbC, MRW, segclust2D): grouping based on two variables.
- **Multivariate clustering** (e.g. K-means, K-medians, GMM, random forest): grouping based on one up to several variables.

K-means & K-medians

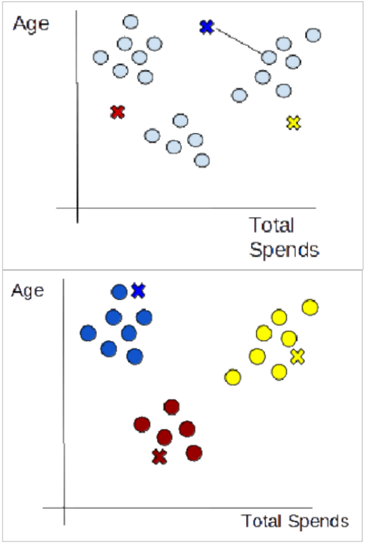
- Based on: multivariate - any numerical variable of choice
- Calculation: K-means clustering is a famous unsupervised machine learning algorithm that tries to group items in clusters minimising squared euclidean distance within cluster. K is the number of groups/clusters. Centroids based on clusters' means or medians get recalculated until assignment to group is stable.
- Parameters to specify: K, matrix of variables

\+ of simple use and interpretation

\+ very flexible and can include many variables

\- results depend on the initial centroids (initialization methods)

\- results depend on the chosen variables in the matrix



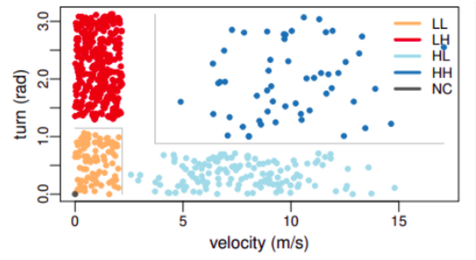
EMbC (Expectation-Maximization Binary Clustering)

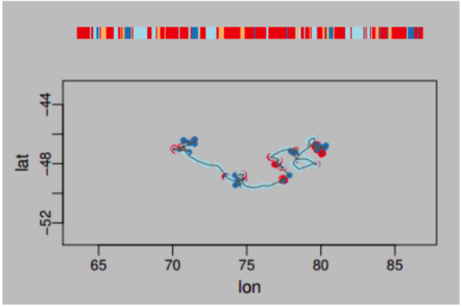
- Based on: bivariate - speed and turning angles by default (others possible)
- Calculation: splits time series into binary categories finding a best split based on machine learning processes

\+ Potential simplicity of the interpretation

\+ Good at finding breaking points

\- Sensitive to initial range of the variable, sometimes variable transformation is needed





Segmentation Methods: BCPA

BCPA (Behavioural Change Point Analysis)

- Based on: univariate - persistence in velocity = speed*cos(turning angle) (others possible)
- Calculation: identifies changes in movement parameter values by sweeping an analysis window over the time series and locating the most likely change point within each window

  Response variable is modeled as a continuous-time autoregressive Gaussian process with mean, standard deviation and characteristic time-scale of autocorrelation, with an unknown number of behavioural states
- Two types:
  - Smooth: estimates parameters for each data point averaged over all windows
  - Flat: finds the most frequently chosen change points and estimates the parameter values within each phase
- Parameters to specify: sensitivity parameter K; window size w; cluster width for flat BCPA

\+ Data can be autocorrelated and gapy/irregular

\+ No need to pre-specify number of states

\+ Calculates an error

\- Lots of tweaking

Segmentation Methods: BPMM

BPMM (Bayesian Partitioning of Markov Models)

- Based on: univariate - step length (others possible)
- Calculation: algorithm uses randomized likelihood-based method to: (i) choose the number of candidate models & (ii) partition the time series into a specific sequence of those models
- **Assumption:** Movement composed of discrete number of homogeneous processes (homogeneous within), and transition between this models is Markovian (only depend on previous state, no long-term memory), mov. variable independent between observations within state
- Parameters to specify: movement variable, set of candidate models (states)
- Track characteristics: track has to be regularized, movement variable should not be autocorrelated

\+ Works excellent if changes are caused by changes in speed/step length

\- Can't deal with autocorrelation. i.e. data need to be independent

\- Can't deal with gapy/irregular data, i.e. data need to be regularized

Segmentation Methods: MRW

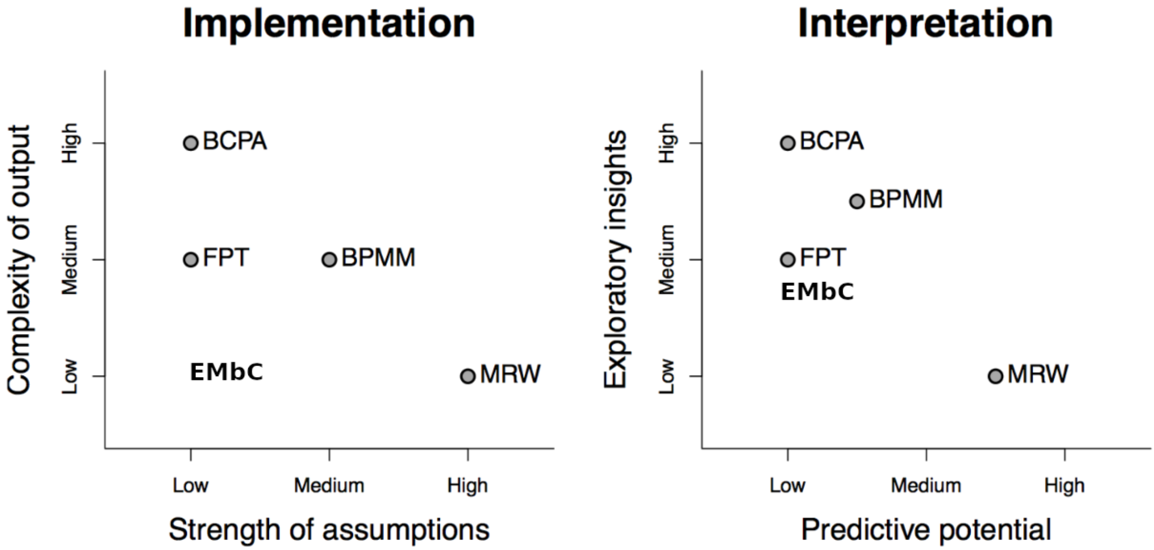
MRW (Multistrate Random Walk)

- Based on: bivariate - step length and turning angles by default (others possible)
- Calculation: correlated random walk (CRW) in which the turning angles are assumed to be drawn from some circular distribution, and step lengths are assumed to be drawn from some unimodal positive distribution
- **Assumption:** animal transitions between several discrete states (can be chosen as independent or Markovian) and each state is associated with unique parameters of a movement model.
- Parameters to specify: number of states, model and distribution of step length and turning angle for each state
- Track characteristics: track has to be regularized, movement variable should not be autocorrelated

+ Can include covariates
+ Can do predictions, simulations

- Strong assumptions of step length and turning angle distribution and independence
- Lots of parametrizations (expectations required and different ways of fitting)

Segmentation methods overview



Many other possibles...

...variables to segment:

- Spatial clustering (latitude and longitude)
- Space-time clustering (lat, long and time)
- Step-length and time
- First Passage Time (FPT)
- Net Squared Displacement (NSD)
- Variables derived from other sensors (e.g. VeDBA)
- Basically any possible combination of variables that best describe the behaviour you want to identify (e.g. climbing rate + turning angle for soaring flight)

Many other possibles...

...segmentation methods:

- GMM() function from package `ClusterR`: Gaussian mixture models are probabilistic models that assumes all the data points are generated from a mixture of a finite number of Gaussian distributions with unknown parameters. A generalization of K-Means based not only on centroids but on distributions (incorporate covariance).
- randomForest() function and package: machine learning algorithm based on decision trees for classification adn regression (training dataset needed).
- lavielle() and segclust2D() functions from package adehabitatLT: allow to perform a non-parametric segmentation (univariate and bivariate respectively) of a time series based on a number of expected segments.
- locmodes() function from package `multimode`: helps locating the location of the modes in a variable with multimodal distribution.
- corridor() function from package `move`: identifies sections in the trajectory characterized by parallel movement[1]. Works best for animals that are predominantly within a home range.

[1] LaPoint et al. (2013). Animal behavior, cost-based corridor models, and real corridors. Landscape Ecology, 28(8), 1615-1630.