A peer-reviewed version of this preprint was published in PeerJ on 9 February 2018.

View the peer-reviewed version (peerj.com/articles/4363), which is the preferred citable publication unless you specifically need to cite this preprint.

Kontopoulos D-G, García-Carreras B, Sal S, Smith TP, Pawar S. 2018. Use and misuse of temperature normalisation in meta-analyses of thermal responses of biological traits. PeerJ 6:e4363
https://doi.org/10.7717/peerj.4363

Use and misuse of temperature normalisation in meta-analyses of thermal responses of biological traits

Dimitrios - Georgios Kontopoulos[1,2,*]

Bernardo García-Carreras[2]

Sofía Sal[2]

Thomas P. Smith[2]

Samraat Pawar[2]

1. Science and Solutions for a Changing Planet DTP;

2. Department of Life Sciences, Imperial College London, Silwood Park, Ascot, Berkshire SL5 7PY, UK;

∗ Corresponding author; e-mail: d.kontopoulos13@imperial.ac.uk.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3068v1 | CC BY 4.0 Open Access | rec: 3 Jul 2017, publ: 3 Jul 2017

Abstract

There is currently unprecedented interest in quantifying variation in thermal physiology among organisms in order to understand and predict the biological impacts of climate change. A key parameter in this quantification of thermal physiology is the performance or value of a trait, across individuals or species, at a common temperature (temperature normalisation). An increasingly popular model for fitting thermal performance curves to data – the Sharpe-Schoolfield equation – can yield strongly inflated estimates of temperature-normalised trait values. These deviations occur whenever a key thermodynamic assumption of the model is violated, i.e. when the enzyme governing the performance of the trait is not fully functional at the chosen reference temperature. Using data on 1,758 thermal performance curves across a wide range of species, we identify the conditions that exacerbate this inflation. We then demonstrate that these biases can compromise tests to detect metabolic cold adaptation, which requires comparison of fitness or trait performance of different species or genotypes at some fixed low temperature. Finally, we suggest alternative methods for obtaining unbiased estimates of temperature-normalised trait values for meta-analyses of thermal performance across species in climate change impact studies.

Keywords: Sharpe-Schoolfield, model, thermal response, trait, rate, physiology, temperature.

Introduction

A rapidly growing body of empirical and theoretical research shows that climate change is likely to influence the dynamics of populations, communities and ecosystems by impacting the thermal physiology of individual organisms (Brown et al. 2004; Pörtner et al. 2006; Dell et al. 2011; Hoffmann and Sgrò 2011; Schulte et al. 2011; Garcia et al. 2014; Pawar et al. 2015). Therefore, to be able to predict climate change impacts, it is important to understand how biological traits such as respiration, photosynthesis, and population growth rate respond to changes in environmental temperature (the thermal performance curve, TPC; fig. 1).

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3068v1 | CC BY 4.0 Open Access | rec: 3 Jul 2017, publ: 3 Jul 2017

The TPCs of fundamental biological rates (traits) are unimodal, and trait-value versus temperature data are typically well-fitted by mathematical models that quantify four key features of the response: the temperature where the performance peaks (T_{pk}), the trait performance at a reference temperature (B_0), typically well below T_{pk} within its operational temperature range (Pawar et al. 2016), the rise of the trait up to T_{pk} (E), and the fall after T_{pk} (E_D) (fig. 1). The normalised trait value B_0 is particularly important, as it allows trait performance to be standardised for comparison across individuals and species (Gillooly et al. 2001). In particular, comparisons of normalised trait values at a reference temperature between species are key for studying metabolic cold adaptation (MCA; e.g., see Seibel et al. 2007; White et al. 2012).

Mechanistic models that explicitly link the trait's value to the temperature-dependence of the underlying biochemical kinetics (e.g., Johnson and Lewin 1946; Sharpe and DeMichele 1977; Schoolfield et al. 1981; Ikemoto 2005; Corkrey et al. 2012; Hobbs et al. 2013; DeLong et al. 2017) are becoming increasingly popular for quantifying empirically observed TPCs (Hochachka and Somero 2002). Among these, the Sharpe-Schoolfield model (Schoolfield et al. 1981) has been frequently used in recent studies to address both ecological and evolutionary questions about the effects of temperature change on individuals, populations, and communities (Barmak et al. 2014; Barneche et al. 2014; Fand et al. 2014; Simoy et al. 2015; Barneche et al. 2016; Padfield et al. 2016; Vimercati et al. 2016). In particular, the B_0 calculated from fitting this model to TPC data has been used to compare the trait performance of different species (e.g., Wohlfahrt et al. 1999), treatments (e.g., Padfield et al. 2016), or developmental stages (e.g., Hopp and Foley 2001) at a reference temperature, T_{ref}. However, the implicit assumption made by these studies, that B_0 is exactly the normalised trait value at T_{ref}, is only valid under certain conditions (see next section), and may in fact heavily overestimate the actual trait value at that temperature (Schoolfield et al. 1981) (fig. 1).

Here, we study the likely incidence of this overestimation of the normalised B_0 obtained by fitting the Sharpe-Schoolfield model to data. To this end, we investigate the conditions under which this overestimation becomes particularly pronounced by analysing 1,758 real thermal performance

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3068v1 | CC BY 4.0 Open Access | rec: 3 Jul 2017, publ: 3 Jul 2017

curves across diverse ectotherm species and traits. We then show how conclusions based upon biased B_0 estimates can compromise the results of an important application of TPC models — detecting metabolic cold adaptation. Finally, we present alternative methods for obtaining realistic estimates of trait performance at a reference temperature under different scenarios of usage of the model.

The Sharpe-Schoolfield model

The Sharpe-Schoolfield model proposes that the effect of temperature on the performance of a biological rate largely reflects the thermal sensitivity of a single rate-limiting enzyme that becomes deactivated at both extreme-high and low temperatures (Schoolfield et al. 1981). Because low-temperature enzyme inactivation is hard to detect, a simpler version of the full model that ignores low-temperature enzyme inactivation is most often used (fig. 1):

$$B(T) = B_0 \cdot \frac{e^{\frac{-E}{k}\cdot\left(\frac{1}{T}-\frac{1}{T_{\text{ref}}}\right)}}{1+e^{\frac{E_{\text{D}}}{k}\cdot\left(\frac{1}{T_{\text{h}}}-\frac{1}{T}\right)}}. \quad (1)$$

Here, B is the value of the trait at a given temperature T (K), E is the activation energy (eV), which controls the rise of the curve up to the peak, E_{D} is the de-activation energy (eV), which sets the rate at which the trait falls after the peak, T_{h} (K) is the temperature at which 50% of the enzyme units are inactive, and k is the Boltzmann constant ($8.617 \cdot 10^{-5}$ eV $\cdot$ K^{-1}). B_0 is the value of the trait at a reference (normalisation) temperature T_{ref} – i.e., $B_0 \approx B(T_{\text{ref}})$ – assuming enzyme units are fully operational at that temperature. The model can also be reformulated without normalisation, but then B_0 would lose any biological meaning (see section A2.1 in Appendix A).

Schoolfield et al. (1981) originally suggested using $T_{\text{ref}} = 25°$C, a choice they considered appropriate for most poikilotherm species. This suggestion has frequently been followed (e.g., Ungerer et al. 1999; Hopp and Foley 2001; Depinay et al. 2004; Barmak et al. 2014; Nealis and Régnière

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3068v1 | CC BY 4.0 Open Access | rec: 3 Jul 2017, publ: 3 Jul 2017

2014; Kang et al. 2015; Simoy et al. 2015; Padfield et al. 2016; see also table A1 and fig. A1 in Appendix A). However, when non-negligible loss of enzyme activity occurs at T_{ref} – e.g., due to denaturation or inactivation of some other component of the metabolic pathway – B_0 overestimates the real value of the trait at that temperature ($B(T_{ref})$) (Ikemoto 2005). This is particularly problematic for comparisons of B_0 across diverse species, as significant temperature-mediated inactivation may begin at very different temperatures, potentially leading to different degrees of inaccuracy in the B_0 estimates.

The inflation of trait value at reference temperature (B_0)

We first consider why B_0 can be biased. For this, in addition to the parameters in eq. (1) (B_0, E, E_D, T_h, T_{ref}), two extra parameters need to be defined to capture all aspects of the shape of the TPC: the temperature at which the TPC peaks (T_{pk}), and the performance at that peak (P_{pk}; see sections A2.2-3 in Appendix A for their derivations). Setting $T = T_{ref}$ in eq. (1) shows that the amount by which B_0 will deviate from $B(T_{ref})$ is equal to the denominator of eq. (1):

$$B(T_{ref}) = B_0 \cdot \frac{1}{1 + e^{\frac{E_D}{k} \cdot \left(\frac{1}{T_h} - \frac{1}{T_{ref}}\right)}} \quad (2)$$

When T_{ref} is much lower than T_h (the temperature at which 50% of the enzyme units become inactive), $B_0 \approx B(T_{ref})$ because the denominator ≈ 1. On the other hand, as the chosen T_{ref} approaches T_h, B_0 will increasingly deviate from $B(T_{ref})$. To explore this behavior numerically across real TPCs of a single biological rate (for consistency reasons), we compiled a dataset of phytoplankton growth rates versus temperature (a combination of the López-Urrutia et al. 2006, Rose and Caron 2007, Bissinger et al. 2008, and Thomas et al. 2012 datasets), containing 672 species/strains. To each TPC in this dataset, we fitted the Sharpe-Schoolfield model across a range of T_{ref} values (-10°C to 30°C) using the nonlinear least-squares method (Levenberg-Marquardt algorithm; com-

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3068v1 | CC BY 4.0 Open Access | rec: 3 Jul 2017, publ: 3 Jul 2017

puter code available at https://github.com/dgkontopoulos/Kontopoulos_et_al_temperature_normalisation_2017). In order to eliminate less reliable fitted parameter estimates, we rejected fits with i) an R^2 below 0.5 or ii) fewer than four data points either before or after T_{pk}.

Plotting the fold increase of B_0 from $B(T_{ref})$ against the difference of T_{ref} from T_h for this dataset reveals a negative, nonlinear relationship (fig. 2A). Moreover, in many circumstances, the deviation of B_0 is extreme, becoming even greater than the trait value at or near optimum temperature, P_{pk} (fig. 2B).

Conditions leading to a severely overestimated B_0

We next determine the characteristics of TPCs (parameter combinations of the Sharpe-Schoolfield model) that lead to a severely overestimated B_0. This is a complex problem and not just a matter of determining the difference between T_h and T_{ref}, because the denominator of eq. (2) also includes the E_D parameter. As E_D influences the relationship between T_h and T_{pk} (see section A2.2 in Appendix A), it is necessary to take into account the interplay of T_h and T_{ref} with T_{pk}. To address this, we use a machine learning approach to determine the TPC model's parameter combinations that lead to strong overestimation, again using a large empirical dataset.

For maximising the power of the machine learning method we used a larger dataset — Biotraits (Dell et al. 2013) combined with additional data extracted from the published literature (see section A4 in Appendix A). We first fitted the Sharpe-Schoolfield model to each empirical TPC in this dataset. As the dataset is very diverse – including, among others, traits from bacteria, macroalgae, and terrestrial plants – we set T_{ref} to 0°C so that we could obtain reasonable estimates (i.e., at a temperature below T_{pk}) of B_0 and $B(T_{ref})$ even for cold-adapted species with low T_{pk} values. In total, 1,758 species/individual curves were produced from this dataset. We did not filter the results based on goodness of fit metrics because we are interested in all the different parameter combinations regardless of how well they describe the data. Only curves which estimated B_0 or $B(T_{ref})$ at values indistinguishable from zero were rejected.

We then analysed this ensemble of fitted curves through the construction of a conditional in-

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3068v1 | CC BY 4.0 Open Access | rec: 3 Jul 2017, publ: 3 Jul 2017

ference tree (Hothorn et al. 2006) from the data. In short, conditional inference trees attempt to classify a response variable based on combinations of other predictor variables. This is done through a tree-like structure (see fig. 3) where terminal nodes (or leaves) are the predicted classifications, whereas all other nodes up to the root of the tree correspond to predictor variables. Essentially, each internal node is a decision point in the model at which two alternative paths can be followed, depending on the value of the corresponding predictor. The topology of the tree is constructed through a statistical significance approach, consisting of non-parametric tests for each particular node, corrected for multiple testing to avoid overfitting. As a result, the predictors are unbiased and the model does not require cross-validation, often performed in other classes of models to prevent overfitting. Overall, conditional inference trees allow identification of complicated nonlinear associations between a response variable and its predictors, which are otherwise quite difficult to distinguish.

In our case, we fitted a conditional inference tree model with a binary response variable: B_0 is above or below P_{pk}. The choice of P_{pk} as the cutoff was due to the very high classification performance of the resulting model, especially when compared to other possible cutoffs (e.g., a three-fold increase from $B(T_{\mathrm{ref}})$) which performed poorly. The predictor variables were the differences between i) T_{pk} and T_{h}, ii) T_{pk} and T_{ref}, and iii) T_{h} and T_{ref} for each fit. The model was constrained by setting the maximum allowed p-value at each internal node below 10^{-10}. Its quality was evaluated with the Matthews correlation coefficient (MCC; Matthews 1975), a metric often used for machine learning models with a binary response. This metric takes values from -1 (complete disagreement with data) to 1 (complete agreement with data) and is considered reliable even when the different response states of the model (in this case $B_0 > P_{\mathrm{pk}}$ and $B_0 < P_{\mathrm{pk}}$) are not evenly sampled. To further ensure that the model was accurate and generalisable, we also estimated its performance against a distinct dataset of 405 TPCs (testing dataset). The data for these curves originated from the literature - similarly to the 1,758 curves - but were not used for training the model.

The resulting conditional inference tree consisted of four terminal nodes, with B_0 being nearly

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3068v1 | CC BY 4.0 Open Access | rec: 3 Jul 2017, publ: 3 Jul 2017

exclusively below or above P_{pk} in each one of them (fig. 3). The model exhibited high performance both on the training dataset (MCC = 0.954) and the testing dataset (MCC = 0.824; section A3 in Appendix A). The sets of thermal response parameters in which B_0 was greater than P_{pk} almost always had either a $T_h - T_{ref}$ difference that was less than 0.6 (relatively narrow curves), or a $T_{pk} - T_{ref}$ difference of 49.1 or lower (relatively wide curves).

Implications of the inflation for investigations of thermal adaptation

Among other ecological and evolutionary questions, the effects of adaptation to different thermal environments on the shape of the TPC (e.g., see Huey and Kingsolver 1989; Angilletta et al. 2003; Angilletta 2009; Angilletta et al. 2010) can be investigated using estimates from the Sharpe-Schoolfield model. For example, a study may aim to uncover whether there are any trade-offs between performance at lower and higher temperatures by correlating B_0 and T_{pk} (e.g., a negative correlation would suggest that high performance at warmer temperatures would come at the cost of lower performance at colder temperatures). Overestimating B_0 – especially for cold-adapted species with a T_h value close to T_{ref} – may potentially introduce such correlations where none existed, serving as false-positive evidence for the MCA hypothesis.

To explore this possible issue, we generated a synthetic dataset of 1,000 negatively skewed TPCs, in which MCA was absent. While a real-world dataset of a trait could also be used for this purpose (e.g., the phytoplankton growth rates dataset in fig. 2), we resorted to a simulation in order to obtain a bigger sample and, more importantly, to ensure that the input data were not even in the slightest an outcome of MCA. To this end, each curve was obtained by sampling from a distinct realisation of the beta distribution, with shape parameters that were in turn sampled from normal distributions (table 1). Curves that were not negatively skewed were removed and new ones were produced in their place. We also randomly varied the width and the height of the curves. In this population of curves, there was no significant association between the performance at a T_{ref} of 7°C

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3068v1 | CC BY 4.0 Open Access | rec: 3 Jul 2017, publ: 3 Jul 2017

(T_{ref} was set below the minimum T_{pk} of 8.23°C, but could be set even lower), and the thermal optimum ($r = -0.03$, 95% CI = -0.09 to 0.03, $p = 0.35$).

We then fitted the Sharpe-Schoolfield model to each synthetic curve and obtained parameter estimates where possible. To assess the impact of the deviation of B_0 from $B(T_{ref})$, we performed two different tests for MCA using B_0 and $B(T_{ref})$ estimates, and compared their results. For the first test, the estimates were split into two groups: i) those originating from curves with a T_{pk} less than 15°C (from hypothetical cold-adapted species) and ii) those with higher T_{pk} values. We next investigated whether the normalised trait distributions varied between the two groups using the two-sample Kolmogorov-Smirnov test (Corder and Foreman 2014). The second test was a simple correlation of normalised trait values with the corresponding T_{pk} values.

In total, we were able to obtain thermal response parameter estimates for 968 simulated curves, as the nonlinear least-squares algorithm failed to converge on solutions for the remaining 32. In the first test for MCA the distributions of B_0 estimates differed between the two groups ($D = 0.18$, $p = 1.7 \cdot 10^{-6}$), with species adapted to colder temperatures having a higher median value of B_0 (fig. 4A, light blue violin plots). In contrast, the two distributions of $B(T_{ref})$ estimates were statistically indistinguishable ($D = 0.07$, $p = 0.21$), as expected (fig. 4A, orange violin plots). The overestimation of B_0 also affected the second MCA test, as a negative correlation between B_0 and T_{pk} was detected, but not between $B(T_{ref})$ and T_{pk} (fig. 4B). These results indicate that the inflation of B_0 can provide false support for the MCA hypothesis, even for datasets with complete absence of this pattern.

Discussion

In this paper we have addressed the consequences of estimating the value of a trait at a reference temperature, B_0, using the Sharpe-Schoolfield model, but without satisfying one of its fundamental assumptions: that the key enzyme – which is responsible for the temperature dependence of the trait – is fully functional at the reference temperature. When this assumption is not met, B_0 will

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3068v1 | CC BY 4.0 Open Access | rec: 3 Jul 2017, publ: 3 Jul 2017

overestimate the real trait performance at the reference temperature, $B(T_{ref})$.

We explain how this discrepancy arises and determine the conditions under which it becomes particularly pronounced using a machine learning approach (fig. 3). This shows that B_0 estimates will generally exceed the trait performance at the peak of the curve (P_{pk}) as long as: i) $T_{pk} - T_h$ is less than $\sim 37.58°C$ and $T_h - T_{ref}$ is less than $\sim 0.6°C$, or ii) $T_{pk} - T_h$ is greater than $\sim 37.58°C$ and $T_{pk} - T_{ref}$ is less than $\sim 49.11°C$. In any other case, B_0 would most likely be smaller than P_{pk}, although its inflation may well still be of concern.

Using a synthetic dataset, we then demonstrate that wrongly assuming $B_0 = B(T_{ref})$ can lead to erroneous conclusions in analyses of thermal adaptation, as the overestimation of B_0 can mimic the effects of metabolic cold adaptation (fig. 4) (a Type I error).

As mentioned before, previous studies have tended to set the T_{ref} – usually at a value of 25°C – while fitting the Sharpe-Schoolfield model without considering the potential inflation of B_0 (table A1 and fig. A1, Appendix A). Whether results of these studies have been compromised by an inappropriate use of T_{ref} is impossible to determine definitively because most of these studies report either T_h or T_{pk} estimates, whereas the machine learning model depends on both (see the 'Conditions leading to a severely overestimated B_0' section), along with the value of T_{ref}. If these data were available, using the machine learning model that we generated would provide a straightforward procedure to identify cases where B_0 is highly likely to be extremely overestimated (i.e, greater than P_{pk}). In fact, the only study where all necessary parameter estimates were reported for all fitted curves was that by Padfield et al. (2016). In that study, the maximum difference of T_h from T_{pk} is 2.49°C, and the minimum difference of T_{ref} from T_h is 5.79°C, which, according to the machine learning model (see fig. 3), are sufficient for the B_0 estimates to be below the P_{pk} ones. Having said that, as we showed in this paper, the fact that the overestimation of B_0 is not extreme does not necessarily rid any drawn conclusions of bias.

Based on all the aforementioned results, we provide the following suggestions for future studies that will be using the Sharpe-Schoolfield model.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3068v1 | CC BY 4.0 Open Access | rec: 3 Jul 2017, publ: 3 Jul 2017

Comparisons of temperature-normalised rates of diverse species

When data span the entire TPC

For studies in which the end goal is to compare the performance of different species at a common temperature, the simplest approach would be to fit the Sharpe-Schoolfield model - with or without normalising B_0 at a reference temperature - and compare estimates of $B(T_{\text{ref}})$, calculated a posteriori. The confidence intervals around $B(T_{\text{ref}})$ can then be estimated by bootstrapping. Other models could also be considered, such as the macromolecular rates model (Hobbs et al. 2013) or the enzyme-assisted Arrhenius model (DeLong et al. 2017).

When data only cover the rising part of the TPC

While the previous solutions are applicable to thermal response datasets that capture either the rise of the curve or its entirety, few studies report temperature performance measurements after the unimodal peak of the response. Therefore, to obtain an estimate of baseline performance from a dataset that only covers the exponential rise component, one could instead fit the Boltzmann-Arrhenius model (e.g., see Gillooly et al. 2001),

$$B(T) = B_0 \cdot e^{\frac{-E}{k} \cdot \left(\frac{1}{T} - \frac{1}{T_{\text{ref}}}\right)}, \tag{3}$$

which does not suffer from the problems of the Sharpe-Schoolfield model, as $B(T_{\text{ref}})$ indeed simplifies to B_0.

A second alternative model is the one that includes the Q_{10} factor (see Gillooly et al. 2001), i.e. the rate of change in trait performance after a temperature rise of 10°C:

$$Q_{10} = \left(\frac{B(T_2)}{B(T_1)}\right)^{\frac{10}{T_2 - T_1}}. \tag{4}$$

In this case, one would first estimate the value of Q_{10} from known trait values at two temperatures,

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3068v1 | CC BY 4.0 Open Access | rec: 3 Jul 2017, publ: 3 Jul 2017

and use it to calculate the trait value at the reference temperature:

$$B(T_{\text{ref}}) = B(T_1) \cdot Q_{10}^{\frac{T_{\text{ref}} - T_1}{10}}. \quad (5)$$

Regardless of which of these two models is chosen, careful attention must be paid to ensure that the biological rate increases exponentially across the entire temperature range, without signs of a plateau being reached. Otherwise, the estimates may yet again be biased.

Using the 'intrinsic optimum temperature' instead of T_{ref}

Alternatively, baseline performance could be defined as the height of the curve at the temperature where the population of the key enzyme is fully active, which should be characteristic for each individual or species. In the Sharpe-Schoolfield model, the denominator indicates the percentage of enzymes that are active. Therefore, the intrinsic optimum temperature could be estimated as the temperature at which this percentage is 100 (or, at least, sufficiently high). Otherwise, this temperature can also be obtained from the Sharpe-Schoolfield-Ikemoto (SSI) model (Ikemoto 2005). This model integrates the law of total effective temperature - often used in studies of arthropod or parasite development - within the Sharpe-Schoolfield model, replacing T_{ref} with the intrinsic optimum temperature. However, this model introduces an extra parameter and is more challenging to fit compared to the original Sharpe-Schoolfield model. To mitigate this problem, software implementations have been developed that reduce the computation time from often more than 3 hours (Ikemoto 2008) down to less than a second (Shi et al. 2011; Ikemoto et al. 2013).

Conclusions

Obtaining accurate estimates of temperature-normalised trait performance is of crucial importance – especially in the face of climate change – for comparisons of the same trait across different organisms, or different traits within an individual. In this context, our study explains why temperature-

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3068v1 | CC BY 4.0 Open Access | rec: 3 Jul 2017, publ: 3 Jul 2017

normalised trait estimates can be strongly exaggerated when one of the assumptions of the Sharpe-Schoolfield model is violated, and gives an example of possible consequences of this exaggeration. The suggestions that we provide to address this issue should be useful to the burgeoning studies on ectotherm thermal performance and climate change, both for performing meta-analyses and for determining appropriate temperature ranges in laboratory experiments.

Acknowledgements

DGK was supported by a Natural Environment Research Council (NERC) Doctoral Training Partnership (DTP) scholarship. TPS was supported by a Biotechnology and Biological Sciences Research Council (BBSRC) DTP scholarship. BGC, SS, and SP were supported by a NERC grant awarded to SP (NE/M004740/1).

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3068v1 | CC BY 4.0 Open Access | rec: 3 Jul 2017, publ: 3 Jul 2017

References

Angilletta, M. J. 2009. Thermal adaptation: a theoretical and empirical synthesis. Oxford University Press.

Angilletta, M. J., R. B. Huey, and M. R. Frazier. 2010. Thermodynamic effects on organismal performance: is hotter better? Physiological and Biochemical Zoology 83:197–206.

Angilletta, M. J., R. S. Wilson, C. A. Navas, and R. S. James. 2003. Tradeoffs and the evolution of thermal reaction norms. Trends in Ecology & Evolution 18:234–240.

Barmak, D. H., C. O. Dorso, M. Otero, and H. G. Solari. 2014. Modelling interventions during a dengue outbreak. Epidemiology and Infection 142:545–561.

Barneche, D., M. Kulbicki, S. Floeter, A. Friedlander, and A. Allen. 2016. Energetic and ecological constraints on population density of reef fishes. Proceedings of the Royal Society of London. Series B, Biological sciences 283:20152186.

Barneche, D. R., M. Kulbicki, S. R. Floeter, A. M. Friedlander, J. Maina, and A. P. Allen. 2014. Scaling metabolism from individuals to reef-fish communities at broad spatial scales. Ecology Letters 17:1067–1076.

Bissinger, J. E., D. J. Montagnes, D. Atkinson, et al. 2008. Predicting marine phytoplankton maximum growth rates from temperature: Improving on the Eppley curve using quantile regression. Limnology and Oceanography 53:487–493.

Brown, J. H., J. F. Gillooly, A. P. Allen, V. M. Savage, and G. B. West. 2004. Toward a metabolic theory of ecology. Ecology 85:1771–1789.

Corder, G. W., and D. I. Foreman. 2014. Nonparametric Statistics: A Step-by-Step Approach, 2nd Edition. John Wiley & Sons.

Corkrey, R., J. Olley, D. Ratkowsky, T. McMeekin, and T. Ross. 2012. Universality of thermodynamic constants governing biological growth rates. PLoS One 7:e32003.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3068v1 | CC BY 4.0 Open Access | rec: 3 Jul 2017, publ: 3 Jul 2017

306 Dell, A. I., S. Pawar, and V. M. Savage. 2011. Systematic variation in the temperature dependence of physiological and ecological traits. Proceedings of the National Academy of Sciences 108:10591–10596.

309 ———. 2013. The thermal dependence of biological traits. Ecology 94:1205–1206.

DeLong, J. P., J. P. Gibert, T. M. Luhring, G. Bachman, B. Reed, A. Neyer, and K. Montooth. 2017. The combined effects of reactant kinetics and enzyme stability explain the temperature 312 dependence of metabolic rates. Ecology and Evolution .

Depinay, J.-M. O., C. M. Mbogo, G. Killeen, B. Knols, J. Beier, J. Carlson, J. Dushoff, P. Billingsley, H. Mwambi, J. Githure, et al. 2004. A simulation model of African *Anopheles* ecology and 315 population dynamics for the analysis of malaria transmission. Malaria Journal 3:1.

Fand, B. B., H. E. Z. Tonnang, M. Kumar, A. L. Kamble, and S. K. Bal. 2014. A temperature-based phenology model for predicting development, survival and population growth potential of 318 the mealybug, *Phenacoccus solenopsis* Tinsley (Hemiptera: Pseudococcidae). Crop Protection 55:98–108.

Garcia, R. A., M. Cabeza, C. Rahbek, and M. B. Araújo. 2014. Multiple dimensions of climate 321 change and their implications for biodiversity. Science 344:1247579.

Gillooly, J. F., J. H. Brown, G. B. West, V. M. Savage, and E. L. Charnov. 2001. Effects of size and temperature on metabolic rate. Science 293:2248–2251.

324 Hobbs, J. K., W. Jiao, A. D. Easter, E. J. Parker, L. A. Schipper, and V. L. Arcus. 2013. Change in heat capacity for enzyme catalysis determines temperature dependence of enzyme catalyzed rates. ACS Chemical Biology 8:2388–2393.

327 Hochachka, P. W., and G. N. Somero. 2002. Biochemical Adaptation: Mechanism and Process in Physiological Evolution. Oxford University Press.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3068v1 | CC BY 4.0 Open Access | rec: 3 Jul 2017, publ: 3 Jul 2017

Hoffmann, A. A., and C. M. Sgrò. 2011. Climate change and evolutionary adaptation. Nature 470:479–485.

Hopp, M. J., and J. A. Foley. 2001. Global-scale relationships between climate and the dengue fever vector, *Aedes aegypti*. Climatic Change 48:441–463.

Hothorn, T., K. Hornik, and A. Zeileis. 2006. Unbiased recursive partitioning: a conditional inference framework. Journal of Computational and Graphical Statistics 15:651–674.

Huey, R. B., and J. G. Kingsolver. 1989. Evolution of thermal sensitivity of ectotherm performance. Trends in Ecology & Evolution 4:131–135.

Ikemoto, T. 2005. Intrinsic optimum temperature for development of insects and mites. Environmental Entomology 34:1377–1387.

———. 2008. Tropical malaria does not mean hot environments. Journal of Medical Entomology 45:963–969.

Ikemoto, T., I. Kurahashi, and P.-J. Shi. 2013. Confidence interval of intrinsic optimum temperature estimated using thermodynamic SSI model. Insect science 20:420–428.

Johnson, F. H., and I. Lewin. 1946. The growth rate of *E. coli* in relation to temperature, quinine and coenzyme. Journal of Cellular and Comparative Physiology 28:47–75.

Johnson, Z. I. 2006. Niche Partitioning Among Prochlorococcus Ecotypes Along Ocean-Scale Environmental Gradients. Science 311:1737–1740.

Kang, S. H., J.-H. Lee, and D.-S. Kim. 2015. Temperature-dependent fecundity of overwintered *Scirtothrips dorsalis* (Thysanoptera: Thripidae) and its oviposition model with field validation. Pest Management Science 71:1441–1451.

López-Urrutia, Á., E. San Martin, R. P. Harris, and X. Irigoien. 2006. Scaling the metabolic balance of the oceans. Proceedings of the National Academy of Sciences 103:8739–8744.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3068v1 | CC BY 4.0 Open Access | rec: 3 Jul 2017, publ: 3 Jul 2017

Matthews, B. W. 1975. Comparison of the predicted and observed secondary structure of T4 phage lysozyme. Biochimica et Biophysica Acta (BBA)-Protein Structure 405:442–451.

Nealis, V. G., and J. Régnière. 2014. An individual-based phenology model for western spruce budworm (Lepidoptera: Tortricidae). The Canadian Entomologist 146:306–320.

Padfield, D., G. Yvon-Durocher, A. Buckling, S. Jennings, and G. Yvon-Durocher. 2016. Rapid evolution of metabolic traits explains thermal adaptation in phytoplankton. Ecology Letters 19:133–142.

Pawar, S., A. I. Dell, and V. M. Savage. 2015. From metabolic constraints on individuals to the dynamics of ecosystems. Pages 3–36 *in* A. Belgrano, G. Woodward, and U. Jacob, eds. Aquatic Functional Biodiversity: An Ecological and Evolutionary Perspective. Elsevier.

Pawar, S., A. I. Dell, V. M. Savage, and J. L. Knies. 2016. Real versus artificial variation in the thermal sensitivity of biological traits. The American Naturalist 187:E41–E52.

Pörtner, H. O., A. F. Bennett, F. Bozinovic, A. Clarke, M. A. Lardies, M. Lucassen, B. Pelster, F. Schiemer, and J. H. Stillman. 2006. Trade-offs in thermal adaptation: The need for a molecular to ecological integration. Physiological and Biochemical Zoology 79:295–313.

Rose, J. M., and D. A. Caron. 2007. Does low temperature constrain the growth rates of heterotrophic protists? Evidence and implications for algal blooms in cold waters. Limnology and Oceanography 52:886–895.

Schoolfield, R., P. Sharpe, and C. Magnuson. 1981. Non-linear regression of biological temperature-dependent rate models based on absolute reaction-rate theory. Journal of Theoretical Biology 88:719–731.

Schulte, P. M., T. M. Healy, and N. A. Fangue. 2011. Thermal performance curves, phenotypic plasticity, and the time scales of temperature exposure. Integrative and Comparative Biology 51:691–702.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3068v1 | CC BY 4.0 Open Access | rec: 3 Jul 2017, publ: 3 Jul 2017

Seibel, B. A., A. Dymowska, and J. Rosenthal. 2007. Metabolic temperature compensation and coevolution of locomotory performance in pteropod molluscs. Integrative and Comparative
378 Biology 47:880–891.

Sharpe, P. J., and D. W. DeMichele. 1977. Reaction kinetics of poikilotherm development. Journal of Theoretical Biology 64:649–670.

381 Shi, P., T. Ikemoto, C. Egami, Y. Sun, and F. Ge. 2011. A modified program for estimating the parameters of the SSI model. Environmental Entomology 40:462–469.

Simoy, M. I., M. V. Simoy, and G. A. Canziani. 2015. The effect of temperature on the population
384 dynamics of *Aedes aegypti*. Ecological Modelling 314:100–110.

Thomas, M. K., C. T. Kremer, C. A. Klausmeier, and E. Litchman. 2012. A global pattern of thermal adaptation in marine phytoplankton. Science 338:1085–1088.

387 Ungerer, M. J., M. P. Ayres, and M. J. Lombardero. 1999. Climate and the northern distribution limits of *Dendroctonus frontalis* Zimmermann (Coleoptera: Scolytidae). Journal of Biogeography 26:1133–1145.

390 Vimercati, L., S. Hamsher, Z. Schubert, and S. Schmidt. 2016. Growth of high-elevation *Cryptococcus* sp. during extreme freeze–thaw cycles. Extremophiles 20:579–588.

White, C. R., L. A. Alton, and P. B. Frappell. 2012. Metabolic cold adaptation in fishes occurs
393 at the level of whole animal, mitochondria and enzyme. Proceedings of the Royal Society of London B: Biological Sciences 279:1740–1747.

Wohlfahrt, G., M. Bahn, E. Haubner, I. Horak, W. Michaeler, K. Rottmar, U. Tappeiner, and
396 A. Cernusca. 1999. Inter-specific variation of the biochemical limitation to photosynthesis and related leaf traits of 30 species from mountain grassland ecosystems under different land use. Plant, Cell & Environment 22:1281–1296.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3068v1 | CC BY 4.0 Open Access | rec: 3 Jul 2017, publ: 3 Jul 2017

Tables

Parameter name	Estimation
α	$\alpha \sim \mathcal{N}(\mu = 10,\ \sigma = 3)$
β	$\alpha - i,\ i \sim \mathcal{N}(\mu = 4,\ \sigma = 2)$
Final curve width	original width $\cdot\ j,\ j \sim \mathcal{N}(\mu = 25,\ \sigma = 4)$
Final curve height	original height $+\ k,\ k \sim \mathcal{N}(\mu = 3,\ \sigma = 0.8)$

Table 1: Parameters for the generation of simulated curves. α and β are shape parameters of the beta distribution, whereas the two other parameters generate variation in the width and the height of the curves. β is constrained to be smaller than α, in order for the resulting curves to be negatively skewed, similarly to the observed thermal response curves of biological traits.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3068v1 | CC BY 4.0 Open Access | rec: 3 Jul 2017, publ: 3 Jul 2017

Figures

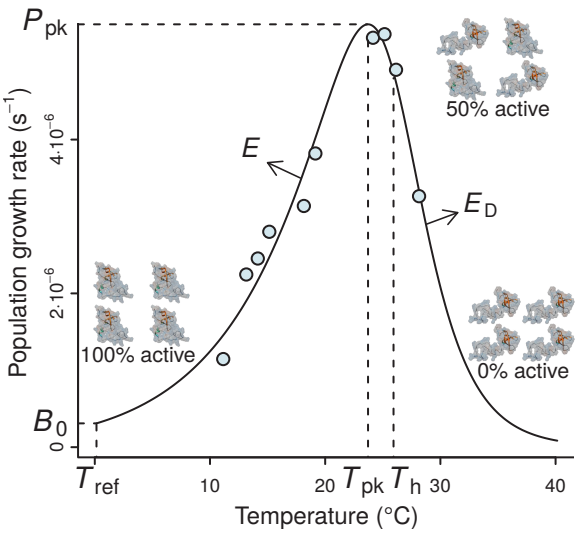

Figure 1: A typical example of the four-parameter Sharpe-Schoolfield model fitted to a thermal performance curve of *Prochlorococcus marinus* strain MIT9515 (Johnson 2006). As depicted, the model assumes that the activity of a single rate-controlling enzyme controls the apparent temperature dependence of the trait. T_h is defined as the temperature (before or after the peak) at which 50% of enzyme units are made inactive. Beyond T_h, an increasing proportion of the enzyme population is deactivated, to the point where all of them become non-functional, and the curve falls to zero. B_0 accurately represents the real trait performance at a reference temperature (T_{ref}), only if the enzyme population is fully functional at this particular T_{ref}, i.e., $T_{ref} \ll T_h$; otherwise, B_0 will necessarily be greater than the real trait value at T_{ref} ($B(T_{ref})$).

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3068v1 | CC BY 4.0 Open Access | rec: 3 Jul 2017, publ: 3 Jul 2017

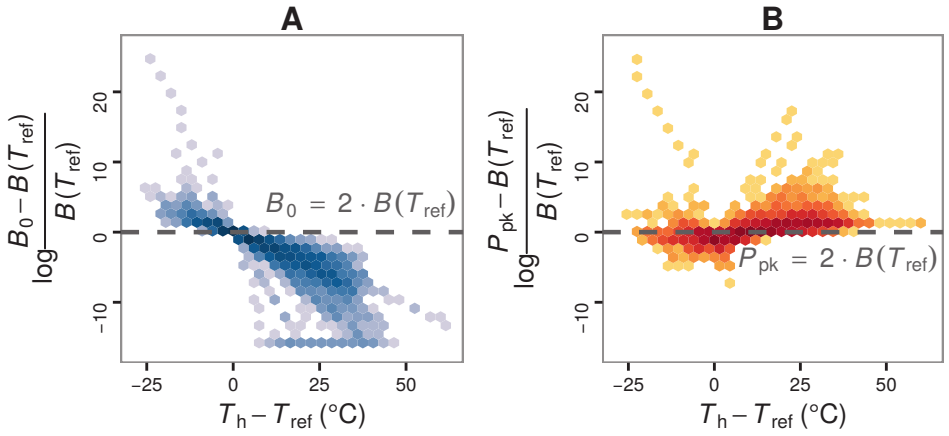

Figure 2: The effect of choice of reference temperature T_{ref} on the deviation of B_0 from $B(T_{ref})$ (panel A) and its relationship with P_{pk} (panel B). Data points were obtained by fitting the Sharpe-Schoolfield model to a dataset of phytoplankton growth rate measurements versus temperature (see main text) across a range of T_{ref} values. The colour depth of each hexagon is proportional to the number of data points at that location in the graph. At the vertical axis of panel A, a value of zero indicates that B_0 is double the real $B(T_{ref})$ value, and is used here as a reference point around and above which B_0 becomes non-negligibly exaggerated. As expected from eq. (2), the deviation of B_0 from $B(T_{ref})$ decreases nonlinearly with the difference between T_h and T_{ref}, to the point where the former asymptotically approaches zero (in linear scale). Towards the left end of the horizontal axis, the values of the estimates of B_0 even exceed those of the trait value at or close to optimum, P_{pk}.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3068v1 | CC BY 4.0 Open Access | rec: 3 Jul 2017, publ: 3 Jul 2017

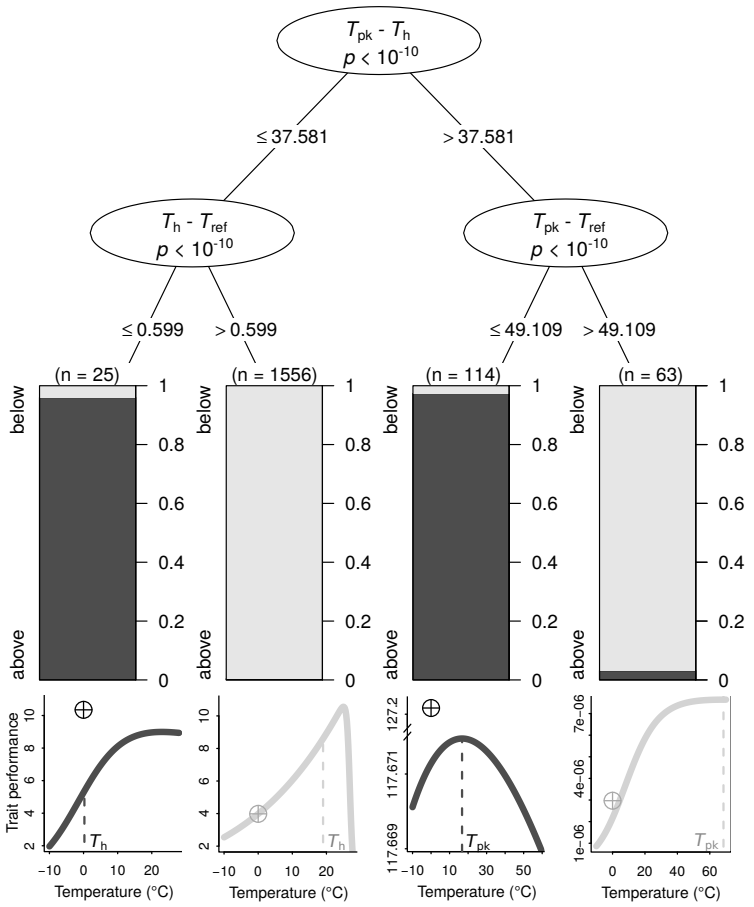

Figure 3: The conditions under which B_0 is highly overestimated (above the peak of the curve; dark grey bars and curves) or less so (below the peak; light grey bars and curves), determined using a conditional inference tree algorithm. Representative examples of thermal performance curves, along with their B_0 estimates (crossed circles; normalised at 0°C for consistency), are shown under each terminal node. The curves are not drawn on the same axes, as their trait performance values are at different orders of magnitude. For a few very wide – and possibly biologically unrealistic – curves (right half), the difference between T_{pk} and T_{ref} determines whether $B_0 > P_{pk}$. In contrast, for the remaining curves, a T_h value that is greater than T_{ref} by more than 0.599°C will always lead to B_0 estimates that are below P_{pk}.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3068v1 | CC BY 4.0 Open Access | rec: 3 Jul 2017, publ: 3 Jul 2017

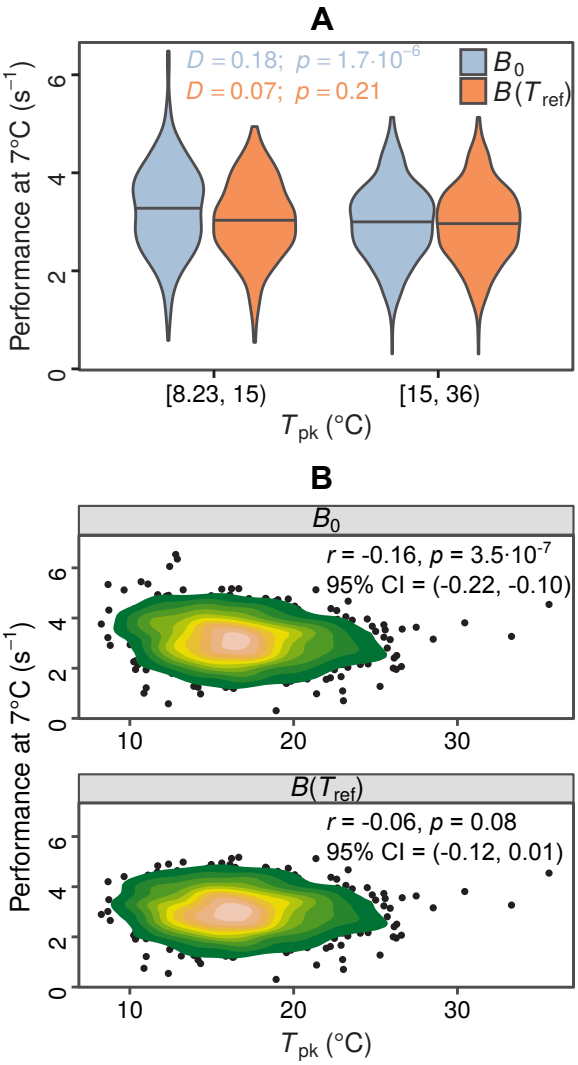

Figure 4: Impacts of exaggerated B_0 estimates on tests for metabolic cold adaptation. A: Violin plots of trait performance at $T_{\text{ref}} = 7°\text{C}$, as estimated using B_0 (light blue) and $B(T_{\text{ref}})$ (orange), for hypothetical cold-adapted species ($T_{\text{pk}} < 15°\text{C}$; left half) and species adapted to higher temperatures (right half). Horizontal lines indicate the median of each distribution. The statistical significance of the difference in performance between the two temperature groups was evaluated according to the two-sample Kolmogorov-Smirnov test. Based purely on the B_0 estimates – which get increasingly inflated at low temperatures as T_{h} approaches T_{ref} – one would mistakenly conclude that metabolic cold adaptation is present in this dataset. B: Correlations of B_0 with T_{pk}, and $B(T_{\text{ref}})$ with T_{pk}. The color surfaces represent the local density of data points. A similar pattern to the previous panel emerges, as the inflated B_0 estimates – in contrast to the true values – suggest that cold adaptation is present, albeit weakly.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3068v1 | CC BY 4.0 Open Access | rec: 3 Jul 2017, publ: 3 Jul 2017