Imperial College
London

# Comparison of Methodologies to Fit Microbial Population Growth Models

Sam Turner

March 2020

*Computational Method in Ecology and Evolution MSc*

3857 words

**Abstract**

The modelling of microbial growth is vital for developing food preservation measures to decrease food waste. By fitting a range of mechanistic and phenomenological models to a large dataset of 305 growth curves, I demonstrate that temperature has a major impact on microbial population dynamics. I also show that model fitting in log space increases the resolution of parameter estimates relative to linear space model fitting, and that different models with equivalent parameterisations can produce systematically biased parameter estimates. This results in recommendations for model fitting, and in particular how to carry out meta analyses where parameter values must be compared between sources.

# 1 Introduction

There is great interest in modelling the growth of microbial populations, such that the biology of microbial growth and its response to external variables can be understood [20]. Through this modelling, food shelf life can be predicted, and preservation methods can be optimised. In this project, I aim to answer four questions:

1. Which models best describe microbial population growth, and how does this vary with temperature?
2. How does choice of model affects parameter estimates?
3. What are the qualitative patterns of how temperature affects microbial growth?
4. Does the choice to fit models in linear or log space affect results?

I aimed to answer these questions by fitting a range of mechanistic and phenomenological models to a large dataset of 305 digitised growth curves. The mechanistic models each attempt to describe the shape of the bacterial growth curve with a mathematical relationship representing a real hypothesis about the dynamics of the microbial population [5], and therefore can be parameterised with meaningful biological quantities. This is in contrast to the phenomenological models, which are simply mathematical functions that we attempt to fit to the data, and can be used as a neutral comparison for the mechanistic models. By comparing the fit of the mechanistic models to each other, and to the phenomenological models, we can determine which underlying biological hypothesis has the best support from the data - and how this support varies with temperature.

I also examine how temperature affects the growth rate and lag time of bacterial populations, the dynamics of which play a central role in food spoilage. It is also important to determine whether the parameter estimates from various mechanistic models with analogous parameterisations vary significantly, as this may invalidate comparison of parameter values from different sources in meta-analyses.

## The Mechanistic Models

The mechanistic models can be split into three four-parameter models, and the three-parameter logistic model. The four-parameter models split the bacterial growth curve into a lag phase, exponential growth phase, and stationary phase, with the three-parameter logistic model discarding the lag phase.

**Logistic** The logistic model [16, 17] assumes that the microbial population immediately starts growing exponentially, with its growth rate falling towards zero as the population size approaches a carrying capacity. This is described by the logistic differential equation:

$$\frac{\mathrm{d}N}{\mathrm{d}x} = \mu_{max} N(1 - N)$$

Which has solutions:

$$N_t = \frac{N_0 N_{max} e^{\mu t}}{N_{max} + N_0(e^{\mu t} - 1)}$$

### 1.0.1 Four-parameter models

The four-parameter models parameterise the sigmoidal growth curve with minimum and maximum population sizes, maximum growth rate, and a time lag before the population starts growing - defined as the intercept between the tangent to the curve at the point of inflection, and the line $N = N_0$ (Figure 2 ). As such, the rate of growth and length of the lag phase can be independently varied (Figure 1 )

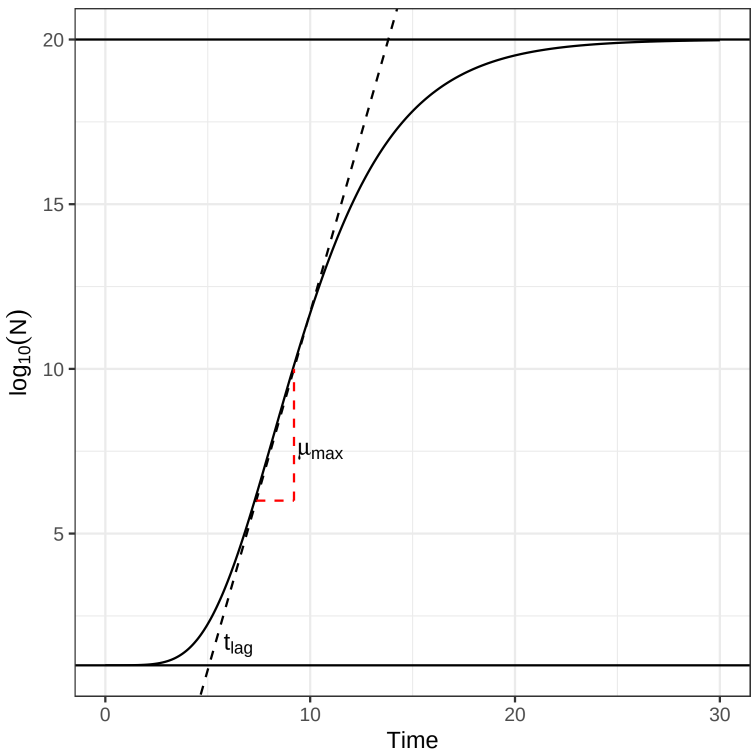

**Figure 1:** Illustation of $\mu_{max}$ and $t_{lag}$ parameters. $\mu_{max}$ is the maximum gradient of the growth curve, and $t_{lag}$ is the intercept between this tangent line and the line $N = N_0$

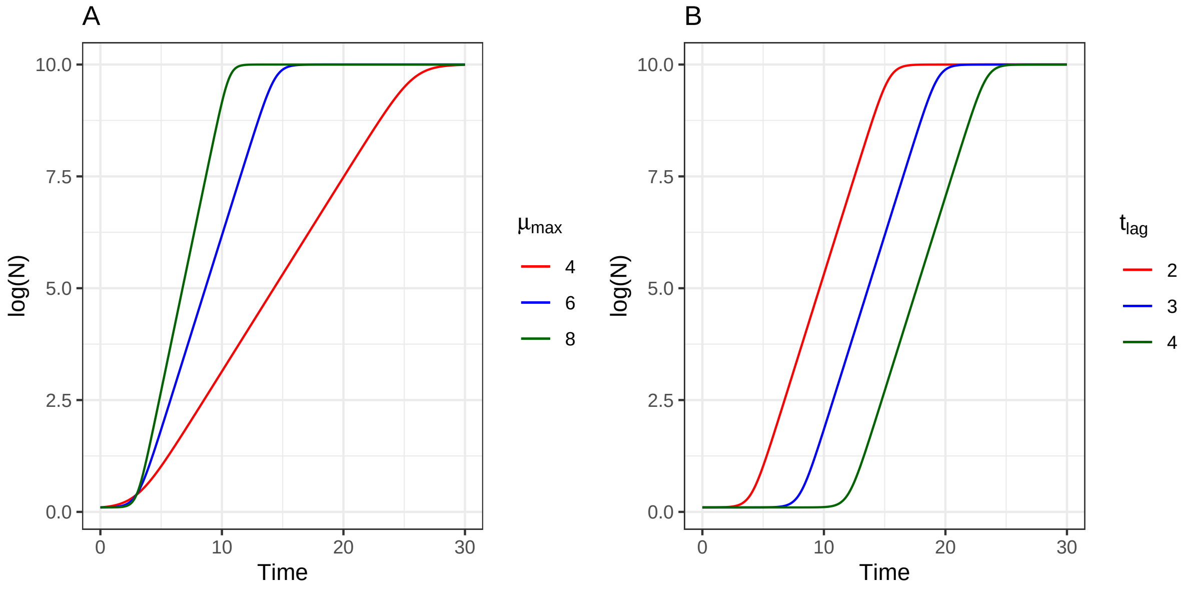

**Figure 2:** $\mu_{max}$ (A) and $t_{lag}$ (B) parameters can be varied independently to control the rate of microbial growth, and the duration of the lag phase.

**Buchanan** The Buchanan model [5] is the simplest implementation of this, assuming that the population has a growth rate of zero during the lag and stationary phases, with a period of exponential growth at a constant rate between this:

$$log(N_t) = \begin{cases} N_0 & \text{if } t \leq t_{lag} \\ N_0 + \mu_{max} \cdot (t - t_{lag}) & \text{if } t_{lag} \leq t \leq t_{max} \\ N_{max} & \text{if } t \geq t_{max} \end{cases}$$

**Gompertz** The Gompertz model[8, 20] similarly describes a sigmoidal growth curve, adding a gradual transition into and out of the exponential phase. The Gompertz model has long been popular for microbial population modelling, as it has empirically been found to fit well in a wide range of circumstances. However, it should be noted that its derivation does not have a mechanistic basis in population growth modelling [9], but rather in describing human mortality rates [8].

$$log_{10}(N_t) = N_0 + (N_{max} - N_0)e^{-e^{\frac{\mu_{max} \ \cdot \ e \ \cdot \ (t_{lag} - t)}{(N_{max} - N_0) \ \cdot \ log(10)} + 1}}$$

**Baranyi** The Baranyi model is an attempt to formulate a truly mechanistic model for microbial population growth [2, **?**, 20], explicitly modelling the internal physiological state and external environmental conditions of the cell. This mechanistic formulation produces a much more flexible and widely applicable model, which can be applied to model population growth curves in time-varying environmental conditions. Here, we use a version of the model in which conditions are homogenous, with the curvature parameter into the exponential phase $\nu$ set to $\mu_{max}$, and curvature parameter out of the exponential phase $m$ set at an empirical estimate of 1:

$$log_{10}(N_t) = N_{max} + log_{10}(\frac{-1 + e^{\mu_{max} \cdot t} + e^{\mu_{max} \cdot t_{lag}}}{-1 + e^{\mu_{max} \cdot t} + e^{\mu_{max} \cdot t_{lag}} \cdot 10^{N_{max} - N_0}})$$

## Phenomenological models

As controls, I decided to fit quadratic and cubic polynomials to the data:

**Quadratic** $N_t = at^2 + bt + c$

**Cubic** $N_t = at^3 + bt^2 + ct + d$

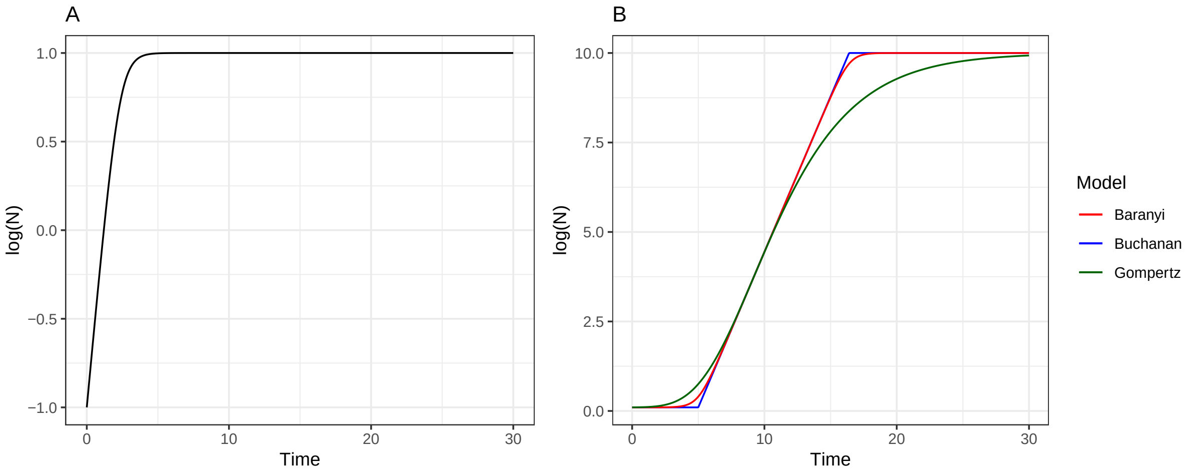

**Figure 3:** Demonstration of the shape of the 4 mechanistic models, plotted in log space. The logistic model (A) immediately starts growing at its maximum growth rate, whilst the three four-parameter models (B) have growth phases that transition into the exponential phase

# 2 Methods

## 2.1 Data

**Data sources** The provided microbial growth data contained 305 population size time series, digitised from graphs in 10 papers from the microbial growth literature. The data set contained time series from 47 different species, 18 media, and 17 temperatures, with population size measured in 4 different units. However, the data from one citation [13] clearly had incorrect time values, with times extending to 13,000 hours. I removed the data from this citation as there was not a clear correct transformation to produce the correct times, with time values appeared to be have multiplied by around 20 relative to the original paper. This left 287 time series for analysis.

**Log transformation** As explained below, I decided to fit the model to $log_{10}(N)$ as well as to linear data. In 5 /287 data sets, there were negative population counts due to numerical error in growth curve digitization. One value of approximately -668, which was removed, whilst other values were greater that -10$^{-2}$ - in these cases, I transformed increased population sizes for that time series by the minimum population size in the time series, discarding the minimum value of 0 before taking the logarithm.

**Time transformation** I transformed each time series such that the first data point was at $t = 0$ for fitting the four-parameter models, as model fits were clearly suboptimal without this transformation. Since $t_{lag}$ directly transforms the curve along the t-axis, $t_{lag}$ could be adjusted back to obtain model parameters which fit the untransformed time series.

## 2.2 Model Fitting

When fitting the models, I used a least squares approach, both to $\log_{10}$ transformed data, and to the original linear data (Figure **??**. This required reformulating the four-parameter models to predict linear population size, by exponentiating the right hand side of the function, and reformulating the logistic model to predict log population size, by taking the logarithm of the right hand side. As such, the parameters of the four-parameter models remained in log space, and the parameters of the Logistic model in linear space.

Principally, the difference between fitting in linear space and log space is whether we are concerned with absolute values or relative values. Log space expands variation at small population sizes, and therefore more strongly highlights the lag phase. This makes estimation of $\mathrm{t_{lag}}$ more robust, and means that models such as the logistic model are more strongly penalised for mischaracterising small population sizes early in the time series due to the absence of a lag phase. By contrast, in linear space model fit is strongly weighted towards large absolute population sizes. For food preservation purposes, this is undesirable, as the the lag phase is more interesting than behaviour near $\mathrm{N_{max}}$, as food is often spoiled long before this.

As such, I decided to use parameter estimates from the log-fitted models for my main analysis (although linear fit results are presented in the Supplementary Figures), also choosing not to linearise residuals, as this resulted in non-sensical fits (demonstrated in Figure S1). These fits can be explained by the fact that, in general:

$$log(f(x)) \neq f(log(x))$$

This means that minimising errors in log space does not result in minimised errors in linear space, so evaluating fit in linear space would not capture the best possible fit for each model, (Figure S1).

### 2.2.1 Fitting linear models

The two polynomial phenomenological models are linear in their parameters, so were fitted using Ordinary Least Squares with the lm() function in R [14].

### 2.2.2 Fitting non-linear models

The mechanistic models are non-linear, so most be fitted using an optimisation approach. The nlsLM function in the minpack.lm [7] package uses the Levenberg-Marquardt algorithm to find parameter values for a specified model. The optimiser requires initial values for the parameters, which are optimised to minimise RSS. If these initial values are far from the true optimal parameters, it is likely that the optimisation will not converge, or will converge to very suboptimal parameters. It is therefore important to estimate reasonable starting values to provide to the optimiser.

**Estimating parameter values** For the logistic model, parameters are in linear space (even though RSS is minimised in log space), whereas for the four-parameter models, the parameters are in linear space. Therefore, the following parameter estimation method was carried out both on population size and log population size.

$\mathrm{N_{min}}$ and $\mathrm{N_{max}}$ were set as the minimum and maximum observed population size respectively. $\mu_{max}$ and $t_{lag}$ estimates were determined using a rolling regression (Figure S2), which computed the regression line for every possible set of 5 adjacent data points. The regression line with the greatest gradient was selected, with the gradient taken as an estimate of $\mu_{max}$, and the $t$ value of the intercept between this selected regression line with the line $N = \min(N)$ taken as the $t_{lag}$ estimate .

**Multiple fit attempts** Even with reasonably chosen parameter estimates, models often fail to converge, or converge to poorly fitting parameter values. In order to increase the proportion of models which fit, and to ensure that parameter values nearer the global optimum were found, I used a grid search to test a range of initial parameter values. I specified a range for each parameter in each model surrounding the estimated parameter value. For each parameter, I choose linearly spaced samples from this range, and tested every possible combination of these parameter values. For example, if I take 3 linearly spaced samples for each of $N_0$ and $N_{max}$, and 5 for $t_{lag}$ and $\mu_{max}$, I fit $3 \times 3 \times 5 \times 5 = 225$ models. From the model fits that converge, I select the model with the smallest RSS. If none of the models converge, I randomly sample parameter values with normally distributed noise around my parameter estimates until a model fits, or until there have been 1,000 failed attempts.

Using this method, I was able to fit all 287 data sets for the Logistic, Gompertz, and

Baranyi models in both linear space and log space. In linear space, there was one failed fit for the Buchanan model, and in log space there were two failed fits for the Buchanan model.

## 2.3 Filtering

Some data sets show no pattern of population growth, or have too few points in the growth phase to constrain the $mu_{max}$ and $t_{lag}$ parameters. The 64 data sets with the lowest maximum $R^2$ value were visualised, of which 20 with no meaningful growth pattern were removed (Figure S3). For analyses involving parameter estimates, I further filtered out datasets with less than 2 points in the growth phase (Figure S4). This involved counting the number of points that have a time value within the predicted exponential phase. The predicted exponential phase was determined using Akaike weighted parameter values. This removed a further 57 data sets, leaving 210 data sets for parameter estimate analysis.

## 2.4 Model Selection

Model fit was assessed using AIC [1] and BIC [15]. With the assumption of normally distributed errors, and simplification to remove constants, these are:

$$\text{AIC} = 2k + n\log(RSS)$$

$$\text{BIC} = n\log\frac{RSS}{n} + k\log n$$

As these are dependent on the number of parameters of the model, they can be used for direct comparison of fit for models with differing numbers of parameters. For each data set, the model with the lowest AIC/BIC value was selected as the best fit model.

## 2.5 Parameter Estimation

Parameter estimation was performed using Akaike weights, which produce a single parameter estimate from a family of models [6], weighted by the model's $\Delta_{AIC}$ value: With $\text{AIC}_{best}$ being the lowest observed AIC values for a dataset, $\Delta_{AIC}$ values are calculated with:

$$\Delta_i = \text{AIC}_i - \text{AIC}_{best}$$

The Akaike weight for the $i^{th}$ model is then:

$$w_i = \frac{e^{-\frac{1}{2}\Delta_i}}{\sum_{r=1}^{r=R} e^{-\frac{1}{2}\Delta_r}}$$

The Akaike weighted parameter estimate for $\theta$, with weights $w_i$ and individual model parameter estimates $\theta_i$ is:

$$\theta_w = \sum_{i=1}^{i=I} \theta_i * w_i$$

## 2.6 Statistical Analysis

I ran linear regressions of $\Delta_{\text{Logistic - Gompertz}}$ against temperature, and $\Delta_{\text{Logistic - Gompertz}}$ against proportion of points in the lag phase. These were carried out using lm() in R [14], with their diagnostic plots examined to confirm the suitability of a linear model.

I ran a mixed effects model of $\mu max$ against temperature, and of $t_{lag}$ against temperature, with species and medium used as random effects in each case. Again, diagnostic plots were used to confirm that residuals were homogeneous across the fitted values.

I ran a paired t-test between the each pair of the four-parameter models to check for systematic differences in their parameter estimates. I check for normality of differences by plotting histograms of these differences.

## 2.7 Computing Tools

### 2.7.1 R 3.6.1 [14]

Used for initial data preparation script, for model fitting, for analysis, and for figure preparation. I chose R for these tasks because the native datastructures are natural tools for handling this type of data. The native support of vectorised operations also makes the code cleaner than it would be in Python. I chose to write my multi-start method to fit models over a range of initial parameter values in R because of the robustness and efficiency of non-linear least squares model fitting in the minpack.lm package.

- **minpack.lm [7]:** Non-linear least squares model fitting using the more robust Levenberg-Marquardt algorithm.
- **ggplot2 [18]:** To more easily create high quality graphics.
- **dplyr [19]:** For some data wrangling tasks.

- **lme4 [3]:** For fitting linear mixed models.
- **toOrdinal [4]:** For creating ordinal numbers for annotation.

**Python 3.7.4 [?]** Used for calculation of initial values.

- **pandas [11]:** For dataframes in python
- **matplotlib [10]:** To create graphics in python
- **scipy [12]:** For linear model fitting in python
- **numpy [12]:** For efficient arrays in python

LaTeXUsed for compilation of write up.

# 3 Results

## 3.1 Model fitting is highly successful

The method of taking the best fitting model from a range of initial parameter values was very effective, with all models converging to every time series when those with no discernible growth pattern were removed, and only 3 unsuccessful fits otherwise (2293 / 2296 successful fits overall). The model fits were visually inspected, and found to describe the data well (Figure 4 for log space fits, Figure S5 for linear space fits). For the reasons mentioned in the Model fitting section of Methods, I discuss the results obtained from fitting models in log space first, before presenting the results from linear space fits at the end of this section.

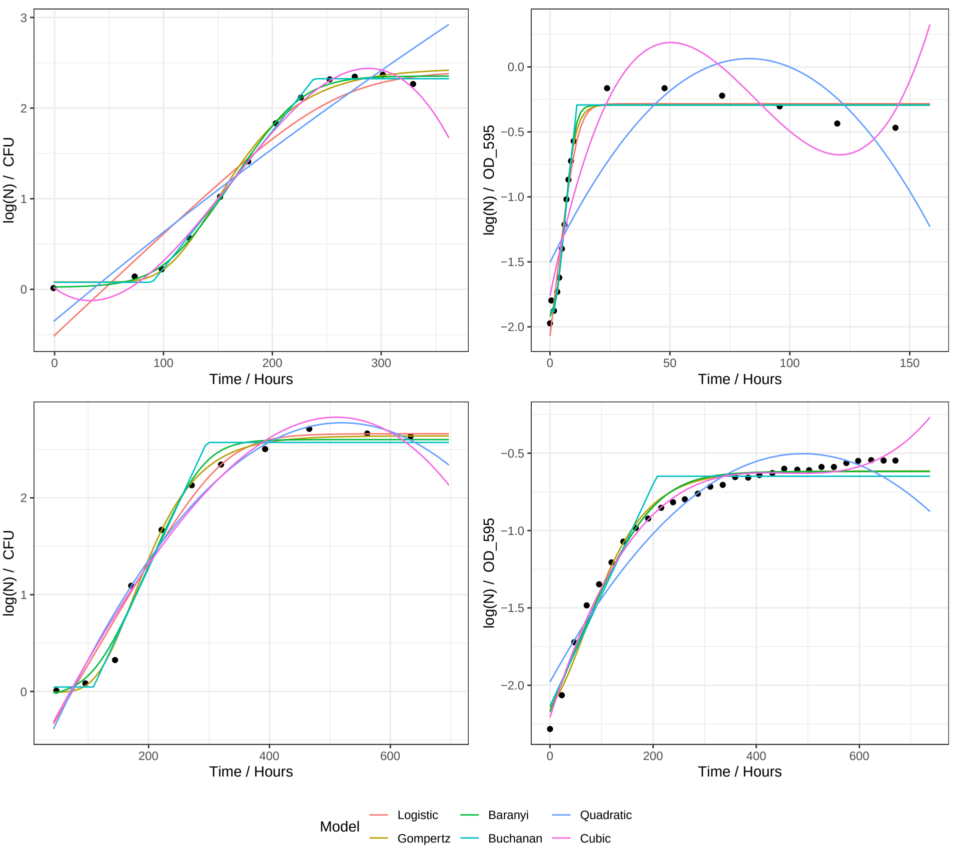

**Figure 4:** Demonstration of model fits in log space. Out of 287 data sets, all fits were successful apart from two fits of the Buchanan model, to datasets which were removed anyway.

### 3.2 Gompertz and Baranyi are best supported

I found that the model which most frequently had the lowest AIC and BIC values was the Gompertz model, which was determined to be the best fit model in 91 of 267 cases with AIC, and 90 cases with BIC (Figure 5 A). Although second most frequent best model is the Logistic model, this should not be interpreted as the Logistic model having the second most support from the data. From the means and distributions of AIC values (Figure 6 and Table 1), we can see that the Baranyi model in fact has a slightly lower mean AIC score and slightly lower mean ranking than even the Gompertz model (Figure 5 B and C), despite being the best fit model on only 35 occasions. This can be explained by the fact that the Baranyi model is often the second best fitting model - on 131 out of 267 occasions, in fact. It therefore must be common that the Gompertz model is the best fitting model with the Baranyi model just behind in second, with many cases of the Baranyi model being first or second with the Gompertz performing much more poorly, such that its AIC and BIC become slightly worse than the Baranyi model's. Indeed, when we include the Baranyi model as the only 4 parameter model, it is the best fitting

215 model on 130 occasions (Figure 5 D)

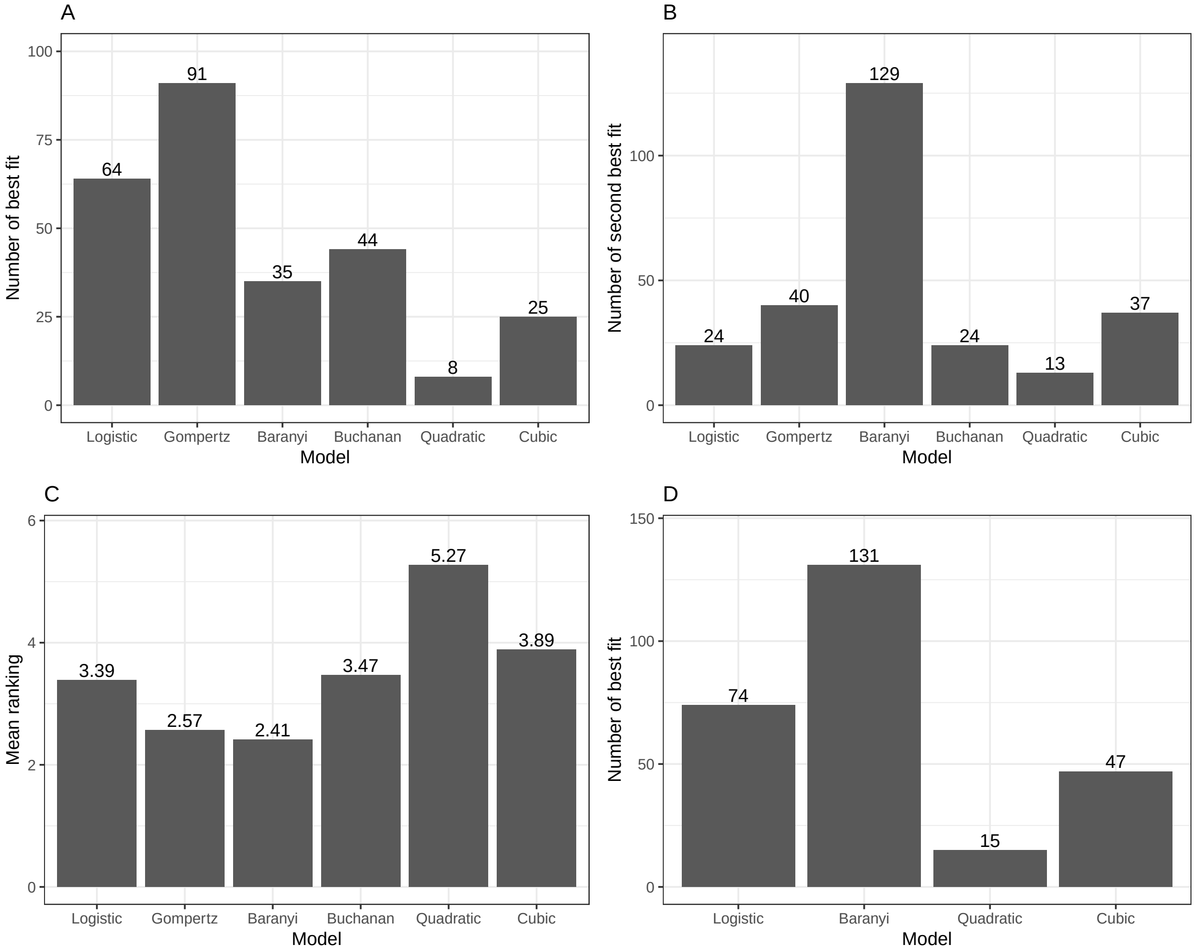

**Figure 5:** Plots showing the rankings of models in log space. (A) shows the number of times each model is the best fitting model for a dataset. (B) shows the number of times each model is the second best fitting model for a dataset. (C) shows the mean ranking for each model. (D) shows the number of times each model is the best fitting model for a dataset if we do not consider the Gompertz and Buchanan models.

| | Logistic | Gompertz | Baranyi | Buchanan | Quadratic | Cubic |
|---|---|---|---|---|---|---|
| AIC | -58.5 | -66.2 | -66.3 | -61.1 | -43.7 | -53.2 |
| BIC | -57.0 | -64.1 | -64.2 | -59.1 | -42.2 | -51.1 |

**Table 1:** Mean AIC / BIC values for each model, fitted in log space.

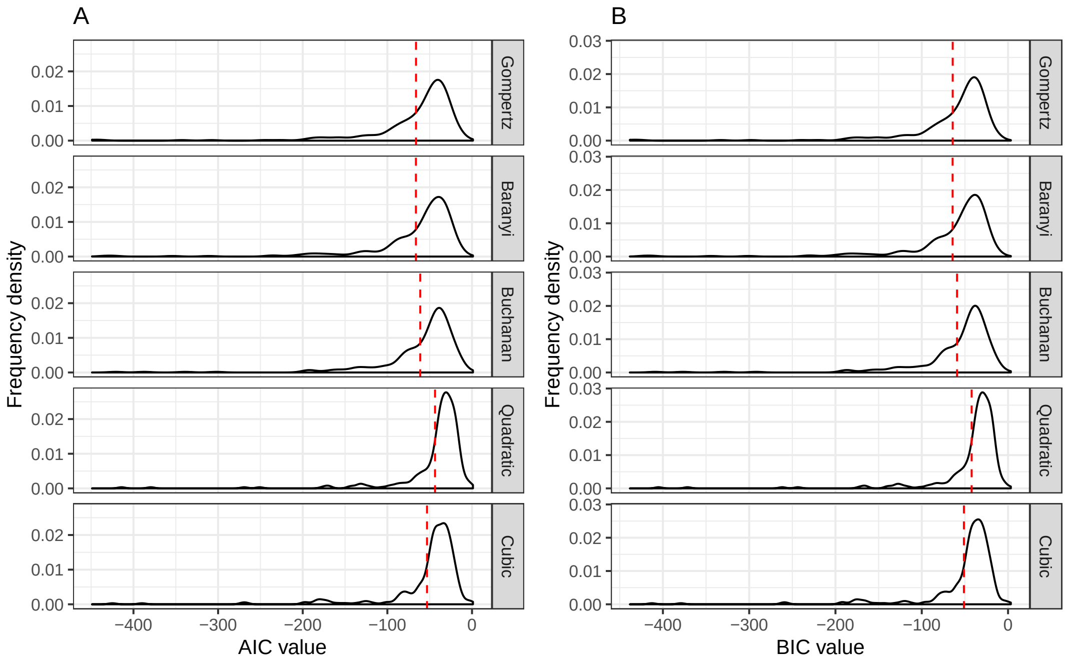

**Figure 6:** Gaussian Kernel Density Estimates for AIC (A) and BIC (B) distributions for each model. We can see that the choice of AIC or BIC makes little difference.

## 3.3 Logistic fit is poor when $\mathbf{t_{lag}}$ is high

It is notable that the Logistic model is the best model second most frequently, but that its mean AIC and BIC scores are only fourth best overall - which implies great variability in its performance. This can be explained by the absence of a lag phase in the model: when the data exhibits a lag phase, the Logistic model fits poorly, but when there is no lag phase, it is able to fit the data well without incurring the AIC/BIC penalty for a fourth parameter. This is evident in the increasing proportion of points in the lag phase as the ranking of the logistic model gets worse (Figure 7 A). More explicitly, we can see that the difference in AIC between the Logistic and Gompertz model increases with proportion of data points in the lag phase (Figure 7 B), indicating that the Gompertz model has an increasing advantage over the Logistic model as the lag phase increases in duration. This relationship is strongly significant, with a linear regression of $\Delta_{\text{Logistic - Gompertz}}$ against proportion of lag phase points produces an estimated increase in $\Delta_{\text{Logistic - Gompertz}}$ of 6.5 for every increase in lag phase proportion of 0.1, with a standard error of 0.49.

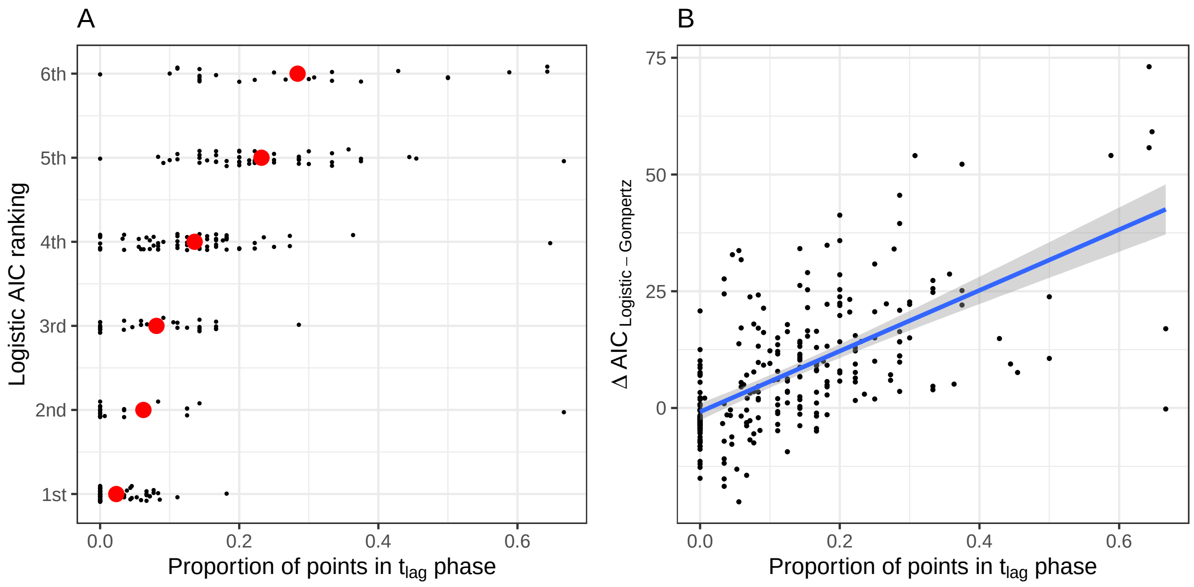

**Figure 7:** (A) is a plot of proportion of points in the lag phase, against the ranking of the Logistic model by AIC. We see that an increasing proportion of points in the lag phase is associated with a poorer ranking - with the mean proportion for each ranking shown by the red circle. (B) shows the difference in AIC score between the Gompertz and Logistic model against this proportion, with a statistically significant slope observed.

## 3.4 Temperature affects $\mu_{max}$ and $t_{lag}$

I ran a linear mixed model to determine the effect of temperature on the $\mu_{max}$ and $t_{lag}$ parameters across the entire data set. Each of these models included species and medium as random effects, and temperature included as a categorical variable as it is unlikely that the response to temperature is linear across the entire temperature range. From this, estimates for $\log(\mu_{max})$ and $t_{lag}$ values could be made for each temperature, which are shown in Figure 8. It is clear that the log of the maximum rate of bacterial growth increases with temperature, approximately linearly across much of the temperature range, indicating an Arrhenius type functional response. $t_{lag}$ decreases with temperature, quickly levelling off at around 10°C.

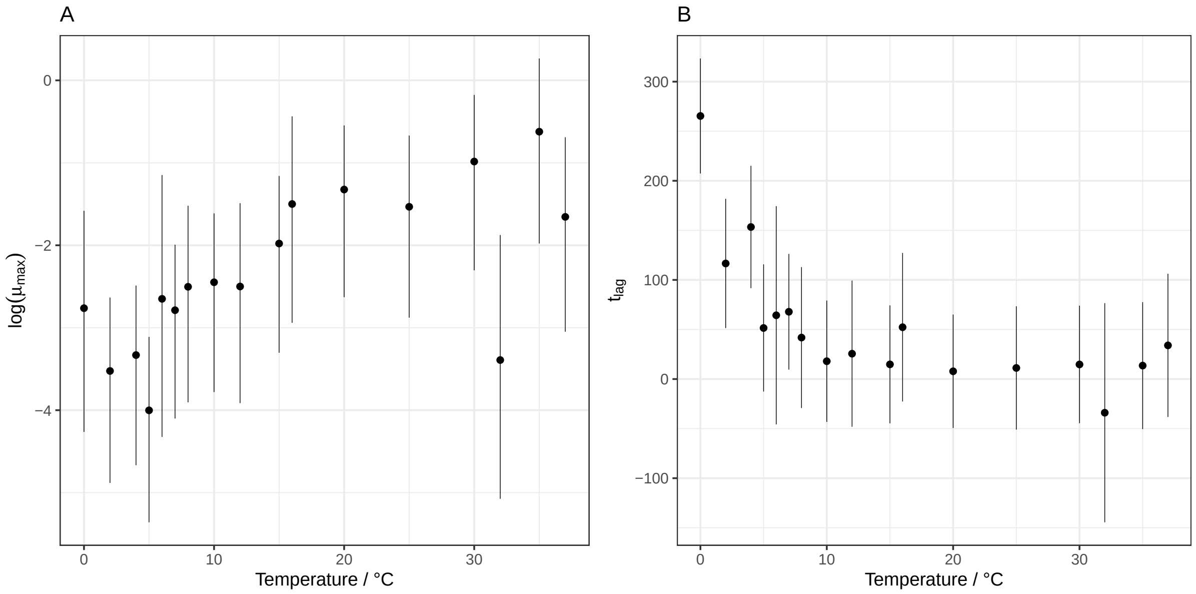

**Figure 8:** (A) shows the relationship between $\mu_{max}$ and temperature, as estimated by a linear mixed model with temperature as a categorical variable. Vertical lines indicate the 95% CI. $\mu_{max}$ increases steadily with temperature over the observed range. Similarly, (A) shows the relationship between $t_{lag}$ and temperature. We see that below 10°C, $t_{lag}$ increases rapidly as temperature decreases.

## 3.5 Logistic model fits better at higher temperatures

The decrease in $t_{lag}$ with temperature implies that the Logistic model, which has better fit with smaller $t_{lag}$, should have better fit with increasing temperature. Indeed, we can see that this is the case in Figure 9 A, which shows that $\Delta_{\text{Logistic - Gompertz}}$ decreases with increasing temperature - implying that the Logistic model gets better support relative to the Gompertz model at higher temperatures. A linear regression suggests a decrease in $\Delta_{\text{Logistic - Gompertz}}$ of 3.30 for every 10°C, with standard error of 0.8.

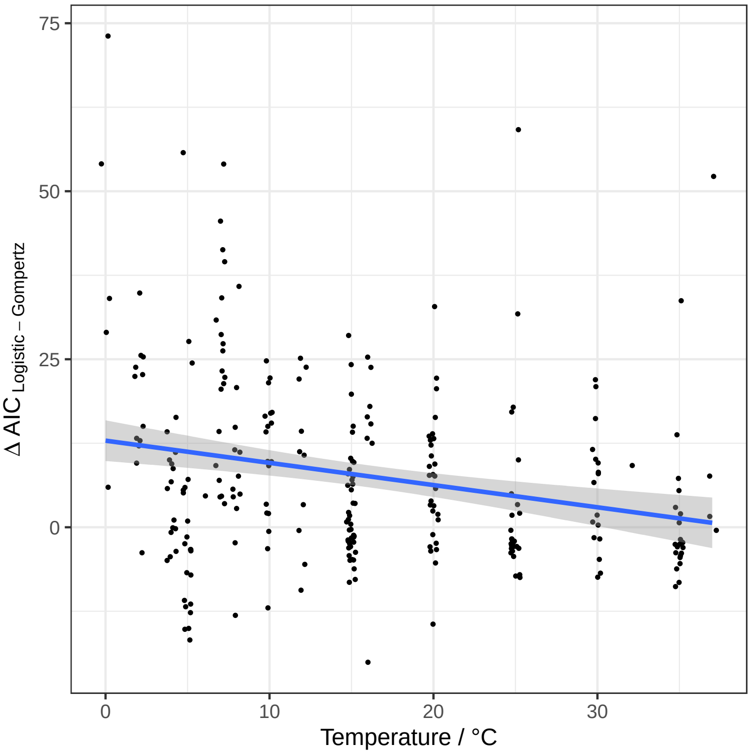

**Figure 9:** $\Delta_{\text{Logistic - Gompertz}}$ is seen to decrease by 3.30 for every 10°C, indicating an improvement in the fit of the Logistic model relative to the Gompertz model as temperature increases.

### 3.6 Model choice biases parameter estimates

Tables 2 and 3 show the mean pairwise differences in $t_{lag}$ and $\mu_{max}$ for the 4-parameter models. We can see that the Buchanan model results in the lowest $t_{lag}$ and lowest $\mu_{max}$ values. This can be explained by the fact that the model assumes that the population growth remains exactly at zero until $t = t_{lag}$, rather than smoothly transitioning to $\mu_{max}$ before this intercept. This means that Buchanan model fits must decrease $t_{lag}$ to capture the first part of the growth phase, consequently lowering $t_{lag}$ and $\mu_{max}$ values (Figure S6).

| | Gompertz | Baranyi | Buchanan |
|---|---|---|---|
| Gompertz | | -5.21 = - 9.8% | 4.23 = 8.0% |
| Baranyi | | | 9.44 = 17.8% |
| Buchanan | | | |

**Table 2:** Mean pairwise differences in $t_{lag}$ estimates for all models, and as proportions of the Akaike weighted mean $t_{lag}$ value. The value in $(i, j) = t^i_{lag} - t^j_{lag}$. The Buchanan model produces lower $t_{lag}$ estimates than either the Baranyi or the Gompertz models. All values are significant at the 5% level.

| | Gompertz | Baranyi | Buchanan |
|---|---|---|---|
| Gompertz | | -0.025 = -5.4% | 0.11 = 23.4% |
| Baranyi | | | 0.14 = 29.3% |
| Buchanan | | | |

**Table 3:** Mean pairwise differences in $\mu_{max}$ estimates for all models, and as proportions of the Akaike weighted mean $t\mu max$ value. The value in $(i, j) = \mu^i_{max} - \mu^j_{max}$. The Buchanan model produces lower $\mu_{max}$ estimates than either the Baranyi or the Gompertz models. All values are significant at the 5% level.

### 3.7 Model fitting in linear space fails to resolve these effects

As explained in Methods, fitting linearised models to minimise linear RSS reduces the ability of the models to resolve the lag phase. As such, the Logistic model is favoured when fitting in linear space despite its lack of a lag phase. When minimising linear residuals, the Logistic model is most frequently the best model and has the lowest mean ranking (Figure S7 and Table S1). This weakened penalisation of the lag phase also means that the relationship between Logistic model fit and proportion of points in the lag phase is much weaker (Figure S9), and the relationship with temperature becomes smaller and loses significance (Figure S9).

## 4 Discussion

### 4.1 The Baranyi model is recommended for microbial odelling

The results are unequivocal about the importance of a lag phase for modelling microbial growth in log space, with the three four-parameter having the lowest AIC and BIC values, and the fit of the Logistic model worsening with increasing lag phase duration. Although I find that the Gompertz model is most frequently the best fitting model, the Baranyi model in fact has very slightly lower mean AIC and mean BIC. Whilst this cannot be directly translated into clear guidelines for which model should be preferred for microbial growth modelling, I propose that other advantages of the Baranyi model make it a preferable choice:

1. Consistency The fact that the Baranyi model achieves lower AIC and BIC scores despite rarely having the best fit implies that it much more consistently fits the data well, and that the Gompertz more often has a poor fit. This implies that the Baranyi model can be deployed ore broadly and reliably than the Gompertz model.

2. Truly mechanistic The Baranyi model has the advantage that it was formulated entirely with microbial growth in mind, giving it a true mechanistic basis in microbial growth

modelling. The full Baranyi model also has extra curvature parameters $\nu$ and $m$, which alter the transition into and out of the exponential phase. As more data becomes available, these parameters can be estimated (rather than pre-set) to form a more detailed understanding of the mechanisms behind microbial growth. Similarly, the Baranyi model's capacity to make predictions with non homogenous environmental conditions with its differential equation form makes it a much more flexible tool for growth modelling.

## 4.2 Lowering temperature can dramatically increase the shelf life of food

I find that $\log(\mu)max)$ increases roughly linearly with temperature, corresponding to an exponential increase in $\mu_{max}$ - which in turn describes the rate of an exponential growth. Similarly, I find that $t_{lag}$ decreases rapidly with increasing temperature, such that the advantage of lower temperatures is almost entirely lost by 10 °C. Clearly, minimising the temperature of food storage can have a dramatic effect in decreasing microbial load - and conversely, increasing the temperature can result in growth that is orders of magnitude faster, perhaps fast enough that even relatively brief exposure to higher temperatures could spoil food. It will therefore be important to see how microbial growth patterns respond to temperatures that vary over time, which the Baranyi model will be very helpful in studying.

## 4.3 Parameter biases necessitate caution in meta-analyses

There is a clear hierarchy in the magnitude of parameter estimates from the three models. The Buchanan model produces the lowest $t_{lag}$ and $\mu_{max}$ estimates, with the Baranyi producing the highest, and Gompetz producing intermediate values. The magnitude of the differences between the Buchanan and Baranyi model reach 29% of the weighted estimate for $\mu_{max}$, and 17% for $t_{lag}$. These are extremely significant differences, so great care, probably involving reevaluation of raw data sets, must be taken to ensure that parameter estimates from different sources are meaningfully compared. A simulation study could be used to determine which model, or combination of models, produces the most reliable parameter estimate. Datasets produced by a variety of simulation mechanisms, with known 'true parameters' could be produced for a range of generative processes, and used to test how robust each of the models is to deviations from their mechanistic assumptions, and to noise in the data.

## 4.4 Fitting in log space is required for food spoilage studies

Fitting in linear space allows the logistic model to become the most frequent best fitting model, and have the lowest mean ranking. This is a strong remit to recommend that model fitting should be carried out and evaluated in log space, as the improved ranking of the Logistic model in linear space fits demonstrates that fitting in this space only weakly enforces accurate modelling of the lag phase. For the purposes of food spoilage work, this is problematic, as maximisation of the lag phase is of central interest in preservation. It is also the case that four-parameter models cannot infer $t_{lag}$ and $\mu_{max}$ with as much confidence when fitted in linear space, as shown by the wider confidence intervals in Figure S10 relative to Figure 8 - because the beginning of the exponential phase is less well differentiated from the lag phase in linear space relative to log space.

As such, this study has demonstrated that the choice of space and model used in microbial growth modelling can have profound effects both on the parameters estimated, and on the biological hypothesis that is given greatest support by the fitting results. Further work examining the reliability of the various models to violations of their assumptions is needed to devise strategies that can provide the most accurate parameter estimates, which will allow increasingly quantitative approaches to food preservation.

# 5 Supplementary Material

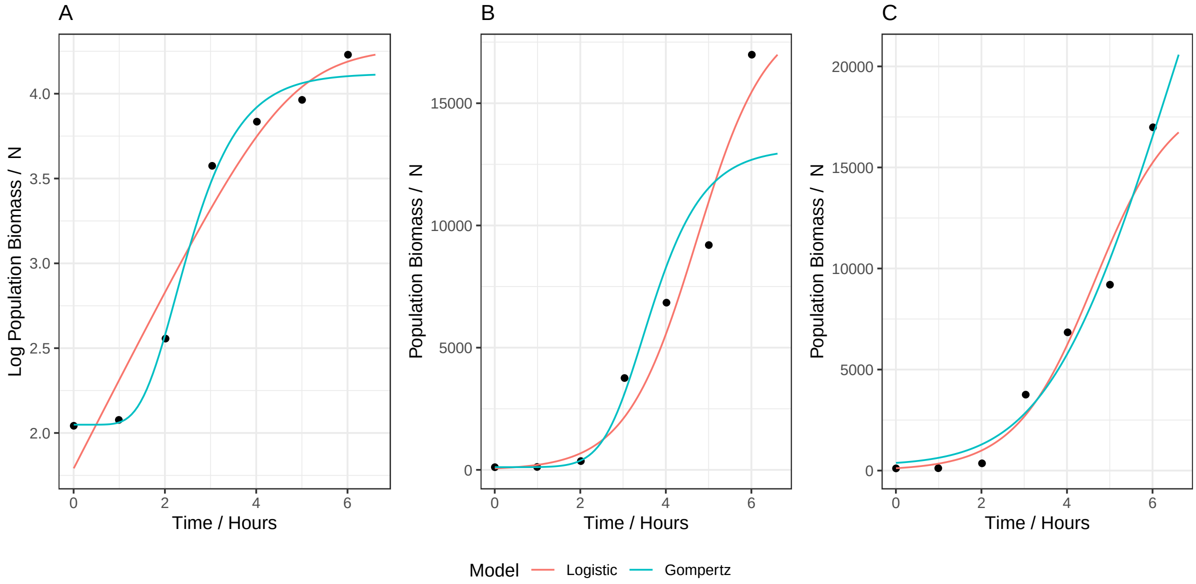

**Figure S1:** Comparison of: (A) fitting and evaluating modelling in log space. (B) fitting model in log space, evaluating fit in linear space. (C) fitting and evaluating model in linear space. When the models are fit and evaluated in log space, it is clear that the Gompertz model has a better fit. However, if evaluation is done in linear space of these fits, the logistic model appears to have a better fit. This is shown to be invalid, as fitting the models in linear space demonstrates that the Gompertz model can in fact achieve better fits in both spaces.

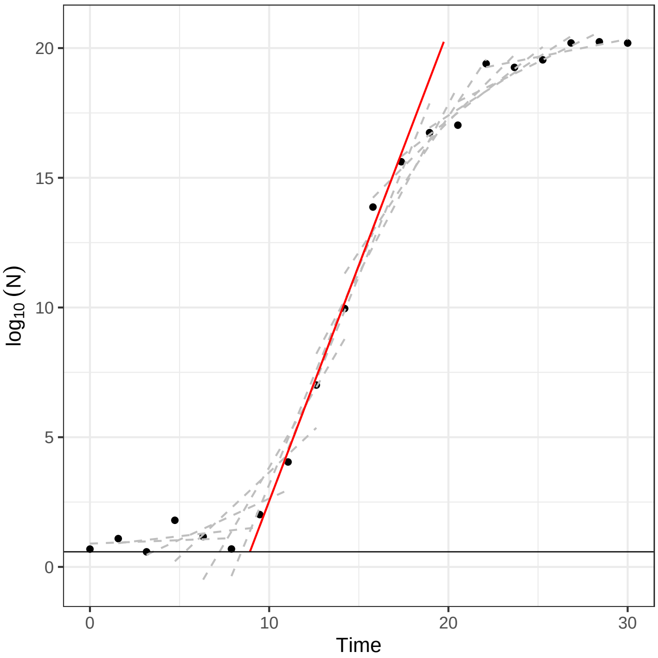

**Figure S2:** Rolling regression was used to produce starting parameters. The regression line with the greatest gradient was selected, with its gradient used as $\mu_{max}$, and its intercept with $N = \min(N)$ used as $t_{lag}$

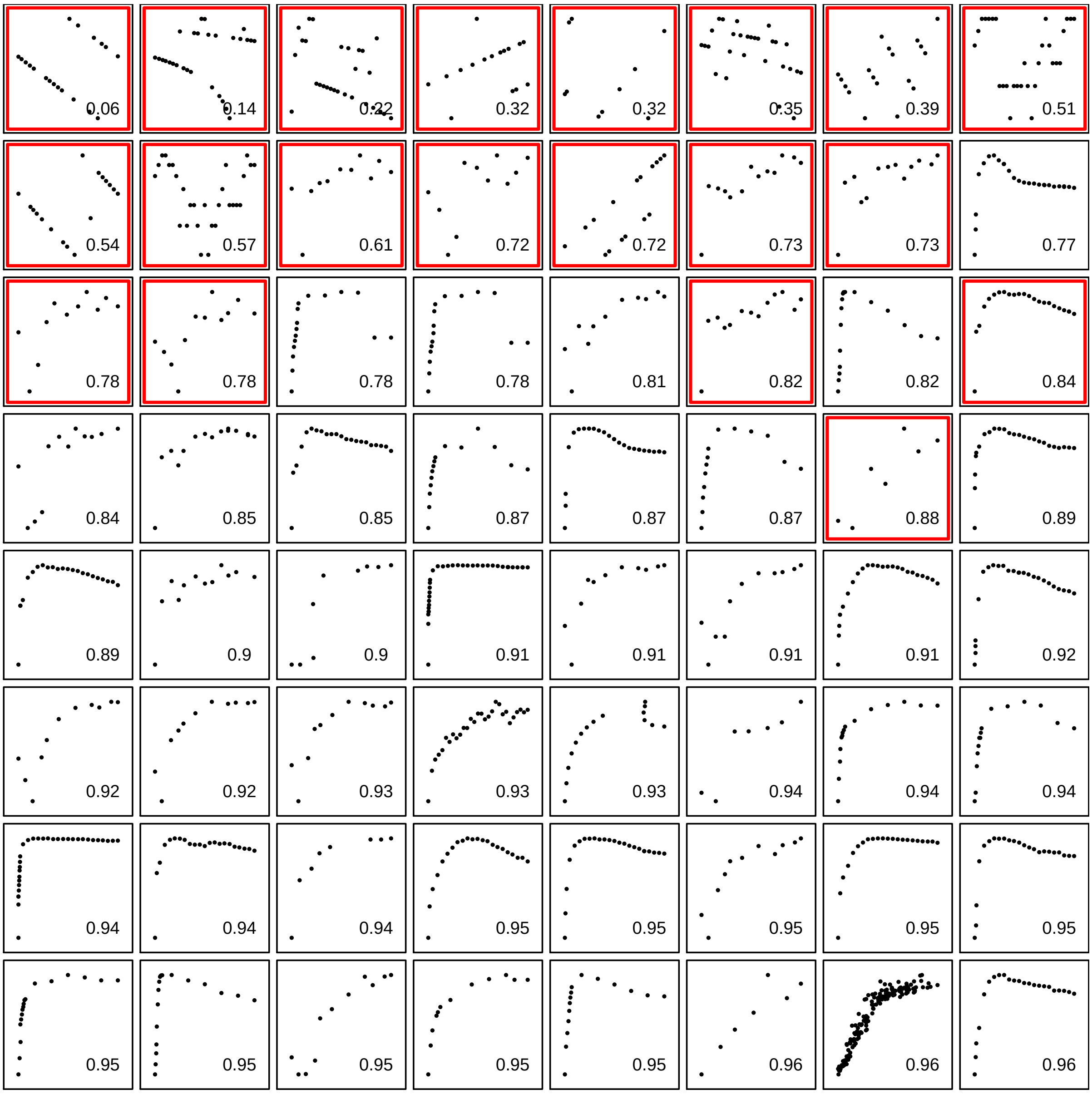

**Figure S3:** The 64 datasets with the lowest maximum $R^2$. Those which were manually removed are highlighted with a red border.

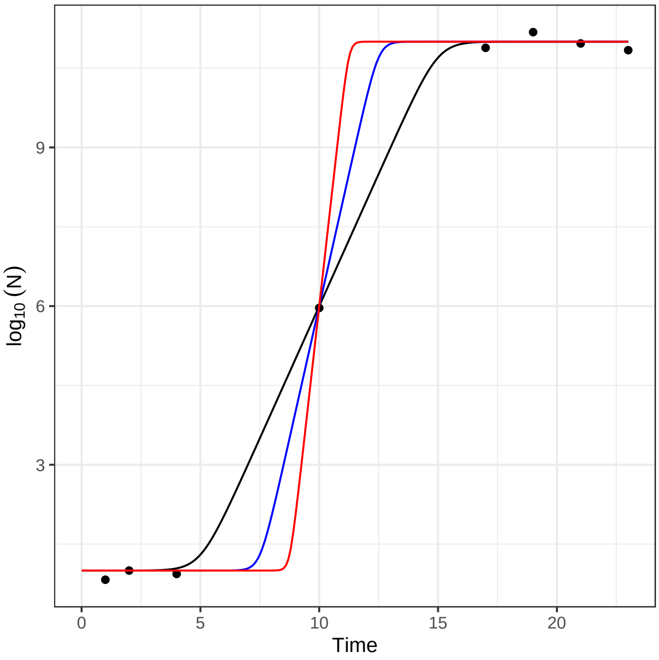

**Figure S4:** If there are too few points in the growth phase, $\mu_{max}$ and $tlag$ will not be well constrained. Here, we can see that there are infinitely many solutions that can fit equally well to the data as there is only one point in the growth phase.

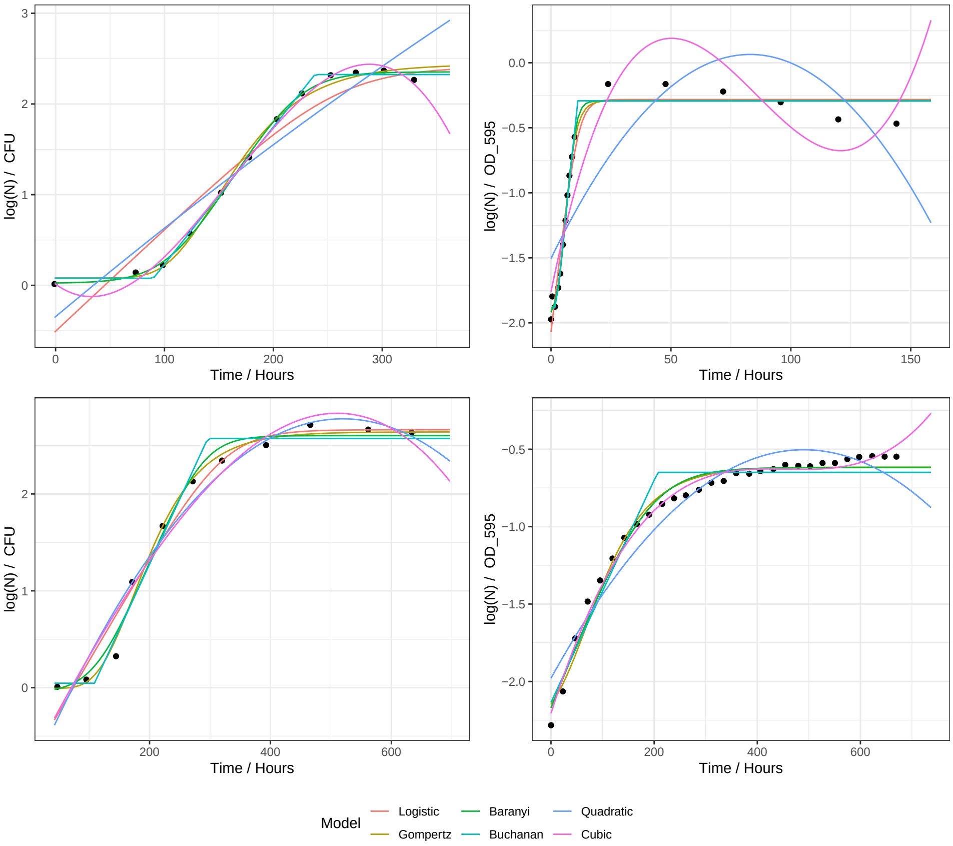

**Figure S5:** Demonstration of model fits in linear space. Out of 287 data sets, all fits were successful apart from a single fit of the Buchanan model, to a dataset which was removed anyway.

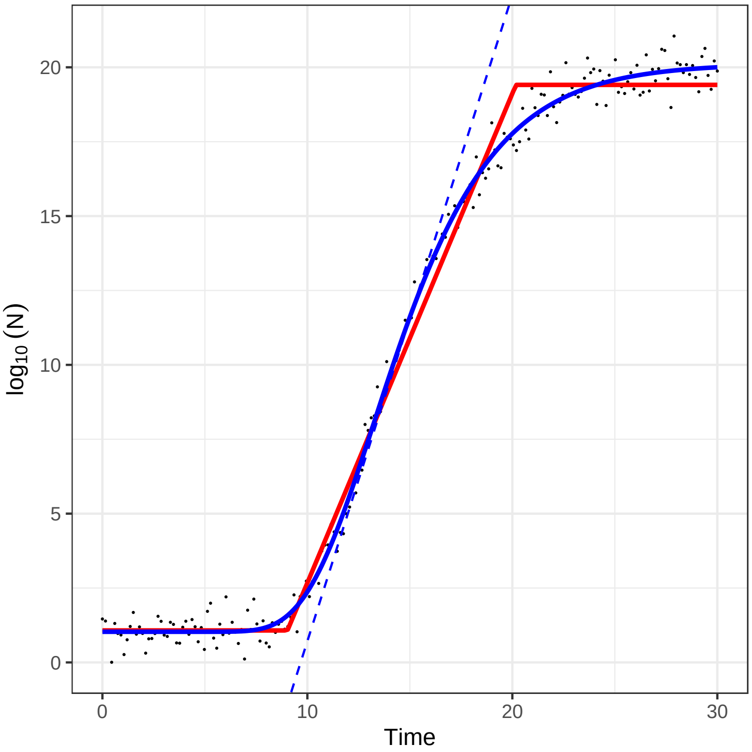

**Figure S6:** The Buchanan and Gompertz models are fit to the same hypothetical growth data, which results in smaller $t_{lag}$ and $\mu_{max}$ estimates from the Buchanan model in order to capture the start of the growth phase. By contrast, there is some growth before $t = t_{lag}$ in the Gompertz model.

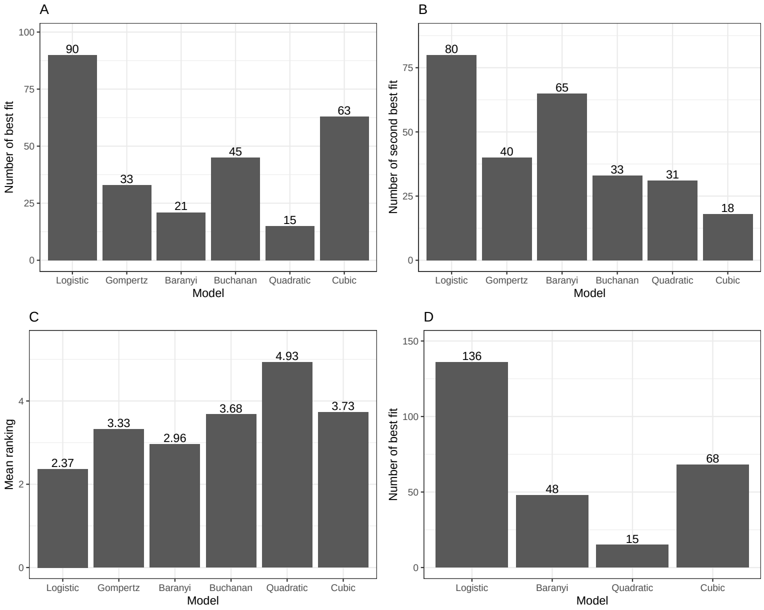

**Figure S7:** Plots showing rankings of models in linear space. (A) shows the number of times each model is the best fitting model for a dataset. (B) shows the number of times each model is the second best fitting model for a dataset. (C) shows the mean ranking for each model. (D) shows the number of times each model is the best fitting model for a dataset if we do not consider the Gompertz and Buchanan models. In linear space, the logistic model performs better than it does in log space, as less emphasis is put on the lag phase.

|  | Logistic | Gompertz | Baranyi | Buchanan | Quadratic | Cubic |
|---|---|---|---|---|---|---|
| AIC | 62.3 | 61.8 | 61.3 | 65.5 | 76.3 | 67.7 |
| BIC | 63.8 | 63.8 | 63.3 | 67.5 | 77.8 | 69.7 |

**Table S1:** Mean AIC / BIC values for each model fitted in linear space

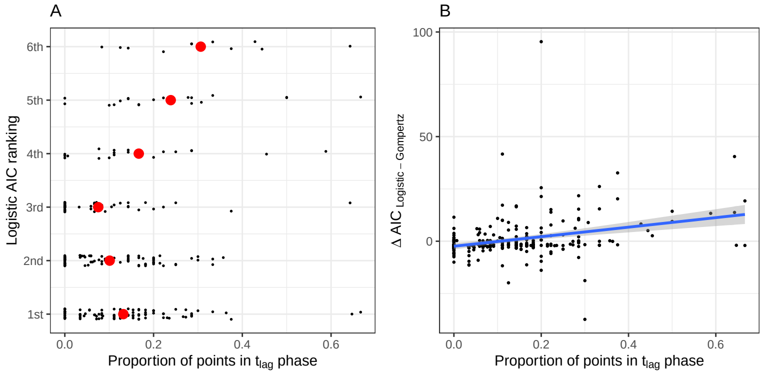

**Figure S8:** (A) is a plot of proportion of points in the lag phase, against the ranking of the Logistic model by AIC in linear space. (B) shows the difference in AIC score between the Gompertz and Logistic model against this proportion, In both cases, the relationship between Logistic model performance and $t_{lag}$ is much weaker than in log space. The difference in AIC between Logistic and Gompertz decreasing by 2.2 for every increase in lag phase proportion by 0.1 (in contrast to 6.5 for log space)

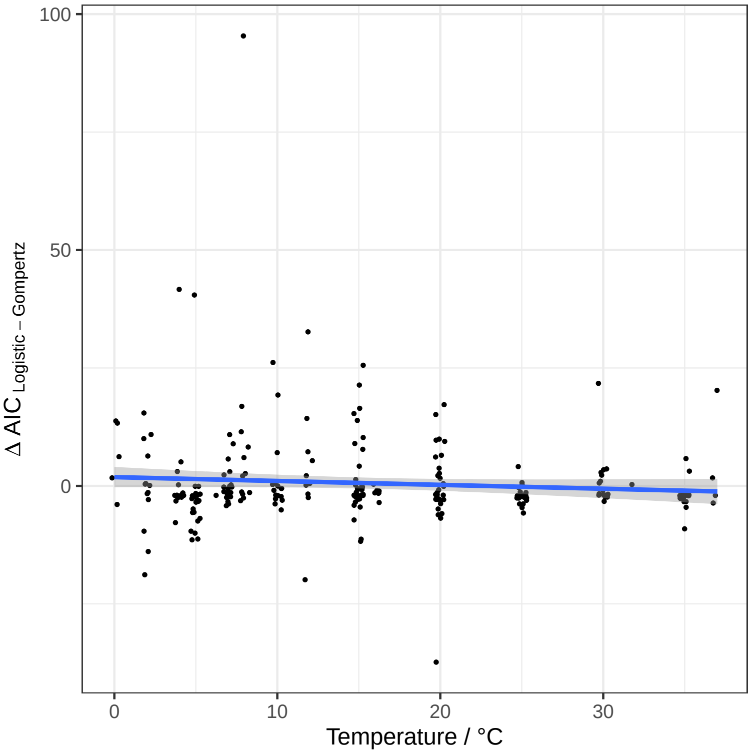

**Figure S9:** $\Delta_{\text{Logistic - Gompertz}}$ no longer has a statistically significant relationship with temperature when models are fit in linear space. This implies that the fit of the Logistic model does not become any weaker realtive to thegompertz model as temperature increases.

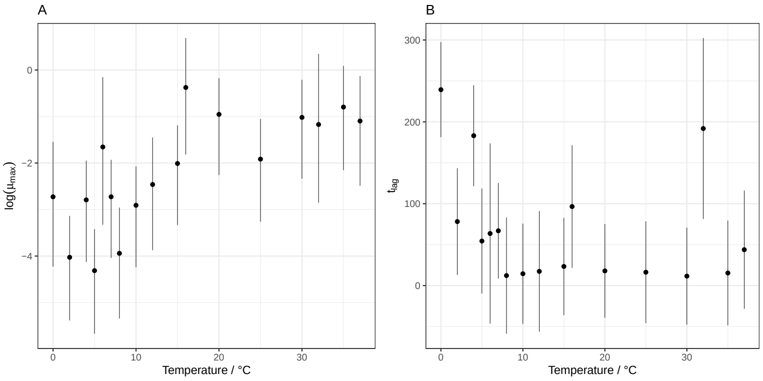

**Figure S10:** The broad patterns of $\mu_{max}$ and $t_{lag}$ are the same when model fitting is in linear space, but the confidence intervals are much broader.

# References

[1] H. Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723, 1974.

[2] Roberts TA. Baranyi J. A dynamic approach to predicting bacterial growth in food. *Int J Food Microbiol.*, 23:277–294, 1994.

[3] Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48, 2015.

[4] Damian W. Betebenner. *toOrdinal: Function for Converting Cardinal to Ordinal Numbers by Adding a Language Specific Ordinal Indicator to the Number*, 2019. R package version 1.1-0.0.

[5] R. C. Whiting Buchanan, R. L. and W. C. Damert. When is simple good enough: A comparison of the gompertz, baranyi, and three-phase linear models for fitting bacterial growth curves. *Food Microbiology*, 14:313–26, 1997.

[6] Anderson David R Burnham, Kenneth P. *Model Selection and Multimodel-Inference.* 2002.

[7] Timur V. Elzhov, Katharine M. Mullen, Andrej-Nikolai Spiess, and Ben Bolker. *minpack.lm: R Interface to the Levenberg-Marquardt Nonlinear Least-Squares Algorithm Found in MINPACK, Plus Support for Bounds*, 2016. R package version 1.2-1.

[8] B. Gompertz. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Phil. Trans. Roy. Soc. London*, 123:513–585, 1832.

[9] K. Grijspeerdt and P. Vanrolleghem. Estimating the parameters of the baranyi model for bacterial growth. *Food Microbiology*, 16:593–605, 1999.

[10] J. D. Hunter. Matplotlib: A 2d graphics environment. *Computing in Science & Engineering*, 9(3):90–95, 2007.

[11] Wes McKinney. Data structures for statistical computing in python. In Stéfan van der Walt and Jarrod Millman, editors, *Proceedings of the 9th Python in Science Conference*, pages 51 – 56, 2010.

[12] Travis Oliphant. NumPy: A guide to NumPy. USA: Trelgol Publishing, 2006–. [Online; accessed ¡today¿].

[13] J.D. Phillips and M.W. Griffiths. The relation between temperature and growth of bacteria in dairy products. *Food Microbiology*, 4(2):173–185, 1987.

[14] R Core Team. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria, 2019.

[15] Gideon E. Schwarz. Estimating the dimension of a model. *Annals of Statistics*, 6 (2):461–464, 1978.

[16] P.-F Verhulst. Recherches mathématiques sur la loi d'accroissement de la population. *Nouv. mém. de l'Academie Royale des Sci. et Belles-Lettres de Bruxelles*, 18:1–41, 1845.

[17] P.-F Verhulst. Deuxième mémoire sur la loi d'accroissement de la population. *Mém. de l'Academie Royale des Sci., des Lettres et des Beaux-Arts de Belgique*, 20:1–32, 1847.

[18] Hadley Wickham. *ggplot2: Elegant Graphics for Data Analysis.* Springer-Verlag New York, 2016.

[19] Hadley Wickham, Romain François, Lionel Henry, and Kirill Müller. *dplyr: A Grammar of Data Manipulation*, 2019. R package version 0.8.3.

[20] I. Jongenburger F. M. Rombouts Zwietering, M. H. and K. Van't Riet. Modeling of the bacterial growth curve. *Applied and Environmental Microbiology*, 56:1875–81, 1990.