Supplementary Material: Parameter identifiability and model selection for sigmoid population growth models.

Matthew J Simpson[a,*], Alexander P Browning[a], David J Warne[a,b], Oliver J Maclaren[c], Ruth E Baker[d]

[a]*School of Mathematical Sciences, Queensland University of Technology (QUT), Brisbane, Australia.*
[b]*Centre for Data Science, QUT, Brisbane, Australia.*
[c]*Department of Engineering Science, University of Auckland, Auckland 1142, New Zealand.*
[d]*Mathematical Institute, University of Oxford, Oxford, UK.*

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

S1. Additional data

To illustrate the generality of the methods outlined in the main manuscript we analyse a second data set describing coral re-growth near Lady Musgrave island, Australia [1]. Figure S1 compares the experimental data from Lady Musgrave Island (site 1, used in the main manuscript) with a second data set describing coral re-growth at Lady Musgrave island (site 3). Here we see that the two data sets cover the same period of time after disturbance, and that the re-growth at site 3 leads to a slightly reduced coral cover than at site 1. Simply visualising these data indicates that the fluctuations in the data at site 3 are greater than the fluctuations in the data at site 1, and we will discuss this point later. We now seek to describe the coral re-growth data at site 3 using the same tools from the main manuscript.

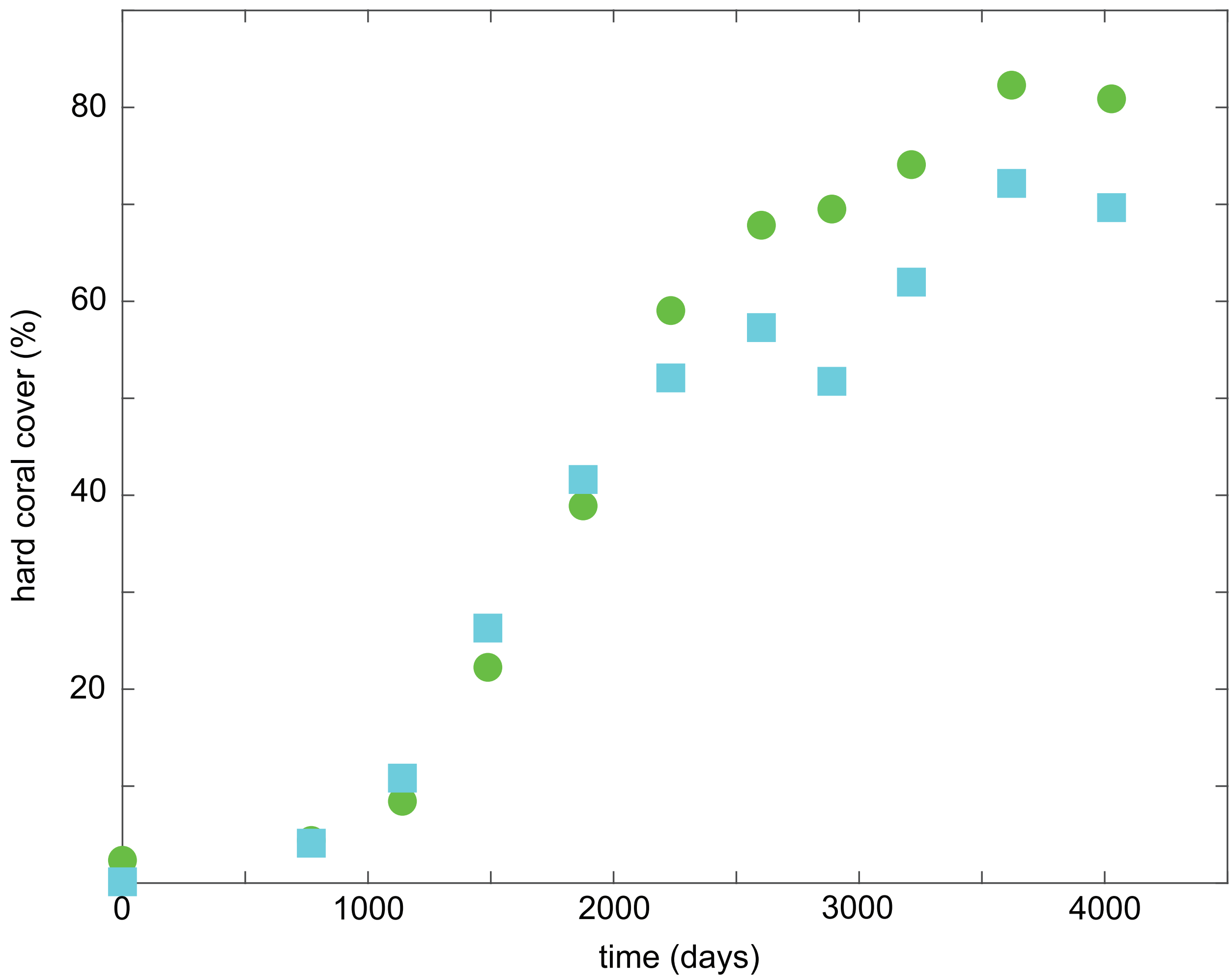

Figure S1: Experimental data showing hard coral re-growth at Lady Musgrave island site 1 (green dots) and site 3 (blue squares) [1].

S1.1. Parameter estimation and model checks

Figure S2(a)–(b) compares the data with the logistic and Gompertz models evaluated at the MLE. Scaled residuals are given in Figure S2(c)–(d), and Table S1 summarises $\widehat{\theta}$ for each model. Visual checks and the Durbin-Watson test suggest that the scaled residuals are relatively uncorrelated in all cases. Our estimates of $\widehat{K}$ for site 3 data are less than $\widehat{K}$ for site 1. This is consistent with our visual comparison of the two data sets in Figure S1. Our estimates of $\widehat{\sigma}$ for site 3 are greater than $\widehat{\sigma}$ for site 1 data, and this is consistent with our visual comparison of the fluctuations in the data in Figure S1.

Note that, for site 3 data, we are unable to work with the Richards' model since the numerical optimisation algorithm, fmincon, fails to converge when we attempt to estimate $\widehat{\theta}$. In this case, convergence issues with the numerical optimisation routine for the most complicated model provides very strong evidence that the Richards' model is unsuitable for this data due to practical identifiability constraints. However, at first glance, comparisons between the calibrated models and data in Figure S2 suggest that both the logistic and Gompertz models appear to provide a good match to the data, and we will now explore whether the same identifiability trends explored in the main document for site 1 are relevant for the site 3 data set.

Table S1: MLE parameter estimates with 95% confidence intervals are given in parentheses for data at Lady Musgrave Island site 3.

	Logistic ($m = 1$)	Gompertz ($m = 2$)
$\widehat{r_m}$	2.2×10^{-3} [1.5×10^{-3}, 3.5×10^{-3}]	1.4×10^{-3} [9.6×10^{-4}, 2.1×10^{-3}]
$\widehat{K}$	67 [61, 75]	70 [64, 80]
$\widehat{C(0)}$	1.3 [0.20, 4.1]	1.9×10^{-2} [1.0×10^{-6}, 5.0×10^{-2}]
$\widehat{\sigma}$	4.3 [3.0, 6.9]	3.6 [2.5, 5.9]

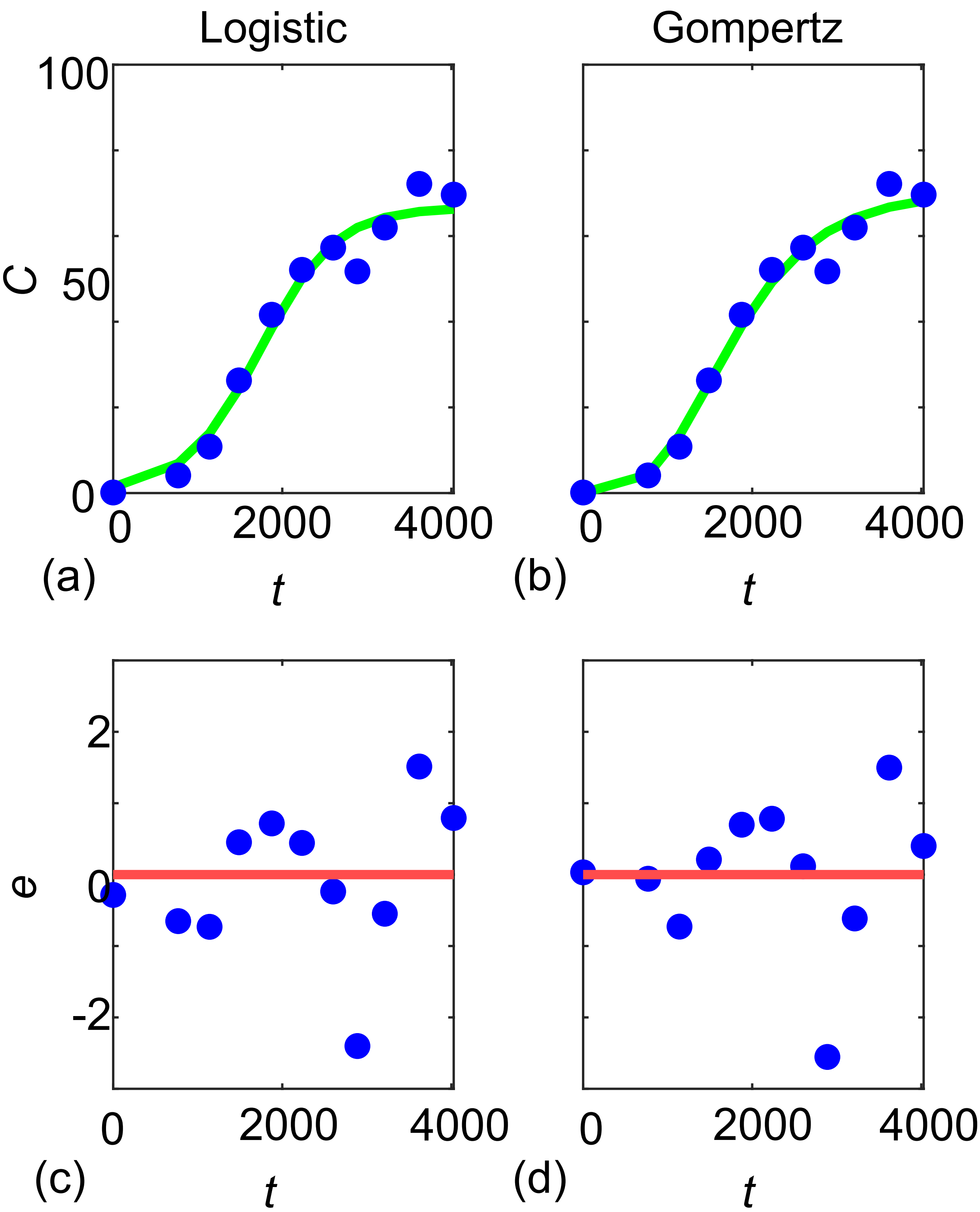

Figure S2: (a)-(c) Observed data (blue discs) superimposed with MLE solutions of the logistic, Gompertz and Richards' models. (d)-(f) scaled residuals for each model. The Durbin-Watson test yields $DW = 1.4$ ($p = 0.14$), and $DW = 1.7$ ($p = 0.30$), for the logistic and Gompertz models, respectively.

S1.2. Identifiability analysis

Figure S3 show the various univariate profiles for the logistic and Gompetz models with site 3 data. Each profile is superimposed with a vertical line at the MLE, and a horizontal line at -1.92 so we can visualise and calculate the width of the confidence intervals reported in Table S1. Just like the profiles for site 1 in the main document, results in Figure S3 for the logistic growth model at site 3 indicate that each parameter is identifiable with relatively narrow profiles (Table S1). In contrast, profiles shows that r_2 and K are identifiable for the Gompertz model, but again $C(0)$ shows signs of practical non-identifiability. In particular, the width of the confidence interval for $C(0)$ spans four orders of magnitude with the lower confidence bound determined by the $C(0) > 0$ constraint, again indicating that this may be at best one-sided identifiable. Therefore, just like the data at site 1 (main document), we conclude that the logistic growth model is preferable over the Gompertz and Richards' growth models because the univariate profile for each parameter is well–defined and the confidence interval is relatively narrow in each case. This result, is of particular interest in the context of modelling coral re-growth since this part of the literature routinely uses the Gompertz growth model without explicitly checking other options, such as the logistic growth model.

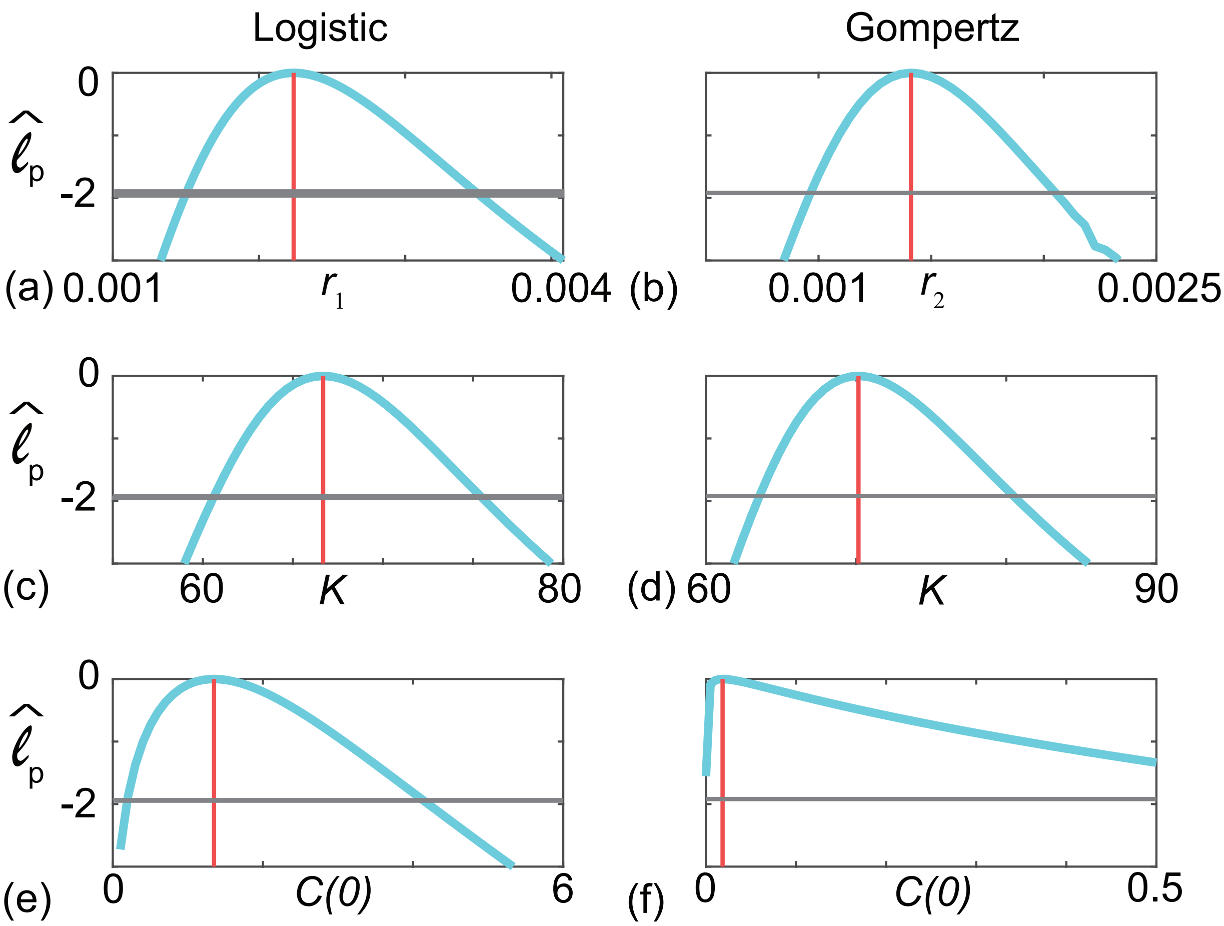

Figure S3: Univariate profile likelihoods for parameters in the logistic, Gompertz and Richards' model, as indicated. In each case the profile is shown (solid blue) together with the MLE (vertical red) and a horizontal line at -1.92.

References

[1] eAtlas: Largest GBR coral reef survey data repository. Retrieved November 2021 eAtlas.