What is the animals doing?

Analytical approaches for inferring behavioral modes from movement data

Elie Gurarie

University of Maryland - College Park

Auke Bay Movement Workshop, Norway - 2016

August 25, 2016

Common and Inconvenient Features of Movement Data

- Multi-dimensional
- Auto-correlated
- Error-ridden
- Irregularly sampled
- **Heterogeneous!**
  • Across space (environmental) • Between individuals (population)
  • **Within an individual (behavioral)**
- **Lack of consensus on appropriate models or analysis methods**

Basic Question:

How do we identify/characterize/model behavioral heterogeneity / complexity in individual animal movement data?

Basic Question:

How do we identify/characterize/model behavioral heterogeneity / complexity in individual animal movement data?

Manuscript Objective:

Actually, many methods exist in the literature! This is confusing. Our goal is to take (some of) these methods and

- *schematize, review, assess, compare and contrast*.
  - strengths and weaknesses
  - assumptions
  - difficulty of implementation
  - interpretabilty of outputs
  - *other criteria?*
- Provide conceptual framework to guide choice of appropriate analysis.

Methods:

1. Hidden Markov modeling / State space modeling of multi-state random walks - SSM/HMM (Morales et al. 2004, Jonsen et al. 2008)
2. First passage time - FPT (Fauchald and Tveraa 2003)
3. Trajectory segmentation (Barraquand Benhamou 2008, Calenge)
4. Behavioral change point analysis - BCPA (Gurarie et al. 2009)
5. *More?!*

Multi-state random walks

Basic Idea:

- Assume the animal has several distinct / discrete behavioral states, each associated with a set of movement behaviors.

Multi-state random walks

Basic Idea:

- Assume the animal has several distinct / discrete behavioral states, each associated with a set of movement behaviors.
- Formulate a model in which the switching between discrete states occurs with some rate (that can be a function of covariates)

Multi-state random walks

Basic Idea:

- Assume the animal has several distinct / discrete behavioral states, each associated with a set of movement behaviors.
- Formulate a model in which the switching between discrete states occurs with some rate (that can be a function of covariates)
- Estimate! (typically using Bayesian MCMC)

Hidden Markov Model: Morales et al. 2004

$$P(y \mid a, b, \mu, \rho) = \prod_{t=1}^{T} W(r_t \mid a_{I_t}, b_{I_t}) C(\phi_t \mid \mu_{I_t}, \rho_{I_t})$$

Uses **Correlated Random Walk** with wrapped-Cauchy and Weibull parameters. I is index of state

$$q_{21} = \frac{\exp\left(\beta_1 + \sum_{h=1}^{H} m_h d_h\right)}{1 + \exp\left(\beta_1 + \sum_{h=1}^{H} m_h d_h\right)}$$

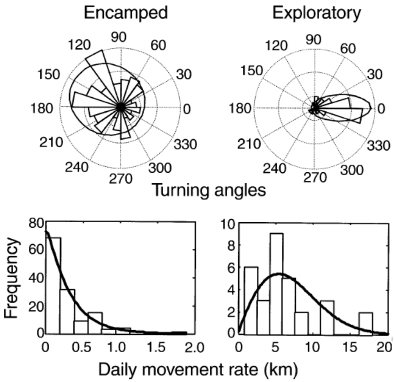
$$q_{11} = 1 - q_{21}$$

Ecology, 85(9), 2004, pp. 2436–2445

EXTRACTING MORE OUT OF RELOCATION DATA: BUILDING MOVEMENT MODELS AS MIXTURES OF RANDOM WALKS

Juan Manuel Morales,[1,4] Daniel T. Haydon,[2] Jacqui Frair,[3] Kent E. Holsinger,[1] and John M. Fryxell[2]

HMM: Morales et al. 2004



Estimated states ...

HMM: Morales et al. 2004

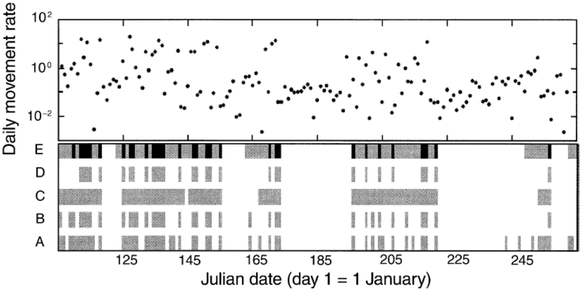


FIG. 2. Activity bar showing the assignment of behavioral states through time for all multiple RW (random walk) models fitted to elk-163: A, "Double"; B, "Double with covariates"; C, "Double switch"; D, "Switch constrained"; E, "Triple switch." Gray bars correspond to the exploratory state, white bars to the encamped state, and black bars to the fastest movement state in the "Triple switch" model. Dots above the activity bars indicate daily movement rate on log scale. Models are defined in *Methods: Models*.

State-space modeling

A SSM consists of coupled stochastic models (Figure I): a process model

$$\boldsymbol{x}_t = g(\boldsymbol{x}_{t-1}, \eta_t) \qquad \text{[Equation Ia]}$$

describing the state of an animal (e.g. position $\boldsymbol{x}_t = \{x_{\text{longitude},t}\, t, x_{\text{latitude},t}\}$) at time t and an observation model

$$\boldsymbol{y}_t = h(\boldsymbol{x}_t, \varepsilon_t) \qquad \text{[Equation Ib]}$$

TREE-894; No of Pages 8 ARTICLE IN PRESS

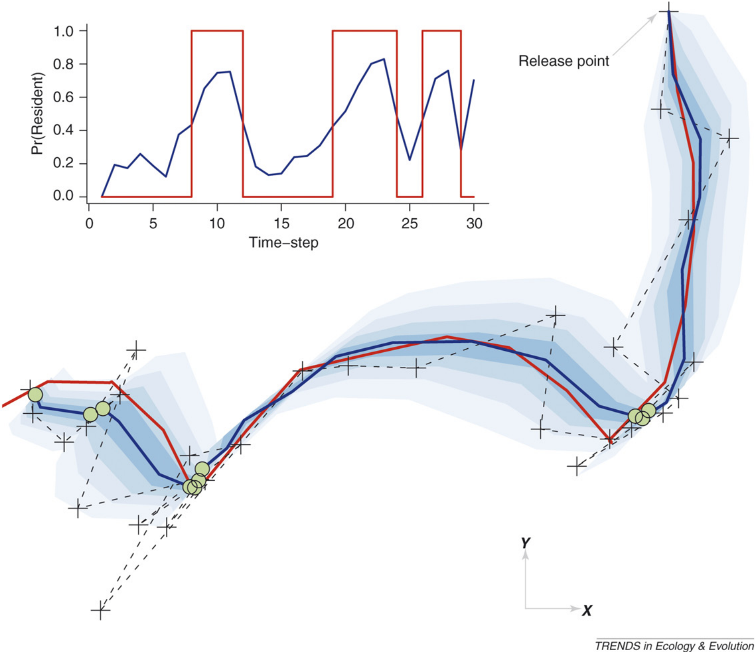
Review

Cell PRESS

State–space models of individual animal movement

Toby A. Patterson[1,2], Len Thomas[3], Chris Wilcox[1], Otso Ovaskainen[4] and Jason Matthiopoulos[5,3]

State-space modeling



Switching SSM: another example

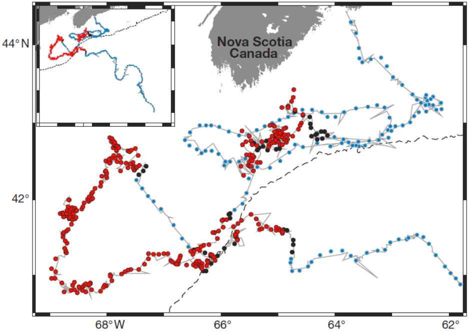


Fig. 1. *Dermochelys coriacea*. State estimates ($\mathbf{x}_t$, filled circles) with associated behavioural mode estimates (blue = transiting, red = foraging, black = uncertain) obtained from the SSSM for a leatherback turtle (B.1) tagged in coastal waters off Nova Scotia, Canada. The full path is shown inset. The underlying grey line indicates the observed Argos positions. The time interval between each $\mathbf{x}_t$ is 6 h. The 1000 m isobath is displayed as a dashed black line

Vol. 337: 255–264, 2007 | MARINE ECOLOGY PROGRESS SERIES
Mar Ecol Prog Ser | Published May 14

Identifying leatherback turtle foraging behaviour from satellite telemetry using a switching state-space model

Ian D. Jonsen*, Ransom A. Myers†, Michael C. James

Department of Biology, Dalhousie University, 1355 Oxford Street, Halifax, Nova Scotia B3H 4J1, Canada

HMM/SSM/SSSM: Components

- Movement model
  - often 2-3-state CRW, but also applications with centers of attraction
- Behavioral switching model
  - Markovian state-transition matrix
  - **Can depend on co-variates**
- Observation model
  - but only if error is important, especially in marine environments

HMM/SSM: Assessment

Strengths

- **Natural implementation of responses to covariates**.
- Ability to estimate a “full” process-based models
- Capable of including “arbitrarily” complex behaviors, e.g. with external biases or (possibly unknown) centers of attraction.
- Potential for hierarchical modeling
- Explicit accounting for observation error.

HMM/SSM: Assessment

Strengths

- **Natural implementation of responses to covariates**.
- Ability to estimate a "full" process-based models
- Capable of including "arbitrarily" complex behaviors, e.g. with external biases or (possibly unknown) centers of attraction.
- Potential for hierarchical modeling
- Explicit accounting for observation error.

Weaknesses

- **Very heavy reliance on a priori assumption of "countable" states**
- Dependence on having some idae
- To date: Fitting of discrete-time models. High reliance on CRW-type models (which do not scale very well, or are easily fitted with irregularly sampled data).
- Computationally very intensive. Sometimes (often) intractable.

First Passage Time Analysis

Ecology, 84(2), 2003, pp. 282–288

USING FIRST-PASSAGE TIME IN THE ANALYSIS OF AREA-RESTRICTED SEARCH AND HABITAT SELECTION

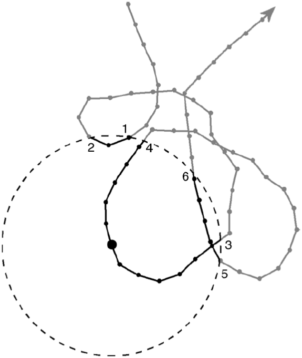
Per Fauchald[1] and Torkild Tveraa

Norwegian Institute for Nature Research, Division of Arctic Ecology, Polar Environmental Center, N-9296 Tromsø, Norway

Basic Idea:
The longer it takes you to leave an area, the longer you are searching.

Basic Method: Compute how long it takes to "leave" a variety range of radii over entire track.

FPT: Implementation

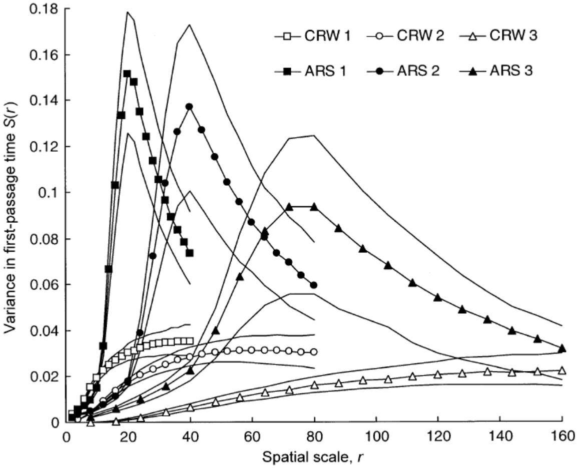


[1]

Compute all times $T(r)$ within a circle of radius r around point i.

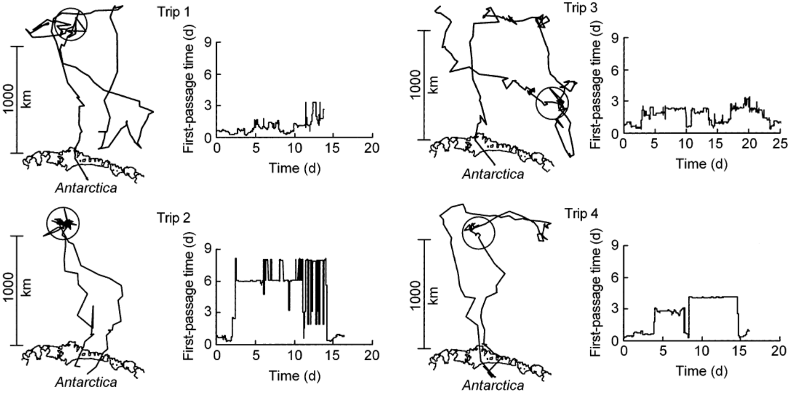
[1] image from Barraquand and Benhamou 2008

FPT: Area Restricted Search



Use of *variance* of first passage time at various spatial scales to identify *area-restricted search*

FPT: Applied to petrel tracks



FPT: Assessment

Strengths

- Few assumptions about movement model
- Insights into scales of overall and intensive searching properties
- Easy to implement

FPT: Assessment

Strengths

- Few assumptions about movement model
- Insights into scales of overall and intensive searching properties
- Easy to implement

Weaknesses

- No "objective" statistical method to identify segmentations
- Relies heavily on assumption that Area Restricted Use is related to intensive foraging
  - Fails if foraging is itself mobile (e.g. fur seals)
  - Fails if resting and prey processing are more stationary than hunting.
- Difficult to implement for irregularly sampled data

Related technique: Spatio-temporal path segmentation

Basic Idea:

- Compute total residence time within a radius
- Apply statistical time-series segmentation techniques (Lavielle 2005) to classify time-series of use-intensity
  - Heuristic selection of number of change points within a time series
  - Implemented by Clément in adehabitatLT

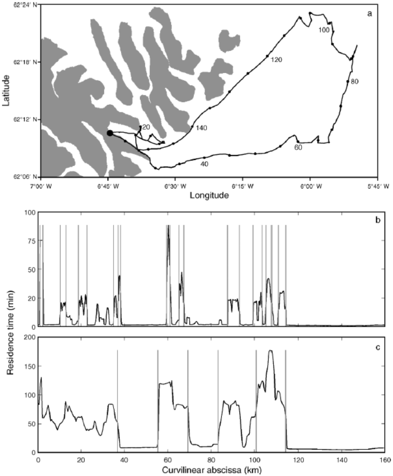
Ecology, 89(12), 2008, pp. 3336–3348
© 2008 by the Ecological Society of America

ANIMAL MOVEMENTS IN HETEROGENEOUS LANDSCAPES: IDENTIFYING PROFITABLE PLACES AND HOMOGENEOUS MOVEMENT BOUTS

Frédéric Barraquand[1] and Simon Benhamou[2]

Centre d'Écologie Fonctionnelle et Évolutive, CNRS Montpellier, France

Related technique: Spatio-temporal path segmentation



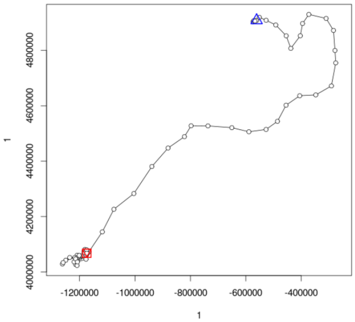
More Segmentation: adehabitatLT

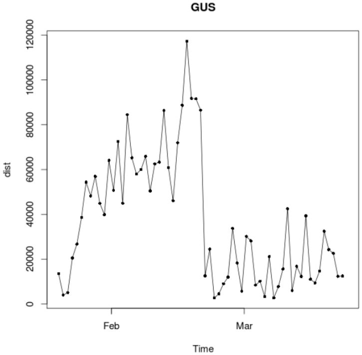
Analysis of Animal Movements in R:
the `adehabitatLT` Package

Clement Calenge,
Office national de la chasse et de la faune sauvage
Saint Benoist – 78610 Auffargis – France.

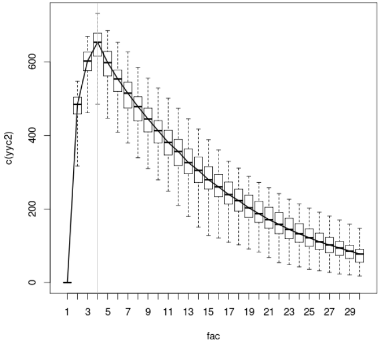
April 2011

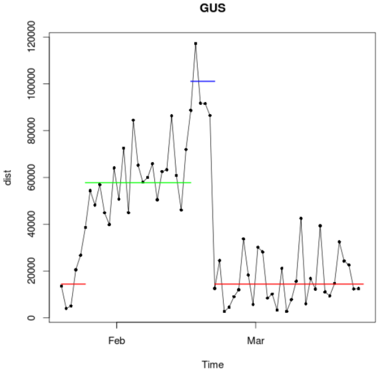
More Segmentation: adehabitatLT





More Segmentation: adehabitatLT





Behavioral Changepoint Analysis (BCPA)

Basic Goal:

Identify behavioral changes with minimal a priori assumptions in a movement dataset:

- that is autocorrelated and can be sampled irregularly or with error;
- that contains an arbitrary number of discrete states, and can accomodate gradual changes in behavior

Behavioral Changepoint Analysis (BCPA)

Basic Goal:

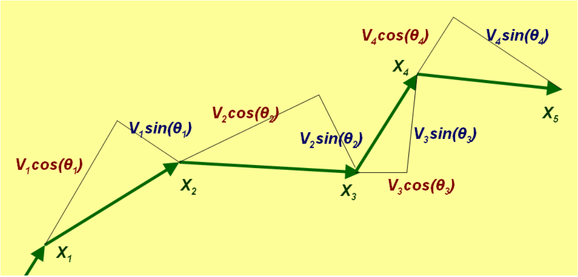
Identify behavioral changes with minimal a priori assumptions in a movement dataset:

- that is autocorrelated and can be sampled irregularly or with error;
- that contains an arbitrary number of discrete states, and can accomodate gradual changes in behavior

Basic Method:

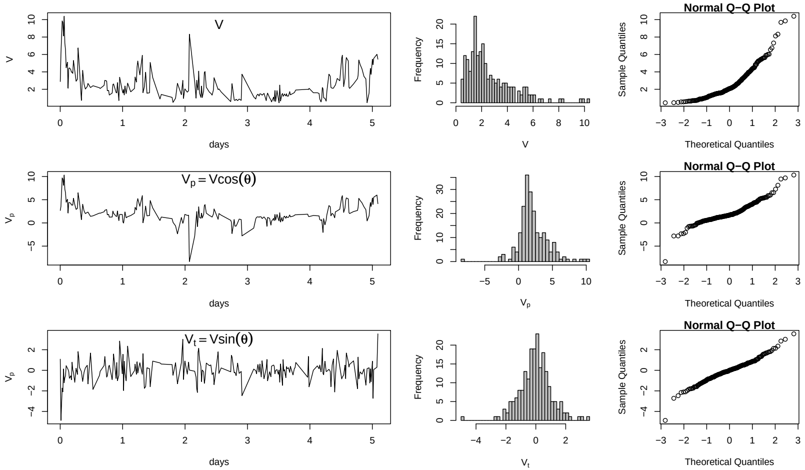
- Define a homogeneous portion of movement from organism's perspective (e.g. velocities and turning angles) as a time-series with a few parameters (μ, σ, ρ)
- Use likelihood methods and selection criteria to see if there is a "significant" shift in the values of the parameters over a portion of the data.
- Scan the complete data with a moving window.

BCPA: Select movement variable



Persistence Velocity Component: $V_p = V\cos(\theta)$

- Captures tendency and speed of persistence:
  - high **mean** = high speed and consistent orientation
  - high **variance** = variable behaviors (stopping and going, slowing down and speeding)
  - high **auto-corellation** = behavioral changes slower than sampling interval



- Stationary
- Gaussian
- Modelable using standard time-series techniques

BCPA: Likelihood Model

$$\begin{aligned} \mathrm{E}\left(X(t)\right) &= \mu \\ \mathrm{Var}\left(X(t)\right) &= \sigma^2 \\ \mathrm{Corr}\left(X(t)\right) X(t-\tau) &= \rho^\tau \end{aligned}$$

$$f(X(t)|X(t-\tau)) \sim \text{Gaussian}\left[\rho^\tau X(t-\tau), \sigma^2(1-\rho^{2\tau})\right]$$

BCPA: Estimating ρ

Conditional Likelihood:

$$L(\rho|\mathbf{X},\mathbf{T}) = \prod_{i=1}^{n} f(X_i|X_{i-1},\tau_i,\rho),$$

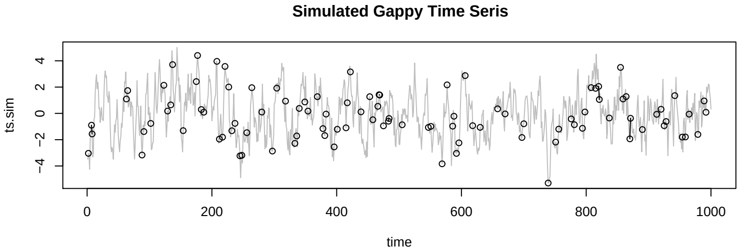
where:

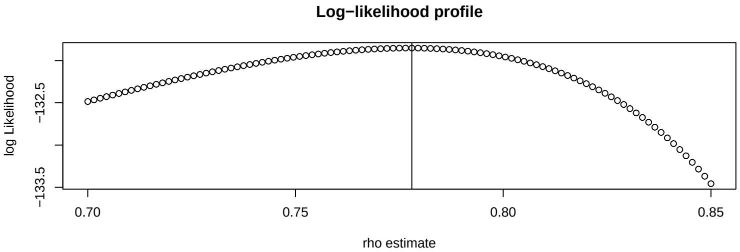
$$f(X_i|X_{i-1}) = \frac{1}{\sigma\sqrt{2\pi(1-\rho^{2\tau_i})}} \exp\left(\frac{(X_i - \rho^{\tau_i}(X_{i-1}-\mu))^2}{2\sigma^2(1-\rho^{2\tau_i})}\right).$$

then:

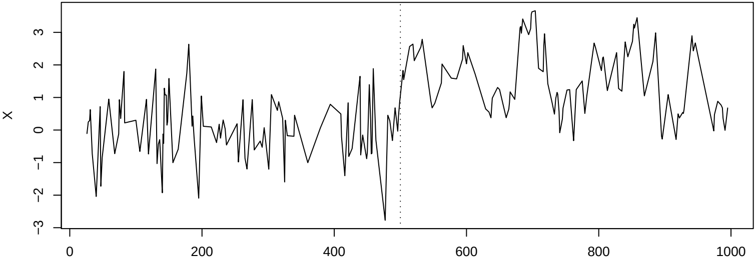
$$\widehat{\rho} = \text{argmax}_\rho \; L(\rho|\mathbf{X},\mathbf{T})$$

Estimating ρ

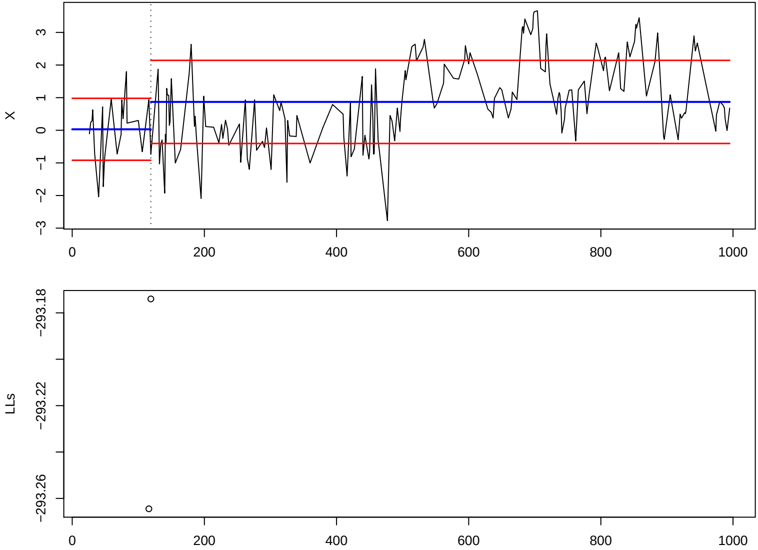




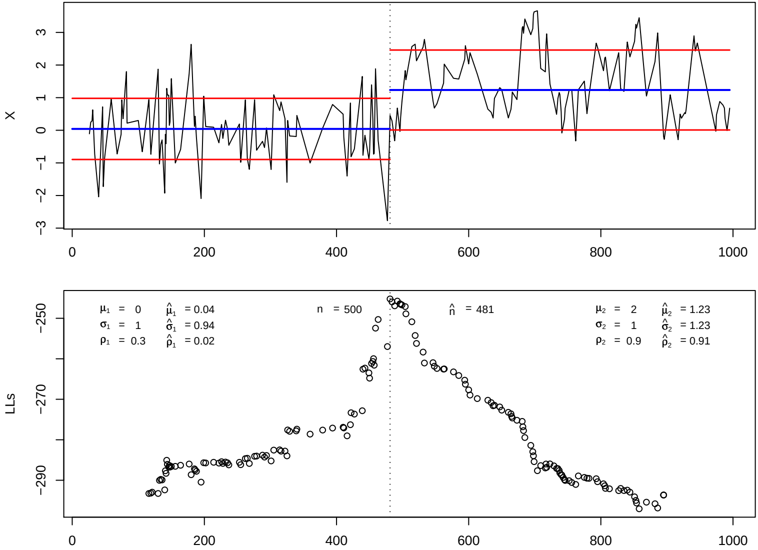
BCPA: Identifying Change Point



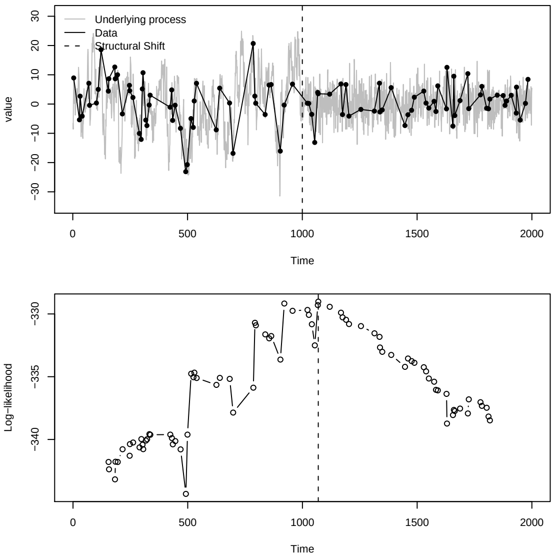
BCPA: Identifying Change Point



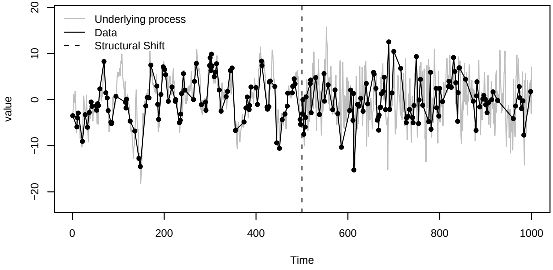
BCPA: Identifying Change Point

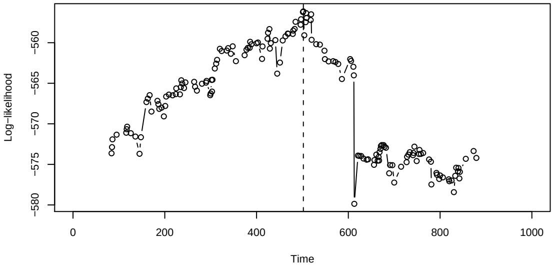


BCPA: Identifying Change Point, sparse data



BCPA: Identifying Change Point, different ρ's





BCPA: Identifying Models

Model 0	$\mu_1 = \mu_2$	$\sigma_1 = \sigma_2$	$\rho_1 = \rho_2$
Model 1	$\mu_1 \neq \mu_2$	$\sigma_1 = \sigma_2$	$\rho_1 = \rho_2$
Model 2	$\mu_1 = \mu_2$	$\sigma_1 \neq \sigma_2$	$\rho_1 = \rho_2$
Model 3	$\mu_1 = \mu_2$	$\sigma_1 = \sigma_2$	$\rho_1 \neq \rho_2$
Model 4	$\mu_1 \neq \mu_2$	$\sigma_1 \neq \sigma_2$	$\rho_1 = \rho_2$
Model 5	$\mu_1 \neq \mu_2$	$\sigma_1 = \sigma_2$	$\rho_1 \neq \rho_2$
Model 6	$\mu_1 = \mu_2$	$\sigma_1 \neq \sigma_2$	$\rho_1 \neq \rho_2$
Model 7	$\mu_1 \neq \mu_2$	$\sigma_1 \neq \sigma_2$	$\rho_1 \neq \rho_2$

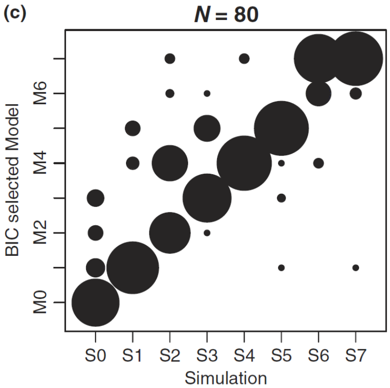
BCPA: Identifying Models

Model 0	$\mu_1 = \mu_2$	$\sigma_1 = \sigma_2$	$\rho_1 = \rho_2$
Model 1	$\mu_1 \neq \mu_2$	$\sigma_1 = \sigma_2$	$\rho_1 = \rho_2$
Model 2	$\mu_1 = \mu_2$	$\sigma_1 \neq \sigma_2$	$\rho_1 = \rho_2$
Model 3	$\mu_1 = \mu_2$	$\sigma_1 = \sigma_2$	$\rho_1 \neq \rho_2$
Model 4	$\mu_1 \neq \mu_2$	$\sigma_1 \neq \sigma_2$	$\rho_1 = \rho_2$
Model 5	$\mu_1 \neq \mu_2$	$\sigma_1 = \sigma_2$	$\rho_1 \neq \rho_2$
Model 6	$\mu_1 = \mu_2$	$\sigma_1 \neq \sigma_2$	$\rho_1 \neq \rho_2$
Model 7	$\mu_1 \neq \mu_2$	$\sigma_1 \neq \sigma_2$	$\rho_1 \neq \rho_2$

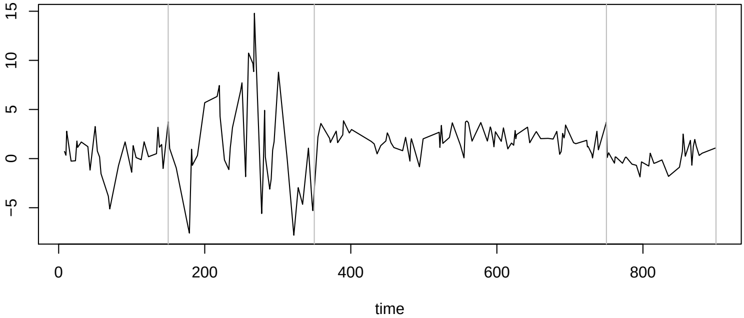
How to choose?

$$\text{BIC} : I_B(\mathbf{X}, \mathbf{T}) \quad = \quad -2n \log\left(L(\hat{\theta}|\mathbf{X}, \mathbf{T})\right) + d \log(n)$$

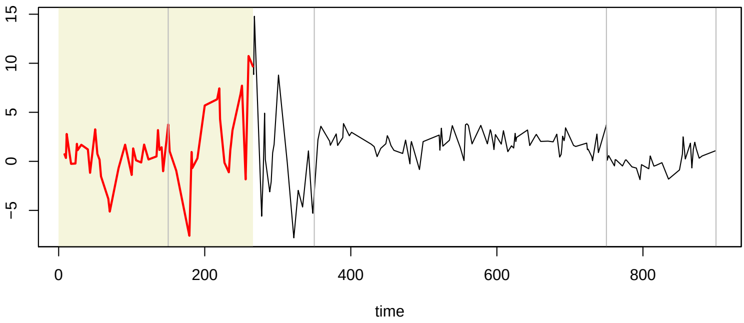
BCPA: Model selection



	μ_1	μ_2	σ_1	σ_2	ρ_1	ρ_2
S0	0	0	1	1	0.5	0.5
S1	**-1**	**1**	1	1	0.5	0.5
S2	0	0	**0.5**	**2**	0.5	0.5
S3	0	0	1	1	**0.2**	**0.9**
S4	**-1**	**1**	**0.5**	**2**	0.5	0.5
S5	**-1**	**1**	1	1	**0.2**	**0.9**
S6	0	0	**0.5**	**2**	**0.2**	**0.9**
S7	**-1**	**1**	**0.5**	**2**	**0.2**	**0.9**

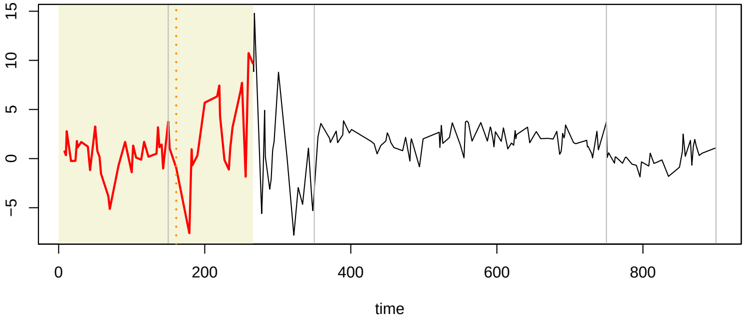


15
10
5
0
−5
0
200
400
600
800
time



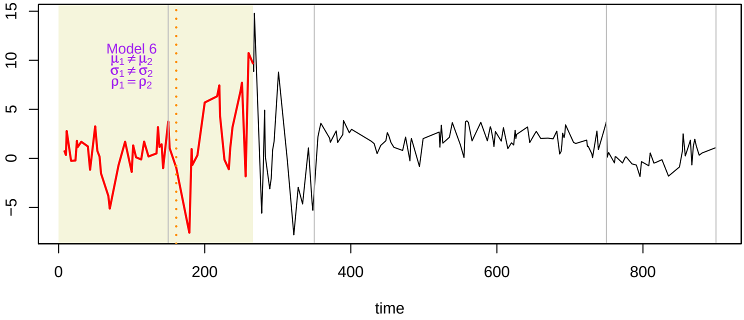
BCPA: Algorithm for Identifying Multiple Changepoints

- Select Window
- Find MLBP
- Identify Model
- Record estimates based on model selected.
- Move window forward and repeat



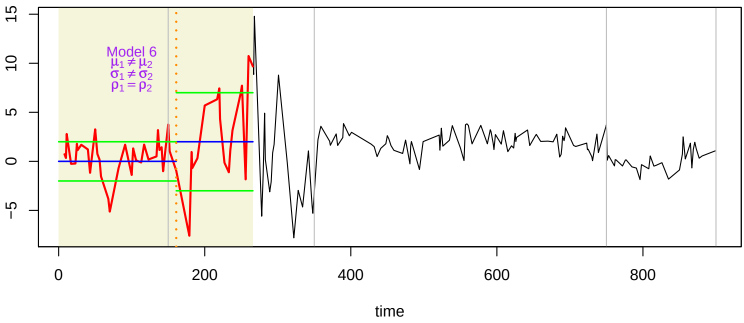
BCPA: Algorithm for Identifying Multiple Changepoints

- Select Window
- Find MLBP
- Identify Model
- Record estimates based on model selected.
- Move window forward and repeat



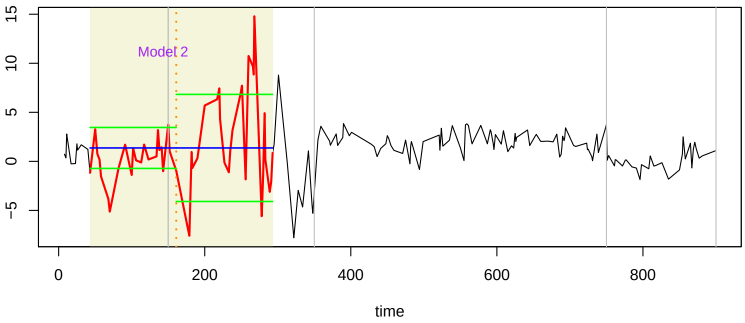
BCPA: Algorithm for Identifying Multiple Changepoints

- Select Window
- Find MLBP
- Identify Model
- Record estimates based on model selected.
- Move window forward and repeat



BCPA: Algorithm for Identifying Multiple Changepoints

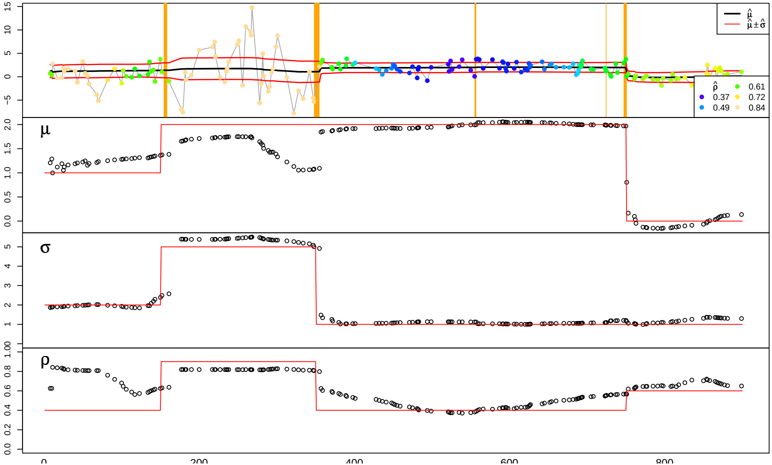
- Select Window
- Find MLBP
- Identify Model
- Record estimates based on model selected.
- Move window forward and repeat



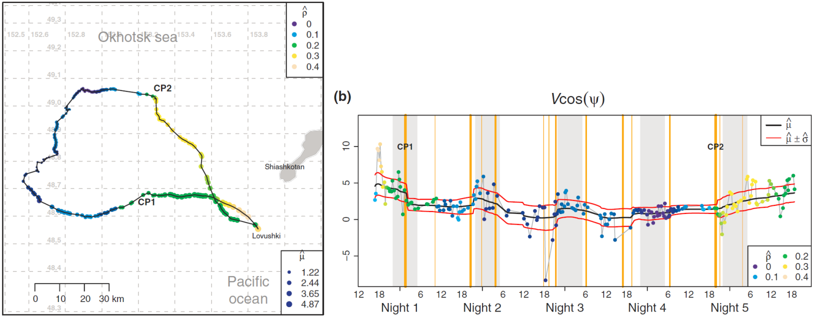
BCPA: Algorithm for Identifying Multiple Changepoints

- Select Window
- Find MLBP
- Identify Model
- Record estimates based on model selected.
- Move window forward and repeat

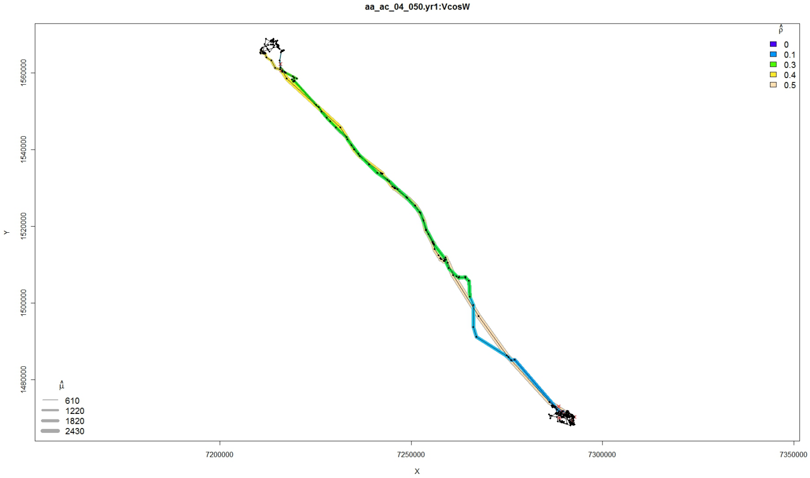
BCPA: Movement analysis output



BCPA: Northern Fur Seal

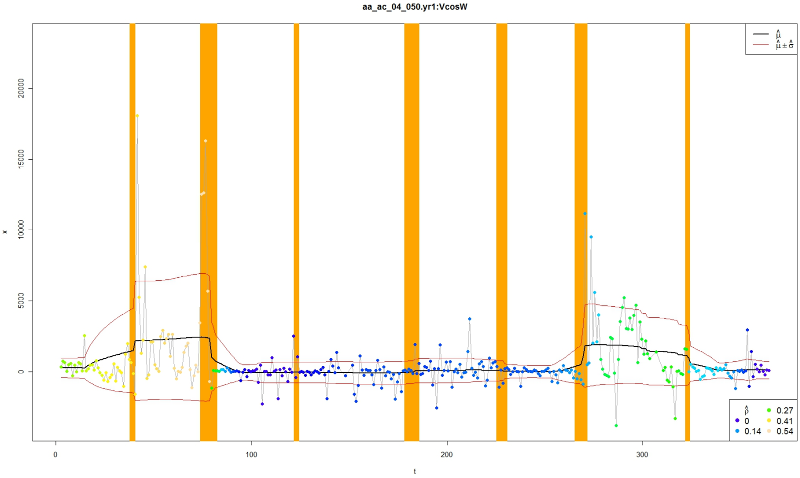


BCPA: Swedish Moose[2]



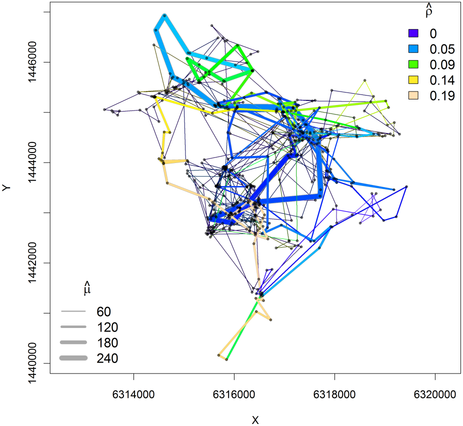
[2]Thanks to Navi Singh

BCPA: Swedish Moose[3]



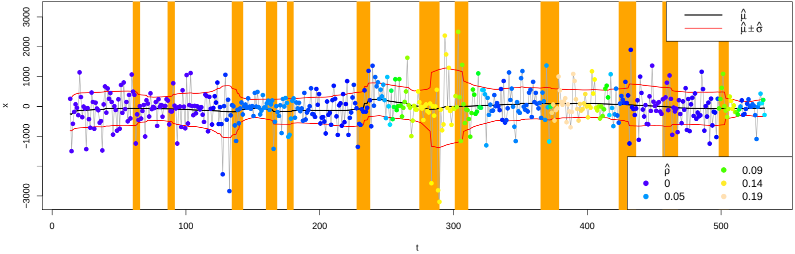
[3] Thanks to Navi Singh

BCPA: Swedish Moose[4]



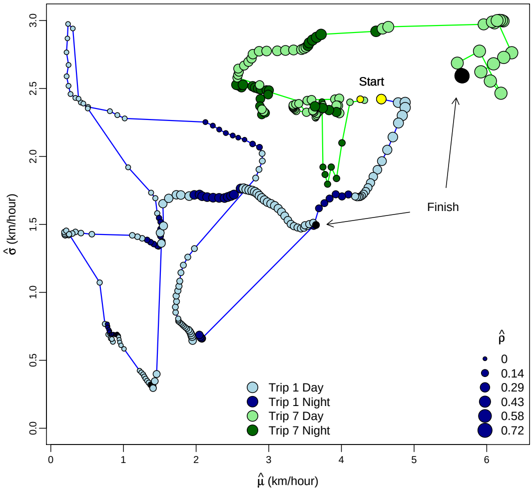
[4] Thanks to Navi Singh

BCPA: Swedish Moose[5]



[5] Thanks to Navi Singh

BCPA: Behavioral phase plot



BCPA: Assessment

Strengths

- Very few a priori assumptions
- Works on continuous time-series
- Relatively easy/fast to implement
- Synthesizes complex/messy data into something digestible
- Can be applied to any time-series derived from movement (speed, persistence, turning, depths)

BCPA: Assessment

Strengths

- Very few a priori assumptions
- Works on continuous time-series
- Relatively easy/fast to implement
- Synthesizes complex/messy data into something digestible
- Can be applied to any time-series derived from movement (speed, persistence, turning, depths)

Weaknesses

- Confusing output.
- Not a "complete" model of a process.
- Requires considerable post hoc effort to interpret results.
- Requires some "tuning" (window-sizes, selection thresholds)
- Joint modeling of multiple time-series?

Discussion point

There seems to be a fundamental distinction between "full modeling" approaches (HMM, SSM) and "synthesizing approaches" (BCPA, segmentation).

Full process modeling

- Can be very powerful for quantifying relationships between (behaviorally heterogeneous) movements and explanatory factors (e.g. habitat), allowing for fully parameterized models, model comparison, etc.
- Can incorporate arbitrarily complex movement models, e.g. with centers of attraction (including unknown ones, e.g. McClintock 2012)
- Relies very heavily on strong a priori assumptions about:
    - a) the movement model, especially the number of states
    - b) the functional form of response to covariates ... which leads to questions about what is "selection", "availability", the role of internal states, etc.
- Can be computationally very intensive. Effort (and tuning) goes into making models converge, independent biological knowledge goes (perhaps) into Bayesian priors.

Full process modeling

- Can be very powerful for quantifying relationships between (behaviorally heterogeneous) movements and explanatory factors (e.g. habitat), allowing for fully parameterized models, model comparison, etc.
- Can incorporate arbitrarily complex movement models, e.g. with centers of attraction (including unknown ones, e.g. McClintock 2012)
- Relies very heavily on strong a priori assumptions about:
  - a) the movement model, especially the number of states
  - b) the functional form of response to covariates ... which leads to questions about what is "selection", "availability", the role of internal states, etc.
- Can be computationally very intensive. Effort (and tuning) goes into making models converge, independent biological knowledge goes (perhaps) into Bayesian priors.

Future directions:

- Account better for continuous time processes
- More flexible definitions, possible numbers of discrete states
- Explicit accounting for relationship between dimensions of movement data
- more?

Synthesizing approaches

- require far fewer a priori assumptions, and can more flexibly accommodate different kinds of movement models (correlated, continuous) or metrics (but probably only as long as they can be "coerced" to be Gaussian).
- Generally, computationally much easier/faster.
- provides NO further suggestion as to how to explain the synthesized output, i.e. requires a whole separate, wide open step of modeling the outcumes (e.g. the times and locations of behavioral change points).
- Generally requires some biologically meaningful/case specific "tuning" of parameters.

Synthesizing approaches

- require far fewer a priori assumptions, and can more flexibly accommodate different kinds of movement models (correlated, continuous) or metrics (but probably only as long as they can be "coerced" to be Gaussian).
- Generally, computationally much easier/faster.
- provides NO further suggestion as to how to explain the synthesized output, i.e. requires a whole separate, wide open step of modeling the outcumes (e.g. the times and locations of behavioral change points).
- Generally requires some biologically meaningful/case specific "tuning" of parameters.

Future directions:

- Formalization of methods/strategies for post hoc analysis?
- Incorporation of multi-dimensional data (multiple time-series)?
- Expansion beyond entirely organism-centric variables?

What is a researcher with data supposed to do?

Obviously this depends on what the QUESTIONS are and what the DATA can support!

Some general (hypothetized) principles:

- Consider carefully (and explore empirically) whether CRW or continuous time model is appropriate, what metrics might be most meaningful, whether measurement error is important.
  - This depends on resolution and regularity of data. Very high temporal resolution and very irregular data, should NOT be modeled with CRW.
  - Discretization of data (e.g. "best daily location") can be a helpful simplification.
- Explore data (with more "synthesizing approaches" and heavy use of visualization) before fitting complex models.
- Be aware of assumptions behind different methods, and test to see if they are supported by the data! (provide concrete examples.)

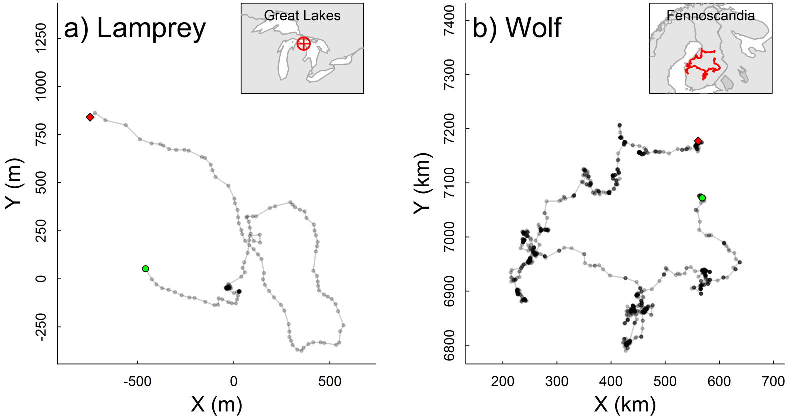
What are we doing?

Exploring/illustrating these principles...

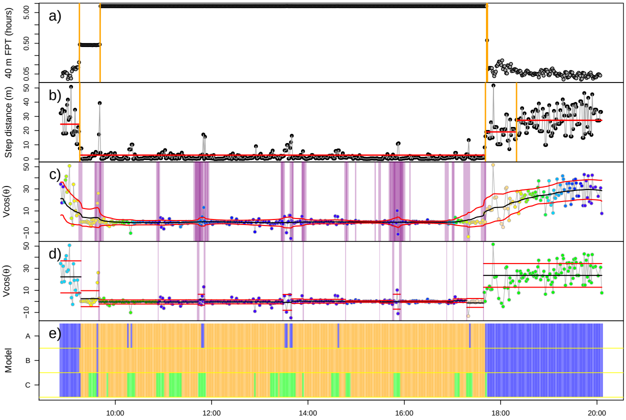
- Picking some data-sets to analyze.
  - Different scales and resolutions?
  - Different richness of covariate data?
  - Simulate some data to illustrate strenghts/weakness of different methods?
  - How many?
- Picking some methods to apply
  - Some or all of these?
  - Assign experts to perform analyses!
- Picking some"objective" criteria to assess/compare outputs?
- Perhaps make an honest attempt to "break" the models - i.e. show how if you don't account for some basic process (violate assumptions) the wrong result appears?

Please, discuss!

Data

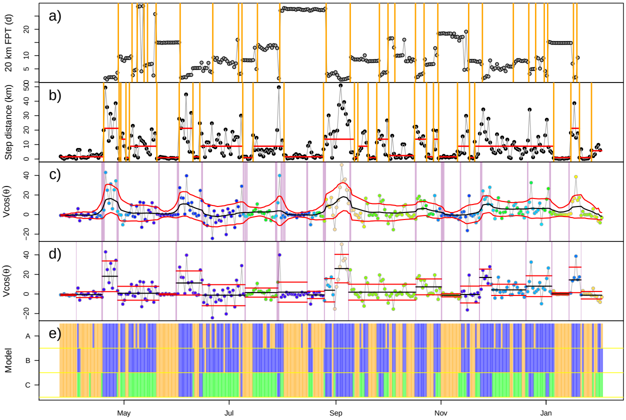

a) Lamprey
Great Lakes
Y (m)
X (m)
1250
1000
750
500
250
0
-250
-500
0
500
b) Wolf
Fennoscandia
Y (km)
X (km)
7400
7300
7200
7100
7000
6900
6800
200
300
400
500
600
700

Results



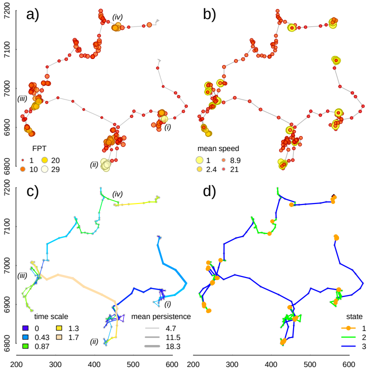
a)
b)
c)
d)
e)
40 m FPT (hours)
Step distance (m)
Vcos(θ)
Vcos(θ)
Model
A
B
C
10:00
12:00
14:00
16:00
18:00
20:00

Results



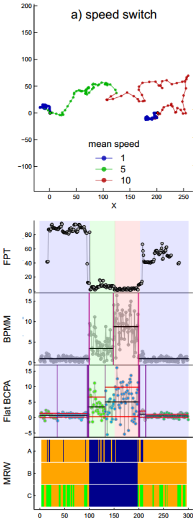
a)
b)
c)
d)
e)
20 km FPT (d)
Step distance (km)
Vcos(θ)
Vcos(θ)
Model
A
B
C
May
Jul
Sep
Nov
Jan

Results



a)
b)
c)
d)
(i)
(ii)
(iii)
(iv)
FPT
1
10
20
29
mean speed
1
2.4
8.9
21
time scale
0
0.43
0.87
1.3
1.7
mean persistence
4.7
11.5
18.3
state
1
2
3
7200
7100
7000
6900
6800
200
300
400
500
600

Change-point analyses[6]



Four phase CVM

Slow Fast Faster Slow

I. First Passage Time

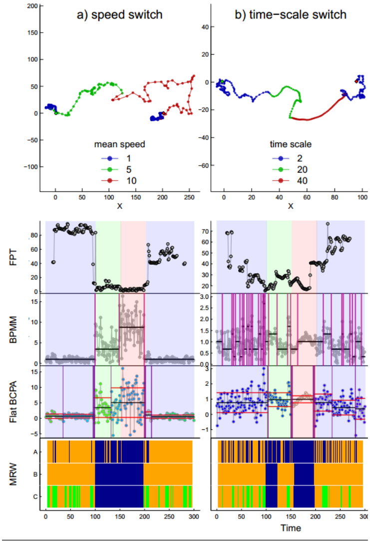
II. Bayesian Partitioning

III. Behavioral Change Point Analysis

IV. Multi-state Random Walks

[6] Gurarie, Bracis, Delgado, Meckley, Kojola and Wagner, Journal of Animal Ecology

Change-point analyses[6]

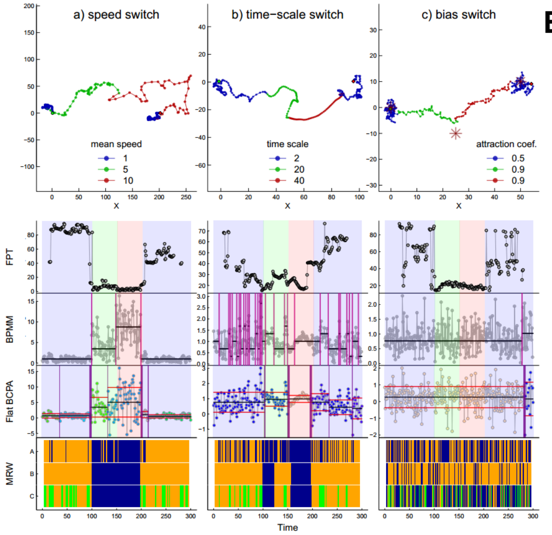


Four phase CVM

tortuous less tortuous
linear tortuous

[6]Gurarie, Bracis, Delgado, Meckley, Kojola and Wagner, Journal of Animal Ecology

Change-point analyses[6]



Biased CRW

three points of attraction

uniform speed and tortuosity

[6]Gurarie, Bracis, Delgado, Meckley, Kojola and Wagner, Journal of Animal Ecology

Broad recommnedations

- Trust biological intuition
- Be aware of the structure of the data
- Consider feasibility
- Be aware of and test assumptions
- Build your analysis off focal individuals
- Assess via simulation
- Combine tools