

IMPERIAL COLLEGE LONDON

COMPUTATIONAL METHOD IN ECOLOGY AND EVOLUTION

Thermal Response Model Selection and Analysis is Dependent on the Studied Traits and Temperature Range

Author: Danica Duan

CID: 01790819

January 22, 2021

Word Count: 2378

Abstract

Understanding thermal responses of metabolic rate is fundamental for studying the impacts of changing temperature from cellular level to the ecosystem. Among a wide range of published models, I fitted 3 phenomenological (cubic, quadratic and Briere) and 1 mechanistic (simplified Schoolfield) model in this project to a subset of bio-trait data. After model selection and comparison, the linear cubic model provides best simulation for temperature response of respiration, given most data collected under regular environmental temperature does not reach the optimum rate temperature for respiration; and the Schoolfield model provides best simulation for photosynthesis. Through parameter estimations of the Schoolfield model, the respiration shows higher sensitivity to high temperature than respiration, and the activation energy for metabolism of aquatic organism is significantly higher than that for terrestrial organisms.

1 Introduction

Since temperature is one of the major factors direct or indirectly affecting a wide range of biological rates, from biochemical reaction in cells to functions of the ecosystem (Brown et al. 2004, Schulte 2015). Understanding thermal responses of metabolic rate is fundamental for studying the impacts of changing temperature on major bioprocesses such as growth, respiration, and photosynthesis.

Chemical reaction rates are known to increase with rising temperature. According to the dominant theory of chemical kinetics — the Arrhenius equation (Menzinger & Wolfgang 1969), the relationship of reaction rate with temperature is governed by the energy barrier for molecule activation known as the activation energy (E_a) and the collision frequency of the activated molecules (A), which are both assumed to be independent of temperature. $k(T) = Ae^{-E_a/kT}$ is the equivalent version of the original Arrhenius equation, in which $k(T)$ is the reaction rate, T is temperature (K), and k is the Boltzmann constant ($8.617 \times 10^{-5} \text{ eV} \cdot \text{K}^{-1}$).

However, bioprocesses are dominated by enzyme reactions which only operates under optimal temperature range and denatures at a certain high and low temperature, therefore cannot be properly described by Arrhenius equation. Hence various mechanistic models have been proposed to describe the unimodal response of multiple bioprocesses with a wider range of temperature (Johnson & Lewin 1946, Rezende & Bozinovic 2019, Schoolfield et al. 1981), and to test and understand the underlying mechanistic theory of thermal performance among different metabolism, taxa, habitat, etc. (Dell et al. 2011) through predicting the activation energies and optimal temperature for metabolic rates. These estimations could vary significantly with habitats due to adaptation (Sage & Kubien 2007); and traits, due to the entirely different pathways and reactions within those two metabolisms (Abedon et al. 2008, Raven et al. 2005).

In addition to mechanistic models, phenomenological models (Boatman et al. 2017, Briere et al. 1999, Ratkowsky et al. 1983) were also widely developed to simulate the unimodal temperature response of bio-traits, which includes applying quadratic, Gaussian and modified Gaussian equations to empirical data (Montagnes et al. 2008).

The aim of this project is to compare and determine the suitability of these models to a subset of field and lab collected data. Research questions are: 1) Are non-linear model better for simulating the thermal responses of the dataset? 2) Is the mechanistic model suitable for simulating this dataset? 3) Do estimated parameter values from the mechanistic model differ with metabolic rates/habitats?

2 Methods

2.1 Data and Model Fitting

The given dataset is a subset of the “BioTraits” database, containing thermal responses of respiration and photosynthesis rates (net photosynthesis = gross photosynthesis - respiration) in different taxa of plant and bacteria from different habitats across the world. The trait values collected were calculated as the difference of measured with the reference data.

I first discarded all negative trait values, since negative metabolic rates of respiration and photosynthesis are not possible under real-life scenario. Then I applied two linear phenomenological models — quadratic and cubic equations, and two non-linear models the Briere model (phenomenological) and the Schoolfield model (mechanistic).

Since R has more statistic and graphing packages and is more well-used in data analysis, I used R for conducting this project. Function `lm` was used for fitting linear models and function `nlsLM` from package `minpack.lm` (Elzhov et al. 2016) was used for fitting non-linear models.

The Briere model (Briere et al. 1999) was developed aiming at reducing parameters in describing the non-linear relationship of traits to temperature.

$$B = \begin{cases} 0 & T \leq T_0 \\ B_0 T (T - T_0) \sqrt{T_m - T} & T_0 \leq T \leq T_m \\ 0 & T \geq T_m \end{cases} \quad (1)$$

In this equation, B represents the trait (such as growth, respiration or photosynthesis) value at given temperature, B_0 is the normalization constant and T_0 and T_m are the minimum and maximum feasible temperatures for the trait.

I used maximum and minimum temperature values of each data subset at each sample ID as their T_m and T_0 starting values. And I obtained starting B_0 values for the Briere model by: 1) input the trait values of data subsets of each sample ID into the Briere model equation; 2) then calculate the

mean value of all B_0 values obtained; 3) use this value as the mean value for a normal distribution with the absolute mean value for the standard deviation (since the differences of trait values are of magnitudes among samples), and sample the starting B_0 value from this normal distribution.

Since the T_m and T_0 values are the maximum and minimum feasible temperature, a wide enough boundary range is given to prevent good but biologically meaningless fits of the model. For the Briere model, I chose -80-40 °C for T_0 and 0-100 °C for T_m .

The Schoolfield model (Schoolfield et al. 1981):

$$B = \frac{B_0 e^{-\frac{E_a}{k}(\frac{1}{T} - \frac{1}{283.15})}}{1 + e^{\frac{E_l}{k}(\frac{1}{T_l} - \frac{1}{T})} + e^{\frac{E_h}{k}(\frac{1}{T_h} - \frac{1}{T})}} \quad (2)$$

In this equation, B_0 is the trait value at reference temperature, E_l and E_h are the low- and high-temperature deactivation energy, T_l and T_h are the low- and high-temperature at which the enzyme is 50% deactivated.

I applied the simplified version of Schoolfield model without low temperature deactivation since the data from low temperature is not sufficient.

$$B = \frac{B_0 e^{-\frac{E_a}{k}(\frac{1}{T} - \frac{1}{283.15})}}{1 + e^{\frac{E_h}{k}(\frac{1}{T_h} - \frac{1}{T})}} \quad (3)$$

The starting values of E_a and E_h for fitting the Schoolfield model were obtained from linear fitting of the log transformed version of Arrhenius equation with data subsets of before (for E_a) and after (for E_h) high temperature deactivation. Considering the trait values were collected by comparing to a reference value at a reference temperature of 10 °C, the log version of Arrhenius equation should be transformed as follow:

$$\ln k(T) = \ln B_0 - \frac{E}{kT} \quad (4)$$

$$\int_{T_{ref}}^T \frac{d \ln k(T)}{dT} = \frac{E}{kT^2} \quad (5)$$

Therefore getting $(k(T))$ as B and T_{ref} as 283.15 K):

$$\ln B = -\frac{E}{k}(\frac{1}{T} - \frac{1}{283.15}) + \ln B_0 \quad (6)$$

By fitting this equation with $\ln B$ value and $(-\frac{1}{kT} + \frac{1}{283.15})$, then getting the slope and intercept values, I got the starting values of both E and $\ln B_0$.

Bounds were selected to be slightly wider than reported normal ranges for activation/deactivation energies and temperature at maximum metabolic rates to eliminate biologically meaningless fittings (Doran 1995, Gillooly et al. 2001, Zwietering et al. 1991), as 273-333 K (0-60 °C) for T_h , 0-2 eV for E_a , and 0-5 eV for E_h .

2.2 Model Assessment

I applied Akaike information criterion (AIC) and Bayesian information criterion (BIC) as selection criteria for model assessment (Johnson & Omland 2004), among which BIC tends to favor simpler models with a punishment for model complexity in the calculation equation (as below). Small sample unbiased AIC (AICc) was inefficient due to the low sample size ($n \leq 5$) of large number of curves. The calculation equations are shown below, with $\ln[L(\hat{\theta}_y)] = -n/2\ln(RSS/n)$ for least squares model fits.

$$AIC = -2\ln[L(\hat{\theta}_y)] + 2p \quad (7)$$

$$BIC = -2\ln[L(\hat{\theta}_y)] + p \cdot \ln(n) \quad (8)$$

The performance of different models to data are evaluated with delta AIC and delta BIC methods, which is calculating the difference between candidate AIC (BIC) with the minimum AIC (BIC) values (the best fit). Since the lower AIC (BIC) the better, in this method, if the difference is less than 2, then it provides substantial evidence for the candidate model to be well supported by the data as well (Burnham & Anderson 2002, Raftery 1995).

2.3 Statistical Analysis of Model Prediction

Activation and deactivation energies, and the temperatures of maximum metabolic rates for the thermal response curves are estimated by fitting the Schoolfield model equation to the given data. These estimations were further tested by ANOVA (with pairwise t-test) and t-test to analyse the mean value and significance of differences among different habitats and between respiration and photosynthesis.

3 Results

Out of all 903 sample sets, 897 remained after discarding all negative metabolic rates of the data subset, 98.2% among which have more than 5 data points. 74.0% of temperature response data are unimodal shaped with a maximum temperature and decreasing rate values after high temperature deactivation, most of the rest are singularly increasing without reaching the half deactivation temperature, several has higher temperature sensitivity thus the metabolic rates were falling right from the start of the measurements.

A good example for simulation of all four models (quadratic, cubic, Briere and Schoolfield) to a unimodal temperature response curve of net photosynthesis rate (sample ID: 341, the shoots stage of a freshwater aquarium plant species *Vallisneria Americana*) is demonstrated in Figure 1. This example also shows the general model assessment pattern found throughout this project that the cubic and Schoolfield models are the best fit for the given data subset.

101 Judging by R square, 63.0% of quadratic models were fitted with R^2 of over 0.9, and 78.7% for
102 cubic models. Generally speaking, of the entire data subset, the performance of all four model equa-
103 tions can be ranked as: Cubic >Schoolfield >Quadratic >Briere. Model selection criteria AIC and
104 BIC gave the almost identical result considering a threshold of delta AIC and delta BIC to be over 2
105 for comparison, the number of best simulations for those four models to the data are 600, 386, 301
106 and 185 curves using AIC, and 553, 386, 305 and 189 curves using BIC for selection, out of all 897
107 sample datasets. (Figure 2)

108 Since BIC has a stricter punishment for model complicity, which is important for model selection,
109 I finally chose BIC model selection method for selecting the best performance of models. As shown in
110 Figure 3, the performance of the fitted four models differs highly between traits. The linear models
111 are largely outperforming the non-linear models in respiration rate data, possibly due to that as
112 observed in the plots of these datasets, a large number of samples were displaying a single increase
113 response, and a few were showing a concave shape with increasing temperature. While the non-
114 linear models performed better in photosynthesis rate datasets, it was especially noticed that the
115 mechanistic Schoolfield model outcompeted all other models when simulating temperature response
116 of photosynthesis.

117 Besides simulating and assessing the models of thermal performance curves, parameters estima-
118 tion from model simulation results is also important for further prediction and understanding of the
119 mechanisms of thermal responses. Here I will use parameters estimation from Schoolfield model fits
120 for a simple analysis on thermal performance.

121 Comparing between the two traits (Figure 4), temperature for maximum respiration rate (average
122 as 302.43 K) is significantly higher than the temperature for maximum photosynthesis rate (average
123 as 298.35 K) with a p value of 0.0007. And activation energy required for enzymatic reactions of
124 respiration is higher than photosynthesis with averages as 0.93 eV and 0.82 eV, with p value of 0.02.
125 However, the estimated deactivation energy cannot be used for comparison due to that large number
126 of respiration curves do not have measurements after deactivation, but the deactivation energy for
127 photosynthesis is estimated to be averagely 2.25 eV.

128 Comparing among habitats (Figure 5), especially between aquatic and terrestrial since the fresh-
129 water/terrestrial subset contains only 4 sample datasets, the activation energy for enzymatic reac-
130 tions of aquatic organisms is significantly higher than terrestrial organisms with average values of
131 0.96 and 0.82 eV (p value: 0.005); and deactivation energy of aquatic organisms is significantly higher
132 than terrestrial organisms, with average values of 2.70 and 1.89 (p value: 1.34e-12); no significant
133 different observed with temperature for maximum rate.

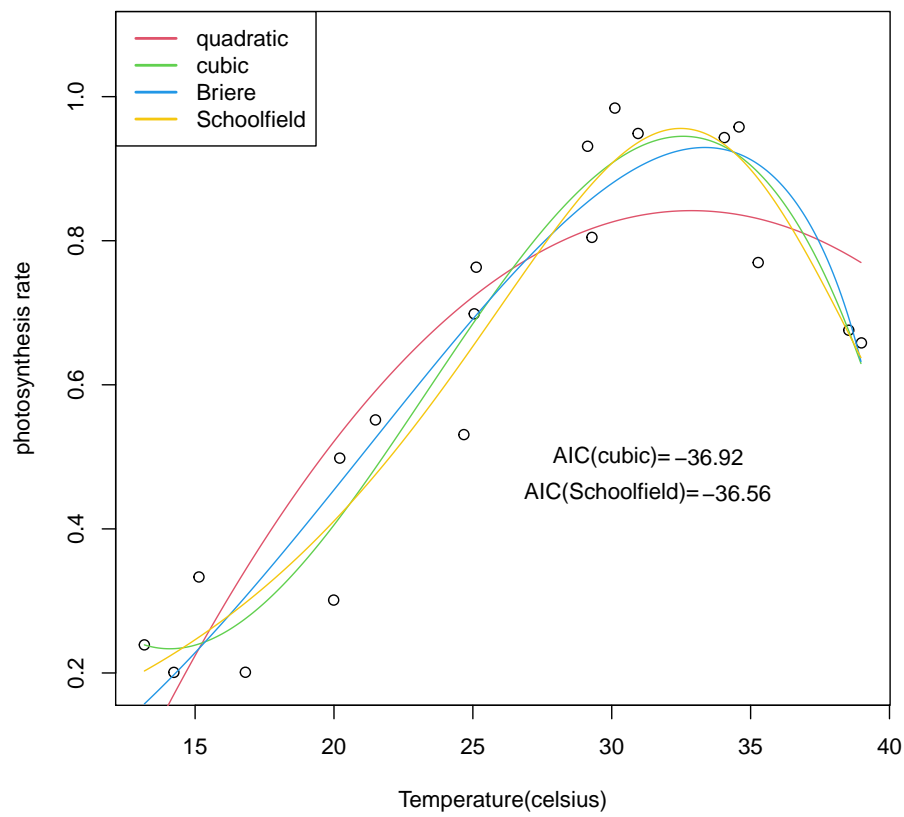


Figure 1: An example of all fitted models to a temperature response curve of net photosynthesis rate. Cubic (linear phenomenological) and Schoolfield (non-linear mechanistic) models are the better fits.

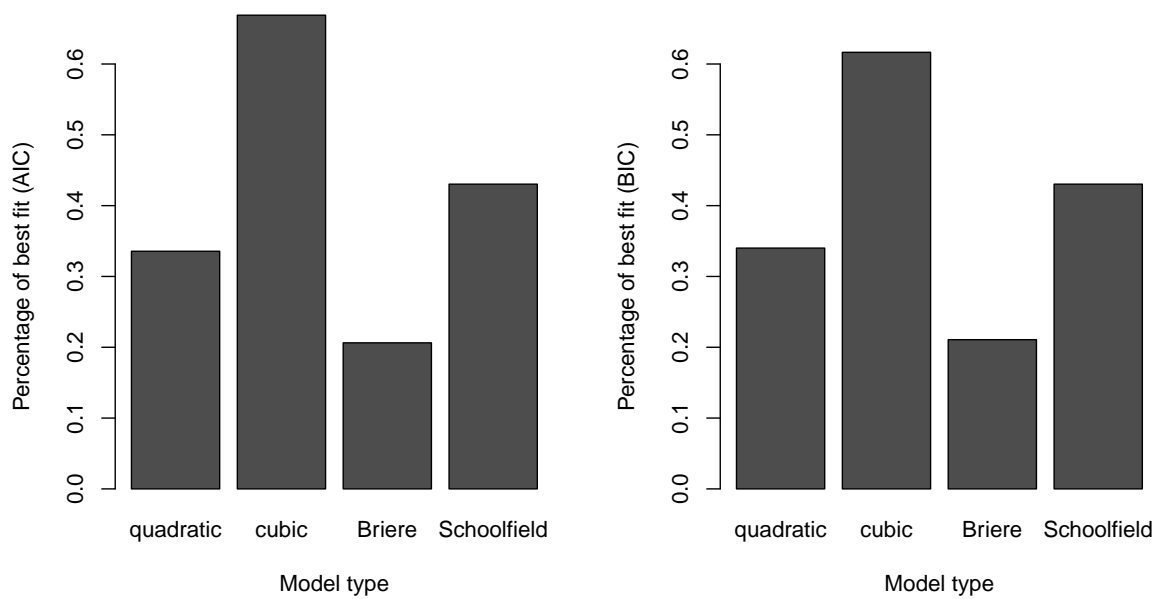


Figure 2: Best fitted models selected based on AIC (left) and BIC (right).

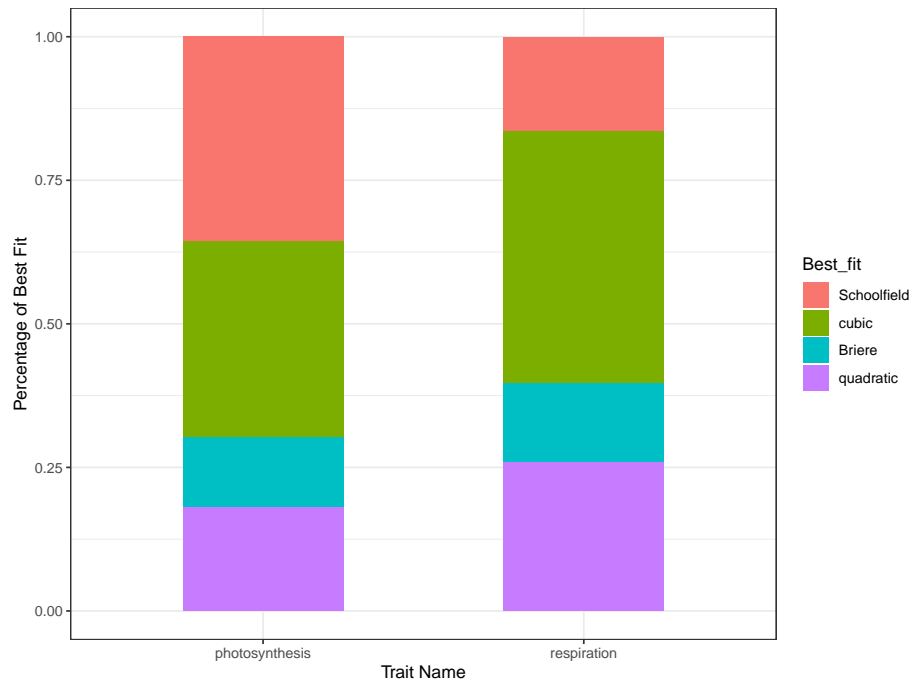


Figure 3: Best fitted models for respiration and photosynthesis, based on BIC selection.

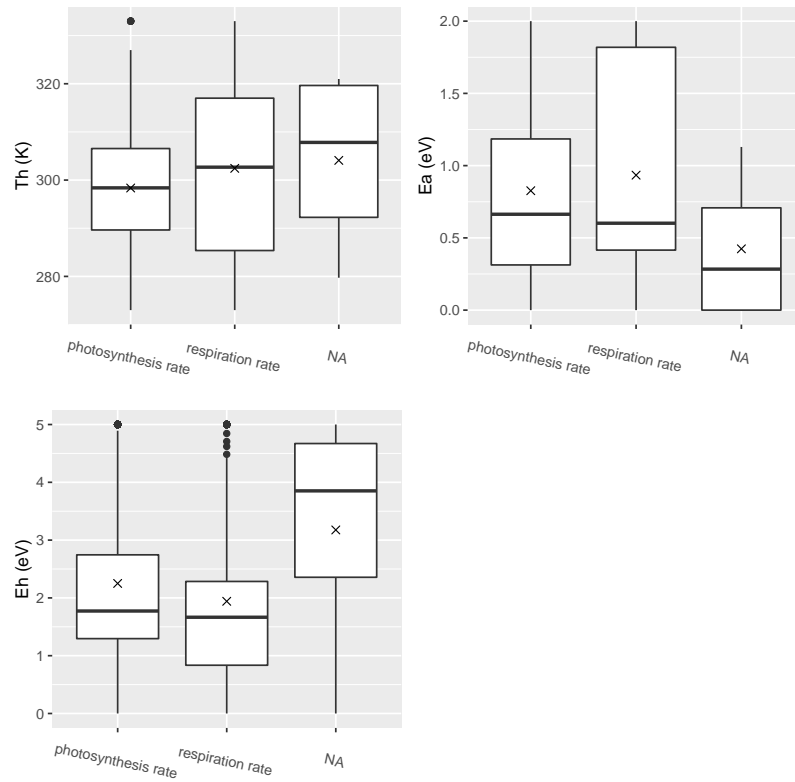


Figure 4: Estimation of optimal temperature, activation and deactivation energy for respiration and photosynthesis by Schoolfield model. (x marking the mean values)

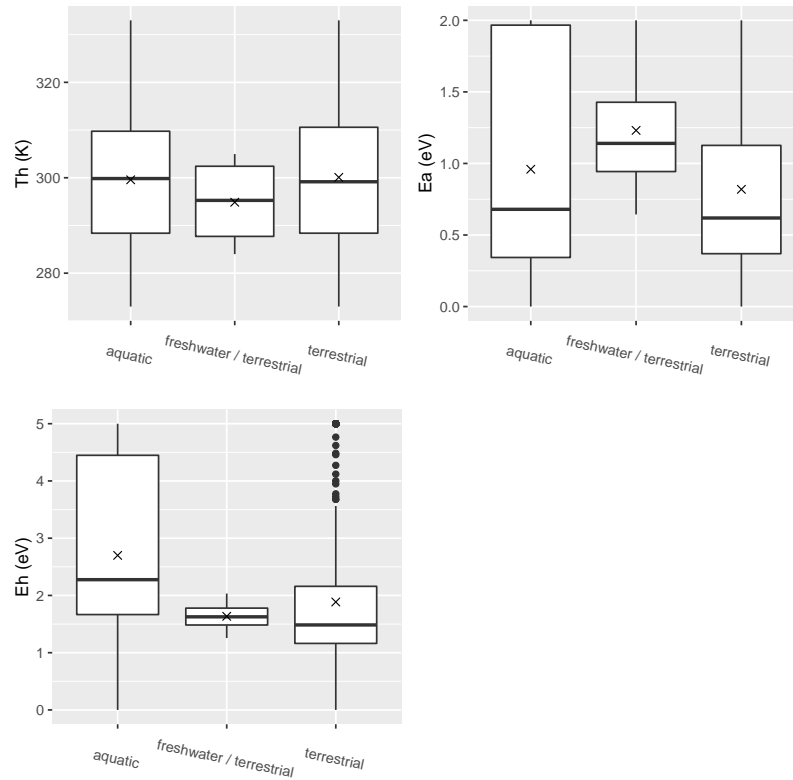


Figure 5: Estimation of optimal temperature, activation and deactivation energy for organisms from different habitats, by Schoolfield model. (x marking the mean values)

4 Discussion

In this project, I have fitted four models — three phenomenological and one mechanistic to collected field data subset to compare and determine the suitability of these models to the data and look for patterns in the estimation made by the mechanistic Schoolfield model.

The thermal performance curves normally show a similar shape that increases with temperature then rapidly decrease after reaching a maximum, this shape is usually resulted from a combination of thermodynamic effects of reaction rates and destabilization effects on intermolecular interactions by temperature (Schulte et al. 2011), so the non-linear models published all tend to describe a unimodal curve with a maximum trait value.

The thermal response of respiration is better simulated by the linear models, since normal respiration activity does not require to be at optimum level, the denaturation temperature for respiration is recorded as approximately 40 °C (Sage & Kubien 2007). However, most measurements conducted are around 20 °C (at normal environmental temperature) for this study, which did not reach the deactivation temperature for respiration activity, resulting in a singularly increasing curve for respiration temperature responses.

However, since temperature for optimum photosynthesis rate is much lower than respiration (Liu et al. 2018), the data collected for photosynthesis rate mostly contain values after the activation,

151 generating the unimodal thermal curves. In this case, the non-linear models prevailed, especially the
152 Schoolfield model.

153 Comparing AIC and BIC selection method, a slight reduction of best fits was observed for the
154 linear models in BIC. Since BIC method has a stronger punishment for model complexity; and by ob-
155 serving the original plots of data, unrealistic concave shapes, singularly falling data and completely
156 messy datasets can be noticed, this could infer an overfitting of the linear models, especially the
157 cubic.

158 Not all parameter estimation results can be adapted for further analysis. The average temper-
159 ature for maximum rate estimated for photosynthesis is approximately 25 °C, though the estimations
160 differ significantly between traits, this significance should not be taken into consideration for com-
161 paring due to the small temperature range of collected data. For the same reason, the deactivation
162 energy should not be considered neither for respiration data, but the activation energy for enzyme
163 deactivation of photosynthesis was estimated as 2.5 eV in this project, lies within the suggested nor-
164 mal range of approximately 1.8 eV to 4.1 eV for enzyme reactions Doran (1995).

165 The activation energy can be adapted, however the average values of activation energy of respi-
166 ration and photosynthesis estimated by the given data are higher than previously recorded values of
167 approximately 0.65 eV and 0.3 eV (Allen et al. 2005, Yvon-Durocher et al. 2012).

168 The general pattern for the estimation is that the activation energy and the maximum rate temper-
169 ature of respiration is significantly higher than that of photosynthesis, that photosynthesis has higher
170 temperature sensitivity, which is consistent with previous studies. And the average activation energy
171 for aquatic organism overall metabolism is significantly higher than that for terrestrial organism, but
172 the maximum rate temperature was not significantly different between these two habitats.

173 Further work should include fitting more appropriate mechanistic models to for better predic-
174 tions and understanding of the biological mechanisms, collecting data from a wider temperature
175 range, and conducting further analysis to discover patterns of temperature response among different
176 taxa or stages for this given data subset.

177 Overall, the linear models (especially cubic) can provide better simulation for temperature re-
178 sponse of respiration rates, however can result in slightly overfitting; and the Schoolfield model
179 simulates better for the photosynthesis rates. The general pattern obtained from Schoolfield model
180 estimation includes that respiration rate has higher temperature sensitivity than respiration, and the
181 activation energy of aquatic organism overall metabolism is significantly higher than that for terres-
182 trial organism.

References

- Abedon, S., Breitenberger, C., Roden, E. & Williams, J. (2008), Respiration, in S. E. Jørgensen & B. D. Fath, eds, 'Encyclopedia of Ecology', Academic Press, Oxford, pp. 3010 – 3020.
URL: <http://www.sciencedirect.com/science/article/pii/B9780080454054002913>
- Allen, A., Gillooly, J. & Brown, J. (2005), 'Linking the global carbon cycle to individual metabolism', *Functional Ecology* **19**(2), 202–213.
- Boatman, T. G., Lawson, T. & Geider, R. J. (2017), 'A key marine diazotroph in a changing ocean: The interacting effects of temperature, co2 and light on the growth of trichodesmium erythraeum', *PloS one* **12**(1), e0168796.
- Briere, J.-F., Pracros, P., Le Roux, A.-Y. & Pierre, J.-S. (1999), 'A novel rate model of temperature-dependent development for arthropods', *Environmental Entomology* **28**(1), 22–29.
- Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M. & West, G. B. (2004), 'Toward a metabolic theory of ecology', *Ecology* **85**(7), 1771–1789.
- Burnham, K. P. & Anderson, D. R. (2002), 'A practical information-theoretic approach', *Model selection and multimodel inference*, 2nd ed. Springer, New York **2**.
- Dell, A. I., Pawar, S. & Savage, V. M. (2011), 'Systematic variation in the temperature dependence of physiological and ecological traits', *Proceedings of the National Academy of Sciences* **108**(26), 10591–10596.
- Doran, P. (1995), 'Homogeneous reactions', *Bioprocess engineering principles*. Academic, San Diego pp. 277–278.
- Elzhov, T. V., Mullen, K. M., Spiess, A.-N., Bolker, B., Mullen, M. K. M. & Suggests, M. (2016), 'Package 'minpack.lm'', Title R Interface to the Levenberg-Marquardt Nonlinear Least-Squares Algorithm Found in MINPACK, Plus Support for Bounds'. Available at: <https://cran.rproject.org/web/packages/minpack.lm/minpack.lm.pdf>.
- Gillooly, J. F., Brown, J. H., West, G. B., Savage, V. M. & Charnov, E. L. (2001), 'Effects of size and temperature on metabolic rate', *science* **293**(5538), 2248–2251.
- Johnson, F. H. & Lewin, I. (1946), 'The growth rate of e. coli in relation to temperature, quinine and coenzyme', *Journal of Cellular and Comparative Physiology* **28**(1), 47–75.
- Johnson, J. B. & Omland, K. S. (2004), 'Model selection in ecology and evolution', *Trends in ecology & evolution* **19**(2), 101–108.

- 213 Liu, Y., He, N., Wen, X., Xu, L., Sun, X., Yu, G., Liang, L. & Schipper, L. A. (2018), 'The optimum tem-
214 perature of soil microbial respiration: Patterns and controls', *Soil Biology and Biochemistry* **121**, 35–
215 42.
- 216 Menzinger, M. & Wolfgang, R. (1969), 'The meaning and use of the arrhenius activation energy',
217 *Angewandte Chemie International Edition in English* **8**(6), 438–444.
- 218 Montagnes, D. J., Morgan, G., Bissinger, J. E., Atkinson, D. & Weisse, T. (2008), 'Short-term temper-
219 ature change may impact freshwater carbon flux: a microbial perspective', *Global Change Biology*
220 **14**(12), 2823–2838.
- 221 Raftery, A. E. (1995), 'Bayesian model selection in social research', *Sociological methodology* pp. 111–
222 163.
- 223 Ratkowsky, D., Lowry, R., McMeekin, T., Stokes, A. & Chandler, R. (1983), 'Model for bacterial
224 culture growth rate throughout the entire biokinetic temperature range.', *Journal of bacteriology*
225 **154**(3), 1222–1226.
- 226 Raven, P. H., Evert, R. F., Eichhorn, S. E. et al. (2005), *Biology of plants*, Macmillan.
- 227 Rezende, E. L. & Bozinovic, F. (2019), 'Thermal performance across levels of biological organization',
228 *Philosophical Transactions of the Royal Society B* **374**(1778), 20180549.
- 229 Sage, R. F. & Kubien, D. S. (2007), 'The temperature response of c3 and c4 photosynthesis', *Plant, cell*
230 *& environment* **30**(9), 1086–1106.
- 231 Schoolfield, R. M., Sharpe, P. & Magnuson, C. E. (1981), 'Non-linear regression of biological
232 temperature-dependent rate models based on absolute reaction-rate theory', *Journal of theoretical*
233 *biology* **88**(4), 719–731.
- 234 Schulte, P. M. (2015), 'The effects of temperature on aerobic metabolism: towards a mechanistic un-
235 derstanding of the responses of ectotherms to a changing environment', *Journal of Experimental*
236 *Biology* **218**(12), 1856–1866.
- 237 Schulte, P. M., Healy, T. M. & Fanguie, N. A. (2011), 'Thermal performance curves, phenotypic plastic-
238 ity, and the time scales of temperature exposure', *Integrative and comparative biology* **51**(5), 691–702.
- 239 Yvon-Durocher, G., Caffrey, J. M., Cescatti, A., Dossena, M., Del Giorgio, P., Gasol, J. M., Montoya,
240 J. M., Pumpanen, J., Staehr, P. A., Trimmer, M. et al. (2012), 'Reconciling the temperature depen-
241 dence of respiration across timescales and ecosystem types', *Nature* **487**(7408), 472–476.
- 242 Zwietering, M., De Koos, J., Hasenack, B., De Witt, J. & Van't Riet, K. (1991), 'Modeling of bacterial
243 growth as a function of temperature.', *Applied and environmental microbiology* **57**(4), 1094–1101.