

The Yoccoz-Birkeland livestock population model coupled with random price dynamics

Riccardo Ceccon* Giulia Livieri† Stefano Marmi‡

23rd May 2022

Abstract

We study a random version of the population-market model proposed by Arlot, Marmi and Papini in Arlot et al. (2019). The latter model is based on the Yoccoz–Birkeland integral equation and describes a time evolution of livestock commodities prices which exhibits endogenous deterministic stochastic behaviour. We introduce a stochastic component inspired from the Black-Scholes market model into the price equation and we prove the existence of a random attractor and of a random invariant measure. We compute numerically the fractal dimension and the entropy of the random attractor and we show its convergence to the deterministic one as the volatility in the market equation tends to zero. We also investigate in detail the dependence of the attractor on the choice of the time-discretization parameter. We implement several statistical distances to quantify the similarity between the attractors of the discretized systems and the original one. In particular, following a work by Cuturi Cuturi (2013), we use the Sinkhorn distance. This is a discrete and penalized version of the Optimal Transport Distance between two measures, given a transport cost matrix.

Keywords: population dynamics, delays dynamical systems, strange attractor, chaotic livestock commodities cycles, Sinkhorn distance, statistical distances

2020 Mathematics Subject Classification: Primary: 37D45; Secondary: 37M05; 37N40; 92D25; 34K60

1 Introduction

One of the most outstanding phenomena in ecology is given by the statistically cyclical variations of small Arctic rodents (see, e.g., Hanski et al. (1993); Andreassen et al. (2021)). The amplitude of these cycles varies widely and seemingly chaotically. In 1998, Yoccoz and Birkeland (Yoccoz and Birkeland (1998))

*Scuola Normale Superiore, Pisa, Italy. Email address: riccardo.ceccon@sns.it

†Scuola Normale Superiore, Pisa, Italy. Email address: giulia.livieri@sns.it

‡Scuola Normale Superiore, Pisa, Italy. Email address: stefano.marmi@sns.it

proposed a model for the evolution of the population of *Microtus epiroticus* (the sibling vole) on the Svalbard Islands in the Arctic Ocean. This species presents a high fertility rate that has a strong dependence on seasonal factors and on the population density. The peculiarity of this model is that it exhibits an *endogenous* chaotic behaviour: it tries to explain the strong annual oscillations presented by the *Microtus epiroticus* population only by exploiting the biological characteristics of the species and its interaction with the environment. Remarkably, these oscillations are determined neither from the lack of food nor from the presence of significant predators, the main one being seagulls.

In Arlot (2012), the authors studied the Yoccoz-Birkeland model via a mathematical analysis and some simulation experiments. Their main results are the following: the Yoccoz-Birkeland model is able to reproduce a complex dynamics with a high sensitivity to initial conditions, only by the combination of density-dependent fertility, the lag due to the maturation age and a periodic seasonality. Nieto et al. (2012), instead, proved the existence of periodic points for the discrete version of the model and performed numerical simulations with a special emphasis to some values for the model parameters.

Persistent approximately periodic fluctuations are also a feature of time-series of livestock commodities prices (Rosen et al. (1994)). Cobweb models show that non-linearities (Chiarella (1988); Hommes (1994)) and simple expectation rules (Hommes (2013)) may lead to chaotic deterministic price fluctuations. In Arlot et al. (2019), the authors coupled the Yoccoz-Birkeland model with a *deterministic* equation modelling the price dynamics of a livestock commodity market. The main idea of this model is the following. A cattle population is split at the birth into reproducing females and cattle for butchery. The splitting “strategy” is determined by very simple naive expectations of the breeder, namely by the spot price of the meat, whereas the logarithmic derivative of the price is driven by the unbalance between the demand and supply of the meat. As in Arlot (2012), the model for the population dynamics accounts explicitly for seasons (or artificial synchronization of births) and maturation lags.

The major outcomes of their study are the following. First, the model gives rise to a chaotic time evolution of price by simply connecting the percentage of reproducing females with the price equation. Second, they show global existence and uniqueness for initial value problems together with some useful estimates on the solutions. Third, they give simple sufficient conditions that ensure the existence of a global attractor containing at least a non-trivial periodic solution. Finally, they perform some numerical experiments in which they show that the attractor has sensitive dependence on initial conditions and non-integer dimension, and that the influence of the market on the model is a crucial factor for the dynamics.

In this paper, we extend the model of Arlot et al. (2019): in particular, the price dynamics of the livestock commodity market is now described by a *Stochastic Differential Equation (SDE, henceforth)*, in a Black-Scholes-like fashion (Bachelier (1900) and Bhattacharya and Majumdar (2007)), instead of a deterministic one. The aim is that of proving also in the new framework the asymptotic results produced in Arlot et al. (2019), in particular the existence of

a global attractor. Thanks to the bounded dependence of the population function with respect to the price, we manage to prove that the population size is bounded and Lipschitz. Nonetheless, the presence of the stochastic component on the price dynamics compromises the usage of the theorem employed in Arlot et al. (2019) to prove the existence of a bounded attractor. Here, instead, we use the concepts of *Random Dynamical System (RDS, henceforth)* and *Random Attractor*, which are based on the so-called *pull-back* approach (see, e.g., Crauel (2002) and Crauel and Flandoli (1994)), to prove the existence of a random attractor on the population component of the phase space. In addition, to prove a more powerful asymptotic result on the population-price dynamical system, we use the concept of *Random (Invariant) Measure* and a version of the Prohorov Theorem for Random Measures (Crauel (2002)) to prove the existence of an invariant random measure. Moreover, we complement the numerical experiments in Arlot et al. (2019): the aim is that of simplifying, i.e., reducing as far as possible the dimension of the phase space of the dynamical system obtained by discretizing the model proposed by Arlot et al. (2019) with the intention of preserving as much as possible the chaotic behaviour observed in the previous study. Toward this end, we allow also for a non-integer number of integration steps. We show graphically the effects of reducing the number of integration steps per year, which appears to be smoothing and simplifying the attractor of the dynamical system. Due to this simplification, we introduce a new set of parameters, such that the plot for a smaller number of integration steps per year changes the least possible. In addition, we employ the concept of Optimal Transport Distance and of Sinkhorn Distance, referring to Cuturi (2013). Thanks to these distances, we are able to formalize and compute explicitly the geometrical difference between the attractors. A plot of this distance, together with the entropy and the fractal dimension, as a function of the number of integration steps per year, is then shown.

The paper is organized in the following way. In Section 2, we recall the Yoccoz-Birkeland model coupled with deterministic price dynamics developed in Arlot et al. (2019). Section 3 describes the random model obtained by adding a diffusion term to the price dynamics and some preliminary estimates on the time evolution of the system are derived in Section 4. The existence of a global random attractor is established in Section 5. Section 6 is devoted to the numerical study of the discretization of the deterministic model and the sensitivity of the attractor on the choice of integration step, whereas Section 7 contains the numerical study of the random attractor. The two appendices contain the fundamental notions on RDS needed and the application to the logistic map of the numerical methodologies used in Appendix B to quantify the distance between attractors.

2 The Yoccoz-Birkeland model and the coupling with price dynamics

This section summarises the model introduced in Arlot et al. (2019) for the time evolution of livestock commodities prices; it hinges on the Yoccoz-Birkeland integral equation Yoccoz and Birkeland (1998) (see Equation (1) below) and on the subsequent analysis in Arlot (2012); Nieto et al. (2012).

A cattle population is divided into two parts: the first one comprises females for reproduction whereas the second one the cattle for butchery (i.e., all the males plus some of the females). Then, the mechanism governing the time evolution of the livestock population and of the meat price is the following:

- (1) At the birth of some babies, part of the newborn females are put in the reproduction line, and the remaining newborn females are put in the butchery line together with all newborn males. The fraction R of newborn females that will reproduce is chosen by the breeder, and is only determined by the price of meat at birth time; the breeder can choose either a short-term strategy or a long-term strategy. Below the dependence upon the price P is denoted by $R(P)$.
- (2) In the reproducing line, females between ages A_0 and A_1 have children. Their fertility can be affected by seasons, or because births are synchronized by the breeder (through a function $m_p(t)$). Reproducing females older than A_1 (hence, non fertile) are not taken into account anywhere in the model - in particular, they are not butchered.
- (3) In the butchery line, cattle can be butchered between ages Ω_0 and Ω_1 . So, only the (alive) butchery population between ages Ω_0 and Ω_1 can count as a “supply” for the market.
- (4) The price evolution is a simple function of the supply (which comes from the butchery line population dynamics) and the demand (which depends only on the price).

Before describing the model, for the reader’s convenience, we sum up in the next subsection the notation and the terminology which will be used throughout this paper. For the sake of clarity and possibility of comparison, we use the notation as in Arlot et al. (2019).

2.1 Terminology and notations

- (N1) t is the time measured in years.
- (N2) $N_r(t)$ is the total population of *mature females* that are in the *reproducing* line and can give birth to pups at time t .
- (N3) $N_b(t)$ is the total population of cattle that is *suitable for butchery* at time t (both males and non-reproducing females, old enough and in the butchery line).

- (N4) $R(P)$ is the fraction of *newborn females* that are put in the reproducing line when the price of meat is P when they are born.
- (N5) A_0 is the age from which *females* can have children (i.e., the age of sexual maturity plus the length of the first gestation).
- (N6) A_1 is the maximal age at which *females* can give birth to children (i.e., the age of sexual infertility plus the length of the last gestation).
- (N7) Ω_0 is the minimal age at which the cattle (*male or female*) can be butchered.
- (N8) Ω_1 is the maximal age at which the cattle (*male or female*) can be butchered.
- (N9) $m(N)$ is the average annual *female* (resp. *male*) fertility of each mature female when the total population is N , i.e., the average number of female (resp. male) babies per year for a single mature female. Typically it is a decreasing function; see Equation (2.3) in Arlot et al. (2019) for a concrete example. We assume a sex ratio equal to $1/2$, i.e. the average number of male babies is equal to the average number of female babies. Hence, $m(N)$ is half of the average annual fertility.
- (N10) $m_\rho(t)$ is the 1-periodic step function (with $\int_0^1 m_\rho(t) dt = 1$) that accounts for a possible modulation of fertility during each year (e.g., births synchronization or seasonal effects).
- (N11) $P(t)$ is the market price of meat at time t .
- (N12) $D(P)$ is the demand of the market (per time unit) when the price of meat is P (typically a decreasing function of P).
- (N13) $S(t)$ is the supply to the market (per time unit) at time t (typically proportional to $N_b(t)$).
- (N14) λ is a “temperature” parameter of the meat market: higher values of λ correspond to bigger price variations in response to the same demand/supply imbalance.
- (N15) $F(D, S)$ is the function of demand and supply that rules the meat price dynamics.
- (N16) $B_f(t)$ is the density of newborn female cattle at time t (i.e., $B_f(t)dt$ females are born between t and $t + dt$).
- (N17) $B_m(t)$ is the density of newborn male cattle at time t .
- (N18) $B_r(t)$ is the density of newborn (female) cattle that are put in the reproducing line at time t .
- (N19) $B_b(t)$ is the density of newborn cattle that are put in the butchery line at time t .

2.2 Deterministic population and price dynamics

The above assumptions on the coupling between the cattle population evolution and the deterministic market model made in Arlot et al. (2019) (see Arlot et al. (2019), Subsections 3.2, 3.3 for a detailed derivation) give rise to the following set of equations for the time evolution of the reproducing females population N_r , of the price P and of the meat supply S :

$$N_r(t) = \int_{A_0}^{A_1} m_\rho(t-a)m(N_r(t-a))N_r(t-a)R(P(t-a))da \quad (1)$$

$$\frac{P'(t)}{P(t)} = \lambda F(D(P(t)), S(t)) \quad (2)$$

$$S(t) = \frac{1}{\Delta\Omega} \int_{\Omega_0}^{\Omega_1} m_\rho(t-a)m(N_r(t-a))N_r(t-a)[2-R(P(t-a))]da \quad (3)$$

with $F(D, S) = \frac{(D-S)}{(D+S)}$, $m : [0, +\infty) \rightarrow [0, +\infty)$, $m_\rho : \mathbb{R} \rightarrow [0, +\infty)$, $R : [0, +\infty) \rightarrow [0, 1]$, $D : [0, +\infty) \rightarrow [0, +\infty)$, $A_1 > A_0 > 0$, $\Omega_1 > \Omega_0 > 0$, $\Delta\Omega = \Omega_1 - \Omega_0$ and $\lambda > 0$.

Notice that the integral evolution Equation (1), which is a standalone equation in the Yoccoz-Birkeland model, is coupled with the *differential* equation in Equation (2) describing the price of a livestock commodity; the latter is driven by the unbalance between its demand and supply.

Under reasonable assumptions on m_ρ , m , R and D (see Arlot et al. (2019), beginning of Section 4, and Equations (4.7)–(4.9)), the authors show that a unique solution to Equations (1)–(3) exists, is globally defined and satisfies some (uniform with respect to the initial condition) estimates: N_r and S are globally bounded and the component N_r is Lipschitz continuous on $[t_0, +\infty)$ for some t_0 . Moreover, they prove the existence of a global attractor and of a non-trivial periodic solution (see Arlot et al. (2019), Theorem A and Theorem B, respectively). Finally, via a numerical investigation, they show that the global attractor is indeed a strange attractor: it has a fractal dimension of ≈ 1.53 and the Kolmogorov-Sinai entropy is positive. Moreover, they showed that the attractor is persistent but its chaotic behaviour depends also on the time evolution of the price in an essential way, a feature which was completely absent in the original Yoccoz-Birkeland model. In particular, they show that if the price dynamic is constant, then the resulting orbit becomes quasi-periodic. Figure 1 displays an example of attractor for the model of Arlot et al. (2019).

3 The Yoccoz-Birkeland model and the coupling with random price dynamics

The model we propose is obtained from Equations (1)–(3) by adding a stochastic Brownian component with a constant volatility $\sigma > 0$ to price dynamics (Equation (2)). The constancy of the volatility is consistent with two popular

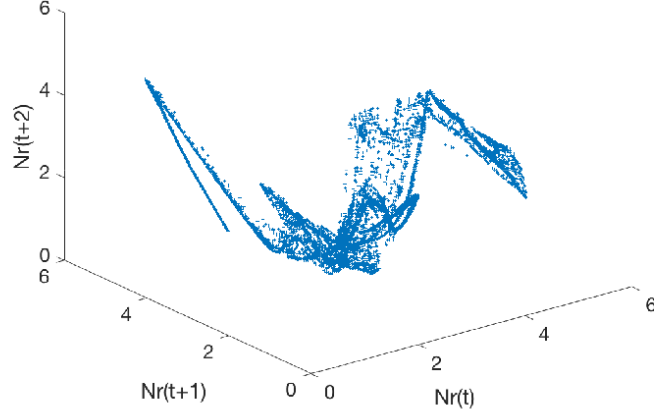


Figure 1: Three dimensional plot of $(N_r(t), N_r(t+1), N_r(t+2))$ with $\rho = 0.30$, $\gamma = 8.25$, $A_0 = 0.18$ and $m_0 = 50$.

models used in the financial literature, i.e., the Bachelier and the Black and Scholes (see Bachelier (1900) and Bhattacharya and Majumdar (2007), respectively). Precisely, the price is described by the following *Stochastic Differential Equation* (SDE):

$$dP(t) = P(t)\lambda F(D(P(t)), S(t)) dt + P(t)\sigma dW_t, \quad (4)$$

or, equivalently, in its logarithmic form:

$$dQ(t) = \left(F(D(P(t)), S(t)) - \frac{\sigma^2}{2} \right) dt + \sigma dW_t. \quad (5)$$

Therefore, the model we are going to study is the following:

$$N_r(t) = \int_{A_0}^{A_1} m_\rho(t-a)m(N_r(t-a))N_r(t-a)R(P(t-a)) da \quad (6)$$

$$dP(t) = P(t)\lambda F(D(P(t)), S(t)) dt + \sigma P(t) dW_t \quad (7)$$

$$S(t) = \frac{1}{\Delta\Omega} \int_{\Omega_0}^{\Omega_1} m_\rho(t-a)m(N_r(t-a))N_r(t-a) [2 - R(P(t-a))] da, \quad (8)$$

where $(W_t)_{t \geq 0}$ is a standard Brownian motion.

In order to analyse the effect of the noise on the attractor of a deterministic dynamical system, we will use the *pull-back approach* (see, e.g., Chekroun et al.

(2011)). To this end, we introduce the two-sided Wiener measure before proceeding with the mathematical analysis of the model. Precisely, on some probability space, we take two independent copies of the Brownian motion $(W_t)_{t \geq 0}$, say $(W_t^{(i)})_{t \geq 0}$ with $i = 1, 2$, and we define the two-sided Brownian motion:

$$W_t = W_t^{(1)} \text{ for } t \geq 0, \quad W_t = W_{-t}^{(2)} \text{ for } t \leq 0.$$

We call \mathcal{P} its law on Borel sets of the space of real-valued continuous functions on \mathbb{R} that are null at zero, i.e. the space $\mathcal{C}_0(\mathbb{R}; \mathbb{R})$. This is the two-sided Wiener measure. Now, consider the canonical space $\Omega = \mathcal{C}_0(\mathbb{R}; \mathbb{R})$ with Borel σ -field and two-sided Wiener measure \mathcal{P} . On $(\Omega, \mathcal{F}, \mathcal{P})$ we consider the canonical two-sided Brownian motion defined as

$$W_t(\omega) = \omega(t) \quad \omega \in \mathcal{C}_0(\mathbb{R}; \mathbb{R})$$

and we interpret the stochastic Equation (7) on the canonical space.

In the next section, we prove some rigorous preliminary results on the Yoccoz-Birkeland model coupled with random price dynamics described by Equations (6)–(8). We will closely follow Section 4 of Arlot et al. (2019), but at the same time account for the stochastic nature of the price dynamics.

4 A first analysis of the model

Following Arlot et al. (2019), we impose the following assumptions on the seasonality m_ρ , the fertility m , the fraction R of newborn females put in the reproducing line and the demand function D :

(A1) $m_\rho : \mathbb{R} \rightarrow \mathbb{R}$ is a non-negative, bounded, 1-periodic function such that $\int_0^1 m_\rho(t) dt = 1$ and we let $m_\rho(t) \leq \mu_{\max}$ and

$$0 < c_0 \leq \int_{A_0}^{A_1} m_\rho(t - a) da \leq c_1 \quad \forall t.$$

(A2) $m : [0, +\infty) \rightarrow \mathbb{R}$ is a continuous function that satisfies

$$\frac{m_0}{2} \min\{1, N^{-\gamma}\} \leq m(N) \leq m_0 \min\{1, N^{-\gamma}\} \quad \forall N > 0$$

with $m_0 > 0$ and $\gamma \geq 1$.

(A3) $R : [0, +\infty) \rightarrow \mathbb{R}$ is a continuous function such that $R_0 \leq R(P) \leq R_1$ for all $P \geq 0$ and some constants $R_1, R_0 > 0$ with $R_1 \leq 1$.

(A4) $D : [0, +\infty) \rightarrow \mathbb{R}$ is a strictly decreasing and locally Lipschitz continuous function such that $D(+\infty) = 0$, and we set $D_0 = D(0)$.

We can now define the phase-space of the dynamical system and the notion of solution of Equations (6)–(8).

We let $T_0 := \min\{A_0, \Omega_0\}$, $T_1 := \max\{A_1, \Omega_1\}$ and we set:

$$\mathcal{X} := \mathcal{X}_1 \times \mathcal{X}_2 := C_{\text{int}}^0([-T_1, 0]; [0, +\infty)) \times C^0([-T_1, 0]; [0, +\infty)), \quad (9)$$

where $C_{\text{int}}^0([-T_1, 0]; [0, +\infty))$ is the space of the positive-valued continuous functions on $[-T_1, 0]$ such that:

$$N_r(0) = \int_{A_0}^{A_1} m_\rho(-a) m(N_r(-a)) N_r(-a) R(P(-a)) da. \quad (10)$$

The space \mathcal{X} is a complete metric space with respect to the distance induced by the norm:

$$\|(N, P)\|_{\mathcal{X}} := \|N\|_{\infty} + \|P\|_{\infty} := \text{esssup}_{s \in [-T_1, 0]} |N(s)| + \sup_{s \in [-T_1, 0]} |P(s)|.$$

Definition 4.1. *Let $(N_r^0, P^0) \in \mathcal{X}$, and $t_0, T \in \mathbb{R}$ with $t_0 < T$. A solution of Equations (6)–(8) with initial data (N_r^0, P^0) is a couple $(N_r, P) : \Omega \times [t_0 - T_1, T) \rightarrow \mathbb{R}^2$ such that $N_r(\omega)|_{[t_0, T)}$ is continuous, $P(\omega)|_{[t_0, T)}$ is continuous, N_r and P satisfy Equations (6)–(8) for $t \in [t_0, T)$, while $N_r(\omega, t_0 + a) = N_r^0(\omega, a)$ and $P(\omega, t_0 + a) = P^0(\omega, a)$ for $a \in [-T_1, 0]$.*

The assumptions (A1)–(A4) guarantee that the model has a unique globally defined solution:

Proposition 4.2 (cfr. Arlot et al. (2019), Proposition 4.2). *Let $(N_r^0, P^0) \in \mathcal{X}$ and $t_0 \in \mathbb{R}$ be given. Then there exists a unique solution pair $(N_r, P) : \Omega \times [-T_1 + t_0, +\infty) \rightarrow \mathbb{R}^2$ of Equations (6)–(8) with initial data (N_r^0, P^0) . Moreover, N_r, P are non-negative and \mathcal{P} -a.s. it holds that:*

$$\begin{aligned} 0 &\leq N_r(t) \leq N_{\max} \quad \forall t \geq t_0 \\ |N_r(t) - N_r(s)| &\leq L_1 |t - s| \quad \forall t, s \geq t_0 \\ 0 &\leq S(t) \leq S_{\max} \quad \forall t \geq t_0, \end{aligned}$$

where:

$$\begin{aligned} N_{\max} &:= m_0 R_1 c_1, \quad L_1 := 2m_0 R_1 \mu_{\max} \\ S_{\max} &:= m_0 \frac{2 - R_0}{\Delta \Omega} \sup_{s \in [0, 1]} \int_{\Omega_0}^{\Omega_1} m_\rho(s - a) da. \end{aligned} \quad (11)$$

Proof. It's a simple adaptation of the proof of the analogous statement given in Arlot et al. (2019), the main difference being that we have to determine P on $[t_0, t_0 + T_0)$ once N_r and S are extended on $[t_0 - T_1, t_0 + T_0)$ in a unique and continuous way. This follows from the stochastic version of the Cauchy-Lipschitz (or Picard-Lindelöf) theorem. \square

Remark 4.3. Equation (6) prescribes the value $N_r(t_0)$ which may be different from $N_r^0(t_0^-)$. Therefore, the solution component N_r may have a jump discontinuity at t_0 even if the initial condition N_r^0 is continuous; this despite the fact that the solution is going to be Lipschitz continuous on $[t_0, \infty)$. This justifies the additional condition we imposed on $N_r^0(0)$ (Equation (10)) while defining the space \mathcal{X} . Should this condition not be imposed, we would have that the shifted solution has a discontinuity at one point. Notice also that if $N_r^0(t) > 0$ for all $t \in [-T_1, 0)$, then $N_r(t)$ is going to be positive for all $t \in [t_0, \infty)$.

The following proposition establishes lower bounds for N_r and S which, together with the estimates in Proposition 4.2, will allow us to define for the deterministic model a compact subset K of \mathcal{X} to which all the solutions belong after a sufficiently big period of time. However, notice that differently from the upper bounds, the time of first entry for the lower bounds is not uniform in the initial data. This will be important when discussing the existence of a random attractor.

Proposition 4.4 (see. Arlot et al. (2019), Proposition 4.4). *Let (N_r, P) be the solution of Equations (6)–(8) with initial data $(N_r^0, P^0) \in \mathcal{X}$ at time t_0 .*

1. *If $N_r(t) \leq N_{\max}$ for a.a. $t \in [\hat{t} - A_1, \hat{t}]$ for some $\hat{t} \geq t_0$, then*

$$N_r(t) \geq \frac{m_0 R_0 c_0}{2} \min \left\{ \inf_{[\hat{t}-A_1, \hat{t}]} N_r, N_{\max}^{1-\gamma} \right\} \quad \forall t \in [\hat{t}, \hat{t} + A_0],$$

In particular this inequality holds for all $t \geq t_0 + A_1$ by Proposition 4.2.

2. *If $m_0 R_0 c_0 > 2$ and $\inf_{[\hat{t}-A_1, \hat{t}]} N_r \geq N_{\max}^{1-\gamma}$ for some $\hat{t} \geq t_0 + A_1$, then*

$$N_r(t) \geq N_{\min} \quad \text{and} \quad S(t) \geq S_{\min} \quad \forall t \geq \hat{t},$$

where:

$$N_{\min} := \frac{m_0 R_0 c_0}{2} N_{\max}^{1-\gamma} \quad \text{and} \quad S_{\min} := m_0 \frac{2 - R_1}{2\Delta\Omega} N_{\max}^{1-\gamma} \inf_{s \in [0, 1]} \int_{\Omega_0}^{\Omega_1} m_\rho(s-a) da. \quad (12)$$

3. *If $m_0 R_0 c_0 > 2$ and $N_0(a) > 0$ for almost all $a \in [-A_1, 0]$, then there exists $t^* \geq t_0$ such that $N_r(t) \geq N_{\min}$ and $S(t) \geq S_{\min}$ for all $t \geq t^*$.*

The fact that there exists a time t^* after which $N_r(t)$ belongs to a compact set K enabled Arlot et al. (2019) to show that a similar property holds also for the price $P(t)$; cfr. Proposition 4.5 and Corollary 4.6 in that paper. However, because of the stochastic nature of our price dynamics, in our case this is no longer true. In particular, this does not allow us to use the same techniques employed in Arlot et al. (2019) for the study of the attractors of the model. For this reason, we need to introduce the concept of *Random Dynamical System* (RDS, henceforth), of *random attractors* and of *random invariant measure*; see Appendix A. In addition, we make the following remark.

Remark 4.5. From here on, the theoretical exposition will be made using $t_0 = 0$. It can be observed that this does not entail any loss in generality, since we can translate a general solution $(N_r(t), P(t))$ at time t_0 to a solution $(N'_r(t), P'(t)) = (N_r(t + t_0), P(t + t_0))$ at time 0, which follows the same model with the seasonal periodic function m_p “shifted” by t_0 . Consequently, the result of the next two sections are true for every $t_0 \in \mathbb{R}$.

5 Existence of a global random attractor

In this section, we prove the existence of a random attractor and of random invariant measure; we refer to Appendix A for a brief review of the relevant concepts. The traditional approach employed to analyse the effect of the noise on the attractor of a deterministic dynamical system is that of finding stationary solutions of the Fokker-Planck equation. Indeed, numerically it is often easier to integrate the system forward in time; see Lasota and Mackey (1998). However, the perturbation of a deterministic system by noise is such that the support of the (numerically) obtained probability density function corresponds to a neighborhood of the deterministic attractor. Whence, it is possible to get statistical information only, without any link with the geometry of the attractor; see Figure 2 for a graphical display of the this claim. We will thus employ the so-called RDS approach: instead of integrating the system forward in time, we will run it from a distant point s in the past until the current time t where the system will be frozen. This approach is also called *pullback approach*. First, we define

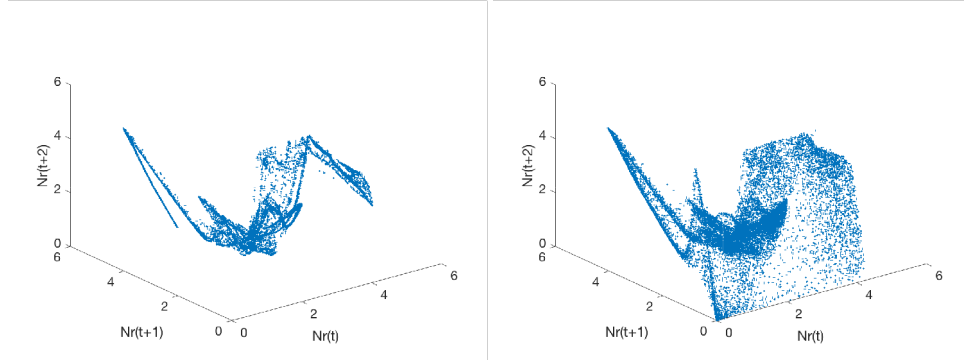


Figure 2: Three dimensional plot of $(N_r(t), N_r(t + 1), N_r(t + 2))$ both for the deterministic (left panel) and the stochastic (right panel) case.

the RDS object of our study (cfr. Definition A.3). The set of times is $T = \mathbb{R}_+$ and the phase-space \mathcal{X} has been defined in Equation (9). The Borel σ -algebra of the phase-space is denoted by \mathcal{B} . The operator $\theta_t(\omega)(s) = \omega(t + s) - \omega(t)$ is the time shift of the Brownian motion, defined $\forall t \in \mathbb{R}$. Therefore, the operator $\varphi(t, \omega)$ acts on an element $(N_r, P) \in \mathcal{X}$ in the following way:

$$\varphi(t, \omega)(N_r, P) = (N_r^t, P^t) \text{ where } N_r^t(s) = N_r(t + s) \text{ and } P^t(s) = P(t + s).$$

The dependence on the parameter ω of the variables on the r . h . s is omitted. By abuse of notation, we will also write $\varphi(t, \omega)N_r = N_r^t$ and $\varphi(t, \omega)P = N_r^t$. We now state and prove the following lemma.

Lemma 5.1. *The RDS $\varphi(t, \omega) : \mathcal{X} \rightarrow \mathcal{X}$ given by the stochastic coupled Birkeland-Yoccoz model in Equations (6)-(8) is continuous.*

Proof. We prove the continuity of the RDS by induction. On the interval $[0, T_0]$ the continuity of the P component is a straightforward consequence of the continuous dependence on the initial condition for the SDE, whereas the continuity of the N_r component is a consequence of the continuity of the integral operator. Inductively, we can suppose that if (N_r^0, P^0) is arbitrarily close to (\hat{N}_r^0, \hat{P}^0) in \mathcal{X} , then $\varphi(t, \omega)(N_r^0, P^0)$ is arbitrarily close to $\varphi(t, \omega)(\hat{N}_r^0, \hat{P}^0)$ for $t \in [(k-1)T_0, kT_0]$. The base case can be thus straightforwardly applied to $\varphi(t, \omega)(N_r^0, P^0)$ and $\varphi(t, \omega)(\hat{N}_r^0, \hat{P}^0)$ to obtain continuity also for $t \in [kT_0, (k+1)T_0]$. \square

In order to find a global attractor for φ , one could be tempted to use Theorem A of Appendix A. However, it cannot be applied directly to our RDS because the price function $P(t)$ is not ω -wise bounded; instead, N_r belongs to the interval $[0, N_{max}] \forall t \geq t_0$. For this reason, we aim at finding a random attractor $A(\omega)$ defined similarly to that in Definition A.6 but with the difference that $A(\omega) = D(\omega) \times \mathcal{X}_2$, where $D(\omega) \subset \mathcal{X}_1$ a compact set on the first component of the space \mathcal{X} . Precisely, we have the following definition:

Definition 5.2 (Globally attracting set). *Let φ be a RDS such that there exists a random set $A(\omega) := D(\omega) \times \mathcal{X}_2$, for some random compact set $D(\omega)$, satisfying to the following conditions:*

1. $\varphi(t, \omega)A(\omega) = A(\theta_t \omega) \forall t > 0$;
2. *A attracts every bounded deterministic set $B \subset \mathcal{X}$.*

Then A is said to be a universally or globally attracting set for φ onto the first component of the phase-space \mathcal{X} .

Notice that that asking that $A(\omega)$ attracts every bounded set $B \subset \mathcal{X}$ is equivalent to asking that $D(\omega)$ attracts every bounded set $B \subset \mathcal{X}_1$. We give the following

Definition 5.3. *Given $B \subset \mathcal{X}_1$, we define $\Omega_B^1(\omega)$ as the projection onto \mathcal{X}_1 of the set $\Omega_{B \times \mathcal{X}_2}(\omega)$.*

In particular, the following proposition holds true.

Proposition 5.4. *Let φ be a RDS on the Polish space \mathcal{X} . Suppose there exists a compact set $K_1(\omega)$ such that $K_1 \times \mathcal{X}_2$ absorbs the product between every bounded non-random set $B \subset \mathcal{X}_1$ and \mathcal{X}_2 . Then the set*

$$A(\omega) = \overline{\bigcup_{B \subset \mathcal{X}_1} \Omega_B^1(\omega) \times \mathcal{X}_2}$$

is a global attractor for φ in the sense of the Definition 5.2.

Proof. See the proof of Theorem 3.11 in Crauel and Flandoli (1994). \square

Thus, we can now define the set

$$K_1 := \{N_r^0 \in \mathcal{X}_1 : 0 \leq N_r^0(t) \leq N_{\max} \text{ for all } t \in [-T_1, 0], \text{ and } N_r^0 \text{ is } 2m_0 R_1 \mu_{\max}\text{-Lipschitz-continuous}\}; \quad (13)$$

see Proposition 4.2. In particular, we have that:

Proposition 5.5. *The set K_1 is a compact set and it satisfies the assumptions of Proposition 5.4. As a consequence, there exists a global attractor $A(\omega)$ onto the first component of the space \mathcal{X} .*

Proof. The set K_1 is compact thanks to Ascoli-Arzelà's Theorem. In addition, thanks to Proposition 4.2 the set $K_1 \times \mathcal{X}_2$ absorbs the product between every bounded nonrandom set $B \subset \mathcal{X}_1$ and \mathcal{X}_2 . Indeed, for every $(N_r^0, P^0) \in \mathcal{X}$ and for all $\omega \in \Omega$ it holds that $\varphi(t, \omega)(N_r^0, P^0)$ belongs to $K_1 \times \mathcal{X}_2$ for every $t \geq T_1$. \square

Notice that the notion of random attractor is *global* and there is no a definition of basin of attraction in the random framework (in the deterministic setup, the attractor's basin of attraction is given by the set of functions N_r that are positive on the interval $[-T_1, 0]$; see Arlot et al. (2019), Section 4).

For this reason, we follow an averaging procedure described in detail in Subsection A.1.2 on invariant measures on random sets. In particular, we want to apply Theorem A.16. To this end, we have to define a suitable probability distribution ν for the initial values $(N_r^0, P^0) \in \mathcal{X}$. We define it component-wise as the conditional product of a measure $\nu_1(N_r^0|P^0)$ on \mathcal{X}_1 and of a measure $\nu_2(P^0)$ on \mathcal{X}_2 . Notice that there is no a standard recipe to construct such measures: in this work, we define them via the Brownian motion's law. For ν_2 we can choose any probability measure on the set of continuous functions: we pick the law of the absolute value of a two-sided Brownian motion starting at $t = -T_1$. In order to define ν_1 , we have to remember that for a fixed $P^0 \in \mathcal{X}_2$ every function $N_r^0 \in \mathcal{X}_1$ satisfies Equation (10), which does not allow for discontinuities at $t = 0$. Thus, we introduce the stochastic process $(N_r^0(t))$ and we set it equal to the absolute value of a two-sided Brownian motion starting at $t = -T_1$ for $t \in [-T_1, -T_0]$ and to a linear interpolant between the point N_{-T_0} and the point

$$N_r^0(0) = \int_{A_0}^{A_1} m_\rho(-a) m(N_r^0(-a)) N_r^0(-a) R(P^0(-a)) ds$$

on $[-T_0, 0]$. We define $\nu_1(\cdot|P^0)$ as the law of such a process on \mathcal{X}_1 and we construct ν in the following way:

$$\nu(A) := \int_{\mathcal{X}_1} \int_{\mathcal{X}_2} I_A(N_r^0, P^0) d\nu_1(N_r^0|P^0) d\nu_2(P^0) \quad \forall A \in \mathcal{B}.$$

At this point, our aim is to show that $\forall \varepsilon > 0$ we can construct a compact set C_ε such that $\Theta_t \nu(\Omega \times C_\varepsilon) > 1 - \varepsilon$. We define $C_\varepsilon := K_1 \times K_2^\varepsilon$, where $K_1 \subset \mathcal{X}_1$ and $K_2^\varepsilon \subset \mathcal{X}_2$ are two suitable compact sets. The set K_1 is the set defined in Equation (13); in particular, Propositions 4.2 and 4.4 ensure that $\varphi(t, \omega) N_r^0 \in K_1 \forall t \geq T_1$ and $\forall \omega \in \Omega$. Therefore, we only have to find the set K_2^ε . Toward this aim, we recall the following two results, which are valid for the deterministic model in Equation (1)-(3).

Proposition 5.6 (cfr. Arlot et al. (2019), Proposition 4.5). *Let (N_r, P) be a solution of (1)-(3) and assume that some $t^* \geq t_0$ exists such that $0 < S_{\min} \leq S(t) \leq S_{\max}$ for all $t \geq t^*$, where we recall that S_{\min} is defined in Proposition 4.4 and S_{\max} is defined in Proposition 4.2.*

1. *Let $P^* \geq 0$ be such that $D(P^*) < S_{\min}$. If $P(\hat{t}) > P^*$ for some $\hat{t} \geq t^*$, then we have*

$$P(t) < P^* \quad \forall t > \hat{t} + \frac{P(\hat{t}) - P^*}{\lambda P^*} \cdot \frac{D_0 + S_{\max}}{S_{\min} - D(P^*)}.$$

2. *Let $P_* \geq 0$ be such that $D(P_*) > S_{\max}$. If $0 < P(\hat{t}) < P_*$ for some $\hat{t} \geq t^*$, then we have*

$$P(t) > P_* \quad \forall t > \hat{t} + \frac{P_* - P(\hat{t})}{\lambda P(\hat{t})} \cdot \frac{D_0 + S_{\max}}{D(P_*) - S_{\max}}.$$

Corollary 5.7 (cfr. Arlot et al. (2019), Corollary 4.6). *Assume that $m_0 R_0 c_0 > 2$, $D_0 > S_{\max}$ and $S_{\min} > 0$ and let $C > 1$ and P_{\min}, P_{\max} be such that $[P_{\min}, P_{\max}] = D^{-1}([S_{\min}, S_{\max}])$.*

1. *If $S(t) \geq S_{\min}$ for all $t \geq t^*$ and $P(\hat{t}) \in [P_{\min}/C, CP_{\max}]$ for some $\hat{t} \geq t^*$, then $P(t) \in [P_{\min}/C, CP_{\max}]$ for all $t \geq \hat{t}$.*
2. *In any case, for every non trivial solution (N_r, P) of (1)-(3) there exists \hat{t} such that $P(t) \in [P_{\min}/C, CP_{\max}]$ for all $t \geq \hat{t}$.*
3. *Moreover, as long as $P(t)$ stays in $[P_{\min}/C, CP_{\max}]$, we have that $|P'(t)| \leq C\lambda P_{\max}$.*

The previous results guarantee that there exists a P^* such that $F(P(t), S(t)) < 0$ whenever $P(t) > P^*$. We define K_2^ε as the set

$$K_2^\varepsilon := \{P^0 \in \mathcal{X}_2 : \sup_{t \in [-T_1, 0]} P^0(t) \leq P^* \text{ for some } P^*\}$$

Without loss of generality, we can assume P^* big enough in order to have that $\nu(K_2^\varepsilon) > 1 - \varepsilon$. In particular, because of the mean-reversion behaviour of the stochastic price process, it is not difficult to prove the following

Proposition 5.8. *Let $P(0) < P^*$. Then, $\forall \varepsilon > 0$ there exists an $M > 0$ such that $\forall t > T_1$ we have that $\mathcal{P}(P(t) > M) < \varepsilon$.*

Proof. We firstly recall that the drift term is smaller than 0 (more precisely, it's smaller than $a < 0$) for $P(t) > P^*$. We introduce two sequences of stopping times to control the behaviour of the stochastic process $P(t)$:

$$\begin{aligned}\tau_0 &:= 0; \\ \sigma_1 &:= \inf\{t \geq 0 : P(t) \geq P^*\}; \\ \tau_n &:= \inf\{t \geq \sigma_n : P(t) \geq P^* + 1\}; \\ \sigma_{n+1} &:= \inf\{t \geq \tau_n : P(t) \leq P^*\};\end{aligned}$$

which are nothing but the “up-crossings” at the levels P^* and $P^* + 1$. We also define the index $N(t) := \inf\{k : \tau_k > t\} - 1$, and the stopping time $\tau_t = \tau_{N(t)}$.

We thus have the following inequalities, for any $M > P^* + 1$:

$$\begin{aligned}\mathbb{P}[P(t) > M] &= \underbrace{\mathbb{P}[P(t) > M, \tau_t = 0]}_{=0} + \underbrace{\mathbb{P}[P(t) > M, \tau_t \neq 0, P(t) < P^* + 1]}_{=0} + \\ &\quad + \mathbb{P}[P(t) > M, \tau_t \neq 0, P(t) > P^* + 1] \\ &\leq \mathbb{P}[P(t) - P(\tau_t) > M - P^* - 1, \tau_t \neq 0] \\ &\leq \mathbb{P}[a(t - \tau_t) + W_t - W_{\tau_t} > M - P^* + 1, \tau_t \neq 0] \\ &\leq \mathbb{P}[a(t - \tau_t) + W_t - W_{\tau_t} > M - P^* + 1]\end{aligned}$$

and the last term is smaller than ϵ for all $t > 0$ if we pick M big enough. The first inequality above is due to the fact that $P(\tau_t)$ is equal to $P^* + 1$, since τ_t is not zero. The second inequality is due to the fact that P remains always above P^* between τ_t and t : if it got below P^* we would have another up-crossing, being $P(t)$ above $P^* + 1$, which would collide with the definition of $N(t)$. Since when $P(t) > P^*$ we have $P(t) - P(\tau_t) < a(t - \tau_t) + W_t - W_{\tau_t}$, we thus conclude the proof. \square

In particular, $\forall t \geq T_1$ we have that:

$$\Theta_t \nu(\Omega \times (K_1 \times K_2^\epsilon)) \geq 1 - \epsilon, \quad (14)$$

which is translated into a bound on $\mu_t(\Omega \times C_\epsilon)$, being the latter the average of $\Theta_t \nu(\Omega \times C_\epsilon)$. The following corollary immediately follows.

Corollary 5.9. *Given the bound in Equation (14), the family $(\mu_t)_{t \geq T_1}$ is tight. As a consequence, there exists an invariant random measure μ which is the limit in the weak topology of a sub-sequence of $(\mu_t)_{t \in T}$ for the model in Equations (6)-(8).*

Proof. The existence follows from TheoremA.16, whereas the invariance from PropositionA.14. \square

In addition, we can show the following proposition.

Proposition 5.10. *It holds that $\mu(\Omega \times (C_{\min} \times \mathcal{X}_2)) = 0$, where*

$$C_{\min} := \left\{ N \in \mathcal{X}_1 : \min_{t \in [-T_1, 0]} N(t) < N_{\min} \right\}.$$

Proof. Since $\mu(\Omega \times (C_{\min} \times \mathcal{X}_2))$ is the limit of (a subsequence of) the averages of $\Theta_t \nu(\Omega \times (C_{\min} \times \mathcal{X}_2))$, it is sufficient to show that these quantities tend to zero as t goes to $+\infty$. Thanks to Proposition 4.4, if $m_0 R_0 c_0 > 2$ we have that

$$\varphi(t, \omega)^{-1}(C_{\min}) \subset \left\{ N_r^0 \in \mathcal{X}_1 : \min_{t \in [-T_1, 0]} N(t) < \eta(t) \right\},$$

for some $\eta(t)$ decreasing function converging to 0 as t goes to infinity. In particular, the measure ν of this set converges to zero; indeed, by definition, ν -almost surely the functions N_r^0 are positive almost everywhere on $[-T_1, 0]$. Thus, $\Theta_t \nu(\Omega \times (C_{\min} \times \mathcal{X}_2)) \rightarrow 0$ as $t \rightarrow \infty$. Thus, we can conclude that $\mu(\Omega \times (C_{\min} \times \mathcal{X}_2)) = 0$ by approximating the indicator function of the set $\Omega \times (C_{\min} \times \mathcal{X}_2)$ with bounded continuous functions f . \square

Proposition 5.11. *It holds that $\mu(\Omega \times (C_{\min} \times \mathcal{X}_2)) = 0$, where*

$$C_{\min} := \left\{ N \in \mathcal{X}_1 : \min_{t \in [-T_1, 0]} N(t) < N_{\min} \right\}.$$

Proof. Since $\mu(\Omega \times (C_{\min} \times \mathcal{X}_2))$ is the limit of (a subsequence of) the averages of $\Theta_t \nu(\Omega \times (C_{\min} \times \mathcal{X}_2))$, it is sufficient to show that these quantities tend to zero as t goes to $+\infty$. Thanks to Proposition 4.4, if $m_0 R_0 c_0 > 2$ we have that

$$\varphi(t, \omega)^{-1}(C_{\min}) \subset \left\{ N_r^0 \in \mathcal{X}_1 : \min_{t \in [-T_1, 0]} N(t) < \eta(t) \right\},$$

for some $\eta(t)$ decreasing function converging to 0 as t goes to infinity. In particular, the measure ν of this set converges to zero; indeed, by definition, ν -almost surely the functions N_r^0 are positive almost everywhere on $[-T_1, 0]$. Thus, $\Theta_t \nu(\Omega \times (C_{\min} \times \mathcal{X}_2)) \rightarrow 0$ as $t \rightarrow \infty$. Thus, we can conclude that $\mu(\Omega \times (C_{\min} \times \mathcal{X}_2)) = 0$ by approximating the indicator function of the set $\Omega \times (C_{\min} \times \mathcal{X}_2)$ with bounded continuous functions f . \square

6 Numerical study of the deterministic model and of its dependence on the integration step

In this section, we extend and complement the numerical study of the deterministic model given by Equations (1)–(3) performed in Arlot et al. (2019). First, we describe the discretization of the model. We will partition the time interval $\Delta t = 1$, which corresponds to one year, with both an integer and a non-integer

number of steps q^1 ; see Subsection 6.1. Second, we describe how we set both the initial conditions and the parameters of the model; see Subsections 6.2 and 6.3, respectively. Third, in Subsection 6.4 we focus on our main objective for this Section and we investigate the dependence of the attractor on the number of integration steps per time unit for two different sets of experimental parameters. The main result of this study is that the chaotic attractor is persistent under the dimensionality reduction process associated to choosing longer and longer integration time steps. In order to make this claim more precise and less qualitative, in Subsection 6.5 we introduce several statistical notions of distance between attractors and we numerically compute them for quantifying the distance of the "asymptotic" attractor (i.e. corresponding to the continuous time system, approximated by a very short integration time) and the one corresponding to the choice of a relatively large integration step.

6.1 Discretization of the model

The phase-space of the original model is infinite dimensional. Here we describe how to numerically integrate it by using an integer number of steps q . The time parameter $t \in [0, +\infty)$ is replaced by indices $i \in \mathbb{N} \setminus \{0\}$; roughly, $i \geq 1$ replaces the interval $\left(\frac{i-1}{q}, \frac{i}{q}\right]$. We now explain the notation we are going to use; notice that the discretized values of the different functions at index i do not correspond to their values on $t = \frac{i}{q}$ but, instead, to their integral or their average on the interval $\left(\frac{i-1}{q}, \frac{i}{q}\right]$.

(D1) $N_{r,i}$ approximates the average of $N_r(t)$ over $t \in \left(\frac{i-1}{q}, \frac{i}{q}\right]$: $N_{r,i} \approx q \int_{\frac{i-1}{q}}^{\frac{i}{q}} N_r(t) dt$.

Similarly, $N_{b,i}$, $m_{\rho,i}$, S_i and P_i are constructed from $N_b(t)$, $m_{\rho}(t)$, $S(t)$ and $P(t)$ as their average over the interval $t \in \left(\frac{i-1}{q}, \frac{i}{q}\right]$.

(D2) $B_{r,i}$ approximates the number of newborn females put in the reproducing line in the interval $t \in \left(\frac{i-1}{q}, \frac{i}{q}\right]$: $B_{r,i} \approx \int_{\frac{i-1}{q}}^{\frac{i}{q}} B_r(t) dt$. The fact that this quantity is not divided by q - differently from $N_r(t)$ - is because $B_r(t)$ is a density (number per unit of time), whereas $N_r(t)$ is the number of individuals.

(D3) The quantities $B_{b,i}$, $B_{f,i}$, $B_{m,i}$ are defined in the same fashion, respectively from $B_b(t)$, $B_f(t)$, $B_m(t)$.

We construct the discretized model at any step $k \geq 1$ by computing

$$Z_k := (N_{r,k}; N_{b,k}; S_k; P_k; B_{r,k}; B_{b,k}) \quad (15)$$

¹In the first case, we will obtain a deterministic dynamical system in a phase-space of dimension $2 \times (T_1 \times q + 1)$. In the second case, instead, the dimension will be $2 \times (T_1 \times [q] + 1)$

from its past values

$$(Z_i)_{k-\max\{k_{A_1}, k_{\Omega_1}\} \leq i \leq k-1},$$

in the following order:

$$N_{r,k} = \sum_{j=k_{A_0}}^{k_{A_1}} B_{r,k-j}, \quad (16)$$

$$N_{b,k} = \sum_{j=k_{\Omega_0}}^{k_{\Omega_1}} B_{b,k-j}, \quad (17)$$

$$S_k = \frac{qN_{b,k}}{k_{\Omega_1} - k_{\Omega_0} + 1}, \quad (18)$$

$$P_k = \max \left\{ 0, P_{k-1} + \frac{\lambda P_{k-1} F(D(P_{k-1}), S_k)}{q} \right\}, \quad (19)$$

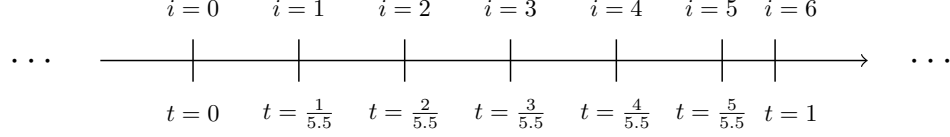
$$B_{r,k} = \frac{m_0}{q} m_{\rho,k} m(N_{r,k}) R(P_k), \quad (20)$$

$$B_{b,k} = \frac{m_0}{q} m_{\rho,k} m(N_{r,k}) (2 - R(P_k)). \quad (21)$$

Notice that the Equations (16)-(17) and (18) depend on the parameters $k_{A_0}, k_{A_1}, k_{\Omega_0}, k_{\Omega_1}$: these are the discretized version of the parameters $A_0, A_1, \Omega_0, \Omega_1$ of the original dynamical system, but with some minor technical modifications needed to ensure the convergence of the numerical scheme (and that the discretization leads to a finite-dimensional dynamical system). Namely, we define:

1. $k_{A_0} = \max\{1, [qA_0]\}$, where $[x]$ is the closest integer to x . Notice that the $\max\{1, \cdot\}$ is because we want $k_{A_0} > 0$ so that $N_{r,k}$ only depends on the past of B_r in Equation (16), otherwise we would have a circular definition.
2. $k_{A_1} = \max\{k_{A_0}, [qA_1] - 1\}$. We take $[qA_1] - 1$ instead of $[qA_1]$ because animals are supposed to die exactly at age A_1 , so individuals of age A_1 at time k/q (counted in $B_{r,k-[qA_1]}$) do not count in $N_{r,k}$. At the same time, we want $k_{A_1} \geq k_{A_0}$ in order to make the sums defining $N_{r,k}$ and $N_{b,k}$ non-empty when A_0 is really close to A_1 .
3. $k_{\Omega_0} = \max\{1, [q\Omega_0]\}$ is defined similarly to k_{A_0} .
4. $k_{\Omega_1} = \max\{k_{\Omega_0}, [q\Omega_1] - 1\}$ is defined similarly to k_{A_1} .

Non-integer number of steps When $q \notin \mathbb{N}$, we partition in $\lceil q \rceil$ parts each interval $[n, n+1]$ for $n \in \mathbb{N}$. The first $\lfloor q \rfloor$ parts are long $\frac{1}{q}$, whereas the last one is long $\frac{\{q\}}{q}$, where $\{q\}$ denotes the fractional part of the number q . An example of this kind of partition for $q = 5.5$ is shown in the following picture:



Also, it is necessary to modify the discretization of the functions N_r, N_b, S, P and m_ρ in (D1). Indeed, for $i \neq 0 \bmod \lceil q \rceil$ the quantity $N_{r,i}$ –and analogously $N_{b,i}, S_i, P_i, m_{\rho,i}$ – is defined as the approximation of the following quantity:

$$N_{r,i} \approx q \int_{t_i - \frac{1}{q}}^{t_i} N_r(t) dt, \quad t_i := \left\lfloor \frac{i}{\lceil q \rceil} \right\rfloor + \left(i - \left\lfloor \frac{i}{\lceil q \rceil} \right\rfloor \right) \cdot \frac{1}{q}.$$

Instead, $i = 0 \bmod \lceil q \rceil$ the quantity $N_{r,i}$ approximates the following integral:

$$N_{r,i} \approx \frac{q}{\{q\}} \int_{t_i - \frac{\{q\}}{q}}^{t_i} N_r(t) dt, \quad t_i := \left\lfloor \frac{i}{\lceil q \rceil} \right\rfloor.$$

Finally, in order to satisfy this approximation requirement, we re-weight the components of Z_k in Equation (15) dividing them by $\{q\}$ whenever k is a multiple of $\lceil q \rceil$.

6.2 Initial conditions

The initial conditions are given by the values of the number of newborns ($B_{r,i}$ and $B_{b,i}$) for time steps $i = 1, \dots, k_{A_1}$ before the start of the simulation, since the discretized model in Subsection 6.1 only needs the past values of B_r and B_b to compute the future dynamics of the model. For each parameter set, we choose an initial condition randomly as follows: $B_{r,1}, \dots, B_{r,k_{A_1}}, B_{b,1}, \dots, B_{b,k_{A_1}}$ are chosen independently with common distribution $\mathcal{U} \left(\left[0, \frac{2}{k_{A_1} - k_{A_0} + 1} \right] \right)$. The reason for this choice is that $k_{A_1} - k_{A_0} + 1$ is the number of time steps corresponding to reproducing ages of females. So, for instance, the reproducing female population at time $k_{A_1} + 1$ is the average of random variables uniform over $[0, 2]$, so it should be close to 1, the threshold value for density-dependence to apply.

6.3 Parameters of the model

We follow Arlot et al. (2019), and we fix the functions m, m_ρ, D and R in such a way that they have to satisfy the requirements of Section 4. In particular in our experiments we make the following choices:

(F1) **The fertility function.** We take:

$$m(N) = m_0 (\max\{N, 1\})^{-\gamma},$$

where γ and m_0 are parameters of the model to be chosen.

(F2) **Seasonality.** We let

$$m_\rho(t) = \frac{1}{1-\rho} \mathbf{1}_{\{t - \lfloor t \rfloor \in [0, 1-\rho)\}}.$$

In this way, we model the fact that the births are synchronized and concentrated in one specific period of the year.

(F3) **Demand function.** We choose the following demand function

$$D(P) = D_{\text{exp}}(P) := D_0 e^{-\alpha_D P}$$

for some parameters $D_0, \alpha_D > 0$.

(F4) **Breeder strategy:** We assume that the breeder follows a counter-cyclical policy and takes R close to 1 when the price is high. In the numerical experiments we choose

$$R(P) = R_{\text{logistic}}(P) := R_0 + (R_1 - R_0) f_d(P/P_0) \quad (22)$$

where

$$f_d(x) = \begin{cases} \frac{x^d}{2} & \text{if } x \in [0, 1) \\ \frac{1}{1 + \exp(-2d(x-1))} & \text{otherwise} \end{cases}$$

and $R_0, R_1 \in [0, 1]$ and $P_0, d > 0$ have to be chosen. This is a sigmoid function between 0 and $+\infty$.

Finally, for the sake of simplicity, we took the parameters Ω_0 and Ω_1 in the supply function to be equal to A_0 and A_1 , respectively. The normalization of the supply is $\Delta\Omega = \Omega_1 - \Omega_0 = A_1 - A_0$.

6.4 Dependence of the attractor on the discretization time step q

The aim of this section is to study the dimensionality reduction of the dynamical system in Equations (1)–(3) as a function of the number q of steps per year. In order to do so, we visualize attractor of the deterministic model by plotting in \mathbb{R}^3 the set $C(t) := (N_r(t), N_r(t-1), N_r(t-2))$, $t \in \mathbb{N}$, for different values of q . The goal will be to see how much the geometric characteristics of the attractor will change by passing from, e.g., $q = 100$ to, e.g., $q = 2$. Should the chaotic dynamics be maintained, this would be a remarkable result. Indeed we would pass from a dynamical system with a dimension of the phase-space equal to 200 to a dynamical system with a dimension of the phase-space equal to 8 and, following Yoccoz and Birkeland (1998), for $t \in [290000, 300000]$. We define the first set of experimental parameters, say \mathcal{H}_1 , in the following way:

(S.1.1) **Population dynamics:** $A_0 = 0.18$, $A_2 = 2.0$, $m_0 = 5.0$, $\gamma = 8.25$ and $\rho = 0.79$.

(S.1.2) **Market dynamics:** $\lambda = 1$ and the demand function is $D = D_{\text{exp}}$ with $D_0 = 5$ and $\alpha_D = 1$.

(S.1.3) **Interaction between population and market:** $\Omega_0 = 0.18$, $\Omega_1 = 2$, and $R = R_{\text{logistic}}$ with $R_0 = 0$ (*minimal value*), $R_1 = 1$ (*maximal value*), $P_0 = 1$ (*price threshold*) and $d = 4$ (*degree of $R(P)$ for small P*).

This set is called SP in Arlot et al. (2019) and is close to the main setting for the Birkeland-Yoccoz model numerically solved in Arlot (2012). The left (resp. right) panel of Figure 3 displays the dynamics of the set $C(t)$ associated to $q = 100$ (resp. to $q = 2$) steps per year. Interestingly, several skeleton of the attractor and some of its structure survives. On the other hand, the geometric structure of the set $C(t)$ is undoubtedly simplified with respect to the set depicted on the left-panel of the same figure. Also, it presents some –a priori unexplained– stronger accumulation towards the value of zero.

some of the geometrical features are preserved; indeed, even with $q = 2$ there are still some of the spikes and of the edges of the three dimensional projection with $q = 100$.

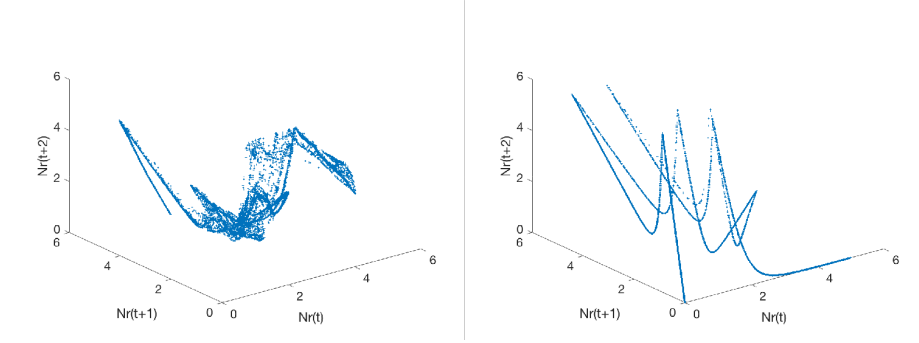


Figure 3: Three dimensional plot of $C(t) = (N_r(t), N_r(t+1), N_r(t+2))$ of the deterministic dynamical system in Equations (1)–(3). Setting \mathcal{H}_1 . *Left Panel:* $q = 100$. *Right Panel:* $q = 2$.

Actually, the dimension of the dynamical system’s phase-space with $q = 2$ can be lowered to 5; indeed, the following lemma holds true.

Lemma 6.1. *The discretized dynamics that arise by choosing $q = 2$ with the set of experimental parameters \mathcal{H}_1 below can be modeled through a dynamical system whose phase-space dimension is equal to 5.*

Proof. First, we notice that $m_{\rho,k}$ consists of a sequence of 2’s for k even and of zeros for k odd. Also, we point out that the values of k_{A_0} and k_{Ω_0} are 1 and the values of k_{A_1} and k_{Ω_1} are 3. As a consequence, in the summations defining $N_{r,k}$ and S_k (see Equations (16) and (18)) the terms $N_{r,i}$ and P_i for i odd do not matter, since they are multiplied by $m_{\rho,i}$ which is zero for k odd. Also, P_k in Equation (19) depends only on P_{k-1} and on S_k , which in turn depends on

the values of $N_{r,i}$ and P_i for even i . Therefore, for $q = 2$, we can construct a dynamical system consisting only of the couple $(N_{r,k}, N_{r,k-2})$ and of the prices (P_k, P_{k-1}, P_{k-2}) whose evolution is given by

$$\begin{pmatrix} N_{r,k-2}, N_{r,k} \\ P_{k-2}, P_{k-1}, P_k \end{pmatrix} \mapsto \begin{pmatrix} N_{r,k}, \frac{m_0}{q} m_{\rho,k} m(N_{r,k} R(P_k)) \\ P_k, 0 \vee P_k + \frac{1}{q} (\lambda P_k F(F(P_k), S_{k+1})), P_{k+2} \end{pmatrix}.$$

Notice that in the previous equation we pointed out that S_{k+1} can be obtained from $N_{r,k}, N_{r,k-2}, P_k$ and P_{k-2} , and that P_{k+2} can be obtained from the value of

$$P_{k+1} = 0 \vee P_{k-1} + \frac{1}{q} (\lambda P_{k-1} F(D(P_{k-1}), S_k))$$

and from that of S_{k+2} , which can be computed from $N_{r,k}$ and P_k . \square

A more interesting phenomenon is shown in Figure 4. The left panel displays the dynamics of the set $C(t)$ associated to $q = 10$ steps per year: as expected, the geometry of the set stands between the one with $q = 2$ and $q = 100$. However, the right panel shows the possibility that $C(t)$ exhibits, for some values of q , only a low-complexity seemingly non-chaotic orbit, very close to being periodic and totally different also from the original Yoccoz-Birkeland attractor. In particular, the bifurcation diagram of $N_r(t)$, $290000 \leq t \leq 300000$, as a function of the time discretization parameter $q \in [1, 40]$ in Figure 5 shows that the presence of a quasi-periodic orbit is more frequent for small values of q .

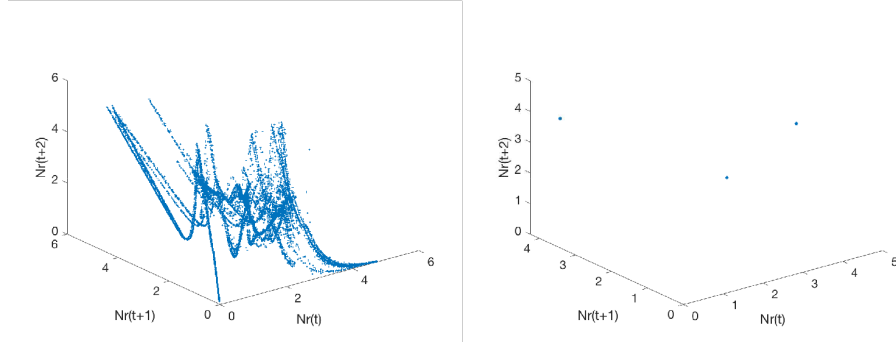


Figure 4: Three dimensional plot of $C(t) = (N_r(t), N_r(t+1), N_r(t+2))$ of the deterministic dynamical system in Equations (1)–(3). Setting \mathcal{H}_1 . Left Panel: $q = 10$. Right Panel: $q = 20$.

6.4.1 Comparison with setting \mathcal{H}_1 : small values of q with a chaotic dynamics

We tune the set of experimental parameters \mathcal{H}_1 until the three-dimensional plots of $(N_r(t), N_r(t+1), N_r(t+2))$ for small and high values of q look very similar. In particular, we name \mathcal{H}_2 this set of new experimental parameters. With respect to \mathcal{H}_1 , we modify the following parameters :

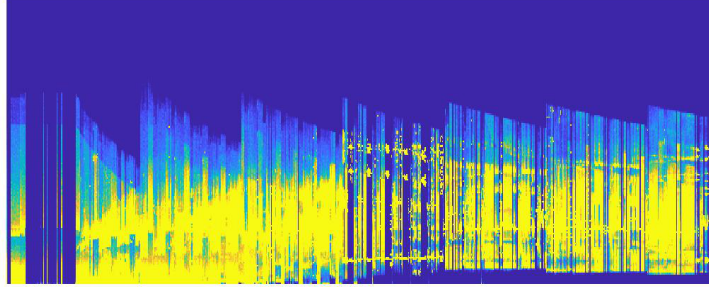


Figure 5: Bifurcation diagram for $N_r(t)$, $290000 \leq t \leq 300000$, as a function of the parameter $q \in [1, 40]$. Setting \mathcal{H}_1 .

- (S.2.1) $A_0 = \Omega_0 = 0.8$, which corresponds to a longer period of growth of the cattle before they reach fertility.
- (S.2.2) $m_0 = 10$.
- (S.2.3) $\gamma = 7$.
- (S.2.4) The degree d of the R_{logistic} function is set equal to 2.
- (S.2.5) The coefficient α_D of the D function is set equal to 2.

The left (resp. right) panel of Figure 6 displays the dynamics of the set $C(t)$ associated to $q = 3$ (resp. to $q = 100$) steps per year in the set of experimental parameters \mathcal{H}_2 . We make the following considerations. First, the accumulation towards the value of zero is preserved for small values of q . Second, even for such a values there exists a non trivial geometric structure that is similar to the one observed for $q = 100$. Also, we find that for only a few values of q the set $C(t)$ exhibits an orbit very close to being periodic; see Figure 7, left panel, for a graphical representation of this statement. As an example, Figure 7, right panel, displays the quasi-periodic orbit when $q = 10.5$.

6.5 Measuring distances between attractors

By comparing Figure 3 with Figure 6, it is visually clear that the reduction in the number of integration steps q distorts less the geometry of $C(t)$ in the case of the set of experimental parameters \mathcal{H}_2 . In order to make this statement more rigorous, we use a number of *Statistical Distances* proposed in the literature, i.e. a number of distances between measures defined on a metric space (X, δ) . In particular, each distance is approximated by the distance between two histograms (i.e. between two vectors of the same length with positive entries that sum-up to one) constructed from the corresponding attractors. Precisely, we construct the histograms by measuring the density of the points constituting the plot of the attractors over a three dimensional grid X of dimension

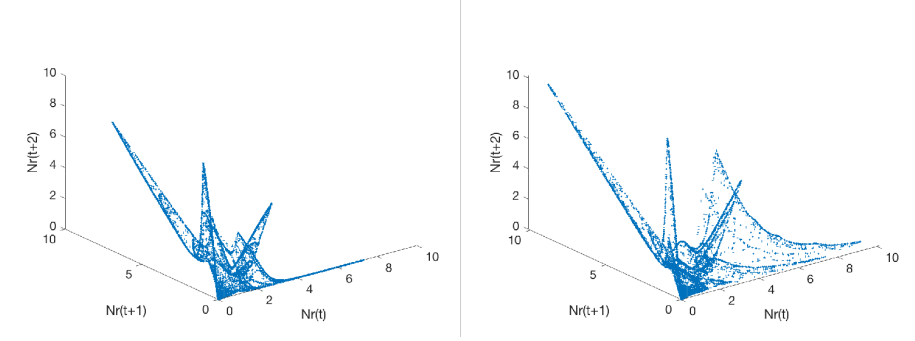


Figure 6: Three dimensional plot of $C(t) = (N_r(t), N_r(t+1), N_r(t+2))$ of the deterministic dynamical system in Equations (1)–(3). Setting \mathcal{H}_2 . Left Panel: $q = 2$. Right Panel: $q = 100$.

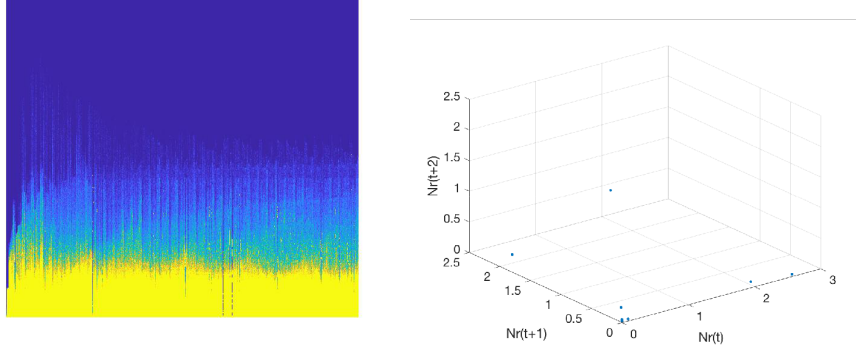


Figure 7: Left Panel: Bifurcation diagram for $N_r(t)$, $290000 \leq t \leq 300000$, as a function of the parameter $q \in [1, 40]$. Setting \mathcal{H}_2 . Right Panel: Three dimensional plot of $(N_r(t), N_r(t+1), N_r(t+2))$ of the deterministic dynamical system in Equations (1)–(3). Setting \mathcal{H}_2 and $q = 10.5$.

$n_g \times n_g \times n_g$: for computational reasons, we used $n_g = 23$. Henceforth, we denote by r the histogram associated to the attractor obtained with $q = 100$ and we call this attractor *asymptotic* attractor, and by c_i , $i = 1, \dots, 100$, the histograms associated to the attractors obtained with $q \in [2, 100]$, on a pseudo-logarithmic scale. We use the following statistical distances between two given histograms P and Q :

(D1) Kullback-Leibler divergence (Kullback and Leibler (1951)). It is defined as

$$D_{KL}(P||Q) := \sum_{x \in X} P(x) \log \left(\frac{P(x)}{Q(x)} \right).$$

- (D2) Jensen-Shannon distance (Wong and You (1985)). It is defined as the smoothed and symmetrized version of the Kullback-Leibler, i.e.:

$$D_{JS}(P||Q) := \frac{1}{2} (D_{KL}(P||M) + D_{KL}(Q||M)) \quad \text{with} \quad M = \frac{1}{2} (P + Q).$$

- (D3) The Wasserstein distance (Ambrosio et al. (2005)). It is defined as:

$$W(P, Q) := \left(\inf_{\gamma \in \Gamma(P, Q)} \int_{X \times X} \delta(x, y)^2 d\gamma(x, y) \right)^{1/2},$$

where $\Gamma(P, Q)$ denotes the collection of all measures on $X \times X$ with marginal distributions P and Q on the first and second factors respectively.

- (D4) A regularized version of the Sinkhorn distance proposed by Cuturi (2013). It is defined in the following way. Let r and c be two given histograms in the simplex $\Sigma_d := \{x \in \mathbb{R}_+^d : x^T \mathbf{1}_d = 1\}$. Let $U(r, c)$ be the transportation polytope of r and c , namely the polyhedral set of $d \times d$ matrices

$$U(r, c) := \{P \in \mathbb{R}_+^{d \times d} \mid P \mathbf{1}_d = r, P^T \mathbf{1}_d = c\},$$

where $\mathbf{1}_d$ is the d dimensional vector of ones. Given a $d \times d$ cost matrix M , the cost of mapping r to c using a transportation matrix P can be quantified as $\langle P, M \rangle, := \text{Tr}(P^T M)$ the Frobenius product between P and M . The following minimization problem:

$$D_M(r, c) := \min_{P \in U(r, c)} \langle P, M \rangle$$

is called the *optimal transportation* problem between r and c , given the cost M , whereas

$$D_M^\lambda(r, c) := \min_{P \in U(r, c)} \langle P, M \rangle - \frac{1}{\lambda} h(P)$$

is an approximated version of the problem above, where h is the Shannon entropy, and $D_M^\lambda(r, c)$ is called *Sinkhorn distance*. Both D_M and D_M^λ are distances between two histograms; see Cuturi (2013). Last but not least, Cuturi (2013) proposes a computationally efficient algorithm –called Sinkhorn-Knopp Algorithm– to compute the distance D_M^λ between a given histogram r and a set of histograms $\{c_i\}_{i \in I}$, as well as a lower bound on the real optimal transport distance.

We discuss now the results. Figure 8, left (resp. right) panel, plots the distances just presented for the set of experimental parameters \mathcal{H}_1 (resp. \mathcal{H}_2). For both settings, we notice that the distance is a decreasing function of q , correspondingly to the qualitative assessment that for higher values of q we obtain figures that are more and more similar to the asymptotic one. However, we sometimes

observe some spikes; they correspond to the values of q for which the orbit is quasi-periodic. In addition, notice that these spikes are more sporadic in the setting \mathcal{H}_2 , thus confirming what we observed in Section 6.4. Moreover, the Sinkhorn distance (i.e. the transportation cost) for low values of q is much higher for the setting \mathcal{H}_1 than for the setting \mathcal{H}_2 . This confirms the visual insight that the latter setting is much more stable than \mathcal{H}_1 for small values of q , i.e. it leads to attractors which have a more gentle dependence on q . These considerations are also corroborated by Figure 9, which reports, as function of q , the entropy and the fractal dimension computed following the same methodology as in Arlot et al. (2019): we postpone the discussion of these quantities to Section 7.1. Finally, we remark that we also test the reliability of the previous methodology on the logistic map; see Appendix B.

7 Numerical study of the random model

In this section, we will consider the random model given by Equations (6)–(8) and we perform some numerical experiments on it. In Section 7.1 we investigate the chaotic nature of the attractor of the random dynamical system by computing the entropy and its fractal dimension. In Section 7.2 we show some plots of the random attractors as in Definition A.6.

In order to numerically solve equations (6)–(8) we need to follow a discretization scheme different from the one described in Subsection 6.1 and used for the deterministic model (1)–(3). Indeed we need to discretize differently the dynamics of the price ($P(t)$): in particular, we use the finite-difference Euler-Maruyama method (Maruyama (1954)) for the numerical solution of the SDE in Equation (7).

7.1 Entropy and fractal dimension

In this section, we investigate how the metric entropy and the fractal dimension of the attractor change as the volatility σ varies when the parameters of the deterministic dynamical system are fixed as both in \mathcal{H}_1 and \mathcal{H}_2 ; see Subsection 6.4. For the sake of clarity, we briefly describe how the entropy and the fractal dimension are computed; cfr. Arlot et al. (2019), Appendix A.3 and Appendix A.4.

In order to compute the entropy, we start from the “continuous” time series of $N_r(t)$ –the reasoning can be replicated also for the price– for $t > T_{\max} - 10000$, where T_{\max} is the total length of the simulation experiment i.e., $T_{\max} = 300000$. Then, we compute the empirical auto-correlation function of N_r with lags $\tau \in [0, 100]$ (τ is expressed in years). We get its first “zero” τ^* as the first point where the empirical auto-correlation crosses zero, and we then check that the value of the empirical auto-correlation at $\tau = \tau^*$ is smaller than an arbitrary threshold (here, 10^{-2}). After that, we consider the sequence $Y(t)$ only at times

t equal to an integer multiple of τ^* , and compute the returns

$$\log_{10} \left(\frac{Y((i+1)\tau^*)}{Y(i\tau^*)} \right)$$

and their respective signs $\varepsilon_i \in \{-1, 1\}$. From this (finite) sequence of signs, we compute for $K = 1, \dots, 12$ the combinatorial entropy $H_K(\tau^*)$ of $((\varepsilon_i)_{k-K+1 \leq i \leq k})_{k \in \mathbb{N}}$:

$$H_K(\tau^*) := - \sum_{x \in \{-1, 1\}^K} p_x \log_2(p_x) \quad \text{where} \quad p_x = \mathbb{P}((\varepsilon_i)_{k-K+1 \leq i \leq k} = x)$$

the latter probability being with respect to k . Finally, we plot $H_K(\tau^*)$ as a function of K and perform a standard robust linear regression in order to estimate its slope: the slope will be eventually the estimate on the entropy of the system.

Instead, in order to compute the fractal dimension of the attractors, we start from their three-dimensional visualization over 200000 years, that is, for N_r for instance:

$$\mathcal{K} = \left\{ (N_r(t), N_r(t+1), N_r(t+2)) , 100\,000 \leq t \leq 300\,000 \right\}.$$

Then, for various values of $\varepsilon > 0$, we compute the number $\tilde{\mathcal{N}}_\varepsilon(\mathcal{K})$ of cubes

$$\mathcal{C}_{i,j,k} = [i\varepsilon; (i+1)\varepsilon] \times [j\varepsilon; (j+1)\varepsilon] \times [k\varepsilon; (k+1)\varepsilon]$$

that contain at least one point of \mathcal{K} . Theoretically, the fractal dimension is the opposite of the slope of this graph at infinity. Here, since \mathcal{K} is finite, $\tilde{\mathcal{N}}_\varepsilon(\mathcal{K})$ is constant equal to $\text{Card}(\mathcal{K})$ for small ε . So, we estimated the slope of the graph only for a limited set of values of ε . Because of the numerous approximations made during this estimation, the precise value of the fractal dimension should not be taken into account too seriously, but its order of magnitude should be correct.

Figure ? displays the results. The plots of the entropy, as expected, confirm that the randomization of the system brings more chaos and also fuzziness and, therefore, an higher dimension. However, the behaviour of the plot for the fractal dimension is not monotone as much as that of the entropy. This can be linked to the lack of precision in the computation of the fractal dimension, or simply to the fact that too much “fuzziness” leads to having many isolated points that do not contribute to a substantial increase in the fractal dimension.

7.2 Plotting the random attractor

In this subsection, we display some plots of the globally attracting set as defined in Definition 5.2. Notice that in Proposition 5.5 we prove the existence of such set, which is *random* and it is constructed by fixing a realization ω of the Brownian motion and then measuring the value $\varphi(t, \omega)(N_r^0, P^0)$ for many different initial values (N_r^0, P^0) . In practice, we fix a large enough time, say $t = 10000$,

we generate a huge amount of random initial values, say 20 000, and we move the system forward in time by 10000 years, while using the same seed for the generation of the Brownian Motion. Because the attractor is high-dimensional, we display in Figure 11 its two-dimensional projection. More precisely, we plot the points $(N_r(t), N_r(t+1))$ for $t = 10000 + i/50$ with $i = 0, 1, \dots, 5$: we observe actually a gradual evolution of the shape. Similarly, Figure 12 shows, instead, the same two-dimensional projection for $t = 10\,000 + i$ with $i = 0, 1, 2, 3$, i.e. for integer values of t . Now, we make the following remarks.

Remark 7.1. *Because of the definition of random attractor, should both the initial states and the path of the Brownian motion be random, we would have obtained the same plot for all the values of t .*

Remark 7.2. *Figures 11 and 12 show not only the two-dimensional projection of the attractor as function of the random parameter ω , but also the projection of the random measure, whose existence has been proved in Corollary 5.9. In particular, the areas with higher density are shown in yellow, while the areas of lower density are shown in blue.*

Finally, we observe that by fixing the seed of the Brownian motion and by letting t vary, we can show every possible instance of the attractors $A(\omega)$ thanks to the following lemma.

Lemma 7.3. *The dynamical system $(\Omega, \mathcal{F}, (\theta(t))_{t \in \mathbb{R}}, \mathbb{P})$ is ergodic.*

Proof. It is a consequence of Kolmogorov's zero-one law. Consider the two-parameter filtration $\{\mathcal{F}_s^t\}_{s \leq t}$ generated by the Brownian motion, and define the σ -algebra

$$\mathcal{T}^\infty = \bigcap_{t \in \mathbb{R}} \mathcal{F}_t^\infty.$$

The independence of the σ -algebras \mathcal{F}_s^u and \mathcal{F}_t^z for all $s < u \leq t < z$ allows to apply Kolmogorov's dichotomy and deduce that the σ -algebra \mathcal{T}^∞ is degenerate, i.e. $\mathbb{P}(A) \in \{0, 1\}$ for all $A \in \mathcal{T}^\infty$. The conclusion follows by observing that every $\theta(t)$ -invariant set is contained in \mathcal{T}^∞ . \square

References

- Ambrosio, L., N. Gigli, and G. Savaré (2005). *Gradient flows: in metric spaces and in the space of probability measures*. Springer Science & Business Media.
- Andreassen, H. P., J. Sundell, F. Ecke, S. Halle, M. Haapakoski, H. Henttonen, O. Huitu, J. Jacob, K. Johnsen, E. Koskela, et al. (2021). Population cycles and outbreaks of small rodents: ten essential questions we still need to solve. *Oecologia* 195(3), 601–622.
- Arlot, S. (2012). Etude d’un modèle de dynamique des populations. *arXiv preprint arXiv:1204.0799*.
- Arlot, S., S. Marmi, and D. Papini (2019). Coupling the yoccoz–birkeland population model with price dynamics: chaotic livestock commodities market cycles. *Nonlinearity* 32(7), 2564.
- Bachelier, L. (1900). Théorie de la spéculation. In *Annales scientifiques de l’École normale supérieure*, Volume 17, pp. 21–86.
- Bhattacharya, R. and M. Majumdar (2007). *Random dynamical systems: theory and applications*. Cambridge University Press.
- Chekroun, M. D., E. Simonnet, and M. Ghil (2011). Stochastic climate dynamics: Random attractors and time-dependent invariant measures. *Physica D: Nonlinear Phenomena* 240(21), 1685–1700.
- Chiarella, C. (1988). The cobweb model: Its instability and the onset of chaos. *Economic modelling* 5(4), 377–384.
- Crauel, H. (2002). *Random probability measures on Polish spaces*, Volume 11. CRC press.
- Crauel, H. and F. Flandoli (1994). Attractors for random dynamical systems. *Probability Theory and Related Fields* 100(3), 365–393.
- Cuturi, M. (2013). Sinkhorn distances: Lightspeed computation of optimal transport. *Advances in neural information processing systems* 26.
- Flandoli, F. and E. Tonello (2021). An introduction to random dynamical systems for climate.
- Hanski, I., P. Turchin, E. Korpimäki, and H. Henttonen (1993). Population oscillations of boreal rodents: regulation by mustelid predators leads to chaos. *Nature* 364(6434), 232–235.
- Hommes, C. (2013). *Behavioral rationality and heterogeneous expectations in complex economic systems*. Cambridge University Press.
- Hommes, C. H. (1994). Dynamics of the cobweb model with adaptive expectations and nonlinear supply and demand. *Journal of Economic Behavior & Organization* 24(3), 315–335.

- Kullback, S. and R. A. Leibler (1951). On information and sufficiency. *The annals of mathematical statistics* 22(1), 79–86.
- Lasota, A. and M. C. Mackey (1998). *Chaos, fractals, and noise: stochastic aspects of dynamics*, Volume 97. Springer Science & Business Media.
- Maruyama, G. (1954). On the transition probability functions of the markov process. *Nat. Sci. Rep. Ochanomizu Univ* 5, 10–20.
- Nieto, J. J., M. J. Pacifico, and J. L. Vieitez (2012). Long-term and short-term dynamics of microtus epiroticus: A yoccoz–birkeland model. *SIAM Journal on Applied Dynamical Systems* 11(4), 1499–1532.
- Rosen, S., K. M. Murphy, and J. A. Scheinkman (1994). Cattle cycles. *Journal of Political Economy* 102(3), 468–492.
- Wong, A. K. and M. You (1985). Entropy and distance of random graphs with application to structural pattern recognition. *IEEE transactions on pattern analysis and machine intelligence* (5), 599–609.
- Yoccoz, J.-C. and H. Birkeland (1998). Informal commentaries on the numerical investigation of the «toy model».

A Appendix A: Random Dynamical Systems, Random Attractors and Random Measures

universal approximation theorem for continuous functions between Euclidean spaces. In this appendix, we remind some fundamental notions and tools from the theory of random dynamical systems we have used in our analysis of the model. We refer the reader to Crauel and Flandoli (1994); Flandoli and Tonello (2021) (which we closely follow) and ? for further information and more details.

Definition A.1 (Non-Autonomous Dynamical Systems (NADS)). *Let T be either the set of real numbers \mathbb{R} or the set of integer number \mathbb{Z} . Let \mathcal{X} be defined as in (9). A NADS with time T is a family of continuous maps $\varphi(s, t) : \mathcal{X} \rightarrow \mathcal{X}$ indexed by two times $s \leq t$ with $s, t \in T$ satisfying to the following rules:*

1. $\varphi(s, s) = Id_{\mathcal{X}} \forall s \in T$.
2. $\varphi(r, t) \circ \varphi(s, r) = \varphi(s, t) \forall s \leq r \leq t$.

Notice that in the autonomous case a one-parameter family of continuous map would be sufficient to entirely determine the evolution of the system because in this case the evolution is invariant with respect to translation in time, i.e. $\varphi(s, t)x = \varphi(t-s)x$. Instead, in the non-autonomous case the time at which the initial data are prescribed is crucial; as a consequence, it is natural to expect that the “fixed points” of the system depend upon the second variable t by letting $s \rightarrow -\infty$. We now give the notion of attractors in the non-autonomous set-up:

Definition A.2 (Pullback attractor). *A family of objects $A(t)$ in a complete metric phase-space (\mathcal{X}, δ) is a pullback attractor for the NADS φ if it satisfies the following two conditions:*

1. *For all t , $A(t)$ is a compact subset of \mathcal{X} and is invariant with respect to the dynamics, namely, $\varphi(s, t)A(s) = A(t) \forall s \leq t$.*
2. *For all bounded sets B , $\forall \varepsilon > 0$ there exists $s_0 < 0$ such that $\forall s < s_0$ we have $\varphi(s, t)(B) \subset \mathcal{U}_\varepsilon(A(t))$; $\mathcal{U}_\varepsilon(A(t))$ denotes the neighbourhood of radius ε around the set $A(t)$, namely:*

$$\mathcal{U}_\varepsilon(A(t)) = \{x \in \mathcal{X} : \inf_{y \in A(t)} \delta(x, y) < \varepsilon\} \quad (23)$$

We remark that the previous definition can be given also by means of the non-symmetric Hausdorff-like distance δ_h between sets. In this case, point 2. becomes: “for all bounded sets B it holds that $\lim_{s \rightarrow -\infty} \delta_h(\varphi(s, t)(B), A(t)) = 0$ ”. Because of the fact that NADS and RDS are closely related, the notion of stochastic attractor will be based on the latter definition.

A.1 Random Dynamical Systems

We recall that a measurable dynamical system is a tuple $((\Omega, \mathcal{F}, \mathcal{P}), (\theta_t)_{t \in T})$ where $(\Omega, \mathcal{F}, \mathcal{P})$ is a probability space and $\{\theta_t : \Omega \rightarrow \Omega\}_{t \in T}$ is a family of measure preserving transformations of the probability space $(\Omega, \mathcal{F}, \mathcal{P})$ such that $(t, \omega) \rightarrow \theta_t \omega$ is measurable, $\theta_0 = \text{Id}$ and $\theta_{t+s} = \theta_t \circ \theta_s \forall t, s \in T$.

Definition A.3 (RDS). *Let $T = \mathbb{R}, \mathbb{R}_+, \mathbb{Z}$ or \mathbb{N} . A RDS with time T on a metric, complete and separable space (\mathcal{X}, δ) with Borel σ -algebra \mathcal{B} over $(\theta_t)_{t \in T}$ on $(\Omega, \mathcal{F}, \mathcal{P})$ is a measurable map*

$$\begin{aligned} \varphi : T \times \mathcal{X} \times \Omega &\rightarrow \mathcal{X} \\ (t, x, \omega) &\rightarrow \varphi(t, \omega)x \end{aligned}$$

such that $\varphi(0, \omega) = \text{id}_{\mathcal{X}}$ and

$$\varphi(t+s, \omega) = \varphi(t, \theta_s \omega) \circ \varphi(s, \omega) \quad (24)$$

$\forall t, s \in T$ and $\forall \omega \in \Omega$. A family of maps $\varphi(t, \omega)$ satisfying (24) is called a cocycle and (24) is the cocycle property.

A RDS is said to be *continuous* or *differentiable* if $\varphi(t, \omega) : \mathcal{X} \rightarrow \mathcal{X}$ is continuous or differentiable, respectively, $\forall t \in T$ outside a \mathcal{P} -nullset. In addition, $\varphi(t, \omega)$ is automatically invertible if $T = \mathbb{R}$ or \mathbb{Z} ; indeed, in this case, we have $\varphi(t, \omega)^{-1} = \varphi(-t, \theta_t \omega)$ for $t \in T$.

The notion of skew product collects all the ω 's in order to define a (measurable) dynamical system on the product space $(\Omega \times \mathcal{X}, \mathcal{F} \otimes \mathcal{B})$:

Definition A.4 (Skew product). *The measurable map*

$$\begin{aligned} \Theta : T \times \Omega \times \mathcal{X} &\rightarrow \Omega \times \mathcal{X} \\ (t, \omega, x) &\rightarrow (\theta(t)\omega, \varphi(t, \omega)x), \end{aligned}$$

is called the skew product flow of the dynamical system $((\Omega, \mathcal{F}, \mathcal{P}), (\theta_t)_{t \in T})$ and of the co-cycle φ .

In particular, the family of mapping $\Theta_t = \Theta(t, \cdot, \cdot)$ with $t \in T$ is the measurable dynamical system on $(\Omega \times \mathcal{X}, \mathcal{F} \otimes \mathcal{B})$ we were referring to. In addition, it holds that $\Theta_0 = \text{Id}_{\Omega \times \mathcal{X}}$ and $\Theta_{t+s}(\omega, x) = \Theta_t \circ \Theta_s(\omega, x) \forall t, s \in T, \omega \in \Omega$ and $x \in \mathcal{X}$.

A.1.1 Attraction and absorption

We here define the notions of attraction and absorption.

Definition A.5. *A random set A is said to attract another random set B if \mathcal{P} -almost surely*

$$\lim_{t \rightarrow \infty} \delta_h(\varphi(t, \theta_{-t} \omega)B(\theta_{-t} \omega), A(\omega)) = 0.$$

Moreover, we have that a random set $K(\omega)$ is said to be (*strictly*) φ -forward invariant if

$$\varphi(t, \omega)K(\omega) \subset K(\theta_t \omega) \quad (\varphi(t, \omega)K(\omega) = K(\theta_t \omega)) \quad \forall t > 0.$$

In addition, the following definitions hold.

Definition A.6 (Globally attracting set). *Suppose φ is a RDS such that there exists a random compact set $A(\omega)$ which satisfies the following conditions:*

1. $\varphi(t, \omega)A(\omega) = A(\theta_t \omega) \forall t > 0$.
2. A attracts every bounded deterministic set $B \subset \mathcal{X}$.

Then, A is said to be a universally or globally attracting set for φ .

Definition A.7 (Absorption time). *If K and B are random sets such that for \mathcal{P} -almost all ω there exists a time $t_B(\omega)$ such that for all $t \geq t_B(\omega)$ we have*

$$\varphi(t, \theta_{-t} \omega)B(\theta_{-t} \omega) \subset K(\omega),$$

then K is said to absorb B and $t_B(\omega)$ is called the absorption time.

Definition A.8 (Ω -limit set). *Given a random set K , the set*

$$\Omega(K, \omega) = \Omega_K(\omega) = \bigcap_{T \geq 0} \overline{\bigcap_{t \geq T} \varphi(t, \theta_{-t} \omega)K(\theta_{-t} \omega)}$$

is said to be the Ω -limit set of K . By definition, $\Omega_K(\omega)$ is closed.

In particular, it is possible to identify $\Omega_K(\omega)$ with

$$\Omega_K(\omega) = \{y \in \mathcal{X} : \exists t_n \rightarrow \infty, x_n \in K(\theta_{-t_n} \omega) : \varphi(t_n, \theta_{-t_n} \omega)x_n \rightarrow y\}.$$

With this identification, the θ -shift of an Ω -limit set is given by:

$$\begin{aligned} \Omega_K(\cdot) \circ \theta_t &= \Omega(K, \theta_t \omega) = \{y \in \mathcal{X} : \exists t_n \rightarrow \infty, \\ &\quad \exists x_n \in K(\theta_{-t_n+t} \omega) \text{ such that } \varphi(t_n, \theta_{-t_n+t} \omega)x_n \rightarrow y\}. \end{aligned}$$

Also, the following theorem holds:

Theorem A (cfr. Crauel and Flandoli (1994), Theorem 3.11). *Suppose φ is an RDS on the Polish space \mathcal{X} and suppose that there exists a compact set $K(\omega)$ absorbing every bounded non-random set $B \subset \mathcal{X}$. Then the set*

$$A(\omega) = \overline{\bigcup_{B \subset \mathcal{X}} \Omega_B(\omega)}$$

is a global attractor for φ .

A.1.2 Invariant measures on random sets

First, we introduce the concept of random (probability) measure on \mathcal{X} .

Definition A.9 (Random (probability) measure). *A map $\mu : \mathcal{B} \times \Omega \rightarrow [0, 1]$, $(B, \omega) \rightarrow \mu_\omega(B)$ satisfying the following two conditions*

1. *for every $B \in \mathcal{B}$, $\omega \rightarrow \mu_\omega(B)$ is measurable,*
2. *for \mathcal{P} -almost every $\omega \in \Omega$, $B \rightarrow \mu_\omega(B)$ is a Borel probability measure,*

is said to be a random (probability) measure on \mathcal{X} .

Hereafter, we will denote by $\mathcal{P}_\Omega(\mathcal{X})$ the set of random measures on \mathcal{X} and by $\mathcal{P}(\mathcal{X})$ the set of canonical probability measures.

Definition A.10 (Invariant measure). *A random (probability) measure $\mu = (\mu_\omega)_{\omega \in \Omega}$ is said to be invariant for the RDS φ if $\forall t \in T$ and for \mathcal{P} -almost every $\omega \in \Omega$ we have*

$$\varphi(t, \omega) \mu_\omega = \mu_{\theta(t)\omega}.$$

In this work, we will aim at finding an invariant measure via an averaging procedure applied to the initial distribution of the data. To this end, let λ_μ be the measure of which μ is the *factorization*, i.e., the measure defined $\forall A \in \mathcal{F} \otimes \mathcal{B}$ as

$$\lambda_\mu(A) = \int_\Omega \left(\int_{\mathcal{X}} I_A(\omega, x) \mu_\omega(dx) \right) \mathcal{P}(d\omega). \quad (25)$$

The concept of invariance for a random (probability) measure μ is related to the invariance of the measure λ_μ with respect to the skew product; the following proposition holds:

Proposition A.11 (cfr. Flandoli and Tonello (2021), Proposition 72). *The random (probability) measure $\mu = (\mu_\omega)_{\omega \in \Omega}$ is invariant for the RDS φ if and only if the measure λ_μ on $(\Omega \times \mathcal{X}, \mathcal{F} \otimes \mathcal{B})$ is invariant for the skew product associated to φ .*

Hereafter, we will denote by $\mathcal{P}_\mathcal{P}(\Omega \times \mathcal{X})$ the set of probability measures on $(\Omega \times \mathcal{X}, \mathcal{F} \otimes \mathcal{B})$ of the form (25) for some random (probability) measure μ . In particular, Θ_t maps $\mathcal{P}_\mathcal{P}(\Omega \times \mathcal{X})$ into itself. Also, it can be shown that all the probability measures on $(\Omega \times \mathcal{X})$ with marginal \mathcal{P} on Ω have a unique random (probability) measure satisfying Equation (25). This leads to a one-to-one correspondence between $\mathcal{P}_\mathcal{P}(\Omega \times \mathcal{X})$ and $\mathcal{P}_\Omega(\mathcal{X})$. In addition, we denote by $\mathcal{I}_\mathcal{P}(\varphi) \subset \mathcal{P}_\mathcal{P}(\Omega \times \mathcal{X})$ the set of measures λ_μ for which the associated random measure μ is invariant. We give now the following definition.

Definition A.12 (cfr. Flandoli and Tonello (2021), Page 51). *We define $L^1_\mathcal{P}(\omega, \mathcal{C}_b(\mathcal{X}))$ as the space of those functions $f : \Omega \rightarrow \mathcal{C}_b(\mathcal{X})$ such that the map $(\omega, x) \rightarrow f(\omega)(x) = f(\omega, x)$ is measurable and the integral*

$$\|f\|_{1,\infty} := \int_\Omega \sup_{x \in \mathcal{X}} |f(\omega, x)| d\mathcal{P}(\omega)$$

is finite.

In particular, we identify two functions f and g if $\mathcal{P}(f(\cdot, \omega) \neq g(\cdot, \omega)) = 0$; the equivalence class of f will be identified with f .

The space $\mathcal{P}_{\mathcal{P}}(\Omega \times \mathcal{X})$ is endowed with the topology of the weak convergence, which is the smallest topology on $\mathcal{P}_{\Omega}(\mathcal{X})$ such that the maps

$$\mu \rightarrow \mu(f) = \int_{\Omega} \int_{\mathcal{X}} f(\omega, x) d\mu_{\omega}(x) d\mathcal{P}(\omega) = \int_{\Omega \times \mathcal{X}} f(\omega, x) d\lambda_{\mu}(\omega, c)$$

are continuous for each $f \in L^1_{\mathcal{P}}(\omega, \mathcal{C}_b(\mathcal{X}))$. At this point, we consider the action of the skew product (see Definition A.4) on functions $f \in L^1_{\mathcal{P}}(\omega, \mathcal{C}_b(\mathcal{X}))$ given by $\Theta_t f = f \circ \Theta_t$; in particular, such a product belongs to $L^1_{\mathcal{P}}(\omega, \mathcal{C}_b(\mathcal{X}))$ too.

Proposition A.13 (cfr. Flandoli and Tonello (2021), Proposition 73). *If φ is a continuous RDS on a Polish space \mathcal{X} , the map $\mu \rightarrow \Theta_t \mu$ on $\mathcal{P}_{\mathcal{P}}(\Omega \times \mathcal{X})$ is affine and continuous. Moreover, the set $\mathcal{I}_{\mathcal{P}}(\varphi)$ is convex and closed.*

We have now all the theoretical instruments to introduce the averaging process mentioned before, which enables us to state the existence of measures in $\mathcal{I}_{\mathcal{P}}(\varphi)$.

Proposition A.14 (cfr. Flandoli and Tonello (2021), Proposition 74). *Let φ be a continuous RDS on a Polish space \mathcal{X} with continuous time T . Let ν be in $\mathcal{P}_{\mathcal{P}}(\Omega \times \mathcal{X})$. For each $t \in T$, $t > 0$, let μ_t be the measure defined as*

$$\mu_t(A) := \frac{1}{t} \int_0^t (\Theta_s \nu)(A) ds, \quad \forall A \in \mathcal{F} \otimes \mathcal{B}. \quad (26)$$

Then, every limit point of $(\mu_t)_t$ for $t \rightarrow \infty$, in the topology of the weak convergence, is in $\mathcal{I}_{\mathcal{P}}(\varphi)$.

Finally, the existence of limit points for the sequence (26), and then of random invariant measures for φ , can be established through an analogous of Prohorov theorem for random measures.

Definition A.15. *A set of measures $\Gamma \subset \mathcal{P}_{\mathbb{P}}(\Omega \times \mathcal{X})$ is said to be tight if for every $\varepsilon > 0$ there exists a compact set $C_{\varepsilon} \subset \mathcal{X}$ such that, for each $\lambda \in \Gamma$ it holds that $\lambda(\Omega \times C_{\varepsilon}) \geq 1 - \varepsilon$. In other words, we must have $\int_{\Omega} \mu_{\omega}(C_{\varepsilon}) \mathbb{P}(d\omega) \geq 1 - \varepsilon$, where μ_{ω} is the factorization of λ_{μ} .*

Theorem A.16 (Prohorov theorem for Random Measures, cfr. Crauel (2002), Theorem 4.4). *If $\Gamma \subset \mathcal{P}_{\mathbb{P}}(\Omega \times \mathcal{X})$ is tight, then every sequence $(\mu_n)_{n \in \mathbb{N}} \subset \Gamma$ admits a convergent sub-sequence.*

B Appendix B: Measuring distances between attractors: the logistic map case

We show the reliability of the methodology employed in Subsection 6.5 to quantify the distance between two attractors by applying it to the logistic map, i.e. the map $T(t) = rt(1 - t)$, with $r \in [3.5, 4]$ and $\#r = 3000$. For each parameter

r , we let the system evolve for 50000 steps forward, and we record the position of the system in 3000 histograms with 500 bins. We then apply the four presented distances to all of the histograms on the first argument, and to the penultimate histogram on the right argument. The choice of the penultimate histogram is due to the fact that the dynamics when the logistic parameter $r = 4$ is somewhat simpler, being conjugated to the tent map. Thus, we picked the histogram of the more “chaotic” logistic map as a second argument. All the plots have a rather uniform shape, with some spikes corresponding to the stable windows in the bifurcation diagram of the logistic map, and the distance gradually decreasing as the parameter approaches the value of 4.

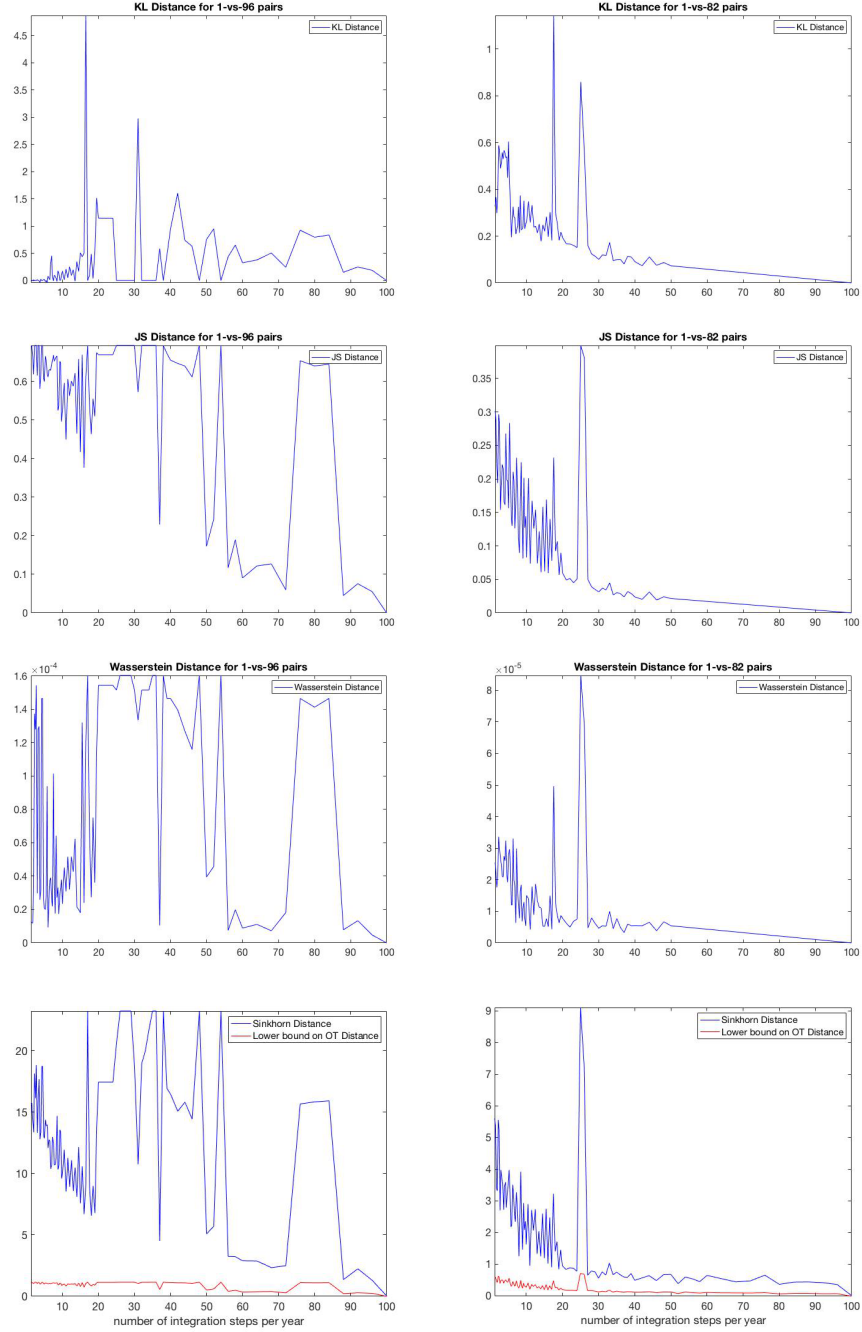


Figure 8: From top to bottom, left panel: distances between the attractors as a function of $q \in [2, 100]$ on a pseudo-logarithmic scale for the set of experimental parameters \mathcal{H}_1 . From top to bottom, right panel: the same quantities for the set of experimental parameters \mathcal{H}_2 . The distances are, in order: the Kullback-Leibler divergence, the Jensen-Shannon distance, the Wasserstein distance and the Sinkhorn distance.

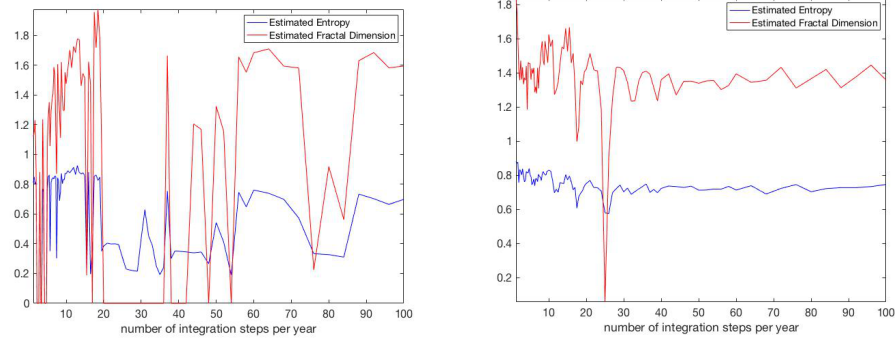
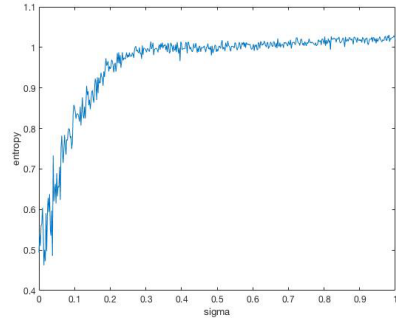
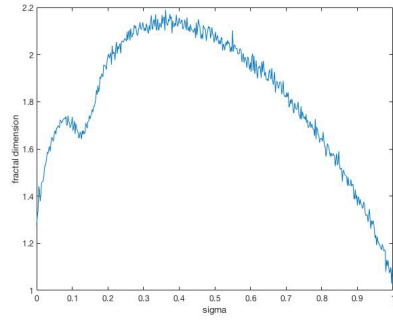


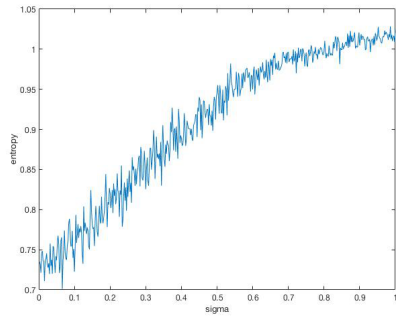
Figure 9: Estimated fractal dimension (*left panel*) and entropy (*right panel*) for different values of $q \in [2, 100]$, on a pseudo-logarithmic scale.



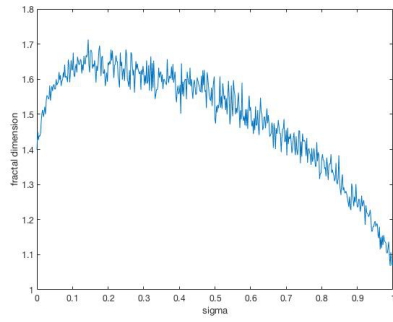
(a) Entropy as function of σ for \mathcal{H}_1



(b) Fractal Dimension in terms of σ for \mathcal{H}_1



(c) Entropy as function of σ for \mathcal{H}_2



(d) Fractal Dimension in terms of σ for \mathcal{H}_2

Figure 10: Estimated fractal dimension and estimated entropy as a function of the parameter σ when the parameters of the deterministic dynamical system are fixed as both in \mathcal{H}_1 (*top panel*) and in \mathcal{H}_2 (*bottom panel*)

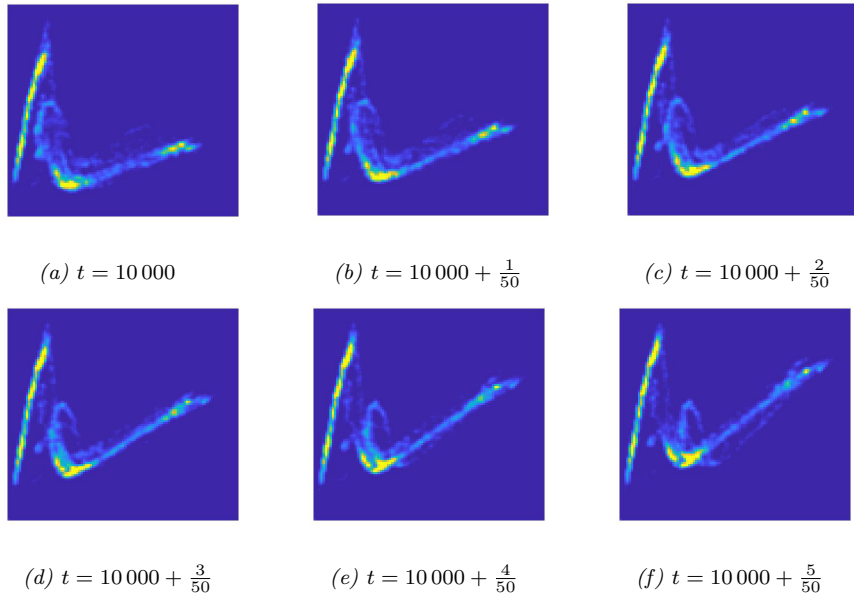
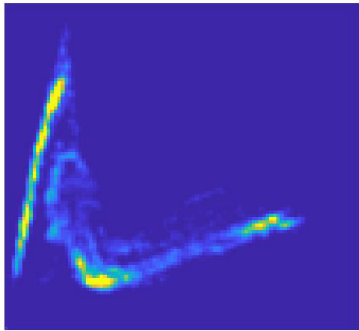
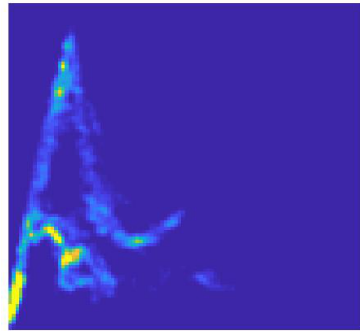


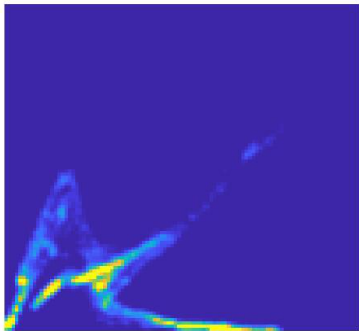
Figure 11: Two dimensional projection of the random attractor.



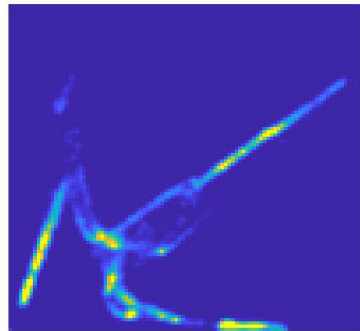
(a) $t = 10\,000$



(b) $t = 10\,001$



(c) $t = 10\,002$



(d) $t = 10\,003$

Figure 12: Two dimensional projection of the random attractor (in the case of consecutive integer values of t).

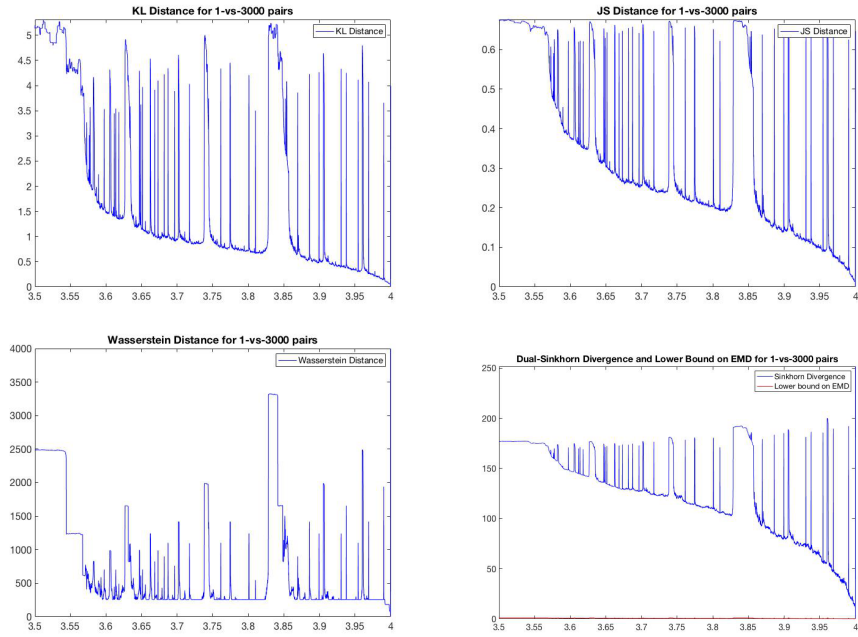


Figure 13: From top to bottom, left panel: distances between the attractors as a function of $r \in [3.5, 4]$ for the logistic map. The distances are, in order: the Kullback-Leibler divergence, the Wasserstein, the Jensen-Shannon distance and the Sinkhorn distance.