

Multi-state Models for Animal Movement Analysis

SSC 2016

Aurelien Nicosia

May 30, 2016

Motivation

In *ecology* understanding and modeling the movement of animals is fundamental:

—> predicting the spacial organization of animals is useful for

1. organizing the reintroduction of animal species;
2. planing a vaccination campaign;
3. or preventing the transport of harmful particles by animals.

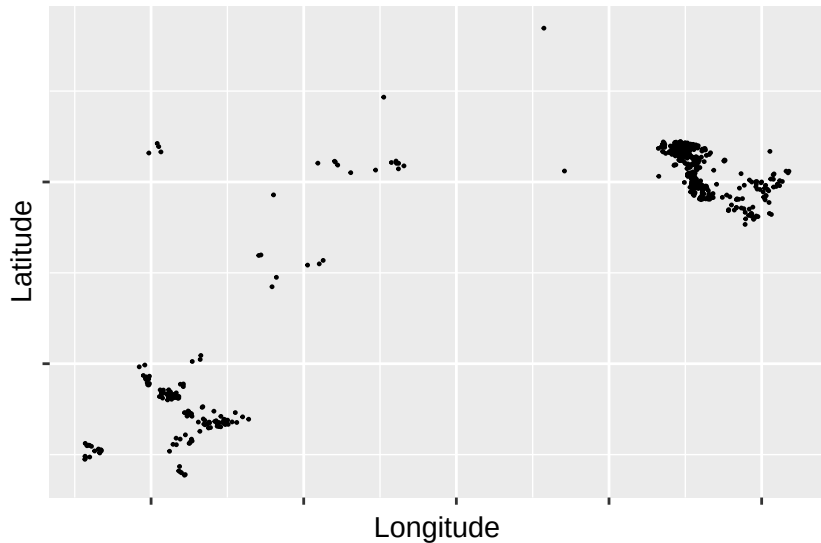
Let me tell you a story

Statistics in perpetual motion to help animals

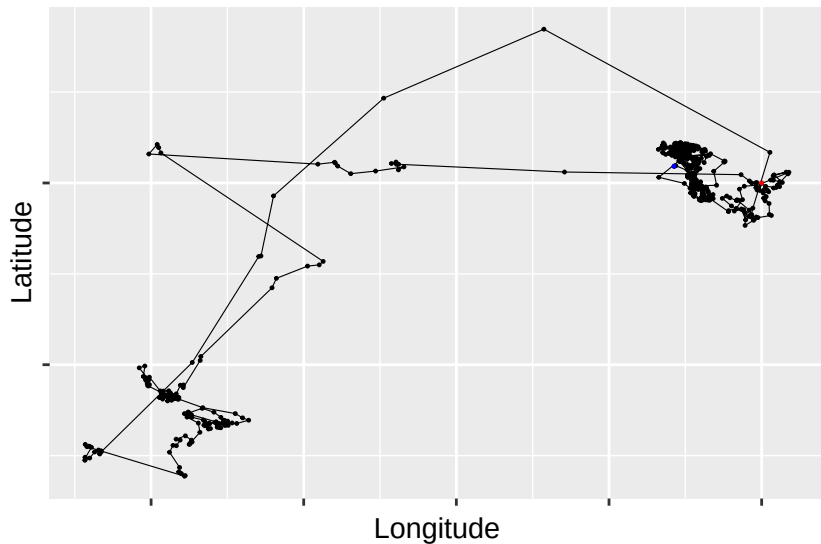
Preface : Available data

- ▶ Collar with Global Positioning System (GPS) positionning
- ▶ Geographic Information System (GIS) for environmental characterization

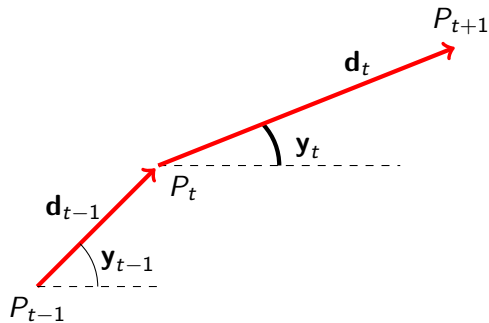
Preface: What do we want to model?



Chapter 1: Movement



Random walk model

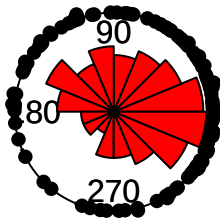


Directional statistics

- ▶ direction = angle = periodic data
- ▶ Directional Statistics, Mardia & Jupp, 2000
- ▶ R: package circular

Directional persistence $y_t - y_{t-1}$ on long traveled distances

```
caribou.long = subset(caribou, d>d.crit)
direc.persis <- circular(diff(as.vector(caribou.long$y)))
plot(direc.persis, stack = TRUE, bins = 300, cex = 1, tcl.tex
rose.diag(direc.persis, bins = 16, col = "red",
          cex = 1, prop = 2.5, add = TRUE, tcl.text = 0.2, u
```



```
mean(direc.persis)[[1]] # different from 0!
```

```
## [1] 0.404587
```

Angular regression model

- ▶ Targets, such as wood cut, road or even previously visited location can influence the trajectory.

→ How to handle several explanatory variables to explain the direction in the trajectory?

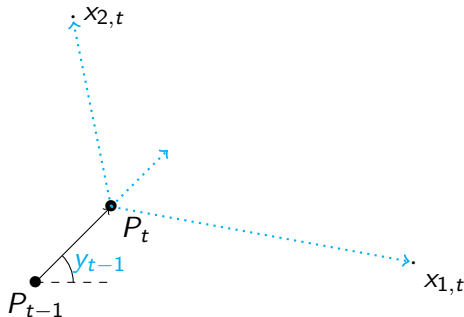
Angular regression model

Let x_t denote potential explanatory targets:

$$y_t = \mu_t(x_t) + \epsilon_t$$

where $\mu_t(x_t)$ is the direction of the vector

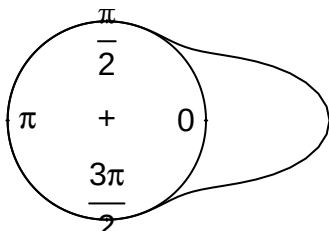
$$\mathbf{v}_t^{(k)} = \kappa_0^{(k)} \begin{pmatrix} \cos(y_{t-1}) \\ \sin(y_{t-1}) \end{pmatrix} + \sum_{i=1}^p \kappa_i^{(k)} z_{it} \begin{pmatrix} \cos(\theta_{it}) \\ \sin(\theta_{it}) \end{pmatrix},$$



What about the error?

The von Mises distribution

```
ff <- function(x) dvonmises(x, mu=circular(0), kappa=10)  
curve.circular(ff, join=TRUE, xlim=c(-2.3, 1), tcl.text=0.2)
```



Result of the regression model

```
consensus(formula = y ~ y.previous + xcut + xcenter,  
           data = na.omit(caribou))
```

```
##
```

```
## Maximum log-likelihood : 11.59544
```

```
##
```

```
## Parameters:
```

##	estimate	stderr	z value	P(z >.)
## y.previous	0.1058897	0.05752115	1.840883	0.06564
## xcut	0.1555772	0.06401522	2.430316	0.01509
## xcenter	0.2771297	0.06412821	4.321494	0.00002

What about the distances

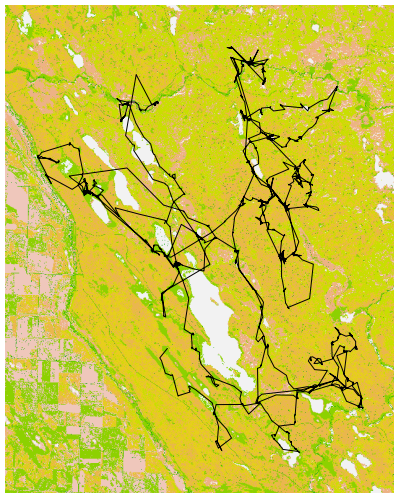
d_t , $t = 1, \dots, T$ are assumed i.i.d. from a gamma distribution:

```
mle.d <- mledist(caribou$d, distr = "gamma");  
mle.d$estimate
```

```
##      shape      rate  
## 0.6774039 2.1709335
```

Chapter 2: Local choice

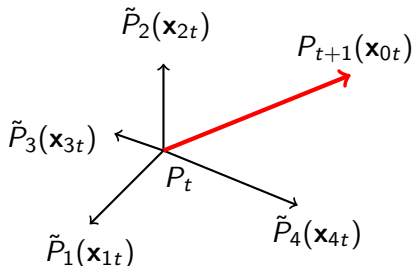
- * In a random walk model the targets are variables of the environment from a *global* point of view.
- * Do animals move to the next location with a more local choice of habitat?



Local discrete choice model

Why does, from the position P_t , the animal choose the position P_{t+1} for the next step instead of other close locations?

What are the “environmental determinants” of this choice?



Conditional logistic regression

Let us consider J control locations at each time step t :

$$q_t = \frac{\exp(\mathbf{x}_{0t}^T \beta)}{\sum_{j=0}^J \exp(\mathbf{x}_{jt}^T \beta)}$$

q_t is the probability that the animal chooses its actual next step.
Two questions remains:

- ▶ The choice of the explanatory variables x
- ▶ How to sample the J control locations?

Application

```
formula1 = Surv(Times, Cote) ~ eau + Dec + meadow +  
  road + strata(idstrate)  
coxph(formula1, data = data)
```

```
## Call:
```

```
## coxph(formula = formula1, data = data)
```

```
##
```

```
##
```

##		coef	exp(coef)	se(coef)	z	p
##	eau1	-1.0241	0.3591	0.1993	-5.14	2.8e-07
##	Dec1	-0.4290	0.6512	0.0574	-7.47	7.9e-14
##	meadow1	2.1278	8.3961	0.0504	42.23	< 2e-16
##	road1	1.0379	2.8234	0.3516	2.95	0.0032

```
##
```

```
## Likelihood ratio test=3078 on 4 df, p=0
```

```
## n= 64512, number of events= 3072
```

Chapter 3: Movement and local choice?

- ▶ How can we handle movement and selection of habitat in the same model?
- ▶ How can the random walk model and the conditional logistic regression model meet?

Help of Monte Carlo!

Distribution from exponential family and by Law of Large Numbers:

$$\begin{aligned}f_X(x|\boldsymbol{\eta}) &= \exp\left(\boldsymbol{\eta}^T \cdot \mathbf{T}(x) - A(\boldsymbol{\eta})\right) \\&= \frac{\exp\left(\boldsymbol{\eta}^T \cdot \mathbf{T}(x)\right)}{\int \exp\left(\boldsymbol{\eta}^T \cdot \mathbf{T}(u)\right) du} \\&\approx \frac{\exp\left(\boldsymbol{\eta}^T \cdot \mathbf{T}(x)\right)}{\frac{1}{J} \sum_{j=1}^J \exp\left(\boldsymbol{\eta}^T \cdot \mathbf{T}(x_j)\right)} \\&\rightarrow \text{it looks like Conditional logistic regression model}\end{aligned}$$

where $x_j, j = 1, \dots, J$ uniformly distributed.

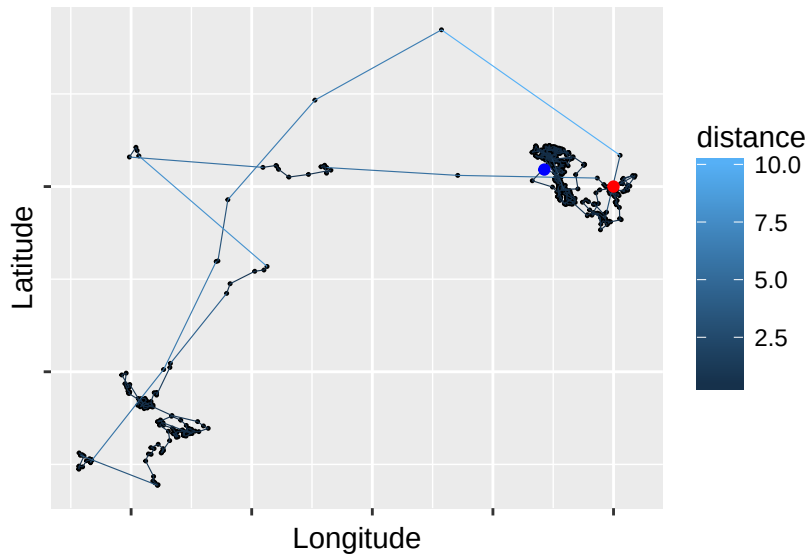
Movement and local choice together

In the conditional logistic regression model, since the random walk model is from the exponential family:

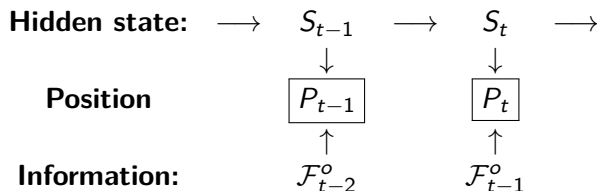
$$q_t = \frac{\exp\left((\cos(y_t - y_{t-1}), \cos(y_t - \theta_t), \log(d_t), -d_t, \mathbf{x}_{0t})^T \beta\right)}{\sum_{j=0}^J \exp\left((\cos(\phi_j - y_{t-1}), \cos(\phi_j - \theta_t), \log(h_j), -h_j, \mathbf{x}_{jt})^T \beta\right)}$$
$$\approx \text{von Mises} \times \text{Gamma} \times \frac{\exp\left(\mathbf{x}_{0t}^T \beta\right)}{\sum_{j=0}^J \exp\left(\mathbf{x}_{jt}^T \beta\right)}$$

when J is large and the control locations uniformly sampled in a large disc.

Chapter 4: Different behaviors?

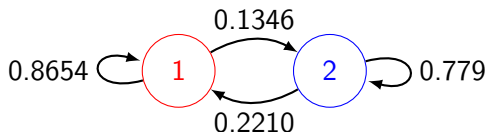


Development



- $\mathcal{F}_t^o$: all the observed information (trajectory of the animal, explanatory variables) gathered from time 0 up to time t .

Multi-state Model for the Bison



	estimate	s.e.	estimate	s.e.
$\beta_{\text{cos.persis}}$	-0.5079	0.0618	0.3362	0.0590
$\beta_{\text{dist.neg}}$	5.5462	0.7397	0.0225	0.0315
$\beta_{\text{dist.log}}$	0.2234	0.1114	-0.5046	0.0550
β_{water}	-0.5236	0.3829	-1.2321	0.4169
β_{dec}	0.1202	0.1266	-0.4898	0.1105
β_{meadow}	0.7637	0.1254	1.6535	0.1059
β_{road}	1.2228	1.7055	1.4853	0.4761

Final chapter

With the mix of these models:

- ▶ random walk = angular regression model for the direction + gamma for the distances
- ▶ local choice = conditional logistic regression model
- ▶ extension to hidden Markov model

We are able to much better understand the multiple behaviors of an animal in its environment.

Bison's behaviors



Figure 1:

Thank you for you attention

Directional statistic

```
library(circular)
plot(caribou$y, stack = TRUE, bins = 300, shrink= 1, cex = 1)

## Warning in plot.window(...): "stack" is not a graphical
## Warning in plot.window(...): "bins" is not a graphical p
## Warning in plot.window(...): "shrink" is not a graphical
## Warning in plot.window(...): "tcl.text" is not a graphi
## Warning in plot.window(...): "units" is not a graphical
## Warning in plot.xy(xy, type, ...): "stack" is not a grap
## Warning in plot.xy(xy, type, ...): "bins" is not a graph
## Warning in plot.xy(xy, type, ...): "shrink" is not a gra
```

Directional statistic

Summary of angular variable:

```
mean(caribou$y)[[1]]
```

```
## [1] 0.07871534
```

```
rho.circular(caribou$y) # dispersed ~ 0; concentrated ~ 1
```

```
## [1] 0.03797862
```

Directional persistence $y_t - y_{t-1}$

```
direc.persis <- circular(diff(as.vector(caribou$y)))  
plot(direc.persis, stack = TRUE, bins = 300, cex = 1, tcl.tex  
rose.diag(direc.persis, bins = 16, col = "red",  
          cex = 1, prop = 2.5, add = TRUE, tcl.text = 0.2, u
```

