Received: 7 June 2022 | Revised: 1 August 2022 | Accepted: 1 August 2022

DOI: 10.1111/ele.14095

ECOLOGY LETTERS cnrs WILEY

METHOD

Uncertainty quantification for ecological models with random parameters

Jody R. Reimer[1,2] | **Frederick R. Adler**[1,2] | **Kenneth M. Golden**[1] | **Akil Narayan**[1,3]

[1]Department of Mathematics, University of Utah, Salt Lake City, Utah, USA

[2]School of Biological Sciences, University of Utah, Salt Lake City, Utah, USA

[3]Scientific Computing and Imaging Institute, University of Utah, Salt Lake City, Utah, USA

Correspondences
Jody R. Reimer, Department of Mathematics and School of Biological Sciences, University of Utah, Salt Lake City, Utah, USA.
Email: reimer@math.utah.edu

Funding information
Division of Mathematical Sciences; Office of Naval Research, Grant/Award Number: N00014-21-1-2909; National Science Foundation, Grant/Award Number: DMS-1715680, DMS-1848508 and DMS-2136198; National Institutes of Health; National Institute of Biomedical Imaging and Bioengineering, Grant/Award Number: U24EB029012

Editor: Robin Snyder

Abstract

There is often considerable uncertainty in parameters in ecological models. This uncertainty can be incorporated into models by treating parameters as random variables with distributions, rather than fixed quantities. Recent advances in uncertainty quantification methods, such as polynomial chaos approaches, allow for the analysis of models with random parameters. We introduce these methods with a motivating case study of sea ice algal blooms in heterogeneous environments. We compare Monte Carlo methods with polynomial chaos techniques to help understand the dynamics of an algal bloom model with random parameters. Modelling key parameters in the algal bloom model as random variables changes the timing, intensity and overall productivity of the modelled bloom. The computational efficiency of polynomial chaos methods provides a promising avenue for the broader inclusion of parametric uncertainty in ecological models, leading to improved model predictions and synthesis between models and data.

KEYWORDS

aleatory uncertainty, bloom dynamics, epistemic uncertainty, global sensitivity, Jensen's inequality, polynomial chaos, random parameters, sea ice algae, uncertainty quantification

INTRODUCTION

Heterogeneity is ubiquitous in ecology. Locations within a landscape can vary dramatically. Individuals within a population may differ. Ecological modellers have long grappled with how to account for the heterogeneous nature of ecological systems and the uncertainty that this induces in model parameters (Regan et al., 2002; Wu & Li, 2006).

Uncertainty about parameters also arises from our inability to accurately measure the natural world, known as epistemic uncertainty. Parameters that are directly measurable may be uncertain due to small sample sizes or measurement errors. Parameters that must be estimated indirectly, such as through model fitting using Markov Chain Monte Carlo methods, are also uncertain, with this uncertainty reflected in the posterior distribution of the estimated parameter(s).

One way of incorporating parametric uncertainty—whether caused by heterogeneity, epistemic uncertainty or both— into models is to treat parameters as random variables rather than fixed quantities. Treating a parameter as a random variable, described by a probability distribution, causes the state variable to also be a random variable, with its own probability distribution. For example, if the growth rate in a population model is treated as a random variable R (random variables are typically denoted with capital letters) rather than a fixed parameter r, then the population size will also be a random variable X.

Interpretation of the state variable distribution depends on the biological motivation for including random parameters. If heterogeneity is the motivation, then the distribution of the state variable describes its predicted heterogeneity across the landscape or population. Statistics from this distribution (e.g., mean, variance, skewness) are biologically meaningful. For example, the mean of the state variable distribution describes the landscape (or population) level prediction, averaged over all locations (or individuals). These statistics can also be compared directly to observations.

If epistemic uncertainty motivates including random parameters, the resulting distribution of the state variable describes the likelihood that a given outcome is realized. In this case, statistics of the distribution cannot be directly compared against observations, because only a single outcome will be realized. Rather, the statistics of the output distribution help us understand how uncertainty in model inputs translates to uncertainty in model outputs.

Model dynamics often differ substantially between a model with fixed versus random parameters. Classic sensitivity analyses, where parameters are shifted slightly but still treated as fixed quantities, cannot provide insight into expected model behaviour or other moments. Failure to fully consider parametric uncertainty may result in biased estimates of the timing and magnitude of ecological phenomena and insufficient consideration of resulting uncertainty in model outputs, with ramifications for conservation (Harper et al., 2011; McCarthy et al., 2011) and fundamental ecological understanding (Cenci & Saavedra, 2018; O'neill et al., 1980; Paine et al., 2012).

In this paper, we explore the impact of including parametric uncertainty on model dynamics and introduce recent advances in uncertainty quantification. We focus on a simplified model of sea ice algae, motivated by the heterogeneity of observed ice algal densities over small spatial scales arising from spatially variable environmental conditions (Cimoli et al., 2017; Eicken et al., 1991). This heterogeneity at small scales creates challenges for modelling bloom dynamics over ecologically meaningful scales and connecting ice algal models to field data (Vancoppenolle et al., 2013). The methods demonstrated in this case study can be applied to any mathematical model with parametric uncertainty.

Case study: Sea ice algae in a heterogeneous environment

Sea ice is a porous material, due to salt expulsion and entrapment during freezing (Petrich & Eicken, 2017). Sea ice algae—specifically, diatoms—are endemic to this extreme environment, living within brine inclusions inside the ice (Arrigo, 2017). These algal populations exist in low densities through the dark winter months and then experience a period of rapid growth in the spring once light is no longer limiting (Leu et al., 2015). Termination of this spring bloom is often linked to nutrient limitation, as high algal concentrations during the bloom deplete the available nutrients faster than they are replenished (Gosselin et al., 1990; Leu et al., 2015). Ice algae provide lipid-rich food for the lowest polar marine trophic levels early in the spring (Leu et al., 2011; Søreide et al., 2010), support overwintering Antarctic krill (Kohlbach et al., 2017), and impact polar biogeochemical fluxes, such as carbon cycling (Boetius et al., 2013; Vancoppenolle et al., 2013).

Heterogeneity of the sea ice ecosystem over meter scales complicates both empirical and modelling studies of sea ice ecology (Cimoli et al., 2017; Mundy et al., 2005; Swadling et al., 1997). Important physical characteristics such as snow depth (Massom et al., 2001; Perovich et al., 1998), ice thickness (Stroeve et al., 2021; Thorndike et al., 1975), and salinity (Tucker III et al., 1984), vary over these small spatial scales, resulting in patchy ice algal distributions (Gosselin et al., 1986; Meiners et al., 2017; Rysgaard et al., 2001). Because of this heterogeneity and the nonlinearity of bloom dynamics, regional mean values of snow depth or ice thickness cannot be expected to accurately predict regional algal dynamics (Abraham et al., 2015; Stroeve et al., 2021).

To understand how spatial heterogeneity impacts algal bloom dynamics, we focus on the ecosystem component of the larger biogeochemical system, ignoring seasonality and other physical forcings. Many of the same fluxes determining pelagic primary production also determine primary production within sea ice (Vancoppenolle & Tedesco, 2017); thus, in the simplest case, ice algae are modelled using so-called *N*-*P* models (Deal et al., 2011; Dupont, 2012; Lavoie et al., 2005). These models describe the changes in the concentration of some limiting nutrient N (often assumed to be silicate in sea ice models [Lavoie et al., 2005; Vancoppenolle et al., 2010; Tedesco & Vichi, 2014]) and primary producers P (here, sea ice algae).

We build on the differential equation model of Huppert et al. (2002),

$$\begin{aligned} \frac{dN}{dt} &= \underbrace{\alpha}_{\text{input}} - \underbrace{\beta NP}_{\text{uptake}} - \underbrace{\eta N}_{\text{loss}} \\ \frac{dP}{dt} &= \underbrace{\gamma \beta NP}_{\text{growth}} - \underbrace{\delta P}_{\text{death}}, \end{aligned} \tag{1}$$

with initial conditions $N(0) = n_0$, $P(0) = p_0$ and units of $\mathrm{kg\,m^{-3}}$ and $\mathrm{mg\,m^{-3}}$ for algal and nutrient concentrations, respectively. This model assumes a constant rate of nutrient input, α, while nutrients and algae are lost proportionally (η, δ) to their concentrations. Nutrient uptake by algae obeys mass action with the uptake rate (β) proportional to both nutrient and algal concentrations. Algal growth is proportional to nutrient uptake, with conversion parameter γ. This model is able to produce features typical of algal bloom dynamics: a period of rapid increase in algae that leads to a depletion of nutrients and subsequent termination of the algal bloom (Figure 1a).

We are interested in algal dynamics at the start of the spring, when light becomes sufficient for photosynthesis. Due to highly variable and unpredictable physical processes occurring throughout the winter, we treat the initial conditions, n_0 and p_0, as parameters that are independent of the other parameters in the model.

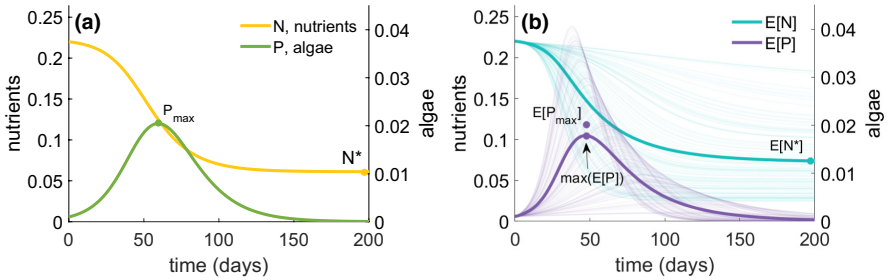

FIGURE 1 (a) Typical algal bloom dynamics, generated by model (1) with a fixed growth parameter. (b) 100 Monte Carlo simulations (thin lines) and mean behaviour (bold lines) of the same algal bloom model as in (a), but with a random variable describing growth (model (2)). The maximum bloom height (P_{max} and $\max_t(\mathbb{E}[P])$) and asymptotic nutrient levels (N^* and $\mathbb{E}[N^*]$) are shown. In (b), the mean of the peak of each Monte Carlo simulation, $\mathbb{E}[P_{max}]$, is also shown for comparison with $\max_t(\mathbb{E}[P])$, the peak of the mean curve. Both are ecologically meaningful, but only the former is analytically computable. In (a), the growth rate is $\beta = 1$, and in (b), the growth rate is a random variable B with a uniform distribution $U(0.6,1.4)$ with a mean of $\mathbb{E}[B] = 1$. Other parameters, $\alpha = 0, \gamma = 0.8, \delta = 0.1, \eta = 0, n_0 = 0.22, p_0 = 0.001$, consistent with the closed model described in the text.

To implicitly incorporate spatial environmental heterogeneity into this model, we first consider the algal growth parameter β at each location in a chosen region as a random variable B with a known distribution. Later, we also consider the initial conditions (n_0, p_0) as random variables (N_0, P_0). We assume $\mathbb{E}[B] = \beta$, and compare the model with a random growth parameter to the same model with the growth parameter fixed at its mean value.

Algal dynamics at any randomly selected location within the chosen region are described by a modified version of (1),

$$\begin{aligned} \frac{dN}{dt} &= \alpha - BNP - \eta N \\ \frac{dP}{dt} &= \gamma BNP - \delta P, \end{aligned} \tag{2}$$

where $N(0) = n_0$ and $P(0) = p_0$. Here, $N(t)$ and $P(t)$ are random variables, and consideration of $\mathbb{E}[N]$ and $\mathbb{E}[P]$—the expected values with respect to the random growth rate B—provides an improved representation of nutrient and algal concentrations over regional scales by averaging the algal dynamics at each individual location (Figure 1b). Consideration of higher moments of N and P provides insight into the variance and skewness we may expect in field data.

Overview

We first present analytic approaches to understanding the biases introduced by treating parameters as fixed rather than random variables, illustrated through the ice algal case study (Section 2). However, analytic techniques are unable to provide insight into other aspects of the model with random parameters, such as the mean and higher moments of the algal concentration (e.g., variance, skewness) throughout the bloom. This more sophisticated understanding is necessary if we are to successfully connect models to field data. While statistics such as the mean or variance can be computed using Monte Carlo sampling, this requires the generation of a large ensemble of model solutions. Generating this ensemble may be undesirable for models that are computationally costly, such as biogeochemical models that model sea ice algae within the context of ocean-ice-atmospheric couplings (Tedesco & Vichi, 2014). In Section 3, we introduce polynomial chaos (PC) techniques for parametric uncertainty quantification, which can compute these statistics to any given level of accuracy with a small fraction of the computational effort required by Monte Carlo methods (Smith, 2013). So-called intrusive PC methods have been used in a few ecological studies to date (Albi et al., 2021; Harman & Johnston, 2016; Hickson & Roberts, 2014; Stanescu & Chen-Charpentier, 2009), mainly focused on infectious disease ecology. The nonintrusive methods we introduce here have only recently been introduced into the mathematical disease ecology literature (Bertaglia & Pareschi, 2021; Zanella et al., 2021), and have the potential to offer the broader community both computational gains and improved ease of implementation.

Using these uncertainty quantification methods in our case study, we treat both model parameters and initial conditions as random variables, exploring the relative contribution of these multiple sources of uncertainty to variability in model outputs through global sensitivity analyses (Section 4). We find that treating the algal growth rate and initial conditions as random variables lowers the regional peak bloom intensity, reduces the total algae present, and suggests earlier bloom initiation. Modelled algal concentrations are positively skewed except at the peak of the bloom; this novel statistical pattern can be compared to field data. Global sensitivity analysis suggests that heterogeneous initial algal concentrations may contribute less than heterogeneous algal growth rates and initial nutrient levels to variability in algae throughout the bloom.

The results from the case study demonstrate that neglecting spatial heterogeneity or other sources of parametric uncertainty when estimating regional dynamics may result in biased estimates of important ecological

phenomena and make it difficult to connect models to data. We hope that these approaches—both analytic and computational—will empower ecological modellers faced with parametric uncertainty.

ANALYTIC RESULTS ON MODEL BIASES IN BLOOM HEIGHT AND TOTAL ALGAE

Many ecological models with fixed parameters can be understood, at least to some extent, analytically. These analytic results can provide insights into the effects of neglecting uncertain parameters, using the approaches demonstrated in this section. We begin by building on known results for our case study model with fixed parameters (model (1)). Huppert et al. (2002) described and analysed two versions of this model: a "closed" model, where nutrients neither leave nor enter the system (i.e. $\alpha = \eta = 0$), and an "open" model, where $\alpha > 0$ and/or $\eta > 0$. Important properties of the bloom can be calculated for the (simpler) closed model, including the maximum bloom height P_{max}, the level of nutrients remaining after the bloom N^*, and the total amount of algae present during the bloom P_{total} (Figure 1a).

The relationship between these bloom features and the growth rate β determines whether considering the algal growth rate as a random versus fixed parameter will result in their over- or under-estimation. This hinges on the relationship described by Jensen's inequality (Jensen, 1906), which tells us whether a nonlinear model that includes a random parameter will overestimate (or underestimate) a model that only includes that parameter's mean value. Jensen's inequality is of relevance for understanding how variable parameters affect predictions in any nonlinear relationship in ecology (Ruel & Ayres, 1999). For example, consider the maximum algal density achieved during the bloom, P_{max}, which depends on the algal growth rate, β (the exact relationship is provided below in [5]). Since P_{max} is a function of β, we can take its derivative(s) with respect to β. P_{max} is concave up in β if its second derivative exists and is positive; it is concave down if this second derivative is negative. Then by Jensen's inequality, if P_{max} is concave up in β, the model with the fixed parameter β will underestimate the mean of the model with the random parameter B, so $P_{max}(\mathbb{E}[B]) \leq \mathbb{E}\left[P_{max}(B)\right]$ (and vice versa if P_{max} is concave down).

The dynamics of the closed model approach a globally asymptotically stable solution, $(N^*, P^*) = (N^*, 0)$, where N^* depends on the initial conditions (Huppert et al., 2002). So regardless of the parameter values, the algae eventually die out (Figure 1). Whether a bloom occurs prior to this die out depends on the model parameters (Huppert et al., 2002; Murray, 1993). At time $t = 0$, algal dynamics follow

$$\left.\frac{dP}{dt}\right|_{t=0} = (\gamma\beta n_0 - \delta)p_0. \tag{3}$$

If $n_0 < \rho$, where

$$\rho = \delta / (\gamma\beta), \tag{4}$$

then algae are lost (δ) faster than they are growing ($\gamma\beta n_0$), so (3) is negative and the algal population decays monotonically, precluding the possibility of a bloom. Because we are interested in bloom dynamics, we assume $n_0 > \rho$ unless stated otherwise.

Analysis of bloom height, asymptotic nutrients and total algae

We now demonstrate what Jensen's inequality tells us about three analytic properties of the closed algal bloom model: peak bloom height, asymptotic nutrient levels and total algae present during the bloom.

Peak bloom height, P_{max}

For the closed model, the peak bloom height can be found analytically (details in S1) to be

$$P_{max}(\beta) = -\frac{\delta}{\beta} + \frac{\delta}{\beta}\ln\left(\frac{\rho}{n_0}\right) + p_0 + \gamma n_0. \tag{5}$$

P_{max} takes a single value if we assume a fixed parameter β, and becomes a random variable if we substitute the random variable B into (5). P_{max} is a twice differentiable function of β, and we find that $\frac{d^2P_{max}}{d\beta^2} > 0$ if the model parameters satisfy $n_0/\rho < \sqrt{\mathrm{e}}$, where e is Euler's number. So by Jensen's inequality,

$$\mathbb{E}\left[P_{max}(B)\right] > P_{max}(\beta), \qquad \text{if } \frac{n_0}{\rho} < \sqrt{\mathrm{e}}. \tag{6}$$

Because we have assumed $n_0 > \rho$, for any parameterization satisfying the relatively narrow bounds of $1 < n_0/\rho < \sqrt{e} \approx 1.65$, the model with the fixed parameter β underestimates the expected peak bloom height (Figure 2a). Conversely, if $n_0/\rho > \sqrt{\mathrm{e}}$ then $\frac{d^2P_{max}}{d\beta^2} < 0$ and so

$$\mathbb{E}\left[P_{max}(B)\right] < P_{max}(\beta), \qquad \text{if } \frac{n_0}{\rho} > \sqrt{\mathrm{e}}, \tag{7}$$

in which case the model with the fixed parameter β overestimates the expected peak bloom height (Figure 2b). Note that these results are about $\mathbb{E}\left[P_{max}\right]$, which is subtly different from the peak of the mean curve, $\max_t(\mathbb{E}[P])$ (Figure 1b) though the general pattern holds (details in S2).

For a given distribution of the random variable B, we can also calculate $\mathbb{E}[P_{max}]$ and use this to confirm that the difference between the peak heights of the models with fixed versus random growth rates increases with the variance in B (Figure 2a,b; details in S2). Additionally, we can explore how a given distribution of B is distorted by the nonlinear model by deriving the full distribution of $P_{max}(B)$ (details in S4).

Nutrients remaining after the bloom, N^*

The asymptotic nutrient concentration in the closed model, $N^*(\beta)$, is the solution to the transcendental equation (derivation in S1),

$$N^*(\beta) = \rho \ln\left(\frac{N^*(\beta)}{n_0}\right) + \frac{p_0}{\gamma} + n_0. \tag{8}$$

Using similar approaches as for P_{max} (details in S5), we find conditions under which the model with a fixed parameter β over- or under-estimates the model with random parameter B. For example, if $N^*(\beta) < \rho/2$, then the model with a fixed growth rate underestimates the model with the random growth parameter (Figure 2c).

Total algae in the bloom, P_{total}

All nutrients that are removed from the closed system must have been removed through uptake by algae. It follows that the total amount of algae in the bloom is $P_{total} = p_0 + \gamma(n_0 - N^*)$, the initial algae concentration plus the change in nutrients times the conversion factor. Because

$$\frac{d^2 P_{total}}{d\beta^2} = -\gamma \frac{d^2 N^*}{d\beta^2}, \tag{9}$$

our understanding of N^* implies that if N^* is sufficiently small ($< \rho/2$), then (9) is negative, so the model with a fixed growth rate β will predict greater cumulative algal biomass compared to the model with a random growth parameter B.

This difference in predicted total algae increases with increased variance in the distribution B (Figure 3).

UNCERTAINTY QUANTIFICATION METHODS FOR PARAMETRIC UNCERTAINTY

The results from the previous sections demonstrate that taking the mean value of a parameter rather than considering its full distribution biases nonlinear model outputs, and this bias typically increases with the variance in the distribution of that parameter. In the closed algal bloom model, simplifying spatially heterogeneous growth rates to their mean value results in an overestimation of both the maximum bloom concentrations P_{max} and the total algae present during the bloom P_{total} for a wide range of parameters.

These general patterns hold even if we no longer restrict ourselves to the closed model and rather consider the open model, where nutrients can enter and leave the system (details in S6). As in the closed model, the ratio of parameters n_0/ρ determines whether the models with a fixed growth parameter over- or under-estimates the bloom height. However, we cannot obtain analytic expressions for the peak bloom height or total algae during the bloom for the open model.

For many ecological models, analytic results using the methods outlined in the previous section may not be possible. Additionally, we may wish to treat several of the model parameters as random variables, complicating analytic approaches. For our case study, the algal and nutrient concentrations at any given spatial location at the start of the bloom are the product of complex and stochastic physical and biological drivers, so we may also

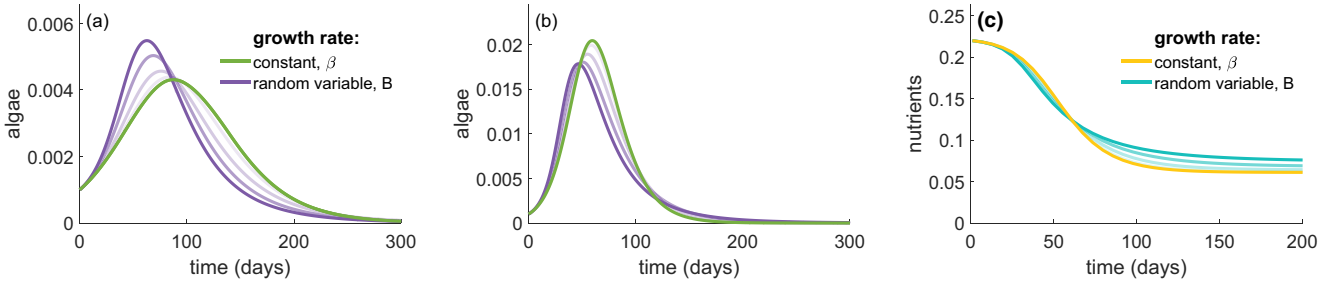

FIGURE 2 Comparison of blooms with different peak heights P_{max} and asymptotic nutrient levels N^*. Green and yellow curves are from model (1) with a fixed growth parameter β. Purple and blue curves show $\mathbb{E}[P]$ and $\mathbb{E}[N]$ from 10,000 Monte Carlo simulations of (2), the model with random growth parameter B. Depending on model parameters, the model with a fixed growth rate underestimates (a) or overestimates (b) the maximum bloom height. (c) The model with a fixed growth rate underestimates the nutrients remaining after the bloom, and thus overestimates the amount of algae present during the bloom. Here B is uniformly distributed with mean β and width 2ν. Curves of increasing colour intensity for both $\mathbb{E}[P]$ and $\mathbb{E}[N]$ correspond to increasing values of ν, ranging through 0.1,0.2, … ,0.4. (a) Parameters: $\alpha = 0, \beta = 1, \gamma = 0.8, \delta = 0.1, \eta = 0, n_0 = 0.16, p_0 = 0.001$. (b, c) Parameters as in (a) except $n_0 = 0.22$. Parameters in (a) and (b) were chosen to satisfy the conditions of (6) and (7), respectively.

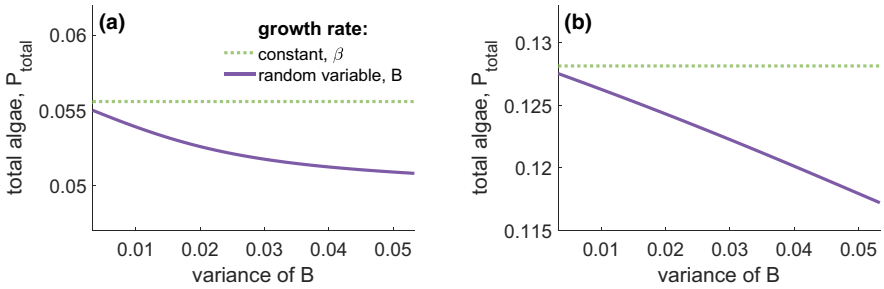

FIGURE 3 (a) Comparison of (expected) total algae present throughout the algal bloom for the model with a fixed algal growth rate, β, and with a random growth rate, B. The difference in total algae between the two models increases as the variance in B increases. Here B is uniformly distributed with mean β and width 2ν; variance values shown correspond to $\nu = 0.1$ to 0.4. Parameters in (a) and (b) correspond to those in Figure 2a,b, respectively.

want to consider the initial conditions (n_0, p_0) as random variables (N_0, P_0).

We now introduce numerical uncertainty quantification methods. Using these methods, we quantify uncertainty (i.e., variability) in the trajectories of N and P resulting from uncertainty in the random parameters, here denoted $\mathbf{y} = (B, N_0, P_0)$. We introduce and investigate three approaches: Monte Carlo, and intrusive and nonintrusive PC methods, using our case study model.

Monte Carlo

The Monte Carlo procedure is the conceptually simplest approach, where statistics of (N, P) are approximated with empirical ensemble estimates. This method requires drawing a large number (M) of independent and identically distributed random samples of the parameters $\boldsymbol{y}_i$ and then solving the deterministic model (1) with these parameters to obtain an ensemble of trajectories $N(t, \boldsymbol{y}_i)$ and $P(t, \boldsymbol{y}_i)$ for $i = 1, \ldots, M$,. This ensemble of trajectories provides an estimate of the full distribution of the random variables $N(t, \boldsymbol{y})$ and $P(t, \boldsymbol{y})$, as well as any statistical properties such as the mean and variance.

However, error estimates for the mean $\mathbb{E}[P(t)]$ scale like $1/\sqrt{M}$, which reveals the relative inefficiency of this approach: a large M is needed to guarantee accuracy of the estimator to several digits. For computationally expensive models, solving the model thousands of times may not be feasible.

Polynomial chaos

PC approaches provide an alternative to Monte Carlo for approximating N and P (Wiener, 1938; Xiu, 2010; Xiu & Karniadakis, 2002). The overall idea is separation of variables, separating the time-dependence of the solutions from their dependence on the random parameters $\boldsymbol{y}$. The approximations to $(N(t, \boldsymbol{y}), P(t, \boldsymbol{y}))$, known as PC expansions, are denoted $(N_V(t, \boldsymbol{y}), P_V(t, \boldsymbol{y}))$ and defined as

$$N_V(t, \boldsymbol{y}) = \sum_{j=1}^{n} \tilde{N}_j(t)\phi_j(\boldsymbol{y}),$$
$$P_V(t, \boldsymbol{y}) = \sum_{j=1}^{n} \tilde{P}_j(t)\phi_j(\boldsymbol{y}). \tag{10}$$

The dependence of N and P on the random parameter(s) $\boldsymbol{y}$ is approximated by a set of polynomials, $\{\phi_j\}_{j=1}^{n}$. The choice of polynomials dictates the ability of the PC expansion to closely approximate the true solutions (N, P). The subscript V here denotes the polynomial subspace spanned by the polynomials $\{\phi_j\}_{j=1}^{n}$, and serves as a reminder that the quality of the approximation depends on the choice of polynomials. For example, if $\boldsymbol{y}$ is a single (scalar) parameter, then it is standard to choose the n functions ϕ_j so that they span polynomials up to degree $n-1$, and if $\boldsymbol{y}$ is a vector-valued parameter, then choosing ϕ_j to span polynomials up to a given degree is a common choice, but in this vector-valued setting there are numerous alternative choices. In this manuscript we follow two conventional assumptions: that $\{\phi_j\}_{j=1}^{n}$ are orthonormal polynomials (i.e., $\mathbb{E}[\phi_j(\boldsymbol{y})\phi_k(\boldsymbol{y})] = \delta_{j,k}$ with $\delta_{j,k}$ the Kronecker delta function), and that $\phi_1(\boldsymbol{y}) \equiv 1$.

The time-dependence of N and P is captured by the Fourier Series-like coefficients $(\tilde{N}_j(t), \tilde{P}_j(t))$, which need to be computed. We provide an overview of both intrusive and nonintrusive methods for computing $(\tilde{N}_j(t), \tilde{P}_j(t))$ in Sections 3.2.1 and 3.2.2.

The convergence of PC approximations to (N, P) as $n \to \infty$ is well understood, provided N and P depend smoothly on $\boldsymbol{y}$ (Ernst et al., 2012). Once we have constructed the coefficients $\tilde{N}_j, \tilde{P}_j$, then the PC approximations to the full solutions (10) can be used to estimate relevant statistics of $(N(t, \boldsymbol{y}), P(t, \boldsymbol{y}))$. For example, the mean and variance of algae concentrations are conveniently given by

$$\mathbb{E}[P(t)] \approx \tilde{P}_1(t), \qquad \mathrm{var}[P(t)] \approx \sum_{j=2}^{n} \tilde{P}_j^2(t). \tag{11}$$

This also yields estimation of more sophisticated statistical quantities; for example, the coefficient of variation (CV), defined as the ratio of the standard deviation to the mean, or higher moments such as skewness.

For the ice algae model with three random parameters, we are also interested in how much of the predicted variance in the algae concentration is the result of variance in algal growth rate versus variance in the initial conditions. PC approaches make it easy to compute partial variances and closely-related global sensitivities. Global sensitivity analysis allows us to ask whether uncertainty in some parameters or initial conditions gets amplified more over time than others.

If P_V (or, similarly, N_V) has finite variance, then it admits the ANOVA decomposition (Sobol, 2001),

$$P_V(t;\boldsymbol{y}) = \sum_{I \in \mathbb{P}(\boldsymbol{y})} P_{V,I}\left(t;\boldsymbol{y}_I\right),$$

where $\mathbb{P}(\boldsymbol{y})$ denotes the power set, the set of all subsets of $\boldsymbol{y}$ (including the empty set). For the ice algae model with $\boldsymbol{y} = \left(B, N_0, P_0\right)$, the power set is

$$\mathbb{P}(\boldsymbol{y}) = \left\{\emptyset, \{B\}, \left\{N_0\right\}, \left\{P_0\right\}, \left\{B, N_0\right\}, \left\{B, P_0\right\}, \left\{N_0, P_0\right\}, \left\{B, N_0, P_0\right\}\right\}.$$

Each $P_{V,I}$ can be defined recursively as,

$$P_{V,\emptyset}(t) = \mathbb{E}\left[P_V(t)\right], P_{V,I}\left(t;\boldsymbol{y}_I\right) = \mathbb{E}\left[P_V(t)\,|\boldsymbol{y}_I\right] - \sum_{\substack{J \subset I \\ J \neq I}} P_{V,J}\left(t;\boldsymbol{y}_J\right),$$

where $\boldsymbol{y}_I$ denotes the set of variables $\boldsymbol{y}$ restricted to the subset I. Under this decomposition, one can express the absolute and relative partial variance attributable to a given non-empty set as $\mathrm{var}_I\left[P_V\right]$ and S_I, respectively, defined as,

$$\mathrm{var}_I\left[P_V(t)\right] = \mathrm{var}\left[P_{V,I}\left(t;\boldsymbol{y}_I\right)\right], \qquad S_I(t) = \frac{\mathrm{var}_I\left[P_V(t)\right]}{\mathrm{var}\left[P_V(t)\right]}. \tag{12}$$

S_I is called the *global sensitivity* associated with subset I, and has the partition of unity property $\Sigma S_I = 1$, if summed over all subsets I. These global sensitivities can be used to assess which parameter sets I contribute most (or least) to variability in the predicted algal concentrations. These partial variances depend on time, and so the contribution of any subset I of parameters to the variance in P (or N) does also. The partial variances, and in turn the global sensitivities S_I, are easily computed from the PC coefficients $\tilde{P}_j$ (e.g., sometimes explicitly, or by numerical integration).

PC approaches provide a powerful strategy for propagating randomness in parameters and initial conditions through a differential equation model. The next critical step is to compute the coefficients $\left(\tilde{N}_j, \tilde{P}_j\right)$ in (10).

Intrusive methods

We introduce an intrusive method (known as the Stochastic Galerkin method) for computing the coefficients $\left(\tilde{N}_j, \tilde{P}_j\right)$ (Xiu, 2010). This approach is *intrusive* in the sense that we must manipulate the original model.

The main idea of this approach is to replace (N, P) in the nutrient-algae model (2) with the approximations $\left(N_V, P_V\right)$, as defined in (10). This results in

$$\begin{aligned} \frac{dN_V}{dt} &= \alpha - BN_V P_V - \eta N_V, & N_V(0) &= N_0 \\ \frac{dP_V}{dt} &= \gamma BN_V P_V - \delta P_V, & P_V(0) &= P_0 \end{aligned} \tag{13}$$

where $\boldsymbol{y} = \left(B, N_0, P_0\right)$ are random variables. For each $j = 1, \ldots, n$, we then multiply both sides of the equations (13) by ϕ_j and take the expected value with respect to the random parameters $\mathbf{y}$, resulting in a differential equation for $d\tilde{N}_j/dt$ and $d\tilde{P}_j/dt$. More technically speaking, since $\left\{\phi_j\right\}$ are orthonormal polynomials, we are projecting (13) onto $V = \mathrm{span}\left\{\phi_j\right\}_{j=1}^{n}$, the polynomial subspace spanned by the polynomials. In doing so for each j value, we replace the *random* differential equations in (13) by a *deterministic* size-$2n$ system of coupled differential equations in each of the coefficients $d\tilde{N}_j/dt$ and $d\tilde{P}_j/dt$ for $j = 1, \ldots, n$ (further details in S7). Numerically solving this system yields the desired coefficients $\tilde{N}_j$ and $\tilde{P}_j$ for each $j = 1, \ldots, n$. Using this method amounts to replacing the random differential equations of our original model (2) with deterministic ones, at the cost of increased dimension of the system, from 2 to $2n$ coupled differential equations. When both the number of random parameters and the dimension n of the polynomial space is small, this intrusive approximation is typically far less computationally demanding than Monte Carlo.

Nonintrusive methods

Unlike intrusive methods, nonintrusive methods do not manipulate the original model, but instead sample the nutrient-algae model (2) for specific fixed values of the random parameters $\boldsymbol{y}$. This makes nonintrusive methods intuitively similar to Monte Carlo and so more transparent to implement than intrusive methods.

A nonintrusive method for building PC approximations is the Stochastic Collocation approach. We discuss the quadrature approach for computing (approximations to) the coefficients $\left(\tilde{N}_j, \tilde{P}_j\right)$, but note that numerous alternative approaches exist (Narayan & Zhou, 2015; Smith, 2013; Xiu, 2010).

We assume the existence of a $\boldsymbol{y}$-quadrature rule $(\boldsymbol{y}_m, w_m)_{m=1}^M$ comprised of deterministic nodes $\boldsymbol{y}_m$ (these are specific values of the random parameters $\mathbf{y}$ dictated by quadrature theory) and corresponding weights w_m such that

$$\mathbb{E}[f(\boldsymbol{y})] \approx \sum_{m=1}^{M} w_m f(\boldsymbol{y}_m), \tag{14}$$

for any function $f(\boldsymbol{y})$ with finite variance. Using this quadrature rule, we can compute

$$\mathbb{E}[\phi_j(\boldsymbol{y})P(t;\boldsymbol{y})] \approx \sum_{m=1}^{M} w_m \phi_j(\boldsymbol{y}_m) P(t;\boldsymbol{y}_m). \tag{15}$$

Since $\mathbb{E}[\phi_j(\boldsymbol{y})P(t;\boldsymbol{y})]$ would be a mean-square optimal choice for $\tilde{P}_j(t)$, then we compute

$$\tilde{P}_j(t) = \sum_{m=1}^{M} w_m \phi_j(\boldsymbol{y}_m) P(t;\boldsymbol{y}_m), \tag{16}$$

and similarly for $\tilde{N}_j$. The approximation above is the nonintrusive PC method using quadrature rules. This procedure requires only the availability of a size-M ensemble of trajectories:

$$N(t;\boldsymbol{y}_m), \quad P(t;\boldsymbol{y}_m), \qquad m = 1, \ldots, M,$$

which are easily constructed if any numerical solver for the original model with fixed parameters (here model (1)) is provided. Thus the major computational bottleneck involved in this approach is construction of the above ensemble, requiring M simulations of the original size-2 system of differential equations.

To compare costs of the nonintrusive and intrusive methods, we describe how M can be chosen: in the case where $\boldsymbol{y}$ is scalar, then accurate choices with $M = n$ exist, where n is the dimension of the polynomial space. If $\boldsymbol{y} = (B, N_0, P_0)$, and the components of $\boldsymbol{y}$ are assumed to be independent, then the expansion with common choices of a total degree polynomial space and tensorized quadrature rule requires $M \approx 6n$ samples. Thus we can directly compare the computational cost of the intrusive versus the nonintrusive methods in terms of the parameter n for the algal bloom model: the intrusive method requires a single solve of a size-$2n$ system of coupled ordinary differential equations, and the nonintrusive method requires $6n$ solves of a size-2 system of ordinary differential equations. Since the complexity of solving coupled differential equations is superlinear in the size of the system, the nonintrusive method is much more computationally efficient here.

UNCERTAINTY QUANTIFICATION RESULTS

All three methods described above—Monte Carlo, intrusive and nonintrusive PC methods—result in bloom dynamics that are indistinguishable from each other (Figure 4).

To understand the ecological implications of incorporating spatial heterogeneity into the algal bloom model, we present results for the nutrient-algae model with the growth rate B and initial conditions (N_0, P_0) treated as random variables. Using the nonintrusive method described above, we can examine the mean, standard deviation, CV and skewness of the algae distribution $P(t)$, as well as global sensitivities throughout the bloom (Figure 5). The mean provides an estimate of regional bloom dynamics, accounting for spatial heterogeneity in growth rates and initial conditions, which can then be compared to the dynamics of the model with fixed parameters. The standard deviation, CV and skewness provide statistical quantities that could qualitatively and quantitatively be compared against field data. The global sensitivities provide insight into the drivers of algal variability. The same analysis could similarly be conducted for $N(t)$.

For comparison with the earlier results in Figures 1–4, we set the mean values of P_0 and N_0 to 0.001 and 0.22,

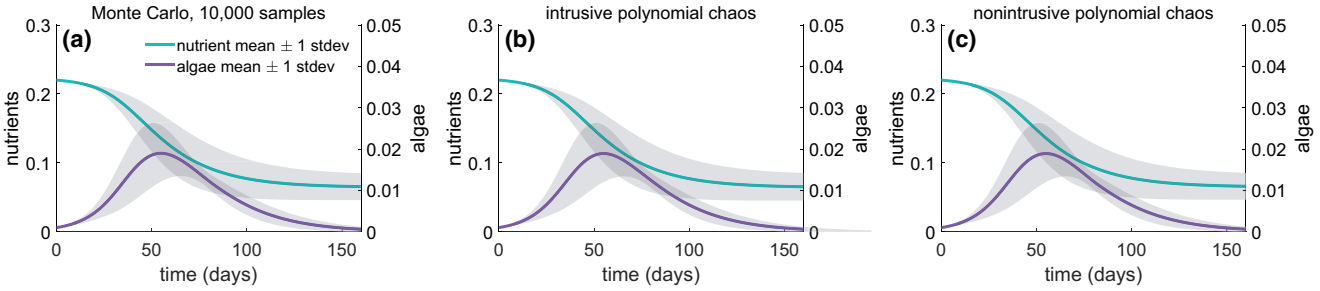

FIGURE 4 Mean and standard deviation of nutrient and algae dynamics approximated using three different uncertainty quantification methods for model (2) with the growth rate B as a random parameter; (a) 10,000 Monte Carlo simulations, (b) intrusive polynomial chaos expansion methods ($n = 6$) and (c) nonintrusive polynomial chaos expansion methods ($M = 7$ samples). Shaded regions show ± 1 standard deviation from the mean. Parameters: $\alpha = 0, \gamma = 0.8, \delta = 0.1, \eta = 0, n_0 = 0.22, p_0 = 0.001$. The growth rate B has a uniform distribution $U(0.8,1.2)$ with a mean of $\mathbb{E}[B] = 1$.

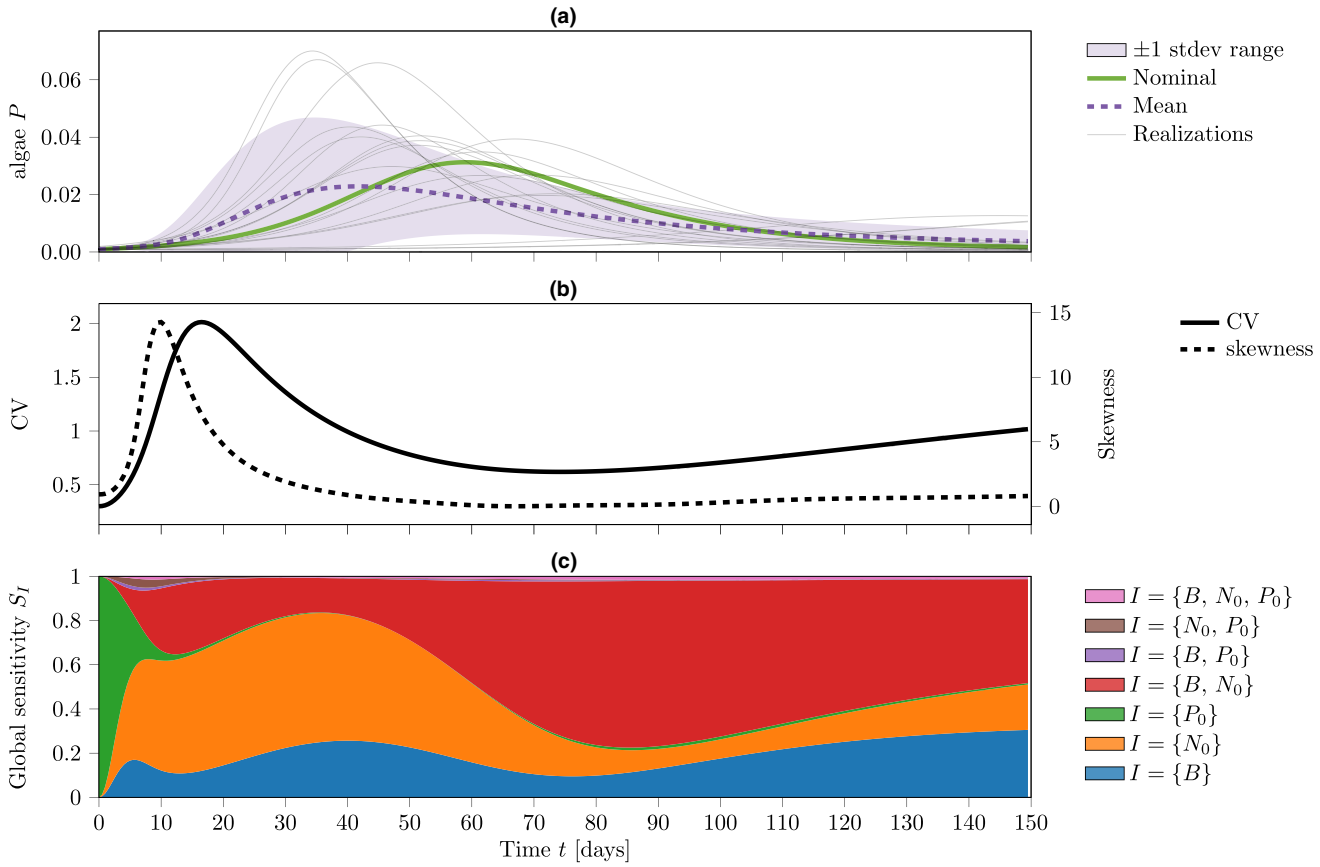

FIGURE 5 Statistics of the algal concentration $P(t)$ obtained using nonintrustive polynomial chaos. (a) Mean and ± 1 standard deviation of $P(t)$, along with the nominal trajectory (all parameters set to their mean values), and several realisations with randomly sampled parameters. (b) Coefficient of variation (CV) and skewness of $P(t)$. (c) Global sensitivities of $P(t)$, indicating the relative contributions of variability in the random parameters to variability in $P(t)$. Most sensitivities *not* involving B and N_0 are visually close to zero for time $t \gtrsim 15$. Parameters: $\alpha = 0.00075, \gamma = 0.8, \delta = 0.1, \eta = 0.0005$. B is uniformly distributed on [0.6,1.4]. N_0 and P_0 are lognormally distributed with respective mean values of 0.22 and 0.001, each with a CV of 0.3. The number of samples is $M = 1331$, as described in S8.

respectively. We choose values from lognormal distributions in order to ensure nonnegative values and consider all combinations (9 total) of three coefficients of variation (CV) for both N_0 and P_0: 0.1, 0.3 and 0.5. We consider B as a uniform random variable on [0.6,1.4], which has CV ≈ 0.23. Technical details (e.g., the specific polynomial bases and quadrature rules) are described in S8.

For all scenarios considered, the model with random parameters shows a regional (mean) algal bloom with a lower and earlier peak than the model with all parameters fixed at their mean values (Figure 5a, Figure S3). The standard deviation of the algae concentration increases as the bloom increases, both in absolute terms (Figure 5a) and relatively, as indicated by high CV values (Figure 5b). Skewness tends to be positive, with the highest values occurring during the first part of the bloom, and then increasing again at the end of the bloom. The decline in skewness near the peak of the bloom is especially pronounced when there is less variability in the initial conditions (i.e., low CV for P_0 and N_0; Figure S4), and corresponds to the time at which roughly half of the locations have already reached their peak and are coming down from it, while the other half are still in the growing phase of their local bloom.

Global sensitivity analysis (Figure 5c, Figure S5) suggests that variability in the initial algae concentration P_0 only contributes to variability in subsequent algal dynamics early on in the bloom. As the bloom progresses, variability in the initial nutrients N_0 and the growth rate B contribute most to variability in algal dynamics.

Our analysis of the closed model (Section 2) with only a single random parameter, B, suggested that the model with fixed parameters will often overestimate the total biomass present throughout the bloom. To see whether this generalizes to the case with variable initial conditions, we consider the average value of both models (with random and fixed parameters). The model with random parameters (Figure 5) had an average algae concentration 11% lower than the model with all parameters fixed at their mean values (0.0110 $\mathrm{kg\,m^{-3}}$ and 0.0122 $\mathrm{kg\,m^{-3}}$, respectively).

DISCUSSION

Parametric uncertainty in ecological models can generate dynamics that differ in important ways from their deterministic counterparts. The effects of treating

uncertain parameters as random variables can, in some circumstances, be understood analytically, as demonstrated in Section 2. When analytic approaches are insufficient or difficult to implement, contemporary uncertainty quantification methods (Section 3) accurately and efficiently compute statistics and global sensitivities of ecological processes subject to parametric uncertainty. We have shown that these powerful methods can and should play an increasingly prominent role in ecological modelling.

In our case study of sea ice algal bloom dynamics, we incorporated small scale heterogeneity into a canonical model of nutrient-dependent algal growth by treating key parameters as random variables. Treating the algal growth rate as a random parameter and applying Jensen's inequality to analytically obtainable bloom features provided insights into conditions under which spatial heterogeneity in growth rates may decrease the maximum regional bloom intensity and the total biomass present during the bloom (Section 2).

We then employed sophisticated uncertainty quantification methods to this system to understand the full dynamics and higher moments of the algal distribution, as well as incorporate spatial heterogeneity in initial conditions (Section 3). We introduced and compared three methods: Monte Carlo, intrusive and nonintrusive PC methods. For this relatively simple system, all three methods produced the same results. The choice of method depends on the trade-off between the time required to code PC approaches and the time required to run Monte Carlo simulations. PC approaches are recommended when computational constraints render Monte Carlo methods impractical. Monte Carlo methods are recommended when each deterministic run is computationally inexpensive, or when there is a large number of random parameters.

Using nonintrusive PC, we demonstrated that including spatial heterogeneity in the algal growth rate and the initial conditions of the full open model reduces the peak bloom height and shifts the timing of the bloom peak. It also reduces the total biomass present throughout the bloom, consistent with the analytic findings of the simpler model in Section 2. The predicted coefficient of variation, skewness and global sensitivities of the algal concentration suggest patterns which may be compared against field data.

We illustrated the analytic approaches using the peak bloom intensity and total biomass, quantities that are analytically tractable (Section 2). Many important quantities of interest are not state variables and cannot be found analytically, such as the timing of the peak algal density. Uncertainty quantification approaches can still be applied in this case, with the same potential to improve computational efficiency.

The (closed) model presented here matches the classic infectious disease model of Kermack and McKendrick (1927) if $\gamma = 1$. Our analytic results in Section 2 suggest that if there is uncertainty in the infection rate (analogous to the algal growth rate), then the expected dynamics of an epidemic may vary considerably from a model that ignores that uncertainty. For example, whether the expected infection peak is over- or underestimated when ignoring parametric uncertainty can be shown to depend on R_0, the basic reproduction number.

Infectious disease ecology is one of the few fields of ecology where PC methods have been introduced, including intrusive methods (Albi et al., 2021; Calatayud et al., 2018; Chen-Charpentier & Stanescu, 2010; Harman & Johnston, 2016; Hickson & Roberts, 2014) and, more recently, nonintrusive approaches (Bertaglia et al., 2021; Bertaglia & Pareschi, 2021; Zanella et al., 2021). The few studies in other areas of ecology have all used an intrusive approach (Calatayud et al., 2020; Stanescu et al., 2009; Stanescu & Chen-Charpentier, 2009). However, nonintrusive methods have the potential to contribute significantly to our understanding of ecological dynamics in more complex situations (e.g., with multiple random parameters or more complicated model structures) under parametric uncertainty, as they are easier to implement and often significantly more computationally efficient. Several uncertainty quantification software packages have been developed in recent years, including packages for R (e.g., `uncertainty`), Matlab (Marelli & Sudret, 2014; Petzke et al., 2020), and Python (Olivier et al., 2020; The SCI Institute, University of Utah, 2020); see Ghanem et al. (2017) for a more comprehensive list. As these uncertainty quantification methods continue to gain popularity, further development of software packages—especially those facilitating nonintrusive PC methods—should empower ecological modellers to regularly include spatial heterogeneity or other forms of parametric uncertainty.

Considering parameters as random variables has implications not only for improving our understanding of how parametric uncertainty affects model dynamics but also for comparing ecological models to data. For example, given the assumptions resulting in Figure 5, we would expect to see a large variance in algal measurements taken at individual locations during the bloom, with a high coefficient of variation being a possible marker of bloom initiation and termination. Field surveys of ice algal densities during the spring consistently find positively skewed distributions, with a high proportion of cores having very little algae (Lange et al., 2016; Meiners et al., 2017). This is consistent with the skewness of $P(t)$ observed in Figure 5. Those results suggest that if skewness values transition from being positive to around 0 during what appears to be the bloom, this may indicate that the regional bloom has peaked and will soon begin to decline.

Historically, ice algae data have been obtained using small numbers of ice cores (Miller et al., 2015). More recently, technological advances in remotely operated vehicles and remote sensing are allowing for data collection

over larger areas (Cimoli et al., 2017; Forrest et al., 2019; Pinkerton & Hayward, 2021). Connecting these data meaningfully to sea ice biogeochemical models remains challenging (Steiner et al., 2016; Vancoppenolle & Tedesco, 2017), but incorporating small-scale heterogeneity into these models, using the framework presented here, is a step in the right direction. To incorporate random parameters into these and other computationally expensive models, we propose nonintrusive PC, as it offers a powerful combination of computational efficiency, relative ease of implementation and computation of global sensitivities.

CONCLUSIONS

Ecological uncertainty can be represented in ecological models by treating parameters as random variables. We have provided a road map for analysing such models, including introducing tools from modern uncertainty quantification such as PC expansions. Using these methods in our case study provides a first step towards incorporating small scale heterogeneity in parameter values into regional models of ice algal blooms. Consideration of this heterogeneity results in changes in estimated bloom phenology and intensity, and should not be neglected if we are to understand this important component of polar ecology and biogeochemistry.

AUTHOR CONTRIBUTION

JRR conceptualized this study and obtained analytic results. JRR and AN conducted the Uncertainty quantification analysis. JRR and AN wrote the first draft of the manuscript. All authors contributed substantially to study design and revisions.

ACKNOWLEDGEMENTS

AN was partially supported by the National Institute of Biomedical Imaging and Bioengineering of the National Institutes of Health under grant number U24EB029012, and the National Science Foundation under award DMS-1848508. KMG acknowledges support from the Office of Naval Research Applied and Computational Analysis Program through Grant N00014-21-1-2909 and the Division of Mathematical Sciences at the U.S. National Science Foundation through Grants DMS-1715680 and DMS-2136198.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/ele.14095.

OPEN RESEARCH BADGES

This article has earned an Open Materials badge for making publicly available the components of the research methodology needed to reproduce the reported procedure and analysis. All materials are available at: 10.5281/zenodo.6924827.

DATA AVAILABILITY STATEMENT

The code used to generate all figures and analysis is freely available at https://doi.org/10.5281/zenodo.6924827.

ORCID

Jody R. Reimer https://orcid.org/0000-0001-7742-2728
Frederick R. Adler https://orcid.org/0000-0002-9022-3157
Kenneth M. Golden https://orcid.org/0000-0002-1688-183X
Akil Narayan https://orcid.org/0000-0002-5914-4207

REFERENCES

Abraham, C., Steiner, N., Monahan, A. & Michel, C. (2015) Effects of subgrid-scale snow thickness variability on radiative transfer in sea ice. *Journal of Geophysical Research: Oceans*, 120(8), 5597–5614.

Albi, G., Pareschi, L. & Zanella, M. (2021) Control with uncertain data of socially structured compartmental epidemic models. *Journal of Mathematical Biology*, 82(7), 1–41.

Arrigo, K.R. (2017) Sea ice as a habitat for primary producers. *Sea Ice*, 352–369.

Bertaglia, G., Boscheri, W., Dimarco, G. & Pareschi, L. (2021) Spatial spread of COVID-19 outbreak in Italy using multiscale kinetic transport equations with uncertainty. *Mathematical Biosciences and Engineering: MBE*, 18(5), 7028–7059.

Bertaglia, G. & Pareschi, L. (2021) Hyperbolic compartmental models for epidemic spread on networks with uncertain data: application to the emergence of covid-19 in Italy. *Mathematical Models and Methods in Applied Sciences*, 31(12), 2495–2531.

Boetius, A., Albrecht, S., Bakker, K., Bienhold, C., Felden, J., Fernández-Méndez, M. et al. (2013) Export of algal biomass from the melting Arctic Sea ice. *Science*, 339(6126), 1430–1432.

Calatayud, J., Cortés, J.C. & Jornet, M. (2020) Improvement of random coefficient differential models of growth of anaerobic photosynthetic bacteria by combining bayesian inference and gpc. *Mathematical Methods in the Applied Sciences*, 43(14), 7885–7904.

Calatayud, J., Cortés, J.C., Jornet, M. & Villanueva, R.J. (2018) Computational uncertainty quantification for random time-discrete epidemiological models using adaptive gpc. *Mathematical Methods in the Applied Sciences*, 41(18), 9618–9627.

Cenci, S. & Saavedra, S. (2018) Uncertainty quantification of the effects of biotic interactions on community dynamics from nonlinear time-series data. *Journal of the Royal Society Interface*, 15(147), 20180695.

Chen-Charpentier, B.M. & Stanescu, D. (2010) Epidemic models with random coefficients. *Mathematical and Computer Modelling*, 52(7–8), 1004–1010.

Cimoli, E., Meiners, K.M., Lund-Hansen, L.C. & Lucieer, V. (2017) Spatial variability in sea-ice algal biomass: an under-ice remote sensing perspective. *Advances in Polar Science*, 28(4), 268–296.

Deal, C., Jin, M., Elliott, S., Hunke, E., Maltrud, M. & Jeffery, N. (2011) Large-scale modeling of primary production and ice algal biomass within Arctic Sea ice in 1992. *Journal of Geophysical Research: Oceans*, 116(C7), C07004.

Dupont, F. (2012) Impact of sea-ice biology on overall primary production in a biophysical model of the pan-Arctic Ocean. *Journal of Geophysical Research: Oceans*, 117(C8), C00D17.

Eicken, H., Lange, M.A. & Dieckmann, G.S. (1991) Spatial variability of sea-ice properties in the northwestern Weddell Sea. *Journal of Geophysical Research: Oceans*, 96(C6), 10603–10615.

Ernst, O.G., Mugler, A., Starkloff, H.-J. & Ullmann, E. (2012) On the convergence of generalized polynomial chaos expansions. *ESAIM: Mathematical Modelling and Numerical Analysis*, 46(2), 317–339.

Forrest, A.L., Lund-Hansen, L.C., Sorrell, B.K., Bowden-Floyd, I., Lucieer, V., Cossu, R. et al. (2019) Exploring spatial heterogeneity of Antarctic Sea ice algae using an autonomous underwater vehicle mounted irradiance sensor. *Frontiers in Earth Science*, 7, 169.

Ghanem, R., Higdon, D. & Owhadi, H. (2017) *Handbook of uncertainty quantification*, Vol. 6. New York: Springer.

Gosselin, M., Legendre, L., Therriault, J.-C. & Demers, S. (1990) Light and nutrient limitation of sea-ice microalgae (Hudson Bay, Canadian Arctic). *Journal of Phycology*, 26(2), 220–232.

Gosselin, M., Legendre, L., Therriault, J.-C., Demers, S. & Rochet, M. (1986) Physical control of the horizontal patchiness of sea-ice microalgae. *Marine Ecology Progress Series*, 29(3), 289–298.

Harman, D.B. & Johnston, P.R. (2016) Applying the stochastic galerkin method to epidemic models with uncertainty in the parameters. *Mathematical Biosciences*, 277, 25–37.

Harper, E.B., Stella, J.C. & Fremier, A.K. (2011) Global sensitivity analysis for complex ecological models: a case study of riparian cottonwood population dynamics. *Ecological Applications*, 21(4), 1225–1240.

Hickson, R. & Roberts, M. (2014) How population heterogeneity in susceptibility and infectivity influences epidemic dynamics. *Journal of Theoretical Biology*, 350, 70–80.

Huppert, A., Blasius, B. & Stone, L. (2002) A model of phytoplankton blooms. *The American Naturalist*, 159(2), 156–171.

Jensen, J.L.W.V. (1906) Sur les fonctions convexes et les inégalités entre les valeurs moyennes. *Acta Mathematica*, 30(1), 175–193.

Kermack, W.O. & McKendrick, A.G. (1927) A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London A*, 115(772), 700–721.

Kohlbach, D., Lange, B.A., Schaafsma, F.L., David, C., Vortkamp, M., Graeve, M. et al. (2017) Ice algae-produced carbon is critical for overwintering of Antarctic krill Euphausia superba. *Frontiers in Marine Science*, 4, 310.

Lange, B.A., Katlein, C., Nicolaus, M., Peeken, I. & Flores, H. (2016) Sea ice algae chlorophyll a concentrations derived from under-ice spectral radiation profiling platforms. *Journal of Geophysical Research: Oceans*, 121(12), 8511–8534.

Lavoie, D., Denman, K. & Michel, C. (2005) Modeling ice algal growth and decline in a seasonally ice-covered region of the Arctic (resolute passage, Canadian archipelago). *Journal of Geophysical Research: Oceans*, 110(C11), C11009.

Leu, E., Mundy, C., Assmy, P., Campbell, K., Gabrielsen, T., Gosselin, M. et al. (2015) Arctic spring awakening–steering principles behind the phenology of vernal ice algal blooms. *Progress in Oceanography*, 139, 151–170.

Leu, E., Søreide, J., Hessen, D., Falk-Petersen, S. & Berge, J. (2011) Consequences of changing sea-ice cover for primary and secondary producers in the European Arctic shelf seas: timing, quantity, and quality. *Progress in Oceanography*, 90(1–4), 18–32.

Marelli, S. & Sudret, B. (2014) UQLab: a framework for uncertainty quantification in Matlab. In: *Vulnerability, Uncertainty, and Risk: Quantification, Mitigation, and Management*, Second International Conference on Vulnerability and Risk Analysis and Management (ICVRAM) and the Sixth International Symposium on Uncertainty, Modeling, and Analysis (ISUMA) July 13-16, 2014. pp. 2554–2563.

Massom, R.A., Eicken, H., Hass, C., Jeffries, M.O., Drinkwater, M.R., Sturm, M. et al. (2001) Snow on Antarctic Sea ice. *Reviews of Geophysics*, 39(3), 413–445.

McCarthy, M.A., Thompson, C.J., Moore, A.L. & Possingham, H.P. (2011) Designing nature reserves in the face of uncertainty. *Ecology Letters*, 14(5), 470–475.

Meiners, K., Arndt, S., Bestley, S., Krumpen, T., Ricker, R., Milnes, M. et al. (2017) Antarctic pack ice algal distribution: floe-scale spatial variability and predictability from physical parameters. *Geophysical Research Letters*, 44(14), 7382–7390.

Miller, L.A., Fripiat, F., Else, B.G., Bowman, J.S., Brown, K.A., Collins, R.E. et al. (2015) Methods for biogeochemical studies of sea ice: the state of the art, caveats, and recommendations. *Elementa: Science of the Anthropocene*, 3, 000038.

Mundy, C., Barber, D. & Michel, C. (2005) Variability of snow and ice thermal, physical and optical properties pertinent to sea ice algae biomass during spring. *Journal of Marine Systems*, 58(3–4), 107–120.

Murray, J.D. (1993) Epidemic models and the dynamics of infectious diseases. In: *Mathematical biology*. New York, NY: Springer, pp. 610–650.

Narayan, A. & Zhou, T. (2015) Stochastic collocation on unstructured multivariate meshes. *Communications in Computational Physics*, 18(1), 1–36 arXiv:1501.05891 [math.NA].

Olivier, A., Giovanis, D., Aakash, B., Chauhan, M., Vandanapu, L. & Shields, M.D. (2020) UQpy: a general purpose python package and development environment for uncertainty quantification. *Journal of Computational Science*, 47, 101204.

O'neill, R., Gardner, R. & Mankin, J. (1980) Analysis of parameter error in a nonlinear model. *Ecological Modelling*, 8, 297–311.

Paine, C.T., Marthews, T.R., Vogt, D.R., Purves, D., Rees, M., Hector, A. et al. (2012) How to fit nonlinear plant growth models and calculate growth rates: an update for ecologists. *Methods in Ecology and Evolution*, 3(2), 245–256.

Perovich, D.K., Roesler, C.S. & Pegau, W.S. (1998) Variability in Arctic Sea ice optical properties. *Journal of Geophysical Research: Oceans*, 103(C1), 1193–1208.

Petrich, C. and Eicken, H. (2017). Overview of sea ice growth and properties. In Thomas, D. N., editor, *Sea Ice*, chapter 1, pages 1–41. Oxford, UK: Wiley Blackwell, 3rd edition.

Petzke, F., Mesbah, A., and Streif, S. (2020). Pocet: a polynomial chaos expansion toolbox for matlab. *arXiv preprint arXiv:2007.05245*.

Pinkerton, M.H. & Hayward, A. (2021) Estimating variability and long-term change in sea ice primary productivity using a satellite-based light penetration index. *Journal of Marine Systems*, 221, 103576.

Regan, H.M., Colyvan, M. & Burgman, M.A. (2002) A taxonomy and treatment of uncertainty for ecology and conservation biology. *Ecological Applications*, 12(2), 618–628.

Ruel, J.J. & Ayres, M.P. (1999) Jensen's inequality predicts effects of environmental variation. *Trends in Ecology & Evolution*, 14(9), 361–366.

Rysgaard, S., Kühl, M., Glud, R.N. & Hansen, J.W. (2001) Biomass, production and horizontal patchiness of sea ice algae in a high-Arctic fjord (young sound, NE Greenland). *Marine Ecology Progress Series*, 223, 15–26.

Smith, R.C. (2013) *Uncertainty quantification: theory, implementation, and applications*. Philadelphia: SIAM-Society for Industrial and Applied Mathematics.

Sobol, I. (2001) Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation*, 55(1–3), 271–280.

Søreide, J.E., Leu, E.V., Berge, J., Graeve, M. & Falk-Petersen, S. (2010) Timing of blooms, algal food quality and calanus glacialis reproduction and growth in a changing Arctic. *Global Change Biology*, 16(11), 3154–3163.

Stanescu, D., Chen-Charpentier, B., Jensen, B.J. & Colberg, P.J. (2009) Random coefficient differential models of growth of anaerobic photosynthetic bacteria. *Electronic Transactions on Numerical Analysis*, 34, 44–58.

Stanescu, D. & Chen-Charpentier, B.M. (2009) Random coefficient differential equation models for bacterial growth. *Mathematical and Computer Modelling*, 50(5–6), 885–895.

Steiner, N., Deal, C., Lannuzel, D., Lavoie, D., Massonnet, F., Miller, L.A. et al. (2016) What sea-ice biogeochemical modellers need from observers. *Elementa: Science of the Anthropocene*, 4, 000084.

Stroeve, J., Vancoppenolle, M., Veyssiere, G., Lebrun, M., Castellani, G., Babin, M. et al. (2021) A multi-sensor and modeling approach for mapping light under sea ice during the ice-growth season. *Frontiers in Marine Science*, 7, 1253.

Swadling, K.M., Gibson, J.A., Ritz, D.A. & Nichols, P.D. (1997) Horizontal patchiness in sympagic organisms of the Antarctic fast ice. *Antarctic Science*, 9(4), 399–406.

Tedesco, L. & Vichi, M. (2014) Sea ice biogeochemistry: a guide for modellers. *PLoS One*, 9(2), e89217.

The SCI Institute, University of Utah (2020). UncertainSCI. https://github.com/SCIInstitute/UncertainSCI.

Thorndike, A.S., Rothrock, D.A., Maykut, G.A. & Colony, R. (1975) The thickness distribution of sea ice. *Journal of Geophysical Research*, 80(33), 4501–4513.

Tucker, W., III, Gow, A. & Richter, J. (1984) On small-scale horizontal variations of salinity in first-year sea ice. *Journal of Geophysical Research: Oceans*, 89(C4), 6505–6514.

Vancoppenolle, M., Goosse, H., De Montety, A., Fichefet, T., Tremblay, B. & Tison, J.-L. (2010) Modeling brine and nutrient dynamics in Antarctic Sea ice: the case of dissolved silica. *Journal of Geophysical Research: Oceans*, 115(C2), C02005.

Vancoppenolle, M., Meiners, K.M., Michel, C., Bopp, L., Brabant, F., Carnat, G. et al. (2013) Role of sea ice in global biogeochemical cycles: emerging views and challenges. *Quaternary Science Reviews*, 79, 207–230.

Vancoppenolle, M. and Tedesco, L. (2017). Numerical models of sea ice biogeochemistry. In Thomas, D. N., editor, *Sea Ice*, chapter 20, 492–515. Oxford, UK: Wiley Blackwell, 3rd edition.

Wiener, N. (1938) The homogeneous chaos. *American Journal of Mathematics*, 60(4), 897–936.

Wu, J. & Li, H. (2006) Uncertainty analysis in ecological studies: an overview. *Scaling and uncertainty analysis in ecology*, 45–66.

Xiu, D. (2010) *Numerical methods for stochastic computations: a spectral method approach*. Princeton, NJ: Princeton University Press.

Xiu, D. & Karniadakis, G.E. (2002) The Wiener–Askey polynomial chaos for stochastic differential equations. *SIAM Journal on Scientific Computing*, 24(2), 619–644.

Zanella, M., Bardelli, C., Dimarco, G., Deandrea, S., Perotti, P., Azzi, M. et al. (2021) A data-driven epidemic model with social structure for understanding the COVID-19 infection on a heavily affected Italian province. *Mathematical Models and Methods in Applied Sciences*, 31(12), 2533–2570.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Reimer, J.R., Adler, F.R., Golden, K.M. & Narayan, A. (2022) Uncertainty quantification for ecological models with random parameters. *Ecology Letters*, 00, 1–13. Available from: https://doi.org/10.1111/ele.14095