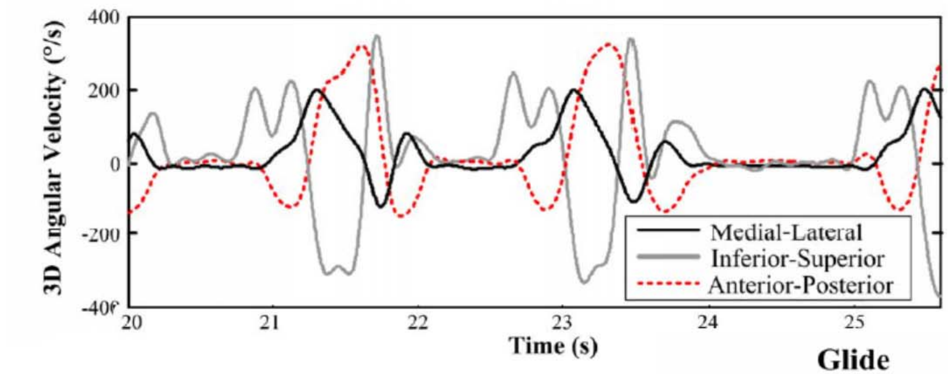
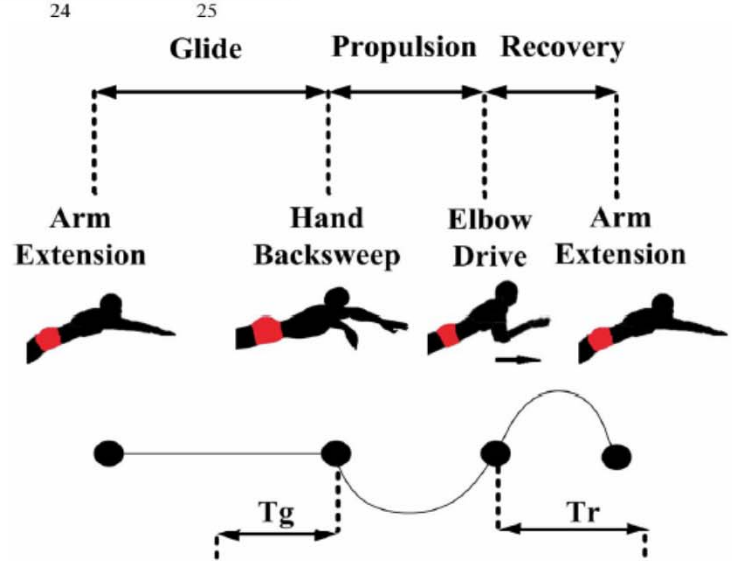
TIME SERIES DATA AND HIDDEN MARKOV MODELS

Paul Wu

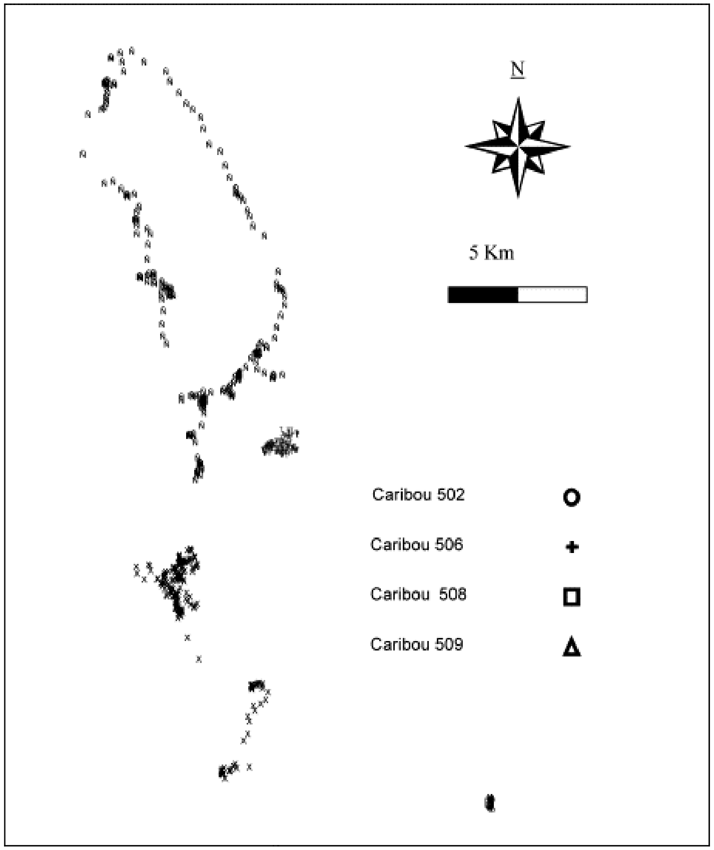
Time Series Analysis: Breaststroke



(Dadashi, 2013)



Spatio-Temporal Analysis



Observations: distance travelled and turn angle

3 behavioural states (bedding, foraging and relocating)

(Franke, 2004)

Hidden Markov Models (HMMs)

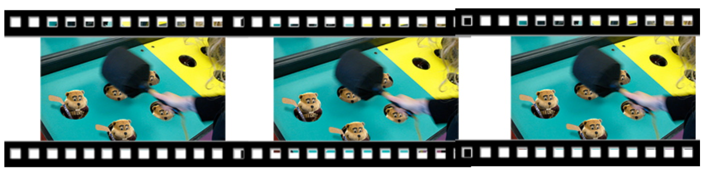
- What characterises these problems?
  - Multivariate time series
  - Data at each point in time is not independent
  - Data distribution may change over time
  - Data sequences not of the same length
- Why HMM?
  - Want to understand some process of a dynamical system
  - Unobserved/latent variables
  - Simple and computational efficiency

Where are HMMs used?

- Widely used (Zucchini, 2016):
  - Recognition (face, handwriting)
  - Bioinformatics
  - Environment (rainfall, earthquakes)
  - Finance
  - Ecology (animal movements)

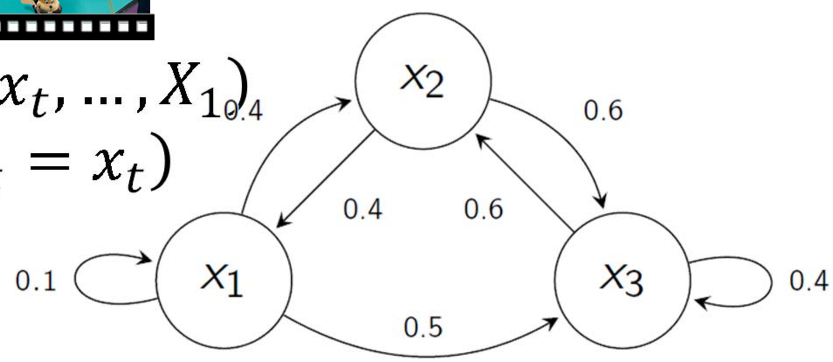
Basics: Markov Chains

Markov chain = sequence of random variables $X_1, X_2, \ldots, X_t$ where:



https://fixhepc.com/support-forum/geopolitics/1250-whack-a-mole.html
https://gophersunlimited.com/trapping-moles/

$$P(X_{t+1} = x_{t+1} | X_t = x_t, \ldots, X_1) = P(X_{t+1} = x_{t+1} | X_t = x_t)$$



(Frazzoli, 2010)

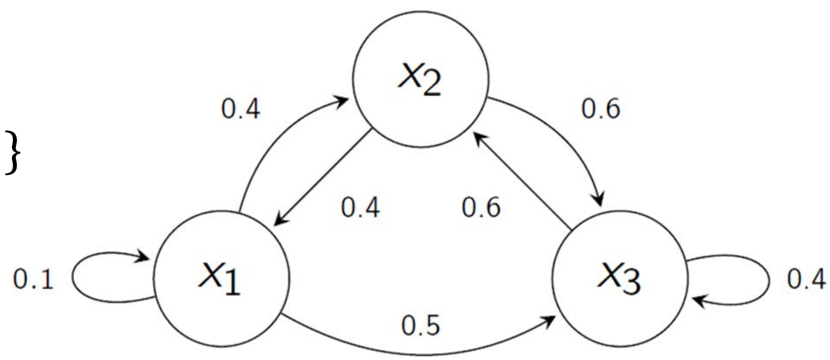
QUT

ACEMS
MATHEMATICAL AND STATISTICAL FRONTIERS

Transition Matrix

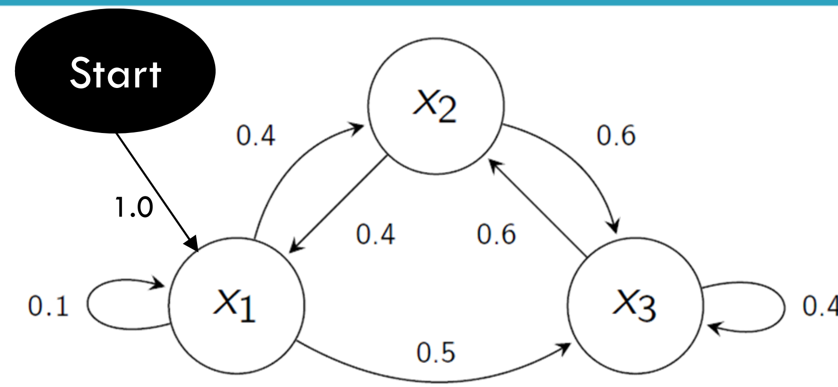
$$T = P\left(x_j \middle| x_i\right) = \begin{array}{cc} \text{from\textbackslash to} & \begin{array}{ccc} x_1 & x_2 & x_3 \end{array} \\ \begin{array}{c} x_1 \\ x_2 \\ x_3 \end{array} & \begin{pmatrix} 0.1 & 0.4 & 0.5 \\ 0.4 & 0 & 0.6 \\ 0 & 0.6 & 0.4 \end{pmatrix} \end{array}$$

Finite set $X = \{x_1, x_2, x_3\}$



Transition Matrix

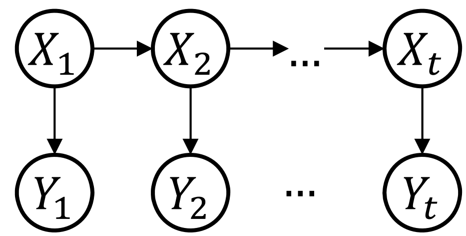
- If we know the state at time step 1 is $X_1 = x_1$, then $p_1 = \pi = (1{,}0{,}0)$.
- $p_2 = T'p_1 = (0.1{,}0.4{,}0.5) = (0.1 \times 1{,}0.4 \times 1{,}0.5 \times 1)$



- $p_3 = (T')^2 p_1 = (0.17{,}0.34{,}0.49) = (0.1 \times 0.1 + 0.4 \times 0.4{,}0.1 \times 0.1 + 0.6 \times 0.5{,}0.1 \times 0.5 + 0.6 \times 0.4 + 0.4 \times 0.5)$
- $p_t = (T')^t p_1 = (0.15{,}0.36{,}0.49)$
- For some systems, stationary distribution $p^\infty = T'p^\infty$

Hidden Markov Model

- States X (finite set)
- Observations Y (finite set)
- T (transition probabilities)
- Q (emission=measurement probabilities)
- π (initial/prior state probability distribution)



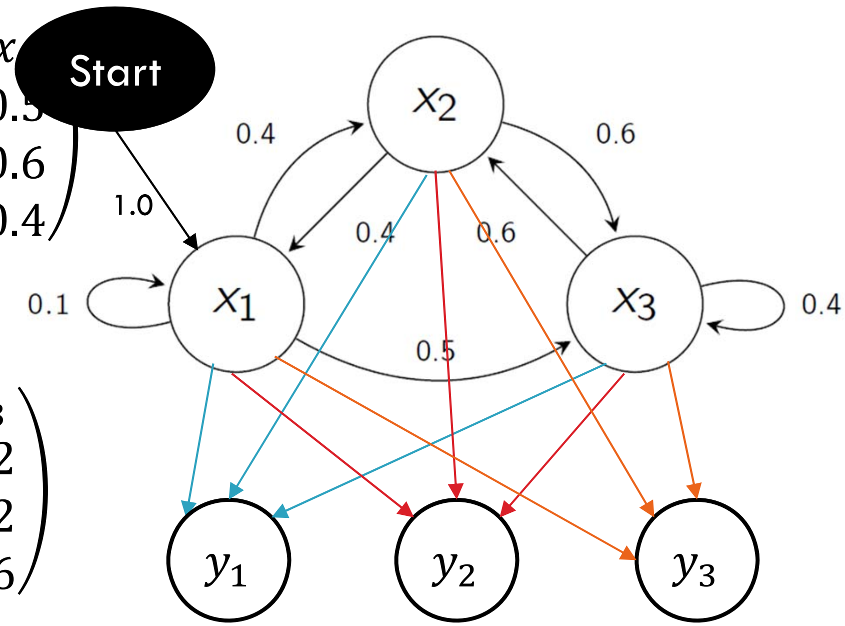
Moles at Night: HMM

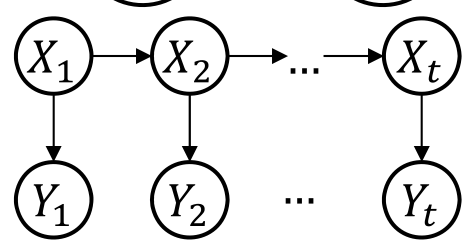
$$T = \begin{array}{c} \\ \\ \\ \end{array}\begin{pmatrix} \text{from\\to} & x_1 & x_2 & x_3 \\ x_1 & 0.1 & 0.4 & 0.5 \\ x_2 & 0.4 & 0 & 0.6 \\ x_3 & 0 & 0.6 & 0.4 \end{pmatrix}$$

$$Q = P(y_j | x_i)$$

$$= \begin{pmatrix} \text{from\\to} & y_1 & y_2 & y_3 \\ x_1 & 0.6 & 0.2 & 0.2 \\ x_2 & 0.2 & 0.6 & 0.2 \\ x_3 & 0.2 & 0.2 & 0.6 \end{pmatrix}$$

$$p_1 = \pi = (1,0,0)$$





HMM: What Can It Do?

- Given observations y up to time t, we can do:
  - Filtering – what is the current state distribution $X_t = p_t$
  - Smoothing – hindsight, what is $X_u, u < t$
  - Prediction – forecast, what is $X_u, u > t$
  - Decoding – find the most likely state history $X_{1,\dots,t}$.

HMM Types

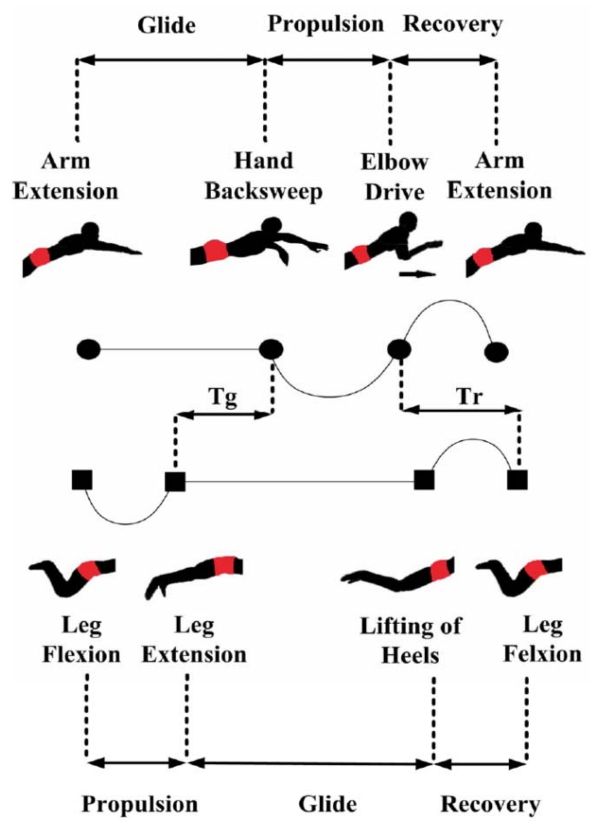
- Discrete time
  - Discrete observations (genetics: observe letters A, G, C, T)
  - Continuous observations - q_{x_i} has a distribution (e.g. Gaussian, Poisson)
  - Covariates – prior, transition
- Continuous time

Approach for Today

- Case study 1: swimming (Dadashi, 2013)
- Case study 2: animal movement (Franke, 2004)
- Practical: response time (Dutilh, 2011)

Case Study 1: Swimming

- 7 swimmers, 2 100Hz IMUs (arm + leg)
- 3x200m trials with target speeds
- Temporal phases of locomotion – arm-leg motor organisation affects speed fluctuations and energy
- Total time gap = Tg + Tr



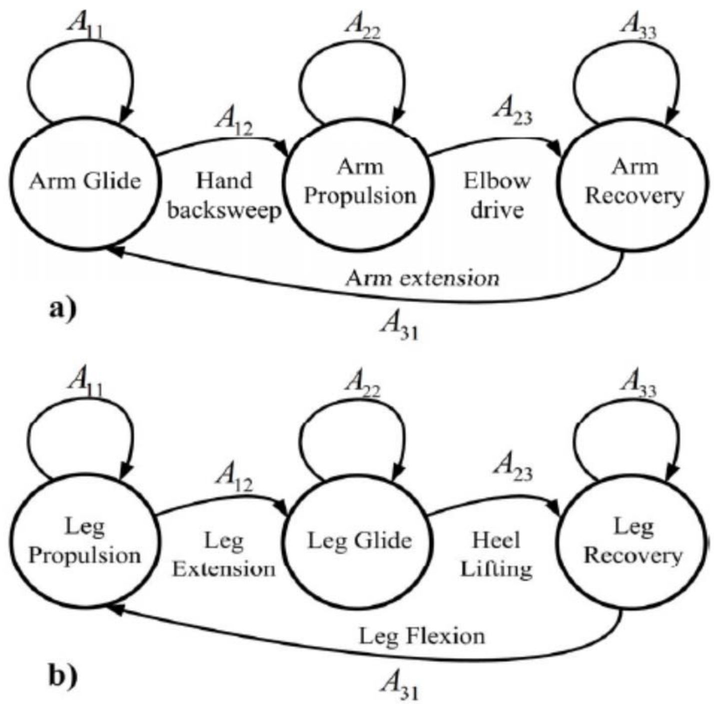
(Dadashi, 2013)

Case Study 1: Data

- Observations: IMU sensors – medial-lateral, inferior-superior, anterior-posterior angular velocities for arm and leg
- States and thus prior and transitions: expert annotated video

Case Study 1: Goal

Find Tg via the most probable sequence of states



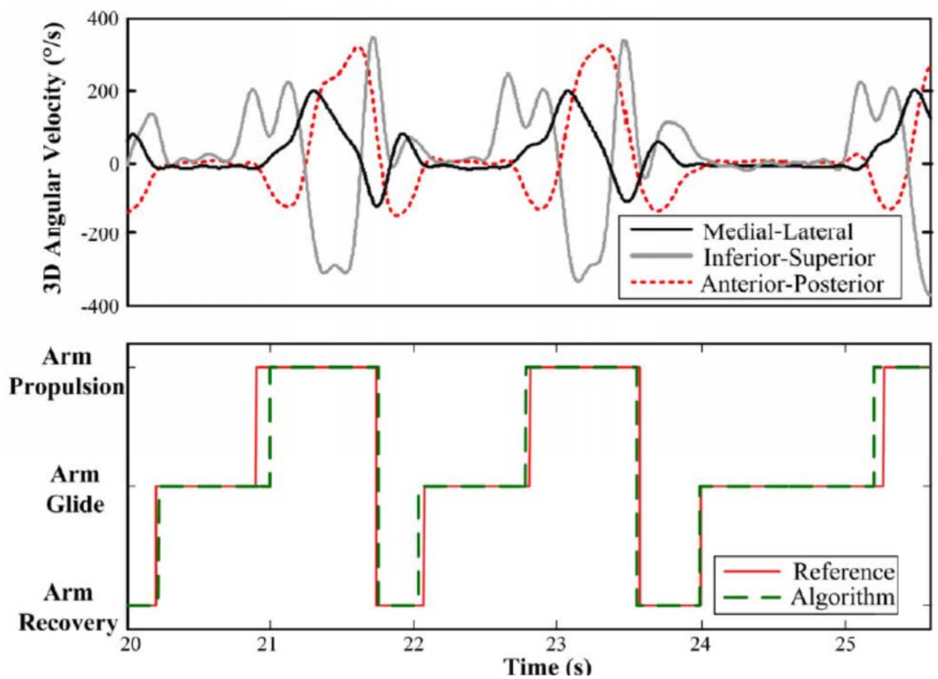
(Dadashi, 2013)

Case Study 1: Estimating Parameters

- π: Prior estimated as frequency of states in training set
- T: Transition probabilities estimated as #samples in training set with transition from state i to j divided by total #samples labelled j.
- Q: Emissions modelled as multi-variate Gaussian mixtures.
  - If Y is discrete, q is estimated #outputs i from state j divided by total #samples of state j.
  - mclust, mixtools

Supervised learning

Results



(Dadashi, 2013)

Results and Learnings

- Error in estimating the state/phase:
  - Sensitivity 93.5% for arm, 94.4% for leg; specificity of 96.2% and 97.2%, respectively
  - Boosted model – use output of classifiers (linear + non-parametric estimator, Gaussian kernel) as input to HMM
- Error in estimating TG: -11 ± 52ms

Case Study 2: Caribou Movements

- Developing a set of interpretable states for caribou behaviours
  - State and transitions ('strategy') not known/defined in advance!
  - Postulate 3 hidden states – bedding, feeding and relocating
- 12 caribou over 10 days with GPS tracking
- Goal: interpret duration of behaviour and transition to different behaviours

QUT ACEMS AUSTRALIAN RESEARCH COUNCIL CENTRE OF EXCELLENCE FOR MATHEMATICAL AND STATISTICAL FRONTIERS

(Franke, 2004)

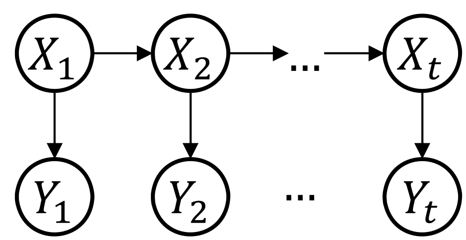
Case Study 2: Data

- Distance moved and turn angle
  - Distance discretised into zero <20m, short 20-100m, medium 100-250m, long >250m
  - Angles discretised: ahead 316-45 °, right 46-135°, back136-225°, left 226-315°.

QUT ACEMS AUSTRALIAN RESEARCH COUNCIL CENTRE OF EXCELLENCE FOR MATHEMATICAL AND STATISTICAL FRONTIERS

Case Study 2: Estimating Parameters

- Expectation maximisation (EM)
  - “Missing data” are the states
  - E: estimate probability of the state sequence given observations
  - M: fit new model parameters
  - R – depmixS4



Case Study 2: Results

Table 4
Multiple-observation HMM for caribou 506

	State $1_{(t+1)}$ (B)	State $2_{(t+1)}$ (F)	State $3_{(t+1)}$ (R)
A: State transitions			
State $1_{(t)}$ (B)	0.87	0.037	0.09
State $2_{(t)}$ (F)	0.90	0.04	0.06
State $3_{(t)}$ (R)	0.14	0.02	0.84

Table 5
Multiple-observation HMM for caribou 508

	State $1_{(t+1)}$ (B)	State $2_{(t+1)}$ (F)	State $3_{(t+1)}$ (R)
A: State transitions			
State $1_{(t)}$ (B)	0.68	0.29	0.02
State $2_{(t)}$ (F)	0.69	0.28	0.03
State $3_{(t)}$ (R)	0.21	0.10	0.69

(Franke, 2004)

Case Study 2: Results

Table 5
Multiple-observation HMM for caribou 508

	State $1_{(t+1)}$ (*B*)	State $2_{(t+1)}$ (*F*)	State $3_{(t+1)}$ (*R*)		
A: State transitions					
State $1_{(t)}$ (*B*)	0.68	0.29	0.02		
State $2_{(t)}$ (*F*)	0.69	0.28	0.03		
State $3_{(t)}$ (*R*)	0.21	0.10	0.69		
	Stationary	Short	Medium	Long	
B1: Distance between locations					
State $1_{(t)}$ (*B*)	0.99	0.01	0.00	0.00	
State $2_{(t)}$ (*F*)	0.91	0.09	0.00	0.00	
State $3_{(t)}$ (*R*)	0.00	0.85	0.11	0.04	
	Stationary	Ahead	Right	Left	Backward
B2: Turn angle					
State $1_{(t)}$ (*B*)	1.00	0.00	0.00	0.00	0.00
State $2_{(t)}$ (*F*)	1.00	0.00	0.00	0.00	0.00
State $3_{(t)}$ (*R*)	0.00	0.50	0.11	0.17	0.22

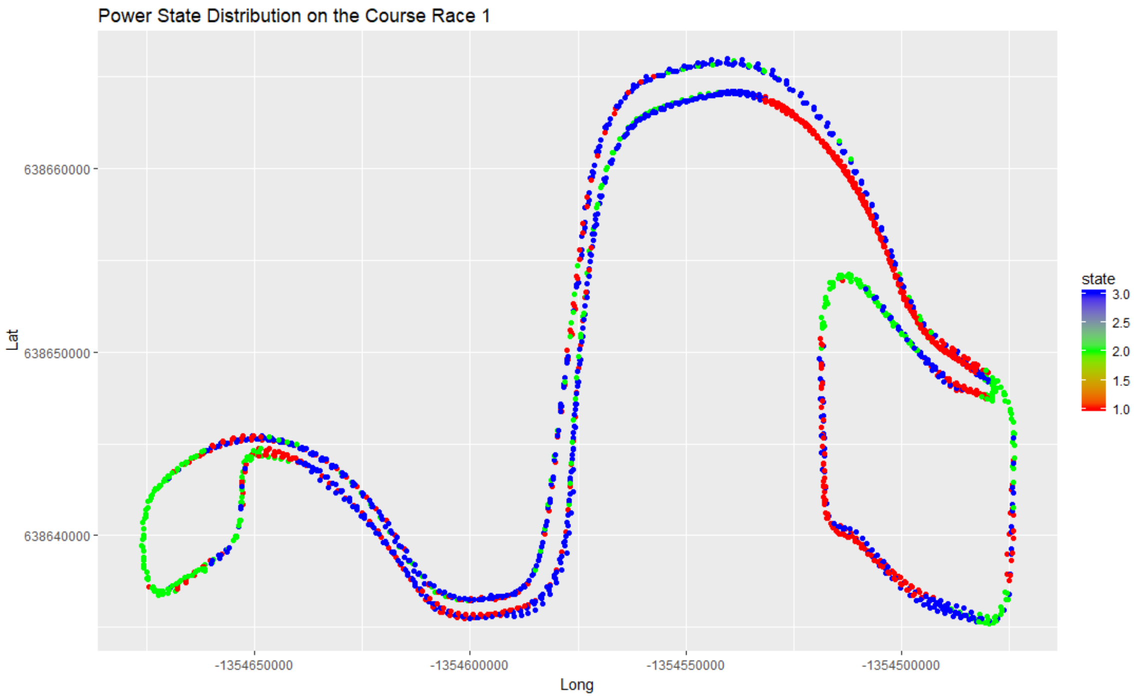
Results and Learnings

- Relating behaviour to habitat and conservation
- Challenges
  - Potential for insertion and deletion of states
  - Initialisation and convergence to local maximum
  - Constraints: challenging for EM and result in incorrect parameter estimates
    - depmixs4: direct optimisation

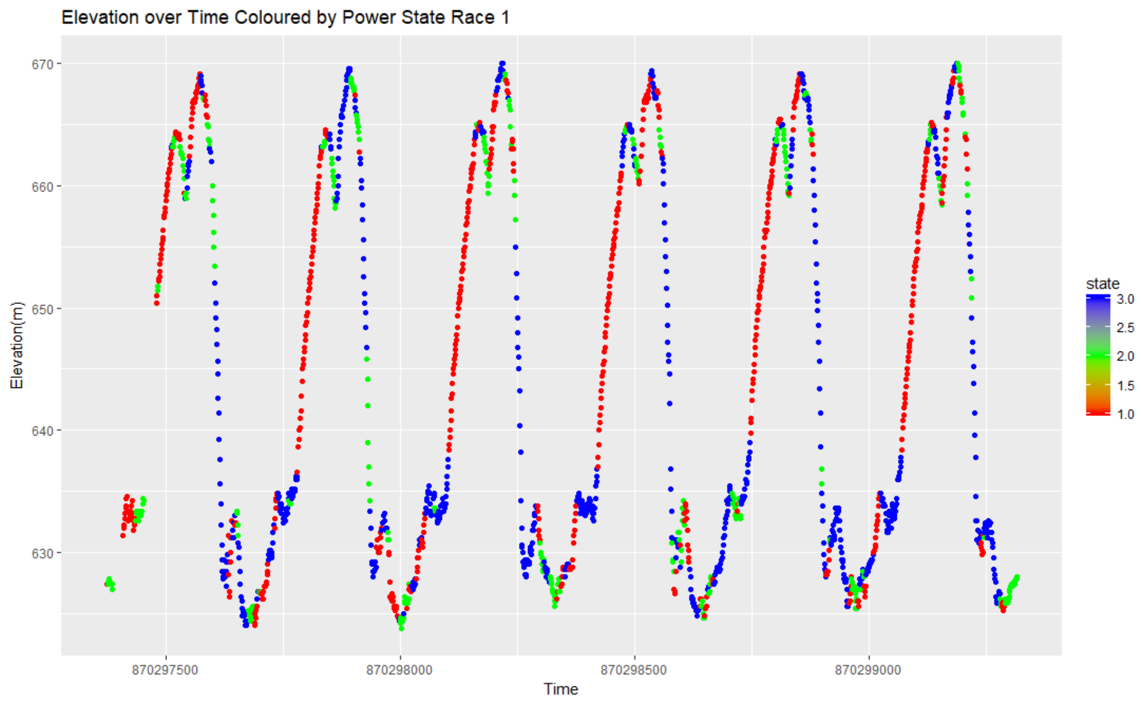
Results and Learnings

- Covariates – affect prior or transition
  - Time of day
  - Season
  - Spatial – terrain and vegetation

QAS Example of HMM



QAS Example for HMM



Reflection

- How might HMMs be used for your work?

Prac: Response Time

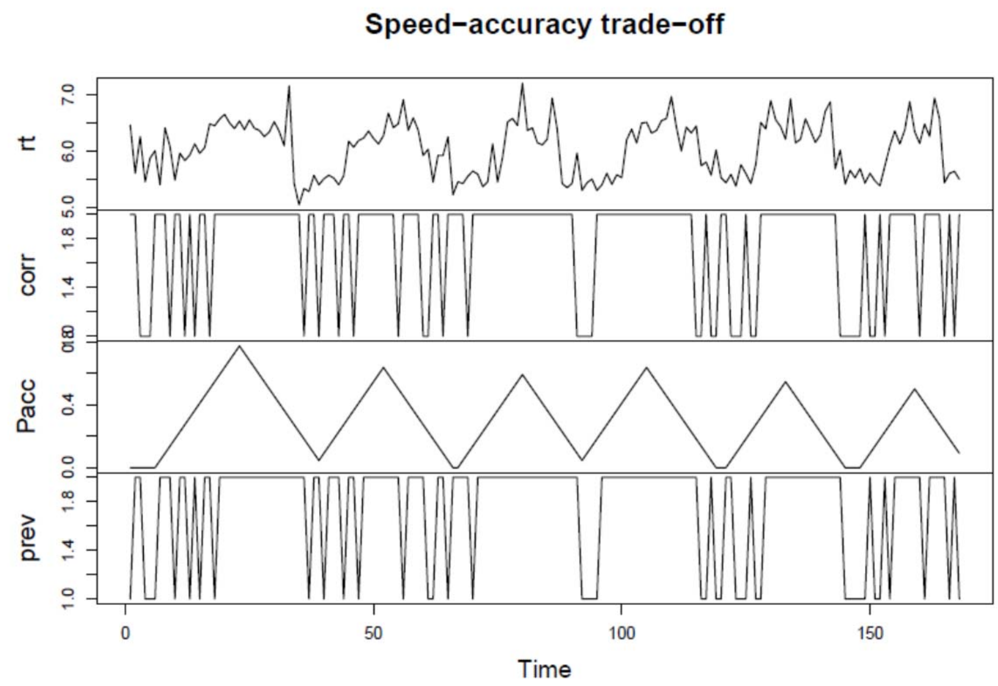
Dutilh, Gilles, et al. "A phase transition model for the speed-accuracy trade-off in response time experiments." *Cognitive Science* 35.2 (2011): 211-250.

QUT

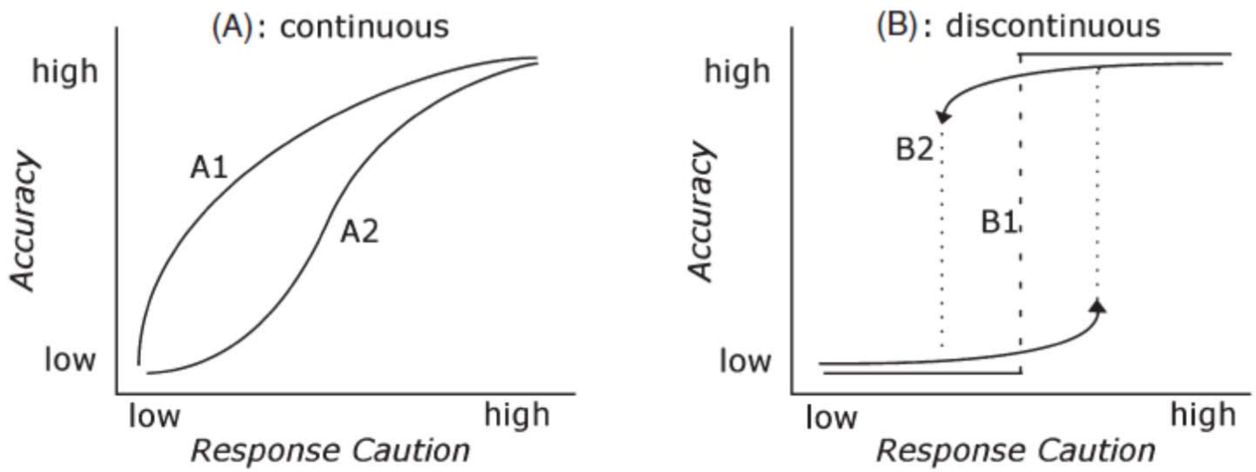
ACEMS

AUSTRALIAN RESEARCH COUNCIL CENTRE OF EXCELLENCE FOR MATHEMATICAL AND STATISTICAL FRONTIERS

Prac: Response Time



Prac: Response Time



Prac: Load Data

- # load the library and data
- install.packages('depmixS4')
- library(depmixS4)
- data(speed)
- # View the data
- View(speed)

Prac: Visualise the Data


```
library(ggplot2)
library(reshape2)
speed = cbind(speed,list(t=1:nrow(speed)))
mspeed = melt(speed,measure.vars=c('rt','corr','Pacc'))
ggplot(mspeed,aes(x=t,y=value,group=variable)) +
  geom_line() +
  facet_wrap(~variable,scales='free',ncol=1)
```

Prac: Fit a HMM

> modsimple = depmix(response=rt~1,data=speed,nstates=2,trstart=runif(4))

> fmsimple = fit(modsimple)

iteration 0 logLik: -305.3318

iteration 5 logLik: -305.3198

iteration 10 logLik: -305.3066

iteration 15 logLik: -305.2831

iteration 40 logLik: -88.92333

iteration 45 logLik: -88.71571

iteration 50 logLik: -88.71502

converged at iteration 52 with logLik: -88.71502

Prac: Fit a HMM


```
> summary(fmsimple)
Initial state probabilties model
pr1  pr2
  1    0

Transition matrix
        toS1   toS2
fromS1 0.916 0.084
fromS2 0.117 0.883
```

```
Response parameters
Resp 1 : gaussian
    Re1.(Intercept) Re1.sd
St1           6.385  0.244
St2           5.510  0.192
```

QUT

ACEMS
MATHEMATICAL AND STATISTICAL FRONTIERS

Prac: Results

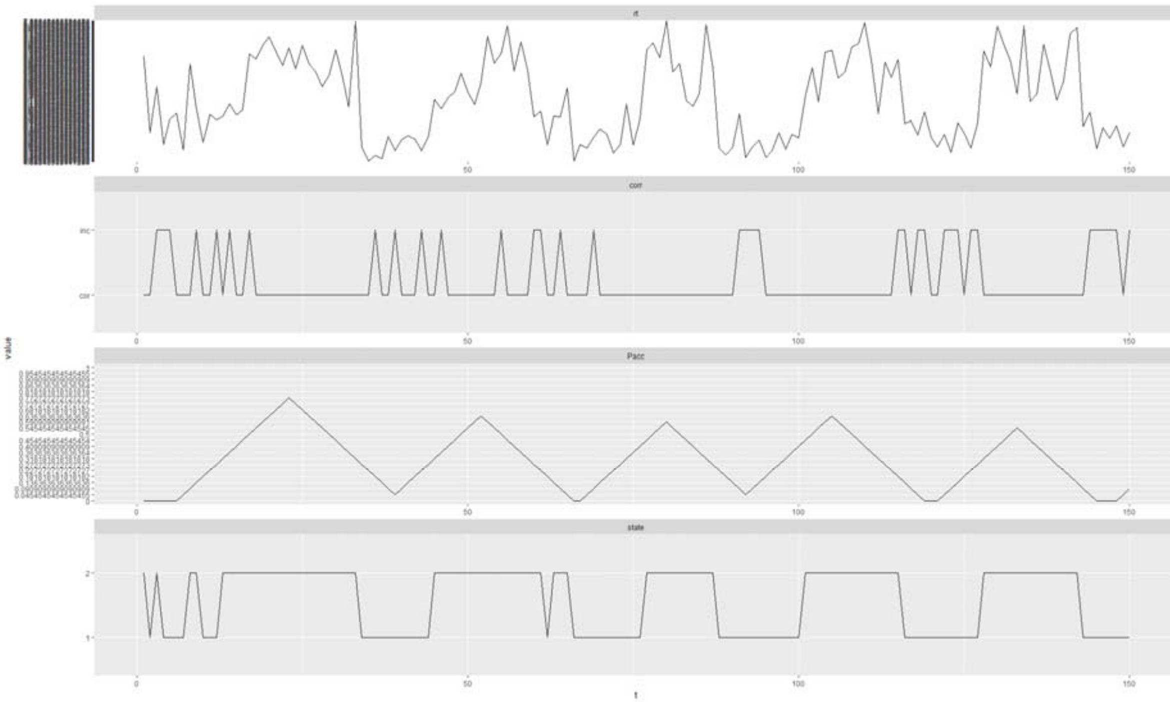

```
> fmsimple
Convergence info: Log likelihood converged to within tol. (relative change)
'log Lik.' -88.71502 (df=7)
AIC:  191.43
BIC:  220.0215
```

```
> head(fmsimple@posterior)
  state          S1         S2
1     1 1.000000000 0.00000000
2     2 0.052835996 0.94716400
3     1 0.993992010 0.00600799
4     2 0.005894012 0.99410599
5     2 0.063159278 0.93684072
6     2 0.457223228 0.54277677
```

Prac: Questions

- Q1: Explain your results
- Q2: What did we learn from this?

Prac: Questions



Extension: Fit with Covariates


```
modtrans =
depmix(response=rt~1,transition=~Pacc,data=speed,nsta
tes=2,instart=runif(2))
fmtrans = fit(modtrans)
```

- Explain your results?
- What is different?
- What additional information can we learn?

Extension : Fit with Covariates


```
> summary(fmtrans)
Initial state probabilties model
pr1 pr2
  0   1
Transition model for state (component) 1
Model of type multinomial (mlogit), formula: ~Pacc
Coefficients:
            St1       St2
(Intercept)   0 -3.128895
Pacc          0  5.971186
Probalities at zero values of the covariates.
0.958069 0.04193098
Transition model for state (component) 2
Model of type multinomial (mlogit), formula: ~Pacc
Coefficients:
            St1        St2
(Intercept)   0  -3.126696
Pacc          0  15.354371
Probalities at zero values of the covariates.
0.9579806 0.04201939

Response parameters
Resp 1 : gaussian
    Re1.(Intercept) Re1.sd
St1           5.507  0.187
St2           6.386  0.242
```

Extension : Fit with Covariates

> fmtrans

Convergence info: Log likelihood converged to within tol. (relative change)

'log Lik.' -44.19949 (df=9)

AIC: 106.399

BIC: 143.1595

Extension2: Multivariate HMM


```
modmulti = depmix(list(rt~1,corr~1),data=speed,nstates=2,
family=list(gaussian(),multinomial('identity')),transition=~scale(Pa
cc), instart=runif(2))

fmmulti = fit(modmulti)
```

Extension2: Multivariate HMM

Initial state probabilties model

pr1 pr2

0 1

Transition model for state (component) 1

Model of type multinomial (mlogit), formula: ~scale(Pacc)

Coefficients:

	St1	St2
(Intercept)	0	-0.9265658
scale(Pacc)	0	1.5984641

Probalities at zero values of the covariates.

0.716378 0.283622

Transition model for state (component) 2

Model of type multinomial (mlogit), formula: ~scale(Pacc)

Coefficients:

	St1	St2
(Intercept)	0	2.401846
scale(Pacc)	0	3.722318

Probalities at zero values of the covariates.

0.08303203 0.916968

Response parameters

Resp 1 : gaussian

Resp 2 : multinomial

	Re1.(Intercept)	Re1.sd	Re2.inc	Re2.cor
St1	5.517	0.197	0.475	**0.525**
St2	6.391	0.239	0.098	**0.902**

QUT

ACEMS

MATHEMATICAL AND STATISTICAL FRONTIERS

Extension2: Multivariate HMM

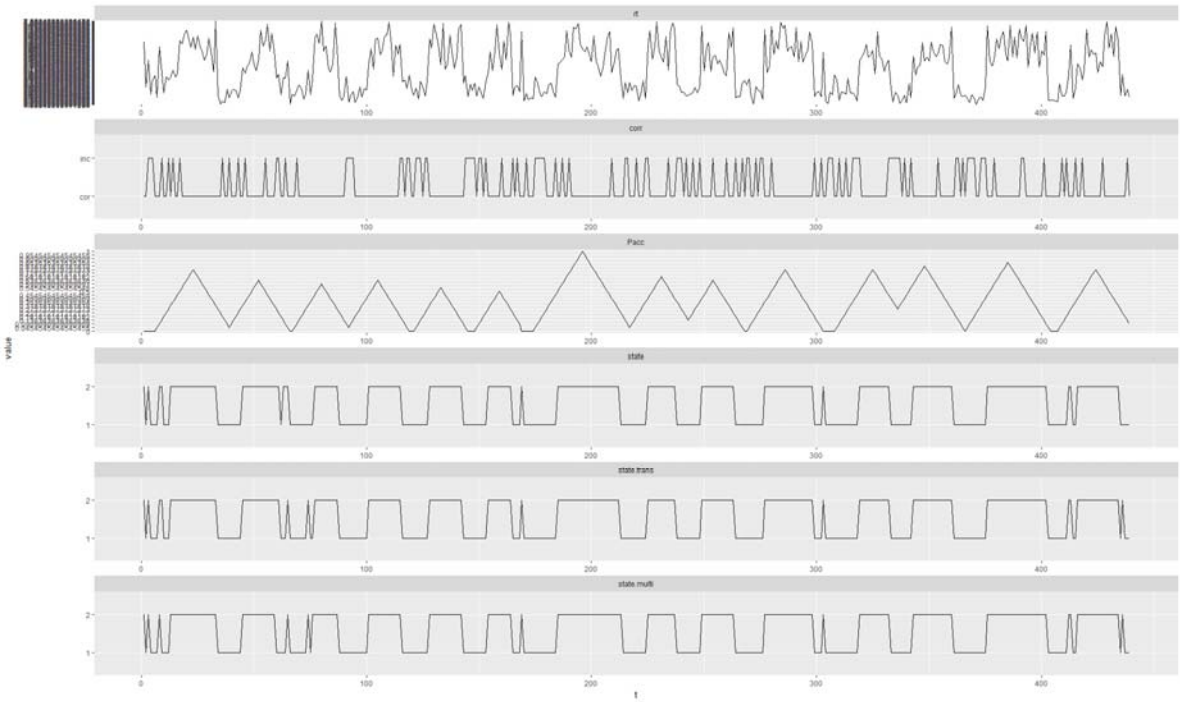
Convergence info: Log likelihood converged to within tol. (relative change)

'log Lik.' -255.5337 (df=11)

AIC: 533.0674

BIC: 577.9969

Extension2: Multivariate HMM



Questions

p.wu@qut.edu.au

Paul Wu

No moles were harmed in the making of this presentation

References

- Dadashi, F., A. Arami, F. Crettenand, G. P. Millet, J. Komar, L. Seifert and K. Aminian (2013). A Hidden Markov Model of the breaststroke swimming temporal phases using wearable inertial measurement units. 2013 IEEE International Conference on Body Sensor Networks.
- Franke, Alastair, Terry Caelli, and Robert J. Hudson. "Analysis of movements and behavior of caribou (Rangifer tarandus) using hidden Markov models." *Ecological Modelling* 173.2-3 (2004): 259-270.

QUT ACEMS AUSTRALIAN RESEARCH COUNCIL CENTRE OF EXCELLENCE FOR MATHEMATICAL AND STATISTICAL FRONTIERS

References

- Zucchini, Walter, Iain L. MacDonald, and Roland Langrock. *Hidden Markov models for time series: an introduction using R*. CRC press, 2016.
- Roohi, N. "HMM: Viterbi Algorithm – a toy example", University of Pennsylvania. http://www.cis.upenn.edu/~cis262/notes/Example-Viterbi-DNA.pdf.
- Rabusseau, G. and Islam, R. "Lecture 9: Hidden Markov Models." McGill University. https://rllabmcgill.github.io/COMP-652/lectures.html.
- Frazzoli, E. "Lecture 20: Intro to Hidden Markov Models.". MIT. https://ocw.mit.edu/courses/aeronautics-and-astronautics/16-410-principles-of-autonomy-and-decision-making-fall-2010/lecture-notes/MIT16_410F10_lec20.pdf.

QUT

ACEMS
MATHEMATICAL AND STATISTICAL FRONTIERS

- Mannini, A. and A. M. Sabatini (2012). "Gait phase detection and discrimination between walking–jogging activities using hidden Markov models applied to foot motion data from a gyroscope." Gait & Posture **36**(4): 657-661.

- Langrock, Roland, et al. "Flexible and practical modeling of animal telemetry data: hidden Markov models and extensions." *Ecology* 93.11 (2012): 2336-2342.

QUT

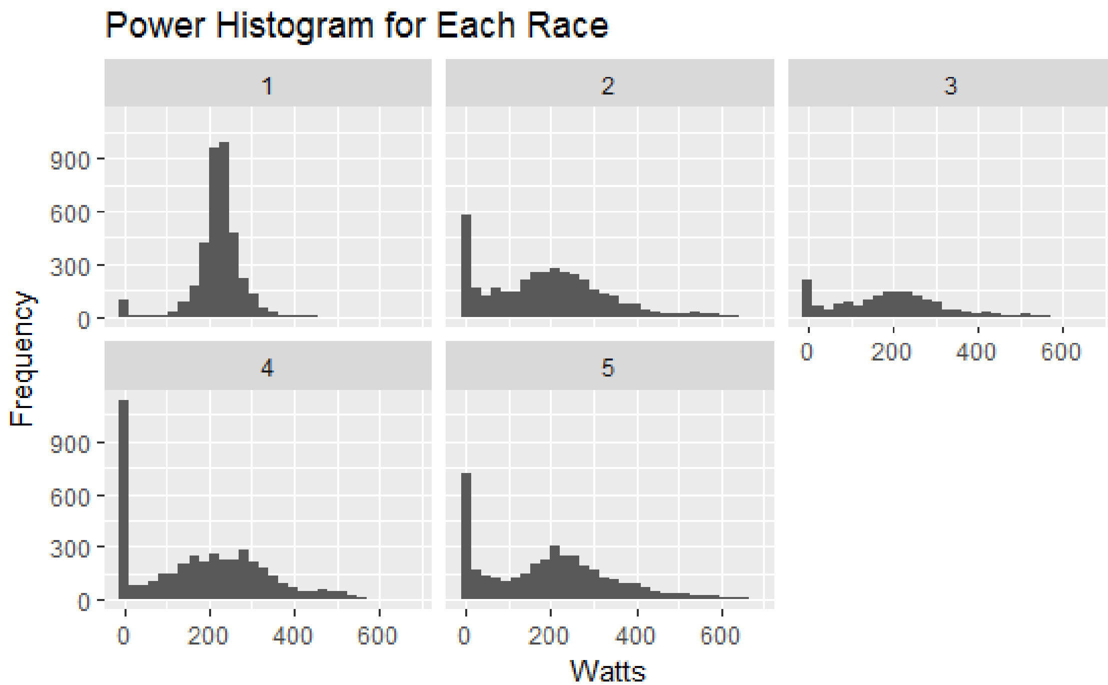
ACEMS

AUSTRALIAN RESEARCH COUNCIL CENTRE OF EXCELLENCE FOR MATHEMATICAL AND STATISTICAL FRONTIERS

Prac Case Study: Cycling Power

- Goal: understand power variability and physiological states in cycling
- Data:
  - GPS: latitude, longitude, elevation
  - Power
  - Speed
  - Cadence

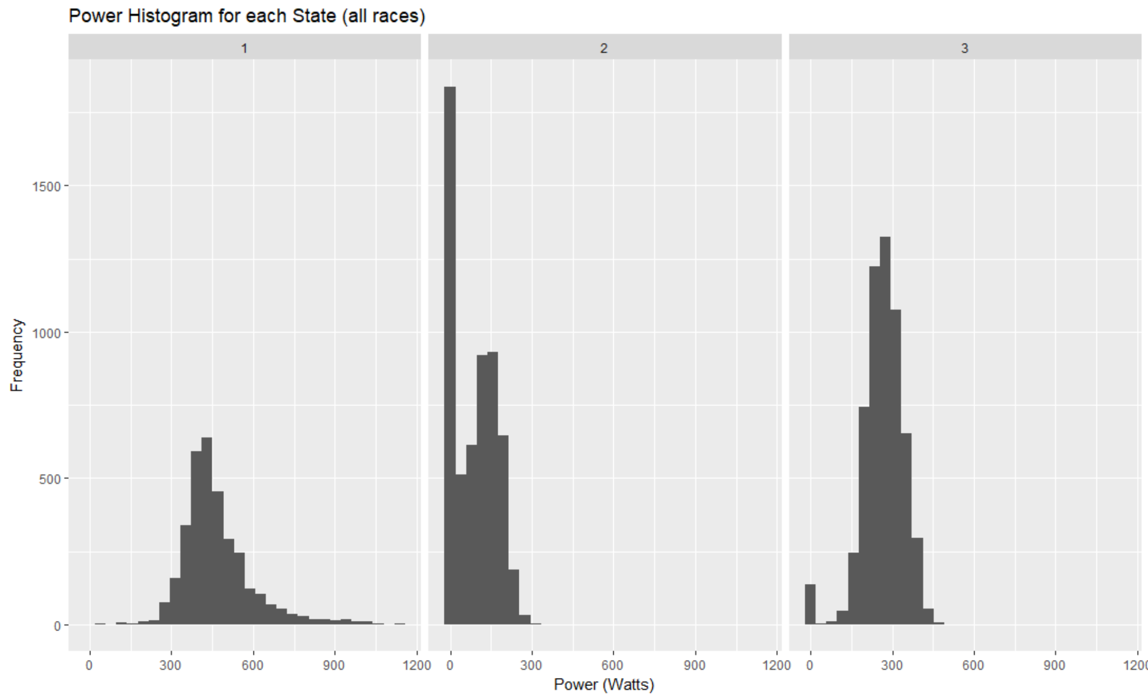
Prac Case Study: Power



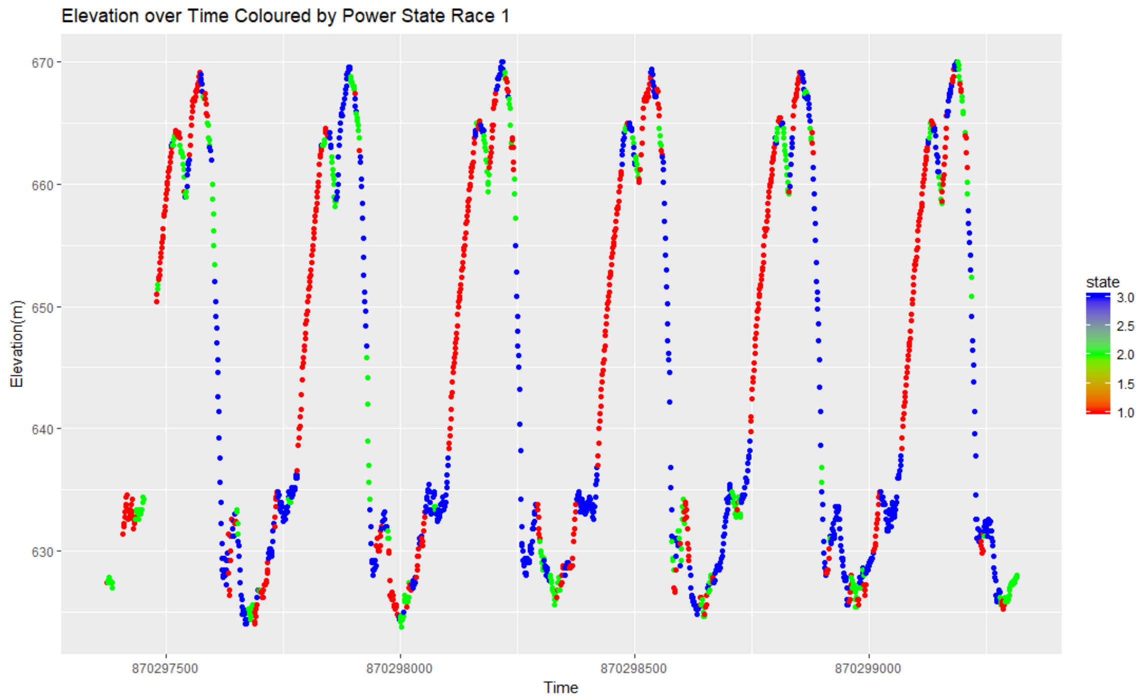
Prac Case Study: Model

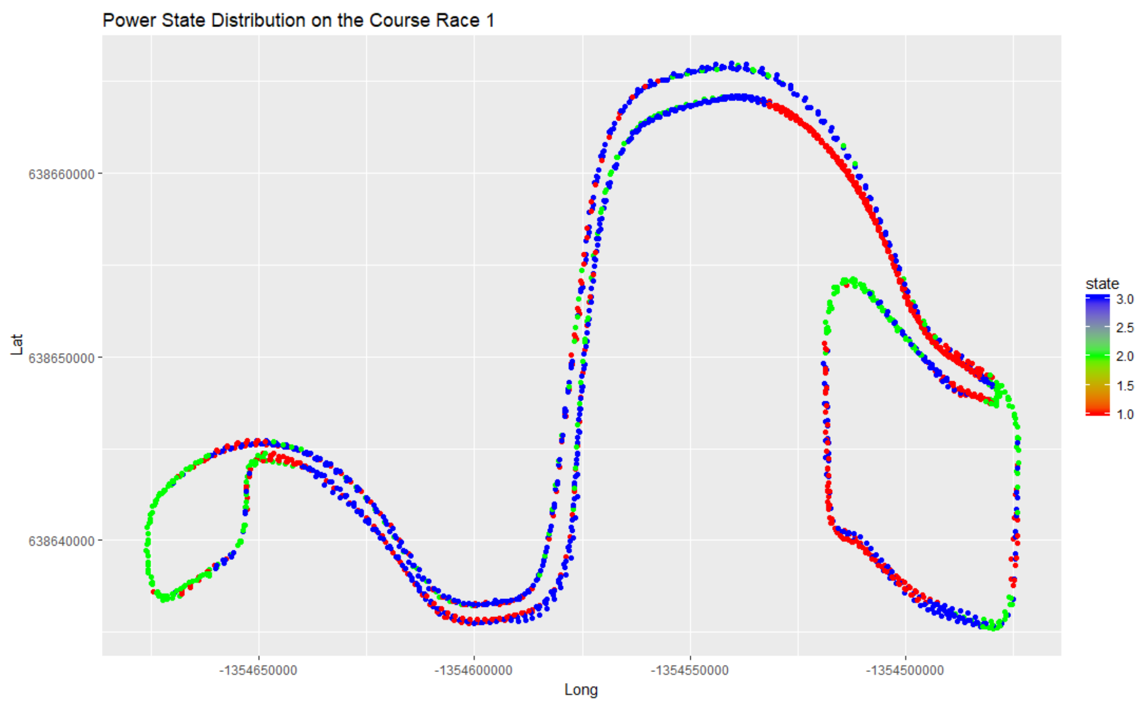

```
mod <- depmix(response = Power ~ 1 + Speed
+ Cadence + Elevation, data = R, nstates =
3)

fm <- fit(mod)
```

Prac Case Study: Results



Prac Case Study: Results




Power State Distribution on the Course Race 1
Lat
638660000
638650000
638640000
-1354650000
-1354600000
-1354550000
-1354500000
Long
state
3.0
2.5
2.0
1.5
1.0