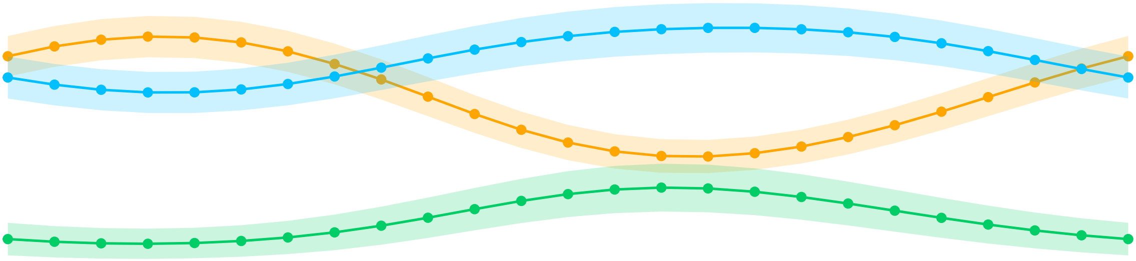
Young Researchers Workshop

Inference on the state process of periodically inhomogeneous hidden Markov models

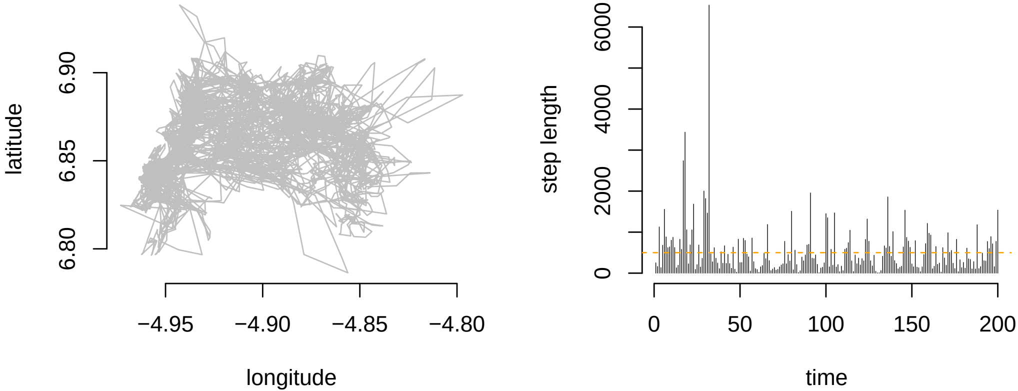
Jan-Ole Koslik

February 16, 2024



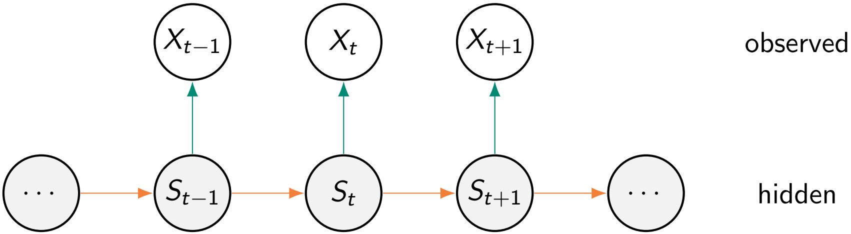
Motivating example

Motivating example



Quick Recap on HMMs

Doubly stochastic process:



- every observation is generated by one of N possible distributions $f_1, \dots, f_N$,
- the state process selects which distribution is active at any given time point

More formal definition of an HMM

An N-state HMM is a (doubly) stochastic process in discrete time, with

- an unobserved **state process** $S_1, S_2, \ldots, S_T$ taking values in $\{1, \ldots, N\}$,
- and an **observation process** $X_1, X_2, \ldots, X_T$,

such that

- $f(s_t \mid s_1, \ldots, s_{t-1}) = f(s_t \mid s_{t-1})$
  (Markov property)
- $f(x_t \mid s_1, \ldots, s_T, x_1, \ldots, x_{t-1}, x_{t+1}, \ldots, x_T) = f(x_t \mid s_t)$
  (conditional independence assumption)

Reminder: Markov chains

The **Markovian state process** is fully characterised by the initial distribution

$$\boldsymbol{\delta}^{(1)} = \big(\Pr(S_1 = 1), \ldots, \Pr(S_1 = N)\big)$$

and the **transition probabilities**

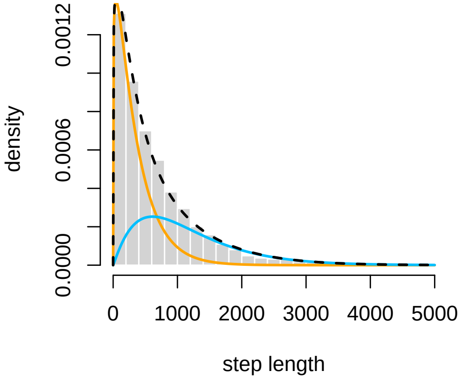
$$\gamma_{ij}^{(t)} = \Pr(S_{t+1} = j \mid S_t = i).$$

which we summarise in the **transition probability matrix** (t.p.m.)

$$\boldsymbol{\Gamma}^{(t)} = (\gamma_{ij}^{(t)})_{i,j=1,\ldots,N}.$$

Example continued

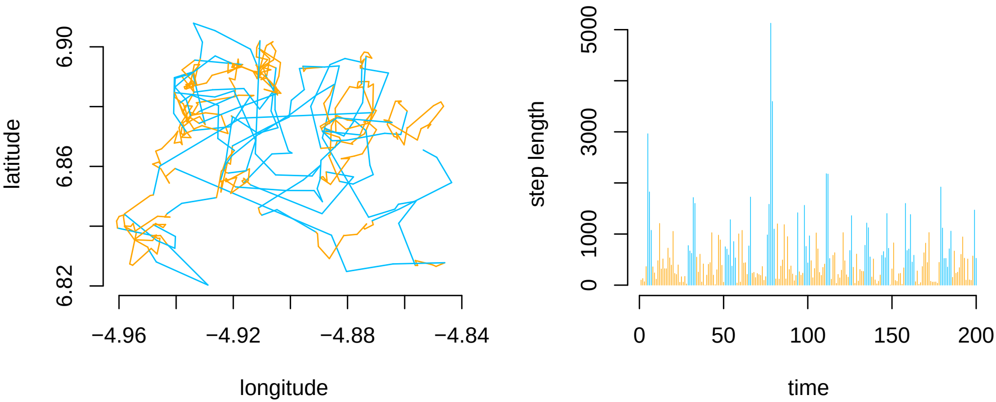
Fitting a basic 2-state HMM to the elephant data yields the following results:



$$\hat{\boldsymbol{\Gamma}} = \begin{pmatrix} 0.81 & 0.19 \\ 0.29 & 0.71 \end{pmatrix}$$

$$\hat{\boldsymbol{\delta}} = (0.6, 0.4)$$

Example continued



Example continued (the problems)

… which is all very nice, **but**

Example continued (the problems)

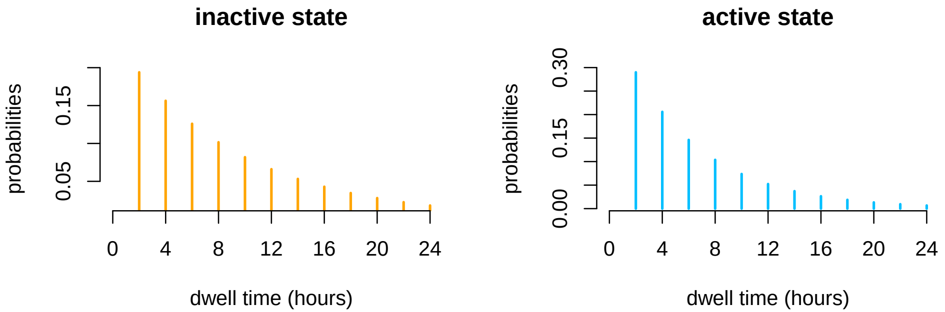
… which is all very nice, **but**

1. our model assumes the unconditional state probabilities are constant,

Example continued (the problems)

... which is all very nice, **but**

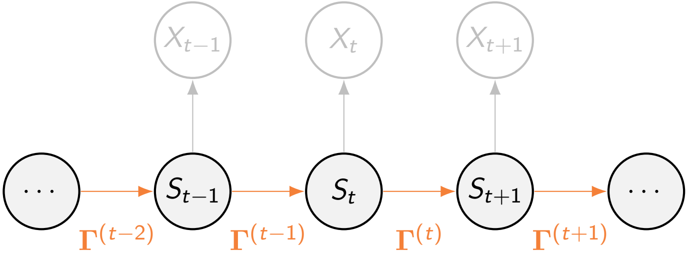
1. our model assumes the unconditional state probabilities are constant,
2. and the distribution of time spent in a state is geometric (often criticised) (e.g. for state 1: $\Pr(R_1 = r) = 0.81^{r-1} 0.19$)



Periodic variation

In the real world, there are many reasons for processes to involve a periodic component. For animal data, this could be

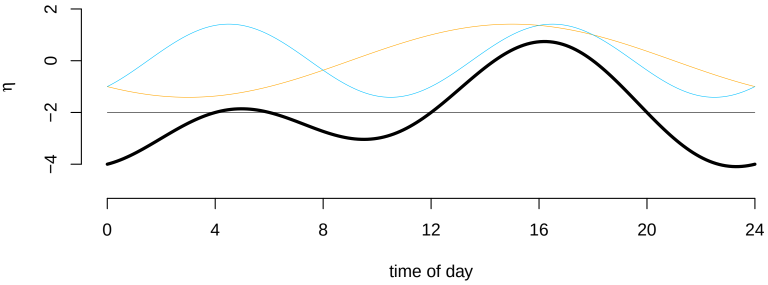
- diurnal/ time-of-day variation,
- seasonal variation,
- migratory behaviour.



How to model periodic variation

Typical parametric[1] model:

$$\gamma_{ij}^{(t)} = \text{mlogit}^{-1}\left(\beta_0^{(ij)} + \sum_{k=1}^{K} \beta_{1k}^{(ij)} \sin\left(\frac{2\pi kt}{L}\right) + \sum_{k=1}^{K} \beta_{2k}^{(ij)} \cos\left(\frac{2\pi kt}{L}\right)\right)$$



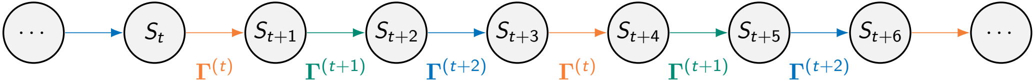
[1]We can also do this non-parametrically (Feldmann et al., 2023).

Periodic variation

For a cycle length of L, both options lead to

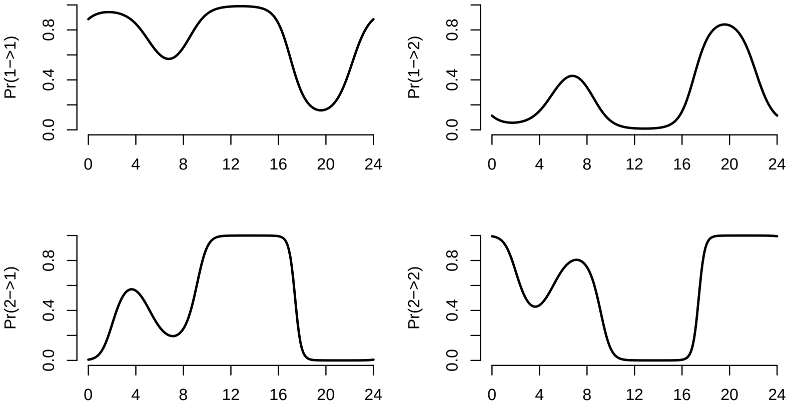
$$\boldsymbol{\Gamma}^{(t)} = \boldsymbol{\Gamma}^{(t+L)} \quad \text{for all } t = 1, \ldots, T, \tag{1}$$

which is what we mean by **periodic variation**.



Bad interpretability

$$\hat{\mathbf{\Gamma}} =$$



Periodic stationarity

- We would like to have $\Pr(S_t = i)$ as a function of the time of day.
- For periodic variation, this is typically approximated[2] by $\boldsymbol{\rho}^{(t)}$ which is the solution to

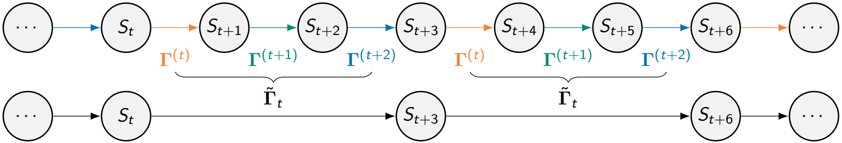
$$\boldsymbol{\rho}^{(t)}\boldsymbol{\Gamma}^{(t)} = \boldsymbol{\rho}^{(t)}, \quad \text{s.t.} \sum_{i=1}^{N} \rho_i^{(t)} = 1, \quad \text{for } t = 1, \ldots, L.$$

- But this estimate is biased, as it ignores the preceding process dynamics.

[2] e.g. Patterson et al., 2009

Periodic stationarity

- Consider for every $t \in \{1, \ldots, L\}$ the thinned Markov chain $S_t, S_{t+L}, S_{t+2L}, \ldots$.

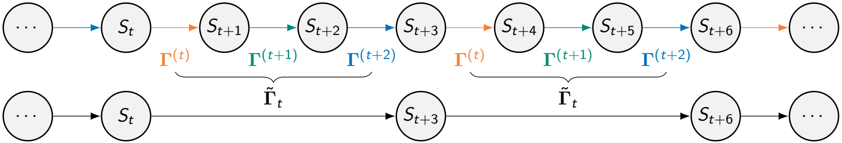


- It has constant t.p.m.

$$\tilde{\Gamma}_t = \Gamma^{(t)} \cdot \Gamma^{(t+1)} \cdot \ldots \cdot \Gamma^{(t+L-1)}.$$

Periodic stationarity

- Consider for every $t \in \{1, \ldots, L\}$ the thinned Markov chain $S_t, S_{t+L}, S_{t+2L}, \ldots$.



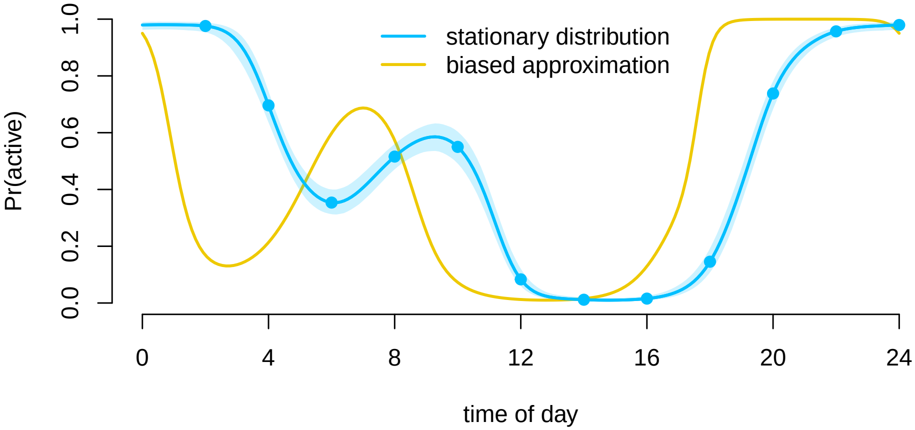
- It has constant t.p.m.

$$\tilde{\boldsymbol{\Gamma}}_t = \boldsymbol{\Gamma}^{(t)} \cdot \boldsymbol{\Gamma}^{(t+1)} \cdot \ldots \cdot \boldsymbol{\Gamma}^{(t+L-1)}.$$

- Thus each thinned chain converges, and we get $\Pr(S_t = i)$ for each t by solving

$$\boldsymbol{\delta}^{(t)} \tilde{\boldsymbol{\Gamma}}_t = \boldsymbol{\delta}^{(t)} \quad \text{s.t.} \sum_{i=1}^{N} \delta_i^{(t)} = 1.$$

Periodically stationary distribution (Elephant example)



Dwell-time distribution(s)

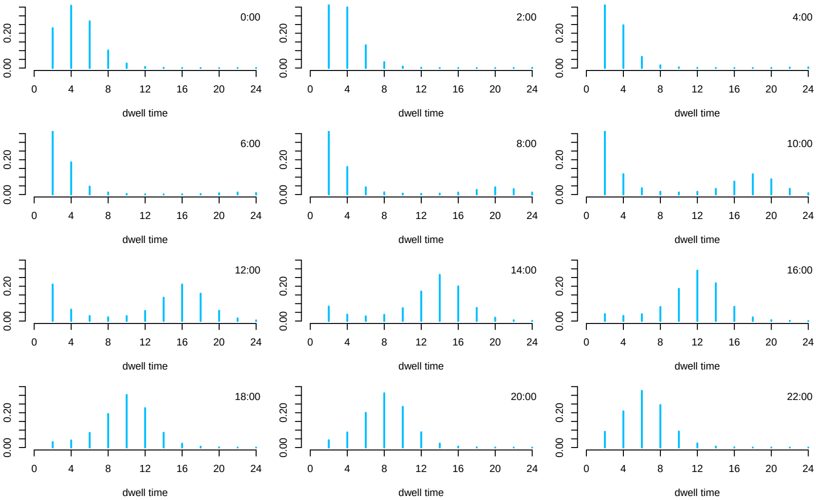
- We added periodic variation to our model, which is much more realistic.
- But what happened to the dwell-time distributions? Did it help?

Time-varying dwell-time distribution

First, consider the dwell-time distribution only at a certain time point t:

$$d_i^{(t)}(r) = \underbrace{\left(1 - \gamma_{ii}^{(t+r-1)}\right)}_{\text{leave}} \cdot \underbrace{\prod_{j=1}^{r-1} \gamma_{ii}^{(t+j-1)}}_{\text{stay r times}}, \qquad r \in \mathbb{N}$$

Time-varying dwell-time distribution (Elephant example)



Time-varying dwell-time distribution

- The **time-varying** dwell-time distributions allow for detailed inference regarding the state process.
- However, it might be cumbersome to interpret all of them (e.g. minute-by-minute data: $L = 1440$).
- Focus of the inference is often the **overall** distribution of time spent in each state...

Overall dwell-time distribution

For a periodically stationary Markov chain defined by $\mathbf{\Gamma}^{(t)}, t = 1, \ldots, L$, the p.m.f. of the dwell-time distribution in state i is

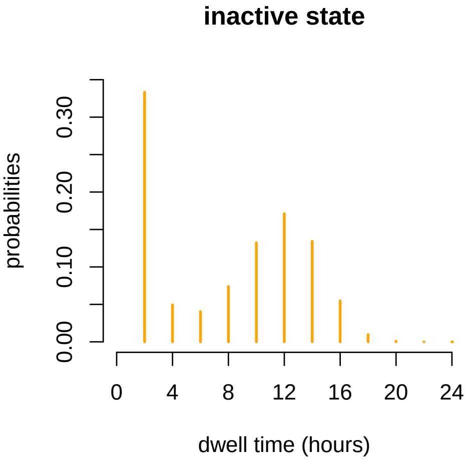
$$d_i(r) = \sum_{t=1}^{L} w_i^{(t)} d_i^{(t)}(r), \qquad r \in \mathbb{N}$$

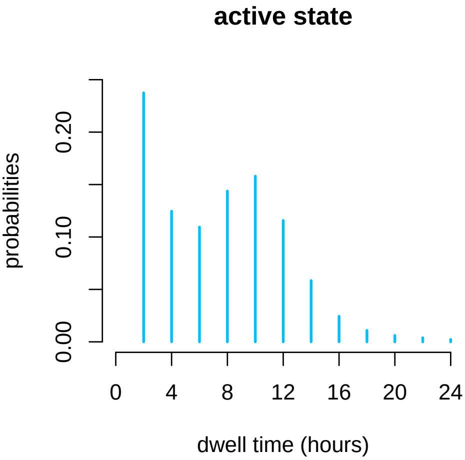
where with the mixture weights defined as

$$w_i^{(t)} = \frac{\sum_{l \neq i} \delta_l^{(t-1)} \gamma_{li}^{(t-1)}}{\sum_{t=1}^{L} \sum_{l \neq i} \delta_l^{(t-1)} \gamma_{li}^{(t-1)}}, \qquad t = 1, \ldots, L,$$

where $\boldsymbol{\delta}^{(t)}$ is the periodically stationary distribution.

Overall dwell-time distribution (Elephant example)





Conclusion and Outlook

- In HMM including periodic variation, dwell-time distributions can be very non-geometric, i.e.
- it alleviates biologically unrealistic consequences of the Markov assumption.
- Thus, criticising the Markov assumption for implying geometric dwell times falls short of such models' actual potential.

Literature

Feldmann, Carlina C et al. (2023). "Flexible modelling of diel and other periodic variation in hidden Markov models". In: *Journal of Statistical Theory and Practice*. in press.

Patterson, Toby A. et al. (2009). "Classifying movement behaviour in relation to environmental conditions using hidden Markov models". In: *Journal of Animal Ecology* 78.6, pp. 1113–1123.

Wikelski M Davidson SC, Kays R. (2024). *Movebank: archive, analysis and sharing of animal movement data. Hosted by the Max Planck Institute of Animal Behavior.* www.movebank.org (Accessed: 30.01.2024).

Thank you very much for your attention!