Dynamical Modeling of Abundance Data in Ecology

Guillaume Franchi

February 14, 2022

1 Abundance Data

The study of abundance of species (plants or animals) is a central theme in ecology, as it is a measure of the wellness of a species in a given ecosystem. This abundance can be defined in two different ways.

Definition 1.1. The **standard abundance** of a species is the total number of individuals of this species.

The **relative abundance** of a species in an ecosystem is the proportion of individuals of the species among all individuals from all species.

Example 1.2. Consider the extremely childish example where we are located in the savanna, and our ecosystem is composed as follow.

Lions	Zebras	Buffalos	Hyenas	Total
12	36	48	24	120
10 %	30%	40%	20%	100%

The vector $(12, 36, 48, 24)$ is the standard abundance of our ecosystem, while the vector $(0.1, 0.3, 0.4, 0.2)$ is the relative abundance.

Remark 1.3. Relative abundance data are therefore compositional data, i.e. taking values in the simplex:

$$\mathcal{S}_{d-1} = \left\{ (y_1, \ldots, y_d) \in]0; 1[^d \mid \sum_{i=1}^{d} y_i = 1 \right\}. \tag{1}$$

The species standard abundance distribution (see McGill et al. (2007) for details) has been particularly studied in the literature: it describes the commonness and rarity of species in an ecosystem.
In this case, count data are considered, and distributions can be modeled for example with Poisson models.

Relative abundance is however an attractive notion since it could be a good measure of the **biodiversity** of an ecosystem. In this case, probability distributions on the simplex can be considered, such as the Dirichlet distribution $\mu_\alpha^{(d)}$ on $\mathbb{R}^d$ with density:

$$f(y_1, \ldots, y_d) = \frac{\Gamma\left(\sum_{i=1}^{d} \alpha_i\right)}{\Gamma(\alpha_1) \ldots \Gamma(\alpha_d)} \prod_{i=1}^{d} y_i^{\alpha_i - 1} \tag{2}$$

for $(y_1, \ldots, y_d) \in \mathcal{S}_{d-1}$.

In all that follows, we will focus on relative abundance.

The abundance of a species can be explained by several exogenous variables. One can mention climatic variables such as the temperature or the precipitation. This leads to regression problems for compositional data, where methods seen in the first lesson can be applied.

However, the abundance of one species is also affected by the abundance of other species, since there always exists a competition between them. Therefore, one can immediately guess that the abundance of species in a given ecosystem is a dynamic process, varying along the time.

Here we will propose time series models for abundance data, in order to explain the dynamic of abundance.
For simplification purpose, we will not consider in our models the impact of exogenous variables on the abundance, but only the impact of species between them.

Just like regression models for the static case, we will present two approaches: data transformation and «Stay in the simplex» approach.

2 General framework

In all that follows, we will study the relative abundance along time $t \in \mathbb{Z}$ of $d \geqslant 2$ studies in a given ecosystem, modeled by a sequence $Y^{(t)} = \left(Y_1^{(t)}, \dots, Y_d^{(t)} \right)$ of random variables, valued in $\mathcal{S}_{d-1}$.

Remark 2.1. If there is a special dynamic for the abundance, the random variables $\left(Y^{(t)} \right)_{t \in \mathbb{Z}}$ are not independent.

We will assume that we have in our possession a sample $\left(y^{(t)} \right)_{0 \leqslant t \leqslant N} \in \mathcal{S}_{d-1}^N$ of this abundance, where $y^{(t)}$ is a realization of the random variable $Y^{(t)}$.

The following example is derived from a study by Svensson (2006).

Example 2.2 (Scandinavian birds)**.** Let's consider an ecosystem composed by three alpine birds species: «Anthus pratensis», «Calcarius lapponicus» and «Oenanthe oenanthe». The relative abundance of this ecosystem has been registered from 1964 to 2001 and can be found in Table 2.

Considering that time $t = 0$ corresponds to the year 1964, we have a sample of $N = 38$ abundances that can be written $y^{(t)} = \left(y_1^{(t)}, y_2^{(t)}, y_3^{(t)} \right)$, $0 \leqslant t \leqslant N$, and a graphical representation of the time series is presented in Figure 1.
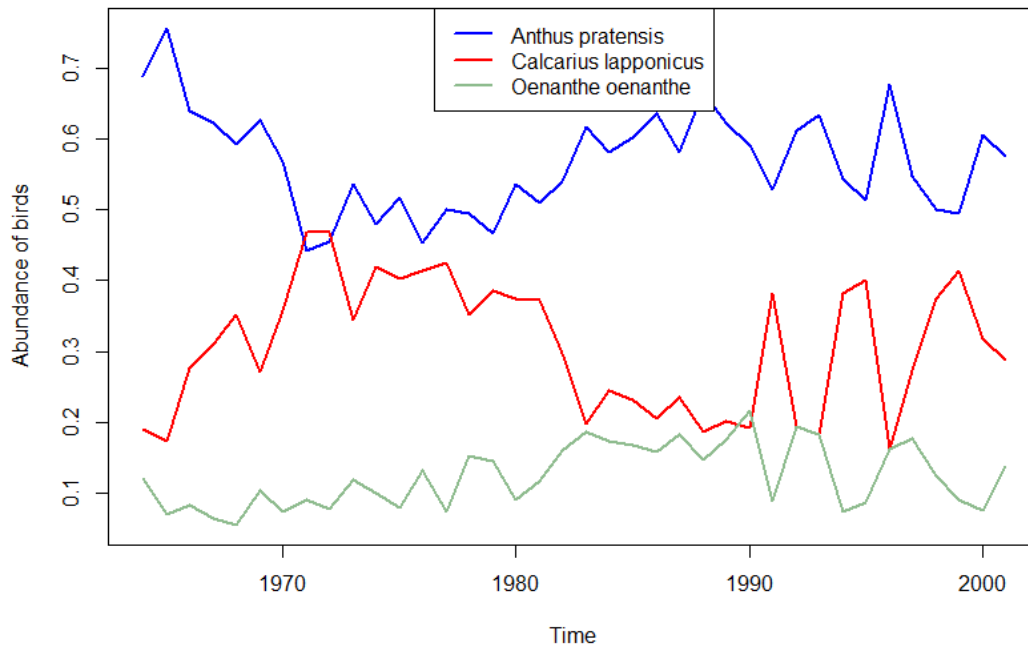

Figure 1 – Graphical Representation of the relative abundance of Scandinavian birds along time.

3 Data transformation

As in the static case, the idea is to use a one-to-one mapping

$$g : \mathcal{S}_{d-1} \longrightarrow \mathbb{R}^k$$

for some k.
We can take for example the **additive log-ratio**:

$$alr : \ \mathcal{S}_{d-1} \longrightarrow \mathbb{R}^{d-1} \tag{3}$$
$$y = (y_1, \dots, y_d) \longmapsto z = \left(\log \left(\frac{y_1}{y_d} \right) \dots, \log \left(\frac{y_{d-1}}{y_d} \right) \right).$$

We thus transform our abundance time series $\left(Y^{(t)} \right)_t$ valued in the simplex into a time series $\left(Z^{(t)} \right)_t$ valued in $\mathbb{R}^{d-1}$ with:

$$\forall t \in \mathbb{Z}, \ Z^{(t)} = alr \left(Y^{(t)} \right). \tag{4}$$

The process $(Z^{(t)})$ is constraint-free, so we can use any classical time series method to model this latter process. Finally, once we have obtained a fitted time series $(\widehat{Z^{(t)}})_t$ on the transformed scale, it is possible to get a fitted time series $(\widehat{Y^{(t)}})_t$ on the original scale with:

$$\forall t \in \mathbb{Z}, \ \widehat{Y^{(t)}} = alr^{-1}\left(\widehat{Z^{(t)}}\right). \tag{5}$$

Remark 3.1. Let us recall that we have:

$$
\begin{aligned}
alr^{-1}: \ & \mathcal{S}_{d-1} \longrightarrow \mathbb{R}^{d-1} \\
& z = (z_1, \ldots, z_{d-1}) \longmapsto y = \left(\frac{\exp(z_1)}{1 + \sum_{j=1}^{d-1}\exp(z_j)}, \ldots, \frac{\exp(z_{d-1})}{1 + \sum_{j=1}^{d-1}\exp(z_j)}, \frac{1}{1 + \sum_{j=1}^{d-1}\exp(z_j)}\right).
\end{aligned} \tag{6}
$$

One of the easiest model for the time series $(Z^{(t)})$ is the **VAR**(p) model, with $p \in \mathbb{N}^*$:

$$\forall t \in \mathbb{Z}, \ Z^{(t)} = c + \sum_{i=1}^{p} \phi_i \cdot Z^{(t-i)} + \varepsilon_t \tag{7}$$

where $\phi_1, \ldots, \phi_p$ are $(d-1) \times (d-1)$ matrices, $c \in \mathbb{R}^{d-1}$ and $(\varepsilon_t)_{t \in \mathbb{Z}}$ is a gaussian white noise in $\mathbb{R}^{d-1}$.

Remark 3.2. Recall that a gaussian white noise $(\varepsilon_t)_t$ is a process of centered, independent and identically distributed gaussian variables.

Proposition 3.3. *For $z \in \mathbb{C}$, let us denote $\phi(z)$ the complex matrix given by:*

$$\phi(z) = I_{d-1} - \sum_{j=1}^{p} z^j \phi_j. \tag{8}$$

If all the solutions of the equation:

$$\det(\phi(z)) = 0$$

have an absolute value strictly larger than 1, then there exists a unique stationary process satisfying equation (7).

Example 3.4 (Scandinavian Birds)**.** Applying the *alr* transformation to the time series of Scandinavian birds given in Table 2, we obtain the time series given in Figure 2 .
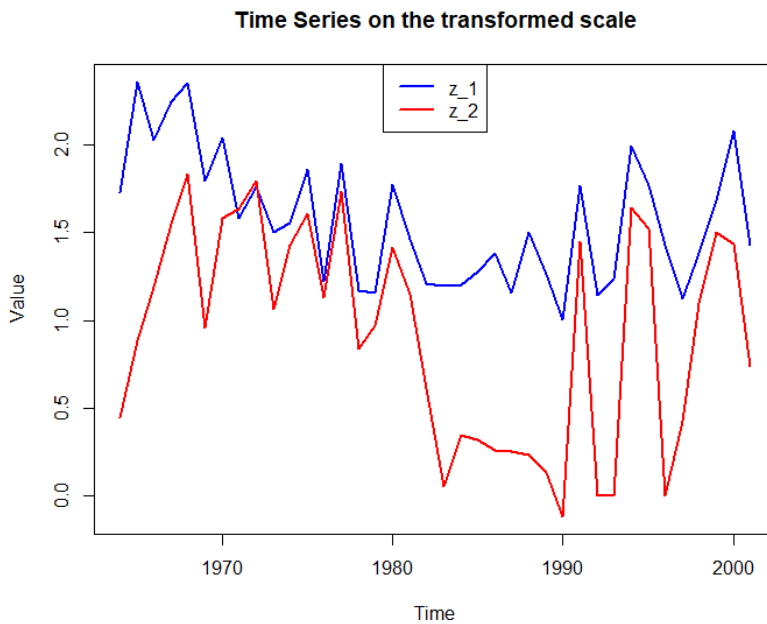

Figure 2 – *alr* transformation of the Scandinavian birds time series.

A visual inspection of the partial autocorrelograms (see Figure 3) leads us to use a VAR(3) model for the time series $\left(Z^{(t)}\right)_t$.
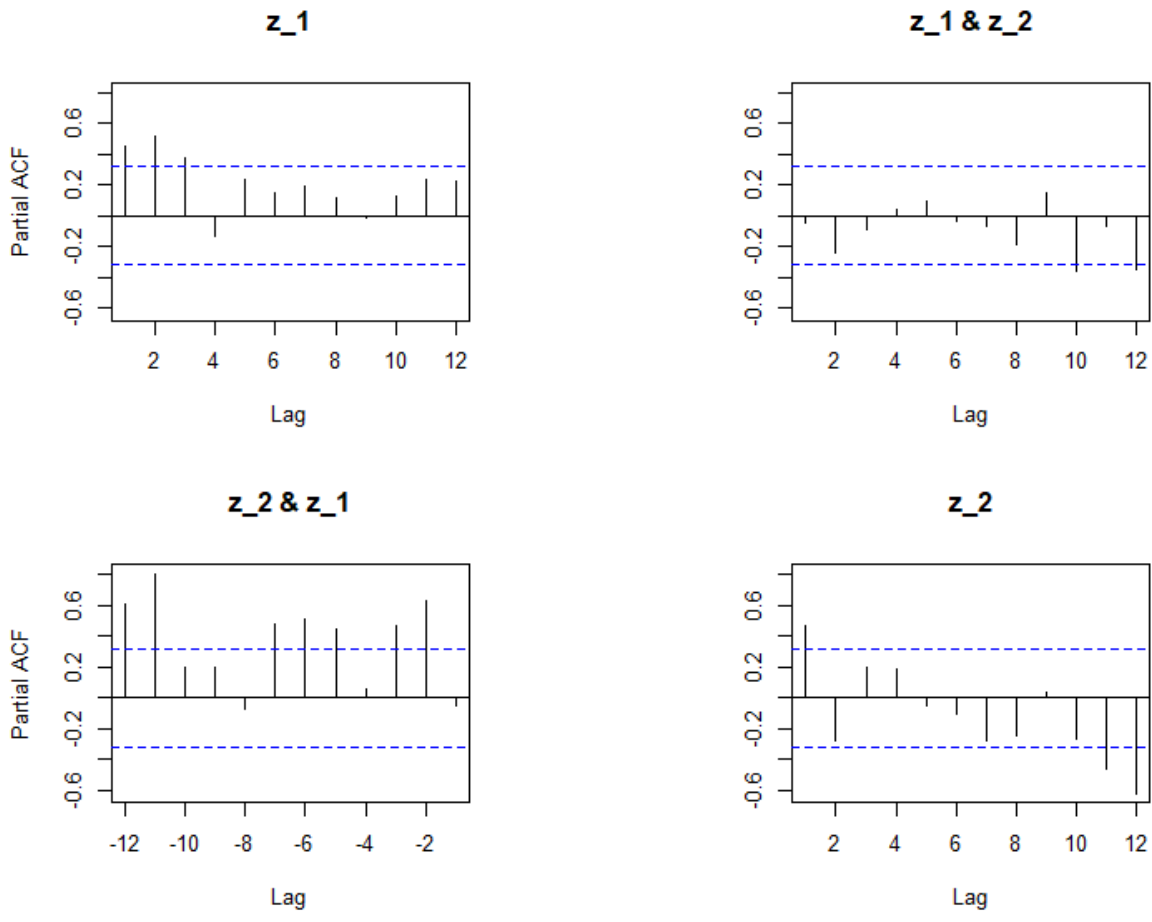

Figure 3 – Partial autocorrelograms of the *alr*-transformed time series for the Scandinavian birds.

The estimates of the model parameters will not be given here, but allow us to compute fitted values on the transform scale.

Finally, a backtransformation of those fitted values gives us fitted values on the original scale (see Figure 4).
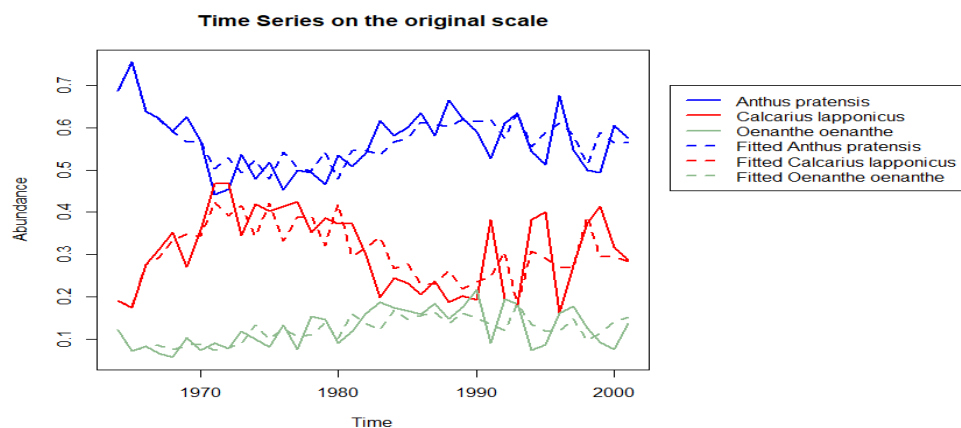

Figure 4 – Scandinavian birds time series and their fitted values with a VAR(3) model.

4 The «Stay in the simplex approach»

Let us recall first some basic ideas of the Dirichlet regression for compositional data.

The response variable Y follows the Dirichlet distribution of mean $\mu = (\mu_1, \ldots, \mu_d)$ and dispersion parameter ϕ with for all $1 \leqslant i \leqslant d-1$:

$$\mu_i = \frac{\exp\left(\beta^{(i)} \cdot X\right)}{1 + \sum_{j=1}^{d-1} \exp\left(\beta^{(j)} \cdot X\right)} \tag{9}$$

where X is the vector of explanatory variables.

Taking into account the dynamic of the process leads to consider naturally the past values of the process as explanatory variables themselves. Because we don't consider exogenous variables in our model, we thus suggest that for all $t \in \mathbb{Z}$, $Y^{(t)}$ follows the Dirichlet distribution with mean $\mu_t = (\mu_{t,1}, \ldots, \mu_{t,d})$ satisfying:

$$\forall 1 \leqslant i \leqslant d-1, \; \mu_{t,i} = \frac{\exp\left(\beta^{(i)} \cdot Y^{(t-1)}\right)}{1 + \sum_{j=1}^{d-1} \exp\left(\beta^{(j)} \cdot Y^{(t-1)}\right)} \tag{10}$$

and constant *(for simplification purpose)* dispersion parameter ϕ.

Remark 4.1. In equation (10), observe that we only took the previous value $Y^{(t-1)}$ as covariate. It is of course possible to take several lag-values $Y^{(t-2)}, Y^{(t-3)}, \ldots$ in this expression.
In the following, we just take one lag-value in order to emphasize the main idea of the model.

4.1 Some properties about Markov Chains

Definition 4.2 (Transition Kernel). Let us consider $(E, \mathcal{E})$ and $(F, \mathcal{F})$ two measurable spaces.
An application $K : E \times \mathcal{F} \longrightarrow [0;1]$ is a **transition kernel** on $E \times \mathcal{F}$ if:

- for all $x \in E$ the application $K(x, \cdot)$ is a probability measure;

- fo rall $B \in \mathcal{F}$, the application $K(\cdot, B)$ is $\mathcal{E}$-measurable.

Definition 4.3 (Markov chain). Consider a stochastic process $\left(Y^{(t)}\right)_{t \in \mathbb{Z}}$ valued in a measurable space $(E, \mathcal{E})$, and denote for all $t \in \mathbb{Z}$:

$$\mathcal{F}_t^- = \sigma\left(Y^{(t-1)}, Y^{(t-2)}, \ldots\right) \tag{11}$$

and:

$$\mathcal{F}_t = \sigma\left(Y^{(t)}\right). \tag{12}$$

The process $Y^{(t)}$ is a **Markov chain** if for any measurable bounded function $f : E \longrightarrow \mathbb{R}$:

$$\mathbb{E}\left(f\left(Y^{(t)}\right) \mid \mathcal{F}_{t-1}^-\right) = \mathbb{E}\left(f\left(Y^{(t)}\right) \mid \mathcal{F}_{t-1}\right). \tag{13}$$

Remark 4.4. Intuitively, this means that knowing the entire history of the process does not bring more information than knowing the last value.

Definition 4.5. Let K be a transition kernel on $E \times \mathcal{E}$.
A Markov chain $\left(Y^{(t)}\right)_{t \in \mathbb{Z}}$ valued in $(E, \mathcal{E}$ has transition kernel K if for all $t \in \mathbb{Z}$ and all $A \in \mathcal{E}$:

$$\mathbb{P}\left(Y^{(t)} \in A \mid Y_{t-1} = y_{t-1}\right) = K(y_{t-1}, A). \tag{14}$$

In this case, the chain is said to be homogeneous.

4.2 Some properties about ergodicity

Definition 4.6. Consider a strongly stationary process $\left(Y^{(t)}\right)_{t \in \mathbb{Z}}$ valued in a measurable space $(E, \mathcal{E})$, of law $\mathbb{P}_Y$.

Let us define the shift operator $\sigma : E^{\mathbb{Z}} \longrightarrow E^{\mathbb{Z}}$ by $\tau\left((y_t)_{t \in \mathbb{Z}}\right) = (y_{t+1})_{t \in \mathbb{Z}}$.

The process $\left(Y^{(t)}\right)_{t \in \mathbb{Z}}$ is **ergodic** if for all τ-invariant event $A \subset E^{\mathbb{Z}}$:

$$\mathbb{P}_Y(A) = 0 \text{ or } 1. \tag{15}$$

Definition 4.7. Let K be a transition kernel on a measurable space $(E, \mathcal{E})$. A probability π on E is said to be $K-$**invariant** if:

$$\pi \cdot K = \pi \tag{16}$$

where:

$$\pi \cdot K(A) = \int K(y, A)\pi(\mathrm{d}y) \tag{17}$$

Theorem 4.8. *If a transition kernel K admits a unique invariant measure π, then there exists a unique strongly stationary Markov chain $\left(Y^{(t)}\right)_t$ with transition kernel K, and it is ergodic.*

4.3 Back to the model

We thus decide to model our abundance times series $\left(Y^{(t)}\right)_{t \in \mathbb{Z}}$ as a Markov chain with Dirichlet transition kernel K with mean μ satisfying equation (10), and unknown dispersion parameter ϕ.

It can be shown that our kernel K satisfies the Doeblin's condition:

Definition 4.9. A transition kernel K defined on $E \times \mathcal{E}$ satisfies the **Doeblin's condition** if there exists a constant $\eta > 0$ and a probability measure λ on E such that:

$$\forall A \in \mathcal{E}, \ \forall y \in E, \ K(y, A) \geqslant \eta \cdot \lambda(A). \tag{18}$$

Lemma 4.10. *Define for all $i \in \{1, \ldots, d-1\}$ and all $y \in \mathcal{S}_{d-1}$:*

$$\mu_i(y) = \frac{\exp\left(\beta^{(i)} \cdot y\right)}{1 + \sum_{j=1}^{d-1} \exp\left(\beta^{(j)} \cdot y\right)} \tag{19}$$

and:

$$\mu_d(y) = \frac{1}{1 + \sum_{j=1}^{d-1} \exp\left(\beta^{(j)} \cdot y\right)}. \tag{20}$$

For all $i \in \{1, \ldots, d\}$, there exists $m_i > 0$ and $M_i > 0$ such that :

$$\forall y \in \mathcal{S}_{d-1}, \ m_i \leqslant \phi\mu_i(y) \leqslant M_i.$$

Proof.

▶ Since for all $i \in \{1, \ldots, d\}$ and all $y \in \mathcal{S}_{d-1}$, $\mu_i(y) \leqslant 1$, we can take $M_i = \phi$.

▶ For all $j \in \{1, \ldots d-1\}$ and all $y \in \mathcal{S}_{d-1}$, we have :

$$\exp\left(\beta^{(j)} \cdot y\right) \leqslant \exp\left(\|\beta^{(j)}\|_1\right).$$

Furthermore, for all $i \in \{1, \ldots d-1\}$ and $y \in \mathcal{S}_{d-1}$, the Cauchy-Schwarz inequality provides :

$$-\|\beta^{(i)}\|_2 \cdot \|y\|_2 \leqslant \beta^{(i)} \cdot y \leqslant \|\beta^{(i)}\|_2 \cdot \|y\|_2.$$

By equivalence of the norms, there exists $K > 0$ independent from y and i such that :

$$-\|\beta^{(i)}\|_2 \cdot \|y\|_2 \geqslant -K\|\beta^{(i)}\|_2 \cdot \|y\|_1 = -K\|\beta^{(i)}\|_2.$$

Hence, for all $i \in \{1, \ldots, d-1\}$ and all $y \in \mathcal{S}_{d-1}$:

$$\phi\mu_i(y)) \geqslant \frac{\phi \exp(-K\|\beta^{(i)}\|_2)}{1 + \sum_{j=1}^{d-1} \exp(\|\beta^{(j)}\|_1)} := m_i.$$

Finally, we naturally have :

$$\phi\mu_d(y) \geqslant \frac{\phi}{1 + \sum_{j=1}^{d-1} \exp(\|\beta^{(j)}\|_1)} := m_d.$$

$\blacksquare$

Proposition 4.11. *The transition kernel K defined by (10) satisfies the Doeblin's condition (18).*

Proof. According to Lemma 4.10, for all $i \in \{1, \dots, d\}$ there exist $m_i, M_i > 0$ such that:

$$\forall y \in \mathcal{S}_{d-1}, \ m_i \leqslant \phi\mu_i(y) \leqslant M_i.$$

Thus, for all $i \in \{1, \dots, d\}$ and all $y \in \mathcal{S}_{d-1}$:

$$
\begin{aligned}
\Gamma(\phi\mu_i(y)) &= \int_0^1 \mathrm{e}^{-t} \cdot t^{\phi\mu_i(y)-1}\mathrm{d}t + \int_1^{+\infty} \mathrm{e}^{-t} \cdot t^{\phi\mu_i(y)-1}\mathrm{d}t \\
&\leqslant \int_0^1 \mathrm{e}^{-t} \cdot t^{m_i-1}\mathrm{d}t + \int_1^{+\infty} \mathrm{e}^{-t} \cdot t^{M_i-1}\mathrm{d}t \\
&\leqslant \Gamma(m_i) + \Gamma(M_i) := \zeta_i.
\end{aligned}
$$

As a result, for all $y \in \mathcal{S}_{d-1}$ and all $A \in \mathcal{B}(\mathcal{S}_{d-1})$, we have:

$$
\begin{aligned}
K(y, A) &= \int \mathbb{1}_A \left(x_1, \dots, x_{d-1}, 1 - \sum_{i=1}^{d-1} x_i \right) \frac{\Gamma(\phi)}{\Gamma(\phi\mu_1(y)) \dots \Gamma(\phi\mu_d(y))} \\
&\quad \times \left(1 - \sum_{i=1}^{d-1} x_i \right)^{\phi\mu_d(y)-1} \prod_{i=1}^{d-1} x_i^{\phi\mu_i(y)-1}\mathrm{d}x_i \\
&\geqslant \int \mathbb{1}_A \left(x_1, \dots, x_{d-1}, 1 - \sum_{i=1}^{d-1} x_i \right) \frac{\Gamma(\phi)}{\zeta_1 \dots \zeta_d} \left(1 - \sum_{i=1}^{d-1} x_i \right)^{M_d-1} \prod_{i=1}^{d-1} x_i^{M_i-1}\mathrm{d}x_i \\
&\geqslant \eta \times \lambda(A)
\end{aligned}
$$

with $\eta = \dfrac{\Gamma(\phi) \times \Gamma(M_1) \dots \Gamma(M_d)}{\zeta_1 \dots \zeta_d \times \Gamma\left(\sum_{i=1}^d M_i \right)}$ and $\lambda = \mathrm{Dir}(M_1, \dots, M_d)$. ∎

The following proposition is then crucial:

Proposition 4.12. *If K satisfies the Doeblin's condition (18), then it admits a unique invariant probability measure π.*

Proof. We actually prove that this unique invariant probability measure satisfies:

$$\forall y \in E, \ d_{TV}(\delta_y K, \pi) \xrightarrow[n \to +\infty]{} 0. \tag{21}$$

▶ Step 1: Assume there exists $\pi \in \mathcal{M}_1(\mathcal{E})$ such that:

$$\forall y \in E, \ d_{TV}(\delta_y K^n, \pi) \xrightarrow[n \to +\infty]{} 0.$$

Considering the limit in total variation, we have:

$$\pi = \lim_{n \to +\infty} \delta_y K^n$$

thus:

$$\pi K = \lim_{n \to +\infty} \delta_y K^{n+1} = \pi.$$

The probability measure π is invariant for K.

▶ Step 2: Assume there exists another invariant measure π' for K:

$$\pi' K = \pi',$$

thus for all $n \in \mathbb{N}$, $\pi' = \pi' K^n$. Yet, we have in total variation distance $\lim_{n \to +\infty} \pi' K^n = \pi$. Indeed, for all $A \in \mathcal{E}$:

$$
\begin{aligned}
|\pi' P^n(A) - \pi(A)| &= \left| \int_E K^n(y, A)\pi'(\mathrm{d}y) - \pi(A) \right| \\
&= \left| \int_E K^n(y, A) - \pi(A)\pi'(\mathrm{d}y) \right| \\
&\leqslant \sup_{y \in E} |K^n(y, A) - \pi(A)| \\
&\leqslant \sup_{y \in E} |\delta_y K^n(A) - \pi(A)| \xrightarrow[n \to +\infty]{} 0.
\end{aligned}
$$

Hence, we have $\pi' = \pi$ and the uniqueness of an invariant measure for K.

▶ Step 3: Let us show at least that there exists a probability measure $\pi \in \mathcal{M}_1(\mathcal{E})$ such that :

$$\sup_{y \in E} d_{TV}(\delta_y K^n, \pi) \underset{n \to +\infty}{\longrightarrow} 0$$

which will give us the desired result accorded to step 1.

Since K satisfies the Doeblin's condition, there exists $\eta > 0$ and $\lambda \in \mathcal{M}_1(\mathcal{E})$ such that:

$$\forall Y \in E, \forall A \in \mathcal{E}, \ K(y, A) \geqslant \eta \cdot \lambda(A).$$

Dobrushin's coefficient for P satisfies then:

$$c(K) = \sup_{y \neq z} d_{TV}(\delta_y K, \delta_z K) \leqslant 1 - \eta < 1.$$

Indeed, for $y \neq z$ and all $A \in \mathcal{E}$, we have :

$$\begin{aligned}
\delta_y K(A) - \delta_z K(A) &= K(y, A) - K(z, A) \\
&= 1 - K(y, A^c) - K(z, A) \\
&\leqslant 1 - \eta \lambda(A^c) - \eta \lambda(A) \\
&\leqslant 1 - \eta.
\end{aligned}$$

In a similar way, we have $\delta_z K(A) - \delta_y K(A) \leqslant 1 - \eta$ and so:

$$|\delta_y K(A) - \delta_z K(A)| \leqslant 1 - \eta.$$

As a result, $d_{TV}(\delta_y K, \delta_z K) \leqslant 1 - \eta$ and $c(K) \leqslant 1 - \eta < 1$. Furthermore, using the fact that:

$$\forall \nu_1, \nu_2 \in \mathcal{M}_1(\mathcal{E}), \ d_{TV}(\nu_1 K, \nu_2 K) \leqslant c(K) d_{TV}(\nu_1, \nu_2)$$

we have for all $y \in E$:

$$\forall n \in \mathbb{N}^*, \ d_{TV}(\delta_y K^n, \delta_y K^{n+1}) \leqslant c(K) d_{TV}(\delta_y K^{n-1}, \delta_y K^n)$$

and by immediate induction over n :

$$\forall n \in \mathbb{N}, \forall k \in \mathbb{N}, \ d_{TV}(\delta_y K^n, \delta_y K^{n+k}) \leqslant c(K)^n d_{TV}(\delta_y, \delta_y K^k) \leqslant 2 \times c(K)^n \underset{n \to +\infty}{\longrightarrow} 0.$$

For all $y \in E$, the sequence $(\delta_y K^n)_{n \in \mathbb{N}}$ is a Cauchy sequence in $(\mathcal{M}_1(\mathcal{E}), d_{TV})$ which is complete, thus:

$$\forall y \in E, \exists \pi_y \in \mathcal{M}_1(\mathcal{E}), \ d_{TV}(\delta_y K^n, \pi_y) \underset{n \to +\infty}{\longrightarrow} 0.$$

Let $y, z \in E$, we have for all $n \in \mathbb{N}$:

$$d_{TV}(\pi_y, \pi_z) \leqslant \underbrace{d_{TV}(\pi_y, \delta_z K^n)}_{\underset{n \to +\infty}{\longrightarrow} 0} + d_{TV}(\delta_y K^n, \delta_z K^n) + \underbrace{d_{TV}(\delta_z K^n, \pi_z)}_{\underset{n \to +\infty}{\longrightarrow} 0}.$$

Yet, for all $n \in \mathbb{N}$, we have:

$$d_{TV}(\delta_y K^n, \delta_z K^n) \leqslant c(K)^n(\delta_y, \delta_z) \leqslant c(K)^n \underset{n \to +\infty}{\longrightarrow} 0.$$

As a result, $d_{TV}(\pi_y, \pi_z) = 0$ and $\pi_y = \pi_z$. Hence, there exists $\pi \in \mathcal{M}_1(\mathcal{E})$ such that:

$$\sup_{y \in E} d_{TV}(\delta_y K^n, \pi) \underset{n \to +\infty}{\longrightarrow} 0$$

which concludes the proof.

∎

Corollary 4.13. *There exists a unique strongly stationary Markov chain $\left(Y^{(t)}\right)_t$ satisfying (10), and it is ergodic.*

Thus, there exists a unique strongly stationary abundance time series satisfying (10), and it is ergodic.

In the following, we assume that our abundance time series is indeed this particular time series.

The model parameters ϕ and $\beta^{(1)}, \ldots, \beta^{(d-1)}$ can be estimated by the maximum of conditional likelihood, and ergodicity is an essential point here, because it ensures the strong consistency of these estimates, as well as their asymptotic normality *under some reasonable additional assumptions).*

Example 4.14 (Scandinavian Birds)**.** If we try to apply our model to the Scandinavian Birds data of Table 2, we obtain the following results:

	$\hat{\phi}$	$\beta_1^{(1)}$	$\beta_2^{(1)}$	$\beta_3^{(1)}$	$\beta_1^{(2)}$	$\beta_2^{(2)}$	$\beta_3^{(2)}$
Estimates	171.61	3.70	1.16	-7.20	1.10	3.91	-7.32

Table 1 – Estimates in the Dirichlet model.

In Figure 5, we present the fitted values of abundance according to our model. Note that here, the previsions are made step by step, i.e. the future value of the abundance $\widehat{Y^{(t+1)}}$ is calculated according to the real abundance observed $Y^{(t)}$.
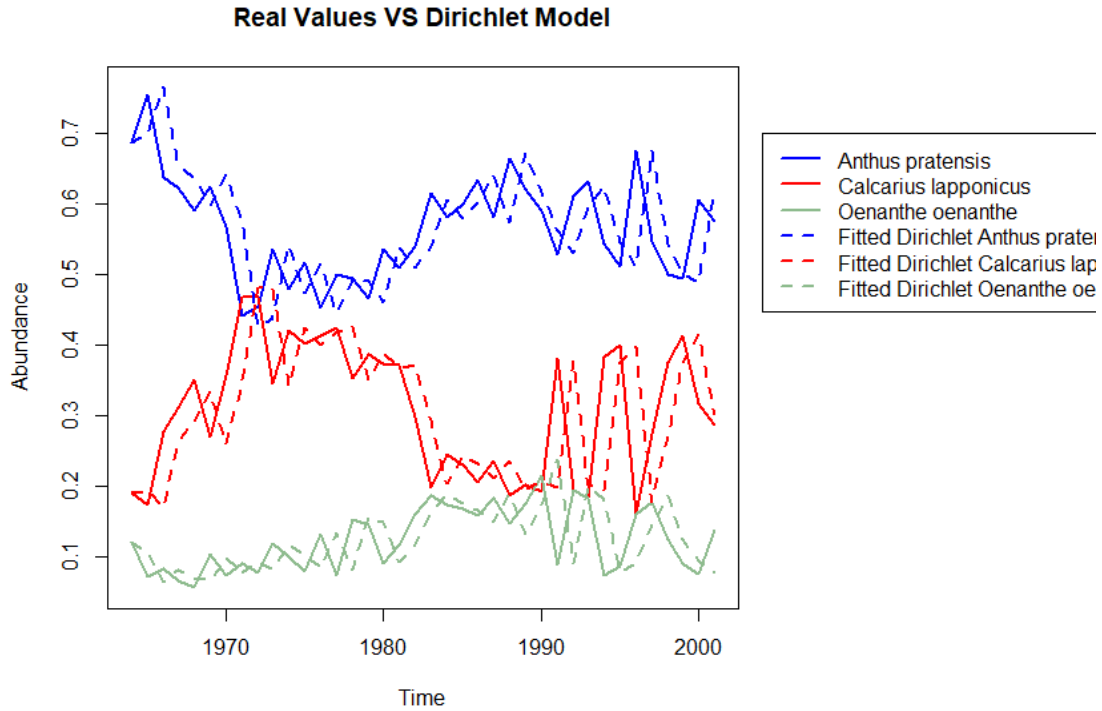

Figure 5 – Fitted Values of abundance of Scandinavian Birds with the Dirichlet model.

What happens here is that there is actually not a lot of variation predicted for the future value, in comparison to the present value *(i.e. the one observed)*. The value of $\hat{\phi}$ is indeed quite high *($\hat{\phi} \approx 171.61$)*. This explains the fact that we have the impression that the abundance predicted is the same as the observed one, but with a lag 1.

However, the observed time series for abundance presents a lot of variability, so there is a big suspicion on the estimates $\hat{\phi}$. Recall that our estimators are asymptotically consistent, and we only had 38 values to compute our estimates.

One final question of interest is the interpretation of the parameters $\beta^{(i)}$ in equation (10).

Remark 4.15. Take again Example 4.14.
Assume that between t and $t+1$, the abundance is evolving according to the following equation:

$$y^{(t+1)} = y^{(t)} + (c, -c, 0) \tag{22}$$

i.e. the proportion of the first species increases by c, at the expense of the second one.

Consider the **power balance** between the two species defined by:

$$\frac{\mu_{1,t}}{\mu_{2,t}}. \tag{23}$$

One can easily show that this balance evolves according to:

$$\frac{\frac{\mu_{1,t+1}}{\mu_{2,t+1}}}{\frac{\mu_{1,t}}{\mu_{2,t}}} = \exp\left(c \cdot (\beta_1^{(1)} + \beta_2^{(2)} - \beta_2^{(1)} - \beta_1^{(2)})\right). \tag{24}$$

9

Once again, just like in the static case, the parameters of the model can be interpreted in terms of power balance.

5 Appendices

5.1 Scandinavian Birds Data

YEAR	Anthus pratensis	Calcarius lapponicus	Oenanthe oenanthe
1964	0.69	0.19	0.12
1965	0.76	0.17	0.07
1966	0.64	0.28	0.08
1967	0.62	0.31	0.07
1968	0.59	0.35	0.06
1969	0.62	0.27	0.10
1970	0.57	0.36	0.07
1971	0.44	0.47	0.09
1972	0.45	0.47	0.08
1973	0.54	0.35	0.12
1974	0.48	0.42	0.10
1975	0.52	0.40	0.08
1976	0.45	0.41	0.13
1977	0.50	0.42	0.07
1978	0.49	0.35	0.15
1979	0.47	0.39	0.15
1980	0.54	0.37	0.09
1981	0.51	0.37	0.12
1982	0.54	0.30	0.16
1983	0.62	0.20	0.19
1984	0.58	0.25	0.17
1985	0.60	0.23	0.17
1986	0.63	0.21	0.16
1987	0.58	0.24	0.18
1988	0.66	0.19	0.15
1989	0.62	0.20	0.18
1990	0.59	0.19	0.22
1991	0.53	0.38	0.09
1992	0.61	0.19	0.19
1993	0.63	0.18	0.18
1994	0.54	0.38	0.07
1995	0.51	0.40	0.09
1996	0.68	0.16	0.16
1997	0.55	0.27	0.18
1998	0.50	0.38	0.12
1999	0.49	0.41	0.09
2000	0.61	0.32	0.08
2001	0.57	0.29	0.14

Table 2 – Table of relative abundance for the example of Scandinavian birds

References

McGill, B. J., Etienne, R. S., Gray, J. S., Alonso, D., Anderson, M. J., Benecha, H. K., Dornelas, M., Enquist, B. J., Green, J. L., He, F., et al. (2007). Species abundance distributions: Moving beyond single prediction theories to integration within an ecological framework. *Ecology letters*, *10*(10), 995–1015.

Svensson, S. (2006). Species composition and population fluctuations of alpine bird communities during 38 years in the scandinavian mountain range. *Ornis Svecica*, *16*(4), 183–210.