

1 Estimating physiological tolerances– a comparison of traditional approaches to nonlinear
2 regression techniques

3

4 Dustin J. Marshall*¹, Michael Bode² & Craig R. White³

5 *Correspondence: Dustin.Marshall@monash.edu

6 ¹ School of Biological Sciences, Monash University, Victoria, 3800,

7 ² Department of Botany, University of Melbourne, Victoria, 3052

8 ³ School of Biological Sciences, University of Queensland, Queensland, 4072

9

10 Summary: Traditionally, physiologists have estimated the ability of organisms to withstand
11 lower partial pressures of oxygen by estimating the partial pressure at which oxygen
12 consumption begins to decrease (known as the 'critical P_{O_2} ' or ' P_C '). For almost 30 years, the
13 principal way in which P_C has been estimated has been via piecewise 'broken stick' regression.
14 Broken stick regression (BSR) was a useful approach when more sophisticated analyses were
15 less available, but BSR makes a number of unsupported assumptions about the underlying form
16 of the relationship between the rate of oxygen consumption and oxygen availability. The BSR
17 approach also distils a range of values into a single point with no estimate of error. In
18 accordance with more general calls to fit functions to continuous data, we propose the use of
19 nonlinear regression (NLR) to fit various curvilinear functions to oxygen consumption data in
20 order to estimate P_C . Importantly, our approach is back-compatible so that estimates using
21 traditional methods in earlier studies can be compared to data estimates from our technique.
22 When we compared the performance of our approach relative to the traditional BSR approach
23 for real world and simulated data, we found that under realistic circumstances, the NLR was
24 more accurate and provided more powerful hypothesis tests. We recommend that future
25 studies make use of NLR to estimate P_C , and also suggest that this approach might be more
26 appropriate for a range of physiological studies that use BSR currently.

27

28

29

Introduction

30

31

32

33

34

35

36

37

Understanding the physiological tolerances of an organism to its environment has long been a focus of ecophysiology. Oxygen is a fundamental requirement of most organisms, and its availability may be limiting across a range of habitats (Portner, 2010; Verberk et al., 2011; Ferguson et al. in press). For over 30 years, measures of oxygen tolerance have accumulated in a wide variety of taxa with variation evident at a range of scales (Greenlee and Harrison, 2004a; Mueller and Seymour, 2011; Lease et al., 2012; Ferguson et al., in press). Surprisingly, no clear consensus over how to estimate oxygen tolerance exists, and for the most part, modern statistical approaches have not been brought to bear on this problem.

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

The most common estimate of oxygen tolerance is the critical partial pressure of oxygen for aerobic metabolism (P_c), which represents the lowest level of oxygen at which aerobic metabolism is independent of the ambient partial pressure of oxygen (P_{O_2} ; Hochachka and Somero, 2002). At levels of P_{O_2} below P_c , either metabolism cannot be supported by aerobic processes entirely, and metabolic rate decreases; and/or anaerobic processes that are relatively inefficient and produce potentially toxic end-products become increasingly important (Hochachka and Somero, 2002). The original method for estimating P_c is the 'broken stick' regression (BSR) approach (Yeager and Ultsch, 1989) – an approach that remains the most common today. The broken stick approach to estimating P_c has been applied in a range of contexts, and been used to demonstrate, for example, that the P_c of a species is generally matched to the minimum oxygen level encountered in the environment in which it lives (Childress and Seibel, 1998; Nilsson, 2007; Ferguson et al., in press), and that mobile species show behavioural avoidance of oxygen levels below their P_c (Burleson et al., 2001). Nonetheless, not all species show clear break points in the relationship between rate of oxygen consumption ($\dot{V}O_2$) and P_{O_2} , which complicates efforts to assess the regulatory ability these species (Mueller and Seymour, 2011).

54 The traditional broken stick regression approach (piecewise linear regression) makes a
55 number of unsupported assumptions about the underlying relationship between oxygen
56 availability and respiration rate. First, it assumes that the functional response of an organism to
57 decreasing partial pressures of oxygen is biphasic, that is, it consists of two linear elements with
58 a clear break between these two phases. Above P_c , $\dot{V}O_2$ is assumed to be characterised by a linear
59 function that is completely flat, while below P_c , $\dot{V}O_2$ decreases linearly with P_{O_2} with an abrupt
60 transition between these two functions (Chiu et al., 2006). Of course, in reality, rates of
61 respiration are likely to be a continuous function between these two phases, and furthermore,
62 concentration-dependent reaction kinetics make a linear relationship between $\dot{V}O_2$ and P_{O_2}
63 highly unlikely. As such, the broken stick regression approach does not reflect the underlying
64 structure of the data, violating the basic assumptions of the regression approach (Quinn and
65 Keough, 2002). Our discussion here should not be taken as a criticism of the original
66 progenitors of these approaches: when they were developed, those analyses represented the
67 best approach available with the statistical and computational tools of the time. Today,
68 however, more sophisticated approaches are available that better reflect the underlying
69 processes that generated the data.

70 Since the development of the BSR technique, there have been a number of other ways in
71 which authors have attempted to estimate P_c . These approaches have largely been subjective
72 and lack repeatability (e.g. Portner et al., 1991; Greenlee and Harrison, 2004a, b, Lease et al.,
73 2012). Ideally, any technique for estimating P_c should both represent the mechanistic process
74 by which the data were generated and be repeatable. More generally, there are compelling
75 reasons for describing continuous traits with functional relationships, rather than taking
76 isolated point measures (Stinchcombe and Kirkpatrick, 2012). For example, Mueller and
77 Seymour (2011) fit a nonlinear function to estimate the ability of organisms to regulate oxygen
78 consumption beyond simple oxy-conformity across a range of oxygen values: they propose a
79 ‘Regulation Index’, which estimates a relative measure of oxyregulatory ability using either

80 linear, quadratic [$y = a + bx + cx^2$] or one-phase association [$y = y_0 + (y_{\max} - y_0)(1 - e^{-kx})$] fits).
81 Here, we apply a similar logic for estimating P_c that meets the above criteria – a nonlinear
82 regression approach coupled with simple differential calculus. We then evaluate the
83 performance of this new method relative to the most widely used approach to date, BSR, across
84 real-world data and explore a broader parameter space using simulations to compare the
85 sensitivity and reliability of these two approaches. Finally, we examine the statistical power of
86 the two techniques to distinguish differences in oxygen tolerance among groups.

87

88 **Methods**

89 *Fits to published data*

90 After examining the form of many published relationships between $\dot{V}O_2$ and P_{O_2} , we
91 settled on six candidate functions that approximated the general relationship well and were
92 reasonably tractable analytically (Table 1). While other forms could also fit the data and we
93 encourage investigators to explore alternatives (e.g. a power function with an intercept), we
94 chose these forms as, on first inspection, they appeared to fit real world data reasonably well
95 and consistently. We used standard nonlinear regression to fit these functions to published data.
96 It is beyond the scope of this paper to describe the general theory and approach of nonlinear
97 regression but we recommend Quinn and Keough (2002; pg 150) for a general introduction and
98 Ritz and Streibig (2008) for an excellent primer in how to implement this analysis in the freely
99 available statistical software R (R Development Core Team 2012). The nonlinear regression
100 analysis provides estimates of between two and four parameters (or even more if a selected
101 function has more): with these values the relationship between $\dot{V}O_2$ and P_{O_2} can be visualised.
102 For ease of use and comparison among studies, however, a single metric that best represents P_c
103 is desirable. Such a metric must be repeatable and objective. We based our metric on the
104 underlying principles of what P_c seeks to describe: the point where $\dot{V}O_2$ is no longer strongly

105 affected by P_{O_2} . In other words, when the slope of the function begins to flatten out and
106 approach zero. As the absolute values of a slope will vary according to the maximum oxygen
107 consumption rate of that particular organism, we first standardised our data by the maximum
108 $\dot{V}O_2$ observed for any set of values ($\dot{V}O_2 \text{ max} = 1$). By standardising, we can choose one slope
109 value and compare where that value occurs among individuals, species and studies.

110

111 The slope across function is, of course, given by the first derivative of the original function and
112 the derivative for each function is shown in Table 1. After rearranging we can solve for the
113 partial pressure P_c , at which any slope value (here, denoted m), and these are shown in Table 1.
114 We explored a range of values of m and found that a slope of $m = 0.065$ best approximates P_c
115 such that the solved values for P_c are shown in Table 2.

116 We coupled this formula to our nonlinear regression (NLR) estimates of a , b , c and d as
117 appropriate to estimate the value of P_c for a range of published data from the literature (see
118 Table 2). NLR relationships with additive error structures were fitted using the 'nls' function in
119 R v2.15.0 (R Development Core Team 2012), and NLR relationships with multiplicative error
120 structures were fitted using the 'gnls' function. The NLR and BSR fits to each data set were then
121 compared on the basis of Akaike's Information Criterion as a measure of model fit; the fit with
122 the lowest AIC was considered the best of the candidate set of models, given the data (Burnham
123 and Anderson 2008). We then also compared the performance of the NLR approach to the
124 broken stick method across a range of parameter values to evaluate how each approach coped
125 with differences in variation and sampling resolution with a simulation approach.

126 *Simulation methods*

127 We explored the influence of a suite of characteristics of the relationship between V_{O_2} and P_{O_2}
128 on NLR- and broken stick regression (BSR) derived estimates of P_c using Monte Carlo
129 simulations. We began by generating a function relating $\dot{V}O_2$ to P_{O_2} that was explicitly biphasic,

130 incorporating two linear portions (one that increased with P_{O_2} until $P_{O_2} = P_c$, and a second that
131 was independent of P_{O_2} at P_{O_2} levels greater than P_c). P_c was set at 5 kPa and $\dot{V}O_2$ above P_c was
132 set at 1. We then repeatedly simulated the process of sampling these data, and estimating the
133 value of P_c by fitting both the BSR and NLR. We elected to use a biphasic function because it
134 reflects the underlying assumption of the BSR approach, and because fitting such data should
135 represent the greatest challenge to NLR. Put simply, if NLR outperforms BSR, even when the
136 function is biphasic, then there can be little justification for preferring BSR over NLR, and as
137 such our approach was highly conservative. We sampled the relationship between $\dot{V}O_2$ and P_{O_2}
138 at a range of resolutions (0.125, 0.25, 0.5, 1 kPa). To each $\dot{V}O_2$, we then added a normal deviate
139 with a mean of zero and coefficient of variation (= s.d. divided by $\dot{V}O_2$) of 0.025, 0.05, 0.1, 0.15, or
140 0.20. These values for the CV of $\dot{V}O_2$ were selected to span the range of observed values of CV in
141 real data (n = 10, mean = 0.09, range: 0.04 – 0.21; See references in Table 1). 1000 such data
142 sets were generated for each combination of sampling resolution and CV, and P_c was estimated
143 using both NLR and BSR techniques. NLR estimates of P_c were derived as described above by
144 fitting a Weibull function to each data set (excluding that of *Cryptobranchus* as the model would
145 not converge) using the 'nls' function in R v2.15.0 (R Development Core Team 2012) to obtain a
146 non-linear least squares fit using a Gauss-Newton algorithm. We used the Weibull function as
147 this function best fit the most published relationships in the literature (see Results). BSR
148 estimates of P_c were obtained by simultaneously fitting two linear regressions constrained to
149 meet at a specified P_{O_2} , P_c . The slope of the linear regression below the specified P_c was a free
150 parameter, whereas the slope above the specified P_{O_2} was set at zero. Beginning at the third-
151 lowest P_{O_2} in the data set, a series of specified values of P_{O_2} were trialled, with each successive
152 P_{O_2} being 0.01 kPa greater than the last until the third highest P_{O_2} in the data set was reached.
153 The value of the P_{O_2} break point that minimised the sum of squared deviations from the
154 biphasic function was considered to represent P_c .

155 *Testing for differences among groups*

156 We compared the ability of our best fitting NLR and BSR to identify differences among groups
157 that differ in P_c with Monte Carlo Simulations. For two groups of 6 simulated datasets, each
158 with P_{O_2} values sampled at 1 kPa resolution, we generated a relationship between $\dot{V}O_2$ and P_{O_2}
159 that included a break point (P_c) at either 6.5 kPa (group 1) or 8.5 kPa (group 2). As a
160 conservative measure, these values of P_c were deliberately chosen to be different from the P_c at
161 which NLR performs best. As above, for each data set $\dot{V}O_2$ increased linearly with P_{O_2} to equal 1
162 when $P_{O_2} = P_c$. For values of P_{O_2} above P_c , $\dot{V}O_2$ was independent of P_{O_2} and equal to 1. To each
163 value of $\dot{V}O_2$ we then added a normal deviate with a mean of zero and coefficient of variation (=
164 s.d. divided by $\dot{V}O_2$) of 0.10. We then estimated P_c for each of the data sets using BSR or NLR with
165 $m = 0.065$, and tested for differences among groups using t tests. We also tested for differences
166 among groups using the best fitting NLR regression by pooling data sets for each group, and
167 testing for the significance of a fixed grouping factor using likelihood ratio tests. This simulation
168 procedure was repeated 1000 times, and statistical power was estimated by calculating the
169 proportion of tests that reported a significant difference among groups with α set at 0.05.

170

171 **Results**

172 *Fits to published data*

173 Both techniques provide a good fit to real-world data (Figure 1), and the predicted P_c
174 based on NLR was strongly correlated with the P_c estimates based on BSR ($R^2 = 0.82$, $n = 13$, $P <$
175 0.001), though the relationship was not precisely one to one (Regression equation: $P_{c(NLR)} =$
176 $0.646 \times P_{c(BSR)} + 1.12$). Neither consistently over- or under-estimated the other, rather the
177 relationship between the two was idiosyncratic (Figure 2). In eight out the 13 cases, the NLR

178 approach provided a better fit to the data than did the BSR approach (Table 2). For the
179 remaining 5 cases where the BSR approach had a lower AIC, the differences in AIC between the
180 BSR and NLR approaches were equivocal (Δ AIC around 2) in two cases. For nine out of thirteen
181 of the cases, the Weibull function (either with or without an intercept) provided a better fit to
182 the data than any other nonlinear function. For the four remaining cases, each of the other four
183 nonlinear functions provided the best fit for one case, suggesting that all should be considered
184 candidate functions in future studies.

185 *Simulations*

186 The best fitting NLR was far more robust to variation in the data than was the BSR
187 (Figure 3). While both performed well when variation was very low and sampling resolution
188 very high, the broken stick technique was far more likely to dramatically misestimate P_c when
189 variance was high. The NLR also coped with lower sampling resolution relative to the BSR: at
190 sampling resolutions greater than 0.5 kPa, the error rate of the BSR increased dramatically with
191 this approach being far more likely to misestimate P_c . For the BSR, the effects of sampling
192 resolution and error interacted such that the most error-prone estimates came from simulated
193 data with high levels of variation and low sampling resolutions (Figure 3). In contrast, neither
194 resolution nor variation had strong effects on the error rate of the NLR technique.

195 *Testing for differences among groups*

196 Power to detect differences among groups was lowest for t -tests comparing BSR-derived
197 estimates of P_c (0.63), and was higher for t -tests comparing best NLR-derived estimates of P_c
198 (0.74); power was also relatively higher for comparisons of pooled data made using likelihood
199 ratio tests (0.83).

200

201

202

Discussion

203

204

205

206

207

208

209

210

211

212

213

214

215

216

217

218

219

220

221

222

223

224

225

226

227

For both real-world and simulated data, a nonlinear regression (NLR) approach to estimating tolerance to decreasing availability to oxygen was superior to a broken stick regression (BSR) approach in almost every regard. The NLR approach was more robust to variation in the data and reductions in sampling resolution, as well as offering a more powerful means of detecting differences between groups. Our simulations suggest that BSR has a tendency to overestimate P_c when data are variable, sampling resolutions are low and the underlying 'true' P_c is low. It is possible therefore that some species may be more tolerant to low oxygen conditions than is currently appreciated. NLR performed less well in simulations when the underlying P_c was close to the highest P_c measured in the study and in this instance, BSR was a more accurate indicator of the P_c . The poor performance of NLR in this parameter space was due to the relationship between $\dot{V}O_2$ and PO_2 being linear in our simulations (recall that we used an artificially biphasic function) and as such, it was inevitable that the nonlinear function fit the data poorly. We therefore favour the use of NLR over BSR despite this shortcoming because in nature, very high P_c s are very rare, and genuinely linear relationships between $\dot{V}O_2$ and PO_2 are similarly rare. Nevertheless, we suggest that visual inspections of the data be conducted before P_c s are estimated and on the rare occasions when the relationship between $\dot{V}O_2$ and PO_2 appears to be linear, linear regression be used instead of NLR. In other words, as for any statistical model fitting exercise, the underlying form of the relationship should dictate the form of the model that is fit (Quinn and Keough, 2002).

The use of NLR to estimate tolerance to low oxygen carries the additional advantage of allowing formal testing for differences among groups within the same statistical framework. Traditionally, comparisons of P_c among species, individuals or groups have involved estimating P_c s within groups using the BSR approach, using these estimates of P_c as individual data points and then using a separate statistical test to detect differences (e.g. Ferguson et al., in press). There are a number of disadvantages to this two-stage approach. First, estimating the P_c as a

228 summary statistic for the entire function results in significant loss of information – essentially, a
229 large amount of data is collected to estimate a single point, and then the bulk of the data are
230 ignored in the hypothesis test (Stinchcombe and Kirkpatrick, 2012). Second, using the P_c
231 summary statistic ignores error and variance associated with this estimate, the subsuming of
232 error will result in hypothesis tests that are more susceptible to Type I errors (Hadfield et al.,
233 2010). In contrast, formal hypothesis tests of differences across entire functions among groups
234 allow the use of the full complement of data that were collected as well as incorporating the
235 appropriate error into the hypothesis test of interest (Stinchcombe and Kirkpatrick, 2012).
236 Furthermore, the NLR method could be extended to include random effects, whereby error
237 associated with different experimental equipment or different measurement days could be
238 partitioned, increasing the sensitivity of this analytical approach and providing much greater
239 scope to incorporate additional factors.

240

241 As well as the statistical advantages described above, the use of NLR affords practical
242 advantages for comparisons among studies and meta-analyses. For example, studies from
243 different research groups may not sample $\dot{V}O_2$ at identical values of P_{O_2} and may sample with
244 different resolutions. Our simulations show that differences in sampling resolution will result in
245 systematic differences in the estimate of P_c and furthermore, the estimate of P_c will be a product
246 of the P_{O_2} s at which $\dot{V}O_2$ was measured. In contrast, an NLR approach is robust to these
247 differences in sampling regime, allowing more comparable estimates of P_c among studies, as
248 well as estimates of $\dot{V}O_2$ across a greater parameter space. Furthermore, no single function fit the
249 data best such that comparisons among studies could be problematic. The retention of an
250 estimate of P_c in the context of our approach provides a means of comparing equivalent slopes
251 among studies that have fit very different functions.

252

253 We chose a point (where $dV/dP = m = 0.065$) on our function that best approximated the
254 P_c estimated by the BSR approach. We chose this point so that future studies can provide
255 estimates of oxygen tolerance that are comparable with traditional measures of P_c . In practice,
256 we recommend that future studies that utilise the NLR approach present estimates of not only
257 P_c , but also estimates of function parameters (i.e. estimates of a , b , c and d , as appropriate), as
258 well as their associated error, as these values will allow comparisons across the entire range of
259 oxygen partial pressures. Standardised estimates of parameters in selection in evolutionary
260 biology have facilitated important insights from meta-analyses (Kingsolver et al., 2001), and we
261 hope similarly valuable insights might become available via the widespread use of the
262 standardised approach we recommend here.

263 The use of approaches such as BSR to estimate biological ‘thresholds’ in nonlinear
264 responses more generally is increasingly being scrutinized. The estimation of ecological
265 thresholds using BSR techniques has been called into question (Toms and Lesperance, 2003), as
266 has the distillation of continuous functions into single points in evolutionary studies
267 (Stinchcombe and Kirkpatrick, 2012). Generally, the failure to use nonlinear functions to
268 describe nonlinear traits results in less powerful, less accurate estimates (Griswold et al., 2008).
269 BSR techniques are also used to estimate changes in developmental rates across temperatures
270 (Stanwell-Smith and Peck, 1998), as well as cardinal temperatures (estimates maximum,
271 minimum and optimal temperature), and we suspect that NLR techniques may be more
272 appropriate in these settings as well. While BSR techniques were once a pragmatic recognition
273 of limited computing power and statistical programming, today we have unprecedented access
274 to powerful statistical estimation techniques with freely available statistical software. We join
275 others (Stinchcombe and Kirkpatrick 2012) in calling for the increased uptake of ‘function
276 valued’ approaches to estimating continuous traits such as the relationship between $\dot{V}O_2$ and PO_2 .

277 **Acknowledgements:** We thank Hayley Cameron, Amy Hooper, Marcelo Lagos Orostica and Matthew Thompson for
278 help in preparing the manuscript. Two anonymous reviewers provided extremely thoughtful and helpful comments
279 that improved the manuscript.

280 Funding: DJM, MB and CRW are all supported by grants from the Australian Research Council.

281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332

Literature Cited

Burleson, M.L., Wilhelm, D.R. & Smatresk, N.J. (2001) The influence of fish size on the avoidance of hypoxia and oxygen selection by largemouth bass. *Journal of Fish Biology*, **59**, 1336-1349.

Burnham, K. P. & D. R. Anderson (2008) *Model selection and multi-model inference: a practical information-theoretic approach*. 2nd edition. Springer, New York

Cassin, S. (1963) Critical oxygen tensions in newborn, young and adult mice. *American Journal of Physiology*, **205**, 325-330.

Childress, J.J. & Seibel, B.A. (1998) Life at stable low oxygen levels: adaptations of animals to oceanic oxygen minimum layers. *Journal of Experimental Biology*, **201**, 1223-1232.

Chiu, G., Lockhart, R. & Routledge, R. (2006) Bent-cable regression theory and applications. *Journal of the American Statistical Association*, **101**, 542-553.

Ferguson, N., White, C.W & Marshall, D.J. (in press) Competition in benthic marine invertebrates: the unrecognised role of exploitative competition for oxygen. *Ecology*.

Greenlee, K.J. & Harrison, J.F. (2004a) Development of respiratory function in the American locust *Schistocerca americana* I. Across-instar effects. *Journal of Experimental Biology*, **207**, 497-508.

Greenlee, K.J. & Harrison, J.F. (2004b) Development of respiratory function in the American locust *Schistocerca americana* II. Within-instar effects. *Journal of Experimental Biology*, **207**, 509-517.

Griswold, C.K., Gomulkiewicz, R. & Heckman, N. (2008) Hypothesis testing in comparative and experimental studies of function-valued traits. *Evolution*, **62**, 1229-1242.

Hadfield, J.D., Wilson, A.J., Garant, D., Sheldon, B.C. & Kruuk, L.E.B. (2010) The Misuse of BLUP in Ecology and Evolution. *American Naturalist*, **175**, 116-125.

Hochachka, P.W. & Somero, G.N. (2002) *Biochemical Adaptation: Mechanism and Process in Physiological Evolution*. Oxford University Press, New York.

Kingsolver, J.G., Hoekstra, H.E., Hoekstra, J.M., Berrigan, D., Vignieri, S.N., Hill, C.E., Hoang, A., Gibert, P. & Beerli, P. (2001) The strength of phenotypic selection in natural populations. *American Naturalist*, **157**, 245-261.

Lease, H.M., Klok, C.J., Kaiser, A. & Harrison, J.F. (2012) Body size is not critical for critical PO₂ in scarabaeid and tenebrionid beetles. *Journal of Experimental Biology*, **215**, 2524-2533.

Mickel, T.J. & Childress, J.J. (1982) Effects of pressure and pressure acclimation on activity and oxygen consumption in the bathypelagic mysid *Gnathophausia ingens*. *Deep-Sea Research Part a-Oceanographic Research Papers*, **29**, 1293-1301.

Mueller, C.A. & Seymour, R.S. (2011) The Regulation Index: a new method for assessing the relationship between oxygen consumption and environmental oxygen. *Physiological and Biochemical Zoology*, **84**, 522-532.

Nilsson, G.E. (2007) Gill remodeling in fish - a new fashion or an ancient secret? *Journal of Experimental Biology*, **210**, 2403-2409.

Peck, L.S., Portner, H.O & Hardewig, I. 2002. Metabolic demand, oxygen supply, and critical temperatures in the antarctic bivalve *Laternula elliptica*. *Physiological and Biochemical Zoology*, **75**, 123-133.

Portner, H.O. (2010) Oxygen- and capacity-limitation of thermal tolerance: a matrix for integrating climate-related stressor effects in marine ecosystems. *Journal of Experimental Biology*, **213**, 881-893.

Portner, H.O., Maclatchy, L.M. & Toews, D.P. (1991) Metabolic responses of the toad *bufo marinus* to environmental hypoxia: an analysis of the critical PO₂. *Physiological Zoology*, **64**, 836-849.

Quinn, G.P. & Keough, M.J. (2002) *Experimental design and data analysis for biologists*. Cambridge University Press, Cambridge.

Ritz, C. & Streibig, J.C. (2008) *Nonlinear regression with R*. Springer Science+Business Media,

- 333 Stanwell-Smith, D. & Peck, L.S. (1998) Temperature and embryonic development in relation to
334 spawning and field occurrence of larvae of three Antarctic echinoderms. *Biological*
335 *Bulletin*, **194**, 44-52.
- 336 Steffensen, J.F., Schurmann, H. & Bushnell, P.G. (1994) Oxygen consumption in four species of
337 teleosts from Greenland: no evidence of metabolic cold adaptation. *Polar Biology*, **14**, 49-
338 54.
- 339 Stinchcombe, J.R. & Kirkpatrick, M. (2012) Genetics and evolution of function-valued traits:
340 understanding environmentally responsive phenotypes. *Trends in Ecology and Evolution*,
341 **27**, 637-647.
- 342 Toms, J.D. & Lesperance, M.L. (2003) Piecewise regression: a tool for identifying ecological
343 thresholds. *Ecology*, **84**, 2034-2041.
- 344 Ultsch, G.R., Ott, M.E & Heisler, N. (1980) Standard metabolic rate, critical oxygen tension and
345 aerobic scope for spontaneous activity of trout (*salmo gairdneri*) and carp (*cyprinus*
346 *carpio*) in acidified water. *Comparative Biochemistry and Physiology Part a-Physiology*,
347 **67**, 329-335.
- 348 Verberk, W., Bilton, D.T., Calosi, P. & Spicer, J.I. (2011). Oxygen supply in aquatic ectotherms:
349 partial pressure and solubility together explain biodiversity and size patterns. *Ecology*,
350 **92**, 1565-1572.
- 351 Yeager, D.P. & Ultsch, G.R. (1989) Physiological regulation and conformation: a basic program
352 for the determination of critical points. *Physiological Zoology*, **62**, 888-907.
- 353

354 Figure 1. Relationship between partial pressure of oxygen (PO_2) and oxygen consumption (VO_2)
355 for a) *Necturus maculosus* (y-axis shows $\mu\text{L}\cdot\text{g}^{-1}\cdot\text{h}^{-1}$) and b) *Crinia georgiani* tadpoles (y-axis
356 shows $\text{nanomoles}\cdot\text{h}^{-1}$). Line shows predicted line of best fit from nonlinear regression (NLR),
357 dashed lines show lines of best fit from broken stick regression (BSR) red and blue lines show
358 critical oxygen value (P_c) as estimated by BSR and NLR approaches respectively.

359

360 Figure 2. The relationship between P_c calculated using nonlinear regression and broken stick
361 regression approaches. Each point represents a separate species.

362

363 Figure 3. Effect of sampling resolution and variation in the underlying data on the standard
364 error of the estimates of P_c for: a) broken stick regression; and b) Nonlinear regression
365 approaches.

366

367

368

369

370

371

372

373

374

375

376 Table 1. Compilation of functions (and their derivatives) fit to relationships between partial
 377 pressure of oxygen (PO_2) and oxygen consumption ($\dot{V}O_2$) where a, b and c are constants to be
 378 estimated. All have approximately similar forms and all fit published data reasonably well. We
 379 also present the function for calculating a particular slope (m – see main text for details) for
 380 calculating critical oxygen value (P_c).

	Function	Derivative	Critical oxygen partial pressure
Michaelis-Menten	$\dot{V}_{O_2} = \frac{aP_{O_2}}{b + P_{O_2}}$	$\frac{d\dot{V}_{O_2}}{dP_{O_2}} = \frac{ab}{(P_{O_2} + b)^2}$	$P_c = \sqrt{\frac{ab}{m}} - b$
Power	$\dot{V}_{O_2} = aP_{O_2}^b$	$\frac{d\dot{V}_{O_2}}{dP_{O_2}} = abP_{O_2}^{b-1}$	$P_c = \left(\frac{m}{ab}\right)^{\frac{1}{b-1}}$
Hyperbola	$\dot{V}_{O_2} = \frac{aP_{O_2}}{b + P_{O_2}} + c$	$\frac{d\dot{V}_{O_2}}{dP_{O_2}} = abP_{O_2}^{-b-1}$	$P_c = \sqrt{\frac{ab}{m}} - b$
Pareto	$\dot{V}_{O_2} = 1 - (aP_{O_2})^b$	$\frac{d\dot{V}_{O_2}}{dP_{O_2}} = a^b b P_{O_2}^{b-1}$	$P_c = \left(\frac{m}{a^b b}\right)^{\frac{1}{b-1}}$
Weibull	$\dot{V}_{O_2} = a \left(1 - e^{-\left(\frac{P_{O_2}}{b}\right)^c}\right)$	$\frac{d\dot{V}_{O_2}}{dP_{O_2}} = a \left(\frac{c}{b}\right) \left(\frac{P_{O_2}}{b}\right)^{(c-1)} e^{-\left(\frac{P_{O_2}}{b}\right)^c}$	See note
Weibull with intercept	$\dot{V}_{O_2} = a \left(1 - e^{-\left(\frac{P_{O_2}}{b}\right)^c}\right) + d$	$\frac{d\dot{V}_{O_2}}{dP_{O_2}} = a \left(\frac{c}{b}\right) \left(\frac{P_{O_2}}{b}\right)^{(c-1)} e^{-\left(\frac{P_{O_2}}{b}\right)^c}$	See note

381 Note: Because there is no unique P_c for any given value of m in the Weibull functions, P_c is estimated
 382 numerically for these functions by finding the value of PO_2 closest to the highest in the data set where the
 383 calculated derivative is equal to m .

384

385

386

387

388 Table 2. Compilation of estimates of fit (AIC values shown in italics) for both broken stick regression and nonlinear regression for a range of
 389 candidate functions. Note that models with both additive or multiplicative error were fit to each relationship and the AIC values for each are shown
 390 here. Columns on the right show the estimated critical oxygen value (P_c) for both the BSR approach and the best fitting NLR approach. Lowest AIC
 391 values are shown in **bold** as are the estimated P_c from the best fitting model. 'NA' indicates that model would not converge and no estimate could be
 392 made.

Species	Study	AIC value												Summary	
		Michaelis-Menten		Power		Hyperbola		Pareto		Weibull		Weibull with intercept		Pcrit	
		Additive error	Multiplicative error	Additive error	Multiplicative error	Additive error	Multiplicative error	Additive error	Multiplicative error	Additive error	Multiplicative error	Additive error	Multiplicative error	BS	BestNLR
<i>Necturus maculosus</i> (salamander)	(Yeager and Ultich 1989)	<i>-97.06</i>	<i>-100.58</i>	<i>-73.42</i>	<i>-90.97</i>	NA	NA	<i>-112.20</i>	<i>-110.63</i>	<i>-123.01</i>	<i>-121.08</i>	<i>-132.30</i>	-133.13	6.40	5.26
<i>Crinia georgiana</i> , after hatching (frog)	(Cassin 1963)	<i>-191.51</i>	<i>-189.94</i>	<i>-161.15</i>	<i>-159.22</i>	NA	NA	<i>-179.91</i>	<i>-185.91</i>	-211.30	<i>-209.48</i>	<i>-210.99</i>	<i>-209.12</i>	14.41	6.27
<i>Gnathopausia ingens</i> (crab)	(Cassin 1963)	<i>-9.55</i>	<i>-18.50</i>	<i>-8.41</i>	<i>-17.33</i>	NA	NA	<i>-20.39</i>	<i>-5.62</i>	<i>-26.15</i>	<i>-46.10</i>	<i>-26.59</i>	-59.15	12.53	11.19
Carp, pH 7.9	(Peck et al., 2002)	NA	NA	<i>-9.75</i>	<i>-29.56</i>	NA	NA	<i>-7.65</i>	<i>-5.69</i>	<i>-25.53</i>	<i>-32.80</i>	<i>-26.15</i>	-53.27	14.62	12.78
Trout, pH 8	(Mueller and Seymour 2011)	<i>-25.66</i>	<i>-24.27</i>	<i>-23.19</i>	<i>-21.86</i>	<i>-24.16</i>	<i>-22.81</i>	<i>-15.19</i>	<i>-13.27</i>	-26.27	<i>-24.49</i>	<i>-25.88</i>	<i>-24.13</i>	10.23	6.86
<i>Cryptobranchus alleganiensis</i> (salamander)	(Mueller and Seymour 2011)	NA	NA	-52.75	<i>-51.04</i>	NA	NA	<i>-6.11</i>	<i>-4.03</i>	NA	NA	NA	NA	16.11	12.19
Adult mouse (25C)	(Mueller and Seymour 2011)	<i>-35.52</i>	<i>-34.92</i>	<i>-20.04</i>	<i>-22.95</i>	<i>-51.35</i>	<i>-52.84</i>	<i>-49.06</i>	<i>-47.21</i>	<i>-65.59</i>	-65.64	<i>-64.67</i>	<i>-64.04</i>	3.67	4.16
Neonate mouse (25C)	(Mueller and Seymour 2011)	<i>-67.89</i>	<i>-66.20</i>	<i>-57.80</i>	<i>-57.87</i>	-73.68	<i>-73.05</i>	<i>-58.02</i>	<i>-55.81</i>	<i>-68.06</i>	<i>-66.71</i>	<i>-71.36</i>	<i>-70.66</i>	7.28	5.22
<i>Laternula elliptica</i> (Antarctic bivalve)	(Mickel and Childress 1982)	-88.87	<i>-88.22</i>	<i>-64.16</i>	<i>-68.69</i>	<i>-87.01</i>	<i>-87.22</i>	<i>-60.68</i>	<i>-63.92</i>	<i>-85.50</i>	<i>-85.92</i>	<i>-83.89</i>	<i>-84.05</i>	3.74	3.10
<i>Pseudophyrne bibroni</i> , before hatching (frog)	(Ultich et al., 1980)	<i>-17.07</i>	<i>-15.21</i>	<i>-4.08</i>	<i>-18.48</i>	<i>-17.38</i>	<i>-19.74</i>	<i>-14.47</i>	<i>-13.22</i>	<i>-35.95</i>	-37.08	<i>-34.80</i>	<i>-35.30</i>	1.72	2.12
<i>Pseudophyrne bibroni</i> , after hatching (frog)	(Ultich et al., 1980)	<i>-19.31</i>	<i>-21.55</i>	<i>-9.25</i>	<i>-18.95</i>	NA	NA	<i>-28.51</i>	-34.25	<i>-32.59</i>	<i>-30.61</i>	<i>-30.80</i>	<i>-31.57</i>	2.42	2.67
<i>Crinia georgiana</i> , before hatching (frog)	(Steffensen et al., 1994)	<i>-32.95</i>	<i>-80.55</i>	<i>-19.60</i>	<i>-62.06</i>	NA	NA	<i>-75.48</i>	<i>-84.15</i>	<i>-106.47</i>	<i>-104.50</i>	-107.71	<i>-106.38</i>	3.08	3.16
Greenland Cod, <i>Gadus ogac</i>	(Yeager and Ultich 1989)	<i>-54.98</i>	<i>-58.88</i>	<i>-37.74</i>	<i>-49.47</i>	NA	NA	<i>-69.47</i>	<i>-74.53</i>	<i>-100.17</i>	<i>-103.49</i>	<i>-98.81</i>	-105.86	7.53	6.63

393

394





