# Modeling abundance time series through a «pseudo-HMM» framework

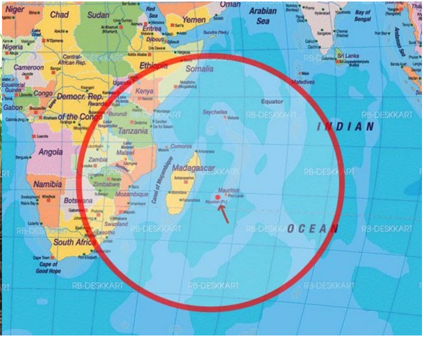
Guillaume Franchi
Joint work with L. Truquet and M-P. Etienne

Ecodep Closing Conference
30 September 2024

# Outlines

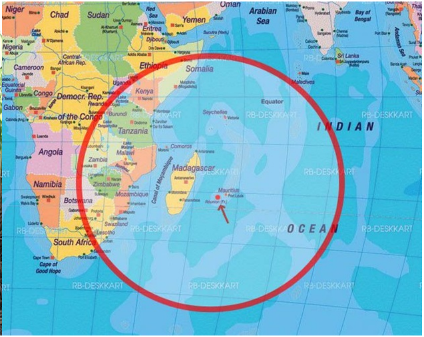
# Insect pests

◎ The objective of the initial study is to control a population of insect pests in a sugar-cane field in La Réunion, while limiting the use of pesticides.



# Insect pests

The objective of the initial study is to control a population of insect pests in a sugar-cane field in La Réunion, while limiting the use of pesticides.




It requires a good understanding of the ecosystem, and the impacts of species on each other.

# The study (1/2)

We focus here on four group of species

(a) Coleoptera

(b) Diptera

(c) Hymenoptera

(d) Oribatida

# The study (1/2)

We focus here on four group of species


(a) Coleoptera


(b) Diptera


(c) Hymenoptera

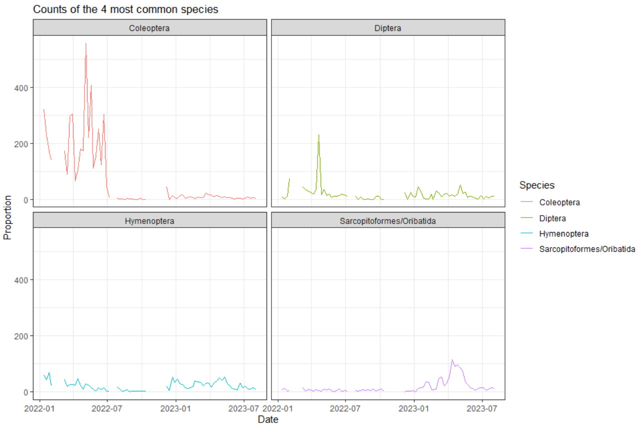

(d) Oribatida

Their population has been observed weekly from January, 2022 to August, 2023, catching the insects into traps.

# The study (2/2)

We thus obtain a time series of counts $(Y_t)_{1\leqslant t\leqslant 82}$ valued in $\mathbb{N}^4$

$$Y_t = (Y_{1,t},\ Y_{2,t},\ Y_{3,t},\ Y_{4,t}),\ 1 \leqslant t \leqslant 82.$$



# The underlying relative abundance

⚠ The time series of counts $(Y_t)_{t\in\mathbb{Z}}$ is just a representation of the entire population.

# The underlying relative abundance

⚠ The time series of counts $(Y_t)_{t\in\mathbb{Z}}$ is just a representation of the entire population.

➜ It may not reflect the exact reality of the ecosystem.

# The underlying relative abundance

⚠ The time series of counts $(Y_t)_{t\in\mathbb{Z}}$ is just a representation of the entire population.

➜ It may not reflect the exact reality of the ecosystem.

Each multivariate count $Y_t$ depend strongly on the overall proportion of each species in the whole ecosystem: the **relative abundance** $X_t$.

# The underlying relative abundance

⚠ The time series of counts $(Y_t)_{t\in\mathbb{Z}}$ is just a representation of the entire population.

➜ It may not reflect the exact reality of the ecosystem.

Each multivariate count $Y_t$ depend strongly on the overall proportion of each species in the whole ecosystem: the **relative abundance** $X_t$.

$X_t$ is the vector of proportions of each species in the whole ecosystem. For $p$ species, $X_t$ is valued in the simplex

$$\mathcal{S}_{p-1} = \left\{ (x_1, \ldots, x_p) \in (0, +\infty)^p : \sum_{1=1}^{p} x_i = 1 \right\}.$$

# The underlying relative abundance

⚠ The time series of counts $(Y_t)_{t\in\mathbb{Z}}$ is just a representation of the entire population.

➜ It may not reflect the exact reality of the ecosystem.

Each multivariate count $Y_t$ depend strongly on the overall proportion of each species in the whole ecosystem: the **relative abundance** $X_t$.

$X_t$ is the vector of proportions of each species in the whole ecosystem. For $p$ species, $X_t$ is valued in the simplex

$$\mathcal{S}_{p-1} = \left\{ (x_1, \ldots, x_p) \in (0, +\infty)^p : \sum_{1=1}^{p} x_i = 1 \right\}.$$

Our goal is to provide a model for the joint process $(X_t, Y_t)_{t\in\mathbb{Z}}$, where $(X_t)_{t\in\mathbb{Z}}$ is not observed.

# Outlines

# Dynamic of the relative abundance $(X_t)_{t\in\mathbb{Z}}$ (1/2)

Let us fix $p$ the number of species in the ecosystem, and consider a process $(Z_t)_{t\in\mathbb{Z}}$ of $k$ exogenous variables.

# Dynamic of the relative abundance $(X_t)_{t\in\mathbb{Z}}$ (1/2)

Let us fix $p$ the number of species in the ecosystem, and consider a process $(Z_t)_{t\in\mathbb{Z}}$ of $k$ exogenous variables.

$(X_t)_{t\in\mathbb{Z}}$ is defined as a Markov chain with a Dirichlet transition kernel

$$\mathbb{P}\left(X_t \mid X_{t-1}, Z_t\right) = \mathrm{Dir}\left(\alpha_t\right),$$

## Dynamic of the relative abundance $(X_t)_{t\in\mathbb{Z}}$ (1/2)

Let us fix $p$ the number of species in the ecosystem, and consider a process $(Z_t)_{t\in\mathbb{Z}}$ of $k$ exogenous variables.

$(X_t)_{t\in\mathbb{Z}}$ is defined as a Markov chain with a Dirichlet transition kernel

$$\mathbb{P}\left(X_t \mid X_{t-1}, Z_t\right) = \mathrm{Dir}\left(\alpha_t\right),$$

where $\alpha_t = \alpha(X_{t-1}, Z_t) \in (0, +\infty)^p$ satisfies $\alpha(X_{t-1}, Z_t) = \varphi_t \cdot \mu_t$ with:

# Dynamic of the relative abundance $(X_t)_{t\in\mathbb{Z}}$ (1/2)

Let us fix $p$ the number of species in the ecosystem, and consider a process $(Z_t)_{t\in\mathbb{Z}}$ of $k$ exogenous variables.

$(X_t)_{t\in\mathbb{Z}}$ is defined as a Markov chain with a Dirichlet transition kernel

$$\mathbb{P}\left(X_t \mid X_{t-1}, Z_t\right) = \mathrm{Dir}\left(\alpha_t\right),$$

where $\alpha_t = \alpha(X_{t-1}, Z_t) \in (0, +\infty)^p$ satisfies $\alpha(X_{t-1}, Z_t) = \varphi_t \cdot \mu_t$ with:

- ➜ $\varphi_t \in \mathbb{R}_+^*$ a dispersion parameter

$$\varphi_t = \exp\left(a_0 + a_1 I_S(X_{t-1})\right),$$

where $I_S$ denotes here the Shannon entropy.

# Dynamic of the relative abundance $(X_t)_{t\in\mathbb{Z}}$ (1/2)

Let us fix $p$ the number of species in the ecosystem, and consider a process $(Z_t)_{t\in\mathbb{Z}}$ of $k$ exogenous variables.

$(X_t)_{t\in\mathbb{Z}}$ is defined as a Markov chain with a Dirichlet transition kernel

$$\mathbb{P}\left(X_t \mid X_{t-1}, Z_t\right) = \mathrm{Dir}\left(\alpha_t\right),$$

where $\alpha_t = \alpha(X_{t-1}, Z_t) \in (0, +\infty)^p$ satisfies $\alpha(X_{t-1}, Z_t) = \varphi_t \cdot \mu_t$ with:

- ➜ $\varphi_t \in \mathbb{R}_+^*$ a dispersion parameter

  $$\varphi_t = \exp\left(a_0 + a_1 I_S(X_{t-1})\right),$$

  where $I_S$ denotes here the Shannon entropy.
- ➜ $\mu_t = (\mu_{1,t}, \ldots, \mu_{p,t}) \in \mathcal{S}_{p-1}$ a mean parameter

  $$\log\left(\frac{\mu_{i,t}}{\mu_{p,t}}\right) = A_0(i) + A_1(i,\cdot)\cdot(X_{1,t-1}, \ldots X_{p-1,t-1})' + B(i,\cdot)\cdot Z_t,\ 1 \leqslant i \leqslant p-1,$$

  where $A_0 \in \mathbb{R}^{p-1}$, $A_1$ is a matrix of dimension $(p-1)\times(p-1)$ and $B$ is a matrix with $(p-1)$ rows.

# Dynamic of the relative abundance $(X_t)_{t\in z}$ (2/2)

The process $(X_t)_{t\in\mathbb{Z}}$ is actually a non-homogeneous Markov chain, with a random transition kernel determined by $(Z_t)_{t\in\mathbb{Z}}$.

# Dynamic of the relative abundance $(X_t)_{t\in z}$ (2/2)

The process $(X_t)_{t\in\mathbb{Z}}$ is actually a non-homogeneous Markov chain, with a random transition kernel determined by $(Z_t)_{t\in\mathbb{Z}}$.

The Dirichlet transtion kernel $P(\cdot \mid x, z) \sim \mathrm{Dir}(\alpha(x, z))$ satisfies a Doeblin condition

$$\forall x \in \mathcal{S}_{p-1},\ \forall z \in \mathbb{R}^k,\ \forall A \in \mathcal{B}\left(\mathcal{S}_{p-1}\right)\ \ P(A \mid x, z) \geqslant \varepsilon(z)\mu(z, A),$$

for some fixed measurable application $\varepsilon$ valued in $(0, 1]$ and a fixed Markov kernel $\mu$ defined on $\mathbb{R}^k \times \mathcal{B}(\mathcal{S}_{p-1})$.

# Dynamic of the relative abundance $(X_t)_{t\in z}$ (2/2)

The process $(X_t)_{t\in\mathbb{Z}}$ is actually a non-homogeneous Markov chain, with a random transition kernel determined by $(Z_t)_{t\in\mathbb{Z}}$.

The Dirichlet transtion kernel $P(\cdot \mid x, z) \sim \mathrm{Dir}(\alpha(x, z))$ satisfies a Doeblin condition

$$\forall x \in \mathcal{S}_{p-1},\ \forall z \in \mathbb{R}^k,\ \forall A \in \mathcal{B}\left(\mathcal{S}_{p-1}\right)\ \ P(A \mid x, z) \geqslant \varepsilon(z)\mu(z, A),$$

for some fixed measurable application $\varepsilon$ valued in $(0, 1]$ and a fixed Markov kernel $\mu$ defined on $\mathbb{R}^k \times \mathcal{B}(\mathcal{S}_{p-1})$.

**Proposition 1**

Assume that $(Z_t)_{t\in\mathbb{Z}}$ is stationary.

# Dynamic of the relative abundance $(X_t)_{t\in z}$ (2/2)

The process $(X_t)_{t\in\mathbb{Z}}$ is actually a non-homogeneous Markov chain, with a random transition kernel determined by $(Z_t)_{t\in\mathbb{Z}}$.

The Dirichlet transtion kernel $P(\cdot \mid x, z) \sim \mathrm{Dir}(\alpha(x, z))$ satisfies a Doeblin condition

$$\forall x \in \mathcal{S}_{p-1},\ \forall z \in \mathbb{R}^k,\ \forall A \in \mathcal{B}\left(\mathcal{S}_{p-1}\right)\ \ P(A \mid x, z) \geqslant \varepsilon(z)\mu(z, A),$$

for some fixed measurable application $\varepsilon$ valued in $(0, 1]$ and a fixed Markov kernel $\mu$ defined on $\mathbb{R}^k \times \mathcal{B}(\mathcal{S}_{p-1})$.

**Proposition 1**

Assume that $(Z_t)_{t\in\mathbb{Z}}$ is stationary.
Then, there exists a stationary process $(X_t, Z_t)_{t\in\mathbb{Z}}$ satisfying our dynamics, and its distribution is unique.

# Dynamic of the relative abundance $(X_t)_{t\in z}$ (2/2)

The process $(X_t)_{t\in\mathbb{Z}}$ is actually a non-homogeneous Markov chain, with a random transition kernel determined by $(Z_t)_{t\in\mathbb{Z}}$.

The Dirichlet transtion kernel $P(\cdot \mid x, z) \sim \mathrm{Dir}(\alpha(x, z))$ satisfies a Doeblin condition

$$\forall x \in \mathcal{S}_{p-1},\ \forall z \in \mathbb{R}^k,\ \forall A \in \mathcal{B}\left(\mathcal{S}_{p-1}\right)\ \ P(A \mid x, z) \geqslant \varepsilon(z)\mu(z, A),$$

for some fixed measurable application $\varepsilon$ valued in $(0, 1]$ and a fixed Markov kernel $\mu$ defined on $\mathbb{R}^k \times \mathcal{B}(\mathcal{S}_{p-1})$.

## Proposition 1

Assume that $(Z_t)_{t\in\mathbb{Z}}$ is stationary.
Then, there exists a stationary process $(X_t, Z_t)_{t\in\mathbb{Z}}$ satisfying our dynamics, and its distribution is unique.
Moreover, if $(Z_t)_{t\in\mathbb{Z}}$ is ergodic, then $(X_t, Z_t)_{t\in\mathbb{Z}}$ is also ergodic.

# Dynamic of the count process $(Y_t)_{t\in\mathbb{Z}}$

The count process $(Y_t)_{t\in\mathbb{Z}}$ is derived from the relative abundance process $(X_t)_{t\in\mathbb{Z}}$ and a process $(N_t)_{t\in\mathbb{Z}}$ valued in $\mathbb{N}$ accounting for the total number of counts at time $t$:

$$Y_t \sim \mathcal{M}_p(N_t, X_t),$$

where $\mathcal{M}_p(N_t, X_t)$ denotes the multinomial distribution.

# Dynamic of the count process $(Y_t)_{t\in\mathbb{Z}}$

The count process $(Y_t)_{t\in\mathbb{Z}}$ is derived from the relative abundance process $(X_t)_{t\in\mathbb{Z}}$ and a process $(N_t)_{t\in\mathbb{Z}}$ valued in $\mathbb{N}$ accounting for the total number of counts at time $t$:

$$Y_t \sim \mathcal{M}_p(N_t, X_t),$$

where $\mathcal{M}_p(N_t, X_t)$ denotes the multinomial distribution.

💡 The Dirichlet distribution and the multinomial distribution are conjugate to each other.

# Dynamic of the count process $(Y_t)_{t\in\mathbb{Z}}$

The count process $(Y_t)_{t\in\mathbb{Z}}$ is derived from the relative abundance process $(X_t)_{t\in\mathbb{Z}}$ and a process $(N_t)_{t\in\mathbb{Z}}$ valued in $\mathbb{N}$ accounting for the total number of counts at time $t$:

$$Y_t \sim \mathcal{M}_p(N_t, X_t),$$

where $\mathcal{M}_p(N_t, X_t)$ denotes the multinomial distribution.

The Dirichlet distribution and the multinomial distribution are conjugate to each other.

Dirichlet density function:

$$f_\alpha(x_1, \dots, x_p) = \frac{\Gamma(\sum \alpha_i)}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_p)} x_1^{\alpha_1 - 1} \cdots x_p^{\alpha_p - 1}.$$

# Dynamic of the count process $(Y_t)_{t\in\mathbb{Z}}$

The count process $(Y_t)_{t\in\mathbb{Z}}$ is derived from the relative abundance process $(X_t)_{t\in\mathbb{Z}}$ and a process $(N_t)_{t\in\mathbb{Z}}$ valued in $\mathbb{N}$ accounting for the total number of counts at time $t$:

$$Y_t \sim \mathcal{M}_p(N_t, X_t),$$

where $\mathcal{M}_p(N_t, X_t)$ denotes the multinomial distribution.

The Dirichlet distribution and the multinomial distribution are conjugate to each other.

- Dirichlet density function:

$$f_\alpha(x_1, \dots, x_p) = \frac{\Gamma(\sum \alpha_i)}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_p)} x_1^{\alpha_1 - 1} \cdots x_p^{\alpha_p - 1}.$$

- Multinomial mass function:

$$f_{N,x}(y_1, \dots y_p) = \frac{N!}{y_1! \cdots y_p!} x_1^{y_1}, \cdots x_p^{y_p}.$$

# «Pseudo-HMM» framework (1/2)

We assume that:

# «Pseudo-HMM» framework (1/2)

We assume that:

- Processes $(X_t)_{t\in\mathbb{Z}}$ and $(N_t)_{t\in\mathbb{Z}}$ are independent;

## «Pseudo-HMM» framework (1/2)

We assume that:

- Processes $(X_t)_{t\in\mathbb{Z}}$ and $(N_t)_{t\in\mathbb{Z}}$ are independent;
- $(N_t)_{t\in\mathbb{Z}}$ is a Markov chain;

# «Pseudo-HMM» framework (1/2)

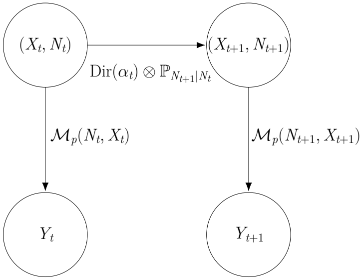
We assume that:

- Processes $(X_t)_{t\in\mathbb{Z}}$ and $(N_t)_{t\in\mathbb{Z}}$ are independent;
- $(N_t)_{t\in\mathbb{Z}}$ is a Markov chain;
- Conditionally on $(X_t, N_t)_{t\in\mathbb{Z}}$, the $Y_t$'s are independent, and depend only on $(X_t, N_t)$.

# «Pseudo-HMM» framework (1/2)

We assume that:

- Processes $(X_t)_{t\in\mathbb{Z}}$ and $(N_t)_{t\in\mathbb{Z}}$ are independent;
- $(N_t)_{t\in\mathbb{Z}}$ is a Markov chain;
- Conditionally on $(X_t, N_t)_{t\in\mathbb{Z}}$, the $Y_t$'s are independent, and depend only on $(X_t, N_t)$.




We thus obtain a «pseudo-HMM» framework, where the underlying process is partially observed.

# «Pseudo-HMM» framework (2/2)

Our framework allows us to consider $(X_t, Y_t)_{t\in\mathbb{Z}}$ as a Markov chain in a random environment, which is determined by $(Z_t, N_t)_{t\in\mathbb{Z}}$.

# «Pseudo-HMM» framework (2/2)

Our framework allows us to consider $(X_t, Y_t)_{t\in\mathbb{Z}}$ as a Markov chain in a random environment, which is determined by $(Z_t, N_t)_{t\in\mathbb{Z}}$.

**Proposition 2**

Assume that $(N_t, Z_t)_{t\in\mathbb{Z}}$ is stationary.

# «Pseudo-HMM» framework (2/2)

Our framework allows us to consider $(X_t, Y_t)_{t \in \mathbb{Z}}$ as a Markov chain in a random environment, which is determined by $(Z_t, N_t)_{t \in \mathbb{Z}}$.

**Proposition 2**

Assume that $(N_t, Z_t)_{t \in \mathbb{Z}}$ is stationary.

Then, there exists a stationary process $(X_t, Y_t, N_t, Z_t)_{t \in \mathbb{Z}}$ satisfying our dynamics, and its distribution is unique.

## «Pseudo-HMM» framework (2/2)

Our framework allows us to consider $(X_t, Y_t)_{t\in\mathbb{Z}}$ as a Markov chain in a random environment, which is determined by $(Z_t, N_t)_{t\in\mathbb{Z}}$.

**Proposition 2**

Assume that $(N_t, Z_t)_{t\in\mathbb{Z}}$ is stationary.

Then, there exists a stationary process $(X_t, Y_t, N_t, Z_t)_{t\in\mathbb{Z}}$ satisfying our dynamics, and its distribution is unique.

Moreover, if $(N_t, Z_t)_{t\in\mathbb{Z}}$ is ergodic, then $(X_t, Y_t, N_t, Z_t)_{t\in\mathbb{Z}}$ is also ergodic.

# Outlines

# Joint log-likelihood

Assume that for some $T \in \mathbb{N}^*, (y_0, n_0), \ldots, (y_T, n_T)$ are observed realizations of our process, corresponding to the unobserved relative abundances $x_0, \ldots, x_T$.

# Joint log-likelihood

Assume that for some $T \in \mathbb{N}^*$, $(y_0, n_0), \dots, (y_T, n_T)$ are observed realizations of our process, corresponding to the unobserved relative abundances $x_0, \dots, x_T$.

The joint log-likelihood of our process is

$$\mathcal{L}_\theta(x_{0:T}, y_{0:T}) = \sum_{t=1}^{T} \Bigg\{ \log(n_t!) + \log(\Gamma(\phi_t)) + \sum_{j=1}^{p} \left[ (\alpha_{j,t} + y_{j,t} - 1) \log(x_{j,t}) - \log(y_{j,t}!) - \log(\Gamma(\alpha_{j,t})) \right] \Bigg\} + \log(n_0!) + \sum_{j=1}^{p} \left[ y_{j,0} \log(x_{j,0}) - \log(y_{j,0}!) \right].$$

# Joint log-likelihood

Assume that for some $T \in \mathbb{N}^*$, $(y_0, n_0), \dots, (y_T, n_T)$ are observed realizations of our process, corresponding to the unobserved relative abundances $x_0, \dots, x_T$.

The joint log-likelihood of our process is

$$\mathcal{L}_\theta(x_{0:T}, y_{0:T}) = \sum_{t=1}^{T} \left\{ \log(n_t!) + \log(\Gamma(\phi_t)) + \sum_{j=1}^{p} \left[ (\alpha_{j,t} + y_{j,t} - 1)\log(x_{j,t}) - \log(y_{j,t}!) \right.\right.$$
$$\left.\left. - \log(\Gamma(\alpha_{j,t})) \right] \right\} + \log(n_0!) + \sum_{j=1}^{p} \left[ y_{j,0}\log(x_{j,0}) - \log(y_{j,0}!) \right].$$

We are interested in the estimation of the true parameter $\theta_0$ of our model, where

$$\theta = (a_0, a_1, A_0(1), \dots, A_0(p-1), A_1(1,1), \dots, A_1(p-1, p-1), B(1,1), \dots B(k, p-1)).$$

# Inference strategy (1/2)

➜ The EM algorithm is a classical manner of estimating $\theta_0$.

# Inference strategy (1/2)

- The EM algorithm is a classical manner of estimating $\theta_0$.
  - Initialize $\theta^{(0)}$ wisely.

# Inference strategy (1/2)

- The EM algorithm is a classical manner of estimating $\theta_0$.
  - Initialize $\theta^{(0)}$ wisely.
  - After $n$ iteration, update the estimate with

$$\theta^{(n+1)} = \underset{\theta}{\operatorname{argmax}} \underbrace{\mathbb{E}_{X_{0:T} \sim \mathbb{P}_{\theta^{(n)}}(\cdot|y_{0:T})} \left[\mathcal{L}_\theta \left(X_{0:T}, y_{0:T}\right)\right]}_{\mathcal{I}\left(\theta^{(n)},\theta\right)}.$$

# Inference strategy (1/2)

➜ The EM algorithm is a classical manner of estimating $\theta_0$.

- Initialize $\theta^{(0)}$ wisely.
- After $n$ iteration, update the estimate with

$$\theta^{(n+1)} = \underset{\theta}{\operatorname{argmax}} \underbrace{\mathbb{E}_{X_{0:T} \sim \mathbb{P}_{\theta^{(n)}}(\cdot|y_{0:T})}\left[\mathcal{L}_\theta\left(X_{0:T}, y_{0:T}\right)\right]}_{\mathcal{I}\left(\theta^{(n)},\theta\right)}.$$

✗ The quantity $\mathcal{I}\left(\theta^{(n)}, \theta\right)$ can not by computed directly.

# Inference strategy (2/2)

➜ Use of a particle filter to tackle this issue.

# Inference strategy (2/2)

➜ Use of a particle filter to tackle this issue.

The idea is to perform a large number $N$ of simulated trajectories $\widetilde{X}_{0:T}^{(1)}, \ldots, \widetilde{X}_{0:T}^{(N)}$, where for each $1 \leqslant i \leqslant N$

$$\widetilde{X}_{0:T}^{(i)} \sim \widetilde{\mathbb{P}}_{\theta^{(n)}}(\cdot \mid y_{0:T})$$

approaches the target distribution $\mathbb{P}_{\theta^{(n)}}(\cdot \mid y_{0:T})$.

# Inference strategy (2/2)

➜ Use of a particle filter to tackle this issue.

The idea is to perform a large number $N$ of simulated trajectories $\widetilde{X}_{0:T}^{(1)}, \dots, \widetilde{X}_{0:T}^{(N)}$, where for each $1 \leqslant i \leqslant N$

$$\widetilde{X}_{0:T}^{(i)} \sim \widetilde{\mathbb{P}}_{\theta^{(n)}}(\cdot \mid y_{0:T})$$

approaches the target distribution $\mathbb{P}_{\theta^{(n)}}(\cdot \mid y_{0:T})$.

Each simulation $\widetilde{X}_{0:T}^{(i)}$ is then weighted by $w_i$, which measures its fit with the observed data.

# Inference strategy (2/2)

➜ Use of a particle filter to tackle this issue.

The idea is to perform a large number $N$ of simulated trajectories $\widetilde{X}^{(1)}_{0:T}, \dots, \widetilde{X}^{(N)}_{0:T}$, where for each $1 \leqslant i \leqslant N$

$$\widetilde{X}^{(i)}_{0:T} \sim \widetilde{\mathbb{P}}_{\theta^{(n)}}(\cdot \mid y_{0:T})$$

approaches the target distribution $\mathbb{P}_{\theta^{(n)}}(\cdot \mid y_{0:T})$.

Each simulation $\widetilde{X}^{(i)}_{0:T}$ is then weighted by $w_i$, which measures its fit with the observed data.

We then approach $\mathcal{I}\left(\theta^{(n)}, \theta\right)$ by the mean

$$\sum_{i=1}^{N} w_i \mathcal{L}_\theta(\widetilde{X}^{(i)}_{0:T}, y_{0:T}).$$

# About the particle filter (1/2)

The proposal distribution chosen here is the one of a Markov chain $\left(\widetilde{X}_t\right)_{t\in\mathbb{Z}}$ with transition kernel

$$\widetilde{P}_{\theta^{(n)}}\left(\widetilde{X}_t \mid \widetilde{X}_{t-1}\right) = \mathrm{Dir}(\alpha_t^{(n)} + y_t).$$

# About the particle filter (1/2)

The proposal distribution chosen here is the one of a Markov chain $\left(\widetilde{X}_t\right)_{t\in\mathbb{Z}}$ with transition kernel

$$\widetilde{P}_{\theta^{(n)}}\left(\widetilde{X}_t \mid \widetilde{X}_{t-1}\right) = \mathrm{Dir}(\alpha_t^{(n)} + y_t).$$

Most of the weights are close to zero.

# About the particle filter (1/2)

The proposal distribution chosen here is the one of a Markov chain $\left(\widetilde{X}_t\right)_{t\in\mathbb{Z}}$ with transition kernel

$$\widetilde{P}_{\theta^{(n)}}\left(\widetilde{X}_t \mid \widetilde{X}_{t-1}\right) = \mathrm{Dir}(\alpha_t^{(n)} + y_t).$$

Most of the weights are close to zero.

➜ Resampling strategy:

# About the particle filter (1/2)

The proposal distribution chosen here is the one of a Markov chain $\left(\widetilde{X}_t\right)_{t\in\mathbb{Z}}$ with transition kernel

$$\widetilde{P}_{\theta^{(n)}}\left(\widetilde{X}_t \mid \widetilde{X}_{t-1}\right) = \mathrm{Dir}(\alpha_t^{(n)} + y_t).$$

Most of the weights are close to zero.

→ Resampling strategy:

- Each simulation $\widetilde{X}_{0:T}^{(i)}$ and weight $w_i$ are computed sequentially.

# About the particle filter (1/2)

The proposal distribution chosen here is the one of a Markov chain $\left(\widetilde{X}_t\right)_{t\in\mathbb{Z}}$ with transition kernel

$$\widetilde{P}_{\theta^{(n)}}\left(\widetilde{X}_t \mid \widetilde{X}_{t-1}\right) = \mathrm{Dir}(\alpha_t^{(n)} + y_t).$$

Most of the weights are close to zero.

Resampling strategy:

- Each simulation $\widetilde{X}_{0:T}^{(i)}$ and weight $w_i$ are computed sequentially.
- Once the **particles** $\widetilde{X}_t^{(1)}, \ldots, \widetilde{X}_t^{(N)}$ and the weights $w_{t,1}, \ldots, w_{t,N}$ are computed, we replace the particles by a new sample, drawn with replacement, from the same set of particles, with probabilities given by the particles weights.

# About the particle filter (2/2)

✔ We avoid the degeneracy of the weights.

# About the particle filter (2/2)

- ✔ We avoid the degeneracy of the weights.
- ⚠ The trajectories obtained are based on a small number of initial particles.

# About the particle filter (2/2)

✔ We avoid the degeneracy of the weights.

⚠ The trajectories obtained are based on a small number of initial particles.

💡 We perform backward smoothing in order to reduce this correlation.

# About the particle filter (2/2)

✔ We avoid the degeneracy of the weights.

⚠ The trajectories obtained are based on a small number of initial particles.

💡 We perform backward smoothing in order to reduce this correlation.

Given the weighted set of particles $\left\{\widetilde{X}_t^{(i)}, w_{t,i} : 0 \leqslant t \leqslant T, 1 \leqslant i \leqslant N\right\}$:

# About the particle filter (2/2)

✔ We avoid the degeneracy of the weights.

⚠ The trajectories obtained are based on a small number of initial particles.

We perform backward smoothing in order to reduce this correlation.

Given the weighted set of particles $\left\{\widetilde{X}_t^{(i)}, w_{t,i} : 0 \leqslant t \leqslant T, 1 \leqslant i \leqslant N\right\}$:

- Set $\xi_T = \widetilde{X}_T^{(i)}$ with probability $w_{i,T}$.

# About the particle filter (2/2)

✔ We avoid the degeneracy of the weights.

⚠ The trajectories obtained are based on a small number of initial particles.

We perform backward smoothing in order to reduce this correlation.

Given the weighted set of particles $\left\{\widetilde{X}_t^{(i)}, w_{t,i} : 0 \leqslant t \leqslant T, 1 \leqslant i \leqslant N\right\}$:

- Set $\xi_T = \widetilde{X}_T^{(i)}$ with probability $w_{i,T}$.
- For $t = T-1, \ldots, 0$:

# About the particle filter (2/2)

✔ We avoid the degeneracy of the weights.

⚠ The trajectories obtained are based on a small number of initial particles.

We perform backward smoothing in order to reduce this correlation.

Given the weighted set of particles $\left\{\widetilde{X}_t^{(i)}, w_{t,i} : 0 \leqslant t \leqslant T, 1 \leqslant i \leqslant N\right\}$:

- Set $\xi_T = \widetilde{X}_T^{(i)}$ with probability $w_{i,T}$.
- For $t = T-1, \ldots, 0$:
  - Compute new weights $w_{i,t|t+1} \propto w_{i,t} \times f_{\alpha_{t+1}(\widetilde{X}_t^{(i)})}(\xi_{t+1})$.

# About the particle filter (2/2)

✔ We avoid the degeneracy of the weights.

⚠ The trajectories obtained are based on a small number of initial particles.

We perform backward smoothing in order to reduce this correlation.

Given the weighted set of particles $\left\{\widetilde{X}_t^{(i)}, w_{t,i} : 0 \leqslant t \leqslant T, 1 \leqslant i \leqslant N\right\}$:

- Set $\xi_T = \widetilde{X}_T^{(i)}$ with probability $w_{i,T}$.
- For $t = T-1, \ldots, 0$:
  - Compute new weights $w_{i,t|t+1} \propto w_{i,t} \times f_{\alpha_{t+1}(\widetilde{X}_t^{(i)})}(\xi_{t+1})$.
  - Set $\xi_t = \widetilde{X}_t^{(i)}$ with probability $w_{i,t|t+1}$.

# About the particle filter (2/2)

✔ We avoid the degeneracy of the weights.

⚠ The trajectories obtained are based on a small number of initial particles.

💡 We perform backward smoothing in order to reduce this correlation.

Given the weighted set of particles $\left\{\widetilde{X}_t^{(i)}, w_{t,i} : 0 \leqslant t \leqslant T, 1 \leqslant i \leqslant N\right\}$:

- Set $\xi_T = \widetilde{X}_T^{(i)}$ with probability $w_{i,T}$.
- For $t = T-1, \ldots, 0$:
  - Compute new weights $w_{i,t|t+1} \propto w_{i,t} \times f_{\alpha_{t+1}(\widetilde{X}_t^{(i)})}(\xi_{t+1})$.
  - Set $\xi_t = \widetilde{X}_t^{(i)}$ with probability $w_{i,t|t+1}$.

The trajectory $\xi_{0:T}$ obtained is distributed with respect to $\mathbb{P}_{\theta^{(n)}}(\cdot \mid y_{0:T})$.

# Outlines

I. The problematic

II. The framework

III. Estimation procedure

IV. Numerical experiments

## Simulations (1/3)

We simulated $N = 100$ trajectories of relative abundances for two species, with one exogenous variable.

# Simulations (1/3)

We simulated $N = 100$ trajectories of relative abundances for two species, with one exogenous variable.

We then applied our estimation strategy to infer the model's parameters.

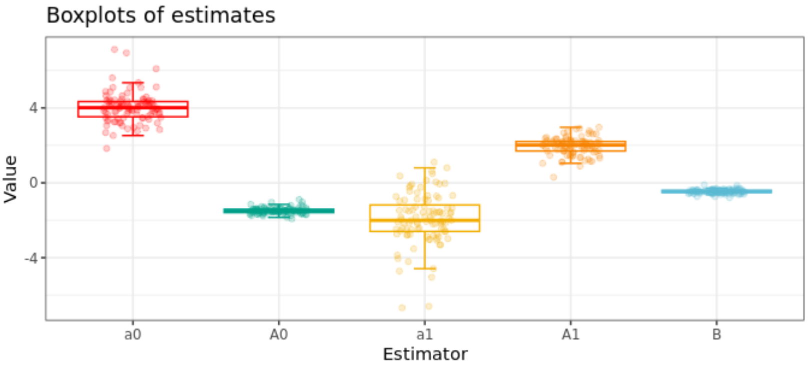
# Simulations (1/3)

We simulated $N = 100$ trajectories of relative abundances for two species, with one exogenous variable.

We then applied our estimation strategy to infer the model's parameters.

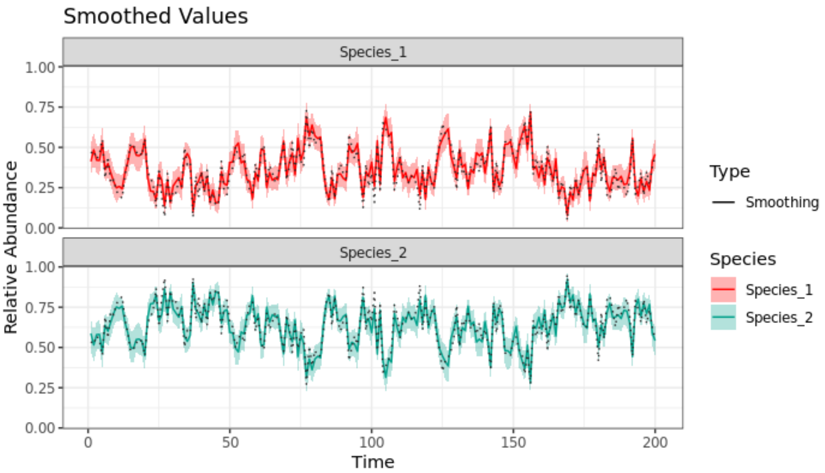
| Parameter | True Value | Mean Estimate | MSE | Bias | Variance |
|---|---|---|---|---|---|
| $a_0$ | 4 | 4.019 | 0.629 | 0.019 | 0.628 |
| $a_1$ | -2 | -1.932 | 1.872 | 0.068 | 1.868 |
| $A_0$ | -1.5 | -1.490 | 0.026 | 0.010 | 0.026 |
| $A_1$ | 2 | 1.959 | 0.220 | -0.041 | 0.219 |
| $B$ | -0.5 | -0.466 | 0.127 | 0.034 | 0.012 |

# Simulations (2/3)



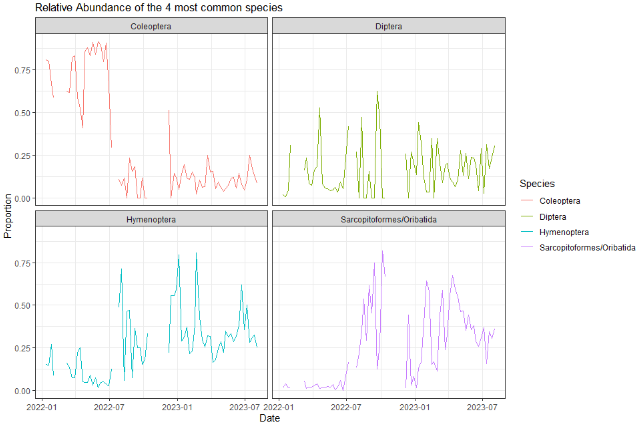
# Simulations (3/3)

Once we obtain estimates of the model's parameters, it is possible to use backward smoothing to recover the hidden process of relative abundance.



# Back in La Réunion (1/2)

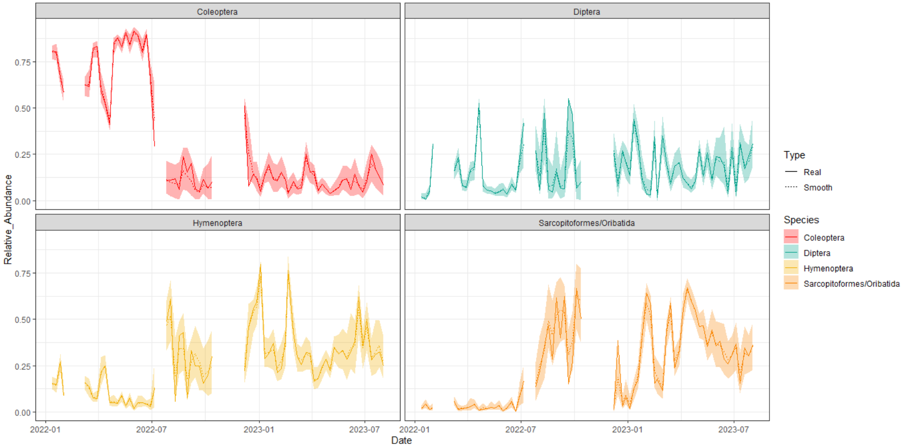
We finally fit our model to the relative abundance of the insects in a sugar-cane field in La Réunion.




We use the temperature and the amount of precipitation as exogenous variables.

# Back in La Réunion (2/2)

Once our estimation is complete, we use backward smoothing in order to recover the correct relative abundance of our species.



# Take away message

➜ An innovative manner to model the Joint Species Distribution in Ecology.

# Take away message

- An innovative manner to model the Joint Species Distribution in Ecology.
- A very interpretable model.

# Take away message

- An innovative manner to model the Joint Species Distribution in Ecology.
- A very interpretable model.
- With good statistical properties.

# Take away message

- ➜ An innovative manner to model the Joint Species Distribution in Ecology.
- ➜ A very interpretable model.
- ➜ With good statistical properties.
- ✘ Estimators have a large variance.

# Take away message

- ➜ An innovative manner to model the Joint Species Distribution in Ecology.
- ➜ A very interpretable model.
- ➜ With good statistical properties.
- ✗ Estimators have a large variance.
- ✗ Would probably fail to recover the hidden relative abundance if its too far from the observed one.

# Take away message

- ➜ An innovative manner to model the Joint Species Distribution in Ecology.
- ➜ A very interpretable model.
- ➜ With good statistical properties.
- ✘ Estimators have a large variance.
- ✘ Would probably fail to recover the hidden relative abundance if its too far from the observed one.
- Add a variable selection.

# Thank you !