PARAMETER AND DENSITY ESTIMATION IN SDES VIA PINNS AND NORMALIZING FLOWS: APPLICATIONS TO ENVIRONMENTAL SCIENCES

Séminaire Geolearning

Lucia Clarotto[1] (joint work with Hugo Gangloff[1], Nicolas Jouvin[1], Sophie Donnet[1], Marie-Pierre Etienne[2] and Adeline Samson[3])

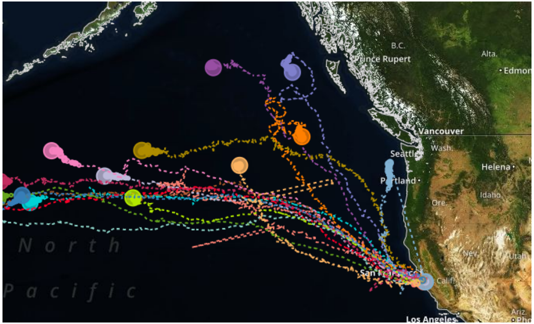
Fréjus, 31/03-03/04 2025

[1]MIA Paris-Saclay, AgroParisTech, INRAE, Université Paris Saclay
[2]ENSAI, Rennes
[3]Jean Kuntzmann Laboratory, Université Grenoble Alpes

GEOLEARNING
CHAIRE /// Data Science for the Environment

INTRODUCTION AND CONTEXT

Time series models, based on stochastic processes, offer a natural framework to describe the dynamics of animal movement, and to capture the strong serial correlation that is often present in such data.



Telemetry data of sea lions' movement in the North Pacific ocean

TRANSITION AND UTILIZATION DISTRIBUTIONS

Two main characteristics of a process (X_t) may be of interest, depending on the aim of the study: the **transition distribution** and the **utilization distribution**.

Transition distribution $p(X_t|X_s)$

Probability density of an animal's location at time t, given its location at time s, describing the short-term dynamics of the movement model.

$\rightarrow$ Lagrangian approach: describes the movement from the point of view of the individual animal

Utilization distribution p

Equilibrium distribution of the process $(X_t)_{t\geq 0}$, i.e., pdf of the animal's location in geographical space in the long-term

$$\Pr(X_t \in A) = \int_A p(x)\,dx$$

$\rightarrow$ Eulerian approach: describes the movement from the point of view of a point in space

$\Rightarrow$ Closely related, because the long-run distribution of space use arises from the accumulation of short-term displacements

CONTINUOUS-TIME MODELS

Continuous-time models consider that telemetry observations arise from a continuous movement process.

- naturally accommodate different temporal scales, and irregular sampling rates
- mostly based on diffusion processes:
  - Ornstein-Uhlenbeck processes (Uhlenbeck and Ornstein (1930))
  - Brownian bridges (Horne et al. (2007))
  - **more complex processes based on potential functions**

DIFFUSION STOCHASTIC DIFFERENTIAL EQUATIONS

DIFFUSION STOCHASTIC DIFFERENTIAL EQUATIONS

$$dX_t = a(t, X_t)\,dt + \sigma(t, X_t)\,dW_t, \quad X_0 = x_0$$

with

- $a(t, X_t)$ the drift coefficient, modeling the direction preference depending on position
- $\sigma(t, X_t)$ the diffusion coefficient, modeling the variability around the mean
- W_t the Wiener process, almost-surely continuous stochastic process, with stationary and independent increments, which satisfy

$$W_{t+\delta} - W_t \sim \mathcal{N}(0, \delta I)$$

Under some boundedness conditions on a and σ, and given an initial condition $X_0 = x_0$, the SDE has a unique solution $(X_t)_{t \geq 0}$.

DIFFUSION STOCHASTIC DIFFERENTIAL EQUATIONS

$$dX_t = a(t, X_t)\,dt + \sigma(t, X_t)\,dW_t, \quad X_0 = x_0$$

with

- $a(t, X_t)$ the drift coefficient, modeling the direction preference depending on position
- $\sigma(t, X_t)$ the diffusion coefficient, modeling the variability around the mean
- W_t the Wiener process, almost-surely continuous stochastic process, with stationary and independent increments, which satisfy

$$W_{t+\delta} - W_t \sim \mathcal{N}(0, \delta I)$$

Under some boundedness conditions on a and σ, and given an initial condition $X_0 = x_0$, the SDE has a unique solution $(X_t)_{t \geq 0}$.

Based on simulation

- Approximated simulation with a numerical scheme
    - Euler-Maruyama (Kloeden and Platen, 1992)
    - Numerical splitting scheme (Buckwar et al., 2022, Pilipovic et al., 2024)
- Approximate Bayesian Computation (ABC)

Based on approximation of the transition density

- Contrast estimator (Gloter, 2006, Ditlevsen and Samson, 2017)
- EM and SAEM algorithm (Beskos et al., 2006, Gloaguen et al., 2018, Ditlevsen and Samson, 2017)

- The marginal density $p(t,x)$ of the process $(X_t)_{t\geq 0}$ is the solution of the following PDE, called Fokker-Planck equation (FPE): for any t

$$\partial_t p(t,x) = \nabla \cdot \left(-a(t,x)p(t,x) + \frac{1}{2}\sigma(t,x)\sigma^\top(t,x)\nabla p(t,x) \right)$$

with the initial condition $p(0,x) = p_0(x_0)$

Fokker-Planck

The Fokker-Planck equation describes the evolution of the probability density $p(t,x)$ from an ensemble of stochastic trajectories initiated with density $p_0(x_0)$.

- The **stationary** (or long-run) distribution $p^{stat}(x)$ of the SDE is the solution of the ODE

$$\nabla \cdot \left(-a(x)p^{stat}(x) + \frac{1}{2}\sigma(x)\sigma^\top(x)\nabla p^{stat}(x) \right) = 0$$

- The marginal density $p(t,x)$ of the process $(X_t)_{t\geq 0}$ is the solution of the following PDE, called Fokker-Planck equation (FPE): for any t

$$\partial_t p(t,x) = \nabla \cdot \left(-a(t,x)p(t,x) + \frac{1}{2}\sigma(t,x)\sigma^\top(t,x)\nabla p(t,x) \right)$$

with the initial condition $p(0,x) = p_0(x_0)$

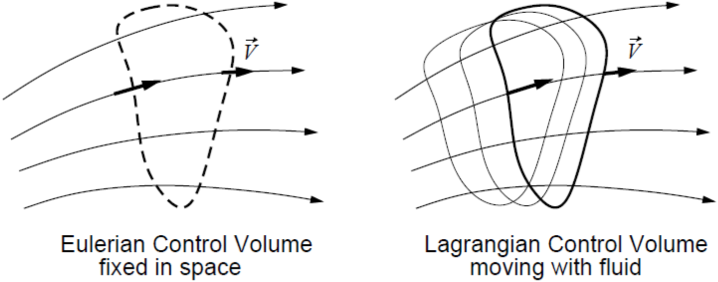
Fokker-Planck

The Fokker-Planck equation describes the evolution of the probability density $p(t,x)$ from an ensemble of stochastic trajectories initiated with density $p_0(x_0)$.

- The **stationary** (or long-run) distribution $p^{stat}(x)$ of the SDE is the solution of the ODE

$$\nabla \cdot \left(-a(x)p^{stat}(x) + \frac{1}{2}\sigma(x)\sigma^\top(x)\nabla p^{stat}(x) \right) = 0$$

SDE AND FPE: EULERIAN AND LAGRANGIAN POINT OF VIEW



Point of view of the SDE
(transition distribution)

Point of view of the FPE
(utilization distribution)

PINNS FOR FPE

The FPE can be viewed as a general PDE with a given differential operator $\mathcal{N}_\nu[u]$ depending on some parameter ν,

$$\mathcal{N}_\nu[u] = 0$$

Neural networks search for parametric $u_\theta \in \mathcal{F} = \{u_\theta,$ NNs with weights and biases $\theta\}$

Define a loss[1] $\rightsquigarrow$ PDE **residuals** $\rightarrow$ empirical version on collocation points: $\{x_i\}_{i=1}^r \subset \Omega$

$$\mathcal{L}_{PDE}(\theta) = \frac{1}{n} \sum_{i=1}^{r} |\mathcal{N}_\nu[u_\theta](x_i)|^2$$

$\rightarrow$ **Forward problem:** solve $\min_\theta \mathcal{L}_{PDE}(\theta)$ via stochastic optimization (mini-batch)
$\rightarrow$ **Inverse problem:** solve $\min_{\theta,\nu} \mathcal{L}_{PDE}(\theta, \nu)$ (include estimation of parameter ν)

[1] easily includes IC/BC and observations with some data-fitting term

The FPE can be viewed as a general PDE with a given differential operator $\mathcal{N}_\nu[u]$ depending on some parameter ν,

$$\mathcal{N}_\nu[u] = 0$$

Neural networks search for parametric $u_\theta \in \mathcal{F} = \{u_\theta,$ NNs with weights and biases $\theta\}$

Define a loss[1] $\rightsquigarrow$ PDE **residuals** $\rightarrow$ empirical version on collocation points: $\{x_i\}_{i=1}^r \subset \Omega$

$$\mathcal{L}_{PDE}(\theta) = \frac{1}{n} \sum_{i=1}^{r} |\mathcal{N}_\nu[u_\theta](x_i)|^2$$

$\rightarrow$ **Forward problem:** solve $\min_\theta \mathcal{L}_{PDE}(\theta)$ via stochastic optimization (mini-batch)

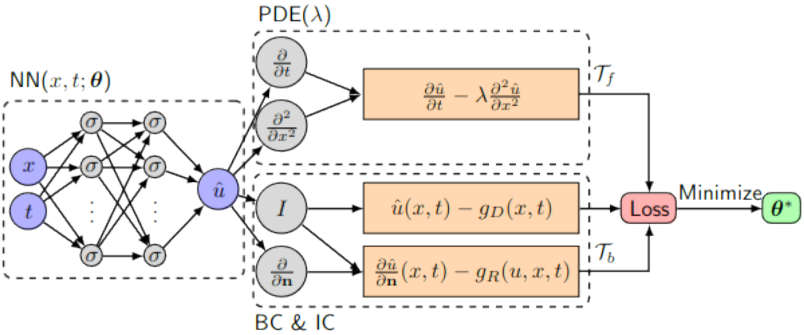
$\rightarrow$ **Inverse problem:** solve $\min_{\theta,\nu} \mathcal{L}_{PDE}(\theta, \nu)$ (include estimation of parameter ν)

[1] easily includes IC/BC and observations with some data-fitting term

WHY NEURAL NETWORKS ?

Neural networks are

1. good for approximating arbitrary complex functions
2. suitable for automatic differentiation $\rightsquigarrow$ easy-and-exact $\mathcal{N}_\nu[u_\theta](x)$ at any x
3. able to account for observations



Graphical representation of a PINN from Lu et al. (2021)

LOSS FUNCTION FOR STATIONARY FPE

The PINN loss for the stationary FPE is given by the sum of the observational and physical losses:

$$\mathcal{L}(p_\theta^{stat}) = \underbrace{-\sum_{i=1}^{n} \log p_\theta^{stat}(X_{t_i})}_{\mathcal{L}_{obs}(p_\theta^{stat})} + \underbrace{\frac{1}{r}\sum_{i=1}^{r}\left[\nabla\cdot\left(-a(x_i)p_\theta^{stat}(x_i) + \frac{1}{2}\sigma(x_i)\sigma^\top(x_i)\nabla p_\theta^{stat}(x_i)\right)\right]^2}_{\mathcal{L}_{PDE}(p_\theta^{stat})}$$

p_θ^{stat} must be a density $\Rightarrow$ we define $p_\theta^{stat} = \frac{\tilde{p}_\theta^{stat}}{\int_\Omega \tilde{p}_\theta^{stat}(x)\,dx} = \frac{1}{C}\tilde{p}_\theta^{stat}$, so that $\tilde{p}_\theta^{stat}$ can be any NN

$\Rightarrow$ $\tilde{p}_\theta^{stat}$ is also solution to the ODE, since

$$\nabla\cdot\left(-a(x)\tilde{p}_\theta^{stat}(x) + \frac{1}{2}\sigma(x)\sigma^\top(x)\nabla\tilde{p}_\theta^{stat}(x)\right) = \nabla\cdot\left(-a(x)Cp_\theta^{stat}(x) + \frac{1}{2}\sigma(x)\sigma^\top(x)\nabla Cp_\theta^{stat}(x)\right) = 0$$

Complete loss function

$$\mathcal{L}(\tilde{p}_\theta^{stat}) = \underbrace{-\frac{1}{n}\sum_{i=1}^{n}\log\tilde{p}_\theta^{stat}(X_{t_i}) + \log\int_\Omega \tilde{p}_\theta^{stat}(x)\,dx}_{\mathcal{L}_{obs}(\tilde{p}^{stat})} + w_{PDE}\mathcal{L}_{PDE}(\tilde{p}_\theta^{stat}) + w_{norm}\underbrace{\left|\int_\Omega \tilde{p}_\theta^{stat}(x)\,dx - 1\right|^2}_{\mathcal{L}_{norm}(\tilde{p}^{stat})}.$$

The PINN loss for the stationary FPE is given by the sum of the observational and physical losses:

$$\mathcal{L}(p_\theta^{stat}) = \underbrace{-\sum_{i=1}^{n} \log p_\theta^{stat}(X_{t_i})}_{\mathcal{L}_{obs}(p_\theta^{stat})} + \underbrace{\frac{1}{r}\sum_{i=1}^{r}\left[\nabla\cdot\left(-a(x_i)p_\theta^{stat}(x_i) + \frac{1}{2}\sigma(x_i)\sigma^\top(x_i)\nabla p_\theta^{stat}(x_i)\right)\right]^2}_{\mathcal{L}_{PDE}(p_\theta^{stat})}$$

p_θ^{stat} must be a density $\Rightarrow$ we define $p_\theta^{stat} = \frac{\tilde{p}_\theta^{stat}}{\int_\Omega \tilde{p}_\theta^{stat}(x)\,dx} = \frac{1}{C}\tilde{p}_\theta^{stat}$, so that $\tilde{p}_\theta^{stat}$ can be any NN

$\Rightarrow$ $\tilde{p}_\theta^{stat}$ is also solution to the ODE, since

$$\nabla\cdot\left(-a(x)\tilde{p}_\theta^{stat}(x) + \frac{1}{2}\sigma(x)\sigma^\top(x)\nabla\tilde{p}_\theta^{stat}(x)\right) = \nabla\cdot\left(-a(x)Cp_\theta^{stat}(x) + \frac{1}{2}\sigma(x)\sigma^\top(x)\nabla Cp_\theta^{stat}(x)\right) = 0$$

Complete loss function

$$\mathcal{L}(\tilde{p}_\theta^{stat}) = \underbrace{-\frac{1}{n}\sum_{i=1}^{n}\log\tilde{p}_\theta^{stat}(X_{t_i}) + \log\int_\Omega \tilde{p}_\theta^{stat}(x)\,dx}_{\mathcal{L}_{obs}(\tilde{p}^{stat})} + w_{PDE}\mathcal{L}_{PDE}(\tilde{p}_\theta^{stat}) + w_{norm}\underbrace{\left|\int_\Omega \tilde{p}_\theta^{stat}(x)\,dx - 1\right|^2}_{\mathcal{L}_{norm}(\tilde{p}^{stat})}.$$

The PINN loss for the stationary FPE is given by the sum of the observational and physical losses:

$$\mathcal{L}(p_\theta^{stat}) = \underbrace{-\sum_{i=1}^{n} \log p_\theta^{stat}(X_{t_i})}_{\mathcal{L}_{obs}(p_\theta^{stat})} + \underbrace{\frac{1}{r}\sum_{i=1}^{r}\left[\nabla \cdot \left(-a(x_i)p_\theta^{stat}(x_i) + \frac{1}{2}\sigma(x_i)\sigma^\top(x_i)\nabla p_\theta^{stat}(x_i)\right)\right]^2}_{\mathcal{L}_{PDE}(p_\theta^{stat})}$$

p_θ^{stat} must be a density $\Rightarrow$ we define $p_\theta^{stat} = \frac{\tilde{p}_\theta^{stat}}{\int_\Omega \tilde{p}_\theta^{stat}(x)\,dx} = \frac{1}{C}\tilde{p}_\theta^{stat}$, so that $\tilde{p}_\theta^{stat}$ can be any NN

$\Rightarrow$ $\tilde{p}_\theta^{stat}$ is also solution to the ODE, since

$$\nabla \cdot \left(-a(x)\tilde{p}_\theta^{stat}(x) + \frac{1}{2}\sigma(x)\sigma^\top(x)\nabla \tilde{p}_\theta^{stat}(x)\right) = \nabla \cdot \left(-a(x)Cp_\theta^{stat}(x) + \frac{1}{2}\sigma(x)\sigma^\top(x)\nabla Cp_\theta^{stat}(x)\right) = 0$$

Complete loss function

$$\mathcal{L}(\tilde{p}_\theta^{stat}) = \underbrace{-\frac{1}{n}\sum_{i=1}^{n} \log \tilde{p}_\theta^{stat}(X_{t_i}) + \log \int_\Omega \tilde{p}_\theta^{stat}(x)\,dx}_{\mathcal{L}_{obs}(\tilde{p}^{stat})} + w_{PDE}\mathcal{L}_{PDE}(\tilde{p}_\theta^{stat}) + w_{norm}\underbrace{\left|\int_\Omega \tilde{p}_\theta^{stat}(x)\,dx - 1\right|^2}_{\mathcal{L}_{norm}(\tilde{p}^{stat})}.$$

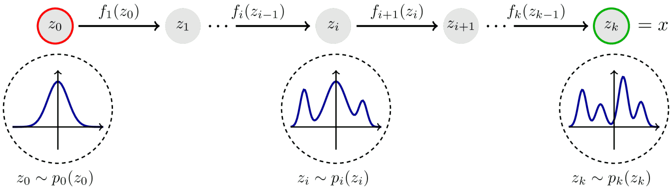
Instead of computing the normalization term in the loss, we test a Normalizing Flow (NF) network architecture for p_θ^{stat}. We take our inspiration from Liu et al. (2023).

Normalizing Flow

Class of generative models that transform a simple probability distribution into a complex one using a **series of invertible functions** (and such that the computation of the determinant of the Jacobian of the inverse is efficient).

Given a latent variable $z \sim p_z(z)$ and an invertible function f_θ, the variable $x = f_\theta(z)$ follows

$$p_x(x) = p_z(z) \left| \det \frac{\partial f_\theta^{-1}(x)}{\partial x} \right|$$



TESTS

Ornstein-Uhlenbeck (OU) SDE (Uhlenbeck and Ornstein, 1930)

$$dX_t = \alpha(\mu - X_t)dt + \sigma dW_t$$

with W_t a 2D Wiener process, $\mu \in \mathbb{R}^2$ and α and σ (2×2) matrices, with α invertible.

- Stationary distribution: $p^{stat} = \mathcal{N}(\mu, \Sigma)$
  with $\Sigma = (\alpha \oplus \alpha)^{-1}\text{vec}(2D)$ and $D = \frac{\sigma\sigma^\top}{2}$ (diffusion tensor)
- Transition probability : $X_t \mid X_s \sim \mathcal{N}(m(t-s), \Sigma_2(t-s))$
  with $m(t-s) = e^{-\alpha(t-s)}X_s + (I - e^{-\alpha t})\mu$ and $\Sigma_2(t-s) = \int_0^t e^{\alpha(s-t)}\sigma\sigma^\top e^{\alpha^\top(s-t)}ds$

- We estimate the stationary pdf (and eventually the parameters $\theta = \{\mu, \Sigma\}$) using a NF-PINN p_θ^{stat} that respects the FPE corresponding to the OU SDE.
- We reparameterize the stationary FPE so that it is defined through Σ instead of α and σ.

Ornstein-Uhlenbeck (OU) SDE (Uhlenbeck and Ornstein, 1930)

$$dX_t = \alpha(\mu - X_t)dt + \sigma dW_t$$

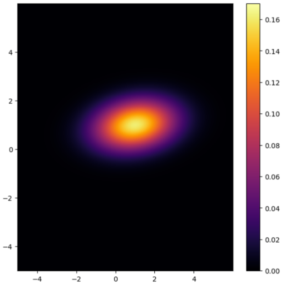
with W_t a 2D Wiener process, $\mu \in \mathbb{R}^2$ and α and σ (2×2) matrices, with α invertible.

- Stationary distribution: $p^{stat} = \mathcal{N}(\mu, \Sigma)$
  with $\Sigma = (\alpha \oplus \alpha)^{-1}\mathrm{vec}(2D)$ and $D = \frac{\sigma\sigma^\top}{2}$ (diffusion tensor)
- Transition probability : $X_t \mid X_s \sim \mathcal{N}(m(t-s), \Sigma_2(t-s))$
  with $m(t-s) = e^{-\alpha(t-s)}X_s + (I - e^{-\alpha t})\mu$ and $\Sigma_2(t-s) = \int_0^t e^{\alpha(s-t)}\sigma\sigma^\top e^{\alpha^\top(s-t)}ds$

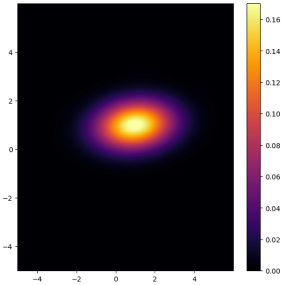
- We estimate the stationary pdf (and eventually the parameters $\theta = \{\mu, \Sigma\}$) using a NF-PINN p_θ^{stat} that respects the FPE corresponding to the OU SDE.
- We reparameterize the stationary FPE so that it is defined through Σ instead of α and σ.

$$dX_t = \alpha(\mu - X_t)\,dt + \sigma\, dW_t$$

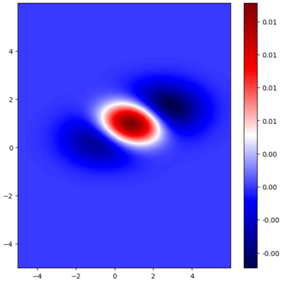
Parameters: $\alpha = \begin{bmatrix} 1 & 0.1 \\ 0.1 & 1 \end{bmatrix}, \quad \sigma = \begin{bmatrix} 2 & 0.2 \\ 0.2 & 1 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 2.003 & 0.175 \\ 0.175 & 0.503 \end{bmatrix}, \quad \mu = [1, 1]^\top$



Stationary distribution $p^{stat} = \mathcal{N}(\mu, \Sigma)$



Estimated distribution $\hat{p}^{stat}$



$\hat{p}^{stat} - p^{stat}$

Estimated parameters (via inverse problem): $\hat{\Sigma} = \begin{bmatrix} 2.095 & 0.240 \\ 0.240 & 0.458 \end{bmatrix}, \quad \hat{\mu} = [1.009, 1.020]^\top$

Langevin movement

Michelot et al. (2019) propose to use the Langevin equation to model the animal movement

$$dX_t = \frac{\gamma^2}{2} \nabla \log p(X_t)\, dt + \gamma\, dW_t, \quad X_0 \sim x_0$$

with γ^2 the speed parameter.

Resource selection function

In order to link the movement of the Langevin process to environmental drivers, Michelot et al. (2019) model the utilization distribution p with a resource selection function (or **potential function**)

$$p(x|\beta) = \frac{1}{C} \underbrace{\exp\left(\sum_{j=1}^{J} \beta_j c_j(x)\right)}_{\text{potential function}}, \quad \nabla \log p(x|\beta) = \sum_{j=1}^{J} \beta_j c_j(x)$$

with J spatial covariates $c_1, \ldots, c_J$ and parameters $\beta_1, \ldots, \beta_J$, and C a normalizing constant.

The potential surface represents the forces of attraction and repulsion driving animals' movement.

Langevin movement

Michelot et al. (2019) propose to use the Langevin equation to model the animal movement

$$dX_t = \frac{\gamma^2}{2} \nabla \log p(X_t)\, dt + \gamma\, dW_t, \quad X_0 \sim x_0$$

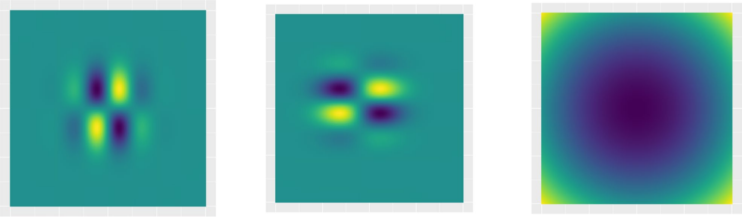
with γ^2 the speed parameter.

Resource selection function

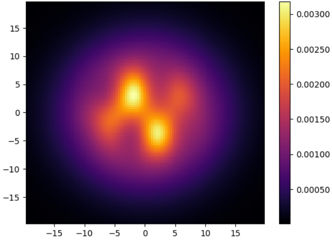
In order to link the movement of the Langevin process to environmental drivers, Michelot et al. (2019) model the utilization distribution p with a resource selection function (or **potential function**)

$$p(x|\beta) = \frac{1}{C} \underbrace{\exp\left(\sum_{j=1}^{J} \beta_j c_j(x)\right)}_{\text{potential function}}, \quad \nabla \log p(x|\beta) = \sum_{j=1}^{J} \beta_j c_j(x)$$

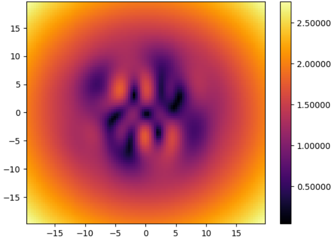
with J spatial covariates $c_1, \ldots, c_J$ and parameters $\beta_1, \ldots, \beta_J$, and C a normalizing constant.
The potential surface represents the forces of attraction and repulsion driving animals' movement.



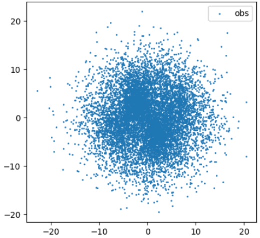
Three spatial covariates c_1, c_2, c_3 of the resource selection function



Stationary distribution p^{stat}

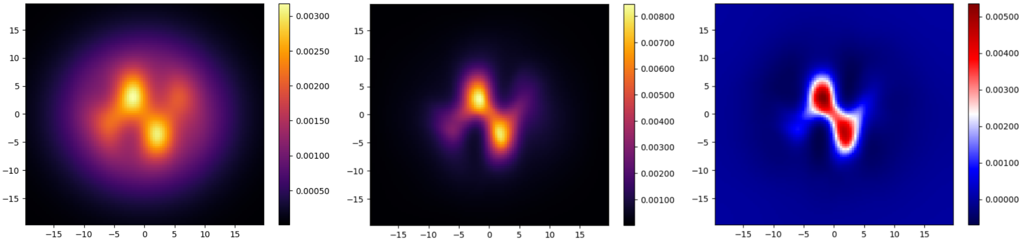


Gradient $\nabla \log p^{stat}$



Observations (Euler-Maruyama)

- We estimate p^{stat} with a NF-PINN p_θ^{stat} from telemetry and habitat data (the covariates c_1, c_2, c_3).
- We focus on one individual animal.



Stationary distribution p^{stat} | Estimated distribution $\hat{p}^{stat}$ | $\hat{p}^{stat} - p^{stat}$

PERSPECTIVES

Issues (work in progress)

- The stationary distribution is not perfectly captured.
- The inverse problem does not retrieve the true covariate parameters $\beta_1, \beta_2, \beta_3$.

Next developments

- Apply to a second case study: randomly generated covariate fields on a discrete grid

$\rightarrow$ more similar to real environmental data.

- Apply to a real dataset (Stellar sea lions in Alaska).
- Compare with results in Michelot et al. (2019).

PERSPECTIVES

Issues (work in progress)

- The stationary distribution is not perfectly captured.
- The inverse problem does not retrieve the true covariate parameters $\beta_1, \beta_2, \beta_3$.

Next developments

- Apply to a second case study: randomly generated covariate fields on a discrete grid

$\rightarrow$ more similar to real environmental data.

- Apply to a real dataset (Stellar sea lions in Alaska).
- Compare with results in Michelot et al. (2019).

Issues of Michelot et al. (2019)

- The movement model was chosen so that the stationary distribution (accounting for environmental covariates) is convenient.
- Handling categorical covariates is challenging (problems in the gradients), yet they are the most common.

Potential of our approach

- Freely choose the SDE to characterize movement and estimate the stationary distribution and parameters.
  - Unless we find a corresponding interesting potential function, we would loose the direct inference of covariate effect.
- Add variability in diffusion coefficient σ based on the environment via categorical parameters.

POTENTIAL BENEFITS

Issues of Michelot et al. (2019)

- The movement model was chosen so that the stationary distribution (accounting for environmental covariates) is convenient.
- Handling categorical covariates is challenging (problems in the gradients), yet they are the most common.

Potential of our approach

- Freely choose the SDE to characterize movement and estimate the stationary distribution and parameters.
    - Unless we find a corresponding interesting potential function, we would loose the direct inference of covariate effect.
- Add variability in diffusion coefficient σ based on the environment via categorical parameters.

Beskos, A., Papaspiliopoulos, O., Roberts, G. O., and Fearnhead, P. (2006). Exact and computationally efficient likelihood-based estimation for discretely observed diffusion processes (with discussion). Journal of the Royal Statistical Society Series B: Statistical Methodology, 68(3):333–382.

Buckwar, E., Samson, A., Tamborrino, M., and Tubikanec, I. (2022). A splitting method for sdes with locally lipschitz drift: Illustration on the fitzhugh-nagumo model. Applied Numerical Mathematics, 179:191–220.

Ditlevsen, P. and Ditlevsen, S. (2023). Warning of a forthcoming collapse of the atlantic meridional overturning circulation. Nature communications, 14(1):4254–12.

Ditlevsen, S. and Samson, A. (2017). Hypoelliptic diffusions: discretization, filtering and inference from complete and partial observations. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 81.

Gloaguen, P., Etienne, M.-P., and Le Corff, S. (2018). Online sequential monte carlo smoother for partially observed diffusion processes. EURASIP Journal on Advances in Signal Processing, 2018.

Gloter, A. (2006). Parameter estimation for a discretely observed integrated diffusion process. Scandinavian Journal of Statistics, 33(1):83–104.

Horne, J. S., Garton, E. O., Krone, S. M., and Lewis, J. S. (2007). Analyzing animal movements using brownian bridges. Ecology, 88(9):2354–2363.

Kloeden, P. E. and Platen, E. (1992). Numerical Solution of Stochastic Differential Equations. Springer Berlin Heidelberg, Berlin, Heidelberg.

Liu, F., Wu, F., and Zhang, X. (2023). Pinf: Continuous normalizing flows for physics-constrained deep learning.

Lu, L., Meng, X., Mao, Z., and Karniadakis, G. E. (2021). Deepxde: A deep learning library for solving differential equations. SIAM review, 63(1):208–228.

Michelot, T., Gloaguen, P., Blackwell, P. G., and Étienne, M.-P. (2019). The langevin diffusion as a continuous-time model of animal movement and habitat selection. Methods in Ecology and Evolution, 10(11):1894–1907.

Pilipovic, P., Samson, A., and Ditlevsen, S. (2024). Parameter estimation in nonlinear multivariate stochastic differential equations based on splitting schemes. The Annals of Statistics, 52(2):842 – 867.

Uhlenbeck, G. E. and Ornstein, L. S. (1930). On the theory of the brownian motion. Phys. Rev., 36:823–841.