Modeling the Training-Performance Relationship Using a Mixed Model in Elite Swimmers

MARTA AVALOS1, PHILIPPE HELLARD2, and JEAN-CLAUDE CHATARD3

1HEUDIASYC Laboratory, UMR CNRS 6599, Compiègne University of Technology, FRANCE; 2Research Department, French Swimming Federation, FRANCE; and 3Laboratory of Physiology, GIP Exercise, Faculty of Medicine, Saint-Etienne University, FRANCE

ABSTRACT

AVALOS, M., P. HELLARD, and J.-C. CHATARD. Modeling the Training-Performance Relationship Using a Mixed Model in Elite Swimmers. Med. Sci. Sports Exerc., Vol. 35, No. 5, pp. 838–846, 2003. Purpose: The aim of this study was to model the relationship between training and performance in 13 competitive swimmers, over three seasons, and to identify individual and group responses to training. Methods: A linear mixed model was used as an alternative to the Banister model. Training effect on performance was studied over three training periods: short-term, the average of training load accomplished during the 2 wk preceding each performance of the studied period; mid-term, the average of training load accomplished during weeks 3, 4, and 5 before each performance; and long-term, weeks 6, 7, and 8. Results: Cluster analysis identified four groups of subjects according to their reactions to training. The first group corresponded to the subjects who responded well to the long-term training period, the second group to the long- and mid-term periods, the third to the short- and mid-term periods, and the fourth to the combined periods. In the model, the intersubject differences and the evolution over the three seasons were statistically significant for the identified groups of swimmers. Influence of short-term training was negative on performance in the four groups, whereas mid- and long-term training had, on the average, a positive effect in three groups out of four. Between seasons 1 and 3, the effect of mid-term training declined, whereas the effect of long-term training increased. The fit between real and modeled performances was significant for all swimmers (0.15 ≤ r² ≤ 0.65; F ≤ 0.01). Conclusion: The mixed model described a significant relationship between training and performance both for individuals and for groups of swimmers. This relationship was different over the 3 yr. Personalized training schedules could be prescribed on the basis of the model results. Key Words: COACHING, EXERCISE, INTENSITY, MATHEMATICAL MODEL.

The training-performance relationship is particularly important for elite sports coaches who search for reproducible phenomena useful for organizing the athlete’s training program. Many authors have studied the relative influence of training (7,22,23,27) and found that reactions to training depend on volume, intensity, and frequency of the training sessions (7,16,23). Others have reported divergent results (4,9), perhaps related to the fact that delayed effects and interindividual differences were not taken into account.

For individual swimmers, mathematical models have been developed to describe the dynamic aspect of training and the consequences of succession of training loads over time (2). The Banister model (2–4) and its modifications (5,6) are based on two antagonistic functions, both calculated from the training impulse. Studies on cellular adaptability reactions to exercise (3) have demonstrated that the negative function can be assimilated to a fatiguing impulse. The positive function can be compared with a fitness impulse resulting from the organism’s adaptation to training. Expressed as an exponential, the functions account for the decreasing impact of the training effect. When iterative training sessions are considered, the time course of performance is described by:

\[
p_t = p_0 + k_a \sum_{i=0}^{n-1} e^{-i\tau_a}w_s - k_f \sum_{i=0}^{n-1} e^{-i\tau_f}w_s \quad (1)
\]

where \(p_t\) is the known performance at week (or day) \(t\); \(w_s\) is the known training load per week (or day) from the first week of training to the week (or day) preceding performance \(p_t\); \(k_a\) and \(k_f\) are the fitness and fatigue multiplying factors, respectively; \(\tau_a\) and \(\tau_f\) are the fitness and fatigue decay time constants, respectively; and \(p_0\) corresponds to an initial basic level of performance.

There is no clear consensus on just how many data points are needed per parameter to ensure a stable solution in a regression analysis. Proposals reported in the literature have ranged from 5 to 50. Stevens (26) recommends a nominal number of 15 observations per parameter (except the inter-
except parameter) for a multiple linear regression. But as the Banister model is a nonlinear model, inference is based on asymptotic theory (10), which implies more data points per parameter than for a linear regression model. This means a large number of observations would be required to obtain precise results and enable pertinent statistical analysis. The Banister model also assumes the parameters remain constant over time, an assumption that is not consistent with observed time-dependent alterations in response to training (3,4,6,22).

When few repeated measurements are available for several subjects, mixed models provide an attractive solution (29). Instead of constructing a personal model for each subject, a model of popular behavior is constructed, allowing parameters to vary from one individual to another, to take into account the heterogeneity between subjects. Particular care in characterizing random variation in the data is required to recognize two levels of variability: random variation among measurements within a given individual (intraindividual variation) and random variation among individuals (interindividual variation) (10). In addition, mixed models analyze responses corresponding to different dose inputs (10), a common situation in swimming as training loads differ with age, specialty, and/or competition level (7,23).

The aim of the present study was to investigate the effect of training on performance of 13 elite swimmers taking into account a) individual profiles, b) subpopulation profiles, and c) time effect over three seasons.

METHODS

Subjects. Thirteen national and international level French swimmers were studied (6 females, 7 males). Their mean (±SD) age, height, and weight were 22 ± 3 yr, 177 ± 7 cm, and 66 ± 10 kg, respectively. Other population characteristics can be found in Table 1. Written informed consent was obtained from the subjects before entering the study. Approval for the project was obtained from the Saint-Etienne University Committee on Human Research. Each swimmer trained according to the program prescribed by the two-team coaches independently of the authors of the present study. Performances were measured during real competitions and expressed as a percentage of the personal record performed by each swimmer during the studied period. The total number of performances was 302, and the mean number of performances per season was 9.5 ± 2.4 (over a total of 13 subjects and 3 seasons). \( P_{ij} \) designated the \( j \)th percentage for the \( i \)th individual, \( j = 1, \ldots, n_i; i = 1, \ldots, N \); \( n_i \) the number of performances for the \( i \)th individual; and \( N = 13 \), number of subjects in the study. A percentage increment indicated a faster performance. Inversely, when percentage decreased, the performance was lower.

Quantification of the training load. For workouts in water, intensity levels were evaluated using the method proposed by Mujika et al. (22). A progressive test was performed at the beginning of the first season to determine blood lactate concentration. During the test, each swimmer performed 200-m swims at a progressively increased percentage of his/her own best competition time in that distance, until exhaustion (failure to follow the required pace). Blood lactate concentration was determined from blood samples taken from a fingertip during the 1-min recovery periods separating the 200-m swims. According to the individual results obtained during this test, all the training performed in water was divided into five intensity levels. Intensities Z1, Z2, and Z3 represented swimming speeds between 70% and 100% of the maximal strength measured (29). Two dry-land intensity levels were defined. Maximal strength exercises, Z6, consisted of exercises performed between 70% and 100% of the maximal strength measured for each exercise. Endurance exercises, Z7, included exercises between 40% and 70% and/or exercises performed with a number of repetitions greater than 20. Dry-land workout was quantified in minutes per week (min-wk\(^{-1}\)). Weekly volumes for each intensity level of every two groups where compared using Friedman analysis of variance (ANOVA).

For each swimmer, each intensity level was then expressed as a percentage of the maximal intensity recorded during the study. This normalization allowed training loads of different units or intensities to be compared using the same scale of values. The global weekly training load, \( w \), was the mean of normalized weekly training intensities.

In swimming, training time and quantity are considerably different for different training regimes. In our study, the quantitative relationship between Z5 training and Z1 training was, approximately, 1-fold to 40-fold. Thus, if the data are not normalized zone by zone, training sessions with a lower quantitative level would be overpowered by training.

### TABLE 1. Population characteristics.

<table>
<thead>
<tr>
<th>Subjects</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
<th>S4</th>
<th>S5</th>
<th>S6</th>
<th>S7</th>
<th>S8</th>
<th>S9</th>
<th>S10</th>
<th>S11</th>
<th>S12</th>
<th>S13</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>F</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td>F</td>
<td>M</td>
<td>M</td>
<td>F</td>
<td>F</td>
<td>M</td>
<td>F</td>
<td>M</td>
<td>F</td>
</tr>
<tr>
<td>Distance</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>400</td>
<td>400</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>200</td>
</tr>
<tr>
<td>Style</td>
<td>FS</td>
<td>ME</td>
<td>FS</td>
<td>FS</td>
<td>FS</td>
<td>BR</td>
<td>BR</td>
<td>FS</td>
<td>BR</td>
<td>BR</td>
<td>BU</td>
<td>BR</td>
<td>BU</td>
</tr>
</tbody>
</table>

Sex: M, male; F, female.
Distance values are expressed in meters.
Style: FS, freestyle; ME, medley; BR, breaststroke; BU, butterfly.
sessions with a higher quantitative level. Other options are possible to consider this fact: Mujika et al. (22) condensed all variables for the different training loads to a more comparable level by multiplying them by their coefficients of energetic intensity. Because it was decided not to make any preliminary hypotheses concerning the impact of training charges, this mode of normalization was preferred.

**Definition of variables.** The total training load of the 8 wk preceding the performance was taken into account: \( w_0, w_1, w_2, w_3, w_4, w_5, w_6, \) and \( w_7 \). Three variables were defined: \( X_1 \), short-term training, was the average of \( w_0 \) and \( w_1 \); \( X_2 \), mid-term training, was the average of \( w_2 \) and \( w_3 \); and \( X_3 \), long-term training, was the average of \( w_4, w_5, \) and \( w_7 \). Variables were denoted \( X_{ij} \), \( X_{2ij} \), and \( X_{3ij} \), to indicate the \( j \)th observation for the \( i \)th subject of aforementioned variables, \( j = 1, \ldots, n_i; i = 1, \ldots, N \).

**Definition of covariates.** Two covariates were chosen to take into account 1) the evolution over the three seasons of the relationship between training and performance and 2) some reactions to training. The covariate season indicated whether the observation belonged to the first, second, or third season. It was parameterized by two indicator variables: \( A_{1ij} \) and \( A_{2ij} \). They were defined to be “1” if the observation \( j \) of subject \( i \) belonged to the first season or the second season, respectively, and “0” otherwise.

The covariate “group” stemmed from a classification study. First, correlation values of short-, mid- and long-term training periods with performance were used to achieve a principal component analysis (PCA) and, then, investigate distances between subjects. Four kinds of reactions to training classes appeared accounting for the first and the second principal component factors. Second, the hypothesis of four well-distinguished classes were tested and significantly accepted \( (P < 0.05) \) from a cluster analysis according to the \( k \)-means method. The four groups were interpreted (relating each group to the others, not in absolute meaning) as follows: 1) poor responders to the short- or mid-term training and neutral to the long-term training: PN \( (S6, S7, S10) \); 2) good responders to the short- or mid-term training as well as to the long-term training: GG \( (S4, S5, S13) \); 3) good responders to the short- or mid-term training and neutral to the long-term training: GN \( (S9, S11) \); and 4) neutral responders to the short- or mid-term training and good to the long-term training: NG \( (S1, S2, S3, S8, S12) \). “Neutral response” expressed a dispersed reaction into the group or a slight effect of the factor. This covariate was parameterized by three indicator variables. \( G_{1i} \), \( G_{2i} \), and \( G_{3i} \) are defined to be “1” if the subject \( i \) belonged to PN, GG, or GN, respectively, and “0” otherwise.

**The mixed model.** A linear mixed-effects model is any model that satisfies (29): \( Y_i = X_i \beta + Z_i b_i + \epsilon_i \) with \( b_i \sim N(0, D) \), \( \epsilon_i \sim N(0, \Sigma) \) and \( b_i \)’s independent of \( \epsilon_i \)’s.

\( Y_i \) represents the performances of subject \( i \). \( X_i \beta \) is the common to the population part of the model. \( Z_i b_i \) is the specific to each subject part of the model, and \( \epsilon_i \) is the random sampling error of the model. \( \beta \) (the fixed effects) represent the population regression coefficients that are applied to \( X_i \), the whole set of variables and covariates (training variables, and season and group covariates) of subject \( i \). \( b_i \) (the random effects) represents the personal regression coefficients for subject \( i \) that are applied to \( Z_i \), the subset of variables and covariates that present a specific-subject profile. \( \epsilon_i \) can be thought of as sampling error, or random perturbations; it is also taken to be normally distributed with an autoregressive covariance matrix, \( \Sigma \). This covariance structure supposes that two observations taken close in time within an individual are more closely correlated than two observations taken far apart in the same individual. \( D \) models the between-subject variance, whereas \( \Sigma \) models the within-subject variance.

Summarizing, the main ideas and hypotheses of mixed-effects modeling that are used in this study are:

i. Mixed effect models have two levels: one level corresponding to the popular behavior and a second level corresponding to the individual behavior. All data are used to construct the common to the population part of the model, but only the observations specific to each individual are used to construct the personal part of the model. The contribution of population information is higher when the individual information is poor, and inversely.

ii. Two levels of variability corresponding to the popular and the individual levels are characterized: the intraindividual and the interindividual variability.

iii. Intraindividual modeling takes into account the closeness in time of performances.

Mathematical descriptions of the particular linear mixed model explaining the relationship between training and performance can be found in the appendix.

**Fitting the model.** For fitting the model and testing suitable hypotheses, the SAS 8.1 MIXED procedure was employed. Estimates and SE for all fixed effects and all variance components were computed. Parameters were calculated from the restricted maximum likelihood estimates, to take into account the loss of the degrees of freedom involved in estimating the fixed effects (29).

Hypotheses about the mean structure had to be tested. These hypotheses concerned, first, the influence of training variables, covariate of evolution through seasons and covariate of reaction to training and, second, differences between every two classes of covariates “season” and “group.” Pertinence of hypotheses about the variance structure had to be tested too. These hypotheses concerned, first, the shape of matrix \( D \) and \( \Sigma \) and, lastly, the assumption of variability in all subject-specific intercepts and slopes.

For each hypothesis, an appropriate statistical test was available in the MIXED procedure (29). For the assumptions of the mean structure, an approximate \( F \)-test, using the containment method to compute degrees of freedom and the sandwich estimate of variance, to ensure robustness, were used. Results were used to reduce the original model to a more parsimonious model. This was done in a hierarchical
way, starting with highest $P$-value term and comparing the model with and without the term. If the presence of the term did not imply a significant improvement in the model, it was deleted. The obtained final model assumed no effect of season on basic level nor on near training.

With regard to variance structure, a likelihood ratio test for nonstandard testing situations was used for testing the need of random effects. The presence of random effects was clearly necessary. Finally, Akaike information criteria and Schwarz information criteria discriminated between different nonnested models of structure of variance. The greatest amount of information was given by models with an autoregressive error covariance. This can be explained because, in addition to seasonal evolution, a continuous evolution through time, caused by a progressive adaptation to training, is likely.

RESULTS

Training volume and intensity. Global training load changed through the studied period according to a wave pattern of four macrocycles in each season training plan (Fig. 1). Each macrocycle consisted of a high-load phase, whereby total training load was gradually increased alternating with regeneration phases (weeks 16, 37, 69, 97, 122) and competition periods: national events (weeks 25, 82, 140), European events (week 50) and Olympic games (week 108).

PN was the group with the highest Z3, Z4, Z6, and Z7 volumes. GG was the group with the highest Z2 volume, and the lowest Z4 and Z5 volumes. GN was the group with the lowest dry-land intensity-level volumes. Finally, NG was the group with the highest Z1 volume and the second highest Z2, Z3, Z4, Z6, and Z7 volumes (Table 2). Weekly Z2 and Z4 volumes were not significantly different between groups. For the other intensity levels, there were some significant differences between some groups.

Cluster analysis and PCA. PCA results are shown in Table 3. Only factors with eigenvalues greater than 1 were retained. Factor 1 accounted for 56% of the total variance. This factor was highly negatively related to the short- and mid-term training period ($r = -0.91$; $r = -0.92$, respectively). Factor 1 is represented in Figure 2 by the horizontal axis. Subjects S6, S7, and S10, on the right, for whom $X_1$ and $X_2$ exerted a negative influence, were separated from subjects S9 and S11, on the left, for whom $X_1$ and $X_2$ exerted a positive influence.

Factor 2 accounted for 37% of the total variance. This factor was highly positively related to long training period ($r = 0.99$). Factor 2 is represented in Figure 2 by the vertical axis. Subjects S4, S5, and S13, at the top, for whom $X_3$ exerted a positive influence, were separated from subjects S9 and S11, at the bottom, for whom $X_3$ exerted a negative influence. For subjects S1, S2, S3, S8, and S12, situated in

<table>
<thead>
<tr>
<th>Intensity Level</th>
<th>PN</th>
<th>GG</th>
<th>NG</th>
<th>PN</th>
<th>GG</th>
<th>NG</th>
<th>PN</th>
<th>GG</th>
<th>NG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z1 (km-wk$^{-1}$)</td>
<td>13543 (7529)</td>
<td>15785 (9477)</td>
<td>15725 (7624)</td>
<td>16201 (8407)</td>
<td>15725 (7624)</td>
<td>16201 (8407)</td>
<td>15725 (7624)</td>
<td>16201 (8407)</td>
<td>16201 (8407)</td>
</tr>
<tr>
<td>Z2 (km-wk$^{-1}$)</td>
<td>17488 (13120)</td>
<td>19018 (16447)</td>
<td>17262 (11857)</td>
<td>17847 (13748)</td>
<td>17262 (11857)</td>
<td>17847 (13748)</td>
<td>17262 (11857)</td>
<td>17847 (13748)</td>
<td>17847 (13748)</td>
</tr>
<tr>
<td>Z3 (km-wk$^{-1}$)</td>
<td>3065 (2507)</td>
<td>2249 (2130)</td>
<td>2249 (2130)</td>
<td>2033 (2061)</td>
<td>2033 (2061)</td>
<td>2033 (2061)</td>
<td>2033 (2061)</td>
<td>2033 (2061)</td>
<td>2033 (2061)</td>
</tr>
<tr>
<td>Z4 (km-wk$^{-1}$)</td>
<td>663 (541)</td>
<td>572 (652)</td>
<td>572 (652)</td>
<td>586 (492)</td>
<td>586 (492)</td>
<td>586 (492)</td>
<td>586 (492)</td>
<td>586 (492)</td>
<td>586 (492)</td>
</tr>
<tr>
<td>Z6 (min-wk$^{-1}$)</td>
<td>4.8 (14.2)</td>
<td>2.5 (13.2)</td>
<td>2.5 (13.2)</td>
<td>1.4 (12)</td>
<td>1.4 (12)</td>
<td>1.4 (12)</td>
<td>1.4 (12)</td>
<td>1.4 (12)</td>
<td>1.4 (12)</td>
</tr>
<tr>
<td>Z7 (min-wk$^{-1}$)</td>
<td>6.4 (15.4)</td>
<td>2.13 (7.5)</td>
<td>2.13 (7.5)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
</tbody>
</table>

Values are mean (SD).

* Significant difference of the first line group from the second line group, for each intensity level ($P \leq 0.05$).

** Significantly different ($P \leq 0.01$).

--- Not significantly different.

FIGURE 1—Global weekly training load (the mean of the seven intensity levels expressed in percentage of their maxima) for all the subjects, over the three seasons. Values are means (SE). Disks in horizontal axis indicate main competitions. Over the three seasons, the 100% value was not attempted: there is no week in which the 100% value was attempted for all intensity levels and for all subjects.
the middle of the figure, X₃ exerted a positive influence and X₁ and X₂ exerted a neutral influence. Cluster analysis divided the swimmers into 4 statistically different groups.

**Fitted model analysis.** After testing, the model was modified and reduced (mathematical descriptions can be found in the appendix). Random effects reflect how much the subject-specific profiles deviates from the overall average profile. In the initial model, four-dimensional random effects were taken into account: the first element was associated with basic level, and the second, third, and fourth elements were associated with training variables X₁, X₂, and X₃, respectively. Nevertheless, tests showed that, even though every combination of two random effects was significant, more than two effects were not necessary. Thus, short-, mid-, and long-term training varied from a subject to another, but this variation was mainly provided by two variables. A supplementary variable did not contribute significantly to describe the subject-specific profiles. So, to ensure a parsimonious model, only random effects providing the most of the information were kept in the final model: random effect associated with basic level and to long-term training. This could mean that groups were homogeneous enough to account for a great part of the behavior of each swimmer. Only little complementary information was needed to describe individual profiles.

From parameter estimates of the mean structure, it was possible to compute each group’s parameter, for each variable X₁, X₂, and X₃, and the basic level. Also, parameters assigned to each of the three seasons and to their average were computed, when effect of the season was significant. Results are given in Table 4. The higher basic level, 98.7%, belonged to the group PN, and the lower, 95.4%, belonged to GG. Short-term training was negative on performance for all the groups (PN: -1.02; GG: -0.74; GN: -0.29; NG: -1.04). Mid-term training was negative for PN and positive for the others (PN: -0.11; GG: 0.63; GN: 0.93; NG: 0.41). Finally, long-term training was negative for GN and positive for the others (PN: 0.60; GG: 0.69; GN: 0.05; NG: 0.54).

The effect of the seasons was not significant for short-term training, whereas for mid- and long-term training, it was statistically different. For mid-term training, the effect of seasons was significantly negatively decreasing on PN (-0.11, -0.16, and -0.78, for the first, second, and third seasons, respectively). For GG and GN, the effect was significantly positively decreasing and for NG, the effect was significantly decreasing from positive to negative values (GG: 0.87, 0.82, 0.20; GN: 1.17, 1.12, 0.49; and NG: 0.74, 0.74, 0.29).

![FIGURE 2 — Geometric representation of subjects (from S1 to S13) from the PCA. Factor 1 was negatively related to correlation between the short- and mid-term training periods with performance. Factor 2 was positively related to correlation between the long training period and performance. Groups: poor responders to the short- or mid-term training and neutral to the long-term training (PN); good responders to the short-, mid-, and long-term training (GG); good responders to the short- or mid-term training and poor to the long-term training (GN); and neutral responders to the short- or mid-term training and good to the long-term training (NG).](http://www.acsm-msse.org)

**TABLE 3. Results of PCA. Weights of principal factors and explained variance.**

<table>
<thead>
<tr>
<th>Factor</th>
<th>( p_1 )</th>
<th>( p_2 )</th>
<th>( p_3 )</th>
<th>EV</th>
<th>%EV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor 1</td>
<td>-0.91</td>
<td>-0.92</td>
<td>-0.02</td>
<td>1.68</td>
<td>0.56</td>
</tr>
<tr>
<td>Factor 2</td>
<td>-0.28</td>
<td>0.25</td>
<td>-0.99</td>
<td>1.12</td>
<td>0.37</td>
</tr>
</tbody>
</table>

\( p_i \) correlation between \( X_i \) and performance, \( i = 1, 2, 3. \)

EV, explained variance.

* Significant value (\( P \leq 0.05 \)).

**TABLE 4. Weights for basic level, \( X_1 \), \( X_2 \), and \( X_3 \), for each group and season.**

<table>
<thead>
<tr>
<th>Group</th>
<th>PN</th>
<th>GG</th>
<th>GN</th>
<th>NG</th>
</tr>
</thead>
<tbody>
<tr>
<td>BL</td>
<td>—</td>
<td>98.70</td>
<td>95.40</td>
<td>97.31</td>
</tr>
<tr>
<td>( X_1 )</td>
<td>—</td>
<td>-1.02</td>
<td>-0.74</td>
<td>-0.29</td>
</tr>
<tr>
<td>( X_2 )</td>
<td>—</td>
<td>0.11</td>
<td>0.87</td>
<td>1.17</td>
</tr>
<tr>
<td>1st</td>
<td>-0.16</td>
<td>0.82</td>
<td>1.12</td>
<td>0.61</td>
</tr>
<tr>
<td>2nd</td>
<td>-0.78</td>
<td>0.20</td>
<td>0.49</td>
<td>-0.02</td>
</tr>
<tr>
<td>3rd</td>
<td>-0.35</td>
<td>0.63</td>
<td>0.93</td>
<td>0.41</td>
</tr>
<tr>
<td>AV</td>
<td>0.17</td>
<td>0.26</td>
<td>-0.48</td>
<td>0.12</td>
</tr>
<tr>
<td>1st</td>
<td>0.59</td>
<td>0.68</td>
<td>-0.07</td>
<td>0.53</td>
</tr>
<tr>
<td>2nd</td>
<td>1.04</td>
<td>1.13</td>
<td>0.39</td>
<td>0.99</td>
</tr>
<tr>
<td>3rd</td>
<td>0.60</td>
<td>0.69</td>
<td>-0.05</td>
<td>0.54</td>
</tr>
</tbody>
</table>

All values are percent of the personal record.

BL, basic level; AV, average of seasonal values.

— Not significantly different between seasons.
0.66, 0.61, −0.02, for the first, second, and third seasons, respectively). For long-term training, the effect was significantly positively increasing for all the groups, except GN. For this group, the effect was significantly increasing from negative to positive values (PN: 0.17, 0.59, 1.04; GG: 0.26, 0.68, 1.13; GN: −0.48, −0.07, 0.39; and NG: 0.12, 0.53, 0.99).

Table 5 presents $r^2$ and average standard error values (ASE) for all subjects. The average $r^2$ was 37.7%, range 15.1–65.5%. The mean ASE was 0.33, range 0.15–0.55. Figure 3 shows real and modeled performance for subject 1 (A) and subject 7 (B). For subject 1, ASE was 0.27, this value indicated an error of 0.32 s for a 200-m event performed in less than 2 min. For subject 7, optimal modeled performance was situated shortly after the main seasonal competition, corresponding to the last event of the first and second season.

**DISCUSSION**

The main findings of this study are that:

1. Reactions to the short-, mid- and long-term training periods ($X_1$, $X_2$, and $X_3$, respectively) were significantly different for swimmers when divided into four groups. Interpretation of groups was based on terms of relative (to the other groups) good, poor, or neutral reaction to training. Swimmers reacting poorly to the short or mid-periods and reacting neutrally to the long-term training period (PN) composed the first group. The second group was composed by those reacting well to the three training periods (GG). The third group was composed by those reacting well to the short or mid-periods and neutrally to the long-term training period (GN). Swimmers reacting neutrally to the short and mid-periods and well to the long-term training period composed the fourth group (NG). This classification of swimmers from cluster and principal component analyses was incorporated into the mixed model.

2. The impact of training on performance changed significantly from the first to the third season, for the mid- and long-term training period. For $X_2$, the effect decreased, inversely for $X_3$.

3. Proposed model fitted significantly training and performance data for all the subjects.

**Specific adaptations to training for each group.**

Reactions to training periods were significantly different according to the four clusters. These responses were independent of training load differences which were integrated into the model (10). Differences between athletes’ reactions to the same training program matched previous studies (23).

| Subject S1 S2 S3 S4 S5 S6 S7 S8 S9 S10 S11 S12 S13 |
|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| $r^2$ (%) | 46.5     | 26.0     | 40.9     | 43.1     | 20.1     | 46.5     | 51.7     | 15.1     | 31.0     | 27.4     | 65.5     | 57.5     | 20.6     |
| ASE for   | 0.27     | 0.28     | 0.44     | 0.23     | 0.34     | 0.38     | 0.44     | 0.32     | 0.15     | 0.55     | 0.17     | 0.38     | 0.32     |

ASE, average standard error expressed in percent of the personal record.
Some swimmers are more tolerant to the training than others (7). The four different adaptation groups found in the present study confirmed studies by Hohmann (16), that differentiated athletes reacting “early” to training from others reacting “late.”

**Effect of short-term training period.** Training exerted a negative effect on PN, GG, GN, and NG, which is consistent with taper studies. Morton et al. (19) and Mujika et al. (22) proposed a 2- to 3-wk taper duration, on average. Taper allows for the recovery of previous accumulated fatigue, and the conservation of acquired capacities over such a period (17,22).

Training exerted the slightest effect on GN. This group was composed by the two youngest females, with the lowest performance level of swimmers in the study. These swimmers performed the lowest volume of dry-land training that implies a lower muscular fatigue (7). Furthermore, optimal recovery length seems to be longer for elite athletes than for sedentary subjects performing a moderate endurance training (5). In addition, adolescents rely less on glycolytic metabolism than adults during intense training (15), implying less fatigue and, so, shorter taper (27).

**Effect of mid-term training period.** Mid-term training had a negative influence on PN and a mean positive influence on the other groups. Several studies (4,11,20,21) confirm the importance of this period during which the increase in training load volume and intensity creates a powerful training impulse and a differed stimulation of biological adaptations through an overcompensation process (11,20,27). Nevertheless, there is a threshold to the impact of training loads on the organism. Furthermore, for certain loads, a greater amount of training implies the risk of overtraining (20,21).

PN was composed by three older male sprinters, specialized in breaststroke (Table 1). For these swimmers, cumulative effects of high Z3, Z4, Z5, and Z6 volumes implies, in all likelihood, a severe impact on the organism and a longer recovery period (7,23). In addition, energy consumption in breaststroke is very high (28), probably causing more fatigue and longer recovery (8). Also, adaptation capacity seems to decrease with age (14).

Mid-term training had a positive influence on GG. Some studies point out the interest of increasing training volume and aerobic loads for this type of swimmer (7). Coyle et al. (9) observed that maximal O₂ uptake, which is the most solicited quality for distance swimmers (28), declined by 7% during the first 3 wk of inactivity.

Training had a moderate mean positive effect on NG. This group was composed by 200-m swimmers: 1 female (S1) and 3 male swimmers (S2, S3, S12), and a female 100-m swimmer (S8). NG had a slight mean weekly Z2, Z3, Z4, Z6, and Z7 volume; Two-hundred-meter races have been described as maximal aerobic and anaerobic metabolism solicitation (28). For this group, aerobic (Z1, Z2), maximal oxygen consumption (Z3), anaerobic (Z4), and dry-land (Z6, Z7) training should probably be maintained for a sufficiently long period before competition to avoid detraining (24). On the other hand, a sufficiently long taper is needed to recover from the cumulative effect of these high training loads, generating severe fatigue and “overreaching” risks (18,21). In this respect, optimal training duration ensures overcompensation without the downfall of overtraining manifestations (11).

**Effect of long-term training period.** Training had a positive effect on PN, GG, and NG. A high training volume has to be maintained over this period to develop main qualities for further specific training periods (8,9). Weeks 5–12 before the performance are the most suitable weeks for hard training (4,19,20). Training had a slight mean negative effect on GN. These results emphasized that, for athletes with a short-term training background, schedules based on a regular and continuous distribution of training loads are more efficient than schedules based on the amount of hard training units, when comparing overcompensation (27).

**Influence of the season on the effect of training.** Subject reactions to mid- and long-term training were significantly modified between the first and third season. This result confirmed that swimmer reaction to training changes when identical training load is reiteratively applied through time (4,6,8). Busso et al. (6) suggested that variations over time in the model parameters appear with training. In the present study, response variations as a function of time were similar to the one found by Busso et al. (6). The application of the same training load through time leads to a negligible decrease of performance, in the mid period (2–3 wk), and a bigger decrease in the long training period (4–6 wk). First, repeated training through time would cause accumulated effects that improve performance (4,6,7,8). On the other hand, performance through years would be improved by a progressive increase of volume and intensity of training (6,12,21). Elite athletes would need to train more than untrained athletes to maintain the same progression (2). Conversely, they need a longer recovery time (19).

Finally, different reactions to training between different groups of subjects are known to be highly individualized (23). These differences can be attributed to genetic factors (30), individual training background (17,22), psychological factors (2), and technical factors (28). Thus, these results suggest that major personalization of the training programs has to be prescribed for each individual swimmer, depend-

### TABLE 6. Individual coefficients and SE for the two considered random effects.

<table>
<thead>
<tr>
<th>Subject</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
<th>S4</th>
<th>S5</th>
<th>S6</th>
<th>S7</th>
<th>S8</th>
<th>S9</th>
<th>S10</th>
<th>S11</th>
<th>S12</th>
<th>S13</th>
</tr>
</thead>
<tbody>
<tr>
<td>b0 (%)</td>
<td>-0.10</td>
<td>0.07</td>
<td>0.21</td>
<td>0.11</td>
<td>0.22</td>
<td>-0.29</td>
<td>-0.07</td>
<td>-0.10</td>
<td>-0.03</td>
<td>0.16</td>
<td>0.03</td>
<td>-0.22</td>
<td>-0.15</td>
</tr>
<tr>
<td>SE</td>
<td>0.28</td>
<td>0.28</td>
<td>0.30</td>
<td>0.29</td>
<td>0.29</td>
<td>0.30</td>
<td>0.30</td>
<td>0.30</td>
<td>0.32</td>
<td>0.31</td>
<td>0.32</td>
<td>0.30</td>
<td>0.30</td>
</tr>
<tr>
<td>b3 (%)</td>
<td>-0.01</td>
<td>0.07</td>
<td>0.16</td>
<td>0.14</td>
<td>0.13</td>
<td>-0.21</td>
<td>-0.08</td>
<td>-0.05</td>
<td>-0.03</td>
<td>0.13</td>
<td>0.03</td>
<td>-0.11</td>
<td>-0.10</td>
</tr>
<tr>
<td>SE</td>
<td>0.15</td>
<td>0.14</td>
<td>0.14</td>
<td>0.15</td>
<td>0.14</td>
<td>0.14</td>
<td>0.14</td>
<td>0.14</td>
<td>0.16</td>
<td>0.15</td>
<td>0.16</td>
<td>0.14</td>
<td>0.14</td>
</tr>
</tbody>
</table>

All values are expressed in percent of the personal record. b0 and b3 reflect deviation of the individual profiles from the population profile.
ing on age, specialty, training background, and individual profiles.

**Fitting the model.** Results of fitting showed that the proposed mathematical model was a correct method to describe the relationship between training and performance. For 13 elite swimmers over three different training seasons, including rest time, the fit between actual and modeled performance was statistically significant. The average $r^2$ was 37.7%, with extreme values ranging from 15.1% to 65.5%. The mean average standard error for estimated performances was 0.33, with extreme values 0.15 and 0.55. From a practical point of view, these values are accurate (for subject 1, ASE was 0.32 s for a 200-m event performed in less than 2 min).

Although transformations through the seasons and common and individual variability were taken into account, $r^2$ were lower than data reported in swimming (22,23) or other sports (4,19). Interpreting and comparing different $r^2$ is difficult. Indeed, the sample size can have a marked effect on the calculated value of the determination coefficient. For small samples, the mean $r^2$ may assume a high value even though there is no relationship between predictor and response variables (1). A longer period of study could also explain these lower fitting values. In addition, even if $r^2$ is one of the most important measures of the adequacy of prediction equations, a high value of $r^2$ does not necessarily guarantee accurate prediction (13). A complementary measure (like confidence interval used in the present study) is needed to indicate accuracy and sensitivity. Furthermore, classification in seven intensity levels may be insufficient when compared with the high number of different training variables used by the coaches in their programs (6,22,23). Psychological, nutritional (2,19), and technical factors (28) are related to performance. Hence, the long-term and cumulative effects of training that may be evidenced some macrocycles after training may not have been taken into account (8).

**CONCLUSION**

Modeling relationships between training and performance by a linear mixed model has made contributions to the Banister model. Identification of immediate and delayed reactions to training for each group as well as individual reactions help to improve the personalized distribution of training load through the 8 wk preceding competitions. Evidence of the different evolutions through the training seasons suggests that multi-annual planning must be made most carefully. The graphical representation of the real performance versus modeled performance may estimate, with a fixed confidence interval, the athlete’s responses to training load week after week. A supplementary study, with a larger number of subjects would be necessary to generalize these results.

**APPENDIX**

The linear mixed model explaining, in particular, the relationship between training and performance, taking into account training variables and season and group covariates can be expressed as follows:

1. $P_{ij} = (\beta_{00} + \beta_{01}G_1 + \beta_{02}G_2 + \beta_{03}G_3 + \beta_{04}A_{1ij} + \beta_{05}A_{2ij} + \beta_{10}G_1 + \beta_{11}G_2 + \beta_{12}G_3 + \beta_{13}G_2 + \beta_{14}A_{1ij} + \beta_{15}A_{2ij})X_{ij} + (\beta_{20} + \beta_{21}G_1 + \beta_{22}G_2 + \beta_{23}G_3 + \beta_{24}A_{1ij} + \beta_{25}A_{2ij})X_{2ij} + (\beta_{30} + \beta_{31}G_1 + \beta_{32}G_2 + \beta_{33}G_3 + \beta_{34}A_{1ij} + \beta_{35}A_{2ij})X_{3ij} + \beta_{0i} + \beta_{1i}X_{ij} + \beta_{2i}X_{2ij} + \beta_{3i}X_{3ij} + \epsilon_{ij}$

2. $\epsilon_i = (\epsilon_{1i}, \epsilon_{2i}, \ldots, \epsilon_{ni}) \sim N(0, \Sigma_i)$

3. $D$ is a diagonal matrix: $D = \text{diag}[\sigma_{01}^2, \sigma_{02}^2, \sigma_{03}^2, \sigma_{04}^2, \sigma_{05}^2, \sigma_{10}^2, \sigma_{11}^2, \sigma_{12}^2, \sigma_{13}^2, \sigma_{14}^2, \sigma_{15}^2, \sigma_{20}^2, \sigma_{21}^2, \sigma_{22}^2, \sigma_{23}^2, \sigma_{24}^2, \sigma_{25}^2, \sigma_{30}^2, \sigma_{31}^2, \sigma_{32}^2, \sigma_{33}^2, \sigma_{34}^2, \sigma_{35}^2]$

4. $\Sigma_i$ is an autoregressive matrix: $\Sigma_i(k, l) = \sigma^2 \gamma_{p}^{(k-l)}$, $k, l = 1, \ldots, n_i$.

After fitting and testing, the model was modified and reduced to:

5. $P_{ij} = (97.52 + 1.18G_1 - 2.12G_2 - 0.20G_3) + (-1.04 + 0.02G_1 + 0.30G_2 + 0.76G_3)X_{ij} + (-0.02 - 0.77G_1 + 0.22G_2 + 0.51G_3 + 0.67A_{1ij} + 0.62A_{2ij})X_{2ij} + (0.99 + 0.06G_1 + 0.15G_2 - 0.60G_3 - 0.87A_{1ij} - 0.46A_{2ij})X_{3ij} + \beta_{0i} + \beta_{1i}X_{ij} + \beta_{2i}X_{2ij} + \beta_{3i}X_{3ij} + \epsilon_{ij}$

6. $\epsilon_i = (\epsilon_{1i}, \epsilon_{2i}, \ldots, \epsilon_{ni}) \sim N(0, \Sigma_i)$, where $D = \text{diag}[0.11, 0.03]$ and $\Sigma_i(k, l) = 2.04 - 0.51^{k-l}$.

Individual coefficients and the associated standard errors are given in Table 6 for the two considered random effects.

The authors thank Frédéric Barale and Lucien Lacoite, coaches of the Pôle France of Toulouse, for their invaluable contribution. Thanks for financial support go to Laboratory of Physiology, GIP Exercise, Faculty of Medicine, Saint-Etienne University, France.


