Profile likelihood-based parameter and predictive interval analysis guides model choice for ecological population dynamics

Matthew J Simpson[a,*], Shannon A Walker[a], Emma N Studerus[a], Scott W McCue[a], Ryan J Murphy[a], Oliver J Maclaren[b]

[a]*School of Mathematical Sciences, Queensland University of Technology, Brisbane, Australia.*
[b]*Department of Engineering Science, University of Auckland, Auckland 1142, New Zealand.*

Abstract

Calibrating mathematical models to describe ecological data provides important insight via parameter estimation that is not possible from analysing data alone. When we undertake a mathematical modelling study of ecological or biological data, we must deal with the trade-off between data availability and model complexity. Dealing with the nexus between data availability and model complexity is an ongoing challenge in mathematical modelling, particularly in mathematical biology and mathematical ecology where data collection is often not standardised, and more broad questions about model selection remain relatively open. Therefore, choosing an appropriate model almost always requires case-by-case consideration. In this work we present a straightforward approach to quantitatively explore this trade-off using a case study exploring mathematical models of coral reef regrowth after some ecological disturbance, such as damage caused by a tropical cyclone. In particular, we compare a simple single species ordinary differential equation (ODE) model approach with a more complicated two-species coupled ODE model. Univariate profile likelihood analysis suggests that the both models are practically identifiable. To provide additional insight we construct and compare approximate prediction intervals using a new parameter-wise prediction approximation, confirming both the simple and complex models perform similarly with regard to making predictions. Our approximate parameter-wise prediction interval analysis provides explicit information about how each parameter affects the predictions of each model. Comparing our approximate prediction intervals with a more rigorous and computationally expensive evaluation of the full likelihood shows that the new approximations are reasonable in this case. All algorithms and software to support this work are freely available as jupyter notebooks on GitHub so that they can be adapted to deal with any other ODE-based models.

Keywords: Model Selection, Identifiability analysis; Population dynamics; Logistic growth; Coral reef; Great Barrier Reef.

*Corresponding author: matthew.simpson@qut.edu.au

1. Introduction

Appropriate model choice is a difficult problem with, in our view, no universal or automatic solution. However, *identifiability analysis* [6, 14, 22, 25, 33–35, 46] provides one set of tools that can help guide model choice in terms of determining an appropriate level of model complexity relative to available data. Formally, a model is *identifiable* when distinct parameter values imply distinct distributions of observations [25]. Less formally, a model is identifiable when it is possible to uniquely determine the model parameters using an infinite amount of ideal data. While a model being non-identifiable does not mean the model is wrong, it is, as a whole, then essentially unfalsifiable, although some aspects of the model may still be falsifiable or otherwise testable. Thus we may decide to prioritise the use of identifiable models over non-identifiable models. Often in the systems biology literature, *identifiability* is also referred to as *structural identifiability* [33–35, 46], and working within the framework of ordinary differential equation (ODE)-based models, it is possible to formally analyse the structural identifiability using a range of algebraic methods [3, 8, 9, 24]. This kind of analysis can be very insightful and aid in the choice of identifiable models over non-identifiable models, or help reformulate non-identifiable models into simpler identifiable models [12, 26, 27].

The focus of the present work is to interpret a set of real ecological field data using different mathematical models framed in terms of ODEs. Specifically, the data we consider describes the temporal increase in the percentage coral cover on a coral reef after some kind of natural disturbance, such as a tropical cyclone [49]. Such coral recovery data can be interpreted in several different ways, and there is no standard approach to interpreting this data. The simplest approach is to work with the evolution of the total hard coral cover as a function of time. Another is to consider dividing the total hard corals into different coral groups and consider modelling the temporal evolution of each group of hard corals. In essence, these approaches amount to working with either: (i) a single ODE model for the total hard coral cover or; (ii) a system of coupled ODE models describing the coral cover associated with each grouping. While structural identifiability analysis can aid model choice in general, here both sets of differential equations are already structurally identifiable [42] and the data are sparse and noisy. Thus we require additional means of guiding model choice and to assess the associated trade-off between model complexity and data availability. Here we focus on *practical identifiability*. While the formal definition of structural identifiability requires an infinite amount of ideal data, the less formal concept of practical identifiability (or *estimability*) seeks to describe whether it is possible to provide reasonably precise estimates using finite, non-ideal data [10, 25, 50]. This question typifies the kind of scenario one can expect to confront when using mathematical models to interpret biological and ecological data more generally.

In the literature, practical identifiability can be assessed in several different ways. In both Bayesian and (standard parametric) frequentist frameworks, identifiability is usually considered a property of the likelihood function, although a related but distinct Bayesian concept of *posterior identifiability* has also been proposed [10]. Computationally, practical non-identifiability can often be detected in a Bayesian setting through Monte Carlo Markov Chain (MCMC) sampling techniques when the samples fail to converge [17, 38, 39]. An alternative approach is to focus

directly on the likelihood function and use profile likelihood and computational optimisation to construct profiles for each parameter, or for combinations of parameters [4, 14, 17, 31, 39]. Such profiles are routinely used to construct frequentist asymptotic confidence intervals for target parameters [31, 36]. Our previous work indicates that profile likelihood analysis can lead to quantitatively similar conclusions as, for example, MCMC-based analysis, with the main practical difference being the profile likelihood analysis is often far less computationally demanding [39]. Other authors have also emphasised the suitability of profile likelihood for practical identifiability analysis relative to other approaches such as MCMC, bootstrap, or Fisher information-based methods [14, 34]. Therefore, as part of our study, we carry out a profile likelihood-based analysis. The computational differences between using MCMC-based sampling methods compared to optimization-based profiling methods is most significant for more computationally demanding models, such as working with systems of nonlinear partial differential equations [39] .

Working in a profile likelihood-based framework, we can assess models in terms of the widths of their associated parameter confidence intervals as an informal strategy for trading off model complexity with data availability. This provides a potentially more insightful means of model comparison relative to overly simplistic information criteria-based methods [1, 5, 20, 37, 42, 48]. However, in addition to constructing and comparing confidence intervals for each parameter in all models considered, we also construct parameter-wise *predictive* intervals for the mean coral cover trajectory using the univariate parameter profile likelihoods. (Note: here, we use the term *predictive* for any output from running the *forward* model given a set of parameters, that is we consider the mean trajectory as a predictive quantity.) This parameter-wise approach to prediction enables us to separate the roles of each parameter in terms of a more tangible model prediction rather than a less tangible parameter estimate [29]. For our models of interest, we show that generating parameter-wise predictions based on profile likelihoods provides an extremely efficient means of carrying out a sensitivity study and constructing approximate predictive intervals for the mean trajectory in the models we consider. Our approximate prediction intervals are analogous to performing a posterior predictive check in a Bayesian framework, except that our approach is likely to be more computationally efficient. This information provides insight that is complementary to the usual parameter confidence intervals considered in a typical practical identifiability analysis and helps guide model choice according to user goals, for example parameter inference or prediction or some combination of these.

Overall, this work provides insight into the question of model choice for scenarios where models are structurally identifiable, but may differ in terms of practical identifiability. Although we motivate this work with an ecological field data set describing coral reef regrowth after a disturbance, our approach can be applied directly to any model based on ordinary differential equations. All of our software is freely available in the open source Julia language on GitHub.

2. Field data and modelling objectives

We work with field data describing regrowth of hard corals off an island on the Great Barrier Reef, Australia. The regrowth process involves coral recolonisation after some external disturbance, such as the impact of a tropical cyclone [13]. Developing the ability to understand the mechanisms that drive this regrowth, and to predict the regrowth process, is extremely important as climate change-related disturbances increase in frequency [18, 49]. For example, if the time scale of regrowth is shorter than the time between disturbances we might anticipate that regrowth and recovery is likely, however if the time scale of regrowth is longer than the time between disturbances we might anticipate that regrowth and recovery is unlikely. One way to develop this understanding of the predictive capability is to calibrate a mathematical model to match the observations.

In this work we consider data describing the temporal regrowth of hard coral cover on a reef near Lady Musgrave Island, Australia (Figure 1(a)). Data in Figure 1(b) shows the temporal evolution of the total percentage area cover of hard corals, $S(t)$ (blue discs). The field methods by which this data are collected are described by Warne et al. [48] and the data is available through the eATLAS repository [13]. In brief, data collection involved repeated human surveys of several sites near Lady Musgrave Island over a period of approximately 10 years after some external disturbance, such as a cyclone. The data presented in Figure 1 here comes from site 1, whereas the data presented in the Supplementary Material document comes from site 3. Data collection involves measuring the proportion of area covered by different types of coral species, and the reported data is averaged over many transects across each site. In this case, the data shows a classical sigmoid growth response with apparently exponential growth at early times, followed by a reduction in net growth as the total hard coral cover approaches some maximum density at later times [42, 49]. Data in Figure 1(c) shows more detailed information where we plot the time evolution of the percentage of the dominant hard coral group called *Acroporidae*, $C_1(t)$ (red diamonds), together with the time evolution of the percentage of all other hard corals grouped together, $C_2(t)$ (green hexagons). For these sites at Lady Musgrave Island the dominant hard coral group is *Acroporidae*, but this may not be the case on other coral reefs in different locations. We also plot the total hard coral cover, $S(t) = C_1(t) + C_2(t)$ (blue discs), and note that the total hard coral cover in Figure 1(b) is identical to the total hard coral cover in Figure 1(c). Plotting the data in these two ways motivates the question of whether we should aim to model the time evolution of $S(t)$ alone, or whether we should aim to model the time evolution of both $C_1(t)$ and $C_2(t)$. This is the main question we will explore through our model-based analysis.

The field data in Figure 1 is associated with monitoring site 1 at Lady Musgrave Island. All algorithms and results used in this study are available on GitHub where we also present a second set of data for monitoring site 3. For completeness, all data and analysis associated with site 3 are given in the Supplementary Material documents. This additional data and the computational results for site 3 are presented in precisely the same format as given here for site 1. As we show, both data sets can be analysed with the same tools developed in this study, so we choose to present data for the site 1 here in the main document whereas the equivalent data and results for site 3 are provided on GitHub in the form of several user-friendly jupyter notebooks. It is worth noting the timescale of the coral regrowth in Figure

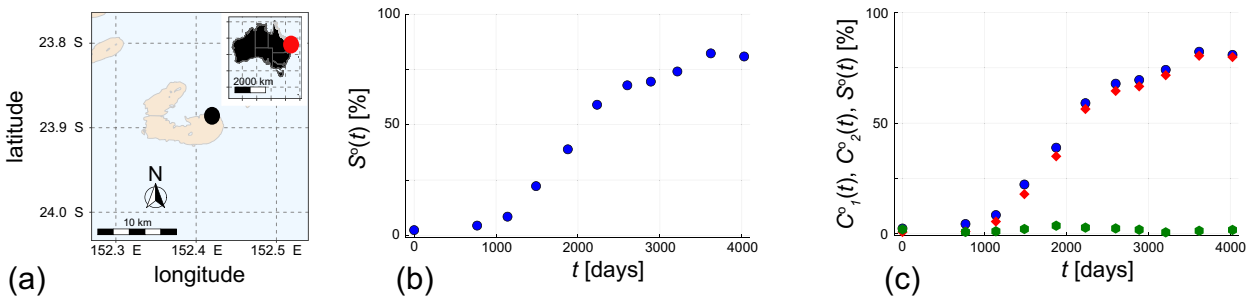

Figure 1: (a) Location of Lady Musgrave Island (black disc) relative to the Australian mainland (inset, red disc). (b) Field data showing the time evolution of the percentage total hard coral cover, $S(t)$ (blue discs) after some disturbance at monitoring site 1. (c) Field data showing the time evolution of the percentage *Acroporidae* coral cover, $C_1(t)$ (red diamonds), the percentage cover of all other hard corals, $C_2(t)$ (green hexagon), and the total hard coral cover, $S(t) = C_1(t) + C_2(t)$ (blue discs) after some disturbance at monitoring site 1. Note that the blue discs in (b) and (c) are identical. Time is measured in days after disturbance.

1 is approximately a decade, which is very different to a typical experiment in cell biology, for example, where cell proliferation takes places over a few days only [15, 19, 48]. Given the relatively rapid timescale of cell proliferation, it is typical for a cell biology experiment to be repeated several times to provide an estimate of the variability in the growth process [19, 48]. Here, in contrast, we deal with a very different type of biological growth process that takes approximately a decade, and so it is infeasible to repeat this experiment to provide a pre-estimate of the variability in the regrowth processes.

Throughout this work, all estimates of hard coral cover are expressed in terms of percentage coverage meaning that all model parameters associated with density are dimensionless. Further, the time-scale of regrowth is measured in days (Figure 1), so all growth and competition rates have dimensions of days^{-1}.

3. Methods

3.0.1. Process models

As outlined in Section 1, we will model the data in Figure 1 using two different approaches. First we will use a very simple logistic growth model to describe the evolution of the total coral cover, $S(t)$,

$$\frac{\mathrm{d}S(t)}{\mathrm{d}t} = \lambda S(t)\left[1 - \frac{S(t)}{K}\right], \tag{1}$$

where $S(t)$ is the percentage total coral cover at time t, $\lambda > 0$ is the growth rate and $K > 0$ is the long-time carrying capacity density. While Equation (1) can be solved exactly, we keep our approach general by solving all ODE models numerically, using standard numerical integration routines available in Julia [32]. The solution of Equation (1) depends upon the initial cover, $S(0)$, which means when we solve (1) we must specify three model parameters: λ, K and $S(0)$. While there are many other forms of single-species logistic growth-type models that we could use to capture this kind of sigmoid growth data [28, 45], our previous analysis of this data indicates that the logistic growth model

performs better than alternative models (e.g. Gompertz model, Richards' model) that encounter issues of practical identifiability [42].

The second approach we take is to divide the total population into two interacting subpopulations, $C_1(t)$ and $C_2(t)$, such that $S(t) = C_1(t) + C_2(t)$, where $C_1(t)$ represents the percentage cover of the dominant hard coral *Acroporidae*, and $C_2(t)$ is the percentage cover of all other hard corals. We propose the following minimal, but realistic, model

$$\frac{\mathrm{d}C_1(t)}{\mathrm{d}t} = \lambda_1 C_1(t)\left[1 - \frac{S(t)}{K}\right], \tag{2}$$

$$\frac{\mathrm{d}C_2(t)}{\mathrm{d}t} = \lambda_2 C_2(t)\left[1 - \frac{S(t)}{K}\right] - \delta C_2(t)\left[\frac{C_1(t)}{K}\right], \tag{3}$$

where $\lambda_1 > 0$ is the low-density growth rate of *Acroporidae*, $\lambda_2 > 0$ is the low-density growth rate of other hard corals, $K > 0$ is the long-time carrying capacity density and $\delta > 0$ describes the rate at which *Acroporidae* competes with the other hard corals. This coupled model is constructed to capture the key features in the measured data since we see that the percentage of *Acroporidae* increases monotonically with time (Figure 1(c)), where as the percentage of other hard corals initially increases and then eventually decreases at later time (Figure 1(c)). These observations motivate the use of a simple logistic-like term governing the evolution of $C_1(t)$ in Equation (2), whereas the initial increase in other hard corals followed by the later decline in $C_2(t)$ motivates a combination of a logistic-like growth term and a competition term in Equation (3) as one straightforward way of capturing these trends. Of course, there are other potential terms that one could include in our model, and we will discuss some other options later in Section 5. However, we choose to work with this relatively straightforward form since it appears to capture the key features of the data and is ecologically plausible. Solutions of Equations (2)–(3) are obtained numerically [32] and require the specification of the initial coverage of both coral types, $C_1(0)$ and $C_2(0)$. Therefore, when we solve (2)–(3) we must specify six parameters: $\lambda_1, \lambda_2, K, \delta, C_1(0)$ and $C_2(0)$.

3.1. Observation model

We assume that observed data $y^{\mathrm{o}}_{i,j}$ are measured at I discrete times, t_i, for $i = 1, 2, 3, \ldots, I$, and that at each time point we record J measurements, for $j = 1, 2, 3, \ldots, J$. Within this framework, when we work with total hard corals we only have $J = 1$, whereas when we divide the total hard corals into *Acroporidae* and all other hard corals we have $J = 2$. We use a superscript 'o' to distinguish the noisy observed data from the model predictions. Regardless of whether we work with Equation (1) or (2)–(3), noise-free model predictions are denoted $y_{i,j}(\theta) = y(t_i \mid \theta)$, where θ is a vector of parameters that we will estimate. We collect the noisy data into a matrix denoted by $y^{\mathrm{o}}_{1:I,1:J}$, where the subscript makes it clear that the data consists of a time series of I observations with J measurements at each time point. Similarly, we denote the process model solution by $y_{1:I,1:J}(\theta)$ for the matrix of predictions at the discrete time points and by $y(\theta)$ for the full (continuous) model trajectory over the time interval of interest [29].

We estimate parameters in the process models by assuming that the observed data are noisy versions of the model solutions of the form

$$y^{\mathrm{o}}_{i,j} \mid \theta \sim \mathcal{N}(y_{i,j}(\theta), \sigma^2(\theta)), \tag{4}$$

where σ^2 is a constant variance that we also estimate from the data. In this work we take the simplest, most fundamental approach by assuming all measurements are subjected to the same noise model with the same constant variance. Our parameter vector, θ, is extended to include this constant variance and hence we write $\sigma^2(\theta)$ for the variance component of θ. Regardless of whether we work with Equation (1) or (2)–(3), we make the assumption that the observation error is additive and normally distributed with zero mean and constant variance σ^2 [42]. Different error models could be used within our likelihood-based framework if supported by the available data, for example working with a time-dependent variance if appropriate [4].

3.2. Parameter estimation

Combining our assumptions about process and noise models means that working with Equation (1) involves estimating four parameters, $\theta = (\lambda, K, S(0), \sigma)$, whereas working with Equations (2)–(3) requires the estimation of seven parameters $\theta = (\lambda_1, \lambda_2, K, \delta, C_1(0), C_2(0), \sigma)$, where we make the simplifying assumption that the same noise model applies to all observations. Again, this assumption could be relaxed within our framework if the observed data justify doing so [4].

Taking a likelihood-based approach to parameter inference and uncertainty quantification, given a time series of observations together with our assumptions about the process and noise models, the log-likelihood function is

$$\ell(\theta \mid y^{\mathrm{o}}_{1:I,1:J}) = \sum_{i=1}^{I} \sum_{j=1}^{J} \log \left[\phi \left(y^{\mathrm{o}}_{i,j}; y_{i,j}(\theta), \sigma^2(\theta) \right) \right], \tag{5}$$

where $\phi(x; \mu, \sigma^2)$ denotes a Gaussian probability density function with mean μ and variance σ^2. Maximum likelihood estimation (MLE) provides an estimate of θ that gives the *best* match to the data (in the sense of assigning the data the highest probability). The MLE is given by

$$\hat{\theta} = \sup_{\theta} \ell(\theta \mid y^{\mathrm{o}}_{1:I,1:J}), \tag{6}$$

subject to bound constraints. The procedure for estimating $\hat{\theta}$ involves numerical maximisation of the log-likelihood, which can be achieved using many different algorithms. In this work, we find that a local optimisation routine from the open-source `NLopt` optimisation package in Julia performs well [21]. In particular, we use the Nelder-Mead optimisation routine within the `NLopt` with the default stopping criteria.

3.3. Practical identifiability analysis and profile predictions

We use a profile likelihood-based approach to explore practical identifiability by working with a normalised (relative) log-likelihood function

$$\hat{\ell}(\theta \mid y^{\mathrm{o}}_{1:I,1:J}) = \ell(\theta \mid y^{\mathrm{o}}_{1:I,1:J}) - \sup_{\theta} \ell(\theta \mid y^{\mathrm{o}}_{1:I,1:J}), \tag{7}$$

which we consider as a function of θ for a fixed set of data $y^{\mathrm{o}}_{1:I,1:J}$. Normalising the log-likelihood ensures that $\hat{\ell}(\hat{\theta} \mid y^{\mathrm{o}}_{1:I,1:J}) = 0$.

3.3.1. Profile likelihood for interest parameters

When the full parameter θ can be partitioned into an interest parameter ψ and nuisance parameter ω, where these may be vector valued in general, we write $\theta = (\psi, \omega)$. More generally, an interest parameter can be defined as any function of the full parameter vector, $\psi = \psi(\theta)$; the nuisance parameter is then the implied complement, allowing the full parameter vector to be reconstructed. For a set of data, $y^{\mathrm{o}}_{1:I,1:J}$, the profile log-likelihood for the interest parameter ψ given the partition (ψ, ω) is defined as [11, 31]

$$\hat{\ell}_p(\psi \,|\, y^{\mathrm{o}}_{1:I,1:J}) = \sup_{\omega \,|\, \psi} \hat{\ell}(\psi, \omega \,|\, y^{\mathrm{o}}_{1:I,1:J}), \tag{8}$$

which indicates that ω is optimised out for each fixed value of ψ. This implicitly gives a function $\omega^*(\psi)$ of optimal values of ω for each value of ψ, and defines a curve (or, for non-scalar interest parameters, hypersurface) with points $(\psi, \omega^*(\psi))$ in parameter space. In the case of an interest parameter defined by a function of the full parameter, the profile (or induced) log-likelihood is defined by the constrained optimisation problem [7, 10, 30]

$$\hat{\ell}_p(\psi \,|\, y^{\mathrm{o}}_{1:I,1:J}) = \sup_{\theta \,|\, \psi(\theta)=\psi} \hat{\ell}(\theta \,|\, y^{\mathrm{o}}_{1:I,1:J}), \tag{9}$$

in which the remaining degrees of freedom (implicit nuisance parameters) in θ, after fixing ψ, are optimised out. Note that if we consider ψ to be a scalar interest parameter then $(\psi, \omega^*(\psi))$ defines a univariate function that we can visualise as a curve. In contrast, if ψ is a pair of interest parameters then $(\psi, \omega^*(\psi))$ is a function of two variables that is often called a bivariate profile that we can visualise as a heat map or contour plot [40].

To demonstrate, consider working with Equations (2)–(3) where $\theta = (\lambda_1, \lambda_2, K, \delta, C_1(0), C_2(0), \sigma)$. If we wish to profile λ_2 we have $\psi(\theta) = \lambda_2$ and $\omega(\theta) = (\lambda_1, K, \delta, C_1(0), C_2(0), \sigma)$ so that

$$\begin{aligned} \hat{\ell}_p(\lambda_2 = \lambda_2^* \,|\, y^{\mathrm{o}}_{1:I,1:J}) &= \sup_{(\lambda_1,\lambda_2,K,\delta,C_1(0),C_2(0),\sigma) \,|\, \lambda_2=\lambda_2^*} \hat{\ell}(\lambda_1, \lambda_2, K, \delta, C_1(0), C_2(0), \sigma \,|\, y^{\mathrm{o}}_{1:I,1:J}) \\ &= \sup_{(\lambda_1,K,\delta,C_1(0),C_2(0),\sigma) \,|\, \lambda_2=\lambda_2^*} \hat{\ell}(\lambda_1, \lambda_2, K, \delta, C_1(0), C_2(0), \sigma \,|\, y^{\mathrm{o}}_{1:I,1:J}), \end{aligned} \tag{10}$$

noting that we could write $\hat{\ell}_p(\lambda_2 = \lambda_2^* \,|\, y^{\mathrm{o}}_{1:I,1:J})$ equivalently as $\hat{\ell}_p(\lambda_2 \,|\, y^{\mathrm{o}}_{1:I,1:J})$. Here we chose the former notation to make it clear that we construct the univariate profile of λ_2 by maximising the log-likelihood function at fixed values of $\lambda_2 = \lambda_2^*$.

In all cases, we implement this numerical optimisation using the same Nelder-Mead routine in `NLopt` that we use to estimate the MLE, $\hat{\theta}$ [21]. Our approach to computing the profile is very straightforward; we simply define a single uniform-spaced mesh across an interval in the interest parameter that contains the MLE approximately in the centre of the interval. Since the computations are relatively fast we can rapidly assess whether our choice of the width of the interval is sufficient. If, for example, we find that the interval leads to a profile that does not intersect the appropriate threshold we simple re-compute the profile on a wider interval. For each profile in this work we use a mesh of 40 uniformly spaced mesh points, and for each numerical optimisation calculation we simply use the MLE as the initial estimate for the iterative solver. The choice of working with a mesh of 40 uniformly spaced points can be varied easily,

and this depends on the shape of the resulting profile. If the profile gradually varies across the interval of interest, as we find in this study, 40 mesh points appears to be sufficient to describe the shape of the profile; however if the profiles are relatively rapidly varying then we can easily repeat the calculations using more mesh points to resolve the shape of the profile. With these profiles, log-likelihood-based confidence intervals can be defined from the profile log-likelihood by an asymptotic approximation in terms of the chi-squared distribution that holds for sufficiently regular problems [31, 36]. For example, 95% confidence intervals for a univariate (scalar) interest parameter correspond to a threshold profile log-likelihood value of -1.92 [36]. Other confidence intervals can be estimated by changing the threshold value accordingly.

3.3.2. Predictive profile likelihood and parameter-wise profile predictions

As discussed in [22, 29, 51], profile likelihoods for *predictive* quantities that are a deterministic function of the full parameter θ, such as model mean trajectories $y(\theta)$, are defined in the same way as for any other function of the full parameter (described in Section 3.3.1). For example, the model mean trajectory has an associated profile likelihood

$$\hat{\ell}_p(y(\theta) = y \,|\, y^{\mathrm{o}}_{1:I,1:J}) = \sup_{\theta \,|\, y(\theta)=y} \hat{\ell}(\theta \,|\, y^{\mathrm{o}}_{1:I,1:J}). \tag{11}$$

This defines the profile likelihood value for a prediction to be equal to the maximum likelihood value across all parameters giving that prediction. In principle, this gives the profile prediction for the full, continuous, model trajectory. However, this can be difficult to implement in practice and the literature typically focuses on a single-time prediction [22, 51].

In the present work, we are interested in some measure of the dependence of predictions on a given interest (target) parameter. However, as discussed in [29], given a partition $\theta = (\psi, \omega)$ and a function $q(\theta)$ of the full parameter, in general the quantity $q(\psi)$ is not well defined, unless q is independent of ω. To overcome this, we consider *parameter-wise profile predictions* [29], rather than the standard predictive profile likelihood discussed above. These consider the dependence of a predictive function $q(\psi, \omega)$ of the full parameter on the interest parameter ψ via its value along the corresponding profile curve, i.e. by considering $q(\psi, \omega^*(\psi))$ where $\omega^*(\psi)$ is the optimal value of the nuisance parameter for a given value of the interest parameter. As discussed in [29], the transformation of profile confidence intervals for ψ into confidence intervals for the predictive quantity of interest $q(\psi, \omega)$ will, in general, only be approximate due to the additional dependence on nuisance parameters. However, we can use these parameter-wise intervals as an intuitive diagnostic tool that can reveal the influence of an interest parameter on predictions, as well as explore their coverage properties via simulation. In contrast, a standard predictive profile cannot reveal the individual influence of particular parameters as all parameters are simultaneously optimised over.

Formally, the associated profile likelihood for $q(\psi, \omega^*(\psi))$ is defined by [29]:

$$\hat{\ell}_p(q(\psi, \omega^*(\psi)) = q \,|\, y^{\mathrm{o}}_{1:I}) = \sup_{\psi \,|\, q(\psi,\omega^*(\psi))=q} \hat{\ell}_p(\psi \,|\, y^{\mathrm{o}}_{1:I}). \tag{12}$$

We compute this quantity by first evaluating the predictive quantity q along the profile curve $(\psi, \omega^*(\psi))$ and assigning it tentative likelihood values equal to the associated profile likelihood value of $(\psi, \omega^*(\psi))$. Then $q(\psi, \omega^*(\psi))$ is assigned

a single profile likelihood value equal to the maximum profile likelihood value over compatible ψ values (the 'sup' step). If $q(\psi, \omega^*(\psi))$ is 1-1 in ψ then this value is simply the profile likelihood value of ψ, but it is well-defined regardless. This is analogous to constructing a profile for an interest predictive quantity q using a standard likelihood in θ but now starting from a profile likelihood for ψ. This makes it much easier to implement computationally – for a one-dimensional interest parameter we only need to evaluate the predictive quantity along a one-dimensional curve embedded in parameter space. For more on this definition, see [29]. Since we are evaluating q along the parameter profile curve $(\psi, \omega^*(\psi))$, the likelihood thresholds defining approximate confidence sets for the predictive quantities are the same as for the parameters.

The above process allows us to construct an approximate interval for the predictive quantity assuming the uncertainty is driven primarily by the parameter of interest. As a sensitivity tool, this analysis illustrates which aspects of a prediction depend on the interest parameter. For example, the early part of the mean trajectory may have a stronger dependence on a given parameter than the later part of the trajectory. This feature of this tool is desirable from the point of view of understanding the impacts of individual parameters on predictions, but also means the individual intervals will in general be expected to have lower coverage for the prediction than the associated profile parameter interval (due to neglecting additional dependence on/uncertainty in the nuisance parameters). To overcome this issue to some degree, when interested in overall uncertainty we can combine the individual intervals to give an approximate overall interval. In particular, given a collection of individual intervals for the same quantity based on different interest parameters, more conservative confidence intervals for that quantity can be constructed by taking the union over all intervals. As an illustration, given two intervals (or sets) $\mathcal{I}_q^K = [L_q^K, U_q^K]$ and $\mathcal{I}_q^\lambda = [L_q^\lambda, U_q^\lambda]$ for a quantity q based on profiles for K and λ, respectively, we can form an interval (or set) $\mathcal{I}_q^{K,\lambda} = \mathcal{I}_q^K \cup \mathcal{I}_q^\lambda$ which has coverage at least as great as the individual intervals. In the case where the two intervals overlap, we have $\mathcal{I}_q^{K,\lambda} = [\min\{L_q^K, L_q^\lambda\}, \max\{U_q^K, U_q^\lambda\}]$. As with the individual intervals, the precise coverage properties of these combined intervals are difficult to establish formally without direct simulation, but such union intervals can provide an approximate picture of the overall variation that is efficient to compute and intuitive to understand.

With these definitions, we will now compute various univariate profile likelihoods, parameter-wise prediction intervals and the union of the parameter-wise prediction intervals for both Equation (1) and Equations (2)–(3).

4. Results

4.1. Single species model

Figure 2(a) compares the solution of Equation (1) evaluated at $\hat{\theta}$ with the data, while in (b)–(c) univariate profiles are shown for each parameter. In each case, the univariate profiles are regularly shaped with clearly defined peaks, indicating that the model is practically identifiable [42]. Asymptotic confidence intervals for each parameter are determined by interpolating the profiles at the threshold of $\ell = -1.92$ and reported in the caption of Figure 2.

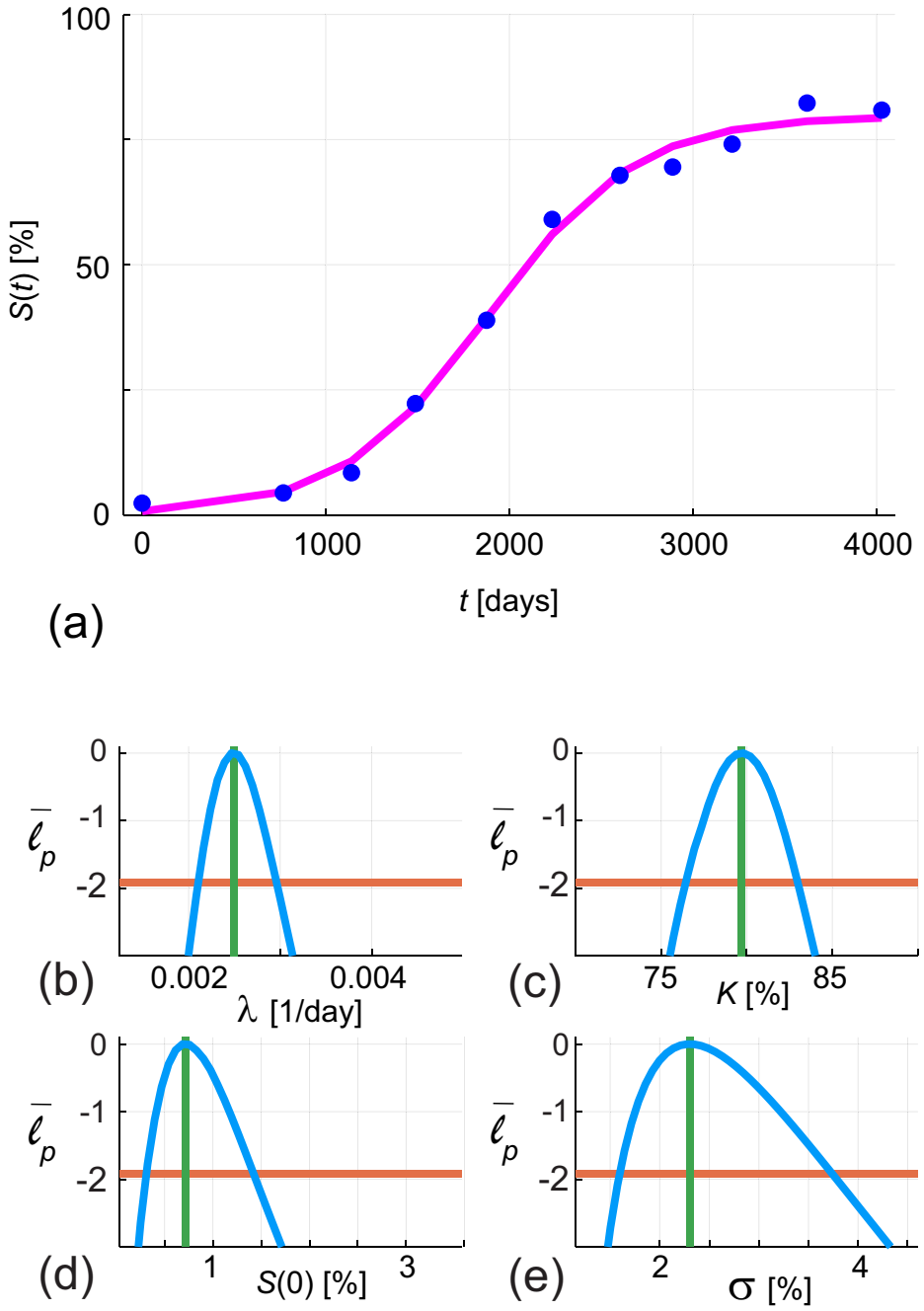

Figure 2: (a) Comparison of data (blue discs) and the solution of (1) evaluated at the MLE, $\hat{\theta} = (\lambda, K, S(0), \sigma) = (0.0025, 80, 0.7, 2.3)$ (pink line). (b) Univariate profile for λ gives $\hat{\lambda} = 0.0025 \in [0.0021, 0.0030]$. (c) Univariate profile for K gives $\hat{K} = 80 \in [76, 83]$. (d) Univariate profile for $S(0)$ gives $S(\hat{0}) = 0.7 \in [0.3, 1.4]$. (e) Univariate profile for σ gives $\hat{\sigma} = 2.2 \in [1.6, 3.7]$. Each profile is shown in solid blue; the MLE as a vertical green line, and the asymptotic log-likelihood threshold as a horizontal orange line.

The profiles in Figure 2 provide important quantitative information about the extent to which the quality and quantity of data identifies the parameters in the mathematical model. For example, the MLE for the carrying capacity density gives $\hat{K} = 80\,\%$, but the profiles allow us to identify the credible interval for this estimate, which tells us that the 95% credible interval is $\hat{K} = 80 \in [76, 83]$ in this case, which provides us with a clear quantitative measure of the uncertainty in our estimate.

With the univariate profiles in Figure 2 we construct parameter-wise profile predictions shown in Figure 3 for λ, K and $S(0)$, respectively. The width of the prediction interval for the parameter is correlated with the strength of the parameter in influencing the prediction of $S(t)$, and the trends we see are intuitively reasonable. For example, the prediction interval associated with K is relatively narrow at early time, but widens considerably at later times. This is entirely reasonable since we know that $S(t) \to K$ as $t \to \infty$ in the model. Similarly, the prediction interval associated with $S(0)$ leads to a relatively wide prediction interval at early time, which is reasonable since the value of $S(0)$ influences the early time dynamics, but the long-time solution of (1) is independent of $S(0)$, since $S(t) \to K$ as $t \to \infty$ regardless of $S(0)$. Note that we do not construct a prediction interval for σ since this is a measure of the noise in the data and the solution of (1) is completely independent of σ [29].

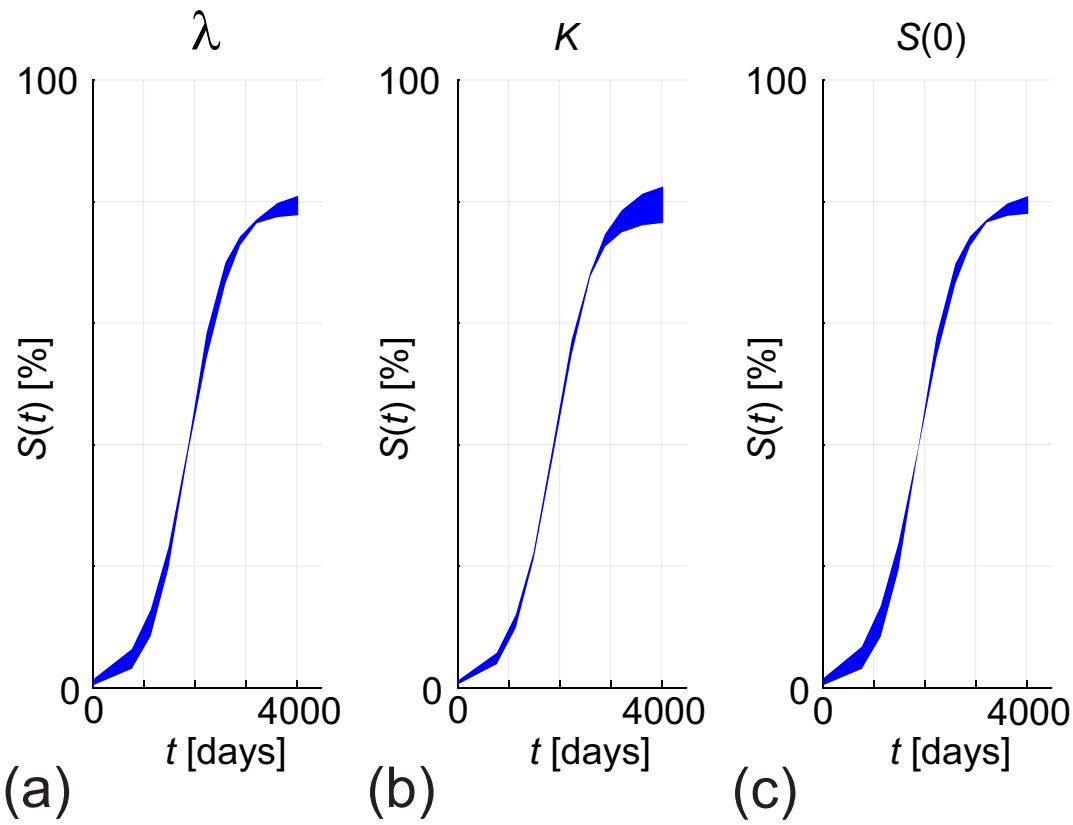

Figure 3: Parameter-wise profile predictions for modelling total hard coral coverage with (1). Results in (a)–(c) show the 95% parameter-wise profile predictions for λ, K and $S(0)$, respectively.

4.2. Two species model

Figure 4(a)–(b) compares the solution of Equations (2)–(3) evaluated at $\hat{\theta}$ with the data Figure 4(c) compares $S(t) = C_1(t) + C_2(t)$ with the data describing the evolution of the total hard corals. Univariate profiles are given for each parameter in Figure 4(d)–(j); again, in each case, the profiles are regularly-shaped with clearly defined peaks, suggesting that this more complicated model is also practically identifiable with the data. Asymptotic confidence intervals for each parameter are determined by interpolating the profiles at the threshold of $\ell = -1.92$ and reported in the caption of Figure 4.

With the univariate profiles in Figure 4, we construct parameter-wise profile predictions, shown in Figure 5, for each parameter in Equations (2)–(3). In this case we can plot prediction intervals for $C_1(t)$, $C_2(t)$ and for $S(t) = C_1(t) + C_2(t)$ as a function of each parameter in the process model. Again, these results are intuitively reasonable since we see, for example, that the prediction interval for λ_2 has a more significant influence on the prediction interval for $C_2(t)$ than $C_1(t)$, and that the prediction interval associated with K has an influence on the late-time behaviour of $C_1(t)$, $C_2(t)$ and $S(t)$ and a negligible influence on the early time prediction intervals for any of these three quantities.

4.3. Comparing the single- and two species modelling approaches

With our profile likelihood results constructed for Equation (1) in Figure 2 and for Equations (2)–(3) in Figure 4 we are able to make several comparisons between the two modelling approaches. For example, we can simply compare parameter point and interval estimates. To illustrate, for Equation (1) we have a point estimate $\bar{\lambda} = 0.0025$ and interval estimate $[0.0021, 0.0030]$, whereas for Equations (2)–(3) we have point estimates $\bar{\lambda_1} = 0.0029$ and $\bar{\lambda_2} = 0.0032$ and interval estimates $[0.0025, 0.0034]$ and $[0.0024, 0.0039]$ for λ_1 and λ_2, respectively. This suggests that the growth rates in the two models are largely statistically indistinguishable, with slightly greater uncertainty in λ_2 in the coupled model.

Comparing the uncertainty, as measured by the widths of confidence intervals, associated with our more complex model can inform our opinion about which modelling approach is preferred [42]. For example, with the two-species model we have $\hat{K} = 79\,\%$ which is close to the MLE for the single species model, but we can also quantitatively compare the widths of the confidence intervals. Here for the two species model we have $\hat{K} = 79 \in [76, 82]$, whereas for the single species model we have $\hat{K} = 80 \in [76, 83]$, from which we can conclude that our estimate of the carrying capacity density, and the uncertainty in that estimate is not very different between the two modelling approaches despite significant differences between the two modelling approaches. However, instead of just comparing models via uncertainties in parameter estimates, it is perhaps more natural in many applications to convert these to associated predictive intervals for a predictive quantity of interest such as $S(t)$, given in Figure 6. This figure shows the union of the various parameter-wise prediction intervals for each model. Here we see, despite the differences in both models, in this application the prediction intervals are not strikingly different. Thus the overall uncertainty in one of the key quantities of interest is largely the same for both models.

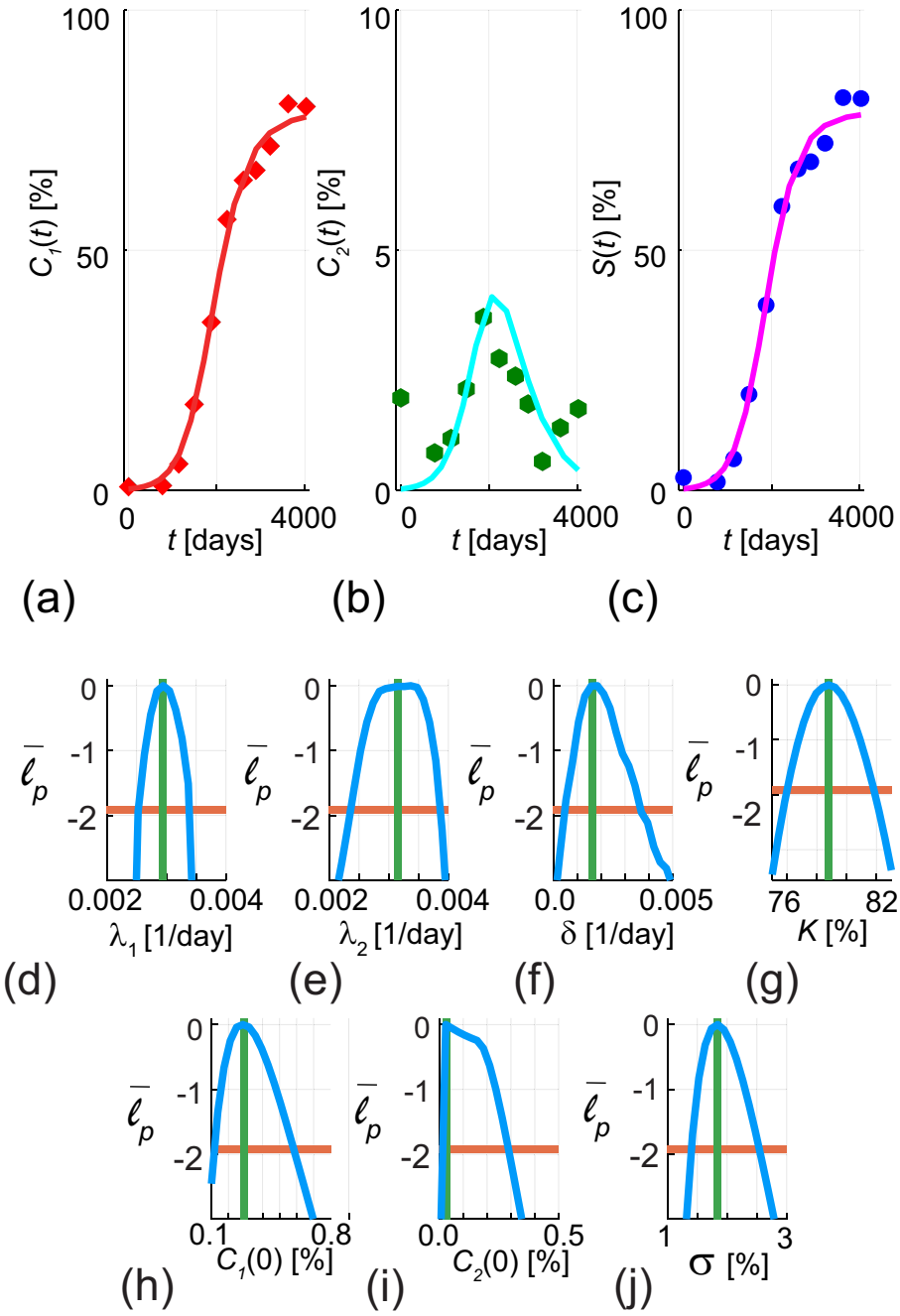

Figure 4: (a)–(b) Comparison of data and the solution of (2)–(3) for $C_1(t)$ and $C_2(t)$, respectively, evaluated at the MLE, $\hat{\theta} = (\lambda_1, \lambda_2, \delta, K, C_1(0), C_2(0), \sigma) = (0.0029, 0.0032, 0.0016, 79, 0.29, 0.03, 1.8)$. Results in (c) compare data and the solution of (2)–(3) in terms of the total hard coral coverage, $S(t)$, evaluated at $\hat{\theta}$. Results in (a) compare the data (red diamonds) with the MLE solution (solid red curve), results in (b) compare the data (green dots) with the MLE solution (solid cyan curve). results in (c) compare the data (blue dots) with the MLE solution (solid pink curve). (d) Univariate profile for λ_1 gives $\hat{\lambda}_1 = 0.0029 \in [0.0025, 0.0034]$. (e) Univariate profile for λ_2 gives $\hat{\lambda}_2 = 0.0032 \in [0.0024, 0.0039]$. (f) Univariate profile for δ gives $\hat{\delta} = 0.0016 \in [0.0005, 0.0037]$. (g) Univariate profile for K gives $\hat{K} = 79 \in [76, 82]$. (h) Univariate profile for $C_1(0)$ gives $\hat{C}_1(0) = 0.29 \in [0.12, 0.58]$. (i) Univariate profile for $C_2(0)$ gives $\hat{C}_2(0) = 0.03 \in [0.01, 0.29]$. (j) Univariate profile for σ gives $\hat{\sigma} = 1.8 \in [1.4, 2.5]$. Note that $C_2(t)$ is plotted on a different vertical scale than $C_1(t)$ or $S(t)$. Each profile in (d)–(j) is shown in solid blue; the MLE as a vertical green line, and the asymptotic log-likelihood threshold as a horizontal orange line.

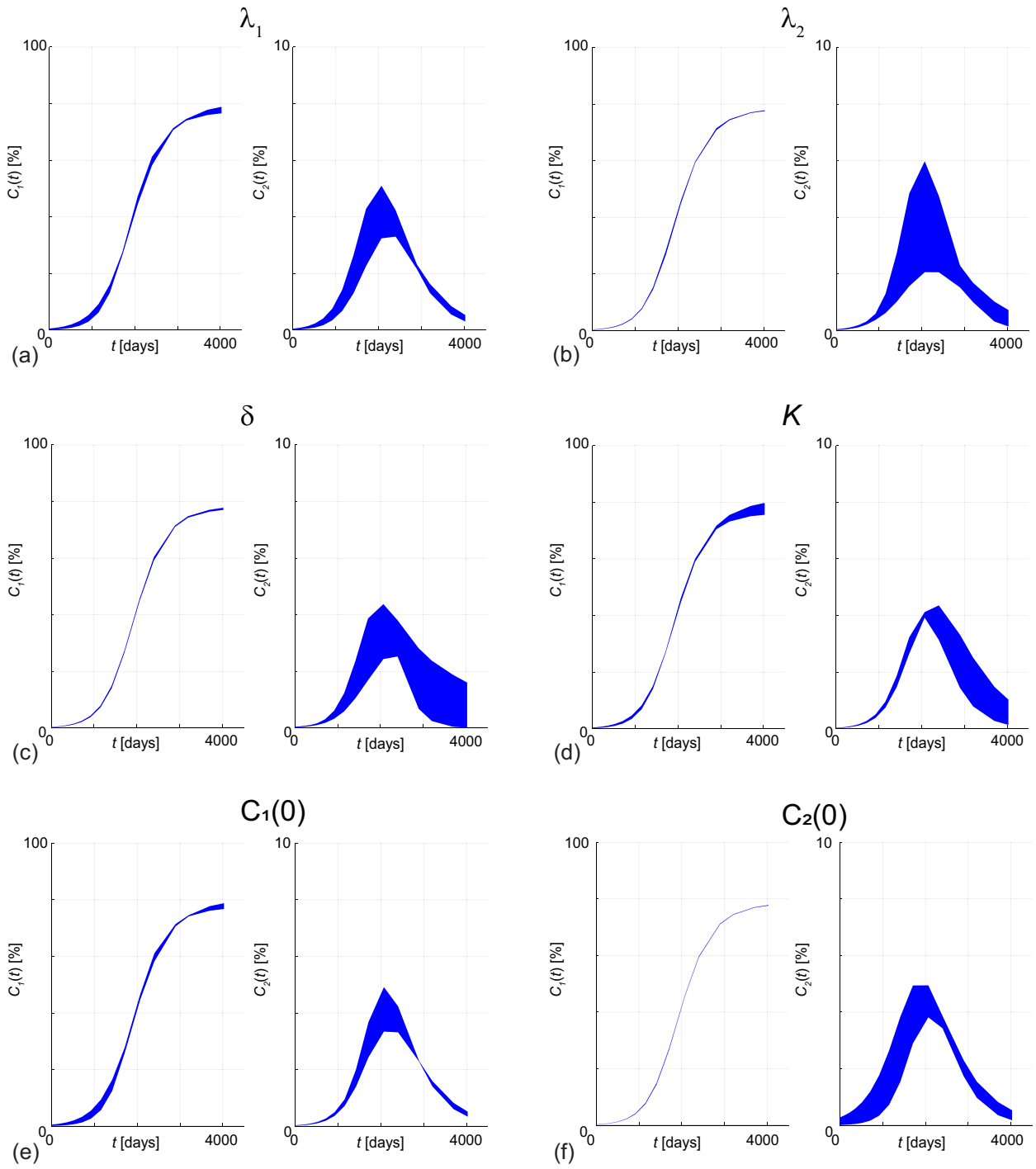

Figure 5: Parameter-wise profile predictions for modelling coral coverage with (2)–(3) in terms of $C_1(t)$, $C_2(t)$ and $S(t) = C_1(t) + C_2(t)$. Results in (a)–(f) show the 95% parameter-wise profile predictions for λ_1, λ_2, δ, K, $C_1(0)$ and $C_2(0)$, respectively. Note that $C_2(t)$ is plotted on a different vertical scale than $C_1(t)$ or $S(t)$.

There are many ways that we can interpret these results. If, for example, the main interest in modelling the coral regrowth is simply to predict the total coral coverage, then our comparisons in Figure 6 suggest that working with Equation (1) may be preferable since the prediction interval for the quantity of interest is not very different between the simpler model and the more complicated model, and of course working with the simpler model has pragmatic advantages. Alternatively, if the aim of the modelling exercise is to use a mathematical model to highlight the distinction between the regrowth of *Acroporidae* and other hard corals in the community, then the comparison in Figure 6 could be used to indicate that working with Equations (2)–(3) is entirely reasonable. For example, despite the relative complexity of the coupled model, working with the coupled model does not lead to significantly wider prediction intervals for the quantity in common with the simple model, although it does provide intervals for the separate components.

As we stated in Section 4.2, parameter-wise prediction intervals are approximate in the sense that the transformation of confidence intervals for ψ into confidence intervals for the predictive quantity is not exact owing to non-trivial dependencies between nuisance parameters. For simple models, such as Equation (1), it is straightforward to explore the impact of this approximation by constructing prediction intervals from the full likelihood and comparing these with the union of parameter-wise prediction intervals. Additional results in the Appendix for Equation (1) show the union of parameter-wise intervals does indeed lead to smaller uncertainty estimates than evaluating the full likelihood directly, although the differences are relatively small in this case and the increase in computational effort for the full likelihood is significant.

In addition to comparing the single species and two species models through the lens of profile likelihoods and parameter wise profile predictions, it is also instructive to compare our approach with more traditional information criteria based analysis [1]. One standard way to compare models is using the Akaike Information Criteria (AIC) [1]

$$\mathrm{AIC} = -2 \sup_{\theta} \ell(\theta|y_i) + 2k, \tag{13}$$

where k is the dimension of θ. Given a set of candidate models describing the same data, the preferred model is the one with the minimum AIC [1]. This kind of analysis rewards goodness of fit (as measured by the log-likelihood function) and penalises model complexity (as measured by the dimensionality of θ). This approach is very easy to implement and very popular and enables us to quantitatively compare two different models of the same data by ranking them in terms of AIC [48]. However, while information criteria-based measures describe the tradeoff between the goodness-of-fit and the dimensionality of the parameter space, estimates of AIC do not provide insight into the shape of the likelihood function, thereby providing no insight into the practical identifiability of the parameters. This is important because, in general, the curvature of the likelihood is related to inferential precision and so our approach provides additional insight by combining information about the goodness of fit through of the model (indicated by the MLE), the shape of the likelihood function (indicated by the width of the profile), as well as the width of the prediction intervals.

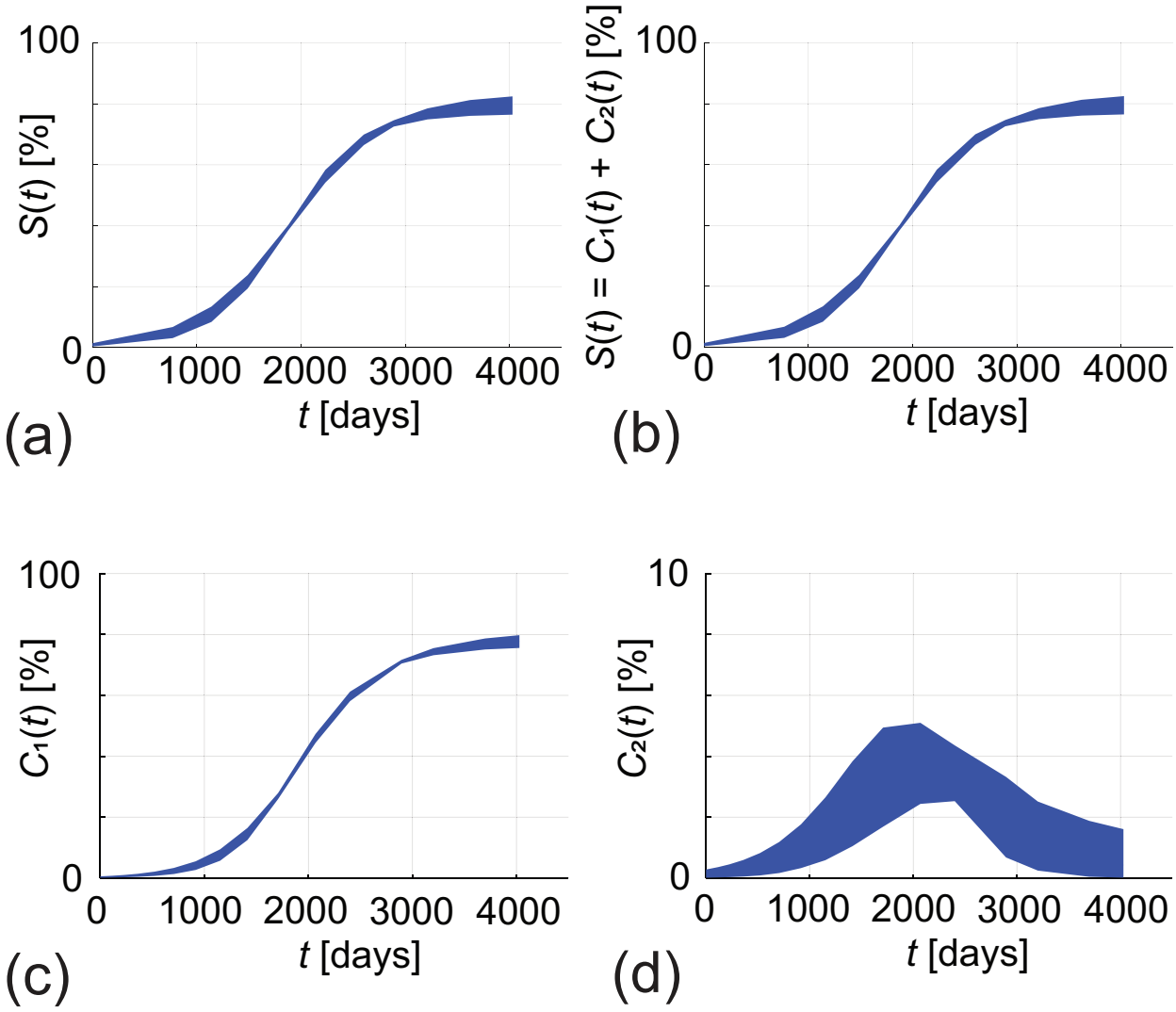

Figure 6: (a)–(b) gives the union of the parameter-wise profile predictions using (1) and (2)–(3) for $S(t)$, respectively. The prediction intervals in (a)–(b) are formed by taking the union of the prediction intervals in Figure 3 and Figure 5, respectively. The prediction intervals in (c)–(d) are formed by taking the union of the prediction intervals in Figure 5 for $C_1(t)$ and $C_2(t)$, respectively. Note that $C_2(t)$ is plotted on a different vertical scale than $C_1(t)$ or $S(t)$.

5. Conclusions and Outlook

In this work we provide a straightforward means of comparing mathematical models of population dynamics, in particular we use a case study describing coral reef regrowth after some disturbance. Our focus is to compare models and data using two different approaches. First, we study the time evolution of the total hard coral cover, $S(t)$. Second, we divide up the total population of hard corals into two groups, the dominant hard coral group *Acroporidae*, $C_1(t)$, and all other hard corals lumped together, $C_2(t)$, such that $S(t) = C_1(t) + C_2(t)$. Our modelling philosophy is always to use a minimal model where possible, so in the first approach where we make no distinction between the different types of hard corals, we model the total hard coral cover, $S(t)$, using a classic logistic growth model. In the second approach, where we explicitly model the regrowth of both types of hard corals, we use a system of coupled ODEs. This system involves logistic growth-like terms for each type of hard coral, together with a competition term describing how the dominant hard coral group competes with the remaining hard corals. Both of these ODE-based models are structurally identifiable, and univariate profile likelihood analysis confirms that, with our data, both the simple single species logistic model and the more complicated two-species models are practically identifiable. This is an important result since it is unclear when we move from the simple four parameter logistic model to the more complicated seven parameter coupled model whether the parameters in the more complicated model are identifiable. To provide more insight into the performance of both modelling approaches we construct and compare parameter-wise prediction intervals for both approaches. Our approach to constructing parameter-wise prediction intervals is novel and computationally straightforward, and reveals how each parameter impacts the predictions of the mathematical models. Understanding how model prediction is impacted by variability in each parameter is insightful since this gives us the ability to understand how uncertainty in each parameter contributes to the overall variability in the quantity of interest. Our results show that both modelling approaches lead to similar prediction outcomes. Therefore, choosing which modelling approach suits our needs depends strongly upon the aim of the modelling study; if the aim of the modelling study is to understand the total hard coral cover then it is entirely reasonable to work with the simpler single-species logistic model for pragmatic reasons. In contrast, if the aim of the modelling study is to understand interactions of different coral groups within the total population of hard corals, then working with the more complicated model does not lead to any parameter identifiability issues, nor is there any significant difference in the predictive capability of either modelling approach.

One of the aims of this work is to produce a fairly straightforward set of algorithms that can be used to guide model choice more broadly in the fields of mathematical biology and mathematical ecology. Model choice remains an on-going challenge in both mathematical biology and mathematical ecology, in part because data collection is not standardised and in part because there are many competing models available to describe similar phenomena [16, 37]. Within the context of modelling coral regrowth there are many other modelling approaches that one could take. For example, in this study we have compared working with total coral cover, $S(t)$, with a more detailed approach where we divide hard corals into two groups: $C_1(t)$ and $C_2(t)$. Another approach would be to divide up the hard corals into

more than two groups, giving $C_j(t)$ for $j = 1, 2, 3, \ldots J$, and to then work with a system of J coupled ODEs to describe the evolution of the community of hard corals. In this work we focus on setting $J = 2$ because it is clear from the field data in Figure 1(c) that the total hard corals is dominated by *Acroporidae*, and so it is pragmatic to set $J = 2$ so we capture the evolution of *Acroporidae* and the evolution of all other hard coral types. In a different scenario, where no single type of hard coral dominates the total population, we might choose to work with $J > 2$, and in this case the tools presented here could be applied in a very similar way. Another extension would be to maintain working with $J = 2$ groups of hard coral in the more complicated mathematical model but to also include an additional competition term,

$$\frac{\mathrm{d}C_1(t)}{\mathrm{d}t} = \lambda_1 C_1(t)\left[1 - \frac{S(t)}{K}\right] - \delta_1 C_1(t)\frac{C_2(t)}{K}, \tag{14}$$

$$\frac{\mathrm{d}C_2(t)}{\mathrm{d}t} = \lambda_2 C_2(t)\left[1 - \frac{S(t)}{K}\right] - \delta_2 C_2(t)\frac{C_1(t)}{K}. \tag{15}$$

This extended model is identical to Equations (2)–(3) with the exception that now we explicitly include the possibility of the other hard corals competing with the *Acroporidae* through the parameter δ_2, whereas in Equations (2)–(3) we implicitly assume that $\delta_2 = 0$ because there is no evidence of $C_1(t)$ decreasing with time in our data. This extended two-species model could be applied to our data using the same methodology and computational tools that we have developed. Of course, this particular extension is just one of many kinds of modelling extensions that we could analyse with our framework. Another extension would be to consider separate carrying capacity densities K_1 and K_2 for each coral type C_1 and C_2, respectively. While, in general, it would be reasonable to assume that different coral types might grow to reach different carrying capacity densities, here in our data we see that $C_1(t)$ dominates at late time with $C_2(t) \to 0$ for large t and so our data does not characterise late time maximum density behaviour for $C_2(t)$, and our previous experience in estimating late-time high-density parameters using early-time, low-density data leads to identifiability issues [47] so we do not pursue this extension here. Another feature of our modelling approach is that we focus on deterministic ODE models with a separate noise model. While this is an important class of widely-used models, other approaches such as working with stochastic process models are also of high interest for different applications [40] and an interesting extension of the current work would be to extend the current framework so that we can deal with stochastic process models. An intermediate approach would be to use a deterministic ODE model but correct the noise model to account for missing stochastic process model effects [41].

In addition to simply applying the tools developed here to different mathematical models, the current work lays the foundation for future theoretical developments. For example, results in the Appendix compare predictions for the simple single-species logistic model, Equation (1), obtained by taking the union of our parameter-wise prediction intervals with the prediction intervals obtained by evaluating the likelihood function directly. While working with a uniform discretisation of the parameter space to evaluate the full likelihood is reasonable for a simple model, this approach rapidly becomes computationally intractable for more complicated models like Equations (2)–(3). This is precisely the scenario where our parameter-wise prediction interval approach is advantageous. Nonetheless, our re-

sults for the simple model indicate that the union of the parameter-wise prediction intervals leads to slightly narrowed prediction intervals relative to evaluating the likelihood directly. This observation suggests further investigation is warranted and one approach would be to construct bivariate and higher order profiles, and take the union of a series of predictions made across these bivariate and higher order profiles [4]. While this approach would incur an increased computational cost over working with univariate profiles only, it would still be more efficient compared to evaluating the likelihood directly. We intend to explore the balance of computation time and accuracy in future work.

Appendix: Comparing predictions from the full and profile likelihood

In this appendix we construct prediction intervals for $S(t)$ using Equation (1), noting that for the simple model the log-likelihood function, $\hat{\ell}(\theta \mid y^{\mathrm{o}}_{1:I,1:J})$ depends on just three parameters in the process model: λ, K and $S(0)$. We evaluate the full log-likelihood function by defining a three-dimensional rectangular prism in $\lambda, K, S(0)$ parameter space such that: $0.004 \leq \lambda \leq 0.001$, $75 \leq K \leq 85$ and $0 \leq S(0) \leq 1$. We will justify the choice of these limits later. We then take a $30 \times 30 \times 30$ uniform discretization of this rectangular prism and evaluate $\hat{\ell}(\theta \mid y^{\mathrm{o}}_{1:I,1:J})$ at each of the 27,000 points in the discretised parameter space. For those mesh points where $\hat{\ell}(\theta \mid y^{\mathrm{o}}_{1:I,1:J}) \geq -3.91$ (corresponding to a 95% asymptotic confidence interval [36]) we plot the evolution of $S(t)$ in Figure 7(a). Note that we chose the extent of the rectangular prism in $\lambda, K, S(0)$ so that it completely enclosed the region where $\hat{\ell}(\theta \mid y^{\mathrm{o}}_{1:I,1:J}) \geq -3.91$.

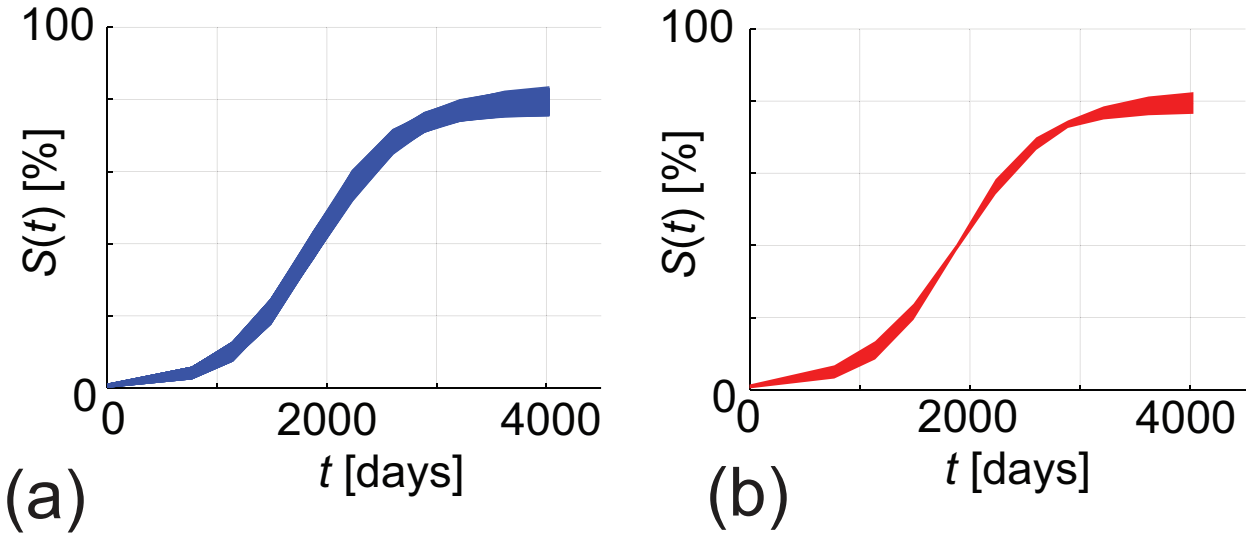

Figure 7: (a) prediction interval for the full likelihood (blue curves). (b) approximate prediction interval constructed from profile likelihoods (red curves).

The prediction intervals in Figure 7(a) for $S(t)$ are generated by evaluating the log-likelihood directly. Visually comparing this collection of predictions with the union of the corresponding three parameter-wise prediction intervals in Figure 7(b) indicates that evaluating the full likelihood leads to similar, albeit slightly more conservative (broader), prediction intervals than those obtained from the parameter-wise approach. Therefore, in this case the parameter-wise intervals are slightly narrower than evaluating the full likelihood. However, the parameter-wise approach is both insightful (in terms of illustrating the impacts of each parameter) and computationally efficient since the computational effort required to compute the prediction intervals scales linearly with the dimension of the parameter space. For example here the prediction intervals are generated simply by considering 3×40 parameter combinations, whereas uniformly evaluating the log-likelihood function required $30 \times 30 \times 30$ parameter combinations. While it is possible to

use more sophisticated evaluation methods (e.g. Latin Hypercube Sampling) to work with the full likelihood function, this approach is likely infeasible for more complicated models where evaluating larger parameter spaces efficiently is a computational challenge. The parameter-wise approach, in contrast, remains computationally efficient even when the dimensionality of the full parameter space increases, since we rely on constructing the prediction intervals using a series of univariate profile likelihood functions only.

Acknowledgements MJS is supported by the Australian Research Council (DP200100177). SW and ES thank the Faculty of Science, Queensland University of Technology, for providing summer vacation research scholarships that supported this study. We thank three referees for helpful suggestions.

References

[1] Akaike H. 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control. 19, 716-723. (doi:10.1109/TAC.1974.1100705).

[2] Bjornstad JF. 1990. Predictive likelihood: A review. Statistical Science. 5, 242-254. (doi:10.1214/ss/1177012175).

[3] Browning AP, Warne DJ, Burrage K, Baker RE, Simpson MJ (2020). Identifiability analysis for stochastic differential equations models in systems biology. Journal of the Royal Society Interface. 17, 20200652. (doi:10.1098/rsif.2020.0652).

[4] Browning AP, Maclaren OJ, Buenzli PR, Lanaro M, Allenby MC, Woodruff MA, Simpson MJ (2021). Model-based data analysis of tissue growth in thin 3D printed scaffolds. Journal of Theoretical Biology. 528, 110852. (doi:10.1016/j.jtbi.2021.110852).

[5] Burnham KP, Anderson DR, Huyvaert KP. 2011. AIC model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons. Behavioral Ecology and Sociobiology. 65, 23-35. (doi:10.1007/s00265-010-1029-6).

[6] Campbell DA, Chkrebtii O. 2013. Maximum profile likelihood estimation of differential equation parameters through model based smoothing state estimate. Mathematical Biosciences. 246, 283-292. (doi:10.1016/j.mbs.2013.03.011).

[7] Casella G, Berger R. 2001. Statistical Inference. Belmont, CA: Duxbury.

[8] Chiş O, Banga JR, Balsa-Canto E. 2011. Structural identifiability of systems biology models: a critical comparison of methods. PLoS One. 6, e27755. (doi:10.1371/journal.pone.0027755).

[9] Chiş O, Banga JR, Balsa-Canto E. 2011. GenSSI: a software toolbox for structural identifiabilty analysis of biological models. Bioinformatics. 18, 2610–2611. (doi:10.1093/bioinformatics/btr431).

[10] Cole D. 2020. Parameter redundancy and identifiability. CRC Press.

[11] Cox DR. 2006. Principles of statistical inference. Cambridge University Press.

[12] Eisenberg MC, Hayashi MAL. 2014. Determining identifiable parameter combinations using subset profiling. Mathematical Biosciences. 256, 115–126. (doi:10.1016/j.mbs.2014.08.008).

[13] eAtlas: Largest GBR coral reef survey data repository. September 2022 eAtlas.

[14] Fröhlich F, Theis FJ, Hasenauer J. 2014. Uncertainty analysis for non-identifiable dynamical systems: Profile likelihoods, bootstrapping and more. International Conference on Computational Methods in Systems Biology. 61-72. Springer.

[15] Gnerucci A, Faraoni P, Sereni E, Ranaldi F. 2020. Scratch assay microscopy. A reactio-diffusion approach for common instruments and data. Mathematical Biosciences. 330, 108482. (doi: 10.1016/j.mbs.2020.108482).

[16] Gerlee P. 2013. The model muddle: In search of tumor growth laws. Cancer Research. 73, 2407-2411. (doi: 10.1158/0008-5472.CAN-12-4355).

[17] Hines KE, Middendorf TR, Aldrich RW. 2014. Determination of parameter identifiability in nonlinear biophysical models: A Bayesian approach. Journal of General Physiology. 143, 401. (doi: 10.1085/jgp.201311116).

[18] Hughes TP, Kerry JT, Baird AH, Connolly SR, Chase TJ, Dietzel A, Hill T, Hoey AS, Hoogenboom MO, Jacobson M, Kerswell A, Madin JS, Mieog A, Paley AS, Pratchett MS, Torda G, Woods RM. 2019. Global warming impairs stock-recruitment dynamics of corals. Nature. 568, 387-390. (doi: 10.1038/s41586-019-1081-y).

[19] Jin W, McCue SW, Simpson MJ. 2018. Extended logistic growth model for heterogeneous populations. Journal of Theoretical Biology. 445, 51–61. (doi:10.1016/j.jtbi.2018.02.027).

[20] Johnson, JB, Omland, KS. 2004. Model selection in ecology and evolution. Trends in Ecology & Evolution. 19(2), 101–108. (doi: 10.1016/j.tree.2003.10.013).

[21] Johnson SG. 2022. The NLopt module for Julia. Retrieved September 2022 NLopt.

[22] Kreutz C, Raue A, Timmer J. 2012. Likelihood based observability analysis and confidence intervals for predictions of dynamics models. BMC Systems Biology. 6, 120. (doi:10.1186/1752-0509-6-120).

[23] Lambert B, MacLean AL, Fletcher AG, Combes AN, Little MH, Byrne HM. 2018. Bayesian inference of agent-based models: a tool for studying kidney branching morphogenesis. Journal of Mathematical Biology. 76, 1673-1697. (doi:10.1007/s00285-018-1208-z).

[24] Ligon TS, Frölich F, Chiş O, Banga JR, Balsa-Canto E, Hasenauer J. 2018. GenSSI 2.0: multi-experimental structural identifiability analysis of SBML models. Bioinformatics. 34, 1421–1423. (doi:10.1093/bioinformatics/btx735).

[25] Maclaren OJ, Nicholson R. 2019. What can be estimated? Identifiabiliy, estimability, casual infernce and ill-posed inverse problems. arXiv. https://arxiv.org/abs/1904.02826.

[26] Meshkat N, Eisenberg M, DiStefano III JJ. 2009. An algorithm for finding globally identifiable parameter combinations of nonlinear ODE models using Gröbner Bases. 222, 61-72. (doi: 10.1016/j.mbs.2009.08.010).

[27] Meshkat N, Sullivant S, Eisenberg M. 2018. Identifiability results for several classes of linear compartment models. Bulletin of Mathematical Biology. 8, 1620-1651. (doi: 10.1007/s11538-015-0098-0).

[28] Murray JD. 2002. Mathematical Biology I: An Introduction. Springer.

[29] Murphy RJ, Maclaren OJ, Calabrese AR, Thomas PB, Warne DJ, Williams ED, Simpson MJ (2022). Computationally efficient framework for diagnosing, understanding, and predicting biphasic population growth. To appear, Journal of the Royal Society Interface. www.biorxiv.org/content/10.1101/2022.07.27.501797v1.

[30] Pace L, Salvan A. 1997. Principles of Statistical Inference from a Neo-Fisherian Perspective. In: Advanced Series on Statistical Science and Applied Probability, vol. 4. World Scientific, Singapore.

[31] Pawitan Y. 2001. In all likelihood: statistical modelling and inference using likelihood. Oxford University Press.

[32] Rackauckas C, Nie Q. 2017. DifferentialEquations.jl - a performant and feature-rich ecosystem for solving differential equations in Julia. Journal of Open Research Software. 5, 15. (doi: 10.5334/jors.151).

[33] Raue A, Kreutz C, Maiwald T, Bachmann J, Schilling M, Klingmüller U, Timmer J. 2009. Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. Bioinformatics. 25, 1923-1929. (doi: 10.1093/bioinformatics/btp358).

[34] Raue A, Kreutz C, Theis FJ, Timmer J. 2013. Joining forces of Bayesian and frequentist methodology: a study for inference in the presence of non-identifiability. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences. 371, 20110544. (doi: 10.1098/rsta.2011.0544).

[35] Raue A, Karlsson J, Saccomani MP, Jirstrand M, Timmer J. 2014. Comparison of approaches for parameter ifentifiability analysis of biological systems. Bioinformatics. 30, 1440-1448. (doi: 10.1093/bioinformatics/btu006).

[36] Royston P. 2007. Profile likelihood for estimation and confidence intervals. The Stata Journal. 7, 376-387. (doi: 10.1177/1536867X0700700305).

[37] Sarapata EA, de Pillis LG. 2014. A comparison and catalog of intrinsic tumor growth models. Bulletin of Mathematical Biology. 76, 2010-2024. (doi: 10.1007/s11538-014-9986-y).

[38] Siekmann I, Sneyd J, Crampin EJ. 2012. MCMC can detect nonidentifiable models. Biophysical Journal. 103, 2275-2286. (doi: 10.1016/j.bpj.2012.10.024).

[39] Simpson MJ, Baker RE, Vittadello ST, Maclaren OJ. 2020. Parameter identifiability analysis for spatiotemporal models of cell invasion. Journal of the Royal Society Interface. 17, 20200055. (doi: 10.1098/rsif.2020.0055).

[40] Simpson MJ, Browning AP, Drovandi C, Carr EJ, Maclaren OJ, Baker RE (2021). Profile likelihood analysis for a stochastic model of diffusion in heterogeneous media. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences. 477, 20210214. (doi:10.1098/rspa.2021.0214).

[41] Simpson MJ, Baker RE, Buenzli PR, Nicholson R, Maclaren, OJ. 2022. Reliable and efficient parameter estimation using approximate continuum limit descriptions of stochastic models. Journal of Theoretical Biology. 549, 111201. (doi: 10.1016/j.jtbi.2022.111201).

[42] Simpson MJ, Browning AP, Warne DJ, Maclaren OJ, Baker RE. 2022. Parameter identifiability and model selection for sigmoid population growth models. Journal of Theoretical Biology. 535, 110998. (doi:10.1016/j.jtbi.2021.110998).

[43] Shmueli G. 2010. To explain or to predict? Statistical Science. 25(3), 289-310.

[44] Toni T, Welch D, Strelkowa N, Ipsen A, Stumpf MPH. 2009. Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. Journal of the Royal Society Interface. 6, 187-202. (doi:10.1098/rsif.2008.0172).

[45] Tsoularis A, Wallace J. 2002. Analysis of logistic growth models. Mathematical Biosciences. 179, 21-55. (doi:10.1016/S0025-5564(02)00096-2).

[46] Villaverde AF, Tsiantis N, Banga JR. 2019. Full observability and estimation of unknown inputs, states and parameters of nonlinear biological models. Journal of the Royal Society Interface. 16, 20190043. (doi:10.1098/rsif.2019.0043).

[47] Warne DJ, Baker RE, Simpson MJ (2017). Optimal quantification of contact inhibition in cell populations. Biophysical Journal. 113, 1920-1934. (doi:10.1016/j.bpj.2017.09.016).

[48] Warne DJ, Baker RE, Simpson MJ. 2019. Using experimental data and information criteria to guide model selection for reaction-diffusion problems in mathematical biology. Bulletin of Mathematical Biology. 81, 1760-1804. (doi:10.1007/s11538-019-00589-x).

[49] Warne DJ, Crossman KA, Jin W, Mengersen KL, Osborne K, Simpson MJ, Thompson AA, Wu P, Ortiz J-C. 2022. Identification of two-phase recovery in hard corals across the Great Barrier Reef. Journal of Applied Ecology. 59, 153-164. (doi:10.1111/1365-2664.14039).

[50] Wieland F-G, Hauber AL, Rosenblatt M, Tönsing C, Timmer J. 2021. On structural and practical identifiability. Current Opinion in Systems Biology. 25, 60-69. (doi:10.1016/j.coisb.2021.03.005).

[51] Wu D, Petousis-Harris H, Paynter J, Suresh V, Maclaren OJ. 2022. Likelihood-based estimation and prediction for a measles outbreak in Samoa. arXiv. (https://arxiv.org/abs/2103.16058).