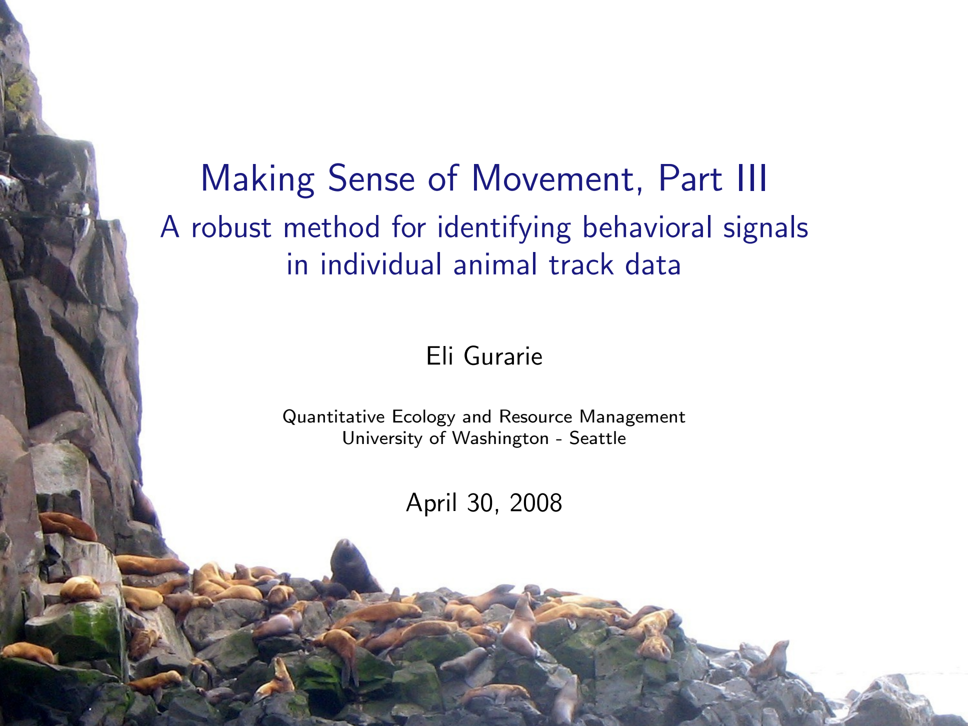


Making Sense of Movement, Part III

A robust method for identifying behavioral signals in individual animal track data

Eli Gurarie

Quantitative Ecology and Resource Management
University of Washington - Seattle

April 30, 2008

Motivation

- Movement is important
- There's lots of data
- There's lots of really interesting questions to ask

Motivation

- Movement is important
- There's lots of data
- There's lots of really interesting questions to ask

Motivation

- Movement is important
- There's lots of data
- There's lots of really interesting questions to ask

Movement Data

I. Mark-Recapture

II. First-Passage Time

III. Tracks

Movement Data

I. Mark-Recapture

II. First-Passage Time

III. Tracks

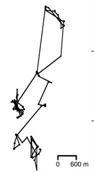
Movement Data

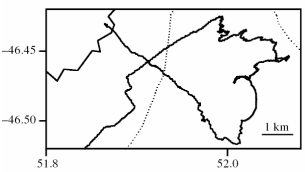
I. Mark-Recapture

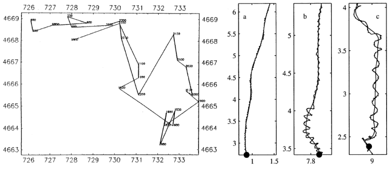
II. First-Passage Time

III. Tracks

Track Data



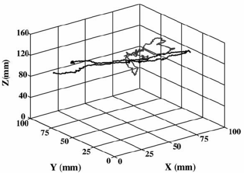


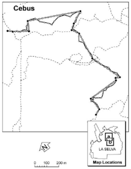


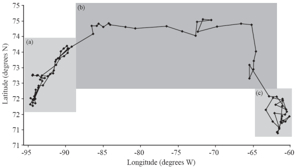
Treecreeper (Doerr 2004)
Heterosigma (Bearon 2003)

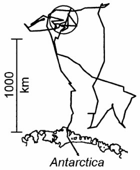
Albatross (Fritz 2002)

Iberian wolf (Bascompte 1997)







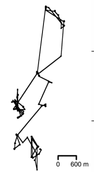


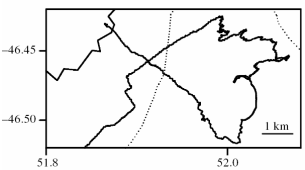
Daphnia Pulex (Uttieri 2005)
Petrel (Fouchauld 2003)

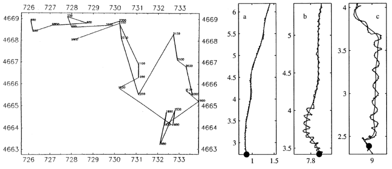
Cebus monkey (Wentz 2003)

Narwhal (Laidre 2004)

Track Data



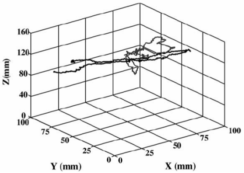


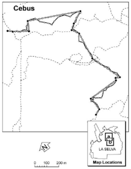


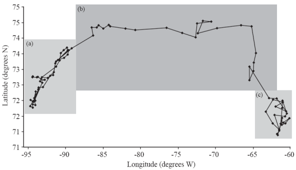
Treecreeper (Doerr 2004)
Heterosigma (Bearon 2003)

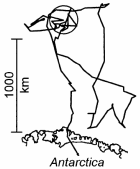
Albatross (Fritz 2002)

Iberian wolf (Bascompte 1997)









Daphnia Pulex (Uttieri 2005)
Petrel (Fouchauld 2003)

Cebus monkey (Wentz 2003)

Narwhal (Laidre 2004)

What kind of information is hiding in this data?

Common and Somewhat Inconvenient Features of Track Data

- Multi-dimensional
- Auto-correlated
- Error-ridden
- Gappy
- Heterogeneous

Common and Somewhat Inconvenient Features of Track Data

- Multi-dimensional
- Auto-correlated
- Error-ridden
- Gappy
- Heterogeneous

Common and Somewhat Inconvenient Features of Track Data

- Multi-dimensional
- Auto-correlated
- Error-ridden
- Gappy
- Heterogeneous

Common and Somewhat Inconvenient Features of Track Data

- Multi-dimensional
- Auto-correlated
- Error-ridden
- Gappy
- Heterogeneous

Common and Somewhat Inconvenient Features of Track Data

- Multi-dimensional
- Auto-correlated
- Error-ridden
- Gappy
- Heterogeneous

The Goal

Develop ways to describe/analyze movement data that:

- are biologically meaningful, versatile, robust;

The Goal

Develop ways to describe/analyze movement data that:

- are biologically meaningful, versatile, robust;
- take into account the *temporal* aspect of movement;

The Goal

Develop ways to describe/analyze movement data that:

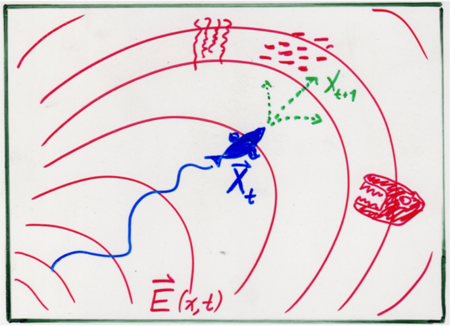
- are biologically meaningful, versatile, robust;
- take into account the *temporal* aspect of movement;
- potentially allow for exploring an organism's *behavioral* response to *environmental/seasonal* cues

The Goal

Develop ways to describe/analyze movement data that:

- are biologically meaningful, versatile, robust;
- take into account the *temporal* aspect of movement;
- potentially allow for exploring an organism's *behavioral* response to *environmental/seasonal* cues

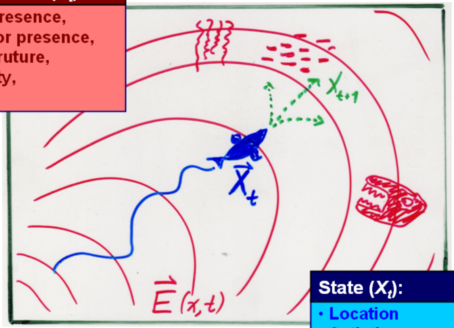
Conceptual model of Behavior



Conceptual model of Behavior

Environment (E_t):

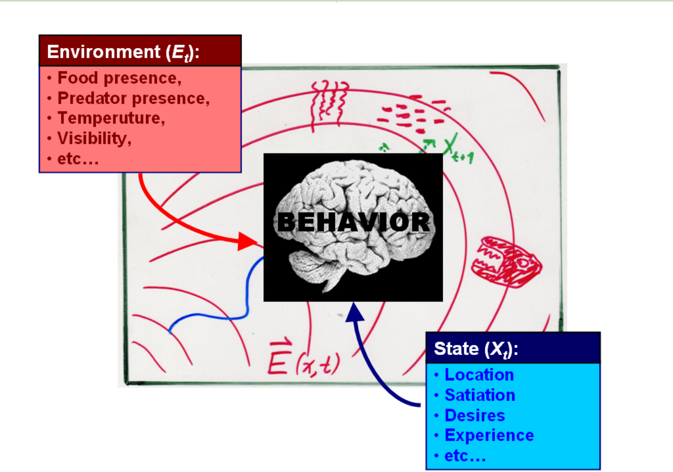
- Food presence,
- Predator presence,
- Temperuture,
- Visibility,
- etc...



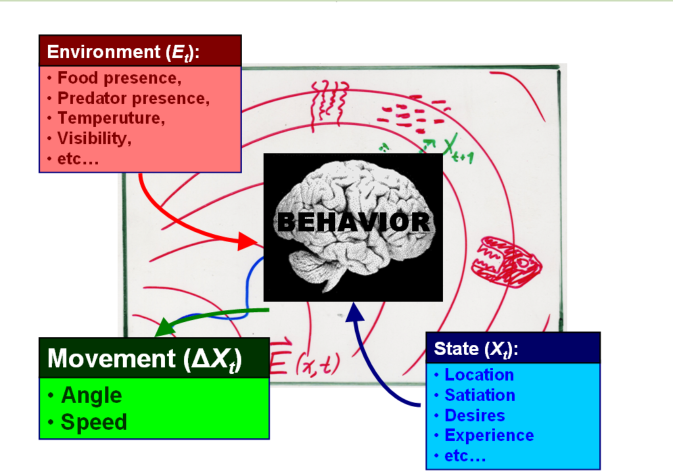
State (X_t):

- Location
- Satiation
- Desires
- Experience
- etc...

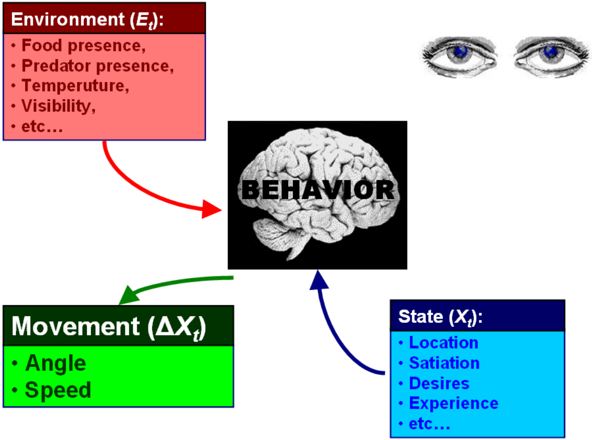
Conceptual model of Behavior



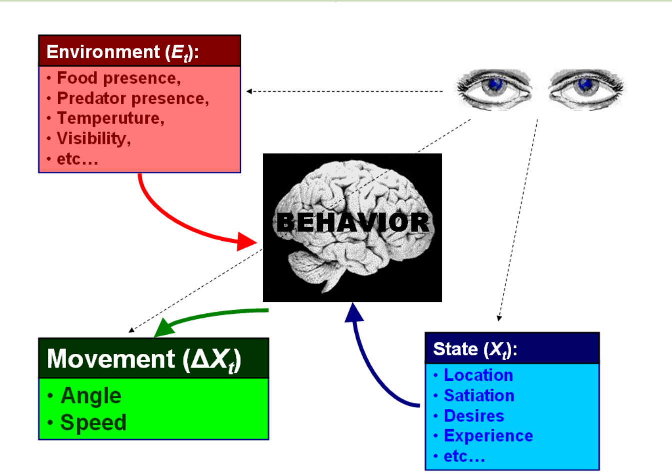
Conceptual model of Behavior



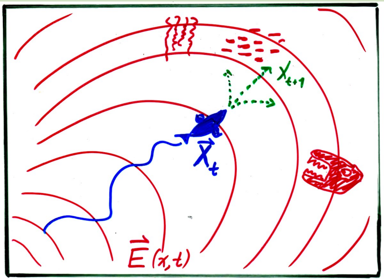
Conceptual model of Behavior



Conceptual model of Behavior



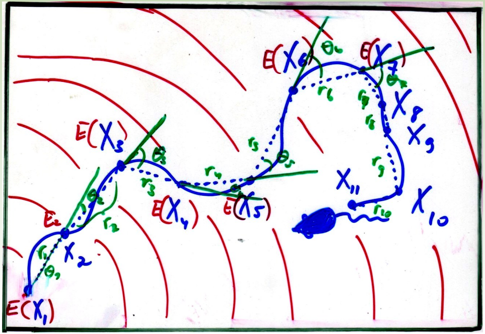
Conceptual model of Behavior



In Math: $\Delta X_t = f(\mathbf{X}_t, \mathbf{E}_t)$

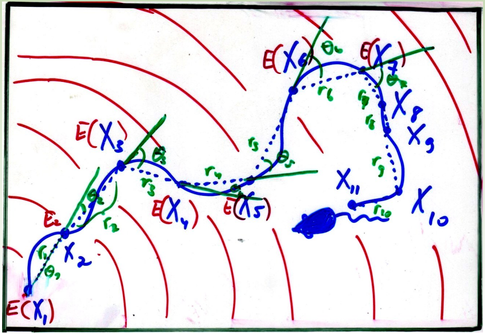
In English: *Behavior* (f) is a process which transforms the state of an organism (X_t) and the the local environment (E_t) into Movement (ΔX_t).

Decomposing Tracks



Strategy:

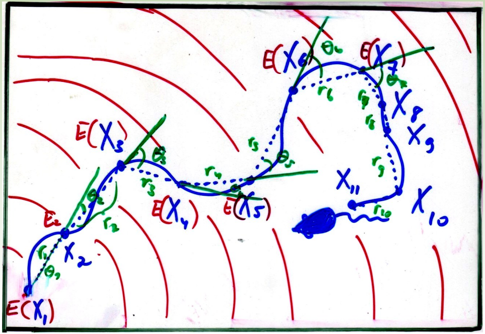
Decomposing Tracks



Strategy:

- Obtain positions X_t and times of observation T_t

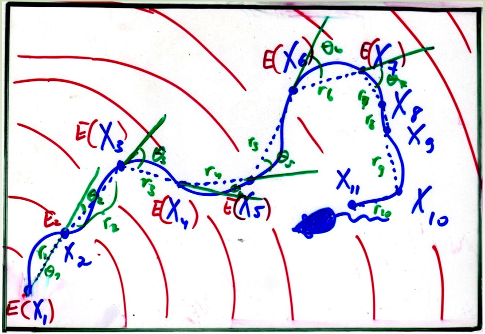
Decomposing Tracks



Strategy:

- Obtain positions X_t and times of observation T_t
- Find series of speeds $V_t = \Delta X / \Delta T$and turning angles θ_t

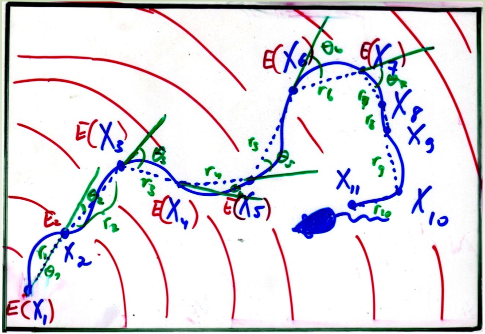
Decomposing Tracks



Strategy:

- Obtain positions X_t and times of observation T_t
- Find series of speeds $V_t = \Delta X / \Delta T$ and turning angles θ_t
- Obtain any environmental/accessory data E_t you can get

Decomposing Tracks



Strategy:

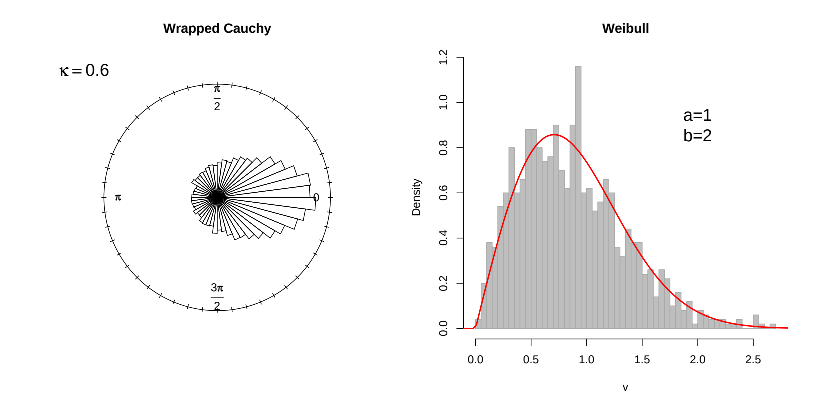
- Obtain positions X_t and times of observation T_t
- Find series of speeds $V_t = \Delta X / \Delta T$ and turning angles θ_t
- Obtain any environmental/accessory data E_t you can get
- Analyze and Interpret!

Correlated Random Walk (CRW) Model

- $\theta_t \sim$ Some Circular Distribution
- $V \sim$ Some Skewed Unimodal Positive Distribution.

Correlated Random Walk (CRW) Model

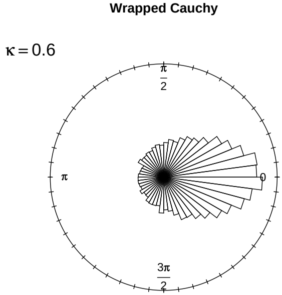
- $\theta_t \sim$ Some Circular Distribution
- $V \sim$ Some Skewed Unimodal Positive Distribution.

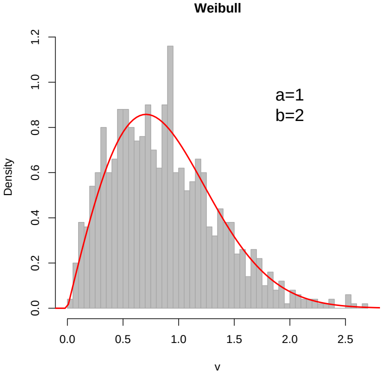


BUT ... these distributions depend on the time interval!

Correlated Random Walk (CRW) Model

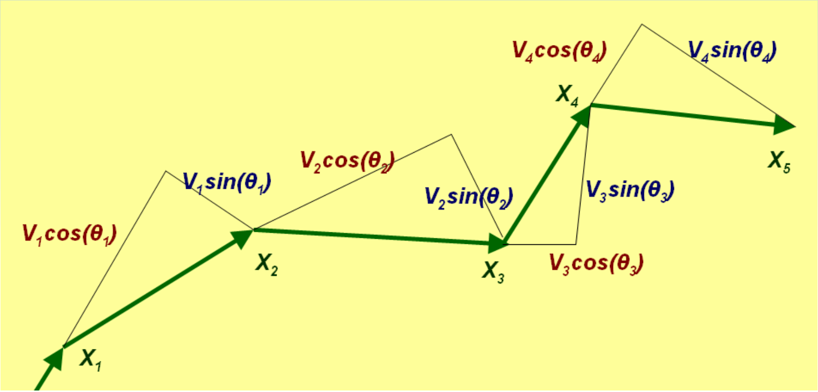
- $\theta_t \sim$ Some Circular Distribution
- $V \sim$ Some Skewed Unimodal Positive Distribution.



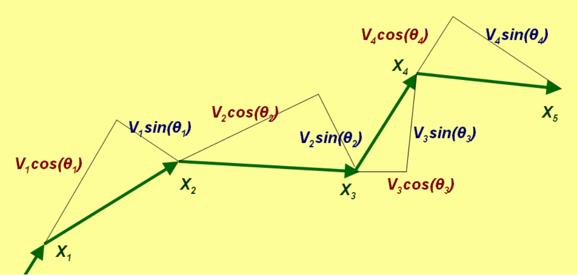


BUT ... these distributions depend on the time interval!
AND ... this is a slightly narrow “rendering” of correlation

Orthogonal decomposition



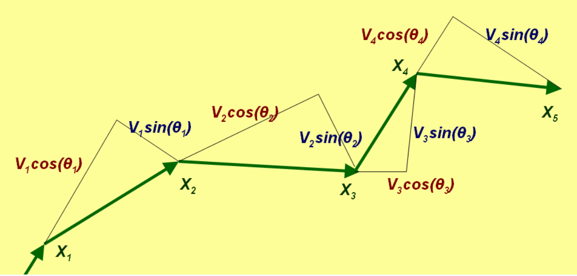
Orthogonal decomposition



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

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

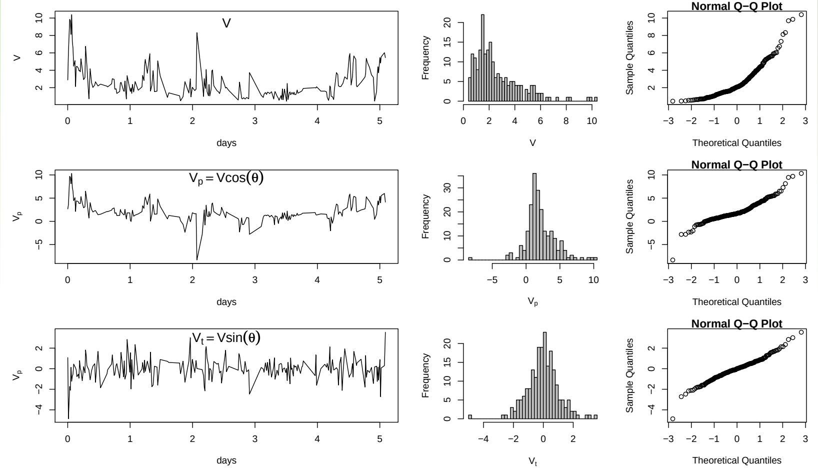
Orthogonal decomposition

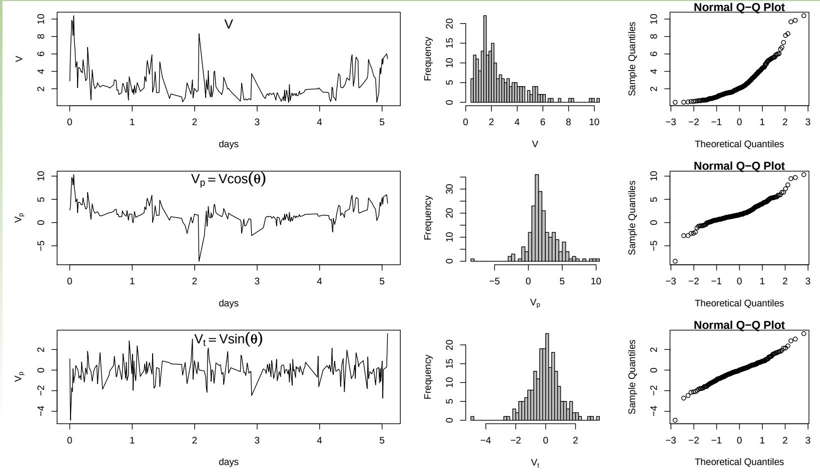


Orthogonal Component of Velocity: $V_t = V\sin(\theta)$

- Captures tendency and speed of turning (displaciement at a perpendicular direction)
  - **mean** can be fixed at zero.
  - high **variance** indicates combination of high speeds and “sharp” turns
  - high **auto-corellation** indicates turning radius larger than the scale of measurement.

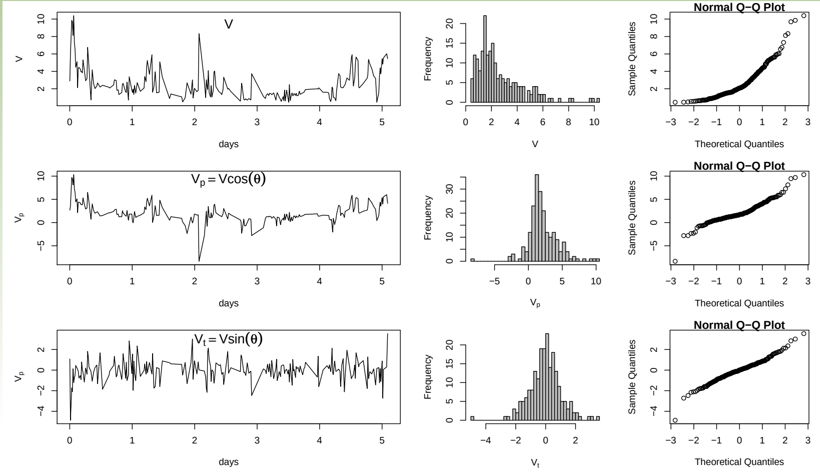
Actual Data Decomposed (northern fur seal)





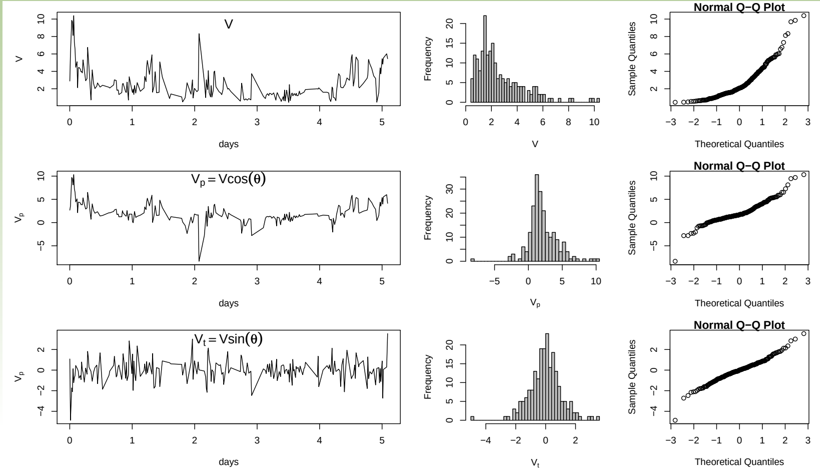
The real trump card

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



The real trump card

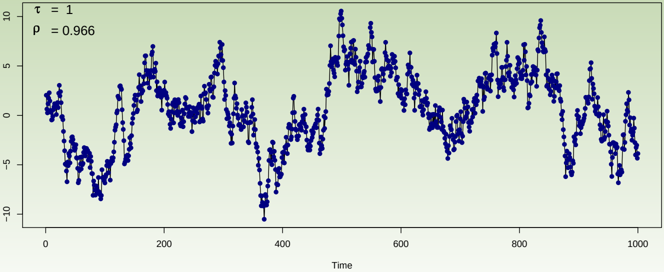
- Stationary
- Gaussian
- Modellable using standard time-series techniques



The real trump card

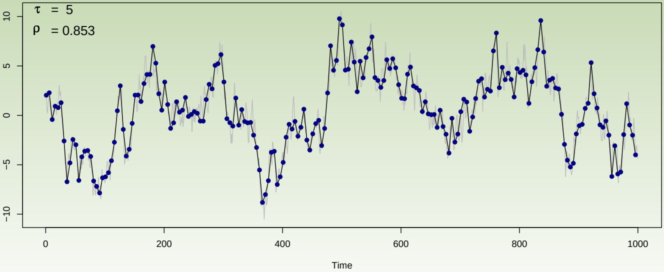
- Stationary
- Gaussian
- Modellable using standard time-series techniques

Properties of AR(1) time-series



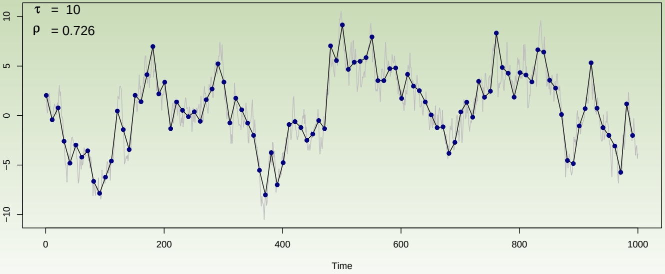
$$
\begin{aligned}
X_t &= \rho\left(X_{t-1} - \mu\right) + \mu + \epsilon \\
\epsilon &\sim \mathrm{N}(0, \sigma^2)
\end{aligned}
$$

Properties of AR(1) time-series



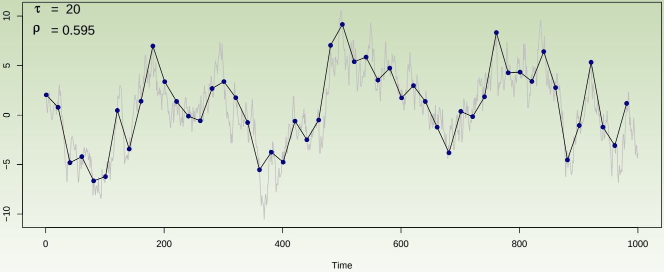
$$
\begin{aligned}
X_t &= \rho\left(X_{t-1} - \mu\right) + \mu + \epsilon \\
\epsilon &\sim \mathrm{N}(0, \sigma^2)
\end{aligned}
$$

Properties of AR(1) time-series



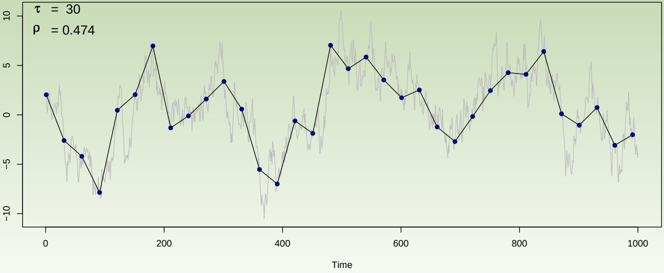
$$
\begin{aligned}
X_t &= \rho\left(X_{t-1} - \mu\right) + \mu + \epsilon \\
\epsilon &\sim N(0, \sigma^2)
\end{aligned}
$$

Properties of AR(1) time-series



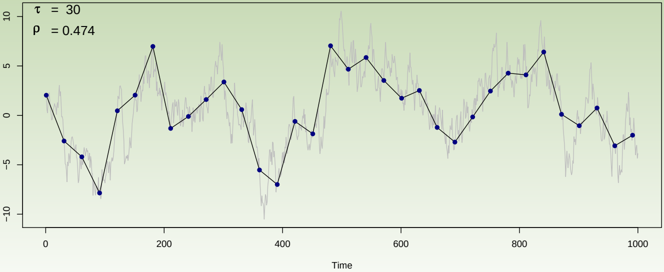
$$X_t = \rho\left(X_{t-1} - \mu\right) + \mu + \epsilon$$
$$\epsilon \sim \mathrm{N}(0, \sigma^2)$$

Properties of AR(1) time-series



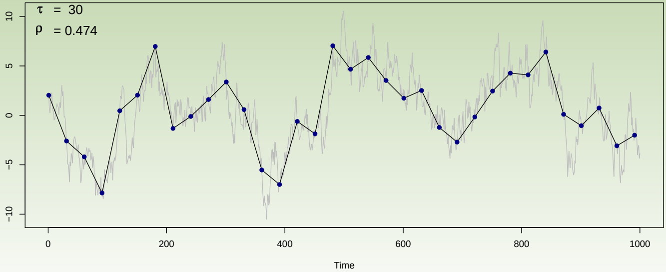
$$X_t = \rho\left(X_{t-1} - \mu\right) + \mu + \epsilon$$
$$\epsilon \sim N(0, \sigma^2)$$

Properties of AR(1) time-series



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

Properties of AR(1) time-series



$$f(X(t)|X(t-\tau)) \sim$$

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

Estimating ρ

Conditional Likelihood:

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

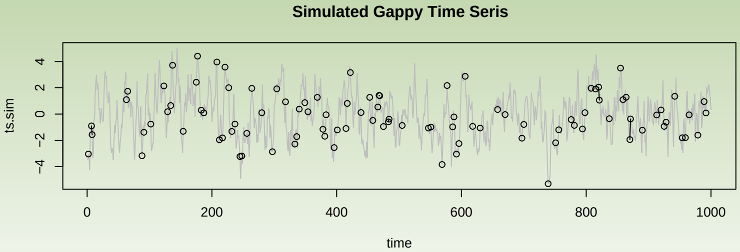
where:

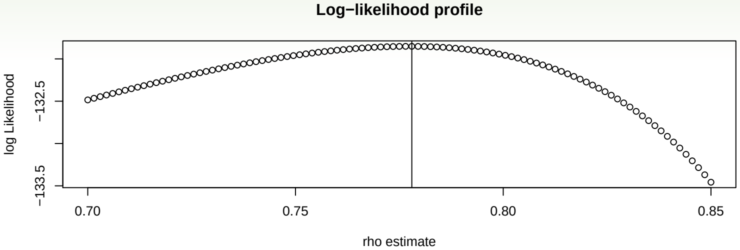
$$f(X_i|X_{i-1}) = \frac{1}{\sigma\sqrt{2\pi(1-\rho^{2\tau_i})}} \exp\left(\frac{(X_i-\rho^{\tau_i}(X_{i-1}-\mu))^2}{2\sigma^2(1-\rho^{2\tau_i})}\right).$$

then:

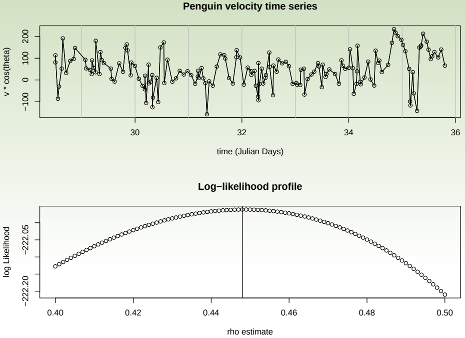
$$\widehat{\rho} = \underset{\rho}{\operatorname{argmax}}\ L(\rho|\mathbf{X},\mathbf{T})$$

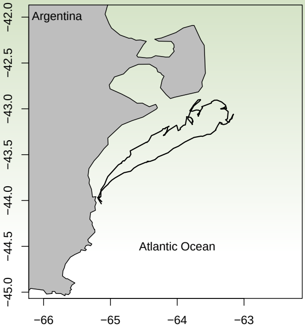
Estimating ρ





Estimating ρ





Structural shifts

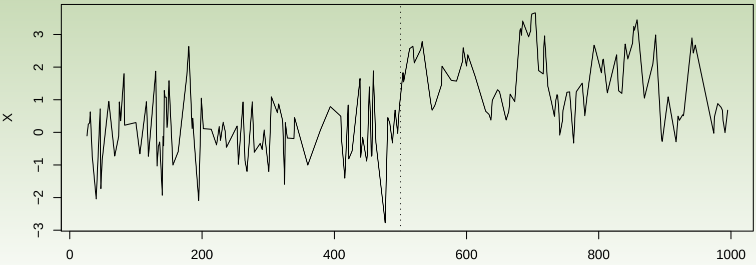
$$\mathbf{\Theta(t)} = \left\{ \begin{array}{lll} \mathbf{\Theta_1} & \text{if} & 0 < t \leq \tau \\ \mathbf{\Theta_2} & \text{if} & \tau < t \leq T \end{array} \right\}$$

$$L(\mathbf{\Theta}|\mathbf{X},\mathbf{T}) = \prod_{i=1}^{n} f(X_i|X_{i-1},\mathbf{\Theta}_1) \prod_{j=n+1}^{N} f(X_j|X_{j-1},\mathbf{\Theta}_2)$$

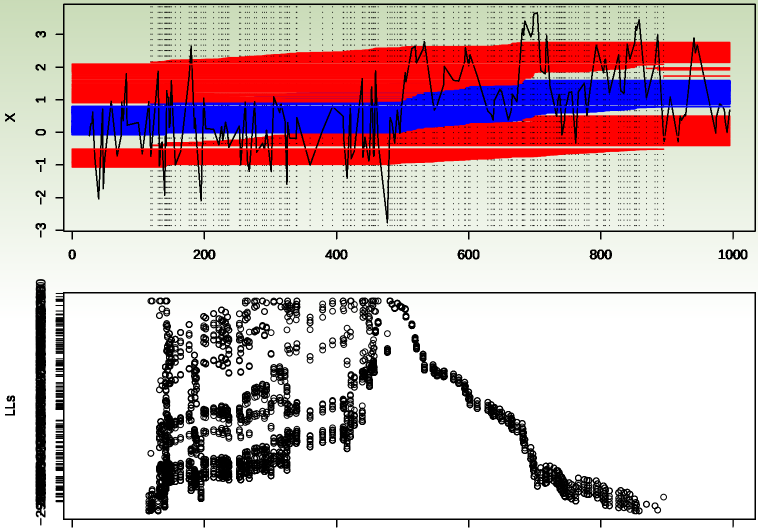
All estimates:

$$\begin{aligned} \widehat{n} &= \underset{n}{\operatorname{argmax}}\ L(\mathbf{\Theta}|\mathbf{X},\mathbf{T}) \\ \hat{\mu}_j &= \bar{X}_j \\ \hat{\sigma}_j &= S_j \\ \hat{\rho}_j &= \underset{\rho}{\operatorname{argmax}}\ L(\rho|\mathbf{X_j},\mathbf{T_j},\hat{\mu}_j,\hat{\sigma}_j) \end{aligned}$$

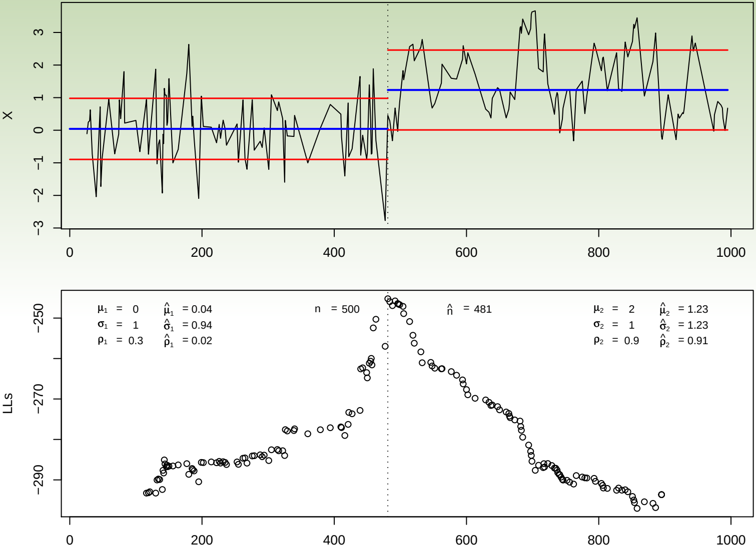
Identifying Change Point



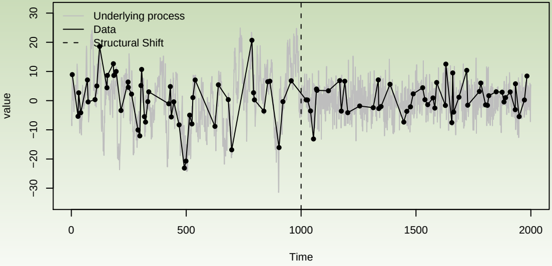
Identifying Change Point

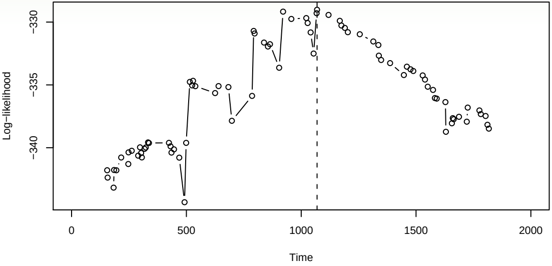


Identifying Change Point

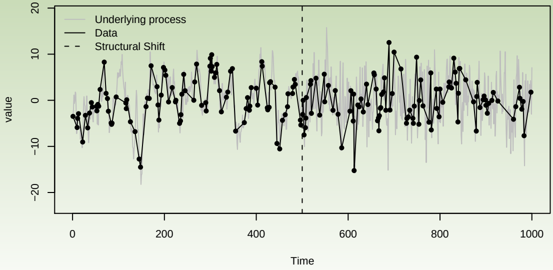


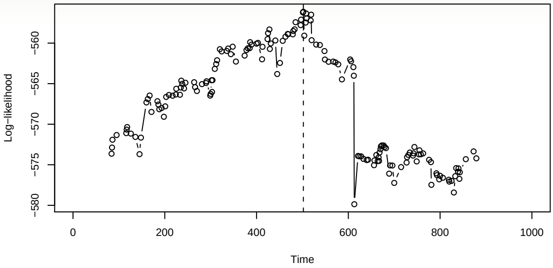
Identifying Change Point, sparce data





Identifying Change Point, different ρ's





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	$\mu_1 \neq \mu_2$	$\sigma_1 \neq \sigma_2$	$\rho_1 \neq \rho_2$

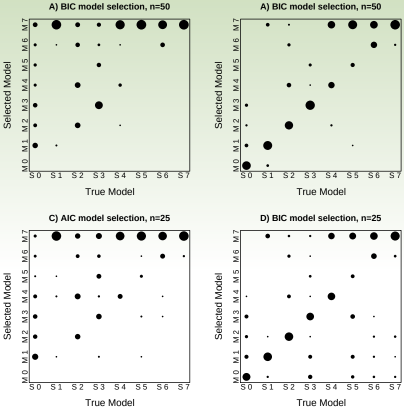
Identifying Models

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	$\mu_1 \neq \mu_2$	$\sigma_1 \neq \sigma_2$	$\rho_1 \neq \rho_2$

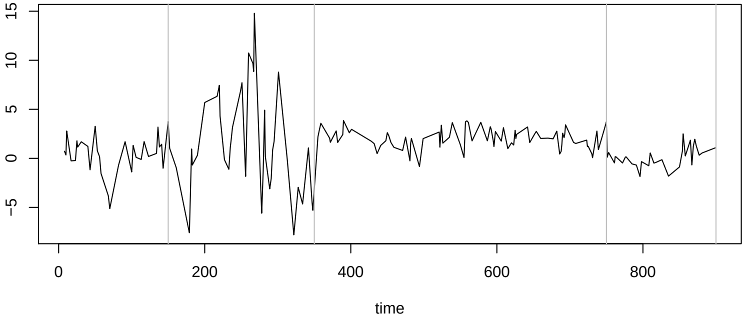
How to choose?

$$
\begin{aligned}
\text{AIC} : I_A(\mathbf{X}, \mathbf{T}) &= -2n \log \left(L(\hat{\theta} | \mathbf{X}, \mathbf{T}) \right) + 2d \\
\text{BIC} : I_B(\mathbf{X}, \mathbf{T}) &= -2n \log \left(L(\hat{\theta} | \mathbf{X}, \mathbf{T}) \right) + d \log(n)
\end{aligned}
$$

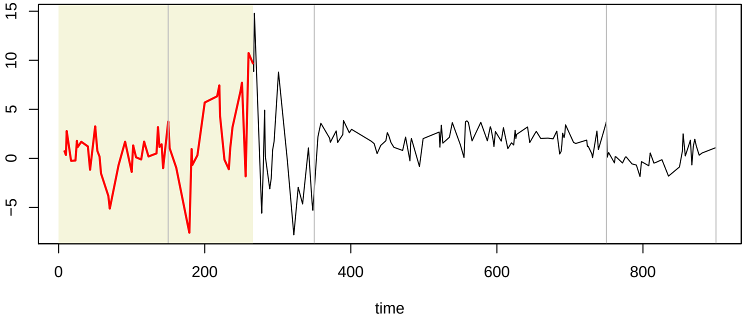
Identifying Models



	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**

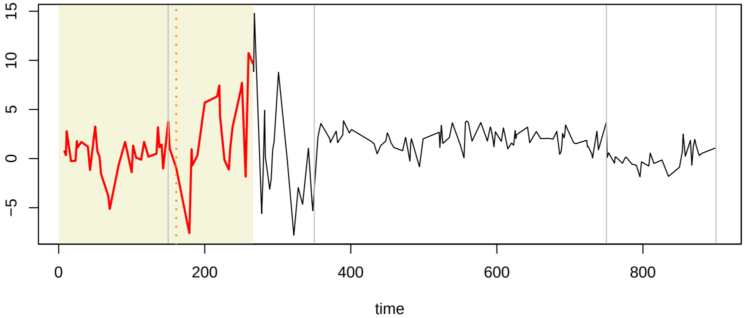


15
10
5
0
−5
0
200
400
600
800
time



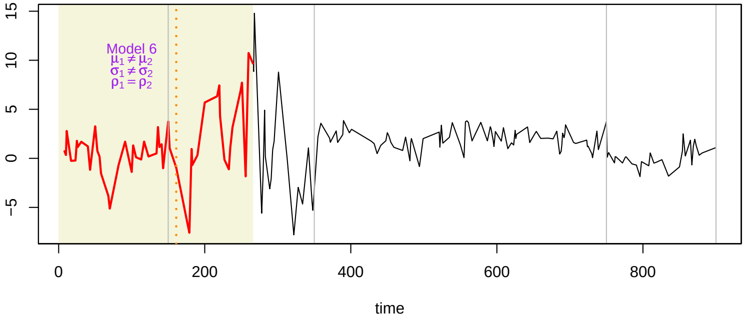
Algorithm for Identifying Multiple Changepoints

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



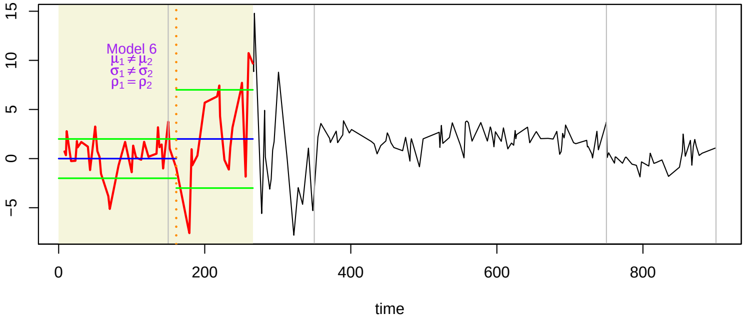
Algorithm for Identifying Multiple Changepoints

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



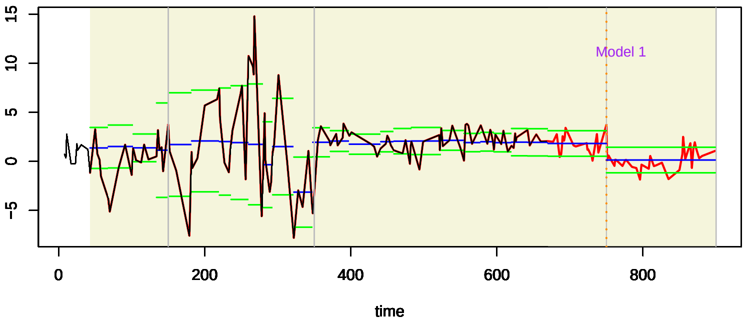
Algorithm for Identifying Multiple Changepoints

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



Algorithm for Identifying Multiple Changepoints

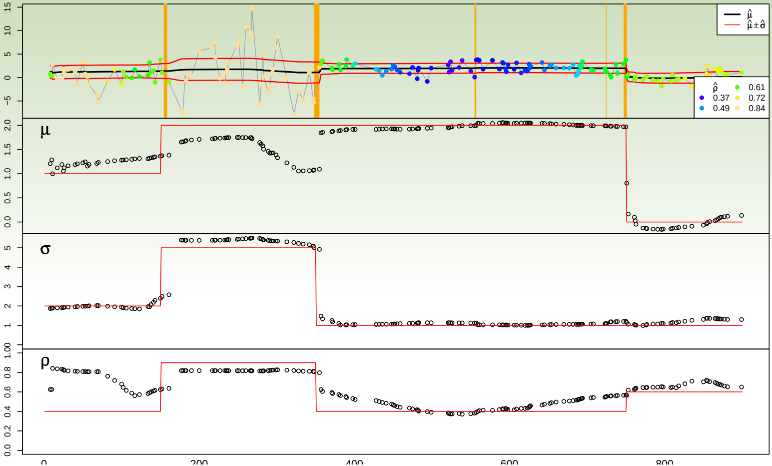
- Select Window
- Find MLBP
- Identify Model
- Record estimates based on model selected.
- Move window forward and repeat



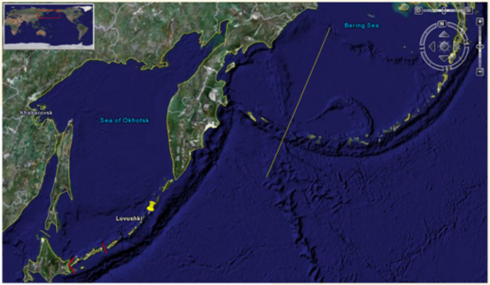
Algorithm for Identifying Multiple Changepoints

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

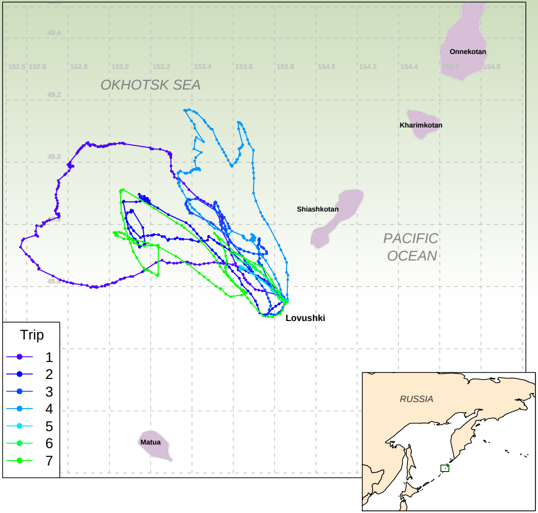
Movement analysis output



Northern Fur Seal (*Callorhinus ursinus*)

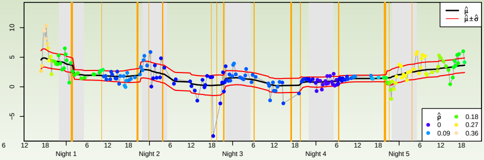


Map of all tracks

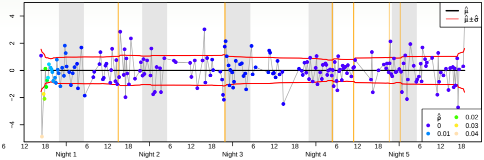


Track Analysis

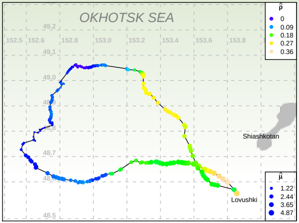
Vcos(θ)



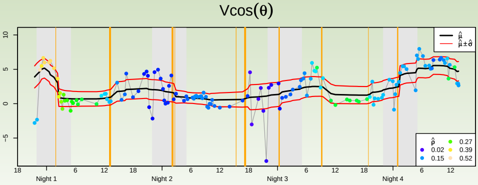
Vsin(θ)

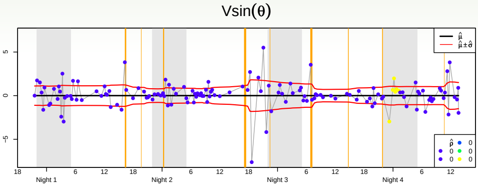


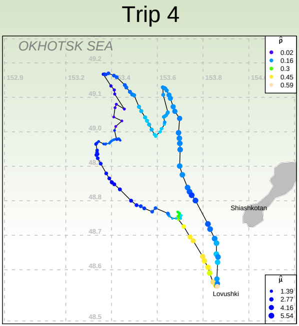
Trip 1



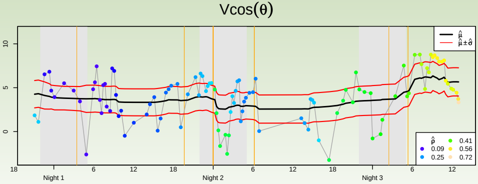
Track Analysis

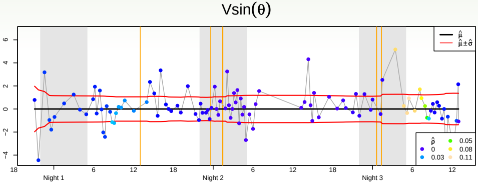


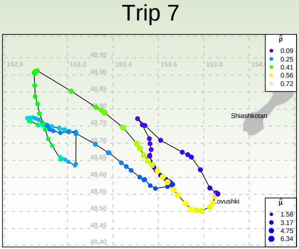




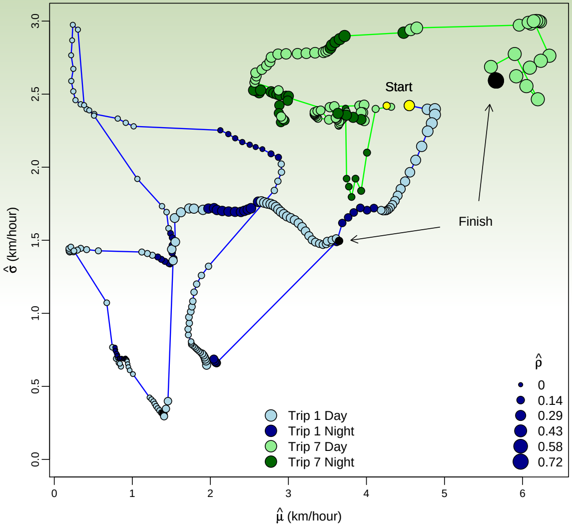
Track Analysis



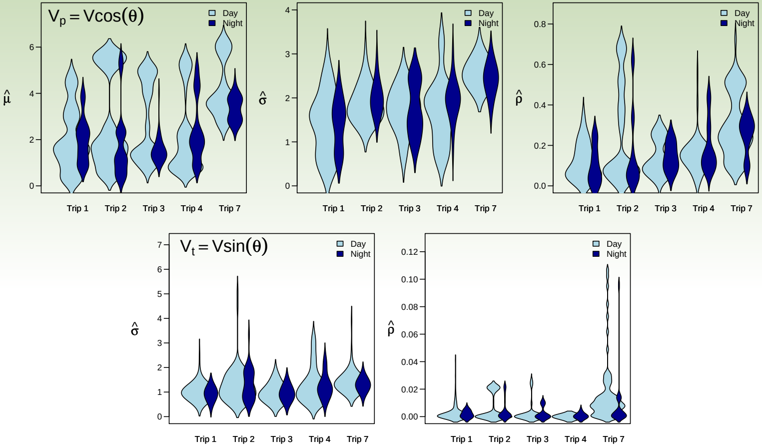




Behavioral Phaseplot



Violin plots



Conclusions:

Conclusions:

- Behavior can be very complex!
- But patterns *can* be robustly picked out of messy data.
- Method suggests the possibility of asking more sophisticated questions.

Conclusions:

- Behavior can be very complex!
- But patterns *can* be robustly picked out of messy data.
- Method suggests the possibility of asking more sophisticated questions.

Conclusions:

- Behavior can be very complex!
- But patterns *can* be robustly picked out of messy data.
- Method suggests the possibility of asking more sophisticated questions.

Acknowledgements

Data:

- Fur Seals: R. Andrews, V. Burkanov, Russian Far East Marine Mammal Project
- Penguin: Elizabeth Skewgar, P. Dee Boersma

Funding:

- University of Washington School of Arts and Sciences
- National Marine Fisheries Service

Inspiration:

- V. Minin, T. Gneiting, D. Grünbaum
- Colleagues and friends at QERM, SAFS and beyond
- Metapopulation Research Group in Helsinki, Finland