Hidden Markov models and sequential Monte Carlo Project :

Estimating behavioral parameters in animal movement models using a state-augmented particle filter

B. Maurel - G. Watkinson - J.Stym-Popper - G. Guaquière

ENSAE Paris

Dowd and Joy [1] proposed a methodology to model the behavior of seals in their natural habitat, using their vertical velocity.

Indeed, depth measurements vary depending on whether the seals are hunting, exploring, moving, and so on ...

They model the animal's behavior with state-space models and estimate the latent behavioral parameters with an algorithm.

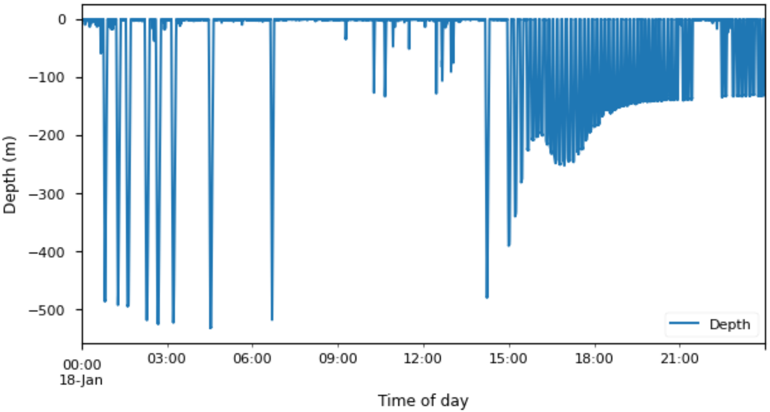

Figure: Evolution of the depth (in m) of the seal on January 18th, 2008, taken at 5 second intervals.

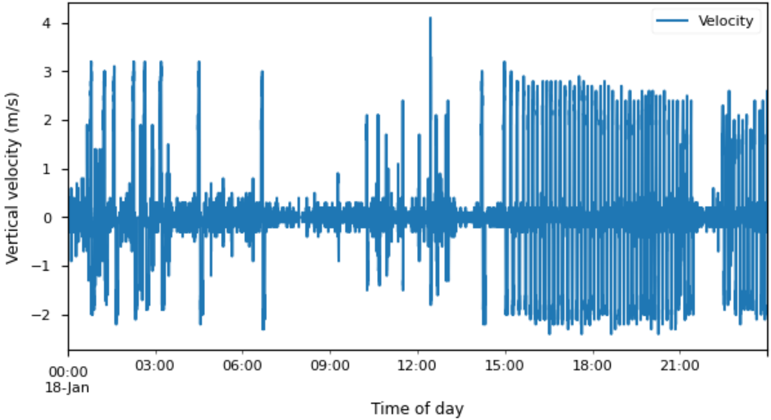

Figure: Vertical velocity (m/s) of a single seal on January 18th, 2008.

We use the following state evolution equation to model a stochastic animal movement :

$$\mathbf{x}_t = \mathbf{f}(\mathbf{x}_{t-1}, \theta_t) + \mathbf{n}_t \tag{1}$$

where :

$\mathbf{x}_t$ = vertical velocity
θ_t = latent behavioral parameters
$\mathbf{f}$ = the movement
$\mathbf{n}_t$ = system noise.

First equations

Observation equation

Equation (1) is not ideal because we would prefer the sought-after value θ_t to follow a Markov process, rather than the observed value $\mathbf{x}_t$

We artificially augment the state space by introducing the variable $X_t = \begin{pmatrix} \mathbf{x}_t \\ \theta_t \end{pmatrix}$.

The observation equation is given by $y_t = HX_t + e_t$, where $H = (1, 0)$

y_t represents the observation of vertical velocity at time t
e_t, the observation error, is a normal noise term.

We seek a Markov process of the form:

$$\begin{pmatrix} \mathbf{x}_t \\ \theta_t \end{pmatrix} = \mathbf{f}\begin{pmatrix} \mathbf{x}_{t-1} \\ \theta_{t-1} \end{pmatrix} + \begin{pmatrix} \mathbf{n}_t \\ \mathbf{v}_t \end{pmatrix} \tag{2}$$

First equations

How to deal with θ_t

Problem : there is no consistent model to represent the evolution of θ_t

Since it represents the behavioral the seal, this parameters is constant over by period. We consider windows of 26 minutes and we make the assumption that θ_t is constant over each window.

To estimate θ_t, we still introduce an arbitrary movement on θ_t:

$$\theta_t = \theta_{t-1} + \nu_t$$

where v_t follows a normal distribution with variance σ_ν^2.

First equations

Definition of the movement model f

f characterizes the movement of the seal

We use an $AR(2)$ to describe the vertical velocity of a seal :

$$z_t = a_1 z_{t-1} + a_2 z_{t-2} + \epsilon_t \tag{3}$$

In order to have a Markov process, we augmented the state space with the dummy variable ζ_t, such that:

$$\begin{pmatrix} \mathbf{z}_t \\ \zeta_t \end{pmatrix} = \begin{pmatrix} a_{1,t} & a_{2,t} \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{z}_{t-1} \\ \zeta_{t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_t \\ 0 \end{pmatrix} \tag{4}$$

Taking into account both state augmentations, we end up with the movement model :

$$\mathbf{X_t} = \begin{pmatrix} \mathbf{z_t} \\ \zeta_\mathbf{t} \\ a_{1,t} \\ a_{2,t} \end{pmatrix} = \begin{pmatrix} a_{1,t-1} & a_{2,t-1} & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{z}_{t-1} \\ \zeta_{t-1} \\ a_{1,t-1} \\ a_{2,t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_t \\ 0 \\ \nu_{1,t} \\ \nu_{2,t} \end{pmatrix} \tag{5}$$

The system noise ε_t takes the form of a normal mixture process that allows for occasional large values,
$\varepsilon_t \sim 0.9\mathcal{N}(0, \sigma_\varepsilon^2) + 0.1\mathcal{N}(0, 10\sigma_\varepsilon^2)$.

Meanwhile, $\nu_{1,t}$ and $\nu_{2,t}$ are random noise for the random walk of the parameters, $\nu_{1,t}, \nu_{2,t} \sim \mathcal{N}(0, \sigma_\nu^2)$.

And the observation model :

$$\mathbf{y}_t = \begin{pmatrix} 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{z_t} \\ \zeta_{\mathbf{t}} \\ a_{1,t} \\ a_{2,t} \end{pmatrix} + \mathbf{e_t} \tag{6}$$

with the observation error $\mathbf{e_t} \sim \mathcal{N}(0, \sigma_o^2)$. It can be rewritten as $\mathbf{y}_t = G_t \mathbf{X_t} + \mathbf{e_t}$.
This leads to the associated Feynman-Kac model where :

- M_t is the Hidden Markov Process described in (Eq. 5).
- G_t is the density of $\mathbf{y_t}$ knowing $\mathbf{X_t}$ (Eq. 6).

The algorithm of the paper

Estimation of $\theta = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ inside a time window

$m \leftarrow 0$ ▷ The iteration step
$\sigma_v^2(0) \leftarrow 0.1$ ▷ The parameter noise variance
$\theta_0(0) \leftarrow 0$ ▷ The initial value of the parameter
while $m < 10$ **do**
    Estimate $\theta_{1,\dots,T}(m)$ using a Particle Filter
    Update $\theta_0(m+1) \leftarrow \frac{1}{T}\sum_{t=1}^{T} \theta_t^m$
    Update $\sigma_v^2(m+1) \leftarrow \alpha \times \sigma_v^2(m)$
end while
Use $\theta_0(10)$ as the estimation for θ_i

Results using the aglorithm of the paper

Results of 20 run of the aglorithm to estimate a_1 and a_2

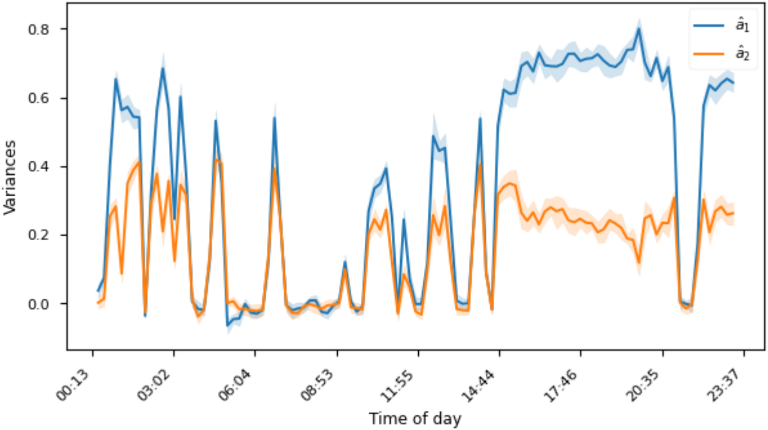

Figure: We thus want latent parameters estimate, that translates those states in a mathematical representation easier to analyze at scale.

How to use SMC^2 within this context

Definition of the Markov-process

In that section, we do not assume anymore that a_1, a_2 follows a random walk on each window. They are now supposed constant.

The Markov process used to define the Feynman-Kac model becomes the following:

$$\mathbf{X_t} = \begin{pmatrix} \mathbf{z_t} \\ \zeta_\mathbf{t} \end{pmatrix} = \begin{pmatrix} a_1 & a_2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{z}_{t-1} \\ \zeta_{t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_t \\ 0 \end{pmatrix} \tag{7}$$

A SMC defined with this Hidden Markov Process will return an estimation of for all $t, \mathbb{P}^{a_1,a_2}(y_{0:T})$.

To be able to have a good approximation of of the latent parameters a_1, a_2, we would like to estimate the density function:

$$(a_1, a_2) \to \mathbb{P}^{a_1,a_2}(y_{0:T})$$

It is possible to do that with a SMC sampler.

How to use SMC^2 within this context

Définition of the Inputs

In this SMC Sampler, the particles will represents (a_1, a_2). The inputs of the algorithm will be:

- $\mathcal{N}(0, \sigma_v)$ the prior distribution
- $\gamma_t = \mathbb{P}^{a1,a2}(y_{0:t})$
- A Random Walk kernel
- The usual input of an SMC algorithm: the number of particles N, the choice of an unbiased resampling scheme, and a threshold ESS min.

$\gamma_t(a_1, a_2) = \mathbb{P}^{a_1,a_2}(y_{0:t})$ would be computed using the particle filter described by the Hidden Markov Process above.

How to use SMC^2 within this context

Results of the SMC^2

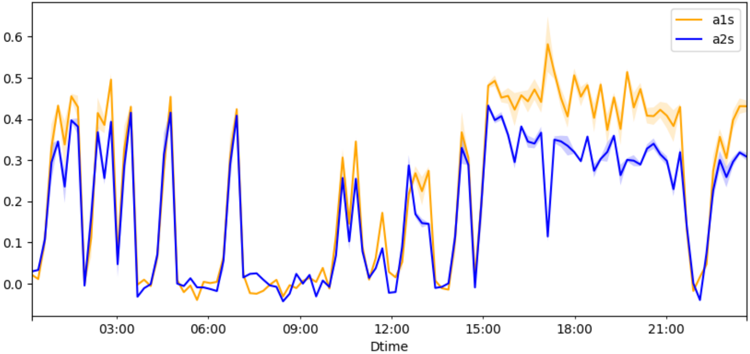

Figure: Results of the SMC2 to estimate a1, a2

Conclusion

We used an augmented state-space model to estimate the behavioral parameters of seals in Alaska.

However, the paper, the hypothesis and the data we worked on have several limits :

1. Window : arbitrary duration of a window (26 minutes), (a_1,a_2) are not likely to be constant over a window period
2. We don't use horizontal velocity to estimate the behavioral parameters of seals.
3. Need recurrence in the behavioral of seals
4. Difficult to use on other data or usecases

Appendix 1

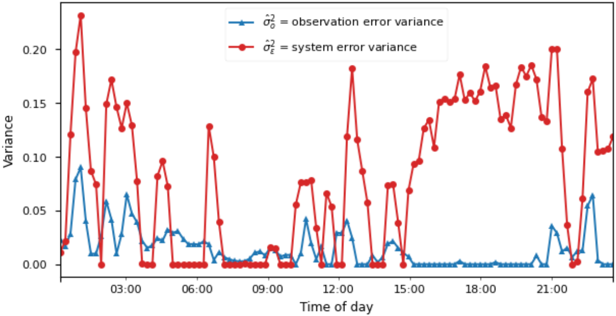

Figure: Offline estimation of the system error and observation error variances, using a quadratic regression on the ACVF for each time window.

Appendix 2

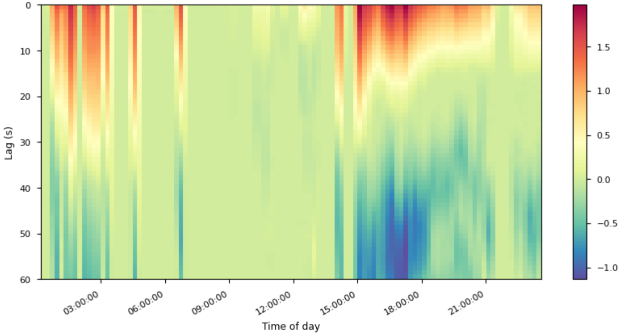

Figure: Estimation of the ACVF on each of the 109 time windows.

Appendix 3: Comparaison of the results

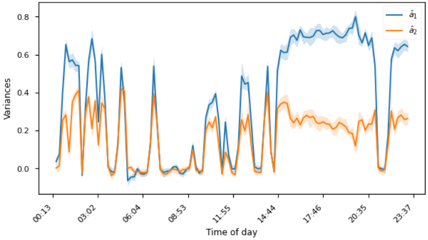

Figure: Results with the method of the article

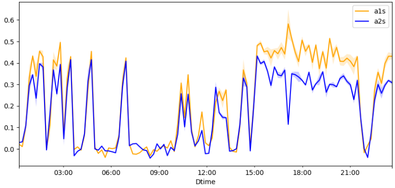

Figure: Results from the SMC^2 model

References

[1] Michael Dowd and Ruth Joy. Estimating behavioral parameters in animal movement models using a state-augmented particle filter. *Ecology*, 92(3):568–575, 2011.

[2] Nicolas Chopin, Omiros Papaspiliopoulos, et al. *An introduction to sequential Monte Carlo*. Springer, 2020.