European Journal of Applied Physiology (2020) 120:765–770
https://doi.org/10.1007/s00421-020-04314-8

ORIGINAL ARTICLE

A regression method for the power–duration relationship when both variables are subject to error

Giovanni Vinetti[1] · Anna Taboni[2] · Guido Ferretti[1,2]

Received: 10 December 2019 / Accepted: 1 February 2020 / Published online: 20 February 2020

Abstract

Purpose The power–duration relationship has been variously modelled, although duration must be acknowledged as the dependent variable and is supposed to represent the only source of experimental error. However, there are certain situations, namely extremely high power outputs or outdoor field conditions, in which the error in power output measurement may not remain negligible. The geometric mean (GM) regression method deals with the assumption that also the independent variable is subject to a certain amount of experimental error, but has never been utilized in this context.

Methods We applied the GM regression method for the two- and three-parameter critical power models and tested it against the usual weighted least square (WLS) procedure with our previous published data.

Results There were no significant differences between parameter estimates of WLS and GM. Bias and limit of agreements between the two methods were low, while correlation coefficients were high (0.85–1.00).

Conclusions GM provided equivalent results with respect to WLS in fitting the critical power model to experimental data and for its conceptual characteristics must be preferred wherever concerns on the precision of *P* measurement are present, such as for in-field power meters.

Keywords Curve fitting · Cycling · Hyperbolic model · Model II nonlinear regression · Reduced major axis regression · Sport

Abbreviations

2-p	Two-parameter critical power model
3-p	Three-parameter critical power model
CP	Critical power
GM	Geometric mean regression method
k	Time asymptote of the critical power model
P	Mechanical power output
P_0	Power limit of the critical power model as time to exhaustion approaches 0 s
SEE	Standard error of the estimate
$T_{\lim}$	Time to exhaustion
W'	Curvature constant of the critical power model
WLS	Weighted least square regression method
ε_{P}	Absolute error in power output measurement
$\varepsilon_{\mathrm{Tlim}}$	Absolute error in time to exhaustion measurement

Communicated by Jean-René Lacour.

✉ Giovanni Vinetti
g.vinetti001@unibs.it

[1] Department of Molecular and Translational Medicine, University of Brescia, Viale Europa 11, 25123 Brescia, Italy

[2] Department of Anesthesiology, Pharmacology, Intensive Care and Emergencies, University of Geneva, 1 rue Michel Servet, 1211 Geneva 4, Switzerland

Introduction

The power–duration relationship is a well-established framework for modelling human performance and in its hyperbolic form is known also as the critical power model (Morton and Hodgson 1996). Although different parameterizations and re-arrangements of the equation are possible, duration, i.e. the time to exhaustion ($T_{\lim}$), must be acknowledged as the dependent variable (Morton and Hodgson 1996; Jones et al. 2010). This has sound (bio)logical foundations: the external power output (P) is the major determinant of the time course of exhaustion-related physiological variables (Poole et al. 1988; Black et al. 2017; Vinetti et al. 2017) and not vice versa. This choice is also justified by statistical theory: due to the intra-individual biological variability of $T_{\lim}$, the absolute random error in P (ε_{P}) could be judged negligible as compared to the absolute error in $T_{\lim}$ ($\varepsilon_{\mathrm{Tlim}}$) (Morton and

Hodgson 1996). Thus, the regression procedure can focus on minimizing the distances along the T_{lim}-axis only. Moreover, the relative error of T_{lim} (i.e., the ratio $\varepsilon_{Tlim}/T_{lim}$) is known to increase with T_{lim} itself (Poole et al. 1988; Hinckson and Hopkins 2005; Faude et al. 2017), thus representing a source of heteroscedasticity. Therefore, the weighted least squares (WLS) regression has been proposed as the most appropriate method to fit the critical power model to experimental data (Morton and Hodgson 1996; Morton 1996).

However, there are situations in which sources of random error must be acknowledged in both T_{lim} and P. In the exemplary case of cycling, power meter technology is characterised by a fairly low relative random error (i.e., a low ε_P/P, good precision) up to 360 W (Maier et al. 2017), but this implies that ε_P increases with P. Moreover, it is not excluded that also relative error increases with higher P or due to environmental factors such as vibrations, external shocks, changing ambient temperature (Maier et al. 2017). For those reasons, the more intense is the cycling burst, the higher is the uncertainty about P. However, if such an extreme P is carried until exhaustion, the error in T_{lim} (ε_{Tlim}) is very low (since it is proportional to T_{lim} itself, which is here very low too), thus the assumption that ε_{Tlim} is disproportionally greater than ε_P is no longer valid and alternative statistical approaches must be adopted.

While ordinary and weighted least squares methods (named Model I regressions) minimize the distance in the dependent variable only, with an implicit assumption that the dependent variable is error-free, Model II regression analysis deals with the assumption that also the independent variable is subject to a certain amount of experimental error, thus minimizing the distance in both axes with several approaches (Ludbrook 2012). Among them, the geometric mean (GM, also known as reduced major axis regression or ordinary least product) regression method, minimizes the sum of the areas determined by the curve and the horizontal and the vertical lines connecting each experimental point to the curve (Brace 1977; Ludbrook 2012) and it can be generalized to nonlinear functions (Ebert and Russell 1994). With respect to other Model II regression methods, GM does not need an arbitrary a priori estimation of the measurements' errors. In fact, GM assumes that the ratio between the magnitude of the absolute error in the independent and the dependent variable is approximately equal to the absolute local slope of the function (Brace 1977). Luckily, the hyperbolic nature of the critical power model is in line with this assumption: when the slope dT_{lim}/dP is high (Fig. 1, point A), the ratio $\varepsilon_{Tlim}/\varepsilon_P$ is high, and vice versa (Fig. 1, point B). In other words in the GM method, when increasing T_{lim} and decreasing P, progressively less weight is given to the minimization of the T_{lim}-axis distance (similarly to the WLS proposed by Morton 1996) with the addition that progressively more weight is given to the minimization of the P-axis distance.

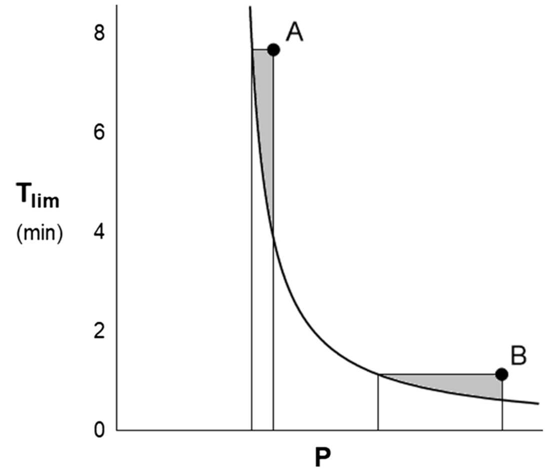

Fig. 1 Geometric mean regression minimizes the sum of the highlighted grey areas. The method's assumption is that the ratio of the error of T_{lim} and P ($\varepsilon_{Tlim}/\varepsilon_P$) approximately equal to the local slope (dT_{lim}/dP) of the hyperbola. This assumption is met. Point A: low P (then low ε_P), high T_{lim} (then high ε_P) and high dT_{lim}/dP, but also high $\varepsilon_{Tlim}/\varepsilon_P$; point B high P (then high ε_P), low T_{lim} (then low ε_{Tlim}) and low dT_{lim}/dP, but also low $\varepsilon_{Tlim}/\varepsilon_P$

With the present report, we sought to illustrate the GM regression method for the two- and three-parameter critical power models and testing its reliability against the WLS method on our previously published P–T_{lim} data (Vinetti et al. 2019).

Methods

Data from Vinetti et al. (2019) were retrospectively fitted by the hyperbolic critical power model by means of nonlinear regression analysis with both the WLS and the GM method. The general form of the model is:

$$T_{lim} = \frac{W'}{P - CP} + k, \tag{1}$$

where W' is the curvature constant, CP (critical power) is the power asymptote, and k is the time asymptote (3-parameter model, 3-p). The theoretical maximal instantaneous power (P_0) was calculated as the T_{lim}-axis intercept of Eq. (1). The two-parameter model (2-p) was obtained by removing the parameter k. The 2-p model was applied to six data points within the severe exercise intensity domain (85–120% of the maximal aerobic power), while the 3-p model included also three points in the extreme domain (150–250% of the maximal aerobic power). The GM method was developed with the same approach of Ebert and Russell (1994) (see Appendix for further details). Briefly, the highlighted areas in Fig. 1 were set as the loss function to be minimized. For the WLS method, the loss function was set as the squared residuals of each ith data point multiplied by the weighting factor $1/T_{\lim(i)}^{2}$ (Morton 1996). The standard error of the parameter estimates (SEE) was calculated by bootstrapping. SEE of P_0 was calculated from the alternative parameterization of Eq. (1) (see Appendix). Paired-sample t test, linear regression and Bland–Altman analysis were used to compare parameter estimates from GM and WLS. Slope and intercept of linear regressions between parameter estimates were also calculated with the GM method as recommended when comparing methods of measurements (Ludbrook 2012). The level of significance was set at $p < 0.05$. The statistical package SPSS (Version 23.00, IBM Corp., Armonk, NY) was used.

Results

All parameter estimates were not significantly different between GM and WLS (Table 1). SEEs were also non-significantly different, except for that of *CP* in the 2-p model, which was lower with GM. Concerning 2-p model, GM yielded *CP* and *W′* identical to WLS, with bias – 0.7 ± 1.6 W and – 0.1 ± 0.5 kJ, respectively, and 95% limits of agreement – 3.8 and 2.4 W, and – 1.1 and 0.9 kJ, respectively (Fig. 2). In 3-p model, *CP* was identical between GM and WLS, with bias – 0.2 ± 2.0 W and 95% limits of agreement – 4.0 and 3.7 W, while *W′*, *k* and P_0 present some marginal, nonsignificant differences, with bias – 0.6 ± 0.9 kJ, 1.7 ± 3.0 s and 56 ± 217 W, respectively, and 95% limits of agreement of – 2.4 and 1.3 kJ, – 4.2 and 7.7 s, and – 370 and 483 W, respectively (Fig. 3).

Table 1 Average parameter estimates and standards errors obtained with the geometric mean (GM) and the weighted least square (WLS) regression methods

Model	Parameter	GM	WLS	*p* value
2-p	*CP* (W)	179 ± 26	179 ± 27	0.20
	SEE	5 ± 2	6 ± 3	0.01
	W′ (kJ)	15.0 ± 3.6	14.9 ± 3.4	0.66
	SEE	1.8 ± 0.7	2.1 ± 0.8	0.18
3-p	*CP* (W)	178 ± 26	178 ± 27	0.85
	SEE	6 ± 3	6 ± 3	0.82
	W′ (kJ)	16.4 ± 4.2	15.8 ± 3.6	0.09
	SEE	2.7 ± 1.1	2.7 ± 0.9	0.88
	k (s)	– 17.7 ± 9.1	– 16.0 ± 7.9	0.10
	SEE	10.9 ± 3.3	12.3 ± 3.4	0.19
	P_0 (W)	1247 ± 368	1303 ± 408	0.43
	SEE	844 ± 768	877 ± 696	0.68

2-p two-parameter, *3-p* three-parameter, *CP* critical power, *k* time asymptote constant, P_0 maximal instantaneous power, SEE standard error of the estimate, *W′* curvature constant

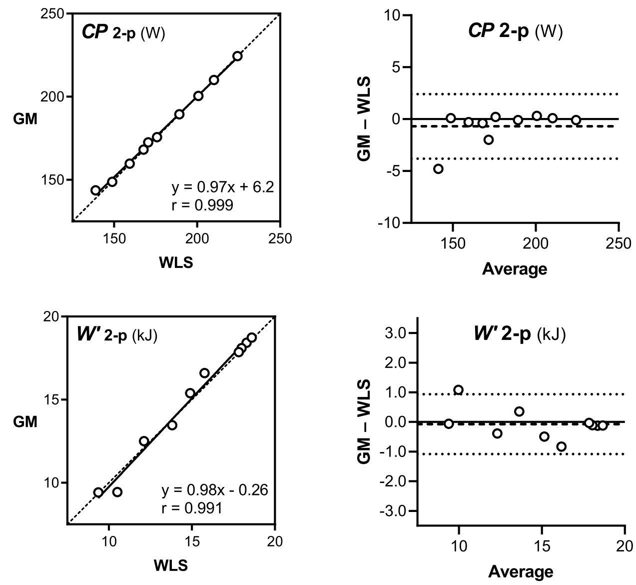

Fig. 2 Comparison of parameter estimates of the 2-p model obtained with GM and WLS regression methods. Left column: regression (continuous) lines with identity (dashed) lines; Right column: Bland–Altman plots including bias (dashed lines) and 95% limits of agreement (dotted lines). *CP* critical power, *W’* curvature constant

Discussion

From a statistical viewpoint, the implemented hyperbolic GM regression method has several advantages over WLS: (1) it progressively accounts also for an error in the P variable when higher P are investigated, (2) it does not require further weighting procedures since it is intrinsically weighted both for P and $T_{\lim}$ and (3) it is independent of whether model’s equation is expressed in terms of $T_{\lim}$ or P. GM belongs to the broader context of errors-in-variable models, mostly confined in econometrics (Schennach 2016)—where large amount of data can be collected and more complex assumptions and analyses are required—and it represents a concise method that is well suited also for the exercise science field.

From an experimental viewpoint, GM was successful in fitting the two- and three-parameter model, leading to results similar to the traditional WLS method. This is particularly evident in the 2-p model, where all parameters were perfectly identical (Fig. 2), conforming the theoretical prediction that when all points in the steep part of the curve and the $\varepsilon_{\mathrm{Tlim}}/\varepsilon_{\mathrm{P}}$ is low (point A of Fig. 1), GM tends to mimic WLS. A slightly lower agreement between the two methods is present in the 3-p model for those parameters influencing the extreme exercise intensity domain (point B of Fig. 1), namely k and P_0 (Fig. 3). Still, there are no statistically significant differences, probably because of the relatively high reliability of the stationary cycle-ergometer used in the study also for extreme values of P (Vinetti et al. 2019). We expect

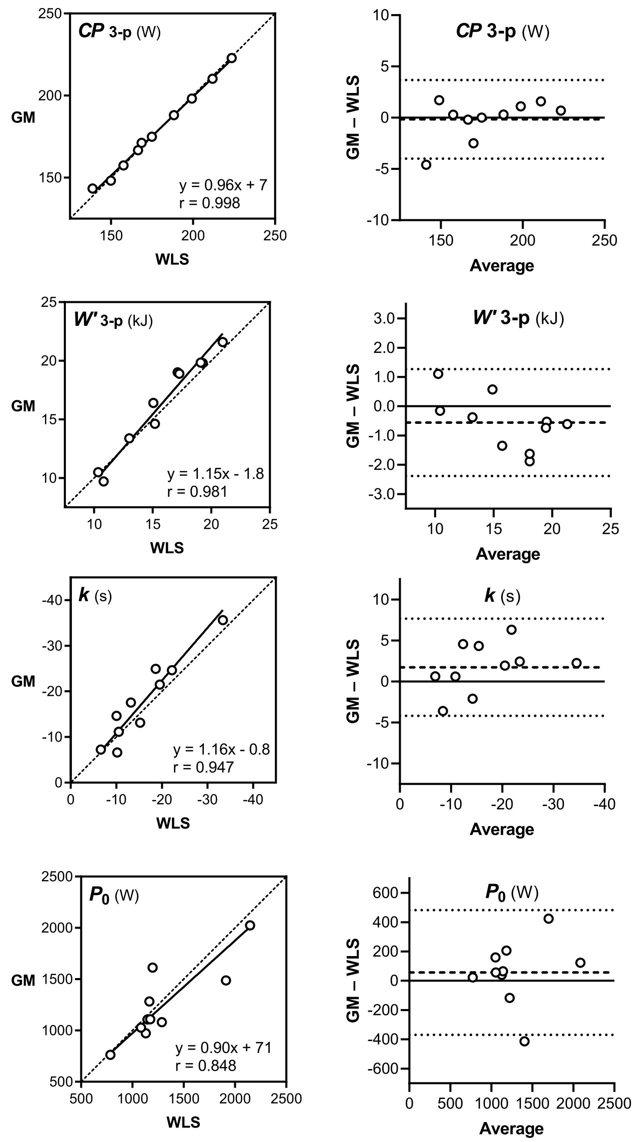

Fig. 3 Comparison of parameter estimates of the 3-p model obtained with GM and WLS regression methods. Left column: regression (continuous) lines with identity (dashed) lines; Right column: Bland–Altman plots including bias (dashed lines) and 95% limits of agreement (dotted lines). *CP* critical power, *W'* curvature constant, *k* time asymptote constant, P_0 maximal instantaneous power

further divergence in the two methods with less precise ergometers: in this case, parameter estimation with the GM method should be preferred.

In this context, it is noteworthy that the choice of the regression model and method should be an a priori decision based on the identification of the sources of experimental error. Not surprisingly, studies using a systematic a posteriori selection based on the lowest SEE for *CP* and *W′* are necessarily biased towards models that erroneously assume T_{lim} as the independent, error-free, variable (see for example Black et al. 2017). In fact, since data points mostly lie in the time window where there is both high dT_{lim}/dP and $\varepsilon_{\text{Tlim}}/\varepsilon_{\text{P}}$ (2–15 min, point A of Fig. 1), a high distance in the T_{lim}-axis corresponds to a small distance in the other axis (either *P* or the total work, *W*); therefore, assuming *P* or *W* instead of T_{lim} as the dependent variable is likely to provide lower residual sum of squares (and SEE) and thus the perception of a better statistical fitting. However, this assumption is not consistent with the process that generated the data, but just with the data itself.

In conclusion, when fitting the critical power model to experimental data with a low error in *P*, the parameters provided by GM do not differ from those provided by WLS. However, for its intrinsic characteristics, GM is conceptually preferable wherever concerns on the precision of *P* measurement are present. Therefore, it should become the method of choice for statistical treatment of critical power data. Future testing of the GM method with more error-prone data sets, such as from in-field power meters, is welcome.

Author contribution statement GV and GF conceived the study; GV analyzed the data; all authors interpreted the results; GV drafted the first version of the manuscript; all authors edited and revised the manuscript and approved the final version before submission.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Appendix

The developments detailed in this Appendix describe how to apply the GM method to the hyperbolic function expressed in Eq. (1). The interested reader can refer to Ludbrook (2012) for practical considerations on linear GM regression and to Ebert and Russell (1994) for an example of adaptation of GM to nonlinear functions, as in the present study.

The assumption behind GM regression is that the ratio between the magnitude of the absolute error in *y* and in *x* is similar to the absolute slope of the fitted curve (Brace 1977). In the case of Eq. (1) we set $y = T_{\text{lim}}$ and $x = P$, thus the slope is equal to:

$$\left|\frac{dT_{\text{lim}}}{dP}\right| = \left|-\frac{W'}{(P-\text{CP})^2}\right| \sim \frac{W'}{P^2} \tag{2}$$

Is this compatible with the ratio between the magnitude of the absolute error in T_{lim} and P ($\varepsilon_{\text{Tlim}}/\varepsilon_{\text{P}}$)? Yes, both for biological and technological reasons. We know in fact that $\varepsilon_{\text{Tlim}}$ and ε_{P} are T_{lim} and in P are proportional to T_{lim} and P, respectively thus:

Table 2 SPSS commands for GM regression

Model	2-p	3-p (original parameterization)	3-p (alternative parameterization)
Dependent variable	T	T	T
Independent variable	P	P	P
Model expression	W/(P-CP)	W/(P-CP)+k	W/(P-CP)-W/(P0-CP)
Parameters (initial values)	CP (1) W (1)	CP (1) W (1) k (-1)	CP (1) W (1) P0 (1)
Loss function	T*(P-CP)-W*(1+LN(T*(P-CP))-LN(W))	(T-k)*(P-CP)-W*(1+LN((T-k)*(P-CP))-LN(W))	(T+W/(P0-CP))*(P-CP)-W*(1+LN((T+W/(P0-CP))*(P-CP))-LN(W))
Constraints	CP <= lowest P − 1 W >= 1	CP <= lowest P − 1 W >= 1 k <= − 0.001	CP <= lowest P − 1 W >= 1 P0 >= highest P + 1

Constraints are added to avoid negative or zero arguments in logarithms and ease the iterative computation

Highest and lowest P represents the highest and lowest power output in a given individual dataset

$$\frac{\varepsilon_{T_{\lim}}}{\varepsilon_P} \sim \frac{T_{\lim}}{P} \tag{3}$$

Inserting (1) into (3):

$$\frac{\varepsilon_{T_{\lim}}}{\varepsilon_P} \sim \frac{W'}{P(P-\mathrm{CP})} + \frac{k}{P} \sim \frac{W'}{P^2} \tag{4}$$

Therefore,

$$\frac{\varepsilon_{T_{\lim}}}{\varepsilon_P} \sim \left|\frac{dT_{\lim}}{dP}\right| \tag{5}$$

Thus, the assumption required for GM regression is met. Since GM is independent of the choice of the dependent variable, the demonstration is still valid when inverting x and y.

We, therefore, define the loss function as the grey area of Fig. 1, that is absolute difference between the rectangles formed by the thin lines and the integral of the function in the same P interval (white area between the curve and the P-axis). For a generic data point with coordinates (P_i, T_i)

$$T_i\left(\frac{W'}{T_i-k} + \mathrm{CP} - P_i\right) - \int_{P_i}^{\frac{W'}{T_i-k}+\mathrm{CP}} \left(\frac{W'}{P-\mathrm{CP}} + k\right) dP \tag{6}$$

It is noteworthy that Eq. (6) in this formulation has always a positive value thus there is not the need of a modulo operator. Solving and simplifying yields to the following loss function:

$$\left(T_i - k\right)\left(P_i - \mathrm{CP}\right) - W'\left\{1 + \ln\left[\left(T_i - k\right)\left(P_i - \mathrm{CP}\right)\right] - \ln W'\right\} \tag{7}$$

By setting $k=0$ we obtain the 2-p version of the loss function, while $k = -\frac{W'}{P_0-\mathrm{CP}}$ leads to the alternative parameterization of Eq. (1) (Morton 1996). The commands for SPSS are derived from this demonstration are reported Table 2, although it must be remembered that equivalent results are granted if starting the demonstration assuming P instead of $T_{\lim}$ as the dependent variable.

References

Black MI, Jones AM, Blackwell JR, Bailey SJ, Wylie LJ, McDonagh STJ, Thompson C, Kelly J, Sumners P, Mileva KN, Bowtell JL, Vanhatalo A (2017) Muscle metabolic and neuromuscular determinants of fatigue during cycling in different exercise intensity domains. J Appl Physiol 122:446–459

Brace RA (1977) Fitting straight lines to experimental data. Am J Physiol 233:R94–R99

Ebert TA, Russell MP (1994) Allometry and model II non-linear regression. J Theor Biol 168:367–372

Faude O, Hecksteden A, Hammes D, Schumacher F, Besenius E, Sperlich B, Meyer T (2017) Reliability of time-to-exhaustion and selected psycho-physiological variables during constant-load cycling at the maximal lactate steady-state. Appl Physiol Nutr Metab 42:142–147

Hinckson EA, Hopkins WG (2005) Reliability of time to exhaustion analyzed with critical-power and log-log modeling. Med Sci Sports Exerc 37:696–701

Jones AM, Vanhatalo A, Burnley M, Morton RH, Poole DC (2010) Critical power: implications for determination of V̇O2max and exercise tolerance. Med Sci Sports Exerc 42:1876–1890

Ludbrook J (2012) A primer for biomedical scientists on how to execute model II linear regression analysis. Clin Exp Pharmacol Physiol 39:329–335

Maier T, Schmid L, Müller B, Steiner T, Wehrlin J (2017) Accuracy of cycling power meters against a mathematical model of treadmill cycling. Int J Sports Med 38:456–461

Morton RH (1996) A 3-parameter critical power model. Ergonomics 39:611–619

Morton RH, Hodgson DJ (1996) The relationship between power output and endurance: a brief review. Eur J Appl Physiol Occup Physiol 73:491–502

Poole DC, Ward SA, Gardner GW, Whipp BJ (1988) Metabolic and respiratory profile of the upper limit for prolonged exercise in man. Ergonomics 31:1265–1279

Schennach SM (2016) Recent advances in the measurement error literature. Annu Rev Econom 8:341–377

Vinetti G, Fagoni N, Taboni A, Camelio S, di Prampero PE, Ferretti G (2017) Effects of recovery interval duration on the parameters of the critical power model for incremental exercise. Eur J Appl Physiol 117:1859–1867

Vinetti G, Taboni A, Bruseghini P, Camelio S, D'Elia M, Fagoni N, Moia C, Ferretti G (2019) Experimental validation of the 3-parameter critical power model in cycling. Eur J Appl Physiol 119:941–949