Continuity of the thermal optimum in mesophilic and psichrophilic *Arthrobacter* species. A multimodel inference approach

Pablo Lechón Alonso

Imperial College London

Word count: 3570

Abstract

We fit seven primary models to a data set of 285 bacterial growth curves and use model selection to address the question of which model performs best. Moreover, we examine the continuity of the thermal optimum (temperature (T) at maximum growth rate ($\hat{\mu}_{max}$)) of seven mesophilic and psychrophilic species from the genus *Arthrobacter*, by fitting the data ($T, \hat{\mu}_{max}$) with a thermal performance curve. To obtain $\hat{\mu}_{max}$ we perform model averaging on the primary models that yield a maximum growth rate estimate. No discontinuity is found between the thermal optimum of psychrophilic and mesophilic bacteria.

Contents

1 Introduction

Bacterial growth modeling dates back to the 1950's [1]. Thus, many models have been conceived to try to capture the different growth/inactivation phases (figure 1). These models are referred to as primary models. Some of them are more phenomenological (Gompertz, polynomials) than others, which are widely considered as more 'mechanistic' (Logistic, Baranyi). In this work, seven models are being fitted (see table 6.1). Moreover, the models can be confronted with each other in different ways (null hypothesis testing, model selection) to determine which one fits the data best. In this work, we address this task via model selection and model averaging [2, 3].

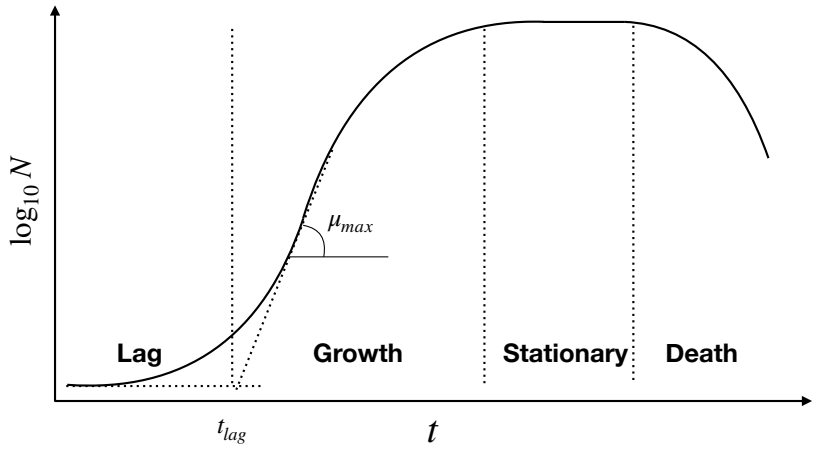

Figure 1: Tipical growth curve showing the four phases: lag phase, exponential growth phase, stationary phase, and death phase. The geometrical interpretations of the growth rate and lag time are also shown.

Biological information can be extracted from the values of the parameters resulting from the model fits. Furthermore, the dependance of these parameters on external factors such as temperature T, water activity a_w, or pH can also be studied by using secondary models. Many studies have been conducted aiming to model the dependance of $\mu_{max} \sim T$ [4–12]. A widely used group of secondary models that model this dependance are the thermal performance curves (TCP)

(see figure 2). In this work, the Lactin-2 model (see equation 17) is used to model the dependance of μ_{max} on T for seven species on the genus *Arthrobacter*.

We aim to address the continuity of psychrophilic and mesophilic growth characteristics in the genus Arthrobacter. Previous work on this topic [14] reported: "[...] no sharp cutoff point between growth-temperature requirements of psychrophilic and mesophilic bacteria". Here, we rigorously show such experimental observation by inspecting the distribution of fitted T_{opt} across species.

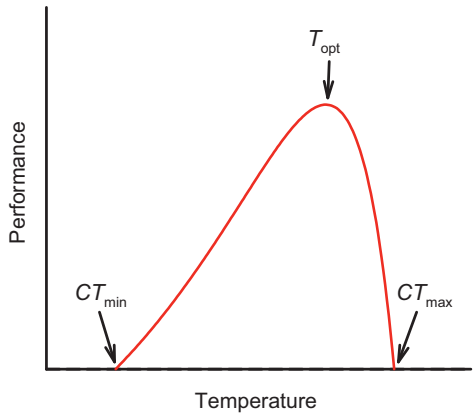

Figure 2: General shape of a thermal performance curve. The thermal optimum T_{opt} defines the temperature at maximum performance. The critical thermal minimum (CT_{min}) and maximum (CT_{max}) are temperatures with thermal performance zero; they define the limit of an organism's thermal niche. The figure is taken from [13].

2 Methods

In this section, we explain the details of the data, present, and briefly discuss the tested models, summarize the workflow of the mini-project code as a whole, and motivate the use of each computing tool we employed.

The data comes from a collection of 10 papers written between 1962 and 2018 [5–7, 9, 14–19] where the population of 45 species of bacteria is measured

at different times (h), temperatures (°C) and mediums. The populations are expressed in different units depending on the paper, i.e, optical density measured at 535 nm (OD_{535}), number of bacteria (N), colony-forming unit (CFU) and dry weight (DryWeight). Grouping datasets of population versus time for each temperature, species and medium yields 301 bacterial growth curves to which we can fit our models.

After taking the $\log_{10}$ of the population values, seven models are fit to the growth curves; exponential (which turns into linear after logging the population), quadratic , cubic, logistic [20,21], Gompertz [22], Baranyi [23] and Buchanan [24] (figure 3). Refer to section 6.1 to see the mathematical form of each model. As seen in figure 3, we are fitting four linear models (linear, quadratic, cubic, Buchanan) and three non-linear models (logistic, Gompertz, Baranyi).

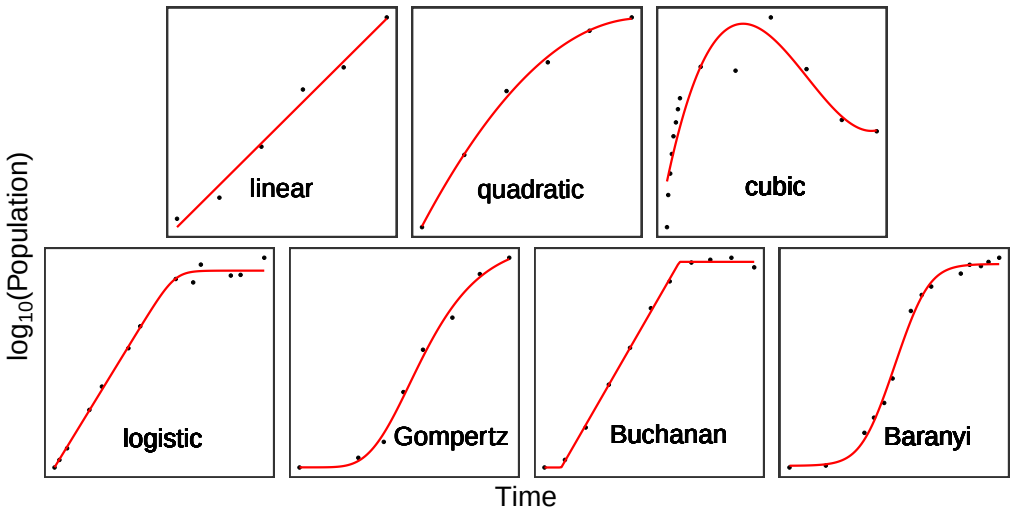

Figure 3: Overview of fit of each model to a growth curve where they perfrorm best (they have the lowest AIC value of out of all models for that curve)

The general workflow of this project has two modules: Module 1 (fitting of all bacterial growth curves and model selection), and Module 2 (analysis of the biological question). Each module has 3 subdivisions: data preparation, fitting and storing results, and plotting. These modules, along with the LaTeX code that

produces the present document are glued together in a script that coordinates all the tasks to run the mini-project smoothly

2.1 Module 1

This module concerns the fitting of all growth curves to all models and the model selection.

The first step in this module is the data preparation (`data_preparation_1.R`), in which we log the data, deal with bad quality data, and repeated measurements. First, we decide to log our data ($N(t) \rightarrow y(t)$) because the fitting performance and number of iterations until convergence are enhanced. This is because taking $\log_{10}$ decreases the scale of population values, particularly those measured in N units, which can be $\sim 10^{10}$. The mathematical form of the models also changes when logging the populations. Second, we eliminate the outliers in the population data (only one value of -682) that don't have any biological meaning. Next, we deal with negative values closer to 0 by finding the minimum of those values and adding it to all population values. This assures that all elements are positive without losing any data points, which are very valuable for the fits. Adding a constant to our population values will modify some of the fit parameter estimations, specifically all the coefficients from the second and third-order degree polynomials, but more importantly, y_0 and y_{max} from the exponential, logistic, Gompertz, Baranyi and Buchanan models. However, that does not affect our model selection, or our biological question, because none of these analyses depend on the value of the affected parameters. Third, we address the repeated measurements, which are only present in the growth curves of the species *Tetraselmis tetrahele*. This data consists of repeated measures of the population at different times. However, the times at which the population is measured slightly differ among each repetition. To deal with this we create the function `binning` (stored in `data_preparation_functions.R`) that groups the slightly different times (within a threshold of 0.25) into averaged time groups,

$\overline{t_i}$. We use the binning for the time vector to bin our population accordingly into a vector of averaged populations, $\overline{y_i}$. This procedure yields a unique averaged dataset $\left(\overline{t_i}, \overline{y_i}\right)$ that substitutes the repeated measurements of growth curves. This is the reason why we fit 285 growth curves instead of 305.

The second step of module 1 is the model fitting and storing of results (`fit_store_models_1.py`). This process consists of (a) loading the data, (b) looping through each growth curve, (c) calculating initial values for all the models, (d) fit each model to the ith curve, and (e) storing results (parameter estimates, fit evaluations, AIC values, Akaike weights, and performance information) in a dictionary for later use in the analysis.

Steps a) and b) are trivial to perform. The initial values (step c) are set to 1 for the polynomial models. The convergence of the non-linear fits depends on the initial parameter values being reasonably close to the actual best-fit parameter values. Thus, the initial estimation for y_0 and y_{max} are the maximum and the minimum of the population data being fitted. We determine μ_{max} and t_{lag} by isolating the growth phase, and fitting a line to it. For a detailed explanation of how do we isolate the growth phase, and why fitting a line to it provides us with good initial parameter estimations, refer to section 6.2. Step d) is implemented using the non-linear least squares (NLLS) fitting method. For a brief discussion on the theory behind this method, refer to section 6.3. The AIC is calculated in step e. The expression for the AIC is

$$AIC = -2\log\left(\mathcal{L}(\boldsymbol{\beta}|x)\right) + 2p \tag{1}$$

where $\mathcal{L}$ is the likelihood of the model, and p the number of parameters. Under the assumptions that (1) we are using least squares to minimize our

residuals, and (2) the errors are normally distributed, the following holds.

$$\log\left(\mathcal{L}(\boldsymbol{\beta}|x)\right) = -\frac{n}{2} log\left(\frac{S}{n}\right) \tag{2}$$

where n is the number of data and S the residual sum of squares. Substituting the latter in equation 1 one obtains the expression used to calculate our AIC values.

$$AIC = n + 2 + n\log\left(\frac{2\pi}{n}\right) + n\log S + 2p \tag{3}$$

The Akaike weights w_i, where $i = (1, 2, ..., R)$ is the number of models, are calculated for the Gompertz, Baranyi and Buchanan models in the following way. [2]

First, we calculate the AIC differences $\Delta_i = AIC_i = AIC_{min}$ for each model. This allows us to calculate the likelihood of the model given the data as

$$\mathcal{L}(g_i|x) \propto \exp\left(-\frac{1}{2}\Delta_i\right) \tag{4}$$

Such likelihoods represent the relative strenght of evidence for each model. Finally, normalizing $\mathcal{L}(g_i|x)$ gives us a set of positive Akaike weights for each model

$$w_i = \frac{\exp\left(-\frac{1}{2}\Delta_i\right)}{\sum_{r=1}^{R} \exp\left(-\frac{1}{2}\Delta_r\right)} \tag{5}$$

The performance information (step e) provides the clasification *best, succesful, poor, bad, fail* in figure 4. Classifying a fit as *best, success, bad, or fail* is trivial, and it is explained in figure 4. To understand how can a fit be classified as *poor*, refer to section 6.4.

The third step of the first module is the plotting phase (`plotting_1.R`). Using the results previously obtained we plot every growth curve in a different

figure and overlay the top four fits to it, according to the AIC. The intensity of the color of each fitting curve is also based on its AIC value (see figure 6). All these plots are saved to the results directory. A fitting performance map (figure 4) is also generated here.

2.2 Module 2

This module deals with the biological question stated in section 1.

In the data preparation step (`data_preparation_2.R`), we select all the species of the genus *Arthrobacter* along with the temperature and the growth rate. Note that most of the fits to *Arthrobacter* species were flagged as *poor*. We perform model averaging to obtain model-averaged estimates of maximum growth rate for each species and temperature, ie

$$\hat{\mu}_{max} = \sum_{i=1}^{R} w_i \mu_{max,i} \tag{6}$$

In the end, we have a data frame with seven species, and five data points (T, $\hat{\mu}_{max}$) each.

Secondly, we perform NLLS fitting (`fit_store_models_2.py`) to each dataset using the thermal performance curve proposed by Lactin et al. in 1995 [10], which is a modification of the 1976 Logan-6 model [12]. For a brief discussion on the Lactin-2 model refer to section 6.5. In this case, the initial values for the fitting were taken from the literature [10]. T_{opt} is calculated for each species according to equation 18.

Third, we plot the results (`plotting_2.R`) to generate figure 5.

2.3 Computing tools

Modules 1 and 2 are written using Python and R. The latter one is used to perform the fitting using the LMFIT Python package [25]. NumPy, Pandas [26] and ProgressBar packages are also used for creating the models, saving the

results, and implementing a progress bar, respectively. R is used to prepare data for analysis and plotting using ggplot2 [27]. RColorBrewer package is loaded to enable the `YlOrRd` and `RdYlBu` color palettes used in the fit plots and the fitting performance tilemap (figure 4. The packages grid, and gridExtra are loaded to arrange multiple plots in one panel (figure 3).

We choose to divide our work like this because Python is a faster language than R in general, so we use it to do the heavy lifting (fitting algorithms). However, R's flexibility and intuitive behavior when it comes to working with data frames are much better than Python's. Moreover, ggplot2 is the most powerful tool to generate publishable plots to our knowledge, which justifies using R for plotting instead of something else.

The bash script `run_MiniProject.sh` is used to glue all the scripts together in a working full version of the mini-project that runs in less than 90 seconds. We keep our work under version control with GitHub.

3 Results

Two categories of results are reported in this section: model selection and biological question results.

Figure 4 summarizes the fitting performance and model selection results. The proportion of convergent fits decreases as the complexity of the model increases. Thus, all of the fits to the linear models converge. On the contrary, all the non-linear models fail once (logistic, Gompertz) or more times (Baranyi). Note that every time a non-linear fit fails, the best fit for that curve is not found among the other non-linear models, but instead in one of the linear models. The fits are generally better (in an AIC sense) in the non-linear models than in the linear ones. Particularly, the Gompertz model is the best most of the time (33%), followed by the Baranyi model (25%). The cubic model had a higher best model percent (18%) compared to the Buchanan (16%) and logistic (4%)

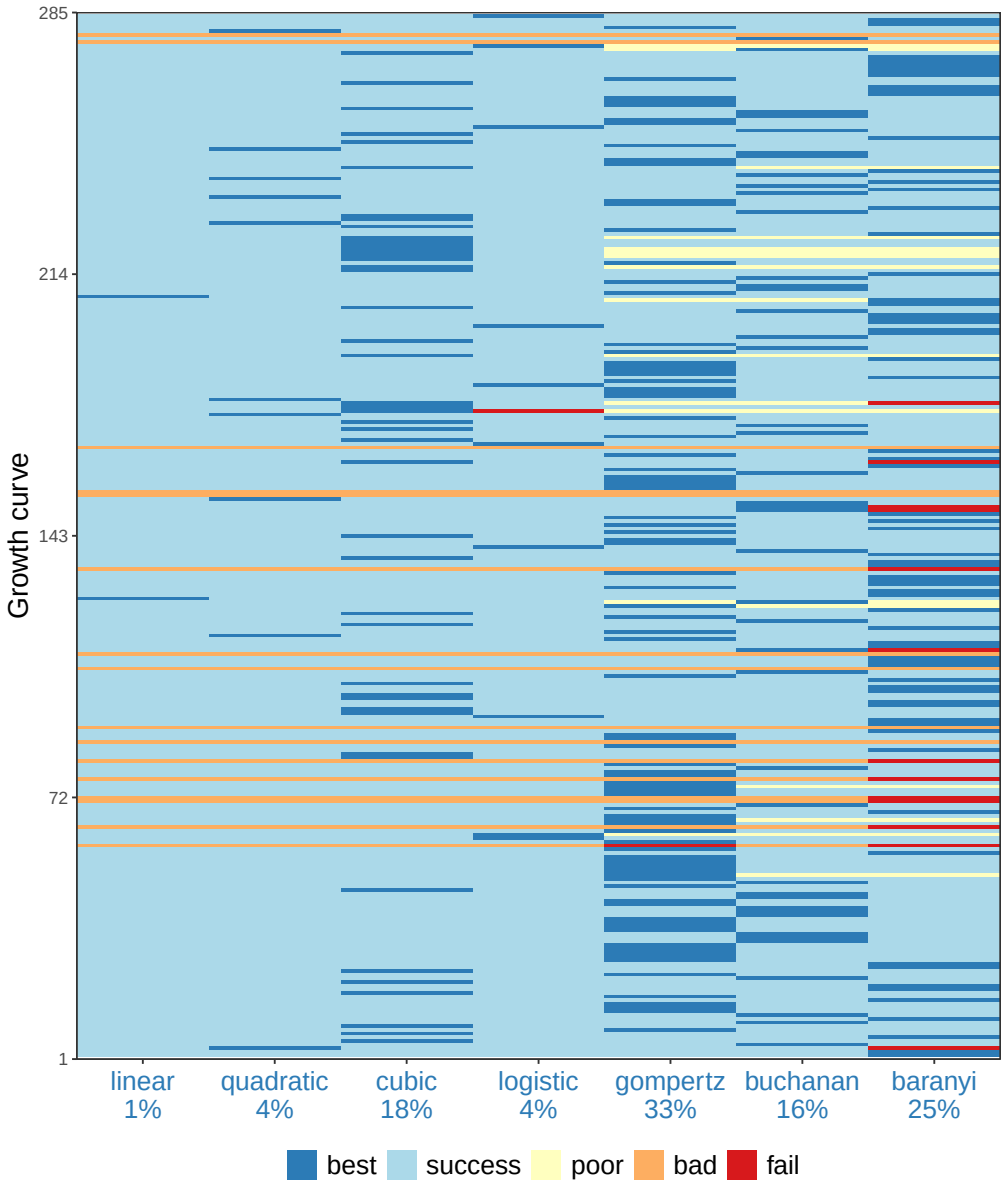

Figure 4: Tile map summarizing the fitting performance and model selection results. Best fits are determined by selecting the model with the lowest AIC value for each growth curve. The amount of times (%) that a model is the best candidate is specified in the x-axis. Success fits are convergent ones. Poor fits deal with data where the growth phase has ≤ 1 data points. Bad fits were deemed as such when no recognizable growth pattern was found upon visual inspection. Fail fits converge to a local minimum instead of the global one, or don't converge at all.

models. The linear (exponential) model was best the least amount of times (1%) and the quadratic model was it 4% of the times. Finally, a small proportion of

the logistic, Gompertz, Buchanan and Baranyi fits are poor fits. Refer to section 6.4 to know how is this determined.

The model-averaged μ_{max} estimations are calculated, and plotted against T.

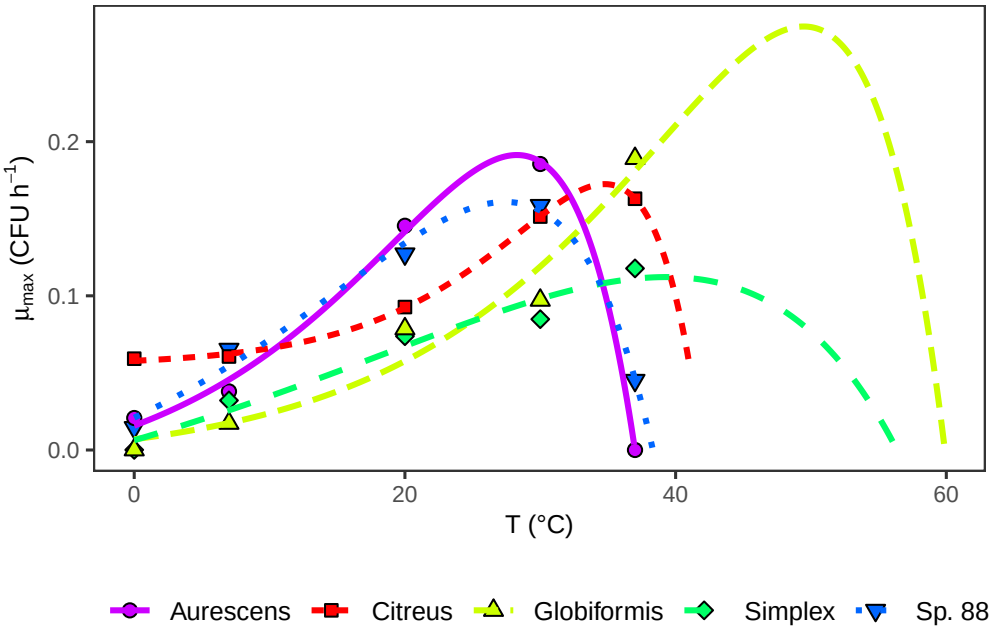

Figure 5: Fits of the Lactin-2 model to data from seven species of the genus *Arthrobacter*. Fits for species *Sp. 77* and *Sp. 62* have been omitted for clarity.

The Lactin-2 model is fitted for the data of each species (see figure 5). Calculating T_{opt} for each species gives 19.2°C for *Sp 62*, 23.9°C for *Sp 77*, 27.4 °C for *Sp 88*, 28.3°C for *Arthrobacter Aurescens*, 34.8°C for *Arthrobacter Citreus*, 39.3°C for *Arthrobacter Simplex*, and 49.5°C for *Arthrobacter Globiformis*.

4 Discussion

If there were no phenomena which were independent of all but a manageably small set of conditions, modeling in biology would be impossible [28]. Many models, each one considering a different set of relevant conditions, can successfully explain a given phenomenon. To obtain the best estimates for the parameters of interest one can test competing models using model selection or null hypothesis

testing. We believe that model selection is a better procedure to test competing hypotheses/models for several reasons. First, it allows for a comparison between more than two models (seven are analyzed in this work). Second, model selection is likelihood-based (as opposed to frequentist statistics based). Therefore, we can explicitly weigh the support for each model and even do model averaging if there is no obvious best model. This is something that cannot be done when using p-values to evaluate hypotheses because a p-value is not the probability that the null hypothesis is true, but instead, the degree of compatibility between a dataset and the null hypothesis [29, 30]. This implies that the winning hypothesis is only implicitly accepted, by explicitly rejecting the null hypothesis. Moreover, the rejection of the null hypothesis is based on an arbitrary threshold of the p-value (0.05, 0.01). Overall, the model selection approach offers a more robust, objective, and meaningful way to test several competing models, than the conventional frequentist statistics method of calculating p-values.

There are several approaches to model selection [3], namely, maximizing the quality of the fit, null hypothesis tests, and model selection criteria. The first one lacks a way of accounting for the complexity of the model, and the second one does not quantify the relative support among competing models, nor does it allow for comparison between more than 2 models. These limitations are overcome in model selection criteria. To determine the quality of the fits under these criteria, the most commonly used method is to calculate AIC or BIC. In this work, model selection was addressed using AIC. We chose to do so because the calculation of BIC assumes (a) the existence of a true model, (b) that the true model is within the examined pool of models, and (c) that each of the modes has the same probability of being the true one [3]. We deny assumption a), the existence of a true model, and consequently also assumptions b) and c) because we argue that *"all models are wrong, but some are useful"* [31, 32]. A true model, if attainable, would have to account for all the physical and biological processes that take place in the studied phenomena. This would make the

model intractable due to its complexity. A good model should be simple enough to capture the behavior of the modeled process in an intuitive way, but not so simple to disregard essential mechanisms [33].

When we calculate AIC values for each fit and model, we surprisingly find that the cubic model performs better than the logistic and Buchanan models. This is because the logistic and Buchanan only model the first three phases of a growth curve (figure 1), and don't focus on a possible death phase. On the contrary, a third-degree polynomial can model decreasing populations after y_{max} has been reached. Fitting a growth model that accounts for a death phase [34, 35] would eliminate this anomalous effect. However, those models are scarce in the food microbiology literature because most foods become inedible even before the mortality phase is reached [36], so there is no scientific incentive to come up with models that account for this.

The logistic model has an unexpected very low *best fit* percent. This result is caused by logging the logistic model. When we do so, the exponential growth that could reproduce the t_{lag} shape transforms into a straight line, which is unable to capture this pattern. From this point of view, it can be argued that fitting the logistic model in a linear scale is conceptually wrong since we are modeling a process of zero growth (lag phase), with slow exponential growth. Taking the logarithm of the data reveals this incongruence. To avoid it, some authors [22, 36] fit logged data with the linear logistic model. This is, in principle, a mistake, because when the data is rescaled the model must be rescaled too. However, the parameters of the logistic model can be recast to introduce a t_{lag} with no mechanistic interpretation [22]. The validity of this resides in how important it is to be able to mechanistically interpret the parameters of a model. This question is an ongoing debate. Some authors even question if any mechanistic insights can be grasped from the value of parameters obtained by curve fitting alone [36]. Being able to mechanistically interpret the parameters of a model depends on the purpose of your fitting. If it is merely for prediction,

the best model, (the Gompertz model in this work), should be the candidate of your choice. But what if there is no clear best model? A novel more robust method of obtaining parameter estimates for this purpose is model averaging. A discussion on why model averaging is used here is presented next, along with some remarks on the continuity of growth characteristics of some species in the genus *Arthrobacter*.

Most of the fits to *Arthrobacter* species are poor. As described in section 6.4, this implies that different models can perform well (have a high $\mathcal{L}(g_i|x)$) and still yield very different values of the parameters. How do we choose which parameter value to use? The answer resides in performing model averaging. Moreover, this approach is consistent with our skepticism about the existence of a true model, since our parameter estimates are a mixture of the parameter estimates of the tested models, weighted with the amount of support that we have for each of the candidates.

The results from figure 5 point out the difficulty of establishing an arbitrary definition of a psychrophile. Among the seven members of the genus *Arthrobacter* studied, no cut off region is apparent between psychrophiles and mesophiles. While we can clearly state that *Sp 77, 88* and *62* represent true mesophiles, according to the definition in [37] and our T_{opt} results. There can also be no doubt that *Simplex* and *Globiformis* are true psychrophiles. However, there *Citreus* and *Aurescens* occupy the middle region of the *Arthrobacter* temperature niche, and therefore, they cannot be categorized under any of the above categories.

The species *Simplex* and *Globiformis* didn't have data in the region $T > T_{opt}$. This caused the fit to yield parameters without biological meaning. To solve this problem, we impose for these two curves that T_{max} has to be no more than 60°C. No mesophile was found in the literature [38] wich survives at higher temperatures than the imposed maximum.

5 Aknowledgements

We would like to thank Sam Turner for fruitful discussions regarding how to summarize the fitting performance and how to define poor fits. We also thank Hovig Artinian for his thoughtful thoughts on model averaging.

6 Appendix

6.1 Models

In this section, we provide all the explicit forms of the fitted models.
The linear, quadratic, and cubic models have the form

$$y = y_0 + \mu_{max} t \quad , \quad y = a + bt + ct^2 \quad , \quad y = \hat{a} + \hat{b}t + \hat{c}t^2 + \hat{d}t^3 \tag{7}$$

The logistic model reads

$$y = \log_{10}\left(\frac{y_0 y_{max}}{y_0 + (y_{max} - y_0)e^{-\mu_{max} t}}\right) \tag{8}$$

For the Gompertz model, we have

$$y = y_0 + (y_{max} - y_0)\exp\left(-\exp\left(e\mu_{max}\ \frac{(t_{lag} - t)}{\log_{10}(y_{max} - y_0)} + 1\right)\right) \tag{9}$$

The Buchanan model can be expressed as

$$y = \begin{cases} y_0 & t < t_lag \\ y_{max} + \mu_{max}(t - t_{lag}) & t_{lag} \leq t \leq t_{max} \\ y_{max} & t \geq t_{max} \end{cases}$$

Finally, the Baranyi model has the form

$$y_{max} + \log_{10}\left(\frac{-1 + e^{\mu_{max} t_{lag}} + e^{\mu_{max} t}}{e^{\mu_{max} t} - 1 + e^{\mu_{max} t_{lag}} 10^{y_{max} - y_0}}\right) \tag{10}$$

6.2 Initial values calculation

Given m data points $D = \{(x_1, y_1), (x_2, y_2), ..., (x_m, y_m)\}$ from a bacterial growth curve, and a tolerance ϵ we perform the following steps to calculate initial estimations for the parameters μ_{max}, t_{lag}. First, we calculate the $m - 1$ dimensional vector gradient between consecutive points as $g_i = y_{i+1} - y_i$. Second, we

find the maximum of that vector, g_{max}. Third, we isloate the growing phase, $D_{grow} = \{(\tilde{x}_1, \tilde{y}_1), (\tilde{x}_2, \tilde{y}_2), ..., (\tilde{x}_n, \tilde{y}_n)\}$ by selecting the datapoints in D whith $g_i \in [g_{max}(1-\epsilon),\ g_{max}(1+\epsilon)]$, with $\epsilon = 0.3$ by default.

Fitting a line to D_{grow} yields the intercept and slope b, m that can be used to calculate the initial parameters. The growth rate can be expressed as the slope of the growing phase scaled by a constant due to having logged the data,

$$\mu_{max} = m \log(10) \tag{11}$$

The time lag has been traditionaly defined as the intersection of the tangents to the growth curve at the lag and exponential growth phases [36]. This can be expressed mathematically, as

$$t_{lag} = \frac{1}{m}(y_0 - b) \tag{12}$$

where y_0 is determined as described in section 2.1

6.3 Non-linear least sqares fitting

Consider m data points $\{(x_1, y_1), (x_2, y_2), ..., (x_m, y_m)\}$, and a model

$$y = f(x, \boldsymbol{\beta}) \tag{13}$$

where $\boldsymbol{\beta} = (\beta_1, \beta_2, ..., \beta_n)$ and $m \geq n$. The aim is to find $\boldsymbol{\beta}$ such as the model y fits best the given data points in the least square sense, i.e., the sum of squares

$$S = \sum_{i=1}^{m} r_i^2 \tag{14}$$

is minimized, where the residuals r_i are given by

$$r_i = y_i - f(x_i, \boldsymbol{\beta}) \tag{15}$$

Finding the minimum value of S is equivalent to solve n equations where the partial derivative of S with respect the parameter β_j, where $j = 1, ..., n$ equals 0. Thus

$$\frac{\partial S}{\partial \beta_j} = 0 \tag{16}$$

This yields a non-linear systems of equations that, in general, does not have a closed solution. Consequently, initial values must be chosen for the parameters, so that they can be iteratively minimized.

6.4 Classifying a fit as *poor*

A fit to a bacterial growth curve with only 1 data point in the growing phase will yield poorly constrained values μ_{max} and t_{lag}, because significantly different values of these parameters for different models will all define a good fit. To ilustrate this, see figure 6

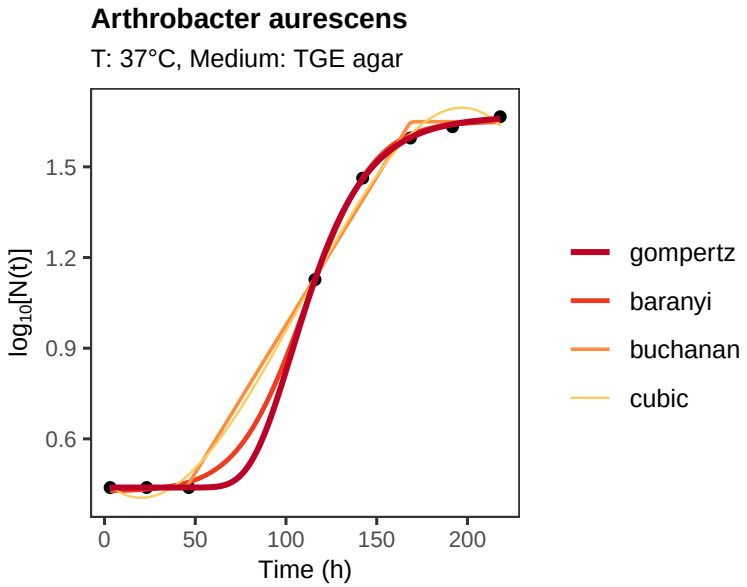

Figure 6: Example fit in which there is only 1 point in the growing phase, according to our method. The fit quality is very high ($R^2 \geq 0.999$). However, the values of μ_{max} and t_{lag} for Gompertz, Buchanan and Baranyi models are, respectively, [0.045, 0.022, 0.061] (CFU h^{-1}) and [80.3, 45.1, 87.8] (h).

Given m data points $D = \{(x_1, y_1), (x_2, y_2), ..., (x_m, y_m)\}$ from a bacterial growth curve, a tolerance ϵ and a bacterial growth model fit to that data, we flag it as *poor* if there are 1 or none data points in the growth phase. To do this, we first isolate the growth phase and then count the elements in it. Given the estimated parameters y_0 and y_{max} from the fits, the growth phase D_{grow} contains points of the form $(x_i, y_i)/y_i \in [y_0 + \epsilon\ |y_0|, y_{max} - \epsilon\ |y_{max}]$.

6.5 Lactin-2 model

The expression for the Lactin-2 model is

$$\mu(T) = \exp(\rho T) - \exp\left(\rho T_{max} - \frac{T_{max} - T}{\Delta T}\right) + \lambda \tag{17}$$

The modifications respect to the Logan-6 model are; first, they omitted the paremeter Ψ which was originally defined as a physiological rate parameter at a given base temperature. Second, they incorporated the intercept parameter λ, which forces the curve to intersect the abcissa at low temperatures and allows the estimation of T_{min}.

To find T_{opt} one finds the maximum of equation 17. Solving $\frac{d\mu}{dT} = 0$ yields the expression

$$T_{opt} = \frac{1}{\rho\Delta T - 1}\left[\rho T_{max}\Delta T - T_{max} + \Delta T \log\left(\frac{1}{\rho\Delta T}\right)\right] \tag{18}$$

References

[1] Moselio Schaechter. A brief history of bacterial growth physiology. *Frontiers in Microbiology*, 2015.

[2] Burnham & Anderson and David R. Anderson. *Model Selection and Inference: a Practical Information-theoretic Approach. New York: Springer.* 2nd edition, 2002.

[3] Jerald B. Johnson and Kristian S. Omland. Model selection in ecology and evolution, 2004.

[4] D. A. Ratkowsky, J. Olley, T. A. McMeekin, and A. Ball. Relationship between temperature and growth rate of bacterial cultures. *Journal of Bacteriology*, 1982.

[5] J. D. Phillips and M. W. Griffiths. The relation between temperature and growth of bacteria in dairy products. *Food Microbiology*, 1987.

[6] P.A.Gibbs C.J.Stannard, A.P.Williams. Temperature/growth relationships for psychrotrophic food-spoilage bacteria. *Food Microbiology*, Volume 2(Issue 2):115–122.

[7] Ana Paula Rosa da Silva, Daniel Angelo Longhi, Francieli Dalcanton, and Gláucia Maria Falcão de Aragão. Modelling the growth of lactic acid bacteria at different temperatures. *Brazilian Archives of Biology and Technology*, 2018.

[8] M. H. Zwietering, J. T. De Koos, B. E. Hasenack, J. C. De Wit, and K. Van 'T Riet. Modeling of bacterial growth as a function of temperature. *Applied and Environmental Microbiology*, 1991.

[9] Joey R. Bernhardt, Jennifer M. Sunday, and Mary I. O'connor. Metabolic theory and the temperature-size rule explain the temperature dependence of population carrying capacity. *American Naturalist*, 2018.

[10] D. J. Lactin, N. J. Holliday, D. L. Johnson, and R. Craigen. Improved rate model of temperature-dependent development by arthropods. *Environmental Entomology*, 1995.

[11] Äubomír Valík, Alžbeta Medveďová, Michal Čižniar, and Denisa Liptáková. Evaluation of temperature effect on growth rate of Lactobacillus rhamnosus GG in milk using secondary models. *Chemical Papers*, 2013.

[12] J. A. Logan, D. J. Wollkind, S. C. Hoyt, and L. K. Tanigoshi. An analytic model for description of temperature dependent rate phenomena in arthropods. *Environmental Entomology*, 1976.

[13] Sascha Krenek, Thomas U. Berendonk, and Thomas Petzoldt. Thermal performance curves of Paramecium caudatum: A model selection approach. *European Journal of Protistology*, 2011.

[14] N G ROTH. WHEATON RB: Continuity of psychrophilic and mesophilic growth characteristics in the genus Arthrobacter. *Journal of bacteriology*, 1962.

[15] Young Min Bae, Ling Zheng, Jeong Eun Hyun, Kyu Seok Jung, Sunggi Heu, and Sun Young Lee. Growth characteristics and biofilm formation of various spoilage bacteria isolated from fresh produce. *Journal of Food Science*, 2014.

[16] Liane Aldrighi Galarz, Gustavo Graciano Fonseca, and Carlos Prentice. Predicting bacterial growth in raw, salted, and cooked chicken breast fillets during storage. *Food Science and Technology International*, 2016.

[17] C. O. Gill and K. M. DeLacy. Growth of Escherichia coli and Salmonella typhimurium on high-pH beef packed under vacuum or carbon dioxide. *International Journal of Food Microbiology*, 1991.

[18] K. Sivonen. Effects of light, temperature, nitrate, orthophosphate, and bacteria on growth of and hepatotoxin production by Oscillatoria agardhii strains. *Applied and Environmental Microbiology*, 1990.

[19] M. H. Zwietering, J. C. De Wit, H. G.A.M. Cuppers, and K. Van't Riet. Modeling of bacterial growth with shifts in temperature. *Applied and Environmental Microbiology*, 1994.

[20] R. Pearl and L. J. Reed. On the Rate of Growth of the Population of the United States since 1790 and Its Mathematical Representation. *Proceedings of the National Academy of Sciences*, 1920.

[21] P F Verhulst. Notice sur la loi que la population suit dans son accroissement. *Correspondance Mathématique et Physique*, 1838.

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

[23] József Baranyi and Terry A. Roberts. A dynamic approach to predicting bacterial growth in food. *International Journal of Food Microbiology*, 1994.

[24] R. L. Buchanan, R. C. Whiting, 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*, 1997.

[25] Matthew Newville, Antonino Ingargiola, Till Stensitzki, and Daniel B. Allen. LMFIT: Non-Linear Least-Square Minimization and Curve-Fitting for Python. *Zenodo*, 2014.

[26] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, and Matt Haberland. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*, 2020.

[27] Hadley Wickham. *ggplot2 Elegant Graphics for Data Analysis*. 2016.

[28] E. P. Wigner. The Unreasonable Effectiveness of Mathematics in the Natural Sciences. In *Philosophical Reflections and Syntheses*. 1995.

[29] Ronald L. Wasserstein and Nicole A. Lazar. ASA Statement on Statistical Significance and P-Values: Context, Process, and Purpose. *The American Statistician*, 70(2):129–133, 2016.

[30] Jeehyoung Kim and Heejung Bang. Three common misuses of P values, 2016.

[31] George E.P. Box. Science and statistics. *Journal of the American Statistical Association*, 1976.

[32] G.E.P. Box. Robustness in the Strategy of Scientific Model Building. In *Robustness in Statistics*. 1979.

[33] Chris Chatfield. Model Uncertainty, Data Mining and Statistical Inference. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 1995.

[34] Micha Peleg, Maria G. Corradini, and Mark D. Normand. Isothermal and non-isothermal kinetic models of chemical processes in foods governed by competing mechanisms. *Journal of Agricultural and Food Chemistry*, 2009.

[35] J. Baranyi, A. Jones, C. Walker, A. Kaloti, T. P. Robinson, and B. M. Mackey. A combined model for growth and subsequent thermal inactivation of Brochothrix thermosphacta. *Applied and Environmental Microbiology*, 1996.

[36] Peleg Micha and Maria G. Corradini. Microbial growth curves: What the models tell us and what they cannot, 2011.

[37] J. L. INGRAHAM. Growth of psychrophilic bacteria. *Journal of bacteriology*, 1958.

[38] Chiara Schiraldi and Mario De Rosa. Mesophilic Organisms. In *Encyclopedia of Membranes*. 2014.