Ecological change points: The strength of density dependence and the loss of history

José M. Ponciano[a,*], Mark L. Taper[b], Brian Dennis[c]

[a]*Department of Biology, University of Florida, Gainesville, FL. 32611.*
[b]*Department of Biology, University of Florida. Carr Hall 309, P.O Box 118525, Gainesville FL 32611*
[c]*Department of Fish and Wildlife Sciences and Department of Statistical Science, University of Idaho, Moscow ID 83844-1136, USA*

Abstract

Change points in the dynamics of animal abundances have extensively been recorded in historical time series records. Little attention has been paid to the theoretical dynamic consequences of such change-points. Here we propose a change-point model of stochastic population dynamics. This investigation embodies a shift of attention from the problem of detecting when a change will occur, to another non-trivial puzzle: using ecological theory to understand and predict the post-breakpoint behavior of the population dynamics. The proposed model and the explicit expressions derived here predict and quantify how density dependence modulates the influence of the pre-breakpoint parameters into the post-breakpoint dynamics. Time series transitioning from one stationary distribution to another contain information about where the process was before the change-point, where is it heading and how long it will take to transition, and here this information is explicitly stated. Importantly, our results provide a direct connection of the strength of density dependence with theoretical properties of dynamic systems, such as the concept of resilience. Finally, we illustrate how to harness such information through maximum likelihood estimation for state-space models, and test the model robustness to widely different forms of compensatory dynamics. The model can be used to estimate important quantities in the theory and practice of population recovery.

Keywords: Strength of density dependence; extinction dynamics; breakpoint; change-point stochastic processes; state-space models

1. Introduction

The strength and effect of density dependence has long been a focus of theoretical and applied population ecology [1, 2]. Although the dependence of the per capita growth rate of a species on its own density or abundance is now widely regarded as a main driver of population dynamics, such proposition still instills rich debates, theoretical problems and practical dilemmas (see citations in [3], [4], [5]). Relatively missing so far in theoretical population ecology is consideration of change in ecological conditions. Various environmental changes impact a population: a new species

*Corresponding author
Email address: josemi@ufl.edu (José M. Ponciano)

can enter the habitat, hunting could be removed (or instituted), a long drought could begin (or end). If such a change point occurred within the monitoring history of a population, the degree and effects of the change should be quantified by explicitly including the change in the population dynamics model. We present a general population model with a change point in the dynamics. At the change point, shifts may occur in any or all of the maximum growth rate, the strength of density dependence and the influence of stochastic environments (one form of process variation).

Ecologists have long used statistical methods to detect changes in abundances and trends in abundances [6]. These statistical methods usually test whether an observed time series of population abundances (or densities) depart from a fixed deterministic equilibrium, or stochastic stationary distribution. In the first approach, the observed population abundances are observations with measurement error (observation error) added to a deterministic equilibrium or trend. In the second approach, stochastic perturbations are included in the dynamic model to describe the natural fluctuations in abundance. These natural fluctuations are also known as "process noise" [7]. With process noise, concepts from deterministic ecological accounts of population regulation take on new forms. For instance, a stable equilibrium of population abundances becomes a stationary probability distribution in stochastic population models [8]. These analyses can be termed static in that the deterministic equilibrium or the stochastic stationary distribution are assumed to be unchanging.

The last twenty years of research in statistical population dynamics concentrated on parameter estimation, model selection and hypothesis testing under one or both of these assumptions [3, 9, 10, 11, 12, 7, 13, 14, 15, 16, 17, 18]. This research program is warranted because the effective statistical coupling of empirical observations with mathematical models is critical for testing hypotheses concerning population regulation [19]. However, the ecological investigation of populations undergoing change has lagged behind the attention to statistical issues [7, 20]. This paper seeks to understand what are the ecological consequences of population processes that are not static, but undergo change.

Statistical time series models with change points in population dynamics are rarely considered in ecological analyses [21], despite being extensively studied in the statistical literature (but see [22, 23, 24, 25]). Asymptotics of maximum likelihood parameter estimates have been derived [26, 27, 28, 29]. Hypotheses tests for the presence of change points have been proposed [30, 31, 32, 33, 34, 35]. Quality control theory features studies of how to incorporate change point warnings in control charts [36, 37, 38]. Although these statistical studies provide substantial underutilized opportunities for ecological data analysis, missing to date is an ecological understanding of the dynamical properties of populations undergoing a change point.

The key idea of this paper is to represent an ecological change point as a saltational change in the parameters of the stochastic population model. Such a change produces both a shift in the stationary distribution of population abundance and a transition to a new equilibrium distribution. We obtain simple formulae characterizing the statistical distribution of the population process during the transition between stationary distributions of abundance. Our study demonstrates that certain historical properties of a population's growth dynamics are determined by the strength of density-dependent processes. In particular, a time series transitioning from one stationary distribution to another

contains information about where the process was before, where is it heading, and how long it is going to take to get there. We illustrate how to harness the information in time series containing transitionary portions via maximum likelihood estimation for state-space models. The model proposed here provides new insights into the role of density dependence in shifting environments. This model also represents a practical tool for detecting and predicting change events in population monitoring.

2. The strength of density dependence in discrete time models

Consider the general discrete time population growth model $n_{t+1} = f(n_t) = \lambda(n_t)n_t$, where n_t is the population density or abundance at time t and $\lambda(n_t)$ is the (density-dependent) per capita growth rate. Assume that $f(n)$, the recruitment map, is continuously differentiable and that n^* is its non-trivial equilibrium abundance (*i.e.* its satisfies $f(n^*) = n^*$).

Three measures of the strength of density dependence have been suggested. The first measure is motivated by thinking of the strength of density dependence simply as the marginal effect on the per capita growth rate of an increase in density [1, 39], which according to our general setting, corresponds to $\partial\,\lambda(n_t)/\partial\,n_t$. This is a measure of the effects of density at the individual -not population- level that has been used, among other things, to phrase an evolutionary perspective of intra-specific competition as the way individual, demographic traits respond to an increase in density [39]. Because for some discrete, density-dependent maps the per capita growth rate $\lambda(n_t)$ is written as an exponential function (as in the Ricker or in the Gompertz equations, see [3, 2]), the marginal effect of an increase in density can be conveniently measured and plotted in the log-scale of the per capita growth rate by computing $\partial \ln \lambda(n_t)/\partial\,n_t$.

The second measure corresponds to the derivative of the recruitment map $f(n)$ at equilibrium. Such measure can be directly read from the graph of $f(n)$ as a function of n as the value of the slope when $f(n)$ crosses the 1:1 line. It is important to note that a consideration of both an individual and a population wide measure of density dependence is of obvious importance in population ecology to explain persistence, average abundance, bounds on temporal stability and in general, for a careful understanding of the unfolding of population dynamics [39].

A third measure of the strength of density dependence in a simple model with no age structure was given by Lande et al (2002), who defined it as the negative elasticity of the per capita population growth rate with respect to change in the population. This amounts to computing $-\frac{n}{\lambda(n)}\frac{\partial\,\lambda(n)}{\partial\,n}\Big|_{n=n^*} = -\frac{1}{\lambda(n)}\frac{\partial\,\lambda(n)}{\partial \ln n}\Big|_{n=n^*} = -\frac{\partial \ln \lambda(n)}{\partial \ln n}\Big|_{n=n^*}$. This definition of the strength of density-dependence is insightful and relevant in population dynamics modeling, not only because it provides a scale-free measure of the change in the per capita growth rate as the (log) population size varies, but also because it is readily extendable to scenarios dealing with more complex life histories [40].

3. Methods

3.1. The discrete Gompertz model

The deterministic skeleton of our stochastic model is the discrete-time Gompertz model [7, 2] is given by:

$$n_t = n_{t-1} e^{a + b \ln n_{t-1}}, \tag{1}$$

where n_t is population abundance at time t and a and b are constants. Note that many formulations of the Gompertz model explicitly include a minus sign before b and thus can look superficially different.

The non-trivial equilibrium density, or carrying capacity is given by $n^* = e^{-a/b}$. This equilibrium is asymptotically stable, provided the absolute value of the derivative of the recruitment map (eq. 1) evaluated at n^* is less than 1 [41], *i.e.*, if $|1 + b| < 1$. For this model, the strength of density dependence measured using the per capita growth rate is $\partial \lambda(n_t)/\partial n_t = b e^{a + b \ln n_{t-1}}/n_t$ which is equal to $b/\exp\{-a/b\}$ at equilibrium. The measure of the strength of density dependence using the logarithm of the per capita growth rate, written as $\partial \ln \lambda(n_t)/\partial n_t = b/n_t$ is again equal to $b/\exp\{-a/b\}$ at equilibrium. Finally, Lande et al's negative elasticity is simply given by $-b$ for all population abundances. Then, the density dependent coefficient b, or its carrying capacity-scaled version $b/\exp\{-a/b\}$, is directly involved in the definition of the strength of density dependence. So following [40, 42, 2], heretofore we identify b as the strength of density dependence, due to its role in the rate of approach to a locally stable equilibrium.

Let $x_t = \ln(n_t)$ and $c = b + 1$. Then, Equation (1) becomes the first order difference equation:

$$\begin{aligned} x_t &= x_{t-1} + a + b x_{t-1} \\ &= a + c x_{t-1}, \end{aligned} \tag{2}$$

whose solution can be found by induction

$$x_t = c^t \left(x_0 - \frac{a}{1-c} \right) + \frac{a}{1-c}, \tag{3}$$

provided $c \neq 1$. If $c = 1$ then $b = 0$ and population growth is density independent. The solution in that case is $x_t = x_0 + at$. If density-dependent processes lead to a stable equilibrium density, *i.e*, if $|c| < 1$, then we may distinguish two cases: $0 < c < +1$ and $-1 < c < 0$. In both cases, as $t \to \infty$, the log population density (abundance) $x_t \to \frac{a}{1-c} = \ln n^* \equiv x^*$. In the first case, x_t approaches x^* directly, either from above or below, depending on initialization. In the second case, the deterministic map exhibits damped oscillatory behavior while approaching the equilibrium.

3.2. The stochastic Gompertz model with environmental noise

A model of Gompertz population growth with environmental stochasticity may be written as

$$N_t = N_{t-1} \lambda(N_{t-1}) W_t = N_{t-1} \exp\{a + b \ln N_{t-1}\} W_t,$$

where the population size N_t is now modeled with a random variable representing environmental variation. A useful model for the noise takes the W_t to be independent and identically distributed (iid) log-normal random variables describing heavy-tailed temporal perturbations of the per-capita growth rate [43]. Accordingly, we set $W_t = e^{E_t}$ where $E_t \sim \mathrm{N}(0, \sigma^2)$ to get

$$N_t = N_{t-1} \exp\{a + b \ln N_{t-1} + E_t\}.$$

On the logarithmic scale, with $X_t = \ln N_t$, this stochastic Gompertz model with environmental noise becomes the well known AR(1) process

$$X_t = a + cX_{t-1} + E_t. \tag{4}$$

Provided $|c| < 1$, this stochastic process has a normal stationary distribution with mean and variance given by

$$\begin{aligned} E[X_\infty] &= \lim_{t\to\infty} E[X_t] = \frac{a}{1-c} \\ Var[X_\infty] &= \lim_{t\to\infty} Var[X_t] = \frac{\sigma^2}{1-c^2}. \end{aligned}$$

Thus, both, the process stationary mean (equivalent to the deterministic stable equilibrium) and its associated variance critically depend on the strength of density dependence. Simply put, although external, environmental variability induces a variability in the growth rate, the magnitude and intensity of the population fluctuations are also influenced by the strength of density dependence. This feature led Ives et al (2003) to a detailed study of the concept of ecological stability applied to stochastic community dynamics models. Before moving on to the breakpoint model, we note that through this article, random variables will denoted with capital letters and realizations of these random variables with lower case letters. Also, random vectors like the multivariate random vector $\mathbf{X} = [X_0, X_1, X_2, \ldots, X_q]'$, where the X_i denote the log-population abundance at time i, will be denoted with capital, boldface letters. Realizations of the random vectors will be written with lower case, bold face letters. Finally note that here, as in all the vectors we define, the length of the vector is the last index (q in this case) plus one.

3.3. The breakpoint Gompertz model

We consider the case where the stochastic process model undergoes a change after a given time time step, τ. Operationally, such change is implemented by assuming that the value of one or more of the stochastic Gompertz model parameters changes after time τ. A general model, from which all of the possible changes can be specified as sub-models, specifies that there are two different sets of parameters for the stochastic Gompertz model before and after the change-point. If Z_t denote $iid\,\mathrm{N}(0,1)$ random variables, a general change-point model is written as

$$N_t = \begin{cases} N_{t-1} \exp\{a_1 + b_1 N_{t-1} + \sigma_1 Z_t\} & \text{if} \quad t \le \tau \\ N_{t-1} \exp\{a_2 + b_2 N_{t-1} + \sigma_2 Z_t\} & \text{if} \quad t > \tau. \end{cases} \tag{5}$$

In the log-scale, the model becomes

$$X_t = \begin{cases} a_1 + c_1 X_{t-1} + \sigma_1 Z_t & \text{if} \quad t \le \tau \\ a_2 + c_2 X_{t-1} + \sigma_2 Z_t & \text{if} \quad t > \tau. \end{cases} \tag{6}$$

Here, the environmental noise process is written as $\sigma_i Z_t$ to show model parameter changes after the breakpoint, as opposed to E_t as above. Also, whether or not $t < \tau$ will determine if $E_t \sim \mathrm{N}(0, \sigma_1^2)$ or $E_t \sim \mathrm{N}(0, \sigma_2^2)$, so no additional subscript or notation will be used from now on to distinguish between the environmental noise random variables before and after the breakpoint. Also, as before, we assume that $|c_2| < 1$. Writing the breakpoint model as in equation 6 is a common practice, for instance, among studies of applications of Random Markov Fields in Biology [44]. However, as we show below, writing the breakpoint model in such form, and/or its equivalent matrix representation falls short of a full analysis of the role of the strength of density dependence in shaping the speed of the change and all statistical properties (means, variances and covariances) of the process. After writing the model in its matrix format, we did the explicit matrix algebra multiplications, simplified the expressions and sought out to write the moments of the process as a function of the strength of density dependence. The resulting formulae as well as their ecological interpretation are the main results of this paper. To begin these calculations, by induction (see Appendix 1), we arrive at a matrix equation for the entire random vector of the log abundances before and after the breakpoint:

$$
\begin{bmatrix} X_0 \\ X_1 \\ X_2 \\ X_3 \\ \vdots \\ X_\tau \\ X_{\tau+1} \\ X_{\tau+2} \\ X_{\tau+3} \\ \vdots \\ X_{\tau+q} \end{bmatrix}
= a_1 \begin{bmatrix} 0 \\ 1 \\ \frac{1-c_1^2}{1-c_1} \\ \frac{1-c_1^3}{1-c_1} \\ \vdots \\ \frac{1-c_1^\tau}{1-c_1} \\ c_2 \frac{1-c_1^\tau}{1-c_1} \\ c_2^2 \frac{1-c_1^\tau}{1-c_1} \\ c_2^3 \frac{1-c_1^\tau}{1-c_1} \\ \vdots \\ c_2^q \frac{1-c_1^\tau}{1-c_1} \end{bmatrix}
+ a_2 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ \frac{1-c_2^2}{1-c_2} \\ \frac{1-c_2^3}{1-c_2} \\ \vdots \\ \frac{1-c_2^q}{1-c_2} \end{bmatrix}
$$

$$
+ \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & \dots & \dots & \dots & \dots & \dots & 0 \\
c_1 & 1 & 0 & 0 & 0 & \dots & \dots & \dots & \dots & \dots & 0 \\
c_1^2 & c_1 & 1 & 0 & 0 & \dots & \dots & \dots & \dots & \dots & 0 \\
c_1^3 & c_1^2 & c_1 & 1 & 0 & \dots & \dots & \dots & \dots & \dots & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & 0 \\
c_1^\tau & c_1^{\tau-1} & c_1^{\tau-2} & \dots & c_1 & 1 & 0 & \dots & \dots & \dots & 0 \\
c_2 c_1^\tau & c_2 c_1^{\tau-1} & c_2 c_1^{\tau-2} & \dots & c_2 c_1 & c_2 & 1 & 0 & \dots & \dots & 0 \\
c_2^2 c_1^\tau & c_2^2 c_1^{\tau-1} & c_2^2 c_1^{\tau-2} & \dots & c_2^2 c_1 & c_2^2 & c_2 & 1 & 0 & \dots & 0 \\
c_2^3 c_1^\tau & c_2^3 c_1^{\tau-1} & c_2^3 c_1^{\tau-2} & \dots & c_2^3 c_1 & c_2^3 & c_2^2 & c_2 & 1 & \dots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\
c_2^q c_1^\tau & c_2^q c_1^{\tau-1} & c_2^q c_1^{\tau-2} & \dots & c_2^q c_1 & c_2^q & \dots & c_2^3 & c_2^2 & c_2 & 1
\end{bmatrix}
\begin{bmatrix} X_0 \\ E_1 \\ E_2 \\ E_3 \\ \vdots \\ E_\tau \\ E_{\tau+1} \\ E_{\tau+2} \\ E_{\tau+3} \\ \vdots \\ E_q \end{bmatrix}
\tag{7}
$$

or

$$\mathbf{X} = a_1\mathbf{g_1} + a_2\mathbf{g_2} + \mathbf{\Gamma Z}. \tag{8}$$

Note that vector $\mathbf{Z}$ is Multivariate Normal (MVN) with mean

$$\mathrm{E}[\mathbf{Z}] = \mathbf{m} = \begin{bmatrix} \frac{a_1}{1-c_1} & 0 & 0 & \dots & 0 \end{bmatrix}'_{1\times(\tau+q+1)}$$

and variance-covariance matrix

$$\mathrm{Var}[\mathbf{Z}] = \mathbf{\Phi} = \begin{bmatrix} \frac{\sigma_1^2}{1-c_1^2} & 0 & 0 & \dots & \dots & \dots & \dots & 0 \\ 0 & \sigma_1^2 & 0 & \dots & \dots & \dots & \dots & 0 \\ \vdots & \vdots & \ddots & \dots & \dots & \dots & \dots & 0 \\ 0 & \dots & 0 & \sigma_1^2 & 0 & \dots & \dots & 0 \\ 0 & \dots & \dots & 0 & \sigma_2^2 & 0 & \dots & 0 \\ 0 & \dots & \dots & \dots & 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & \dots & \dots & \dots & \dots & \dots & 0 & \sigma_2^2 \end{bmatrix},$$

where the changes in the diagonal occur in the 2nd and in the $(\tau+2)^{\text{th}}$ elements. Because a linear transformation of a multivariate normal distribution is again multivariate normal (Theorem 3.3.5 p. 101 Graybill 1976), it immediately follows that

$$\mathbf{X} \sim MVN\left(\boldsymbol{\mu_\tau} = a_1\mathbf{g_1} + a_2\mathbf{g_2} + \mathbf{\Gamma m}, \mathbf{\Sigma_\tau} = \mathbf{\Gamma\Phi\Gamma}'\right), \tag{9}$$

where the subscript $\boldsymbol{\tau}$ indexes the mean and the variance of the process with a change-point, to explicitly differentiate them from the mean $\boldsymbol{\mu}$ and variance $\mathbf{\Sigma}$ of the process without a change-point (see Appendix 1).

3.4. Parameter estimation

A state space model where sampling error is compounded to the process error can be easily specified, and that is useful for parameter estimation. We call this model the Break-Point Gompertz State-Space (BPGSS) model. Using equation 9 and a lognormal sampling error process (see the discussion of [7] for a justification of such model), the multivariate distribution of the log-observations with added sampling error $\mathbf{Y} = \mathbf{X} + \mathbf{F}$ is

$$\mathbf{Y} \sim MVN\left(\boldsymbol{\mu_\tau} = a_1\mathbf{g_1} + a_2\mathbf{g_2} + \mathbf{\Gamma m}, \mathbf{V_\tau} = \mathbf{\Gamma\Phi\Gamma}' + \nu^2\mathbf{I}\right). \tag{10}$$

This MVN distribution for the observations follows since both, the pdf for the process X and for the observation error F are both MVN, just as in the model of [7]. Parameter estimation via Maximum Likelihood (ML) for this state space model then proceeds by maximizing this multivariate pdf over the model parameters. This pdf can be

readily programmed as it is written above, but the computations can be rendered much more efficient when the matrix multiplications in expression (10) are carefully simplified (see Appendices 1 and 2). Sample code for parameter estimation is uploaded as supplementary material.

3.5. Testing the robustness of the BPGSS model

Clearly, the BPGSS model is a simplified representation of a population dynamics undergoing a drastic change. One of the most important questions for practicing ecologists is to what extent these simple population dynamics models can be used to infer intrinsic properties of the system that can be of immediate utility. Therefore, we performed a validation effort via simulations that aimed at testing the robustness of the quantitative and qualitative model inferences derived from this model. To do that, we simulated breakpoint dynamics by relaxing most of the major simplifying assumptions made by the BPGSS model and adding biological complexity. Then, the simulated dynamics (*i.e.* time series) were saved and used to confront the BPGSS model with different ecological scenarios of interest.

The discrete time, discrete state stochastic process used to simulate population dynamics was a Markov process that, besides incorporating sampling error, accounted for demographic stochasticity and different forms of environmental variability. Most of the environmental noise models, like the BPGSS model, incorporate stochasticity in the maximum growth rate. Such noise is multiplicative in the scale of population abundances, and additive in the scale of the logarithmic per capita growth rate. Another possibility, however, is to assume temporal randomness in the strength of density dependence. This approach assumes that the environmental forces shaping the population dynamics affect the intensity of density-dependence rather than the maximum growth rate and to date, remains a rarely explored model alternative (but see [45, 46].

Our discrete-time Markov model of reproduction and survival followed the general discrete time model construction presented by [47, 48] and later revisited and expanded by [49] and [46]. A detailed explanation of such construction is given in the supplemental material of [49]. Briefly, population growth from one generation to the other is modeled as a two-stage, stochastic process: at time t the n_t adults present in the population give birth each to a random number of offspring (*i.e.* $0, 1, 2, 3, \ldots$). Then, we assume that both adults and offspring survive to time $t+1$ according to a density dependent survival probability, which takes on one of the well-known discrete-time density-dependent maps (Ricker, Hassell, Gompertz, Below, etc...). This reproduction and survival process considers only demographic stochasticity. To include environmental stochasticity, [47] and [48] have previously assumed before that the mean of the offspring distribution varies randomly from one time step to the next according to a stochastic environmental process. Such variability mimics the random, temporal fluctuation in the quality of the environment that leads to temporal fluctuations of the average number of offspring produced by every adult. As recently shown by [46], this is but one way of incorporating environmental variability. Here, we assumed that the environment randomly perturbs both the mean of the offspring distribution and the density dependent survival probability. For our simulations, we assumed that the random (over time) environmental process, denoted Λ, had independent and identically distributed realizations from a $\mathrm{Gamma}(k, \alpha)$ distribution. The results we present here use the value of the scale parameter α set to 1000 and

the shape parameter k set to 1500, but hold for other parameter values as well.

Accordingly, conditional on the realized value λ of the random environmental process, each individual, independently from each other, was assumed to have a $\text{Poisson}\,(\lambda)$ offspring distribution. As a result, from time t to time $t+1$, the (conditional) number of potential recruits was distributed $\text{Poisson}\,(n_t\lambda)$. We then modeled offspring survival with a density dependent, binomial thinning process. The survival probability of the offspring (potential recruits) was taken to be an explicit function of the population size (density) at time t, denoted as n_t by writing it as $p_t = \frac{1}{1+(\lambda-1)(n_t/K)^{\beta}}$. This function for the offspring survival probability p_t corresponds to Below's recruitment map [50]. The parameter β controls the type of compensatory dynamics (under compensatory ($\beta < 1$), compensatory ($\beta = 1$), over compensatory ($\beta > 1$) and effectively changes the shape and curvature of the recruitment map. Accounting for binomial survival when the number of potential recruits is Poisson distributed is well known to be a special case of a "randomly stopped sum" and results in a Poisson distributed total number of recruits from one time to the next [47]. Here, the (random) total number of recruits N for the next generation, is denoted as N_{t+1}. From the previous assumptions, it follows that $N_{t+1}|(N_t = n_t, \Lambda = \lambda) \sim \text{Poisson}\,(\lambda n_t p_t)$. Note again that because the realized value of the environmental process λ modifies the density dependent term in Below's model, the effect of the stochastic environmental process we specified is qualitatively different from the environmental noise formulation proposed in our BPGSS model. This stochastic model formulation where the stochastic environmental process affects the density-dependent coefficient was recently studied by [46]. Introducing this type of variability changes the well-known scaling relation between the population variance and population abundances in stochastic growth models, and may be important in many animal populations [46]. Averaging the environmental process over time results in the marginal transition pdf of the process. This transition, however, does not result in the well-known demographic and environmental negative binomial process [47], because of the complex environmental process effects.

To incorporate different density-dependent scenarios, we used three values for β: 0.5,1 and 1.2 . These values represent respectively under-compensatory, compensatory and over-compensatory dynamics. This axis of variation allowed us to test the robustness to different forms of density dependence. The simulations were started at the value of the deterministic skeleton's carrying capacity in the log-scale (14) and a change point of dynamics was introduced at time $t = 40$. The change consisted of a change in the deterministic skeleton carrying capacity, from 14 to 4.48 (in the log-scale). Although we introduced a single change at a single time point, for each value of β it took some time for the time series to "stabilize" around the second deterministic carrying capacity. The time it takes for the average of the process to reach the arithmetic average of the two (log) deterministic carrying capacities is heretofore referred to as the "half-life of the change". The resulting simulated population dynamics was sampled 1000 times per value of β using a multiplicative lognormal sampling error with parameters $\mu = 0$ and $\nu^2 = 0.2315$. This sampling resulted in 1000 time series of length 200 per value of β that were taken as data sets to which the BPGSS model was fitted. Furthermore, for each value of β, the "true" half-life of the change point was computed numerically as the average half-life in 50000 simulated time series. We then computed the relative bias of our half-life estimator for each one of

the 1000 simulated time series to which the BPGSS model was fitted.

4. Results

Understanding the breakpoint dynamics

The vector expressions in the multivariate normal distribution of the log-normal abundances (eq. 9) conceal explicit, simple expressions linking the dynamics before and after the change. In particular, immediately after the change, there is a transitional period. The time-varying distribution of this transition contains information about both, the past stationary mean and variance, as well as the future stationary mean and variance. This information is exposed by partitioning the vector $\mathbf{X}$ of log-population abundances over time, into a vector of pre-breakpoint log-abundances $\mathbf{X_1} = \begin{bmatrix} X_0 & X_1 & X_2 & \dots & X_\tau \end{bmatrix}'$ and another of post-breakpoint log-abundances $\mathbf{X_2} = \begin{bmatrix} X_{\tau+1} & X_{\tau+2} & \dots & X_{\tau+q} \end{bmatrix}'$. In what follows, we state six main results along with three corollaries that follow from this partitioned structure. These results follow from a detailed algebraic examination of the vector and matrix calculations resulting from the model formulation in equations 7-9. These algebraic examinations, necessary to fully understand the results, are presented in appendices 1 and 2. Each one of these results brings a different dimensionality to the understanding of the transitional dynamics in ecological terms. At the end, we also list the result regarding the robustness tests of the BPGSS model.

Result 1: *In transition, the mean of the (log) population process becomes a weighted average of the stationary mean before the breakpoint, and the stationary mean predicted by the new set of dynamical parameters. The weights of this weighted average are given by c_2^t and $1 - c_2^t$ respectively, where t is the number of time steps after the break.*

The result follows directly from the expression for the mean of the multivariate distribution in eq. 9. By developing and explicitly writing the entries of the vector and matrix calculations in this expression, we found that the mean vector of the first partition $\mathbf{X_1}$ is given by $\boldsymbol{\mu_1} = \frac{a_1}{1-c_1}\mathbf{j}_{(\tau+1)\times 1}$, where $\mathbf{j}_{(\tau+1)\times 1}$ is a vector of ones, whereas the mean of the second partition is given by the time-varying weighted average:

$$\boldsymbol{\mu_2} = \begin{bmatrix} \frac{a_1}{1-c_1}c_2 + \frac{a_2}{1-c_2}(1-c_2) \\ \frac{a_1}{1-c_1}c_2^2 + \frac{a_2}{1-c_2}\left(1-c_2^2\right) \\ \frac{a_1}{1-c_1}c_2^3 + \frac{a_2}{1-c_2}\left(1-c_2^3\right) \\ \vdots \\ \frac{a_1}{1-c_1}c_2^q + \frac{a_2}{1-c_2}(1-c_2^q) \end{bmatrix}. \tag{11}$$

The first element in the vector above (eq. 11) is the mean of the transition process one time step after the breakpoint. This element is a weighted average of the first stationary mean $\mu_1 = a_1/(1-c_1)$ and the second stationary mean $\mu_2 = a_2/(1-c_2)$. The weights are given by the strength of density-dependence after the breakpoint, c_2. Thus, this coefficient completely determines how important the mean of the first phase of the process is right after the change. Two time-steps after the breakpoint, the weights are given by c_2^2 and $1 - c_2^2$ respectively. Again, because $0 < c_2 < 1$,

$c_2^2 < c_2$ and hence, the influence of the first stationary means decays two time steps after the breakpoint. t time units after the breakpoint, the weights become c_2^t and $1 - c_2^t$. As a result, as the number of time steps t after the breakpoint grows large, the weight for the first stationary mean goes to 0 while the weight of the second stationary mean converges to 1. It then follows that the mean of the post-breakpoint process truly represents a transitionary trend.

Result 2: *The speed of the change of the mean of the log-population process (eq. 9) is determined by the strength of density dependence. The half life of the change in means, or the average time $\bar{t}$ it takes for the average of the process to reach $(\mu_1 + \mu_2)/2$ from μ_1, is a quantity that directly depends on the strength of density dependence post breakpoint. Thus, any property of this dynamic system that depends on the waiting time to complete the transition, also depends on the strength of density dependence.*

An explicit expression for the half life of the transition in means is readily found by setting

$$\frac{a_1}{1-c_1} c_2^{\bar{t}} + \frac{a_2}{1-c_2}\left(1 - c_2^{\bar{t}}\right) = \frac{a_1}{1-c_1}\frac{1}{2} + \frac{a_2}{1-c_2}\frac{1}{2}$$

and solving for $\bar{t}$, which yields

$$\bar{t} = \ln(1/2)/\ln c_2. \tag{12}$$

Thus, the speed of the change from one stationary mean to the other one is determined by the strength of density dependence after the breakpoint. For example, if c_2 is close to 0, then the transition weights c_2^t and $1 - c_2^t$ (see eq. 11) converge fast to 0 and 1, respectively and the process approaches the second stationary mean faster. A detailed characterization of the speed of the change as a function of c_2 is given in Figs. 1 and 2. From this characterization it is readily apparent that when $1/2 < c_2 < 1$ the strength of density dependence can be thought as weak; strong when $-1/2 < c_2 < 1/2$ and as very strong when $-1 < c_2 < -1/2$ (see legend in Fig. 2).

Result 3: *The covariance between a population abundance before the breakpoint and after the breakpoint decays geometrically over time in powers of c_2 to the right of the breakpoint and in powers of c_1 to the left of the breakpoint. The intensity of such covariance is always proportional to the variance of the stationary distribution of the first partition.*

The partitioned representation of the log-population process $\mathbf{X}$ is completed by writing its variance covariance matrix as a block matrix:

$$\boldsymbol{\Sigma_\tau} = \left[\begin{array}{c|c} \boldsymbol{\Sigma_{11}}_{(\tau+1)\times(\tau+1)} & \boldsymbol{\Sigma_{12}}_{(\tau+1)\times q} \\ \hline \boldsymbol{\Sigma_{21}}_{q\times(\tau+1)} & \boldsymbol{\Sigma_{22}}_{q\times q} \end{array}\right], \tag{13}$$

where the diagonal matrices $\boldsymbol{\Sigma_{ii}}$, $i = 1, 2$ are the covariance matrices for the pre and post-breakpoint elements of $\mathbf{X}$ respectively. Accordingly, the sub-matrix $\boldsymbol{\Sigma_{11}}$ specifies the variance of the process at every time point before the breakpoint and the covariances among these points. Likewise, the sub-matrix $\boldsymbol{\Sigma_{22}}$ specifies the variances at every point after the breakpoint, along with the covariances among these points. As we'll see below, such variances and

covariances also bear information about the process before the breakpoint. Finally, the sub-matrix $\mathbf{\Sigma_{21}}$, which is equal to $\mathbf{\Sigma_{12}}'$, specifies the covariance $Cov(\mathbf{X_2}, \mathbf{X_1})$, between any time point before the breakpoint and any point after it. Remarkably, these unwieldy expressions reduce to simple mathematical forms that can be easily interpreted. To find these explicit expressions, we first recall that $\mathbf{\Sigma_\tau} = \mathbf{\Gamma\Phi\Gamma}'$, and thus that it is convenient to also partition the matrices $\mathbf{\Gamma}$ and $\mathbf{\Phi}$ into four sub-matrices matching the dimensions of the $\mathbf{\Sigma_\tau}$ blocks. That is,

$$\mathbf{\Gamma} = \left[\begin{array}{c|c} \mathbf{\Gamma_{11}}_{(\tau+1)\times(\tau+1)} & \mathbf{\Gamma_{12}}_{(\tau+1)\times q} \\ \hline \mathbf{\Gamma_{21}}_{q\times(\tau+1)} & \mathbf{\Gamma_{22}}_{q\times q} \end{array}\right], \text{ and } \quad \mathbf{\Phi} = \left[\begin{array}{c|c} \mathbf{\Phi_{11}}_{(\tau+1)\times(\tau+1)} & \mathbf{\Phi_{12}}_{(\tau+1)\times q} \\ \hline \mathbf{\Phi_{21}}_{q\times(\tau+1)} & \mathbf{\Phi_{22}}_{q\times q} \end{array}\right].$$

Next, noting that $\mathbf{\Gamma_{12}} = \mathbf{0}$, the variance-covariance of the process can be re-written as

$$\mathbf{\Sigma_\tau} = \mathbf{\Gamma\Phi\Gamma}' = \left[\begin{array}{c|c} \mathbf{\Gamma_{11}\Phi_{11}\Gamma_{11}}'_{(\tau+1)\times(\tau+1)} & \mathbf{\Gamma_{11}\Phi_{11}\Gamma_{21}}'_{(\tau+1)\times q} \\ \hline \mathbf{\Gamma_{21}\Phi_{11}\Gamma_{11}}'_{q\times(\tau+1)} & \mathbf{\Gamma_{22}\Phi_{22}\Gamma_{22}}'_{q\times q} + \mathbf{\Gamma_{21}\Phi_{11}\Gamma_{21}}'_{q\times q} \end{array}\right]. \tag{14}$$

Elementary matrix multiplications and an examination of the power patterns (see Appendix 2) leads to a simple expression of the elements of the matrix $\mathbf{\Sigma_{12}} = \mathbf{\Gamma_{11}\Phi_{11}\Gamma_{21}}'$, which represent the covariance between any process realization before the breakpoint, $X_0, X_1, \ldots, X_\tau$, and any process realization after the breakpoint, $X_{\tau+1}, X_{\tau+2}, \ldots, X_{\tau+q}$:

$$Cov(X_i, X_j) = \frac{\sigma_1^2}{1-c_1^2} c_1^{\tau-i} c_2^{j-\tau},\ i = 0, 1, \ldots, \tau, \text{ and } j = \tau+1, \tau+2, \ldots, \tau+q. \tag{15}$$

Thus, these covariances are proportional to the variance of the stationary distribution of the first partition and their strength increases from the left of the breakpoint in powers of c_1, and to the right of such point, it decreases in powers of c_2.

Result 4: *The variance-covariance matrix $\mathbf{\Sigma_{22}}$ can be decomposed into two variance component matrices. An examination of these components reveals exact analytical expressions detailing information about the past and the future stationary distribution of the population process concealed in the post-breakpoint time series.*

While the sub-matrix $\mathbf{\Sigma_{11}}$ is identical in form to the variance covariance matrix of the process without a breakpoint (eq. A.6, except that with dimensions $(\tau+1) \times (\tau+1)$), the sub-matrix $\mathbf{\Sigma_{22}}$ is the sum of two matrices, $\mathbf{\Sigma_{22}}^{(1)} = \mathbf{\Gamma_{21}\Phi_{11}\Gamma_{21}}'_{q\times q}$ and $\mathbf{\Sigma_{22}}^{(2)} = \mathbf{\Gamma_{22}\Phi_{22}\Gamma_{22}}'_{q\times q}$.

Result 4.1 *The first variance component of the matrix $\mathbf{\Sigma_{22}}$, given by $\mathbf{\Sigma_{22}}^{(1)} = \mathbf{\Gamma_{21}\Phi_{11}\Gamma_{21}}'_{q\times q}$, is proportional to the variance of the stationary distribution in the first partition. The weight of such stationary variance decays over time, in powers of c_2.*

After performing the matrix multiplications in the right hand side (RHS) of $\mathbf{\Sigma_{22}}^{(1)} = \mathbf{\Gamma_{21}\Phi_{11}\Gamma_{21}}'_{q\times q}$ and simplifying the resulting expressions (see Appendix 2) the elements of $\mathbf{\Sigma_{22}}^{(1)}$ are found to be:

$$\left\{\mathbf{\Sigma_{22}}^{(1)}\right\}_{i,j} = \frac{\sigma_1^2}{1-c_1^2}c_2^{i+j}, \text{ where } i,j = 1,\ldots,q. \tag{16}$$

Thus, the first variance component of the second time series partition is proportional to the variance of the stationary distribution in the first partition, $\frac{\sigma_1^2}{1-c_1^2}$. The influence of such stationary variance decays over time, in powers of the second dynamics' density-dependent effect, c_2.

Result 4.2 *The second variance component is proportional to the stationary variance of the second partition. These covariances measure the departure of the variance-covariance matrix of the second partition from its limiting form at stationarity. The magnitude of the departure from the limiting covariances is controlled by the strength of density dependence in the second partition, modulated by the positioning of the observations relative to each other and to the location of the breakpoint.*

The elements of the second matrix $\mathbf{\Sigma_{22}}^{(2)} = \mathbf{\Gamma_{22}\Phi_{22}\Gamma_{22}}'_{q\times q}$ are given by

$$\left\{\mathbf{\Sigma_{22}}^{(2)}\right\}_{i,j} = \frac{\sigma_2^2}{1-c_2^2}\left(c_2^{|i-j|} - c_2^{i+j}\right), \text{ where } i,j = 1,\ldots,q. \tag{17}$$

To understand this formula, note that $\frac{\sigma_2^2}{1-c_2^2}c_2^{|i-j|}$ is the covariance between two observations $|i-j|$ time units apart, had the second partition arisen directly from the stationary distribution of the process parameterized with the second set of parameters. Thus, this second matrix is a measure of the departure of the variance-covariance matrix of the second partition from what the variance-covariance matrix would be if the process, right after the breakpoint, had been drawn from the second stationary distribution. Thus, it follows that the matrix $\mathbf{\Sigma_{22}}$ can be expressed as a sum of two variance components, one directly related with the stationary variance of the process from the first partition, and another measuring the divergence from stationarity in the second partition in the variance-covariance space. In the particular case when $i = j$ (*i.e.* for the diagonal elements of $\mathbf{\Sigma_{22}}$), this sum gives the variance of the elements after the second partition. This brings us to our result:

Result 4.3 *The variances of the log-observations after the breakpoint (the diagonal elements of the variance-covariance matrix $\mathbf{\Sigma_{22}}$) can be written as a weighted average of the stationary distribution before and after the breakpoint.*

The developed representation of the transitional variances using the variance decomposition from "Result 4.1 and 4.2" turns out to be analogous to the the equation for the transtional means (eq. 11). Specifically, these variances can

be written as a weighted average of the variance of the stationary distribution before and after the breakpoint:

$$\{\mathbf{\Sigma_{22}}\}_{i,i} = \frac{\sigma_1^2}{1-c_1^2}c_2^{2i} + \frac{\sigma_2^2}{1-c_2^2}\left(1-c_2^{2i}\right). \tag{18}$$

The half life of the change in variance, denoted $\tilde{t}$ and found by setting

$$\frac{\sigma_1^2}{1-c_1^2}c_2^{2\tilde{t}} + \frac{\sigma_2^2}{1-c_2^2}\left(1-c_2^{2\tilde{t}}\right) = \frac{\sigma_1^2}{1-c_1^2}\frac{1}{2} + \frac{\sigma_2^2}{1-c_2^2}\frac{1}{2},$$

and solving for $\tilde{t}$ is given by

$$\tilde{t} = \frac{\ln 1/2}{2\ln c_2} = \frac{\bar{t}}{2}. \tag{19}$$

Therefore, the variance of the process transitions twice as fast as the mean of the process from one dynamical scenario to the next.

Result 5: *The variance-covariance matrix of the observations* $\mathbf{Y}$*, the process with added sampling error, is written as the sum of the variance-covariance of the process and the variance-covariance matrix of the observation error model.*

According to the sampling error setting (see *Parameter Estimation* subsection above),

$$\begin{aligned}\mathbf{V_\tau} &= \left[\begin{array}{c|c}\mathbf{V_{11}}_{(\tau+1)\times(\tau+1)} & \mathbf{V_{12}}_{(\tau+1)\times q}\\ \hline \mathbf{V_{21}}_{q\times(\tau+1)} & \mathbf{V_{22}}_{q\times q}\end{array}\right]\\ &= \left[\begin{array}{c|c}\mathbf{\Gamma_{11}\Phi_{11}\Gamma_{11}}' + \nu^2\mathbf{I} & \mathbf{\Gamma_{11}\Phi_{11}\Gamma_{21}}'\\ \hline \mathbf{\Gamma_{21}\Phi_{11}\Gamma_{11}}' & \mathbf{\Gamma_{21}\Phi_{11}\Gamma_{21}}' + \mathbf{\Gamma_{22}\Phi_{22}\Gamma_{22}}' + \nu^2\mathbf{I}\end{array}\right].\end{aligned} \tag{20}$$

From this development, it follows that the log-likelihood needed for parameter estimation accounting for a change-point in the dynamics occurring between times τ and $\tau+1$ is simply the probability density function (pdf) of such MVN distribution evaluated at the observations:

$$\ln L(a_1, c_1, \sigma_1^2, a_2, c_2, \sigma_2^2, \nu^2) = -\frac{\tau+q+1}{2}\ln(2\pi) - \frac{1}{2}\ln|\mathbf{V_\tau}| - \frac{1}{2}(\mathbf{y}-\boldsymbol{\mu_\tau})'\mathbf{V_\tau}^{-1}(\mathbf{y}-\boldsymbol{\mu_\tau}). \tag{21}$$

A by-product of this multivariate normal formulation is that the problem of estimation in the presence of missing observations in the time series is readily solved (see Dennis and Ponciano 2014): suppose that a subset s of elements of the vector $\mathbf{Y}$ is missing. Then, the likelihood function of the remaining observations is simply written as in eq. 21 but by removing the sth elements in the mean vector $\boldsymbol{\mu_\tau}$ and the sth rows and columns of the matrix $\mathbf{V_\tau}$.

Result 6: *Despite our model being a simplified representation of a change in population dynamics, statistical inferences from this model are robust to biologically realistic departures of the model assumptions.*

After simulating from a model with both, demographic and environmental stochasticities as well as sampling error, the half-life estimates obtained from fitting the BPGSS model appear reliable and with almost no bias, even at the most extreme level of simulated compensatory dynamics (see Fig. 3). The computer code in R used to do the simulations, parameter estimation and plotting of each one of the three figures is provided as an online Supplementary file.

5. Discussion

In this study, we show that after a saltational change in population dynamics' parameters triggered by a change in environmental conditions, the transition to new equilibrial conditions is governed by the strength of density dependence. In response to a change in environmental conditions, the stationary probability distribution of a species' population size undergoes a transition and asymptotically approaches a new equilibrium distribution. The statistical properties, means, variances and covariances, respond to change in different ways (Results 1 through 6 above). Remarkably, in all of them it is density dependence what governs the speed of change. In all of these statistical properties, the history of the past persists and the future is anticipated.

According to our second result, the strength of density dependence controls the average of a first passage time, the average time it takes to complete the transition from one stationary mean to the other. This result automatically connects the strength of density dependence with any theoretical dynamic concept. Take for instance the concept of "potential functions". For continuous time, deterministic population dynamics models of the form $dn/dt = m(n)$, the potential function is defined as the function $u(n)$ satisfying $\frac{du(n)}{dn} = m(n)$ [51]. This function is deeply connected with deterministic waiting times. As Dennis et al [51] point out, the waiting time $t(n)$ needed to reach abundance n from an initial abundance x can be written as a function of the potential function as follows

$$t(n) = \int_x^n \frac{dy}{m(y)} = \int_x^n \frac{dy}{-u'(y)}. \tag{22}$$

Thus, the formulation in equation 22 aids ecological interpretation because it gives to the potential function a direct connection with $t(n)$.

The famous "marble in a bowl" conceptual image has been extremely fruitful in Ecology to explain classical dynamic ideas like resilience [52]. The potential function $u(n)$ is the mathematical representation of the "bowl" in such image. Just as the potential function is connected to deterministic waiting times (equation 22), one could expect our average first passage times, that depend on the strength of density dependence, to be connected with a stochastic formulation of the potential well. In fact, as depicted in ecological textbooks, the "bowl" is often shaken (*e.g.* Figure 6.15 in [52]). This stochastic formulation would then be the mathematical representation of such perturbations. Importantly, following our results this representation of a stochastic well will not only depend on the strength of density-dependence but also on the variance properties of the system, which we have concisely described in this paper. We are developing a full treatment of the stochastic potential well in a separate manuscript.

Despite our model being a simplified representation of a change in population dynamics, statistical inferences from this model are robust to biologically realistic departures of the model assumptions. Indeed, here we show that the

inference via ML seems to be robust to the inclusion of an atypical yet realistic form of environmental noise: temporal variability in the strength of density dependence, recently shown to be important in animal populations [46] (Fig. 3).This robustness confers the model the generality that is needed for sound hypothesis-driven learning in population dynamics modeling. After simulating from a model with both, demographic and environmental stochasticities as well as sampling error, the half-life estimates obtained from fitting the BPGSS model appear reliable and with almost no bias, even at the most extreme level of simulated compensatory dynamics (see Fig. 3). This bias in ML estimates of population dynamic models has been previously reported and remediated using REML estimation. In our applied conservation biology work (not shown here) using the GSS model, we have successfully applied the REML bias reduction technique (for an introduction to this topic in population dynamics models see [7]). A full exploration of this bias-reduction issue warrants an independent research study. Therefore, we believe that after sufficient investigation and testing our model may turn out to be sufficient for understanding, management and importantly, prediction. The latter is the focus of our ongoing Population Viability Monitoring (VPM) programatic research [53].

A more complete characterization through simulation of the statistical properties of the estimators of this model is in progress and will be reported elsewhere. In particular, a computational study of the properties of the change point parameter estimate and of the biological parameters as a function of the length of the time series before and after the break are still needed. Given the covariance structure of the model's second partition (*i.e.* the time series post-change point), we expect the length of the first partition, as well as the size of the change to be critical to be able to correctly estimate the covariance structure in the second partition. In that respect, the location of the change point within the time series is expected to play a critical role. We also expect that these statistical studies will reveal a suitable approximation for the usually less biased restricted maximum likelihood (REML) estimators. In any case, however, if a change point is suspected to be present in a particular data set, extensive statistical tests should be carried to confront our change-point model to a null model of no change, and possibly, to other types of dynamics (see for instance [54]). Interestingly, because the covariance structure of the second partition also contains information about the parameters in the first partition, one could envision estimating the pre-breakpoint parameters even if only post-breakpoint time series data immediately after the breakpoint is available.

The original motivation for this paper was an empirical study of the population dynamics of a salmonid species. In particular, we were aiming at modeling the reaction of the population dynamics of a Bull Trout population (Salvelinus confluentus) to a drastic change in the community composition [55]. Such change was noticeable in the annual counts of adult females carried by Montana Fish and Wildlife Service: around 1991, a marked drop in abundances consistently occurred in different tributaries of the affected river and lake system. This scenario led us to specify a natural model candidate: a population dynamics model with a change-point. Rather than a phenomenological description, we sought a mechanistic process that with correct partitioning of the variance components of the process would allow us to answer imporant questions: how was the dynamics before the change affecting the dynamics after the change, and the extent of the change? After so many years fluctuating at low abundances, is the population expected to recover? If so,

how long does the recovery process would take? Progress, as it were, did not reside per se in obtaining good estimators of the model parameters. Rather, parameter estimation turned out to be a by-product of the process of scientific inquiry of what a stochastic model is really implying.

The principles driving the inferences from these models are applicable to a wide array of systems, from macroscopic to microscopic scales. The advent of DNA-based technology, for instance, has made it possible to explore the population dynamics of entire microbial communities, whose signature trends exhibit sudden changes. Being able to achieve a process-based understanding of such changes is at the center of very active research in microbial ecology [56], because different microbial community compositions and structures are associated with different pathological states of the hosts. It follows that understanding the drivers of the "stability" of the community composition is key to understand how to manage changes in a microbial community. Our model can be readily extended to a multi-species autoregressive format (e.g. [11]) that can serve as the basis to move the research from the study of patterns of change in abundance to the study of patterns of change in the nature and intensity of ecological interactions. Such move would make readily accessible an ecological, process-based understanding of change. Finally, the results of this paper can be immediately extended to switch from a saltational parameter change to model population dynamics under gradual parameter change, and that is also, research material for upcoming studies on this topic.

Much recent work in ecology and other scientific fields aims at developing statistical tests to anticipate changes of state [57]. One of the latest examinations of the problem [58] emphasizes the importance of model-based approaches to find reliable indicators of impending regime shifts. In this paper, we shift the attention from the problem of detecting when a change will occur, to another non-trivial puzzle: understanding and predicting the post-breakpoint behavior of the population dynamics. In this work we show that population time series transitioning from one stationary distribution to another contain explicit information about where the process was before the change-point, where is it heading and how long it will take to transition. We view the current contribution's explicit modeling as a step in understanding the relationship between stochasticity, density dependence, population dynamics shifts, and the influence of the past dynamics on future dynamics.

Acknowledgements

JMP and MLT were funded by both NIH R01-GM103604 and by a Montana Fish, Wildlife, and Parks contract 060327 to M. L. Taper.

[1] R. D. Holt, Density-independent mortality, non-linear competitive interactions, and species coexistence, Journal of Theoretical Biology 116 (1985) 479–493.

[2] J.-D. Lebreton, O. Gimenez, Detecting and estimating density dependence in wildlife populations, The Journal of Wildlife Management 77 (2013) 12–23.

[3] B. Dennis, M. L. Taper, Density dependence in time series observations of natural populations: estimation and testing, Ecological Monographs 64 (1994) 205–224.

[4] B. Dennis, W. Kemp, M. Taper, Joint density dependence, Ecology 79 (1998) 426–441.

[5] R. P. Freckleton, A. R. Watkinson, R. E. Green, W. J. Sutherland, Census error and the detection of density dependence, Journal of Animal Ecology 75 (2006) 837–851.

[6] M. Bulmer, The statistical analysis of density dependence, Biometrics (1975) 901–911.

[7] B. Dennis, J. M. Ponciano, S. R. Lele, M. L. Taper, D. F. Staples, Estimating density dependence, process noise, and observation error, Ecological Monographs 76 (2006) 323–341.

[8] B. Dennis, G. Patil, The gamma distribution and weighted multimodal gamma distributions as models of population abundance, Mathematical Biosciences 68 (1984) 187–212.

[9] Z. Zeng, R. M. Nowierski, M. L. Taper, B. Dennis, W. P. Kemp, Complex population dynamics in the real world: modeling the influence of time-varying parameters and time lags, Ecology 79 (1998) 2193–2209.

[10] P. de Valpine, A. Hastings, Fitting population models incorporating process noise and observation error, Ecological Monographs 72 (2002) 57–76.

[11] A. Ives, B. Dennis, K. Cottingham, S. Carpenter, Estimating community stability and ecological interactions from time-series data, Ecological Monographs 73 (2003) 301–330.

[12] J. S. Clark, O. N. Bjørnstad, Population time series: process variability, observation errors, missing values, lags, and hidden states, Ecology 85 (2004) 3140–3150.

[13] S. R. Lele, B. Dennis, F. Lutscher, Data cloning: easy maximum likelihood estimation for complex ecological models using bayesian markov chain monte carlo methods, Ecology Letters 10 (2007) 551–563.

[14] J. M. Ponciano, M. L. Taper, B. Dennis, S. R. Lele, Hierarchical models in ecology: confidence intervals, hypothesis testing, and model selection using data cloning, Ecology 90 (2009) 356–362.

[15] L. Polansky, P. De Valpine, J. O. Lloyd-Smith, W. M. Getz, Likelihood ridges and multimodality in population growth rate models, Ecology 90 (2009) 2313–2320.

[16] B. Dennis, J. M. Ponciano, M. L. Taper, Replicated sampling increases efficiency in monitoring biological populations, Ecology 91 (2010) 610–620.

[17] P. de Valpine, Frequentist analysis of hierarchical models for population dynamics and demographic data, Journal of Ornithology 152 (2012) 393–408.

[18] B.-E. Sæther, T. Coulson, V. Grøtan, S. Engen, R. Altwegg, K. B. Armitage, C. Barbraud, P. H. Becker, D. T. Blumstein, F. S. Dobson, et al., How life history influences population dynamics in fluctuating environments, The American Naturalist 182 (2013) 743–759.

[19] S. R. Lele, M. L. Taper, Statistical inference in ecology, in: International Encyclopedia of Statistical Science, Springer, 2011, pp. 1430–1433.

[20] J. Knape, Estimability of density dependence in models of time series data, Ecology 89 (2008) 2994–3000.

[21] J. R. Thomson, W. J. Kimmerer, L. R. Brown, K. B. Newman, R. M. Nally, W. A. Bennett, F. Feyrer, E. Fleishman, Bayesian change point analysis of abundance trends for pelagic fishes in the upper san francisco estuary, Ecological Applications 20 (2010) 1431–1448.

[22] C. Packer, R. Hilborn, A. Mosser, B. Kissui, M. Borner, G. Hopcraft, J. Wilmshurst, S. Mduma, A. R. Sinclair, Ecological change, group territoriality, and population dynamics in serengeti lions, Science 307 (2005) 390–393.

[23] M. B. Hooten, C. K. Wikle, Shifts in the spatio-temporal growth dynamics of shortleaf pine, Environmental and Ecological Statistics 14 (2007) 207–227.

[24] T. Andersen, J. Carstensen, E. Hernandez-Garcia, C. M. Duarte, Ecological thresholds and regime shifts: approaches to identification, Trends in Ecology & Evolution 24 (2009) 49–57.

[25] A. R. Ives, V. Dakos, Detecting dynamical changes in nonlinear time series using locally linear state-space models, Ecosphere 3 (2012) art58.

[26] B. E. Hansen, The likelihood ratio test under nonstandard conditions: testing the markov switching model of gnp, Journal of Applied Econometrics 7 (1992) S61–S82.

[27] R. Garcia, Asymptotic null distribution of the likelihood ratio test in markov switching models, International Economic Review (1998) 763–788.

[28] D. H. Timmer, J. J. Pignatiello, Change point estimates for the parameters of an ar (1) process, Quality and Reliability Engineering International 19 (2003) 355–369.

[29] V. Jandhyala, S. Fotopoulos, I. MacNeill, P. Liu, Inference for single and multiple change-points in time series, Journal of Time Series Analysis (2013).

[30] D. Picard, Testing and estimating change-points in time series, Advances in Applied Probability (1985) 841–867.

[31] D. W. Andrews, Tests for parameter instability and structural change with unknown change point, Econometrica: Journal of the Econometric Society (1993) 821–856.

[32] J. Bai, P. Perron, Estimating and testing linear models with multiple structural changes, Econometrica (1998) 47–78.

[33] M. Lavielle, E. Moulines, Least-squares estimation of an unknown number of shifts in a time series, Journal of Time Series Analysis 21 (2000) 33–59.

[34] L. Perreault, J. Bernier, B. Bobée, E. Parent, Bayesian change-point analysis in hydrometeorological time series. part 1. the normal model revisited, Journal of Hydrology 235 (2000) 221–241.

[35] F. Yuan, Tests for Change-Point of the AR (1) Model, Ph.D. thesis, University of Minnesota, 2013.

[36] M. B. Perry, Identifying the time of polynomial drift in the mean of autocorrelated processes, Quality and Reliability Engineering International 26 (2010) 399–415.

[37] A. Amiri, S. Allahyari, Change point estimation methods for control chart postsignal diagnostics: a literature review, Quality and Reliability Engineering International 28 (2012) 673–685.

[38] T. Zhou, Bayesian approach to Phase II statistical process control for time series, Ph.D. thesis, University of Minnesota, 2013.

[39] R. Holt, M. Barfield, Trait-mediated effects, density dependence and the dynamic stability of ecological systems, Trait-Mediated Indirect Interactions: Ecological and Evolutionary Perspectives (2012) 89.

[40] R. Lande, S. Engen, B.-E. Sæther, Estimating density dependence in time–series of age–structured populations, Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences 357 (2002) 1179–1184.

[41] M. Kot, Elements of mathematical ecology, Cambridge University Press, 2001.

[42] S. Herrando-Pérez, S. Delean, B. W. Brook, C. J. Bradshaw, Density dependence: an ecological tower of babel, Oecologia 170 (2012) 585–603.

[43] B. Dennis, G. Patil, Applications in ecology, Lognormal Distributions: Theory and Applications (1988) 303–330.

[44] H. Rue, L. Held, Gaussian Markov random fields: theory and applications, CRC Press, 2005.

[45] A. Berryman, M. Lima, Deciphering the effects of climate on animal populations: diagnostic analysis provides new interpretation of soay sheep dynamics, The American Naturalist 168 (2006) 784–795.

[46] J. M. Ferguson, J. M. Ponciano, Evidence and implications of higher order scaling in the environmental variation of animal population growth, Proceedings of the National Academy of Sciences In press (2015).

[47] J. M. Cushing, R. F. Costantino, B. Dennis, R. Desharnais, S. M. Henson, Chaos in ecology: experimental nonlinear dynamics, volume 1, Elsevier, 2002.

[48] B. A. Melbourne, A. Hastings, Extinction risk depends strongly on factors contributing to stochasticity, Nature 454 (2008) 100–103.

[49] J. M. Ferguson, J. M. Ponciano, Predicting the process of extinction in experimental microcosms and accounting for interspecific interactions in single-species time series, Ecology Letters 17 (2014) 251–259.

[50] M. Schwager, K. Johst, F. Jeltsch, Does red noise increase or decrease extinction risk? single extreme events versus series of unfavorable conditions, The American Naturalist 167 (2006) 879–888.

[51] B. Dennis, L. Assas, S. Elaydi, E. Kwessi, G. Livadiotis, Allee effects and resilience in stochastic populations, Theoretical Ecology 9 (2016) 323–335.

[52] T. J. Case, Illustrated guide to theoretical ecology, Ecology 80 (2000) 2848–2848.

[53] D. F. Staples, M. L. Taper, B. B. Shepard, Risk-based viable population monitoring, Conservation Biology 19 (2005) 1908–1916.

[54] B. Connors, A. Cooper, R. Peterman, N. Dulvy, The false classification of extinction risk in noisy environments, Proceedings of the Royal Society B: Biological Sciences 281 (2014) 20132935.

[55] C. C. Muhlfeld, M. L. Taper, D. F. Staples, B. B. Shepard, Observer error structure in bull trout redd counts in montana streams: implications for inference on true redd numbers, Transactions of the American Fisheries Society 135 (2006) 643–654.

[56] J. Ravel, R. M. Brotman, P. Gajer, B. Ma, M. Nandy, D. W. Fadrosh, J. Sakamoto, S. S. Koenig, L. Fu, X. Zhou, et al., Daily temporal dynamics of vaginal microbiota before, during and after episodes of bacterial vaginosis, Microbiome 1 (2013) 29.

[57] J. M. Drake, Early warning signals of stochastic switching, Proceedings of the Royal Society B: Biological Sciences 280 (2013) 20130686.

[58] C. Boettiger, A. Hastings, Quantifying limits to detection of early warning for critical transitions, Journal of The Royal Society Interface 9 (2012) 2527–2539.

6. Figures

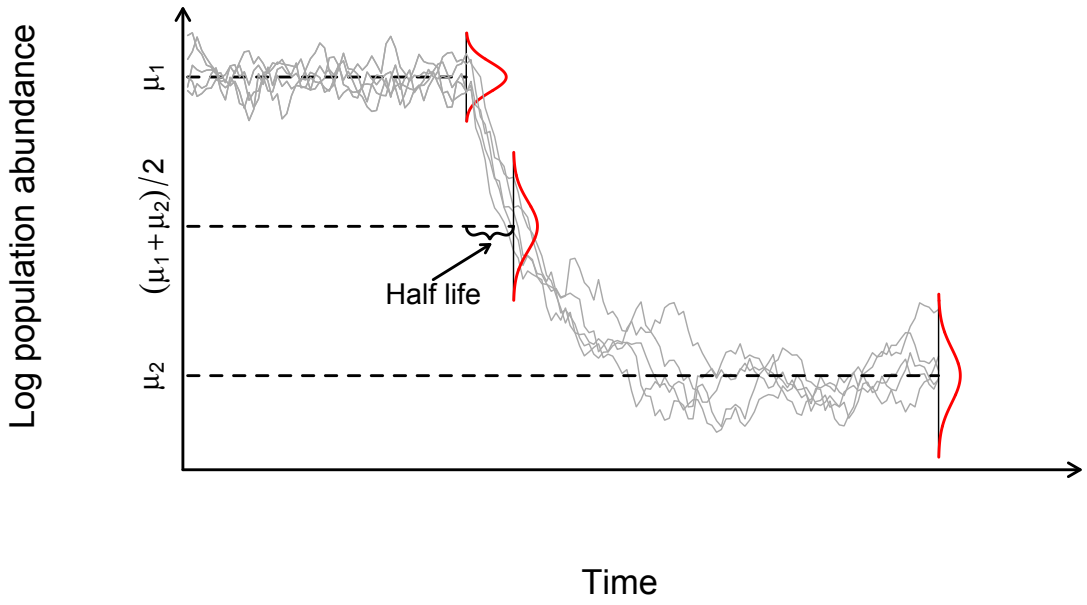

Figure 1: The stochastic Gompertz breakpoint process and the half life of the change in the mean of the dynamical process, $\bar{t} = \ln(1/2)/\ln c_2$. Plotted are 5 realizations of the Stochastic Gompertz breakpoint process. Dotted lines mark the process mean before the breakpoint (μ_1), the mean after the breakpoint (μ_2), and the arithmetic average of both means ($(\mu_1 + \mu_2)/2$). The time at which such arithmetic average is reached is the half life of the process, $\bar{t}$.

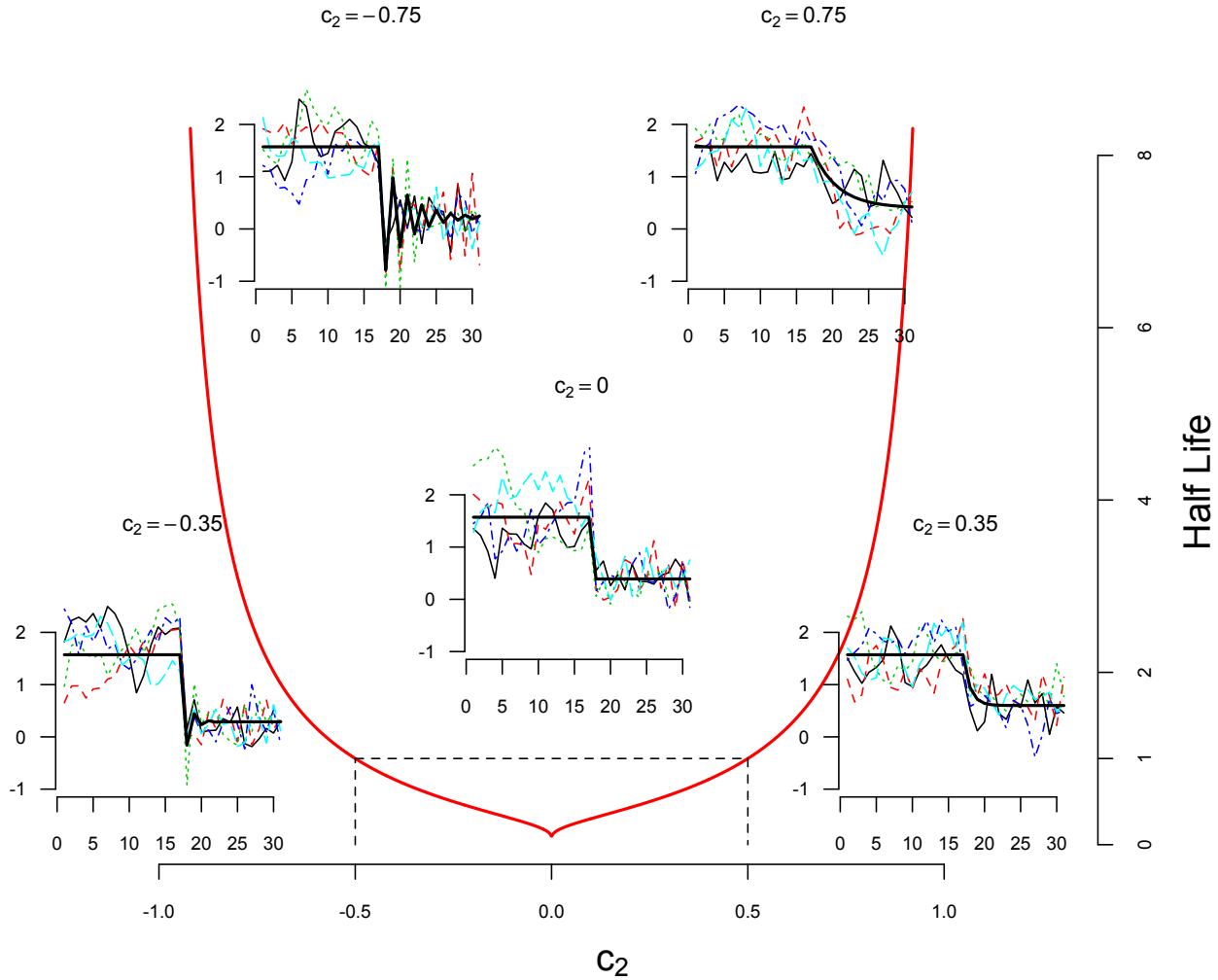

Figure 2: Half life of the change in the mean of the dynamics, $\bar{t} = \ln(1/2)/\ln c_2$, plotted as a function of the strength of density dependence c_2. If the strength of density-dependence after the breakpoint c_2 is close to 1 (*i.e.* b is close to 0 and hence growth is close to density independence), then change occurs slowly. As c_2 approaches 0.5 from the right, the speed of the change increases. When $c_2 = 1/2$, the change occurs within one time step (*e.g.* a year). If c_2 is exactly 0, then the change is immediate. Between 0 and $-1/2$ the change still occurs within a single time step, but weak damped oscillations towards the equilibrium are produced. Finally, when $-1 < c_2 < -1/2$, the speed of the change is slow again, with the added complexity that such change is accompanied with strong, damped oscillations towards the equilibrium. Thus, the non-stationary trend contains information about *a)* the location of the past stationary mean, *b)* the location of the future stationary mean, and *c)* the duration of the transitionary period.

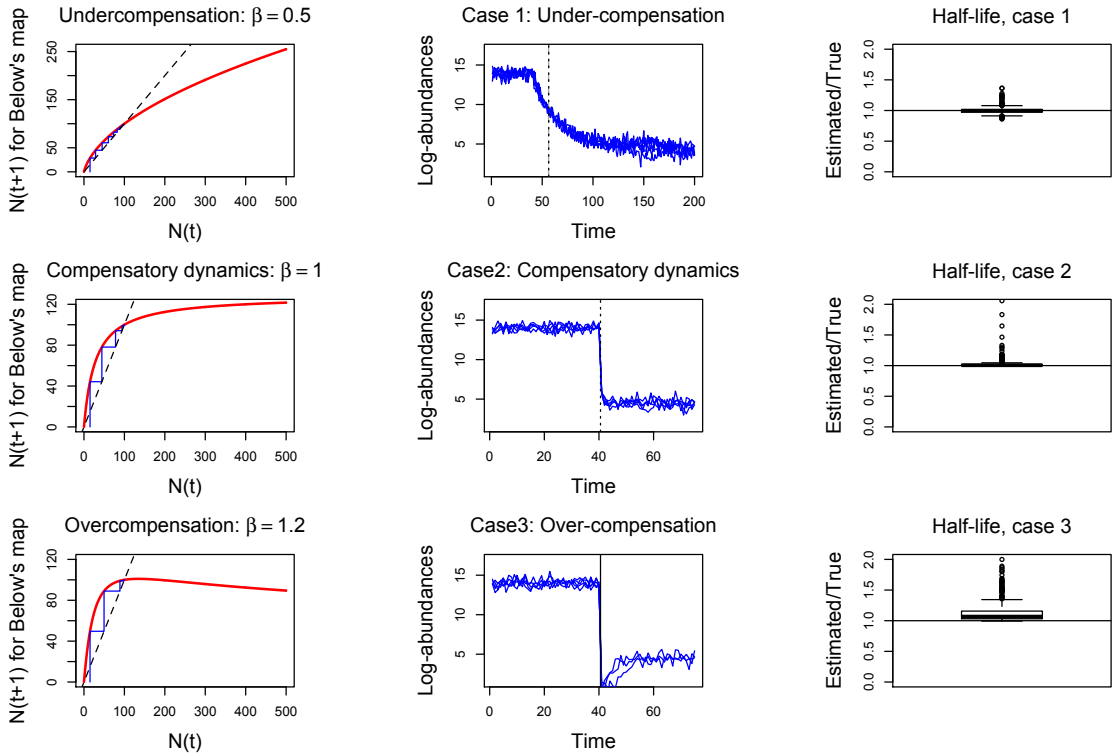

Figure 3: Testing the robustness of the breakpoint Stochastic Gompertz model. A biologically realistic change-point stochastic process was used to test the robustness of the Gompertz model to unaccounted biological complexities. Accordingly, for three different types of compensatory dynamics, 1000 simulations from a population growth model with demographic variability, demographic noise, sampling error and a change-point were used to estimate the parameters of the BPSG model with added sampling error, along with the half life of the change. In the figure, each row, from left to right, shows the type of compensatory dynamics used in the simulations (left panels), 10 samples of the simulated time series of length 200 with the true half life $\bar{t}$ (center panels and solid vertical line) and the estimated half life (dotted vertical line), and a box plot of the 1000 half life estimates (right panels). For the compensatory and over compensatory dynamics, only the first 75 time steps of the simulated time series are shown for clarity.

AppendixA.

AppendixA.1. Matrix formulation of the breakpoint process

Here we derive by induction a matrix formulation for the (random) discrete map with a breakpoint. To do that, we derive first the multivariate distribution of the process without a breakpoint, initialized at time 0 and stopped at time q. Although this derivation has been shown before (Dennis et al 2006), presenting it here illuminates the formulation of the multivariate distribution with a change-point. To find the joint distribution of $[X_0, X_1, X_2, \dots X_q]'$ we first iterate the stochastic Gompertz map in eq. (4):

$$\begin{aligned}
X_0 &= X_0 \\
X_1 &= a + cX_0 + E_1 \\
X_2 &= a + cX_1 + E_2 = a(1+c) + c^2 X_0 + cE_1 + E_2 \\
X_3 &= a + cX_2 + E_3 = a(1+c+c^2) + c^3 X_0 + c^2 E_1 + cE_2 + E_3,
\end{aligned}$$

Then, continuing in this fashion and using vector notation we get

$$X_q = \frac{a(1-c^q)}{1-c} + c^t X_0 + \begin{bmatrix} 1 & c & c^2 & \dots & c^{q-1} \end{bmatrix} \begin{bmatrix} E_q \\ E_{q-1} \\ E_{q-2} \\ \vdots \\ E_1 \end{bmatrix}, \tag{A.1}$$

where we used the well known formula for the partial sum of a geometric series. Upon examination of this iteration, it is seen that the vector of log population sizes through time, $\mathbf{X} = \begin{bmatrix} X_0 & X_1 & X_2 & \dots & X_q \end{bmatrix}'$ can be written as:

$$\begin{aligned}
\mathbf{X} &= a \begin{bmatrix} 0 \\ 1 \\ \frac{1-c^2}{1-c} \\ \frac{1-c^2}{1-c} \\ \vdots \\ \frac{1-c^q}{1-c} \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ c & 1 & 0 & 0 & \dots & 0 \\ c^2 & c & 1 & 0 & \dots & 0 \\ c^3 & c^2 & c & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ c^q & c^{q-1} & c^{q-2} & c^{q-2} & \dots & 1 \end{bmatrix} \begin{bmatrix} X_0 \\ E_1 \\ E_2 \\ E_3 \\ \vdots \\ E_q \end{bmatrix} \\
&= a\mathbf{g}_{(q+1)\times 1} + \mathbf{C}_{(q+1)\times(q+1)}\mathbf{E}_{(q+1)\times 1}.
\end{aligned} \tag{A.2}$$

Assuming that the first observation arises from the Gaussian stationary distribution with mean and variance $a/(1-$

c) and $\sigma^2/(1-c^2)$ respectively, we then have that:

$$
\begin{aligned}
E[\mathbf{E}] &= \begin{bmatrix} \frac{a}{1-c} & 0 & 0 & \dots & 0 \end{bmatrix}' = \mathbf{h} \\
Var[\mathbf{E}] &= \sigma^2 \begin{bmatrix} \frac{1}{1-c^2} & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix} = \mathbf{\Psi},
\end{aligned} \tag{A.3}
$$

and

$$
\begin{aligned}
E[\mathbf{X}] &= a\mathbf{g} + \mathbf{C}\mathbf{h} \\
Var[\mathbf{X}] &= \mathbf{C}\mathbf{\Psi}\mathbf{C}'.
\end{aligned} \tag{A.4}
$$

Using this multivariate mean and variance and the properties of the Multivariate Normal (MVN) distribution (Johnson and Wichern 2002, chapter 4), it follows that

$$
\mathbf{X} \sim MVN\left(\boldsymbol{\mu} = \frac{a}{1-c}\mathbf{j}_{(q+1)\times 1}, \mathbf{\Sigma} = \mathbf{C}\mathbf{\Psi}\mathbf{C}'\right), \tag{A.5}
$$

where

$$
\mathbf{\Sigma} = \frac{\sigma^2}{1-c^2} \begin{bmatrix} 1 & c & c^2 & \dots & c^q \\ c & 1 & c & \dots & c^{q-1} \\ c^2 & c & 1 & \dots & c^{q-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c^q & c^{q-1} & c^{q-2} & \dots & 1 \end{bmatrix}, \tag{A.6}
$$

and $\mathbf{j}_{(q+1)\times 1}$ is a vector of ones of size $q+1$.

Now, to derive the multivariate distribution of the process with a breakpoint, we start by writing expressions for

the log-population process, from time τ (right before the breakpoint) and onwards:

$$
\begin{aligned}
X_\tau &= \frac{a_1(1-c_1^\tau)}{1-c_1} + \begin{bmatrix} c_1^\tau & c_1^{\tau-1} & c_1^{\tau-2} & \dots & c_1 & 1 \end{bmatrix} \begin{bmatrix} X_0 \\ E_1 \\ E_2 \\ \vdots \\ E_{\tau-1} \\ E_\tau \end{bmatrix}, \\
X_{\tau+1} &= a_2 + c_2\left(X_\tau\right) + E_{\tau+1} \\
&= a_2 + a_1 c_2 \frac{1-c_1^\tau}{1-c_1} + \begin{bmatrix} c_2 c_1^\tau & c_2 c_1^{\tau-1} & c_2 c_1^{\tau-2} & \dots & c_2 c_1 & c_2 & 1 \end{bmatrix} \begin{bmatrix} X_0 \\ E_1 \\ E_2 \\ \vdots \\ E_{\tau-1} \\ E_\tau \\ E_{\tau+1} \end{bmatrix},
\end{aligned}
$$

and continuing in this fashion, a geometric series for c_2 appears in the solution:

$$
\begin{aligned}
X_{\tau+2} &= a_2 + c_2\left(X_{\tau+1}\right) + E_{\tau+2} \\
&= a_2(1+c_2) + c_2^2 a_1 \frac{1-c_1^\tau}{1-c_1} + \begin{bmatrix} c_2^2 c_1^\tau & c_2^2 c_1^{\tau-1} & c_2^2 c_1^{\tau-2} & \dots & c_2^2 c_1 & c_2^2 & c_2 & 1 \end{bmatrix} \begin{bmatrix} X_0 \\ E_1 \\ E_2 \\ \vdots \\ E_{\tau-1} \\ E_\tau \\ E_{\tau+1} \\ E_{\tau+2} \end{bmatrix}, \\
X_{\tau+3} &= a_2 + c_2\left(X_{\tau+2}\right) + E_{\tau+3} \\
&= a_2(1+c_2+c_2^2) + c_2^3 a_1 \frac{1-c_1^\tau}{1-c_1} + \begin{bmatrix} c_2^3 c_1^\tau & c_2^3 c_1^{\tau-1} & c_2^3 c_1^{\tau-2} & \dots & c_2^3 c_1 & c_2^3 & c_2^2 & c_2 & 1 \end{bmatrix} \begin{bmatrix} X_0 \\ E_1 \\ E_2 \\ \vdots \\ E_{\tau-1} \\ E_\tau \\ E_{\tau+1} \\ E_{\tau+2} \\ E_{\tau+3} \end{bmatrix}, \\
&\vdots \\
&\vdots \\
X_{\tau+q} &= a_2 + c_2\left(X_{\tau+q-1}\right) + E_{\tau+q} \\
&= a_2(1+c_2+c_2^2+\ldots+c_2^{q-1}) + \ldots + c_2^q a_1 \frac{1-c_1^\tau}{1-c_1} \\
&\quad + \begin{bmatrix} c_2^q c_1^\tau & \dots & c_2^q c_1 & c_2^q & c_2^{q-1} & \dots & 1 \end{bmatrix} \begin{bmatrix} X_0 \\ E_1 \\ \vdots \\ E_{\tau+q-1} \\ E_{\tau+q} \end{bmatrix}. \qquad \text{(A.7)}
\end{aligned}
$$

The induction patterns revealed by this iterated map can then be readily expressed in the matrix form shown in eq. (7) in the main text.

AppendixA.2. Derivation of the variance-covariance matrix of the breakpoint process

Here we derive the four sub-matrices of $\boldsymbol{\Sigma_\tau}$. We recall that we also partitioned the matrices $\boldsymbol{\Gamma}$ and $\boldsymbol{\Phi}$ into four sub-matrices matching the dimensions of the $\boldsymbol{\Sigma_\tau}$ blocks. For $\boldsymbol{\Gamma}$ these matrices are

$$\boldsymbol{\Gamma_{11}}_{(\tau+1)\times(\tau+1)} = \begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ c_1 & 1 & 0 & 0 & \dots & 0 \\ c_1^2 & c_1 & 1 & 0 & \dots & 0 \\ c_1^3 & c_1^2 & c_1 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ c_1^\tau & c_1^{\tau-1} & c_1^{\tau-2} & \dots & c_1 & 1 \end{bmatrix}, \quad \boldsymbol{\Gamma_{12}}_{(\tau+1)\times q} = \mathbf{0},$$

$$\boldsymbol{\Gamma_{21}}_{q\times(\tau+1)} = \begin{bmatrix} c_2c_1^\tau & c_2c_1^{\tau-1} & c_2c_1^{\tau-2} & \dots & c_2c_1 & c_2 \\ c_2^2c_1^\tau & c_2^2c_1^{\tau-1} & c_2^2c_1^{\tau-2} & \dots & c_2^2c_1 & c_2^2 \\ c_2^3c_1^\tau & c_2^3c_1^{\tau-1} & c_2^3c_1^{\tau-2} & \dots & c_2^3c_1 & c_2^3 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ c_2^qc_1^\tau & c_2^qc_1^{\tau-1} & c_2^qc_1^{\tau-2} & \dots & c_2^qc_1 & c_2^q \end{bmatrix}, \quad \boldsymbol{\Gamma_{22}}_{q\times q} = \begin{bmatrix} 1 & 0 & \dots & \dots & 0 \\ c_2 & 1 & 0 & \dots & 0 \\ c_2^2 & c_2 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ c_2^{q-1} & c_2^3 & c_2^2 & c_2 & 1 \end{bmatrix}$$

and for $\boldsymbol{\Phi}$ these are

$$\boldsymbol{\Phi_{11}}_{(\tau+1)\times(\tau+1)} = \sigma_1^2 \begin{bmatrix} \frac{1}{1-c_1^2} & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}, \quad \boldsymbol{\Phi_{22}}_{q\times q} = \sigma_2^2\mathbf{I},$$

and

$$\boldsymbol{\Phi_{12}}_{(\tau+1)\times q} = \boldsymbol{\Phi_{21}}'_{q\times(\tau+1)} = \mathbf{0}.$$

With these matrices in place, the matrices in eqs. (16) and (17) are obtained by simple matrix multiplication and elementary algebraic manipulations:

$$\boldsymbol{\Gamma_{21}}\boldsymbol{\Phi_{11}}\boldsymbol{\Gamma_{21}}'_{q\times q} = \frac{\sigma_1^2}{1-c_1^2}\boldsymbol{\Omega_1}, \text{ where} \tag{A.8}$$

$$\boldsymbol{\Omega_1} = \begin{bmatrix} c_2^2 & c_2^3 & \dots & c_2^{q+1} \\ c_2^3 & c_2^4 & \dots & c_2^{q+2} \\ \vdots & \vdots & \ddots & \vdots \\ c_2^{q+1} & c_2^{q+2} & \dots & c_2^{2q} \end{bmatrix}.$$

whereas

$$\mathbf{\Gamma_{22}\Phi_{22}\Gamma_{22}}'_{q\times q} = \sigma_2^2\mathbf{\Omega_2}, \text{ where} \tag{A.9}$$

$$\mathbf{\Omega_2} = \begin{bmatrix} 1 & c_2 & c_2^2 & \dots & c_2^{q-2} & c_2^{q-1} \\ c_2 & c_2^2+1 & c_2^3+c_2 & \dots & c_2^{q-1}+c_2^{q-3} & c_2^q+c_2^{q-2} \\ c_2^2 & c_2^3+c_2 & c_2^4+c_2^2+1 & \dots & c_2^q+c_2^{q-2}+c_2^{q-4} & c_2^{q+1}+c_2^{q-1}+c_2^{q-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ c_2^{q-2} & c_2^{q-1}+c_2^{q-3} & c_2^q+c_2^{q-2}+c_2^{q-4} & \dots & c_2^{2(q-2)}+c_2^{2(q-3)}+\dots+c_2^2+1 & c_2^{2q-3}+c_2^{2q-5}+\dots+c_2^3+c_2 \\ c_2^{q-1} & c_2^q+c_2^{q-2} & c_2^{q+1}+c_2^{q-1}+c_2^{q-3} & \dots & c_2^{2q-3}+c_2^{2q-5}+\dots+c_2^3+c_2 & c_2^{2(q-1)}+c_2^{2(q-2)}+\dots+c_2^2+1 \end{bmatrix}.$$

Noting that

$$\sigma_2^2\left[\frac{1}{1-c_2^2}\mathbf{\Omega_1}+\mathbf{\Omega_2}\right] = \frac{\sigma_2^2}{1-c_2^2}\begin{bmatrix} 1 & c_2 & c_2^2 & \dots & c_2^{q-1} \\ c_2 & 1 & c_2 & \dots & c_2^{q-2} \\ c_2^2 & c_2 & 1 & \dots & c_2^{q-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_2^{q-1} & c_2^{q-2} & c_2^{q-3} & \dots & 1 \end{bmatrix}$$

is identical to the variance covariance of the second partition, had it arisen directly from the stationary distribution of the second set of parameters, and solving in that equation for $\sigma_2^2\mathbf{\Omega_2}$ leads to the expression in eq. (17).

Next, we compute $\mathbf{\Sigma_{12}} = Cov(\mathbf{X_1}, \mathbf{X_2})$, which is equal to the transpose of $\mathbf{\Sigma}_{21}$. Straightforward matrix multiplications yield

$$\mathbf{\Sigma_{12}} = \mathbf{\Gamma_{11}\Phi_{11}\Gamma_{21}}'_{(\tau+1)\times q} = \begin{bmatrix} c_2\mathbf{U} & c_2^2\mathbf{U} & \dots & c_2^q\mathbf{U} \end{bmatrix}, \tag{A.10}$$

where

$$\mathbf{U}_{(\tau+1)\times 1} = \sigma_1^2\begin{bmatrix} \frac{c_1^\tau}{1-c_1^2} \\ \frac{c_1^{\tau+1}}{1-c_1^2}+c_1^{\tau-1} \\ \frac{c_1^{\tau+2}}{1-c_1^2}+c_1^\tau+c_1^{\tau-2} \\ \frac{c_1^{\tau+3}}{1-c_1^2}+c_1^{\tau+1}+c_1^{\tau-1}+c_1^{\tau-3} \\ \vdots \\ \frac{c_1^{2\tau-1}}{1-c_1^2}+c_1^{2\tau-3}+c_1^{2\tau-5}+\dots+c_1^3+c_1 \\ \frac{c_1^{2\tau}}{1-c_1^2}+c_1^{2\tau-2}+c_1^{2\tau-4}+\dots+c_1^2+1 \end{bmatrix} = \frac{\sigma_1^2}{1-c_1^2}\begin{bmatrix} c_1^\tau \\ c_1^{\tau-1} \\ c_1^{\tau-2} \\ c_1^{\tau-3} \\ \vdots \\ c_1 \\ 1 \end{bmatrix}. \tag{A.11}$$

A simple examination of eq. (A.10) and eq. (A.11) then leads to eq. 15 in the main text.