Journal of Sports Sciences

ISSN: 0264-0414 (Print) 1466-447X (Online) Journal homepage: http://www.tandfonline.com/loi/rjsp20

Assessing the limitations of the Banister model in monitoring training

Philippe Hellard , Marta Avalos , Lucien Lacoste , Frederic Barale , Jean-Claude Chatard & Gregoire P. Millet

To cite this article: Philippe Hellard , Marta Avalos , Lucien Lacoste , Frederic Barale , Jean-Claude Chatard & Gregoire P. Millet (2006) Assessing the limitations of the Banister model in monitoring training, Journal of Sports Sciences, 24:05, 509-520, DOI: 10.1080/02640410500244697

To link to this article: http://dx.doi.org/10.1080/02640410500244697

Published online: 18 Feb 2007.

Journal of Sports Sciences, May 2006; 24(5): 509 – 520

Taylor & Francis
Taylor & Francis Group

Assessing the limitations of the Banister model in monitoring training

PHILIPPE HELLARD[1], MARTA AVALOS[2], LUCIEN LACOSTE[1], FREDERIC BARALE[1], JEAN-CLAUDE CHATARD[3], & GREGOIRE P. MILLET[4]

[1]*Département d'Etudes et Recherches, Fédération Française de Natation, Paris,* [2]*INSERM EO338 Biostatistics, ISPED, Université Victor Segalen Bordeaux 2,* [3]*Laboratoire de Physiologie, GIP Exercice, Faculté de Médecine de Saint-Étienne, and* [4]*Faculté des Sciences du Sport, 34090 Montpellier, France*

(Accepted 1 July 2005)

Abstract

The aim of this study was to carry out a statistical analysis of the Banister model to verify how useful it is in monitoring the training programmes of elite swimmers. The accuracy, the ill-conditioning and the stability of this model were thus investigated. The training loads of nine elite swimmers, measured over one season, were related to performances with the Banister model. First, to assess accuracy, the 95% bootstrap confidence interval (95% CI) of parameter estimates and modelled performances were calculated. Second, to study ill-conditioning, the correlation matrix of parameter estimates was computed. Finally, to analyse stability, iterative computation was performed with the same data but minus one performance, chosen at random. Performances were related to training loads for all participants ($R^2 = 0.79 \pm 0.13$, $P < 0.05$) and the estimation procedure seemed to be stable. Nevertheless, the range of 95% CI values of the most useful parameters for monitoring training was wide: $\tau_a = 38$ (17, 59), $\tau_f = 19$ (6, 32), $t_n = 19$ (7, 35), $t_g = 43$ (25, 61). Furthermore, some parameters were highly correlated, making their interpretation worthless. We suggest possible ways to deal with these problems and review alternative methods to model the training – performance relationships.

Keywords: *Bootstrap method, inaccuracy, instability, training loads, swimming*

Introduction

The relationships between training and performance can be of practical use to sports elite coaches when organizing their athletes' training programmes. These relationships are known to be highly individualized (Avalos, Hellard, & Chatard, 2003; Mujika, Chatard, Busso, & Geyssant, 1996b). These differences can be attributed to genetic factors (Wolfarth *et al.*, 2000), individual training background (Avalos *et al.*, 2003; Mujika *et al.*, 1996a), psychological factors (Banister, Calvert, Savage, & Bach, 1975), technical factors (Toussaint & Hollander, 1994; Wakayoshi, D'Acquisto, Cappaert, & Troup, 1995) and speciality (Avalos *et al.*, 2003; Mujika *et al.*, 1996a, 1996b; Stewart & Hopkins, 2000). There is general consensus that modelling the training – performance relationship provides pertinent information about inter-individual differences that allows the construction of highly individualized training programmes (Avalos *et al.*, 2003; Banister *et al.*, 1975; Banister, Carter, & Zarkadas, 1999; Morton, 1991, 1997; Morton, Fitz-Clarke, & Banister, 1990; Mujika *et al.*, 1996a, 1996b). The model proposed by Banister *et al.* (1975) and its extensions (Busso, 2003; Busso, Benoit, Bonnefoy, Feasson, & Lacour, 2002; Busso, Denis, Bonnefoy, Geyssant, & Lacour, 1997; Calvert, Banister, & Savage, 1976) aimed to relate training loads to performance, taking into account the dynamic and temporal characteristics of training and therefore the effects of load sequences over time. These effects may be described by two antagonistic transfer functions: first, a positive influence that synthesizes all the positive effects leading to an increase in performance, and second, a negative function that synthesizes all the negative effects leading to short- or long-term fatigue and that has a negative influence on performance (Busso, Candau, & Lacour, 1994; Mujika *et al.*, 1996a). The function is as follows (Banister *et al.*, 1975; Busso *et al.*, 1994):

$$p_t = p_0 + k_a \sum_{s=0}^{t-1} e^{-(t-s)/\tau_a} w_s - k_f \sum_{s=0}^{t-1} e^{-(t-s)/\tau_f} w_s$$

Correspondence: M. Avalos, INSERM EO338 Biostatistics, ISPED, Université Victor Segalen Bordeaux 2, 146 rue Léo Saignat, 33076 Bordeaux Cedax, France. E-mail: marta.avalos@isped.u-bordeaux2.fr

ISSN 0264-0414 print/ISSN 1466-447X online © 2006 Taylor & Francis
DOI: 10.1080/02640410500244697

where p_t is the modelled performance at time t; p_0 is the initial performance level; k_a and k_f are the fitness and fatigue magnitude factor, respectively; τ_a and τ_f are the fitness and fatigue decay time constant, respectively; and w_t is the known training load per week (or day) from the first week of training to the week (or day) preceding the performance. These parameters were interpreted as individual response profiles (Mujika *et al.*, 1996a), presented as usable within the training prescription; for example, t_n, the time to recover performance and t_g, the time to peak performance after the completion of training (Fitz-Clarke, Morton, & Banister, 1991). However, several authors (Busso *et al.*, 1990; Mujika *et al.*, 1996a; Taha & Thomas, 2003) reported that the practical interpretation of the positive and negative influences might be difficult. For example, Busso *et al.* (1990) reported a positive correlation between testosterone concentration and the function of fatigue, whereas a negative relationship was anticipated. Recently, Taha and Thomas (2003) criticized the different models stemming from the original Banister model (Banister *et al.*, 1975), stressing: (1) the inability of the model to predict with accuracy future performance; (2) the differences between the estimated time course of change in performance and experimental observations; and (3) the fact that most of these models were poorly corroborated by physiological mechanisms.

Moreover, quality measures of models were commonly limited to the determination coefficient (R^2) associated with the F-ratio test. R^2 evaluates the goodness of fit but does not necessarily guarantee accurate prediction (Atkinson & Nevill, 1998; Bartlett, 1997). With the F-ratio test, the hypothesis H_0: $k_a = 0$ and $k_f = 0$ and $\tau_a = 0$ and $\tau_f = 0$ is tested versus H_1: $k_a \neq 0$ and/or $k_f \neq 0$ and/or $\tau_a \neq 0$ and/or $\tau_f \neq 0$ (Sen & Shrivastava, 1990). A significant F-test implies that the Banister model is better than the constant model, $p_t = p_0$, in which training load does not affect performance. Complementary measures are needed to indicate accuracy and sensitivity (Efron & Tibshirani, 1993; Wetherill, Duncombe, & Kenward, 1986). Thus the 95% CI quantifies accuracy of parameter estimates and modelled performances. The asymptotic correlation matrix of parameter estimates assesses ill-conditioning and variability, which affect the accuracy and precision of parameter estimates (Arsac, Thiaudière, Diolez, & Gerville-Réache, 2004; Bates & Watts, 1988; Belsley, 1991). A different criterion of quality is stability of the estimation procedure. A method is unstable if small perturbations in the data are able to cause significant changes in the estimations (Breiman, 1996).

Also of importance is the number of data points required per parameter to perform appropriate statistical analyses. For multiple linear regression, a nominal number of 15 observations per parameter (except the intercept parameter) is recommended (Stevens, 1986). But since the Banister model is non-linear, inference is based on asymptotic theory (Bates & Watts, 1988; Davidian & Giltinan, 1995; Huet, Bouvier, Gruet, & Jolivet, 1996; Sen & Shrivastava, 1990), which implies more data points per parameter than for a linear regression model. In fact, the studies that have modelled "real" performances in elite sport did not exceed more than 20 performances per year (Millet *et al.*, 2002; Mujika *et al.*, 1996a).

The problem of accuracy when modelling the training–performance relationship is crucial in athletes of a high standard. The higher the standard of performance, the smaller the difference in performance. For example, during the Athens Olympic Games, the mean difference in the swimming finals was $2.2 \pm 0.8\%$ between the Olympic Champion and the poorest (8th) performance, while for the national championship this difference was $6.6 \pm 2.3\%$.

The aims of the present study were: (1) to assess the goodness of fit, accuracy, ill-conditioning and stability of the Banister model for real data, and (2) to review and suggest alternative methods to model the training–performance relationship.

Methods

Participants

Nine elite swimmers (five females, four males) participated in the present study. All swimmers were of international standard (i.e. qualified as junior or elite in the national team for the European, World Championships or Olympic Games). Written informed consent was obtained from all participants. Their training characteristics and performances were analysed over the course of a season consisting of a training period (52 weeks) and a rest period (8 weeks) (see Table I). An individualized training programme was prescribed by the coaches for each swimmer, based on age, training background, individual profiles and speciality. The proportion of the training performed at each intensity was compared between the group of sprint swimmers specializing in the 50 and 100 m events, the group of intermediate-distance swimmers specialising in the 100 and 200 m events, and the group of middle-distance swimmers specialising in the 200 and 400 m events. The season comprised four training cycles. A linear model of periodization characterized the training cycles (American College of Sports Medicine, 2002; Fry, Morton, & Keast, 1992): each training cycle, lasting between 8 and 14 weeks, commenced at a high training volume and low intensity. As training progressed, volume decreased and intensity increased. The final 3 weeks before competition was defined as the taper phase.

Training and performance

An incremental stepwise test to exhaustion (6×200 m) was performed four times during the season to determine the relationship between blood lactate concentration and swimming speed. Then, intensities of swim workouts [below ($I_1 \approx 2\,\text{mmol} \cdot l^{-1}$), equal to ($I_2 \approx 4\,\text{mmol} \cdot l^{-1}$), slightly above ($I_3 \approx 6\,\text{mmol} \cdot l^{-1}$) the onset of blood lactate accumulation, respectively; $I_4 \approx 10\,\text{mmol} \cdot l^{-1}$; I_5 = maximal intensity] were determined as proposed by Mujika *et al.* (1996a). Intensities I_6 and I_7 were equivalent to 40–70% and 70–100% of one-repetition maximum strength training, respectively.

For the group as a whole, the contents of the volume, intensity and taper phases of the last training cycle before the main competitions of the year were compared (Table II). Quantification of the training load was performed as indicated by Avalos *et al.* (2003). Briefly, it was expressed as a percentage of the maximal volume measured at each intensity throughout the study period for each participant.

For each swimmer, performances were measured during actual competition for the same event, throughout the study period. Since performance, in the case of swimming, is represented by a time, it is simpler to operate as the percentage of the best performance $\min_t (P_t)$ achieved in the course of the

Table I. Selected characteristics of the participants training over the one-year period studied.

	Sex	Age (years)	Height (m)	Mass (kg)	Training (km)	Event	Best performance	Number of performances	CV of perf. (%)
1	F	24	1.68	61	1402	100 m freestyle	00:55:65	11	2.3
2	F	21	1.73	62	1856	200 m butterfly	02:10:8	12	1.6
3	F	26	1.79	59	1677	200 m freestyle	01:59:86	21	2.2
4	M	27	1.85	84	1751	200 m medley	02:01:83	12	1.9
5	M	23	1.81	81	1340	100 m breaststroke	01:03:51	14	1.8
6	F	26	1.68	50	1477	200 m backstroke	02:15:00	18	2.4
7	M	20	1.86	80	1815	100 m freestlye	00:51:5	12	2.6
8	F	19	1.67	52	1916	200 m freestyle	02:03:51	11	1.6
9	M	23	1.88	84	1843	400 m freestyle	03:53:42	13	1.5
Mean		23.2	1.77	68.1	1675			13.2	2.0
s		2.8	0.08	14.0	215			2.4	0.4

Note: CV = coefficient of variation. Training (km) = total km swum during the period studied.

Table II. Annual swimming volume percentage and annual dry land training percentage (including rest periods) for each intensity and group.

	SS		IS		MDS	
Intensity	IS	MDS	SS	MDS	SS	IS
I_1	*	–	*	–	–	–
(% · year^{-1})	69.3 (5.4)		57.3 (4.9)		48.1 (4.1)	
I_2	*	–	*	–	–	–
(% · year^{-1})	21.6 (2.2)		34.6 (4.5)		44.1 (5.4)	
I_3	–	–	–	–	–	–
(% · year^{-1})	4.6 (2.2)		4.7 (0.4)		3.8 (1.1)	
I_4	–	–	–	#	–	#
(% · year^{-1})	2.3 (0.2)		2.5 (0.2)		3.3 (0.3)	
I_5	*	–	*	–	–	–
(% · year^{-1})	2.3 (0.2)		0.9 (0.1)		0.7 (0.2)	
I_6	–	–	–	–	–	–
(% · year^{-1})	41.7 (20.2)		37.8 (26.7)		29.6 (0.7)	
I_7	–	–	–	-	–	–
(% · year^{-1})	58.5 (20.2)		62.2 (27.7)		70.4 (0.7)	

Note: Values are mean (*s*). SS = sprint swimmers, IS = intermediate swimmers, MDS = middle-distance swimmers. *Significant difference between SS and IS, for each intensity ($P < 0.05$). #Significant difference between IS and MDS, for each intensity ($P < 0.05$). –, not significantly different. Mann-Whitney *U*-test for non-parametric distribution was used.

study period. Performance can thus be computed in the following manner:

$$p_t = \frac{\min_t (P_t)}{P_t} * 100$$

Performances achieved during the competition were compared with those achieved 3 weeks earlier during a preparatory competition.

Fitting the model

The model proposed by Banister *et al.* (1975) was used. Model parameters were estimated for each participant using the non-linear least squares iterative method, by minimizing the residual sum of quadratic differences between the real and the modelled performances (RSS) with a Gauss-Newton type algorithm (Bates & Watts, 1988; Davidian & Giltinan, 1995; Huet *et al.*, 1996; Sen & Shrivastava, 1990). The values of parameters have generally been reported as $\tau_a = 45$ days, $\tau_f = 15$ days, $k_a = 1$ arbitrary units (a.u.) and $k_f = 2$ a.u. (Morton *et al.*, 1990). However, in swimming, with similar swimmers and daily training loads, these values were $\tau_a = 41.4$ days, $\tau_f = 12.4$ days, $k_a = 0.128$ a.u. and $k_f = 0.055$ a.u. (Mujika *et al.*, 1996a). The latter were therefore used as the initial values in the iterative procedure for each participant in the present study. Computations were completed using Matlab 2000 (version 6.0, Optimization Toolbox, Mathworks Eds). After testing for the normality and homoscedasticity of the residuals, 95% CIs were calculated for the estimated parameters. The determination coefficient was calculated as: $R^2 = 1 - (\text{RSS} / \text{TSS})$, where TSS is the total sum of squares. In addition, the statistical significance of the fit was tested using analysis of variance (*F*-ratio test). The mean standard errors (MSE) were computed as the mean difference between modelled and real performances.

Calculation of t_n and t_g

The time to recover performance, t_n, and the time to peak performance after the completion of training, t_g, were calculated as

$$t_n = \frac{\tau_1 \tau_2}{\tau_1 - \tau_2} \ln \frac{k_2}{k_1} \quad \text{and} \quad t_g = \frac{\tau_1 \tau_2}{\tau_1 - \tau_2} \ln \left(\frac{k_2}{k_1} \frac{\tau_1}{\tau_2} \right)$$

(subscripts "1" and "2" correspond to "*a*" and "*f*") (Fitz-Clarke *et al.*, 1991).

Estimating the change in performance during the rest period

The time it took each swimmer to reach peak performance was evaluated during the rest period (during the final 8 weeks of the season) from modelled performances.

Accuracy

The bootstrap method was used to calculate the 95% CI of parameter estimates and modelled performances. Bootstrap is a powerful method for estimating test statistics like confidence intervals, especially in small samples (Efron & Tibshirani, 1993; Huet *et al.*, 1996). Briefly, the procedure consisted of resampling the original data set with replacement to create 1000 "bootstrap replicate" data sets of the same size as the original data set. A random number generator was used to determine which data from the original data set to include in a replicate data set. Therefore, a given datum could be used more than once in the replicate data set, or not at all. This was repeated 1000 times. For each parameter, the estimates that fell between the 2.5th and the 97.5th percentiles of the 1000 estimates were used to construct a 95% CI for parameter estimates (τ_a, τ_f, k_a, k_f, t_n, t_g) and modelled performances. In addition, the coefficient of variation (CV) was calculated as the ratio of the standard deviation over the mean of the 1000 replicates.

Ill-conditioning

Ill-conditioning is the non-linear generalization of the collinearity problem in linear regression. The asymptotic correlation matrix of parameter estimates was checked to determine whether any parameters were excessively highly correlated, since high correlations can reveal ill-conditioning problems, indicating that the model is over-parameterized for the data set (Bates & Watts, 1988). The correlation matrix was obtained as follows. Let x_i be the explanatory variable and y_i the response variable, $i = 1, \ldots, n$, where n is the number of observations. Suppose the relationship $y_i = \mathrm{f}(x_i, \theta) + \varepsilon_i$, where the function f is the deterministic part of the model, depending on the explanatory variable and some parameters $\theta = (\theta_1, \ldots, \theta_p)$, and ε_i is the random part of the model: $\varepsilon_i \sim \mathrm{N}(0, \sigma^2)$. Let $\hat{\theta}$ be the non-linear least squares estimation. Then, asymptotically, $\hat{\theta} \sim \mathrm{N}(\theta, \hat{\sigma}^2 \mathrm{V})$, where $\boldsymbol{V} = (\mathrm{D}\hat{\mathrm{f}}, \mathrm{D}\hat{\mathrm{f}})^{-1}$, $\mathrm{D}\hat{\mathrm{f}}$ is the derivative matrix of $\hat{\mathrm{f}}$ with respect to θ, $\hat{\mathrm{f}}$ is the estimation of f, and $\hat{\sigma}^2$ is the standard unbiased estimator of the error variance. Elements of the correlation matrix, ρ_{jk}, $j,k = 1, \ldots, p$, are then obtained from the covariance matrix as $\rho_{jk} = (\boldsymbol{V})_{jk} / [(\boldsymbol{V})_{jj} (\boldsymbol{V})_{kk}]^{1/2}$ (Bates & Watts, 1988; Davidian & Giltinan, 1995; Huet *et al.*, 1996; Sen & Shrivastava, 1990).

Stability

The stability of the Banister model fitted by minimizing the RSS value was examined as follows.

For each participant, iterative computation was performed with the same data, but minus one performance that was chosen at random. The initial parameters were the same as those presented above.

Statistical analysis

All values are reported as the mean ± standard deviation (*s*). For all variables, the hypothesis of a normal distribution was tested with the Shapiro-Wilks *W*-test for small samples (performances) and the Kolmogorov test for the large samples (training loads) (Sen & Shrivastava, 1990; Wetherill *et al.*, 1986). The variations in performance and in the content between training phases were evaluated with a multiple paired *t*-test with Bonferonni correction. The training content was compared between training groups using a non-parametric Mann-Whitney *U*-test. Bartlett's test was used to verify homoscedasticity of performances. All statistical analyses were performed using Statistica 5.1 (Statsoft, Tulsa, OK, USA). Statistical significance was set at $P < 0.05$.

Results

Training characteristics and performances

For the group of swimmers as a whole, training volume measured during a season was 1675 ± 215 km. Over the entire study period, 13 ± 2 performances were recorded for each swimmer (see Table I). The mean coefficient of variation of the actual performances was 2.0 ± 0.4%. The sprint swimmers performed more I_1 training and less I_2 training than the intermediate-distance swimmers ($P < 0.05$). The middle-distance swimmers performed more I_4 training than the intermediate-distance swimmers ($P < 0.05$) (see Table II). Training volume equal to I_2 decreased between the volume and the intensity phase ($P < 0.05$), whereas training volume equal to I_4 and strength training (I_7) increased. Total metres swum, training volumes equal to I_1 and I_2 and strength training (I_6, I_7) decreased between the intensity and the taper phase ($P < 0.05$), whereas training volumes equal to I_3, I_4 and the maximal intensity (I_5) remained constant (see Table III).

Performances improved 2.24 ± 1.24% ($P < 0.05$) during the taper phase. The patterns of training load and performance during a season exhibited an undulating trend (Figure 1). The predicted peak performance was in weeks 61 and 62 – that is, 3 or 4 weeks after the end of the taper period and after complete cessation of training.

Performances estimation during the rest period

For the group as a whole, the time to peak performance was 17 ± 9 days (range 7 – 35 days) after the end of the training period.

Goodness of fit and accuracy

The relationship between training and performance (using the original estimation method) was significant for all participants: $R^2 = 0.79 \pm 0.13$, $P < 0.05$. The mean standard error for all participants was 1.05 ± 0.63%. The 95% CI of the parameters and the mean 95% CI width for the modelled performances are presented in Table IV. The 95% CI of the parameters t_n and t_g are presented in Table V. The mean coefficients of variation were 32 ± 14, 42 ± 16, 64 ± 22, 98 ± 32, 44 ± 22 and 41 ± 22% for τ_a, τ_f, k_a, k_f, t_n and t_g, respectively.

Ill-conditioning

Correlations between parameter estimates (mean absolute values ± standard deviation) are shown in

Table III. Respective contents of the volume, intensity and taper phases during the fourth training cycle (before the main competition of the year).

	Volume phase	Intensity phase	Taper phase
Total (m)	55 200 ± 8772	49 500 ± 7340**	28 900 ± 3780++
I_1 (m)	19 108 ± 3375	24 637 ± 2920	18 841 ± 2941+
I_2 (m)	33 800 ± 6229†	21 550 ± 5340**	7 000 ± 973++
I_3 (m)	1 333 ± 472	1 125 ± 382	1 366 ± 502
I_4 (m)	533 ± 227†	1 650 ± 443*	1 291 ± 483
I_5 (m)	425 ± 197	537 ± 176	478 ± 181
I_6 (min)	12 ± 4	8 ± 3**	3 ± 1++
I_7 (min)	15 ± 5†	28 ± 7*	12 ± 4++

Note: Intensities for swim workouts: below ($I_1 \approx 2\,\text{mmol}\cdot\text{l}^{-1}$), equal to ($I_2 \approx 4\,\text{mmol}\cdot\text{l}^{-1}$), slightly above ($I_3 \approx 6\,\text{mmol}\cdot\text{l}^{-1}$) the onset of blood lactate accumulation, respectively; $I_4 \approx 10\,\text{mmol}\cdot\text{l}^{-1}$; I_5 = maximal intensity; I_6 and I_7 consisted of 40 – 70% and 70 – 100% of one-repetition maximum strength training. †Significant differences between volume and intensity phases ($P < 0.05$). +, ++ Significant differences between volume and taper phases ($P < 0.05$ and $P < 0.01$, respectively). *,**Significant differences between intensity and taper phases ($P < 0.05$ and $P < 0.01$, respectively).

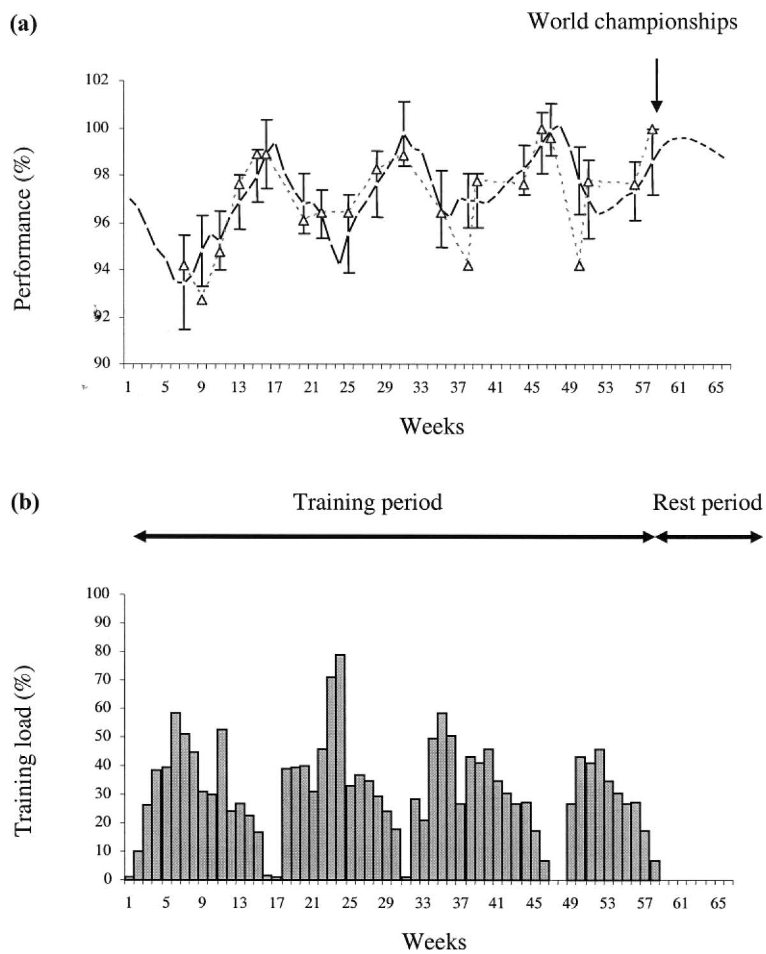

Figure 1. Example for participant 3 (Olympic finalist). (a) Modelled (line) and actual performances (dotted line with triangles). 95% confidence intervals for modelled performances are also presented. Performances on vertical axis were expressed as a percentage of the personal record $\min_t(P_t)$ and computed as $p_t = \frac{\min_t(P_t)}{P_t} * 100$. (b) Training loads on vertical axis are expressed as a percentage of the maximal training load performed by the participant during the course of the study. Time on horizontal axis is expressed in weeks.

Table VI. Parameters τ_a and τ_f were highly correlated (0.99 ± 0.01). A high correlation was also found between k_a and k_f (0.91 ± 0.13). Finally, the fatigue magnitude factor was correlated with the fitness and fatigue decay time constants (0.75 ± 0.30 and 0.76 ± 0.27, respectively).

Stability

The values of the Banister model parameters computed for each athlete with all performances minus one are shown in Table VII. No differences were found between the results listed in Table IV (all performances) and Table VII (all performances minus one).

Discussion

The main findings of the present study were as follows:

1. For the group as a whole, the mean time to peak performance was 17 days after the end of the training period and the main competition of the season.
2. Banister model parameters exhibited wide variability. For example, the mean 95% CI for t_n and t_g was (7, 35) and (25, 61) days, respectively. In addition, the coefficients of variation for the parameter estimates calculated from 1000 bootstrap replicates were greater than 30%. Conversely, the variability in modelled performances was quite small.
3. The Banister model appears to be stable, since no differences in the parameters were found when the computation was performed with one performance less.

In the present study, the characteristics of training and performances were similar to those reported previously (Avalos *et al.*, 2003; Mujika *et al.*, 1995,

Table IV. Banister model parameters measured over the one-year period studies, estimated by non-linear least squares.

	p_0	CI p_0	k_a	CI k_a	k_f	CI k_f	τ_a	CI τ_a	τ_f	CI τ_f	R^2	MIW
1	0.92	0.89, 0.95	0.002	−0.040, 0.044	0.016	−0.018, 0.051	40	8, 71	5	−6, 16	0.69*	3.17
2	0.95	0.92, 0.98	0.106	0.022, 0.189	0.129	0.071, 0.185	13	−6, 33	11	7, 15	0.84†	1.99
3	0.97	0.93, 1.01	0.039	0.006, 0.071	0.048	0.016, 0.081	33	8, 57	27	8, 46	0.65*	2.69
4	0.97	0.95, 0.98	0.050	−0.029, 0.068	0.068	0.048, 0.088	27	14, 40	20	15, 25	0.97§	0.80
5	0.98	0.95, 1.01	0.003	−0.040, 0.046	0.022	−0.012, 0.057	41	11, 70	9	−1, 19	0.78†	1.80
6	0.90	0.84, 0.94	0.009	−0.014, 0.034	0.016	−0.009, 0.041	45	18, 71	18	4, 32	0.61§	2.04
7	0.90	0.78, 1.02	0.018	−0.012, 0.046	0.028	−0.003, 0.057	57	39, 75	31	17, 46	0.95§	1.55
8	0.95	0.92, 0.97	0.083	0.019, 0.148	0.112	0.048, 0.176	23	11, 34	16	9, 24	0.92§	1.58
9	0.93	0.90, 0.96	0.010	−0.024, 0.046	0.012	−0.016, 0.042	65	50, 81	38	3, 64	0.73*	2.44
Mean	0.94	0.90, 0.98	0.036	−0.012, 0.077	0.050	0.014, 0.086	38	17, 59	19	6, 32	0.79	2.01
s	0.03		0.038		0.044		16		11		0.13	0.70

Note: p_0 = the initial performance level (as a percentage of the best performance); k_a and k_f = the fitness and fatigue magnitude factors, respectively (in arbitrary units); τ_a and τ_f = the fitness and fatigue decay time constants, respectively (in days). CI = Bootstrap 95% confidence interval of estimated parameters. If the confidence interval lower bound is negative, truncate the estimate and report the lower bound to zero. MIW = Mean interval width of 95% confidence interval of modelled performances. *$P < 0.05$; †$P < 0.01$; §$P < 0.001$.

Table V. Model parameters t_n and t_g computed from k_a and k_f (the fitness and fatigue magnitude factors) and τ_a and τ_f (the fitness and fatigue decay time constants), estimated by bootstrap resampling.

	t_n	CI t_n	t_g	CI t_g
1	7	−2, 15	15	−3, 34
2	10	1, 19	21	9, 32
3	31	25, 77	61	20, 96
4	25	21, 28	48	43, 53
5	24	6, 40	41	16, 65
6	12	5, 19	35	20, 50
7	29	8, 51	71	57, 85
8	17	12, 20	36	28, 41
9	15	−14, 45	64	34, 95
Mean	19	7, 35	43	25, 61
s	9		19	

Note: t_n = the time to recover performance, t_g = the time to peak performance after training completion. CI = Bootstrap 95% confidence interval of estimated parameters. If the confidence interval lower bound is negative, truncate the estimate and report the lower bound to zero.

Table VI. Correlation between parameter estimates for the Banister model (n = 9; mean absolute value $\pm$ s).

Parameters	Correlation
$p_0 - k_a$	0.49 ± 0.27
$p_0 - k_f$	0.31 ± 0.25
$p_0 - \tau_a$	0.41 ± 0.27
$p_0 - \tau_f$	0.41 ± 0.27
$k_a - k_f$	0.91 ± 0.13
$k_a - \tau_a$	0.69 ± 0.26
$k_a - \tau_f$	0.69 ± 0.26
$k_f - \tau_a$	0.75 ± 0.30
$k_f - \tau_f$	0.76 ± 0.27
$\tau_a - \tau_f$	0.99 ± 0.01

Note: p_0 = the initial performance level; k_a and k_f = the fitness and fatigue magnitude factors, respectively; τ_a and τ_f = the fitness and fatigue decay time constants, respectively.

1996a; Stewart & Hopkins, 2000). The mean positive and negative decay time constants ($\tau_a = 38 \pm 16$ days; $\tau_f = 19 \pm 11$ days) were close to those reported for endurance athletes (Banister *et al.*, 1999; Morton, 1997; Mujika *et al.*, 1996a). It is noteworthy that the parameter values and the determination coefficients are similar to those reported by Mujika *et al.* (1995, 1996a, 1996b) for similar athletes (international-standard swimmers) over the same length of time (one season). However, these values were also similar to those reported in studies with different types of athletes. For example, the time decay constant for the fitness function in the present study is similar to those (~40 days) calculated for less-fit athletes (Busso, Carasso, & Lacour, 1991; Morton *et al.*, 1990). It is therefore unlikely that the range of the parameter values would be similar, irrespective of the quality of the athlete. The mean standard errors for all participants were less than the difference in performance time between the first and the last swimmer in the Athens Olympic Games finals. However, they were less accurate than those reported by Avalos *et al.* (2003) who used a linear mixed model (1.05 ± 0.63% vs. 0.33 ± 0.11%).

Training and performance quantification

Other methods were applied to quantify the training load and performance. The total training load was also expressed as the summation of the volume at each training level weighted by a coefficient according to the relationship between swimming speed and lactate concentration (Mujika *et al.*, 1996a, 1996b). Performance was also modelled using a logarithm transformation, expressed as a percentage of the

Table VII. Banister model parameters measured over the one-year period studied with one performance, chosen at random, removed to assess model stability.

	p_0	k_a	k_f	τ_a	τ_f
1	0.93	0.002	0.077	39	2
2	0.95	0.023	0.054	17	7
3	0.98	0.032	0.040	33	26
4	0.97	0.035	0.054	28	19
5	0.98	0.003	0.021	41	10
6	0.90	0.053	0.067	47	26
7	0.90	0.020	0.030	55	32
8	0.94	0.079	0.106	23	16
9	0.92	0.011	0.012	62	36
Mean	0.94	0.029	0.051	38	19
s	0.03	0.025	0.029	15	12

Note: p_0 = the initial performance level; k_a and k_f = the fitness and fatigue magnitude factors, respectively; τ_a and τ_f = the fitness and fatigue decay time constants, respectively.

world record as well as the criterion points scale proposed by Morton *et al.* (1990). These different methods produced a less reliable adjustment and a change in the k_a parameter that expressed the sensitivity of the model to the scale used.

Performance prediction during the rest period

For the group as a whole, the modelled performance peaks occurred approximately 17 ± 9 days (range 7–35 days) after training ended. Indeed, these results are consistent with those of Morton (1991) who, using the original dose–response model of Calvert *et al.* (1976), studied by simulation the effects of variation in ten parameters (four from the dose–response model and six describing the differing shape of several seasonal training profiles). These authors reported that peak performance occurs a mean of 23 ± 34 days after the end of training. As the season ended with the most important competition of the year, these results are undesirable from a practical point of view: it would have been more logical for performance to peak during the key competition, and not some 17 days later. These results cannot be attributed to poor tapering. The 3 week taper phase was characterized by an approximate 50% decline in training volume and by a decrease in low-intensity and dry-land training, but maintenance of high-intensity training as recommended in the literature (Mujika & Padilla, 2003). In addition, performances improved 2.2 ± 1.2% during the taper phase, which is equivalent to the 2.2 ± 1.5% reported by Mujika, Padilla and Pyne (2002) during the final 3 weeks of training leading up to the Sydney Olympic Games and greater than the 1.5% and 2.1% reported by Bonifazi, Saldella and Luppo (2000) for male swimmers during two seasons. Moreover, the results of the present study are not in agreement with work carried out on short-term detraining (stopping training for 2–4 weeks), which generates a rapid decline in maximal oxygen uptake (Mujika & Padilla, 2000) and a decrease in swim power – that is, the ability to apply force during swimming (Neufer, Costill, Fielding, Flynn, & Kirwan, 1987). It is noteworthy, however, that these qualities have been shown to be strongly linked to swimming performance (Hawley *et al.*, 1992; Wakayoshi *et al.*, 1995). That changes in the parameters of the present model were not taken into account (time-unvarying model) could also explain why the modelled performance peaks occurred so long after training had ended. With a time-varying model, Busso *et al.* (2002) reported a decrease in t_n and t_g when training frequency was reduced. The time needed to recover performance after a training session increased from 0.9 ± 2.1 days at the end of low-frequency training to 3.6 ± 2.0 days at the end of high-frequency training. Busso (2003) used a non-linear model and introduced a variable to account for training-related changes in the magnitude and duration of exercise-induced fatigue. There was a decrease in time to peak performance when the training load was reduced from 37 days for a daily training load of 500 arbitrary units (a.u.) to 7 days for 300 a.u. However, in the latter two models, estimates must be provided for six parameters, implying the need for at least 15 performances per parameter (Stevens, 1986) – that is, 90 performances, which is wholly unworkable under real sporting conditions (Avalos *et al.*, 2003; Millet *et al.*, 2002; Mujika *et al.*, 1996a, 1996b).

Another problem concerns the method used to calculate the global training load, based on a summation of the different types of training. Training for elite swimmers comprises a wide range of diverse exercises (aerobic and anaerobic training, speed exercises, strength training). These types of training have to be individualized based on age, training background, individual profiles and speciality (Avalos *et al.*, 2003; Mujika *et al.*, 1996a; Stewart & Hopkins, 2000). For example, in the present study, sprint swimmers performed a larger proportion of speed training than middle-distance swimmers, who had a higher proportion of $\approx 4\ \text{mmol} \cdot \text{l}^{-1}$ training. The nature of the immediate and long-term training effects of these different exercises on the organism are so diverse that grouping them together or considering them as making up one single training stimulus would be unrealistic (Banister *et al.*, 1975). A similar overall training load could correspond to two very different types of training. (There is a compensation in training volume at each intensity.) Accordingly, Taha and Thomas (2003) argued that the Banister model implicitly assumes that the performance activity matches the training activity and therefore does not consider the specificity of training. Furthermore, classification into five intensities might be insufficient considering the large number of training variables used by coaches in their programmes, such as arm and leg training, training in the four styles and technical training (Avalos *et al.*, 2003; Mujika *et al.*, 1996b; Stewart & Hopkins, 2000). In addition, training for start and turn, which account for about 30% of event time (Thomson, Haljand, & McLaren, 2000), was not taken into consideration when calculating load for the Banister model. In fact, several authors have pointed out the multi-faceted nature of performance excellence (Avalos *et al.*, 2003; Banister *et al.*, 1975; Morton *et al.*, 1990; Mujika *et al.*, 1996a). Psychological, nutritional (Banister *et al.*, 1975; Morton *et al.*, 1990) and technical factors (Toussaint & Hollander, 1994; Wakayoshi *et al.*, 1995) also affect performance. It has been demonstrated that swimming efficiency, which cannot be quantified in overall load, is a major factor in swimming performance. Nevertheless, the impacts of the various types of training load on performance have an upper limit above which training does not elicit further adaptation of the athlete (Morton, 1997). Hence, the long-term and cumulative effects of training, which become clear some macro-cycles later, might not have been taken into account (Avalos *et al.*, 2003; Counsilman & Counsilman, 1991; Werchoschanski, 1992).

Variability in the parameters of the Banister model

The present study showed that the 95% CI in all parameters was large, especially for t_n and t_g [19 (7, 35) and 43 (25, 61) days, respectively]. The 95% CI for the modelled performances was reasonably small ($2001 \pm 0.70\%$), representing from a practical point of view ~ 1.2 s for a 100 m event completed in 55 s. The CI provides an interval of reasonable estimates, where the width of the interval is determined by the uncertainty in the point estimate. The greater the uncertainty, the wider the CI (Efron & Tibshirani, 1993; Huet *et al.*, 1996). In addition, the coefficients of variation for the parameter estimates were high (greater than 30%) (Arsac *et al.*, 2004). Parameter variability was much greater than the variability in time-dependent parameters reported by Busso *et al.* (1997) and Busso (2003), suggesting, unlike the study of Arsac *et al.*, that some methodological issues could be involved.

Study of the correlation matrix of parameter estimates revealed ill-conditioning problems, which are known to affect the accuracy of parameter estimates. Ill-conditioning and its effects are well known in linear regression (collinearity) (Sen & Shrivastava, 1990; Wetherill *et al.*, 1986). In non-linear regression, the problem is more complex, and different types of ill-conditioning can be identified (Bates & Watts, 1988; Belsley, 1991). Thus, in the linear case, the inaccuracies of parameter estimates and estimated responses are generally comparable, whereas in the non-linear case, inaccuracy of parameter estimates (estimator conditioning) and inaccuracy of response estimation (data conditioning) can differ (Belsley, 1991). In the present study, the 95% CI for parameters was large, whereas the 95% CI for the modelled performances was reasonably small.

The ill-conditioning in the present study can be explained by:

1. *A poor sample size*. The most direct and obvious means for improving conditioning is through the collection and use of additional data. Unfortunately, obtaining new data is usually not possible in studies of elite athletes, because of the few observations made. Furthermore, even if new data were obtainable, there is often no guarantee that they will be consistent with the original data or that they will indeed provide independent information (Belsley, 1991).
2. *Interactions between parameters*. If parameters are inter-dependent, their interpretation is conditioned and their practical use becomes meaningless. Several sets of parameters can be the (best) solution for a given load and performances data set.
3. *Misspecification of the model* (Bates & Watts, 1988; Huet *et al.*, 1996). Some hypotheses are supposed in the regression problem: normality, homoscedasticity and independence of errors (regarding the random structure), and the Banister model function (regarding the deterministic structure). If one of these hypotheses is false, the model will be misspecified. As a consequence of ill-conditioning, the estimation of the parameters t_n and t_g was inaccurate. This is supported by the findings of Fitz-Clarke *et al.* (1991), who reported that small (10–15%) changes in τ_f, k_a and k_f induced large variations in t_n and t_g, (41% and 21%, respectively). Thus for elite swimmers, the use of these parameters to provide valuable information for the understanding of individual responses to training and to develop individual training schedules from observational data appears to be hazardous.

Possible solutions to variability

Penalization techniques, based on decreasing variability to improve accuracy, are used extensively to resolve ill-conditioning problems in linear regression (Sen & Shrivastava, 1990; Wetherill *et al.*, 1986). Subset selection and ridge regression are the two main penalization procedures. The former selects a subset of the most relevant variables (Breiman, 1996; Sen & Shrivastava, 1990). Its adaptation to the present problem consists in selecting a subset of parameters. The studies of Busso *et al.* (1991) and Busso (2003) can be understood from a parameter subset selection viewpoint. Thus, the goodness of fit of models with different numbers of components (i.e. different numbers of parameters) was examined.

Ridge regression imposes a constraint on parameters (Breiman, 1996; Sen & Shrivastava, 1990; Wetherill *et al.*, 1986). To date, ridge regression has not been applied to sport data and only a few studies have dealt with the adaptation of ridge regression to non-linear regression (see, for example, Minor, Namini, & Watson, 1996; O'Sullivan & Saha, 1999; Zhou, Huang, Bergsneider, & Wong, 2002).

Short of new data, the introduction of appropriate prior information is another available solution to the ill-conditioning problem. Procedures to introduce prior information include, for example, mixed-estimation techniques (Belsley, 1991; Davidian & Giltinan, 1995). An application of mixed models to sport data was proposed by Avalos *et al.* (2003).

The re-parameterization of the Banister model can improve conditioning (Bates & Watts, 1988; Huet *et al.*, 1996). Since a high correlation between the fitness and fatigue decay time constants and between the fitness and fatigue magnitude factors was found, dissymmetrizing the fitness and fatigue functions could overcome the parameters' correlation. For example, one of the decay time constants and one of the magnitude factors could be inversed: $\upsilon_a = 1/\tau_a$, $l_a = 1/k_a$. These operations do not affect the interpretation of parameters (as initial parameters can be easily computed from the new ones); however, they can be numerically advantageous.

Misspecification problems are probably involved in the present study: (1) The Banister model does not take account of the possible dependence between performances, a typical problem in longitudinal studies (Avalos *et al.*, 2003). (2) In the Banister model, training impulses are proportional to the training loads; thus, higher loads induce more fitness and fatigue acquisition. But previous studies have reported that the impact of training loads on performance has an upper limit, above which training does not elicit further adaptation of the organism (Fry *et al.*, 1992; Morton, 1997). (3) The procedure assumes the parameters remain constant over time, an assumption that is not consistent with observed time-dependent alterations in responses to training (Avalos *et al.*, 2003; Busso, 2003; Busso *et al.*, 1997).

Stability

The Banister model estimated by the non-linear least squares method was stable, since no significant differences were found when the computation was performed with one performance less. The change in the temporal parameters was about 1 – 3 days for most of the participants. However, for participant 6 the difference in τ_f was 8 days. Moreover, the values of t_n and t_g are greatly modified by little change in τ_a or τ_f. For example, for participant 1, the values of t_n and t_g were decreased from 12 to 7 days and from 24 to 15 days, respectively, when τ_f decreased by 3 days; for participant 9, the values of t_n and t_g were increased from 7 to 17 days and 54 to 66 days, respectively, when τ_a decreased by 3 days and τ_f decreased by 2 days (estimated by non-linear least squares). These changes may have important practical implications in planning training loads.

Thus, proposing a single set of parameters might be dubious. Breiman (1996) showed that the averaged estimator of the bootstrap estimations is more stable than the original one. Stabilizing non-linear methods can give non-linear estimators with improved accuracy. This technique was used in the present study but a large variability was also observed in all parameters of the model and accuracy was not improved.

Perspectives for future study and use of the Banister model

Further studies should be conducted to determine whether the parameter estimation of the Banister model would be more accurate under standardized experimental conditions. Such standardized conditions would allow multiple recordings of performance and the use of a single type of exercise, as in the study of Busso (2003), where exercise was limited to cycle ergometry. The accuracy and the stability of the performance estimations observed in the present study suggest that the pattern of performance changes could be assessed as a function of training load. Nevertheless, the mean error of performance estimates was greater than reported by Avalos *et al.* (2003), suggesting that the results need to be validated experimentally by another complementary method.

Other procedures

Another possibility is to use non-parametric regression. Edelmann-Nusser, Hohmann and Henneberg (2002) suggested applying a non-parametric model (multi-layer perceptron neural networks) to model training load – performance relationships. These authors used an unconventional method that gave a surprisingly small prediction error (0.04%). Indeed, the model was fitted with data pertaining to one swimmer and then used to predict the performance of a different swimmer, despite the consensus that exists regarding the singularity of training responses (Avalos *et al.*, 2003; Busso, 2003; Busso *et al.*, 1997; Millet *et al.*, 2002; Mujika *et al.*, 1996a, 1996b).

Neural networks are particularly useful when the primary goal is outcome prediction, but these techniques are a "black box" and have a limited ability to explicitly identify possible causal relationships. Thus, the interpretation of the results obtained is not straightforward. However, in any given parametric problem, the parameters have meaningful interpretations (Hastie, Tibshirani, & Friedman, 2001).

Recently, Perl (2002) developed a dynamic meta model based on two antagonistic systems (two internal buffer potentials: one positive and one negative, which influence the performance potential alternately). This meta model seems conceptually very rich, because it takes into account the collapse effect in the wake of an overloaded training period, atrophy following a period of detraining, and the long-term behaviour of the training–performance relationship (Perl, Dausher, & Hawlitzky, 2001). However, to date, no statistical study has validated the quality of this meta model.

Finally, few authors have used multiple regression to create a model for the relationship between training and performance. Mujika *et al.* (1996b) used stepwise regression and reported a very close match with the Banister model. Multiple regressions make it possible to integrate training loads as independent variables and can take into account the effects of load sequences over time with short-term deferred effects during the 3 weeks that precede the performance (weeks 0, −1, −2), intermediate effects (weeks −3, −4 and −5) and long-term effects (weeks −6, −7, −8), as indicated by Avalos *et al.* (2003). The results we obtained with this method, using data from the present study, showed improved statistical accuracy in estimated parameters and modelled performances (more accurate CI). Moreover, in multiple regression, each training variable could be transformed for instance by a quadratic function (or higher-order function) to take into account a potential parabolic relationship between the quantity of training loads and performance (Sen & Shrivastava, 1990). It is also possible to take into account effects of interaction by associating the different input variables (Sen & Shrivastava, 1990).

Conclusion

The aim of this study was to assess whether the Banister model could be used to monitor the training process in the "real world" – that is, using the performances of international athletes. The present study assessed the goodness of fit, accuracy and stability of the Banister model as applied to training loads and performances in elite swimmers. The model showed substantial variability in its parameters, making it imprecise. To conclude, it would appear to be inappropriate to use these parameters to monitor the training process in elite swimmers. Nevertheless, the variability in modelled performances was reasonably small and the Banister model was stable. Further research should be conducted to determine whether associating these Banister model qualities with other methods of modelling could provide pertinent information to monitor training.

References

American College of Sports Medicine (2002). Progression models in resistance training for healthy adults. *Medicine and Science in Sports and Exercise*, *34*, 364–380.

Arsac, L. M., Thiaudière, E., Diolez, P., & Gerville-Réache, L. (2004). Parameter estimation in modeling phosphocreatine recovery in human skeletal muscle. *European Journal of Applied Physiology*, *91*, 419–424.

Atkinson, G., & Nevill, A. M. (1998). Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Medicine*, *26*, 217–238.

Avalos, M., Hellard, P., & Chatard, J. C. (2003). Modeling the training–performance relationship using a mixed model in elite swimmers. *Medicine and Science in Sports and Exercise*, *35*, 838–846.

Banister, E. W., Calvert, T. W., Savage, M. V., & Bach, A. (1975). A system model of training for athletic performance. *Australian Journal of Sports Medicine*, *7*, 170–176.

Banister, E. W., Carter, J. B., & Zarkadas, P. C. (1999). Training theory and taper: Validation in triathlon athletes. *European Journal of Applied Physiology*, *79*, 182–191.

Bartlett, R. (1997). The use and abuse of statistics in sport and exercise sciences. *Journal of Sports Sciences*, *15*, 1–2.

Bates, D. M., & Watts D. G. (1988). *Nonlinear regression analysis and its applications.* New York: Wiley.

Belsley, D. A. (1991). *Conditioning diagnostics: Collinearity and weak data in regression* (pp. 342–378). New York: Wiley.

Bonifazi, M., Saldella, F., & Luppo, C. (2000). Preparatory versus main competitions: Differences in performances, lactate responses and pre-competition plasma cortisol concentrations in elite male swimmers. *European Journal of Applied Physiology*, *82*, 368–373.

Breiman, L. (1996). Heuristics of instability and stabilization in model selection. *Annals of Statistics*, *24*, 2350–2383.

Busso, T. (2003). Variable dose–response relationship between exercise training and performance. *Medicine and Science in Sports and Exercise*, *35*, 1188–1195.

Busso, T., Benoit, H., Bonnefoy, R., Feasson, L., & Lacour, J. R. (2002). Effects of training frequency on the dynamics of performance response to a single training bout. *Journal of Applied Physiology*, *92*, 572–580.

Busso, T., Candau, R., & Lacour, J. R. (1994). Fatigue and fitness modelled from the effects of training on performance. *European Journal of Applied Physiology*, *69*, 50–54.

Busso, T., Carasso, C., & Lacour, J. R. (1991). Adequacy of a systems structure in the modeling of training effects on performance. *Journal of Applied Physiology*, *71*, 2044–2049.

Busso, T., Denis, C., Bonnefoy, R., Geyssant, A., & Lacour, J. R. (1997). Modeling of adaptations to physical-training by using a recursive least-squares algorithm. *Journal of Applied Physiology*, *82*, 1685–1693.

Busso, T., Hakkinen, K., Pakarinen, A., Carasso, C., Lacour, J. R., Komi, P. V. *et al*. (1990). A systems model of training responses and its relationship to hormonal responses in elite weight-lifters. *European Journal of Applied Physiology*, *61*, 48–54.

Calvert, T. W., Banister, E. W., & Savage, M. V. (1976). A systems model of the effects of training on physical performance. *SMC Systems, Man and Cybernetics*, *2*, 94–102.

Counsilman, B. E., & Counsilman, J. E. (1991). The residual effects of training. *Journal of Swimming Research*, *7*, 5–12.

Davidian, M., & Giltinan, D. M. (1995). *Nonlinear models for repeated measurement data*. London: Chapman & Hall.

Edelmann-Nusser, J., Hohmann, A., & Henneberg, B. (2002). Modeling and prediction of competitive performance in swimming upon neural networks. *European Journal of Sport Science*, *2*, 1–12.

Efron, B., & Tibshirani, R. (1993). *An introduction to the bootstrap*. Monographs on Statistics and Applied Probability. New York: Chapman & Hall.

Fitz-Clarke, J. R., Morton, R. H., & Banister, E. W. (1991). Optimizing athletic performance by influence curves. *Journal of Applied Physiology*, *71*, 1151–1158.

Fry, R. W., Morton, A. R., & Keast, D. (1992). Periodisation of training stress: A review. *Canadian Journal of Sport Sciences*, *17*, 234–240.

Hastie, T., Tibshirani, R., & Friedman, J. (2001). *The elements of statistical learning: Data mining, inference and prediction*. Springer Series in Statistics. New York: Springer.

Hawley, J. A., Williams, M. M., Vickovic, M. M., & Handcock, P. S. (1992). Muscle power predicts freestyle swimming performance. *British Journal of Sports Medicine*, 26, 151–155.

Huet, S., Bouvier, A., & Gruet, M. A., & Jolivet, E. (1996). *Statistical tools for nonlinear regression: A practical guide with S-PLUS examples*. New York: Springer.

Millet, G. P., Candau, R. B., Barbier, B., Busso, T., Rouillon, J. D., & Chatard, J. C. (2002). Modelling the transfers of training effects on performance in elite triathletes. *International Journal of Sports Medicine*, *23*, 55–63.

Minor, J. M., Namini, H., & Watson, G. A. (1996). Generalized ridge analysis with application to population pharmacokinetics/dynamics. *Journal of Biopharmaceutical Statistics*, *6*, 105–114.

Morton, R. H. (1991). The quantitative periodization of athletic training: A model study. *Sports Medicine Training and Rehabilitation*, *3*, 19–28.

Morton, R. H. (1997). Modelling training and overtraining. *Journal of Sports Sciences*, *15*, 335–340.

Morton, R. H., Fitz-Clarke, J. R., & Banister, E. W. (1990). Modelling human performance in running. *Journal of Applied Physiology*, *90*, 1171–1177.

Mujika, I., Busso, T., Lacoste, L., Barale, F., Geyssant, A., & Chatard, J. C. (1996a). Modelled responses to training and taper in competitive swimmers. *Medicine and Science in Sports and Exercise*, *28*, 251–258.

Mujika, I., Chatard, J. C., Busso, T., & Geyssant, A. (1996b). Use of swim-training profiles and performance data to enhance training effectiveness. *Journal of Swimming Research*, *11*, 23–29.

Mujika, I., Chatard, J. C., Busso, T., Geyssant, A., Barale, F., & Lacoste, L. (1995). Effects of training on performance in competitive swimming. *Canadian Journal of Applied Physiology*, *20*, 395–406.

Mujika, I., & Padilla, S. (2000). Detraining: Loss of training-induced physiological and performance adaptations. Part I: Short term insufficient training stimulus. *Sports Medicine*, *30*, 79–87.

Mujika, I., & Padilla, S. (2003). Scientific bases for precompetition tapering strategies. *Medicine and Sciences in Sports and Exercise*, *35*, 1182–1187.

Mujika, I., Padilla, S., & Pyne, D. (2002). Swimming performance changes during the final 3 weeks of training leading to the Sydney 2000 Olympic Games. *International Journal of Sports Medicine*, *23*, 582–587.

Neufer, P. D., Costill, D. L., Fielding, R. A., Flynn, M. G., & Kirwan, J. P. (1987). Effect of reduced training on muscular strength and endurance in competitive swimmers. *Medicine and Science in Sports and Exercise*, *19*, 486–490.

O'Sullivan, F., & Saha, A. (1999). Use of ridge regression for improved estimation of kinetic constants from PET data. *IEEE Transactions on Medical Imaging*, *18*, 115–125.

Perl, J., Dausher, P., & Hawlitzky, M. (2001). On the long term behaviour of the performance potential-metamodel Perpot. *International Journal of Computer Science in Sport*, *1*, 56–57.

Perl, J. (2002). Adaptation, antagonism and system dynamics. In G. Ghent, D. Kluka, & D. Jones (Eds.), *The multidisciplinary series of physical education and sport science* (Vol. 4, pp. 105–125). Oxford: Meyer & Meyer Sport.

Sen, A., & Shrivastava, M. (1990). *Regression analysis: Theory, methods, and applications.* New York: Springer.

Stevens, J. (1986). *Applied multivariate statistics for the social sciences.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Stewart, A. M., & Hopkins, W. G. (2000). Seasonal training and performance of competitive swimmers. *Journal of Sports Sciences*, *18*, 873–884.

Taha, T., & Thomas, S. G. (2003). Systems modeling of the relationship between training and performance. *Sports Medicine*, *33*, 1061–1073.

Thomson, K. G., Haljand, R., & McLaren, D. P. (2000). An analysis of selected kinematic variables in national and elite male and female 100-m and 200-m breastroke swimmers. *Journal of Sports Sciences*, *18*, 421–431.

Toussaint, H. M., & Hollander, A. P. (1994). Energetics of competitive swimming: Implications for training-programs. *Sports Medicine*, *18*, 384–405.

Wakayoshi, K., D'Acquisto, L. J., Cappaert, J. M., & Troup, J. P. (1995). Relationship between oxygen uptake, stroke rate and swimming velocity in competitive swimming. *International Journal of Sports Medicine*, *16*, 19–23.

Werchoschanski, J. W. (1992). *Efficient training (L'entraînement efficace)*. Paris: Presses Universitaires de France.

Wetherill, G. B., Duncombe, P., & Kenward, M. (1986). *Regression analysis with applications.* Monographs on Statistics and applied probability. London: Chapman & Hall.

Wolfarth, B., Rivera, M. A., Oppert, J. M., Boulay, M. R., Dionne, F. T., Chagnon, M. *et al*. (2000). A polymorphism in the alpha 2a-adrenoceptor gene and endurance athlete status. *Medicine and Sciences in Sports and Exercise*, *32*, 1709–1712.

Zhou, Y., Huang, S. C., Bergsneider, M., & Wong, D. F. (2002). Improved parametric image generation using spatial-temporal analysis of dynamic PET studies. *NeuroImage*, *15*, 697–707.