

Quantifying Environmental and Demographic Stochasticity in Population Viability Analysis: The Case of *Rissa tridactyla*

Riddhimann Chattopadhyay

National Institute of Technology

Richik Bandyopadhyay

bandopadhyayrichik@gmail.com

Indian Statistical Institute

Tanoy Mukherjee

Zoological Survey of India

Research Article

Keywords: Population dynamics, Logistic growth, Demographic stochasticity, Environmental stochasticity, Maximum likelihood estimation, Population viability analysis

Posted Date: September 21st, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10804982/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Quantifying Environmental and Demographic Stochasticity in Population Viability Analysis: The Case of *Rissa tridactyla*

Riddhimann Chattopadhyay^a, Richik Bandyopadhyay^{b,*}, Tanoy Mukherjee^c

^a*Department of Computer Applications, National Institute of Technology, Tiruchirappalli-620015, India*

^b*Agricultural and Ecological Research Unit, Indian Statistical Institute, Barrackpore Trunk Road, Baranagar, Kolkata-700108, West Bengal, India*

^c*Estuarine Biology Regional Centre, Zoological Survey of India, Gopalpur-on-Sea, Ganjam, Odisha-761002, India*

Abstract

Mathematical models that incorporate stochasticity are essential for understanding population dynamics and assessing extinction risk in fluctuating environments. In this study, we investigate the extinction risk of the Black-legged Kittiwake (*Rissa tridactyla*) using deterministic and stochastic population growth models. Population data obtained from the Global Population Dynamics Database (GPDD) were fitted using exponential, Gompertz and logistic growth models, with model selection based on the Akaike Information Criterion (AIC). The logistic model provided the best description of the observed population dynamics and was subsequently extended to incorporate demographic and environmental stochasticity through stochastic differential equations. Model parameters were estimated using maximum likelihood estimation, which identified demographic stochasticity as the dominant source of population variability. Future population trajectories were projected using Monte Carlo simulations, yielding an estimated 53.1% probability of quasi-extinction within the next 30 years. Confidence and prediction intervals demonstrated good agreement between model predictions and observed data, supporting the reliability of the proposed framework. These findings emphasize the importance of incorporating stochastic processes into population viability analyses and provide valuable quantitative insights for assessing extinction risk and informing conservation strategies for vulnerable seabird populations.

Keywords: Population dynamics, Logistic growth, Demographic stochasticity, Environmental stochasticity, Maximum likelihood estimation, Population viability analysis

1. Introduction

Understanding population growth and extinction risk is a central objective of population ecology and conservation biology, particularly under increasing environmental variability and anthropogenic pressure. Mathematical models provide an essential framework for describing population dynamics, predicting future population trajectories and assessing long-term persistence. Among these, the logistic growth model introduced by Verhulst and related nonlinear growth models, including the Gompertz and θ -logistic models, remain among the most widely used deterministic models for describing density-dependent population growth ([14, 30]). Owing to their simplicity, analytical tractability

*Corresponding author

Email address: bandopadhyayrichik@gmail.com (Richik Bandyopadhyay)

and biological interpretability, these models have served as the foundation for numerous studies in ecology, conservation biology, and population management ([9, 17]).

Natural populations, however, rarely evolve deterministically. Population growth is influenced by stochastic fluctuations arising from both demographic and environmental sources [31]. Demographic stochasticity results from the inherent randomness of individual birth and death events and is particularly important in small populations, whereas environmental stochasticity reflects temporal variation in external conditions, such as climate, habitat quality, and resource availability, that simultaneously affect all individuals within a population ([20, 22]). These stochastic processes can substantially alter population trajectories, reduce population persistence, and increase extinction risk even when the corresponding deterministic model predicts long-term survival.

To account for these uncertainties, stochastic extensions of classical population growth models have been developed by incorporating random perturbations into the governing dynamics through stochastic differential equations. In demographic stochasticity models, the variance of the stochastic fluctuations scales linearly with population size, whereas under environmental stochasticity it scales with the square of the population size, reflecting fluctuations that simultaneously influence all individuals in the population ([20, 12]). Such stochastic formulations provide a more realistic description of natural populations and have become indispensable tools for estimating extinction probabilities, persistence times, and long-term population viability under uncertainty.

A widely used analytical framework for studying extinction risk is the diffusion approximation, whereby discrete birth-death processes are approximated by continuous-time stochastic differential equations ([8, 32]). This approach facilitates analytical or semi-analytical estimates of extinction probabilities while preserving the essential effects of stochastic fluctuations. However, nonlinear density dependence, parameter uncertainty, and multiple sources of stochasticity often make analytical solutions difficult or impossible to obtain. Consequently, simulation-based approaches, particularly Monte Carlo simulations, have become indispensable for forecasting population trajectories, quantifying uncertainty, and evaluating extinction risk in population viability analyses ([24, 15, 4]).

Reliable prediction of future population dynamics also requires quantifying uncertainty associated with model parameters and stochastic processes. Statistical inference methods, particularly maximum likelihood estimation (MLE), provide a rigorous framework for estimating model parameters from observed time-series data while accounting for process variability. Coupled with variance-covariance analysis and first-order Taylor approximations, MLE enables the construction of confidence intervals for the mean population trajectory and prediction intervals for future observations, thereby providing a quantitative assessment of forecasting uncertainty ([15, 25]).

In this study, we investigate both deterministic and stochastic frameworks to model the population dynamics of the Black-legged Kittiwake (*Rissa tridactyla*), a colonial seabird that has experienced substantial regional population declines due to climate change, reduced prey availability, and other anthropogenic pressures ([29, 23]). In section 2, Using annual population time-series data obtained from the **Global Population Dynamics Database (GPDD)**, we first fit exponential, logistic, Gompertz, and θ -logistic growth models to characterize the observed population trajectory. Model parameters are estimated using nonlinear least-squares methods, and the competing models are compared using standard goodness-of-fit criteria.

In section 3, building upon the best-performing deterministic model, we have developed the mathematical methodology. We first incorporate both demographic and environmental stochasticity into the logistic growth framework and estimate the associated stochastic parameters using maximum likelihood estimation. We subsequently construct confidence intervals for the mean population trajectory and prediction intervals for future population sizes by propagating parameter uncertainty through the stochastic model using the estimated variance-covariance matrix and first-order Taylor approximations. Future population trajectories are generated through stochastic simulations, and the empirical coverage of the prediction intervals is evaluated by comparing simulated trajectories with the derived intervals.

Finally, extinction probabilities are estimated using both diffusion approximations and Monte Carlo simulations, allowing a direct comparison between analytical and simulation-based approaches under demographic and environmental stochasticity. By integrating deterministic growth modeling, statistical inference, stochastic simulation, uncertainty quantification, and extinction analysis, this study provides a comprehensive framework for assessing population persistence under uncertainty. The proposed methodology illustrates the importance of explicitly accounting for stochasticity in ecological forecasting and offers a practical approach to inform conservation and management strategies for vulnerable species in changing environments.

2. Data

2.1. Data Source and Pre-processing

This study uses long-term population time-series data obtained from the **Global Population Dynamics Database (GPDD)**¹, one of the most comprehensive repositories of ecological population records. Developed and maintained by the Centre for Population Biology at Imperial College London, the GPDD contains population abundance data for numerous species spanning diverse ecosystems and geographical regions. In addition to population counts, each record includes associated metadata such as species identity, geographic location, sampling units, temporal resolution, and unique study identifiers (MainID). The database is organized in a wide-format structure in which successive observations are recorded across columns, facilitating the analysis of long-term population dynamics.

The complete GPDD dataset was initially imported into the analysis environment and subsequently filtered to retain records corresponding to **MainID = 6633**, which represents a long-term monitoring study of the **Black-legged Kittiwake** (*Rissa tridactyla*). From the filtered dataset, only the **SampleYear** and **Population** variables were extracted to construct the annual population time series. Records containing missing or non-numeric values in either variable were removed during pre-processing to ensure data integrity and consistency.

The final dataset consists of annual population counts spanning the period **1949–1984**, providing a continuous empirical time series for model calibration and validation. Throughout this study, the observed relationship between time and population size forms

¹The dataset is publicly available at <https://knb.ecoinformatics.org/>.

the basis for estimating deterministic growth dynamics and incorporating stochastic effects arising from demographic and environmental variability.

Year	Population Size
1949	8
1950	22
1951	24
1952	27
1953	32
⋮	⋮
1980	173
1981	179
1982	141
1983	175
1984	138

Table 1: Annual population counts of the Black-legged Kittiwake (*Rissa tridactyla*) extracted from the GPDD (MainID = 6633).

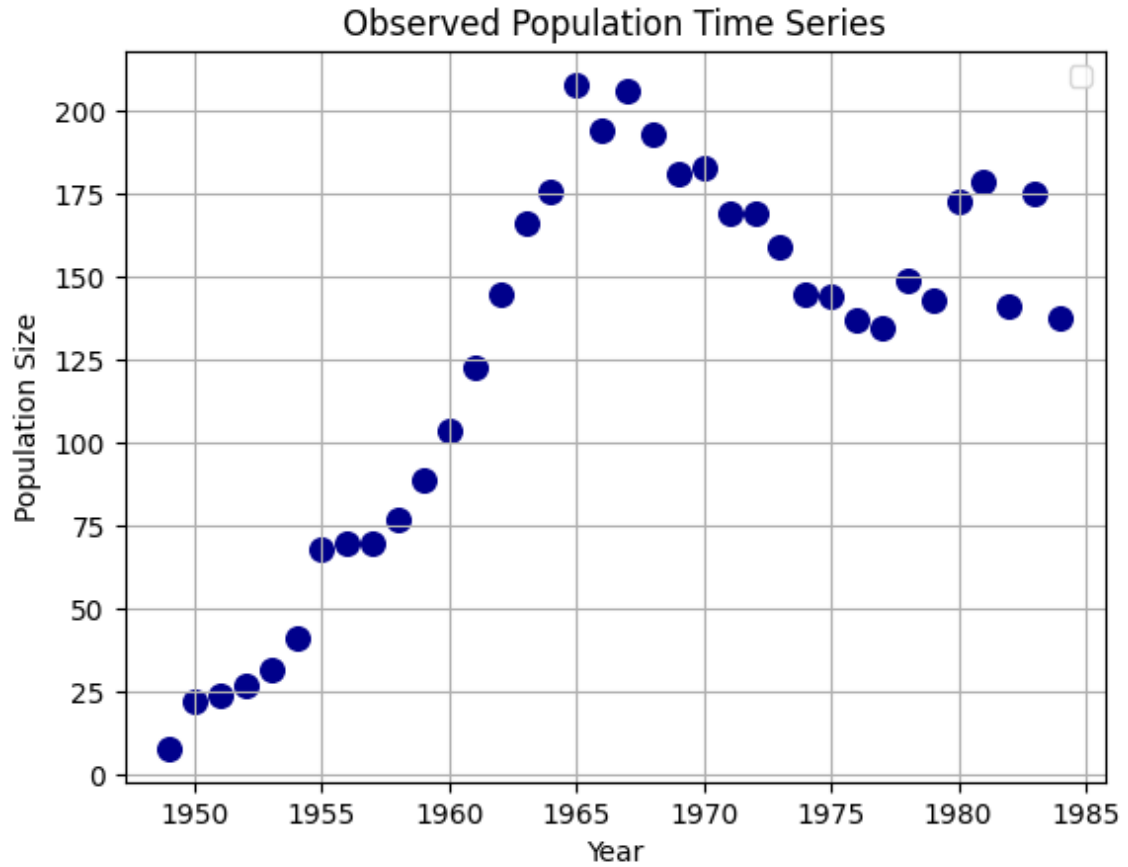


Figure 1: Observed population time series for the Black-legged Kittiwake (GPDD ID 6633). Each point represents the recorded population size in a given year. The trend exhibits initial growth followed by saturation and fluctuations, characteristic of logistic population dynamics

2.2. Species Profile: Black-legged Kittiwake (*Rissa tridactyla*)

The ‘Black-legged Kittiwake’ (*Rissa tridactyla*) is a long-lived pelagic seabird belonging to the family ‘Laridae’ and is one of the most extensively studied seabird species in the Northern Hemisphere ([6]). It breeds colonially on steep coastal cliffs throughout the North Atlantic and North Pacific Oceans, with colony sizes ranging from a few dozen to many thousands of breeding pairs. Outside the breeding season, individuals disperse widely across offshore marine environments, spending the majority of their annual cycle at sea. This strong dependence on pelagic ecosystems makes the species particularly sensitive to changes in oceanographic conditions, marine productivity, and climate variability ([10]).

Black-legged Kittiwakes exhibit a classic slow life-history strategy characterized by delayed sexual maturity (typically 3-5 years), low annual reproductive output, high adult survival and strong breeding-site fidelity ([6]). Consequently, long-term population persistence depends more on the maintenance of high adult survival and consistent juvenile recruitment than on rapid intrinsic population growth. Although this life-history strategy buffers populations against short-term environmental fluctuations, prolonged periods of poor breeding success or reduced juvenile survival can have persistent demographic consequences because recruitment into the breeding population is inherently slow ([16]).

During the breeding season, Black-legged Kittiwakes function as central-place foragers, repeatedly travelling between nesting colonies and offshore feeding grounds to provision their chicks. Their diet consists predominantly of small pelagic forage fish, including sandeels (*Ammodytes* spp.), capelin (*Mallotus villosus*), juvenile herring (*Clupea harengus*), and sprat (*Sprattus sprattus*), although crustaceans and zooplankton may become important alternative prey when fish abundance is low ([11]). Numerous long-term ecological studies have demonstrated that reductions in forage-fish availability translate directly into decreased chick growth, reduced breeding success, lower fledging rates, diminished juvenile recruitment and ultimately population decline ([10, 3]).

Because its demographic performance responds rapidly to fluctuations in marine food-web structure, prey availability, ocean productivity and climatic variability, the Black-legged Kittiwake is widely recognised as a sentinel or indicator species of marine ecosystem health ([27]). Breeding success, reproductive output, and adult survival are strongly influenced by environmental drivers such as sea-surface temperature, changes in plankton communities, fisheries-induced reductions in forage-fish stocks and marine heatwaves, all of which affect both the quantity and quality of prey available during the energetically demanding breeding season ([10]). In addition, recent studies have highlighted the importance of carry-over effects operating across the annual cycle, demonstrating that environmental conditions experienced during the non-breeding season influence adult survival, body condition, and subsequent reproductive success ([28]). Consequently, both breeding and non-breeding environments contribute significantly to the long-term population dynamics of the species.

Despite its broad circumpolar distribution, many regional populations of the Black-legged Kittiwake have undergone substantial declines over recent decades. These declines are attributed to multiple interacting stressors, including reduced forage-fish availability, climate-driven changes in marine ecosystems, overexploitation of fish stocks, increasing frequency of extreme weather events, pollution, habitat degradation, offshore development and other anthropogenic disturbances ([3]). Reflecting these persistent declines,

the species is currently classified as Vulnerable on the ‘IUCN Red List’ (BirdLife International, 2019 [2]).

The ecological characteristics of ‘*Rissa tridactyla*’-its long lifespan, low reproductive rate, strong dependence on variable marine resources and pronounced sensitivity to environmental stochasticity-make it an ideal model species for population viability analysis, stochastic demographic modelling and climate-change ecology ([16]). Long-term monitoring data, when integrated with deterministic and stochastic population models, provide a powerful framework for quantifying extinction risk, identifying the demographic processes that most strongly influence population persistence and evaluating the potential effectiveness of conservation and management strategies. Consequently, the Black-legged Kittiwake has become an important model organism for understanding the demographic consequences of environmental variability and for developing predictive tools that support evidence-based conservation under ongoing climate change.

3. Methodology and Results

3.1. Model Selection and AIC Based Comparison

To model the population dynamics of the selected time series, we considered three classical deterministic growth models: the **Exponential, Logistic, and Gompertz** models. Each of these models captures distinct assumptions about population growth and regulation mechanisms.

Exponential: One of the primary good candidates for the study of population dynamics remains the exponential law of population growth ([13, 1]). It Assumes constant per capita growth rate without environmental limits. The population density of the population at any time t is given as: $N(t) = N(t_0)e^{r(t-t_0)}$, where r is the intrinsic growth rate and $N(t_0)$ is the initial population density.

Logistic: The first step to enhance the biological realism of the exponential growth model, density-dependent regulation must be incorporated. The logistic growth model provides the simplest and most widely used framework for this ([26, 18]). It introduces a carrying capacity K that limits population growth. Consequently the growth rate decreases as the population size approaches K , ultimately vanishing when the carrying capacity is reached. The population density of the population at any time t is given as :

$$N(t) = \frac{K}{1 + e^{-r(t-t_0)}}$$

Gompertz: The Gompertz model is a commonly used sigmoidal growth model for populations exhibiting a gradual decline in growth rate ([7, 19]). It assumes that the growth rate decays exponentially as the population approaches its carrying capacity, providing greater flexibility than the logistic model in many ecological applications. The model is defined as, $N(t) = Ke^{-e^{-r(t-t_0)}}$.

We fitted each model to the empirical population data using nonlinear least squares optimization (curve fit from `scipy.optimize` in Python). Initial parameter values were chosen based on observed data properties (e.g., initial population size, median year, and estimated growth peaks). For the Gompertz model, we included parameter bounds to ensure numerical stability and ecological realism.

To compare model fits, we computed the Akaike Information Criterion (AIC), which penalizes model complexity while rewarding goodness-of-fit. For a model with k parameters and residual variance σ^2 , the AIC is given by: $AIC = 2k - 2\ln L$

where $\ln L$ is the log-likelihood under Gaussian noise assumption. The model with the lower AIC is selected. Below are the AIC scores for the models obtained after fitting:

Table 2: AIC scores for the fitted models

Model	AIC
Logistic	328.69
Gompertz	334.24
Exponential	380.30

Among the three, the Logistic model achieved the lowest AIC, making it the most appropriate choice for this dataset. Its characteristic sigmoidal shape aligns with the ecological assumption of early growth followed by saturation due to resource limitation.

This conclusion is further validated by the visual comparison in Figure 3 where the Logistic model closely follows the observed trajectory, especially around the inflection point and plateau. The Gompertz model slightly underestimates late growth, while the Exponential model overpredicts it due to its lack of a carrying capacity.

3.2. Incorporating Stochasticity: Demographic, Environmental, and Combined Noise

While the deterministic logistic model captures general population trends, real-world population dynamics are inherently noisy. Stochasticity can arise from two primary sources:

Demographic noise: This results from the probabilistic nature of birth and death events, which are especially influential in small populations. This type of noise scales with population size.

Environmental noise: It reflects fluctuations in external factors (e.g., food supply, weather) that affect the entire population. Its effect scales with the square of the population size.

Combined noise: This incorporates both demographic and environmental variability simultaneously, offering a more realistic approximation in complex ecosystems.

To formally incorporate these sources, we extend the logistic model by modelling population size N_t at time t using the stochastic form:

$$N(t+1) = N(t) + r_m N(t) \left(1 - \frac{N(t)}{K}\right) + \text{Noise}.$$

3.2.1. Likelihood Formulation

Each stochastic model assumes Gaussian noise with zero mean and a variance structure specific to the noise type.

Model 1: Demographic Noise

First, we incorporate only the demographic noise. The stochastic model to be considered is given below:

$$dN(t) = r_m N(t) \left(1 - \frac{N(t)}{K}\right) dt + \sigma_d \sqrt{N(t)} dW(t),$$

where r_m is the intrinsic growth rate, K is the carrying capacity, σ_d is the strength of demographic noise and $W(t)$ is the standard Wiener process.

item Noise scales with $\sqrt{N(t)}$: fluctuations are larger in small populations.

235 item This formulation defines a stochastic differential equation (SDE) with multiplica-
 236 tive noise.

237 To approximate this SDE in discrete time, we apply the ‘Euler-Maruyama method’
 238 with time step Δt :

$$N(t + \Delta t) = N(t) + r_m N(t) \left(1 - \frac{N(t)}{K} \right) \Delta t + \epsilon(t),$$

239 where

$$\epsilon(t) \sim \mathcal{N} \left(0, \sigma_d^2 N(t) \Delta t \right).$$

240 Hence,

$$N(t + \Delta t) \sim \mathcal{N} \left(\mu(t), \sigma_d^2 N(t) \Delta t \right),$$

241 where

$$\mu(t) = N(t) + r_m N(t) \left(1 - \frac{N(t)}{K} \right) \Delta t.$$

242 Assuming that the process satisfies the Markov property, the likelihood function can
 243 be written as,

$$\begin{aligned} L(r_m, K, \sigma_d^2 \mid N(t_0), \dots, N(t_n)) &= \mathbb{P} \left(N(t_0), \dots, N(t_n) \mid r_m, K, \sigma_d^2 \right) \\ &= \mathbb{P}(N(t_0)) \prod_{i=1}^n \mathbb{P}(N(t_i) \mid N(t_{i-1})). \end{aligned}$$

244 From the discretized model,

$$N(t_i) \mid N(t_{i-1}) \sim \mathcal{N} \left(\mu_{t_i}, \sigma_d^2 N(t_{i-1}) \Delta t \right),$$

245 where

$$\mu_{t_i} = N(t_{i-1}) + r_m N(t_{i-1}) \left(1 - \frac{N(t_{i-1})}{K} \right) \Delta t.$$

246 Therefore, the corresponding conditional density is

$$\mathbb{P} \left(N(t_i) \mid N(t_{i-1}) \right) = \frac{1}{\sqrt{2\pi\sigma_d^2 N(t_{i-1}) \Delta t}} \exp \left[-\frac{(N(t_i) - \mu_{t_i})^2}{2\sigma_d^2 N(t_{i-1}) \Delta t} \right].$$

247 Hence, the full likelihood becomes

$$L(r_m, K, \sigma_d^2) = \mathbb{P}(N(t_0)) \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma_d^2 N(t_{i-1}) \Delta t}} \exp \left[-\frac{(N(t_i) - \mu_{t_i})^2}{2\sigma_d^2 N(t_{i-1}) \Delta t} \right].$$

248 Taking logarithms yields the log-likelihood function as following:

$$\ell(r_m, K, \sigma_d^2) = -\frac{1}{2} \sum_{i=1}^n \left[\log(2\pi\sigma_d^2 N(t_{i-1})\Delta t) + \frac{(N(t_i) - \mu_{t_i})^2}{\sigma_d^2 N(t_{i-1})\Delta t} \right].$$

Model 2: Environmental Noise

To incorporate environmental stochasticity, we consider the stochastic logistic growth model

$$dN(t) = r_m N(t) \left(1 - \frac{N(t)}{K} \right) dt + \sigma_e N(t) dW_t,$$

where σ_e^2 denotes the environmental noise intensity. Since the diffusion term is proportional to the population size, the variance of the stochastic fluctuations increases quadratically with population density. This formulation captures random environmental perturbations that simultaneously affect all individuals in the population.

Applying the Euler-Maruyama discretization with time step Δt yields

$$N(t + \Delta t) = N(t) + r_m N(t) \left(1 - \frac{N(t)}{K} \right) \Delta t + \epsilon(t),$$

where

$$\epsilon(t) \sim \mathcal{N}(0, \sigma_e^2 N(t)^2 \Delta t).$$

Hence,

$$N(t + \Delta t) \sim \mathcal{N}(\mu(t), \sigma_e^2 N(t)^2 \Delta t),$$

where

$$\mu(t) = N(t) + r_m N(t) \left(1 - \frac{N(t)}{K} \right) \Delta t.$$

Assuming that the population process satisfies the Markov property, the likelihood function can be expressed as,

$$L = \mathbb{P}(N(t_0)) \prod_{i=1}^n \mathbb{P}(N(t_i) | N(t_{i-1})).$$

From the discretized model,

$$N(t_i) | N(t_{i-1}) \sim \mathcal{N}(\mu_{t_i}, \sigma_e^2 N(t_{i-1})^2 \Delta t),$$

where

$$\mu_{t_i} = N(t_{i-1}) + r_m N(t_{i-1}) \left(1 - \frac{N(t_{i-1})}{K} \right) \Delta t.$$

Therefore, the conditional probability density is

$$\mathbb{P}(N(t_i) | N(t_{i-1})) = \frac{1}{\sqrt{2\pi\sigma_e^2 N(t_{i-1})^2 \Delta t}} \exp \left[-\frac{(N(t_i) - \mu_{t_i})^2}{2\sigma_e^2 N(t_{i-1})^2 \Delta t} \right].$$

Hence, the log-likelihood function is

$$\ell(r_m, K, \sigma_e^2) = -\frac{1}{2} \sum_{i=1}^n \left[\log(2\pi\sigma_e^2 N(t_{i-1})^2 \Delta t) + \frac{(N(t_i) - \mu_{t_i})^2}{\sigma_e^2 N(t_{i-1})^2 \Delta t} \right].$$

The variance of the stochastic increment is: $\text{Var}(\epsilon(t)) = \sigma_e^2 N(t)^2 \Delta t$, which increases quadratically with population size, reflecting the fact that environmental fluctuations simultaneously influence all individuals in the population.

Model 3: Combined Demographic and Environmental Noise

To account for both demographic and environmental stochasticity, we consider the stochastic logistic model

$$dN(t) = r_m N(t) \left(1 - \frac{N(t)}{K} \right) dt + \sqrt{\sigma_d^2 N(t) + \sigma_e^2 N(t)^2} dW_t,$$

where σ_d^2 and σ_e^2 denote the demographic and environmental noise intensities, respectively. The demographic component is proportional to the population size and therefore has a stronger influence at low population densities, whereas the environmental component scales with the square of the population size and becomes increasingly important for larger populations. Consequently, this model simultaneously incorporates both sources of stochasticity.

Applying the Euler-Maruyama discretization with time step Δt gives

$$N(t + \Delta t) = N(t) + r_m N(t) \left(1 - \frac{N(t)}{K} \right) \Delta t + \epsilon(t),$$

where

$$\epsilon(t) \sim \mathcal{N}(0, (\sigma_d^2 N(t) + \sigma_e^2 N(t)^2) \Delta t).$$

Hence,

$$N(t + \Delta t) \sim \mathcal{N}(\mu(t), (\sigma_d^2 N(t) + \sigma_e^2 N(t)^2) \Delta t),$$

where

$$\mu(t) = N(t) + r_m N(t) \left(1 - \frac{N(t)}{K} \right) \Delta t.$$

Therefore, the conditional distribution is

$$N(t_i) | N(t_{i-1}) \sim \mathcal{N}(\mu_{t_i}, (\sigma_d^2 N(t_{i-1}) + \sigma_e^2 N(t_{i-1})^2) \Delta t),$$

with

$$\mu_{t_i} = N(t_{i-1}) + r_m N(t_{i-1}) \left(1 - \frac{N(t_{i-1})}{K} \right) \Delta t.$$

284 The corresponding conditional probability density is

$$\mathbb{P}(N(t_i) | N(t_{i-1})) = \frac{1}{\sqrt{2\pi (\sigma_d^2 N(t_{i-1}) + \sigma_e^2 N(t_{i-1})^2) \Delta t}} \exp \left[-\frac{(N(t_i) - \mu_{t_i})^2}{2 (\sigma_d^2 N(t_{i-1}) + \sigma_e^2 N(t_{i-1})^2) \Delta t} \right].$$

285 Assuming conditional independence of successive observations, the likelihood function
286 is

$$L = \mathbb{P}(N(t_0)) \prod_{i=1}^n \mathbb{P}(N(t_i) | N(t_{i-1})).$$

287 Taking logarithms yields the log-likelihood function:

$$\ell(r_m, K, \sigma_d^2, \sigma_e^2) = -\frac{1}{2} \sum_{i=1}^n \left[\log (2\pi (\sigma_d^2 N(t_{i-1}) + \sigma_e^2 N(t_{i-1})^2) \Delta t) + \frac{(N(t_i) - \mu_{t_i})^2}{(\sigma_d^2 N(t_{i-1}) + \sigma_e^2 N(t_{i-1})^2) \Delta t} \right].$$

288 The variance of the stochastic increment is therefore

$$\text{Var}(\epsilon(t)) = (\sigma_d^2 N(t) + \sigma_e^2 N(t)^2) \Delta t,$$

289 which is the sum of the demographic and environmental contributions, reflecting the
290 combined effects of both sources of stochasticity on population fluctuations.

291 3.2.2. Parameter Estimation and Model Comparison

292 To estimate the parameters of each stochastic model, we employed Maximum Like-
293 lihood Estimation (MLE). Specifically, we defined a custom log-likelihood function for
294 each noise type (demographic, environmental, and combined) and minimized its negative
295 using the `minimize` function from the `scipy.optimize` module in Python. This numer-
296 ical optimization procedure identifies the parameter values that maximize the likelihood
297 of observing the data under the specified stochastic model.

298 Initial parameter values were selected using data-driven heuristics, including the empir-
299 ical growth rate, the maximum observed population size, and the variance of population
300 increments. Appropriate parameter bounds were imposed to ensure biologically mean-
301 ingful estimates, such as positive growth rates, carrying capacities, and noise intensities.

302 Model comparison was performed using the Akaike Information Criterion (AIC).

Table 3: Maximum likelihood estimates (MLE) and AIC values for the demographic, environmental, and combined noise models.

Model	r_m	K	σ_d	σ_e	LogLik	AIC
Demographic	0.339670	168.58	1.43	—	-143.35	292.70
Environmental	0.504405	163.52	—	0.26	-165.29	336.57
Combined	0.339670	168.58	1.43	0.00	-143.35	294.70

As shown in Table 3, the demographic noise model yields the lowest AIC value, indicating that it provides the best balance between model fit and complexity. This suggests that intrinsic stochasticity arising from individual birth-death events is the dominant source of variability in the observed population dynamics. Consequently, we adopt the demographic noise model as the stochastic framework for subsequent simulations and extinction risk analyses.

3.2.3. Comparison of Curve-Fit and MLE-Based Logistic Models

To assess the performance of the deterministic logistic model relative to the stochastic MLE-based framework, we compared two different parameter estimation approaches.

1. **Nonlinear Curve Fitting:** The classical logistic growth model was fitted directly to the observed population data using the `curve_fit` routine from the `scipy.optimize` package. This method assumes deterministic logistic growth with independently and normally distributed residuals, and estimates the parameters K , r , and t_0 by minimizing the residual sum of squares.
2. **MLE-Based Stochastic Simulation:** The maximum likelihood estimates of the intrinsic growth rate (r_m) and carrying capacity (K), obtained under the demographic stochasticity model, were used to generate population trajectories by numerically integrating the stochastic logistic equation using the Euler-Maruyama scheme.

The resulting trajectories were then compared with the observed population time series. As shown in Figure 4, both approaches capture the overall growth dynamics and provide comparable fits to the data, accurately reproducing the initial exponential growth phase as well as the subsequent saturation towards the carrying capacity. The MLE-based stochastic trajectory (red solid line) shows slightly improved agreement with the observations in the vicinity of the inflection point, which may be attributed to its likelihood-based estimation procedure and explicit incorporation of demographic stochasticity. In contrast, the deterministic logistic fit obtained via nonlinear least squares (green dashed line) produces a smooth growth curve but is less effective in capturing the variability observed during the later stages of population growth.

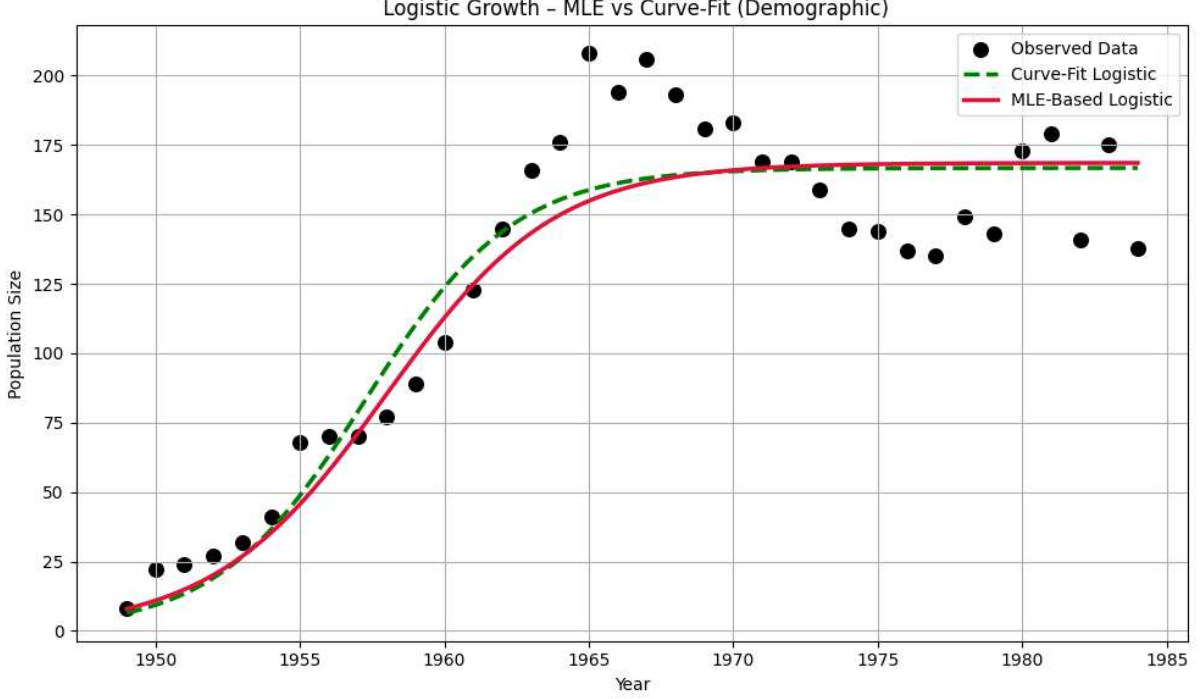


Figure 2: Comparison of logistic growth curves obtained via nonlinear curve fitting (green dashed line) and MLE-based simulation (red solid line) under the demographic noise model.

3.2.4. Confidence and Prediction Intervals for the Demographic Model

Following parameter estimation by maximum likelihood estimation (MLE), the uncertainty associated with the fitted population trajectory was quantified by constructing both confidence intervals (CIs) and prediction intervals (PIs). Confidence intervals describe the uncertainty in the estimated mean trajectory resulting from the estimation of model parameters, whereas prediction intervals additionally account for the inherent stochastic variability of the population process, thereby providing a probabilistic range for future population observations.

The fitted mean trajectory was obtained by numerically integrating the deterministic component of the stochastic logistic model using the MLE estimates of the parameters $\hat{r}_m$, $\hat{K}$, and $\hat{\sigma}_d$. The deterministic dynamics were approximated using the Euler discretization, $N(t + \Delta t) = N(t) + r_m N(t) \left(1 - \frac{N(t)}{K}\right) \Delta t$.

The uncertainty in the fitted trajectory was estimated using the first-order delta method. Specifically, the gradient of the predicted trajectory with respect to the model parameters was evaluated and combined with an approximate covariance matrix of the MLEs to propagate parameter uncertainty to the fitted trajectory. For computational simplicity, the covariance matrix was assumed to be diagonal, thereby neglecting correlations among the parameter estimates. The resulting confidence intervals characterize the uncertainty in the estimated mean trajectory, while the prediction intervals additionally incorporate the effects of demographic stochasticity, providing a more comprehensive measure of the uncertainty associated with future population dynamics.

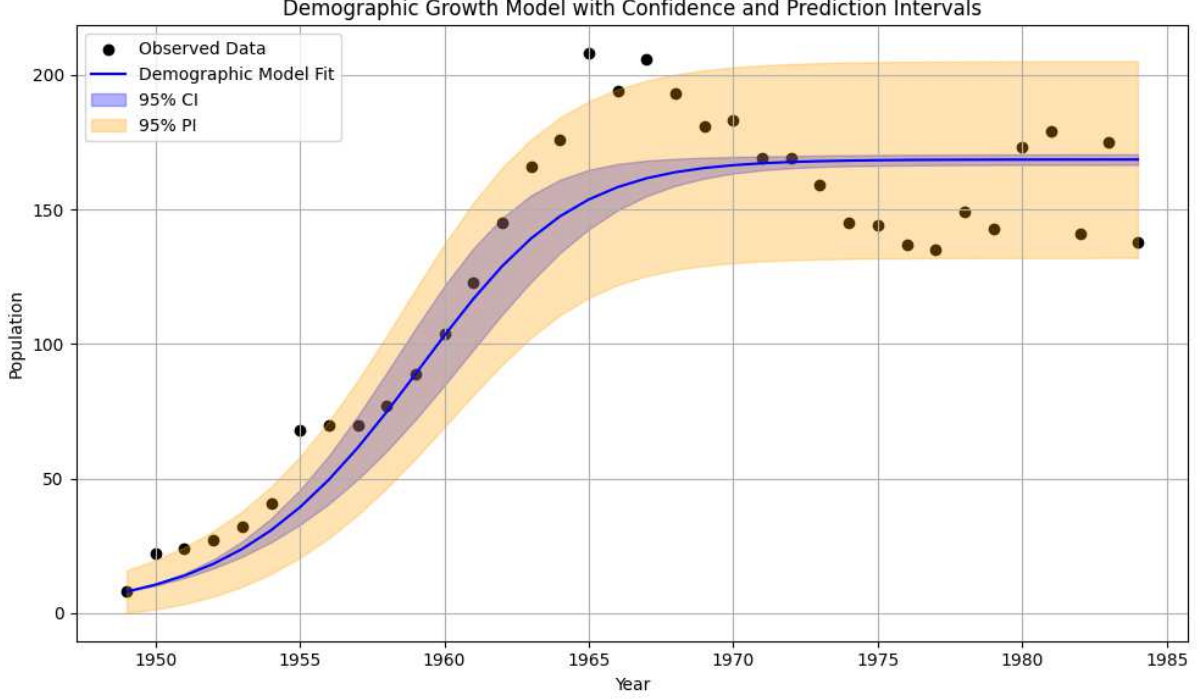


Figure 3: Fitted demographic logistic model (blue curve) with 95% confidence intervals (blue shaded region) and 95% prediction intervals (orange shaded region). The confidence interval reflects uncertainty in the estimated mean trajectory due to parameter estimation, whereas the prediction interval additionally incorporates demographic process noise. Most observed data points (black circles) lie within the prediction interval, indicating an adequate model fit.

The 95% confidence interval was constructed as

$$\mu(t) \pm 1.96\sqrt{\text{Var}_{\text{param}}(N(t))},$$

where $\mu(t)$ denotes the predicted mean population size and $\text{Var}_{\text{param}}(N(t))$ represents the variance attributable to parameter uncertainty.

The 95% prediction interval was computed as

$$\mu(t) \pm 1.96\sqrt{\text{Var}_{\text{param}}(N(t)) + \sigma_d^2 N(t)\Delta t},$$

where the second term accounts for demographic process noise.

Overall, the demographic stochastic logistic model accurately reproduces the observed population dynamics, including the initial exponential growth, the intermediate growth phase, and the eventual approach to carrying capacity. Most observations lie within the prediction interval, indicating that the model adequately captures the stochastic variability present in the data. Furthermore, the widening of the prediction interval over time reflects the accumulation of uncertainty due to demographic stochasticity.

3.3. Forecasting and Extinction Risk under Demographic Stochasticity

Assessing extinction risk is a central objective in population ecology and conservation biology. Although deterministic population models describe the expected trajectory of a population, they fail to capture extinction events arising solely from demographic

stochasticity, which can be particularly important for small populations. Consequently, populations with positive expected growth rates may nevertheless face a substantial risk of extinction due to random fluctuations in births and deaths.

In the present study, extinction was defined using a quasi-extinction threshold of 10 individuals, representing a population size below which recovery is considered unlikely because of demographic stochasticity, inbreeding depression, or Allee-type effects. The probability of extinction was estimated using Monte Carlo simulations of the stochastic logistic model over a forecasting horizon of 30 years. For each simulation, the population was initialized at the observed initial abundance and evolved according to $N(t + \Delta t) = N(t) + r_m N(t) \left(1 - \frac{N(t)}{K}\right) \Delta t + \sigma_d \sqrt{N(t)} \epsilon(t)$, where $\epsilon(t) \sim \mathcal{N}(0, \Delta t)$. A simulation was classified as an extinction event if the population size fell below the quasi-extinction threshold at any time during the simulation period. This procedure was repeated for 1000 independent realizations, and the extinction probability was estimated as the proportion of simulations that resulted in quasi-extinction.

3.3.1. Results and Interpretation

Monte Carlo simulations of the demographic stochastic logistic model yielded an estimated quasi-extinction probability of 53.1% over a 30-year time horizon, indicating that more than half of the simulated population trajectories declined below the quasi-extinction threshold of 10 individuals within the study period. Although the corresponding deterministic logistic model predicts convergence towards the carrying capacity, the incorporation of demographic stochasticity substantially increased the probability of population decline, demonstrating that random demographic fluctuations alone can generate considerable extinction risk. Such effects are particularly pronounced when population sizes are small, where stochastic variation in births and deaths may produce trajectories that deviate markedly from the deterministic expectation. Because simulations were terminated once the population size fell below the quasi-extinction threshold, trajectories crossing this threshold were classified as extinction events and were not permitted to recover, thereby providing an estimate of quasi-extinction risk under the assumed modelling framework.

The predicted extinction risk is consistent with the life-history characteristics of the Black-legged Kittiwake (*Rissa tridactyla*), a relatively long-lived seabird characterized by low fecundity and high adult survival. In such species, long-term population persistence depends more strongly on sustained recruitment and survival than on rapid intrinsic population growth. Consequently, stochastic variation in reproductive success or juvenile survival can have lasting demographic consequences despite comparatively stable adult survival ([16]). The substantial quasi-extinction probability obtained in the present study therefore agrees with the well-established sensitivity of long-lived seabirds to demographic stochasticity. This interpretation is further supported by the ecological characteristics of Black-legged Kittiwake populations. Breeding success is strongly influenced by the availability of energy-rich forage fish, particularly sandeels. Numerous long-term field studies have shown that reductions in prey abundance are associated with lower breeding success, reduced fledging rates, diminished juvenile recruitment, and subsequent declines in colony size ([11, 3]). Because individuals typically require several years to reach reproductive maturity, repeated breeding failures can generate prolonged recruitment deficits whose demographic consequences may persist for many breeding seasons. Although these eco-

logical mechanisms were not explicitly represented in the present demographic stochastic model, they provide a biologically plausible explanation for why stochastic demographic variation can substantially increase extinction risk in natural populations.

It is important to recognize that the logistic model assumes a constant carrying capacity, whereas the effective carrying capacity of natural Black-legged Kittiwake populations is likely to vary through time owing to fluctuations in prey availability, ocean productivity, fisheries pressure and climatic conditions. Consequently, the estimated extinction probability should be interpreted as representing a stationary environmental scenario. Continued climate-driven changes in marine ecosystems may further reduce habitat quality and prey availability, potentially increasing extinction risk beyond that predicted by the present model. Similarly, the species' strong fidelity to natal and breeding colonies, while promoting colony stability under favourable conditions, may limit rapid colonization of newly suitable habitats as environmental conditions shift, thereby increasing the risk of local extinction even when the broader metapopulation persists ([10, 27]).

From a conservation perspective, an estimated quasi-extinction probability of 53.1 % within three decades represents a substantial level of demographic risk for a species that has experienced widespread population declines across much of its breeding range. Although the present model does not explicitly evaluate management interventions, the results suggest that conservation strategies aimed at maintaining breeding productivity, enhancing juvenile recruitment, preserving high adult survival, and protecting critical breeding and marine foraging habitats may contribute to reducing long-term extinction risk. Ecosystem-based management of forage-fish resources and continued long-term demographic monitoring are likewise expected to improve population persistence ([21, 5]). Because the Black-legged Kittiwake is widely recognized as an indicator species of marine ecosystem health, improvements in its demographic performance would also provide evidence of broader recovery in marine food-web integrity.

The present study demonstrates that demographic stochasticity alone can generate substantial extinction risk even when deterministic population dynamics predict persistence. Incorporating additional ecological processes, including environmental stochasticity, density-dependent demographic variation, Allee effects, habitat degradation, and explicit management interventions, would provide a more comprehensive assessment of long-term population viability and represents an important direction for future research.

4. Discussion

This study presented a comprehensive modelling framework for analysing the population dynamics of a long-term ecological time series, using the Black-legged Kittiwake (GPDD ID 6633) as a case study. Employing empirical observations from the Global Population Dynamics Database (GPDD), we investigated the species' population growth through both deterministic and stochastic modelling approaches. By integrating statistical model selection, parameter estimation, stochastic simulations and extinction risk analysis within a single workflow, the proposed framework provides a systematic methodology for understanding the mechanisms governing population persistence and decline.

A key strength of the proposed framework is its methodology for selection of suitable model for study. Primarily, three deterministic growth models were compared using the Akaike Information Criterion (AIC) to identify the most appropriate representation

of the observed population dynamics. The three candidate models chosen are the exponential, logistic, and Gompertz model. Likewise, different sources of stochasticity—namely demographic, environmental, and combined stochasticity—were evaluated using likelihood-based model selection to determine the dominant source of uncertainty underlying the data. This strategy allows us to identify and select the most appropriate model which adequately describes the observed deterministic and stochastic dynamics. This particular objective model-selection strategy reduces subjective model choice and enhances the reliability of subsequent forecasting and extinction risk assessment.

The analysis identified the logistic growth model as the most suitable deterministic description of the observed population dynamics. This finding suggests that density-dependent regulation plays an important role in governing the long-term behaviour of the Black-legged Kittiwake population. Unlike the exponential model, which assumes unlimited resources and unrestricted growth, the logistic model accounts for the gradual reduction in population growth as the population approaches its carrying capacity. Such behaviour is biologically consistent with natural populations, where competition for food resources, breeding sites and nesting habitat imposes constraints on population expansion. Although the Gompertz model is also capable of describing density-dependent growth, the lower AIC support obtained in the present study indicates that the logistic formulation provides a better representation of the available observations. After identifying the logistic growth model with demographic stochasticity as the best-supported model, both short and long-term population forecasts were generated together with confidence and prediction intervals. The incorporation of prediction intervals provides considerably more information than deterministic trajectories alone by explicitly quantifying the uncertainty associated with future population sizes. Such probabilistic forecasts are particularly valuable for ecological management, where conservation decisions must frequently be made despite substantial uncertainty regarding future environmental conditions and demographic processes.

The next stepping stone of this framework was to estimate extinction probabilities under demographic stochasticity at a fixed 30-year horizon as well as over time. The results indicate that the population faces a considerable risk of quasi-extinction, with demographic fluctuations during the early stages playing a dominant role in determining long-term persistence. These findings demonstrate that even when deterministic dynamics predict long-term persistence, stochastic fluctuations may substantially increase extinction risk by driving populations below biologically meaningful thresholds. In short, the prediction of extinction probabilities were less accurately predicted by the deterministic framework alone, particularly for populations experiencing substantial demographic variability.

The identification of demographic stochasticity as the dominant source of variability also provides important ecological insight. Demographic stochasticity originates from random variation in individual birth and death events and is generally expected to exert stronger effects when population sizes are relatively small. The present results therefore suggest that intrinsic demographic fluctuations contribute more substantially to the observed population variability than environmental stochasticity over the available observation period. This highlights the importance of incorporating stochastic demographic processes into population viability analyses and emphasizes that random individual-level processes can significantly influence long-term ecological outcomes.

An another important strength of this analysis is its highly generalized structure enabling it to be applicable to other species data-sets ranging from different taxa and life-history strategies without modifying the underlying implementation. This broad range of applicability of this analysis comes from its reliability on standard statistical estimation techniques and flexible stochastic differential equation models. Consequently, the framework provides a unified, reproducible and computationally efficient tool for population modelling, forecasting and extinction risk assessment across a broad spectrum of ecological systems.

Overall, the proposed framework demonstrates the advantages of integrating deterministic growth models with stochastic population dynamics in a statistically rigorous manner. By combining model selection, parameter estimation, uncertainty quantification, forecasting and extinction probability estimation within a unified computational framework, the methodology provides a practical approach for analysing long-term ecological time-series data. Such integrated analyses have considerable potential to support conservation planning, population viability assessment and evidence-based ecological decision-making under uncertainty.

5. Limitations and Future Research

While the proposed framework captures several important aspects of population dynamics, it has certain limitations that suggest promising avenues for future research.

One limitation of the present study is the relatively small size of the available population time series, which may reduce the precision and stability of parameter estimation under classical frequentist inference. In such situations, Bayesian inference provides a natural alternative by incorporating prior biological knowledge into the estimation process, thereby improving parameter regularization and reducing estimation uncertainty. Furthermore, Bayesian methods yield full posterior distributions of model parameters, allowing a more comprehensive quantification of uncertainty and facilitating probabilistic predictions for ecological systems with limited observational data.

Another limitation arises from the reliance on a single best-fitting model selected using the Akaike Information Criterion (AIC). Although AIC is a widely used model selection criterion, choosing only the model with the minimum AIC implicitly assumes that it provides an adequate representation of the underlying population dynamics, even when several competing models receive comparable empirical support. Future studies could account for model-selection uncertainty through multi-model inference, including model averaging or ensemble modelling, in which predictions from multiple plausible candidate models are combined to obtain more robust forecasts. In addition, recent advances in machine learning, such as random forests and Gaussian process regression, offer flexible data-driven alternatives for capturing complex nonlinear population dynamics while requiring fewer structural assumptions than conventional mechanistic models.

The proposed stochastic density-dependent framework captures the essential features of population regulation but does not explicitly incorporate several biologically relevant processes that may influence long-term dynamics. These include Allee effects, age or stage structure, spatial heterogeneity, dispersal, immigration and emigration, harvesting, species interactions, and climate-driven environmental variability. Incorporating these ecological mechanisms into the present framework would enhance its biological realism

and improve its predictive capability for assessing long-term population persistence, resilience and extinction risk under changing environmental conditions.

Acknowledgement

The first author Riddhiman Chattopadhyay gratefully acknowledges Dr. Amiya Ranjan Bhowmick (ICTS, Mumbai) for suggesting this research problem. His insightful suggestion significantly enhanced the quality of this work. The author Richik Bandyopadhyay is greatly acknowledged to Council of Scientific and Industrial Research (CSIR), India for funding Senior Research Fellowship. (Grant Number: 09/0093(13690)/2022-EMR-I)

Appendix: Python Code for Simulation Studies/Results

```
import pandas as pd
import numpy as np
from scipy.optimize import minimize
from scipy.optimize import curve_fit
import matplotlib.pyplot as plt
```

1: Relevant Packages Used

```
# Load the Excel file
file_path = "C:\\Users\\riddh\\Desktop\\Project\\Black-legged Kittiwake
(Rissa tridactyla, GPDD Main ID 6633).xlsx"
df = pd.read_excel(file_path, sheet_name='Sheet1')

# Extract and clean the relevant columns
df = df[["SampleYear", "Population"]].dropna()
df["SampleYear"] = pd.to_numeric(df["SampleYear"], errors="coerce")
df["Population"] = pd.to_numeric(df["Population"], errors="coerce")
df = df.dropna().reset_index(drop=True)
```

2: Code for Section 2.1: Data Extraction

```
plt.scatter(years, N_data, color='darkblue', s=70,)
plt.legend()
plt.title("Observed Population Time Series")
plt.xlabel("Year")
plt.ylabel("Population Size")
plt.grid(True)
```

3: Code for Section 2.1: Figure 1

```
# --- Define Models ---
def logistic(t, K, r, t0):
    return K / (1 + np.exp(-r * (t - t0)))

def gompertz(t, K, r, t0):
    return K * np.exp(-np.exp(-r * (t - t0)))

def exponential(t, N0, r):
    return N0 * np.exp(r * (t - t[0]))
```

```

5881 # --- AIC Calculator ---
5892 def compute_aic(y_obs, y_pred, k):
5903     resid = y_obs - y_pred
5914     sigma2 = np.var(resid, ddof=k)
5925     loglik = -0.5 * len(y_obs) * (np.log(2 * np.pi * sigma2) + 1)
5936     return 2 * k - 2 * loglik
5947
5958 # --- Fit Logistic ---
5969 try:
5970     p0_logistic = [max(N_data), 0.1, np.median(years)]
5981     params_log, _ = curve_fit(logistic, years, N_data, p0=p0_logistic,
599     maxfev=10000)
6002     pred_log = logistic(years, *params_log)
6013     aic_log = compute_aic(N_data, pred_log, k=3)
6024 except:
6035     pred_log = np.full_like(N_data, np.nan)
6046     aic_log = np.inf
6057
6068 # --- Fit Gompertz with bounds and better init ---
6079 try:
6080     K0 = max(N_data)
6091     r0 = 0.01
6102     t0_0 = years[np.argmax(np.diff(N_data))] if len(N_data) > 1 else np
611     .median(years)
6123     bounds = ([0.8*K0, 0.0001, min(years)], [1.2*K0, 2, max(years)])
6134     p0_gom = [K0, r0, t0_0]
6145     params_gom, _ = curve_fit(gompertz, years, N_data, p0=p0_gom,
615     bounds=bounds, maxfev=50000)
6166     pred_gom = gompertz(years, *params_gom)
6177     aic_gom = compute_aic(N_data, pred_gom, k=3)
6188 except:
6199     pred_gom = np.full_like(N_data, np.nan)
6200     aic_gom = np.inf
6211
6222 # --- Fit Exponential ---
6233 try:
6244     p0_exp = [N_data[0], 0.1]
6255     params_exp, _ = curve_fit(exponential, years, N_data, p0=p0_exp,
626     maxfev=10000)
6276     pred_exp = exponential(years, *params_exp)
6287     aic_exp = compute_aic(N_data, pred_exp, k=2)
6298 except:
6309     pred_exp = np.full_like(N_data, np.nan)
6310     aic_exp = np.inf
6321
6332 # --- Determine Best Fit ---
6343 aic_dict = {
6354     "Logistic": aic_log,
6365     "Gompertz": aic_gom,
6376     "Exponential": aic_exp
6387 }
6398 best_model = min(aic_dict, key=aic_dict.get)
6409
6410 # --- Print Results ---
6421 print("\nAIC Comparison:")

```

```

643 2 for model, aic in aic_dict.items():
644 3     print(f"{model:12} AIC = {aic:.2f}")
645 4 print(f"\nâĀĲ Best fitting model: {best_model}")

```

4: Code for Section 3.1: Model Selection (Table 2)

```

646 1 # Define a smooth time range for plotting the fitted curves
647 2 t_vals = np.linspace(min(years), max(years), 300)
648 3
649 4 # Generate model predictions using previously fitted parameters
650 5 logistic_curve = logistic(t_vals, *params_log)
651 6 gompertz_curve = gompertz(t_vals, *params_gom)
652 7 exp_curve = exponential(t_vals, *params_exp)
653 8
654 9 # Create the plot
655 0 plt.figure(figsize=(10, 6))
656 1 plt.scatter(years, N_data, color='black', s=70, label="Observed Data")
657 2 plt.plot(t_vals, logistic_curve, label="Logistic", linewidth=2.5,
658         linestyle='--', color='green')
659 3 plt.plot(t_vals, gompertz_curve, label="Gompertz", linewidth=2.5,
660         linestyle='-.', color='blue')
661 4 plt.plot(t_vals, exp_curve, label="Exponential", linewidth=2.5,
662         linestyle=':', color='orange')
663 5
664 6 # Customize plot
665 7 plt.title("Growth Model Comparison: Logistic vs Gompertz vs Exponential
666         ")
667 8 plt.xlabel("Year")
668 9 plt.ylabel("Population Size")
669 0 plt.legend()
670 1 plt.grid(True)
671 2 plt.tight_layout()
672 3 plt.show()

```

5: Code for Section 3.1: Comparison for Fitted Models (Figure 3)

```

673 1 # Define helper function for initial guesses
674 2 def generate_initial_guesses(N_data, model="demographic"):
675 3     r_guess = 0.1
676 4     K_guess = 1.2 * np.max(N_data)
677 5     sigma_guess = np.std(np.diff(N_data)) / np.mean(N_data)
678 6     sigma_guess = np.clip(sigma_guess, 0.05, 5.0)
679 7     if model == "demographic":
680 8         return [r_guess, K_guess, sigma_guess]
681 9     elif model == "environmental":
682 0         return [r_guess, K_guess, sigma_guess]
683 1     elif model == "combined":
684 2         return [r_guess, K_guess, sigma_guess, sigma_guess]
685 3     else:
686 4         raise ValueError("Model must be one of: 'demographic', '
687         environmental', 'combined'")
688 5
689 6 # Define log-likelihood functions
690 7 def demographic_log_likelihood(params, N, dt):
691 8     r_m, K, sigma_d = params
692 9     if sigma_d <= 0 or K <= 0:

```

```

6930         return np.inf
6941     loglik = 0
6952     for i in range(1, len(N)):
6963         N_prev = N[i - 1]
6974         mu = N_prev + r_m * N_prev * (1 - N_prev / K) * dt
6985         var = sigma_d**2 * N_prev * dt
6996         if var <= 0:
7007             return np.inf
7018         loglik += -0.5 * np.log(2 * np.pi * var) - ((N[i] - mu)**2) /
702     (2 * var)
7039     return -loglik
7040
7051 def environmental_log_likelihood(params, N, dt):
7062     r_m, K, sigma_e = params
7073     if sigma_e <= 0 or K <= 0:
7084         return np.inf
7095     loglik = 0
7106     for i in range(1, len(N)):
7117         N_prev = N[i - 1]
7128         mu = N_prev + r_m * N_prev * (1 - N_prev / K) * dt
7139         var = sigma_e**2 * N_prev**2 * dt
7140         if var <= 0:
7151             return np.inf
7162         loglik += -0.5 * np.log(2 * np.pi * var) - ((N[i] - mu)**2) /
717     (2 * var)
7183     return -loglik
7194
7205 def combined_log_likelihood(params, N, dt):
7216     r_m, K, sigma_d, sigma_e = params
7227     if sigma_d <= 0 or sigma_e <= 0 or K <= 0:
7238         return np.inf
7249     loglik = 0
7250     for i in range(1, len(N)):
7261         N_prev = N[i - 1]
7272         mu = N_prev + r_m * N_prev * (1 - N_prev / K) * dt
7283         var = (sigma_d**2 * N_prev + sigma_e**2 * N_prev**2) * dt
7294         if var <= 0:
7305             return np.inf
7316         loglik += -0.5 * np.log(2 * np.pi * var) - ((N[i] - mu)**2) /
732     (2 * var)
7337     return -loglik
7348
7359 # Fit models with data-driven guesses
7360 init_dem = generate_initial_guesses(N_data, model="demographic")
7371 init_env = generate_initial_guesses(N_data, model="environmental")
7382 init_comb = generate_initial_guesses(N_data, model="combined")
7393
7404 bounds_dem = [(0, 1), (1e-5, np.max(N_data)*10), (1e-5, 100)]
7415 bounds_env = [(0, 1), (1e-5, np.max(N_data)*10), (1e-5, 100)]
7426 bounds_comb = [(0, 1), (1e-5, np.max(N_data)*10), (1e-5, 100), (1e-5,
743     100)]
7447
7458 res_dem = minimize(demographic_log_likelihood, init_dem, args=(N_data,
746     dt), bounds=bounds_dem)
7479 res_env = minimize(environmental_log_likelihood, init_env, args=(N_data

```

```

748     , dt), bounds=bounds_env)
749 res_comb = minimize(combined_log_likelihood, init_comb, args=(N_data,
750     dt), bounds=bounds_comb)
751
752 mle_dem = res_dem.x
753 mle_env = res_env.x
754 mle_comb = res_comb.x
755
756 # Extract log-likelihoods and AIC
757 loglik_dem = -res_dem.fun
758 loglik_env = -res_env.fun
759 loglik_comb = -res_comb.fun
760
761 aic_dem = 2 * 3 - 2 * loglik_dem
762 aic_env = 2 * 3 - 2 * loglik_env
763 aic_comb = 2 * 4 - 2 * loglik_comb
764
765 # Print results
766 print("=== MLE and AIC Results ===\n")
767 print("Demographic Model:")
768 print(f"   r_m      = {mle_dem[0]:.6f}")
769 print(f"   K        = {mle_dem[1]:.2f}")
770 print(f"   sigma_d  = {mle_dem[2]:.2f}")
771 print(f"   LogLik   = {loglik_dem:.2f}")
772 print(f"   AIC      = {aic_dem:.2f}\n")
773
774 print("Environmental Model:")
775 print(f"   r_m      = {mle_env[0]:.6f}")
776 print(f"   K        = {mle_env[1]:.2f}")
777 print(f"   sigma_e  = {mle_env[2]:.2f}")
778 print(f"   LogLik   = {loglik_env:.2f}")
779 print(f"   AIC      = {aic_env:.2f}\n")
780
781 print("Combined Model:")
782 print(f"   r_m      = {mle_comb[0]:.6f}")
783 print(f"   K        = {mle_comb[1]:.2f}")
784 print(f"   sigma_d  = {mle_comb[2]:.2f}")
785 print(f"   sigma_e  = {mle_comb[3]:.2f}")
786 print(f"   LogLik   = {loglik_comb:.2f}")
787 print(f"   AIC      = {aic_comb:.2f}")

```

6: Code for Section 3.2.2: Parameter Estimation and Model Comparison (Table 3)

```

788 # Define logistic function for curve fitting
789 def logistic_function(t, K, r, t0):
790     return K / (1 + np.exp(-r * (t - t0)))
791
792 # Fit logistic function directly to observed data
793 p0 = [max(N_data), 0.1, np.median(years)] # initial guess: K, r, t0
794 params_curvefit, _ = curve_fit(logistic_function, years, N_data, p0=p0,
795     maxfev=10000)
796
797 # Evaluate curve-fit logistic
798 t_vals = np.linspace(min(years), max(years), 300)
799 logistic_curvefit = logistic_function(t_vals, *params_curvefit)
800

```

```

8013 # Define MLE-based logistic function
8024 def logistic_mle(t, r, K, N0):
8035     N = np.zeros_like(t)
8046     N[0] = N0
8057     dt = t[1] - t[0]
8068     for i in range(1, len(t)):
8079         N[i] = N[i-1] + r * N[i-1] * (1 - N[i-1] / K) * dt
8080     return N
8091
8102 # Evaluate MLE-based logistic
8113 r_m_mle, K_mle, _ = mle_dem
8124 N0 = N_data[0]
8135 logistic_mle_curve = logistic_mle(t_vals, r_m_mle, K_mle, N0)
8146
8157 # Plot bot
8168 plt.figure(figsize=(10, 6))
8179 plt.scatter(years, N_data, color='black', s=70, label="Observed Data")
8180 plt.plot(t_vals, logistic_curvefit, color='green', linestyle='--',
819         linewidth=2.5, label="Curve-Fit Logistic")
8201 plt.plot(t_vals, logistic_mle_curve, color='crimson', linewidth=2.5,
821         label="MLE-Based Logistic")
8222 plt.title("Logistic Growth vs MLE vs Curve-Fit (Demographic)")
8233 plt.xlabel("Year")
8244 plt.ylabel("Population Size")
8255 plt.legend()
8266 plt.grid(True)
8277 plt.tight_layout()
8288 plt.show()

```

7: Code for Section 3.2.3: Comparison of Curve-Fit and MLE based Logistic Models (Figure 4)

```

8291 # Observed data
8302 years = df["SampleYear"].values
8313 N_data = df["Population"].values
8324 dt = 1
8335
8346 # Demographic model likelihood
8357 def demographic_log_likelihood(params, N, dt):
8368     r, K, sigma = params
8379     if sigma <= 0 or K <= 0:
8380         return np.inf
8391     loglik = 0
8402     for i in range(1, len(N)):
8413         N_prev = N[i - 1]
8424         mu = N_prev + r * N_prev * (1 - N_prev / K) * dt
8435         var = sigma**2 * N_prev * dt
8446         if var <= 0:
8457             return np.inf
8468         loglik += -0.5 * np.log(2 * np.pi * var) - ((N[i] - mu)**2) /
847         (2 * var)
8489     return -loglik
8490
8501 #MLE estimation
8512 initial_guess = [0.1, max(N_data), 1.0]
8523 bounds = [(0, 1), (1e-5, np.max(N_data) * 10), (1e-5, 100)]
8534 result = minimize(demographic_log_likelihood, initial_guess, args=(

```

```

854     N_data, dt), bounds=bounds)
855 r_mle, K_mle, sigma_mle = result.x
856 print(f"MLE Estimates:\n r = {r_mle:.4f}, K = {K_mle:.2f}, sigma = {
857     sigma_mle:.4f}")
858
859 # Euler simulation for mean prediction
860 def demographic_mean_prediction(t_vals, r, K, N0):
861     N_pred = [N0]
862     for _ in range(1, len(t_vals)):
863         N_next = N_pred[-1] + r * N_pred[-1] * (1 - N_pred[-1] / K) *
864         dt
865         N_pred.append(N_next)
866     return np.array(N_pred)
867
868 # Variance for CI/PI using delta method (Taylor expansion)
869 t_vals = years # only use observed time range
870 N0 = N_data[0]
871 pred_vals = demographic_mean_prediction(t_vals, r_mle, K_mle, N0)
872
873 # Approximate parameter covariance (use inverse Hessian or identity if
874     unavailable)
875 # For simplicity, we'll use identity matrix scaled by small variance (
876     approximate)
877 param_cov = np.diag([0.0005, 1.0, 0.01]) # user-tunable for rough CI
878
879 # Approximate gradients for delta method
880 epsilon = 1e-5
881 def gradient(func, params, idx):
882     params_eps = np.array(params)
883     params_eps[idx] += epsilon
884     return (func(params_eps) - func(params)) / epsilon
885
886 grads = np.zeros((len(t_vals), 3))
887 for i in range(3):
888     def temp_fn(p):
889         return demographic_mean_prediction(t_vals, p[0], p[1], N0)
890     grads[:, i] = gradient(temp_fn, [r_mle, K_mle, sigma_mle], i)
891
892 # Confidence and prediction intervals
893 conf_var = np.sum(grads @ param_cov * grads, axis=1)
894 pred_var = conf_var + sigma_mle**2 * pred_vals * dt
895 ci95 = 1.96 * np.sqrt(conf_var)
896 pi95 = 1.96 * np.sqrt(pred_var)
897
898 #Plot
899 plt.figure(figsize=(10, 6))
900 plt.scatter(years, N_data, color='black', label='Observed Data')
901 plt.plot(t_vals, pred_vals, color='blue', label='Demographic Model Fit',
902     )
903 plt.fill_between(t_vals, pred_vals - ci95, pred_vals + ci95, color='
904     blue', alpha=0.3, label='95% CI')
905 plt.fill_between(t_vals, pred_vals - pi95, pred_vals + pi95, color='
906     orange', alpha=0.3, label='95% PI')
907 plt.xlabel("Year")
908 plt.ylabel("Population")

```

```

909 2 plt.title("Demographic Growth Model with Confidence and Prediction
910      Intervals")
911 3 plt.legend()
912 4 plt.grid(True)
913 5 plt.tight_layout()
914 6 plt.show()

```

8: Code for Section 3.2.4: Confidence and Prediction Intervals (Figure 5)

```

915 1 def extinction_probability_given_years(years, x0=N_data[0], r=r_m_mle,
916      K=K_mle, sigma=sigma_d_mle, n_sim=1000, dt=1):
917 2
918 3     """
919 4     Computes extinction probability under demographic noise for a given
920 5     number of years.
921 6
922 7     Parameters:
923 8         years      : Time horizon (e.g., 30)
924 9         x0         : Initial population size
925 10        r          : Growth rate (MLE)
926 11        K          : Carrying capacity (MLE)
927 12        sigma      : Std. dev. of demographic noise (MLE)
928 13        n_sim      : Number of simulation runs (default: 10000)
929 14        dt         : Time step (default: 1 year)
930 15
931 16     Returns:
932 17         extinction_prob : Estimated extinction probability within given
933 18         years
934 19     """
935 20
936 21     time_steps = int(years / dt)
937 22     extinctions = 0
938 23
939 24     for _ in range(n_sim):
940 25         N = x0
941 26         for _ in range(time_steps):
942 27             growth = r * N * (1 - N / K) * dt
943 28             noise = sigma * np.sqrt(N) * np.random.normal(0, np.sqrt(dt)
944 29             ))
945 30             N = max(N + growth + noise, 0)
946 31             if N <= 10:
947 32                 extinctions += 1
948 33                 break
949 34
950 35     return extinctions / n_sim
951 36
952 37 ext_prob = extinction_probability_given_years(30)
953 38 print(f"Extinction probability in 30 years: {ext_prob:.4f}")

```

9: Code for Section 3.3.1: Probability of Extinction

954 References

- 955 [1] Alan A Berryman. *Principles of population dynamics and their application*. Garland
956 Science, 2020.

- [2] BirdLife International. *Rissa tridactyla*, 2024. e.T22694927A263770181. Accessed 21 July 2026.
- [3] Matthew J Carroll, Mark Bolton, Ellie Owen, Guy QA Anderson, Elizabeth K Mackley, Euan K Dunn, and Robert W Furness. Kittiwake breeding success in the southern north sea correlates with prior sandeel fishing mortality. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 27(6):1164–1175, 2017.
- [4] Hal Caswell. *Matrix population models*, 2001.
- [5] OSPAR Commission et al. Ospar cemp guideline. common indicator: Condition of benthic habitat communities (bh2)–common approach. ospar agreement 2018-06. updated 2022/2023. 2023.
- [6] John Coulson. *The kittiwake*. A&C Black, 2011.
- [7] Brian Dennis and José Miguel Ponciano. Density-dependent state-space model for population-abundance data with unequal time intervals. *Ecology*, 95(8):2069–2076, 2014.
- [8] Brian Dennis, Patricia L Munholland, and J Michael Scott. Estimation of growth and extinction parameters for endangered species. *Ecological monographs*, 61(2):115–143, 1991.
- [9] Stephen P Ellner and John Guckenheimer. *Dynamic models in biology*. Princeton University Press, 2006.
- [10] Morten Frederiksen, Michael P Harris, Francis Daunt, Peter Rothery, and Sarah Wanless. Scale-dependent climate signals drive breeding phenology of three seabird species. *Global Change Biology*, 10(7):1214–1221, 2004.
- [11] Morten Frederiksen, Robert W Furness, and Sarah Wanless. Regional variation in the role of bottom-up and top-down processes in controlling sandeel abundance in the north sea. *Marine Ecology Progress Series*, 337:279–286, 2007.
- [12] Abhik Ghosh and Ayanendranath Basu. Robust and efficient estimation in the parametric proportional hazards model under random censoring. *Statistics in Medicine*, 38(27):5283–5299, 2019.
- [13] Lev R Ginzburg. The theory of population dynamics: I. back to first principles. *Journal of Theoretical Biology*, 122(4):385–399, 1986.
- [14] Benjamin Gompertz. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. in a letter to francis baily, esq. frs &c. by benjamin gompertz, esq. fr s. In *Abstracts of the Papers Printed in the Philosophical Transactions of the Royal Society of London*, number 2, pages 252–253. The Royal Society London, 1833.
- [15] R Hilborn and M Mangel. *The ecological detective: Confronting models with data* princeton univ, 1997.

- [16] Catharine Horswill and Robert A Robinson. Review of seabird demographic rates and density dependence. jncc report no. 552. 2015.
- [17] Md Aktar Ul Karim, Vikram Aithal, and Amiya Ranjan Bhowmick. Random variation in model parameters: A comprehensive review of stochastic logistic growth equation. *Ecological Modelling*, 484:110475, 2023.
- [18] Sharon E Kingsland. *Modeling nature*. University of Chicago Press, 1995.
- [19] Jonas Knape, Panagiotis Besbeas, and Perry de Valpine. Using uncertainty estimates in analyses of population time series. *Ecology*, 94(9):2097–2107, 2013.
- [20] Russell Lande, Steinar Engen, and Bernt-Erik Saether. *Stochastic population dynamics in ecology and conservation*. OUP Oxford, 2003.
- [21] Robin M Little. The status of listing global bird species. *Ostrich*, 95(1):68–69, 2024.
- [22] Brett A Melbourne and Alan Hastings. Extinction risk depends strongly on factors contributing to stochasticity. *Nature*, 454(7200):100–103, 2008.
- [23] Julie AO Miller, Robert W Furness, Mark Trinder, and Jason Matthiopoulos. The sensitivity of seabird populations to density-dependence, environmental stochasticity and anthropogenic mortality. *Journal of Applied Ecology*, 56(9):2118–2130, 2019.
- [24] William F Morris and Daniel F Doak. Buffering of life histories against environmental stochasticity: accounting for a spurious correlation between the variabilities of vital rates and their contributions to fitness. *The American Naturalist*, 163(4):579–590, 2004.
- [25] Yudi Pawitan. *In All Likelihood: Statistical Modelling and Inference Using Likelihood*. Oxford University Press, Oxford, 2001.
- [26] Raymond Pearl and Lowell J Reed. On the rate of growth of the population of the united states since 1790 and its mathematical representation. *Proceedings of the national academy of sciences*, 6(6):275–288, 1920.
- [27] JF Piatt, WJ Sydeman, WJ Sydeman, JF Piatt, and HI Browman. Seabirds as indicators of marine ecosystems. *Marine Ecology Progress Series*, 352:199, 2007.
- [28] Tone K Reiertsen, Kjell E Erikstad, Tycho Anker-Nilssen, Robert T Barrett, Thierry Boulinier, Morten Frederiksen, Jacob González-Solís, David Gremillet, David Johns, Børge Moe, et al. Prey density in non-breeding areas affects adult survival of black-legged kittiwakes *rissa tridactyla*. *Marine Ecology Progress Series*, 509:289–302, 2014.
- [29] Hanno Sandvik, Tone K Reiertsen, Kjell Einar Erikstad, Tycho Anker-Nilssen, Robert T Barrett, Svein-Håkon Lorentsen, Geir Helge Systad, and Mari S Myksvoll. The decline of norwegian kittiwake populations: modelling the role of ocean warming. *Climate Research*, 60:91–102, 2014.

- 1030 [30] Anastasios Tsoularis and James Wallace. Analysis of logistic growth models. *Math-*
1031 *ematical biosciences*, 179(1):21–55, 2002.
- 1032 [31] Peter Turchin. *Complex population dynamics: a theoretical/empirical synthesis*
1033 *(MPB-35)*. Princeton university press, 2013.
- 1034 [32] Michael Turelli. Random environments and stochastic calculus. *Theoretical Popula-*
1035 *tion Biology*, 12(2):140–178, 1977.