Switching state-space models for modeling penguin population dynamics

Yousef El-Laham[1] · Mónica Bugallo[1] · Heather J. Lynch[2]

Abstract

Tracking individual animals through time using mark-recapture methods is the gold standard for understanding how environmental conditions influence demographic rates, but applying such tags is often infeasible due to the difficulty of catching animals or attaching marks/tags without influencing behavior or survival. Due to the logistical challenges and emerging ethical concerns with flipper banding penguins, relatively little is known about spatial variation in demographic rates, spatial variation in demographic stochasticity, or the role that stochasticity may play in penguin population dynamics. Here we describe how adaptive importance sampling can be used to fit age-structured population models to time series of point counts. While some demographic parameters are difficult to learn through point counts alone, others can be estimated, even in the face of missing data. Here we demonstrate the application of adaptive importance sampling using two case studies, one in which we permit immigration and another permitting regime switching in reproductive success. We apply these methods to extract demographic information from several time series of observed abundance in gentoo and Adélie penguins in Antarctica. Our method is broadly applicable to time series of abundance and provides a feasible means of fitting age-structured models without marking individuals.

✉ Heather J. Lynch
heather.lynch@stonybrook.edu

Yousef El-Laham
yousef-ellaham@stonybrook.com

[1] Department of Electrical & Computer Engineering, Stony Brook University, Stony Brook, NY 11794, USA

[2] Institute for Advanced Computational Science, Stony Brook University, Stony Brook, NY 11794, USA

1 Introduction

Understanding the survival and reproduction of organisms has traditionally required the tracking of individuals through time, usually through bands, tags, or the addition of distinguishing marks that persist through time (e.g., marks on turtle scutes) (Lebreton et al. 1992; Williams et al. 2002). Mark-recapture methods are often challenging or infeasible, either because capturing animals is difficult or because the marks themselves negatively impact fitness. Methods that allow for the estimation of demographic parameters, such as age-structured survival and reproduction, without the need to mark (or track by any means) individual animals would greatly expand our capacity to understand the population biology of many organisms. Point count data (e.g., abundance) of key life stages (e.g., adults, chicks/pups) are often available over large spatial scales and can be collected more cheaply than mark-recapture data. This has led to a growing interest in using this kind of data to extract more detailed demographic information.

Adélie (*Pygoscelis adeliae*) and gentoo (*P. papua*) penguins breed throughout the Antarctic continent and associated sub-Antarctic islands. Despite their remoteness, Adélie and gentoo penguins are among the world's most well-studied seabirds (see Borboroglu and Boersma 2013 and references therein), and researchers have studied their population dynamics, behavior, foraging, breeding success, as well as the impacts of tourism and other human activities, fishing and climate change. Despite this intense focus over the last 40 years, our understanding of age-structured survival and reproduction is unavoidably limited to a small number of locations where long-term banding or tagging has been permitted (e.g., Lescroël et al. 2009 and Dugger et al. 2010). Mark-recapture data are extraordinarily valuable and provide the basis for much of our understanding of penguin life history, but the intense logistics required for banding, combined with concerns that flipper bands increase drag and thus decrease survival (Culik et al. 1993; Putman 1995; Dugger et al. 2010; Dann et al. 2014), have sharply limited its use.

On the other hand, time series of abundance are widely available. Penguin biologists have been recording abundance data since the earliest days of Antarctic exploration, and compilations of such data (e.g., (Croxall and Kirkwood 1979; Woehler and Croxall 1997; Lynch et al. 2013) have been critical to our current understanding of how populations have changed over time. Abundance can now be estimated indirectly from satellite imagery of penguin guano stains (Lynch et al. 2012b; LaRue et al. 2014) and directly from unmanned aerial drones (Shah et al. 2020). Such abundance data are relatively inexpensive to obtain and compilations of count data can be assembled from multiple research teams working independently across the continent (Humphries et al. 2017). Here we report on a method to infer age- or stage-structured demographic rates from time series of abundance, a method illustrated here in the context of Antarctic penguins but widely applicable to other animals with annual reproduction.

One way to model state-structured penguin abundance is to use a state-space model (SSM) (Patterson et al. 2008). SSMs represent the relationship between time-varying latent states (e.g., true abundance) and observed time-series (e.g., collected point count data). SSMs can be described using probability distributions and are parameterized by a set of model parameters representing demographic rates. In an inference task involving SSMs, the goal is to jointly infer the unknown latent states and the parameters of the

model (see Kantas et al. 2015 for a survey of approaches). Among the set of available approaches, particle Markov chain Monte Carlo (PMCMC) sampling is one that can be used for the Bayesian learning of the unknowns in an SSM (Andrieu et al. 2010), which involves inferring the joint posterior distribution of the latent states and model parameters. In this work, our focus will be on Bayesian learning for SSMs using importance sampling (IS)-based methods (Tokdar and Kass 2010; Robert and Casella 2013) for the purposes of identifying those demographic rates that can be successfully learned from time series of abundance.

Our inverse modeling efforts are similar in spirit to the efforts by Gonzales et al. (2016) to jointly estimate the size structure and abundance of a population without individual-level demographic information. Whereas Gonzales et al. (2016) considered a population structured by size (a continuous variable) and assumed prior information on the size structure of the population, here we focus on estimating demographic rates in a discrete age-structured population with no auxiliary information on the population age structure. We explore two case studies to illustrate the approach, one involving immigration and the other involving regime switching in reproductive success, in order to evaluate which demographic parameters are strongly informed by data (rather than the prior) and the sensitivity of these estimates to the amount of missing data in each time series.

2 Methods

2.1 State-space models (SSMs)

SSMs describe how a latent state $\mathbf{x}_t \in \mathbb{R}^{d_x}$ is related to an observation $\mathbf{y}_t \in \mathbb{R}^{d_y}$ at a time instant t over a fixed time horizon T. More specifically, SSMs are described by a set of probability distributions,

$$\mathbf{x}_0 \sim p(\mathbf{x}_0) \tag{1}$$

$$\mathbf{x}_t \sim p(\mathbf{x}_t|\mathbf{x}_{t-1}, \boldsymbol{\theta}), t = 1, \ldots, T, \tag{2}$$

$$\mathbf{y}_t \sim p(\mathbf{y}_t|\mathbf{x}_t, \boldsymbol{\theta}), \ t = 1, \ldots, T, \tag{3}$$

where $p(\mathbf{x}_0)$ is the distribution of the initial state $\mathbf{x}_0$, $p(\mathbf{x}_t|\mathbf{x}_{t-1}, \boldsymbol{\theta})$ is the state transition distribution describing how the latent states evolve from one time instant to the next, $p(\mathbf{y}_t|\mathbf{x}_t, \boldsymbol{\theta})$ is the observation distribution describing how observations $\mathbf{y}_t$ are distributed with respect to the true abundance $\mathbf{x}_t$, and $\boldsymbol{\theta} \in \mathbb{R}^{d_\theta}$ is a vector of model parameters that parameterize both the state transition distribution and the observation distribution. The goal is to infer the unknown states $\mathbf{x}_{0:T} = \{\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_T\}$ and the model parameters $\boldsymbol{\theta}$ of the SSM using the observations $\mathbf{y}_{1:T} = \{\mathbf{y}_1, \ldots, \mathbf{y}_T\}$ under the Bayesian paradigm. In other words, we aim to estimate $p(\mathbf{x}_{0:T}, \boldsymbol{\theta}|\mathbf{y}_{1:T})$, the posterior distribution of $\mathbf{x}_{0:T}$ and $\boldsymbol{\theta}$ given $\mathbf{y}_{1:T}$.

This joint distribution can be written as the following product:

$$p(\mathbf{x}_{0:T}, \boldsymbol{\theta}|\mathbf{y}_{1:T}) = p(\mathbf{x}_{0:T}|\mathbf{y}_{1:T}, \boldsymbol{\theta})p(\boldsymbol{\theta}|\mathbf{y}_{1:T}), \tag{4}$$

where $p(\mathbf{x}_{0:T}|\mathbf{y}_{1:T}, \boldsymbol{\theta})$ is the conditional posterior of the latent states given the model parameters, and $p(\boldsymbol{\theta}|\mathbf{y}_{1:T})$ is the marginal posterior of the model parameters.

Under the assumption that the model parameters are known, the conditional posterior distribution of the states $\mathbf{x}_{0:T}$ given $\mathbf{y}_{1:T}$ and $\boldsymbol{\theta}$ can be determined by Bayes' theorem as

$$p(\mathbf{x}_{0:T}|\mathbf{y}_{1:T}, \boldsymbol{\theta}) = \frac{p(\mathbf{x}_{0:T}, \mathbf{y}_{1:T}|\boldsymbol{\theta})}{p(\mathbf{y}_{1:T}|\boldsymbol{\theta})} \propto p(\mathbf{x}_{0:T}, \mathbf{y}_{1:T}|\boldsymbol{\theta}). \tag{5}$$

The structure of SSMs allows us to easily obtain an expression for the joint distribution $p(\mathbf{x}_{0:T}, \mathbf{y}_{1:T}|\boldsymbol{\theta})$

$$p(\mathbf{x}_{0:T}, \mathbf{y}_{1:T}|\boldsymbol{\theta}) = p(\mathbf{x}_0) \prod_{t=1}^{T} p(\mathbf{y}_t|\mathbf{x}_t, \boldsymbol{\theta}) p(\mathbf{x}_t|\mathbf{x}_{t-1}, \boldsymbol{\theta}). \tag{6}$$

Furthermore, the marginal likelihood $p(\mathbf{y}_{1:T}|\boldsymbol{\theta})$ can be obtained by integrating out the states from the joint distribution

$$p(\mathbf{y}_{1:T}|\boldsymbol{\theta}) = \int \cdots \int \left(p(\mathbf{x}_0) \prod_{t=1}^{T} p(\mathbf{y}_t|\mathbf{x}_t, \boldsymbol{\theta}) p(\mathbf{x}_t|\mathbf{x}_{t-1}, \boldsymbol{\theta}) \right) d\mathbf{x}_0 \cdots d\mathbf{x}_T. \tag{7}$$

If the interest is solely in the parameters of the SSM, then the goal is to obtain the marginal posterior distribution of $\boldsymbol{\theta}$ given $\mathbf{y}_{1:T}$. This can also be determined by applying Bayes' rule:

$$p(\boldsymbol{\theta}|\mathbf{y}_{1:T}) = \frac{p(\mathbf{y}_{1:T}|\boldsymbol{\theta}) p(\boldsymbol{\theta})}{p(\mathbf{y}_{1:T})}. \tag{8}$$

Unfortunately, a tractable solution to (7) only exists for certain classes of SSMs (e.g., linear and Gaussian SSMs Carter and Kohn 1994) and hence, the conditional posterior $p(\mathbf{x}_{0:T}|\mathbf{y}_{1:T}, \boldsymbol{\theta})$ and the marginal posterior $p(\boldsymbol{\theta}|\mathbf{y}_{1:T})$ cannot generally be analytically computed. In the following, we discuss a class of numerical integration techniques that use importance sampling (IS) (Robert and Casella 2013) to obtain a sample-based approximation to the posterior distribution of interest.

2.2 Importance sampling

Under the assumption that the likelihood can be analytically computed (e.g., linear Gaussian case), IS techniques are sufficient for approximation of the marginal posterior $\varphi(\boldsymbol{\theta}) \triangleq p(\boldsymbol{\theta}|\mathbf{y}_{1:T})$, also called the target distribution. IS is a Monte Carlo sampling method that can be used to approximate expectation values with respect to the distribution $\varphi(\boldsymbol{\theta})$, i.e.,

$$\mathbb{E}_{\varphi}[f(\boldsymbol{\theta})] = \int \varphi(\boldsymbol{\theta}) f(\boldsymbol{\theta}) d\boldsymbol{\theta} \tag{9}$$

without the need to draw samples from $\varphi(\boldsymbol{\theta})$ directly. In IS, samples $\{\boldsymbol{\theta}^{(m)}\}_{m=1}^{M}$ are drawn from a proposal distribution $q(\boldsymbol{\theta}; \boldsymbol{\lambda})$, where $\boldsymbol{\lambda}$ denotes the parameters of that proposal. Samples are then weighted according to

$$\tilde{w}^{(m)} = \frac{\tilde{\varphi}(\boldsymbol{\theta}^{(m)})}{q(\boldsymbol{\theta}^{(m)}; \boldsymbol{\lambda})}, \quad m = 1, \dots, M, \tag{10}$$

where $\tilde{\varphi}(\boldsymbol{\theta}) \triangleq p(\mathbf{y}_{1:T}|\boldsymbol{\theta}) p(\boldsymbol{\theta})$. The collection of samples and importance weights can be used to obtain a numerical approximation to integrals of the form in (9) as

$$\mathbb{E}_{\varphi}[f(\boldsymbol{\theta})] \approx \sum_{m=1}^{M} w^{(m)} f(\boldsymbol{\theta}^{(m)}), \tag{11}$$

where $w^{(m)}$ denotes the normalized weight of the mth sample $\boldsymbol{\theta}^{(m)}$ determined as

$$w^{(m)} = \frac{\tilde{w}^{(m)}}{\sum_{j=1}^{M} \tilde{w}^{(j)}}. \tag{12}$$

A nice property of the estimator in Eq. (11) is that it converges to the true value of $\mathbb{E}_{\varphi}[f(\boldsymbol{\theta})]$ as the number of samples M tends to infinity (Tokdar and Kass 2010). The variance of the estimator depends on how "close" the proposal distribution is to the target distribution, and so a poorly chosen proposal (far from the target) can lead to high variance estimators (Owen and Zhou 2000). When the dimension of $\boldsymbol{\theta}$ is large, it is difficult to choose the proposal parameters $\boldsymbol{\lambda}$ to obtain a good fit to the target. In these scenarios, it is often beneficial to use adaptive IS (AIS) methods (see Bugallo et al. (2017) for a review), which can iteratively adapt the proposal parameters to construct a better fit to the target distribution $\varphi(\boldsymbol{\theta})$. At the ith iteration of an AIS method, samples are drawn from a proposal $q(\boldsymbol{\theta}; \boldsymbol{\lambda}_i)$, i.e.,

$$\boldsymbol{\theta}_i^{(m)} \sim q(\boldsymbol{\theta}; \boldsymbol{\lambda}_i), \quad m = 1, \dots, M, \tag{13}$$

which are then weighted accordingly

$$\tilde{w}_i^{(m)} = \frac{\tilde{\varphi}(\boldsymbol{\theta}_i^{(m)})}{q(\boldsymbol{\theta}_i^{(m)}; \boldsymbol{\lambda}_i)}, \quad m = 1, \dots, M. \tag{14}$$

After the samples are drawn and weighted, the proposal parameter $\boldsymbol{\lambda}_i$ are adapted using some rule. For example, if the proposal parameters $\boldsymbol{\lambda}_i$ refer to the mean vector $\boldsymbol{\mu}_i$ and the covariance matrix $\boldsymbol{\Sigma}_i$ of the proposal, then one possible method for updating $\boldsymbol{\mu}_i$ and $\boldsymbol{\Sigma}_i$ is to use the following rule:

$$\boldsymbol{\mu}_{i+1} = \eta_{1,i} \boldsymbol{\mu}_i + (1 - \eta_{1,i}) \sum_{m=1}^{M} w_i^{(m)} \boldsymbol{\theta}_i^{(m)}, \tag{15}$$

$$\boldsymbol{\Sigma}_{i+1} = \eta_{2,i} \boldsymbol{\Sigma}_i + (1 - \eta_{2,i}) \sum_{m=1}^{M} w_i^{(m)} (\boldsymbol{\theta}_i^{(m)} - \boldsymbol{\mu}_{i+1})(\boldsymbol{\theta}_i^{(m)} - \boldsymbol{\mu}_{i+1})^{\top}, \tag{16}$$

where $0 < \eta_{1,i}, \eta_{2,i} < 1$ and $w_i^{(m)}$ denotes the normalized importance weight of the mth sample drawn in the ith iteration.

2.3 Particle adaptive importance sampling

If the marginal likelihood in Eq. (7) cannot computed analytically, then it is impossible to evaluate the importance weights in an AIS scheme, and the use of standard AIS techniques to approximate the marginal posterior $p(\boldsymbol{\theta}|\mathbf{y}_{1:T})$ becomes impractical. Alternatively, one can consider using unbiased estimates of the marginal likelihood obtained through some numerical approach. For SSMs, a straightforward way to obtain an unbiased estimate of Eq. (7) is to use particle filtering (PF) methods (Djuric et al. 2003), which are sometimes more generally referred to as sequential Monte Carlo (SMC) methods (Doucet et al. 2001). With PF methods, one can obtain an IS-based approximation to $p(\mathbf{x}_{0:T}|\mathbf{y}_{1:T}, \boldsymbol{\theta})$ in a recursive manner. Given this approximation of $p(\mathbf{x}_{0:T}|\mathbf{y}_{1:T}, \boldsymbol{\theta})$, one can use the collected importance weights to obtain an unbiased estimator of the marginal likelihood as (Doucet and Johansen 2009):

$$p(\mathbf{y}_{1:T}|\boldsymbol{\theta}) \approx \hat{Z} = \prod_{t=1}^{T} \left(\frac{1}{N} \sum_{n=1}^{N} \tilde{v}_t^{(n)} \right), \tag{17}$$

where $\tilde{v}_t^{(n)}$ is the unnormalized importance weight of the nth sampled particle stream and N is total number of samples in the PF algorithm. Furthermore, one can obtain a sample of the trajectory $\mathbf{x}_{0:T}$ from the conditional posterior distribution $p(\mathbf{x}_{0:T}|\mathbf{y}_{1:T}, \boldsymbol{\theta})$ using PF, i.e.,

$$\mathbf{x}_{0:T,i}^{(m)} \sim \hat{p}(\mathbf{x}_{0:T}|\mathbf{y}_{1:T}, \boldsymbol{\theta}_i^{(m)}), \tag{18}$$

Additional details on the PF method, how the marginal likelihood estimator is derived, and how latent state trajectories are sampled are provided in the Supplementary Materials.

Algorithm 1 shows a general implementation of the particle AIS (PAIS) method used in this work to estimate the joint posterior distribution of the unknown states $\mathbf{x}_{0:T}$ and the model parameters $\boldsymbol{\theta}$. The main differences between Algorithm 1 and a standard AIS algorithm are in the sampling and weighting steps. In the sampling step, M model parameters $\{\boldsymbol{\theta}_i^{(m)}\}_{m=1}^{M}$ are drawn from a proposal distribution $q(\boldsymbol{\theta}; \boldsymbol{\lambda}_i)$. Then, for each of the sampled model parameters $\boldsymbol{\theta}_i^{(m)}$, a PF scheme is employed to sample a corresponding state trajectory $\mathbf{x}_{0:T,i}^{(m)}$ and obtain an unbiased estimate of the marginal likelihood $Z_i^{(m)}$. In the weighting step, instead of evaluating $\tilde{\varphi}(\boldsymbol{\theta}_i^{(m)})$, one uses an unbiased approximation of $\tilde{\varphi}(\boldsymbol{\theta}_i^{(m)})$ given by $\gamma_i^{(m)} = \hat{Z}_i^{(m)} p(\boldsymbol{\theta}_i^{(m)})$, where $\hat{Z}_i^{(m)}$ is the unbiased marginal likelihood estimate obtained from running a PF scheme

---

Algorithm 1 Particle Adaptive Importance Sampling (PAIS)

1: **Initialization:** Set the initial proposal parameters $\boldsymbol{\lambda}_1$.
2: **for** $i = 1, \ldots, I$ **do**
3:    Draw M samples from the proposal distribution,

$$\boldsymbol{\theta}_i^{(m)} \sim q(\boldsymbol{\theta}; \boldsymbol{\lambda}_i), \quad m = 1, \ldots, M.$$

4:    **for** $m = 1, \ldots, M$ **do in parallel**
5:       Sample $\mathbf{x}_{0:T,i}^{(m)}$ conditioned on $\boldsymbol{\theta}_i^{(m)}$ using a PF method,

$$\mathbf{x}_{0:T,i}^{(m)} \sim \hat{p}(\mathbf{x}_{0:T}|\boldsymbol{\theta}_i^{(m)}, \mathbf{y}_{1:T}).$$

6:       Set $\gamma_i^{(m)} = \hat{Z}_i^{(m)} p(\boldsymbol{\theta}_i^{(m)})$.
7:    **end for**
8:    Compute the approximated importance weights as

$$\tilde{w}_i^{(m)} = \frac{\gamma_i^{(m)}}{q(\boldsymbol{\theta}_i^{(m)}; \boldsymbol{\lambda}_i)}.$$

9:    Adapt the proposal parameters, $\boldsymbol{\lambda}_{i+1} \leftarrow \boldsymbol{\lambda}_i$ using some update rule.
10: **end for**

---

conditioned on $\boldsymbol{\theta}_i^{(m)}$. Indeed, since $\mathbb{E}[\hat{Z}_i^{(m)}] = p(\mathbf{y}_{1:T}|\boldsymbol{\theta}_i^{(m)})$, then

$$\begin{aligned}
\mathbb{E}[\gamma_i^{(m)}] &= \mathbb{E}[\hat{Z}_i^{(m)} p(\boldsymbol{\theta}_i^{(m)})] \\
&= \mathbb{E}[\hat{Z}_i^{(m)}] p(\boldsymbol{\theta}_i^{(m)}) \\
&= p(\mathbf{y}_{1:T}|\boldsymbol{\theta}_i^{(m)}) p(\boldsymbol{\theta}_i^{(m)}) \\
&= \tilde{\varphi}(\boldsymbol{\theta}_i^{(m)}).
\end{aligned}$$

A flow diagram summarizing the basic steps of a PAIS method is shown in Fig. 1.

2.3.1 Computational complexity

Here, we briefly discuss the computational complexity of Algorithm 1. At each iteration of the algorithm, M samples are drawn from the proposal distribution. For each of these samples, a particle filter is run to compute the marginal likelihood, where for simplicity we assume N particles are used in each particle filter. For each run of the particle filters, N particles are drawn at each time step for $(T+1)$ total time steps. This sampling process is repeated for I iterations of the algorithm. In summary, the number of samples drawn in the algorithm is $M_{\text{total}} = MN(T+1)I$. Given the per-sample complexity of the AIS method and the PF method (which varies by method), we can get a sense of the total computational complexity of the overall PAIS algorithm. A more complete discussion of the computational complexity of AIS and PF methods may be found in Bugallo et al. (2017) and Doucet and Johansen (2009), respectively.

Depending on the chosen AIS and PF schemes, PAIS can become too computationally expensive to run in practice. To address this, we exploit the embarrassingly
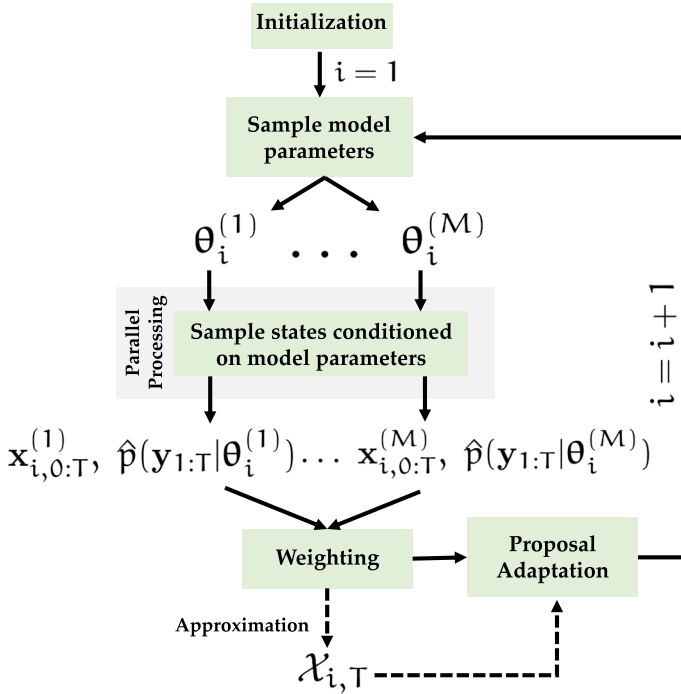
Fig. 1 Schematic of particle adaptive importance sampling

parallel nature of the PF step of the algorithm. At each PAIS iteration, since the M runs of the PF scheme are independent of one another, one can easily parallelize this process across computing resources. To that end, we utilize a computational cluster to efficiently sample the states, which allows the algorithm to become practically feasible.

2.4 Population data

We apply our model to data on gentoo and Adélie penguin population dynamics (Humphries et al. 2017; original data plotted in Supplementary Materials Figures 1 and 2). These time series include counts on the number of nests (equivalent to the number of breeding pairs) and chicks at nine breeding colonies spread across Antarctica. The available time series are highly patchy, and there are relatively few instances in which both a nest count and a chick count are available in the same year. In many years, there is no information on either nests or chicks. While observation error for each count is often estimated by the researchers collecting the data, we found that using this information made model convergence more difficult (see Discussion) and so we assumed a 20% observation error for all nests and chick counts.
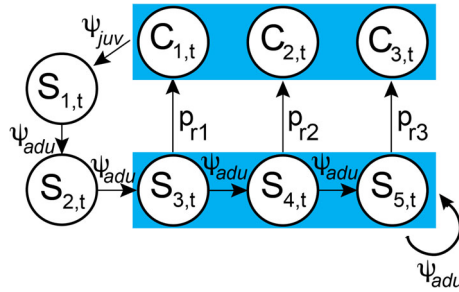
Fig. 2 Basic life cycle diagram with $J = 5$ stages, excluding any immigration. There are two non-breeding stages before penguins mature into the S_3 stage, and penguins are assumed to follow a juvenile survival rate (ψ_{juv}) in the first year and an adult survival rate (ψ_{adu}) in all following years. Breeding success p_r is allowed to vary by stage. Counts of abundance are agnostic to the age of breeders and therefore counts of nests represent the sum of S_3 through S_5 (lower blue box) and counts of chicks represent the sum of C_3 through C_5 (upper blue box)

2.4.1 Gentoo population dynamics with immigration

At each breeding population in year t, the observations $\mathbf{y}_t$ reflect the total number of adult breeding penguins and chicks, denoted $\tilde{S}_{b,t}$ and $\tilde{C}_t$, respectively. To illustrate the estimation of demographic rates in a realistic scenario, we fit the model described in Fig. 2 with the addition of immigration to the S_3 stage. This model tracks female breeders only and therefore we divide the total number of chicks in year $t - 1$ by 2 to calculate the number of S_1 females in year t (Eq. 19). We assume that immigration in year t is a function of the total number of breeding females in year $t - 1$ (Eq. 21) and that immigrants arrive at the S_3 stage and are therefore added to the number of individuals aging into the S_3 stage from the local population (Eq. 22). We assume that reproductive success is related to penguin age (Eq. 25), consistent with observation, and that each breeding female lays two eggs and therefore has an upper bound of two chicks (Eq. 26). We assume that the year-specific observation error for both nests and chicks is Normally distributed with the standard deviation being some fraction of the counted abundance (Eq. 27, 28). While, in principle, the observation error could be fit as an additional free parameter, here we constrain observation error to $\pm 20\%$. The model is described by the following equations

$$S_{1,t} \sim \text{Binomial}\left(\frac{C_{t-1}}{2}, \psi_{juv}\right), \tag{19}$$

$$S_{2,t} \sim \text{Binomial}\left(S_{1,t-1}, \psi_{adu}\right), \tag{20}$$

$$S_{\text{im},t} \sim \text{Poisson}(\alpha_{im} + \beta_{im} \times S_{t-1}), \tag{21}$$

$$S_{3,t} = \bar{S}_{3,t} + S_{im,t}, \quad \bar{S}_{3,t} \sim \text{Binomial}\left(S_{2,t-1}, \psi_{adu}\right), \tag{22}$$

$$S_{j,t} \sim \text{Binomial}\left(S_{j-1,t-1}, \psi_{adu}\right), \quad j = 4, \ldots, J - 1 \tag{23}$$

$$S_{J,t} \sim \text{Binomial}\left(S_{J-1,t-1} + S_{J,t-1}, \psi_{adu}\right), \tag{24}$$

$$p_{r,j} = \text{InvLogit}(\alpha_{rs} + \beta_{rs} \times j), \quad j = 1, \ldots, J - 2, \tag{25}$$

$$C_{j,t} \sim \text{Binomial}\left(2S_{j+2,t}, p_{r,j}\right), \quad j = 1, \ldots, J - 2, \tag{26}$$

$$\tilde{S}_t \sim \mathcal{N}\left(S_t, (\sigma_{s,t} S_t)^2\right), \quad S_t = \sum_{j=3}^{J} S_{j,t}, \tag{27}$$

$$\tilde{C}_t \sim \mathcal{N}\left(C_t, (\sigma_{c,t} C_t)^2\right), \quad C_t = \sum_{j=1}^{J-2} C_{j,t}, \tag{28}$$

where the latent states are $\mathbf{x}_t = [S_{im,t}, S_{1,t}, \ldots, S_{J,t}, C_{1,t}, \ldots, C_{J-2,t}]^\mathsf{T}$. $S_{j,t}$ denotes the number of stage j female penguins, $C_{j,t}$ the number of chicks from stage $j+2$ female penguins, $S_{im,t}$ denotes the number of females immigrating into the third stage, and J denotes the total number of non-chick age classes (here $J = 5$). The unknown model parameters are $\boldsymbol{\theta} = [\psi_{juv}, \psi_{adu}, \alpha_{im}, \beta_{im}, \alpha_{rs}, \beta_{rs}]^\mathsf{T}$, where ψ_{juv} denotes the juvenile survivorship, ψ_{adu} denotes the adult survivorship, α_{im} and β_{im} denote the intercept and slope of a linear model for the number of immigrants, respectively, and α_{rs} and β_{rs} denote the slope and intercept of a logistic model for the stage-structured reproductive success, respectively. We note that the stage-structured reproductive rates $p_{r,j}$ can be extracted from the model using the inverse logit (i.e., sigmoid) transformation.

2.4.2 Adélie population dynamics with regime switching

As a second illustration of these methods, we expand our basic (no immigration) model (Fig. 2) to permit switching between two regimes. This represents one manifestation of regime-switching state-space models (RS-SSM) that have been explored in other contexts (Ghahramani and Hinton 1996; Kim and Nelson 1999). RS-SSMs augment a discrete-valued latent state (called a regime) to an SSM and allows for parameters to switch from one time-instant to another depending on the regime state. In the Adélie model, the intercept of the reproductive rate (α_{rs}) switches between two values, depending on the state of the regime in year t (r_t) as follows:

$$r_t \sim \text{Bernoulli}(1 - \gamma), \quad r_t \in \{0, 1\}, \tag{29}$$

$$p_{rs,j,t} = \text{InvLogit}(\alpha_{rs,r_t} + \beta_{rs} \times j), \quad j = 1, \ldots, J-2 \tag{30}$$

where the parameter $\gamma = \mathbb{P}(r_t = 0)$ denotes the probability of being in the $r_t = 0$ regime governed by intercept $\alpha_{rs,0}$.

2.4.3 Model priors and prior-posterior overlap

Prior distributions for model parameters are summarized in Table 1. For switching reproductive success parameters in the Adélie model ($\alpha_{rs,0}$ and $\alpha_{rs,1}$), we define the prior of the smaller reproductive rate $\alpha_{rs,0}$ and the difference between the larger one and the smaller one as $\phi = \alpha_{rs,1} - \alpha_{rs,0}$. This parameterization, along with the choice of prior on ϕ restricting it to positive values, adds a monotonicity constraint in the reproductive success parameters and ensures that the parameters are identifiable.

Table 1 Prior distributions for each unknown parameter in the gentoo and Adélie models

	Gentoo	Adélie
ψ_{juv}	Beta(43, 57)	Beta(3, 3)
ψ_{adu}	Beta(82, 18)	Beta(3, 3)
α_{im}	Uniform(0, 50)	–
β_{im}	Uniform(−0.5, 0.5)	–
$\alpha_{rs}/\alpha_{rs,0}$	Normal(0, 10)	Normal(0, 10)
ϕ	–	Uniform(0, 3)
β_{rs}	Normal(0.5, 0.04)	Normal(0.5, 0.04)
γ	–	Beta(1, 9)

Our prior distributions were only weakly informative and based on the species' life history (e.g., known clutch size). The only prior distributions that are more strongly constrained are for juvenile and adult survival of the gentoo penguin and the rate of increase for reproductive success with age (β_{rs}), and these priors were selected consistent with published literature

To assess how much the parameter posteriors are informed by the data (as opposed to simply reflecting the prior), we use a metric called the prior-posterior overlap (PPO), which reflects the percentage of overlap between the prior and posterior distributions for each parameter (Gimenez et al. 2009). If the PPO is large (close to 1), then the posterior is nearly identical to the prior and the parameter is considered uninformed by the data (either because the parameter is structurally unidentifiable or because the data are insufficient). If the PPO is small, then the posterior is very different from the prior and we consider the parameter 'learnable'.

3 Results

For the gentoo time-series, the PPO is generally larger for the survivorship parameters and the slope of reproductive success (Table 2). Indeed, strong priors based on Ainley (2002) and Hinke (2012) were assumed for both juvenile and adult survivorship, and the PPO was larger than 0.90 for each dataset run (including the synthetic data). In contrast, the PPOs for the immigration-related parameters (α_{im} and β_{im}) were quite small, implying that they can be learned even with such short and patchy time-series. The intercept of reproductive success was also a learnable parameter in the model and, unsurprisingly, its estimation is sensitive to the amount of chick data available. For example, there was only a single chick count at Vernadsky Station and since this chick count was quite low (relative to the number of nests), our estimate of average reproductive success was also low. While this could reflect increased risk of egg and chick predation at smaller colonies, more chick data would allow us to estimate reproductive success with more confidence.

Unlike the gentoo model, the PPO for the survivorship parameters in the Adélie model are quite small (Table 3). Because this model does not include immigration, fluctuations in the nest time-series are more informative of survivorship. The intercept of reproductive success (for regime $r_t = 0$) is also a learnable parameter, as the

Table 2 Posterior means, credible intervals, and prior-posterior overlaps for the gentoo datasets. For the synthetic dataset, the true values of the parameters were $\psi_{juv} = 0.43$, $\psi_{adu} = 0.82$, $\alpha_{rs} = 0.5$, $\beta_{rs} = 0.5$, $\alpha_{im} = 20$, and $\beta_{im} = 0.02$

	Synthetic	Vernadsky Station	Orne Is.	Moot Pt.	Biscoe Pt.
ψ_{juv}	0.42 (0.33, 0.52) [0.96]	0.43 (0.33, 0.53) [0.98]	0.44 (0.34, 0.53) [0.94]	0.42 (0.33, 0.52) [0.91]	0.43 (0.33, 0.52) [0.96]
ψ_{adu}	0.82 (0.74, 0.88) [0.97]	0.81 (0.73, 0.88) [0.93]	0.82 (0.74, 0.89) [0.97]	0.82 (0.74, 0.89) [0.96]	0.81 (0.73, 0.88) [0.90]
α_{rs}	0.48 (− 0.08, 1.08) [0.07]	− 2.89 (− 3.73, − 1.93) [0.09]	− 0.54 (-1.20, 0.18) [0.08]	− 0.24 (− 0.94, 0.57) [0.09]	1.97 (− 0.55, 5.77) [0.28]
β_{rs}	0.56 (0.20, 0.94) [0.89]	0.50 (0.15, 0.88) [0.96]	0.52 (0.15, 0.90) [0.97]	0.48 (0.16, 0.87) [0.92]	0.51 (0.16, 0.89) [0.97]
α_{im}	22.48 (15.73, 29.35) [0.32]	2.29 (0.07, 9.65) [0.20]	1.48 (0.11, 3.83) [0.10]	46.06 (38.41, 49.87) [0.26]	13.95 (7.14, 21.40) [0.33]
β_{im}	0.02 (− 0.02, 0.06) [0.00]	0.26 (0.13, 0.37) [0.01]	0.12 (0.07, 0.17) [0.00]	0.04 (− 0.01, 0.10) [0.00]	0.10 (0.05, 0.17) [0.00]
# of Years	20	16	21	14	19
% Missing (nests, chicks)	(0.45, 0.70)	(0.69, 0.94)	(0.48, 0.76)	(0.43, 0.71)	(0.11, 1.00)

Table 3 Posterior means, credible intervals, and prior-posterior overlaps for the Adélie datasets. For the synthetic dataset, the true values of the parameters were $\psi_{juv} = 0.45$, $\psi_{adu} = 0.85$, $\alpha_{rs,0} = -0.5$, $\phi = 1.5$, $\beta_{rs} = 0.5$, and $\gamma = 0.2$

	Synthetic	Port Charcot	Petermann Is.	Litchfield Is.	Cormorant Is.	Bechervaise Is.
ψ_{juv}	0.52 (0.36, 0.71) [0.64]	0.54 (0.17, 0.87) [0.92]	0.25 (0.08, 0.50) [0.43]	0.48 (0.21, 0.76) [0.85]	0.21 (0.06, 0.42) [0.31]	0.62 (0.34, 0.89) [0.71]
ψ_{adu}	0.82 (0.76, 0.87) [0.09]	0.75 (0.67, 0.85) [0.21]	0.83 (0.75, 0.90) [0.12]	0.77 (0.71, 0.85) [0.16]	0.87 (0.80, 0.93) [0.07]	0.85 (0.80, 0.91) [0.07]
$\alpha_{rs,0}$	− 0.19 (− 0.88, 0.52) [0.22]	− 1.16 (− 2.17, − 0.17) [0.28]	− 0.89 (− 2.36, 0.16) [0.34]	− 3.17 (− 3.85, − 2.66) [0.12]	− 1.85 (− 2.78, − 0.89) [0.24]	− 2.10 (− 2.89, − 1.35) [0.20]
ϕ	2.31 (1.56, 2.93) [0.50]	0.64 (0.06, 1.58) [0.52]	0.80 (0.08, 2.00) [0.64]	2.92 (2.72, 3.00) [0.12]	2.02 (1.31, 2.82) [0.52]	1.15 (0.59, 1.68) [0.39]
β_{rs}	0.50 (0.12, 0.87) [0.96]	0.52 (0.20, 0.89) [0.96]	0.52 (0.20, 0.89) [0.94]	0.40 (0.14, 0.79) [0.75]	0.49 (0.15, 0.86) [0.94]	0.51 (0.18, 0.89) [0.96]
γ	0.09 (0.02, 0.23) [0.74]	0.09 (0.00, 0.32) [0.99]	0.08 (0.00, 0.30) [0.91]	0.14 (0.05, 0.27) [0.58]	0.11 (0.02, 0.27)[0.76]	0.16 (0.02, 0.37) [0.65]
# of Years	30	36	36	30	26	12
% Missing (nests, chicks)	(0.50, 0.27)	(0.64, 0.78)	(0.44, 0.58)	(0.10, 0.43)	(0.73, 0.35)	(0.00, 0.00)

PPO for that parameter was small (< 0.3) for all sites. We also can extract some information about the difference between the reproductive success of the two regimes (i.e., $\phi = \alpha_{rs,1} - \alpha_{rs,0}$), as we obtain moderate values for PPO for that site. Litchfield Island's small PPO may reflect the comprehensiveness of the time series available. Similar to the gentoo model, the slope of reproductive success had high PPO (> 0.9) for all sites, meaning that inference on that parameter heavily relies on prior information. Regarding the regime-switching probability γ, the PPO is quite large for that parameter, especially when there is a lot of missing chick data (e.g., Port Charcot, which has 78% missing chick data). In contrast, sites with a lot of chick data, such as Litchfield Island, Cormorant Island, and Bechervaise Island, which have a moderate amounts of chick data, have a smaller PPO for γ.

We provide additional results for both the gentoo and Adélie models in the Supplementary Material, including results regarding the predictive performance of each model, the estimated average reproductive success for each site, and posterior histograms for a subset of the parameters. We also provide results on the stochasitic sensitivity of the PAIS algorithm over multiple realizations.

4 Discussion

Despite the challenge of estimating demographic parameters from point counts of total abundance, especially in the face of missing data, several demographic parameters were informed by the data. The immigration component of the gentoo penguin time series was estimable from the data but the addition of immigration made it difficult to estimate either juvenile or adult survival. In contrast, with no assumption of immigration, the model for Adélie penguins did permit an estimation of both juvenile and adult survival even with the inclusion of two reproductive regimes. As would be expected for a long-lived seabird like the Adélie, our estimates suggest greater variation among sites for juvenile survival than for adult survival. Bechervaise Island had the highest estimated juvenile survival among the sites included in the study. In this context, its worth noting that Bechervaise Island specifically (and the surrounding region of Eastern Antarctica more generally) has an increasing Adélie population in contrast to the other sites, all located on the Antarctic Peninsula, where persistent population declines are well documented (Lynch et al. 2012a).

Regarding the sampler, we encountered several challenges. In particular, the convergence of the sampling algorithm was sensitive to the initial values of the latent states (i.e., the penguin abundances). This was not an issue for modeling gentoo penguins, since those time series involved increasing populations starting from a very small numbers of penguins. For the Adélie time series, initial values needed to be well-calibrated for the algorithm to converge. This is especially difficult for sites with large populations, as there was a correspondingly large range of possible initial values. In order to overcome this issue, we opted to calibrate the initial values using the data in the first year. In particular, under the assumption that the relative distribution of abundance (across stages) is known *a priori*, one can use the observations of total abundance in the first year (i.e., $\tilde{S}_1$ and $\tilde{C}_1$) to select appropriate priors for the initial latent states. To clarify this procedure, let us consider an example. If it is known that

the ratio between stage 4 penguins and all breeding penguins is 1:10, then one can approximate an initial value as $\hat{S}_{4,0} = 0.1\tilde{S}_1$ and add some noise to it before feeding it to the PF algorithm. In our implementation, we added noise that perturbed the approximated initial value up to 10% using a uniform distribution, i.e., each particle for stage 4 penguins would be initialized as,

$$S_{4,0}^{(n)} \sim \text{Uniform}(0.9\hat{S}_{4,0}, 1.1\hat{S}_{4,0}), \quad n = 1, \ldots, N. \tag{31}$$

Therefore, for initial value calibration, one needs to rely on domain expert knowledge of stage distribution, which can be assumed prior to running the analysis. Model convergence required a larger uncertainty in the observations (20% error in both the nest and chick counts) than was recorded alongside the initial abundance counts, because more precise abundance estimates led to an insufficient number of particles with non-trivial weight. Model convergence was assessed using the Pareto smoothed importance sampling (PSIS) diagnostic (Vehtari et al. 2021). The PSIS diagnostic provides a means to assess the degeneracy in the importance sampling weights by fitting them to the shape parameter of a generalized Pareto distribution. Theoretical findings in Vehtari et al. (2021) suggest that if the PSIS diagnostic is less than 0.7, then the importance sampling algorithm produces reliable estimates. In the case of our experiments, we were only able to achieve PSIS diagnostics less than 0.7 if we assumed the observation error was at least 20%, which is larger (in almost all cases) than the nominal observation error.

The time series used for this analysis were relatively short and the missing data, while an unavoidable feature of many Antarctic time series, had a major impact on our ability to estimate our model's parameters. Additional data collected over time will generate longer time series from which even more precise estimates of demographic parameters can be extracted, and efforts to estimate abundance using archived satellite imagery may reduce the amount of missing data in the existing time series. One of the advantages of using this approach is that it permits the inclusion of environmental conditions (e.g., sea ice concentration) as covariates on specific demographic transitions, which will allow for a more direct and biologically-interpretable approach than linking environmental conditions to changes in population growth rates. Moreover, this approach provides a natural link to Integrated Population Models and the explicit inclusion of auxiliary data on demographic parameters such as reproductive success (Besbeas et al. 2002).

While aggregated point count data have been used to infer regional patterns of population change (e.g., Che-Castaldo et al. 2017), such data have not been used to estimate demographic rates such as survivorship and reproduction. The use of advanced computational methods allow us to learn something about the underlying demographic rates using simple point counts of breeding animals and, in doing so, greatly expand our capacity for linking environmental conditions to population dynamics in a way that is both highly cost-effective and scalable.

Data Availability All penguin time series data are available at www.penguinmap.com.

Code availability All code is available at https://github.com/yellaham/blp.

Declarations

Conflicts of interest/Competing interests Not applicable.

References

Ainley D (2002) The Adélie penguin: Bellwether of change. Columbia University Press, USA

Andrieu C, Doucet A, Holenstein R (2010) Particle Markov chain Monte Carlo methods. J R Stat Soc 72(3):269–342

Besbeas P, Freeman S, Morgan B, Catchpole E (2002) Integrating mark-recapture-recovery and census data to estimate animal abundance and demographic parameters. Biometrics 58:540–547

Borboroglu P, Boersma P (2013) Penguins: Natural History and Conservation. University of Washington Press, Seattle

Bugallo MF, Elvira V, Martino L, Luengo D, Miguez J, Djuric PM (2017) Adaptive importance sampling: the past, the present, and the future. IEEE Sig Process Mag 34(4):60–79

Carter CK, Kohn R (1994) On Gibbs sampling for state space models. Biometrika 81(3):541–553

Che-Castaldo C, Jenouvrier S, Youngflesh C, Shoemaker KT, Humphries G, McDowall P, Landrum L, Holland MM, Li Y, Ji R, Lynch HJ (2017) Pan-Antarctic analysis aggregating spatial estimates of Adélie penguin abundance reveals robust dynamics despite stochastic noise. Nat Commun 8(832)

Croxall J, Kirkwood E (1979) The distribution of penguins on the Antarctic Peninsula and islands of the Scotia Sea. British Antarctic Survey

Culik B, Wilson R, Bannasch R (1993) Flipper-bands on penguins: What is the cost of a life-long commitment? Mar Ecol Prog Ser 98:209–214

Dann P, Sidhu LA, Jessop R, Renwick L, Healy M, Dettmann B, Baker B, Catchpole EA (2014) Effects of flipper bands and injected transponders on the survival of adult little penguins Eudyptula Minor. Ibis 156(1):73–83

Djuric PM, Kotecha JH, Zhang J, Huang Y, Ghirmai T, Bugallo MF, Miguez J (2003) Particle filtering. IEEE Sig Process Mag 20(5):19–38

Doucet A, Johansen A (2009) A tutorial on particle filtering and smoothing: fifteen years later. Handbook of Nonlinear Filtering 12

Doucet A, De Freitas N, Gordon N (2001) An introduction to sequential Monte Carlo methods. In: Sequential Monte Carlo Methods in Practice, Springer, pp 3–14

Dugger KM, Ainley DG, Lyver PO, Barton K, Ballard G (2010) Survival differences and the effect of environmental instability on breeding dispersal in an Adélie penguin meta-population. Proc Nat Acad Sci 107(27):12375–12380

Ghahramani Z, Hinton GE (1996) Switching state-space models. Tech. rep., King's College Road, Toronto M5S 3H5

Gimenez O, Morgan BJ, Brooks SP (2009) Weak identifiability in models for mark-recapture-recovery data. In: Modeling Demographic Processes in Marked Populations, Springer, pp 1055–1067

Gonzales E, Martorell C, Bolker B (2016) Inverse estimation of integral projection model parameters using time series of population-level data. Meth Ecol Evol 7:147–156

Hinke J (2012) Over-winter behavior and annual survival of Pygoscelid penguins in the South Shetland Islands. PhD thesis, University of California, San Diego

Humphries G, Naveen R, Schwaller M, Che-Castaldo C, McDowall P, Schrimpf M, Lynch H (2017) Mapping Application for Penguin Populations and Projected Dynamics (MAPPPD): data and tools for dynamic management and decision support. Polar Rec 53(2):160–166

Kantas N, Doucet A, Singh SS, Maciejowski J, Chopin N (2015) On particle methods for parameter estimation in state-space models. Stat Sci 30(3):328–351

Kim C, Nelson CR (1999) State-space models with regime switching: classical and gibbs-sampling approaches with applications. MIT Press Books, USA

LaRue MA, Lynch HJ, Lyver P, Barton K, Ainley DG, Pollard AM, Ballard G (2014) Establishing a method to estimate Adélie penguin populations using remotely-sensed imagery. Polar Biol 37:507–517

Lebreton JD, Burnham K, Clobert J, Anderson D (1992) Modeling survival and testing biological hypotheses using marked animals: a unified approach with case studies. Ecol Monograph 62(1):67–118

Lescroël A, Dugger K, Ballard G, Ainley D (2009) Effects of individual quality, reproductive success and environmental variability on survival of a long-lived seabird. J Animal Ecol 78:798–806

Lynch H, Naveen R, Casanovas P (2013) Antarctic Site Inventory breeding bird survey data: 1994–2013. Ecology 94(11):2653–2653

Lynch H, Naveen R, Trathan P, Fagan W (2012a) Spatially integrated assessment reveals widespread changes in penguin populations on the antarctic peninsula. Ecology 93(6):1367–1377

Lynch HJ, White R, Black AD, Naveen R (2012b) Detection, differentiation, and abundance estimation of penguin species by high-resolution satellite imagery. Polar Biol 35:963–968

Owen A, Zhou Y (2000) Safe and effective importance sampling. J Am Statist Associat 95(449):135–143

Patterson TA, Thomas L, Wilcox C, Ovaskainen O, Matthiopoulos J (2008) State-space models of individual animal movement. Trends in Ecol Evol 23(2):87–94

Putman R (1995) Ethical considerations and animal welfare in ecological field studies. Biodiv Conservation 4(8):903–915

Robert C, Casella G (2013) Monte Carlo statistical methods. Springer Science & Business Media, Berlin

Shah K, Ballard G, Schmidt A, Schwager M (2020) Multidrone aerial surveys of penguin colonies in Antarctica. Sci Robot 5(47):eabc3000

Tokdar ST, Kass RE (2010) Importance sampling: a review. Wiley Interdisciplinary Reviews: Computational Statistics 2(1):54–60

Vehtari A, Simpson D, Gelman A, Yao Y, Gabry J (2021) Pareto smoothed importance sampling

Williams BK, Nichols JD, Conroy MJ (2002) Analysis and management of animal populations. Academic Press, Cambridge

Woehler E, Croxall JP (1997) The status and trends of antarctic and sub-antarctic seabirds. Mar Ornithol 25:43–66

Yousef El-Laham received his Ph.D. in Electrical and Computer Engineering from Stony Brook University and is currently an AI Research Scientist at J.P. Morgan Chase & Company. His interests are in deep learning, computational statistics, Bayesian filtering, machine learning, stochastic optimization, and variational inference.

Mónica F. Bugallo received her Ph. D. in computer science and engineering from the University of A Coruña, Spain. She is a Professor of Electrical and Computer Engineering and the Vice Provost for Faculty Affairs, and Diversity, Equity and Inclusion at Stony Brook University, NY, USA. Bugallo is the current Chair of the IEEE SPS Signal Processing Theory and Methods Technical Committee, Senior Associate Editor of the IEEE Signal Processing Letters, and Associate Editor of the IEEE Transactions on Signal Processing. Her research is focused on statistical signal processing, with emphasis on the theory of Monte Carlo methods and its application to different disciplines, including biomedicine, ecology, sensor networks, and finance.

Heather J. Lynch received her M.A. in Physics and Ph. D. in Organismic and Evolutionary Biology from Harvard University. She is the Institute for Advanced Computational Science Endowed Chair for Ecology & Evolution at Stony Brook University, NY, USA. Her research is focused on the distribution and abundance of Antarctic wildlife, the use of satellite imagery for wildlife surveys, and methods relating to time-series modeling and improved ecological forecasts.