

Journal of Thermal Biology 31 (2006) 287-291

Journal of THERMAL BIOLOGY

www.elsevier.com/locate/jtherbio

Cautionary notes on the descriptive analysis of performance curves in reptiles

Grégory Bulté, Gabriel Blouin-Demers*

Department of Biology, University of Ottawa, 30 Marie-Curie, Ottawa, Ont., Canada K1N 6N5

Received 12 October 2005; accepted 29 November 2005

Abstract

Organismal performance curves are important functions for the study of reptilian ecology and evolution, but their interpretation can be affected strongly by the choice of analytical approach. We first use an example from the literature to demonstrate that excluding biologically meaningful data alters the description of performance and leads to non-sensical inferences. We then use fictional data to show that linear models (ANOVA) commonly used in the descriptive analysis of performance curves can also be biologically misleading or can lack biological relevance. Our examples demonstrate that fitting non-linear curves to performance data is more meaningful and avoids erroneous representation and interpretation of these important biological functions.

Keywords: Performance curve; ANOVA; Optimal temperature; Thermal reaction norm; Reptiles

1. Introduction

Thermal sensitivity in reptiles is studied broadly (Huey, 1982; Angilletta et al., 2002b). The relationship between body temperature (T_b) and performance (i.e., the thermal reaction norm) is often used to evaluate and predict the effects of environmental temperature on reptile ecology (Tracy and Christian, 1986; Huey, 1991; Weatherhead and Robertson, 1992; Angilletta et al., 2002a) and to study evolutionary patterns in thermal physiology (Huey and Bennett, 1987; Huey and Kingsolver, 1989; Angilletta et al., 2002b; Blouin-Demers et al., 2003; Zhang and Ji, 2004). Appropriate description and analysis of the relationship between T_b and performance is therefore paramount to interpret properly evolutionary physiology.

Over 25 years ago, Huey and Stevenson (1979) published a discussion of approaches to study thermal sensitivity in ectotherms. One of their goals was to stimulate research in whole animal performance. In recent years, several laboratory studies have quantified the thermal reaction norms of organismal traits in reptiles (Ji et al., 1993; Witz and Lawrence, 1993; Scribner and Weatherhead, 1995; Ji et al., 1996; Du et al., 2000; Angilletta et al., 2002a; Blouin-Demers et al., 2003; Chen et al., 2003; Elsworth et al., 2003; Zhang and Ji, 2004). Despite Huey and Stevenson's (1979) comprehensive discussion, however, we believe that the descriptive analysis of thermal reaction norms of reptiles in recent papers often provides incomplete information and, in some cases, is biologically misleading. Therefore, our aim here is to demonstrate the limitations of current methods and to suggest alternatives that circumvent those limitations.

We will first comment and re-analyze a study of thermal sensitivity of appetite in lizards (McConnachie and Alexander, 2004) in which a misconception of the thermal reaction norm has led to inappropriate data analysis. Second, we will demonstrate that the approach most commonly used to analyze thermal sensitivity in recent studies, ANOVA with $T_{\rm b}$ as a categorical independent variable and performance as a continuous dependent variable, can lead to misrepresentation of the biological reality.

^{*}Corresponding author. Tel.: +613 562 5800 6749; fax: +613 562 5486. *E-mail address:* gblouin@uottawa.ca (G. Blouin-Demers).

 $^{0306\}text{-}4565/\$$ - see front matter @ 2006 Elsevier Ltd. All rights reserved. doi:10.1016/j.jtherbio.2005.11.030

2. Re-anal sis of McConnachie and Ale ander (2004)

McConnachie and Alexander (2004) measured appetite as the mass and number of mealworms consumed by *Cordylus* lizards during 14-day trials at T_b of 20, 22, 25, 32, and 35 °C. Fig. 3a and b of their article illustrate those relationships and show that appetite gradually increases between 20 and 32 °C, but decreases rapidly between 32 and 35 °C. Therefore, the relationship between appetite and T_b is curvilinear. The authors, however, ignored the 35 °C trials in their analyses and fitted a linear regression to the data. Their argument for this exclusion was that lizards were showing signs of distress at 35 °C and were eating little, thus forcing the arrest of the trials after 3 days instead of 14.

The exclusion of the trials at 35 °C from the analyses constitutes an important conceptual mistake leading to a misrepresentation of the biological reality. By rejecting the 35 °C trials, the authors ignore the natural shape of the $T_{\rm b}$ performance function. Performance of ectotherms is an asymmetric (skewed to the left) function of temperature defined by a lower and an upper critical temperature $(CT_{min}\ and\ CT_{max}\!\!:$ the tolerance range), an optimum temperature (T_0) and a performance breadth (e.g., range at which performance is $\geq 80\%$) (Huey, 1982; Stevenson et al., 1985; Angilletta et al., 2002b). CT_{min} and CT_{max} are the minimal and maximal T_b at which performance is still possible and $T_{\rm o}$ is the $T_{\rm b}$ at which performance is maximized. Therefore, the decrease observed at 35 °C is a normal biological response when $T_{\rm b}$ approaches $CT_{\rm max}$ and cannot be ignored in performance analysis if one wants to get the best mathematical representation of the biological reality. Doing otherwise and using linear regression suggests that appetite increases continuously with $T_{\rm b}$, even at $T_{\rm b}$ lethal for the species. When all appetite responses to $T_{\rm b}$ are included, the relationship is clearly curvilinear and, therefore, the most appropriate description is a non-linear equation such as those presented by Stevenson et al. (1985) and Huey (1982).

We extracted the data from Fig. 3a and b of McConnachie and Alexander (2004) using the software GraphClick (version 2.4.3) and reanalyzed the data including the 35 °C trials. We first fitted a linear regression to compare the fit of this line to the one obtained by McConnachie and Alexander that excluded the 35 °C trials. We then fitted the non-linear equation provided by Stevenson et al. (1985)

Performance =
$$S[1/(1 + K_1 e^{-K^2} (T_b - CT_{min}))]$$

 $\times [1 - e^{K^3 (T_b - CT_{max})}],$

where CT_{max} , CT_{min} , and T_b are defined as above and K_1 , K_2 , and K_3 are constants and S is a scaling factor. CT_{max} and CT_{min} were unavailable from McConnachie and Alexander (2004), but they could be estimated from their data. We estimated CT_{max} at 38 °C because lizards were showing signs of distress sufficient to justify the arrest of



Fig. 1. Mass of mealworms consumed by *Cordylus m. melanolotus* as a function of body temperature (data obtained from McConnachie and Alexander (2004)). The curve is the fitted non-linear performance curve. The solid line is the regression line of the original study (excluding the 35 °C data). The dashed line is the regression line including the 35 °C data. When included in the regression, the 35 °C trials make the relationship non-significant.

the experiment at 35 °C and CT_{min} was set at 12 °C based on the distribution of the data and because very little food consumption occurred at 20 °C.

Linear regressions including the 35°C trials led to different statistical conclusions than the original regressions. Both relationships became non-significant (number of worms: $R^2 = 0.49$, p = 0.12 versus $R^2 = 0.91$, p = 0.01; mass of worms: $R^2 = 0.56$, p = 0.084 versus $R^2 = 0.88$, p = 0.02, Fig. 1) outlining that performance is not a linear function of temperature. On the other hand, the non-linear curve fitted the data well. The mean deviation of the predicted values from the actual values was of 0.85 worms for the number of worms ingested and of 0.6 g for the mass of worm ingested (Fig. 1). The major problem with the initial analytical approach is that it implies that performance increases monotonically with $T_{\rm b}$. Therefore, one would conclude, based on their relationship, that performance is greater at 40 °C than at 30 °C, even if 30 °C is the optimal temperature. Moreover, the authors did not specify that their strong positive relationship was restricted to the 20–32 °C range.

The reasoning for eliminating data from the analysis was that the 35 °C trials were shorter. In this case, however, the shorter duration was a normal biological response to increased temperature. Had the authors continued the trials for 14 days (as for the others T_b 's), the lizards would have been unlikely to increase their food consumption because of thermal stress.

3. Shortfalls of the ANOVA for describing thermal performance curves

In the above example, the analytical problem stemmed from a misconception of the relationship between $T_{\rm b}$ and performance. Although this example is an exception in the thermal performance literature, we believe that the analytical approach most commonly used to describe performance in reptiles (ANOVA) can also provide an erroneous representation of the thermal reaction norm.

ANOVA is often used to analyse the effect of $T_{\rm b}$ on performance in reptiles (Ji et al., 1993, 1996; Du et al., 2000; Angilletta et al., 2002a; Chen et al., 2003; Elsworth et al., 2003; Zhang and Ji, 2004). In general, the performance of several individuals is measured at various $T_{\rm b}$ (usually 4-6) encompassing the range of $T_{\rm b}$ experienced by the animal (i.e., within the CT). The mean performance (response variable) at each experimental $T_{\rm b}$ (factor) is then compared using a factorial ANOVA and post-hoc multiple comparisons tests. If the variation in performance within a $T_{\rm b}$ is significantly less than the variation between $T_{\rm b}$, then the mean performances between the two (or more) $T_{\rm b}$ are considered to be statistically different. If there is no statistical difference in mean performance between two or more consecutive $T_{\rm b}$, then performance is considered to be equal between those $T_{\rm b}$. $T_{\rm o}$ is thus considered to be bounded by consecutive and statistically similar $T_{\rm b}$ at which performance is maximized (Huey and Stevenson, 1979; Angilletta et al., 2002a). A range including T_{0} , rather than a single value, is thus obtained. We believe that this approach to describing performance curves can be biologically misleading.

Before we use an example to illustrate how ANOVA can be misleading in the description of performance curves, we need to identify two characteristics of performance studies that should be taken into account in ANOVA. First, $T_{\rm b}$ is a continuous variable. In a standard ANOVA design, however, the factor (i.e., experimental $T_{\rm b}$) is considered an unordered categorical variable. At the very least, $T_{\rm b}$ should be considered an ordered categorical variable if ANOVA is to be used to describe performance curves. Second, in performance studies the same individuals are usually measured at each $T_{\rm b}$. Therefore, the appropriate analysis is a repeated measures ANOVA (Potvin et al., 1990; Angilletta et al., 2002a). The majority of recent studies we examined, however, used a standard ANOVA instead of a repeated measures ordered ANOVA. In addition, ANOVA can produce results that are not representative of the real reaction norm. We illustrate how with fictional data that we generated from actual performance curves.

We generated our data from swimming speed data of northern watersnakes (*Nerodia sipedon*) (Blouin-Demers et al., 2003). We first created a population A in which there was little variation in T_o and maximum performance attained. From this population, we created two scenarios of inter-individual variation. In population B, we introduced variation in the maximum performance attained (*Y* axis only) by varying the scaling factor (*S*) of Stevenson et al.'s (1985) equation so that maximum performance ranged between 0,75 and 1,95 (undefined units of performance), but the general shape of the curve remained unaffected. Consequently, individuals vary in the maximum performance they can achieve, but they all perform best at the same $T_{\rm b}$ (Fig. 2). In population C, we created variation in T_{o} (X axis only) by varying K_{1} , K_{2} , K_{3} , or all three at the same time. Therefore, all individuals achieved the same maximum value, but at different $T_{\rm b}$ (introducing variation in T_{0}). In all cases, we assumed that CT remained the same across all individuals because mean CT are used to fit the curves in most studies. In theory, however, CT could also vary between individuals and this would produce more variation in T_{o} . We then compared T_{o} determined with ANOVA to T_{0} obtained with non-linear curve fitting. For ANOVA, we used the testing temperatures as the factor (ordered) and performance as the response in a repeated measures model. We used Tukey's HSD test to identify the $T_{\rm o}$ range. Then, we fitted a non-linear curve for each individual with the equation described above and calculated the mean T_{0} . We also determined the mean range of the 80% performance breadth, which is indicative of the width of the performance plateau.

For populations B and C, there was no significant difference between 25 and 35 °C with the ANOVA (Fig. 2). The biological interpretation would be that T_o is bounded by those T_b for both populations. With the curve fitting approach, the mean T_o was 31.7 °C (\pm 0.3) for population B and 31.9 °C (\pm 2.0) for population C. Those values are close to one another and fall in the middle of the T_o range depicted by the ANOVA, but the ANOVA does not tell us that those two scenarios represent populations that have very distinct patterns of inter-individual variation. On the other hand, the error term associated with the mean T_o of the curve fitting approach tells us that T_o is variable in population C, but not in population B.

A major problem with the ANOVA approach is that the T_{0} range depends on the inter-individual variation in performance at each experimental $T_{\rm b}$. Consequently, the greater the inter-individual variation in performance (such as in population B), the wider the T_{o} range. This becomes apparent when we compare populations A and B (Fig. 2A, B) that have the same T_{0} , but population B has more individual variation in performance. In population A, all experimental $T_{\rm b}$ are significantly different from one another. Therefore, one would interpret T_{o} to be near 30 °C because it is the experimental $T_{\rm b}$ at which performance is statistically greater than at any other T_b . For population B, however, the T_{o} range is between 25 and $35 \,^{\circ}\text{C}$ because performance at those T_{b} is not significantly different based on ANOVA. This inference is erroneous because performance at 25 °C in population B is not equal to performance at 31 °C (T_{0}), it is 20% lower. In some species, such as water snakes, a 5 $^{\circ}$ C deviation from T_{o} decreases performance by 20% (Blouin-Demers et al., 2003). Comparing populations A and B using ANOVA led to the conclusion that population A has a narrow T_0 range and population B has a wide T_{o} range when their thermal reaction norms are identical; they only differ in the variation in individual maximum performance. The wide $T_{\rm o}$ range in population B is a consequence of the variation in individual maximum performance attained. With such variability, the ANOVA will always lead to a wide $T_{\rm o}$



range. Large individual variability is expected especially when performance is not normalized (Huey and Stevenson, 1979) (i.e., expressed as a percentage of individual maxima) as it is the case in several recent studies (Ji et al., 1993, 1996; Du et al., 2000; Chen et al., 2003; McConnachie and Alexander, 2004; Zhang and Ji, 2004). Normalizing by individual controls for factors that systematically influence absolute performance (such as size or sex) and, thus, eliminates overall inter-individual variability.

Another problem with the ANOVA approach, originally recognized by Huey and Stevenson (1979), arises when the actual T_o is variable in the population, such as in population C. In population C, individual T_o range from 29.5 to 34.2 °C. ANOVA leads to the conclusion that the population has a broad T_o range when, in reality, T_o is variable (Fig. 2C). For population C, T_o obtained by curve fitting is also misleading, but the error term associated with T_o is a warning about variability and, therefore, allows cautious interpretation of the biological significance of T_o in this population. Moreover, variability is an important biological metric because it is one requirement for natural selection to act on a trait, and variance in T_o informs about individual variability.

4. Concluding remarks

We have shown that, with ANOVA, similar biological interpretations can be obtained despite different patterns of individual variation. It has been argued that the ANOVA approach is conservative when the performance curve shows a wide plateau because T_{o} must be included in the range. We believe, however, that a wide range of T_{o} does not provide much biological information. Moreover, it is often impossible from ANOVA tables and plots to evaluate what patterns of variation led to a wide T_{o} range (i.e., variation in T_o or variation in maximum performance). Mean T_{0} and performance breadth with their error terms are better representations of the thermal reaction norm of a species and are essential to compare performance statistically for different groups or for different species. For instance, Zhang and Ji (2004) compared locomotor and digestive performance between three Takydromus after separate ANOVA for each species, but their comparisons had to remain qualitative. Had Zhang and Ji (2004) fitted curves to the data for each individual, they would have been able to compare mean T_{o} and performance breadth statistically between the three species.

Fig. 2. Performance curves of three hypothetical populations. Population A has little inter-individual variation in maximum performance or optimal temperature. Population B has inter-individual variation in maximum performance, but no variation in optimal temperature. Population C has little inter-individual variation in maximum performance, but has inter-individual variation in optimal temperature. The grey lines are the individual curves and the dashed line is the population mean. The error bars indicate one standard deviation in performance at the body temperatures used in the ANOVA.

If from the shape of the curve it is clear that a $T_{\rm o}$ range rather than a single $T_{\rm o}$ value is more reasonable for a species, one can easily calculate a specified performance breadth (e.g., 99% or 95%) from a curve (or, alternatively, from the polygon method proposed by van Berkum, 1986). Such a breadth is more easily interpretable biologically than a range obtained from a post-hoc test following ANOVA.

The ANOVA approach can be informative about the variability in performance at a given $T_{\rm b}$, which may be of biological interest also. In such case, a repeated measures design with ordered factors ANOVA must be used. When the goal is to describe performance curves, however, ANOVA should be avoided. Huey and Stevenson (1979) stressed that non-linear curves should be fitted to performance data whenever possible. We reiterate their suggestion. Stevenson et al. (1985) and Huey (1982) provide several examples of curves that can be used to describe thermal sensitivity.

In our example, we used a logistic-exponential curve fitting approach because it fitted our data well. We do not contend that this form of equation is necessarily the best for all situations. Polynomial curve fitting or the minimum convex polygon method (van Berkum, 1986) can also be used to extract descriptive statistics from thermal performance curves. The latter polygon method is a good alternative when a curve cannot be run through several points. In addition to avoiding erroneous biological interpretations and loss of information (e.g., due to an unordered ANOVA design), the approaches suggested here are more consistent with the concept of thermal reaction norm because performance is treated as a continuous function of temperature. Finally, if the goal of the study is beyond the general description of the curve, more elaborate methods have been proposed recently (Izem et al., 2003; Izem and Kingsolver, 2005). Those approaches allow quantifying the variation in the shape of performance curves and partitioning environmental from genetic effects.

Acknowledgments

Funding for this study was provided by the University of Ottawa, a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada to GBD and a graduate studies scholarship from the Fonds Québecois de Recherche sur la Nature et les Technologies to GB.

References

- Angilletta, M.J., Hill, T., Robson, M.A., 2002a. Is physiological performance optimized by thermoregulatory behavior?: a case study of the eastern fence lizard, *Sceloporus undulatus*. J. Therm. Biol. 27, 199–204.
- Angilletta, M.J., Niewiarowski, P.H., Navas, C.A., 2002b. The evolution of thermal physiology in ectotherms. J. Therm. Biol. 27, 249–268.

- Blouin-Demers, G., Weatherhead, P.J., McCracken, H.A., 2003. A test of the thermal coadaptation hypothesis with black rat snakes (*Elaphe obsoleta*) and northern water snakes (*Nerodia sipedon*). J. Therm. Biol. 28, 331–340.
- Chen, X.J., Xu, X.F., Ji, X., 2003. Influence of body temperature on food assimilation and locomotor performance in white-striped grass lizards, *Takydromus wolteri* (Lacertidae). J. Therm. Biol. 28, 385–391.
- Du, W.G., Yan, S.J., Ji, X., 2000. Selected body temperature, thermal tolerance and thermal dependence of food assimilation and locomotor performance in adult blue-tailed skinks, *Eumeces elegans*. J. Therm. Biol. 25, 197–202.
- Elsworth, P.G., Seebacher, F., Franklin, C.E., 2003. Sustained swimming performance in crocodiles (*Crocodylus porosus*): effects of body size and temperature. J. Herpetol. 37, 363–368.
- Huey, R.B., 1982. Temperature, physiology, and the ecology of reptiles. In: Gans, C., Pough, F.C. (Eds.), Biology of the Reptilia. Academic Press, New York, pp. 25–74.
- Huey, R.B., 1991. Physiological consequences of habitat selection. Am. Nat. 137, S91–S115.
- Huey, R.B., Bennett, A.F., 1987. Phylogenetic studies of coadaptation preferred temperatures versus optimal performance temperatures of lizards. Evolution 41, 1098–1115.
- Huey, R.B., Kingsolver, J.G., 1989. Evolution of thermal sensitivity of ectotherm performance. Trends Ecol. Evol. 4, 131–135.
- Huey, R.B., Stevenson, R.D., 1979. Integrating thermal physiology and ecology of ectotherms—discussion of approaches. Am. Zool. 19, 357–366.
- Izem, R., Kingsolver, J.G., 2005. Variation in continuous reaction norms: quantifying directions of biological interest. Am. Nat. 166, 277–289.
- Izem, R., Marron, J.S., Kingsolver, J.G., 2003. Analyzing variation in thermal performance curves. Integrative Comparative Biol. 43, 929.
- Ji, X., Zhou, W.H., He, G.B., Gu, H.Q., 1993. Food-intake, assimilation efficiency, and growth of juvenile lizards *Takydromus septentrionalis*. Comp. Biochem. Physiol. A 105, 283–285.
- Ji, X., Du, W.G., Sun, P.Y., 1996. Body temperature, thermal tolerance and influence of temperature on sprint speed and food assimilation in adult grass lizards, *Takydromus septentrionalis*. J. Therm. Biol. 21, 155–161.
- McConnachie, S., Alexander, G.J., 2004. The effect of temperature on digestive and assimilation efficiency, gut passage time and appetite in an ambush foraging lizard, *Cordylus melanotus melanotus*. J. Comp. Physiol. B 174, 99–105.
- Potvin, C., Lechowicz, M.J., Tardif, S., 1990. The statistical analysis of ecophysiological response curves obtained from experiments involving repeated measures. Ecology 71, 1389–1400.
- Scribner, S.J., Weatherhead, P.J., 1995. Locomotion and antipredator behavior in three species of semiaquatic snakes. Can. J. Zool. 73, 321–329.
- Stevenson, R.D., Peterson, C.R., Tsuji, J.S., 1985. The thermal dependence of locomotion, tongue flicking, digestion, and oxygen consumption in the wandering garter snake. Physiol. Zool. 58, 46–57.
- Tracy, C.R., Christian, K.A., 1986. Ecological relations among space, time, and thermal niche axes. Ecology 67, 609–615.
- van Berkum, F.H., 1986. Evolutionary patterns of the thermal sensitivity of sprint speed in *Anolis* lizards. Evolution 40, 594–604.
- Weatherhead, P.J., Robertson, I.C., 1992. Thermal constraints on swimming performance and escape response of northern water snakes (*Nerodia sipedon*). Can. J. Zool. 70, 94–98.
- Witz, B.W., Lawrence, J.M., 1993. Nutrient absorption efficiencies of the lizard, *Cnemidophorus sexlineatus* (Sauria, Teiidae). Comp. Biochem. Physiol. A 105, 151–155.
- Zhang, Y.P., Ji, X.A., 2004. The thermal dependence of food assimilation and locomotor performance in southern grass lizards, *Takydromus sexlineatus* (Lacertidae). J. Therm. Biol. 29, 45–53.