

# Model complexity and model choice for animal movement models

Ben Bolker, McMaster University

Departments of Mathematics & Statistics and Biology

Guelph Biomathematics & Biostatistics Symposium

9 June 2016

# Outline

- 1 Animal movement
- 2 Florida panthers
- 3 Hidden Markov models
- 4 Basic analysis (van de Kerk et al., 2015)
- 5 Incorporating diurnal variation (Li, 2015)
- 6 Broader issues/outlook

# Acknowledgements

**People** Michael Li, Madelon van de Kerk,  
Dave Onorato, Madan Oli

**Agencies** US Fish and Wildlife Service, US Geological Survey,  
US National Park Service

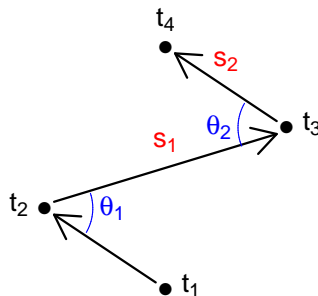
**Funding** NSERC Discovery grant, NSF IGERT program

# Outline

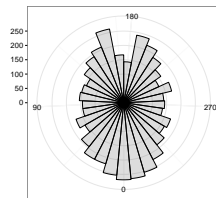
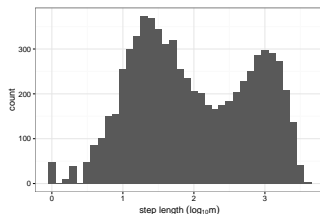
- 1 Animal movement
- 2 Florida panthers
- 3 Hidden Markov models
- 4 Basic analysis (van de Kerk et al., 2015)
- 5 Incorporating diurnal variation (Li, 2015)
- 6 Broader issues/outlook

# Animal movement: data

- observations:  
e.g. mass mark-recapture,  
longitudinal density, direct  
observation, telemetry  
(VHF, GPS)
- most methods provide a  
sequence of times and  
locations for each individual



- summaries:
  - home range  
(convex hull, kernel density estimate, etc.)
  - root-mean-squared displacement
  - step length and turning angle
- covariates:
  - e.g. habitat map,
  - individual characteristics  
(sex, age, weight ...)



# Animal movement: questions

- simple description
- how do animals' movements change as a function of their (internal or external) environment?  
what does that tell us about their biology?
- how might animals' distributions, etc. change when conditions (density, habitat, ...) change?

# Outline

- 1 Animal movement
- 2 Florida panthers**
- 3 Hidden Markov models
- 4 Basic analysis (van de Kerk et al., 2015)
- 5 Incorporating diurnal variation (Li, 2015)
- 6 Broader issues/outlook



# Biological/conservation issues

- Florida panther: *Puma concolor coryi*
- endangered subspecies
- severely reduced habitat
- small, isolated population
- currently recovering



[www.peer.org](http://www.peer.org)



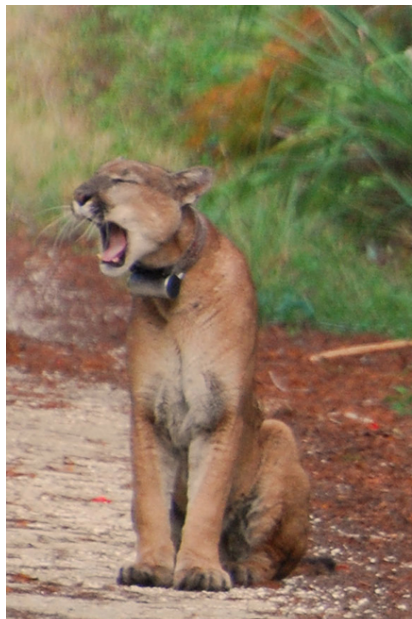
# Panther movement questions

- movement variation by sex and life history stage (juvenile, adult, mom with kittens . . . )
- effects of movement on threats (intraspecific aggression, roadkill) ?
- predicting the effects of future changes in population density / population structure / habitat

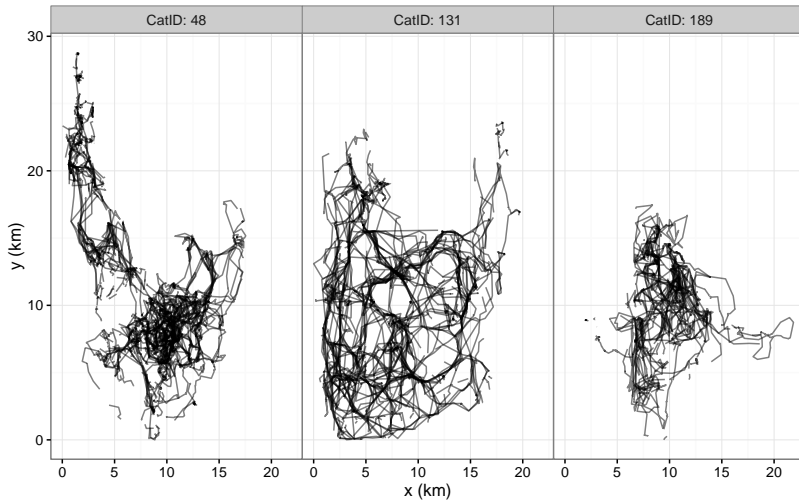


## Panther movement data

- panthers tracked, captured
- GPS collars
- 18 males (13 male, 5 female, 1-15 years old)
- 3200 panther days, hourly/bihourly; 49000 locations
- ?? per panther



# example movement tracks

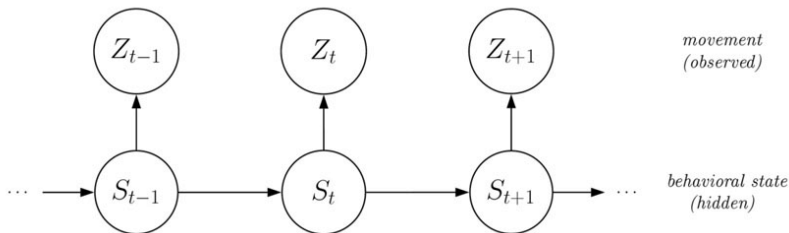


# Outline

- 1 Animal movement
- 2 Florida panthers
- 3 Hidden Markov models**
- 4 Basic analysis (van de Kerk et al., 2015)
- 5 Incorporating diurnal variation (Li, 2015)
- 6 Broader issues/outlook

# Hidden Markov models

- finite mixture model with temporal dependence
- discrete time steps
- discrete latent state; *transition matrix*
- observations from *emission distributions*  
(continuous or discrete, univariate or multivariate)
- **multiphasic movement** (Fryxell et al., 2008; Langrock et al., 2012)



# Hidden Markov models (cont.)

**state:**

$$S_t \sim \text{Multinomial}(S_{t-1}, \mu_{S,t})$$
$$\mu_{S,t} = \text{multi-logistic}(\mathbf{X}_{S,t} \boldsymbol{\beta}_S)$$

**emission:**

$$\mathbf{Z}_t \sim \{\text{Dist}_1(\mu_{Z_1,S_t}), \dots, \text{Dist}_n(\mu_{Z_n,S_t})\}$$
$$\mu_{Z_i,S_t} = g^{-1}(\mathbf{X}_{Z_i,t} \boldsymbol{\beta}_{Z_i,S_t})$$

# Hidden Markov models (part 3)

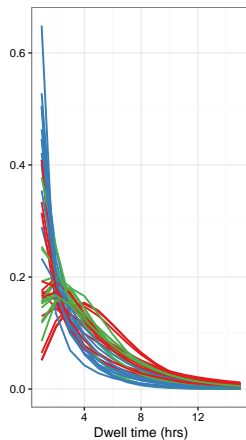
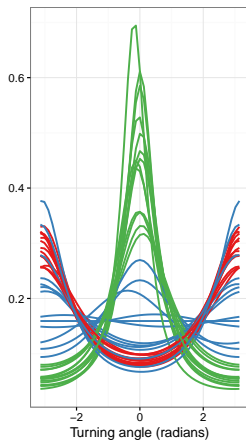
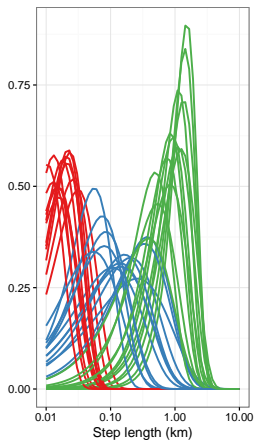
- *forward-backward algorithm* for estimating parameters
- *Viterbi algorithm* for estimating most probable state sequences
- depmixS4 package (Visser and Speekenbrink, 2010) (also moveHMM (Michelot et al., 2016))
- hidden *semi-Markov* models: allow for non-geometric *dwell distributions* (Langrock, 2011; Augustine, 2016): move.HMM



# Outline

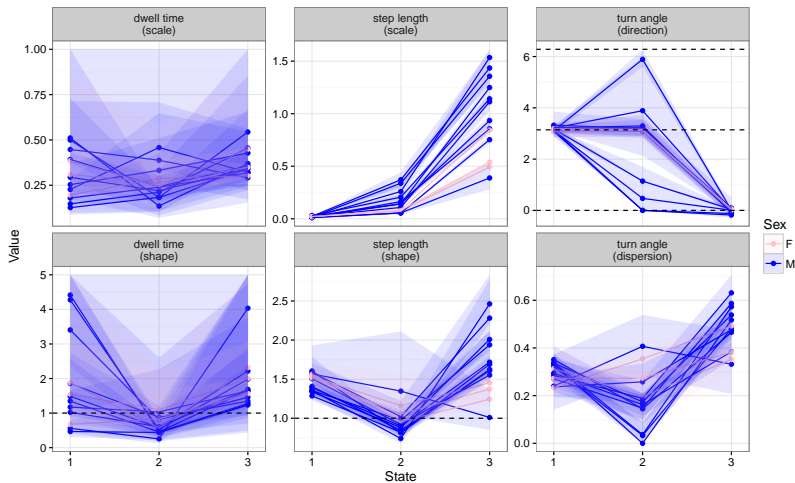
- 1 Animal movement
- 2 Florida panthers
- 3 Hidden Markov models
- 4 Basic analysis (van de Kerk et al., 2015)**
- 5 Incorporating diurnal variation (Li, 2015)
- 6 Broader issues/outlook

# State distributions

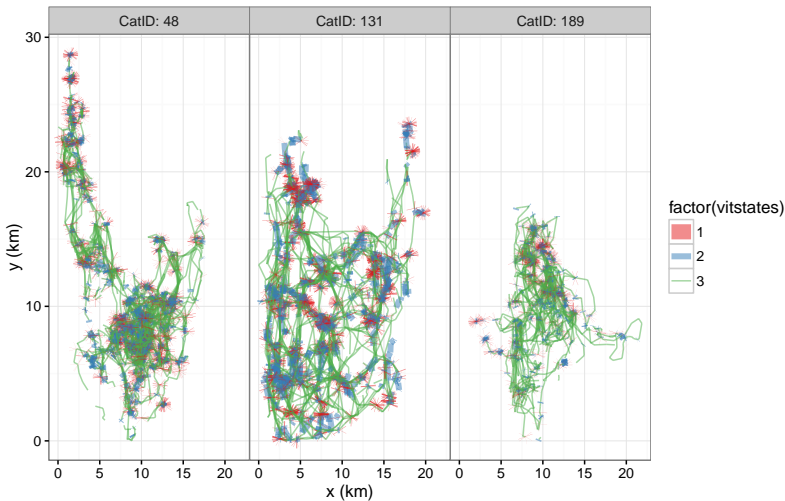


State — 1 — 2 — 3

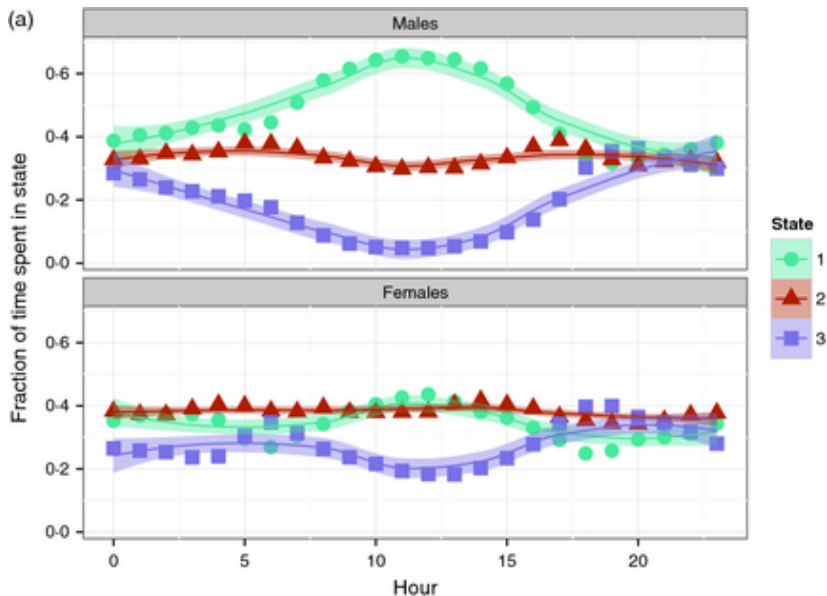
# Parameter estimates



# Tracks with Viterbi estimates



# Diurnal variation



# what can we conclude so far?

## good news

- basic biology: males move faster, farther
- three states are identifiable, sensible
- dwell distributions approximately geometric (HSMM  $\rightarrow$  HMM)

## bad news

- diurnal variation in Viterbi results - but it's not in the model!
- estimates of model complexity are too high

# what can we conclude so far?

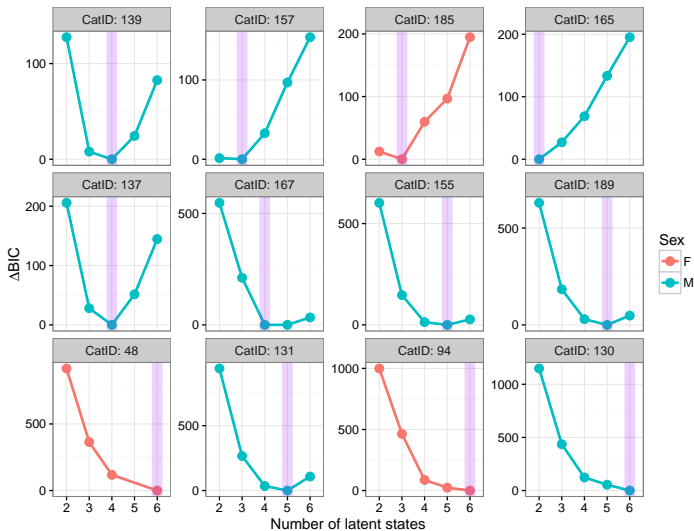
## good news

- basic biology: males move faster, farther
- three states are identifiable, sensible
- dwell distributions approximately geometric (HSMM  $\rightarrow$  HMM)

## bad news

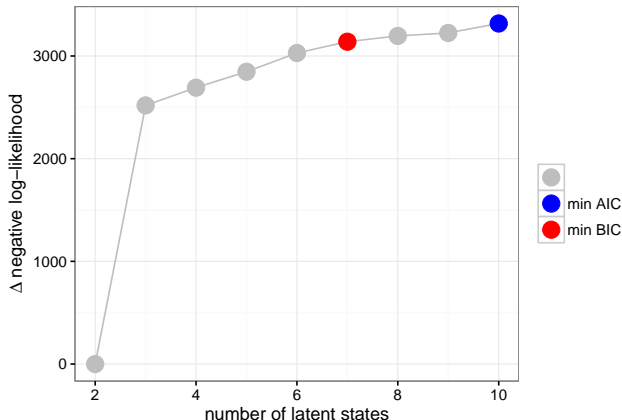
- diurnal variation in Viterbi results - but it's not in the model!
- estimates of model complexity are too high

# Model complexity (bad news)





# Model complexity (Manx shearwaters, Dean et al. (2013))



# Outline

- 1 Animal movement
- 2 Florida panthers
- 3 Hidden Markov models
- 4 Basic analysis (van de Kerk et al., 2015)
- 5 Incorporating diurnal variation (Li, 2015)**
- 6 Broader issues/outlook

# Expanding the model

Attempting to fix these problems:

- extend the model to allow covariates
- specifically, allow for diurnal variation
  - simplify model (log-Normal step length only)
  - *fixed* state-specific emissions parameters (step length mean and std dev)
  - time-varying transition parameters
  - also try *finite mixture models* (independent occupancy)
- how much does this help?

# Expanding the model

Attempting to fix these problems:

- extend the model to allow covariates
- specifically, allow for diurnal variation
  - simplify model (log-Normal step length only)
  - *fixed* state-specific emissions parameters (step length mean and std dev)
  - time-varying transition parameters
  - also try *finite mixture models* (independent occupancy)
- how much does this help?

Animal movement  
ooo

Panthers  
oooo

HMM  
ooo

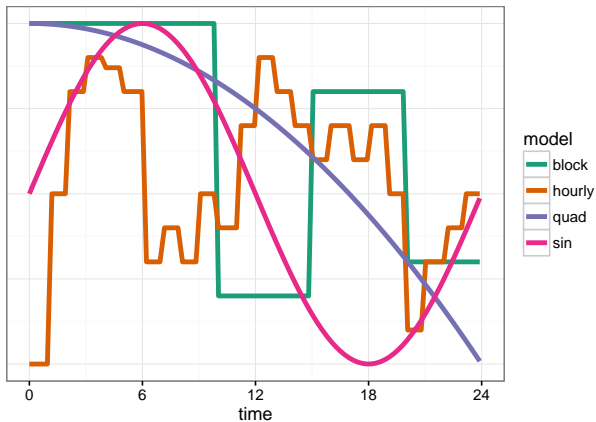
Basic analysis  
ooooooo

Diurnal model  
o●oooo

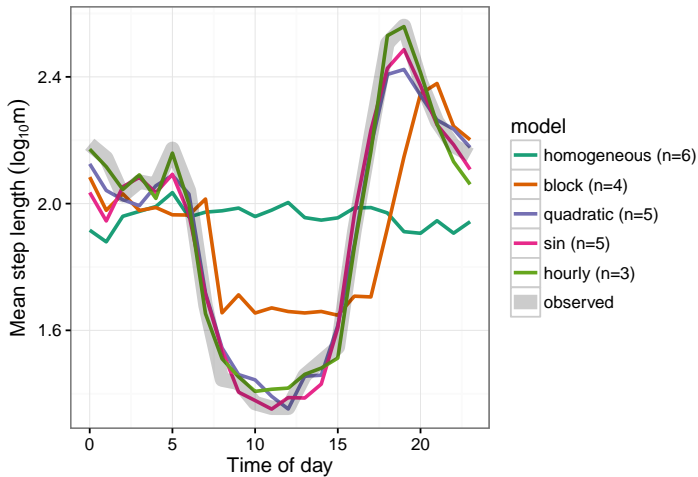
Broader issues/outlook  
oooo

References

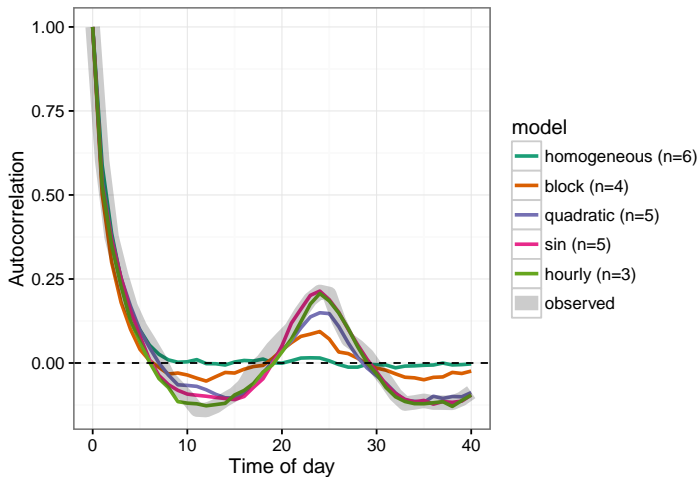
# Temporal models



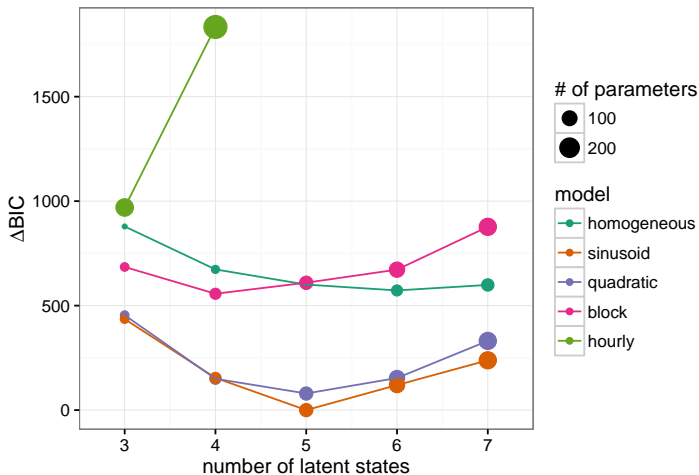
# Temporal patterns (step length)



# Temporal patterns (autocorrelation)



# Goodness of fit/model complexity





# Diurnal model: conclusions

- diurnal structure greatly improves fit ( $\Delta\text{BIC} \approx 500$ )
- slightly improves latent-state issue ( $n = 6 \rightarrow 5$ )
- lots left to do!
  - seasonal variation
  - incorporate habitat, home range behaviour
  - etc. etc. etc.

# Outline

- 1 Animal movement
- 2 Florida panthers
- 3 Hidden Markov models
- 4 Basic analysis (van de Kerk et al., 2015)
- 5 Incorporating diurnal variation (Li, 2015)
- 6 Broader issues/outlook**

# Big data and small models

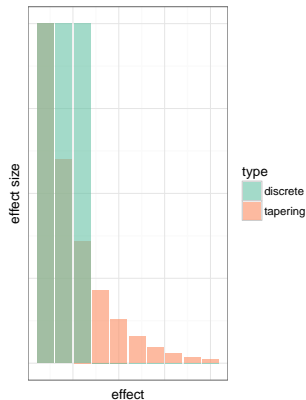
- simple model families +  
model misspecification →  
overparameterization

- Gelman: “Sample sizes are never large”: ([blog post](#))

*N is never enough because if it were “enough” you’d already be on to the next problem for which you need more data.*

# An aside on AIC vs BIC

- “should I use AIC or BIC? I heard that AIC is inconsistent ...”
- complexity penalty = 2 (AIC) vs  $\log(n)$  (BIC)
- best prediction vs. model identification (Yang, 2005)
- *effect size spectrum*: tapering or discrete?



# Animal movement: open challenges

- Cognition/memory (Bracis et al., 2015)
- Intraspecific interaction/collective movement (Delgado et al., 2014)
- Continuous-time movement models (Calabrese et al., 2016)
- Edges, barriers, and corridors (Beyer et al., 2016)
- Efficient (big-data) approaches (Brillinger et al., 2008)
- Putting it all together ...

# Tools needed

- cross-validation (Wenger and Olden, 2012)
- protocols and tools for model checking (Potts et al., 2014);  
score tests?
- flexible computational frameworks  
(ecologists can't afford consultants/  
there are too many species out there)

## References

- Augustine, B., 2016. Flexible, user-friendly hidden (semi) Markov models for animal movement data.
- Beyer, H.L., Gurarie, E., et al., 2016. *Journal of Animal Ecology*, 85(1):43–53. ISSN 00218790. doi:10.1111/1365-2656.12275.
- Bracis, C., Gurarie, E., et al., 2015. *PLOS ONE*, 10(8):e0136057. ISSN 1932-6203. doi:10.1371/journal.pone.0136057.
- Brillinger, D.R., Stewart, B.S., et al., 2008. In *Probability and statistics: Essays in honor of David A. Freedman*, pages 246–264. Institute of Mathematical Statistics.
- Calabrese, J.M., Fleming, C.H., and Gurarie, E., 2016. *Methods in Ecology and Evolution*, pages n/a–n/a. ISSN 2041-210X. doi:10.1111/2041-210X.12559.
- Dean, B., Freeman, R., et al., 2013. *Journal of The Royal Society Interface*, 10(78):20120570. ISSN 1742-5689, 1742-5662. doi:10.1098/rsif.2012.0570.
- Delgado, M.d.M., Penteriani, V., et al., 2014. *Methods in Ecology and Evolution*, 5(2):183–189.
- Fryxell, J.M., Hazell, M., et al., 2008. *Proceedings of the National Academy of Sciences*, 105(49):19114–19119. ISSN 0027-8424, 1091-6490. doi:10.1073/pnas.0801737105.
- Langrock, R., 2011. *Computational Statistics and Data Analysis*, 55(1):715–724. ISSN 01679473.
- Langrock, R., King, R., et al., 2012. *Ecology*, 93(11):2336–2342. ISSN 0012-9658. doi:10.1890/11-2241.1.
- Li, M., 2015. *Incorporating Temporal Heterogeneity in Hidden Markov Models For Animal Movement*. Master's thesis.
- Michelot, T., Langrock, R., and Patterson, T.A., 2016. *Methods in Ecology and Evolution*, in press. doi:10.1111/2041-210X.12578.
- Potts, J.R., Auger-MÃ©thÃ©, M., et al., 2014. *Methods in Ecology and Evolution*, 5(10):1012–1022.
- van de Kerk, M., Onorato, D.P., et al., 2015. *Journal of Animal Ecology*, 84(2):576–585.
- Visser, I. and Spekenbrink, M., 2010. *Journal of Statistical Software*, 36(7):1–21.
- Wenger, S.J. and Olden, J.D., 2012. *Methods in Ecology and Evolution*, 3(2):260–267. ISSN 2041210X. doi:10.1111/j.2041-210X.2011.00170.x.
- Yang, Y., 2005. *Biometrika*, 92(4):937–950. doi:10.1093/biomet/92.4.937.