

METHODS & TECHNIQUES

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

Dustin J. Marshall^{1,*}, Michael Bode² and Craig R. White³

¹School of Biological Sciences, Monash University, Victoria 3800, Australia, ²Department of Botany, University of Melbourne, Victoria 3052, Australia and ³School of Biological Sciences, University of Queensland, Queensland 4072, Australia

*Author for correspondence (Dustin.Marshall@monash.edu)

SUMMARY

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

Key words: metabolism, oxygen consumption, statistics, meta-analysis, oxygen availability.

Received 21 January 2013; Accepted 28 February 2013

INTRODUCTION

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 (Pörtner, 2010; Verberk et al., 2011; Ferguson et al., 2013). 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., 2013). 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.

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 , 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., 2013), and that mobile species show behavioural avoidance of oxygen levels below their P_c (Burlinson et al., 2001). Nonetheless, not all species show clear break points in the relationship between the 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).

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

Since the development of the BSR technique, there have been a number of other ways in which authors have attempted to

estimate P_c . These approaches have largely been subjective and lack repeatability (e.g. Portner et al., 1991; Greenlee and Harrison, 2004a; Greenlee and Harrison, 2004b; Lease et al., 2012). Ideally, any technique for estimating P_c should both represent the mechanistic process by which the data were generated and be repeatable. More generally, there are compelling reasons for describing continuous traits with functional relationships, rather than taking isolated point measures (Stinchcombe and Kirkpatrick, 2012). For example, Mueller and Seymour (Mueller and Seymour, 2011) fit a nonlinear function to estimate the ability of organisms to regulate oxygen consumption beyond simple oxy-conformity across a range of oxygen values: they propose a ‘regulation index’, which estimates a relative measure of oxyregulatory ability using either linear, quadratic ($y=a+bx+cx^2$) or one-phase association [$y=y_0+(y_{max}-y_0)(1-e^{-kx})$] fits. Here, we apply a similar logic for estimating P_c that meets the above criteria – a nonlinear regression approach coupled with simple differential calculus. We then evaluate the performance of this new method relative to the most widely used approach to date, BSR, across real-world data and explore a broader parameter space using simulations to compare the sensitivity and reliability of these two approaches. Finally, we examine the statistical power of the two techniques to distinguish differences in oxygen tolerance among groups.

MATERIALS AND METHODS
Fits to published data

After examining the form of many published relationships between \dot{V}_{O_2} and P_{O_2} , we settled on six candidate functions that approximated the general relationship well and were reasonably tractable analytically (Table 1). While other forms could also fit the data and we encourage investigators to explore alternatives (e.g. a power function with an intercept), we chose these forms as, on first inspection, they appeared to fit real-world data reasonably well and consistently. We used standard nonlinear regression (NLR) to fit these functions to published data. It is beyond the scope of this paper to describe the general theory and approach of NLR, but we

recommend Quinn and Keough [see p.150 (Quinn and Keough, 2002)] for a general introduction and Ritz and Streibig (Ritz and Streibig, 2008) for an excellent primer on how to implement this analysis in the freely available statistical software R (R Development Core Team, 2012). The NLR analysis provides estimates of between two and four parameters (or even more if a selected function has more): with these values the relationship between \dot{V}_{O_2} and P_{O_2} can be visualised. For ease of use and comparison among studies, however, a single metric that best represents P_c is desirable. Such a metric must be repeatable and objective. We based our metric on the underlying principles of what P_c seeks to describe: the point at which \dot{V}_{O_2} is no longer strongly affected by P_{O_2} . In other words, when the slope of the function begins to flatten out and approach zero. As the absolute values of a slope will vary according to the maximum oxygen consumption rate of that particular organism, we first standardised our data by the maximum \dot{V}_{O_2} observed for any set of values ($\dot{V}_{O_2,max}=1$). By standardising, we can choose one slope value and compare where that value occurs among individuals, species and studies.

The slope across function is, of course, given by the first derivative of the original function and the derivative for each function is shown in Table 1. After rearranging we can solve for P_c , at any slope value (here, denoted m), and these are shown in Table 1.

We explored a range of values of m and found that a slope of $m=0.065$ best approximates P_c such that the solved values for P_c are shown in Table 2.

We coupled this formula to our NLR estimates of a , b , c and d as appropriate to estimate the value of P_c for a range of published data from the literature (see Table 2). NLR relationships with additive error structures were fitted using the ‘nls’ function in R v2.15.0 (R Development Core Team, 2012), and NLR relationships with multiplicative error structures were fitted using the ‘gnls’ function. The NLR and BSR fits to each data set were then compared on the basis of Akaike’s information criterion (AIC) as a measure of model fit; the fit with the lowest AIC was considered the best of the candidate set of models, given the data (Burnham and Anderson,

Table 1. Compilation of functions (and their derivatives) fit to relationships between partial pressure of oxygen (P_{O_2}) and oxygen consumption (\dot{V}_{O_2}), where a , b and c are constants to be estimated

	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

All have approximately similar forms and all fit published data reasonably well. We also present the function for calculating a particular slope (m ; see main text for details) for calculating the critical oxygen value (P_c).

Note: Because there is no unique P_c for any given value of m in the Weibull functions, P_c is estimated numerically for these functions by finding the value of P_{O_2} closest to the highest in the data set where the calculated derivative is equal to m .

Table 2. Compilation of estimates of fit for both broken stick regression (BSR) and nonlinear regression (NLR) for a range of candidate functions

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

Note that models with both additive or multiplicative error were fit to each relationship and Akaike's information criterion (AIC) values for each are shown 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 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 made.

2008). We then also compared the performance of the NLR approach with the BSR method across a range of parameter values to evaluate how each approach coped with differences in variation and sampling resolution with a simulation approach.

Simulation methods

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

Testing for differences among groups

We compared the ability of our best-fitting NLR and BSR to identify differences among groups that differ in P_c with Monte Carlo simulations. For two groups of six simulated data sets, each with P_{O_2} values sampled at 1 kPa resolution, we generated a relationship between \dot{V}_{O_2} and P_{O_2} that included a break point (P_c) at either 6.5 kPa (group 1) or 8.5 kPa (group 2). As a conservative measure, these values of P_c were deliberately chosen to be different from the P_c at which NLR performs best. As above, for each data set \dot{V}_{O_2} increased linearly with P_{O_2} to equal 1 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 value of \dot{V}_{O_2} we then added a normal deviate with a mean of zero and CV of 0.10. We then estimated P_c for each of the data sets using BSR or NLR with $m=0.065$, and tested for differences among groups using t -tests. We also tested for differences among groups using the

best-fitting NLR regression by pooling data sets for each group, and testing for the significance of a fixed grouping factor using likelihood ratio tests. This simulation procedure was repeated 1000 times, and statistical power was estimated by calculating the proportion of tests that reported a significant difference among groups with α set at 0.05.

RESULTS

Fits to published data

Both techniques provide a good fit to real-world data (Fig. 1), and the predicted P_c based on NLR was strongly correlated with the P_c estimates based on BSR ($R^2=0.82$, $N=13$, $P<0.001$), though the relationship was not precisely one to one (regression equation: $P_{c(NLR)}=0.646 \times P_{c(BSR)} + 1.12$). Neither consistently over- or underestimated the other; rather, the relationship between the two was idiosyncratic (Fig. 2). In eight out of the 13 cases, the NLR approach provided a better fit to the data than did the BSR approach (Table 2). For the remaining five cases where the BSR approach had a lower AIC, the differences in AIC between the BSR and NLR approaches were equivocal (ΔAIC around 2) in two cases. For nine out of 13 cases, the Weibull function (either with or without an intercept) provided a better fit to the data than any other nonlinear function. For the four remaining cases, each of the other four

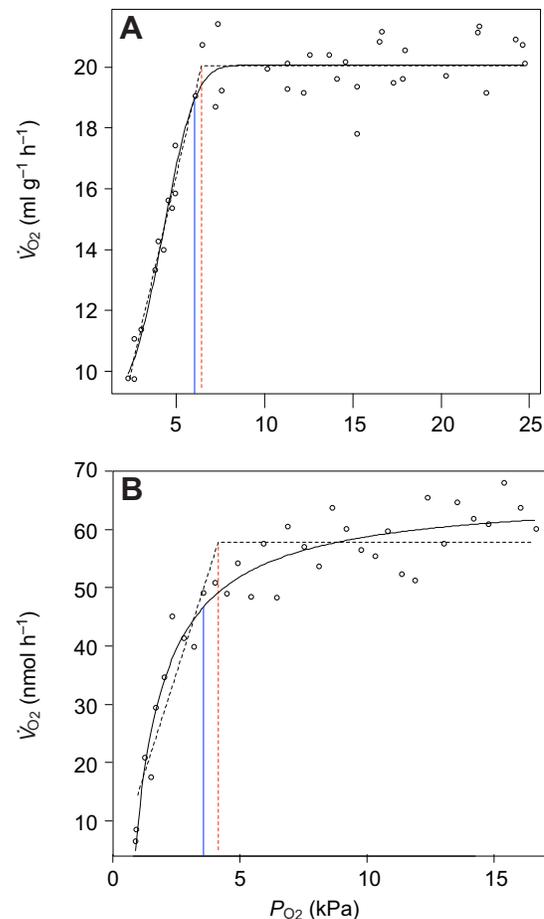


Fig. 1. Relationship between partial pressure of oxygen (P_{O_2}) and oxygen consumption (\dot{V}_{O_2}) for (A) *Necturus maculosus* and (B) *Crinia georgiani* tadpoles. Black lines show predicted lines of best fit from nonlinear regression (NLR); dashed lines show lines of best fit from broken stick regression (BSR); red and blue lines show critical oxygen value (P_c) as estimated by BSR and NLR approaches, respectively.

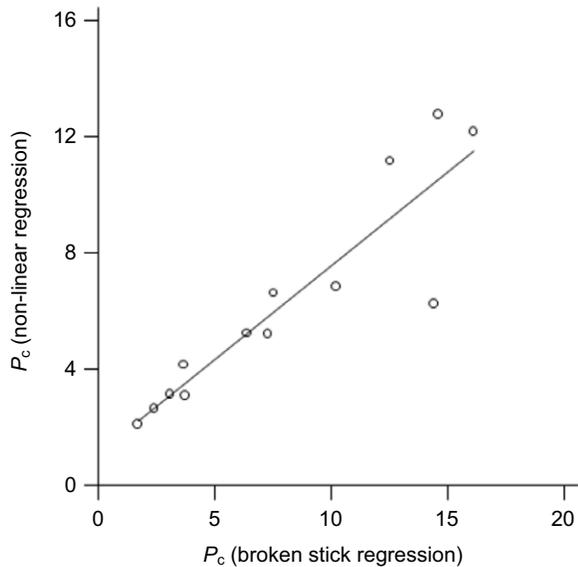


Fig. 2. Relationship between P_c calculated using nonlinear regression and broken stick regression approaches. Each point represents a separate species from Table 2.

nonlinear functions provided the best fit for one case, suggesting that all should be considered candidate functions in future studies.

Simulations

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

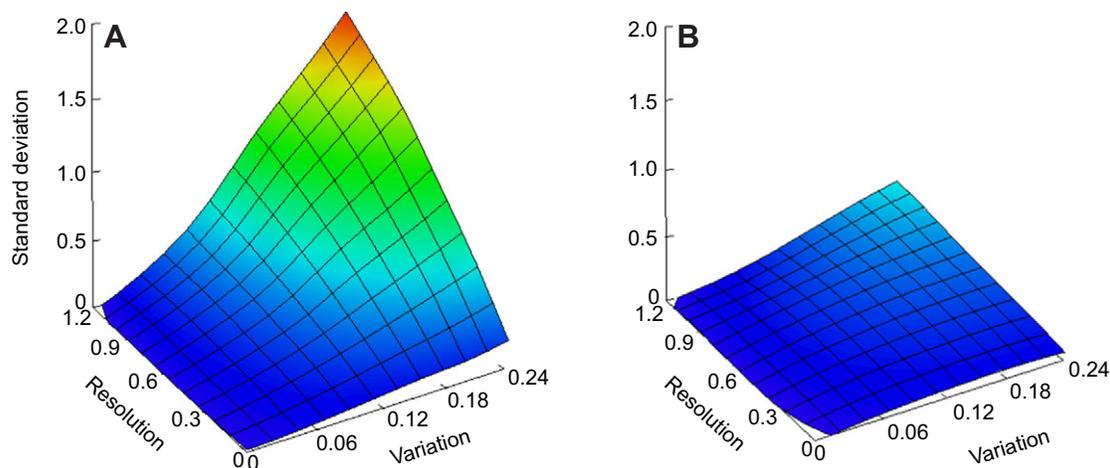


Fig. 3. Effect of sampling resolution and variation in the underlying data on the standard error of the estimates of P_c for (A) broken stick regression and (B) nonlinear regression approaches.

Testing for differences among groups

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

DISCUSSION

For both real-world and simulated data, an NLR approach to estimating tolerance to decreasing availability to oxygen was superior to a 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 P_{O_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 values of P_c are very rare, and genuinely linear relationships between \dot{V}_{O_2} and P_{O_2} are similarly rare. Nevertheless, we suggest that visual inspections of the data be conducted before P_c is estimated, and on the rare occasions when the relationship between \dot{V}_{O_2} and P_{O_2} appears to be linear, linear regression be used instead of NLR. In other words, 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 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.,

2013). There are a number of disadvantages to this two-stage approach. First, estimating the P_c as a summary statistic for the entire function results in significant loss of information – essentially, a large amount of data is collected to estimate a single point, and then the bulk of the data are ignored in the hypothesis test (Stinchcombe and Kirkpatrick, 2012). Second, using the P_c summary statistic ignores error and variance associated with this estimate, and the subsuming of error will result in hypothesis tests that are more susceptible to Type I errors (Hadfield et al., 2010). In contrast, formal hypothesis tests of differences across entire functions among groups allow the use of the full complement of data that were collected as well as incorporating the appropriate error into the hypothesis test of interest (Stinchcombe and Kirkpatrick, 2012). Furthermore, the NLR method could be extended to include random effects, whereby error associated with different experimental equipment or different measurement days could be partitioned, increasing the sensitivity of this analytical approach and providing much greater scope to incorporate additional factors.

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

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

The use of approaches such as BSR to estimate biological ‘thresholds’ in nonlinear responses more generally is increasingly being scrutinized. The estimation of ecological thresholds using BSR techniques has been called into question (Toms and Lesperance, 2003), as has the distillation of continuous functions into single points in evolutionary studies (Stinchcombe and Kirkpatrick, 2012). Generally, the failure to use nonlinear functions to describe nonlinear traits results in less powerful, less accurate estimates (Griswold et al., 2008). BSR techniques are also used to estimate changes in developmental rates across temperatures (Stanwell-Smith and Peck, 1998), as well as cardinal temperatures (estimates of maximum, minimum and optimal temperature), and we suspect that NLR techniques may be more appropriate in these settings as well. While BSR techniques were

once a pragmatic recognition of limited computing power and statistical programming, today we have unprecedented access to powerful statistical estimation techniques with freely available statistical software. We join others (Stinchcombe and Kirkpatrick, 2012) in calling for the increased uptake of ‘function-valued’ approaches to estimating continuous traits such as the relationship between \dot{V}_{O_2} and P_{O_2} .

ACKNOWLEDGEMENTS

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

AUTHOR CONTRIBUTIONS

D.J.M., M.B. and C.R.W. devised the idea and wrote the manuscript, M.B. calculated the derivatives and mathematical functions, and C.R.W. did the simulations.

COMPETING INTERESTS

No competing interests declared.

FUNDING

D.J.M., M.B. and C.R.W. are all supported by grants from the Australian Research Council.

REFERENCES

- Burleson, M. L., Wilhelm, D. R. and Smatresk, N. J. (2001). The influence of fish size on the avoidance of hypoxia and oxygen selection by largemouth bass. *J. Fish Biol.* **59**, 1336-1349.
- Burnham, K. P. and Anderson, D. R. (2008). *Model Selection and Multi-Model Inference: a Practical Information-Theoretic Approach*, 2nd edn. New York, NY: Springer.
- Cassin, S. (1963). Critical oxygen tensions in newborn, young and adult mice. *Am. J. Physiol.* **205**, 325-330.
- Childress, J. J. and Seibel, B. A. (1998). Life at stable low oxygen levels: adaptations of animals to oceanic oxygen minimum layers. *J. Exp. Biol.* **201**, 1223-1232.
- Chiu, G., Lockhart, R. and Routledge, R. (2006). Bent-cable regression theory and applications. *J. Am. Stat. Assoc.* **101**, 542-553.
- Ferguson, N., White, C. R. and Marshall, D. J. (2013). Competition in benthic marine invertebrates: the unrecognized role of exploitative competition for oxygen. *Ecology* **94**, 126-135.
- Greenlee, K. J. and Harrison, J. F. (2004a). Development of respiratory function in the American locust *Schistocerca americana*. I. Across-instar effects. *J. Exp. Biol.* **207**, 497-508.
- Greenlee, K. J. and Harrison, J. F. (2004b). Development of respiratory function in the American locust *Schistocerca americana*. II. Within-instar effects. *J. Exp. Biol.* **207**, 509-517.
- Griswold, C. K., Gomulkiewicz, R. and 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. and Kruuk, L. E. B. (2010). The misuse of BLUP in ecology and evolution. *Am. Nat.* **175**, 116-125.
- Hochachka, P. W. and Somero, G. N. (2002). *Biochemical Adaptation: Mechanism and Process in Physiological Evolution*. New York, NY: Oxford University Press.
- Kingsolver, J. G., Hoekstra, H. E., Hoekstra, J. M., Berrigan, D., Vignieri, S. N., Hill, C. E., Hoang, A., Gibert, P. and Beerli, P. (2001). The strength of phenotypic selection in natural populations. *Am. Nat.* **157**, 245-261.
- Lease, H. M., Klok, C. J., Kaiser, A. and Harrison, J. F. (2012). Body size is not critical for critical P_{O_2} in scarabaeid and tenebrionid beetles. *J. Exp. Biol.* **215**, 2524-2533.
- Mickel, T. J. and Childress, J. J. (1982). Effects of pressure and pressure acclimation on activity and oxygen consumption in the bathypelagic mysid *Gnathopausia ingens*. *Deep Sea Res. A* **29**, 1293-1301.
- Mueller, C. A. and Seymour, R. S. (2011). The regulation index: a new method for assessing the relationship between oxygen consumption and environmental oxygen. *Physiol. Biochem. Zool.* **84**, 522-532.
- Nilsson, G. E. (2007). Gill remodeling in fish – a new fashion or an ancient secret? *J. Exp. Biol.* **210**, 2403-2409.
- Peck, L. S., Pörtner, H. O. and Hardewig, I. (2002). Metabolic demand, oxygen supply, and critical temperatures in the Antarctic bivalve *Laternula elliptica*. *Physiol. Biochem. Zool.* **75**, 123-133.
- Pörtner, H. O. (2010). Oxygen- and capacity-limitation of thermal tolerance: a matrix for integrating climate-related stressor effects in marine ecosystems. *J. Exp. Biol.* **213**, 881-893.
- Pörtner, H. O., Maclatchy, L. M. and Toews, D. P. (1991). Metabolic responses of the toad bufo marinus to environmental hypoxia: an analysis of the critical P_{O_2} . *Physiol. Zool.* **64**, 836-849.
- Quinn, G. P. and Keough, M. J. (2002). *Experimental Design and Data Analysis for Biologists*. Cambridge: Cambridge University Press.
- R Development Core Team (2012). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

- Ritz, C. and Streibig, J. C.** (2008). *Nonlinear Regression with R*. New York, NY: Springer Science and Business Media.
- Stanwell-Smith, D. and Peck, L. S.** (1998). Temperature and embryonic development in relation to spawning and field occurrence of larvae of three Antarctic echinoderms. *Biol. Bull.* **194**, 44-52.
- Steffensen, J. F., Schurmann, H. and Bushnell, P. G.** (1994). Oxygen consumption in four species of teleosts from Greenland: no evidence of metabolic cold adaptation. *Polar Biol.* **14**, 49-54.
- Stinchcombe, J. R., Function-valued Traits Working Group and Kirkpatrick, M.** (2012). Genetics and evolution of function-valued traits: understanding environmentally responsive phenotypes. *Trends Ecol. Evol.* **27**, 637-647.
- Toms, J. D. and Lesperance, M. L.** (2003). Piecewise regression: a tool for identifying ecological thresholds. *Ecology* **84**, 2034-2041.
- Ultsch, G. R., Ott, M. E. and Heisler, N.** (1980). Standard metabolic rate, critical oxygen tension and aerobic scope for spontaneous activity of trout (*Salmo gairdneri*) and carp (*Cyprinus carpio*) in acidified water. *Comp. Biochem. Physiol.* **67A**, 329-335.
- Verberk, W. C., Bilton, D. T., Calosi, P. and Spicer, J. I.** (2011). Oxygen supply in aquatic ectotherms: partial pressure and solubility together explain biodiversity and size patterns. *Ecology* **92**, 1565-1572.
- Yeager, D. P. and Ultsch, G. R.** (1989). Physiological regulation and conformation: a basic program for the determination of critical points. *Physiol. Zool.* **62**, 888-907.